

```
2020-01-31 01:38:46,999 INFO org.apache.hadoop.yarn.client.RMPProxy (main): Connecting to
ResourceManager at ip-172-31-11-22.ec2.internal/172.31.11.22:8032
2020-01-31 01:38:49,254 INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat (main):
Total input files to process : 1
2020-01-31 01:38:49,266 INFO com.hadoop.compression.lzo.GPLNativeCodeLoader (main): Loaded
native gpl library
2020-01-31 01:38:49,268 INFO com.hadoop.compression.lzo.LzoCodec (main): Successfully loaded
& initialized native-lzo library [hadoop-lzo rev 5f788d5e8f90539ee331702c753fa250727128f4]
2020-01-31 01:38:49,580 INFO org.apache.hadoop.mapreduce.JobSubmitter (main): number of
splits:20
2020-01-31 01:38:49,674 INFO org.apache.hadoop.mapreduce.JobSubmitter (main): Submitting
tokens for job: job_1580434644593_0001
2020-01-31 01:38:50,131 INFO org.apache.hadoop.yarn.client.api.impl.YarnClientImpl (main):
Submitted application application_1580434644593_0001
2020-01-31 01:38:50,254 INFO org.apache.hadoop.mapreduce.Job (main): The url to track the
job: http://ip-172-31-11-22.ec2.internal:20888/proxy/application_1580434644593_0001/
2020-01-31 01:38:50,255 INFO org.apache.hadoop.mapreduce.Job (main): Running job:
job_1580434644593_0001
2020-01-31 01:38:57,330 INFO org.apache.hadoop.mapreduce.Job (main): Job
job_1580434644593_0001 running in uber mode : false
2020-01-31 01:38:57,331 INFO org.apache.hadoop.mapreduce.Job (main): map 0% reduce 0%
2020-01-31 01:39:14,421 INFO org.apache.hadoop.mapreduce.Job (main): map 7% reduce 0%
2020-01-31 01:39:15,427 INFO org.apache.hadoop.mapreduce.Job (main): map 13% reduce 0%
2020-01-31 01:39:16,441 INFO org.apache.hadoop.mapreduce.Job (main): map 18% reduce 0%
2020-01-31 01:39:17,444 INFO org.apache.hadoop.mapreduce.Job (main): map 20% reduce 0%
2020-01-31 01:39:18,448 INFO org.apache.hadoop.mapreduce.Job (main): map 59% reduce 0%
2020-01-31 01:39:21,458 INFO org.apache.hadoop.mapreduce.Job (main): map 60% reduce 0%
2020-01-31 01:39:23,465 INFO org.apache.hadoop.mapreduce.Job (main): map 64% reduce 0%
2020-01-31 01:39:24,468 INFO org.apache.hadoop.mapreduce.Job (main): map 77% reduce 0%
2020-01-31 01:39:26,475 INFO org.apache.hadoop.mapreduce.Job (main): map 80% reduce 0%
2020-01-31 01:39:27,478 INFO org.apache.hadoop.mapreduce.Job (main): map 83% reduce 0%
2020-01-31 01:39:29,484 INFO org.apache.hadoop.mapreduce.Job (main): map 88% reduce 0%
2020-01-31 01:39:30,487 INFO org.apache.hadoop.mapreduce.Job (main): map 93% reduce 0%
2020-01-31 01:39:31,491 INFO org.apache.hadoop.mapreduce.Job (main): map 100% reduce 0%
2020-01-31 01:39:32,494 INFO org.apache.hadoop.mapreduce.Job (main): map 100% reduce 6%
2020-01-31 01:39:33,496 INFO org.apache.hadoop.mapreduce.Job (main): map 100% reduce 18%
2020-01-31 01:39:34,499 INFO org.apache.hadoop.mapreduce.Job (main): map 100% reduce 21%
2020-01-31 01:39:35,503 INFO org.apache.hadoop.mapreduce.Job (main): map 100% reduce 55%
2020-01-31 01:39:38,512 INFO org.apache.hadoop.mapreduce.Job (main): map 100% reduce 64%
2020-01-31 01:39:39,516 INFO org.apache.hadoop.mapreduce.Job (main): map 100% reduce 91%
2020-01-31 01:39:41,523 INFO org.apache.hadoop.mapreduce.Job (main): map 100% reduce 100%
2020-01-31 01:39:42,535 INFO org.apache.hadoop.mapreduce.Job (main): Job
job_1580434644593_0001 completed successfully
2020-01-31 01:39:42,758 INFO org.apache.hadoop.mapreduce.Job (main): Counters: 56
File System Counters
    FILE: Number of bytes read=63297557
    FILE: Number of bytes written=162978648
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=2060
    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=1319475045
    S3: Number of bytes written=67641452
    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=11
    Data-local map tasks=20
    Total time spent by all maps in occupied slots (ms)=50143776
    Total time spent by all reduces in occupied slots (ms)=21623616
    Total time spent by all map tasks (ms)=522331
    Total time spent by all reduce tasks (ms)=112623
    Total vcore-milliseconds taken by all map tasks=522331
    Total vcore-milliseconds taken by all reduce tasks=112623
    Total megabyte-milliseconds taken by all map tasks=1604600832
    Total megabyte-milliseconds taken by all reduce tasks=691955712
Map-Reduce Framework
    Map input records=85331845
    Map output records=85331845
    Map output bytes=961483442
    Map output materialized bytes=94415552
    Input split bytes=2060
    Combine input records=85331845
    Combine output records=15362582
    Reduce input groups=6626985
    Reduce shuffle bytes=94415552
    Reduce input records=15362582
    Reduce output records=6626985
    Spilled Records=30725164
    Shuffled Maps =220
    Failed Shuffles=0
    Merged Map outputs=220
    GC time elapsed (ms)=13435
    CPU time spent (ms)=410490
    Physical memory (bytes) snapshot=27920310272
    Virtual memory (bytes) snapshot=173288689664
    Total committed heap usage (bytes)=28577890304
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=1319475045
File Output Format Counters
    Bytes Written=67641452
```