

Administrator Guideline for Map-Reduce Facility

Zichang Feng, Siyuan Zhou

zfeng, siyuanz

System Requirements:

JVM 1.7 or above

AFS similar file system to store jar file

Configurations:

There are two files to configure: conf/dfs.cnf and conf/mapred.cnf.

conf/dfs.cnf:

```
fs.default.name=dfs//<name node host name>:<name node port>
dfs.name.dir=<name node image file directory>
dfs.data.dir=<data node block directory>
dfs.block.size=<maximum block size in bytes>
dfs.replication=<replication factor>
```

conf/mapred.cnf:

```
jobtracker.port=<job tracker port>
tasktracker.tasks.maximum=<max simultaneous task on a task
tracker>
tasktracker.heartbeat.interval=<period of heartbeat in
milliseconds>
mapred.tmp.dir=<temporary directory for map reduce tasks>
fs.default.name=<name node host name>:<name node port>
```

To Compile the System:

```
cd src  
make
```

To Start the System:

```
cd bin  
./start-all.sh
```

To Confirm the System is Running

```
jps
```

There should be four java program running: Namenode, Datanode, JobTracker and TaskTracker.

To Submit a Job:

Submit by jar file directly through command line by

```
java -cp . mapredClient.Client -jar <jar file path> -input  
<input path in DFS> -output <output path in DFS> -reducenum  
<number of reduce tasks> -delim <field delimiter>
```

Or write a java program that calls our API. Please refer to our API manual for more information.

To Stop the System:

```
./stop-all.sh
```