Creating Multinode in a cluster

1. Create two folders master and slave1, slave2, and slave3 in VM directory.

2.Goto VMWorkstation, right click the existing VM - > Manage ->clone -> Next-> Current state in the virtual machine -> create a full clone - > name(hadoopmaster) and select the directory.

3.open hadoopmaster VM

$ifconfig

4.Note the IP of the hadoopmaster - 192.168.174.129

192.168.174.130 hadoopslave1

192.168.174.131 hadoopslave2

192.168.174.132 hadoopslave3

5.Open /etc/hosts file and enter this IP address and host name of the all the VM's

$sudo i

$gedit /etc/hosts

192.168.174.129 hadoopmaster

192.168.174.130 hadoopslave1

192.168.174.131 hadoopslave2

192.168.174.132 hadoopslave3

Save and close

6.Edit /etc/hostname file.

$gedit /etc/hostname

change ubuntu to hadoopmaster

In slave machine1, change the slave1 to hadoopslave1

In slave machine2, change the slave2 to hadoopslave2

In slave machine3, change the slave3 to hadoopslave3

Save and close

7. Goto /etc/hadoop directory (/home/hadoop/HadoopEcoSystem/hadoop-2.9.0/etc/hadoop/) and edit core-site.xml file in all the 3 machines. Update it to hadoopmaster.

<property>

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

<value>**hdfs://localhost:9000**</value>

</property>

Rename - localhost to hadoopmaster

8.Open hdfs-site.xml file

<property>

<name>dfs.replication</name>

<value>1</value>

</property>

Set the replication value to 3.

9. Open yarn-site.xml file and add three properties

<property>

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

<value>hadoopmaster:8025</value>

</property>

<property>

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

<value>hadoopmaster:8030</value>

</property>

<property>

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

<value>hadoopmaster:8050</value>

</property>

10. Open mapred-site.xml in hadopomaster and change the mapreduce.framework.name to mapred.job.tracker and change yarn to hadoopmaster:54311

<property>

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

<value>hadoopmaster:54311</value>

</property>

11. Create three slave machines using clone and IP address will be updated as shown above.

12. Open hadoop master

open hadoop/etc/hadoop/masters file and update hadoopmaster

open hadoop/etc/hadoop/slaves file and update

hadoopslave1

hadoopslave2

hadoopslave3

13. In hadoopmaster, open hdfs-site.xml and copy datanode property

($sudo rm - rf /usr/local/hadoop/hadoop\_data)

root@hadoopmaster:~# sudo rm -rf /home/hadoop/HadoopEcoSystem/Hadoop\_MetaStore/

root@hadoopmaster:~# sudo mkdir -p /home/hadoop/HadoopEcoSystem/hdfs/namenode

root@hadoopmaster:~# sudo chown -R hadoop:hadoop /home/hadoop

14. In slave1, slave2 and slave3

$sudo rm -rf /home/hadoop/HadoopEcoSystem/Hadoop\_MetaStore/

$sudo mkdir -p /home/hadoop/HadoopEcoSystem/hdfs/datanode

$ sudo chown -R hadoop:hadoop /home/hadoop

15. Setup the password less login

ssh-copy-id remote-user@remote-server-ip

ssh-copy-id -i ~/.ssh/id\_rsa.pub hadoop@hadoopmaster

ssh-copy-id -i ~/.ssh/id\_rsa.pub hadoop@hadoopslave1

ssh-copy-id -i ~/.ssh/id\_rsa.pub hadoop@hadoopslave2

ssh-copy-id -i ~/.ssh/id\_rsa.pub hadoop@hadoopslave3

16. Test the connectivity from Master to Slave

from hadoopmaster, to test the connectivity...

$ssh hadoopslave1

$ssh hadoopslave2

$ssh hadoopslave3

$hadoop namenode -format

17. To setup FTP with Windows

gedit /etc/ssh/sshd\_config

Change PasswordAuthentication and ChallengeResponseAuthentication to yes