MapReduce Join 수행시간

file 1

17:53:26 - 17:56:22   **2분 56초**

```
2023-11-23 17:53:26,459 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes wh
2023-11-23 17:53:27,308 INFO client.DefaultNoHARMFailoverProxyProvider: Connecting to ResourceManager at /0.0.0.0:8032
2023-11-23 17:53:28,330 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface
2023-11-23 17:53:28,346 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/bokyung/.staging/jo
2023-11-23 17:53:29,718 INFO input.FileInputFormat: Total input files to process : 4
2023-11-23 17:53:30,433 INFO mapreduce.JobSubmitter: number of splits:9
2023-11-23 17:53:31,132 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1700729551664_0001
2023-11-23 17:53:31,134 INFO mapreduce.JobSubmitter: Executing with tokens: []
2023-11-23 17:53:31,431 INFO conf.Configuration: resource-types.xml not found
2023-11-23 17:53:31,432 INFO resource.ResourceUtils: Unable to find 'resource-types.xml'.
2023-11-23 17:53:31,781 INFO impl.YarnClientImpl: Submitted application application_1700729551664_0001
2023-11-23 17:53:31,855 INFO mapreduce.Job: The url to track the job: http://bokyung-MacBookAir.local:8088/proxy/application_1700729551664_
2023-11-23 17:53:31,856 INFO mapreduce.Job: Running job: job_1700729551664_0001
2023-11-23 17:53:52,417 INFO mapreduce.Job: Job job_1700729551664_0001 running in uber mode : false
2023-11-23 17:53:52,425 INFO mapreduce.Job:  map 0% reduce 0%
2023-11-23 17:54:10,880 INFO mapreduce.Job:  map 22% reduce 0%
2023-11-23 17:54:11,904 INFO mapreduce.Job:  map 33% reduce 0%
2023-11-23 17:54:24,095 INFO mapreduce.Job:  map 44% reduce 0%
2023-11-23 17:54:27,150 INFO mapreduce.Job:  map 67% reduce 0%
2023-11-23 17:54:37,375 INFO mapreduce.Job:  map 89% reduce 0%
2023-11-23 17:54:40,415 INFO mapreduce.Job:  map 100% reduce 0%
2023-11-23 17:54:54,640 INFO mapreduce.Job:  map 100% reduce 71%
2023-11-23 17:55:00,722 INFO mapreduce.Job:  map 100% reduce 76%
2023-11-23 17:55:06,808 INFO mapreduce.Job:  map 100% reduce 82%
2023-11-23 17:55:12,900 INFO mapreduce.Job:  map 100% reduce 87%
2023-11-23 17:55:18,991 INFO mapreduce.Job:  map 100% reduce 93%
2023-11-23 17:55:25,079 INFO mapreduce.Job:  map 100% reduce 98%
2023-11-23 17:55:27,151 INFO mapreduce.Job:  map 100% reduce 100%
2023-11-23 17:55:28,257 INFO mapreduce.Job: Job job_1700729551664_0001 completed successfully
2023-11-23 17:55:28,499 INFO mapreduce.Job: Counters: 51
  File System Counters
    FILE: Number of bytes read=1703860708
    FILE: Number of bytes written=2647434405
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=822811536
    HDFS: Number of bytes written=420394832
    HDFS: Number of read operations=32
    HDFS: Number of large read operations=0
    HDFS: Number of write operations=2
    HDFS: Number of bytes read erasure-coded=0
  Job Counters
    Killed map tasks=1
    Launched map tasks=9
    Launched reduce tasks=1
    Data-local map tasks=9
    Total time spent by all maps in occupied slots (ms)=250782
    Total time spent by all reduces in occupied slots (ms)=194492
    Total time spent by all map tasks (ms)=125391
    Total time spent by all reduce tasks (ms)=48623
    Total vcore-milliseconds taken by all map tasks=125391
    Total vcore-milliseconds taken by all reduce tasks=48623
    Total megabyte-milliseconds taken by all map tasks=256800768
    Total megabyte-milliseconds taken by all reduce tasks=199159808
  Map-Reduce Framework
    Map input records=6000004
    Map output records=5999995
    Map output bytes=924788613
    Map output materialized bytes=940789062
    Input split bytes=1053
    Combine input records=0
    Combine output records=0
    Reduce input groups=3999997
    Reduce shuffle bytes=940789062
    Reduce input records=5999995
    Reduce output records=1999998
```

```
      Spilled Records=15537809
      Shuffled Maps =9
      Failed Shuffles=0
      Merged Map outputs=9
      GC time elapsed (ms)=2543
      CPU time spent (ms)=0
      Physical memory (bytes) snapshot=0
      Virtual memory (bytes) snapshot=0
      Total committed heap usage (bytes)=9906946048
    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=822810483
    File Output Format Counters
      Bytes Written=420394832
2023-11-23 17:55:28,550 INFO client.DefaultNoHARMFailoverProxyProvider: Connecting to ResourceManager at /0.0.0.0:8032
2023-11-23 17:55:28,608 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface
2023-11-23 17:55:28,632 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/bokyung/.staging/jo
2023-11-23 17:55:29,418 INFO input.FileInputFormat: Total input files to process : 1
2023-11-23 17:55:29,930 INFO mapreduce.JobSubmitter: number of splits:4
2023-11-23 17:55:30,440 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1700729551664_0002
2023-11-23 17:55:30,441 INFO mapreduce.JobSubmitter: Executing with tokens: []
2023-11-23 17:55:30,576 INFO impl.YarnClientImpl: Submitted application application_1700729551664_0002
2023-11-23 17:55:30,605 INFO mapreduce.Job: The url to track the job: http://bokyung-MacBookAir.local:8088/proxy/application_1700729551664_
2023-11-23 17:55:30,606 INFO mapreduce.Job: Running job: job_1700729551664_0002
2023-11-23 17:55:48,107 INFO mapreduce.Job: Job job_1700729551664_0002 running in uber mode : false
2023-11-23 17:55:48,111 INFO mapreduce.Job:  map 0% reduce 0%
2023-11-23 17:56:04,456 INFO mapreduce.Job:  map 75% reduce 0%
2023-11-23 17:56:11,576 INFO mapreduce.Job:  map 100% reduce 0%
2023-11-23 17:56:21,719 INFO mapreduce.Job:  map 100% reduce 100%
2023-11-23 17:56:22,767 INFO mapreduce.Job: Job job_1700729551664_0002 completed successfully
2023-11-23 17:56:22,821 INFO mapreduce.Job: Counters: 50
  File System Counters
    FILE: Number of bytes read=838710915
    FILE: Number of bytes written=1268498523
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=420407620
    HDFS: Number of bytes written=387394866
    HDFS: Number of read operations=17
    HDFS: Number of large read operations=0
    HDFS: Number of write operations=2
    HDFS: Number of bytes read erasure-coded=0
  Job Counters
    Launched map tasks=4
    Launched reduce tasks=1
    Data-local map tasks=4
    Total time spent by all maps in occupied slots (ms)=94076
    Total time spent by all reduces in occupied slots (ms)=55700
    Total time spent by all map tasks (ms)=47038
    Total time spent by all reduce tasks (ms)=13925
    Total vcore-milliseconds taken by all map tasks=47038
    Total vcore-milliseconds taken by all reduce tasks=13925
    Total megabyte-milliseconds taken by all map tasks=96333824
    Total megabyte-milliseconds taken by all reduce tasks=57036800
  Map-Reduce Framework
    Map input records=1999998
    Map output records=1999998
    Map output bytes=422394906
    Map output materialized bytes=428395108
    Input split bytes=500
    Combine input records=0
    Combine output records=0
    Reduce input groups=1000000
    Reduce shuffle bytes=428395108
    Reduce input records=1999998
    Reduce output records=1000000
    Spilled Records=5915532
    Shuffled Maps =4
    Failed Shuffles=0
    Merged Map outputs=4
    GC time elapsed (ms)=1051
```

```
      CPU time spent (ms)=0
      Physical memory (bytes) snapshot=0
      Virtual memory (bytes) snapshot=0
      Total committed heap usage (bytes)=4895801344
    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=420407120
    File Output Format Counters
      Bytes Written=387394866
```

file 2

18:14:31 - 18:19:10 **4분 39초**

```
2023-11-23 18:14:31,117 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes wh
2023-11-23 18:14:31,999 INFO client.DefaultNoHARMFailoverProxyProvider: Connecting to ResourceManager at /0.0.0.0:8032
2023-11-23 18:14:33,045 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface
2023-11-23 18:14:33,081 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/bokyung/.staging/jo
2023-11-23 18:14:34,267 INFO input.FileInputFormat: Total input files to process : 4
2023-11-23 18:14:34,930 INFO mapreduce.JobSubmitter: number of splits:14
2023-11-23 18:14:35,688 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1700729551664_0005
2023-11-23 18:14:35,691 INFO mapreduce.JobSubmitter: Executing with tokens: []
2023-11-23 18:14:35,991 INFO conf.Configuration: resource-types.xml not found
2023-11-23 18:14:35,991 INFO resource.ResourceUtils: Unable to find 'resource-types.xml'.
2023-11-23 18:14:36,120 INFO impl.YarnClientImpl: Submitted application application_1700729551664_0005
2023-11-23 18:14:36,174 INFO mapreduce.Job: The url to track the job: http://bokyung-MacBookAir.local:8088/proxy/application_1700729551664_
2023-11-23 18:14:36,175 INFO mapreduce.Job: Running job: job_1700729551664_0005
2023-11-23 18:14:53,587 INFO mapreduce.Job: Job job_1700729551664_0005 running in uber mode : false
2023-11-23 18:14:53,594 INFO mapreduce.Job:  map 0% reduce 0%
2023-11-23 18:15:14,093 INFO mapreduce.Job:  map 21% reduce 0%
2023-11-23 18:15:30,446 INFO mapreduce.Job:  map 29% reduce 0%
2023-11-23 18:15:32,482 INFO mapreduce.Job:  map 43% reduce 0%
2023-11-23 18:15:47,771 INFO mapreduce.Job:  map 50% reduce 0%
2023-11-23 18:15:48,796 INFO mapreduce.Job:  map 64% reduce 0%
2023-11-23 18:16:03,122 INFO mapreduce.Job:  map 71% reduce 0%
2023-11-23 18:16:04,150 INFO mapreduce.Job:  map 79% reduce 0%
2023-11-23 18:16:08,223 INFO mapreduce.Job:  map 85% reduce 0%
2023-11-23 18:16:09,253 INFO mapreduce.Job:  map 86% reduce 0%
2023-11-23 18:16:15,335 INFO mapreduce.Job:  map 93% reduce 0%
2023-11-23 18:16:21,445 INFO mapreduce.Job:  map 100% reduce 0%
2023-11-23 18:16:32,601 INFO mapreduce.Job:  map 100% reduce 67%
2023-11-23 18:16:38,724 INFO mapreduce.Job:  map 100% reduce 69%
2023-11-23 18:16:43,808 INFO mapreduce.Job:  map 100% reduce 72%
2023-11-23 18:16:49,924 INFO mapreduce.Job:  map 100% reduce 74%
2023-11-23 18:16:56,004 INFO mapreduce.Job:  map 100% reduce 77%
2023-11-23 18:17:02,073 INFO mapreduce.Job:  map 100% reduce 79%
2023-11-23 18:17:08,159 INFO mapreduce.Job:  map 100% reduce 82%
2023-11-23 18:17:14,208 INFO mapreduce.Job:  map 100% reduce 85%
2023-11-23 18:17:20,304 INFO mapreduce.Job:  map 100% reduce 87%
2023-11-23 18:17:26,361 INFO mapreduce.Job:  map 100% reduce 90%
2023-11-23 18:17:32,414 INFO mapreduce.Job:  map 100% reduce 92%
2023-11-23 18:17:38,477 INFO mapreduce.Job:  map 100% reduce 95%
2023-11-23 18:17:44,545 INFO mapreduce.Job:  map 100% reduce 98%
2023-11-23 18:17:50,637 INFO mapreduce.Job:  map 100% reduce 100%
2023-11-23 18:17:51,726 INFO mapreduce.Job: Job job_1700729551664_0005 completed successfully
2023-11-23 18:17:51,923 INFO mapreduce.Job: Counters: 51
  File System Counters
    FILE: Number of bytes read=3697090842
    FILE: Number of bytes written=5582811700
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=1645585414
    HDFS: Number of bytes written=840791066
```

```
      HDFS: Number of read operations=47
      HDFS: Number of large read operations=0
      HDFS: Number of write operations=2
      HDFS: Number of bytes read erasure-coded=0
    Job Counters
      Killed map tasks=1
      Launched map tasks=14
      Launched reduce tasks=1
      Data-local map tasks=14
      Total time spent by all maps in occupied slots (ms)=454816
      Total time spent by all reduces in occupied slots (ms)=379476
      Total time spent by all map tasks (ms)=227408
      Total time spent by all reduce tasks (ms)=94869
      Total vcore-milliseconds taken by all map tasks=227408
      Total vcore-milliseconds taken by all reduce tasks=94869
      Total megabyte-milliseconds taken by all map tasks=465731584
      Total megabyte-milliseconds taken by all reduce tasks=388583424
    Map-Reduce Framework
      Map input records=12000004
      Map output records=12000004
      Map output bytes=1849542982
      Map output materialized bytes=1881543859
      Input split bytes=1638
      Combine input records=0
      Combine output records=0
      Reduce input groups=7999996
      Reduce shuffle bytes=1881543859
      Reduce input records=12000004
      Reduce output records=4000007
      Spilled Records=35727341
      Shuffled Maps =14
      Failed Shuffles=0
      Merged Map outputs=14
      GC time elapsed (ms)=4220
      CPU time spent (ms)=0
      Physical memory (bytes) snapshot=0
      Virtual memory (bytes) snapshot=0
      Total committed heap usage (bytes)=14933295104
    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=1645583776
    File Output Format Counters
      Bytes Written=840791066
2023-11-23 18:17:51,981 INFO client.DefaultNoHARMFailoverProxyProvider: Connecting to ResourceManager at /0.0.0.0:8032
2023-11-23 18:17:52,129 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface
2023-11-23 18:17:52,148 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/bokyung/.staging/jo
2023-11-23 18:17:53,099 INFO input.FileInputFormat: Total input files to process : 1
2023-11-23 18:17:54,107 INFO mapreduce.JobSubmitter: number of splits:7
2023-11-23 18:17:54,246 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1700729551664_0006
2023-11-23 18:17:54,247 INFO mapreduce.JobSubmitter: Executing with tokens: []
2023-11-23 18:17:54,559 INFO impl.YarnClientImpl: Submitted application application_1700729551664_0006
2023-11-23 18:17:54,600 INFO mapreduce.Job: The url to track the job: http://bokyung-MacBookAir.local:8088/proxy/application_1700729551664_
2023-11-23 18:17:54,600 INFO mapreduce.Job: Running job: job_1700729551664_0006
2023-11-23 18:18:12,027 INFO mapreduce.Job: Job job_1700729551664_0006 running in uber mode : false
2023-11-23 18:18:12,035 INFO mapreduce.Job:  map 0% reduce 0%
2023-11-23 18:18:29,311 INFO mapreduce.Job:  map 43% reduce 0%
2023-11-23 18:18:43,660 INFO mapreduce.Job:  map 86% reduce 0%
2023-11-23 18:18:51,772 INFO mapreduce.Job:  map 100% reduce 0%
2023-11-23 18:19:03,899 INFO mapreduce.Job:  map 100% reduce 87%
2023-11-23 18:19:08,980 INFO mapreduce.Job:  map 100% reduce 100%
2023-11-23 18:19:10,025 INFO mapreduce.Job: Job job_1700729551664_0006 completed successfully
2023-11-23 18:19:10,095 INFO mapreduce.Job: Counters: 50
    File System Counters
      FILE: Number of bytes read=1677423503
      FILE: Number of bytes written=2536443244
      FILE: Number of read operations=0
      FILE: Number of large read operations=0
      FILE: Number of write operations=0
      HDFS: Number of bytes read=840816517
      HDFS: Number of bytes written=775790921
      HDFS: Number of read operations=26
      HDFS: Number of large read operations=0
```

```
   HDFS: Number of write operations=2
   HDFS: Number of bytes read erasure-coded=0
 Job Counters
   Launched map tasks=7
   Launched reduce tasks=1
   Data-local map tasks=7
   Total time spent by all maps in occupied slots (ms)=185228
   Total time spent by all reduces in occupied slots (ms)=86164
   Total time spent by all map tasks (ms)=92614
   Total time spent by all reduce tasks (ms)=21541
   Total vcore-milliseconds taken by all map tasks=92614
   Total vcore-milliseconds taken by all reduce tasks=21541
   Total megabyte-milliseconds taken by all map tasks=189673472
   Total megabyte-milliseconds taken by all reduce tasks=88231936
 Map-Reduce Framework
   Map input records=4000007
   Map output records=4000007
   Map output bytes=844791235
   Map output materialized bytes=856791708
   Input split bytes=875
   Combine input records=0
   Combine output records=0
   Reduce input groups=2999998
   Reduce shuffle bytes=856791708
   Reduce input records=4000007
   Reduce output records=2999998
   Spilled Records=11831182
   Shuffled Maps =7
   Failed Shuffles=0
   Merged Map outputs=7
   GC time elapsed (ms)=1886
   CPU time spent (ms)=0
   Physical memory (bytes) snapshot=0
   Virtual memory (bytes) snapshot=0
   Total committed heap usage (bytes)=8331460608
 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=840815642
 File Output Format Counters
   Bytes Written=775790921
```

file3

18:34:41 - 18:38:42 **4분 1초**

```
2023-11-23 18:34:41,215 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes wh
2023-11-23 18:34:41,759 INFO client.DefaultNoHARMFailoverProxyProvider: Connecting to ResourceManager at /0.0.0.0:8032
2023-11-23 18:34:42,428 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface
2023-11-23 18:34:42,438 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/bokyung/.staging/jo
2023-11-23 18:34:43,341 INFO input.FileInputFormat: Total input files to process : 4
2023-11-23 18:34:44,428 INFO mapreduce.JobSubmitter: number of splits:20
2023-11-23 18:34:44,655 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1700732034480_0001
2023-11-23 18:34:44,655 INFO mapreduce.JobSubmitter: Executing with tokens: []
2023-11-23 18:34:44,843 INFO conf.Configuration: resource-types.xml not found
2023-11-23 18:34:44,843 INFO resource.ResourceUtils: Unable to find 'resource-types.xml'.
2023-11-23 18:34:45,081 INFO impl.YarnClientImpl: Submitted application application_1700732034480_0001
2023-11-23 18:34:45,147 INFO mapreduce.Job: The url to track the job: http://bokyung-MacBookAir.local:8088/proxy/application_1700732034480_
2023-11-23 18:34:45,149 INFO mapreduce.Job: Running job: job_1700732034480_0001
2023-11-23 18:34:59,565 INFO mapreduce.Job: Job job_1700732034480_0001 running in uber mode : false
2023-11-23 18:34:59,575 INFO mapreduce.Job:  map 0% reduce 0%
2023-11-23 18:35:11,851 INFO mapreduce.Job:  map 15% reduce 0%
2023-11-23 18:35:20,991 INFO mapreduce.Job:  map 30% reduce 0%
2023-11-23 18:35:31,156 INFO mapreduce.Job:  map 45% reduce 0%
2023-11-23 18:35:41,302 INFO mapreduce.Job:  map 60% reduce 0%
```

```
2023-11-23 18:35:52,542 INFO mapreduce.Job:  map 75% reduce 0%
2023-11-23 18:36:01,732 INFO mapreduce.Job:  map 80% reduce 0%
2023-11-23 18:36:06,794 INFO mapreduce.Job:  map 85% reduce 0%
2023-11-23 18:36:07,823 INFO mapreduce.Job:  map 90% reduce 0%
2023-11-23 18:36:09,850 INFO mapreduce.Job:  map 95% reduce 0%
2023-11-23 18:36:11,894 INFO mapreduce.Job:  map 100% reduce 0%
2023-11-23 18:36:26,116 INFO mapreduce.Job:  map 100% reduce 67%
2023-11-23 18:36:32,184 INFO mapreduce.Job:  map 100% reduce 70%
2023-11-23 18:36:38,255 INFO mapreduce.Job:  map 100% reduce 72%
2023-11-23 18:36:44,317 INFO mapreduce.Job:  map 100% reduce 75%
2023-11-23 18:36:50,377 INFO mapreduce.Job:  map 100% reduce 78%
2023-11-23 18:36:56,485 INFO mapreduce.Job:  map 100% reduce 81%
2023-11-23 18:37:02,553 INFO mapreduce.Job:  map 100% reduce 84%
2023-11-23 18:37:08,626 INFO mapreduce.Job:  map 100% reduce 87%
2023-11-23 18:37:14,681 INFO mapreduce.Job:  map 100% reduce 90%
2023-11-23 18:37:20,730 INFO mapreduce.Job:  map 100% reduce 93%
2023-11-23 18:37:26,784 INFO mapreduce.Job:  map 100% reduce 96%
2023-11-23 18:37:32,848 INFO mapreduce.Job:  map 100% reduce 99%
2023-11-23 18:37:35,920 INFO mapreduce.Job:  map 100% reduce 100%
2023-11-23 18:37:36,994 INFO mapreduce.Job: Job job_1700732034480_0001 completed successfully
2023-11-23 18:37:37,166 INFO mapreduce.Job: Counters: 51
  File System Counters
    FILE: Number of bytes read=5618496397
    FILE: Number of bytes written=8446646968
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=2468369045
    HDFS: Number of bytes written=1261177842
    HDFS: Number of read operations=65
    HDFS: Number of large read operations=0
    HDFS: Number of write operations=2
    HDFS: Number of bytes read erasure-coded=0
  Job Counters
    Killed map tasks=1
    Launched map tasks=20
    Launched reduce tasks=1
    Data-local map tasks=20
    Total time spent by all maps in occupied slots (ms)=381442
    Total time spent by all reduces in occupied slots (ms)=336748
    Total time spent by all map tasks (ms)=190721
    Total time spent by all reduce tasks (ms)=84187
    Total vcore-milliseconds taken by all map tasks=190721
    Total vcore-milliseconds taken by all reduce tasks=84187
    Total megabyte-milliseconds taken by all map tasks=390596608
    Total megabyte-milliseconds taken by all reduce tasks=344829952
  Map-Reduce Framework
    Map input records=18000004
    Map output records=18000004
    Map output bytes=2774301422
    Map output materialized bytes=2822302746
    Input split bytes=2340
    Combine input records=0
    Combine output records=0
    Reduce input groups=11999992
    Reduce shuffle bytes=2822302746
    Reduce input records=18000004
    Reduce output records=6000011
    Spilled Records=53859887
    Shuffled Maps =20
    Failed Shuffles=0
    Merged Map outputs=20
    GC time elapsed (ms)=3922
    CPU time spent (ms)=0
    Physical memory (bytes) snapshot=0
    Virtual memory (bytes) snapshot=0
    Total committed heap usage (bytes)=20130037760
  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=2468366705
  File Output Format Counters
    Bytes Written=1261177842
```

```
2023-11-23 18:37:37,201 INFO client.DefaultNoHARMFailoverProxyProvider: Connecting to ResourceManager at /0.0.0.0:8032
2023-11-23 18:37:37,323 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface
2023-11-23 18:37:37,342 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/bokyung/.staging/jo
2023-11-23 18:37:38,355 INFO input.FileInputFormat: Total input files to process : 1
2023-11-23 18:37:39,297 INFO mapreduce.JobSubmitter: number of splits:10
2023-11-23 18:37:39,818 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1700732034480_0002
2023-11-23 18:37:39,819 INFO mapreduce.JobSubmitter: Executing with tokens: []
2023-11-23 18:37:39,922 INFO impl.YarnClientImpl: Submitted application application_1700732034480_0002
2023-11-23 18:37:39,942 INFO mapreduce.Job: The url to track the job: http://bokyung-MacBookAir.local:8088/proxy/application_1700732034480_0
2023-11-23 18:37:39,942 INFO mapreduce.Job: Running job: job_1700732034480_0002
2023-11-23 18:37:52,251 INFO mapreduce.Job: Job job_1700732034480_0002 running in uber mode : false
2023-11-23 18:37:52,259 INFO mapreduce.Job:  map 0% reduce 0%
2023-11-23 18:38:02,472 INFO mapreduce.Job:  map 30% reduce 0%
2023-11-23 18:38:10,639 INFO mapreduce.Job:  map 40% reduce 0%
2023-11-23 18:38:11,656 INFO mapreduce.Job:  map 60% reduce 0%
2023-11-23 18:38:19,798 INFO mapreduce.Job:  map 80% reduce 0%
2023-11-23 18:38:20,826 INFO mapreduce.Job:  map 90% reduce 0%
2023-11-23 18:38:24,889 INFO mapreduce.Job:  map 100% reduce 0%
2023-11-23 18:38:36,998 INFO mapreduce.Job:  map 100% reduce 90%
2023-11-23 18:38:40,054 INFO mapreduce.Job:  map 100% reduce 100%
2023-11-23 18:38:42,187 INFO mapreduce.Job: Job job_1700732034480_0002 completed successfully
2023-11-23 18:38:42,249 INFO mapreduce.Job: Counters: 51
  File System Counters
    FILE: Number of bytes read=2516126559
    FILE: Number of bytes written=3804368893
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=1261215956
    HDFS: Number of bytes written=1164177615
    HDFS: Number of read operations=35
    HDFS: Number of large read operations=0
    HDFS: Number of write operations=2
    HDFS: Number of bytes read erasure-coded=0
  Job Counters
    Killed map tasks=1
    Launched map tasks=10
    Launched reduce tasks=1
    Data-local map tasks=10
    Total time spent by all maps in occupied slots (ms)=146088
    Total time spent by all reduces in occupied slots (ms)=72560
    Total time spent by all map tasks (ms)=73044
    Total time spent by all reduce tasks (ms)=18140
    Total vcore-milliseconds taken by all map tasks=73044
    Total vcore-milliseconds taken by all reduce tasks=18140
    Total megabyte-milliseconds taken by all map tasks=149594112
    Total megabyte-milliseconds taken by all reduce tasks=74301440
  Map-Reduce Framework
    Map input records=6000011
    Map output records=6000011
    Map output bytes=1267178080
    Map output materialized bytes=1285178757
    Input split bytes=1250
    Combine input records=0
    Combine output records=0
    Reduce input groups=4999996
    Reduce shuffle bytes=1285178757
    Reduce input records=6000011
    Reduce output records=4999996
    Spilled Records=17746833
    Shuffled Maps =10
    Failed Shuffles=0
    Merged Map outputs=10
    GC time elapsed (ms)=1535
    CPU time spent (ms)=0
    Physical memory (bytes) snapshot=0
    Virtual memory (bytes) snapshot=0
    Total committed heap usage (bytes)=11370758144
  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=1261214706
```

```
    File Output Format Counters
      Bytes Written=1164177615
```

file 4

20:19:53 - 20:32:08 **12분 15초**

```
2023-11-23 20:19:53,966 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes wh
2023-11-23 20:19:54,827 INFO client.DefaultNoHARMFailoverProxyProvider: Connecting to ResourceManager at /0.0.0.0:8032
2023-11-23 20:19:55,905 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface
2023-11-23 20:19:55,969 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/bokyung/.staging/jo
2023-11-23 20:19:57,549 INFO input.FileInputFormat: Total input files to process : 4
2023-11-23 20:19:58,770 INFO mapreduce.JobSubmitter: number of splits:25
2023-11-23 20:19:59,088 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1700738312419_0001
2023-11-23 20:19:59,089 INFO mapreduce.JobSubmitter: Executing with tokens: []
2023-11-23 20:19:59,376 INFO conf.Configuration: resource-types.xml not found
2023-11-23 20:19:59,377 INFO resource.ResourceUtils: Unable to find 'resource-types.xml'.
2023-11-23 20:19:59,899 INFO impl.YarnClientImpl: Submitted application application_1700738312419_0001
2023-11-23 20:19:59,952 INFO mapreduce.Job: The url to track the job: http://bokyung-MacBookAir.local:8088/proxy/application_1700738312419_
2023-11-23 20:19:59,953 INFO mapreduce.Job: Running job: job_1700738312419_0001
2023-11-23 20:20:20,520 INFO mapreduce.Job: Job job_1700738312419_0001 running in uber mode : false
2023-11-23 20:20:20,525 INFO mapreduce.Job:  map 0% reduce 0%
2023-11-23 20:20:39,953 INFO mapreduce.Job:  map 2% reduce 0%
2023-11-23 20:20:45,036 INFO mapreduce.Job:  map 4% reduce 0%
2023-11-23 20:21:03,245 INFO mapreduce.Job:  map 6% reduce 0%
2023-11-23 20:21:09,343 INFO mapreduce.Job:  map 7% reduce 0%
2023-11-23 20:21:11,421 INFO mapreduce.Job:  map 8% reduce 0%
2023-11-23 20:21:23,630 INFO mapreduce.Job:  map 12% reduce 0%
2023-11-23 20:21:35,807 INFO mapreduce.Job:  map 16% reduce 0%
2023-11-23 20:21:48,028 INFO mapreduce.Job:  map 20% reduce 0%
2023-11-23 20:22:00,244 INFO mapreduce.Job:  map 24% reduce 0%
2023-11-23 20:22:12,417 INFO mapreduce.Job:  map 28% reduce 0%
2023-11-23 20:22:24,583 INFO mapreduce.Job:  map 32% reduce 0%
2023-11-23 20:22:36,740 INFO mapreduce.Job:  map 36% reduce 0%
2023-11-23 20:22:48,948 INFO mapreduce.Job:  map 40% reduce 0%
2023-11-23 20:23:01,135 INFO mapreduce.Job:  map 44% reduce 0%
2023-11-23 20:23:13,260 INFO mapreduce.Job:  map 48% reduce 0%
2023-11-23 20:23:25,447 INFO mapreduce.Job:  map 52% reduce 0%
2023-11-23 20:23:37,681 INFO mapreduce.Job:  map 56% reduce 0%
2023-11-23 20:23:49,871 INFO mapreduce.Job:  map 60% reduce 0%
2023-11-23 20:24:03,060 INFO mapreduce.Job:  map 64% reduce 0%
2023-11-23 20:24:15,212 INFO mapreduce.Job:  map 68% reduce 0%
2023-11-23 20:24:27,369 INFO mapreduce.Job:  map 72% reduce 0%
2023-11-23 20:24:39,529 INFO mapreduce.Job:  map 76% reduce 0%
2023-11-23 20:24:50,648 INFO mapreduce.Job:  map 80% reduce 0%
2023-11-23 20:25:02,784 INFO mapreduce.Job:  map 84% reduce 0%
2023-11-23 20:25:14,901 INFO mapreduce.Job:  map 88% reduce 0%
2023-11-23 20:25:27,023 INFO mapreduce.Job:  map 92% reduce 0%
2023-11-23 20:25:37,127 INFO mapreduce.Job:  map 96% reduce 0%
2023-11-23 20:25:47,285 INFO mapreduce.Job:  map 100% reduce 0%
2023-11-23 20:26:04,446 INFO mapreduce.Job:  map 100% reduce 19%
2023-11-23 20:26:10,489 INFO mapreduce.Job:  map 100% reduce 32%
2023-11-23 20:26:16,545 INFO mapreduce.Job:  map 100% reduce 67%
2023-11-23 20:26:22,584 INFO mapreduce.Job:  map 100% reduce 68%
2023-11-23 20:26:28,626 INFO mapreduce.Job:  map 100% reduce 69%
2023-11-23 20:26:34,692 INFO mapreduce.Job:  map 100% reduce 71%
2023-11-23 20:26:40,772 INFO mapreduce.Job:  map 100% reduce 72%
2023-11-23 20:26:46,816 INFO mapreduce.Job:  map 100% reduce 73%
2023-11-23 20:26:52,907 INFO mapreduce.Job:  map 100% reduce 74%
2023-11-23 20:26:58,952 INFO mapreduce.Job:  map 100% reduce 76%
2023-11-23 20:27:03,995 INFO mapreduce.Job:  map 100% reduce 77%
2023-11-23 20:27:10,060 INFO mapreduce.Job:  map 100% reduce 78%
2023-11-23 20:27:16,095 INFO mapreduce.Job:  map 100% reduce 79%
2023-11-23 20:27:22,134 INFO mapreduce.Job:  map 100% reduce 81%
2023-11-23 20:27:28,184 INFO mapreduce.Job:  map 100% reduce 82%
2023-11-23 20:27:34,234 INFO mapreduce.Job:  map 100% reduce 83%
2023-11-23 20:27:40,295 INFO mapreduce.Job:  map 100% reduce 85%
2023-11-23 20:27:46,365 INFO mapreduce.Job:  map 100% reduce 86%
2023-11-23 20:27:52,412 INFO mapreduce.Job:  map 100% reduce 87%
2023-11-23 20:27:58,480 INFO mapreduce.Job:  map 100% reduce 88%
```

```
2023-11-23 20:28:04,521 INFO mapreduce.Job:  map 100% reduce 90%
2023-11-23 20:28:10,569 INFO mapreduce.Job:  map 100% reduce 91%
2023-11-23 20:28:16,623 INFO mapreduce.Job:  map 100% reduce 92%
2023-11-23 20:28:22,680 INFO mapreduce.Job:  map 100% reduce 94%
2023-11-23 20:28:28,727 INFO mapreduce.Job:  map 100% reduce 95%
2023-11-23 20:28:34,770 INFO mapreduce.Job:  map 100% reduce 96%
2023-11-23 20:28:40,811 INFO mapreduce.Job:  map 100% reduce 97%
2023-11-23 20:28:46,862 INFO mapreduce.Job:  map 100% reduce 99%
2023-11-23 20:28:52,928 INFO mapreduce.Job:  map 100% reduce 100%
2023-11-23 20:28:56,060 INFO mapreduce.Job: Job job_1700738312419_0001 completed successfully
2023-11-23 20:28:56,243 INFO mapreduce.Job: Counters: 51
  File System Counters
    FILE: Number of bytes read=7526136756
    FILE: Number of bytes written=11296445351
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=3291159957
    HDFS: Number of bytes written=1681540214
    HDFS: Number of read operations=80
    HDFS: Number of large read operations=0
    HDFS: Number of write operations=2
    HDFS: Number of bytes read erasure-coded=0
  Job Counters
    Killed map tasks=1
    Launched map tasks=25
    Launched reduce tasks=1
    Data-local map tasks=25
    Total time spent by all maps in occupied slots (ms)=1203320
    Total time spent by all reduces in occupied slots (ms)=741712
    Total time spent by all map tasks (ms)=300830
    Total time spent by all reduce tasks (ms)=185428
    Total vcore-milliseconds taken by all map tasks=300830
    Total vcore-milliseconds taken by all reduce tasks=185428
    Total megabyte-milliseconds taken by all map tasks=1232199680
    Total megabyte-milliseconds taken by all reduce tasks=759513088
  Map-Reduce Framework
    Map input records=24000004
    Map output records=23999979
    Map output bytes=3699066676
    Map output materialized bytes=3763068363
    Input split bytes=2925
    Combine input records=0
    Combine output records=0
    Reduce input groups=15999989
    Reduce shuffle bytes=3763068363
    Reduce input records=23999979
    Reduce output records=7999990
    Spilled Records=71999937
    Shuffled Maps =25
    Failed Shuffles=0
    Merged Map outputs=25
    GC time elapsed (ms)=5243
    CPU time spent (ms)=0
    Physical memory (bytes) snapshot=0
    Virtual memory (bytes) snapshot=0
    Total committed heap usage (bytes)=29441916928
  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=3291157032
  File Output Format Counters
    Bytes Written=1681540214
2023-11-23 20:28:56,315 INFO client.DefaultNoHARMFailoverProxyProvider: Connecting to ResourceManager at /0.0.0.0:8032
2023-11-23 20:28:56,683 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface
2023-11-23 20:28:56,698 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/bokyung/.staging/jo
2023-11-23 20:28:58,271 INFO input.FileInputFormat: Total input files to process : 1
2023-11-23 20:28:59,238 INFO mapreduce.JobSubmitter: number of splits:13
2023-11-23 20:28:59,756 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1700738312419_0002
2023-11-23 20:28:59,759 INFO mapreduce.JobSubmitter: Executing with tokens: []
2023-11-23 20:29:00,105 INFO impl.YarnClientImpl: Submitted application application_1700738312419_0002
2023-11-23 20:29:00,147 INFO mapreduce.Job: The url to track the job: http://bokyung-MacBookAir.local:8088/proxy/application_1700738312419_
2023-11-23 20:29:00,147 INFO mapreduce.Job: Running job: job_1700738312419_0002
```

```
2023-11-23 20:29:19,597 INFO mapreduce.Job: Job job_1700738312419_0002 running in uber mode : false
2023-11-23 20:29:19,607 INFO mapreduce.Job:  map 0% reduce 0%
2023-11-23 20:29:31,777 INFO mapreduce.Job:  map 8% reduce 0%
2023-11-23 20:29:41,905 INFO mapreduce.Job:  map 15% reduce 0%
2023-11-23 20:29:52,053 INFO mapreduce.Job:  map 23% reduce 0%
2023-11-23 20:30:02,201 INFO mapreduce.Job:  map 31% reduce 0%
2023-11-23 20:30:12,360 INFO mapreduce.Job:  map 38% reduce 0%
2023-11-23 20:30:22,469 INFO mapreduce.Job:  map 46% reduce 0%
2023-11-23 20:30:32,586 INFO mapreduce.Job:  map 54% reduce 0%
2023-11-23 20:30:42,706 INFO mapreduce.Job:  map 62% reduce 0%
2023-11-23 20:30:52,804 INFO mapreduce.Job:  map 69% reduce 0%
2023-11-23 20:31:02,912 INFO mapreduce.Job:  map 77% reduce 0%
2023-11-23 20:31:13,016 INFO mapreduce.Job:  map 85% reduce 0%
2023-11-23 20:31:23,203 INFO mapreduce.Job:  map 92% reduce 0%
2023-11-23 20:31:31,364 INFO mapreduce.Job:  map 100% reduce 0%
2023-11-23 20:31:48,550 INFO mapreduce.Job:  map 100% reduce 70%
2023-11-23 20:31:54,605 INFO mapreduce.Job:  map 100% reduce 80%
2023-11-23 20:32:00,658 INFO mapreduce.Job:  map 100% reduce 90%
2023-11-23 20:32:06,737 INFO mapreduce.Job:  map 100% reduce 100%
2023-11-23 20:32:08,830 INFO mapreduce.Job: Job job_1700738312419_0002 completed successfully
2023-11-23 20:32:08,898 INFO mapreduce.Job: Counters: 51
  File System Counters
    FILE: Number of bytes read=3354805520
    FILE: Number of bytes written=5072245985
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=1681590991
    HDFS: Number of bytes written=1552540376
    HDFS: Number of read operations=44
    HDFS: Number of large read operations=0
    HDFS: Number of write operations=2
    HDFS: Number of bytes read erasure-coded=0
  Job Counters
    Killed map tasks=1
    Launched map tasks=13
    Launched reduce tasks=1
    Data-local map tasks=13
    Total time spent by all maps in occupied slots (ms)=473188
    Total time spent by all reduces in occupied slots (ms)=139124
    Total time spent by all map tasks (ms)=118297
    Total time spent by all reduce tasks (ms)=34781
    Total vcore-milliseconds taken by all map tasks=118297
    Total vcore-milliseconds taken by all reduce tasks=34781
    Total megabyte-milliseconds taken by all map tasks=484544512
    Total megabyte-milliseconds taken by all reduce tasks=142462976
  Map-Reduce Framework
    Map input records=7999990
    Map output records=7999990
    Map output bytes=1689540504
    Map output materialized bytes=1713541335
    Input split bytes=1625
    Combine input records=0
    Combine output records=0
    Reduce input groups=6999992
    Reduce shuffle bytes=1713541335
    Reduce input records=7999990
    Reduce output records=6999992
    Spilled Records=23662524
    Shuffled Maps =13
    Failed Shuffles=0
    Merged Map outputs=13
    GC time elapsed (ms)=2207
    CPU time spent (ms)=0
    Physical memory (bytes) snapshot=0
    Virtual memory (bytes) snapshot=0
    Total committed heap usage (bytes)=12675186688
  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=1681589366
```

```
      File Output Format Counters
        Bytes Written=1552540376
```

file 5

20:37:31 - 20:52:12 **14분 41초**

```
2023-11-23 20:37:31,141 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes wh
2023-11-23 20:37:32,013 INFO client.DefaultNoHARMFailoverProxyProvider: Connecting to ResourceManager at /0.0.0.0:8032
2023-11-23 20:37:33,224 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface
2023-11-23 20:37:33,236 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/bokyung/.staging/jo
2023-11-23 20:37:34,484 INFO input.FileInputFormat: Total input files to process : 4
2023-11-23 20:37:34,754 INFO mapreduce.JobSubmitter: number of splits:32
2023-11-23 20:37:35,028 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1700739301538_0001
2023-11-23 20:37:35,029 INFO mapreduce.JobSubmitter: Executing with tokens: []
2023-11-23 20:37:35,318 INFO conf.Configuration: resource-types.xml not found
2023-11-23 20:37:35,319 INFO resource.ResourceUtils: Unable to find 'resource-types.xml'.
2023-11-23 20:37:35,668 INFO impl.YarnClientImpl: Submitted application application_1700739301538_0001
2023-11-23 20:37:35,742 INFO mapreduce.Job: The url to track the job: http://bokyung-MacBookAir.local:8088/proxy/application_1700739301538_
2023-11-23 20:37:35,743 INFO mapreduce.Job: Running job: job_1700739301538_0001
2023-11-23 20:37:56,357 INFO mapreduce.Job: Job job_1700739301538_0001 running in uber mode : false
2023-11-23 20:37:56,364 INFO mapreduce.Job:  map 0% reduce 0%
2023-11-23 20:38:09,686 INFO mapreduce.Job:  map 3% reduce 0%
2023-11-23 20:38:21,855 INFO mapreduce.Job:  map 6% reduce 0%
2023-11-23 20:38:34,046 INFO mapreduce.Job:  map 9% reduce 0%
2023-11-23 20:38:47,223 INFO mapreduce.Job:  map 13% reduce 0%
2023-11-23 20:38:58,420 INFO mapreduce.Job:  map 16% reduce 0%
2023-11-23 20:39:10,583 INFO mapreduce.Job:  map 19% reduce 0%
2023-11-23 20:39:23,767 INFO mapreduce.Job:  map 22% reduce 0%
2023-11-23 20:39:35,955 INFO mapreduce.Job:  map 25% reduce 0%
2023-11-23 20:39:48,131 INFO mapreduce.Job:  map 28% reduce 0%
2023-11-23 20:40:00,343 INFO mapreduce.Job:  map 31% reduce 0%
2023-11-23 20:40:12,525 INFO mapreduce.Job:  map 34% reduce 0%
2023-11-23 20:40:24,685 INFO mapreduce.Job:  map 38% reduce 0%
2023-11-23 20:40:36,822 INFO mapreduce.Job:  map 41% reduce 0%
2023-11-23 20:40:48,944 INFO mapreduce.Job:  map 44% reduce 0%
2023-11-23 20:41:01,149 INFO mapreduce.Job:  map 47% reduce 0%
2023-11-23 20:41:13,317 INFO mapreduce.Job:  map 50% reduce 0%
2023-11-23 20:41:24,431 INFO mapreduce.Job:  map 53% reduce 0%
2023-11-23 20:41:35,569 INFO mapreduce.Job:  map 56% reduce 0%
2023-11-23 20:41:47,755 INFO mapreduce.Job:  map 59% reduce 0%
2023-11-23 20:41:59,931 INFO mapreduce.Job:  map 63% reduce 0%
2023-11-23 20:42:12,068 INFO mapreduce.Job:  map 66% reduce 0%
2023-11-23 20:42:24,254 INFO mapreduce.Job:  map 69% reduce 0%
2023-11-23 20:42:37,441 INFO mapreduce.Job:  map 72% reduce 0%
2023-11-23 20:42:48,573 INFO mapreduce.Job:  map 75% reduce 0%
2023-11-23 20:42:59,692 INFO mapreduce.Job:  map 78% reduce 0%
2023-11-23 20:43:10,830 INFO mapreduce.Job:  map 81% reduce 0%
2023-11-23 20:43:22,025 INFO mapreduce.Job:  map 84% reduce 0%
2023-11-23 20:43:38,230 INFO mapreduce.Job:  map 86% reduce 0%
2023-11-23 20:43:43,299 INFO mapreduce.Job:  map 88% reduce 0%
2023-11-23 20:43:59,458 INFO mapreduce.Job:  map 89% reduce 0%
2023-11-23 20:44:05,518 INFO mapreduce.Job:  map 90% reduce 0%
2023-11-23 20:44:07,601 INFO mapreduce.Job:  map 91% reduce 0%
2023-11-23 20:44:18,738 INFO mapreduce.Job:  map 94% reduce 0%
2023-11-23 20:44:28,941 INFO mapreduce.Job:  map 97% reduce 0%
2023-11-23 20:44:40,084 INFO mapreduce.Job:  map 100% reduce 0%
2023-11-23 20:44:57,271 INFO mapreduce.Job:  map 100% reduce 26%
2023-11-23 20:45:03,343 INFO mapreduce.Job:  map 100% reduce 33%
2023-11-23 20:45:15,486 INFO mapreduce.Job:  map 100% reduce 67%
2023-11-23 20:45:21,535 INFO mapreduce.Job:  map 100% reduce 68%
2023-11-23 20:45:27,589 INFO mapreduce.Job:  map 100% reduce 69%
2023-11-23 20:45:33,653 INFO mapreduce.Job:  map 100% reduce 70%
2023-11-23 20:45:38,707 INFO mapreduce.Job:  map 100% reduce 71%
2023-11-23 20:45:44,770 INFO mapreduce.Job:  map 100% reduce 72%
2023-11-23 20:45:50,816 INFO mapreduce.Job:  map 100% reduce 73%
2023-11-23 20:45:56,866 INFO mapreduce.Job:  map 100% reduce 74%
2023-11-23 20:46:02,942 INFO mapreduce.Job:  map 100% reduce 75%
2023-11-23 20:46:08,996 INFO mapreduce.Job:  map 100% reduce 77%
2023-11-23 20:46:15,053 INFO mapreduce.Job:  map 100% reduce 78%
2023-11-23 20:46:21,117 INFO mapreduce.Job:  map 100% reduce 79%
2023-11-23 20:46:27,172 INFO mapreduce.Job:  map 100% reduce 80%
```

```
2023-11-23 20:46:33,234 INFO mapreduce.Job:  map 100% reduce 81%
2023-11-23 20:46:39,332 INFO mapreduce.Job:  map 100% reduce 82%
2023-11-23 20:46:45,380 INFO mapreduce.Job:  map 100% reduce 83%
2023-11-23 20:46:51,441 INFO mapreduce.Job:  map 100% reduce 84%
2023-11-23 20:46:57,498 INFO mapreduce.Job:  map 100% reduce 85%
2023-11-23 20:47:03,543 INFO mapreduce.Job:  map 100% reduce 86%
2023-11-23 20:47:09,601 INFO mapreduce.Job:  map 100% reduce 87%
2023-11-23 20:47:15,664 INFO mapreduce.Job:  map 100% reduce 88%
2023-11-23 20:47:21,783 INFO mapreduce.Job:  map 100% reduce 89%
2023-11-23 20:47:27,824 INFO mapreduce.Job:  map 100% reduce 91%
2023-11-23 20:47:33,898 INFO mapreduce.Job:  map 100% reduce 92%
2023-11-23 20:47:39,934 INFO mapreduce.Job:  map 100% reduce 93%
2023-11-23 20:47:45,982 INFO mapreduce.Job:  map 100% reduce 94%
2023-11-23 20:47:52,034 INFO mapreduce.Job:  map 100% reduce 95%
2023-11-23 20:47:58,062 INFO mapreduce.Job:  map 100% reduce 96%
2023-11-23 20:48:04,106 INFO mapreduce.Job:  map 100% reduce 97%
2023-11-23 20:48:10,138 INFO mapreduce.Job:  map 100% reduce 98%
2023-11-23 20:48:16,179 INFO mapreduce.Job:  map 100% reduce 99%
2023-11-23 20:48:21,243 INFO mapreduce.Job:  map 100% reduce 100%
2023-11-23 20:48:23,316 INFO mapreduce.Job: Job job_1700739301538_0001 completed successfully
2023-11-23 20:48:23,516 INFO mapreduce.Job: Counters: 51
  File System Counters
    FILE: Number of bytes read=9309757722
    FILE: Number of bytes written=14022764778
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=4113939254
    HDFS: Number of bytes written=2101926767
    HDFS: Number of read operations=101
    HDFS: Number of large read operations=0
    HDFS: Number of write operations=2
    HDFS: Number of bytes read erasure-coded=0
  Job Counters
    Killed map tasks=1
    Launched map tasks=32
    Launched reduce tasks=1
    Data-local map tasks=32
    Total time spent by all maps in occupied slots (ms)=1510128
    Total time spent by all reduces in occupied slots (ms)=882364
    Total time spent by all map tasks (ms)=377532
    Total time spent by all reduce tasks (ms)=220591
    Total vcore-milliseconds taken by all map tasks=377532
    Total vcore-milliseconds taken by all reduce tasks=220591
    Total megabyte-milliseconds taken by all map tasks=1546371072
    Total megabyte-milliseconds taken by all reduce tasks=903540736
  Map-Reduce Framework
    Map input records=30000004
    Map output records=29999972
    Map output bytes=4623815458
    Map output materialized bytes=4703817546
    Input split bytes=3744
    Combine input records=0
    Combine output records=0
    Reduce input groups=19999987
    Reduce shuffle bytes=4703817546
    Reduce input records=29999972
    Reduce output records=9999985
    Spilled Records=87669502
    Shuffled Maps =32
    Failed Shuffles=0
    Merged Map outputs=32
    GC time elapsed (ms)=6076
    CPU time spent (ms)=0
    Physical memory (bytes) snapshot=0
    Virtual memory (bytes) snapshot=0
    Total committed heap usage (bytes)=35841376256
  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=4113935510
  File Output Format Counters
    Bytes Written=2101926767
```

```
2023-11-23 20:48:23,574 INFO client.DefaultNoHARMFailoverProxyProvider: Connecting to ResourceManager at /0.0.0.0:8032
2023-11-23 20:48:23,725 WARN mapreduce.JobResourceUploader: Hadoop command-line option parsing not performed. Implement the Tool interface
2023-11-23 20:48:23,746 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/bokyung/.staging/jo
2023-11-23 20:48:24,758 INFO input.FileInputFormat: Total input files to process : 1
2023-11-23 20:48:25,265 INFO mapreduce.JobSubmitter: number of splits:16
2023-11-23 20:48:25,361 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1700739301538_0002
2023-11-23 20:48:25,362 INFO mapreduce.JobSubmitter: Executing with tokens: []
2023-11-23 20:48:25,477 INFO impl.YarnClientImpl: Submitted application application_1700739301538_0002
2023-11-23 20:48:25,512 INFO mapreduce.Job: The url to track the job: http://bokyung-MacBookAir.local:8088/proxy/application_1700739301538_
2023-11-23 20:48:25,512 INFO mapreduce.Job: Running job: job_1700739301538_0002
2023-11-23 20:48:47,084 INFO mapreduce.Job: Job job_1700739301538_0002 running in uber mode : false
2023-11-23 20:48:47,097 INFO mapreduce.Job:  map 0% reduce 0%
2023-11-23 20:48:58,437 INFO mapreduce.Job:  map 6% reduce 0%
2023-11-23 20:49:07,527 INFO mapreduce.Job:  map 13% reduce 0%
2023-11-23 20:49:18,698 INFO mapreduce.Job:  map 19% reduce 0%
2023-11-23 20:49:28,842 INFO mapreduce.Job:  map 25% reduce 0%
2023-11-23 20:49:38,985 INFO mapreduce.Job:  map 31% reduce 0%
2023-11-23 20:49:49,108 INFO mapreduce.Job:  map 38% reduce 0%
2023-11-23 20:49:59,238 INFO mapreduce.Job:  map 44% reduce 0%
2023-11-23 20:50:09,368 INFO mapreduce.Job:  map 50% reduce 0%
2023-11-23 20:50:19,494 INFO mapreduce.Job:  map 56% reduce 0%
2023-11-23 20:50:29,607 INFO mapreduce.Job:  map 63% reduce 0%
2023-11-23 20:50:39,731 INFO mapreduce.Job:  map 69% reduce 0%
2023-11-23 20:50:49,867 INFO mapreduce.Job:  map 75% reduce 0%
2023-11-23 20:50:59,989 INFO mapreduce.Job:  map 81% reduce 0%
2023-11-23 20:51:10,162 INFO mapreduce.Job:  map 88% reduce 0%
2023-11-23 20:51:20,317 INFO mapreduce.Job:  map 94% reduce 0%
2023-11-23 20:51:29,402 INFO mapreduce.Job:  map 100% reduce 0%
2023-11-23 20:51:45,556 INFO mapreduce.Job:  map 100% reduce 67%
2023-11-23 20:51:51,615 INFO mapreduce.Job:  map 100% reduce 74%
2023-11-23 20:51:57,666 INFO mapreduce.Job:  map 100% reduce 82%
2023-11-23 20:52:03,728 INFO mapreduce.Job:  map 100% reduce 90%
2023-11-23 20:52:09,783 INFO mapreduce.Job:  map 100% reduce 98%
2023-11-23 20:52:11,811 INFO mapreduce.Job:  map 100% reduce 100%
2023-11-23 20:52:12,882 INFO mapreduce.Job: Job job_1700739301538_0002 completed successfully
2023-11-23 20:52:12,946 INFO mapreduce.Job: Counters: 51
  File System Counters
    FILE: Number of bytes read=4283856384
    FILE: Number of bytes written=6430519242
    FILE: Number of read operations=0
    FILE: Number of large read operations=0
    FILE: Number of write operations=0
    HDFS: Number of bytes read=2101990207
    HDFS: Number of bytes written=1940927009
    HDFS: Number of read operations=53
    HDFS: Number of large read operations=0
    HDFS: Number of write operations=2
    HDFS: Number of bytes read erasure-coded=0
  Job Counters
    Killed map tasks=1
    Launched map tasks=16
    Launched reduce tasks=1
    Data-local map tasks=16
    Total time spent by all maps in occupied slots (ms)=579416
    Total time spent by all reduces in occupied slots (ms)=168304
    Total time spent by all map tasks (ms)=144854
    Total time spent by all reduce tasks (ms)=42076
    Total vcore-milliseconds taken by all map tasks=144854
    Total vcore-milliseconds taken by all reduce tasks=42076
    Total megabyte-milliseconds taken by all map tasks=593321984
    Total megabyte-milliseconds taken by all reduce tasks=172343296
  Map-Reduce Framework
    Map input records=9999985
    Map output records=9999985
    Map output bytes=2111927142
    Map output materialized bytes=2141928186
    Input split bytes=2000
    Combine input records=0
    Combine output records=0
    Reduce input groups=8999987
    Reduce shuffle bytes=2141928186
    Reduce input records=9999985
    Reduce output records=8999987
    Spilled Records=29999955
    Shuffled Maps =16
    Failed Shuffles=0
    Merged Map outputs=16
    GC time elapsed (ms)=2745
```

```
    CPU time spent (ms)=0
    Physical memory (bytes) snapshot=0
    Virtual memory (bytes) snapshot=0
    Total committed heap usage (bytes)=18122539008
  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=2101988207
  File Output Format Counters
    Bytes Written=1940927009
```