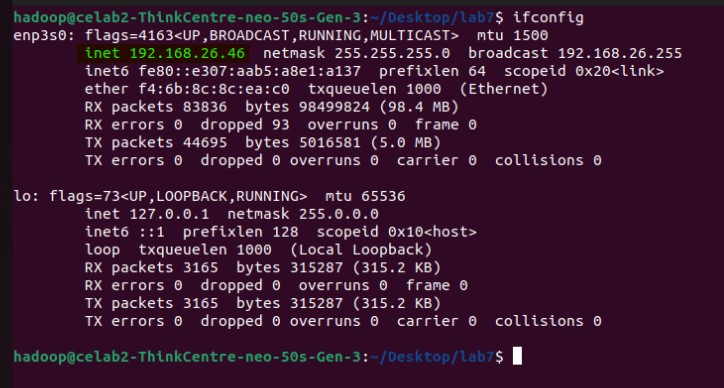
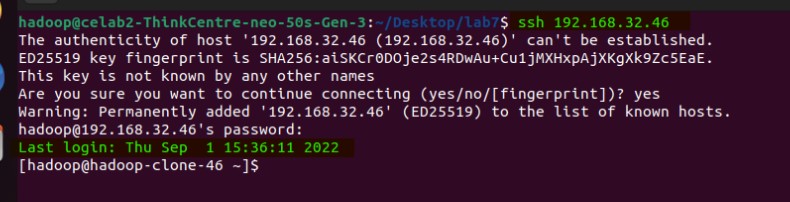
LAB - 7

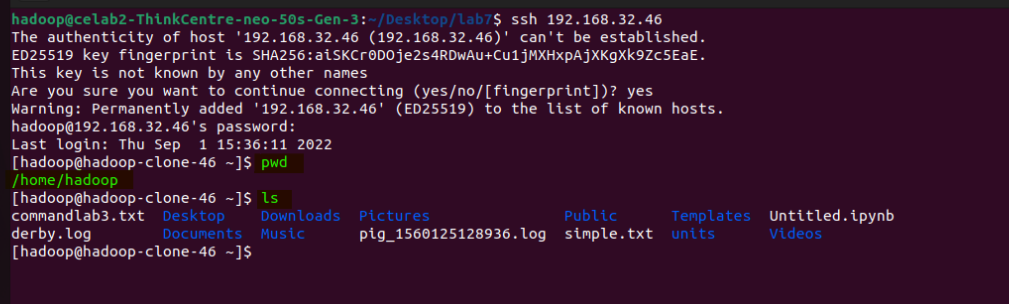
|  |  |
| --- | --- |
| Name | Keval D Gandevia |
| Roll Number | CE046 |
| ID | 19CEUEG017 |
| Subject | Big Data Analytics |

**Aim:** SSH, User Equivalence, Clusters / Cloud / Horizontal Scaling. Hadoop cluster set up.

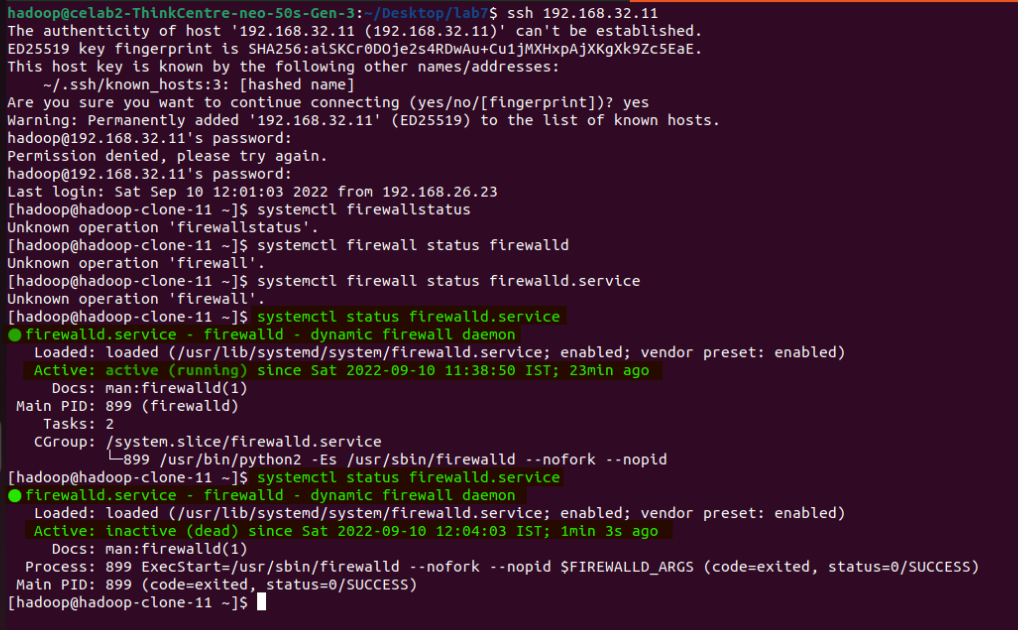
* **SSH:**
* SSH command is used for connecting the other machines through IP address.
* **Checking my IP address through ‘ifconfig’ command:**



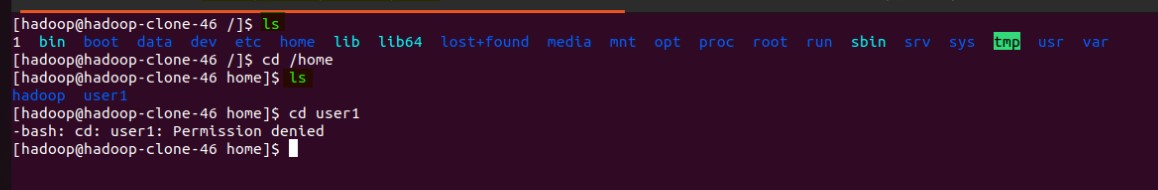
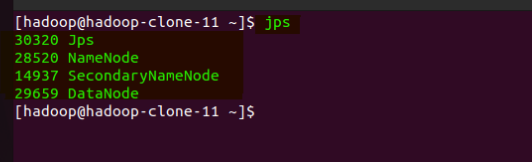
* **Connecting to another machine:**
* Now, we have access to that machine. Hence, we can run any command that we want.

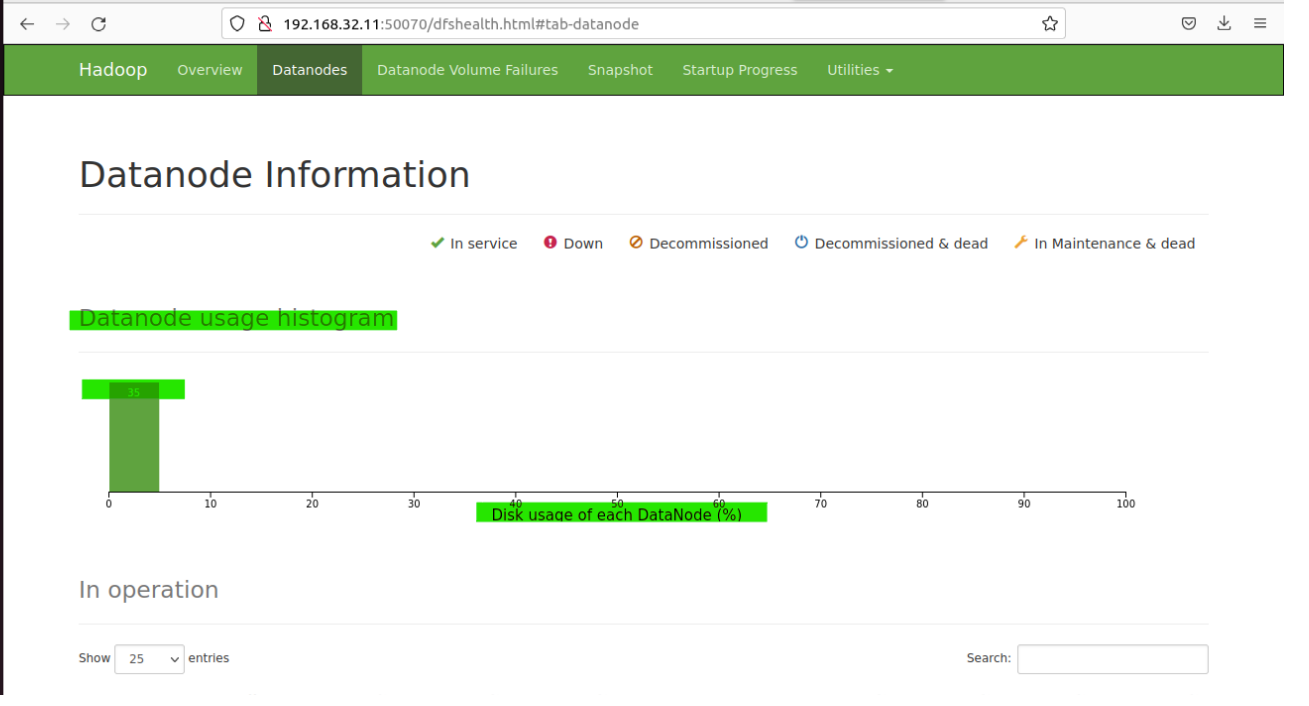


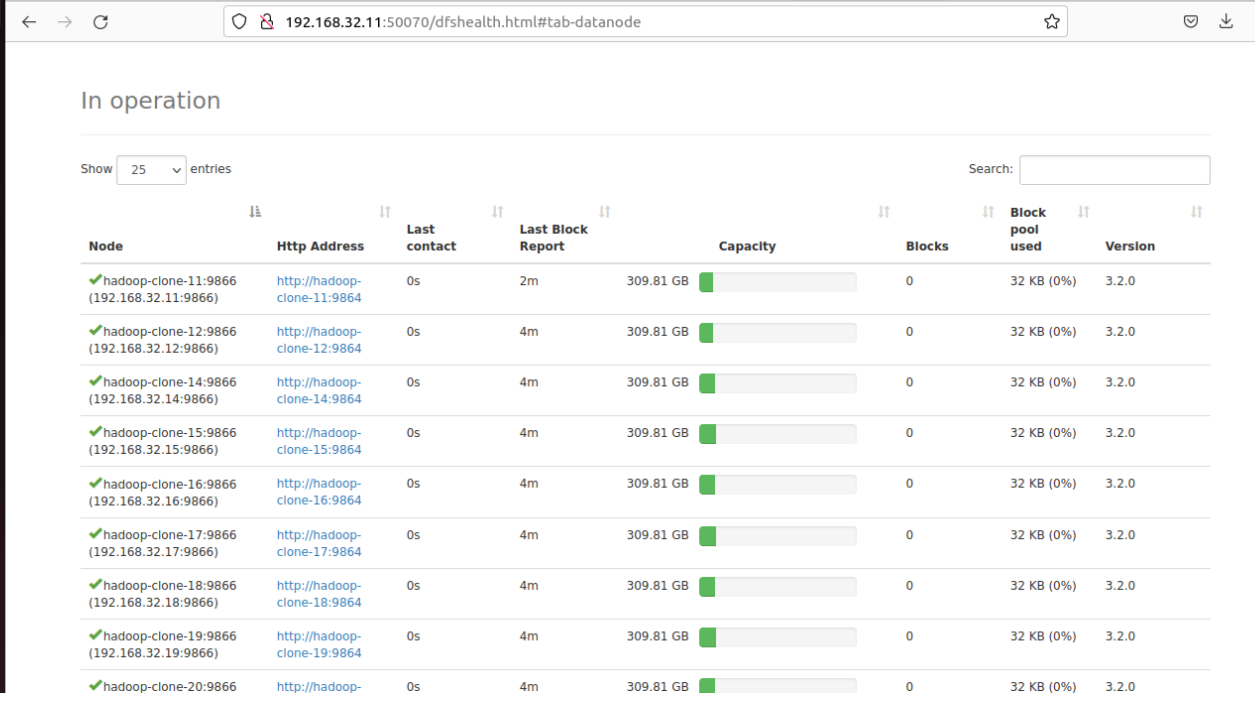
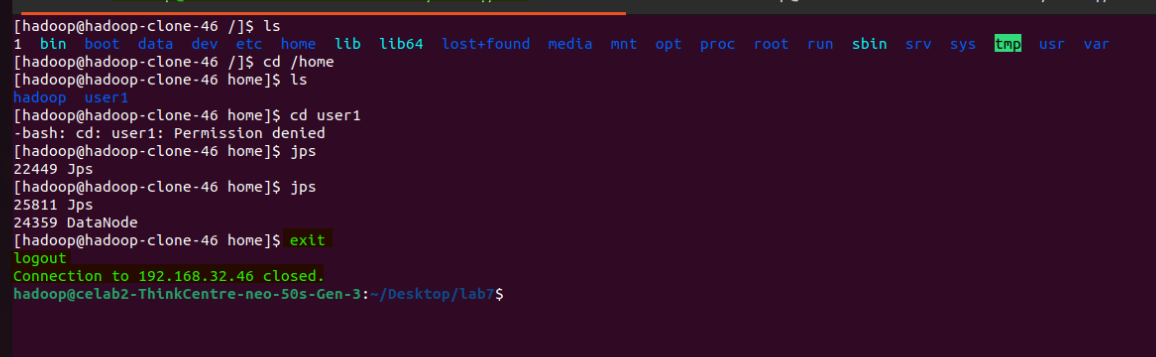
* **Checking the status of the firewall. And, it is active right now.**



* **Purpose of the Hadoop cluster:**
* Hadoop cluster is a collection of the various commodity hardware. This hardware components work together as a single unit.
* In the Hadoop cluster, there are lots of nodes that contain the masters and slaves.
* The NameNode and ResourceManager work as master node and data node, NodeManager works as a slave.
* It is designed to store and analyze the large amount of structured and unstructured data in distributed computing environment.
* **Configuring the Hadoop cluster:**







* **Application of Hadoop cluster: (Sports)**
* On such scenario where cluster can be used is sports.
* When any league auction comes, then each team have to decide which player they can have in their team or which player will perform best for their team.
* When there is little performance data available to train our model on, we can have advantage of clustering. In such case, we can find the similar players using some characteristics.