

```

hduser@jasdeep-hacBookPro /usr/local/hadoop/hadoop-2.9.0$ bin/hadoop jar /usr/local/hadoop/hadoop-2.9.0/share/hadoop/tools/lib/hadoop-streaming-2.9.0.jar
-input /input/assignment51 -output /output/output1 -mapper /home/hduser/learning_files/assignment/mapper1.py -reducer /home/hduser/learning_files/as
signment/reducer1.py
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/hadoop-2.9.0/share/hadoop/com
mon/lib/hadoop-auth-2.9.0.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
18/05/01 15:06:20 INFO Configuration.deprecation: session.id is deprecated. Instead, use dfs.metrics.session.id
18/05/01 15:06:20 INFO jvm.JvmMetrics: Initializing JVM Metrics with processName=JobTracker, sessionId=
18/05/01 15:06:20 INFO jvm.JvmMetrics: Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
18/05/01 15:06:20 INFO mapred.FileInputFormat: Total input files to process : 2
18/05/01 15:06:20 INFO mapreduce.JobSubmitter: number of splits:2
18/05/01 15:06:21 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_local1627431328_0001
18/05/01 15:06:21 INFO mapreduce.Job: The url to track the job: http://localhost:8080/
18/05/01 15:06:21 INFO mapred.LocalJobRunner: OutputCommitter set in config null
18/05/01 15:06:21 INFO mapred.LocalJobRunner: OutputCommitter is org.apache.hadoop.mapred.FileOutputCommitter
18/05/01 15:06:21 INFO mapreduce.Job: Running job: job_local1627431328_0001
18/05/01 15:06:21 INFO output.FileOutputCommitter: FileOutputCommitter Algorithm version is 1
18/05/01 15:06:21 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders under output directory:false, ignore cleanup fat
lures: False
18/05/01 15:06:21 INFO mapred.LocalJobRunner: Waiting for map tasks
18/05/01 15:06:21 INFO mapred.LocalJobRunner: Starting task: attempt_local1627431328_0001_m_000000_0
18/05/01 15:06:21 INFO output.FileOutputCommitter: FileOutputCommitter Algorithm version is 1
18/05/01 15:06:21 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders under output directory:false, ignore cleanup fat
lures: False
18/05/01 15:06:21 INFO mapred.Task: Using ResourceCalculatorProcessTree : [ ]
18/05/01 15:06:21 INFO mapred.MapTask: Processing split: hdfs://localhost:54310/input/assignment51/robinson.0+624405
18/05/01 15:06:21 INFO mapred.MapTask: numReduceTasks: 1
18/05/01 15:06:21 INFO mapred.MapTask: (EQUATOR) 0 kvi 26214396(104857584)
18/05/01 15:06:21 INFO mapred.MapTask: mapreduce.task.sort.mb: 100
18/05/01 15:06:21 INFO mapred.MapTask: soft limit at 83886080
18/05/01 15:06:21 INFO mapred.MapTask: bufstart = 0; bufvoid = 104857600
18/05/01 15:06:21 INFO mapred.MapTask: kvstart = 26214396; length = 6553600
18/05/01 15:06:21 INFO mapred.MapTask: Map output collector class = org.apache.hadoop.mapred.MapTask$MapOutputBuffer
18/05/01 15:06:21 INFO streaming.PipeMapRed: PipeMapRed exec [/home/hduser/learning_files/assignment/mapper1.py]
18/05/01 15:06:21 INFO Configuration.deprecation: mapred.work.output.dir is deprecated. Instead, use mapreduce.task.output.dir
18/05/01 15:06:21 INFO Configuration.deprecation: map.input.start is deprecated. Instead, use mapreduce.map.input.start
18/05/01 15:06:21 INFO Configuration.deprecation: mapred.task.ismap is deprecated. Instead, use mapreduce.task.ismap
18/05/01 15:06:21 INFO Configuration.deprecation: mapred.task.id is deprecated. Instead, use mapreduce.task.attempt.id
18/05/01 15:06:21 INFO Configuration.deprecation: mapred.tip.id is deprecated. Instead, use mapreduce.task.id
18/05/01 15:06:21 INFO Configuration.deprecation: mapred.local.dir is deprecated. Instead, use mapreduce.cluster.local.dir
18/05/01 15:06:21 INFO Configuration.deprecation: map.input.file is deprecated. Instead, use mapreduce.map.input.file
18/05/01 15:06:21 INFO Configuration.deprecation: mapred.skip.on is deprecated. Instead, use mapreduce.job.skiprecords
18/05/01 15:06:21 INFO Configuration.deprecation: map.input.length is deprecated. Instead, use mapreduce.map.input.length
18/05/01 15:06:21 INFO Configuration.deprecation: mapred.job.id is deprecated. Instead, use mapreduce.job.id
18/05/01 15:06:21 INFO Configuration.deprecation: user.name is deprecated. Instead, use mapreduce.job.user.name
18/05/01 15:06:21 INFO Configuration.deprecation: user.name is deprecated. Instead, use mapreduce.job.user.name
18/05/01 15:06:21 INFO Configuration.deprecation: mapred.task.partition is deprecated. Instead, use mapreduce.task.partition
18/05/01 15:06:21 INFO streaming.PipeMapRed: R/W/S=1/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 15:06:21 INFO streaming.PipeMapRed: R/W/S=10/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 15:06:21 INFO streaming.PipeMapRed: R/W/S=100/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 15:06:22 INFO streaming.PipeMapRed: R/W/S=1000/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 15:06:22 INFO mapreduce.Job: Job job_local1627431328_0001 running in uber mode : false
18/05/01 15:06:22 INFO mapreduce.Job: map 0% reduce 0%
18/05/01 15:06:22 INFO streaming.PipeMapRed: Records R/W=2182/1
18/05/01 15:06:22 INFO streaming.PipeMapRed: R/W/S=10000/40950/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 15:06:22 INFO streaming.PipeMapRed: MRErrorThread done
18/05/01 15:06:22 INFO streaming.PipeMapRed: mapRedFinished
18/05/01 15:06:22 INFO mapred.LocalJobRunner:
18/05/01 15:06:22 INFO mapred.MapTask: Starting flush of map output
18/05/01 15:06:22 INFO mapred.MapTask: Spilling map output
18/05/01 15:06:22 INFO mapred.MapTask: bufstart = 0; bufend = 3351742; bufvoid = 104857600
18/05/01 15:06:22 INFO mapred.MapTask: kvstart = 26214396(104857584); kvend = 25989848(103959392); length = 224549/6553600
18/05/01 15:06:23 INFO mapred.MapTask: Finished spill 0
18/05/01 15:06:23 INFO mapred.Task: Task:attempt_local1627431328_0001_m_000000_0 is done. And is in the process of committing
18/05/01 15:06:23 INFO mapred.LocalJobRunner: Records R/W=2182/1
18/05/01 15:06:23 INFO mapred.Task: Task:attempt_local1627431328_0001_m_000000_0' done
18/05/01 15:06:23 INFO mapred.LocalJobRunner: Finishing task: attempt_local1627431328_0001_m_000000_0
18/05/01 15:06:23 INFO mapred.LocalJobRunner: Starting task: attempt_local1627431328_0001_m_000001_0
18/05/01 15:06:23 INFO output.FileOutputCommitter: FileOutputCommitter Algorithm version is 1
18/05/01 15:06:23 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders under output directory:false, ignore cleanup fat
lures: False
18/05/01 15:06:23 INFO mapred.Task: Using ResourceCalculatorProcessTree : [ ]
18/05/01 15:06:23 INFO mapred.MapTask: Processing split: hdfs://localhost:54310/input/assignment51/wild.0+175998
18/05/01 15:06:23 INFO mapred.MapTask: numReduceTasks: 1
18/05/01 15:06:23 INFO mapred.MapTask: (EQUATOR) 0 kvi 26214396(104857584)
18/05/01 15:06:23 INFO mapred.MapTask: mapreduce.task.sort.mb: 100
18/05/01 15:06:23 INFO mapred.MapTask: soft limit at 83886080
18/05/01 15:06:23 INFO mapred.MapTask: bufstart = 0; bufvoid = 104857600
18/05/01 15:06:23 INFO mapred.MapTask: kvstart = 26214396; length = 6553600
18/05/01 15:06:23 INFO mapred.MapTask: Map output collector class = org.apache.hadoop.mapred.MapTask$MapOutputBuffer
18/05/01 15:06:23 INFO streaming.PipeMapRed: PipeMapRed exec [/home/hduser/learning_files/assignment/mapper1.py]
18/05/01 15:06:23 INFO streaming.PipeMapRed: R/W/S=1/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 15:06:23 INFO streaming.PipeMapRed: R/W/S=10/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 15:06:23 INFO streaming.PipeMapRed: R/W/S=100/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 15:06:23 INFO streaming.PipeMapRed: R/W/S=1000/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 15:06:23 INFO streaming.PipeMapRed: Records R/W=2070/1
18/05/01 15:06:23 INFO streaming.PipeMapRed: MRErrorThread done
18/05/01 15:06:23 INFO streaming.PipeMapRed: mapRedFinished
18/05/01 15:06:23 INFO mapred.LocalJobRunner:
18/05/01 15:06:23 INFO mapred.MapTask: Starting flush of map output
18/05/01 15:06:23 INFO mapred.MapTask: Spilling map output
18/05/01 15:06:23 INFO mapred.MapTask: bufstart = 0; bufend = 950161; bufvoid = 104857600
18/05/01 15:06:23 INFO mapred.MapTask: kvstart = 26214396(104857584); kvend = 26146440(104585760); length = 67957/6553600
18/05/01 15:06:23 INFO mapred.MapTask: Finished spill 0

```

MapReduce task2

```

18/05/01 15:06:23 INFO mapred.Task: Task: attempt_local1627431328_0001_m_000001_0 is done. And is in the process of committing
18/05/01 15:06:23 INFO mapred.LocalJobRunner: Records R/W=2070/1
18/05/01 15:06:23 INFO mapred.Task: Task 'attempt_local1627431328_0001_m_000001_0' done.
18/05/01 15:06:23 INFO mapred.LocalJobRunner: Finishing task: attempt_local1627431328_0001_m_000001_0
18/05/01 15:06:23 INFO mapred.LocalJobRunner: map task executor complete.
18/05/01 15:06:23 INFO mapred.LocalJobRunner: Waiting for reduce tasks
18/05/01 15:06:23 INFO mapred.LocalJobRunner: Starting task: attempt_local1627431328_0001_r_000000_0
18/05/01 15:06:23 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 1
18/05/01 15:06:23 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders under output directory:false, ignore cleanup fa
lures: false
18/05/01 15:06:23 INFO mapred.Task: Using ResourceCalculatorProcessTree : [ ]
18/05/01 15:06:23 INFO mapred.ReduceTask: Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@13356eb0
18/05/01 15:06:23 INFO mapreduce.Job: map 100% reduce 0%
18/05/01 15:06:23 INFO reduce.MergeManagerImpl: MergerManager: memoryLimit=375809632, maxSingleShuffleLimit=93952408, mergeThreshold=248034368, toSort
actor=10, nanoToMemMergeOutputsThreshold=10
18/05/01 15:06:23 INFO reduce.EventFetcher: attempt_local1627431328_0001_r_000000_0 Thread started: EventFetcher for fetching Map Completion Events
18/05/01 15:06:23 INFO reduce.LocalFetcher: localfetcher#1 about to shuffle output of map attempt_local1627431328_0001_m_000000_0 decomp: 3464020 len:
3464024 to MEMORY
18/05/01 15:06:23 INFO reduce.InMemoryMapOutput: Read 3464020 bytes from map-output for attempt_local1627431328_0001_m_000000_0
18/05/01 15:06:23 INFO reduce.MergeManagerImpl: closeInMemoryFile -> map-output of size: 3464020, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, u
sedMemory -> 3464020
18/05/01 15:06:23 INFO reduce.LocalFetcher: localfetcher#1 about to shuffle output of map attempt_local1627431328_0001_m_000001_0 decomp: 984143 len:
984147 to MEMORY
18/05/01 15:06:23 INFO reduce.InMemoryMapOutput: Read 984143 bytes from map-output for attempt_local1627431328_0001_m_000001_0
18/05/01 15:06:23 INFO reduce.MergeManagerImpl: closeInMemoryFile -> map-output of size: 984143, inMemoryMapOutputs.size() -> 2, commitMemory -> 34640
0, usedMemory -> 4448163
18/05/01 15:06:23 INFO reduce.EventFetcher: EventFetcher is interrupted.. Returning
18/05/01 15:06:23 INFO mapred.LocalJobRunner: 2 / 2 copied.
18/05/01 15:06:23 INFO reduce.MergeManagerImpl: FinalMerge called with 2 in-memory map-outputs and 0 on-disk map-outputs
18/05/01 15:06:23 INFO mapred.Merger: Merging 2 sorted segments
18/05/01 15:06:23 INFO mapred.Merger: Down to the last merge-pass, with 2 segments left of total size: 4448048 bytes
18/05/01 15:06:24 INFO reduce.MergeManagerImpl: Merged 2 segments, 4448163 bytes to disk to satisfy reduce memory limit
18/05/01 15:06:24 INFO reduce.MergeManagerImpl: Merging 1 files, 4448165 bytes from disk
18/05/01 15:06:24 INFO reduce.MergeManagerImpl: Merging 0 segments, 0 bytes from memory into reduce
18/05/01 15:06:24 INFO mapred.Merger: Merging 1 sorted segments
18/05/01 15:06:24 INFO mapred.Merger: Down to the last merge-pass, with 1 segments left of total size: 4448105 bytes
18/05/01 15:06:24 INFO mapred.LocalJobRunner: 2 / 2 copied.
18/05/01 15:06:24 INFO streaming.PipeMapRed: PipeMapRed exec [/home/hduser/learning_files/assignment/reducer1.py]
18/05/01 15:06:24 INFO Configuration.deprecation: mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address
18/05/01 15:06:24 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps
18/05/01 15:06:24 INFO streaming.PipeMapRed: R/W/S=1/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 15:06:24 INFO streaming.PipeMapRed: R/W/S=10/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 15:06:24 INFO streaming.PipeMapRed: R/W/S=100/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 15:06:24 INFO streaming.PipeMapRed: R/W/S=1000/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 15:06:24 INFO streaming.PipeMapRed: Records R/W=2216/1
18/05/01 15:06:24 INFO streaming.PipeMapRed: R/W/S=10000/2140/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 15:06:24 INFO streaming.PipeMapRed: MRErrorThread done
18/05/01 15:06:24 INFO streaming.PipeMapRed: mapRedFinished
18/05/01 15:06:24 INFO streaming.PipeMapRed: mapRedFinished
18/05/01 15:06:24 INFO mapred.Task: Task: attempt_local1627431328_0001_r_000000_0 is done. And is in the process of committing
18/05/01 15:06:24 INFO mapred.LocalJobRunner: 2 / 2 copied.
18/05/01 15:06:24 INFO mapred.Task: Task: attempt_local1627431328_0001_r_000000_0 is allowed to commit now
18/05/01 15:06:24 INFO output.FileOutputCommitter: Saved output of task 'attempt_local1627431328_0001_r_000000_0' to hdfs://localhost:54310/output/outp
ut1/_temporary/0/task_local1627431328_0001_r_000000
18/05/01 15:06:24 INFO mapred.LocalJobRunner: Records R/W=2216/1 > reduce
18/05/01 15:06:24 INFO mapred.Task: Task 'attempt_local1627431328_0001_r_000000_0' done
18/05/01 15:06:24 INFO mapred.LocalJobRunner: Finishing task: attempt_local1627431328_0001_r_000000_0
18/05/01 15:06:24 INFO mapred.LocalJobRunner: reduce task executor complete.
18/05/01 15:06:24 INFO mapreduce.Job: map 100% reduce 100%
18/05/01 15:06:24 INFO mapreduce.Job: Job: job_local1627431328_0001 completed successfully
18/05/01 15:06:24 INFO mapreduce.Job: Counters: 35

File System Counters
  FILE: Number of bytes read=9303089
  FILE: Number of bytes written=18649854
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
  HDFS: Number of bytes read=2225211
  HDFS: Number of bytes written=1018194
  HDFS: Number of read operations=22
  HDFS: Number of large read operations=0
  HDFS: Number of write operations=5

Map-Reduce Framework
  Map input records=13490
  Map output records=73128
  Map output bytes=4301903
  Map output materialized bytes=4448171
  Input split bytes=200
  Combine input records=0
  Combine output records=0
  Reduce input groups=17084
  Reduce shuffle bytes=4448171
  Reduce input records=73128
  Reduce output records=17084
  Spilled Records=146256
  Shuffled Maps =2
  Failed Shuffles=0
  Merged Map outputs=2
  GC time elapsed (ms)=20
  Total committed heap usage (bytes)=707788800

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=800403
  wrong_reco...

File Input Format Counters
  Bytes Read=800403

File Output Format Counters
  Bytes Written=1018194
18/05/01 15:06:24 INFO streaming.StreamJob: Output directory: /output/output1
hduser@jasdeep-MacBookPro: /usr/local/hadoop/hadoop-2.9.0$ b.user.nampill 0
b.user.nampill: command not found
hduser@jasdeep-MacBookPro: /usr/local/hadoop/hadoop-2.9.0$ █

```

```
hduser@jasdeep-MacBookPro: /usr/local/hadoop/hadoop-2.9.0$ bin/hadoop jar /usr/local/hadoop/hadoop-2.9.0/share/hadoop/tools/lib/hadoop-streaming-2.9.0.jar -input /output/output1/part-00000 -output /output/output2 -mapper /home/hduser/learning_files/assignment/mapper2.py -reducer /home/hduser/learning_files/assignment/reducer2.py
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/hadoop-2.9.0/share/hadoop/common/lib/hadoop-auth-2.9.0.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
18/05/01 15:30:45 INFO Configuration.deprecation: session.id is deprecated. Instead, use dfs.metrics.session-id
18/05/01 15:30:45 INFO jvm.JvmMetrics: Initializing JVM Metrics with processName=JobTracker, sessionId=
18/05/01 15:30:45 INFO jvm.JvmMetrics: Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
18/05/01 15:30:45 INFO mapred.FileInputFormat: Total input files to process : 1
18/05/01 15:30:45 INFO mapreduce.JobSubmitter: number of splits:1
18/05/01 15:30:46 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_local705825676_0001
18/05/01 15:30:46 INFO mapreduce.Job: The url to track the job: http://localhost:8080/
18/05/01 15:30:46 INFO mapred.LocalJobRunner: OutputCommitter set in config null
18/05/01 15:30:46 INFO mapreduce.Job: Running job: job_local705825676_0001
18/05/01 15:30:46 INFO mapred.LocalJobRunner: OutputCommitter is org.apache.hadoop.mapred.FileOutputCommitter
18/05/01 15:30:46 INFO output.FileOutputCommitter: File output Committer Algorithm version is 1
18/05/01 15:30:46 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders under output directory:false, ignore cleanup failures: false
18/05/01 15:30:46 INFO mapred.LocalJobRunner: Waiting for map tasks
18/05/01 15:30:46 INFO mapred.LocalJobRunner: Starting task: attempt_local705825676_0001_m_000000_0
18/05/01 15:30:46 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 1
18/05/01 15:30:46 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders under output directory:false, ignore cleanup failures: false
18/05/01 15:30:46 INFO mapred.Task: Using ResourceCalculatorProcessTree : [ ]
18/05/01 15:30:46 INFO mapred.MapTask: Processing split: hdfs://localhost:54310/output/output1/part-00000_0+1018194
18/05/01 15:30:46 INFO mapred.MapTask: numReduceTasks: 1
18/05/01 15:30:46 INFO mapred.MapTask: (EQUATOR) 0 kvi. 26214396(104857584)
18/05/01 15:30:46 INFO mapred.MapTask: mapreduce.task.io.sort.mb: 100
18/05/01 15:30:46 INFO mapred.MapTask: soft limit at 83886080
18/05/01 15:30:46 INFO mapred.MapTask: bufstart = 0; bufvoid = 104857600
18/05/01 15:30:46 INFO mapred.MapTask: kvstart = 26214396; length = 6553600
18/05/01 15:30:46 INFO mapred.MapTask: Map output collector class = org.apache.hadoop.mapred.MapTask$MapOutputBuffer
18/05/01 15:30:46 INFO streaming.PipeMapRed: PipeMapRed exec [/home/hduser/learning_files/assignment/mapper2.py]
18/05/01 15:30:46 INFO Configuration.deprecation: mapred.work.output.dir is deprecated. Instead, use mapreduce.task.output.dir
18/05/01 15:30:46 INFO Configuration.deprecation: map.input.start is deprecated. Instead, use mapreduce.map.input.start
18/05/01 15:30:46 INFO Configuration.deprecation: mapred.task.id is deprecated. Instead, use mapreduce.task.ismap
18/05/01 15:30:46 INFO Configuration.deprecation: mapred.task.id is deprecated. Instead, use mapreduce.task.attempt.id
18/05/01 15:30:46 INFO Configuration.deprecation: mapred.task.id is deprecated. Instead, use mapreduce.task.id
18/05/01 15:30:46 INFO Configuration.deprecation: mapred.local.dir is deprecated. Instead, use mapreduce.cluster.local.dir
18/05/01 15:30:46 INFO Configuration.deprecation: map.input.file is deprecated. Instead, use mapreduce.map.input.file
18/05/01 15:30:46 INFO Configuration.deprecation: mapred.skip.on is deprecated. Instead, use mapreduce.job.skiprecords
18/05/01 15:30:46 INFO Configuration.deprecation: map.input.length is deprecated. Instead, use mapreduce.map.input.length
18/05/01 15:30:46 INFO Configuration.deprecation: mapred.job.id is deprecated. Instead, use mapreduce.job.id
18/05/01 15:30:46 INFO Configuration.deprecation: mapred.job.id is deprecated. Instead, use mapreduce.job.id
18/05/01 15:30:46 INFO Configuration.deprecation: user.name is deprecated. Instead, use mapreduce.job.user.name
18/05/01 15:30:46 INFO Configuration.deprecation: mapred.task.partition is deprecated. Instead, use mapreduce.task.partition
18/05/01 15:30:46 INFO streaming.PipeMapRed: R/W/S=1/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 15:30:46 INFO streaming.PipeMapRed: R/W/S=10/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 15:30:46 INFO streaming.PipeMapRed: R/W/S=100/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 15:30:46 INFO streaming.PipeMapRed: R/W/S=1000/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 15:30:47 INFO streaming.PipeMapRed: Records R/W=2209/1
18/05/01 15:30:47 INFO streaming.PipeMapRed: R/W/S=10000/7621/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 15:30:47 INFO streaming.PipeMapRed: MRErrorThread done
18/05/01 15:30:47 INFO streaming.PipeMapRed: mapRedFinished
18/05/01 15:30:47 INFO mapred.LocalJobRunner:
18/05/01 15:30:47 INFO mapred.MapTask: Starting flush of map output
18/05/01 15:30:47 INFO mapred.MapTask: Spilling map output
18/05/01 15:30:47 INFO mapred.MapTask: bufstart = 0; bufend = 1018194; bufvoid = 104857600
18/05/01 15:30:47 INFO mapred.MapTask: kvstart = 26214396(104857584); kvend = 26140604(104584256); length = 68333/6553600
18/05/01 15:30:47 INFO mapred.MapTask: Finished spill 0
18/05/01 15:30:47 INFO mapred.Task: Task:attempt_local705825676_0001_m_000000_0 is done. And is in the process of committing
18/05/01 15:30:47 INFO mapred.LocalJobRunner: Records R/W=2209/1
18/05/01 15:30:47 INFO mapred.Task: Task 'attempt_local705825676_0001_m_000000_0' done.
18/05/01 15:30:47 INFO mapred.LocalJobRunner: Finishing task: attempt_local705825676_0001_m_000000_0
18/05/01 15:30:47 INFO mapred.LocalJobRunner: map task executor complete.
18/05/01 15:30:47 INFO mapred.LocalJobRunner: Waiting for reduce tasks
18/05/01 15:30:47 INFO mapred.LocalJobRunner: Starting task: attempt_local705825676_0001_r_000000_0
18/05/01 15:30:47 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 1
18/05/01 15:30:47 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders under output directory:false, ignore cleanup failures: false
18/05/01 15:30:47 INFO mapred.Task: Using ResourceCalculatorProcessTree : [ ]
18/05/01 15:30:47 INFO mapred.ReduceTask: Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle@140619cc
18/05/01 15:30:47 INFO reduce.MergeManagerImpl: MergerManager: memoryLimit=375809632, maxSingleShuffleLimit=93952408, mergeThreshold=248034368, ioSortFactor=10, mergeMemoryMapOutputsThreshold=10
18/05/01 15:30:47 INFO reduce.EventFetcher: attempt_local705825676_0001_r_000000_0 Thread started: EventFetcher for Fetching Map Completion Events
18/05/01 15:30:47 INFO reduce.LocalFetcher: localFetcher#1 about to shuffle output of map attempt_local705825676_0001_m_000000_0 decomp: 1052364 len: 1052368 to MEMORY
18/05/01 15:30:47 INFO reduce.InMemoryMapOutput: Read 1052364 bytes from map-output for attempt_local705825676_0001_m_000000_0
18/05/01 15:30:47 INFO reduce.MergeManagerImpl: closeInMemoryFile -> map-output of size: 1052364, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, usedMemory -> 1052364
18/05/01 15:30:47 INFO reduce.EventFetcher: EventFetcher is interrupted. Returning
18/05/01 15:30:47 INFO mapred.LocalJobRunner: 1 / 1 copied
18/05/01 15:30:47 INFO reduce.MergeManagerImpl: finalMerge called with 1 in-memory map-outputs and 0 on-disk map-outputs
18/05/01 15:30:47 INFO mapred.Merger: Merging 1 sorted segments
18/05/01 15:30:47 INFO mapred.Merger: Down to the last merge-pass, with 1 segments left of total size: 1052311 bytes
18/05/01 15:30:47 INFO mapreduce.Job: Job: job_local705825676_0001 running in uber mode : false
18/05/01 15:30:47 INFO mapreduce.Job: map 100% reduce 0%
18/05/01 15:30:47 INFO reduce.MergeManagerImpl: Merged 1 segments, 1052364 bytes to disk to satisfy reduce memory limit
18/05/01 15:30:47 INFO reduce.MergeManagerImpl: Merging 1 files, 1052368 bytes from disk
18/05/01 15:30:47 INFO reduce.MergeManagerImpl: Merging 0 segments, 0 bytes from memory into reduce
18/05/01 15:30:47 INFO mapred.Merger: Merging 1 sorted segments
18/05/01 15:30:47 INFO mapred.Merger: Down to the last merge-pass, with 1 segments left of total size: 1052311 bytes
18/05/01 15:30:47 INFO mapred.LocalJobRunner: 1 / 1 copied
18/05/01 15:30:47 INFO streaming.PipeMapRed: PipeMapRed exec [/home/hduser/learning_files/assignment/reducer2.py]
```

```

18/05/01 15:30:47 INFO streaming.PipeMapRed: PipeMapRed exec [/home/hduser/learning_files/assignment/reducer2.py]
18/05/01 15:30:47 INFO Configuration.deprecation: mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address
18/05/01 15:30:47 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps
18/05/01 15:30:47 INFO streaming.PipeMapRed: R/W/S=1/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 15:30:47 INFO streaming.PipeMapRed: R/W/S=10/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 15:30:47 INFO streaming.PipeMapRed: R/W/S=100/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 15:30:47 INFO streaming.PipeMapRed: R/W/S=1000/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 15:30:47 INFO streaming.PipeMapRed: R/W/S=10000/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 15:30:47 INFO streaming.PipeMapRed: Records R/W=17084/1
18/05/01 15:30:47 INFO streaming.PipeMapRed: MRErrorThread done
18/05/01 15:30:47 INFO streaming.PipeMapRed: mapRedFinished
18/05/01 15:30:47 INFO mapred.Task: Task:attempt_local705825676_0001_r_000000_0 is done. And is in the process of committing
18/05/01 15:30:47 INFO mapred.LocalJobRunner: 1 / 1 copied
18/05/01 15:30:47 INFO mapred.Task: Task:attempt_local705825676_0001_r_000000_0 is allowed to commit now
18/05/01 15:30:47 INFO output.FileOutputCommitter: Saved output of task 'attempt_local705825676_0001_r_000000_0' to hdfs://localhost:54310/output/output2/_temporary/0/task_local705825676_0001_r_000000
18/05/01 15:30:47 INFO mapred.LocalJobRunner: Records R/W=17084/1 > reduce
18/05/01 15:30:47 INFO mapred.Task: Task:attempt_local705825676_0001_r_000000_0 done
18/05/01 15:30:47 INFO mapred.LocalJobRunner: Finishing task: attempt_local705825676_0001_r_000000_0
18/05/01 15:30:47 INFO mapred.LocalJobRunner: reduce task executor complete.
18/05/01 15:30:48 INFO mapreduce.Job: map 100% reduce 100%
18/05/01 15:30:48 INFO mapreduce.Job: Job:job_local705825676_0001 completed successfully
18/05/01 15:30:48 INFO mapreduce.Job: Counters: 35
  File System Counters
    FILE: Number of bytes read=2375352
    FILE: Number of bytes written=4379788
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=2036388
    HDFS: Number of bytes written=1120698
    HDFS: Number of read operations=13
    HDFS: Number of large read operations=0
    HDFS: Number of write operations=4
  Map-Reduce Framework
    Map input records=17084
    Map output records=17084
    Map output bytes=1018194
    Map output materialized bytes=1052368
    Input split bytes=100
    Combine input records=0
    Combine output records=0
    Reduce input groups=2
    Reduce shuffle bytes=1052368
    Reduce input records=17084
    Reduce output records=17084
    Spilled Records=34168
    Shuffled Maps =1
    Failed Shuffles=0
    Merged Map outputs=1
    GC time elapsed (ms)=36
    Total committed heap usage (bytes)=471859200
  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=1018194
  File Output Format Counters
    Bytes Written=1120698
18/05/01 15:30:48 INFO streaming.StreamJob: Output directory: /output/output2
hduser@jasdeep-MacBookPro: /usr/local/hadoop/hadoop-2.9.0$

```


MapReduce Task3

```
hduser@jasdeep-AsusBookPro: /usr/local/hadoop/hadoop-2.9.0$ bin/hadoop jar /usr/local/hadoop/hadoop-2.9.0/share/hadoop/tools/lib/hadoop-streaming-2.9.0.jar -input /output/output2/part-00000 -output /output/output3 -mapper /home/hduser/learning_files/assignment/mapper3.py --reducer /home/hduser/learning_files/assignment/reducer3.py
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/hadoop-2.9.0/share/hadoop/common/lib/hadoop-auth-2.9.0.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
18/05/01 16:19:49 INFO Configuration.deprecation: session.id is deprecated. Instead, use dfs.metrics.session-id
18/05/01 16:19:49 INFO jvm.JvmMetrics: Initializing JVM Metrics with processName=JobTracker, sessionId=
18/05/01 16:19:49 INFO jvm.JvmMetrics: Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
18/05/01 16:19:50 INFO mapred.FileInputFormat: Total input files to process : 1
18/05/01 16:19:50 INFO mapreduce.JobSubmitter: number of splits:1
18/05/01 16:19:50 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_local208982540_0001
18/05/01 16:19:50 INFO mapreduce.Job: The url to track the job: http://localhost:8080/
18/05/01 16:19:50 INFO mapred.LocalJobRunner: OutputCommitter set in config null
18/05/01 16:19:50 INFO mapreduce.Job: Running job: job_local208982540_0001
18/05/01 16:19:50 INFO mapred.LocalJobRunner: OutputCommitter is org.apache.hadoop.mapred.FileOutputCommitter
18/05/01 16:19:50 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 1
18/05/01 16:19:50 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders under output directory:false, ignore cleanup failures: false
18/05/01 16:19:51 INFO mapred.LocalJobRunner: Waiting for map tasks
18/05/01 16:19:51 INFO mapred.LocalJobRunner: Starting task: attempt_local208982540_0001_m_000000_0
18/05/01 16:19:51 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 1
18/05/01 16:19:51 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders under output directory:false, ignore cleanup failures: false
18/05/01 16:19:51 INFO mapred.Task: Using ResourceCalculatorProcessTree : [ ]
18/05/01 16:19:51 INFO mapred.MapTask: Processing split: hdfs://localhost:54310/output/output2/part-00000_0+1120698
18/05/01 16:19:51 INFO mapred.MapTask: numReduceTasks: 1
18/05/01 16:19:51 INFO mapred.MapTask: (EQUATOR) 0 kvi 26214396(104857584)
18/05/01 16:19:51 INFO mapred.MapTask: mapreduce.task.sort.mb: 100
18/05/01 16:19:51 INFO mapred.MapTask: soft limit at 83886080
18/05/01 16:19:51 INFO mapred.MapTask: bufstart = 0; bufvoid = 104857600
18/05/01 16:19:51 INFO mapred.MapTask: kvstart = 26214396; length = 6553600
18/05/01 16:19:51 INFO mapred.MapTask: Map output collector class = org.apache.hadoop.mapred.MapTask$MapOutputBuffer
18/05/01 16:19:51 INFO streaming.PipeMapRed: PipeMapRed exec [/home/hduser/learning_files/assignment/mapper3.py]
18/05/01 16:19:51 INFO Configuration.deprecation: mapred.work.output.dir is deprecated. Instead, use mapreduce.task.output.dir
18/05/01 16:19:51 INFO Configuration.deprecation: map.input.start is deprecated. Instead, use mapreduce.map.input.start
18/05/01 16:19:51 INFO Configuration.deprecation: mapred.task.is.map is deprecated. Instead, use mapreduce.task.ismap
18/05/01 16:19:51 INFO Configuration.deprecation: mapred.task.id is deprecated. Instead, use mapreduce.task.attempt.id
18/05/01 16:19:51 INFO Configuration.deprecation: mapred.tip.id is deprecated. Instead, use mapreduce.task.id
18/05/01 16:19:51 INFO Configuration.deprecation: mapred.local.dir is deprecated. Instead, use mapreduce.cluster.local.dir
18/05/01 16:19:51 INFO Configuration.deprecation: map.input.file is deprecated. Instead, use mapreduce.map.input.file
18/05/01 16:19:51 INFO Configuration.deprecation: mapred.skip.on is deprecated. Instead, use mapreduce.job.skiprecords
18/05/01 16:19:51 INFO Configuration.deprecation: map.input.length is deprecated. Instead, use mapreduce.map.input.length
18/05/01 16:19:51 INFO Configuration.deprecation: mapred.job.id is deprecated. Instead, use mapreduce.job.id
18/05/01 16:19:51 INFO Configuration.deprecation: user.name is deprecated. Instead, use mapreduce.job.user.name
18/05/01 16:19:51 INFO Configuration.deprecation: mapred.task.partition is deprecated. Instead, use mapreduce.task.partition
18/05/01 16:19:51 INFO streaming.PipeMapRed: R/W/S=1/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 16:19:51 INFO streaming.PipeMapRed: R/W/S=10/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 16:19:51 INFO streaming.PipeMapRed: R/W/S=100/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 16:19:51 INFO streaming.PipeMapRed: R/W/S=1000/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 16:19:51 INFO streaming.PipeMapRed: Records R/W=1957/1
18/05/01 16:19:51 INFO streaming.PipeMapRed: R/W/S=10000/8866/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 16:19:51 INFO streaming.PipeMapRed: MRErrorThread done
18/05/01 16:19:51 INFO streaming.PipeMapRed: mapRedFinished
18/05/01 16:19:51 INFO mapred.LocalJobRunner:
18/05/01 16:19:51 INFO mapred.MapTask: Starting flush of map output
18/05/01 16:19:51 INFO mapred.MapTask: Spilling map output
18/05/01 16:19:51 INFO mapred.MapTask: bufstart = 0; bufend = 1154866; bufvoid = 104857600
18/05/01 16:19:51 INFO mapred.MapTask: kvstart = 26214396(104857584); kvend = 26146064(104584256); length = 68333/6553600
18/05/01 16:19:51 INFO mapred.MapTask: Finished spill 0
18/05/01 16:19:51 INFO mapred.Task: Task: attempt_local208982540_0001_m_000000_0 is done. And is in the process of committing
18/05/01 16:19:51 INFO mapred.LocalJobRunner: Records R/W=1957/1
18/05/01 16:19:51 INFO mapred.Task: Task: attempt_local208982540_0001_m_000000_0' done
18/05/01 16:19:51 INFO mapred.LocalJobRunner: Finishing task: attempt_local208982540_0001_m_000000_0
18/05/01 16:19:51 INFO mapred.LocalJobRunner: map task executor complete
18/05/01 16:19:51 INFO mapred.LocalJobRunner: Waiting for reduce tasks
18/05/01 16:19:51 INFO mapred.LocalJobRunner: Starting task: attempt_local208982540_0001_r_000000_0
18/05/01 16:19:51 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 1
18/05/01 16:19:51 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders under output directory:false, ignore cleanup failures: false
18/05/01 16:19:51 INFO mapred.Task: Using ResourceCalculatorProcessTree : [ ]
18/05/01 16:19:51 INFO mapred.ReduceTask: Using ShuffleConsumerPlugin: org.apache.hadoop.mapreduce.task.reduce.Shuffle$5aa761d7
18/05/01 16:19:51 INFO reduce.MergeManagerImpl: MergerManager: memoryLimit=375809632, maxSingleShuffleLimit=93952408, mergeThreshold=248034368, ioSortFactor=10, memToMemMergeOutputsThreshold=10
18/05/01 16:19:51 INFO reduce.EventFetcher: attempt_local208982540_0001_r_000000_0 Thread started: EventFetcher for fetching Map Completion Events
18/05/01 16:19:51 INFO mapreduce.Job: Job job_local208982540_0001 running in uber mode : false
18/05/01 16:19:51 INFO mapreduce.Job: map 100% reduce 0%
18/05/01 16:19:51 INFO reduce.LocalFetcher: localFetcher#1 about to shuffle output of map attempt_local208982540_0001_m_000000_0 decomp: 1189036 len: 1
189040 to MEMORY
18/05/01 16:19:51 INFO reduce.InMemoryMapOutput: Read 1189036 bytes from map-output for attempt_local208982540_0001_m_000000_0
18/05/01 16:19:51 INFO reduce.MergeManagerImpl: closeInMemoryFile -> map-output of size: 1189036, inMemoryMapOutputs.size() -> 1, commitMemory -> 0, us
edMemory ->1189036
18/05/01 16:19:51 WARN to ReadaheadPool: Failed readahead on tfile
EBADF: Bad file descriptor
    at org.apache.hadoop.mapreduce.TaskIO$NativeIO$POSIX.postixFadvise(Native Method)
    at org.apache.hadoop.mapreduce.TaskIO$NativeIO$POSIX.postixFadviseIfPossible(NativeIO.java:267)
    at org.apache.hadoop.mapreduce.TaskIO$NativeIO$POSIX$CacheManipulator.postixFadviseIfPossible(NativeIO.java:146)
    at org.apache.hadoop.mapreduce.TaskIO$ReadaheadPool$ReadaheadRequestImpl.run(ReadaheadPool.java:208)
    at java.base/java.util.concurrent.ThreadPoolExecutor.runWorker(ThreadPoolExecutor.java:1167)
    at java.base/java.util.concurrent.ThreadPoolExecutor$Worker.run(ThreadPoolExecutor.java:641)
    at java.base/java.lang.Thread.run(Thread.java:844)
18/05/01 16:19:51 INFO reduce.EventFetcher: EventFetcher is interrupted. Returning
18/05/01 16:19:52 INFO mapred.LocalJobRunner: 1 / 1 copied
18/05/01 16:19:52 INFO reduce.MergeManagerImpl: FinalMerge called with 1 in-memory map-outputs and 0 on-disk map-outputs
18/05/01 16:19:52 INFO mapred.Merger: Merging 1 sorted segments
18/05/01 16:19:52 INFO mapred.Merger: Down to the last merge-pass, with 1 segments left of total size: 1189031 bytes
18/05/01 16:19:52 INFO reduce.MergeManagerImpl: Merged 1 segments, 1189036 bytes to disk to satisfy reduce memory limit
18/05/01 16:19:52 INFO reduce.MergeManagerImpl: Merging 1 files, 1189040 bytes from disk
```

```
18/05/01 16:19:52 INFO reduce.MergeManagerImpl: Merging 1 files, 1189040 bytes from disk
18/05/01 16:19:52 INFO reduce.MergeManagerImpl: Merging 0 segments, 0 bytes from memory into reduce
18/05/01 16:19:52 INFO mapred.Merger: Merging 1 sorted segments
18/05/01 16:19:52 INFO mapred.Merger: Down to the last merge-pass, with 1 segments left of total size: 1189031 bytes
18/05/01 16:19:52 INFO mapred.LocalJobRunner: 1 / 1 copied
18/05/01 16:19:52 INFO streaming.PipeMapRed: PipeMapRed exec [/home/hduser/learning_files/assignment/reducer3.py]
18/05/01 16:19:52 INFO Configuration.deprecation: mapred.job.tracker is deprecated. Instead, use mapreduce.job.tracker address
18/05/01 16:19:52 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps
18/05/01 16:19:52 INFO streaming.PipeMapRed: R/W/S=1/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 16:19:52 INFO streaming.PipeMapRed: R/W/S=10/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 16:19:52 INFO streaming.PipeMapRed: R/W/S=100/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 16:19:52 INFO streaming.PipeMapRed: R/W/S=1000/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 16:19:52 INFO streaming.PipeMapRed: R/W/S=10000/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 16:19:52 INFO streaming.PipeMapRed: Records R/W=17084/1
18/05/01 16:20:02 INFO streaming.PipeMapRed: Records R/W=17084/6639
18/05/01 16:20:03 INFO mapred.LocalJobRunner: Records R/W=17084/6639 > reduce
18/05/01 16:20:03 INFO mapreduce.Job: map 100% reduce 100%
18/05/01 16:20:09 INFO mapred.LocalJobRunner: Records R/W=17084/6639 > reduce
18/05/01 16:20:12 INFO streaming.PipeMapRed: Records R/W=17084/13523
18/05/01 16:20:14 INFO streaming.PipeMapRed: MRErrorThread done
18/05/01 16:20:14 INFO streaming.PipeMapRed: mapRedFinished
18/05/01 16:20:14 INFO mapred.Task: Task attempt_local208982540_0001_r_000000_0 is done. And is in the process of committing
18/05/01 16:20:14 INFO mapred.LocalJobRunner: Records R/W=17084/6639 > reduce
18/05/01 16:20:14 INFO mapred.Task: Task attempt_local208982540_0001_r_000000_0 is allowed to commit now
18/05/01 16:20:14 INFO output.FileOutputCommitter: Saved output of task 'attempt_local208982540_0001_r_000000_0' to hdfs://localhost:54310/output/output3/_temporary/0/task_local208982540_0001_r_000000
18/05/01 16:20:14 INFO mapred.LocalJobRunner: Records R/W=17084/13523 > reduce
18/05/01 16:20:14 INFO mapred.Task: Task 'attempt_local208982540_0001_r_000000_0' done.
18/05/01 16:20:14 INFO mapred.LocalJobRunner: Finishing task: attempt_local208982540_0001_r_000000_0
18/05/01 16:20:14 INFO mapred.LocalJobRunner: reduce task executor complete.
18/05/01 16:20:14 INFO mapreduce.Job: Job job_local208982540_0001 completed successfully
18/05/01 16:20:14 INFO mapreduce.Job: Counters: 35
  File System Counters
    FILE: Number of bytes read=2648696
    FILE: Number of bytes written=4789804
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=2241396
    HDFS: Number of bytes written=987165
    HDFS: Number of read operations=13
    HDFS: Number of large read operations=0
    HDFS: Number of write operations=4
  Map-Reduce Framework
    Map input records=17084
    Map output records=17084
    Map output bytes=1154866
    Map output materialized bytes=1189040
    Input split bytes=100
    Combine input records=0
    Combine output records=0
    Combine input bytes=0
    Combine output records=0
    Reduce input groups=14559
    Reduce shuffle bytes=1189040
    Reduce input records=17084
    Reduce output records=14559
    Spilled Records=34168
    Shuffled Maps=1
    Failed Shuffles=0
    Merged Map outputs=1
    GC time elapsed (ms)=26
    Total committed heap usage (bytes)=471859200
  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=1120698
  File Output Format Counters
    Bytes Written=987165
18/05/01 16:20:14 INFO streaming.StreamJob: Output directory: /output/output3
hduser@jasdeep-MacBookPro: /usr/local/hadoop/hadoop-2.9.0$
```

Task #4

```
hduser@jasdeep - MacBookPro: /usr/local/hadoop/hadoop-2.9.0$ bin/hadoop jar /usr/local/hadoop/hadoop-2.9.0/share/hadoop/tools/lib/hadoop-streaming-2.9.0.jar
ar -numReduceTasks 0 -input /output/output3/part-00000 -output /output/output4 -mapper /home/hduser/learning_files/assignment/mapper4.py
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/usr/local/hadoop/hadoop-2.9.0/share/hadoop/conf/lib/hadoop-auth-2.9.0.jar) to method sun.security.krb5.Config.getInstance()
WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
WARNING: All illegal access operations will be denied in a future release
18/05/01 16:37:57 INFO Configuration.deprecation: session id is deprecated. Instead, use dfs.metrics.session-id
18/05/01 16:37:57 INFO jvm.JvmMetrics: Initializing JVM Metrics with processName=JobTracker, sessionId=
18/05/01 16:37:57 INFO jvm.JvmMetrics: Cannot initialize JVM Metrics with processName=JobTracker, sessionId= - already initialized
18/05/01 16:37:57 INFO mapred.FileInputFormat: Total input files to process : 1
18/05/01 16:37:57 INFO mapreduce.JobSubmitter: number of splits:1
18/05/01 16:37:58 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_local288370450_0001
18/05/01 16:37:58 INFO mapreduce.Job: The url to track the job: http://localhost:8080/
18/05/01 16:37:58 INFO mapred.LocalJobRunner: OutputCommitter set in config null
18/05/01 16:37:58 INFO mapreduce.Job: Running job: job_local288370450_0001
18/05/01 16:37:58 INFO mapred.LocalJobRunner: OutputCommitter is org.apache.hadoop.mapred.FileOutputCommitter
18/05/01 16:37:58 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 1
18/05/01 16:37:58 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders under output directory:false, ignore cleanup failures: false
18/05/01 16:37:58 INFO mapred.LocalJobRunner: Waiting for map tasks
18/05/01 16:37:58 INFO mapred.LocalJobRunner: Starting task: attempt_local288370450_0001_m_000000_0
18/05/01 16:37:58 INFO output.FileOutputCommitter: File Output Committer Algorithm version is 1
18/05/01 16:37:58 INFO output.FileOutputCommitter: FileOutputCommitter skip cleanup _temporary folders under output directory:false, ignore cleanup failures: false
18/05/01 16:37:58 INFO mapred.Task: Using ResourceCalculatorProcessTree : [ ]
18/05/01 16:37:58 INFO mapred.MapTask: Processing split: hdfs://localhost:54310/output/output3/part-00000_0+987165
18/05/01 16:37:58 INFO mapred.MapTask: numReduceTasks: 0
18/05/01 16:37:59 INFO streaming.PipeMapRed: PipeMapRed exec [/home/hduser/learning_files/assignment/mapper4.py]
18/05/01 16:37:59 INFO Configuration.deprecation: mapred work output dir is deprecated. Instead, use mapreduce task output.dir
18/05/01 16:37:59 INFO Configuration.deprecation: map input start is deprecated. Instead, use mapreduce task input.start
18/05/01 16:37:59 INFO Configuration.deprecation: mapred task is map is deprecated. Instead, use mapreduce task.ismap
18/05/01 16:37:59 INFO Configuration.deprecation: mapred task id is deprecated. Instead, use mapreduce task.attempt.id
18/05/01 16:37:59 INFO Configuration.deprecation: mapred tip id is deprecated. Instead, use mapreduce task.id
18/05/01 16:37:59 INFO Configuration.deprecation: mapred local.dir is deprecated. Instead, use mapreduce cluster.local.dir
18/05/01 16:37:59 INFO Configuration.deprecation: map input file is deprecated. Instead, use mapreduce map.input.file
18/05/01 16:37:59 INFO Configuration.deprecation: mapred skip on is deprecated. Instead, use mapreduce job.skiprecords
18/05/01 16:37:59 INFO Configuration.deprecation: map input length is deprecated. Instead, use mapreduce map.input.length
18/05/01 16:37:59 INFO Configuration.deprecation: mapred job id is deprecated. Instead, use mapreduce job.id
18/05/01 16:37:59 INFO Configuration.deprecation: user name is deprecated. Instead, use mapreduce job.user.name
18/05/01 16:37:59 INFO Configuration.deprecation: mapred task.partition is deprecated. Instead, use mapreduce task.partition
18/05/01 16:37:59 INFO streaming.PipeMapRed: R/W/S=1/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 16:37:59 INFO streaming.PipeMapRed: R/W/S=10/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 16:37:59 INFO streaming.PipeMapRed: R/W/S=100/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 16:37:59 INFO streaming.PipeMapRed: R/W/S=1000/0/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 16:37:59 INFO streaming.PipeMapRed: Records R/W=1943/1
18/05/01 16:37:59 INFO streaming.PipeMapRed: R/W/S=10000/8549/0 in:NA [rec/s] out:NA [rec/s]
18/05/01 16:37:59 INFO streaming.PipeMapRed: MRErrorThread done
18/05/01 16:37:59 INFO streaming.PipeMapRed: mapredFinished
18/05/01 16:37:59 INFO mapred.LocalJobRunner:
18/05/01 16:37:59 INFO mapred.LocalJobRunner:
18/05/01 16:37:59 INFO mapred.Task: Task: attempt_local288370450_0001_m_000000_0 is done. And is in the process of committing
18/05/01 16:37:59 INFO mapred.LocalJobRunner:
18/05/01 16:37:59 INFO mapred.Task: Task: attempt_local288370450_0001_m_000000_0 is allowed to commit now
18/05/01 16:37:59 INFO output.FileOutputCommitter: Saved output of task 'attempt_local288370450_0001_m_000000_0' to hdfs://localhost:54310/output/output4/_temporary/0/task_local288370450_0001_m_000000
18/05/01 16:37:59 INFO mapred.LocalJobRunner: Records R/W=1943/1
18/05/01 16:37:59 INFO mapred.Task: Task: attempt_local288370450_0001_m_000000_0 done.
18/05/01 16:37:59 INFO mapred.LocalJobRunner: Finishing task: attempt_local288370450_0001_m_000000_0
18/05/01 16:37:59 INFO mapred.LocalJobRunner: map task executor complete.
18/05/01 16:37:59 INFO mapreduce.Job: Job: job_local288370450_0001 running in uber mode : false
18/05/01 16:37:59 INFO mapreduce.Job: map 100% reduce 0%
18/05/01 16:37:59 INFO mapreduce.Job: Job: job_local288370450_0001 completed successfully
18/05/01 16:37:59 INFO mapreduce.Job: Counters: 20
    File System Counters
      FILE: Number of bytes read=135292
      FILE: Number of bytes written=618374
      FILE: Number of read operations=0
      FILE: Number of large read operations=0
      FILE: Number of write operations=0
      HDFS: Number of bytes read=987165
      HDFS: Number of bytes written=1094481
      HDFS: Number of read operations=7
      HDFS: Number of large read operations=0
      HDFS: Number of write operations=3
    Map-Reduce Framework
      Map input records=14559
      Map output records=14559
      Input split bytes=100
      Spilled Records=0
      Failed Shuffles=0
      Merged Map outputs=0
      GC time elapsed (ms)=12
      Total committed heap usage (bytes)=130023424
    File Input Format Counters
      Bytes Read=987165
    File Output Format Counters
      Bytes Written=1094481
18/05/01 16:37:59 INFO streaming.StreamJob: Output directory: /output/output4
hduser@jasdeep - MacBookPro: /usr/local/hadoop/hadoop-2.9.0$
```

