

LOG FILE

Cloning into 'CloudComputing_HW4'...

Username for 'https://github.uc.edu': chintasd

Password for 'https://chintasd@github.uc.edu':

remote: Counting objects: 9, done.

remote: Compressing objects: 100% (7/7), done.

Unpacking objects: 100% (9/9), done.

remote: Total 9 (delta 1), reused 0 (delta 0), pack-reused 0

Checking connectivity... done.

chintasd@hadoop-gate-0:~\$ cd CloudComputing_HW4

chintasd@hadoop-gate-0:~/CloudComputing_HW4\$ hadoop jar /usr/hdp/2.6.3.0-235/hadoop-mapreduce/hadoop-streaming-2.7.3.2.6.3.0-235.jar -file mapper.py -mapper mapper.py -file reducer.py -reducer reducer.py -input /data/nyc/nyc-traffic.csv -output HW4

18/04/05 22:15:09 WARN streaming.StreamJob: -file option is deprecated, please use generic option -files instead.

packageJobJar: [mapper.py, reducer.py] [/usr/hdp/2.6.3.0-235/hadoop-mapreduce/hadoop-streaming-2.7.3.2.6.3.0-235.jar] /tmp/streamjob4297324824089782634.jar tmpDir=null

18/04/05 22:15:10 INFO client.RMProxy: Connecting to ResourceManager at hadoop2-0-0.cscloud.ceas.uc.edu/192.168.2.20:8050

18/04/05 22:15:10 INFO client.AHSPProxy: Connecting to Application History server at hadoop2-0-0.cscloud.ceas.uc.edu/192.168.2.20:10200

18/04/05 22:15:10 INFO client.RMProxy: Connecting to ResourceManager at hadoop2-0-0.cscloud.ceas.uc.edu/192.168.2.20:8050

18/04/05 22:15:10 INFO client.AHSPProxy: Connecting to Application History server at hadoop2-0-0.cscloud.ceas.uc.edu/192.168.2.20:10200

18/04/05 22:15:11 INFO mapred.FileInputFormat: Total input paths to process : 1

18/04/05 22:15:11 INFO net.NetworkTopology: Adding a new node: /default-rack/192.168.2.19:50010

18/04/05 22:15:11 INFO net.NetworkTopology: Adding a new node: /24-core/192.168.2.24:50010

18/04/05 22:15:11 INFO net.NetworkTopology: Adding a new node: /24-core/192.168.2.25:50010

18/04/05 22:15:11 INFO net.NetworkTopology: Adding a new node: /16-core/192.168.2.22:50010

18/04/05 22:15:11 INFO mapreduce.JobSubmitter: number of splits:2

18/04/05 22:15:11 INFO mapreduce.JobSubmitter: Submitting tokens for job:
job_1521804639914_0598

18/04/05 22:15:11 INFO impl.YarnClientImpl: Submitted application application_1521804639914_0598

18/04/05 22:15:11 INFO mapreduce.Job: The url to track the job: http://hadoop2-0-0.csccloud.ceas.uc.edu:8088/proxy/application_1521804639914_0598/

18/04/05 22:15:11 INFO mapreduce.Job: Running job: job_1521804639914_0598

18/04/05 22:15:18 INFO mapreduce.Job: Job job_1521804639914_0598 running in uber mode : false

18/04/05 22:15:18 INFO mapreduce.Job: map 0% reduce 0%

18/04/05 22:15:30 INFO mapreduce.Job: map 19% reduce 0%

18/04/05 22:15:31 INFO mapreduce.Job: map 39% reduce 0%

18/04/05 22:15:33 INFO mapreduce.Job: map 50% reduce 0%

18/04/05 22:15:34 INFO mapreduce.Job: map 61% reduce 0%

18/04/05 22:15:35 INFO mapreduce.Job: map 100% reduce 0%

18/04/05 22:15:46 INFO mapreduce.Job: map 100% reduce 98%

18/04/05 22:15:47 INFO mapreduce.Job: map 100% reduce 100%

18/04/05 22:15:47 INFO mapreduce.Job: Job job_1521804639914_0598 completed successfully

18/04/05 22:15:47 INFO mapreduce.Job: Counters: 49

File System Counters

FILE: Number of bytes read=40558930

FILE: Number of bytes written=81577879

FILE: Number of read operations=0

FILE: Number of large read operations=0

FILE: Number of write operations=0

HDFS: Number of bytes read=176307514

HDFS: Number of bytes written=314

HDFS: Number of read operations=9

HDFS: Number of large read operations=0

HDFS: Number of write operations=2

Job Counters

Launched map tasks=2

Launched reduce tasks=1

Other local map tasks=2

Total time spent by all maps in occupied slots (ms)=55044

Total time spent by all reduces in occupied slots (ms)=19168

Total time spent by all map tasks (ms)=27522

Total time spent by all reduce tasks (ms)=9584

Total vcore-milliseconds taken by all map tasks=27522

Total vcore-milliseconds taken by all reduce tasks=9584

Total megabyte-milliseconds taken by all map tasks=42273792

Total megabyte-milliseconds taken by all reduce tasks=19628032

Map-Reduce Framework

Map input records=924098

Map output records=1808611

Map output bytes=36941702

Map output materialized bytes=40558936

Input split bytes=240

Combine input records=0

Combine output records=0

Reduce input groups=17

Reduce shuffle bytes=40558936

Reduce input records=1808611

Reduce output records=17

Spilled Records=3617222

Shuffled Maps =2

Failed Shuffles=0

Merged Map outputs=2

GC time elapsed (ms)=761

CPU time spent (ms)=41300

Physical memory (bytes) snapshot=2702442496

Virtual memory (bytes) snapshot=10290552832

Total committed heap usage (bytes)=2685403136

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=176307274

File Output Format Counters

Bytes Written=314

18/04/05 22:15:47 INFO streaming.StreamJob: Output directory: HW4

chintasd@hadoop-gate-0:~/CloudComputing_HW4\$ ^C

chintasd@hadoop-gate-0:~/CloudComputing_HW4\$