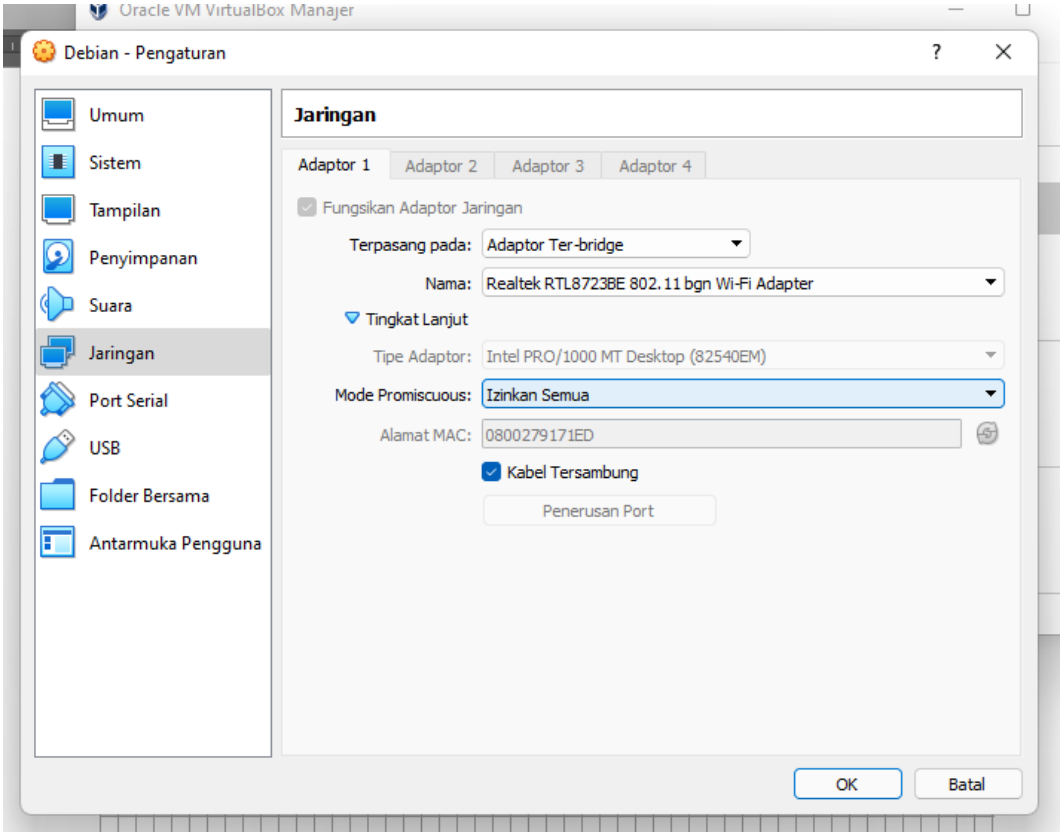
JOBSHEET 7
HADOOP MULTICLUSTER

Arranged By:

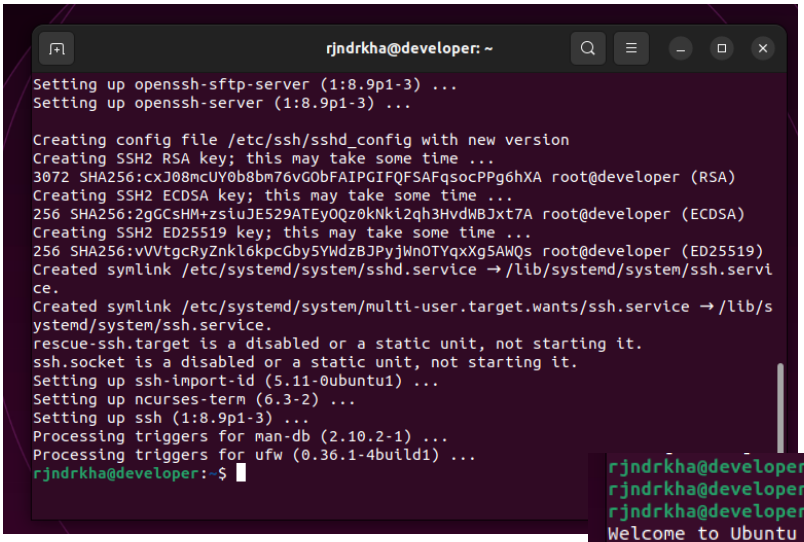
Rajendra Rakha Arya Prabaswara
(1941720080/20)

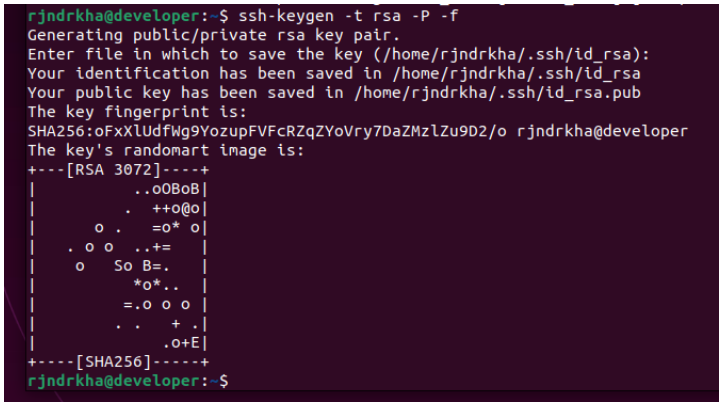
PROGRAM STUDI D-IV TEKNIK INFORMATIKA
JURUSAN TEKNOLOGI INFORMASI
POLITEKNIK NEGERI MALANG

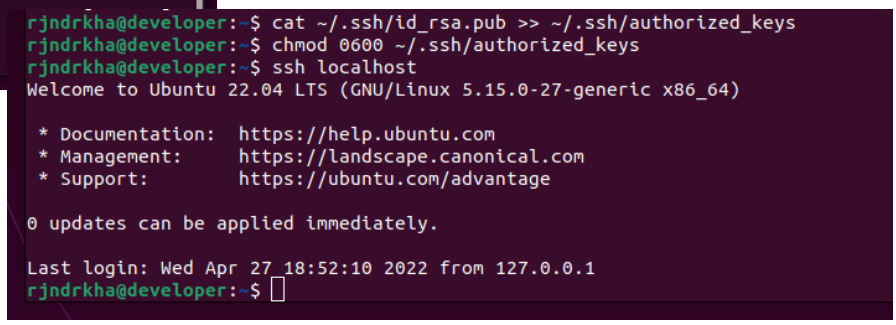
1. Konfigurasi Jaringan di VM



2. Instalasi SSH dan Konfigurasi SSH







3. **Instalasi JDK 8**

```
rjndrkha@developer:~$ java -version
openjdk version "1.8.0_312"
OpenJDK Runtime Environment (build 1.8.0_312-8u312-b07-0ubuntu1-b07)
OpenJDK 64-Bit Server VM (build 25.312-b07, mixed mode)
rjndrkha@developer:~$
```

4. **Download & Configuration Hadoop**
   - **Konfigurasi home direktori java untuk Hadoop**

```
##
## {YARN_xyz|HDFS_xyz} > HADOOP_xyz > hard-coded defaults
##

# Many of the options here are built from the perspective that users
# may want to provide OVERWRITING values on the command line.
# For example:
#
export JAVA_HOME=/usr/lib/jvm/java-8-openjdk-amd64/
#
# Therefore, the vast majority (BUT NOT ALL!) of these defaults
# are configured for substitution and not append.  If append

^G Help      ^O Write Out ^W Where Is  ^K Cut       ^T Execute   ^C Location
^X Exit      ^R Read File ^\ Replace   ^U Paste     ^J Justify   ^/ Go To Line
```

   - **Pindahkan folder Hadoop ke /usr/local/hadoop**

```
rjndrkha@Developer:~$ sudo mv hadoop /usr/local/hadoop
[sudo] password for rjndrkha:
rjndrkha@Developer:~$
```

   - **Tambahkan path untuk Hadoop dan Java**

```
rjndrkha@Developer: ~
  GNU nano 6.2                    /etc/environment *
PATH="/usr/local/sbin:/usr/local/bin:/usr/sbin:/usr/bin:/sbin:/bin:/usr/games:/>
JAVA_HOME="/usr/lib/jvm/java-8-openjdk-amd64/jre"
```

   - **Menambah user hadoopuser**

```
rjndrkha@Developer:~$ sudo adduser hadoopuser
Adding user `hadoopuser' ...
Adding new group `hadoopuser' (1001) ...
Adding new user `hadoopuser' (1001) with group `hadoopuser' ...
Creating home directory `/home/hadoopuser' ...
Copying files from `/etc/skel' ...
New password:
BAD PASSWORD: The password fails the dictionary check - it is based on a diction
ary word
Retype new password:
passwd: password updated successfully
Changing the user information for hadoopuser
Enter the new value, or press ENTER for the default
        Full Name []:
        Room Number []:
        Work Phone []:
        Home Phone []:
        Other []:
Is the information correct? [Y/n] y
rjndrkha@Developer:~$
```

- **Membuat Nama Group**

```
rjndrkha@Developer:~$ sudo usermod -aG hadoopuser hadoopuser
rjndrkha@Developer:~$ sudo chown hadoopuser:root -R /usr/local/hadoop/
rjndrkha@Developer:~$ sudo adduser hadoopuser sudo
Adding user `hadoopuser' to group `sudo' ...
Adding user hadoopuser to group sudo
Done.
rjndrkha@Developer:~$
```

- **Mapping name node dan data node**

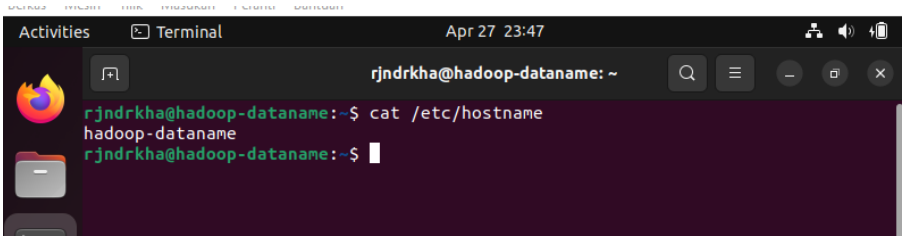
```
rjndrkha@Developer: ~
.
Unpacking net-tools (1.60+git20181103.0eebece-1ubuntu5) ...
Setting up net-tools (1.60+git20181103.0eebece-1ubuntu5) ...
Processing triggers for man-db (2.10.2-1) ...
rjndrkha@Developer:~$ ifconfig
enp0s3: flags=4163<UP,BROADCAST,RUNNING,MULTICAST>  mtu 1500
        inet 192.168.1.13  netmask 255.255.255.0  broadcast 192.168.1.255
        inet6 fe80::71bd:f3a:2e42:ee27  prefixlen 64  scopeid 0x20<link>
        ether 08:00:27:35:96:0a  txqueuelen 1000  (Ethernet)
        RX packets 501438  bytes 722186841 (722.1 MB)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 247142  bytes 16766532 (16.7 MB)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0

lo: flags=73<UP,LOOPBACK,RUNNING>  mtu 65536
        inet 127.0.0.1  netmask 255.0.0.0
        inet6 ::1  prefixlen 128  scopeid 0x10<host>
        loop  txqueuelen 1000  (Local Loopback)
        RX packets 17727  bytes 1601992 (1.6 MB)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 17727  bytes 1601992 (1.6 MB)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0

rjndrkha@Developer:~$
```

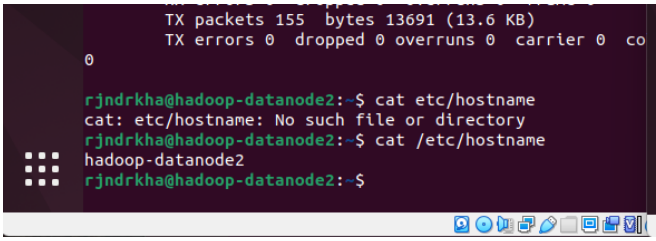
IPADDR : 192.168.1.13

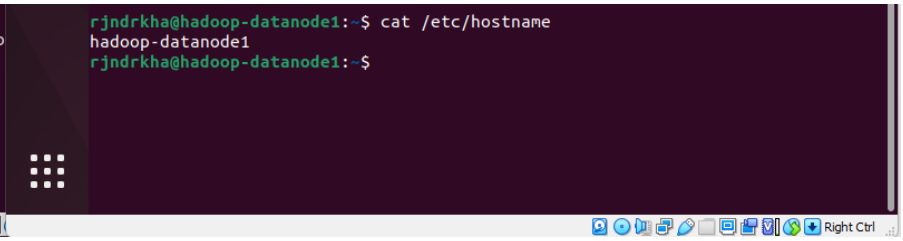
```
rjndrkha@Developer: ~
  GNU nano 6.2                        /etc/hosts *
127.0.0.1       localhost
127.0.1.1       Developer

192.168.1.13 Developer
192.168.1.14 data-node
192.168.1.15 data-node

# 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
```

- **Clone VM name-node**

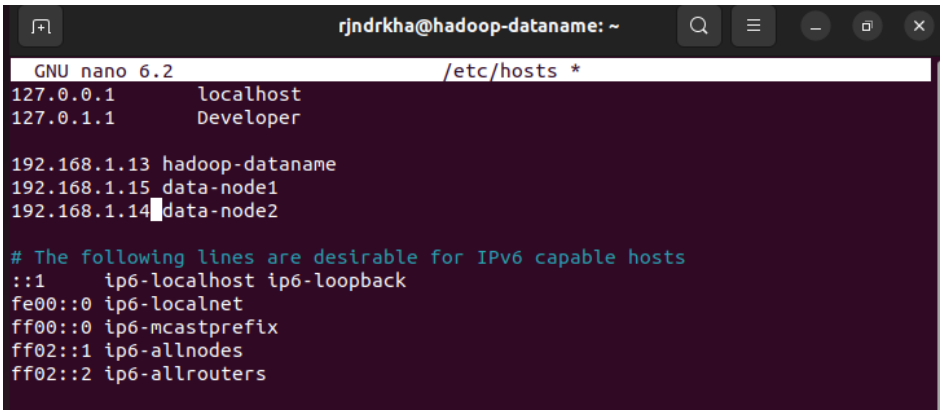
- **Konfigurasi hostname**

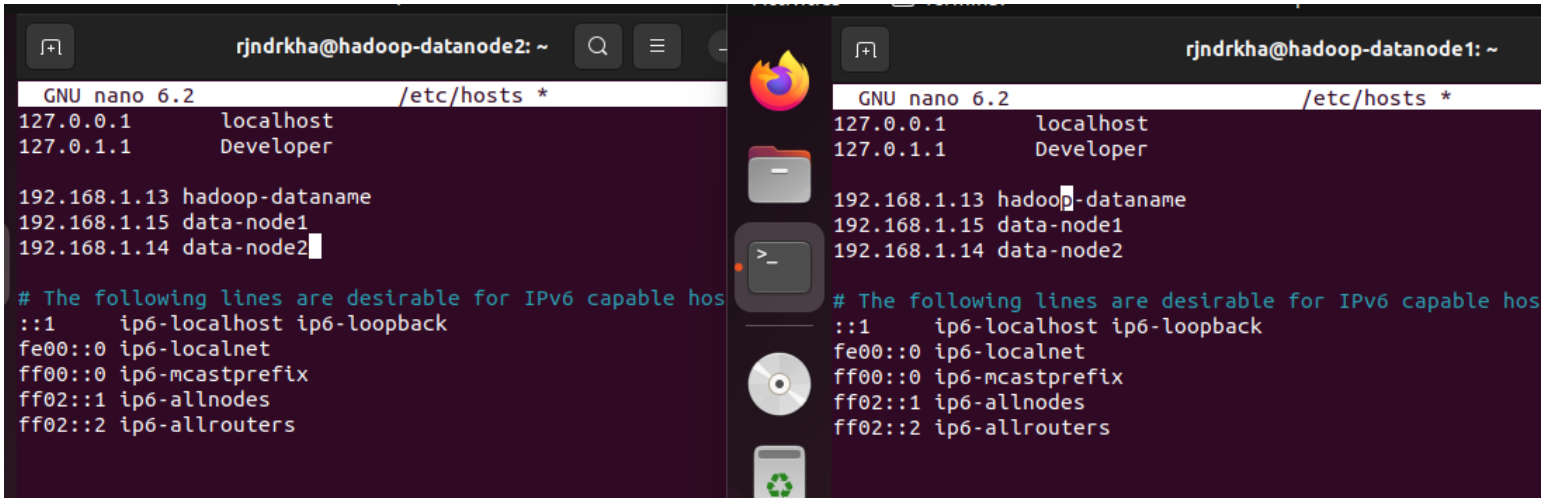






- **Konfigurasi IP baru pada semua vm**





5. Konfigurasi User Hadoop

```
hadoopuser@hadoop-dataname: ~
rjndrkha@hadoop-dataname:~$ su - hadoopuser
Password:
To run a command as administrator (user "root"), use "sudo <command>".
See "man sudo_root" for details.

hadoopuser@hadoop-dataname:~$
```

Generating SSH KEY

```
hadoopuser@hadoop-dataname:~$ ssh-keygen -t rsa
Generating public/private rsa key pair.
Enter file in which to save the key (/home/hadoopuser/.ssh/id_rsa):
Created directory '/home/hadoopuser/.ssh'.
Enter passphrase (empty for no passphrase):
Enter same passphrase again:
Your identification has been saved in /home/hadoopuser/.ssh/id_rsa
Your public key has been saved in /home/hadoopuser/.ssh/id_rsa.pub
The key fingerprint is:
SHA256:O26AxyFFMsUS+8o7u0JsLXaDOOkxxPgTz8NoXZURyXs hadoopuser@hadoop-dataname
The key's randomart image is:
+---[RSA 3072]----+
|    +=o.o=       |
|    .+o =        |
|o   .o . .       |
|.o. ..o . E      |
|.= X =..S.       |
|+o%.@.+  .       |
|.*o+o+ .o        |
| .. .. ...       |
|   .++ ..        |
+----[SHA256]-----+
hadoopuser@hadoop-dataname:~$
```

COPY SSH KEY

```
+----[SHA256]-----+
hadoopuser@hadoop-dataname:~$ ssh-copy-id hadoopuser@hadoop-dataname
/usr/bin/ssh-copy-id: INFO: Source of key(s) to be installed: "/home/hadoopuser
/.ssh/id_rsa.pub"
The authenticity of host 'hadoop-dataname (192.168.1.13)' can't be established.
ED25519 key fingerprint is SHA256:e/KXVne/6R1zKlawpxUjw8TCiDIlWy1X1ARPGTfoYDE.
This key is not known by any other names
Are you sure you want to continue connecting (yes/no/[fingerprint])? 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 promp
ted now it is to install the new keys
hadoopuser@hadoop-dataname's password:
```

```
hadoopuser@hadoop-dataname:~$ ssh-copy-id hadoopuser@hadoop-datanode1
/usr/bin/ssh-copy-id: INFO: Source of key(s) to be installed: "/home/hadoopuser
/.ssh/id_rsa.pub"
The authenticity of host 'hadoop-datanode1 (192.168.1.15)' can't be established
.
ED25519 key fingerprint is SHA256:e/KXVne/6R1zKlawpxUjw8TCiDIlWy1X1ARPGTfoYDE.
This host key is known by the following other names/addresses:
    ~/.ssh/known_hosts:1: [hashed name]
Are you sure you want to continue connecting (yes/no/[fingerprint])? 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 promp
ted now it is to install the new keys
hadoopuser@hadoop-datanode1's password:

Number of key(s) added: 1

Now try logging into the machine, with:   "ssh 'hadoopuser@hadoop-datanode1'"
and check to make sure that only the key(s) you wanted were added.

hadoopuser@hadoop-dataname:~$
```

```
hadoopuser@hadoop-dataname:~$ ssh-copy-id hadoopuser@hadoop-datanode2
/usr/bin/ssh-copy-id: INFO: Source of key(s) to be installed: "/home/hadoopuser
/.ssh/id_rsa.pub"
The authenticity of host 'hadoop-datanode2 (192.168.1.14)' can't be established
.
ED25519 key fingerprint is SHA256:e/KXVne/6R1zKlawpxUjw8TCiDIlWy1X1ARPGTfoYDE.
This host key is known by the following other names/addresses:
    ~/.ssh/known_hosts:1: [hashed name]
    ~/.ssh/known_hosts:4: [hashed name]
Are you sure you want to continue connecting (yes/no/[fingerprint])? 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 promp
ted now it is to install the new keys
hadoopuser@hadoop-datanode2's password:

Number of key(s) added: 1

Now try logging into the machine, with:   "ssh 'hadoopuser@hadoop-datanode2'"
and check to make sure that only the key(s) you wanted were added.
```

6. Konfigurasi Hadoop

- **Konfigurasi file core-site.xml**

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

<configuration>
<property>
  <name>fs.defaulFS</name>
  <value>hdfs://hadoop-dataname:9000</value>
</property>
</configuration>
```

- **Konfigurasi lokasi direktori data Hadoop dan jumlah cluster**

```
<configuration>
<property>
  <name>dfs.namenode.name.dir</name>
  <value>/usr/local/hadoop/data/nameNode</value>
</property>
<property>
  <name>dfs.datanode.data.dir</name>
  <value>/usr/local/hadoop/data/dataNode</value>
</property>
<property>
  <name>dfs.replication</name>
  <value>2</value>
</property>
</configuration>
```

- **Konfigurasi mengatur nama worker datanode yang digunakan pada file workers**

```
  GNU nano 6.2                /usr/local/
localhost

hadoop-datanode1
hadoop-datanode2
```

- **Copy Konfigurasi ke semua cluster**

```
                                                hadoopuser@hadoop-dataname: ~
hadoopuser@hadoop-dataname:~$ scp /usr/local/hadoop/etc/hadoop/* hadoop-datanode1:/usr/local/hadoop/etc/hadoop/
capacity-scheduler.xml                        100% 9213   266.3KB/s   00:00
configuration.xsl                             100% 1335    33.1KB/s   00:00
container-executor.cfg                        100% 2567     1.0MB/s   00:00
core-site.xml                                 100%  869   690.7KB/s   00:00
hadoop-env.cmd                                100% 3999     2.1MB/s   00:00
hadoop-env.sh                                 100%   16KB 402.0KB/s   00:00
hadoop-metrics2.properties                    100% 3321     1.8MB/s   00:00
hadoop-policy.xml                             100%   11KB   6.7MB/s   00:00
hadoop-user-functions.sh.example              100% 3414     1.0MB/s   00:00
hdfs-rbf-site.xml                             100%  683   183.5KB/s   00:00
hdfs-site.xml                                 100% 1065   749.3KB/s   00:00
httpfs-env.sh                                 100% 1484   767.5KB/s   00:00
httpfs-log4j.properties                       100% 1657   977.3KB/s   00:00
httpfs-site.xml                               100%  620   408.9KB/s   00:00
kms-acls.xml                                  100% 3518     2.1MB/s   00:00
kms-env.sh                                    100% 1351   163.0KB/s   00:00
kms-log4j.properties                          100% 1860   593.8KB/s   00:00
kms-site.xml                                  100%  682   497.4KB/s   00:00
log4j.properties                              100%   13KB 757.2KB/s   00:00
mapred-env.cmd                                100%  951   377.6KB/s   00:00
mapred-env.sh                                 100% 1764   466.6KB/s   00:00
mapred-queues.xml.template                    100% 4113     2.3MB/s   00:00
mapred-site.xml                               100%  758   441.1KB/s   00:00
/usr/local/hadoop/etc/hadoop/shellprofile.d: not a regular file
ssl-client.xml.example                        100% 2316     1.1MB/s   00:00
ssl-server.xml.example                        100% 2697     1.0MB/s   00:00
user_ec_policies.xml.template                 100% 2681     1.7MB/s   00:00
workers                                       100%   45    38.8KB/s   00:00
yarn-env.cmd                                  100% 2250    64.1KB/s   00:00
yarn-env.sh                                   100% 6329     3.2MB/s   00:00
yarnservice-log4j.properties                  100% 2591     1.6MB/s   00:00
yarn-site.xml                                 100%  690    86.7KB/s   00:00
hadoopuser@hadoop-dataname:~$
```

```
hadoopuser@hadoop-dataname: ~
hadoopuser@hadoop-dataname:~$ scp /usr/local/hadoop/etc/hadoop/* hadoop-datanode2:/usr/local/hadoop/etc/hadoop/
capacity-scheduler.xml                                        100% 9213   644.7KB/s   00:00
configuration.xsl                                             100% 1335   204.0KB/s   00:00
container-executor.cfg                                        100% 2567     2.2MB/s   00:00
core-site.xml                                                 100%  869   812.4KB/s   00:00
hadoop-env.cmd                                                100% 3999   783.3KB/s   00:00
hadoop-env.sh                                                 100%   16KB   1.9MB/s   00:00
hadoop-metrics2.properties                                    100% 3321     2.2MB/s   00:00
hadoop-policy.xml                                             100%   11KB   9.2MB/s   00:00
hadoop-user-functions.sh.example                              100% 3414     2.6MB/s   00:00
hdfs-rbf-site.xml                                             100%  683   606.6KB/s   00:00
hdfs-site.xml                                                 100% 1065     1.0MB/s   00:00
httpfs-env.sh                                                 100% 1484     1.3MB/s   00:00
httpfs-log4j.properties                                       100% 1657     1.2MB/s   00:00
httpfs-site.xml                                               100%  620   530.5KB/s   00:00
kms-acls.xml                                                  100% 3518     1.0MB/s   00:00
kms-env.sh                                                    100% 1351     1.3MB/s   00:00
kms-log4j.properties                                          100% 1860     1.7MB/s   00:00
kms-site.xml                                                  100%  682   659.6KB/s   00:00
log4j.properties                                              100%   13KB   6.1MB/s   00:00
mapred-env.cmd                                                100%  951   393.0KB/s   00:00
mapred-env.sh                                                 100% 1764     1.3MB/s   00:00
mapred-queues.xml.template                                    100% 4113     3.2MB/s   00:00
mapred-site.xml                                               100%  758   759.0KB/s   00:00
/usr/local/hadoop/etc/hadoop/shellprofile.d: not a regular file
ssl-client.xml.example                                        100% 2316     1.7MB/s   00:00
ssl-server.xml.example                                        100% 2697     2.0MB/s   00:00
user_ec_policies.xml.template                                 100% 2681   961.0KB/s   00:00
workers                                                       100%   45    34.9KB/s   00:00
yarn-env.cmd                                                  100% 2250     1.7MB/s   00:00
yarn-env.sh                                                   100% 6329     4.3MB/s   00:00
yarnservice-log4j.properties                                  100% 2591     2.1MB/s   00:00
yarn-site.xml                                                 100%  690   563.5KB/s   00:00
hadoopuser@hadoop-dataname:~$
```

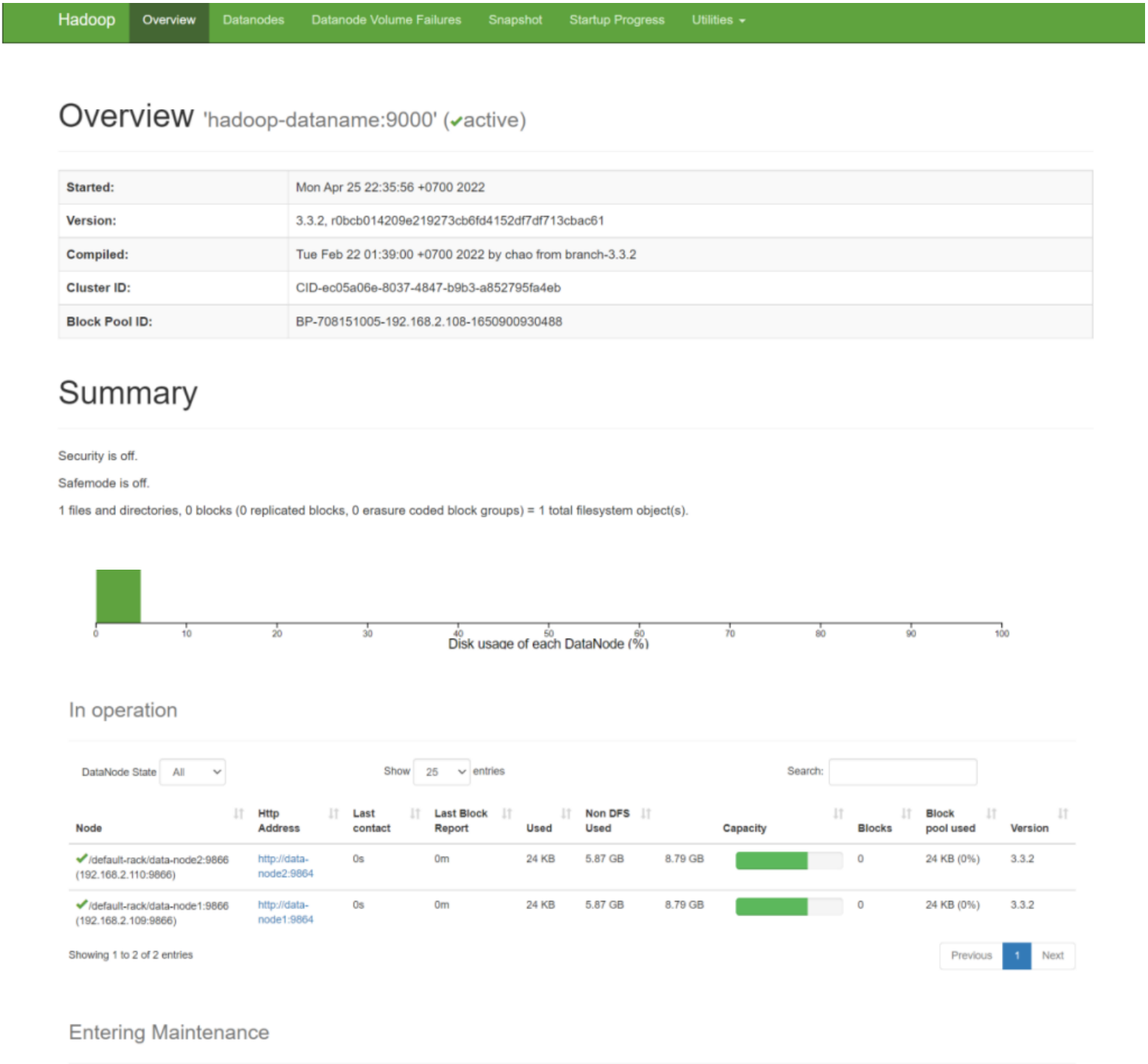
- **Code agar mengalami perubahan pada environment**

```
Try: sudo apt install <deb name>
hadoopuser@hadoop-dataname:/usr/local/hadoop/bin$ cd
hadoopuser@hadoop-dataname:~$ /usr/local/hadoop/bin/hadoop namenode -format
WARNING: Use of this script to execute namenode is deprecated.
WARNING: Attempting to execute replacement "hdfs namenode" instead.

WARNING: /usr/local/hadoop/logs does not exist. Creating.
2022-04-28 01:20:56,100 INFO namenode.NameNode: STARTUP_MSG:
/************************************************************
STARTUP_MSG: Starting NameNode
STARTUP_MSG:   host = hadoop-dataname/127.0.1.1
STARTUP_MSG:   args = [-format]
STARTUP_MSG:   version = 3.3.2
STARTUP_MSG:   classpath = /usr/local/hadoop/etc/hadoop:/usr/local/hadoop/share/hadoop/common/lib/hadoop-auth-3.3.2.jar:/usr/local/hadoop/sha
re/hadoop/common/lib/jackson-xc-1.9.13.jar:/usr/local/hadoop/share/hadoop/common/lib/kerb-common-1.0.1.jar:/usr/local/hadoop/share/hadoop/com
mon/lib/metrics-core-3.2.4.jar:/usr/local/hadoop/share/hadoop/common/lib/jetty-xml-9.4.43.v20210629.jar:/usr/local/hadoop/share/hadoop/common
/lib/javax.servlet-api-3.1.0.jar:/usr/local/hadoop/share/hadoop/common/lib/commons-codec-1.11.jar:/usr/local/hadoop/share/hadoop/common/lib/j
```

- **Menjalankan Hadoop**

```
hadoopuser@hadoop-dataname:/usr/local/hadoop/sbin$ source $HOME/.bashrc
hadoopuser@hadoop-dataname:/usr/local/hadoop/sbin$ start-dfs.sh
Starting namenodes on [hadoop-dataname]
Starting datanodes
localhost: Warning: Permanently added 'localhost' (ED25519) to the list of known hosts.
hadoop-datanode2: WARNING: /usr/local/hadoop/logs does not exist. Creating.
hadoop-datanode1: WARNING: /usr/local/hadoop/logs does not exist. Creating.
Starting secondary namenodes [hadoop-dataname]
hadoopuser@hadoop-dataname:/usr/local/hadoop/sbin$
```

7. RESULT


Hadoop
Overview
Datanodes
Datanode Volume Failures
Snapshot
Startup Progress
Utilities
Overview 'hadoop-dataname:9000' (active)
Started: Mon Apr 25 22:35:56 +0700 2022
Version: 3.3.2, r0bcb014209e219273cb6fd4152df7df713cbac61
Compiled: Tue Feb 22 01:39:00 +0700 2022 by chao from branch-3.3.2
Cluster ID: CID-ec05a06e-8037-4847-b9b3-a852795fa4eb
Block Pool ID: BP-708151005-192.168.2.108-1650900930488
Summary
Security is off.
Safemode is off.
1 files and directories, 0 blocks (0 replicated blocks, 0 erasure coded block groups) = 1 total filesystem object(s).
0 10 20 30 40 50 60 70 80 90 100
Disk usage of each DataNode (%)
In operation
DataNode State All
Show 25 entries
Search:
Node
Http Address
Last contact
Last Block Report
Used
Non DFS Used
Capacity
Blocks
Block pool used
Version
/default-rack/data-node2:9866 (192.168.2.110:9866)
http://data-node2:9864
0s
0m
24 KB
5.87 GB
8.79 GB
0
24 KB (0%)
3.3.2
/default-rack/data-node1:9866 (192.168.2.109:9866)
http://data-node1:9864
0s
0m
24 KB
5.87 GB
8.79 GB
0
24 KB (0%)
3.3.2
Showing 1 to 2 of 2 entries
Previous
1
Next
Entering Maintenance