```
/// ***************************************************************************** ///
>>> start-all.sh - 하둡 데몬들 실행시킵니다.
>>> jps - 실행중인 하둡 데몬들(자바 프로세서) 확인합니다.
/// ***************************************************************************** ///
chosun@chosun-VirtualBox:~/hadoop$ start-all.sh
starting namenode, logging to /home/chosun/hadoop-1.2.1/libexec/../logs/hadoop-chosun-namenode-chosun-VirtualBox.out
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/home/chosun/hadoop-1.2.1/
hadoop-core-1.2.1.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
localhost: starting datanode, logging to /home/chosun/hadoop-1.2.1/libexec/../logs/hadoop-chosun-datanode-chosun-VirtualBox.out
localhost: WARNING: An illegal reflective access operation has occurred
localhost: WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/home/chosun/hadoop-1.2.1/
hadoop-core-1.2.1.jar) to method sun.security.krb5.Config.getInstance()
localhost: WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
localhost: WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
localhost: WARNING: All illegal access operations will be denied in a future release
localhost: starting secondarynamenode, logging to /home/chosun/hadoop-1.2.1/libexec/../logs/hadoop-chosun-secondarynamenode-
chosun-VirtualBox.out
localhost: WARNING: An illegal reflective access operation has occurred
localhost: WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/home/chosun/hadoop-1.2.1/
hadoop-core-1.2.1.jar) to method sun.security.krb5.Config.getInstance()
localhost: WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
localhost: WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
localhost: WARNING: All illegal access operations will be denied in a future release
starting jobtracker, logging to /home/chosun/hadoop-1.2.1/libexec/../logs/hadoop-chosun-jobtracker-chosun-VirtualBox.out
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/home/chosun/hadoop-1.2.1/
hadoop-core-1.2.1.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
localhost: starting tasktracker, logging to /home/chosun/hadoop-1.2.1/libexec/../logs/hadoop-chosun-tasktracker-chosun-VirtualBox.out
localhost: WARNING: An illegal reflective access operation has occurred
localhost: WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/home/chosun/hadoop-1.2.1/
hadoop-core-1.2.1.jar) to method sun.security.krb5.Config.getInstance()
localhost: WARNING: Please consider reporting this to the maintainers of org.apache.hadoop.security.authentication.util.KerberosUtil
localhost: WARNING: Use --illegal-access=warn to enable warnings of further illegal reflective access operations
localhost: WARNING: All illegal access operations will be denied in a future release
chosun@chosun-VirtualBox:~/hadoop$
chosun@chosun-VirtualBox:~/hadoop$ jps
15510 Jps
14855 DataNode
15080 SecondaryNameNode
15209 JobTracker
15420 TaskTracker
4286 org.eclipse.equinox.launcher_1.5.300.v20190213-1655.jar
14670 NameNode
chosun@chosun-VirtualBox:~/hadoop$
/// ***************************************************************************** ///
>>> HWEx7_2MapSideJoin.jar은 맵 사이드 조인을 수행하는 jar파일입니다.
>>> HWEx7_3ReduceSideJoin.jar은 리듀스 사이드 조인을 수행하는 jar파일입니다.
/// ***************************************************************************** ///
chosun@chosun-VirtualBox:~/hadoop$ ls
1987.csv   HWEx7_3ReduceSideJoin.jar   HomeWorkEX6_2_Seq.jar   README.txt   conf   hadoop-ant-1.2.1.jar
hadoop-minicluster-1.2.1.jar   ivy.xml   pids   test.py   2008.csv   HomeWorkEX6_1.jar   HomeWorkEX6_3_Total.jar   bin   contrib
hadoop-client-1.2.1.jar   hadoop-test-1.2.1.jar   lib   sbin   webapps   CHANGES.txt   HomeWorkEX6_2_Map.jar   LICENSE.txt
build.xml   dataexpo   hadoop-core-1.2.1.jar   hadoop-tools-1.2.1.jar   libexec   share   HWEx7_2MapSideJoin.jar
HomeWorkEX6_2_Search.jar   NOTICE.txt   c++   docs   hadoop-examples-1.2.1.jar   ivy   logs   src
```

chosun@chosun-VirtualBox:~/hadoop$

>>> wget http://stat-computing.org/dataexpo/2009/carriers.csv에서 항공사 코드 데이터를 내려받습니다.
>>> 내려받은 carriers.csv파일의 내용 중 첫 번째 줄의 칼럼 정보를 제거합니다.
chosun@chosun-VirtualBox:~/hadoop$ wget http://stat-computing.org/dataexpo/2009/carriers.csv
--2019-05-13 19:57:31-- http://stat-computing.org/dataexpo/2009/carriers.csv
Resolving stat-computing.org (stat-computing.org)... 52.218.241.91
접속 stat-computing.org (stat-computing.org)|52.218.241.91|:80... 접속됨.
HTTP request sent, awaiting response... 200 OK
Length: 43758 (43K) [text/csv]
Saving to: 'carriers.csv'
carriers.csv
100%
[==>]
42.73K   177KB/s   in 0.2s
2019-05-13 19:57:32 (177 KB/s) - 'carriers.csv' saved [43758/43758]
chosun@chosun-VirtualBox:~/hadoop$
chosun@chosun-VirtualBox:~/hadoop$ ls
1987.csv   HWEx7_3ReduceSideJoin.jar   HomeWorkEX6_2_Seq.jar   README.txt   carriers.csv   docs
hadoop-examples-1.2.1.jar   ivy   logs   src   2008.csv   HomeWorkEX6_1.jar   HomeWorkEX6_3_Total.jar   bin   conf
hadoop-ant-1.2.1.jar   hadoop-minicluster-1.2.1.jar   ivy.xml   pids   test.py   CHANGES.txt   HomeWorkEX6_2_Map.jar   LICENSE.txt
build.xml   contrib   hadoop-client-1.2.1.jar   hadoop-test-1.2.1.jar   lib   sbin   webapps   HWEx7_2MapSideJoin.jar
HomeWorkEX6_2_Search.jar   NOTICE.txt   c++   dataexpo   hadoop-core-1.2.1.jar   hadoop-tools-1.2.1.jar   libexec   share
chosun@chosun-VirtualBox:~/hadoop$
chosun@chosun-VirtualBox:~/hadoop$ sed -e '1d' carriers.csv > carriers_new.csv
chosun@chosun-VirtualBox:~/hadoop$
chosun@chosun-VirtualBox:~/hadoop$ ls
1987.csv   HWEx7_3ReduceSideJoin.jar   HomeWorkEX6_2_Seq.jar   README.txt   carriers.csv   dataexpo
hadoop-core-1.2.1.jar   hadoop-tools-1.2.1.jar   libexec   share   2008.csv   HomeWorkEX6_1.jar   HomeWorkEX6_3_Total.jar   bin
carriers_new.csv   docs   hadoop-examples-1.2.1.jar   ivy   logs   src   CHANGES.txt   HomeWorkEX6_2_Map.jar   LICENSE.txt
build.xml   conf   hadoop-ant-1.2.1.jar   hadoop-minicluster-1.2.1.jar   ivy.xml   pids   test.py   HWEx7_2MapSideJoin.jar
HomeWorkEX6_2_Search.jar   NOTICE.txt   c++   contrib   hadoop-client-1.2.1.jar   hadoop-test-1.2.1.jar   lib   sbin   webapps
chosun@chosun-VirtualBox:~/hadoop$
>>>>>>>>>>>> 원래 예제를 실행하려면 carriers_new.csv파일의 따옴표(" ")도 제거해줘야 합니다. <<<<<<<<<<<<<<<
>>>>>>>>>>>>>>>>>>>>>>>>> clc에 올라온 carriers_new2.csv 파일을 업로드 하겠습니다.<<<<<<<<<<<<<<<<<<<<<<<<<<<
>>> carriers_new2.csv 파일을 hdfs에 업로드합니다.
>>> 먼저 HW_0513_meta 디렉터리를 hdfs에 생성하고, 그 디렉터리에 carriers_new2.csv 파일을 carriers_new.csv로 업로드합니다.
chosun@chosun-VirtualBox:~/hadoop$ hadoop fs -ls
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/home/chosun/hadoop-1.2.1/
hadoop-core-1.2.1.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
Found 4 items
drwxr-xr-x   - chosun supergroup   0 2019-05-12 13:28 /user/chosun/HW_0512
drwxr-xr-x   - chosun supergroup   0 2019-05-12 14:09 /user/chosun/HW_Mapfile
drwxr-xr-x   - chosun supergroup   0 2019-05-12 14:13 /user/chosun/HW_Sequencefile
drwxr-xr-x   - chosun supergroup   0 2019-05-12 14:14 /user/chosun/HW_total
chosun@chosun-VirtualBox:~/hadoop$
chosun@chosun-VirtualBox:~/hadoop$ hadoop fs -mkdir HW_0513_meta
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/home/chosun/hadoop-1.2.1/
hadoop-core-1.2.1.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
chosun@chosun-VirtualBox:~/hadoop$

```
chosun@chosun-VirtualBox:~/hadoop$ hadoop fs -ls
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/home/chosun/hadoop-1.2.1/
hadoop-core-1.2.1.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
Found 5 items
drwxr-xr-x   - chosun supergroup          0 2019-05-12 13:28 /user/chosun/HW_0512
drwxr-xr-x   - chosun supergroup          0 2019-05-13 19:58 /user/chosun/HW_0513_meta
drwxr-xr-x   - chosun supergroup          0 2019-05-12 14:09 /user/chosun/HW_Mapfile
drwxr-xr-x   - chosun supergroup          0 2019-05-12 14:13 /user/chosun/HW_Sequencefile
drwxr-xr-x   - chosun supergroup          0 2019-05-12 14:14 /user/chosun/HW_total
chosun@chosun-VirtualBox:~/hadoop$
chosun@chosun-VirtualBox:~/hadoop$ hadoop fs -put carriers_new2.csv HW_0513_meta/carriers_new.csv
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/home/chosun/hadoop-1.2.1/
hadoop-core-1.2.1.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
chosun@chosun-VirtualBox:~/hadoop$
chosun@chosun-VirtualBox:~/hadoop$ hadoop fs -ls HW_0513_meta
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/home/chosun/hadoop-1.2.1/
hadoop-core-1.2.1.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
Found 1 items
-rw-r--r--   1 chosun supergroup      43740 2019-05-13 19:59 /user/chosun/HW_0513_meta/carriers_new.csv
chosun@chosun-VirtualBox:~/hadoop$
/// ************************************************************************************ ///
>>> [Ex7.2]
>>> jar 파일을 실행합니다.
>>> HW_0513_meta/carriers_new.csv와 HW_0512를 처리한 것을 map_join디렉토리에 저장합니다.
>>> MapSideJoin은 클래스 이름입니다.
/// ************************************************************************************ ///
chosun@chosun-VirtualBox:~/hadoop$ hadoop jar HWEx7_2MapSideJoin.jar MapSideJoin HW_0513_meta/carriers_new.csv HW_0512 map_join
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/home/chosun/hadoop-1.2.1/
hadoop-core-1.2.1.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
19/05/13 20:06:04 INFO input.FileInputFormat: Total input paths to process : 1
19/05/13 20:06:04 INFO util.NativeCodeLoader: Loaded the native-hadoop library
19/05/13 20:06:04 WARN snappy.LoadSnappy: Snappy native library not loaded
19/05/13 20:06:05 INFO mapred.JobClient: Running job: job_201905131956_0001              // job 시작
19/05/13 20:06:06 INFO mapred.JobClient:  map 0% reduce 0%                               // 맵 시작, 리듀스 --- x
19/05/13 20:06:16 INFO mapred.JobClient:  map 18% reduce 0%                              // ~~ ing
19/05/13 20:06:21 INFO mapred.JobClient:  map 36% reduce 0%
19/05/13 20:06:27 INFO mapred.JobClient:  map 54% reduce 0%
19/05/13 20:06:32 INFO mapred.JobClient:  map 72% reduce 0%
19/05/13 20:06:37 INFO mapred.JobClient:  map 90% reduce 0%                              // ~~ ing
19/05/13 20:06:40 INFO mapred.JobClient:  map 100% reduce 0%                             // 맵-리듀스 끝
19/05/13 20:06:41 INFO mapred.JobClient: Job complete: job_201905131956_0001            // job 끝
19/05/13 20:06:41 INFO mapred.JobClient: Counters: 19
19/05/13 20:06:41 INFO mapred.JobClient:   Map-Reduce Framework
19/05/13 20:06:41 INFO mapred.JobClient:     Spilled Records=0                           // 작업 중에 디스크에 기록 된 총 레코드 수
19/05/13 20:06:41 INFO mapred.JobClient:     Virtual memory (bytes) snapshot=22930333696     // 모든 작업에 사용된 총 가상 메모리
```

```
19/05/13 20:06:41 INFO mapred.JobClient:       Map input records=7009728                        // 맵에 입력되는 데이터의 수
19/05/13 20:06:41 INFO mapred.JobClient:       SPLIT_RAW_BYTES=1265                  // 맵에서 가져온 입력 분할 개체의 바이트 수
19/05/13 20:06:41 INFO mapred.JobClient:       Map output records=7009728                   // 맵의 출력으로 생성된 데이터의 수
19/05/13 20:06:41 INFO mapred.JobClient:       Physical memory (bytes) snapshot=1690693632     // 모든 작업에 사용되는 총 물리적 메모리
19/05/13 20:06:41 INFO mapred.JobClient:       CPU time spent (ms)=35830                              // 소요된 cpu시간
19/05/13 20:06:41 INFO mapred.JobClient:       Total committed heap usage (bytes)=961544192        // JVM에 사용할 수 있는 총 메모리 양
19/05/13 20:06:41 INFO mapred.JobClient:     File Input Format Counters
19/05/13 20:06:41 INFO mapred.JobClient:       Bytes Read=689454004              // 모든 파일 시스템에 대해 모든 테스크에서 읽은 데이터 양
19/05/13 20:06:41 INFO mapred.JobClient:     FileSystemCounters
19/05/13 20:06:41 INFO mapred.JobClient:       HDFS_BYTES_READ=689455269                    // HDFS에서 읽은 데이터 양
19/05/13 20:06:41 INFO mapred.JobClient:       FILE_BYTES_WRITTEN=642775               // 로컬 파일 시스템에 기록된 데이터 양
19/05/13 20:06:41 INFO mapred.JobClient:       HDFS_BYTES_WRITTEN=900468672                  // HDFS에 기록된 데이터 양
19/05/13 20:06:41 INFO mapred.JobClient:     Job Counters
19/05/13 20:06:41 INFO mapred.JobClient:       Launched map tasks=11                                   // 실행된 총 맵 작업 수
19/05/13 20:06:41 INFO mapred.JobClient:       SLOTS_MILLIS_REDUCES=0                   // 모든 리듀스 작업에 의해 소요된 총 시간(ms)
19/05/13 20:06:41 INFO mapred.JobClient:       Total time spent by all reduces waiting after reserving slots (ms)=0
19/05/13 20:06:41 INFO mapred.JobClient:       SLOTS_MILLIS_MAPS=51496                   // 모든 맵 작업에 의해 사용한 총 시간(ms)
19/05/13 20:06:41 INFO mapred.JobClient:       Total time spent by all maps waiting after reserving slots (ms)=0
19/05/13 20:06:41 INFO mapred.JobClient:       Data-local map tasks=11              // 필요한 데이터가 포함된 노드에서 시작된 맵 테스크 수
19/05/13 20:06:41 INFO mapred.JobClient:     File Output Format Counters
19/05/13 20:06:41 INFO mapred.JobClient:       Bytes Written=900468672            // 모든 파일 시스템에 대해 모든 태스크에서 작성된 데이터 양
MR-Job Result:0
chosun@chosun-VirtualBox:~/hadoop$
/// ******************************************************************************************** ///
>>> 실행 결과 ::: 리듀서가 없는 맵리듀스 잡이므로 매퍼의 출력 데이터가 모두 생성되었습니다.
>>> 만약 리듀서를 정의했다면 part-m-xxxxx파일이 아니라, 리듀서의 출력 데이터인 part-r-xxxxx파일들이 생성됩니다.
>>> 첫 번째 출력 데이터인 part-m-00000을 조회해본 결과,
    항공사 이름과 해당 항공사에 대한 운항 통계 데이터가 정상적으로 저장돼 있음을 확인했습니다.
/// ******************************************************************************************** ///
chosun@chosun-VirtualBox:~/hadoop$ hadoop fs -ls map_join
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/home/chosun/hadoop-1.2.1/
hadoop-core-1.2.1.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
Found 13 items
-rw-r--r--   1 chosun supergroup          0 2019-05-13 20:06 /user/chosun/map_join/_SUCCESS
drwxr-xr-x   - chosun supergroup          0 2019-05-13 20:06 /user/chosun/map_join/_logs
-rw-r--r--   1 chosun supergroup   87401259 2019-05-13 20:06 /user/chosun/map_join/part-m-00000
-rw-r--r--   1 chosun supergroup   87045872 2019-05-13 20:06 /user/chosun/map_join/part-m-00001
-rw-r--r--   1 chosun supergroup   88157694 2019-05-13 20:06 /user/chosun/map_join/part-m-00002
-rw-r--r--   1 chosun supergroup   88954716 2019-05-13 20:06 /user/chosun/map_join/part-m-00003
-rw-r--r--   1 chosun supergroup   87780517 2019-05-13 20:06 /user/chosun/map_join/part-m-00004
-rw-r--r--   1 chosun supergroup   87481529 2019-05-13 20:06 /user/chosun/map_join/part-m-00005
-rw-r--r--   1 chosun supergroup   87195853 2019-05-13 20:06 /user/chosun/map_join/part-m-00006
-rw-r--r--   1 chosun supergroup   87122067 2019-05-13 20:06 /user/chosun/map_join/part-m-00007
-rw-r--r--   1 chosun supergroup   86756011 2019-05-13 20:06 /user/chosun/map_join/part-m-00008
-rw-r--r--   1 chosun supergroup   89365213 2019-05-13 20:06 /user/chosun/map_join/part-m-00009
-rw-r--r--   1 chosun supergroup   23207941 2019-05-13 20:06 /user/chosun/map_join/part-m-00010
chosun@chosun-VirtualBox:~/hadoop$
chosun@chosun-VirtualBox:~/hadoop$ hadoop fs -cat map_join/part-m-00000 | head -10
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/home/chosun/hadoop-1.2.1/
hadoop-core-1.2.1.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
WN      Southwest Airlines Co.2008,1,3,4,2003,1955,2211,2225,WN,335,N712SW,128,150,116,-14,8,IAD,TPA,810,4,8,0,,0,NA,NA,NA,NA,NA
WN      Southwest Airlines Co.2008,1,3,4,754,735,1002,1000,WN,3231,N772SW,128,145,113,2,19,IAD,TPA,810,5,10,0,,0,NA,NA,NA,NA,NA
WN      Southwest Airlines Co.2008,1,3,4,628,620,804,750,WN,448,N428WN,96,90,76,14,8,IND,BWI,515,3,17,0,,0,NA,NA,NA,NA,NA
```

WN   Southwest Airlines Co.2008,1,3,4,926,930,1054,1100,WN,1746,N612SW,88,90,78,−6,−4,IND,BWI,515,3,7,0,,0,NA,NA,NA,NA,NA

WN   Southwest Airlines Co.2008,1,3,4,1829,1755,1959,1925,WN,3920,N464WN,90,90,77,34,34,IND,BWI,515,3,10,0,,0,2,0,0,0,32

WN   Southwest Airlines Co.2008,1,3,4,1940,1915,2121,2110,WN,378,N726SW,101,115,87,11,25,IND,JAX,688,4,10,0,,0,NA,NA,NA,NA,NA

WN   Southwest Airlines Co.2008,1,3,4,1937,1830,2037,1940,WN,509,N763SW,240,250,230,57,67,IND,LAS,1591,3,7,0,,0,10,0,0,0,47

WN   Southwest Airlines Co.2008,1,3,4,1039,1040,1132,1150,WN,535,N428WN,233,250,219,−18,−1,IND,LAS,1591,7,7,0,,0,NA,NA,NA,NA,NA

WN   Southwest Airlines Co.2008,1,3,4,617,615,652,650,WN,11,N689SW,95,95,70,2,2,IND,MCI,451,6,19,0,,0,NA,NA,NA,NA,NA

WN   Southwest Airlines Co.2008,1,3,4,1620,1620,1639,1655,WN,810,N648SW,79,95,70,−16,0,IND,MCI,451,3,6,0,,0,NA,NA,NA,NA,NA

cat: Unable to write to output stream.
chosun@chosun−VirtualBox:~/hadoop$
/// *** ///
>>> [Ex7.3]
>>> jar 파일을 실행합니다.
>>> HW_0513_meta/carriers_new.csv와 HW_0512를 처리한 것을 reduce_join_output 디렉토리에 저장합니다.
>>> ReduceSideJoin은 클래스 이름입니다.
/// *** ///
chosun@chosun−VirtualBox:~/hadoop$ hadoop jar HWEx7_3ReduceSideJoin.jar ReduceSideJoin HW_0513_meta/carriers_new.csv HW_0512 reduce_join_output
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/home/chosun/hadoop−1.2.1/hadoop−core−1.2.1.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
19/05/13 20:10:50 INFO input.FileInputFormat: Total input paths to process : 1
19/05/13 20:10:50 INFO util.NativeCodeLoader: Loaded the native−hadoop library
19/05/13 20:10:50 WARN snappy.LoadSnappy: Snappy native library not loaded
19/05/13 20:10:50 INFO input.FileInputFormat: Total input paths to process : 1
19/05/13 20:10:50 INFO mapred.JobClient: Running job: job_201905131956_0002   // job 시작
19/05/13 20:10:51 INFO mapred.JobClient: map 0% reduce 0%   // 맵−리듀스 시작
19/05/13 20:11:09 INFO mapred.JobClient: map 1% reduce 0%   // ~~ ing
19/05/13 20:11:11 INFO mapred.JobClient: map 9% reduce 0%
19/05/13 20:11:12 INFO mapred.JobClient: map 16% reduce 0%
19/05/13 20:11:19 INFO mapred.JobClient: map 33% reduce 0%
19/05/13 20:11:20 INFO mapred.JobClient: map 33% reduce 2%
19/05/13 20:11:23 INFO mapred.JobClient: map 33% reduce 11%
19/05/13 20:11:27 INFO mapred.JobClient: map 41% reduce 11%
19/05/13 20:11:28 INFO mapred.JobClient: map 50% reduce 11%
19/05/13 20:11:35 INFO mapred.JobClient: map 66% reduce 16%
19/05/13 20:11:44 INFO mapred.JobClient: map 83% reduce 19%
19/05/13 20:11:47 INFO mapred.JobClient: map 91% reduce 22%
19/05/13 20:11:48 INFO mapred.JobClient: map 100% reduce 22%
19/05/13 20:11:50 INFO mapred.JobClient: map 100% reduce 33%
19/05/13 20:11:53 INFO mapred.JobClient: map 100% reduce 70%
19/05/13 20:11:56 INFO mapred.JobClient: map 100% reduce 73%
19/05/13 20:11:59 INFO mapred.JobClient: map 100% reduce 77%
19/05/13 20:12:02 INFO mapred.JobClient: map 100% reduce 80%
19/05/13 20:12:05 INFO mapred.JobClient: map 100% reduce 83%
19/05/13 20:12:08 INFO mapred.JobClient: map 100% reduce 86%
19/05/13 20:12:11 INFO mapred.JobClient: map 100% reduce 89%
19/05/13 20:12:14 INFO mapred.JobClient: map 100% reduce 92%
19/05/13 20:12:17 INFO mapred.JobClient: map 100% reduce 96%
19/05/13 20:12:20 INFO mapred.JobClient: map 100% reduce 99%   // ~~ ing
19/05/13 20:12:22 INFO mapred.JobClient: map 100% reduce 100%   // 맵−리듀스 끝
19/05/13 20:12:22 INFO mapred.JobClient: Job complete: job_201905131956_0002   // job 끝
19/05/13 20:12:22 INFO mapred.JobClient: Counters: 29
19/05/13 20:12:22 INFO mapred.JobClient: Map−Reduce Framework
19/05/13 20:12:22 INFO mapred.JobClient: Spilled Records=20662279   // 작업 중에 디스크에 기록 된 총 레코드 수
19/05/13 20:12:22 INFO mapred.JobClient: Map output materialized bytes=773579277   // 실제로 디스크에 기록된 맵 출력 바이트 수
19/05/13 20:12:22 INFO mapred.JobClient: Reduce input records=7011219   // 리듀스에 입력된 데이터의 수
19/05/13 20:12:22 INFO mapred.JobClient: Virtual memory (bytes) snapshot=27140116480   // 모든 작업에 사용된 총 가상 메모리
19/05/13 20:12:22 INFO mapred.JobClient: Map input records=7011219   // 맵에 입력되는 데이터의 수

```
19/05/13 20:12:22 INFO mapred.JobClient:        SPLIT_RAW_BYTES=2995                    // 맵에서 가져온 입력 분할 개체의 바이트 수
19/05/13 20:12:22 INFO mapred.JobClient:        Map output bytes=759556767              // 모든 맵에 의해 생성된 압축되지 않은 출력 바이트수
19/05/13 20:12:22 INFO mapred.JobClient:        Reduce shuffle bytes=773579277          // 리듀스 단계에 셔플된 맵 출력의 바이트 수
19/05/13 20:12:22 INFO mapred.JobClient:        Physical memory (bytes) snapshot=3207630848    // 모든 작업에 사용되는 총 물리적 메모리
19/05/13 20:12:22 INFO mapred.JobClient:        Reduce input groups=1491                // 리듀스 단계에서 처리한 개별 키 그룹의 수
19/05/13 20:12:22 INFO mapred.JobClient:        Combine output records=0                           // 출력 레코드 수
19/05/13 20:12:22 INFO mapred.JobClient:        Reduce output records=7009708           // 리듀스의 출력으로 생성된 데이터의 수
19/05/13 20:12:22 INFO mapred.JobClient:        Map output records=7011219                // 맵의 출력으로 생성된 데이터의 수
19/05/13 20:12:22 INFO mapred.JobClient:        Combine input records=0                           // 입력 레코드 수
19/05/13 20:12:22 INFO mapred.JobClient:        CPU time spent (ms)=87640                           // 소요된 cpu시간
19/05/13 20:12:22 INFO mapred.JobClient:        Total committed heap usage (bytes)=2125463552    // JVM에 사용할 수 있는 총 메모리 양
19/05/13 20:12:22 INFO mapred.JobClient:     File Input Format Counters
19/05/13 20:12:22 INFO mapred.JobClient:        Bytes Read=0                          // 모든 파일 시스템에 대해 모든 테스크에서 읽은 데이터 양
19/05/13 20:12:22 INFO mapred.JobClient:     FileSystemCounters
19/05/13 20:12:22 INFO mapred.JobClient:        HDFS_BYTES_READ=689494775                           // HDFS에서 읽은 데이터 양
19/05/13 20:12:22 INFO mapred.JobClient:        FILE_BYTES_WRITTEN=2280337398           // 로컬 파일 시스템에 기록된 데이터 양
19/05/13 20:12:22 INFO mapred.JobClient:        FILE_BYTES_READ=1505990302              // 로컬 파일 시스템에서 읽은 데이터 양
19/05/13 20:12:22 INFO mapred.JobClient:        HDFS_BYTES_WRITTEN=900468672                       // HDFS에 기록된 데이터 양
19/05/13 20:12:22 INFO mapred.JobClient:     Job Counters
19/05/13 20:12:22 INFO mapred.JobClient:        Launched map tasks=12                           // 실행된 총 맵 작업 수
19/05/13 20:12:22 INFO mapred.JobClient:        Launched reduce tasks=1                           // 실행된 총 리듀스 작업 수
19/05/13 20:12:22 INFO mapred.JobClient:        SLOTS_MILLIS_REDUCES=48340              // 모든 리듀스 작업에 의해 소요된 총 시간(ms)
19/05/13 20:12:22 INFO mapred.JobClient:        Total time spent by all reduces waiting after reserving slots (ms)=0
19/05/13 20:12:22 INFO mapred.JobClient:        SLOTS_MILLIS_MAPS=57017                 // 모든 맵 작업에 의해 사용한 총 시간(ms)
19/05/13 20:12:22 INFO mapred.JobClient:        Total time spent by all maps waiting after reserving slots (ms)=0
19/05/13 20:12:22 INFO mapred.JobClient:        Data-local map tasks=12                 // 필요한 데이터가 포함된 노드에서 시작된 맵 테스크 수
19/05/13 20:12:22 INFO mapred.JobClient:     File Output Format Counters
19/05/13 20:12:22 INFO mapred.JobClient:        Bytes Written=900468672                 // 모든 파일 시스템에 대해 모든 태스크에서 작성된 데이터 양
MR-Job Result:0
chosun@chosun-VirtualBox:~/hadoop$
/// ****************************************************************************************** ///
>>> 출력 경로를 조회해본 결과, 로그 파일과 출력 데이터가 생성되었음을 확인했습니다.
/// ****************************************************************************************** ///
chosun@chosun-VirtualBox:~/hadoop$ hadoop fs -ls reduce_join_output
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/home/chosun/hadoop-1.2.1/
hadoop-core-1.2.1.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
Found 3 items
-rw-r--r--   1 chosun supergroup          0 2019-05-13 20:12 /user/chosun/reduce_join_output/_SUCCESS
drwxr-xr-x   - chosun supergroup          0 2019-05-13 20:10 /user/chosun/reduce_join_output/_logs
-rw-r--r--   1 chosun supergroup  900468672 2019-05-13 20:11 /user/chosun/reduce_join_output/part-r-00000
chosun@chosun-VirtualBox:~/hadoop$
/// ****************************************************************************************** ///
>>> 저장된 텍스트 파일을 조회해본 결과, 항공사 이름과 항공 운항 통계 데이터가 정상적으로 출력된 것을 확인했습니다.
/// ****************************************************************************************** ///
chosun@chosun-VirtualBox:~/hadoop$ hadoop fs -cat reduce_join_output/part-r-00000 | head -10
WARNING: An illegal reflective access operation has occurred
WARNING: Illegal reflective access by org.apache.hadoop.security.authentication.util.KerberosUtil (file:/home/chosun/hadoop-1.2.1/
hadoop-core-1.2.1.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
9E      Pinnacle Airlines Inc.     2008,12,26,5,1418,1250,1549,1432,9E,5860,91819E,91,102,66,77,88,ATL,PIT,526,3,22,0,,0,0,0,77,0,0
9E      Pinnacle Airlines Inc.     2008,12,28,7,1408,1250,1541,1432,9E,5860,96049E,93,102,65,69,78,ATL,PIT,526,8,20,0,,0,69,0,0,0,0
9E      Pinnacle Airlines Inc.2008,12,29,1,1247,1250,1418,1432,9E,5860,96019E,91,102,72,-14,-3,ATL,PIT,526,6,13,0,,0,NA,NA,NA,NA,NA
9E      Pinnacle Airlines Inc.2008,12,30,2,1245,1250,1419,1432,9E,5860,91669E,94,102,69,-13,-5,ATL,PIT,526,5,20,0,,0,NA,NA,NA,NA,NA
9E      Pinnacle Airlines Inc.     2008,12,3,3,1453,1440,1547,1530,9E,5860,91819E,54,50,30,17,13,ATL,TYS,152,2,22,0,,0,0,0,4,0,13
9E      Pinnacle Airlines Inc.     2008,12,5,5,1448,1440,1604,1530,9E,5860,91479E,76,50,29,34,8,ATL,TYS,152,5,42,0,,0,0,0,26,0,8
```

9E   Pinnacle Airlines Inc.   2008,12,1,1,1156,1049,1511,1355,9E,5860,96049E,135,126,109,76,67,AUS,ATL,813,9,17,0,,0,0,0,9,0,67
9E   Pinnacle Airlines Inc.   2008,12,2,2,1115,1049,1431,1355,9E,5860,91879E,136,126,117,36,26,AUS,ATL,813,9,10,0,,0,0,0,10,0,26
9E   Pinnacle Airlines Inc.   2008,12,3,3,1054,1049,1413,1355,9E,5860,91819E,139,126,110,18,5,AUS,ATL,813,15,14,0,,0,0,0,18,0,0
9E   Pinnacle Airlines Inc.   2008,12,4,4,1057,1049,1426,1355,9E,5860,96059E,149,126,120,31,8,AUS,ATL,813,12,17,0,,0,8,0,23,0,0
cat: Unable to write to output stream.
chosun@chosun-VirtualBox:~/hadoop$
/// *** ///
>>> 실행중인 하둡 데몬 종료합니다.
/// *** ///
chosun@chosun-VirtualBox:~/hadoop$ stop-all.sh
stopping jobtracker
localhost: stopping tasktracker
stopping namenode
localhost: stopping datanode
localhost: stopping secondarynamenode
chosun@chosun-VirtualBox:~/hadoop$
chosun@chosun-VirtualBox:~/hadoop$ jps
14467 Jps
4286 org.eclipse.equinox.launcher_1.5.300.v20190213-1655.jar
chosun@chosun-VirtualBox:~/hadoop$
chosun@chosun-VirtualBox:~/hadoop$

/// *** ///
[정리]
조인을 단독으로 처리하는 맵 사이드 방식에 비해 리듀스 사이드 방식이 번거로워 보이지만,
조인할 데이터의 규모가 클 경우 리듀스 사이드 방식으로 조인을 수행해야 합니다.
/// *** ///