

Cluster Setting

컴퓨터공학과

2017103948 고다경

1. upload the arbitrary file to HDFS snap.stanford.edu/data

(wget <https://snap.stanford.edu/data/soc-LiveJournal1.txt.gz>)

- (1) wget <https://snap.stanford.edu/data/soc-LiveJournal1.txt.gz>
- (2) gunzip soc-LiveJournal1.txt.gz
-> soc-LiveJournal1.txt 로 바뀜

```
hadoop@master:~/hadoop-2.7.6 $ ls
LICENSE.txt  README.txt  etc        lib        logs  share          tmp
NOTICE.txt   bin         include    libexec    sbin  soc-LiveJournal1.txt
hadoop@master:~/hadoop-2.7.6 $ ./bin/hadoop fs -put soc-LiveJournal1.txt input.txt
```

soc-LiveJournal1.txt 를 input.txt 란 이름으로 put.

2. hadoop jar share/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.5.jar wordcount input output

- (1) ./bin/hadoop jar share/Hadoop/mapreduce/hadoop-mapreduce-example-2.7.6.jar wordcount input.txt output-soc

Capture1

```

[hadoop@master:~/hadoop-2.7.6 $ ./bin/hadoop jar share/hadoop/mapreduce/hadoop-mapreduce-examples-2.7.6.jar wordcount input.txt output-soc
21/05/21 16:49:03 INFO client.RMProxy: Connecting to ResourceManager at master/192.168.0.17:8032
21/05/21 16:49:05 INFO input.FileInputFormat: Total input paths to process : 1
21/05/21 16:49:05 INFO mapreduce.JobSubmitter: number of splits:8
21/05/21 16:49:05 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1621566532759_0002
21/05/21 16:49:05 INFO impl.YarnClientImpl: Submitted application application_1621566532759_0002
21/05/21 16:49:05 INFO mapreduce.Job: The url to track the job: http://master:50077/proxy/application_1621566532759_0002/
21/05/21 16:49:05 INFO mapreduce.Job: Running job: job_1621566532759_0002
21/05/21 16:49:15 INFO mapreduce.Job: Job job_1621566532759_0002 running in uber mode : false
21/05/21 16:49:15 INFO mapreduce.Job: map 0% reduce 0%
21/05/21 16:49:39 INFO mapreduce.Job: map 2% reduce 0%
21/05/21 16:49:41 INFO mapreduce.Job: map 4% reduce 0%
21/05/21 16:49:42 INFO mapreduce.Job: map 6% reduce 0%
21/05/21 16:49:44 INFO mapreduce.Job: map 9% reduce 0%
21/05/21 16:49:45 INFO mapreduce.Job: map 10% reduce 0%
21/05/21 16:49:47 INFO mapreduce.Job: map 13% reduce 0%
21/05/21 16:49:50 INFO mapreduce.Job: map 14% reduce 0%
21/05/21 16:50:03 INFO mapreduce.Job: map 16% reduce 0%
21/05/21 16:50:06 INFO mapreduce.Job: map 18% reduce 0%
21/05/21 16:50:09 INFO mapreduce.Job: map 19% reduce 0%
21/05/21 16:50:21 INFO mapreduce.Job: map 20% reduce 0%
21/05/21 16:50:22 INFO mapreduce.Job: map 21% reduce 0%
21/05/21 16:50:25 INFO mapreduce.Job: map 25% reduce 0%
21/05/21 16:50:28 INFO mapreduce.Job: map 29% reduce 0%
21/05/21 16:50:31 INFO mapreduce.Job: map 30% reduce 0%
21/05/21 16:50:43 INFO mapreduce.Job: map 31% reduce 0%
21/05/21 16:50:46 INFO mapreduce.Job: map 33% reduce 0%
21/05/21 16:50:49 INFO mapreduce.Job: map 35% reduce 0%
21/05/21 16:50:52 INFO mapreduce.Job: map 37% reduce 0%
21/05/21 16:50:55 INFO mapreduce.Job: map 39% reduce 0%
21/05/21 16:50:56 INFO mapreduce.Job: map 40% reduce 0%
21/05/21 16:50:59 INFO mapreduce.Job: map 41% reduce 0%
21/05/21 16:51:02 INFO mapreduce.Job: map 42% reduce 0%
21/05/21 16:51:05 INFO mapreduce.Job: map 43% reduce 0%
21/05/21 16:51:08 INFO mapreduce.Job: map 46% reduce 0%
21/05/21 16:51:11 INFO mapreduce.Job: map 47% reduce 0%
21/05/21 16:51:20 INFO mapreduce.Job: map 48% reduce 0%
21/05/21 16:51:23 INFO mapreduce.Job: map 50% reduce 0%
21/05/21 16:51:24 INFO mapreduce.Job: map 51% reduce 0%
21/05/21 16:51:26 INFO mapreduce.Job: map 53% reduce 0%
21/05/21 16:51:29 INFO mapreduce.Job: map 55% reduce 0%
21/05/21 16:51:32 INFO mapreduce.Job: map 56% reduce 0%
21/05/21 16:51:36 INFO mapreduce.Job: map 58% reduce 0%
21/05/21 16:51:39 INFO mapreduce.Job: map 60% reduce 0%
21/05/21 16:51:41 INFO mapreduce.Job: map 61% reduce 0%
21/05/21 16:51:42 INFO mapreduce.Job: map 62% reduce 0%
21/05/21 16:51:45 INFO mapreduce.Job: map 67% reduce 0%
21/05/21 16:51:46 INFO mapreduce.Job: map 69% reduce 0%
21/05/21 16:51:46 INFO mapreduce.Job: map 69% reduce 0%
21/05/21 16:51:55 INFO mapreduce.Job: map 70% reduce 0%
21/05/21 16:51:58 INFO mapreduce.Job: map 71% reduce 0%
21/05/21 16:52:01 INFO mapreduce.Job: map 72% reduce 0%
21/05/21 16:52:02 INFO mapreduce.Job: map 73% reduce 0%
21/05/21 16:52:05 INFO mapreduce.Job: map 74% reduce 0%
21/05/21 16:52:08 INFO mapreduce.Job: map 75% reduce 0%
21/05/21 16:52:11 INFO mapreduce.Job: map 76% reduce 0%
21/05/21 16:52:13 INFO mapreduce.Job: map 76% reduce 13%
21/05/21 16:52:16 INFO mapreduce.Job: map 77% reduce 13%
21/05/21 16:52:19 INFO mapreduce.Job: map 81% reduce 13%
21/05/21 16:52:22 INFO mapreduce.Job: map 81% reduce 17%
21/05/21 16:52:29 INFO mapreduce.Job: map 82% reduce 17%
21/05/21 16:52:35 INFO mapreduce.Job: map 83% reduce 17%
21/05/21 16:52:43 INFO mapreduce.Job: map 86% reduce 17%
21/05/21 16:52:45 INFO mapreduce.Job: map 89% reduce 17%
21/05/21 16:52:47 INFO mapreduce.Job: map 89% reduce 21%
21/05/21 16:52:48 INFO mapreduce.Job: map 93% reduce 21%
21/05/21 16:52:50 INFO mapreduce.Job: map 94% reduce 21%
21/05/21 16:52:51 INFO mapreduce.Job: map 98% reduce 21%
21/05/21 16:52:52 INFO mapreduce.Job: map 100% reduce 21%
21/05/21 16:52:53 INFO mapreduce.Job: map 100% reduce 36%
21/05/21 16:52:56 INFO mapreduce.Job: map 100% reduce 70%
21/05/21 16:52:59 INFO mapreduce.Job: map 100% reduce 83%
21/05/21 16:53:02 INFO mapreduce.Job: map 100% reduce 94%
21/05/21 16:53:03 INFO mapreduce.Job: map 100% reduce 100%
21/05/21 16:53:03 INFO mapreduce.Job: Job job_1621566532759_0002 completed successfully
21/05/21 16:53:03 INFO mapreduce.Job: Counters: 51
File System Counters
  FILE: Number of bytes read=629910501
  FILE: Number of bytes written=857382962
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
  HDFS: Number of bytes read=1080627554
  HDFS: Number of bytes written=49941834
  HDFS: Number of read operations=27
  HDFS: Number of large read operations=0
  HDFS: Number of write operations=2
Job Counters
  Killed map tasks=3
  Launched map tasks=11
  Launched reduce tasks=1
  Data-local map tasks=9
  Rack-local map tasks=2
  Total time spent by all maps in occupied slots (ms)=1580332
  Total time spent by all reduces in occupied slots (ms)=80082
  Total time spent by all map tasks (ms)=1580332
  Total time spent by all reduce tasks (ms)=80082
  Total vcore-milliseconds taken by all map tasks=1580332
  Total vcore-milliseconds taken by all reduce tasks=80082
  Total megabyte-milliseconds taken by all map tasks=1618259968
  Total megabyte-milliseconds taken by all reduce tasks=82083968
Map-Reduce Framework

```

Capture3

```

Map-Reduce Framework
  Map input records=68993777
  Map output records=137987572
  Map output bytes=1563554552
  Map output materialized bytes=227266798
  Input split bytes=840
  Combine input records=167666175
  Combine output records=46329993
  Reduce input groups=4847593
  Reduce shuffle bytes=227266798
  Reduce input records=16651390
  Reduce output records=4847593
  Spilled Records=62981383
  Shuffled Maps =8
  Failed Shuffles=0
  Merged Map outputs=8
  GC time elapsed (ms)=15897
  CPU time spent (ms)=330400
  Physical memory (bytes) snapshot=1947459584
  Virtual memory (bytes) snapshot=17392619520
  Total committed heap usage (bytes)=1237962752
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=1080626714
File Output Format Counters
  Bytes Written=49941834
[hadoop@master:~/hadoop-2.7.6 $ ]

```

