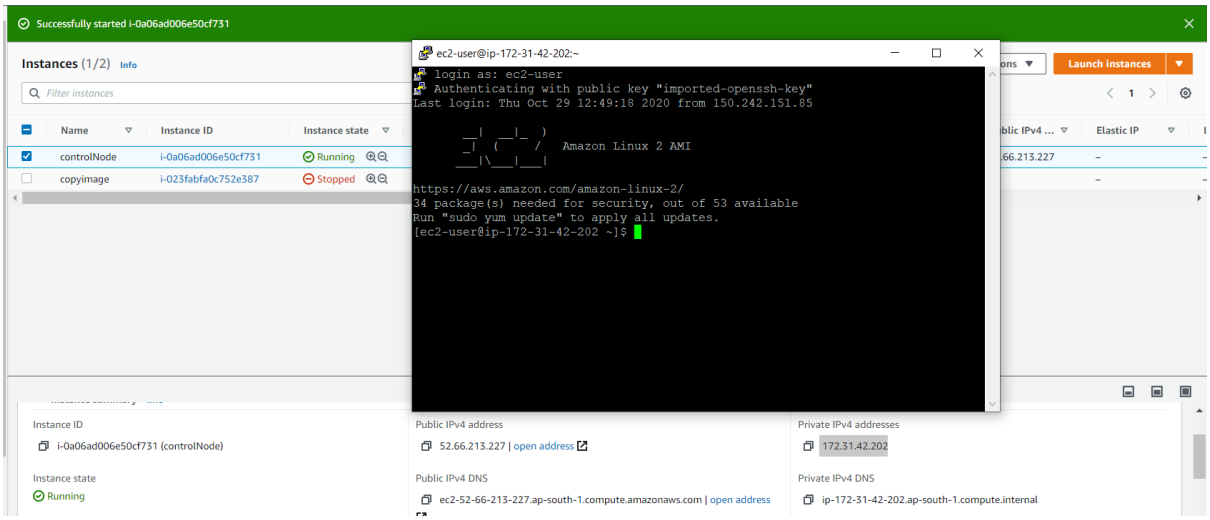
12/22/2020

RH294 Ansible Automation, configuring Hadoop Cluster.

akshay anil
AKSHAYANIL1080.GITHUB.IO/MYWEBSITE

Login into ec2-instance via putty



1.Install ansible

#sudo amazon-linux-extras install ansible2

```
[root@ip-172-31-42-202 ~]# sudo  amazon-linux-extras install ansible2
Installing ansible
Loaded plugins: extras_suggestions, langpacks, priorities, update-motd
Cleaning repos: amzn2-core amzn2extra-ansible2 amzn2extra-docker
12 metadata files removed
4 sqlite files removed
0 metadata files removed
Loaded plugins: extras_suggestions, langpacks, priorities, update-motd
amzn2-core
amzn2extra-ansible2
amzn2extra-docker
(1/7): amzn2-core/2/x86_64/group_gz
(2/7): amzn2-core/2/x86_64/updateinfo
(3/7): amzn2extra-docker/2/x86_64/primary_db
(4/7): amzn2extra-ansible2/2/x86_64/updateinfo
(5/7): amzn2extra-ansible2/2/x86_64/primary_db
```

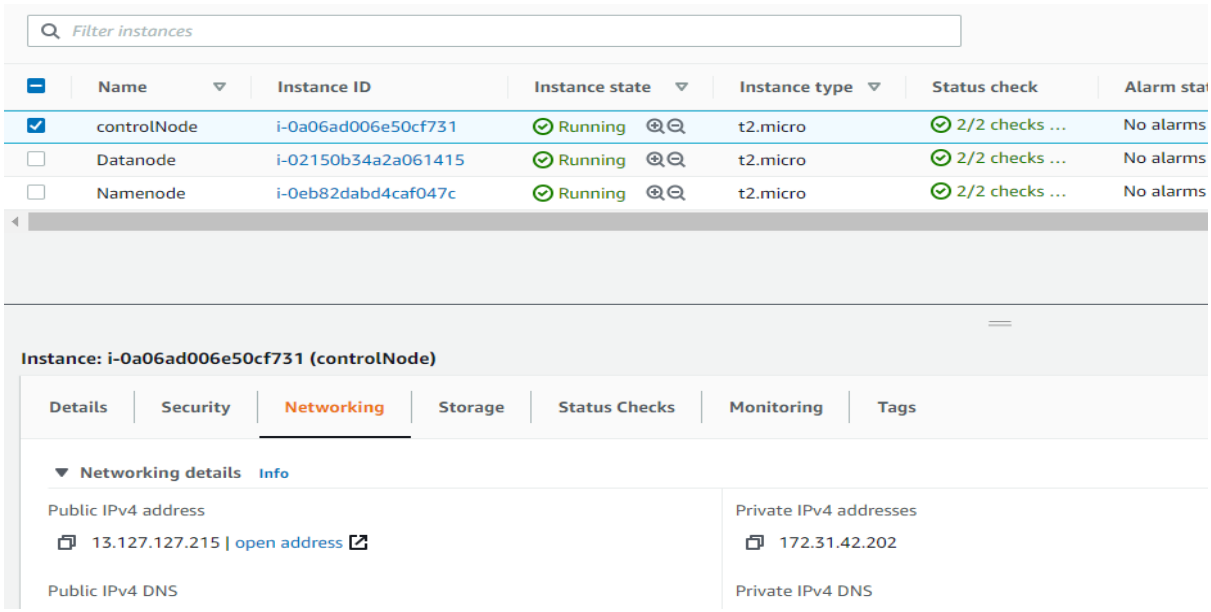
Checking the verison of ansible:

#ansible –version

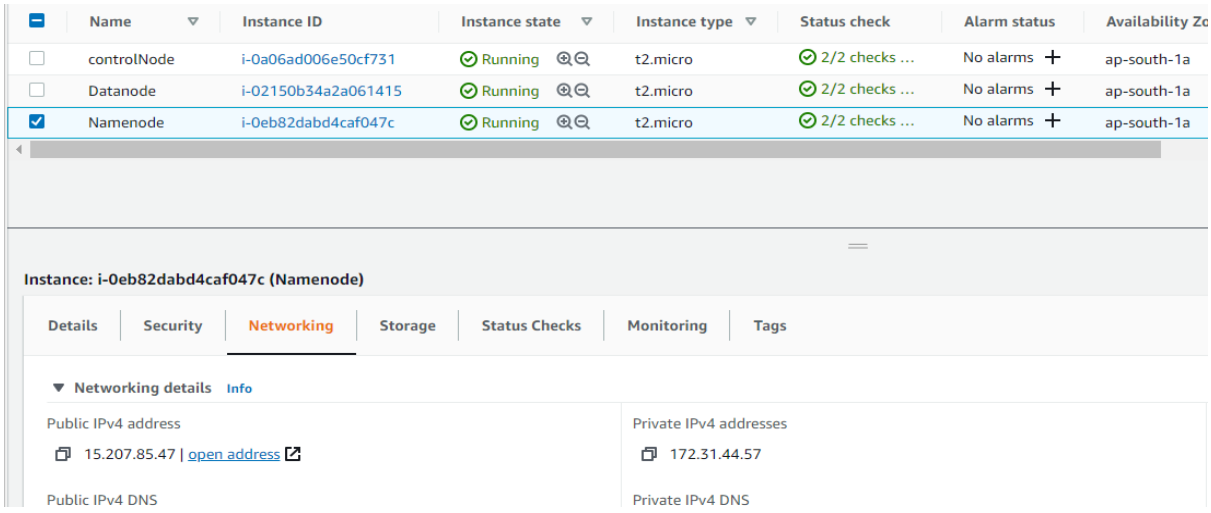
```
[root@ip-172-31-42-202 ~]# ansible --version
ansible 2.9.13
  config file = /etc/ansible/ansible.cfg
  configured module search path = [u'/root/.ansible/plugins/modules', u'/usr/sha
re/ansible/plugins/modules']
  ansible python module location = /usr/lib/python2.7/site-packages/ansible
  executable location = /usr/bin/ansible
  python version = 2.7.18 (default, Aug 27 2020, 21:22:52) [GCC 7.3.1 20180712 (
Red Hat 7.3.1-9)]
```

MY INSTANCES:

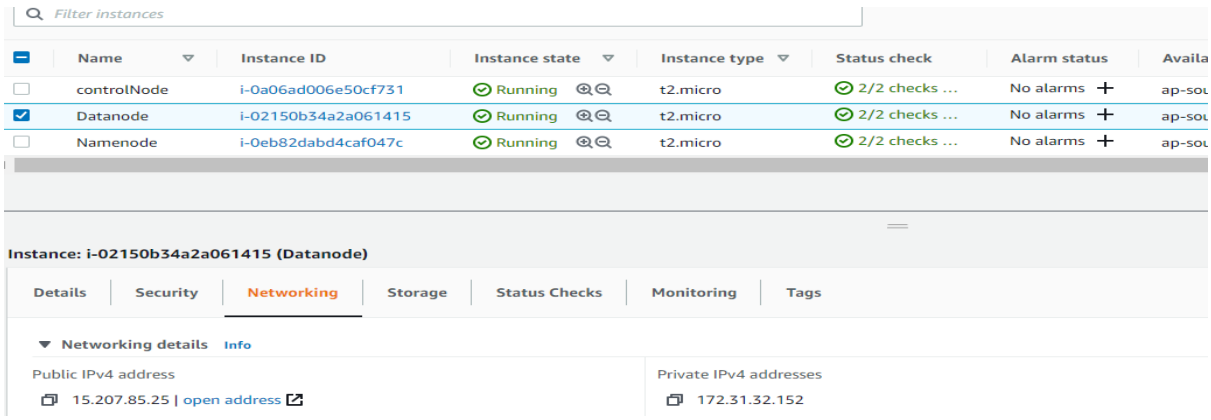
MY CONTROL NODE IP: 13.127.127.215



NAME NODE IP: 15.207.85.47



DATA NODE IP: 15.207.85.25



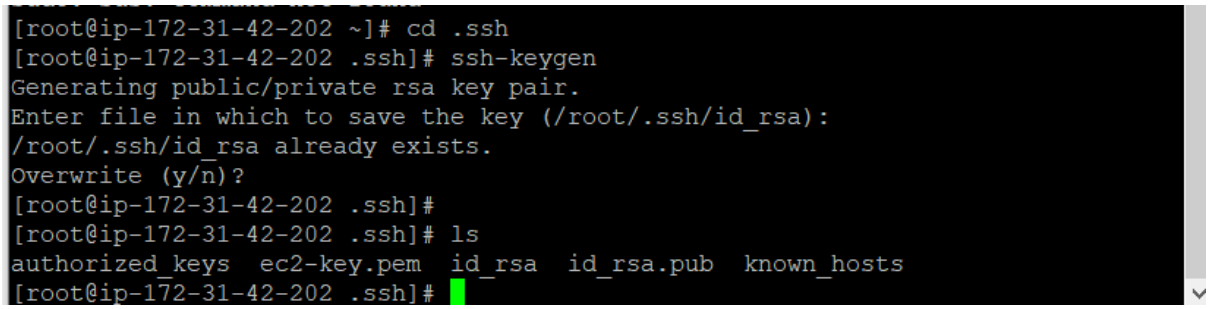
2. ENABLING THE SSH BETWEEN CONROL AND TARGET NODE.

Open control node ,type the following commands.

Cd .ssh

Ssh-keygen

Ls



Do the same with all taerget nodes.

Open the id_rsa.pub file from control node and copy the content into the authrization_keys of the target node .

This will enablet ssh between control and target/managed node.

3. WRITING THE CODE.

```
- hosts: all
  tasks:

  - name: "sending jdk software to targetNode"
    copy:
           dest: "/root"
           src: "/home/ec2-user/jdk-8u171-linux-x64.rpm"

  - name: "sending hadoop software to targetNode"
    copy:
           dest: "/root"
           src: "/home/ec2-user/hadoop-1.2.1-1.x86_64.rpm"

  - name: "installing Jdk software"
    command: "rpm -i jdk-8u171-linux-x64.rpm"
    ignore_errors: yes


  - name: "installing hadoop software"
    command: "rpm -i hadoop-1.2.1-1.x86_64.rpm --force"
    ignore_errors: yes


    #NAME NODE SETUP

- hosts: mynamenode

  tasks:
  - name: "deleting the previous directory"
    shell: "rm -rf /nn"
    ignore_errors: yes
  - name: "create a directory"
```

```
  - name: "create a directory"
    file:
           state: directory
           path: "/nn"

  - name: "hdfs-site file in namenode"
    template:
           dest: "/etc/hadoop/hdfs-site.xml"
           src: "nn_hdfs.xml"

  - name: " core-site file in namenode"
    template:
           dest: "/etc/hadoop/core-site.xml"
           src: "nn_core.xml"
  - name: "formatting the namemode"
    shell: "echo Y  |  hadoop  namenode  -format"


  - name: "stopping namenode service"
    command: "hadoop-daemon.sh stop  namenode"

  - name: "starting namenode service"
    command: "hadoop-daemon.sh start namenode"

     #DATA NODE
- hosts: mydatanode

  tasks:

  - name: "deleting the previous directory"
    shell: "rm -rf /dn"
    ignore_errors: yes

  - name: "craete a directory"
    file:
           state: directory
```

```
          path: "/dn"

  - name: "hdfs-site file in datanode"
    template:
          dest: "/etc/hadoop/hdfs-site.xml"
          src: "dn_hdfs.xml"

  - name: " core-sitee file in datanode"
    template:
          dest: "/etc/hadoop/core-site.xml"
          src: "dn_core.xml"
  - name: "stopping  datanode service"
    command: "hadoop-daemon.sh stop  datanode"
    ignore_errors: yes

  - name: "starting datanode service"
    command: "hadoop-daemon.sh start datanode"
```

Running

```
[root@ip-172-31-42-202 ws]# ansible-playbook preq.yml

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

TASK [Gathering Facts] ***********************************************
**************
[WARNING]: Platform linux on host 15.207.85.47 is using the discov
ered Python
interpreter at /usr/bin/python, but future installation of another
 Python
interpreter could change this. See https://docs.ansible.com/ansibl
e/2.9/referen
ce_appendices/interpreter_discovery.html for more information.
ok: [15.207.85.47]
[WARNING]: Platform linux on host 15.207.85.25 is using the discov
ered Python
interpreter at /usr/bin/python, but future installation of another
 Python
interpreter could change this. See https://docs.ansible.com/ansibl
e/2.9/referen
ce_appendices/interpreter_discovery.html for more information.
ok: [15.207.85.25]

TASK [sending jdk software to targetNode] ****************************
**************
ok: [15.207.85.25]
ok: [15.207.85.47]

TASK [sending hadoop software to targetNode] *************************
**************
ok: [15.207.85.25]
ok: [15.207.85.47]
```

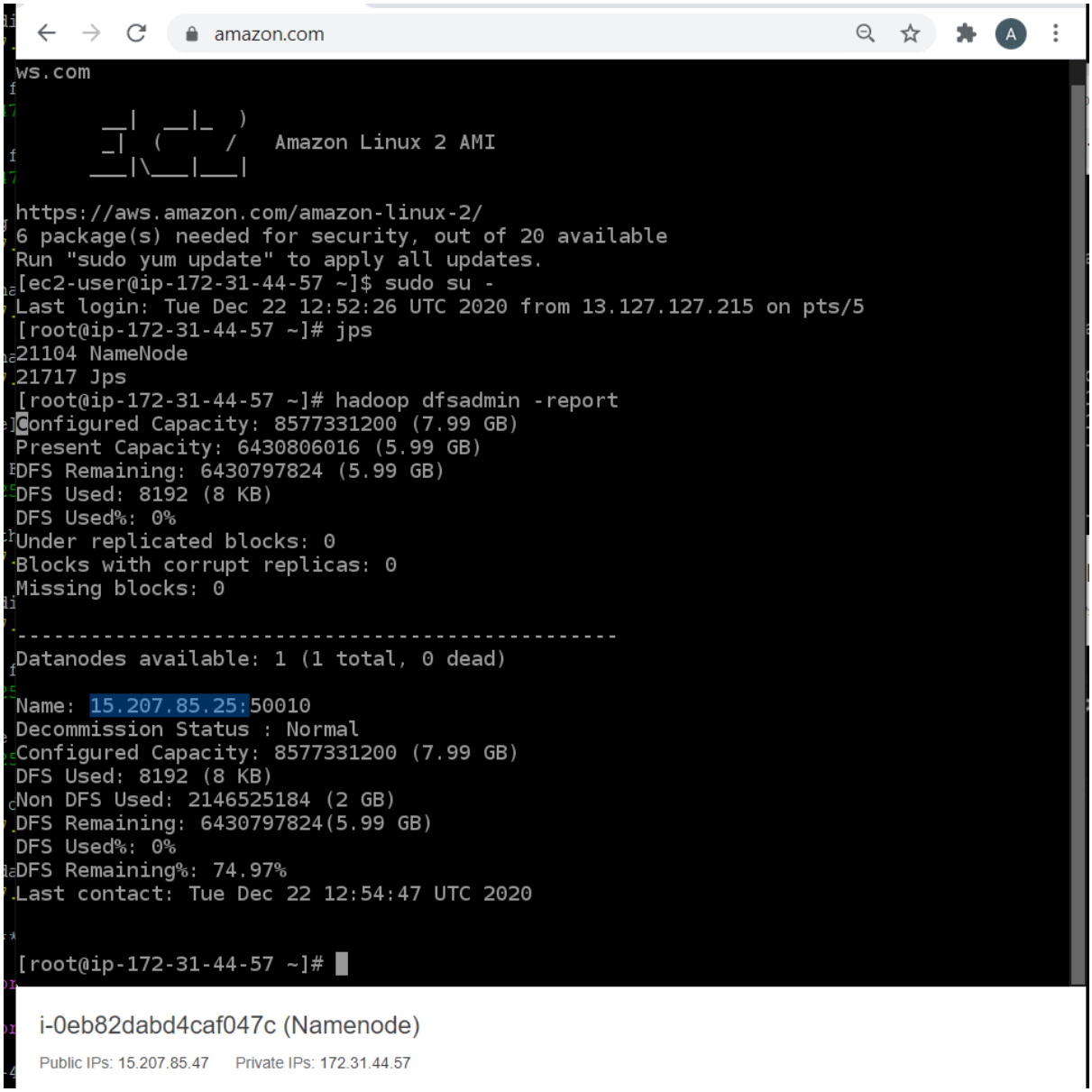
```
TASK [installing hadoop software] ************************************
**************
changed: [15.207.85.47]
changed: [15.207.85.25]

PLAY [mynamenode] ****************************************************
**************

TASK [Gathering Facts] ***********************************************
**************
ok: [15.207.85.47]

TASK [deleting the previous directory] ********************************
**************
[WARNING]: Consider using the file module with state=absent rather
 than running
'rm'.  If you need to use command because file is insufficient you
 can add
'warn: false' to this command task or set 'command_warnings=False'
 in
ansible.cfg to get rid of this message.
changed: [15.207.85.47]

TASK [create a directory] *********************************************
**************
changed: [15.207.85.47]

TASK [hdfs-site file in namenode] ************************************
```

```
TASK [core-site file in namenode] *********************************
**************
ok: [15.207.85.47]

TASK [formatting the namemode] ************************************
**************
changed: [15.207.85.47]

TASK [stopping namenode service] **********************************
**************
changed: [15.207.85.47]

TASK [starting namenode service] **********************************
**************
changed: [15.207.85.47]

PLAY [mydatanode] *************************************************
**************

TASK [Gathering Facts] ********************************************
**************
ok: [15.207.85.25]

TASK [deleting the previous directory] ****************************
**************
changed: [15.207.85.25]

TASK [craete a directory] *****************************************
**************
changed: [15.207.85.25]

TASK [hdfs-site file in datanode] *********************************
**************
```

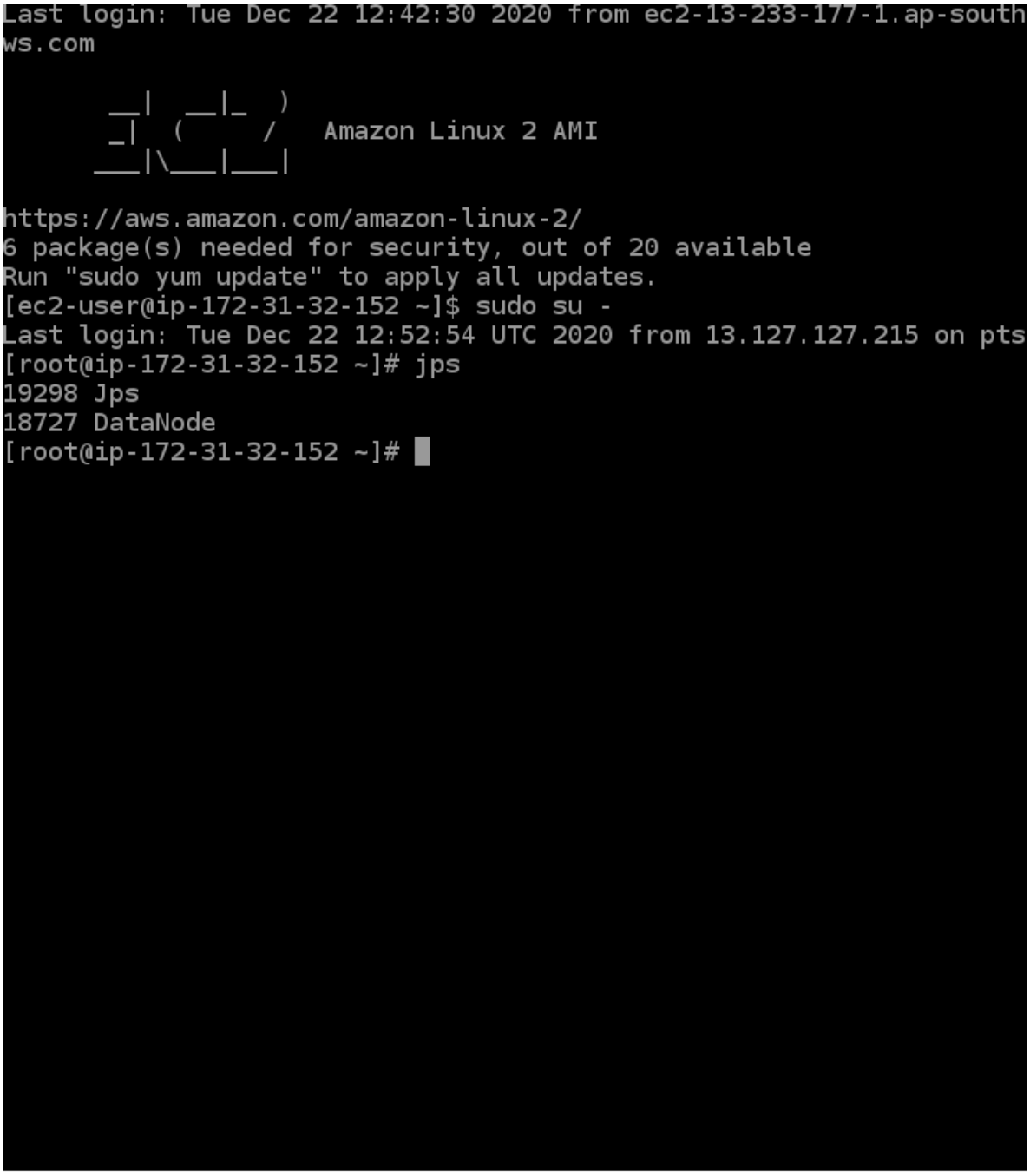
```
PLAY RECAP ********************************************************
**************
15.207.85.25               : ok=12   changed=6    unreachable=0
 failed=0    skipped=0    rescued=0    ignored=1
15.207.85.47               : ok=13   changed=7    unreachable=0
 failed=0    skipped=0    rescued=0    ignored=1
```

VERIFICATION

Name node status



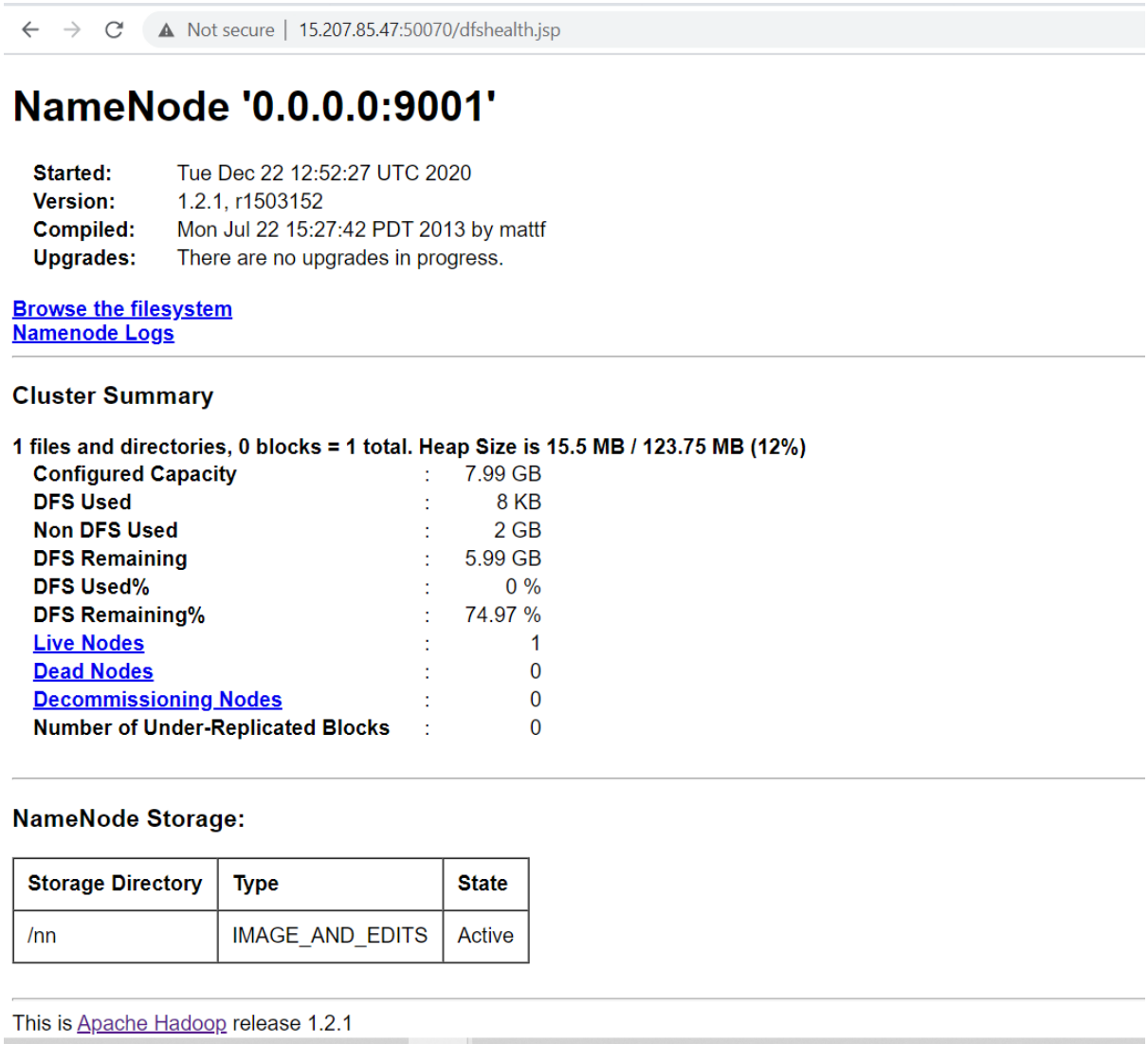
Data Node status



i-02150b34a2a061415 (Datanode)

Public IPs: 15.207.85.25 Private IPs: 172.31.32.152

AWS GUI Checking



ALL RUNNING WELL.....