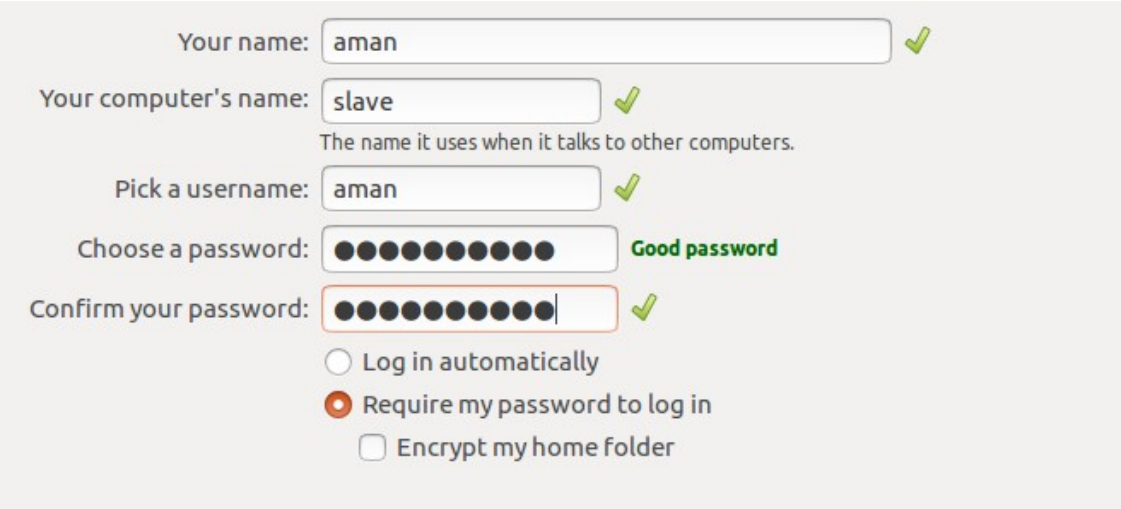
NAME:-AMAN SHAH **BE CSE IBM,BDA-1**
UID:-17BCS3784 **GROUP-A**

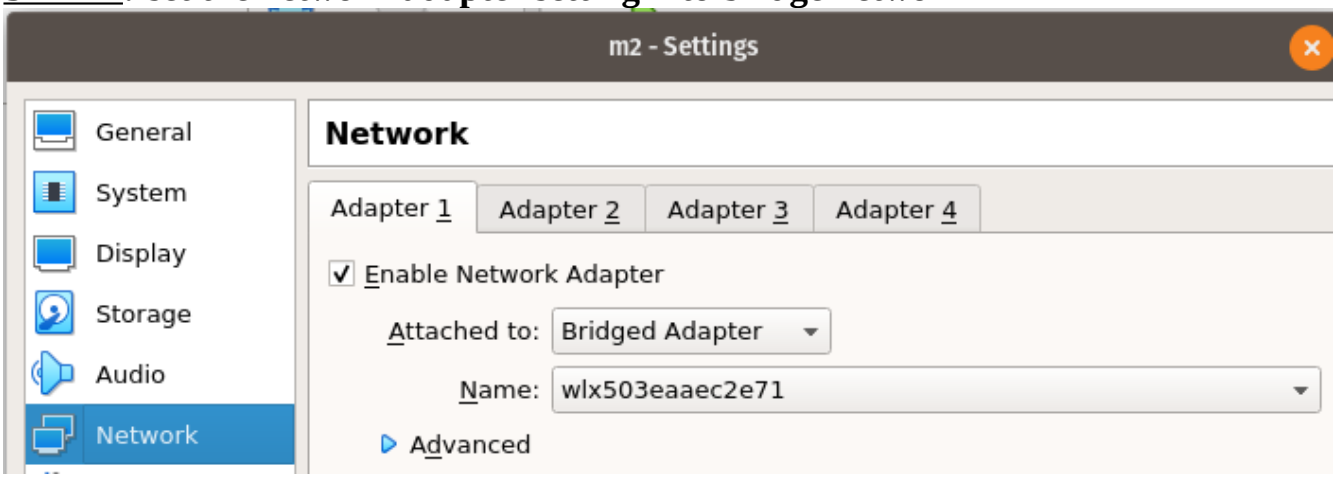
AIM:- To create the multinode cluster in Hadoop 2.x

REQUIREMENT:-i.Virtual Box
ii.2 Ubuntu systems

STEP 1:- Created 2 vms naming aman@master and aman@slave.



STEP 2:-set the network adapter setting into bridge network



STEP 3:-Check the ip address of two machines using ifconfig

command:- ifconfig

```
aman@slave: ~
aman@slave:~$ ifconfig
enp0s3    Link encap:Ethernet  HWaddr 08:00:27:a0:ff:3f
          inet addr:192.168.2.208  Bcast:192.168.2.255  Mask:255.255.255.0
          inet6 addr: fe80::925f:210c:28bf:65f7/64 Scope:Link
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:1307 errors:0 dropped:0 overruns:0 frame:0
          TX packets:607 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1000
          RX bytes:1814242 (1.8 MB)  TX bytes:49373 (49.3 KB)
```

STEP 4:- register one's ip address into another

command:- i.sudo su

ii.vim /*etc*/hosts

```
127.0.0.1       localhost
127.0.1.1       master
192.168.2.208       slave
# The following lines are desirable for IPv6 capable hosts
```

STEP 5:- verify the communication between two machines

command:-ping 192.168.2.206

```
aman@slave:~$ ping 192.168.2.206
PING 192.168.2.206 (192.168.2.206) 56(84) bytes of data.
64 bytes from 192.168.2.206: icmp_seq=1 ttl=63 time=1.92 ms
64 bytes from 192.168.2.206: icmp_seq=2 ttl=63 time=1.57 ms
```

command:-ping 192.168.2.208

```
aman@master:~$ ping 192.168.2.208
PING 192.168.2.208 (192.168.2.208) 56(84) bytes of data.
64 bytes from 192.168.2.208: icmp_seq=1 ttl=64 time=0.338 ms
64 bytes from 192.168.2.208: icmp_seq=2 ttl=64 time=1.11 ms
```

STEP 6:-download openssh-server package

command:-sudo apt-get install openssh-server

```
aman@master:~$ sudo apt-get install openssh-server
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following additional packages will be installed:
  ncurses-term openssh-client openssh-sftp-server ssh-import-id
Suggested packages:
```

STEP 7:-download package called wget
command:-sudo apt-get install wget

```
aman@master:~$ sudo apt-get install wget
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following packages will be upgraded:
```

STEP 8:-disable firewall to open all the ports of our machine
command:- sudo ufw disable

```
aman@slave: ~
aman@slave:~$ sudo ufw disable
[sudo] password for aman:
Firewall stopped and disabled on system startup
aman@slave:~$
```

STEP 9:-download jdk from oracle website i.e. Linux-64 compresses archive
STEP 10:-download hadoop from archive.apache.org(hadoop-2.6.5.tar.gz(190M))
STEP 11:-i.Goto root user and make jvm directory within *usr*/lib/

command:- sudo su
mkdir jvm

```
root@master:/usr/lib# mkdir jvm
```

ii.goto jvm and untar the zip file

```
root@master:/usr/lib# cd jvm/
root@master:/usr/lib/jvm# ls
root@master:/usr/lib/jvm# tar -xvf /home/aman/Downloads/jdk-8u241-linux-x64.tar.
gz
```

```
root@master:/usr/lib/jvm# ls
jdk1.8.0_241
```

STEP 12:-add java to path and update bashrc
command:-i.vim .bashrc
ii.source .bashrc

```
root@master: ~
# ~/.bashrc: executed by bash(1) for non-login shells.
# see /usr/share/doc/bash/examples/startup-files (in the package bash-doc)
# for examples
export JAVA_HOME=/usr/lib/jvm/jdk1.8.0_241
export PATH=$PATH:$JAVA_HOME/bin
# If not running interactively, don't do anything
[ -z "$PS1" ] && return

# don't put duplicate lines in the history. See bash(1) for more options
# ... or force ignoredups and ignorespace
HISTCONTROL=ignoredups:ignorespace

# append to the history file, don't overwrite it
shopt -s histappend

# for setting history length see HISTSIZE and HISTFILESIZE in bash(1)
HISTSIZE=1000
HISTFILESIZE=2000

# check the window size after each command and, if necessary,
# update the values of LINES and COLUMNS.
shopt -s checkwinsize

:wq
```

STEP 13:-similarly untar hadoop file in */usr/*local/

```
root@master:/usr/local# ls -all
total 44
drwxr-xr-x 11 root root 4096 Apr 10 23:53 .
drwxr-xr-x 11 root root 4096 Feb 27  2019 ..
drwxr-xr-x  2 root root 4096 Feb 27  2019 bin
drwxr-xr-x  2 root root 4096 Feb 27  2019 etc
drwxr-xr-x  2 root root 4096 Feb 27  2019 games
drwxrwxr-x  9 aman aman 4096 Oct  3  2016 hadoop-2.6.5
drwxr-xr-x  2 root root 4096 Feb 27  2019 include
drwxr-xr-x  4 root root 4096 Feb 27  2019 lib
lrwxrwxrwx  1 root root    9 Apr 10 12:49 man -> share/man
drwxr-xr-x  2 root root 4096 Feb 27  2019 sbin
drwxr-xr-x  8 root root 4096 Feb 27  2019 share
drwxr-xr-x  2 root root 4096 Feb 27  2019 src
```

STEP 14:-Simplify hadoop-2.6.5 into hadoop

command:-ln -s hadoop2.6.5 hadoop

```
root@master:/usr/local# ln -s hadoop-2.6.5 hadoop
root@master:/usr/local# ls -all
total 44
drwxr-xr-x 11 root root 4096 Apr 10 23:57 .
drwxr-xr-x 11 root root 4096 Feb 27  2019 ..
drwxr-xr-x  2 root root 4096 Feb 27  2019 bin
drwxr-xr-x  2 root root 4096 Feb 27  2019 etc
drwxr-xr-x  2 root root 4096 Feb 27  2019 games
lrwxrwxrwx  1 root root   12 Apr 10 23:57 hadoop -> hadoop-2.6.5
drwxrwxr-x  9 aman aman 4096 Oct  3  2016 hadoop-2.6.5
```

STEP 15:-change the ownership to the hadoop file.

Command:-chown -R aman:aman hadoop*

```
root@master: /usr/local/hadoop
root@master:/usr/local# chown -R aman:aman hadoop*
root@master:/usr/local# cd hadoop
root@master:/usr/local/hadoop# ls
bin  include  libexec      NOTICE.txt  sbin
etc  lib      LICENSE.txt  README.txt  share
root@master:/usr/local/hadoop# pwd
/usr/local/hadoop
root@master:/usr/local/hadoop#
```

STEP 16:-provide path for hadoop and update it

command: i.vim .bashrc

ii.source .bashrc

```
root@master: ~
# ~/.bashrc: executed by bash(1) for non-login shells.
# see /usr/share/doc/bash/examples/startup-files (in the package bash-doc)
# for examples
export JAVA_HOME=/usr/lib/jvm/jdk1.8.0_241
export PATH=$PATH:$JAVA_HOME/bin
export HADOOP=/usr/local/hadoop
export PATH=$PATH:$HADOOP/bin
export PATH=$PATH:$HADOOP/sbin

# If not running interactively, don't do anything
[ -z "$PS1" ] && return

# don't put duplicate lines in the history. See bash(1) for more options
# ... or force ignoredups and ignorespace
HISTCONTROL=ignoredups:ignorespace

# append to the history file, don't overwrite it
shopt -s histappend

# for setting history length see HISTSIZE and HISTFILESIZE in bash(1)
HISTSIZE=1000
HISTFILESIZE=2000

-- INSERT --                                              8,28          Top
```

STEP 17:-provide root privilege to a particular user
command:-visudo

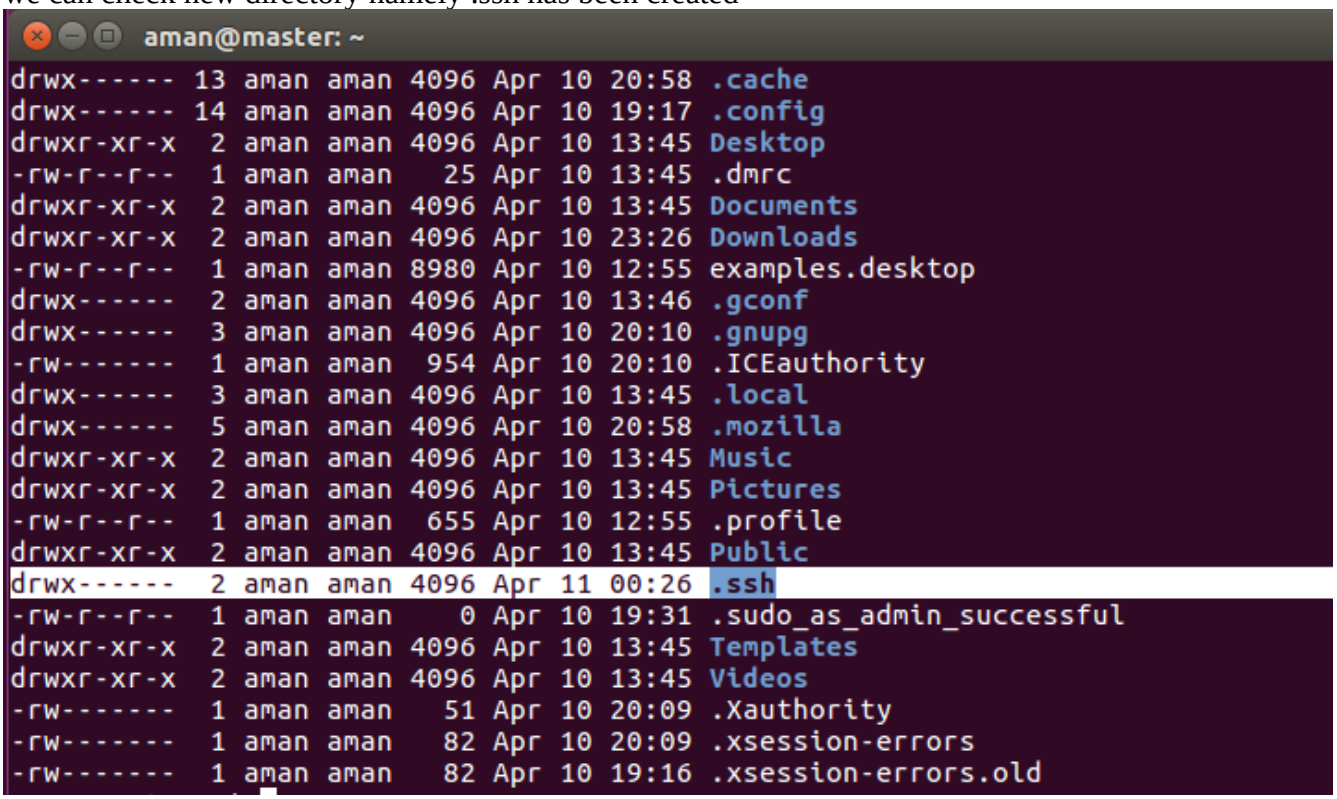
```
root@master:~# visudo
root@master:~# sudo aman
sudo: aman: command not found
root@master:~# su aman
aman@master:/root$
```

```
root@master: ~
  GNU nano 2.5.3            File: /etc/sudoers.tmp                  Modified

Defaults        secure_path="/usr/local/sbin:/usr/local/bin:/usr/sbin:/usr/bin:$

# Host alias specification

# User alias specification

# Cmnd alias specification

# User privilege specification
root    ALL=(ALL:ALL) ALL
aman    ALL=(ALL:ALL) ALL

# Members of the admin group may gain root privileges
%admin ALL=(ALL) ALL

# Allow members of group sudo to execute any command
%sudo   ALL=(ALL:ALL) ALL

# See sudoers(5) for more information on "#include" directives:
File Name to Write: /etc/sudoers.tmp
^G Get Help     M-D DOS Format  M-A Append      M-B Backup File
^C Cancel       M-M Mac Format  M-P Prepend     ^T To Files
```

STEP 18:-generate the keys
command:- ssh-keygen -t rsa -P””

```
aman@master: /root
aman@master:/root$ ssh-keygen -t rsa -P ""
Generating public/private rsa key pair.
Enter file in which to save the key (/home/aman/.ssh/id_rsa):
Created directory '/home/aman/.ssh'.
Your identification has been saved in /home/aman/.ssh/id_rsa.
Your public key has been saved in /home/aman/.ssh/id_rsa.pub.
The key fingerprint is:
SHA256:+kEa/WMHm9veVnlYHnJ9L46E77xWUHda8INCJceu4gg aman@master
The key's randomart image is:
+---[RSA 2048]----+
|           ooo.. |
|          . o.o.+|
|           ..o =+|
|        .    oo.o=|
|       . S ....o+=|
|      E = o.=. +.=|
|       + + Bo.+ o.|
|        o + *+.o  |
|         . .+*o.  |
+----[SHA256]-----+
```

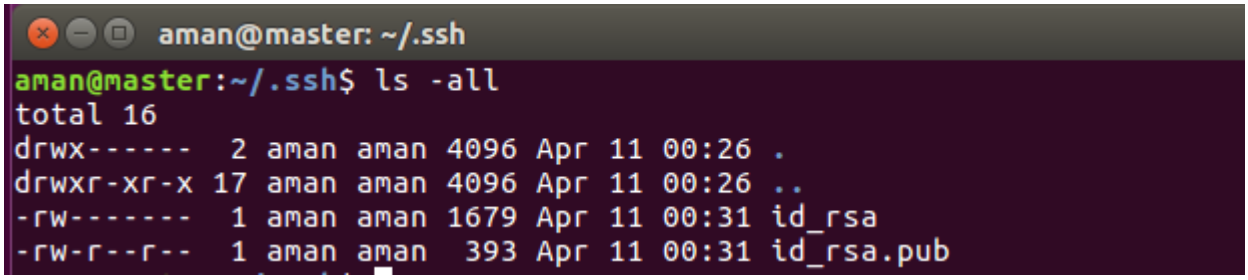
we can check new directory namely .ssh has been created



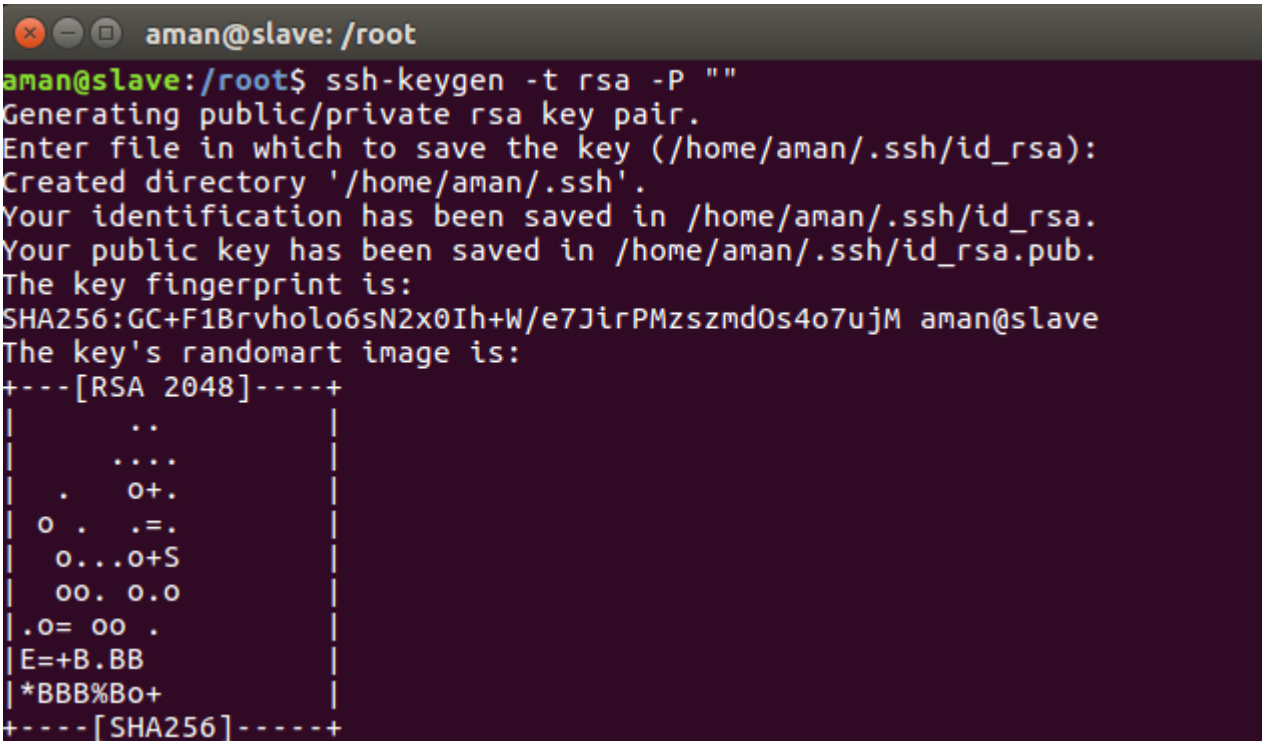
STEP 19:-check the public key of master machine

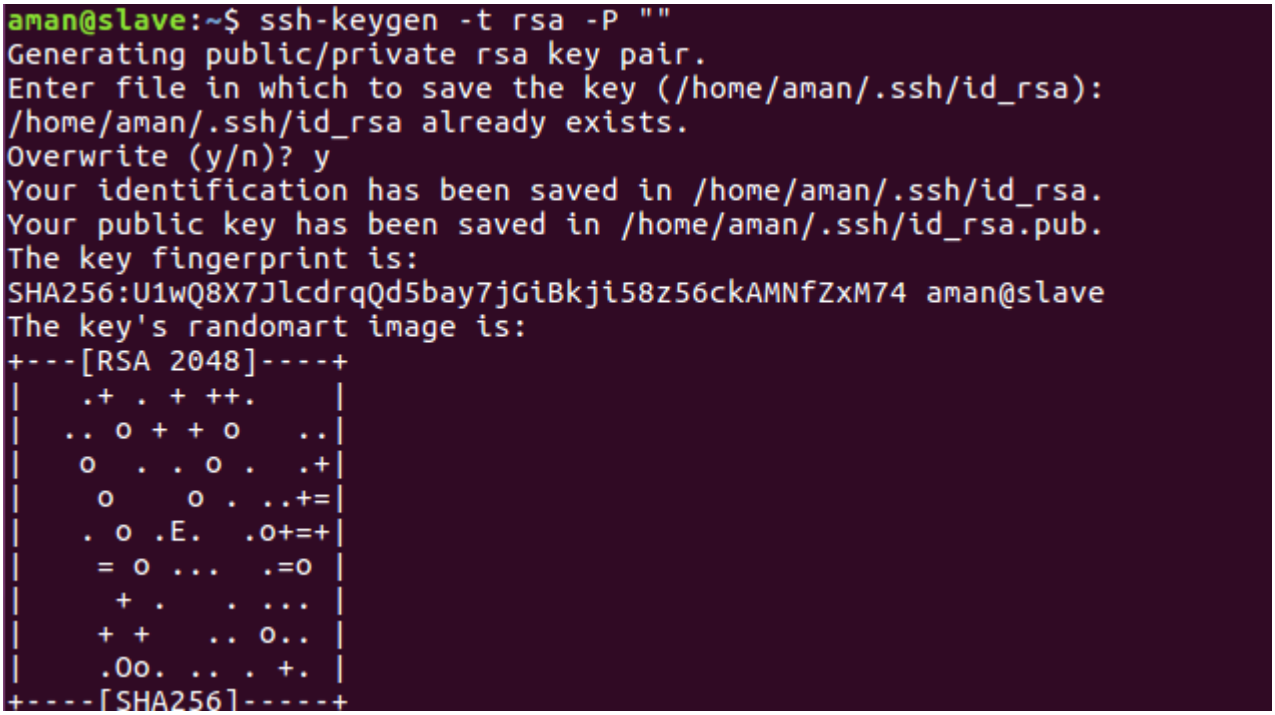
command:-i.cd .ssh

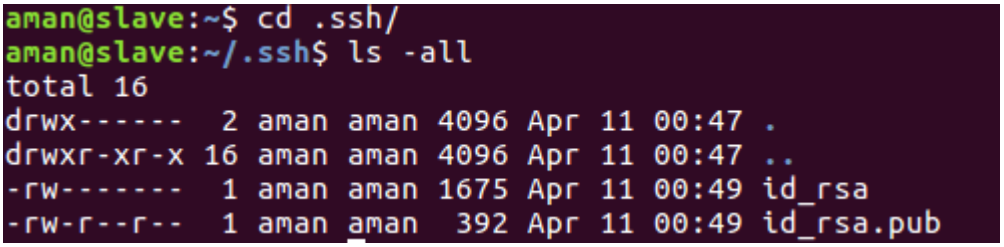
ii.ls -all



STEP 20:-similarly,generate keys in slave machine







STEP 21:-register the public key of a machine to another

command:-ssh-copy-id -i $HOME/.ssh/id_rsa.pub aman@master

```
aman@slave:~/.ssh$ cd
aman@slave:~$ ssh-copy-id -i $HOME/.ssh/id_rsa.pub aman@master
/usr/bin/ssh-copy-id: INFO: Source of key(s) to be installed: "/home/aman/.ssh/i
d_rsa.pub"
The authenticity of host 'master (192.168.2.206)' can't be established.
ECDSA key fingerprint is SHA256:ECOe8QlLya7UJpz+oCGzu664t+Bs4I9kKxOEdIdvKYw.
Are you sure you want to continue connecting (yes/no)? yes
/usr/bin/ssh-copy-id: INFO: attempting to log in with the new key(s), to filter
out any that are already installed
/usr/bin/ssh-copy-id: INFO: 1 key(s) remain to be installed -- if you are prompt
ed now it is to install the new keys
aman@master's password:

Number of key(s) added: 1

Now try logging into the machine, with:   "ssh 'aman@master'"
and check to make sure that only the key(s) you wanted were added.
```

```
aman@slave: ~/.ssh
aman@slave:~$ cd .ssh/
aman@slave:~/.ssh$ ls -all
total 20
drwx------  2 aman aman 4096 Apr 11 00:55 .
drwxr-xr-x 16 aman aman 4096 Apr 11 00:47 ..
-rw-------  1 aman aman 1675 Apr 11 00:49 id_rsa
-rw-r--r--  1 aman aman  392 Apr 11 00:49 id_rsa.pub
-rw-r--r--  1 aman aman  444 Apr 11 00:55 known_hosts
```

STEP 22:-register the public key of a machine to another

command:-ssh-copy-id -i $HOME/.ssh/id_rsa.pub aman@slave

```
aman@master:~/.ssh$ ls -all
total 20
drwx------  2 aman aman 4096 Apr 11 00:55 .
drwxr-xr-x 17 aman aman 4096 Apr 11 00:26 ..
-rw-------  1 aman aman  392 Apr 11 00:55 authorized_keys
-rw-------  1 aman aman 1679 Apr 11 00:31 id_rsa
-rw-r--r--  1 aman aman  393 Apr 11 00:31 id_rsa.pub
```

```
aman@master:~$ ssh-copy-id -i $HOME/.ssh/id_rsa.pub aman@slave
/usr/bin/ssh-copy-id: INFO: Source of key(s) to be installed: "/home/aman/.ssh/i
d_rsa.pub"
The authenticity of host 'slave (192.168.2.208)' can't be established.
ECDSA key fingerprint is SHA256:GX7NjX+iJvb1e9HCujJY3qwD+UJ7CbXxwB6axoy/QAE.
Are you sure you want to continue connecting (yes/no)? yes
/usr/bin/ssh-copy-id: INFO: attempting to log in with the new key(s), to filter
out any that are already installed
/usr/bin/ssh-copy-id: INFO: 1 key(s) remain to be installed -- if you are prompt
ed now it is to install the new keys
aman@slave's password:

Number of key(s) added: 1

Now try logging into the machine, with:   "ssh 'aman@slave'"
and check to make sure that only the key(s) you wanted were added.
```

```
aman@master: ~/.ssh
aman@master:~$ cd .ssh
aman@master:~/.ssh$ ls -all
total 24
drwx------  2 aman aman 4096 Apr 11 01:03 .
drwxr-xr-x 17 aman aman 4096 Apr 11 00:26 ..
-rw-------  1 aman aman  392 Apr 11 00:55 authorized_keys
-rw-------  1 aman aman 1679 Apr 11 00:31 id_rsa
-rw-r--r--  1 aman aman  393 Apr 11 00:31 id_rsa.pub
-rw-r--r--  1 aman aman  444 Apr 11 01:03 known_hosts
```

```
aman@slave:~/.ssh$ ls -all
total 24
drwx------  2 aman aman 4096 Apr 11 01:03 .
drwxr-xr-x 16 aman aman 4096 Apr 11 00:47 ..
-rw-------  1 aman aman  393 Apr 11 01:03 authorized_keys
-rw-------  1 aman aman 1675 Apr 11 00:49 id_rsa
-rw-r--r--  1 aman aman  392 Apr 11 00:49 id_rsa.pub
-rw-r--r--  1 aman aman  444 Apr 11 00:55 known_hosts
```

STEP 23:-append public key to authorised key in another machine

command:- cat */home/*aman/.ssh/id_rsa.pub >> */home*aman/.ssh/authorized_keys

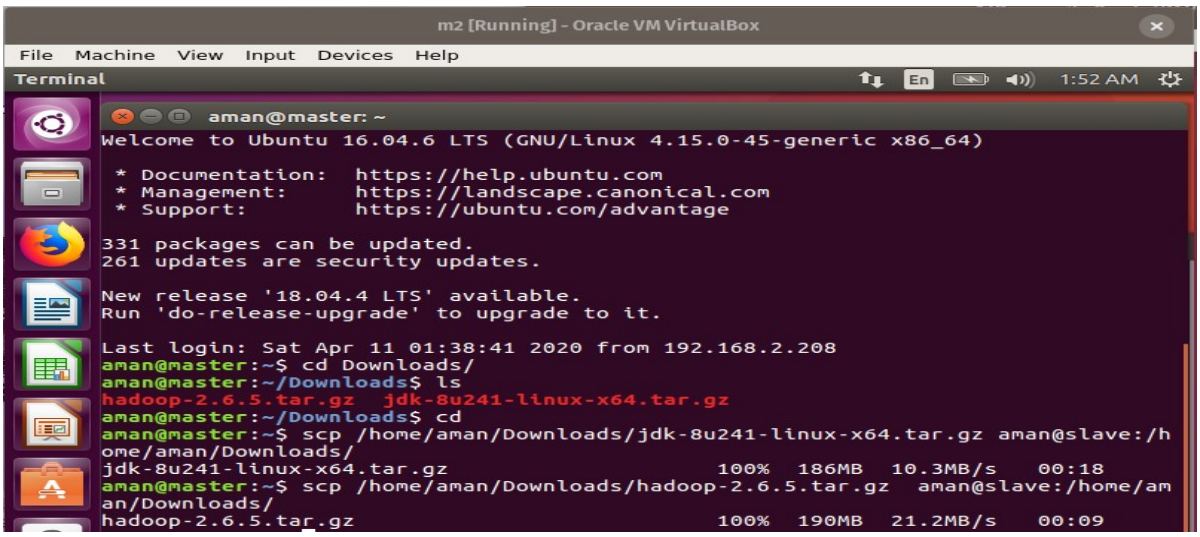
```
aman@master:~/.ssh$ cat /home/aman/.ssh/id_rsa.pub >> /home/aman/.ssh/authorized
_keys
```

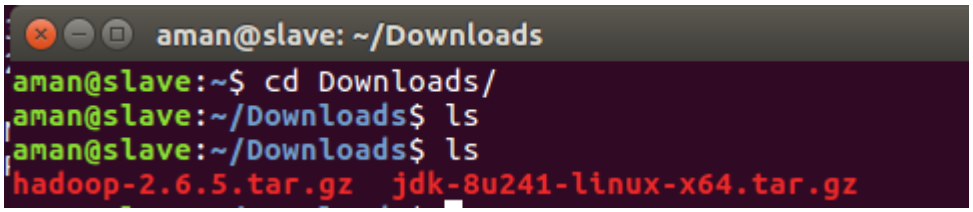
STEP 24:-verify ssh

command:-ssh slave

```
aman@master:~$ ssh slave
Welcome to Ubuntu 16.04.6 LTS (GNU/Linux 4.15.0-45-generic x86_64)

 * Documentation:  https://help.ubuntu.com
 * Management:     https://landscape.canonical.com
 * Support:        https://ubuntu.com/advantage

331 packages can be updated.
261 updates are security updates.

New release '18.04.4 LTS' available.
Run 'do-release-upgrade' to upgrade to it.
```

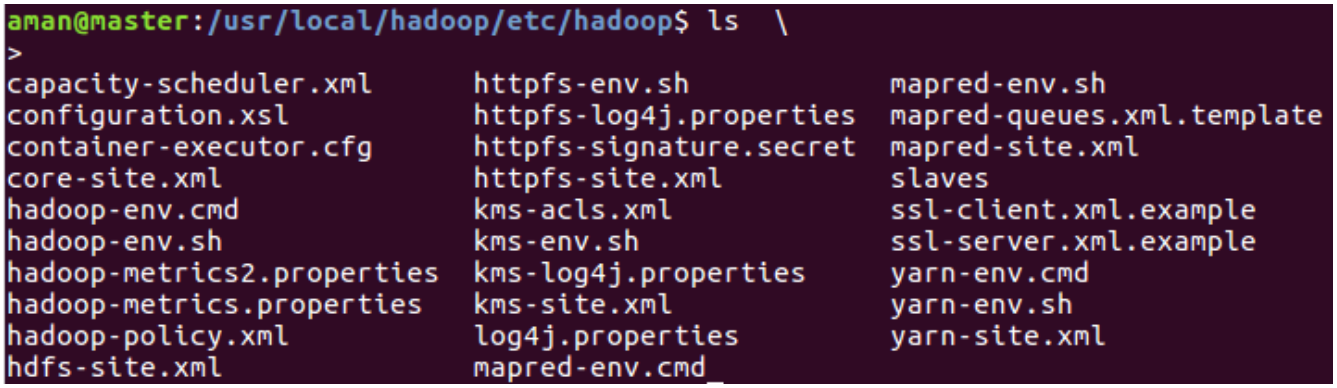
STEP 25:-transfer jdk and hadoop from master to slave.
Command:- scp */home/*aman/Downloads/hadoop-2.6.5.tar.gz aman@slave:/home/aman/Downloads/





STEP 26:-Repeat the same steps as above to install java and hadoop in slave machine

STEP 27:-configuring the hadoop files properties



STEP 28:-edit core-site.xml

command:-vim core-site.xml

```
aman@master: /usr/local/hadoop/etc/hadoop
<?xml version="1.0" encoding="UTF-8"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<!--
  Licensed under the Apache License, Version 2.0 (the "License");
  you may not use this file except in compliance with the License.
  You may obtain a copy of the License at

    http://www.apache.org/licenses/LICENSE-2.0

  Unless required by applicable law or agreed to in writing, software
  distributed under the License is distributed on an "AS IS" BASIS,
  WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
  See the License for the specific language governing permissions and
  limitations under the License. See accompanying LICENSE file.
-->

<!-- Put site-specific property overrides in this file. -->

<configuration>
 <property>
    <name>fs.defaultFS</name>
    <value>hdfs://master:9000</value>
    </property>
```

STEP 29:-rename mapred-site.xml.template to mapred-site.xml

command:-mv mapred-site.xml.template mapred-site.xml

```
aman@master:/usr/local/hadoop/etc/hadoop$ mv mapred-site.xml.template mapred-sit
e.xml
```

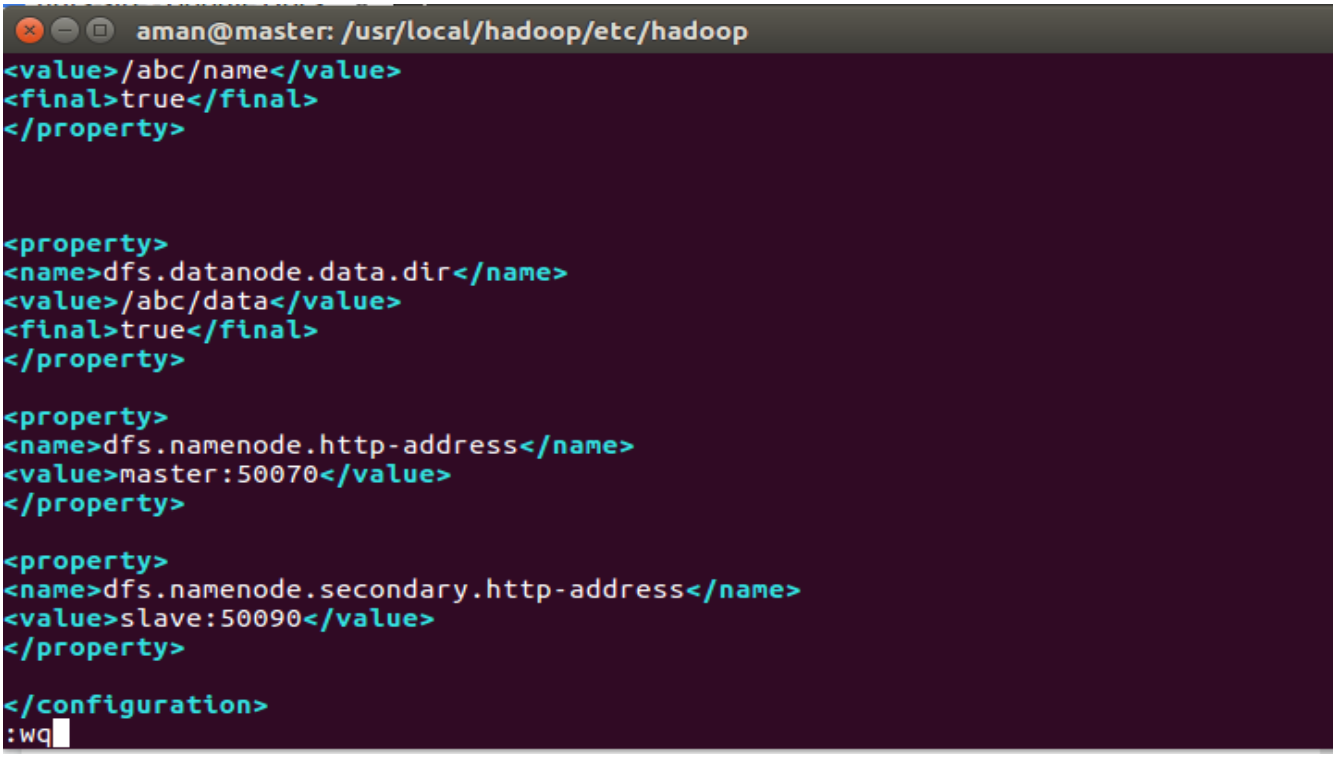
STEP 30:-edit mapred file and mention framework i.e. yarn

command:-vim mapred-site.xml

```
aman@master: /usr/local/hadoop/etc/hadoop
<?xml version="1.0"?>
<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>
<!--
  Licensed under the Apache License, Version 2.0 (the "License");
  you may not use this file except in compliance with the License.
  You may obtain a copy of the License at

    http://www.apache.org/licenses/LICENSE-2.0

  Unless required by applicable law or agreed to in writing, software
  distributed under the License is distributed on an "AS IS" BASIS,
  WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied.
  See the License for the specific language governing permissions and
  limitations under the License. See accompanying LICENSE file.
-->

<!-- Put site-specific property overrides in this file. -->

<configuration>
<property>
<name>mapreduce.framework.name</name>
<value>yarn</value>
</property>
"mapred-site.xml" 25L, 839C                                   1,1           Top
```

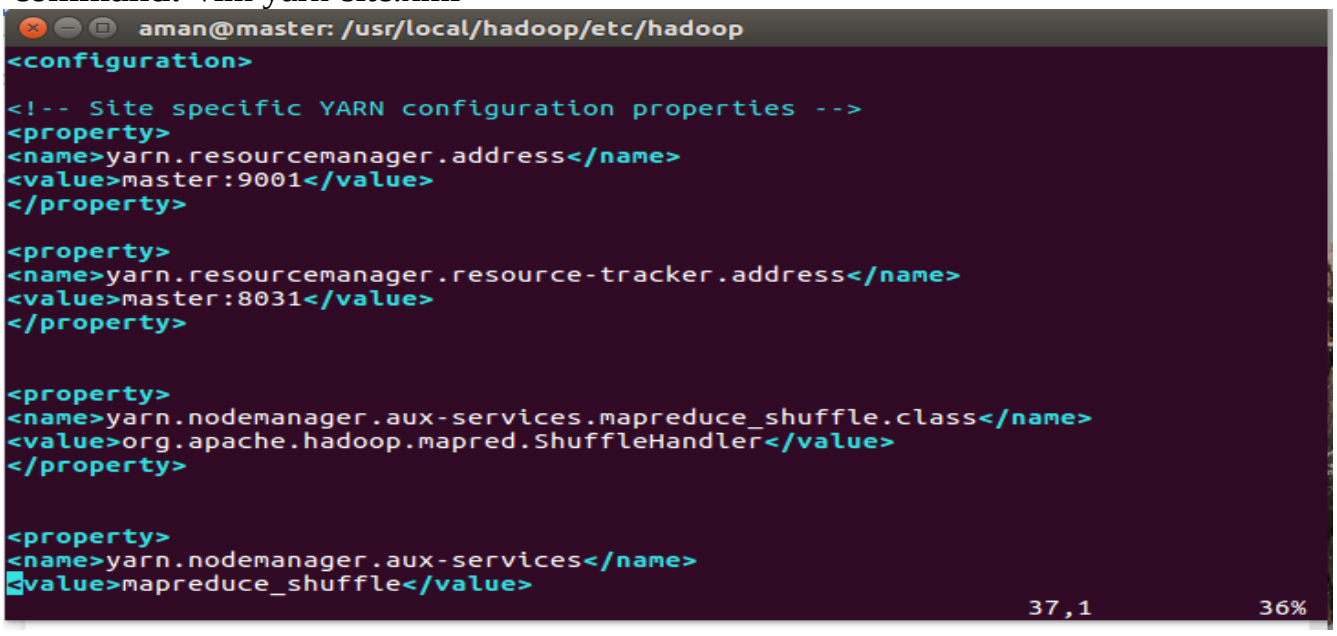
STEP 30:-edit hdfs file

command:-vim hdfs-site.xml



STEP 30:-edit yarn file

command:-vim yarn-site.xml



STEP 31:-configure slave file and mention both macines i.e. master and slave

command:-vim slaves

```
aman@master:/usr/local/hadoop/etc/hadoop$ ls  \
>
capacity-scheduler.xml      httpfs-env.sh            mapred-env.sh
configuration.xsl           httpfs-log4j.properties  mapred-queues.xml.template
container-executor.cfg      httpfs-signature.secret  mapred-site.xml
core-site.xml               httpfs-site.xml          slaves
hadoop-env.cmd              kms-acls.xml             ssl-client.xml.example
hadoop-env.sh               kms-env.sh               ssl-server.xml.example
hadoop-metrics2.properties  kms-log4j.properties     yarn-env.cmd
hadoop-metrics.properties   kms-site.xml             yarn-env.sh
hadoop-policy.xml           log4j.properties         yarn-site.xml
hdfs-site.xml               mapred-env.cmd
```

```
aman@master: /usr/local/hadoop/etc/hadoop
master
slave
~
~
~
~
~
~
~
~
~
~
~
~
~
~
~
~
~
~
:wq
```

STEP 31:-copy all the hadoop files of machine namely master to another machine namely slave

command:-scp */usr*/local/hadoop/* aman@slave:/usr/local/hadoop/etc/hadoop/

```
aman@master:~$ scp /usr/local/hadoop/etc/hadoop/* aman@slave:/usr/local/hadoop/e
tc/hadoop/
capacity-scheduler.xml                        100% 4436     4.3KB/s   00:00
configuration.xsl                             100% 1335     1.3KB/s   00:00
container-executor.cfg                        100%  318     0.3KB/s   00:00
core-site.xml                                 100%  871     0.9KB/s   00:00
hadoop-env.cmd                                100% 3670     3.6KB/s   00:00
hadoop-env.sh                                 100% 4224     4.1KB/s   00:00
hadoop-metrics2.properties                    100% 2598     2.5KB/s   00:00
hadoop-metrics.properties                     100% 2490     2.4KB/s   00:00
hadoop-policy.xml                             100% 9683     9.5KB/s   00:00
hdfs-site.xml                                 100% 1247     1.2KB/s   00:00
httpfs-env.sh                                 100% 1449     1.4KB/s   00:00
httpfs-log4j.properties                       100% 1657     1.6KB/s   00:00
httpfs-signature.secret                       100%   21     0.0KB/s   00:00
httpfs-site.xml                               100%  620     0.6KB/s   00:00
kms-acls.xml                                  100% 3523     3.4KB/s   00:00
kms-env.sh                                    100% 1325     1.3KB/s   00:00
kms-log4j.properties                          100% 1631     1.6KB/s   00:00
kms-site.xml                                  100% 5511     5.4KB/s   00:00
log4j.properties                              100%   11KB  11.0KB/s   00:00
mapred-env.cmd                                100%  938     0.9KB/s   00:00
```

STEP 32:-edit the hdfs file in slave machine and remove the namenode property and remove the mapred-site-xm.template

command:-i.vim hdfs-site.xml

ii.rm mapred-site-xm.template

STEP 33:-formatting with the metadata to start hadoop

command:-hdfs namenode -formatting

```
aman@master:~$ hdfs namenode -formatting
20/04/11 17:55:35 INFO namenode.NameNode: STARTUP_MSG:
/************************************************************
STARTUP_MSG: Starting NameNode
STARTUP_MSG:   host = master/127.0.1.1
STARTUP_MSG:   args = [-formatting]
STARTUP_MSG:   version = 2.6.5
STARTUP_MSG:   classpath = /usr/local/hadoop-2.6.5/etc/hadoop:/usr/local/hadoop
```

```
aman@master: ~
STARTUP_MSG:   java = 1.8.0_241
************************************************************/
20/04/11 18:01:47 INFO namenode.NameNode: registered UNIX signal handlers for [T
ERM, HUP, INT]
20/04/11 18:01:47 INFO namenode.NameNode: createNameNode [-formatting]
Usage: java NameNode [-backup] |
        [-checkpoint] |
        [-format [-clusterid cid ] [-force] [-nonInteractive] ] |
        [-upgrade [-clusterid cid] [-renameReserved<k-v pairs>] ] |
        [-upgradeOnly [-clusterid cid] [-renameReserved<k-v pairs>] ] |
        [-rollback] |
        [-rollingUpgrade <rollback|downgrade|started> ] |
        [-finalize] |
        [-importCheckpoint] |
        [-initializeSharedEdits] |
        [-bootstrapStandby] |
        [-recover [ -force] ] |
        [-metadataVersion ]  ]

20/04/11 18:01:48 INFO namenode.NameNode: SHUTDOWN_MSG:
/************************************************************
SHUTDOWN_MSG: Shutting down NameNode at master/127.0.1.1
************************************************************/
```

STEP 34:-start all the daemons of the hadoop

command:-start-all.sh

```
aman@master: ~
aman@master:~$ start-all.sh
This script is Deprecated. Instead use start-dfs.sh and start-yarn.sh
Starting namenodes on [master]
The authenticity of host 'master (127.0.1.1)' can't be established.
ECDSA key fingerprint is SHA256:ECOe8QlLya7UJpz+oCGzu664t+Bs4I9kKxOEdIdvKYw.
Are you sure you want to continue connecting (yes/no)? yes
master: Warning: Permanently added 'master' (ECDSA) to the list of known hosts.
master: starting namenode, logging to /usr/local/hadoop-2.6.5/logs/hadoop-aman-n
amenode-master.out
master: starting datanode, logging to /usr/local/hadoop-2.6.5/logs/hadoop-aman-d
atanode-master.out
slave: starting datanode, logging to /usr/local/hadoop-2.6.5/logs/hadoop-aman-da
tanode-slave.out
Starting secondary namenodes [slave]
slave: starting secondarynamenode, logging to /usr/local/hadoop-2.6.5/logs/hadoo
p-aman-secondarynamenode-slave.out
starting yarn daemons
starting resourcemanager, logging to /usr/local/hadoop-2.6.5/logs/yarn-aman-reso
urcemanager-master.out
master: starting nodemanager, logging to /usr/local/hadoop-2.6.5/logs/yarn-aman-
nodemanager-master.out
slave: starting nodemanager, logging to /usr/local/hadoop-2.6.5/logs/yarn-aman-n
odemanager-slave.out
```

STEP 35:-check the hadoop daemons

command:-jps

```
aman@master:~$ jps
5424 ResourceManager
5845 Jps
5541 NodeManager
5210 DataNode
```

STEP 36:-open browser and type master:8088/cluster to view the cluster