Multi node cluster set up:

1. Ssh-keygen

- 2. Cd .ssh/
- 3. Is

```
The programs included with the Ubuntu system are free software;
the exact distribution terms for each program are described in the
individual files in /usr/share/doc/*/copyright.
Ubuntu comes with ABSOLUTELY NO WARRANTY, to the extent permitted
applicable law.
To run a command as administrator (user "root"), use "sudo <commandata"
See "man sudo root" for details.
ubuntu@ip-172-31-26-77:~$ ssh-keygen
Generating public/private rsa key pair.
Enter file in which to save the key (/home/ubuntu/.ssh/id_rsa):
Enter passphrase (empty for no passphrase):
Enter same passphrase again:
Your identification has been saved in /home/ubuntu/.ssh/id rsa
Your public key has been saved in /home/ubuntu/.ssh/id rsa.pub
The key fingerprint is:
SHA256:hdN1+/YC7FWJy8ULonrC/PIcoNAMPUE60dmDVHAxBd8 ubuntu@ip-172-
The key's randomart image is:
+---[RSA 3072]----+
    0 = .0Eo..o +.|
         ^{\text{o}+}
+----[SHA256]----+
ubuntu@ip-172-31-26-77:~$
+----[SHA256]----+
ubuntu@ip-172-31-26-77:~$ cd .ssh/
ubuntu@ip-172-31-26-77:~/.ssh$ ls
authorized_keys id_rsa id_rsa.pub
```

4:

• cat id rsa.pub >> authorized keys

ubuntu@ip-172-31-26-77:~/.ssh\$

- Cd
- Ssh localhost

```
authorized keys id rsa id rsa.pub
ubuntu@ip-172-31-26-77:~/.ssh$ cat id_rsa.pub >> authorized_keys
ubuntu@ip-172-31-26-77:~/.ssh$ cd
ubuntu@ip-172-31-26-77:~$ ssh localhost
The authenticity of host 'localhost (127.0.0.1)' can't be established.
ECDSA key fingerprint is SHA256:vcQy/f1aBlz5dPvYWX10Paf/FeX6MbZxwkc4cOu7GDo.
Are you sure you want to continue connecting (yes/no/[fingerprint])? yes
Warning: Permanently added 'localhost' (ECDSA) to the list of known hosts.
Welcome to Ubuntu 20.04.6 LTS (GNU/Linux 5.15.0-1036-aws x86_64)

* Documentation: https://help.ubuntu.com
```

5: exit 6: ls 7: wget. 8: chmod 400

```
ubuntu@ip-172-31-26-77:~$ ls
nw-nishant.pem
ubuntu@ip-172-31-26-77:~$ chmod 400 nw-nishant.pem
ubuntu@ip-172-31-26-77:~$ sudo nano /etc/hosts
ubuntu@ip-172-31-26-77:~$ cat /etc/hosts
172.31.26.77 nn
172.31.24.222 dn
# The following lines are desirable for IPv6 capable hosts
::1 ip6-localhost ip6-loopback
fe00::0 ip6-localnet
ff00::0 ip6-mcastprefix
ff02::1 ip6-allnodes
ff02::2 ip6-allrouters
ff02::3 ip6-allhosts
ubuntu@ip-172-31-26-77:~$
```

9:

```
ff02::1 ip6-allnodes
ff02::2 ip6-allrouters
ff02::3 ip6-allrouters
ff02::3 ip6-allhosts
ubuntug1p-172-31-26-77:~$ ssh -i nw-nishant.pem dn
The authenticity of host 'dn (172.31.24.222)' can't be established.
ECDSA key fingerprint is SHA256:tNU57Rtu3Rv0iIS25CmfL2IGnC6AYRBJj52iin192L8.
Are you sure you want to continue connecting (yes/no/[fingerprint])? yes
Warning: Permanently added 'dn,172.31.24.222' (ECDSA) to the list of known hosts.
Welcome to Ubuntu 20.04.6 LTS (GNU/Linux 5.15.0-1036-aws x86_64)

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

System information as of Thu Oct 5 13:53:47 UTC 2023
```

10: Scp -i nw-nishant.pem .ssh/id_rsa .ssh/id_rsa.pub dn:/home/ubuntu/.ssh

```
dbuntu8ip-172-31-24-222:~$ exit
logout
Connection to dn closed.
dbuntu8ip-172-31-26-77:~$ is
nw-nishant.pem
dbuntu8ip-172-31-26-77:~$ scp .bashrc dn:/home/ubuntu
dbuntu8ip-172-31-26-77:~$ scp .bashrc dn:/home/ubuntu
dbuntu8ip-172-31-26-77:~$ scp -i nw-nishant.pem .ssh/id_rsa .ssh/id_rsa.pub dn:/home/ubuntu/.ssh
id_rsa
id_rsa.pub
dbuntu8ip-172-31-26-77:~$ |
```

11: ls.

```
id_rsa.pub
ubuntu@ip-172-31-26-77:~$ ls
nw-nishant.pem
ubuntu@ip-172-31-26-77:~$ ssh -i nw-nishant.pem dn
Welcome to Ubuntu 20.04.6 LTS (GNU/Linux 5.15.0-1036-aws x86_64)

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

System information as of Thu Oct 5 13:59:39 UTC 2023

System load: 0.0 Processes: 98
```

12:

```
ubuntu@ip-172-31-24-222:~$ 1s
ubuntu@ip-172-31-24-222:~$ cd .ssh/
ubuntu@ip-172-31-24-222:~\.ssh$ 1s
authorized_keys id_rsa id_rsa.pub
ubuntu@ip-172-31-24-222:~\.ssh$ cat id_rsa.pub >> authorized_keys
ubuntu@ip-172-31-24-222:~\.ssh$ csh localhost
The authenticity of host 'localhost (127.0.0.1)' can't be established.
ECDSA key fingerprint is SHA256:tMU57Rtu3RvOiIS25CmfLz1GnC6AYRBJj52iin192L8.
Are you sure you want to continue connecting (yes/no/[fingerprint])? yes
Warning: Permanently added 'localhost' (ECDSA) to the list of known hosts.
Welcome to Ubuntu 20.04.6 LTS (GNU/Linux 5.15.0-1036-aws x86_64)
```

13.

```
ubuntu@ip-172-31-24-222:~$ exit
logout
Connection to localhost closed.
ubuntu@ip-172-31-24-222:~/.ssh$ exit
logout
Connection to dn closed.
ubuntu@ip-172-31-26-77:~$ ls
nw-nishant.pem
ubuntu@ip-172-31-26-77:~$ rm nw-nishant.pem
rm: remove write-protected regular file 'nw-nishant.pem'? yes
ubuntu@ip-172-31-26-77:~$ ls
ubuntu@ip-172-31-26-77:~$ |
```

```
ubuntu@ip-172-31-24-222:~$ exit
logout
Connection to dn closed.
ubuntu@ip-172-31-26-77:~$ ls
ubuntu@ip-172-31-26-77:~$ sudo apt update^C
ubuntu@ip-172-31-26-77:~$ sudo apt update
```

```
Reading package lists... Done
Building dependency tree
Reading state information... Done
111 packages can be upgraded. Run 'apt list --upgradable' to see them.
ubuntu@ip-172-31-26-77:~$ sudo apt install dsh -y
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following additional packages will be installed.
```

```
Processing triggers for marcob [2.5.1-1] ...

Processing triggers for libc-bin (2.31-Oubuntu9.9) ...

ubuntu8ip-172-31-26-77:-$ sudo nano /etc/dsh/machines.list

ubuntu8ip-172-31-26-77:-$ dsh -uptime

dsh: invalid option -- 'u'

ubuntu8ip-172-31-26-77:-$ dsh -a uptime

The authenticity of host 'nn (172.31.26.77)' can't be established.

ECDSA key fingerprint is SHA256:vcQy/fiaBlz5dPvYWX10Paf/FeX6MbZxwkc4cOu7GDo.

Are you sure you want to continue connecting (yes/no/[fingerprint])? yes

Warning: Permanently added 'nn,172.31.26.77' (ECDSA) to the list of known hosts.

14:12:43 up 31 min, 1 user, load average: 0.01, 0.09, 0.05

14:12:44 up 31 min, 0 users, load average: 0.00, 0.00, 0.00

ubuntu8ip-172-31-26-77:-$
```

17.

```
Processing triggers for libgdk-pixbuf2.0-0:amd64 (2.40.0+dfsg-ubuntu@ip-172-31-26-77:~$ dsh -a javac -version javac 17.0.8.1 javac 17.0.8.1 ubuntu@ip-172-31-26-77:~$
```

18.

```
javac 17.0.8.1
javac 17.0.8.1
ubuntu@ip-172-31-26-77:~$ dsh -a wget https://archive.apache.org/dist/hadoop/core/hadoop-2.5.0/hadoop-2.5.0.tar.gz
```

19.

```
nadoop-2.5.0/bin/yarn.cmd
ubuntu@ip-172-31-26-77:~$ dsh -a ls
nadoop-2.5.0
nadoop-2.5.0.tar.gz
nadoop-2.5.0
nadoop-2.5.0
ubuntu@ip-172-31-26-77:~$
```

20.

```
hadoop-2.5.0
hadoop-2.5.0.tar.gz
ubuntu@ip-172-31-26-77:~$ dsh -a sudo mv hadoop-2.5.0 /usr/local/hadoop/
ubuntu@ip-172-31-26-77:~$ dsh -a ls
hadoop-2.5.0.tar.gz
hadoop-2.5.0.tar.gz
ubuntu@ip-172-31-26-77:~$ [
```

```
hbuntu@ip-172-31-26-77:/usr/local/hadoop/etc$ cd hadoop/
ubuntu@ip-172-31-26-77:/usr/local/hadoop/etc$ cd hadoop/
ubuntu@ip-172-31-26-77:/usr/local/hadoop/etc$ cd hadoop/
ubuntu@ip-172-31-26-77:/usr/local/hadoop/etc/hadoop$ ls

capacity-scheduler.xml hadoop-entrics.properties hadoop-metrics2.properties httpfs-log4j.properties mapred-env.sm ssi-server.xml.example httpfs-log4j.properties mapred-env.sm ssi-server.xml.example httpfs-log4j.properties mapred-env.sm ssi-server.xml.example httpfs-log4j.properties slaves

salpado-env.cmd hdfs-site.xml log4j.properties slaves
```

23.

hadoop-env.cmd	hdfs-site.xml	log4j.properties	slaves
ubuntu@ip-172-31-26-77:	/usr/local/hadoop/etc/hadoop\$	nano hadoop-env.sh	
ubuntu@ip-172-31-26-77:	/usr/local/hadoop/etc/hadoop\$	nano hadoop-env.sh	
ubuntu@ip-172-31-26-77:	/usr/local/hadoop/etc/hadoop\$	nano core-site.xml	
ubuntu@ip-172-31-26-77:	/usr/local/hadoop/etc/hadoop\$	nano hdfs-site.xml	
ubuntu@ip-172-31-26-77:,	/usr/local/hadoop/etc/hadoop\$	nano mapred-site.xml.te	mplate
ubuntu@ip-172-31-26-77:	/usr/local/hadoop/etc/hadoop\$		

24.

```
ubuntu@ip-172-31-26-77:/usr/local/hadoop/etc/hadoop% nano mapred-site.xml.template
ubuntu@ip-172-31-26-77:/usr/local/hadoop/etc/hadoop% nano slaves
ubuntu@ip-172-31-26-77:/usr/local/hadoop/etc/hadoop% nano masters
ubuntu@ip-172-31-26-77:/usr/local/hadoop/etc/hadoop%
```

25.

26.

```
Last login: Thu Oct 5 14:04:47 2023 from 172.31.26.77

ubuntu@itp-172-31-24-222:* cd /usr/local/hadoop/

ubuntu@itp-172-31-24-222:/usr/local/hadoop/

starting yarn daemons

starting yarn daemons

starting resourcemanager, logging to /usr/local/hadoop//logs/yarn-ubuntu-resourcemanager-ip-172-31-24-222.out

dn: ssh: Could not resolve hostname dn: Temporary failure in name resolution

ubuntu@ip-172-31-24-222:/usr/local/hadoop$
```

```
dn: ssh: Could not resolve hostname dn: Temporary failure in name resolution

ubuntu@ip-172-31-24-222:/usr/local/hadoop$ sudo nano /etc/hosts

ubuntu@ip-172-31-24-222:/usr/local/hadoop$ cat /etc/hosts

172.31.26.77 nn

172.31.24.222 dn

# The following lines are desirable for IPv6 capable hosts
::1 ip6-localhost ip6-loopback
fe00::0 ip6-localnet
ff00::0 ip6-mcastprefix
ff02::1 ip6-allnodes
ff02::2 ip6-allrouters
ff02::3 ip6-allhosts
ubuntu@ip-172-31-24-222:/usr/local/hadoop$
```

```
starting resourcemanager, logging to /usr/local/hadoop//logs/yarn-ub
dn: starting nodemanager, logging to /usr/local/hadoop/logs/yarn-ubu
ubuntu@ip-172-31-24-222:/usr/local/hadoop$ jps
19346 NodeManager
19460 Jps
ubuntu@ip-172-31-24-222:/usr/local/hadoop$
```

30.

```
ubuntu9ip=772-31-26-77/Masty local/hadoop$ | montuple |
```

31.

```
| 19071 | Jps | 18751 | NameNode | 19760 | Jps | 18969 | SecondaryNameNode | 18751 | NameNode | 18751 | NameNode | 19782 | Jps | 19645 | DataNode | 18751 | DataNode
```

32.

```
19782 DataNode

ubuntudip-172-31-26-77:/usr/local/hadoop% hadoop jar %HADOOP_PREFIX/share/hadoop/mapreduce/hadoop-*examples*.jar pi 10 100000

Number of Maps = 10

Samples per Hap = 1000000

Wrote input for Map #0

Wrote input for Map #1

Wrote input for Map #2

Wrote input for Map #3

Wrote input for Map #3

Wrote input for Map #4

Wrote input for Map #4

Wrote input for Map #5

Wrote input for Map #6

Wrote input for Map #8

Starting Job

Sta
```

```
Physical memory (bytes) snapshot=0
              Virtual memory (bytes) snapshot=0
              Total committed heap usage (bytes) = 1841168384
       Shuffle Errors
              BAD ID=0
              CONNECTION=0
              IO ERROR=0
               WRONG LENGTH=0
              WRONG_MAP=0
WRONG_REDUCE=0
       File Input Format Counters
              Bytes Read=1180
       File Output Format Counters
              Bytes Written=97
Job Finished in 2.897 seconds
ubuntu@ip-172-31-26-77:/usr/local/hadoop$ 🛮
```

```
File Output Format Counters
Bytes Writteney7

Job Finished in 2.897 seconds
Estimated value of Pi is 3.141552000000000000000000000

dubuntudis-1-2-3-1-26-772/usr/local/hadoop?dahare/hadoop/mapreduce/
ubuntudis-1-72-31-26-772/usr/local/hadoop?ahare/hadoop/mapreducei is

ubuntudis-1-72-31-26-772/usr/local/hadoop?ahare/hadoop/mapreducei is

ubuntudis-1-72-31-36-772/usr/local/hadoop/share/hadoop/mapreducei is
```

35.

```
Estimated value of Pi is 3.1415820000000000000000

ubuntudip-172-31-26-771/usr/local/hadoop/mapreduce/
ubuntudip-172-31-26-771/usr/local/hadoop/share/hadoop/mapreduce8 | 18

subspicious situation of the state of t
```

36.

```
HDFS: Number of bytes written=528709900
                      HDFS: Number of read operations=4
HDFS: Number of large read operations=0
HDFS: Number of write operations=3
           Map-Reduce Framework
                      Map input records=5287099
                      Map output records=5287099
                      Input split bytes=82
Spilled Records=0
                      Failed Shuffles=0
Merged Map outputs=0
                      GC time elapsed (ms)=91
                      CPU time spent (ms)=0
Physical memory (bytes) snapshot=0
                      Virtual memory (bytes) snapshot=0
Total committed heap usage (bytes)=29032448
           org.apache.hadoop.examples.terasort.TeraGen$Counters
CHECKSUM=11352382061958180
           File Input Format Counters
                     Bytes Read=0
           File Output Format Counters
                      Bytes Written=528709900
ubuntu@ip-172-31-26-77:/usr/local/hadoop/share/hadoop/mapreduce$
```

```
Bytes Written=528709900

23/10/05 15:14:38 INFO terasort.TeraSort: done
ubuntu@ip-172-31-26-77:/usr/local/hadoop/share/hadoop/mapreduce$ hadoop dfs -ls input/
DEFRECATED: Use of this script to execute hdfs command is deprecated.
Instead use the hdfs command for it.

Found 2 items
-rw-r--r-- 3 ubuntu supergroup 0 2023-10-05 15:11 input/_SUCCESS
-rw-r--r-- 3 ubuntu supergroup 528709900 2023-10-05 15:11 input/part-m-00000
ubuntu@ip-172-31-26-77:/usr/local/hadoop/share/hadoop/mapreduce$
```

40.

