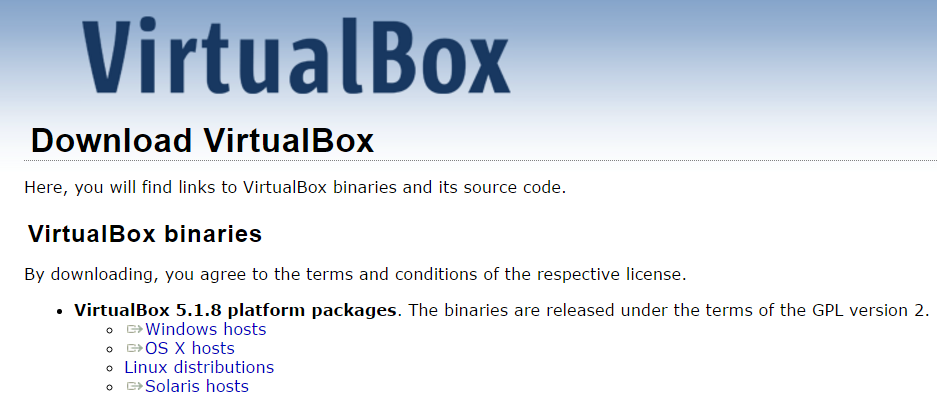
**1. Installing Virtual Box**

Download **VirtualBox-5.1.8-111374-Win** from below site.

<https://www.virtualbox.org/wiki/Downloads>



Click on “**Windows hosts**” to download **Oracle** **VirtualBox** for Windows. Double click on **VirtualBox-5.1.8-111374-Win.exe** and install.

**2. Download CentOS iso image**

Download **CentOS-7-x86\_64-Minimal-1610-01.iso** from below site.

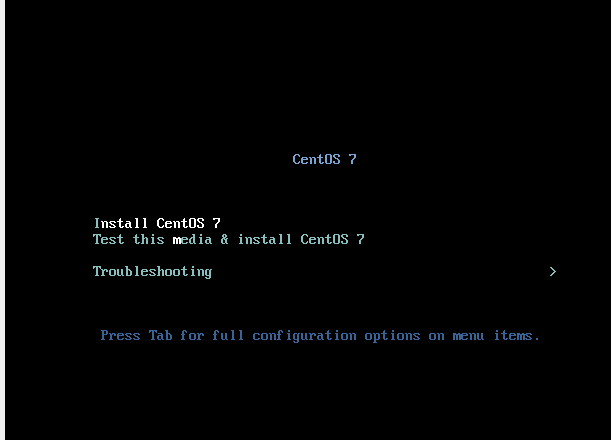
<http://buildlogs.centos.org/rolling/7/isos/x86_64/>

**3. Creating Virtual Machine**

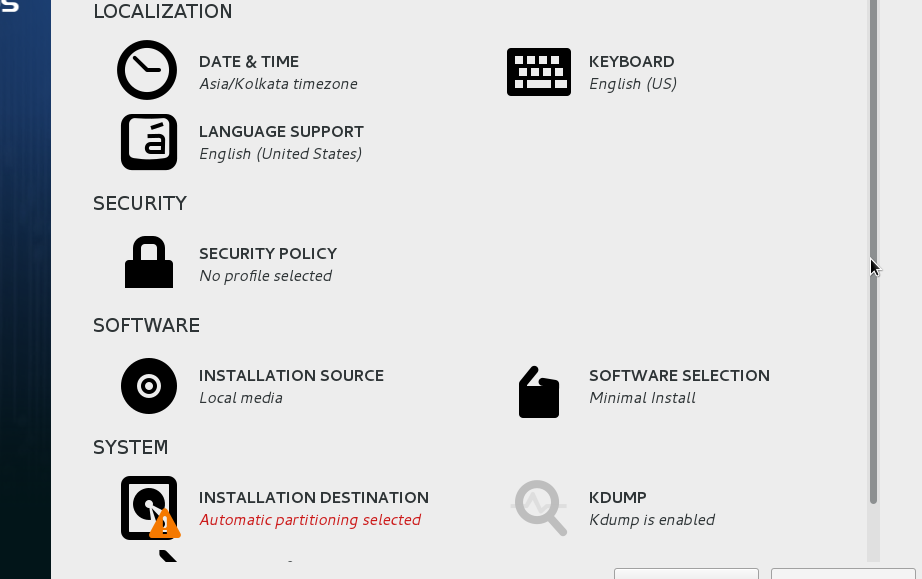
* Open **Oracle VM VirtualBox** **Manager**
* Click on **Group → New Machine**
* Provide the **Name** you wanted give to your VM (Ex: Slave 7)
* Select **Type** as **Linux**
* Select **Version** as **Red Hat (64-bit)** and click **Next**
* Give **Memory Size** of **2048 MB** (which is 2GB of RAM) and click **Next**
* Select **Create a virtual hard disk now** option and click **Create**
* Select **VDI (VirtualBox Disk Image)** and click **Next**
* Select **Dynamically allocated** option and click **Next**
* Select **8.00 GB** and click **Create**. This will create an VM (Slave 7) in left hand side of your VirtualBox Manager
* Right click on **Slave 7** VM and select **Settings → Storage**
* Click on **Empty** under **controller: IDE**
* Click on CD icon under Attributes: and select **CentOS-7-x86\_64-Minimal-1610-01.iso** from browser and click on **OK**
* Right click on **Slave 7** VM and select **Start → Normal Start**
* This will start installation of **CentOS**
* Complete the installation by following the instructions in 4th step
* Click on **Restart** once installation is completed and your VM is ready to use

**4. CentOS 7 Installation**

This is the first screen I got when my machine started. Well you can straight away select “**Install CentOS7”**

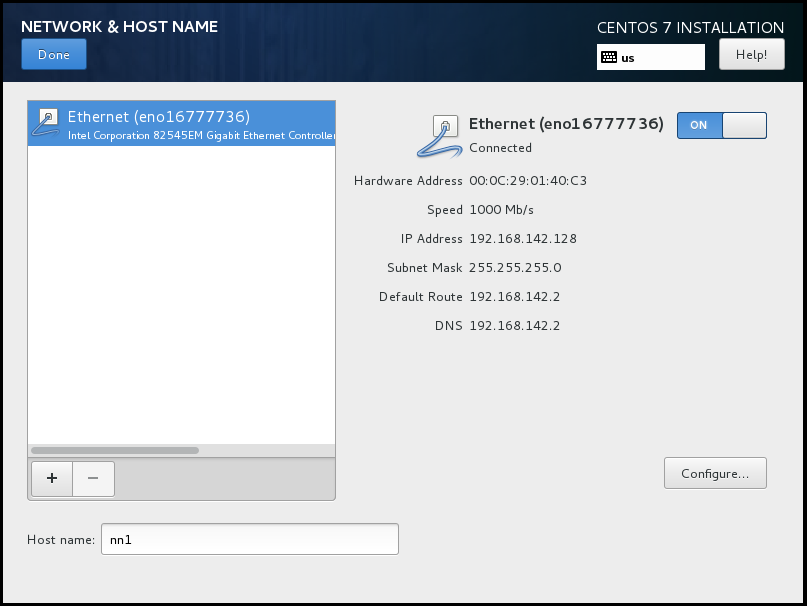


Note: In VirtualBox if you click inside the VM screen your mouse pointer will be captured inside. It usually gives instructions on how to release the pointer from the VM but we rarely read the instructions. The Default release button is **Right Cntrl** to release your pointer from the VM. The release key is usually displayed in the Right Bottom corner of your screen. (For VMWare the default is **Cntrl + Alt**). Once you select **Install CentOS7** the next couple of screens are self-explanatory. It will ask you for your Time Zone Location, Keyboard Layout, Security Settings, Disk Partitions and Network Settings etc.



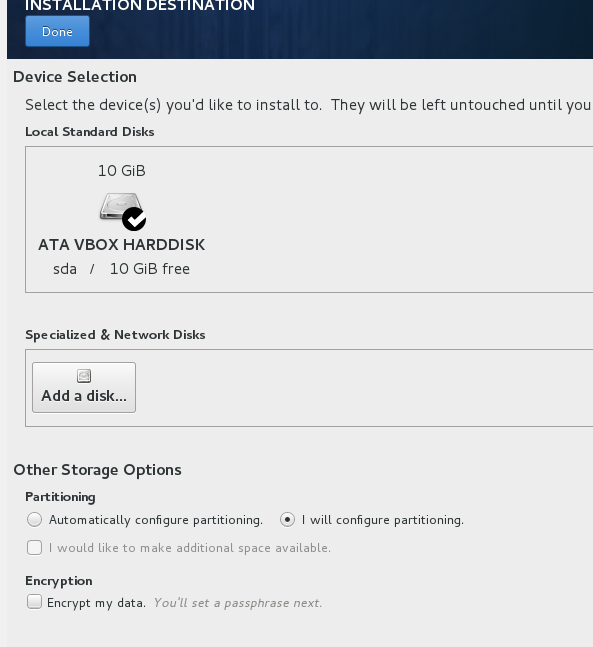
For now just do the below changes

* Change the **Date & Time** to India. (Just select on India on the Map).
* Change the **Language Support** to English India if you want.
* Change the “**Network & Hostname**” state to “Connected”. We can configure the ip address and stuffs later. In the same tab you can change the Hostname to nn1. Well we can do this as well later so not to worry. By Default it has DHCP enabled which is fine for now. We will configure the network according to our requirement later.

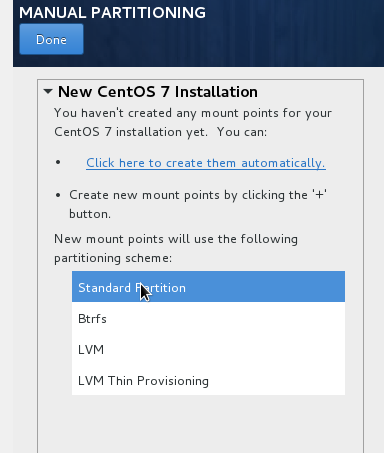


In the **“Installation Destination”** select the Disk that we have added if not already selected. Select ‘I will configure partitioning’ since CentOS7 by default will create the partitions in LVM which we don’t require here.

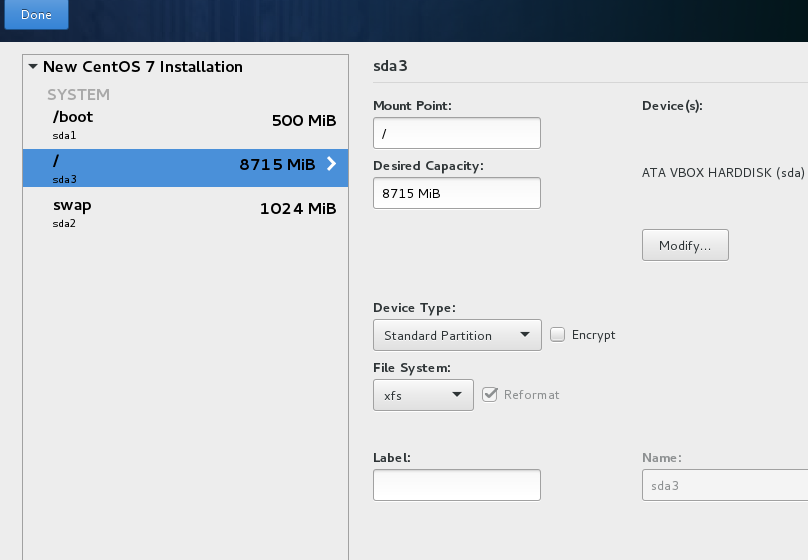
Click **DONE** so that it will take you to the page where you can define the partitions you want.



Now by Default the Partitioning Scheme will populate LVM. Change it to ‘Standard Partition’ as shown below and then click on ‘Click here to create them automatically’



This will create all the default Linux partitions in Standard format as shown in the below screenshot



Here you can see three partitions have been created (which is default for Linux). Ideally we should create a separate partition for Hadoop but I will be doing that later by adding a separate disk. You can see other information such as the name of the partition (sda3) and the filesystem which is xfs. You can change these values if you want but don’t do anything now.

Click **DONE** and then click on ‘**Begin Installation’**. Finally it would ask you to provide the root password and to create one user account. Once this is done the OS will start installing which might take some time. Once the installation is complete it will ask to Reboot the system. Now the machine will reboot from the hard Disk into your CentOS login screen.

**5. Set Static IP (with internet if possible) for any Host/VM**

In CentOS7, **ifconfig** command will not come as a default command. You need to install rpm for that.

**yum install net-tools**

You need to perform below steps to allow host internet connection to guest VM

* Right click on VM and go to **Settings -> Network**
* Select “**Attached to:** ” to “**Bridged Adapter**”

To set a static IP, we need to change DNS settings as per your router/ISP.

To check your router DNS settings, login to [http://192.168.1.1](http://192.168.0.1/) with user ID as **admin** and password as **admin** and get DNS info. If you don't know the login credentials of your router, please check with your Internet Service Provider (ISP) for the DNS info.

Note: At times, few routers will allow only a range of IP addresses. For Ex: My router allows IP range from 192.168.1.100 to 192.168.1.199 only. So, my new static IP should be between 100 and 199 only. This range may differ with router to router.

Also, this login information will be different from Router to Router based on manufacturer.

Get VM MAC ID from below command.

**cat /sys/class/net/enp0s3/address**

Delete all entries and change ifcfg-enp0s3 file as shown below.

**sudo vi /etc/sysconfig/network-scripts/ifcfg-enp0s3**

TYPE="Ethernet"

BOOTPROTO="none" --> should be none for static IP

NAME="enp0s3"

UUID=d42b0589-5212-444a-b1bf-40ebd2acbc7d --> this will come by default

DEVICE="enp0s3"

ONBOOT="yes"

HWADDR=FC:AA:14:8E:BB:BC --> output of above MAC command

NM\_CONTROLLED="yes"

IPADDR=192.168.1.110 --> your static IP

NETMASK=255.255.255.0 --> your net mask is 255.255.255.0 by default

GATEWAY=192.168.1.1 --> your gateway

DNS1=202.83.21.43 --> your internet service provider DNS

Restart the network now.

**systemctl restart network**

**6. Change other files in VM as below**

Check whether you are able to access internet or not.

**ping google.com**

Press Ctrl+C to come out of ping

Change **resolve.conf** as below.

**sudo vi /etc/resolv.conf**

Delete if you have any other info and add below line only

**nameserver 202.83.21.43**

Change **/etc/sysconfig/network** as below.

**sudo vi /etc/sysconfig/network**

Delete all entries in this and add below.

**NETWORKING=yes**

**NETWORKING\_IPV6=no**

**HOSTNAME=slave.hadoop.com**

Run below to change it to static:

**hostnamectl status**

**hostnamectl set-hostname "slave.hadoop.com" --static**

**hostnamectl status**

Note: Change “**HOSTNAME**” value as per your requirement.

If you don’t have DNS configuration, add below to **/etc/hosts** file.

**sudo vi /etc/hosts**

Add below here:

**127.0.0.1 localhost.hadoop.com localhost**

**#Hadoop Master & Slave Base Hosts**

**192.168.1.201 master.hadoop.com master**

**192.168.1.202 slave.hadoop.com slave**

**# Cloudera Installation Hosts**

**192.168.1.111 cm.hadoop.com cm**

**192.168.1.112 cnn.hadoop.com cnn**

**192.168.1.113 dn1.hadoop.com dn1**

**192.168.1.114 dn2.hadoop.com dn2**

**192.168.1.115 dn3.hadoop.com dn3**

**192.168.1.116 dn4.hadoop.com dn4**

**# Hadoop 1.X Installation Hosts**

**192.168.1.151 nn11.hadoop.com nn11**

**192.168.1.152 dn11.hadoop.com dn11**

**192.168.1.153 dn12.hadoop.com dn12**

**192.168.1.154 dn13.hadoop.com dn13**

**192.168.1.155 dn14.hadoop.com dn14**

**192.168.1.156 snn.hadoop.com snn**

**# Hadoop 2.X Installation Hosts with Journal Nodes**

**192.168.1.171 jnn1.hadoop.com jnn1**

**192.168.1.172 jnn2.hadoop.com jnn2**

**192.168.1.173 jdn1.hadoop.com jdn1**

**192.168.1.174 jdn2.hadoop.com jdn2**

**192.168.1.175 jdn3.hadoop.com jdn3**

**192.168.1.176 jedge.hadoop.com jedge**

Restart the VM.

**7. Install Hadoop Mandatory RPMs**

**cd /etc/yum.repos.d**

**sudo** [**echo never > /sys/kernel/mm/transparent\_hugepage/defrag**](http://cm.hadoop.com:7180/cmf/express-wizard/wizard)

**sudo yum -y install wget**

**sudo yum -y install createrepo**

**sudo yum -y install yum-utils createrepo**

**sudo yum -y install MySQL-python\***

**sudo yum -y install python\***

**sudo yum -y install httpd**

**sudo yum -y install telnet**

**sudo yum -y install bind\***

**sudo yum -y install openssh\***

**sudo yum -y install rpmdevtools**

**sudo yum -y install ntp\***

**sudo yum -y install redhat-lsb\***

**sudo yum -y install cyrus\***

**sudo yum -y install mod\_ssl\***

**sudo yum -y install portmap\***

**sudo yum -y install openssl\***

**sudo yum -y install mlocate\*  
sudo yum -y install sshpass**

**sudo wget http://repos.fedorapeople.org/repos/dchen/apache-maven/epel-apache-maven.repo -O /etc/yum.repos.d/epel-apache-maven.repo**

**sudo yum -y install apache-maven**

**sudo updatedb**

**sudo systemctl disable firewalld.service**

**sudo systemctl enable httpd  
sudo systemctl enable ntpd  
sudo systemctl enable ntpdate**

**sudo systemctl status firewalld.service**

**sudo systemctl status httpd**

**sudo systemctl status ntpd  
sudo systemctl status ntpdate**

**8. Prepare Node Prerequisites**

Check for the below mentioned components and make sure their status is as required:

* **SELinux** : Should be disabled

**sudo sestatus**

if it is enabled,

**sudo vi /etc/selinux/config**

set, **selinux=disabled**

* **IPv6** : Should be disabled

**sudo cat /proc/sys/net/ipv6/conf/all/disable\_ipv6**

if output is 1 means disabled, 0 means enabled. If enabled, add below lines in **/etc/sysctl.conf** to disable

**sudo vi /etc/sysctl.conf**

**# disable ipv6**

**net.ipv6.conf.all.disable\_ipv6 = 1**

**net.ipv6.conf.default.disable\_ipv6 = 1**

**net.ipv6.conf.lo.disable\_ipv6 = 1**

**vm.swappiness = 10**

* **Iptables**: Should be disabled

**sudo systemctl status firewalld.service**

Output for above command should show “**Active: inactive (dead)**”. If no,

**sudo systemctl disable firewalld.service**

* **Fastestmirror**: Should be enabled

**sudo cat /etc/yum/pluginconf.d/fastestmirror.conf**

Value of enabled should be 1. If 0 or any other value change to 1.

* Comment CentOS Time Sync Servers in **/etc/ntp.conf**

**sudo sed -i 's/^server /#server /g' /etc/ntp.conf**

* Disable sleep mode of VM

**sudo vi /etc/profile**

Add below

**setterm -blank 0 -powersave off -powerdown 0**

<http://askubuntu.com/questions/47311/how-do-i-disable-my-system-from-going-to-sleep>

**9. Install JAVA**

To download latest JAVA RPM,

**cd /etc/yum.repos.d**

**sudo wget --no-cookies --no-check-certificate --header "Cookie: gpw\_e24=http%3A%2F%2Fwww.oracle.com%2F; oraclelicense=accept-securebackup-cookie" "**[**http://download.oracle.com/otn-pub/java/jdk/8u77-b03/jdk-8u77-linux-x64.rpm**](http://download.oracle.com/otn-pub/java/jdk/8u77-b03/jdk-8u77-linux-x64.rpm)**"**

To install Oracle JDK, in root login:

**cd /etc/yum.repos.d**

**sudo rpm -ivh jdk-8u77-linux-x64.rpm**

**java -version**

If you have multiple java version, use below command tt the latest one.

**sudo alternatives --config java**

**10. Install MySQL JAVA Connector**

To install MySQL Java Connector

**sudo yum -y install mysql-connector-java**

After installation, check whether **mysql-connector-java.jar** is present in **/usr/share/java/**.

**11. Install rpmdevtools**

**Install & Configure CLUSTERSHELL**

This is to run a common command on all or given set of nodes. To install this, you need to have EPEL Repository & Python installed on your host. To execute clush commands, we need to have passwordless ssh for all the nodes in the cluster. So it is better to install in CM / Master node and is not required to install in all nodes. You need to have passwordless SSH to the all hosts from master node.

To create EPEL Repo

**sudo yum -y install epel-release**

Check whether **epel.repo** is create in **/etc/yum.repos.d** folder

Follow step 2 if you need to install PYTHON

To install Clustershell

**sudo yum -y install \*clustershell\***

Delete all entries from below file

**sudo vi /etc/clustershell/groups.d/local.cfg**

**12. Install Postgresql (Optional – Required only for Cloudera Installation)**

For initial installation of Cloudera Manager, we need to have Postgresql to be installed on CM host and later this can be modified to MySQL.

**sudo yum install -y postgresql-server postgresql-contrib**

**sudo postgresql-setup initdb  
sudo systemctl start postgresql**

**sudo systemctl enable postgresql  
sudo systemctl status postgresql**

**13. Install MySQL (Optional – Required only for Cloudera Installation)**

Hadoop supports InnodDB engine for all its databases. InnoDB engine is coming as default engine from latest versions only. To install mysql, download **mysql57-community-release-el7-9.noarch.rpm** from [**https://dev.mysql.com/downloads/repo/yum/**](https://dev.mysql.com/downloads/repo/yum/). Copy this to your VM under **/var/www/html/hadoop\_tools/** by using WinSCP.

**cd /var/www/html/hadoop\_tools/**

**rpm -ivh** **mysql57-community-release-el7-9.noarch.rpm**

Above command will create 2 repo files in **/etc/yum.repos.d** directory

**mysql-community.repo**

**mysql-community-source.repo**

To install mysql, move any existing CentOS related repos from **/etc/yum.repos.d** folder to **/etc/yum.repos.d/baserepos**.

**cd /etc/yum.repos.d**

**mkdir -p /etc/yum.repos.d/baserepos**

**mv CentOS\* /etc/yum.repos.d/baserepos**

**mv epel\* /etc/yum.repos.d/baserepos**

**yum clean all**

**yum install -y mysql mysql-server**

Once the installation is completed, we need to set the password for root login.

**service mysqld status**

**service mysqld start**

**service mysqld stop**

**mysqld\_safe --skip-grant-tables &**

Press ENTER key to start mysql in safe mode

Login to mysql without password.

**mysql -u root**

**update mysql.user set authentication\_string=password('welcome1') where user='root';**

**ALTER USER 'root'@'localhost' IDENTIFIED BY 'welcome1';**

**FLUSH PRIVILEGES;**

Logout from mysql and stop mysql safemode.

**service mysqld stop**

**service mysqld start**

**mysql -u root -p**

**uninstall plugin validate\_password;**

Logout from mysql.

To setup root password:

**$ sudo /usr/bin/mysql\_secure\_installation**

[...]

Enter current password for root (enter for none):

OK, successfully used password, moving on...

[...]

Set root password? [Y/n] y

New password:

Re-enter new password:

Remove anonymous users? [Y/n] Y

[...]

Disallow root login remotely? [Y/n] N

[...]

Remove test database and access to it [Y/n] Y

[...]

Reload privilege tables now? [Y/n] Y

**All done!**

Place below content in my.cnf

**vi /etc/my.cnf**

copy from here:

**[mysqld]**

**transaction-isolation=READ-COMMITTED**

**# Disabling symbolic-links is recommended to prevent assorted security risks;**

**# to do so, uncomment this line:**

**# symbolic-links=0**

**key\_buffer\_size = 32M**

**max\_allowed\_packet = 16M**

**thread\_stack = 256K**

**thread\_cache\_size = 64**

**query\_cache\_limit = 8M**

**query\_cache\_size = 64M**

**query\_cache\_type = 1**

**# Important: see Configuring the Databases and Setting max\_connections**

**max\_connections = 550**

**# log-bin should be on a disk with enough free space**

**#log-bin=/var/log/mysql/mysql\_binary\_log**

**# For MySQL version 5.1.8 or later. Comment out binlog\_format for older versions.**

**binlog\_format = mixed**

**read\_buffer\_size = 2M**

**read\_rnd\_buffer\_size = 16M**

**sort\_buffer\_size = 8M**

**join\_buffer\_size = 8M**

**# InnoDB settings**

**innodb\_file\_per\_table = 1**

**innodb\_flush\_log\_at\_trx\_commit = 2**

**innodb\_log\_buffer\_size = 64M**

**innodb\_buffer\_pool\_size = 4G**

**innodb\_thread\_concurrency = 8**

**innodb\_flush\_method = O\_DIRECT**

**innodb\_log\_file\_size = 512M**

**[mysqld\_safe]**

**log-error=/var/log/mysqld.log**

**pid-file=/var/run/mysqld/mysqld.pid**

Create bin-log directory: (ignore this step if you are not using bin log)

**sudo mkdir /var/log/mysql/**

**sudo chown -R mysql:mysql /var/log/mysql/**

Start mysql now.

**sudo systemctl enable mysqld**

**sudo systemctl stop mysqld  
sudo systemctl start mysqld  
sudo systemctl status mysqld**

**14. Creating default Group & User**

In root login:

**groupadd -g 2000 hdpadmin**

**useradd -u 2000 hdpuser -g hdpadmin**

To set password:

**# passwd hdpuser**

Add these users to sudoers

**sudo visudo**

Add below lines at the end to the **sudoers**

**hdpuser ALL=(ALL) NOPASSWD: ALL**

**%cloudera-scm ALL=(ALL) NOPASSWD: ALL**

**root ALL=(ALL) NOPASSWD: ALL**

**Defaults !secure\_path**

Change below lines in the sudoers

**Defaults !requiretty**

**15. Creating project specific Groups & Users**

**sudo userdel blruser**

**sudo userdel hyduser**

**sudo userdel adam**

**sudo userdel allen**

**sudo userdel bond**

**sudo userdel james**

**sudo userdel donald**

**sudo userdel cloudera-scm**

**sudo groupadd -g 2001 blrproj**

**sudo groupadd -g 2002 hydproj**

**sudo useradd -u 2101 blruser -g blrproj**

**sudo useradd -u 2102 adam -g blrproj**

**sudo useradd -u 2103 allen -g blrproj**

**sudo useradd -u 2104 bond -g blrproj**

**sudo useradd -u 2201 hyduser -g hydproj**

**sudo useradd -u 2202 james -g hydproj**

**sudo useradd -u 2203 donald -g hydproj**

**sudo passwd blruser**

**sudo passwd adam**

**sudo passwd allen**

**sudo passwd bond**

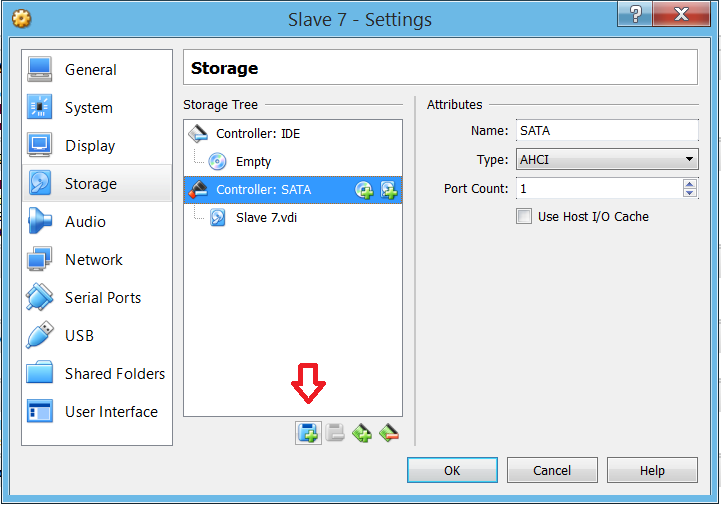
**sudo passwd hyduser**

**sudo passwd james**

**sudo passwd donald**

**16. Adding new Hard Disks:**

* Shutdown the VM
* Open **Oracle VM VirtualBox** **Manager**
* Right click on **Slave 7** VM and select **Settings → Storage**
* Click on **controller: SATA**
* Click on below symbol and select “**Add Hard Disk**”



* Click on **Create new disk** and enter the **File location** as **Disk1** and click on **Create**
* Click on OK
* Start your VM and execute below commands

**mkdir -p /mnt/disk1**

**mkdir -p /mnt/disk2**

**mkfs.ext4 /dev/sdb**

**mount -t ext4 /dev/sdb /mnt/disk1**

**ls -lrt /mnt/disk1**

**rm -rf /mnt/disk1/lost+found/**

**ls -lrt /mnt/disk1**

**mkfs.ext4 /dev/sdc**

**mount -t ext4 /dev/sdc /mnt/disk2**

**ls -lrt /mnt/disk2**

**rm -rf /mnt/disk2/lost+found/**

**ls -lrt /mnt/disk2**

**fdisk -l**

**vi /etc/fstab**

**Add below.**

**/dev/sdb /mnt/disk1/ ext4 defaults 0 0**

**/dev/sdc /mnt/disk2/ ext4 defaults 0 0**

To see the partitions,

**df -h**

**17. Cloning Virtual Machines**

* Open **Oracle VM VirtualBox** **Manager**
* Right click on **Slave 7** VM and select **Clone**
* Enter new VM name in **New machine name** window. (Ex: NN)
* Select check box for **Reinitialize the MAC address of all network cards** and click **Next**
* Select **Full clone** option and click on **Clone** button.

For any questions, email to **narasimha.v.rao.b@gmail.com**