```
2020-02-15 00:37:54,292 INFO org.apache.hadoop.yarn.client.RMProxy (main): Connecting to
ResourceManager at ip-172-31-65-49.ec2.internal/172.31.65.49:8032
2020-02-15 00:37:57,242 INFO org.apache.hadoop.mapreduce.lib.input.FileInputFormat (main):
Total input files to process : 1
2020-02-15 00:37:57,253 INFO com.hadoop.compression.lzo.GPLNativeCodeLoader (main): Loaded
native gpl library
2020-02-15 00:37:57,255 INFO com.hadoop.compression.lzo.LzoCodec (main): Successfully loaded
& initialized native-lzo library [hadoop-lzo rev 5f788d5e8f90539ee331702c753fa250727128f4]
2020-02-15 00:37:57,563 INFO org.apache.hadoop.mapreduce.JobSubmitter (main): number of
splits:20
2020-02-15 00:37:57,672 INFO org.apache.hadoop.mapreduce.JobSubmitter (main): Submitting
tokens for job: job_1581726996807_0001
2020-02-15 00:37:57,990 INFO org.apache.hadoop.yarn.client.api.impl.YarnClientImpl (main):
Submitted application application_1581726996807_0001
2020-02-15 00:37:58,032 INFO org.apache.hadoop.mapreduce.Job (main): The url to track the
job: http://ip-172-31-65-49.ec2.internal:20888/proxy/application_1581726996807_0001/
2020-02-15 00:37:58,032 INFO org.apache.hadoop.mapreduce.Job (main): Running job:
job_1581726996807_0001
2020-02-15 00:38:06,107 INFO org.apache.hadoop.mapreduce.Job (main): Job
job_1581726996807_0001 running in uber mode : false
2020-02-15 00:38:06,108 INFO org.apache.hadoop.mapreduce.Job (main):  map 0% reduce 0%
2020-02-15 00:39:33,473 INFO org.apache.hadoop.mapreduce.Job (main):  map 5% reduce 0%
2020-02-15 00:39:36,482 INFO org.apache.hadoop.mapreduce.Job (main):  map 10% reduce 0%
2020-02-15 00:39:38,489 INFO org.apache.hadoop.mapreduce.Job (main):  map 15% reduce 0%
2020-02-15 00:39:41,497 INFO org.apache.hadoop.mapreduce.Job (main):  map 20% reduce 0%
2020-02-15 00:39:51,524 INFO org.apache.hadoop.mapreduce.Job (main):  map 20% reduce 1%
2020-02-15 00:39:56,537 INFO org.apache.hadoop.mapreduce.Job (main):  map 20% reduce 2%
2020-02-15 00:40:39,637 INFO org.apache.hadoop.mapreduce.Job (main):  map 25% reduce 2%
2020-02-15 00:40:40,639 INFO org.apache.hadoop.mapreduce.Job (main):  map 35% reduce 2%
2020-02-15 00:40:41,641 INFO org.apache.hadoop.mapreduce.Job (main):  map 45% reduce 2%
2020-02-15 00:40:44,648 INFO org.apache.hadoop.mapreduce.Job (main):  map 45% reduce 3%
2020-02-15 00:40:45,650 INFO org.apache.hadoop.mapreduce.Job (main):  map 45% reduce 4%
2020-02-15 00:40:46,652 INFO org.apache.hadoop.mapreduce.Job (main):  map 60% reduce 4%
2020-02-15 00:40:48,657 INFO org.apache.hadoop.mapreduce.Job (main):  map 65% reduce 4%
2020-02-15 00:40:49,660 INFO org.apache.hadoop.mapreduce.Job (main):  map 75% reduce 4%
2020-02-15 00:40:50,662 INFO org.apache.hadoop.mapreduce.Job (main):  map 75% reduce 5%
2020-02-15 00:40:51,664 INFO org.apache.hadoop.mapreduce.Job (main):  map 85% reduce 6%
2020-02-15 00:40:52,666 INFO org.apache.hadoop.mapreduce.Job (main):  map 90% reduce 7%
2020-02-15 00:40:53,668 INFO org.apache.hadoop.mapreduce.Job (main):  map 95% reduce 7%
2020-02-15 00:40:54,670 INFO org.apache.hadoop.mapreduce.Job (main):  map 100% reduce 7%
2020-02-15 00:40:55,673 INFO org.apache.hadoop.mapreduce.Job (main):  map 100% reduce 64%
2020-02-15 00:40:56,676 INFO org.apache.hadoop.mapreduce.Job (main):  map 100% reduce 73%
2020-02-15 00:40:58,680 INFO org.apache.hadoop.mapreduce.Job (main):  map 100% reduce 91%
2020-02-15 00:40:59,683 INFO org.apache.hadoop.mapreduce.Job (main):  map 100% reduce 100%
2020-02-15 00:41:00,688 INFO org.apache.hadoop.mapreduce.Job (main): Job
job_1581726996807_0001 completed successfully
2020-02-15 00:41:00,800 INFO org.apache.hadoop.mapreduce.Job (main): Counters: 56
        File System Counters
                FILE: Number of bytes read=2260
                FILE: Number of bytes written=5278392
                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=27705144113
                S3: Number of bytes written=50
                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
                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)=278389248
                Total time spent by all reduces in occupied slots (ms)=55186560
                Total time spent by all map tasks (ms)=2899888
                Total time spent by all reduce tasks (ms)=287430
                Total vcore-milliseconds taken by all map tasks=2899888
                Total vcore-milliseconds taken by all reduce tasks=287430
                Total megabyte-milliseconds taken by all map tasks=8908455936
                Total megabyte-milliseconds taken by all reduce tasks=1765969920
        Map-Reduce Framework
                Map input records=85331845
                Map output records=3297
                Map output bytes=36267
                Map output materialized bytes=5684
                Input split bytes=2240
                Combine input records=0
                Combine output records=0
                Reduce input groups=1
                Reduce shuffle bytes=5684
                Reduce input records=3297
                Reduce output records=1
                Spilled Records=6594
                Shuffled Maps =220
                Failed Shuffles=0
                Merged Map outputs=220
                GC time elapsed (ms)=52157
                CPU time spent (ms)=2723480
                Physical memory (bytes) snapshot=32384008192
                Virtual memory (bytes) snapshot=173366706176
                Total committed heap usage (bytes)=31178358784
        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=1319507613
        File Output Format Counters
                Bytes Written=50
```