**Install GUI on CentOS7:**

Yum group list

yum groupinstall "GNOME Desktop" "Graphical Administration Tools"

ln -sf /lib/systemd/system/runlevel5.target /etc/systemd/system/default.target

reboot

**Basic Setup:**

VM:edit->virtual network editor-> remove and add VMNET8……

……..

**Note** remove and then add back the network adaptors of copied VMs(after the next step).

Remote transfer:

Scp -rp <dir/filename> <hostname/host ip>:<destination dir>

**Network configuration:**

Vim /etc/sysconfig/network-scripts/ifcfg-eth0 (if there isn’t one then create one)

**Replace with these lines:**

DEVICE="eth0"

BOOTPROTO="static"

NM\_CONTROLLED="yes"

ONBOOT="yes"

TYPE="Ethernet"

IPADDR=192.168.80.10 (current host address)

NETMASK=255.255.255.0

GATEWAY=192.168.80.2 (gateway find in prev step)

DNS1=8.8.8.8 (a valid DNS)

**Changing CentOS ens33 to eth0(if needed):**

1. Edit **/etc/default/grub**
2. At the end of **GRUB\_CMDLINE\_LINUX** line append "**net.ifnames=0 biosdevname=0**"
3. Save the file
4. Type "**grub2-mkconfig -o /boot/grub2/grub.cfg**"
5. Type "**reboot**"

Then ifconfig and see if eth0 is up, if not then:

ifconfig eth0 up

reboot

Finally Ping and Curl to check network.

**Hosts configuration:**

Vim /etc/hosts

**Append all the ip and hostnames:**

192.168.80.10 main

192.168.80.11 sub1

192.168.80.12 sub2

Vim /etc/sysconfig/network

**Append these:**

NETWORKING="yes"

HOSTNAME=main (current machine hostname e.g. main/sub1/sub2)

**Stop FireWall/iptables on CentOS7:**

sudo firewall-cmd –state

sudo systemctl stop firewalld (continue for permanent)

sudo systemctl disable firewalld

sudo systemctl mask --now firewalld

ADDITIONALLY:

Setenforce 0 (use getenforce and see permissive means success)

Google for iptable configs.

**SSH Configuration**

Ssh-keygen

Press return until a graph appears.

Cd ~./ssh

Copy pub keys of all machines into authorized\_keys and distribute it to all machines.

**JAVA\_HOME/Path(after installing jdk)**

Vim ~/.bashrc

Append following lines:

export JAVA\_HOME=/usr/local/src/jdk1.8.0\_291 (this is the path of jdk)

export CLASSPATH=.:$CLASSPATH:$JAVA\_HOME/lib

export PATH=$PATH:$JAVA\_HOME/bin

**Hadoop Starts here:**

Enable shared folder in vmware.

Cd /mnt/hgfs

Here is the shared folder and can cp contents to destinations. (here is to transfer Hadoop tarball)

Cp Hadoop tar to /usr/local/src

Tar zxvf <tarball name>

**More Java\_homes**

Edit: export JAVA\_HOME=/usr/local/src/jdk1.8.0\_291 (this is the path of jdk) in these files:

Vim /usr/local/src/hadoop-2.6.5/etc/hadoop/hadoop-env.sh

Vim /usr/local/src/hadoop-2.6.5/etc/hadoop/yarn-env.sh

Worker nodes

Vim /usr/local/src/hadoop-2.6.5/etc/hadoop/slaves

Add all the hostnames of the worker nodes

**Bulk property configuration**

**Vim /usr/local/src/hadoop-2.6.5/etc/hadoop/core-site.xml**

<configuration>

<property>

<name>fs.defaultFS</name>

<value>hdfs://main:9000</value>

</property>

<property>

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

<value>/usr/local/src/hadoop-2.6.5/tmp</value>

</property>

</configuration>

**Cd /usr/local/src/hadoop-2.6.5**

**# mkdir tmp**

**# mkdir -p dfs/name**

**# mkdir -p dfs/data**

**Vim /usr/local/src/hadoop-2.6.5/etc/hadoop/hdfs-site.xml**

<configuration>

<property>

<name>dfs.namenode.secondary.http-address</name>

<value>main:9001</value>

</property>

<property>

<name>dfs.namenode.name.dir</name>

<value>file:/usr/local/src/hadoop-2.6.5/dfs/name</value>

</property>

<property>

<name>dfs.namenode.data.dir</name>

<value>file:/usr/local/src/hadoop-2.6.5/dfs/data</value>

</property>

<property>

<name>dfs.replication</name>

<value>1</value>

</property>

</configuration>

**Cd /usr/local/src/hadoop-2.6.5/etc/Hadoop**

**cp mapred-site.xml.template mapred-site.xml**

**vim mapred-site.xml**

<configuration>

<property>

<name>mapreduce.framework.name</name>

<value>yarn</value>

</property>

</configuration>

**Vim /usr/local/src/hadoop-2.6.5/etc/hadoop/yarn-site.xml**

<configuration>

!-- Site specific YARN configuration properties -->

<property>

<name>yarn.nodemanager.aux-services</name>

<value>mapreduce\_shuffle</value>

</property>

<property>

<name>yarn.nodemanager.aux-services.mapreduce.shuffle.class</name>

<value>org.apache.hadoop.mapred.ShuffleHandler</value>

</property>

<property>

<name>yarn.resourcemanager.address</name>

<value>main:8032</value>

</property>

<property>

<name>yarn.resourcemanager.scheduler.address</name>

<value>main:8030</value>

</property>

<property>

<name>yarn.resourcemanager.resource-tracker.address</name>

<value>main:8035</value>

</property>

<property>

<name>yarn.resourcemanager.admin.address</name>

<value>main:8033</value>

</property>

<property>

<name>yarn.resourcemanager.webapp.address</name>

<value>main:8088</value>

</property>

</configuration>

**Scp the Hadoop dir to all worker machines**

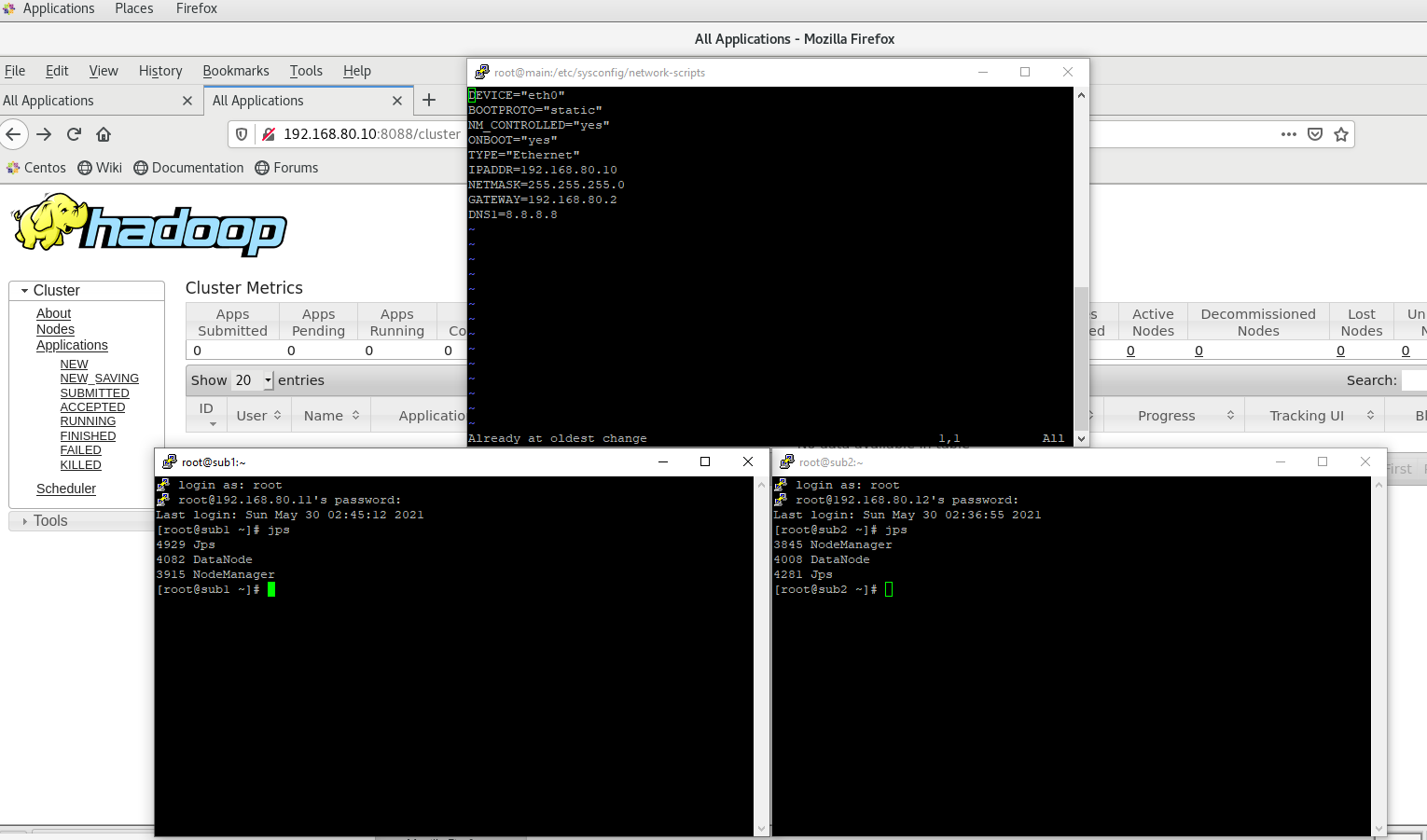
**cd /usr/local/src/hadoop-2.6.5**

**./bin/hadoop namenode -format**

**./sbin/start-dfs.sh + ./sbin/start-yarn.sh**

**Or just ./sbin/start-all.sh**

**Run jps and should see this:**



Data Node disappear when formatting namenode multiple times:

<https://www.programmersought.com/article/35474698451/>

If shared folder not working properly:

First of all you need to create a folder where the mount should be pointed to.

sudo mkdir /your\_shared\_folder

Then you should see your vmware shared folder using this command:

sudo vmware-hgfsclient

If you do see your shared folder then running this should mount your vmware shared folder to the folder you previously created:

sudo vmhgfs-fuse /your\_shared\_folder