

```
2020-02-15 05:23:59,733 INFO org.apache.hadoop.yarn.client.RMPProxy (main): Connecting to
ResourceManager at ip-172-31-76-158.ec2.internal/172.31.76.158:8032
2020-02-15 05:24:01,857 INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat (main):
Total input files to process : 1
2020-02-15 05:24:01,868 INFO com.hadoop.compression.lzo.GPLNativeCodeLoader (main): Loaded
native gpl library
2020-02-15 05:24:01,870 INFO com.hadoop.compression.lzo.LzoCodec (main): Successfully loaded
& initialized native-lzo library [hadoop-lzo rev 5f788d5e8f90539ee331702c753fa250727128f4]
2020-02-15 05:24:02,235 INFO org.apache.hadoop.mapreduce.JobSubmitter (main): number of
splits:20
2020-02-15 05:24:02,329 INFO org.apache.hadoop.mapreduce.JobSubmitter (main): Submitting
tokens for job: job_1581744132020_0001
2020-02-15 05:24:02,642 INFO org.apache.hadoop.yarn.client.api.impl.YarnClientImpl (main):
Submitted application application_1581744132020_0001
2020-02-15 05:24:02,800 INFO org.apache.hadoop.mapreduce.Job (main): The url to track the
job: http://ip-172-31-76-158.ec2.internal:20888/proxy/application_1581744132020_0001/
2020-02-15 05:24:02,817 INFO org.apache.hadoop.mapreduce.Job (main): Running job:
job_1581744132020_0001
2020-02-15 05:24:09,900 INFO org.apache.hadoop.mapreduce.Job (main): Job
job_1581744132020_0001 running in uber mode : false
2020-02-15 05:24:09,902 INFO org.apache.hadoop.mapreduce.Job (main): map 0% reduce 0%
2020-02-15 05:25:40,306 INFO org.apache.hadoop.mapreduce.Job (main): map 10% reduce 0%
2020-02-15 05:25:41,309 INFO org.apache.hadoop.mapreduce.Job (main): map 15% reduce 0%
2020-02-15 05:25:47,328 INFO org.apache.hadoop.mapreduce.Job (main): map 20% reduce 0%
2020-02-15 05:25:48,332 INFO org.apache.hadoop.mapreduce.Job (main): map 25% reduce 0%
2020-02-15 05:25:50,337 INFO org.apache.hadoop.mapreduce.Job (main): map 30% reduce 0%
2020-02-15 05:25:52,342 INFO org.apache.hadoop.mapreduce.Job (main): map 32% reduce 0%
2020-02-15 05:25:56,354 INFO org.apache.hadoop.mapreduce.Job (main): map 36% reduce 1%
2020-02-15 05:25:57,356 INFO org.apache.hadoop.mapreduce.Job (main): map 36% reduce 3%
2020-02-15 05:26:02,368 INFO org.apache.hadoop.mapreduce.Job (main): map 36% reduce 4%
2020-02-15 05:26:05,375 INFO org.apache.hadoop.mapreduce.Job (main): map 36% reduce 5%
2020-02-15 05:26:48,471 INFO org.apache.hadoop.mapreduce.Job (main): map 38% reduce 5%
2020-02-15 05:26:49,477 INFO org.apache.hadoop.mapreduce.Job (main): map 43% reduce 5%
2020-02-15 05:26:50,479 INFO org.apache.hadoop.mapreduce.Job (main): map 43% reduce 6%
2020-02-15 05:26:52,483 INFO org.apache.hadoop.mapreduce.Job (main): map 48% reduce 6%
2020-02-15 05:26:54,490 INFO org.apache.hadoop.mapreduce.Job (main): map 60% reduce 6%
2020-02-15 05:26:55,492 INFO org.apache.hadoop.mapreduce.Job (main): map 64% reduce 6%
2020-02-15 05:26:56,494 INFO org.apache.hadoop.mapreduce.Job (main): map 64% reduce 7%
2020-02-15 05:26:57,497 INFO org.apache.hadoop.mapreduce.Job (main): map 64% reduce 9%
2020-02-15 05:26:58,499 INFO org.apache.hadoop.mapreduce.Job (main): map 71% reduce 9%
2020-02-15 05:27:00,504 INFO org.apache.hadoop.mapreduce.Job (main): map 73% reduce 9%
2020-02-15 05:27:01,507 INFO org.apache.hadoop.mapreduce.Job (main): map 74% reduce 9%
2020-02-15 05:27:02,509 INFO org.apache.hadoop.mapreduce.Job (main): map 74% reduce 10%
2020-02-15 05:27:03,512 INFO org.apache.hadoop.mapreduce.Job (main): map 84% reduce 11%
2020-02-15 05:27:04,515 INFO org.apache.hadoop.mapreduce.Job (main): map 86% reduce 11%
2020-02-15 05:27:07,520 INFO org.apache.hadoop.mapreduce.Job (main): map 90% reduce 11%
2020-02-15 05:27:08,522 INFO org.apache.hadoop.mapreduce.Job (main): map 90% reduce 12%
2020-02-15 05:27:09,524 INFO org.apache.hadoop.mapreduce.Job (main): map 90% reduce 13%
2020-02-15 05:27:10,526 INFO org.apache.hadoop.mapreduce.Job (main): map 93% reduce 15%
2020-02-15 05:27:11,528 INFO org.apache.hadoop.mapreduce.Job (main): map 93% reduce 17%
2020-02-15 05:27:13,533 INFO org.apache.hadoop.mapreduce.Job (main): map 96% reduce 17%
2020-02-15 05:27:14,535 INFO org.apache.hadoop.mapreduce.Job (main): map 96% reduce 20%
2020-02-15 05:27:15,538 INFO org.apache.hadoop.mapreduce.Job (main): map 96% reduce 21%
2020-02-15 05:27:17,542 INFO org.apache.hadoop.mapreduce.Job (main): map 96% reduce 22%
2020-02-15 05:27:18,543 INFO org.apache.hadoop.mapreduce.Job (main): map 96% reduce 24%
2020-02-15 05:27:19,545 INFO org.apache.hadoop.mapreduce.Job (main): map 96% reduce 27%
2020-02-15 05:27:28,564 INFO org.apache.hadoop.mapreduce.Job (main): map 96% reduce 29%
2020-02-15 05:29:06,740 INFO org.apache.hadoop.mapreduce.Job (main): map 100% reduce 29%
2020-02-15 05:29:07,742 INFO org.apache.hadoop.mapreduce.Job (main): map 100% reduce 40%
2020-02-15 05:29:08,745 INFO org.apache.hadoop.mapreduce.Job (main): map 100% reduce 92%
2020-02-15 05:29:15,757 INFO org.apache.hadoop.mapreduce.Job (main): map 100% reduce 100%
2020-02-15 05:29:16,762 INFO org.apache.hadoop.mapreduce.Job (main): Job
job_1581744132020_0001 completed successfully
2020-02-15 05:29:16,859 INFO org.apache.hadoop.mapreduce.Job (main): Counters: 56
```

#### File System Counters

FILE: Number of bytes read=4456182  
FILE: Number of bytes written=14527927  
FILE: Number of read operations=0  
FILE: Number of large read operations=0  
FILE: Number of write operations=0  
HDFS: Number of bytes read=2240  
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=27705130793  
S3: Number of bytes written=53  
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=13  
Data-local map tasks=20  
Total time spent by all maps in occupied slots (ms)=284181504  
Total time spent by all reduces in occupied slots (ms)=380735424  
Total time spent by all map tasks (ms)=2960224  
Total time spent by all reduce tasks (ms)=1982997  
Total vcore-milliseconds taken by all map tasks=2960224  
Total vcore-milliseconds taken by all reduce tasks=1982997  
Total megabyte-milliseconds taken by all map tasks=9093808128  
Total megabyte-milliseconds taken by all reduce tasks=12183533568

#### Map-Reduce Framework

Map input records=85331845  
Map output records=7234833  
Map output bytes=79583163  
Map output materialized bytes=4460108  
Input split bytes=2240  
Combine input records=0  
Combine output records=0  
Reduce input groups=1  
Reduce shuffle bytes=4460108  
Reduce input records=7234833  
Reduce output records=1  
Spilled Records=14469666  
Shuffled Maps =260  
Failed Shuffles=0  
Merged Map outputs=260  
GC time elapsed (ms)=58152  
CPU time spent (ms)=2843650  
Physical memory (bytes) snapshot=33269178368  
Virtual memory (bytes) snapshot=188111032320  
Total committed heap usage (bytes)=31505514496

#### 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=1319494293

#### File Output Format Counters

Bytes Written=53