**Hadoop Path - A Multi node cluster Setup(Without External Databases)**

**Instance creation**

**Step-1**

Create an account at <https://console.cloud.google.com>

Provide your credit or Debit card details

**Step-2**

In Google cloud select Top-Left menu

Select option Compute engine

You should see create instances active on right side panel of screen

**Step-3**

Select Create instance option

Name Node

Give name as NameNode for first instance(Any name is fine.For our convenience)

Select Zone according to your requirement(You can select any one in list.)

(Make sure that no two instances fall in the same region)

Select machine type according to your requirement(2 cpus will be fine)

Select your convenient operating system(Make sure you don’t forget the selected operating system and its version)

Check the boxes Allow http and https boxes respectively

Click on create

DataNode

Select Create instance option

Follow the steps same as the Name node and create three(Can create any number ) instances with different names as datanode1,danode2,datanode3(again any name is fine.For convenience) respectively.

**Configuring the Nodes**

Click on **SSH** option of every instance you have created

**Step-1(Root login and Password setup)**

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Do the following on every instance(Node) \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Type command **su root** now you have moved to root

Type command **passwd** to change your root password

Enter any password and re-enter it

**Step-2(Passwordless ssh)**

This step is used for communication between Namenode and datanodes without any authorization.For which we generate the public key in Namenode and copy the same as authorized\_keys in our data nodes.

Now move to home directory in every instance with **su ~**

In Every instance type command **ssh-keygen** hit enter at every stop

In NameNode

change directory with **cd .ssh**

Now open public key file using **vi id\_rsa.pub**

Copy the data present in opened file(Warning:Copy the file properly,if not you will face errors)

In Every Data Node

change directory with **cd .ssh**

Create a file using **vi authorized\_keys**

Now press **i** to use insert option

Now paste the public key copied from NameNode(Warning:Paste properly make sure it is same as the id\_rsa.pub of NameNode)

**Step-3(Changing SSHD configuration)**

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*In all Instances \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

use **cd ~**  to move to home

Now type **vi /etc/ssh/sshd\_config**

Now hit **i** to move to insert mode

Now parse through the file and perform the below changes

Password authentication : yes

Permit root login : yes

Now hit **Escape(Esc)** and type **:wq**  and hit enter

Now restart service using **service sshd restart**

**Step-4(Disabling Firewall)**

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*In all Instances \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

use **cd ~**  to move to home

Check if firewall is active using **service firewalld status**

If firewall is active,using **service firewalld stop** stop the firewall

To make it always off use **chkconfig firewalld off**

**Step-5(Installing Server using HTTPD)**

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*In all Instances \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

use **cd ~**  to move to home

Using **yum -y install httpd** install the httpd

Now start httpd service using **service httpd start**\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*In all Instances \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

use **cd ~**  to move to home

To make it always ON use **chkconfig httpd on**

**Step-6(Disabling SeLinux)**

Now type **vi /etc/sysconfig/SeLinux**

Hit **i**  to move to insert mode and perform below changes

enforcing---> disabled

Now hit **Escape(Esc)** and type **:wq**  and hit enter

**Step-7(Checking password less login)**

In **NAMENODE**

use **cd ~**  to move to home

Now type **ssh datanode1** and hit enter at every stop.

You should be able to login to datanode1 without any password

\*\*\*\*\*Check the above steps in **NAMENODE** for every **DataNode** you have created\*\*\*\*\*\*

\*\*\*\*\*If any ERROR in the logging in then re do the **STEP-2** properly\*\*\*\*\*\*\*\*

(during copy paste in **step-2** sometimes you may encounter **new lines.**So,make sure they are same**)**

**Step-8(Creating Firewall Rule)**

In your google console page,to the top left corner hit on Menu

Select Networking option

In left side of your page hit on firewall rules

In right side of your page now select create firewall rule

Perform following changes

Name:test

Allow from any source (0.0.0.0/0)

Allowed protocol ports:tcp:0-65000

And save the firewall rule

**Step-9(Installing Cloudera on NameNode)**

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*only on NameNode\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

use **cd ~**  to move to home

Now install wget using **yum -y install wget**

Now download the Latest hadoop CDH using below command

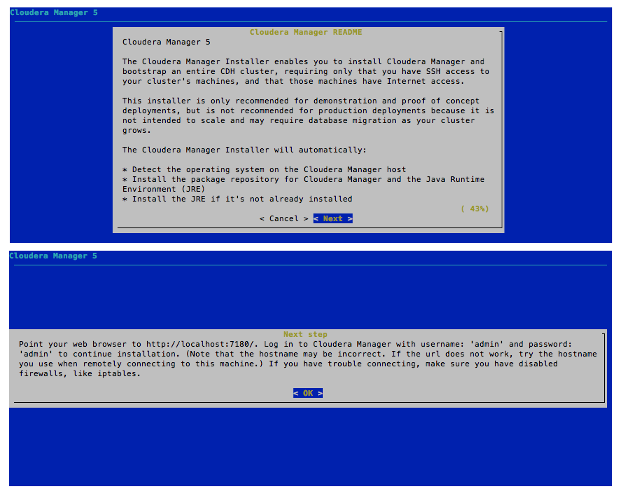
***wget*** [***https://archive.cloudera.com/cm5/installer/latest/cloudera-manager-installer.bin***](https://archive.cloudera.com/cm5/installer/latest/cloudera-manager-installer.bin)

Change the permissions of above file using ***chmod u+x cloudera-manager-installer.bin***

Now install the file using below command

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

If Everything Works Well you see below screen



Hit Next and then ok.Wait until it completes the installation.

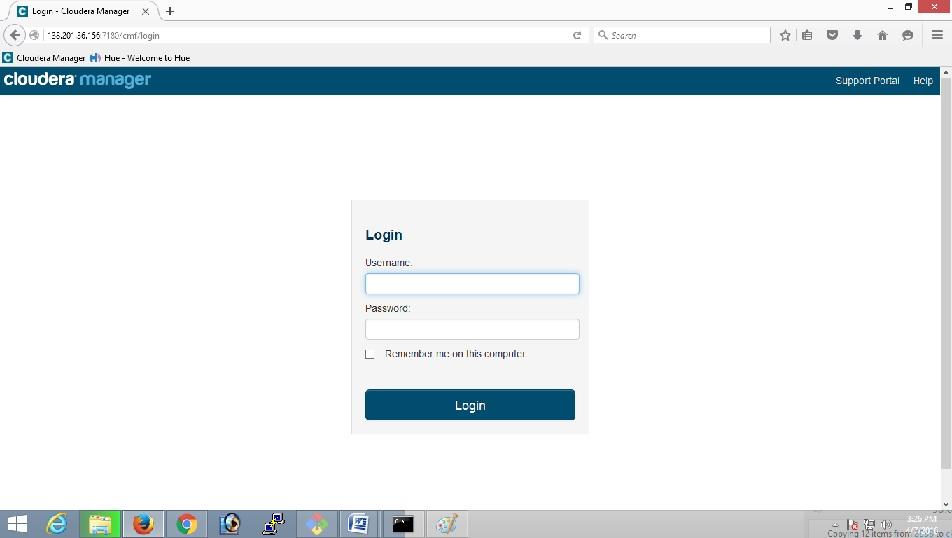
**Step-9(Logging in to your Cloudera Manager Admin Console)**

In NameNode use **sudo service cloudera-scm-server start** to start the cloudera server.

Now move to your Google Console Instances page and copy the external ip of your Namenode

Now in your browser open new window type **http://(paste copied external ip):7180** and hit enter

You should see the below screen



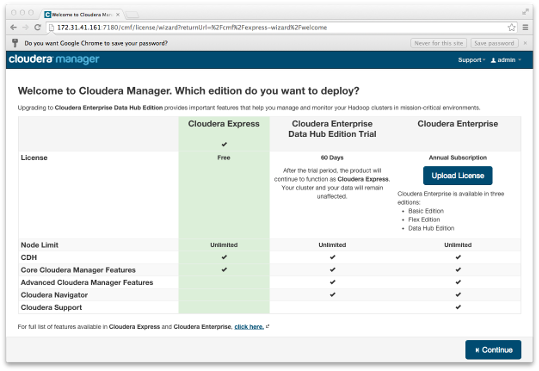
Now Input

Username:admin

Password:admin

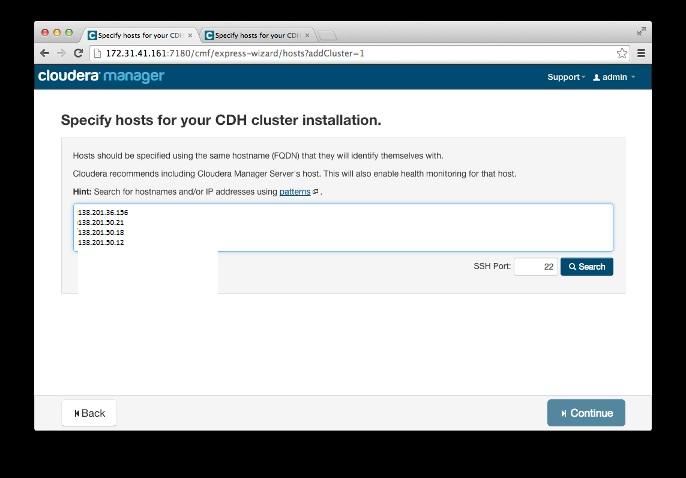
And hit login

Select the plan as required



And hit continue

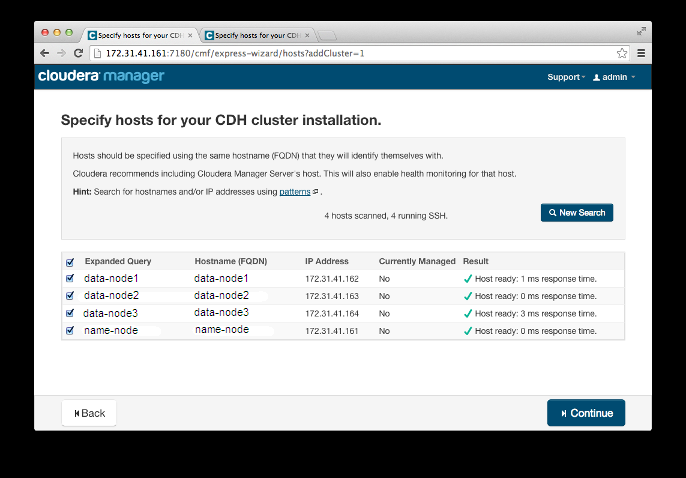
Now screen like below appears



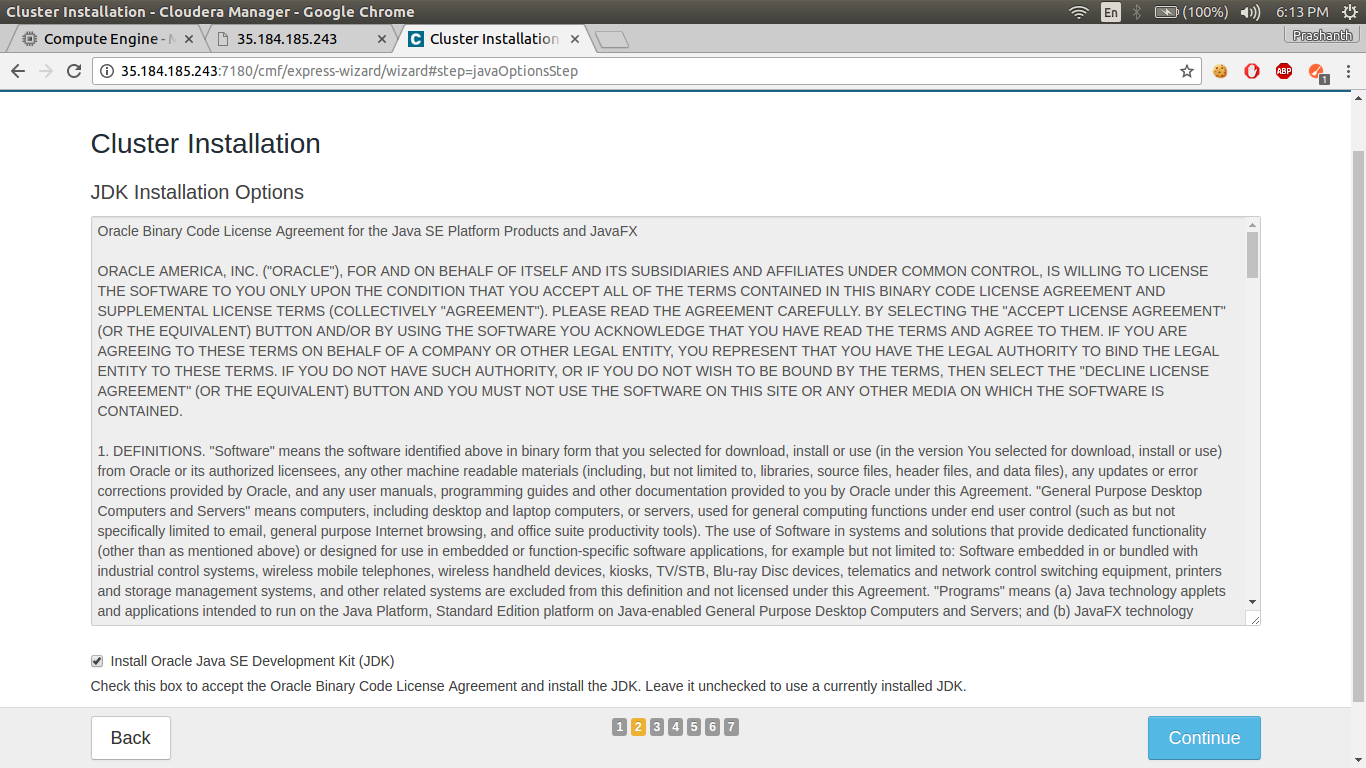
In the above window enter all your instance names or their external ip address

Now hit search,you must be able to see all your Namenode and DataNodes.

Select everything and hit continue

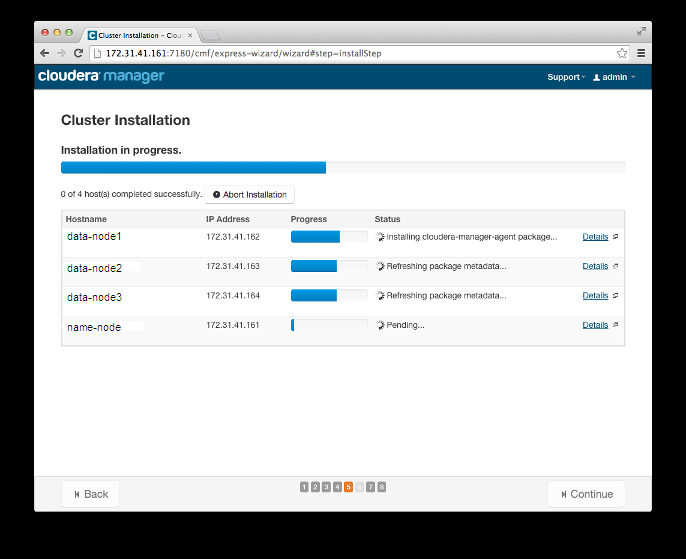


Select everything and hit continue.Continue until you see below screen



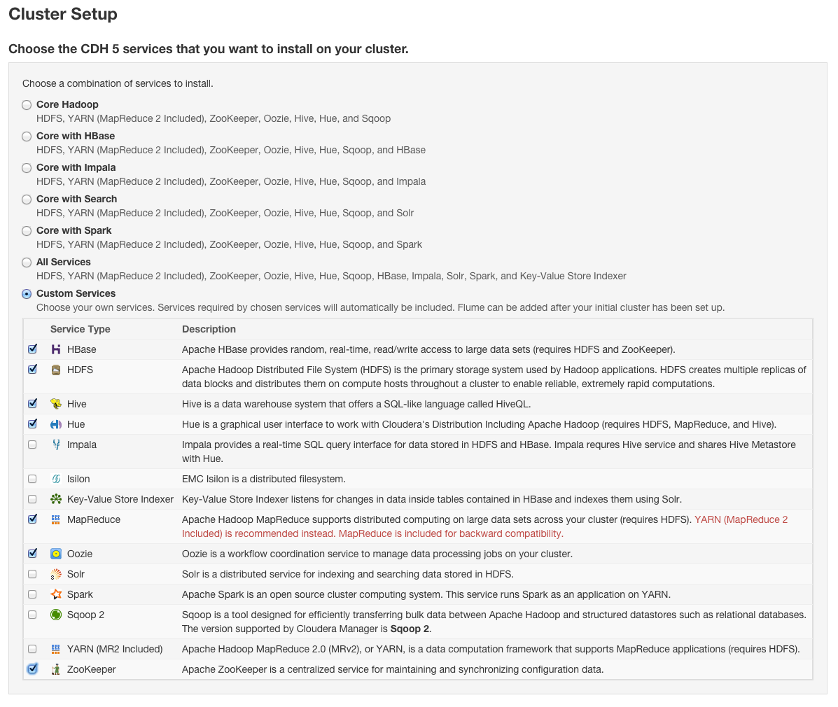
In the above screen select installing oracle jdk

Hit continue until you see below screen(don’t select or make any changes)



Wait until the above installation is complete and hit continue until you see the below screen

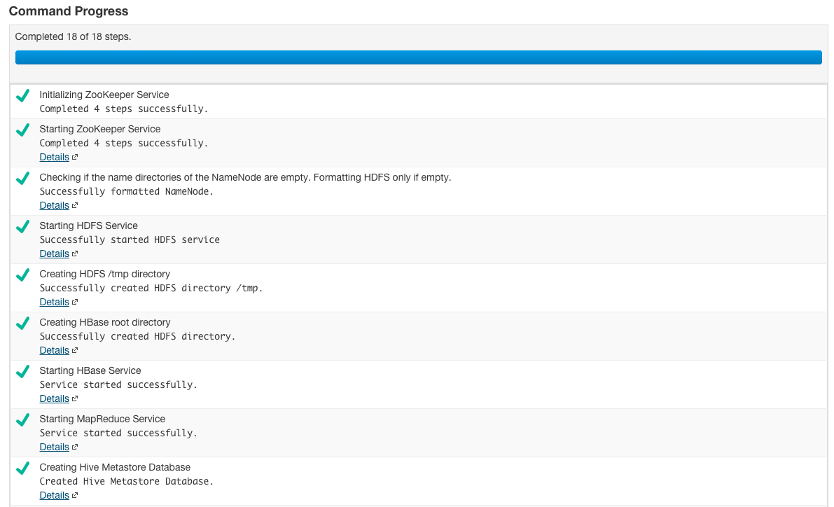
In the below screen select Core Hadoop option or You may customize your services.



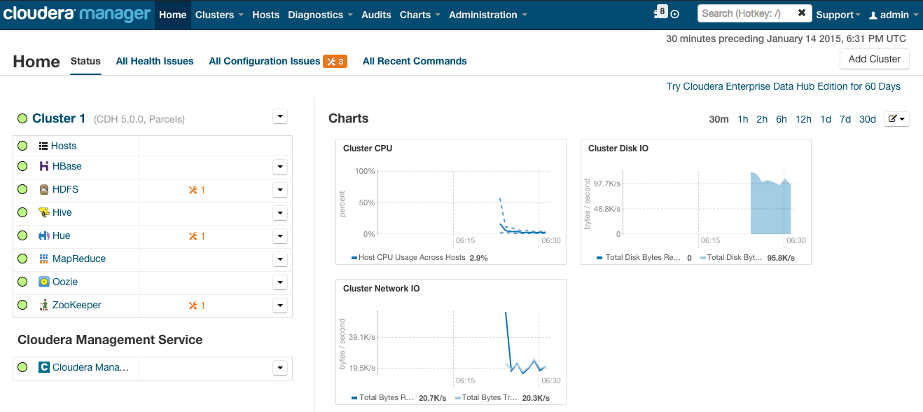
Now hit continue until you see your Database setup screen

In database setup screen hit Test Connection and wait until the test finishes(it May skip few)

Hit continue until you see the below screen.Wait until it finishes.



Once the above steps are done you see the below screen



\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*WELCOME TO YOUR CLOUDERA ADMIN CONSOLE\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*