**Creating another Slave Node:**

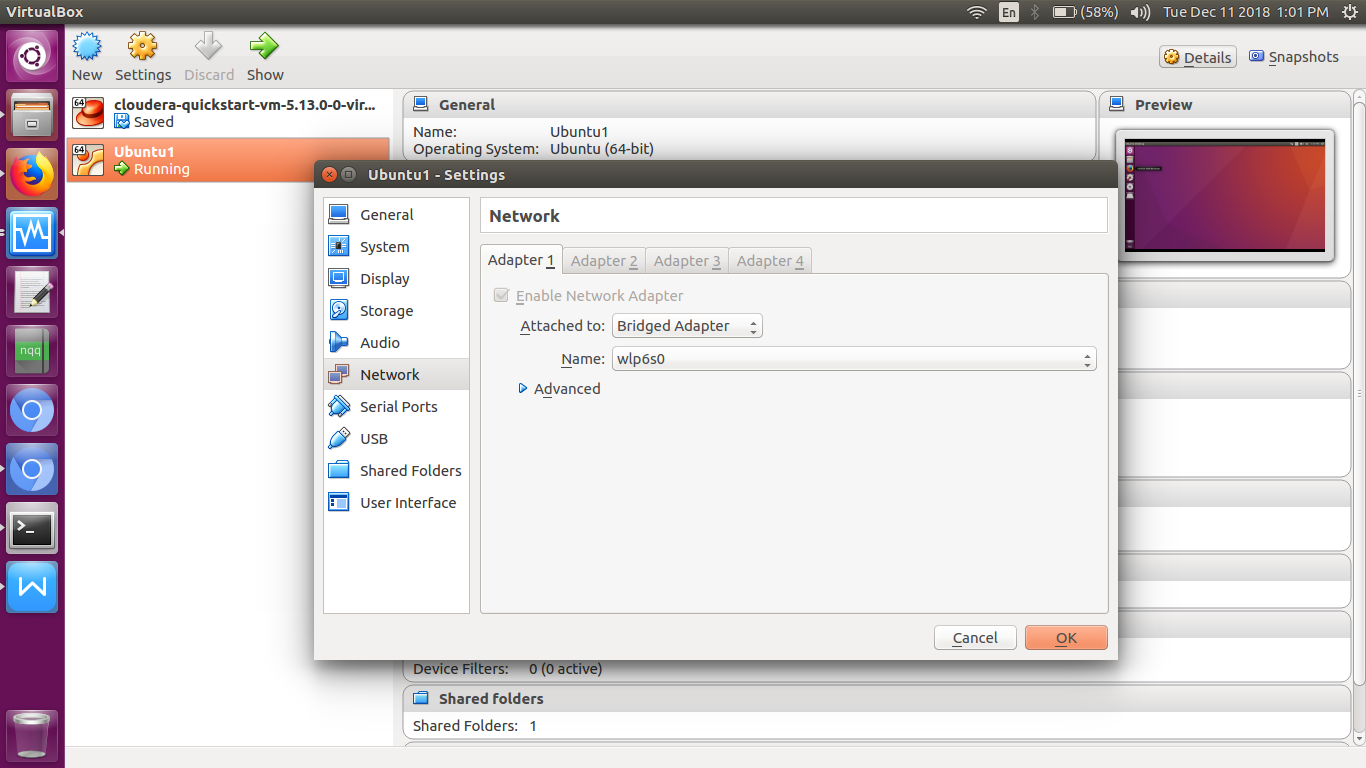
1. Get Ubuntu.iso into virtual box.

2. Create shared folder to copy setups.

**3. Allowing bridge Adapter for machine communications through IP of Ping check:**

Go to Virtual box --> Right Click on Ubuntu1 machine --> Setting --> Network --> TAB- Adapter1 --> Change option: Attached to **Bridged Adapter** --> Press OK

Check below image:

****

**Now check IP in this virtual machine**

**Terminal >** ifconfig

Check **inet addr**: 192.168.43.40

Now check outer OS Terminal> ifconfig

Output: **wlp6s0** Link encap:Ethernet HWaddr 4c:bb:58:ef:dd:f2

**inet addr:192.168.43.37**

wlp6s0 is a adapter name from virtual box setting done above.

**Testing:**

Now ping from Outer OS:

Terminal > ping 192.168.43.37

**OR**

Terminal > ping mypc-VirtualBox

Also check from virtual machine ubuntu:

Terminal > ping 192.168.43.40

**OR**

Terminal > ping sachin-pc

**4.** We have two machines (**Master and Slave/Data Node**) with IP:

Master IP: 192.168.43.37

Slave IP: 192.168.43.40

To Verifying OS details then **Terminal** > cat /etc/\*ease\*

5.(If any **Optional**) Disable the firewall restrictions.

Terminal: service iptables stop

Terminal: sudo chkconfig iptables off

**6.** Open hosts file to add master and data node with their respective IP addresses.

To check hostname/computer name on master and slave use below command:

**Terminal >** hostname

**Terminal >** sudo gedit /etc/hosts

Add following:

*#Hadoop Master-slave*

*192.168.43.37 sachin-pc*

*192.168.43.40 mypc-VirtualBox*

**Same properties will be displayed in the master and slave hosts files (in Ubuntu1).**

**7**. Restart the sshd service.

**Terminal** > service sshd restart

In host if this command fails then

**Terminal** > sudo apt-get install openssh-server

**8.** Create the SSH Key in the master node. (Press enter button when it asks you to enter a filename to save the key).

**Terminal** > ssh-keygen -t rsa -P “”

**9.** Copy the generated ssh key to master node’s authorized keys.

**Terminal** > sudo cat $HOME/.ssh/id\_rsa.pub >> $HOME/.ssh/authorized\_keys

**Terminal** > chmod 600 .ssh/id\_rsa

10. Copy the master node’s ssh key to slave’s authorized keys. (**Nice one**)

**Terminal >** ssh-add

**Terminal** >sudo ssh-copy-id -i $HOME/.ssh/id\_rsa.pub mypc@mypc-VirtualBox -f

Above mypc@mypc-VirtualBox is virtual machine name in virtual box ubuntu.

Now we can login to this slave host here directly.. Magic!!

**Terminal** > ssh 'mypc@mypc-VirtualBox'

And we are in this os now...... Magic!!

1. Make sure java is installed in slave and hadoop extracted there with same version we have in Master:

**Java Install:**

**Terminal >** sudo apt-get install oracle-java8-installer

Use shared folder for copying hadoop.tar.gz and extract to home of slave.

**12.** Add the Hadoop and Java paths in the bash file (.bashrc) on all nodes.

**Better copy same thing from Master and hadoop folder also...**

Open. bashrc file. Now, add Hadoop and Java Path as shown below:

Terminal > sudo gedit ~/.bashrc

*#Hadoop Settings:*

*export HADOOP\_PREFIX=/home/****mypc****/hadoop-2.7.7*

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

*export PATH=$PATH:$HADOOP\_PREFIX/sbin*

*export HADOOP\_MAPRED\_HOME=${HADOOP\_PREFIX}*

*export HADOOP\_COMMON\_HOME=${HADOOP\_PREFIX}*

*export HADOOP\_HDFS\_HOME=${HADOOP\_PREFIX}*

*export YARN\_HOME=${HADOOP\_PREFIX}*

*export HADOOP\_COMMON\_LIB\_NATIVE\_DIR=$HADOOP\_PREFIX/lib/native*

*export HADOOP\_OPTS="-Djava.library.path=$HADOOP\_PREFIX/lib/native"*

**Save it and Terminal >**source .bashrc

Check now Terminal >hadoop version

Terminal > java -version

**13.** Now edit the configuration files in hadoop-2.7.7/etc/hadoop directory.

Create masters file and edit as follows in **both master and slave** machines as below:

**Terminal** > sudo gedit hadoop-2.7.7/etc/hadoop/masters

Add following and save.

*sachin-pc*

Edit slaves file in **master** machine as follows:

**Terminal** > sudo gedit hadoop-2.7.7/etc/hadoop/slaves

Add following and save.

*sachin-pc*

*mypc-VirtualBox*

Edit slaves file in **slave** machine as follows:

**Terminal** > sudo gedit hadoop-2.7.7/etc/hadoop/slaves

Add following and save.

*mypc-VirtualBox*

**14**. Edit core-site.xml on both **master and slave** machines as follows:

**Terminal** > sudo gedit hadoop-2.7.7/etc/hadoop/core-site.xml

*<configuration>*

*<property>*

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

*<value>hdfs://sachin-pc:9000</value>*

*</property>*

*</configuration>*

**15.** Edit hdfs-site.xml on **master** as follows:

**Terminal** > sudo gedit hadoop-2.7.7/etc/hadoop/hdfs-site.xml

*<configuration>*

*<property>*

*<name>dfs.replication</name>*

*<value>2</value>*

*</property>*

*<property>*

*<name>dfs.permissions</name>*

*<value>false</value>*

*</property>*

*<property>*

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

*<value>/home/sachin/hadoop-2.7.7/namenode</value>*

*</property>*

*<property>*

*<name>dfs.datanode.data.dir</name>*

*<value>/home/sachin/hadoop-2.7.7/datanode</value>*

*</property>*

*</configuration>*

**Terminal > sudo chmod 755** *home/sachin/hadoop-2.7.7/datanod*

**15.** Edit hdfs-site.xml on **slave** as follows:

**Terminal** > sudo gedit hadoop-2.7.7/etc/hadoop/hdfs-site.xml

*<configuration>*

*<property>*

*<name>dfs.replication</name>*

*<value>2</value>*

*</property>*

*<property>*

*<name>dfs.permissions</name>*

*<value>false</value>*

*</property>*

*<property>*

*<name>dfs.datanode.data.dir</name>*

*<value>/home/mypc/hadoop-2.7.7/datanode</value>*

*</property>*

*</configuration>*

**16.** No change with **mapred-site.xml and yarn-site.xml** as we have same copied from master to slave. Code is there already in that file.

**17**. Format the namenode (Only on master machine).

**Terminal** > hadoop namenode -format

**18**. Start all daemons (Only on master machine).

**Terminal > Start-all.sh**

**\*\*Start Data node separately if not started...**

**Terminal >** hadoop-daemon.sh start datanode

**19. Goto Datanode now here in virtual machine:**

**Terminal >** /home/mypc/hadoop-2.7.7/bin/hadoop-daemon.sh start datanode

**OR**

**Terminal >** hadoop-daemon.sh start datanode

**20. Check status:**

**On Master browser:** sachin-pc:50070/dfshealth.html

**OR** localhost:50070/dfshealth.html