

// Solution1 report

Compilation

```
bigdata@bigdata-VirtualBox:~$ export HADOOP_CLASSPATH=$(HADOOP_HOME/bin/hadoop
classpath)
bigdata@bigdata-VirtualBox:~$ echo $HADOOP_CLASSPATH
/usr/share/hadoop/etc/hadoop:/usr/share/hadoop-2.7.3/share/hadoop/common/lib/*:/
usr/share/hadoop-2.7.3/share/hadoop/common/*:/usr/share/hadoop-2.7.3/share/hadoo
p/hdfs:/usr/share/hadoop-2.7.3/share/hadoop/hdfs/lib/*:/usr/share/hadoop-2.7.3/s
hare/hadoop/hdfs/*:/usr/share/hadoop-2.7.3/share/hadoop/yarn/lib/*:/usr/share/ha
doo-2.7.3/share/hadoop/yarn/*:/usr/share/hadoop-2.7.3/share/hadoop/mapreduce/li
b/*:/usr/share/hadoop-2.7.3/share/hadoop/mapreduce/*:/usr/share/hadoop/contrib/c
apacity-scheduler/*.jar
bigdata@bigdata-VirtualBox:~$ javac -cp $HADOOP_CLASSPATH solution1.java
```

creation of jar file

```
bigdata@bigdata-VirtualBox:~$ jar cvf solution1.jar solution1*.class
added manifest
adding: solution1.class(in = 1345) (out= 691)(deflated 48%)
```

```
sol1_in1.txt
sol1_in1.txt~
sol1_in2.txt
sol1_in2.txt~
solution1.class
solution1.jar
solution1.java
```

uploading to HDFS two small files for testing

```
bigdata@bigdata-VirtualBox:~$ $HADOOP_HOME/bin/hadoop fs -put sol1_in1.txt
bigdata@bigdata-VirtualBox:~$ $HADOOP_HOME/bin/hadoop fs -put sol1_in2.txt
bigdata@bigdata-VirtualBox:~$ $HADOOP_HOME/bin/hadoop fs -ls
Found 3 items
drwxr-xr-x - bigdata supergroup          0 2017-07-03 01:33 .sparkStaging
-rw-r--r--  1 bigdata supergroup         47 2023-10-19 19:13 sol1_in1.txt
-rw-r--r--  1 bigdata supergroup         45 2023-10-19 19:13 sol1_in2.txt
```

processing of the application

```
bigdata@bigdata-VirtualBox:~$ $HADOOP_HOME/bin/hadoop jar /home/bigdata/solution
1.jar solution1 sol1_in1.txt sol1_in2.txt sol1_output.txt
bigdata@bigdata-VirtualBox:~$ $HADOOP_HOME/bin/hadoop fs -ls
Found 4 items
drwxr-xr-x - bigdata supergroup          0 2017-07-03 01:33 .sparkStaging
-rw-r--r--  1 bigdata supergroup         47 2023-10-19 19:13 sol1_in1.txt
-rw-r--r--  1 bigdata supergroup         45 2023-10-19 19:13 sol1_in2.txt
-rw-r--r--  1 bigdata supergroup        92 2023-10-19 19:14 sol1_output.txt
```

an evidence that two files uploaded into HDFS has been successful merges in one file in HDFS.

```
bigdata@bigdata-VirtualBox:~$ $HADOOP_HOME/bin/hadoop fs -cat sol1_in1.txt
This is a test input file1 that store in hdfs.
bigdata@bigdata-VirtualBox:~$ $HADOOP_HOME/bin/hadoop fs -cat sol1_in2.txt
This is test input file2 that store in hdfs.
bigdata@bigdata-VirtualBox:~$ $HADOOP_HOME/bin/hadoop fs -cat sol1_output.txt
This is a test input file1 that store in hdfs.
This is test input file2 that store in hdfs.
```

//Solution2 report

Compilation

```
bigdata@bigdata-VirtualBox:~$ export HADOOP_CLASSPATH=$(HADOOP_HOME/bin/hadoop
classpath)
bigdata@bigdata-VirtualBox:~$ echo $HADOOP_CLASSPATH
/usr/share/hadoop/etc/hadoop:/usr/share/hadoop-2.7.3/share/hadoop/common/lib/*:/
usr/share/hadoop-2.7.3/share/hadoop/common/*:/usr/share/hadoop-2.7.3/share/hadoo
p/hdfs:/usr/share/hadoop-2.7.3/share/hadoop/hdfs/lib/*:/usr/share/hadoop-2.7.3/s
hare/hadoop/hdfs/*:/usr/share/hadoop-2.7.3/share/hadoop/yarn/lib/*:/usr/share/ha
doo-2.7.3/share/hadoop/yarn/*:/usr/share/hadoop-2.7.3/share/hadoop/mapreduce/li
b/*:/usr/share/hadoop-2.7.3/share/hadoop/mapreduce/*:/usr/share/hadoop/etc/hadoo
p:/usr/share/hadoop-2.7.3/share/hadoop/common/lib/*:/usr/share/hadoop-2.7.3/shar
e/hadoop/common/*:/usr/share/hadoop-2.7.3/share/hadoop/hdfs:/usr/share/hadoop-2.
7.3/share/hadoop/hdfs/lib/*:/usr/share/hadoop-2.7.3/share/hadoop/hdfs/*:/usr/sha
re/hadoop-2.7.3/share/hadoop/yarn/lib/*:/usr/share/hadoop-2.7.3/share/hadoop/yar
n/*:/usr/share/hadoop-2.7.3/share/hadoop/mapreduce/lib/*:/usr/share/hadoop-2.7.3
/share/hadoop/mapreduce/*:/usr/share/hadoop/contrib/capacity-scheduler/*.jar:/us
r/share/hadoop/contrib/capacity-scheduler/*.jar
bigdata@bigdata-VirtualBox:~$ javac -cp $HADOOP_CLASSPATH solution2.java
```

creating jar file

```
bigdata@bigdata-VirtualBox:~$ jar cvf solution2.jar solution2*.class
added manifest
adding: solution2.class(in = 1468) (out= 798)(deflated 45%)
adding: solution2$TextOverSpeedReducer.class(in = 1808) (out= 794)(deflated 56%)
adding: solution2$TokenizerMapper.class(in = 1840) (out= 804)(deflated 56%)
```

```
solution2.class
solution2.jar
solution2.java
solution2.java~
solution2$TextOverSpeedReducer.class
solution2$TokenizerMapper.class
```

listing of your input file with the speed measurements.

```
bigdata@bigdata-VirtualBox:~$ $HADOOP_HOME/bin/hadoop fs -put SpeedCamera.txt
bigdata@bigdata-VirtualBox:~$ $HADOOP_HOME/bin/hadoop fs -ls
Found 5 items
drwxr-xr-x   - bigdata supergroup          0 2017-07-03 01:33 .sparkStaging
-rw-r--r--   1 bigdata supergroup        545 2023-10-19 19:57 SpeedCamera.txt
-rw-r--r--   1 bigdata supergroup         47 2023-10-19 19:13 sol1_in1.txt
-rw-r--r--   1 bigdata supergroup         45 2023-10-19 19:13 sol1_in2.txt
-rw-r--r--   1 bigdata supergroup         92 2023-10-19 19:14 sol1_output.txt
```

```
bigdata@bigdata-VirtualBox:~$ $HADOOP_HOME/bin/hadoop fs -cat SpeedCamera.txt
PKR856 14-NOV-2021 140
UPS234 20-FEB-2022 180
PKR856 20-MAR-2020 90
PKR856 17-JUN-2021 70
UPS234 22-SEP-2022 200
UPS234 03-AUG-2020 130
PKR856 14-NOV-2021 140
UPS234 20-FEB-2022 180
PKR856 20-MAR-2020 90
PKR856 17-JUN-2021 70
UPS234 22-SEP-2022 200
UPS234 03-AUG-2020 130
PKR856 14-NOV-2021 140
UPS234 20-FEB-2022 180
PKR856 20-MAR-2020 90
PKR856 17-JUN-2021 70
UPS234 22-SEP-2022 200
UPS234 03-AUG-2020 130
PKR856 14-NOV-2021 140
UPS234 20-FEB-2022 180
PKR856 20-MAR-2020 90
PKR856 17-JUN-2021 70
UPS234 22-SEP-2022 200
UPS234 03-AUG-2020 130
```

displaying the results of processing solution2.java

```
bigdata@bigdata-VirtualBox:~$ $HADOOP_HOME/bin/hadoop jar /home/bigdata/solution2.jar solution2 SpeedCamera.txt sol2_output
```

```
bigdata@bigdata-VirtualBox:~$ $HADOOP_HOME/bin/hadoop fs -cat sol2_output/*
PKR856 100
UPS234 170
```


//Solution3 report

Compilation creating jar file

```
bigdata@bigdata-VirtualBox:~$ javac -cp $HADOOP_CLASSPATH solution3.java
bigdata@bigdata-VirtualBox:~$ jar cvf solution3.jar solution3*.class
added manifest
adding: solution3.class(in = 1473) (out= 797)(deflated 45%)
adding: solution3$IntMinMaxAveSumReducer.class(in = 1812) (out= 786)(deflated 56%)
adding: solution3$TokenizerMapper.class(in = 1836) (out= 799)(deflated 56%)
```

processing, and displaying the results of processing solution3.java

```
bigdata@bigdata-VirtualBox:~$ $HADOOP_HOME/bin/hadoop fs -cat sol3_output/*
bolt      51
bolt      17
drill      2
drill      1
nail       5
nail       5
screw     168
screw      56
washer    66
washer    22
```