

```
2020-02-15 00:50:24,873 INFO org.apache.hadoop.yarn.client.RMPProxy (main): Connecting to
ResourceManager at ip-172-31-8-181.ec2.internal/172.31.8.181:8032
2020-02-15 00:50:27,104 INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat (main):
Total input files to process : 1
2020-02-15 00:50:27,114 INFO com.hadoop.compression.lzo.GPLNativeCodeLoader (main): Loaded
native gpl library
2020-02-15 00:50:27,117 INFO com.hadoop.compression.lzo.LzoCodec (main): Successfully loaded
& initialized native-lzo library [hadoop-lzo rev 5f788d5e8f90539ee331702c753fa250727128f4]
2020-02-15 00:50:27,424 INFO org.apache.hadoop.mapreduce.JobSubmitter (main): number of
splits:20
2020-02-15 00:50:27,528 INFO org.apache.hadoop.mapreduce.JobSubmitter (main): Submitting
tokens for job: job_1581727705049_0001
2020-02-15 00:50:27,982 INFO org.apache.hadoop.yarn.client.api.impl.YarnClientImpl (main):
Submitted application application_1581727705049_0001
2020-02-15 00:50:28,093 INFO org.apache.hadoop.mapreduce.Job (main): The url to track the
job: http://ip-172-31-8-181.ec2.internal:20888/proxy/application_1581727705049_0001/
2020-02-15 00:50:28,098 INFO org.apache.hadoop.mapreduce.Job (main): Running job:
job_1581727705049_0001
2020-02-15 00:50:35,174 INFO org.apache.hadoop.mapreduce.Job (main): Job
job_1581727705049_0001 running in uber mode : false
2020-02-15 00:50:35,175 INFO org.apache.hadoop.mapreduce.Job (main): map 0% reduce 0%
2020-02-15 00:52:08,645 INFO org.apache.hadoop.mapreduce.Job (main): map 5% reduce 0%
2020-02-15 00:52:09,649 INFO org.apache.hadoop.mapreduce.Job (main): map 10% reduce 0%
2020-02-15 00:52:17,673 INFO org.apache.hadoop.mapreduce.Job (main): map 11% reduce 0%
2020-02-15 00:52:20,682 INFO org.apache.hadoop.mapreduce.Job (main): map 16% reduce 0%
2020-02-15 00:52:29,718 INFO org.apache.hadoop.mapreduce.Job (main): map 17% reduce 0%
2020-02-15 00:52:30,721 INFO org.apache.hadoop.mapreduce.Job (main): map 20% reduce 0%
2020-02-15 00:52:31,724 INFO org.apache.hadoop.mapreduce.Job (main): map 20% reduce 1%
2020-02-15 00:52:45,766 INFO org.apache.hadoop.mapreduce.Job (main): map 20% reduce 2%
2020-02-15 00:53:13,843 INFO org.apache.hadoop.mapreduce.Job (main): map 25% reduce 2%
2020-02-15 00:53:14,845 INFO org.apache.hadoop.mapreduce.Job (main): map 30% reduce 2%
2020-02-15 00:53:16,850 INFO org.apache.hadoop.mapreduce.Job (main): map 40% reduce 2%
2020-02-15 00:53:17,853 INFO org.apache.hadoop.mapreduce.Job (main): map 45% reduce 3%
2020-02-15 00:53:18,856 INFO org.apache.hadoop.mapreduce.Job (main): map 65% reduce 3%
2020-02-15 00:53:19,859 INFO org.apache.hadoop.mapreduce.Job (main): map 75% reduce 4%
2020-02-15 00:53:21,864 INFO org.apache.hadoop.mapreduce.Job (main): map 75% reduce 5%
2020-02-15 00:53:23,870 INFO org.apache.hadoop.mapreduce.Job (main): map 80% reduce 7%
2020-02-15 00:53:24,873 INFO org.apache.hadoop.mapreduce.Job (main): map 85% reduce 7%
2020-02-15 00:53:25,876 INFO org.apache.hadoop.mapreduce.Job (main): map 90% reduce 7%
2020-02-15 00:53:27,881 INFO org.apache.hadoop.mapreduce.Job (main): map 95% reduce 8%
2020-02-15 00:53:30,889 INFO org.apache.hadoop.mapreduce.Job (main): map 100% reduce 26%
2020-02-15 00:53:31,892 INFO org.apache.hadoop.mapreduce.Job (main): map 100% reduce 75%
2020-02-15 00:53:32,895 INFO org.apache.hadoop.mapreduce.Job (main): map 100% reduce 91%
2020-02-15 00:53:35,903 INFO org.apache.hadoop.mapreduce.Job (main): map 100% reduce 100%
2020-02-15 00:53:36,910 INFO org.apache.hadoop.mapreduce.Job (main): Job
job_1581727705049_0001 completed successfully
2020-02-15 00:53:37,004 INFO org.apache.hadoop.mapreduce.Job (main): Counters: 55
File System Counters
    FILE: Number of bytes read=961569
    FILE: Number of bytes written=7196954
    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=27705128295
    S3: Number of bytes written=52
    S3: Number of read operations=0
    S3: Number of large read operations=0
    S3: Number of write operations=0
Job Counters
```

```
Killed map tasks=3
Launched map tasks=20
Launched reduce tasks=11
Data-local map tasks=20
Total time spent by all maps in occupied slots (ms)=288041472
Total time spent by all reduces in occupied slots (ms)=54491712
Total time spent by all map tasks (ms)=3000432
Total time spent by all reduce tasks (ms)=283811
Total vcore-milliseconds taken by all map tasks=3000432
Total vcore-milliseconds taken by all reduce tasks=283811
Total megabyte-milliseconds taken by all map tasks=9217327104
Total megabyte-milliseconds taken by all reduce tasks=1743734784
Map-Reduce Framework
  Map input records=85331845
  Map output records=1560888
  Map output bytes=17169768
  Map output materialized bytes=964911
  Input split bytes=2240
  Combine input records=0
  Combine output records=0
  Reduce input groups=1
  Reduce shuffle bytes=964911
  Reduce input records=1560888
  Reduce output records=1
  Spilled Records=3121776
  Shuffled Maps =220
  Failed Shuffles=0
  Merged Map outputs=220
  GC time elapsed (ms)=59558
  CPU time spent (ms)=2822460
  Physical memory (bytes) snapshot=31764369408
  Virtual memory (bytes) snapshot=173268332544
  Total committed heap usage (bytes)=30263476224
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=1319491795
File Output Format Counters
  Bytes Written=52
```