Run successfully

启动MapReduce服务

查看进程,验证是否成功启动服务

```
syx@syx-OptiPlex-7050:~/hadoop-1.2.1$ ./bin/start-mapred.sh
starting jobtracker, logging to /home/syx/hadoop-1.2.1/libexec/../logs/hadoop-sy
x-jobtracker-syx-OptiPlex-7050.out
localhost: starting tasktracker, logging to /home/syx/hadoop-1.2.1/libexec/../lo
gs/hadoop-syx-tasktracker-syx-OptiPlex-7050.out
syx@syx-OptiPlex-7050:~/hadoop-1.2.1$ jps
10784 Jps
9904 SecondaryNameNode
9476 NameNode
10697 TaskTracker
10474 JobTracker
9690 DataNode
```

提交MapReduce应用程序

提交jar命令并查看运行结果 运行grep示例:

```
-OptiPlex-7050:~/hadoop-1.2.1$ ./bin/hadoop jar hadoop-examples-1.2.1.jar grep /user/syx/input/ /user/syx/output/grep
 'dfs[a-z.]+'
19/09/26 19:32:40 INFO util.NativeCodeLoader: Loaded the native-hadoop library
19/09/26 19:32:40 WARN snappy.LoadSnappy: Snappy native library not loaded
19/09/26 19:32:40 INFO mapred.FileInputFormat: Total input paths to process:
Map output materialized bytes=242
                                                                  Virtual memory (bytes) snapshot=35139354624
Map input records=969
SPLIT_RAW_BYTES=1817
                                                                  Reduce shuffle bytes=242
Physical memory (bytes) snapshot=3662336000
Map input bytes=34596
                                                                  CPU time spent (ms)=4590
Total committed heap usage (bytes)=3146252288
 19/09/26 19:32:56 INFO mapred.JobClient:
 19/09/26 19:32:56 INFO mapred.JobClient:
                                                               File Input Format Counters
 19/09/26 19:32:56 INFO mapred.JobClient:
19/09/26 19:32:56 INFO mapred.JobClient:
19/09/26 19:32:56 INFO mapred.JobClient:
                                                                  Bytes Read=34596
                                                               FileSystemCounters
HDFS_BYTES_READ=36413
              19:32:56
                           INFO mapred.JobClient:
                                                                  FILE BYTES WRITTEN=1003316
```

运行 wordcount 示例, 并且查看系统执行该任务过程中启动的进程:

```
<mark>syx@syx-OptiPlex-7050:~/hadoop-1.2.1</mark>$ ./bin/hadoop jar hadoop-examples-1.2.1.jar wordcount /user/syx/ir
put/test2.8G.text /user/syx/output/wordcount
19/09/26 19:44:39 INFO input.FileInputFormat: Total input paths to process : 1
19/09/26 19:44:39 INFO util.NativeCodeLoader: Loaded the native-hadoop library
19/09/26 19:44:39 WARN snappy.LoadSnappy: Snappy native library not loaded
19/09/26 19:44:39 INFO mapred.JobClient: Running job: job_201909261834_0021
19/09/26 19:44:40 INFO mapred.JobClient: map 0% reduce 0%
19/09/26 19:44:46 INFO mapred.JobClient:
                                             map 4% reduce 0%
19/09/26 19:44:50 INFO mapred.JobClient:
                                             map 6% reduce 0%
19/09/26 19:44:51 INFO mapred.JobClient:
                                             map 9% reduce 0%
19/09/26 19:44:54 INFO mapred.JobClient:
                                             map 11% reduce 0%
19/09/26 19:44:55 INFO mapred.JobClient:
                                             map 13% reduce 0%
19/09/26 19:45:01 INFO mapred.JobClient:
                                             map 15% reduce 0%
19/09/26 19:45:02 INFO mapred.JobClient:
                                             map 16% reduce 0%
19/09/26 19:45:03 INFO mapred.JobClient:
                                             map 16% reduce 4%
19/09/26 19:45:05 INFO mapred.JobClient:
                                             map 17% reduce 4%
19/09/26 19:45:10 INFO mapred.JobClient:
                                             map 18% reduce 4%
                                             map 20% reduce 4%
19/09/26 19:45:11 INFO mapred.JobClient:
19/09/26 19:45:12 INFO mapred.JobClient:
                                             map 23% reduce 4%
19/09/26 19:45:13 INFO mapred.JobClient:
                                             map 25% reduce 4%
19/09/26 19:45:15 INFO mapred.JobClient:
                                             map 25% reduce 8%
19/09/26 19:45:16 INFO mapred.JobClient:
                                             map 27% reduce 8%
19/09/26 19:45:20 INFO mapred.JobClient:
19/09/26 19:45:22 INFO mapred.JobClient:
                                             map 29% reduce 8%
                                             map 30% reduce 8%
19/09/26 19:45:23 INFO mapred.JobClient:
                                             map 32% reduce 8%
19/09/26 19:45:24 INFO mapred.JobClient:
                                             map 32% reduce 9%
19/09/26 19:45:30 INFO mapred.JobClient:
                                             map 34% reduce 9%
19/09/26 19:45:31 INFO mapred.JobClient:
                                             map 34% reduce 10%
19/09/26 19:45:32 INFO mapred.JobClient:
                                             map 37% reduce 10%
19/09/26 19:45:33 INFO mapred.JobClient:
                                             map 39% reduce 10%
19/09/26 19:45:34 INFO mapred.JobClient:
                                             map 39% reduce 12%
19/09/26 19:45:38 INFO mapred.JobClient:
                                             map 41% reduce 12%
19/09/26 19:45:39 INFO mapred.JobClient:
                                             map 43% reduce 12%
19/09/26 19:45:42 INFO mapred.JobClient:
                                             map 44% reduce 13%
19/09/26 19:45:43 INFO mapred.JobClient:
                                             map 46% reduce 13%
19/09/26 19:45:46 INFO mapred.JobClient:
                                             map 48% reduce 13%
19/09/26 19:45:48 INFO mapred.JobClient:
                                             map 51% reduce 16%
                                             map 53% reduce 16%
19/09/26 19:45:53 INFO mapred.JobClient:
19/09/26 19:45:55 INFO mapred.JobClient:
                                             map 54% reduce 17%
19/09/26 19:46:01 INFO mapred.JobClient:
                                             map 55% reduce 17%
19/09/26 19:46:04 INFO mapred.JobClient: map 56% reduce 17%
```

```
9/09/26 19:48:55 INFO mapred.JobClient:
                                          map 100% reduce 95%
19/09/26 19:48:58 INFO mapred.JobClient:
                                         map 100% reduce 96%
.9/09/26 19:49:01 INFO mapred.JobClient: map 100% reduce 98%
9/09/26 19:49:05 INFO mapred.JobClient:
                                         map 100% reduce 99%
                                         map 100% reduce 100%
19/09/26 19:49:06 INFO mapred.JobClient:
9/09/26 19:49:06 INFO mapred.JobClient: Counters: 29
.9/09/26 19:49:06 INFO mapred.JobClient:
                                           Map-Reduce Framework
                                             Spilled Records=51673334
9/09/26 19:49:06 INFO mapred.JobClient:
9/09/26 19:49:06 INFO mapred.JobClient:
                                             Map output materialized bytes=2954567651
9/09/26 19:49:06 INFO mapred.JobClient:
                                             Reduce input records=13448299
.9/09/26 19:49:06 INFO mapred.JobClient:
                                             Virtual memory (bytes) snapshot=86031462400
9/09/26 19:49:06 INFO mapred.JobClient:
                                             Map input records=5011909
9/09/26 19:49:06 INFO mapred.JobClient:
                                             SPLIT RAW BYTES=4945
9/09/26 19:49:06 INFO mapred.JobClient:
                                             Map output bytes=2938284221
9/09/26 19:49:06 INFO mapred.JobClient:
                                             Reduce shuffle bytes=2954567651
9/09/26 19:49:06 INFO mapred.JobClient:
                                             Physical memory (bytes) snapshot=10760355840
.9/09/26 19:49:06 INFO mapred.JobClient:
.9/09/26 19:49:06 INFO mapred.JobClient:
                                             Reduce input groups=10105942
                                             Combine output records=13448299
9/09/26 19:49:06 INFO mapred.JobClient:
                                             Reduce output records=10105942
9/09/26 19:49:06 INFO mapred.JobClient:
                                             Map output records=15035725
9/09/26 19:49:06 INFO mapred.JobClient:
                                             Combine input records=15035725
9/09/26 19:49:06 INFO mapred.JobClient:
                                             CPU time spent (ms)=246080
.9/09/26 19:49:06 INFO mapred.JobClient:
.9/09/26 19:49:06 INFO mapred.JobClient:
                                             Total committed heap usage (bytes)=8610906112
                                           File Input Format Counters
9/09/26 19:49:06 INFO mapred.JobClient:
                                             Bytes Read=2868289536
9/09/26 19:49:06 INFO mapred.JobClient:
                                           FileSystemCounters
.9/09/26 19:49:06 INFO mapred.JobClient:
                                             HDFS_BYTES_READ=2868294481
.9/09/26 19:49:06 INFO mapred.JobClient:
                                             FILE_BYTES_WRITTEN=11354922200
                                             FILE_BYTES_READ=8397904355
HDFS_BYTES_WRITTEN=2844000718
9/09/26 19:49:06 INFO mapred.JobClient:
.9/09/26 19:49:06 INFO mapred.JobClient:
.9/09/26 19:49:06 INFO mapred.JobClient:
                                           Job Counters
                                             Launched map tasks=43
9/09/26 19:49:06 INFO mapred.JobClient:
9/09/26 19:49:06 INFO mapred.JobClient:
                                             Launched reduce tasks=1
9/09/26 19:49:06 INFO mapred.JobClient:
                                             SLOTS_MILLIS_REDUCES=256010
19/09/26 19:49:06 INFO mapred.JobClient:
                                             Total time spent by all reduces waiting after reserving sl
ts (ms)=0
19/09/26 19:49:06 INFO mapred.JobClient:
                                             SLOTS_MILLIS_MAPS=320169
9/09/26 19:49:06 INFO mapred.JobClient:
                                             Total time spent by all maps waiting after reserving slots
(ms)=0
.9/09/26 19:49:06 INFO mapred.JobClient:
                                             Data-local map tasks=43
19/09/26 19:49:06 INFO mapred.JobClient:
                                           File Output Format Counters
9/09/26 19:49:06 INFO mapred.JobClient:
                                             Bytes Written=2844000718
syx@syx-OptiPlex-7050:~/hadoop-1.2.1$
```

jps:

```
syx@syx-OptiPlex-7050:~$ jps
9904 SecondaryNameNode
24963 Child
9476 NameNode
10697 TaskTracker
10474 JobTracker
27626 Jps
9690 DataNode
24574 RunJar
syx@syx-OptiPlex-7050:~$
```

启动MapReduce服务

- 启动YARN命令
- 开启历史服务器
- 启动HDFS服务

```
syx@syx-OptiPlex-7050:~/hadoop-2.9.2$ ./sbin/start-yarn.sh
                                                                # 启动YARN
starting yarn daemons
starting resourcemanager, logging to /home/syx/hadoop-2.9.2/logs/yarn-syx-resour
cemanager-syx-OptiPlex-7050.out
localhost: starting nodemanager, logging to /home/syx/hadoop-2.9.2/logs/yarn-syx
-nodemanager-syx-OptiPlex-7050.out
syx@syx-OptiPlex-7050:~/hadoop-2.9.2$ ./sbin/mr-jobhistory-daemon.sh start histo
ryserver
starting historyserver, logging to /home/syx/hadoop-2.9.2/logs/mapred-syx-histor
yserver-syx-OptiPlex-7050.out
syx@syx-OptiPlex-7050:~/hadoop-2.9.2$ ./sbin/start-dfs.sh
Starting namenodes on [localhost]
localhost: starting namenode, logging to /home/syx/hadoop-2.9.2/logs/hadoop-syx-
namenode-syx-OptiPlex-7050.out
localhost: starting datanode, logging to /home/syx/hadoop-2.9.2/logs/hadoop-syx-
datanode-syx-OptiPlex-7050.out
Starting secondary namenodes [0.0.0.0]
0.0.0.0: starting secondarynamenode, logging to /home/syx/hadoop-2.9.2/logs/hado
op-syx-secondarynamenode-syx-OptiPlex-7050.out
syx@syx-OptiPlex-7050:~/hadoop-2.9.2$ jps
8304 NodeManager
8932 DataNode
7942 ResourceManager
9322 Jps
8698 NameNode
8476 JobHistoryServer
9181 SecondaryNameNode
syx@syx-OptiPlex-7050:~/hadoop-2.9.2$
运行MapReduce应用程序
```

o 提交jar命令并查看运行结果运行grep示例,结果如下:

```
syx@syx-OptiPlex-7050:~/hadoop-2.9.2$ ./bin/yarn jar ./share/hadoop/mapreduce/ha
doop-mapreduce-examples-2.9.2.jar grep /user/syx/input/ /user/syx/output/grep 'd
fs[a-z.]+'
19/09/26 20:33:38 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0
:8032
19/09/26 20:33:38 INFO input.FileInputFormat: Total input files to process : 0
19/09/26 20:33:38 INFO mapreduce.JobSubmitter: number of splits:0
19/09/26 20:33:38 INFO Configuration.deprecation: yarn.resourcemanager.system-me
trics-publisher.enabled is deprecated. Instead, use yarn.system-metrics-publishe
r.enabled
19/09/26 20:33:38 INFO mapreduce.JobSubmitter: Submitting tokens for job: job 15
69500919351 0002
19/09/26 20:33:39 INFO impl.YarnClientImpl: Submitted application application 15
69500919351 0002
19/09/26 20:33:39 INFO mapreduce.Job: The url to track the job: http://syx-OptiP
lex-7050:8088/proxy/application 1569500919351 0002/
19/09/26 20:33:39 INFO mapreduce.Job: Running job: job_1569500919351_0002
19/09/26 20:33:44 INFO mapreduce.Job: Job job 1569500919351 0002 running in uber
mode : false
19/09/26 20:33:44 INFO mapreduce.Job: map 0% reduce 0%
19/09/26 20:33:49 INFO mapreduce.Job: map 0% reduce 100%
19/09/26 20:33:49 INFO mapreduce.Job: Job job 1569500919351 0002 completed succe
ssfully
19/09/26 20:33:49 INFO mapreduce.Job: Counters: 38
        File System Counters
                 FILE: Number of bytes read=0
                 FILE: Number of bytes written=198785
                 FILE: Number of read operations=0
                 FILE: Number of large read operations=0
                 FILE: Number of write operations=0 HDFS: Number of bytes read=0
                 HDFS: Number of bytes written=86
                 HDFS: Number of read operations=3
                 HDFS: Number of large read operations=0
                 HDFS: Number of write operations=2
        Job Counters
                 Launched reduce tasks=1
                 Total time spent by all maps in occupied slots (ms)=0
                 Total time spent by all reduces in occupied slots (ms)=1459
                 Total time spent by all reduce tasks (ms)=1459
                 Total vcore-milliseconds taken by all reduce tasks=1459
                 Total megabyte-milliseconds taken by all reduce tasks=1494016
```

```
Total time spent by all reduce tasks (ms)=1459
                Total vcore-milliseconds taken by all reduce tasks=1459
                Total megabyte-milliseconds taken by all reduce tasks=1494016
        Map-Reduce Framework
                Combine input records=0
                Combine output records=0
                Reduce input groups=0
                Reduce shuffle bytes=0
                Reduce input records=0
                Reduce output records=0
                Spilled Records=0
                Shuffled Maps =0
                Failed Shuffles=0
                Merged Map outputs=0
                GC time elapsed (ms)=34
                CPU time spent (ms)=280
                Physical memory (bytes) snapshot=201277440
                Virtual memory (bytes) snapshot=2015866880
                Total committed heap usage (bytes)=124256256
        Shuffle Errors
                BAD ID=0
                CONNECTION=0
                IO ERROR=0
                WRONG_LENGTH=0
WRONG_MAP=0
WRONG_REDUCE=0
        File Output Format Counters
                Bytes Written=86
19/09/26 20:33:49 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0
:8032
19/09/26 20:33:49 INFO input.FileInputFormat: Total input files to process : 1
19/09/26 20:33:49 INFO mapreduce.JobSubmitter: number of splits:1
19/09/26 20:33:50 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_15
69500919351 0003
19/09/26 20:33:50 INFO impl.YarnClientImpl: Submitted application application_15
69500919351 0003
19/09/26 20:33:50 INFO mapreduce.Job: The url to track the job: http://syx-OptiP
lex-7050:8088/proxy/application 1569500919351 0003/
19/09/26 20:33:50 INFO mapreduce.Job: Running job: job_1569500919351_0003
19/09/26 20:33:59 INFO mapreduce.Job: Job job_1569500919351_0003 running in uber
mode : false
19/09/26 20:33:59 INFO mapreduce.Job: map 0% reduce 0%
19/09/26 20:34:03 INFO mapreduce.Job: map 100% reduce 0%
```

```
Launched reduce tasks=1
                Data-local map tasks=1
                Total time spent by all maps in occupied slots (ms)=1347
                Total time spent by all reduces in occupied slots (ms)=1472
                Total time spent by all map tasks (ms)=1347
                Total time spent by all reduce tasks (ms)=1472
                Total vcore-milliseconds taken by all map tasks=1347
                Total vcore-milliseconds taken by all reduce tasks=1472
                Total megabyte-milliseconds taken by all map tasks=1379328
                Total megabyte-milliseconds taken by all reduce tasks=1507328
        Map-Reduce Framework
                Map input records=0
                Map output records=0
                Map output bytes=0
                Map output materialized bytes=6
                Input split bytes=128
                Combine input records=0
                Combine output records=0
                Reduce input groups=0
                Reduce shuffle bytes=6
                Reduce input records=0
                Reduce output records=0
                Spilled Records=0
                Shuffled Maps =1
                Failed Shuffles=0
                Merged Map outputs=1
                GC time elapsed (ms)=85
                CPU time spent (ms)=670
                Physical memory (bytes) snapshot=512491520
                Virtual memory (bytes) snapshot=4024590336
                Total committed heap usage (bytes)=330301440
        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=86
        File Output Format Counters
                Bytes Written=0
syx@syx-OptiPlex-7050:~/hadoop-2.9.2$
```

运行 wordcount 示例

```
syx@syx-OptiPlex-7050:~/hadoop-2.9.2$ ./bin/yarn jar ./share/hadoop/mapreduce/ha
doop-mapreduce-examples-2.9.2.jar wordcount /user/syx/input/test2.8G.text /user/
svx/output/wordcount
19/09/26 20:40:04 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0
:8032
19/09/26 20:40:04 INFO input.FileInputFormat: Total input files to process : 1
19/09/26 20:40:04 INFO mapreduce.JobSubmitter: number of splits:22
19/09/26 20:40:05 INFO Configuration.deprecation: yarn.resourcemanager.system-me
trics-publisher.enabled is deprecated. Instead, use yarn.system-metrics-publishe
r.enabled
19/09/26 20:40:05 INFO mapreduce.JobSubmitter: Submitting tokens for job: job 15
69500919351 0005
19/09/26 20:40:05 INFO impl.YarnClientImpl: Submitted application application 15
69500919351 0005
19/09/26 20:40:05 INFO mapreduce.Job: The url to track the job: http://syx-OptiP
lex-7050:8088/proxy/application_1569500919351_0005/
19/09/26 20:40:05 INFO mapreduce.Job: Running job: job_1569500919351_0005
mode : false
19/09/26 20:40:09 INFO mapreduce.Job: map 0% reduce 0%
19/09/26 20:40:33 INFO mapreduce.Job:
                             map 14% reduce 0%
19/09/26 20:40:39 INFO mapreduce.Job:
                             map 17% reduce 0%
19/09/26 20:40:45 INFO mapreduce.Job:
                             map 18% reduce 0%
19/09/26 20:40:51 INFO mapreduce.Job:
                             map 20% reduce 0%
19/09/26 20:40:57 INFO mapreduce.Job:
                             map 22% reduce 0%
19/09/26 20:41:03 INFO mapreduce.Job:
                             map 23% reduce 0%
19/09/26 20:41:09 INFO mapreduce.Job:
                             map 25% reduce 0%
19/09/26 20:41:15 INFO mapreduce.Job:
                             map 26% reduce 0%
19/09/26 20:41:20 INFO mapreduce.Job:
                             map 27% reduce 0%
19/09/26 20:41:50 INFO mapreduce.Job:
                             map 36% reduce 0%
19/09/26 20:42:02 INFO mapreduce.Job: map 41% reduce 0%
19/09/26 20:42:08 INFO mapreduce.Job:
                             map 43% reduce 0%
19/09/26 20:42:33 INFO mapreduce.Job: map 48% reduce 2%
```

```
Launched map tasks=23
                Launched reduce tasks=1
                Data-local map tasks=23
                Total time spent by all maps in occupied slots (ms)=1626180
                Total time spent by all reduces in occupied slots (ms)=406368
                Total time spent by all map tasks (ms)=1626180
                Total time spent by all reduce tasks (ms)=406368
                Total vcore-milliseconds taken by all map tasks=1626180
                Total vcore-milliseconds taken by all reduce tasks=406368
                Total megabyte-milliseconds taken by all map tasks=1665208320
                Total megabyte-milliseconds taken by all reduce tasks=416120832
       Map-Reduce Framework
                Map input records=5011909
                Map output records=15035725
                Map output bytes=2938284221
                Map output materialized bytes=2944210375
                Input split bytes=2530
                Combine input records=15035725
                Combine output records=12757848
                Reduce input groups=10105942
                Reduce shuffle bytes=2944210375
                Reduce input records=12757848
                Reduce output records=10105942
                Spilled Records=46026946
                Shuffled Maps =22
                Failed Shuffles=0
                Merged Map outputs=22
                GC time elapsed (ms)=8128
                CPU time spent (ms)=274170
                Physical memory (bytes) snapshot=6686240768
                Virtual memory (bytes) snapshot=46184607744
                Total committed heap usage (bytes)=4518313984
       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=2868203520
        File Output Format Counters
                Bytes Written=2844000718
syx@syx-OptiPlex-7050:~/hadoop-2.9.2$
```

Mistakes

```
1. FS
    syx@syx-OptiPlex-7050:~/hadoop-1.2.1$ ./bin/hadoop fs -put ./conf /user/syx/inpu
t/
    put: Target /user/syx/input/conf is a directory
    solu:
    syx@syx-OptiPlex-7050:~/hadoop-1.2.1$ ./bin/hadoop fs -put ./conf /user/syx/inpu
t/11
    hadoop's FS is different from linux's FS. I have tried
```

```
sudo rm -rf /user/syx/input
```

but this "input" is not the one in the hadoop FS.