Task 11:

Configure the Hadoop cluster and start the Hadoop cluster using the ansible playbook.

Sol:

Steps →

First we need to copy the Hadoop and jdk rpm file in the managed node i.e 172.20.10.14 & 172.20.10.9

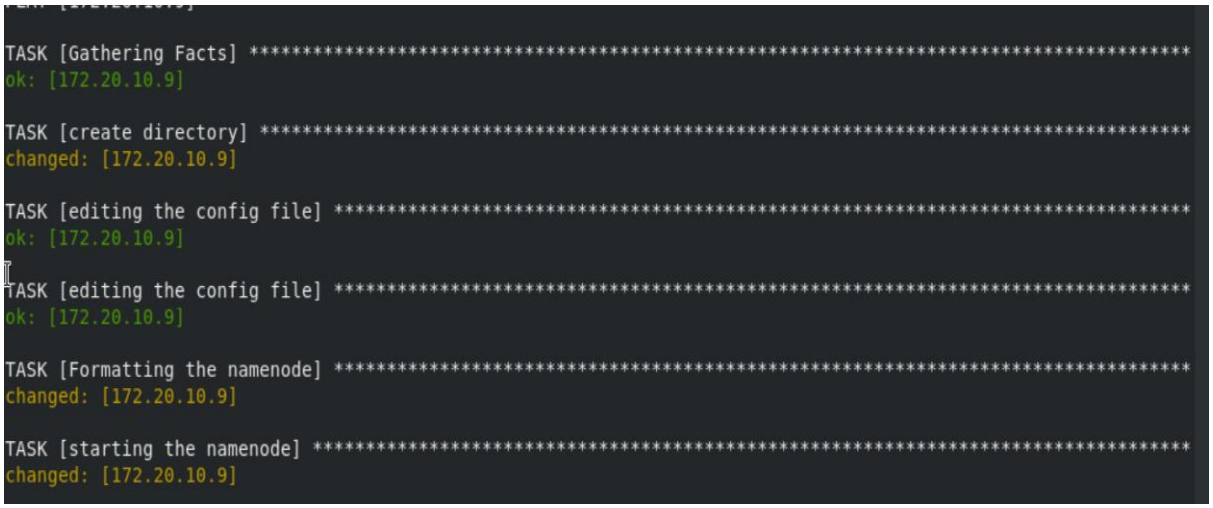
After copying the file Now we need to install both the software and also turn off the firewall or we can make the rule in firewall

```
[root@localhost aH]# vim con.yml
[root@localhost aH]# ansible-playbook con.yml

PLAY [all] *********************************************************************************************************

TASK [Gathering Facts] *********************************************************************************************
ok: [172.20.10.14]
ok: [172.20.10.9]

TASK [copying the hadoop and jdk from local to remote] *************************************************************
ok: [172.20.10.9] => (item=hadoop-1.2.1-1.x86_64.rpm)
ok: [172.20.10.14] => (item=hadoop-1.2.1-1.x86_64.rpm)
ok: [172.20.10.14] => (item=jdk-8u171-linux-x64.rpm)
ok: [172.20.10.9] => (item=jdk-8u171-linux-x64.rpm)

TASK [installation of jdk] *****************************************************************************************
changed: [172.20.10.14]
changed: [172.20.10.9]

TASK [Installation of the hadoop] **********************************************************************************
changed: [172.20.10.9]
changed: [172.20.10.14]

TASK [stop firewalld] **********************************************************************************************
ok: [172.20.10.14]
ok: [172.20.10.9]
```

Step →

After successfully installation Now we are going to configure the Hadoop NameNode on IP 172.20.10.9 and to configure we have to edit two file i.e core-site.xml and hdfs-site.xml

Here we have created a directory and need to format that directory. The created directory can be format using cmd: Hadoop namenode -format and we have used this command on the playbook

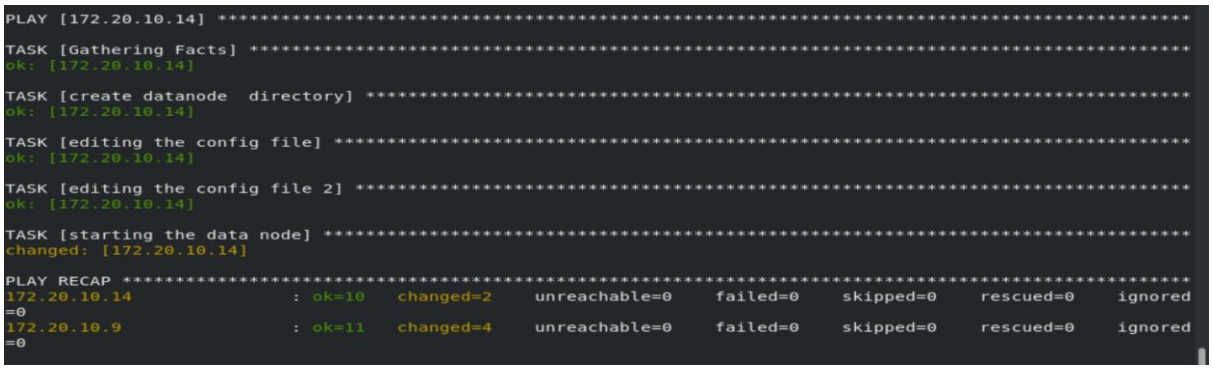


Now we have successfully edited the both config file and the directory is also formatted .After completion of these steps we can start our NameNode service of the Hadoop cluster IP 172.20.10.9 here we can see the task “starting the namenode”

Steps →

Our namenode is configured now we have to configure tha data node

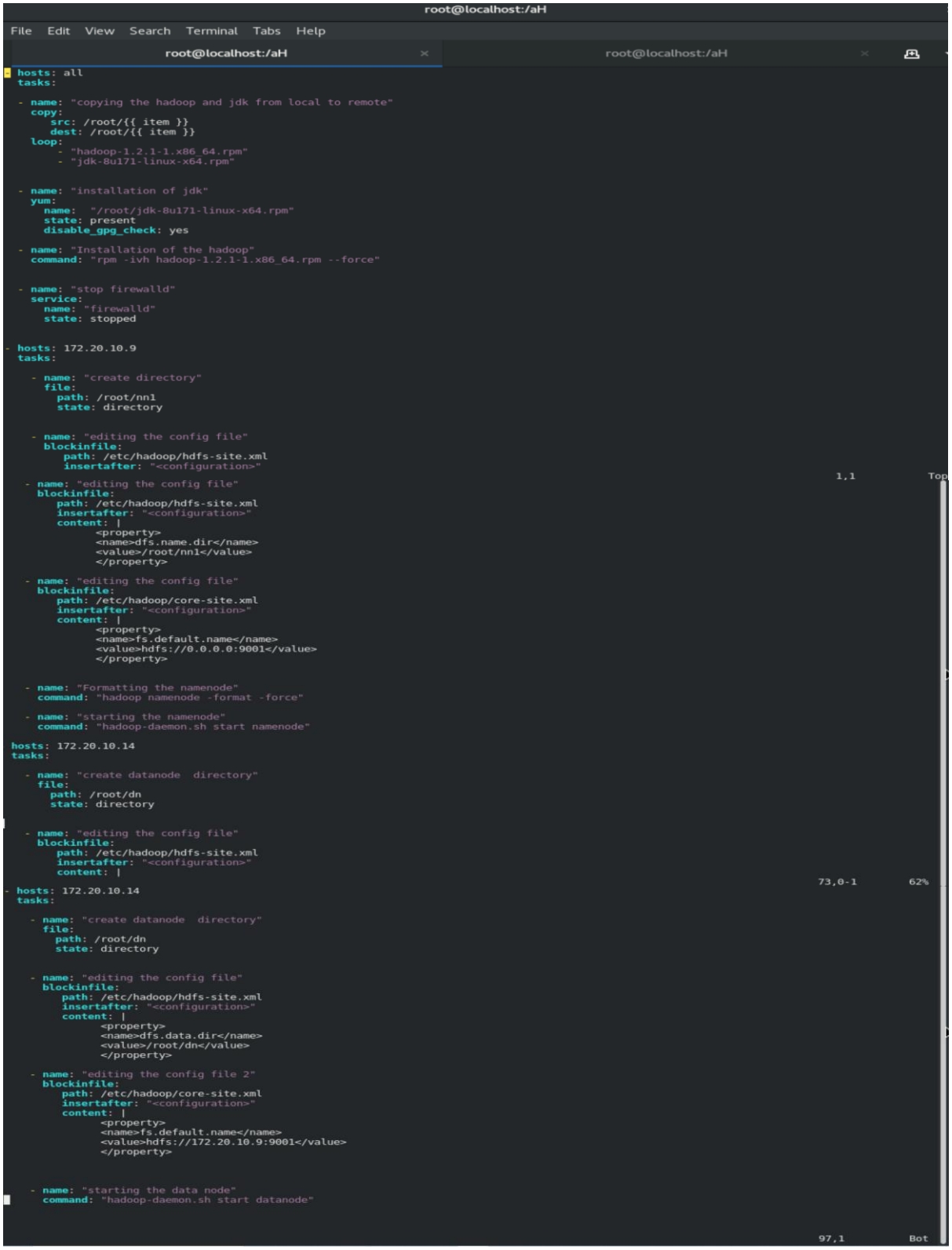
Again here we have to configure the file of 172.20.10.14 and start the datanode services.



OUTPUT:

```
[root@localhost ~]# hadoop dfsadmin -report
Configured Capacity: 53660876800 (49.98 GB)
Present Capacity: 50389725184 (46.93 GB)
DFS Remaining: 50389716992 (46.93 GB)
DFS Used: 8192 (8 KB)
DFS Used%: 0%
Under replicated blocks: 0
Blocks with corrupt replicas: 0
Missing blocks: 0

-------------------------------------------------
Datanodes available: 1 (1 total, 0 dead)

Name: 172.20.10.14:50010
Decommission Status : Normal
Configured Capacity: 53660876800 (49.98 GB)
DFS Used: 8192 (8 KB)
Non DFS Used: 3271151616 (3.05 GB)
DFS Remaining: 50389716992(46.93 GB)
DFS Used%: 0%
DFS Remaining%: 93.9%
Last contact: Sat Mar 27 14:55:14 IST 2021


[root@localhost ~]# _
```

ANSIBLE PLAYBOOK→ →


root@localhost:/aH
File Edit View Search Terminal Tabs Help
root@localhost:/aH
root@localhost:/aH
- hosts: all
tasks:
- name: "copying the hadoop and jdk from local to remote"
copy:
src: /root/{{ item }}
dest: /root/{{ item }}
loop:
- "hadoop-1.2.1-1.x86_64.rpm"
- "jdk-8u171-linux-x64.rpm"
- name: "installation of jdk"
yum:
name: "/root/jdk-8u171-linux-x64.rpm"
state: present
disable_gpg_check: yes
- name: "Installation of the hadoop"
command: "rpm -ivh hadoop-1.2.1-1.x86_64.rpm --force"
- name: "stop firewalld"
service:
name: "firewalld"
state: stopped
- hosts: 172.20.10.9
tasks:
- name: "create directory"
file:
path: /root/nn1
state: directory
- name: "editing the config file"
blockinfile:
path: /etc/hadoop/hdfs-site.xml
insertafter: "<configuration>"
1,1 Top
- name: "editing the config file"
blockinfile:
path: /etc/hadoop/hdfs-site.xml
insertafter: "<configuration>"
content: |
<property>
<name>dfs.name.dir</name>
<value>/root/nn1</value>
</property>
- name: "editing the config file"
blockinfile:
path: /etc/hadoop/core-site.xml
insertafter: "<configuration>"
content: |
<property>
<name>fs.default.name</name>
<value>hdfs://0.0.0.0:9001</value>
</property>
- name: "Formatting the namenode"
command: "hadoop namenode -format -force"
- name: "starting the namenode"
command: "hadoop-daemon.sh start namenode"
hosts: 172.20.10.14
tasks:
- name: "create datanode directory"
file:
path: /root/dn
state: directory
- name: "editing the config file"
blockinfile:
path: /etc/hadoop/hdfs-site.xml
insertafter: "<configuration>"
content: |
73,0-1 62%
- hosts: 172.20.10.14
tasks:
- name: "create datanode directory"
file:
path: /root/dn
state: directory
- name: "editing the config file"
blockinfile:
path: /etc/hadoop/hdfs-site.xml
insertafter: "<configuration>"
content: |
<property>
<name>dfs.data.dir</name>
<value>/root/dn</value>
</property>
- name: "editing the config file 2"
blockinfile:
path: /etc/hadoop/core-site.xml
insertafter: "<configuration>"
content: |
<property>
<name>fs.default.name</name>
<value>hdfs://172.20.10.9:9001</value>
</property>
- name: "starting the data node"
command: "hadoop-daemon.sh start datanode"
97,1 Bot