

```
2019-09-21 16:41:25,512 INFO org.apache.hadoop.yarn.client.RMProxy (main): Connecting to
ResourceManager at ip-172-31-70-185.ec2.internal/172.31.70.185:8032
2019-09-21 16:41:30,786 INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat (main):
Total input files to process : 1
2019-09-21 16:41:30,800 INFO com.hadoop.compression.lzo.GPLNativeCodeLoader (main): Loaded
native gpl library
2019-09-21 16:41:30,803 INFO com.hadoop.compression.lzo.LzoCodec (main): Successfully
loaded & initialized native-lzo library [hadoop-lzo rev
7e6c862e89bc8db32c064454a55af74ddff73bae]
2019-09-21 16:41:31,211 INFO org.apache.hadoop.mapreduce.JobSubmitter (main): number of
splits:20
2019-09-21 16:41:31,396 INFO org.apache.hadoop.mapreduce.JobSubmitter (main): Submitting
tokens for job: job_1569084022349_0001
2019-09-21 16:41:32,062 INFO org.apache.hadoop.yarn.client.api.impl.YarnClientImpl (main):
Submitted application application_1569084022349_0001
2019-09-21 16:41:32,432 INFO org.apache.hadoop.mapreduce.Job (main): The url to track the
job: http://ip-172-31-70-185.ec2.internal:20888/proxy/application_1569084022349_0001/
2019-09-21 16:41:32,432 INFO org.apache.hadoop.mapreduce.Job (main): Running job:
job_1569084022349_0001
2019-09-21 16:41:43,680 INFO org.apache.hadoop.mapreduce.Job (main): Job
job_1569084022349_0001 running in uber mode : false
2019-09-21 16:41:43,681 INFO org.apache.hadoop.mapreduce.Job (main): map 0% reduce 0%
2019-09-21 16:42:10,883 INFO org.apache.hadoop.mapreduce.Job (main): map 3% reduce 0%
2019-09-21 16:42:12,893 INFO org.apache.hadoop.mapreduce.Job (main): map 7% reduce 0%
2019-09-21 16:42:22,952 INFO org.apache.hadoop.mapreduce.Job (main): map 8% reduce 0%
2019-09-21 16:42:23,957 INFO org.apache.hadoop.mapreduce.Job (main): map 15% reduce 0%
2019-09-21 16:42:24,962 INFO org.apache.hadoop.mapreduce.Job (main): map 25% reduce 0%
2019-09-21 16:42:25,971 INFO org.apache.hadoop.mapreduce.Job (main): map 33% reduce 0%
2019-09-21 16:42:26,975 INFO org.apache.hadoop.mapreduce.Job (main): map 36% reduce 0%
2019-09-21 16:42:29,991 INFO org.apache.hadoop.mapreduce.Job (main): map 41% reduce 0%
2019-09-21 16:42:30,997 INFO org.apache.hadoop.mapreduce.Job (main): map 48% reduce 0%
2019-09-21 16:42:32,008 INFO org.apache.hadoop.mapreduce.Job (main): map 52% reduce 0%
2019-09-21 16:42:33,013 INFO org.apache.hadoop.mapreduce.Job (main): map 55% reduce 0%
2019-09-21 16:42:36,032 INFO org.apache.hadoop.mapreduce.Job (main): map 56% reduce 0%
2019-09-21 16:42:37,039 INFO org.apache.hadoop.mapreduce.Job (main): map 60% reduce 0%
2019-09-21 16:42:38,043 INFO org.apache.hadoop.mapreduce.Job (main): map 63% reduce 0%
2019-09-21 16:42:47,106 INFO org.apache.hadoop.mapreduce.Job (main): map 67% reduce 0%
2019-09-21 16:42:50,118 INFO org.apache.hadoop.mapreduce.Job (main): map 68% reduce 0%
2019-09-21 16:42:51,124 INFO org.apache.hadoop.mapreduce.Job (main): map 77% reduce 0%
2019-09-21 16:42:52,128 INFO org.apache.hadoop.mapreduce.Job (main): map 82% reduce 0%
2019-09-21 16:42:53,134 INFO org.apache.hadoop.mapreduce.Job (main): map 83% reduce 0%
2019-09-21 16:42:56,150 INFO org.apache.hadoop.mapreduce.Job (main): map 85% reduce 0%
2019-09-21 16:42:57,158 INFO org.apache.hadoop.mapreduce.Job (main): map 87% reduce 0%
2019-09-21 16:42:58,164 INFO org.apache.hadoop.mapreduce.Job (main): map 92% reduce 0%
2019-09-21 16:42:59,170 INFO org.apache.hadoop.mapreduce.Job (main): map 93% reduce 0%
2019-09-21 16:43:00,177 INFO org.apache.hadoop.mapreduce.Job (main): map 97% reduce 0%
2019-09-21 16:43:01,182 INFO org.apache.hadoop.mapreduce.Job (main): map 98% reduce 0%
2019-09-21 16:43:02,187 INFO org.apache.hadoop.mapreduce.Job (main): map 100% reduce 0%
2019-09-21 16:43:07,205 INFO org.apache.hadoop.mapreduce.Job (main): map 100% reduce 11%
2019-09-21 16:43:11,221 INFO org.apache.hadoop.mapreduce.Job (main): map 100% reduce 33%
2019-09-21 16:43:15,233 INFO org.apache.hadoop.mapreduce.Job (main): map 100% reduce 56%
2019-09-21 16:43:16,236 INFO org.apache.hadoop.mapreduce.Job (main): map 100% reduce 78%
2019-09-21 16:43:18,245 INFO org.apache.hadoop.mapreduce.Job (main): map 100% reduce 89%
2019-09-21 16:43:19,252 INFO org.apache.hadoop.mapreduce.Job (main): map 100% reduce 100%
```

2019-09-21 16:43:20,261 INFO org.apache.hadoop.mapreduce.Job (main): Job
job_1569084022349_0001 completed successfully
2019-09-21 16:43:20,411 INFO org.apache.hadoop.mapreduce.Job (main): Counters: 56

File System Counters

FILE: Number of bytes read=62489917
FILE: Number of bytes written=160349521
FILE: Number of read operations=0
FILE: Number of large read operations=0
FILE: Number of write operations=0
HDFS: Number of bytes read=2120
HDFS: Number of bytes written=0
HDFS: Number of read operations=20
HDFS: Number of large read operations=0
HDFS: Number of write operations=0
S3: Number of bytes read=1319431946
S3: Number of bytes written=87522407
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
Killed reduce tasks=1
Launched map tasks=20
Launched reduce tasks=10
Data-local map tasks=20
Total time spent by all maps in occupied slots (ms)=58147344
Total time spent by all reduces in occupied slots (ms)=13898592
Total time spent by all map tasks (ms)=1211403
Total time spent by all reduce tasks (ms)=144777
Total vcore-milliseconds taken by all map tasks=1211403
Total vcore-milliseconds taken by all reduce tasks=144777
Total megabyte-milliseconds taken by all map tasks=1860715008
Total megabyte-milliseconds taken by all reduce tasks=444754944

Map-Reduce Framework

Map input records=85331845
Map output records=85331845
Map output bytes=961483442
Map output materialized bytes=92935036
Input split bytes=2120
Combine input records=85331845
Combine output records=15362582
Reduce input groups=6626985
Reduce shuffle bytes=92935036
Reduce input records=15362582
Reduce output records=6626985
Spilled Records=30725164
Shuffled Maps =180
Failed Shuffles=0
Merged Map outputs=180
GC time elapsed (ms)=21322
CPU time spent (ms)=530930
Physical memory (bytes) snapshot=20123680768
Virtual memory (bytes) snapshot=108195426304
Total committed heap usage (bytes)=17896046592

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=1319431946
File Output Format Counters
  Bytes Written=87522407
```