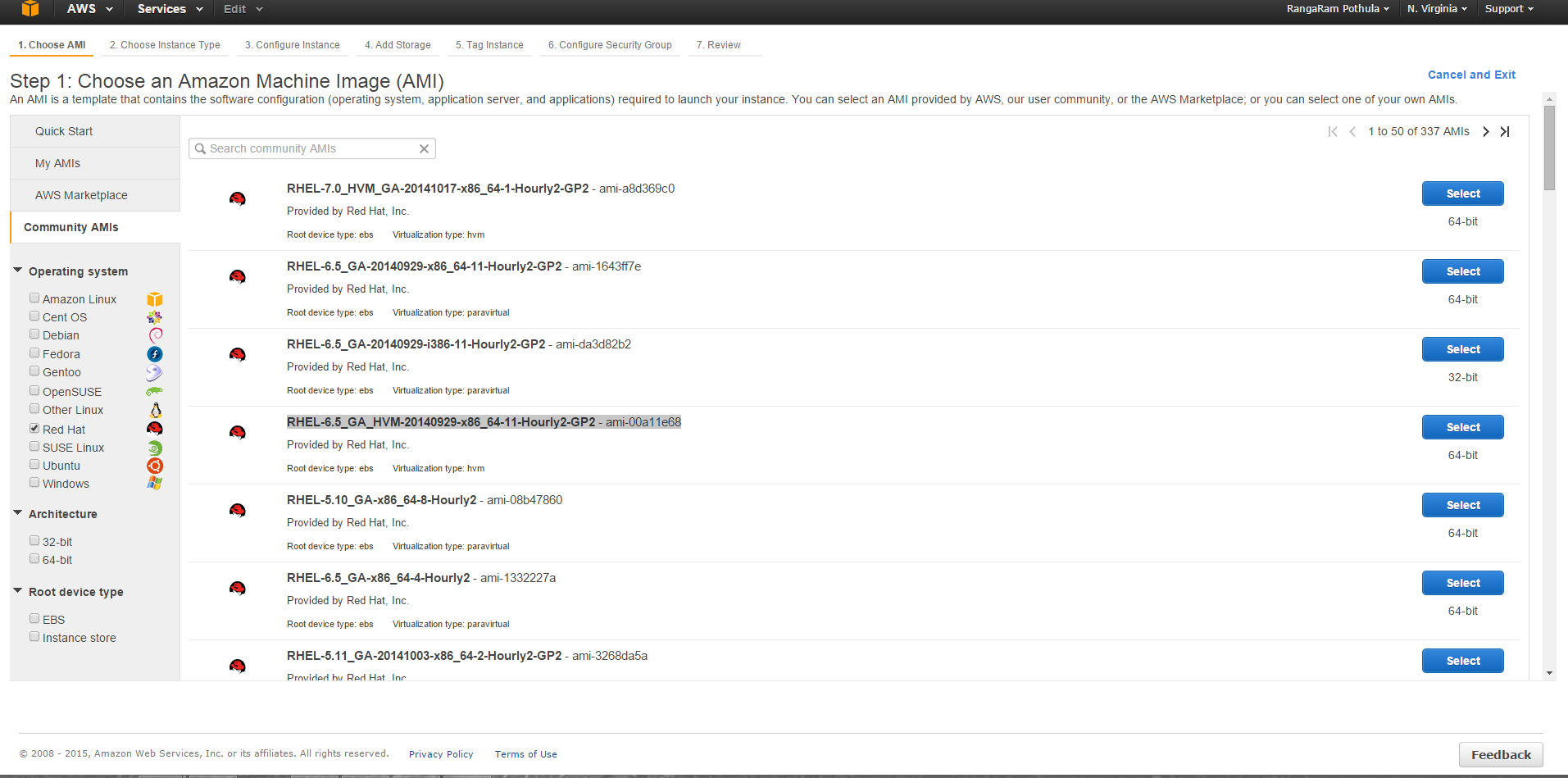
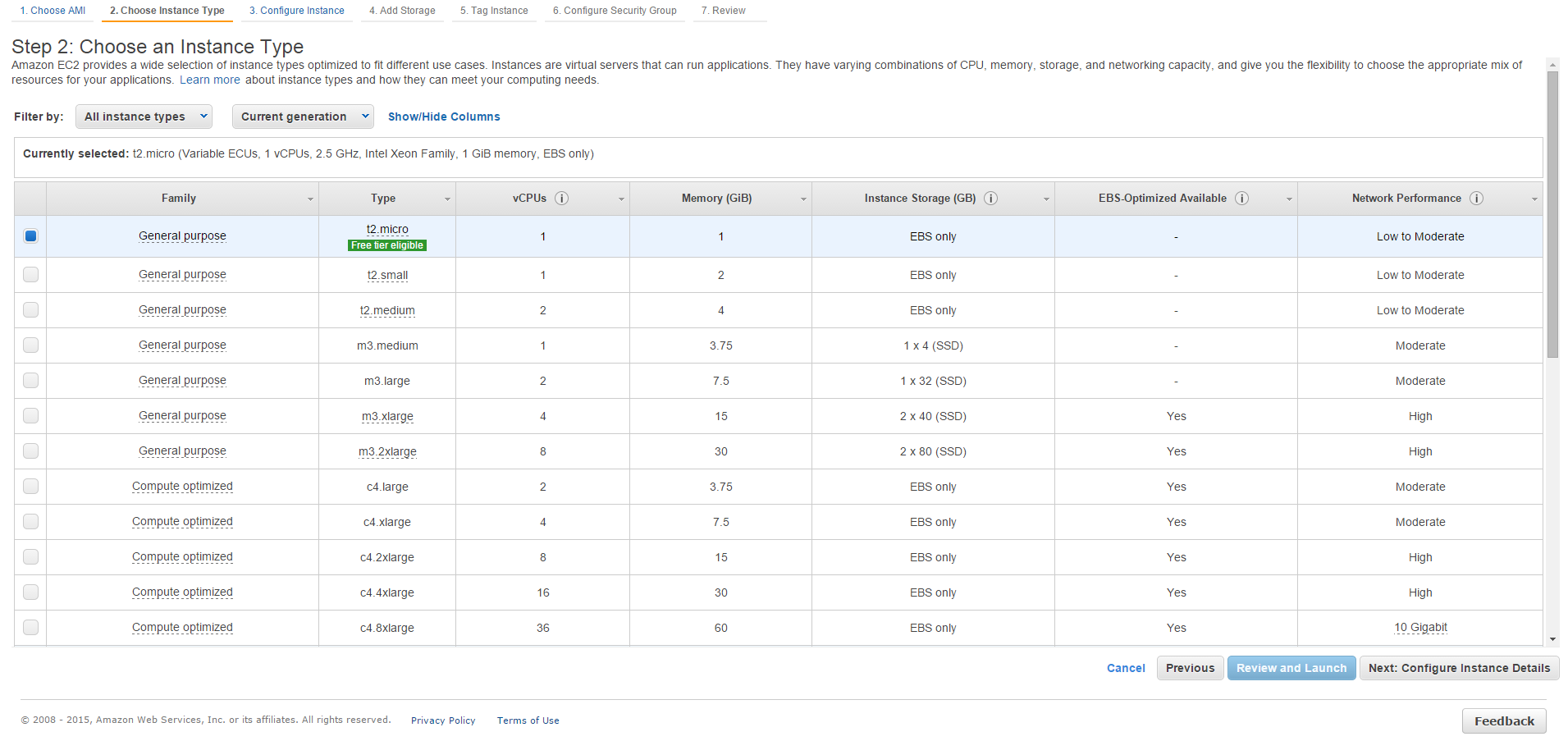
**Hadoop multi node cluster installation**

**Open EC2-Dashboard and launch new instances:**

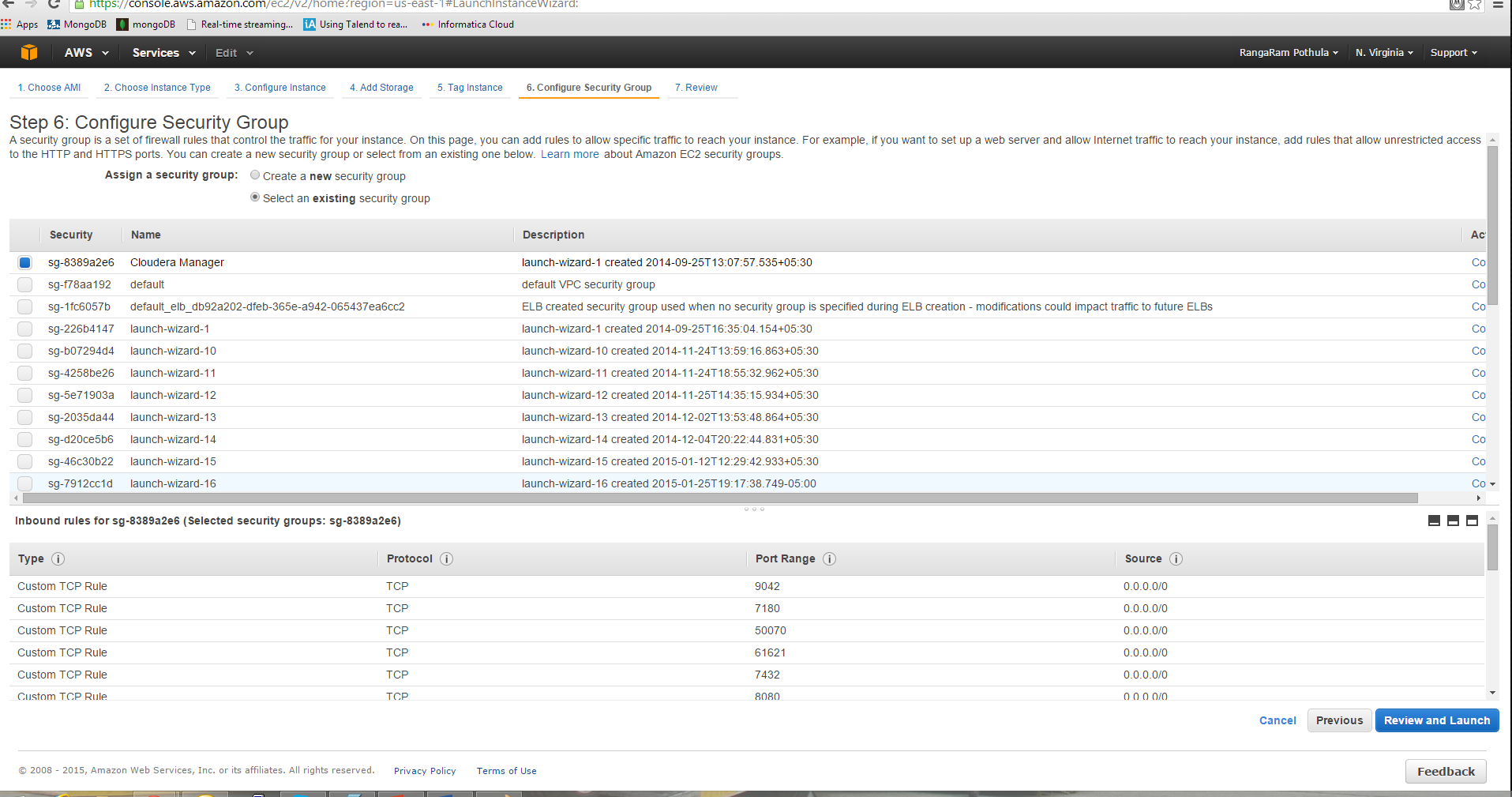
**RED-HAT 64BIT OS (RHEL-6.5\_GA\_HVM-20140929-x86\_64-11-Hourly2-GP2** - ami-00a11e68)

****

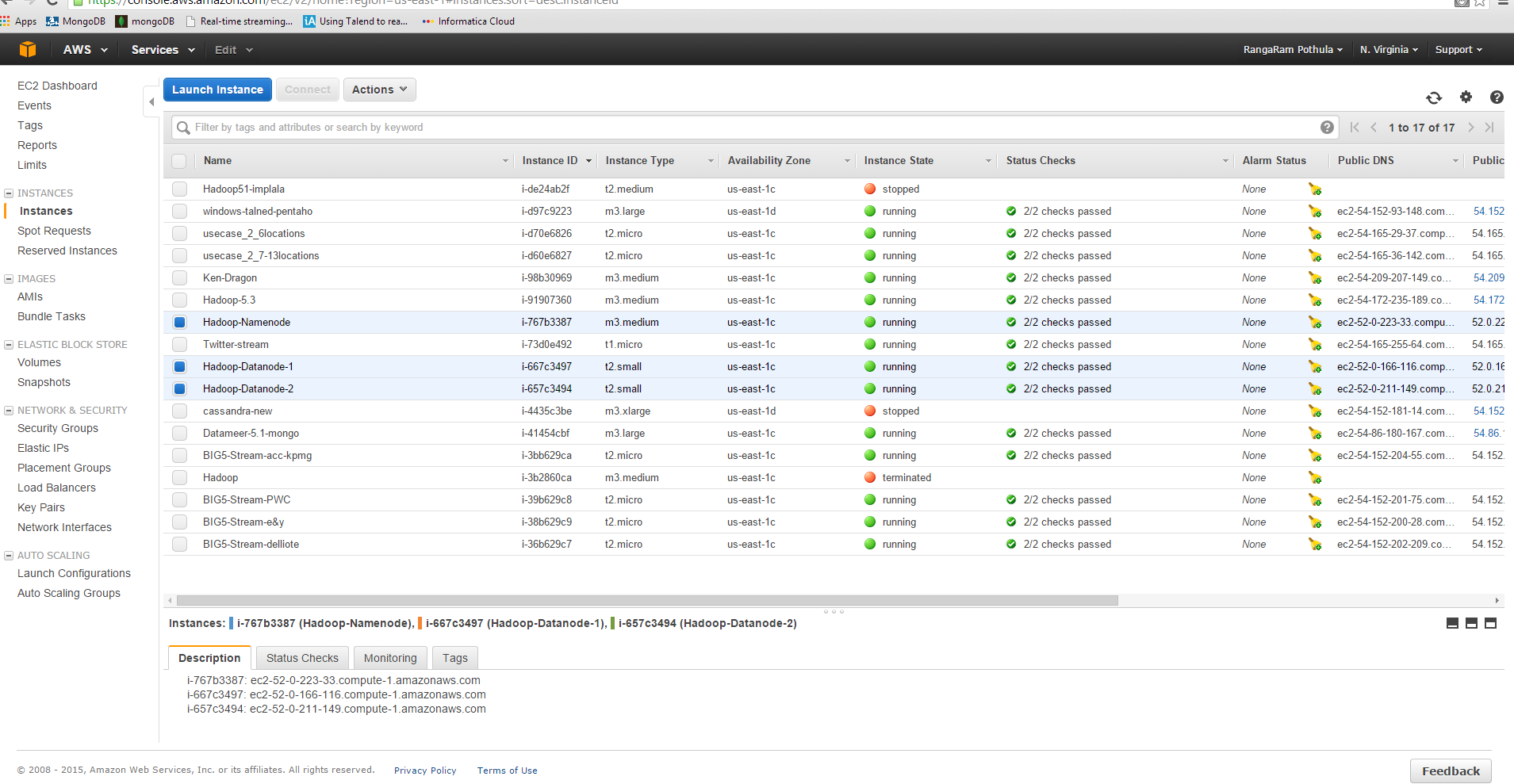
**Select RAM and processers for Namenode and datanode. Selct Namenode is Medium and datanodes small instances:**

****

**Proceed until the configuration section by default please select cloudera manager which is already existing security configurations:**

****

**And then launch instances:**

****

**Start cluster install using cloudera manager before processed please install Java and ssh configuration in all 3 instances and make them to communicate each other:**

**Install java in all 3 instances:**

**Please follow and set env variables in ~/.bashrc :**

[**http://tecadmin.net/steps-to-install-java-on-centos-5-6-or-rhel-5-6/**](http://tecadmin.net/steps-to-install-java-on-centos-5-6-or-rhel-5-6/)

**Edit the following file**vi /etc/sysconfig/network —add the following in this file  
NETWORKING=yes  
HOSTNAME=www. n1.com  
GATEWAY=192.168.1.1

**Disable Selinux an all nodes**  
vi /etc/selinux/config  
SELINUX=disabled —- # — change the value from enforcing to disabled

**Run the following command**  
chkconfig iptables off

Edit your hosts file  
vi /edit/hosts ——-add your hosts in the file



Use private ip and private DNS on every instance and same like add those details in all instances

**Setup SSH:**

**Please generate SSH in namenode and copy to all datanodes:**

**Do it this step under ec2-user:**

Now check that you can ssh to the localhost without a passphrase:

$ ssh localhost

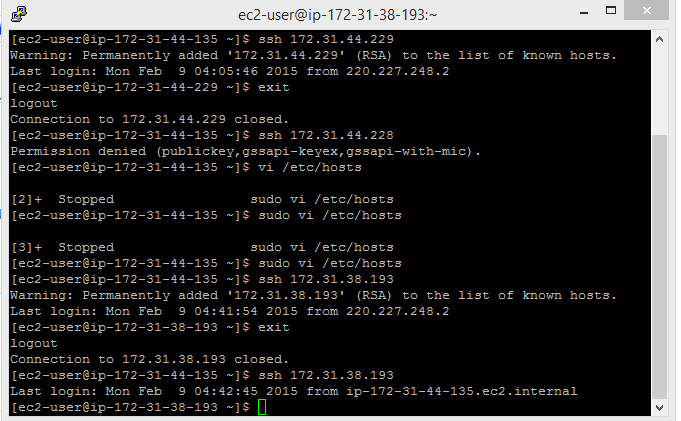
If you cannot ssh to localhost without a passphrase, execute the following commands:

$ ssh-keygen -t dsa -P '' -f ~/.ssh/id\_dsa

$ cat ~/.ssh/id\_dsa.pub >> ~/.ssh/authorized\_keys

Copy ssh pub and id\_dsa files from namenode to all datanodes. /home/ec2-user/.ssh/

And check ssh <datanode ip> if it communicates fine from name-node its good.

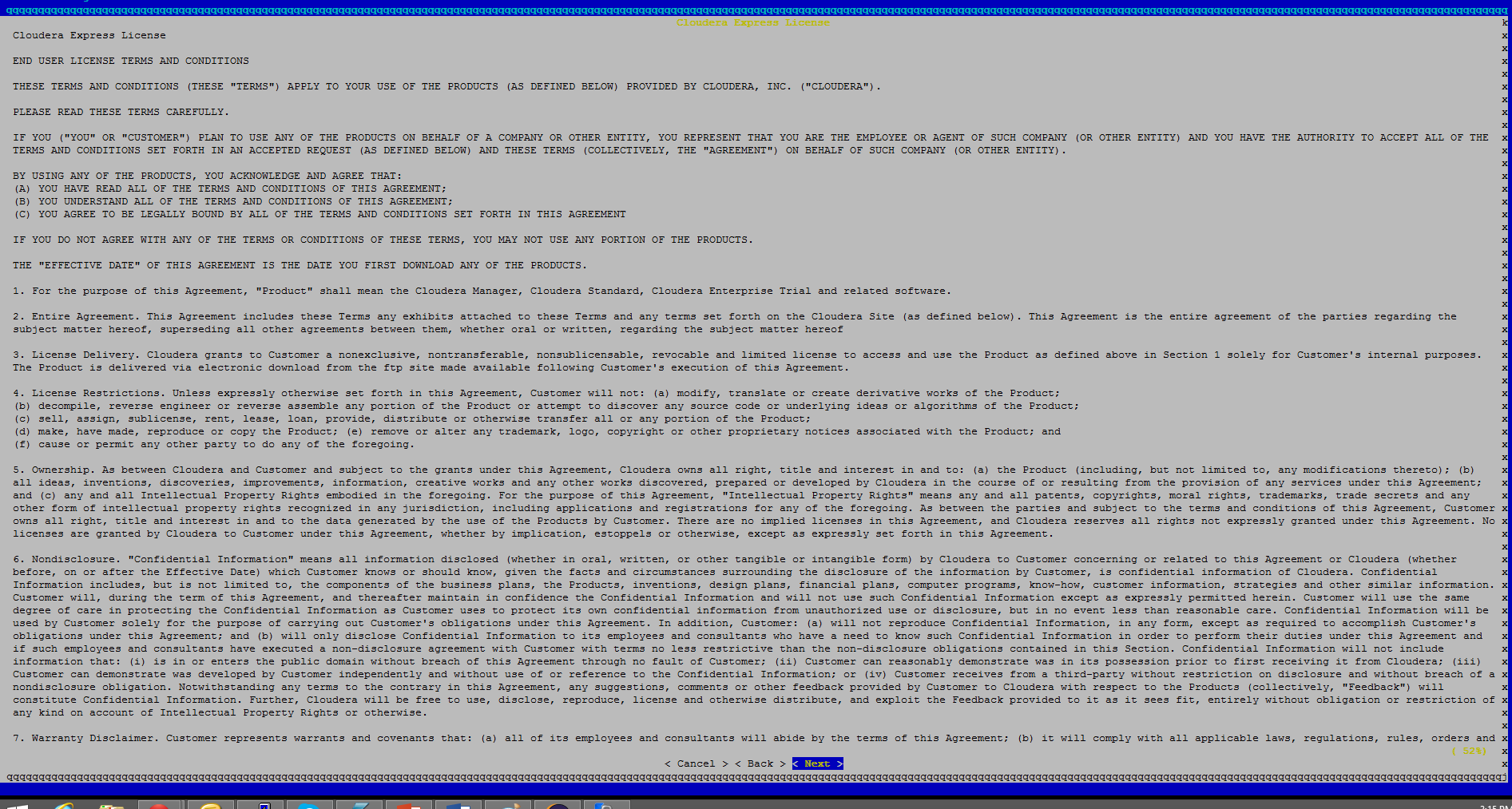


**Now install cloudera manager in namenode:**

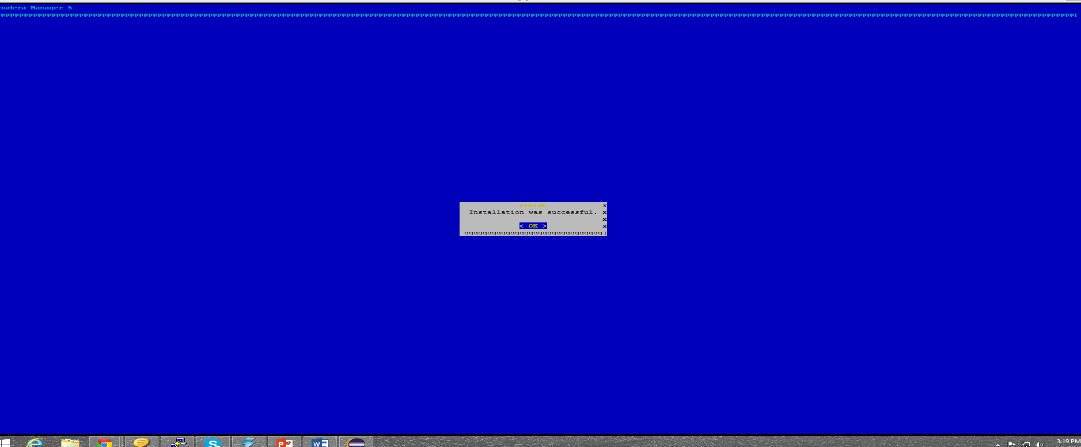
$ wget http://archive.cloudera.com/cm5/installer/latest/cloudera-manager-installer.bin

$ chmod u+x cloudera-manager-installer.bin

$ sudo ./cloudera-manager-installer.bin



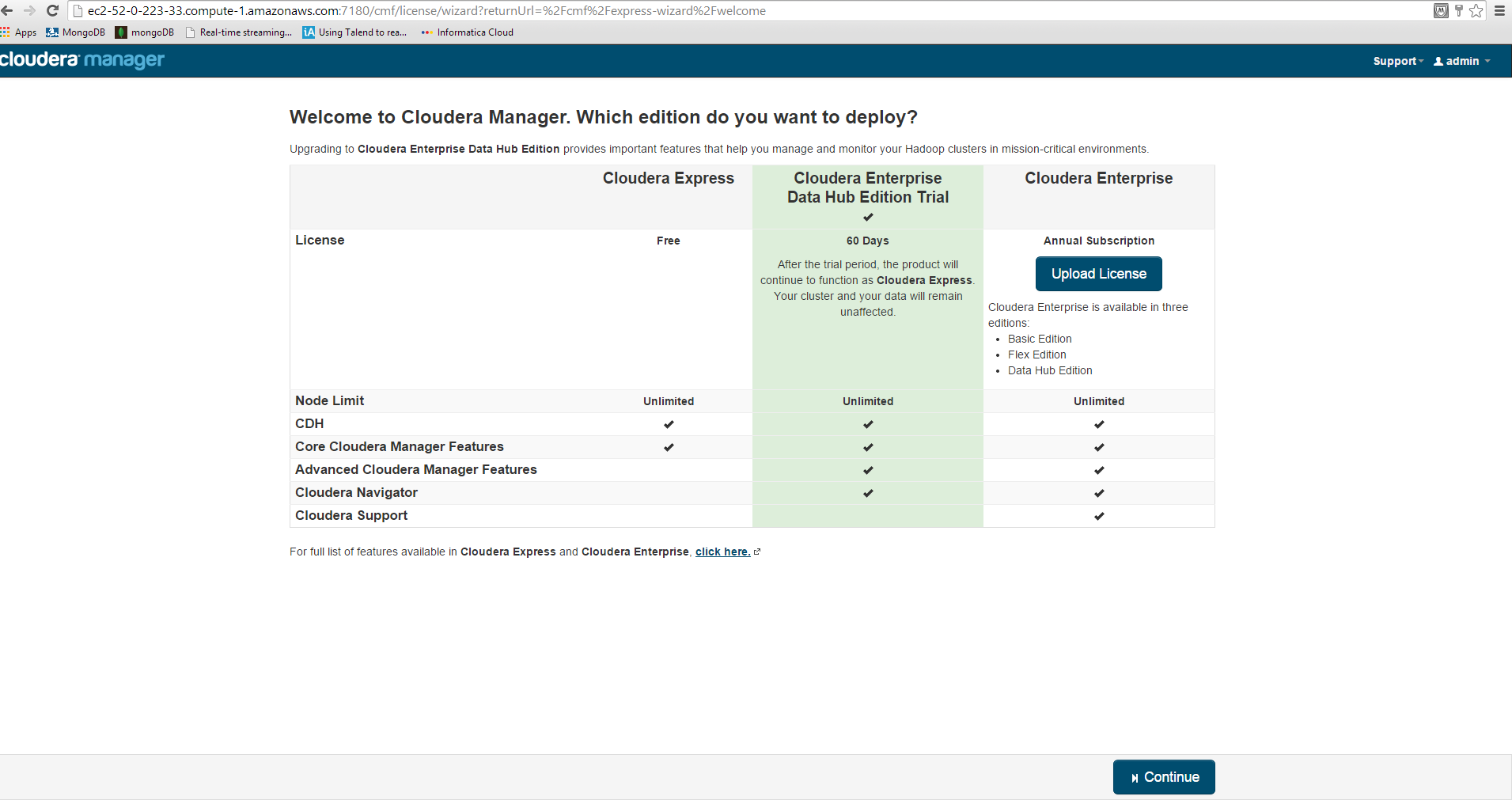
**Click ok and complete installation**

****

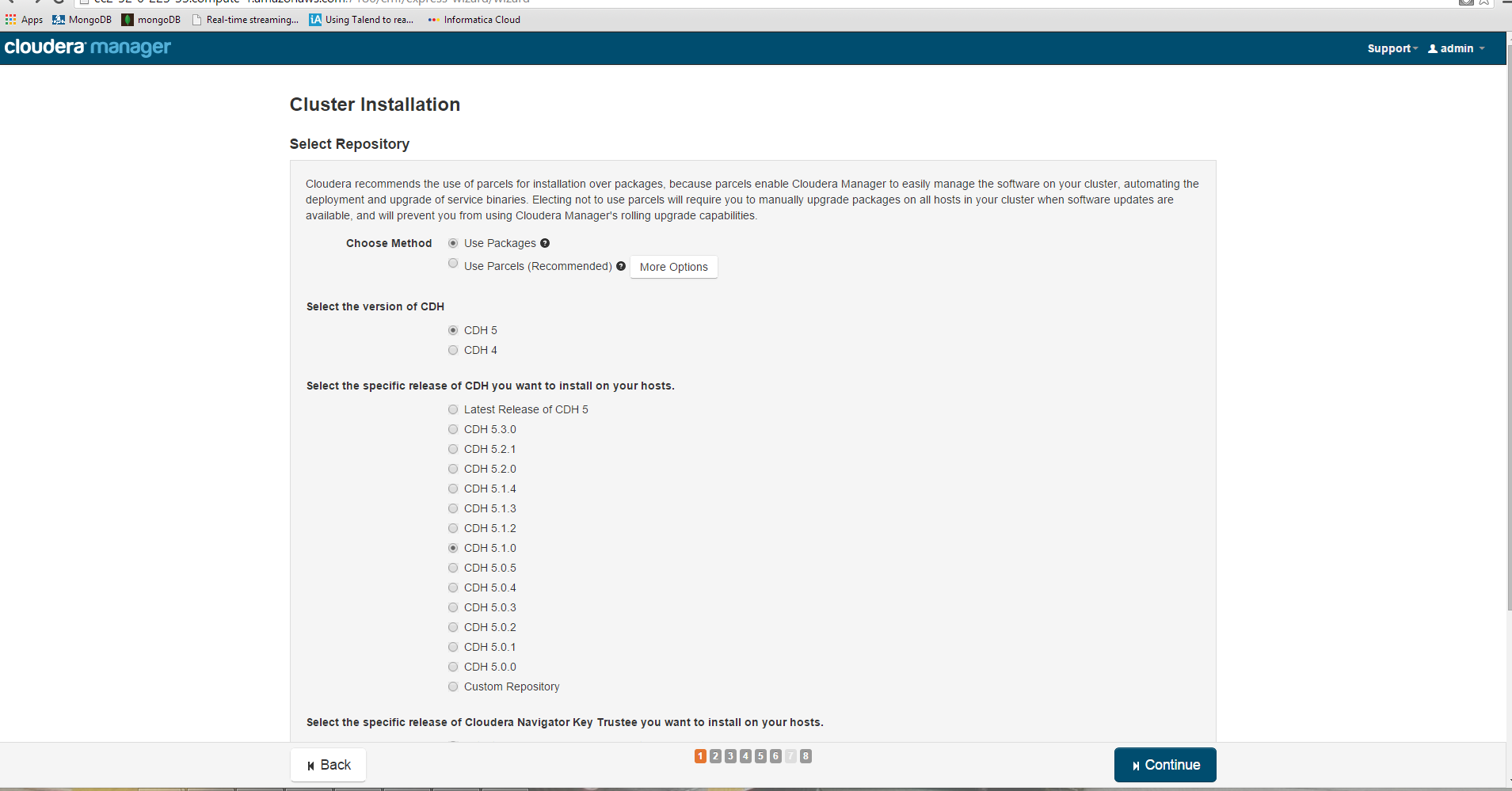
**Now install cloudera cluster:**

[**http://ec2-52-0-223-33.compute-1.amazonaws.com:7180/cmf/login**](http://ec2-52-0-223-33.compute-1.amazonaws.com:7180/cmf/login)

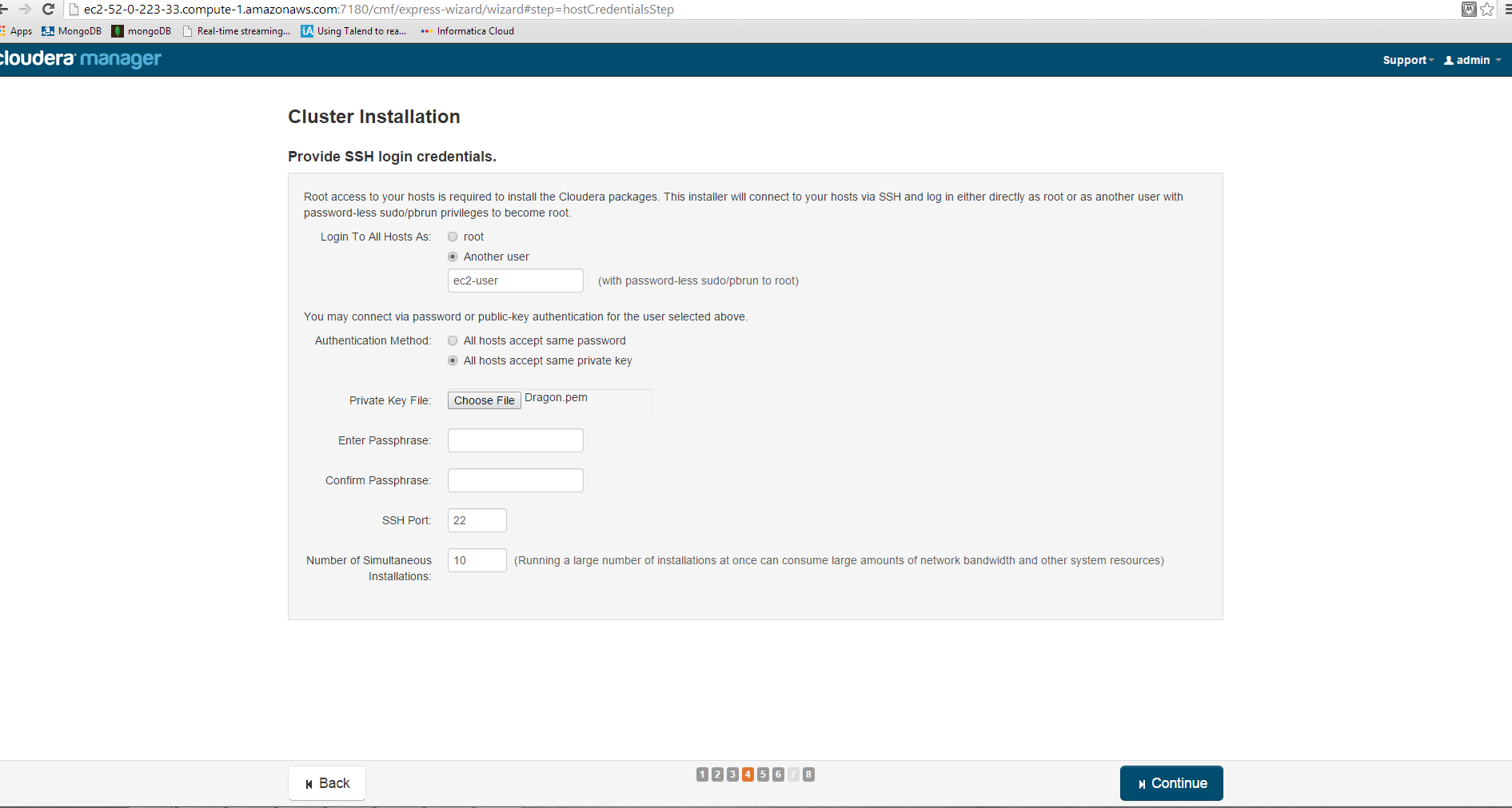
**use: admin/admin**

****

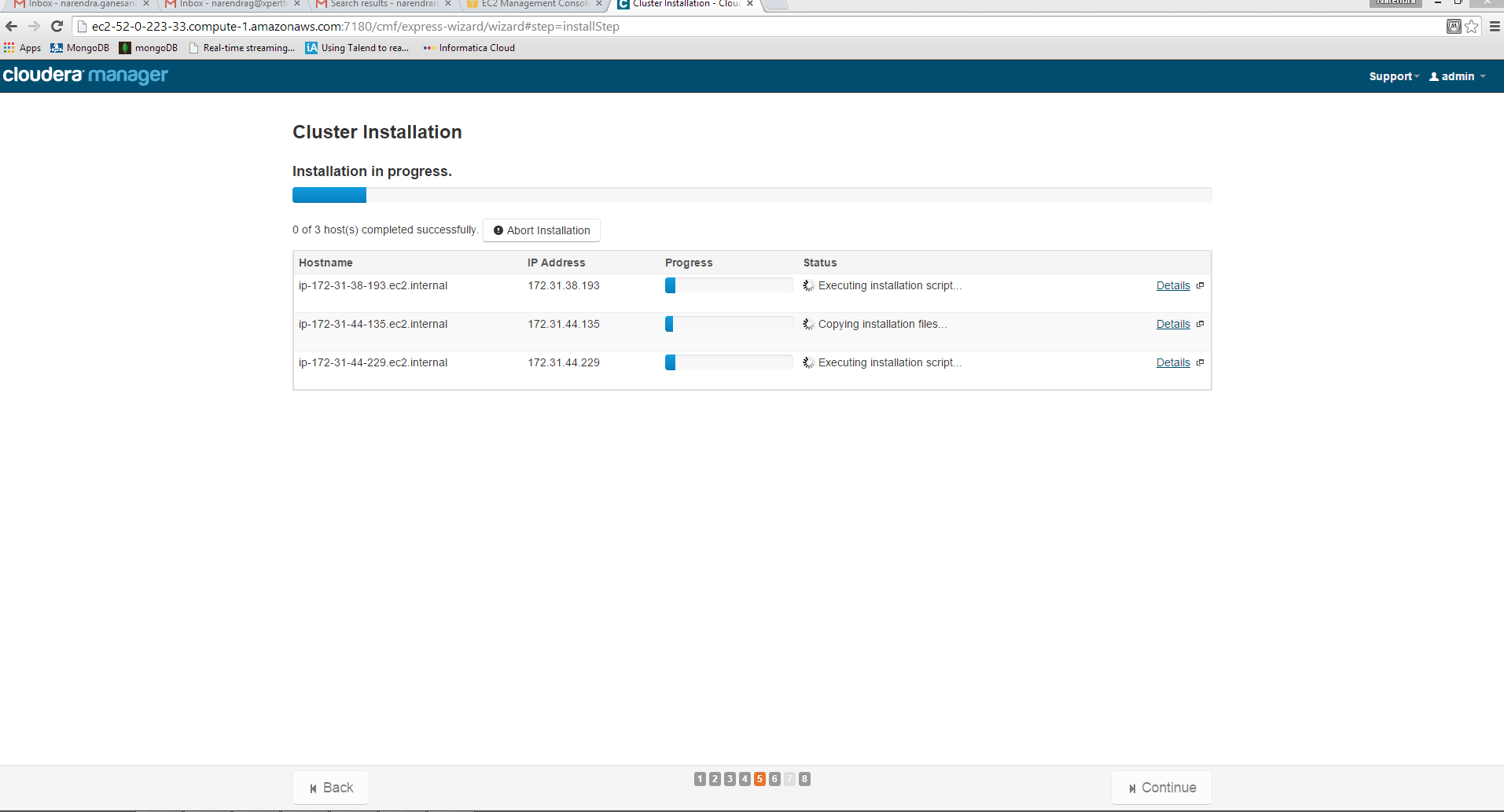
**Enter private ips of all 3 nodes**

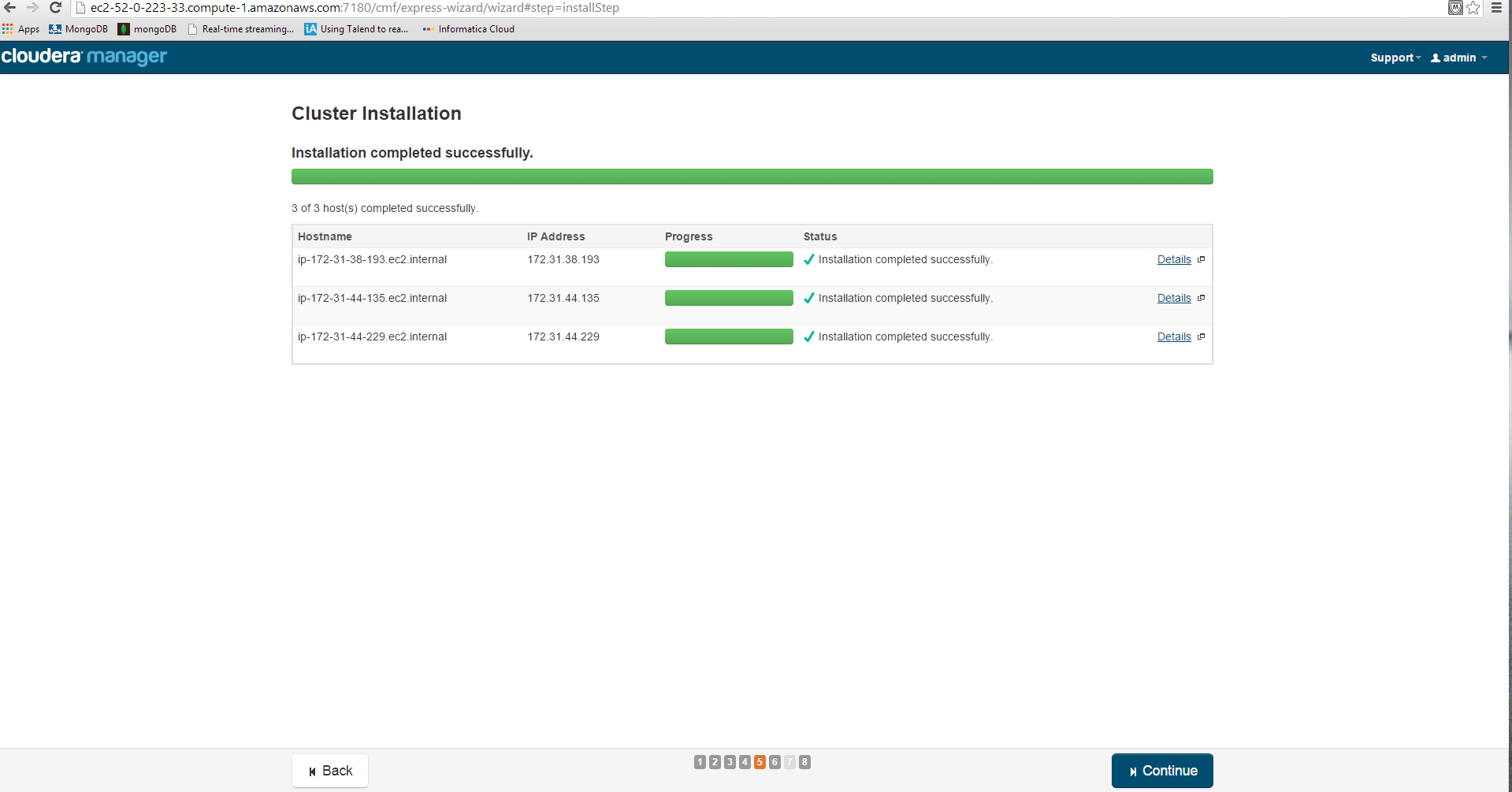
****

**Select cdh 5.1.0 , uncheck jdk, uncheck single user**

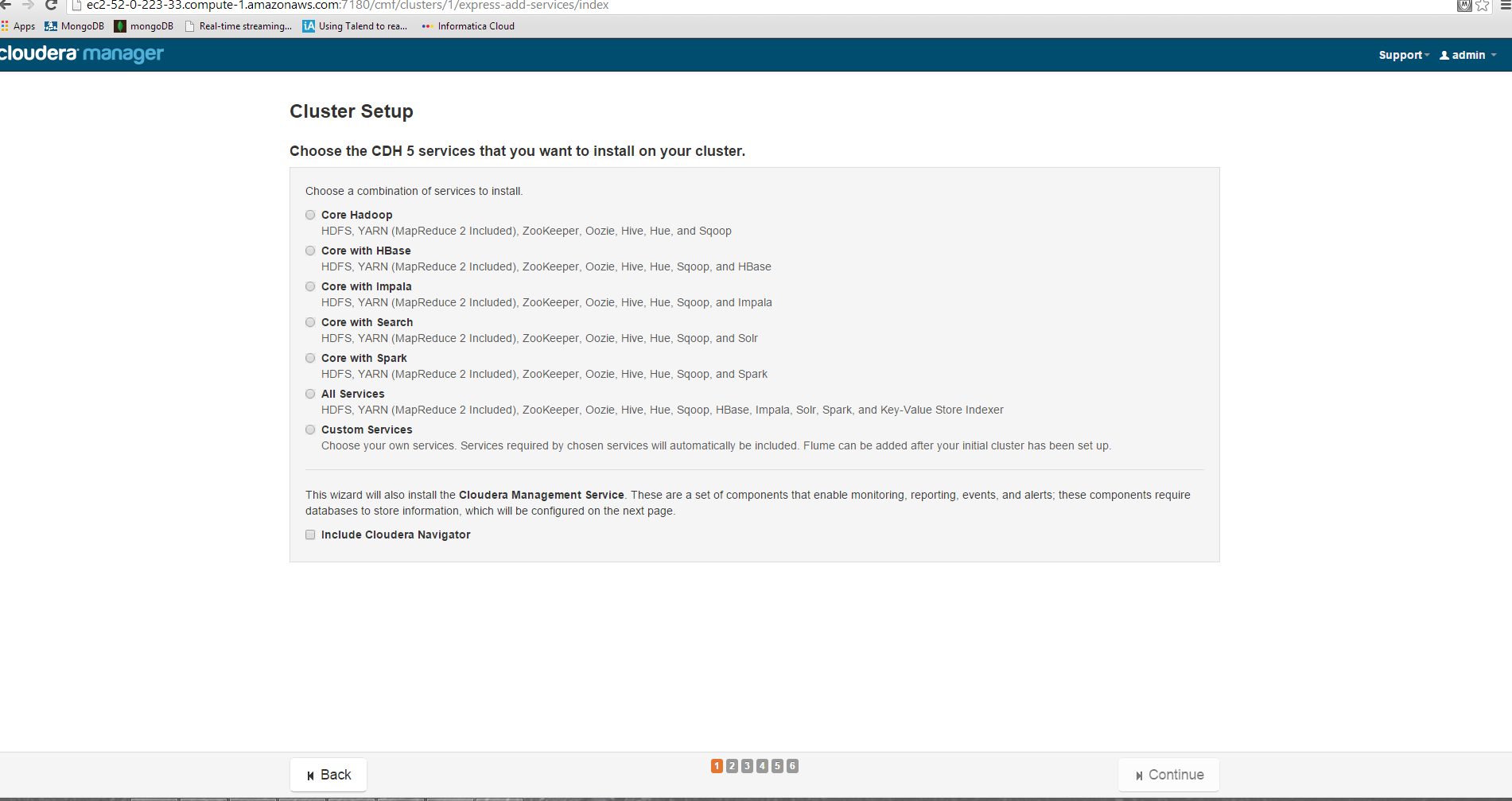
****

**Select user: ec2-user and pem file attach**

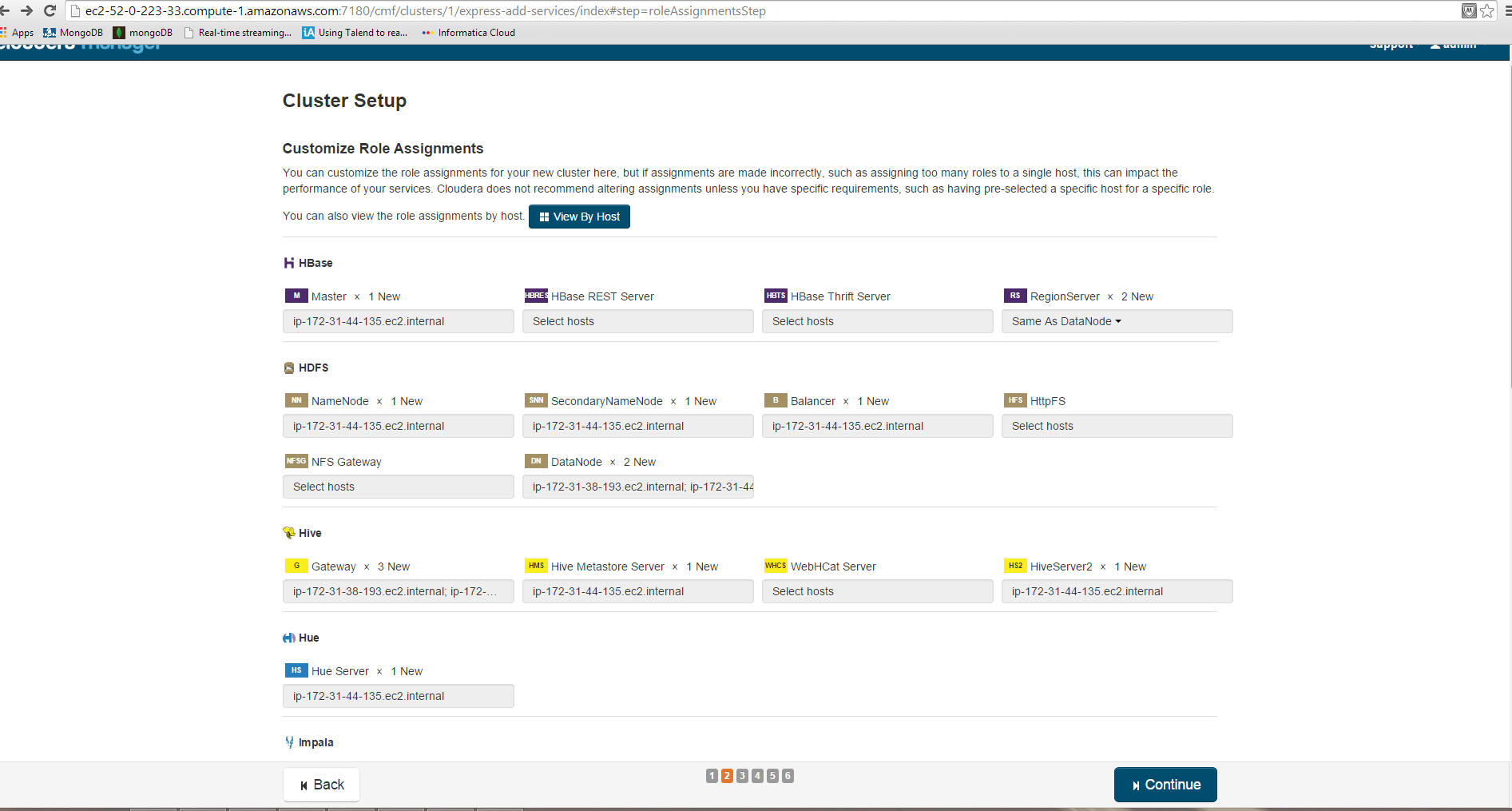
****

****

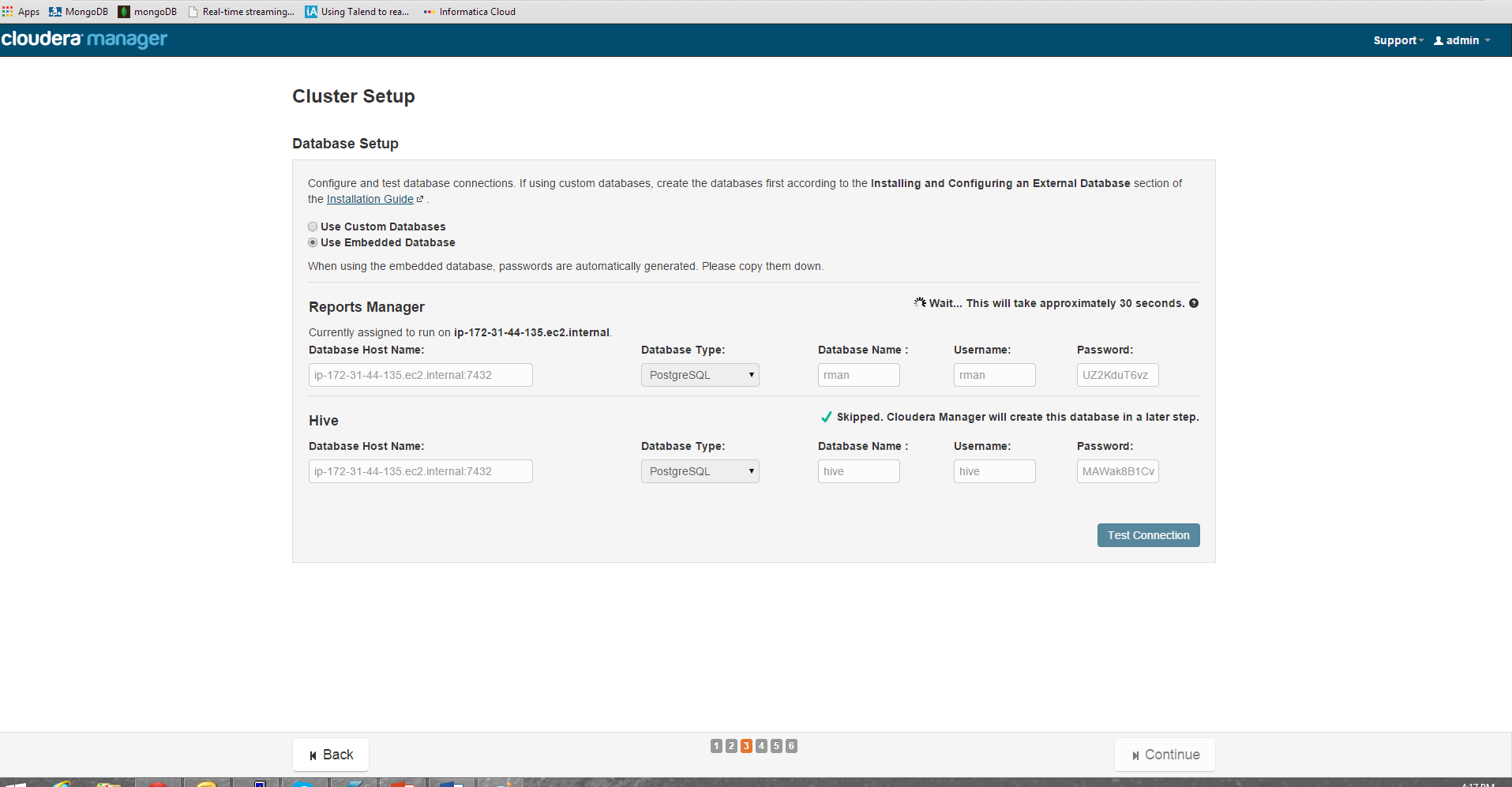
**And then continue till the following screen will appear**

****

**Choose all services and then continue**

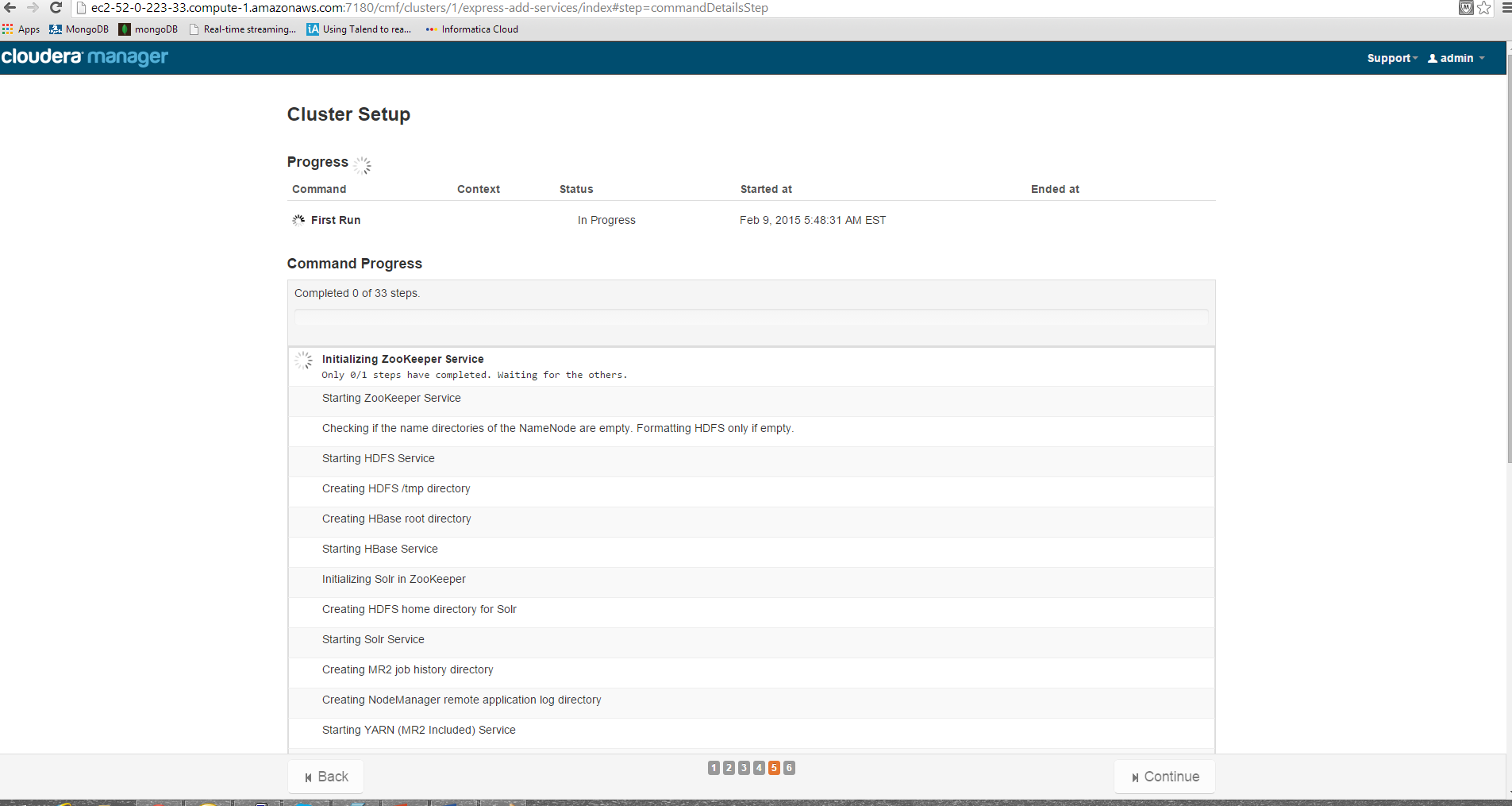
****

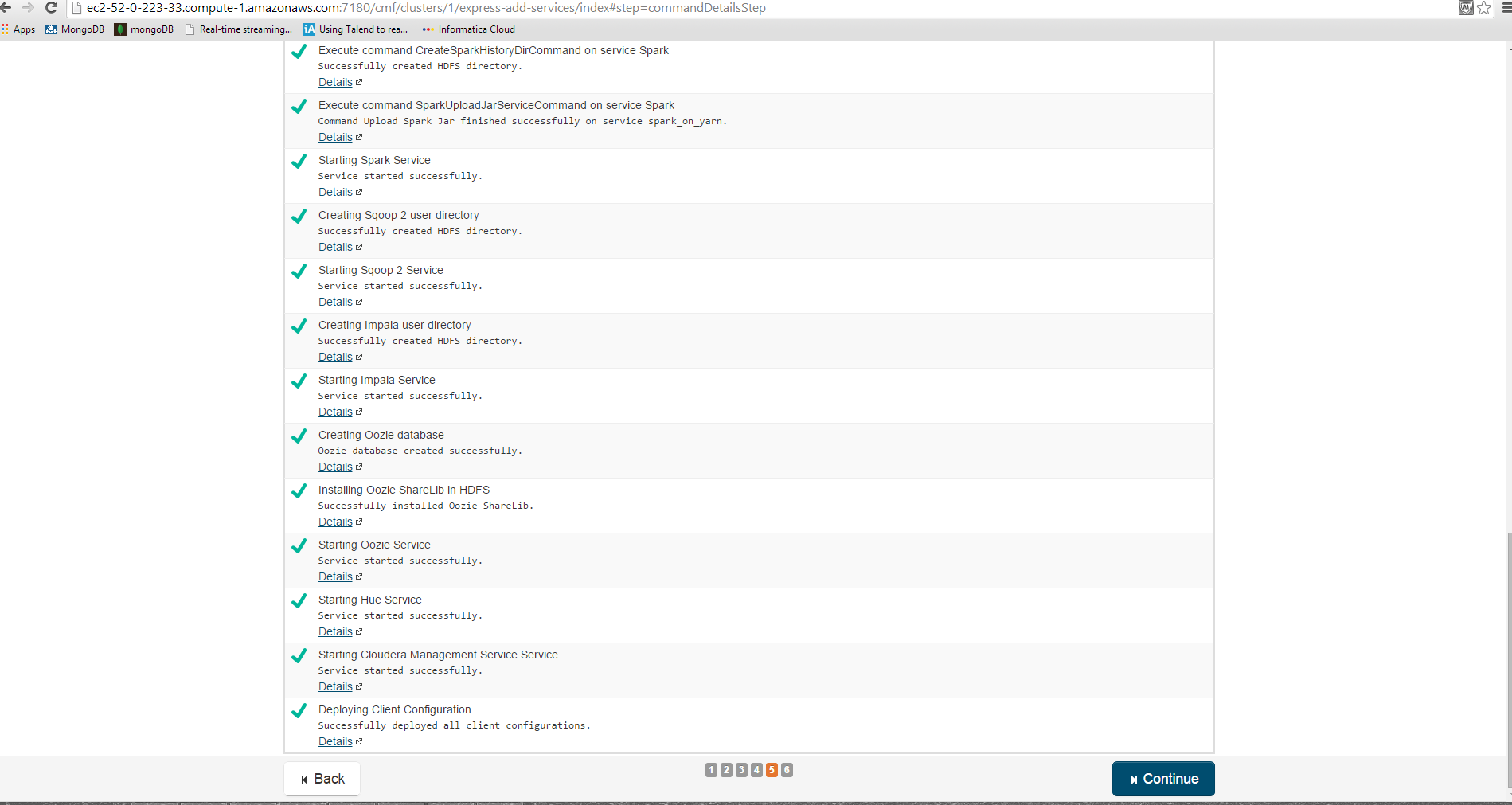
**Check all ips are correctly places to namenode and datanodes as well**

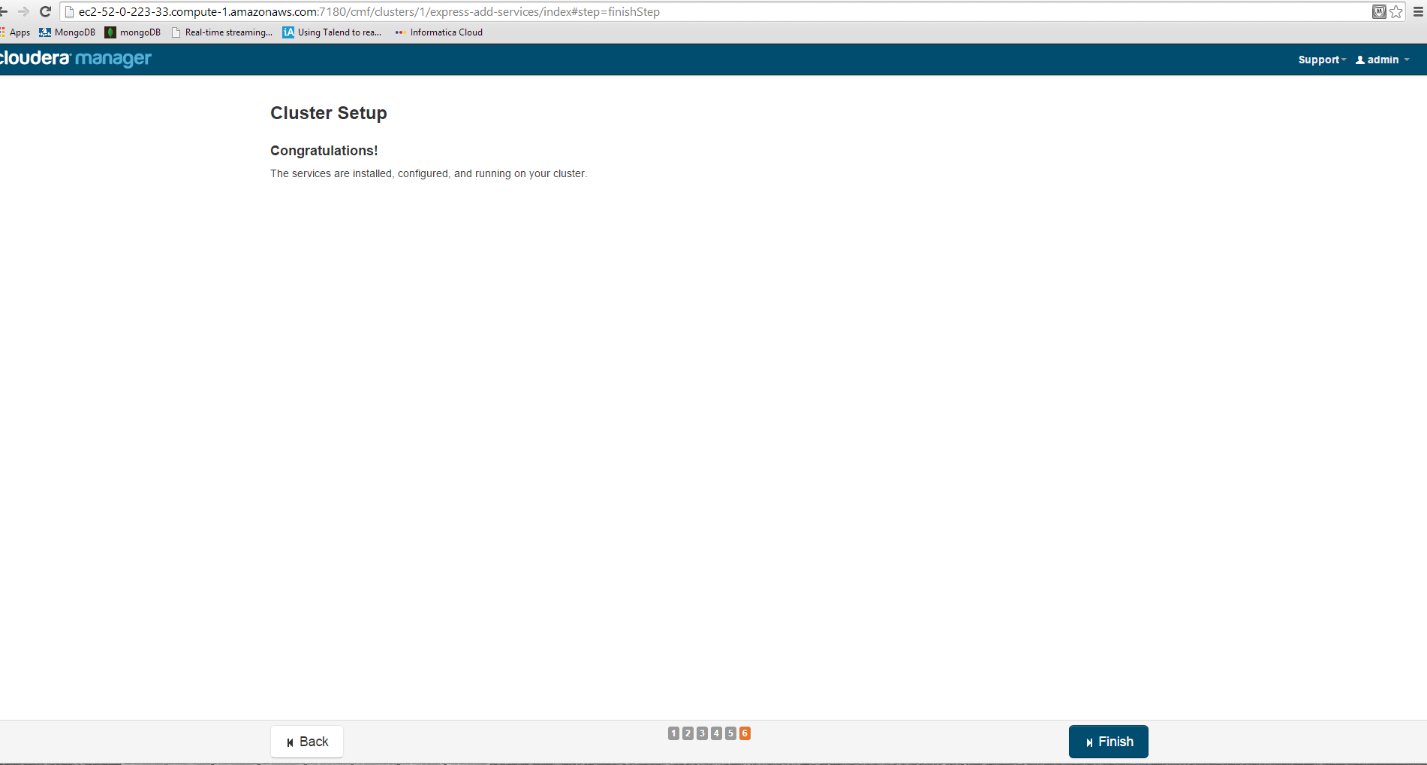
****

**Test connection to metadata it’s correctly working or not**

**Continue until the below screen will appear**

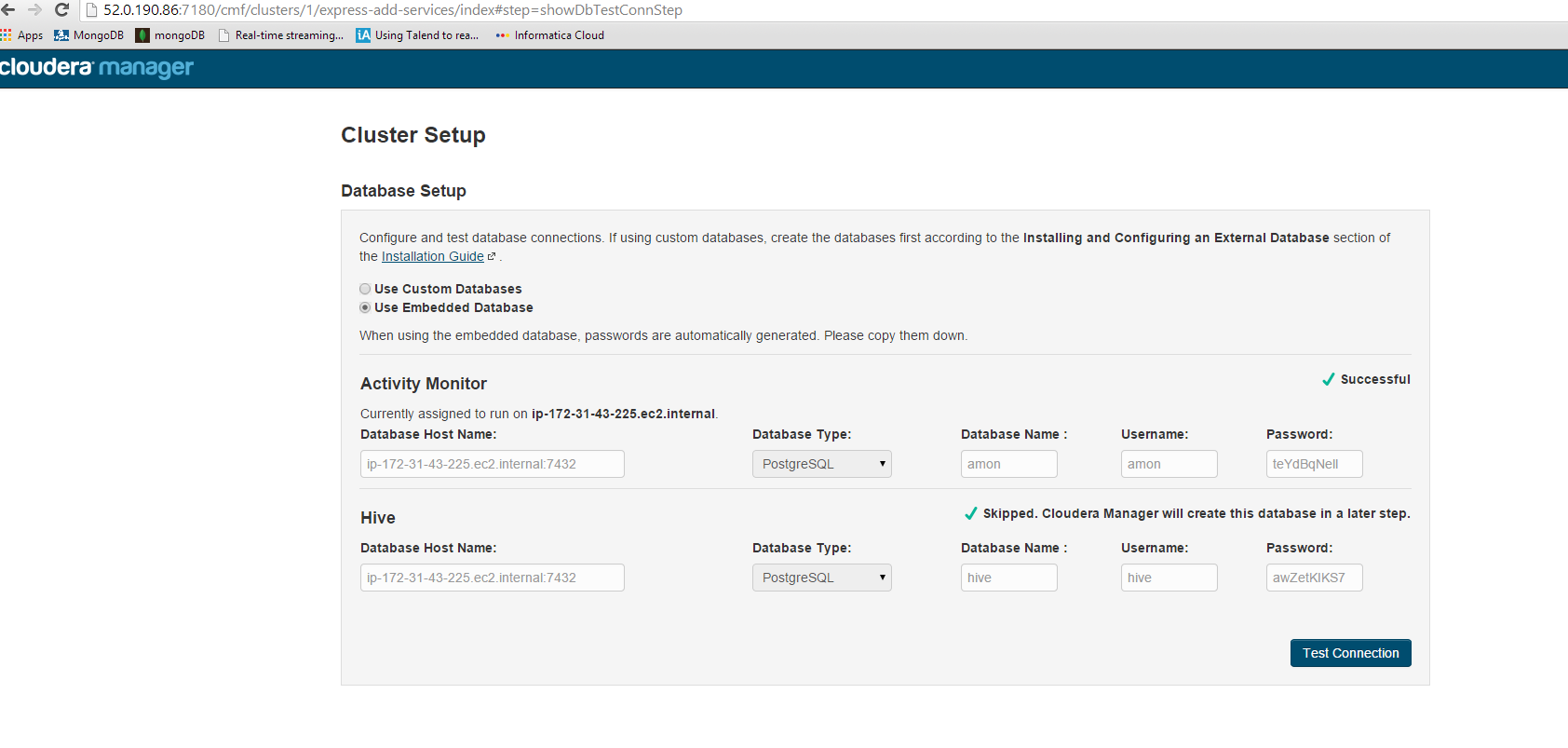
****

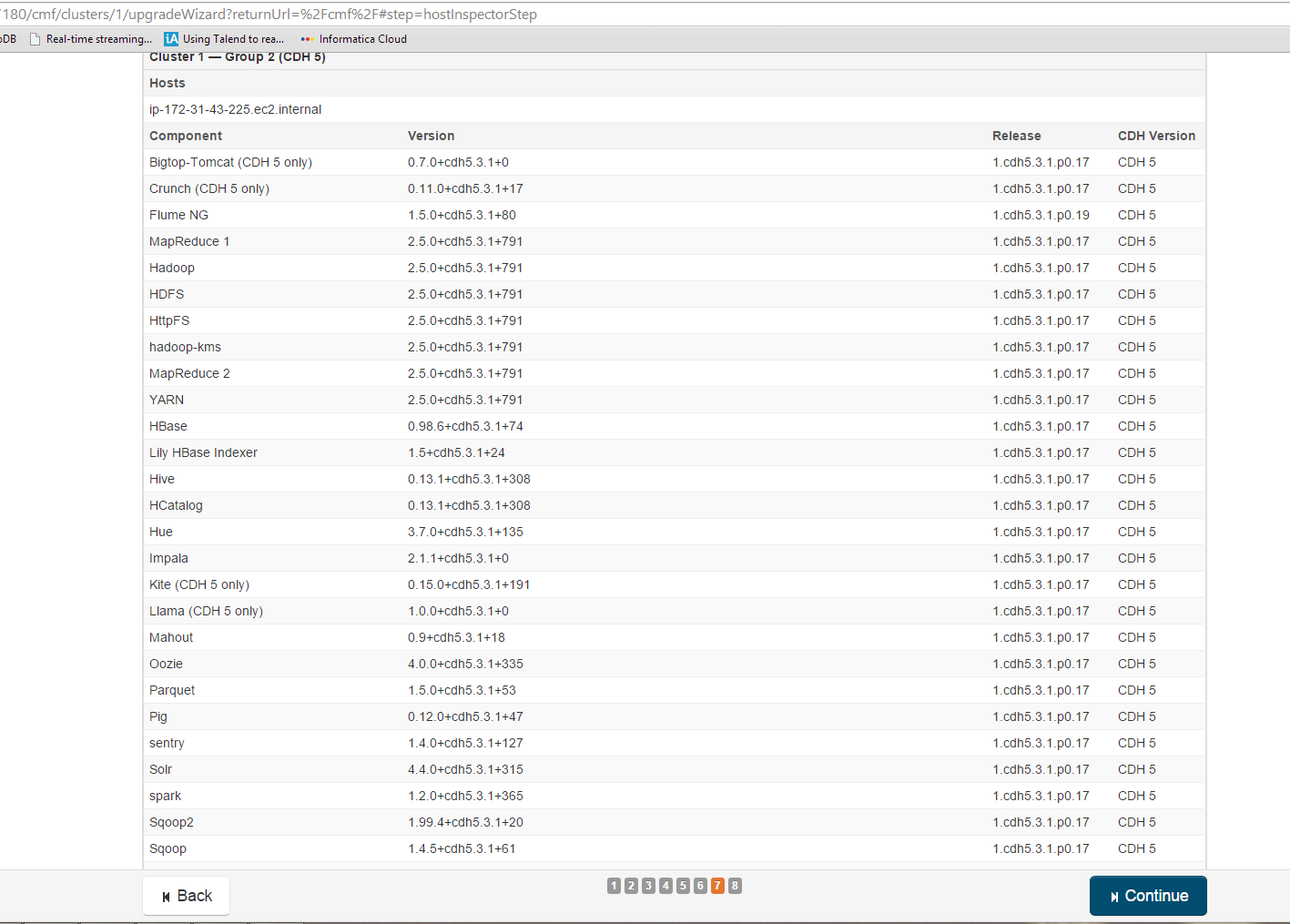
****

****

**Test installation by using below default mapreduce program:**

sudo -u hdfs hadoop jar /usr/lib/hadoop-mapreduce/hadoop-mapreduce-examples.jar pi 10 100

****

****