**System Administrator Document**

We need a system configuration files for the system(which in our example is default.conf). In the configuration, we know where the master is, where the RMI server for Namenode and MasterNode is, and how many mappers per host, how many reducers per host, replica factor.

For the system administrator, we should start a master node(Name node is also there), several task nodes and several data nodes before the user use JobClient to submit job.

If each part of the system get correctly connected, the master node will get every information.

The master and task node will show information on the machine in real time.

When you want to shut down the system ,just tape “ctrl”+”c” in each system.

After that ,just wait for the client to submit jobs.

In bin folder, type “java ha.mapreduce.MasterNode” + system configuration file will start the master and name nodes.

For task nodes, type java ha.mapreduce.TaskNodes + system configuration file + your host+ your port.

For data nodes, type java ha.mapreduce.DataNodes + system configuration file + your host+ your port.

For clients, you should cd to root folder and type in : java –classpath bin ha.mapreduce.JobClient client.conf host:port.

Just a FYI that all host:ports typed in when starting task nodes, data nodes and clients can be random but should be the full path and real host. For example, you can start a task node in ghc70:8000 so you will type in

java ha.mapreduce.TaskNodes ../sample.conf [ghc70@ghc.andrew.cmu.edu:8000](mailto:ghc70@ghc.andrew.cmu.edu:8000).

For detail use, please see the last paragraph of reports.