

1) Single node cluster

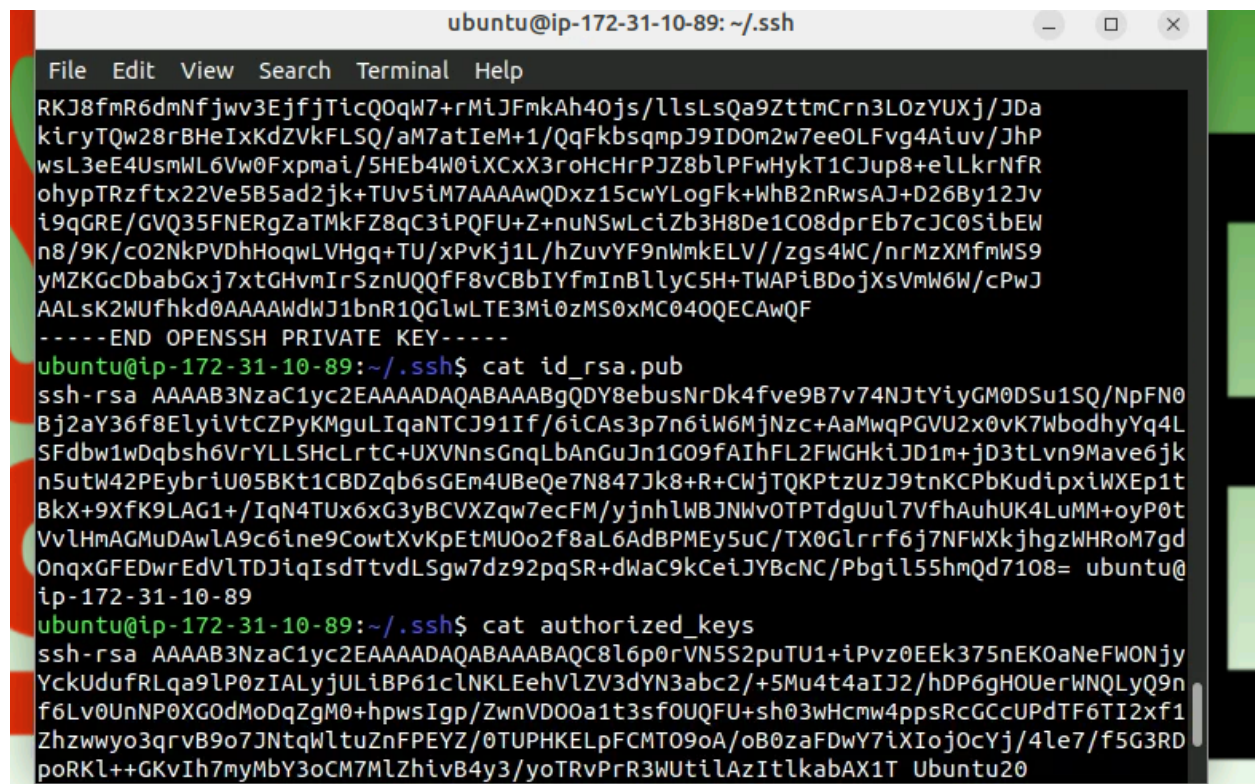
Sudo apt update

```
netto python
ubuntu@ip-172-31-10-89:~$ sudo apk update
sudo: apk: command not found
ubuntu@ip-172-31-10-89:~$ sudo apt update
Hit:1 http://ap-south-1.ec2.archive.ubuntu.com/ubuntu jammy InRelease
Get:2 http://ap-south-1.ec2.archive.ubuntu.com/ubuntu jammy-updates InRelease [1
19 kB]
```

ssh-keygen

cat id_rsa.pub >> authorizedkeys

ssh localhost



The screenshot shows a terminal window titled 'ubuntu@ip-172-31-10-89: ~/.ssh'. The terminal output shows the execution of 'ssh-keygen' which generates an RSA key pair. The public key is then appended to the 'authorized_keys' file. The terminal text is as follows:

```
File Edit View Search Terminal Help
RKJ8fmR6dmNfjwv3EjffjTicQ0qW7+rMiJFmkAh40js/llsLsQa9ZttmCrn3L0zYUXj/JDa
kirYtQw28rBHeIxKdZVkfLSQ/aM7atIeM+1/QqFkbsqmpJ9ID0m2w7eeOLFvg4Aiuv/JhP
wsL3eE4UsmWL6Vw0Fxpmai/5HEb4W0iXCxX3roHcHrPJZ8bLPFwHykT1CJup8+eLLkrNfR
ohypTRzftx22Ve5B5ad2jk+TUV5iM7AAAAwQDxz15cwYLogFk+WhB2nRwsAJ+D26By12Jv
i9qGRE/GVQ35FNERgZaTMkFZ8qC3iPQFU+Z+nuNSwLciZb3H8De1C08dprEb7cJC0SibEW
n8/9K/c02NkPVDhHoqWLVHgg+TU/xPvKj1L/hZuvYF9nWmkELV//zgs4WC/nrMzXMfmWS9
yMZKGcDbabGxj7xtGHvmIrSznUQQfF8vCBbIYfmInBllyC5H+TWAPiBDojXsVmW6W/cPwJ
AALsK2WufhkD0AAAAwDwJ1bnR1QGlWLT3Mi0zMS0xMC040QECAwQF
-----END OPENSSH PRIVATE KEY-----
ubuntu@ip-172-31-10-89:~/.ssh$ cat id_rsa.pub
ssh-rsa AAAAB3NzaC1yc2EAAAADAQABAAQGDY8ebusNrDk4fve9B7v74NJtYiyGM0DSu1S0/NpFN0
Bj2aY36f8ELyivtCZPyKMguLIqaNTCJ91If/6iCAs3p7n6iW6mJNzc+AaMwqPGVU2x0vK7WbodhyYq4L
SFdbw1wDqbsh6VrYLLSHcLrtC+UXVNnsGnqLbAnGuJn1G09fAIhFL2FWGHkiJD1m+jD3tLvN9Mave6jk
n5utW42PEybriu05BKt1CB0Zqb6sGEm4UBeQe7N847Jk8+R+CWjTQKPtZUzJ9tnKCPbKudipxiWxEp1t
BkX+9XfK9LAG1+/IqN4TUx6xG3yBCVXZqw7ecFM/yjnhlWBjNwvOTPTdgUul7VfhAuhUK4LuMM+oyP0t
VvlHmAGMuDAwLA9c6ine9COWtXvKpEtMU0o2f8aL6AdBPMEy5uC/TX0GLrrf6j7NFWXkjhgZWHRoM7gd
OnqxGFEDwrEdVlTDJiQIsdTvdLSgw7dz92pqSR+dWac9kCeIJYBcNC/Pbgil55hmQd7108= ubuntu@
ip-172-31-10-89
ubuntu@ip-172-31-10-89:~/.ssh$ cat authorized_keys
ssh-rsa AAAAB3NzaC1yc2EAAAADAQABAAQGDY8ebusNrDk4fve9B7v74NJtYiyGM0DSu1S0/NpFN0
YckUdufRLqa9lP0zIALyJULiBP61clNkLEehVLZV3dYN3abc2/+5Mu4t4aIJ2/hDP6gH0UerWNQLyQ9n
f6Lv0UnNP0XG0dMoDqZgM0+hpwsIgp/ZwnVD00a1t3sf0UQFU+sh03wHcmw4ppsRcGCCUPdTF6TI2xf1
ZhzwWyo3qrVb9o7JntqWltZnFPEYZ/0TUPHKElPFCMT09oA/oB0zaFDwY7iXIojoCyj/4le7/f5G3RD
poRKL++GKVIh7myMbY3oCM7MLzhivB4y3/yoTrvPrR3WUtiAzItIkabAX1T Ubuntu20
```

Checking localhost

```
ubuntu@ip-172-31-10-89: ~  
File Edit View Search Terminal Help  
nSutW42PEybrIU05Bkt1CBDZqb6sGEm4UBeQe7N847Jk8+R+CWjTQKPtZUzJ9tnKCPbKudipxiWxEp1t  
BkX+9XfK9LAG1+/IqN4TUX6xG3yBCVXZqw7ecFM/yjnhlWBjNWv0TPTdgUul7VfhAuhUK4LuMM+oyP0t  
VvLHmAGMuDAwLA9c6ine9COWtXvKpEtMU0o2f8aL6AdBPMEy5uC/TX0GLrrf6j7NFWXkjhgZWHRoM7gd  
OnqxGFEDwrEdVlTDJiQIsdTtvdLSgw7dz92pqSR+dWac9kCeIJYBcNC/Pbgil55hmQd7108= ubuntu@  
ip-172-31-10-89  
ubuntu@ip-172-31-10-89:~/.ssh$ ssh localhost  
Welcome to Ubuntu 22.04.1 LTS (GNU/Linux 5.15.0-1028-aws x86_64)  
  
* Documentation:  https://help.ubuntu.com  
* Management:    https://landscape.canonical.com  
* Support:       https://ubuntu.com/advantage  
  
System information as of Mon Sep 25 10:37:58 UTC 2023  
  
System load:  0.0390625      Processes:            263  
Usage of /:   65.1% of 38.58GB Users logged in:       0  
Memory usage: 43%           IPv4 address for docker0: 172.17.0.1  
Swap usage:   0%            IPv4 address for ens5:   172.31.10.89  
  
=> There is 1 zombie process.  
  
* Ubuntu Pro delivers the most comprehensive open source security and
```

Downloading hadoop

```
ubuntu@ip-172-31-10-89:~$ wget https://archive.apache.org/dist/hadoop/core/hadoop-2.5.0/hadoop-2.5.0.tar.gz
```

Extracting tar files

```
ubuntu@ip-172-31-10-89:~$ tar -xvzf hadoop-2.5.0.tar.gz hadoop-2.5.0/
```

Confirm the files

```
ubuntu@ip-172-31-10-89:~$ ls  
Desktop          Templates        'python git'  
Documents        Videos          'python git 1'  
Downloads        aws              'python git 2'  
HR_Hiring_details_transaction.csv eclipse          sample.txt  
Music            eclipse-workspace snap  
Pictures         hadoop-2.5.0.tar.gz thinclient_drives  
Public           kafka_2.13-3.5.1.tgz  
PycharmProjects pyt
```

Moving downloaded file to hadoop

```
ubuntu@ip-172-31-10-89:~$ sudo mv hadoop-2.5.0 /usr/local/hadoop
```

Checking the path

```
ubuntu@ip-172-31-10-89:~$ cd /usr/local/hadoop/
```

Configure hadoop.env.sh

```
# The directory where pid files are stored. /tmp by default.
# NOTE: this should be set to a directory that can only be written to by
#       the user that will run the hadoop daemons. Otherwise there is the
#       potential for a symlink attack.
export HADOOP_PID_DIR=${HADOOP_PID_DIR}
export HADOOP_SECURE_DN_PID_DIR=${HADOOP_PID_DIR}

# A string representing this instance of hadoop. $USER by default.
export HADOOP_IDENT_STRING=$USER

export JAVA_HOME=/usr/lib/jvm/java-11-openjdk-amd64
export HADOOP_OPTS=-Djava.net.preferIPv4Stack=true
```

Installing Checking Java version:

```
ubuntu@ip-172-31-10-89: /usr/lib/jvm/java-11-openjdk-amd64
File Edit View Search Terminal Help

ubuntu@ip-172-31-10-89:~$ ls
Desktop          PycharmProjects  pyt
Documents        Templates         'python git'
Downloads        Videos           'python git 1'
HR_Hiring_details_transaction.csv  aws              snap
Music            eclipse           thinclient_drives
Pictures         eclipse-workspace
Public           hadoop-1.2.1.tar.gz

ubuntu@ip-172-31-10-89:~$ whereia jdk
Command 'whereia' not found, did you mean:
  command 'whereis' from deb util-linux (2.37.2-4ubuntu3)
Try: sudo apt install <deb name>
ubuntu@ip-172-31-10-89:~$ whereis jdk
jdk:
ubuntu@ip-172-31-10-89:~$ whereis java
java: /usr/bin/java /usr/share/java /usr/share/man/man1/java.1.gz
ubuntu@ip-172-31-10-89:~$ cd /usr/lib/jvm/
ubuntu@ip-172-31-10-89:/usr/lib/jvm$ ls
default-java  java-1.11.0-openjdk-amd64  java-11-openjdk-amd64  openjdk-11
ubuntu@ip-172-31-10-89:/usr/lib/jvm$ cd java-11-openjdk-amd64/
ubuntu@ip-172-31-10-89:/usr/lib/jvm/java-11-openjdk-amd64$ ls
bin  conf  docs  include  jmods  legal  lib  man  release
ubuntu@ip-172-31-10-89:/usr/lib/jvm/java-11-openjdk-amd64$
```

Configure all files:

core-site.xml

<property>

<name>fs.default.name</name>

<value>hdfs://localhost:9000</value>

</property>

<property>

<name>hadoop.tmp.dir</name>

<value>/usr/local/hadoop/tmp</value>

</property>

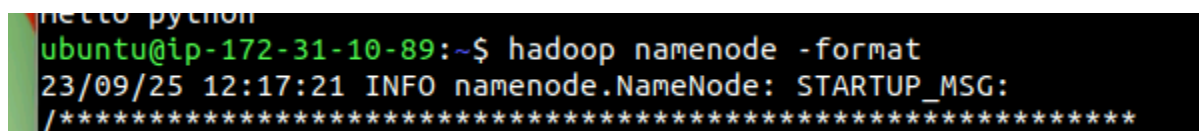
hdfs-site.xml

```
<property>
<name>dfs.replication</name>
<value>1</value>
</property>
```

mapred.xml

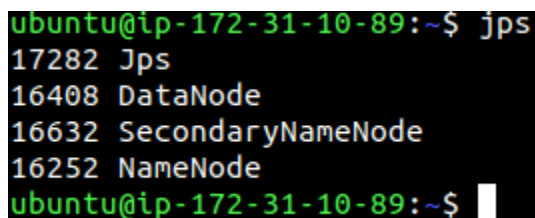
```
<property>
<name>mapred.job.tracker</name>
<value>hdfs://localhost:9001</value>
</property>
```

Starting namenode



```
ubuntu@ip-172-31-10-89:~$ hadoop namenode -format
23/09/25 12:17:21 INFO namenode.NameNode: STARTUP_MSG:
/*****
```

Checking demons after (start-all.sh)



```
ubuntu@ip-172-31-10-89:~$ jps
17282 Jps
16408 DataNode
16632 SecondaryNameNode
16252 NameNode
ubuntu@ip-172-31-10-89:~$
```

Checking localhost(127.0.0.01:50070)

NameNode 'localhost:9000' (active)

Started: Fri Oct 06 04:58:02 UTC 2023
Version: 2.5.0, 1616291
Compiled: 2014-08-06T17:31Z by jenkins from branch-2.5.0
Cluster ID: CID-c97b2ee0-cab4-4224-a725-d6ba0d823e88
Block Pool ID: BP-669782240-172.31.10.89-1696511311332

[Browse the filesystem](#)
[NameNode Logs](#)

Cluster Summary

Security is *OFF*

1 files and directories, 0 blocks = 1 total.

Heap Memory used 41.60 MB is 37% of Committed Heap Memory 110 MB. Max Heap Memory is 1000 MB.

Non Heap Memory used 32.06 MB is 91% of Committed Non Heap Memory 35.19 MB. Max Non Heap Memory is -1 B.

Configured Capacity	: 38.58 GB
DFS Used	: 28 KB
Non DFS Used	: 29.99 GB