

```
zyh14848@cluster-de78-m:~/mapreduce$ hadoop jar /usr/lib/hadoop/hadoop-streaming.jar -file mapper.py -mapper 'python mapper.py' -file reducer.py -reducer 'python reducer.py' -input /data/* -output /output1
2021-11-09 01:57:10,579 WARN streaming.StreamJob: -file option is deprecated, please use generic option -files instead.
packageJobJar: [mapper.py, reducer.py] (/usr/lib/hadoop/hadoop-streaming-3.2.2-jar1 /tmp/streamjob1396670542475255459.jar tmpDir=null)
2021-11-09 01:57:11,421 INFO client.RMProxy: Connecting to ResourceManager at cluster-de78-m/10.142.0.3:8032
2021-11-09 01:57:11,663 INFO client.AHSProxy: Connecting to Application History server at cluster-de78-m/10.142.0.3:10200
2021-11-09 01:57:12,110 INFO client.RMProxy: Connecting to ResourceManager at cluster-de78-m/10.142.0.3:8032
2021-11-09 01:57:12,111 INFO client.AHSProxy: Connecting to Application History server at cluster-de78-m/10.142.0.3:10200
2021-11-09 01:57:12,278 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/zyh14848/.staging/job_1636421783896_0002
2021-11-09 01:57:12,619 INFO mapred.FileInputFormat: Total input files to process : 2
2021-11-09 01:57:12,686 INFO mapreduce.JobSubmitter: number of splits:22
2021-11-09 01:57:12,834 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1636421783896_0002
2021-11-09 01:57:12,835 INFO mapreduce.JobSubmitter: Executing with tokens: []
2021-11-09 01:57:13,062 INFO conf.Configuration: resource-types.xml not found
2021-11-09 01:57:13,062 INFO resource.ResourceCellitis: Unable to find 'resource-types.xml'.
2021-11-09 01:57:13,118 INFO impl.YarnClientImpl: Submitted application application_1636421783896_0002
2021-11-09 01:57:13,154 INFO mapreduce.Job: The url to track the job: http://cluster-de78-m:8088/proxy/application_1636421783896_0002/
2021-11-09 01:57:13,155 INFO mapreduce.Job: Running job: job_1636421783896_0002
2021-11-09 01:57:13,350 INFO mapreduce.Job: Job job_1636421783896_0002 running in uber mode : false
2021-11-09 01:57:13,350 INFO mapreduce.Job: map 0% reduce 0%
2021-11-09 01:57:28,445 INFO mapreduce.Job: map 14% reduce 0%
2021-11-09 01:57:30,456 INFO mapreduce.Job: map 32% reduce 0%
2021-11-09 01:57:35,486 INFO mapreduce.Job: map 45% reduce 0%
2021-11-09 01:57:39,508 INFO mapreduce.Job: map 64% reduce 0%
2021-11-09 01:57:41,519 INFO mapreduce.Job: map 73% reduce 0%
2021-11-09 01:57:42,524 INFO mapreduce.Job: map 77% reduce 0%
2021-11-09 01:57:44,533 INFO mapreduce.Job: map 82% reduce 0%
2021-11-09 01:57:46,543 INFO mapreduce.Job: map 91% reduce 0%
2021-11-09 01:57:47,548 INFO mapreduce.Job: map 100% reduce 0%
2021-11-09 01:57:53,561 INFO mapreduce.Job: map 100% reduce 14%
2021-11-09 01:57:54,586 INFO mapreduce.Job: map 100% reduce 29%
2021-11-09 01:57:55,591 INFO mapreduce.Job: map 100% reduce 43%
2021-11-09 01:57:56,595 INFO mapreduce.Job: map 100% reduce 71%
2021-11-09 01:57:57,600 INFO mapreduce.Job: map 100% reduce 86%
2021-11-09 01:57:58,604 INFO mapreduce.Job: map 100% reduce 100%
2021-11-09 01:58:01,524 INFO mapreduce.Job: Job job_1636421783896_0002 completed successfully
2021-11-09 01:58:01,707 INFO mapreduce.Job: Counters: 55

File System Counters
  FILE: Number of bytes read=192588
  FILE: Number of bytes written=7586827
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
  HDFS: Number of bytes read=1861002
  HDFS: Number of bytes written=9111
  HDFS: Number of read operations=101
  HDFS: Number of large read operations=0
  HDFS: Number of write operations=21
  HDFS: Number of bytes read erasure-coded=0

Job Counters
  Killed map tasks=1
  Launched map tasks=22
  Launched reduce tasks=7
  Data-local map tasks=22
  Total time spent by all maps in occupied slots (ms)=444727740
  Total time spent by all reduces in occupied slots (ms)=109001928
  Total time spent by all map tasks (ms)=140915
  Total time spent by all reduce tasks (ms)=34538
  Total vcore-milliseconds taken by all map tasks=140915
  Total vcore-milliseconds taken by all reduce tasks=34538
  Total megabyte-milliseconds taken by all map tasks=444727740
  Total megabyte-milliseconds taken by all reduce tasks=109001928

Map-Reduce Framework
  Map input records=13130
  Map output records=13129
  Map output bytes=166288
  Map output materialized bytes=193470
  Input split bytes=1914
  Combine input records=0
  Combine output records=0
  Reduce input groups=730
  Reduce shuffle bytes=193470
  Reduce input records=13129
  Reduce output records=730
  Spilled Records=26258
  Shuffled Maps =154
  Failed Shuffles=0
  Merged Map outputs=154
  GC time elapsed (ms)=4231
  CPU time spent (ms)=26730
  Physical memory (bytes) snapshot=12929708032
  Virtual memory (bytes) snapshot=129467588608
  Total committed heap usage (bytes)=130537222624
  Peak Map Physical memory (bytes)=539942912
  Peak Map Virtual memory (bytes)=4465610752
  Peak Reduce Physical memory (bytes)=288968704
  Peak Reduce Virtual memory (bytes)=4471607296

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=1859088

File Output Format Counters
  Bytes Written=9111
2021-11-09 01:58:01,708 INFO streaming.StreamJob: Output directory: /output1
zyh14848@cluster-de78-m:~/mapreduce$
```