

```
2022-05-01 06:31:42,375 INFO org.apache.hadoop.yarn.client.RMPProxy (main): Connecting to
ResourceManager at ip-172-31-23-208.ec2.internal/172.31.23.208:8032
2022-05-01 06:31:44,498 INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat (main):
Total input files to process : 1
2022-05-01 06:31:44,522 INFO com.hadoop.compression.lzo.GPLNativeCodeLoader (main): Loaded
native gpl library
2022-05-01 06:31:44,525 INFO com.hadoop.compression.lzo.LzoCodec (main): Successfully loaded
& initialized native-lzo library [hadoop-lzo rev 29dbe82f1ecfd8384b89f1a32e5e2d9e5dfd7724]
2022-05-01 06:31:45,134 INFO org.apache.hadoop.mapreduce.JobSubmitter (main): number of
splits:7
2022-05-01 06:31:45,309 INFO org.apache.hadoop.mapreduce.JobSubmitter (main): Submitting
tokens for job: job_1651386557939_0001
2022-05-01 06:31:45,848 INFO org.apache.hadoop.yarn.client.api.impl.YarnClientImpl (main):
Submitted application application_1651386557939_0001
2022-05-01 06:31:45,953 INFO org.apache.hadoop.mapreduce.Job (main): The url to track the
job: http://ip-172-31-23-208.ec2.internal:20888/proxy/application_1651386557939_0001/
2022-05-01 06:31:45,966 INFO org.apache.hadoop.mapreduce.Job (main): Running job:
job_1651386557939_0001
2022-05-01 06:31:57,179 INFO org.apache.hadoop.mapreduce.Job (main): Job
job_1651386557939_0001 running in uber mode : false
2022-05-01 06:31:57,181 INFO org.apache.hadoop.mapreduce.Job (main): map 0% reduce 0%
2022-05-01 06:32:07,264 INFO org.apache.hadoop.mapreduce.Job (main): map 14% reduce 0%
2022-05-01 06:32:08,278 INFO org.apache.hadoop.mapreduce.Job (main): map 86% reduce 0%
2022-05-01 06:32:09,284 INFO org.apache.hadoop.mapreduce.Job (main): map 100% reduce 0%
2022-05-01 06:32:09,292 INFO org.apache.hadoop.mapreduce.Job (main): Job
job_1651386557939_0001 completed successfully
2022-05-01 06:32:09,468 INFO org.apache.hadoop.mapreduce.Job (main): Counters: 36
```

#### File System Counters

```
FILE: Number of bytes read=0
FILE: Number of bytes written=1175524
FILE: Number of read operations=0
FILE: Number of large read operations=0
FILE: Number of write operations=0
HDFS: Number of bytes read=686
HDFS: Number of bytes written=0
HDFS: Number of read operations=7
HDFS: Number of large read operations=0
HDFS: Number of write operations=0
S3: Number of bytes read=441669447
S3: Number of bytes written=8794495
S3: Number of read operations=0
S3: Number of large read operations=0
S3: Number of write operations=0
```

#### Job Counters

```
Killed map tasks=1
Launched map tasks=7
Data-local map tasks=7
Total time spent by all maps in occupied slots (ms)=2864928
Total time spent by all reduces in occupied slots (ms)=0
Total time spent by all map tasks (ms)=59686
Total vcore-milliseconds taken by all map tasks=59686
Total megabyte-milliseconds taken by all map tasks=91677696
```

#### Map-Reduce Framework

```
Map input records=165660
Map output records=3299
Input split bytes=686
Spilled Records=0
Failed Shuffles=0
Merged Map outputs=0
GC time elapsed (ms)=1295
CPU time spent (ms)=25630
Physical memory (bytes) snapshot=3350851584
Virtual memory (bytes) snapshot=23104659456
Total committed heap usage (bytes)=3099066368
```

## File Input Format Counters

Bytes Read=441669447

## File Output Format Counters

Bytes Written=8794495

2022-05-01 06:32:09,566 INFO org.apache.hadoop.yarn.client.RMPProxy (main): Connecting to ResourceManager at ip-172-31-23-208.ec2.internal/172.31.23.208:8032

2022-05-01 06:32:10,068 INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat (main): Total input files to process : 1

2022-05-01 06:32:10,186 INFO org.apache.hadoop.mapreduce.JobSubmitter (main): number of splits:7

2022-05-01 06:32:10,251 INFO org.apache.hadoop.mapreduce.JobSubmitter (main): Submitting tokens for job: job\_1651386557939\_0002

2022-05-01 06:32:10,283 INFO org.apache.hadoop.yarn.client.api.impl.YarnClientImpl (main): Submitted application application\_1651386557939\_0002

2022-05-01 06:32:10,288 INFO org.apache.hadoop.mapreduce.Job (main): The url to track the job: http://ip-172-31-23-208.ec2.internal:20888/proxy/application\_1651386557939\_0002/

2022-05-01 06:32:10,288 INFO org.apache.hadoop.mapreduce.Job (main): Running job: job\_1651386557939\_0002

2022-05-01 06:32:21,370 INFO org.apache.hadoop.mapreduce.Job (main): Job job\_1651386557939\_0002 running in uber mode : false

2022-05-01 06:32:21,370 INFO org.apache.hadoop.mapreduce.Job (main): map 0% reduce 0%

2022-05-01 06:32:34,330 INFO org.apache.hadoop.mapreduce.Job (main): map 14% reduce 0%

2022-05-01 06:32:38,348 INFO org.apache.hadoop.mapreduce.Job (main): map 86% reduce 0%

2022-05-01 06:32:39,352 INFO org.apache.hadoop.mapreduce.Job (main): map 100% reduce 0%

2022-05-01 06:32:40,355 INFO org.apache.hadoop.mapreduce.Job (main): map 100% reduce 100%

2022-05-01 06:32:40,361 INFO org.apache.hadoop.mapreduce.Job (main): Job job\_1651386557939\_0002 completed successfully

2022-05-01 06:32:40,397 INFO org.apache.hadoop.mapreduce.Job (main): Counters: 57

## File System Counters

FILE: Number of bytes read=112

FILE: Number of bytes written=1364311

FILE: Number of read operations=0

FILE: Number of large read operations=0

FILE: Number of write operations=0

HDFS: Number of bytes read=1736

HDFS: Number of bytes written=0

HDFS: Number of read operations=17

HDFS: Number of large read operations=0

HDFS: Number of write operations=2

S3: Number of bytes read=441673791

S3: Number of bytes written=0

S3: Number of read operations=0

S3: Number of large read operations=0

S3: Number of write operations=0

## Job Counters

Killed map tasks=1

Launched map tasks=7

Launched reduce tasks=1

Data-local map tasks=7

Total time spent by all maps in occupied slots (ms)=3946848

Total time spent by all reduces in occupied slots (ms)=329088

Total time spent by all map tasks (ms)=82226

Total time spent by all reduce tasks (ms)=3428

Total vcore-milliseconds taken by all map tasks=82226

Total vcore-milliseconds taken by all reduce tasks=3428

Total megabyte-milliseconds taken by all map tasks=126299136

Total megabyte-milliseconds taken by all reduce tasks=10530816

## Map-Reduce Framework

Map input records=165660

Map output records=313290

Map output bytes=3428160

Map output materialized bytes=294

Input split bytes=1736

Combine input records=313290

Combine output records=14

```
Reduce input groups=2
Reduce shuffle bytes=294
Reduce input records=14
Reduce output records=0
Spilled Records=28
Shuffled Maps =7
Failed Shuffles=0
Merged Map outputs=7
GC time elapsed (ms)=1999
CPU time spent (ms)=86530
Physical memory (bytes) snapshot=6079045632
Virtual memory (bytes) snapshot=27699249152
Total committed heap usage (bytes)=6011486208

Shuffle Errors
BAD_ID=0
CONNECTION=0
IO_ERROR=0
WRONG_LENGTH=0
WRONG_MAP=0
WRONG_REDUCE=0

decisiontree.DecisionTreeTest$Counter
Accuracy=89
Record_Count=165660

File Input Format Counters
Bytes Read=0

File Output Format Counters
Bytes Written=0

2022-05-01 06:32:40,407 INFO decisiontree.DecisionTreeTest (main):
*****
2022-05-01 06:32:40,407 INFO decisiontree.DecisionTreeTest (main): 165660
2022-05-01 06:32:40,407 INFO decisiontree.DecisionTreeTest (main):
*****
```