

**Multi Node Installation**

sudo apt-get update

----------\*\*upload .pem key\*\*----------

windows : (open cmd)

pscp -i a.ppk a.pem ubuntu@52.18.223.48:/home/ubuntu/.ssh

linux : (terminal)

scp -i a.pem a.pem ubuntu@52.18.223.48:/home/ubuntu/.ssh

chmod 400 /home/ubuntu/.ssh/a.pem

-------\*\*configuring .profile\*\*-----------

nano ~/.profile

eval `ssh-agent` ssh-add /home/ubuntu/.ssh/a.pem

source ~/.profile

-------\*\*adding hosts\*\*----

sudo nano /etc/hosts

172.31.23.8 ip-172-31-23-8.eu-west-1.compute.internal nn

172.31.23.7 ip-172-31-23-7.eu-west-1.compute.internal snn

172.31.23.10 ip-172-31-23-10.eu-west-1.compute.internal dn1

172.31.23.9 ip-172-31-23-9.eu-west-1.compute.internal dn2

ssh snn

sudo nano /etc/hosts

172.31.23.8 ip-172-31-23-8.eu-west-1.compute.internal nn

172.31.23.7 ip-172-31-23-7.eu-west-1.compute.internal snn

172.31.23.10 ip-172-31-23-10.eu-west-1.compute.internal dn1

172.31.23.9 ip-172-31-23-9.eu-west-1.compute.internal dn2

exit

ssh dn1

sudo nano /etc/hosts

172.31.23.8 ip-172-31-23-8.eu-west-1.compute.internal nn

172.31.23.7 ip-172-31-23-7.eu-west-1.compute.internal snn

172.31.23.10 ip-172-31-23-10.eu-west-1.compute.internal dn1

172.31.23.9 ip-172-31-23-9.eu-west-1.compute.internal dn2

exit

ssh dn2

sudo nano /etc/hosts

172.31.23.8 ip-172-31-23-8.eu-west-1.compute.internal nn

172.31.23.7 ip-172-31-23-7.eu-west-1.compute.internal snn

172.31.23.10 ip-172-31-23-10.eu-west-1.compute.internal dn1

172.31.23.9 ip-172-31-23-9.eu-west-1.compute.internal dn2

exit

--------\*\*installing and configuring dsh\*\*--------------

sudo apt-get install dsh

cd /etc/dsh

sudo nano machines.list

#localhost

nn

snn

dn1

dn2

---------------------------\*\*configuring bashrc linux env setup\*\*---------------

nano ~/.bashrc

export HADOOP\_PREFIX=/usr/local/hadoop/

export PATH=$PATH:$HADOOP\_PREFIX/bin

export JAVA\_HOME=/usr/lib/jvm/java-7-openjdk-amd64

export PATH=$PATH:$JAVA\_HOME

-----------------\*\*scp .pem .profile & .bashrc to all hosts\*\*------------------

scp .bashrc .profile ubuntu@snn:/home/ubuntu

scp .bashrc .profile ubuntu@dn1:/home/ubuntu

scp .bashrc .profile ubuntu@dn2:/home/ubuntu

scp .ssh/a.pem ubuntu@snn:/home/ubuntu/.ssh

scp .ssh/a.pem ubuntu@dn1:/home/ubuntu/.ssh

scp .ssh/a.pem ubuntu@dn2:/home/ubuntu/.ssh

dsh -a bash

dsh -a source ~/.profile

sudo hostname nn

bash

ssh snn

sudo hostname snn

bash

ssh dn1

sudo hostname dn1

bash

ssh dn2

sudo hostname dn2

bash

ssh nn

-----------------------------\*\*downloading hadoop\*\*------------------------

dsh -a wget http://apache.mirror.gtcomm.net/hadoop/common/hadoop-1.2.1/hadoop-1.2.1.tar.gz

dsh -a tar -xzvf hadoop-1.2.1.tar.gz

dsh -a sudo mv hadoop-1.2.1 /usr/local/hadoop

----------\*\*downloading and installing java (Hadoop requires a working Java 1.6+)\*\*--------

dsh -a sudo apt-get update

dsh -a sudo apt-get install openjdk-7-jdk -y

------\*\*setting hadoop env\*\*------

nano /usr/local/hadoop/conf/hadoop-env.sh

export JAVA\_HOME=/usr/lib/jvm/java-7-openjdk-amd64

export HADOOP\_OPTS=-Djava.net.preferIPV4Stack=true

-------\*\*configuring xml's\*\*------

nano /usr/local/hadoop/conf/core-site.xml

<property>

<name>fs.default.name</name>

<value>nn:9000</value>

</property>

<property>

<name>hadoop.tmp.dir</name>

<value>/usr/local/hadoop/tmp</value>

</property>

nano /usr/local/hadoop/conf/hdfs-site.xml

<property>

<name>dfs.replication</name>

<value>2</value>

</property>

<property>

<name>dfs.permissions</name>

<value>false</value>

</property>

nano /usr/local/hadoop/conf/mapred-site.xml

<property>

<name>mapred.job.tracker</name>

<value>hdfs://nn:9001</value>

</property>

--------------\*\* scp xml’s to all hosts\*\*----------

cd /usr/local/hadoop/conf/

scp hadoop-env.sh core-site.xml hdfs-site.xml mapred-site.xml ubuntu@snn:/usr/local/hadoop/conf/

scp hadoop-env.sh core-site.xml hdfs-site.xml mapred-site.xml ubuntu@dn1:/usr/local/hadoop/conf/

scp hadoop-env.sh core-site.xml hdfs-site.xml mapred-site.xml ubuntu@dn2:/usr/local/hadoop/conf/

---------------\*\*configuring masters and slaves\*\*-------------

nano /usr/local/hadoop/conf/masters

snn

nano /usr/local/hadoop/conf/slaves

dn1

dn2

----------\*\*making tmp dir\*\*---------

dsh -a mkdir /usr/local/hadoop/tmp

-------------\*\*exec bash\*\*-------

dsh -a exec bash

----------------\*\*formatting hadoop namenode\*\*--------------

hadoop namenode -format

---------------\*\*starting hadoop services(daemons)\*\*-------------------

start-all.sh

jps