Single Node Hadoop Cluster Creation

Download Hadoop and Java

$ tar -zxvf hadoop-1.2.1.tar.gz   **(extract the tar file)**

$ sh jdk-6u45-linux-x64.bin   **(unpack the bin file)**

$ sudo apt-get install vim   **(vi editer install)**

$ vi .bashrc   **(set the java home path)**

export JAVA_HOME=/home/students/jdk1.6.0_38

$ source .bashrc   **(execute the bashrc file)**

$ echo $JAVA_HOME   **(check the java path)**

 (Modify Hadoop Configuration Files)

NAMENODE	---->	core-site.xml
JOBTRACKER	---->	mapperd-site.xml
SECONDARYNAMENODE	---->	mastors
DATANODE	---->	
TASKTRECKER	---->	slaves

Ports used by Hadoop Daemons

Daemon	RPC	WEB
NameNode	8020 (50000*)	50070
SecondaryNameNode		50090
JobTracker	8021(50001*)	50030
TaskTracker	50020	50060
DataNode	50010	50075

$ vi hadoop-env.sh

export JAVA_HOME=/home/students/jdk1.6.0_38

$ vi core-site.xml/

```
<property>
<name>fs.default.name</name>
<value>hdfs://localhost:50000</value>
</property>
```

$ vi mapred-site.xml

```
<property>
<name>mapred.job.tracker</name>
<value>hdfs://localhost:50001</value>
</property>
```

$ vi hdfs-site.xml

```
<property>
<name>dfs.name.dir</name>
<value>/home/students/hadoop-dir/name-dir</value>
</property>
<property>
<name>dfs.data.dir</name>
<value>/home/students/hadoop-dir/data-dir</value>
</property>
<property>
<name>dfs.permissions</name>
<value>false</value>
</property>
```

$ vi masters

localhost

$ vi slaves

localhost

Install the ssh key

(generates, manages and converts authentication keys)

$ ssh-keygen -t rsa

(Setup passwordless ssh to localhost and to slaves)

$ sudo apt-get install openssh-server

$ cd .ssh

$ ls

$ cat id_rsa.pub >> authorized_keys   (copy the .pub)

(Copy the id_rsa.pub from NameNode to authorized_keys in all machines)

$ ssh localhost

(no password)

$ cd hadoop-1.2.1/

$ bin/hadoop namenode -format   **(Format Hadoop NameNode)**

$ bin/start-all.sh   **(Start All Hadoop Related Services)**

$ jps   **(java prroces status)**

             (java 5 hadoop daemons display)

             (Browse NameNode and JobTracker Web GUI)

             NameNode - http://localhost:50070/

             JobTracker - http://localhost:50030/

$ bin/stop-all.sh   **(Stop All Hadoop Related Services)**