



­

Assignment Document:

Cloudera

**NOTE: WE NEED TO FOLLOW STEPS 1 – 6 FOR ALL THE HOSTS IN THE CLUSTER.**

**# Cloudera Installation Hosts**

|  |  |  |
| --- | --- | --- |
| **IP Address** | **FQDN** | **Hostname** |
| 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 |

**1. Preparing VMs for hadoop cluster installation**

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 service iptables status**

Output for above command should show “**firewall is not running**”. If no,

**sudo chkconfig iptables off**

* 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.

**2. Creating Groups & Users on all hosts**

In root login:

**groupadd -g 2000 cloudera-scm**

**useradd -u 2000 cloudera-scm -g cloudera-scm**

To set password:

**# passwd cloudera-scm**

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

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>

You need to repeat these steps in all hosts. From now, you need to login to **hadoop** user in all hosts and perform below steps.

**4. Install JAVA (Optional Step)**

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>**"**

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**

**5.1 Install PostgreSQL for default CM database**

**5.2 Install MySQL Database with InnoDB engine support**

**6. Install MySQL JAVA Connector in CM Node**

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/**. If no, execute below steps.

**sudo mkdir -p /usr/share/java/**

**sudo cp mysql-connector-java-5.1.17/mysql-connector-java-5.1.17-bin.jar /usr/share/java/mysql-connector-java.jar**

**7. Create local repository in CM node**

Download all RPMs for java and hadoop components and keep them in **/var/www/html/hadoop** in the VM in which you want to create your local repository. **/var/www/html** already exists, **hadoop** directory needs to be created. Any other name can also be given. In this we’ll follow **/var/www/html/hadoop**.

After copying all the RPMs to the above specified directory, do:

**cd /var/www/html/cloudera-cdh5**

**sudo createrepo .**

After the above command is executed a directory **repodata** will be made in the directory, which will contain **repomd.xml**

Now we need to create a .repo file in order to enable yum to access our repository.

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

**sudo vi cloudera-cdh5.repo** (any name can be given, we have taken local\_hadoop)

add below content to the file:

**[cloudera-cdh5.repo]**

**enabled = 1**

**name = "My Local Repository for Hadoop"**

**baseurl = http://192.168.1.111/cloudera-cdh5/**

**gpgcheck = 0**

Or copy from

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

**sudo cp baserepos/cloudera-cdh5.repo .**

**cat cloudera-cdh5.repo**

**8. Install Cloudera Manager in CM node**

Download **cloudera-manager-installer.bin** using below.

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

**sudo yum clean all**

**sudo cp baserepos/cloudera-manager-installer.bin .**

**sudo chmod u+x cloudera-manager-installer.bin**

**sudo ./cloudera-manager-installer.bin --skip\_repo\_package=1**

**sudo service cloudera-scm-server status**

**sudo service cloudera-scm-server-db status**

**sudo telnet cm.hadoop.com 7180**

This will start CM with postgresql. To see the logs,

**sudo ls -ltr /var/log/cloudera-manager-installer**

**sudo vi /var/log/cloudera-scm-server/cloudera-scm-server.log**

**9. To uninstall,**

**sudo bash /usr/share/cmf/uninstall-cloudera-manager.sh**

**sudo rm -rf /etc/cloudera-scm-server**

**sudo rm -rf /var/log/cloudera-manager-installer**

**sudo rm -rf /var/log/cloudera-scm-server**

**sudo rm -rf /var/run/cloudera-scm-server.pid**

**sudo rm -rf /var/spool/mail/cloudera-scm**

**10. To change to MySQL from PostgreSQL, Create DB Schema in CM Node:**

To delete any pre-existing databases and users:

**mysql -u root -pwelcome1 -e "show databases;"**

**mysql -u root -pwelcome1 -e "select user from mysql.user;"**

**mysql -u root -pwelcome1 -e "drop database amon;"**

**mysql -u root -pwelcome1 -e "drop database hmon;"**

**mysql -u root -pwelcome1 -e "drop database rman;"**

**mysql -u root -pwelcome1 -e "drop database smon;"**

**mysql -u root -pwelcome1 -e "drop database nav;"**

**mysql -u root -pwelcome1 -e "drop database navms;"**

**mysql -u root -pwelcome1 -e "drop database metastore;"**

**mysql -u root -pwelcome1 -e "drop database sentry;"**

**mysql -u root -pwelcome1 -e "drop database oozie;"**

**mysql -u root -pwelcome1 -e "drop database hue;"**

**mysql -u root -pwelcome1 -e "drop database scm;"**

**mysql -u root -pwelcome1 -e "drop user amon;"**

**mysql -u root -pwelcome1 -e "drop user hmon;"**

**mysql -u root -pwelcome1 -e "drop user rman;"**

**mysql -u root -pwelcome1 -e "drop user smon;"**

**mysql -u root -pwelcome1 -e "drop user nav;"**

**mysql -u root -pwelcome1 -e "drop user navms;"**

**mysql -u root -pwelcome1 -e "drop user hive;"**

**mysql -u root -pwelcome1 -e "drop user sentry;"**

**mysql -u root -pwelcome1 -e "drop user oozie;"**

**mysql -u root -pwelcome1 -e "drop user hue;"**

**mysql -u root -pwelcome1 -e "drop user 'scm'@'localhost';"**

**mysql -u root -pwelcome1 -e "show databases;"**

**mysql -u root -pwelcome1 -e "select user from mysql.user;"**

To create new databases and users:

**mysql -u root -pwelcome1**

In mysql prompt, run below commands.

**create database amon DEFAULT CHARACTER SET utf8;**

**create database hmon DEFAULT CHARACTER SET utf8;**

**create database rman DEFAULT CHARACTER SET utf8;**

**create database smon DEFAULT CHARACTER SET utf8;**

**create database nav DEFAULT CHARACTER SET utf8;**

**create database navms DEFAULT CHARACTER SET utf8;**

**create database metastore DEFAULT CHARACTER SET utf8;**

**create database sentry DEFAULT CHARACTER SET utf8;**

**create database oozie DEFAULT CHARACTER SET utf8;**

**create database hue DEFAULT CHARACTER SET utf8;**

**grant all on \*.\* TO 'root'@'jedge.hadoop.com' IDENTIFIED BY 'welcome1';   
grant all on \*.\* TO 'root'@'localhost' IDENTIFIED BY 'welcome1';**

**grant all on \*.\* TO 'root'@'%' IDENTIFIED BY 'welcome1';**

**flush privileges;**

**create user 'amon'@'%' identified by 'amon123';**

**grant all on amon.\* TO 'amon'@'%' IDENTIFIED BY 'amon123';**

**create user 'hmon'@'%' identified by 'hmon123';**

**grant all on hmon.\* TO 'hmon'@'%' IDENTIFIED BY 'hmon123';**

**create user 'rman'@'%' identified by 'rman123';**

**grant all on rman.\* TO 'rman'@'%' IDENTIFIED BY 'rman123';**

**create user 'smon'@'%' identified by 'smon123';**

**grant all on smon.\* TO 'smon'@'%' IDENTIFIED BY 'smon123';**

**create user 'nav'@'%' identified by 'nav123';**

**grant all on nav.\* TO 'nav'@'%' IDENTIFIED BY 'nav123';**

**create user 'navms'@'%' identified by 'navms123';**

**grant all on navms.\* TO 'navms'@'%' IDENTIFIED BY 'navms123';**

**create user 'hive'@'%' identified by 'hive123';**

**grant all on metastore.\* TO 'hive'@'%' IDENTIFIED BY 'hive123';**

**create user 'sentry'@'%' identified by 'sentry123';**

**grant all on sentry.\* TO 'sentry'@'%' IDENTIFIED BY 'sentry123';**

**create user 'oozie'@'%' identified by 'oozie123';**

**grant all on oozie.\* TO 'oozie'@'%' IDENTIFIED BY 'oozie123';**

**create user 'hue'@'%' identified by 'hue123';**

**grant all on hue.\* TO 'hue'@'%' IDENTIFIED BY 'hue123';**

**flush privileges;**

**mysql -u root -pwelcome1 -e "show databases;"**

**mysql -u root -pwelcome1 -e "select user from mysql.user;"**

**mysql -u root -pwelcome1 -e "select user,** **Select\_priv, from mysql.user;"**

To create schema for SCM:

**sudo service cloudera-scm-server stop**

**sudo service cloudera-scm-server-db stop**

Login to the node where your MYSQL is installed.

**cd /usr/share/cmf/schema/**

**sudo ./scm\_prepare\_database.sh mysql -h localhost -u root -pwelcome1 --scm-host localhost scm scm scm123**

This will create a file as below

**sudo cat /etc/cloudera-scm-server/db.properties**

**sudo cat /etc/cloudera-scm-server/db.mgmt.properties**

After successfully creating SCM schema, login to mysql and create below DBs for Cloudera installation. Delete existing properties in **/etc/cloudera-scm-server/db.mgmt.properties** and place new values as below.

**sudo vi /etc/cloudera-scm-server/db.mgmt.properties**

**com.cloudera.cmf.ACTIVITYMONITOR.db.type=mysql**

**com.cloudera.cmf.ACTIVITYMONITOR.db.host=cm.hadoop.com:3306**

**com.cloudera.cmf.ACTIVITYMONITOR.db.name=amon**

**com.cloudera.cmf.ACTIVITYMONITOR.db.user=amon**

**com.cloudera.cmf.ACTIVITYMONITOR.db.password=amon123**

**com.cloudera.cmf.REPORTSMANAGER.db.type=mysql**

**com.cloudera.cmf.REPORTSMANAGER.db.host=cm.hadoop.com:3306**

**com.cloudera.cmf.REPORTSMANAGER.db.name=rman**

**com.cloudera.cmf.REPORTSMANAGER.db.user=rman**

**com.cloudera.cmf.REPORTSMANAGER.db.password=rman123**

**com.cloudera.cmf.NAVIGATOR.db.type=mysql**

**com.cloudera.cmf.NAVIGATOR.db.host=cm.hadoop.com:3306**

**com.cloudera.cmf.NAVIGATOR.db.name=nav**

**com.cloudera.cmf.NAVIGATOR.db.user=nav**

**com.cloudera.cmf.NAVIGATOR.db.password=nav123**

**com.cloudera.cmf.NAVIGATORMETASERVER.db.type=mysql**

**com.cloudera.cmf.NAVIGATORMETASERVER.db.host=cm.hadoop.com:3306**

**com.cloudera.cmf.NAVIGATORMETASERVER.db.name=navms**

**com.cloudera.cmf.NAVIGATORMETASERVER.db.user=navms**

**com.cloudera.cmf.NAVIGATORMETASERVER.db.password=navms123**

After this, restart cloudera-scm-server

**sudo chkconfig cloudera-scm-server on**

**sudo chkconfig cloudera-scm-server-db on**

**sudo service cloudera-scm-server restart**

**sudo service cloudera-scm-server-db restart**

**sudo service cloudera-scm-server status**

**sudo service cloudera-scm-server-db status**

**sudo telnet cm.hadoop.com 7180**

[http://cm.hadoop.com:7180](http://http://cm.hadoop.com:7180)

If you want to install Cloudera using parcels, run below command. Otherwise continue with next step.

**sudo cp /var/www/html/cloudera-parcel/\* /opt/cloudera/parcel-repo**

**11. Setup Password-less SSH for hadoop cluster installation**

Configure passwordless SSH between the master and slave nodes. Namenode should be able to ssh to other datanodes without providing password. Make sure only namenode is able to ssh other datanodes without providing password. For this, login to Namenode.

mkdir -p ~/.ssh

cd ~/.ssh

sudo rm -f \*

ssh-keygen -t rsa

Press **Enter** without providing any password for each step.

Above step will create 2 files in **.ssh** folder. Copy public key to to authorized keys.

cd ~/.ssh

cat id\_rsa.pub >> authorized\_keys

sudo chmod go-w $HOME $HOME/.ssh

sudo chmod 600 $HOME/.ssh/authorized\_keys

sudo chown hdpuser:hdpadmin -R $HOME/.ssh/

cd

ssh-copy-id -i ~/.ssh/id\_rsa.pub 192.168.1.111

ssh-copy-id -i ~/.ssh/id\_rsa.pub 192.168.1.112

ssh-copy-id -i ~/.ssh/id\_rsa.pub 192.168.1.113

ssh-copy-id -i ~/.ssh/id\_rsa.pub 192.168.1.114

ssh-copy-id -i ~/.ssh/id\_rsa.pub 192.168.1.115

12. To configure Clustershell, go to

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

cm: 192.168.1.111

nn: 192.168.1.112

dn: 192.168.1.113 192.168.1.114 192.168.1.115

all: 192.168.1.111 192.168.1.112 192.168.1.113 192.168.1.114 192.168.1.115

13. To configure NTPD Service

clush -g all -b "sudo sed -i 's/^server /#server /g' /etc/ntp.conf"

clush -g all -b "sudo sed -i '/^#server 192.168/d' /etc/ntp.conf"

clush -g all -x 192.168.1.112 -b "echo 'server 192.168.1.112 prefer' | sudo tee -a /etc/ntp.conf > /dev/null 2>&1"

Add below line to NTP server (Without internet access):

sudo vi /etc/ntp.conf

server 127.127.1.0

fudge 127.127.1.0 stratum 10

Add below line to NTP server (With internet access):

sudo vi /etc/ntp.conf

restrict 192.168.1.0 netmask 255.255.255.0 nomodify notrap

clush -g all -b "sudo /etc/init.d/ntpd restart"

clush -g all -x 192.168.1.112 -b "/usr/sbin/ntpdate -d 192.168.1.112"

clush -g all -x 192.168.1.112 -b "/usr/sbin/ntpq -p"

clush -g all -b "date"

14. Open Cloudera Manager in browser by typing [http://cm.hadoop.com:7180](http://http://cm.hadoop.com:7180)

Enter user and password as **admin** and **admin**

Start installing Cloudera :

Custom Repository Location:

http://192.168.1.111/cloudera-cdh5/

Custom GPG Key URL:

http://192.168.1.111/cloudera-cdh5/RPM-GPG-KEY-cloudera

15. If you see any error in Host Inspector, run below.

sudo sh -c "echo never > /sys/kernel/mm/transparent\_hugepage/defrag"

sudo -u hdfs hdfs dfsadmin -report