

## Assignment No 6 Hadoop–

Q) Create single node hadoop cluster having the machine name as your name and upload the project data in HDFS. Upload all the screenshots that would describe the steps taken in completing this assignment.

Complete Process of Hadoop single node cluster Setup

1) First install vmware in local system.

2) After installation of vmware step up any operating system in vmware workstation(for ex: linux , microsoft).

3) Java Setup in Linux virtual operating system:

Steps for Java Setup:

Step 1: start a virtual operating system in VMware.

Step 2: start ip address of virtual operating system.

commands for start ip address:

```
nmcli conn show
```

```
nmcli conn up device_name
```

for ip check:

```
ip addr show
```

Step 3: Transfer Java tar file to virtual operating system using any file transfer tool.

Step 4: Extract the tar file using below command.

```
tar -xvf file_name.
```

Step 5: Move and rename the extracted file /use/local location using the below command.

```
mv file_name /use/local/new_file_name
```

Step 6: create an environment variable for Java in /etc/profile location using command vi /etc/profile.

Step 7: Refresh the desktop with the help of command source /etc/profile.

Step 8: Type Jps and Java version commands for checking setup of Java.

Step 9: Go to /etc/hostname location and change machine host name.

Step 10: Then go to /etc/hosts location and change ip address and machine name.

#### 4) Process of Hadoop Setup in Linux:

This process is start after complete setup of java in linux.

Steps for setup of Hadoop:

Step 1: Transfer hadoop tar file to virtual operating system using any file transfer tool.

Step 2: Extract tar file using `tar -xvf file_name` command

Step 3: Move and rename the extracted file to `/usr/local` location using command `mv file_name /usr/local/new_file_name`.

Step 4: Create an environment variable for hadoop setup in Linux in `/etc/profile` directory.

Step 5: Set the path of Java file in the hadoop environment file. Location of the hadoop environment file: `/usr/local/hadoop/etc/hadoop`

File name: `hadoop-env.sh`

Step 6: rewrite XML files in hadoop.

Files location: `usr/local/hadoop/etc/hadoop`

Files: `core-site.xml`, `hdfs-site.xml`, `mapred-site.xml`, `yarn-site.xml`.

Step 7: Refresh the desktop using the source command.

`source /etc/profile`

Step 8: Before setup of hadoop namenode change machine name in `/usr/local/hadoop/etc/hadoop/slaves` file location.

Step 9: Format the name node of hadoop using `hdfs namenode -format` command.

Step 10: Start all Storage and Processing tools of hadoop using the below commands.

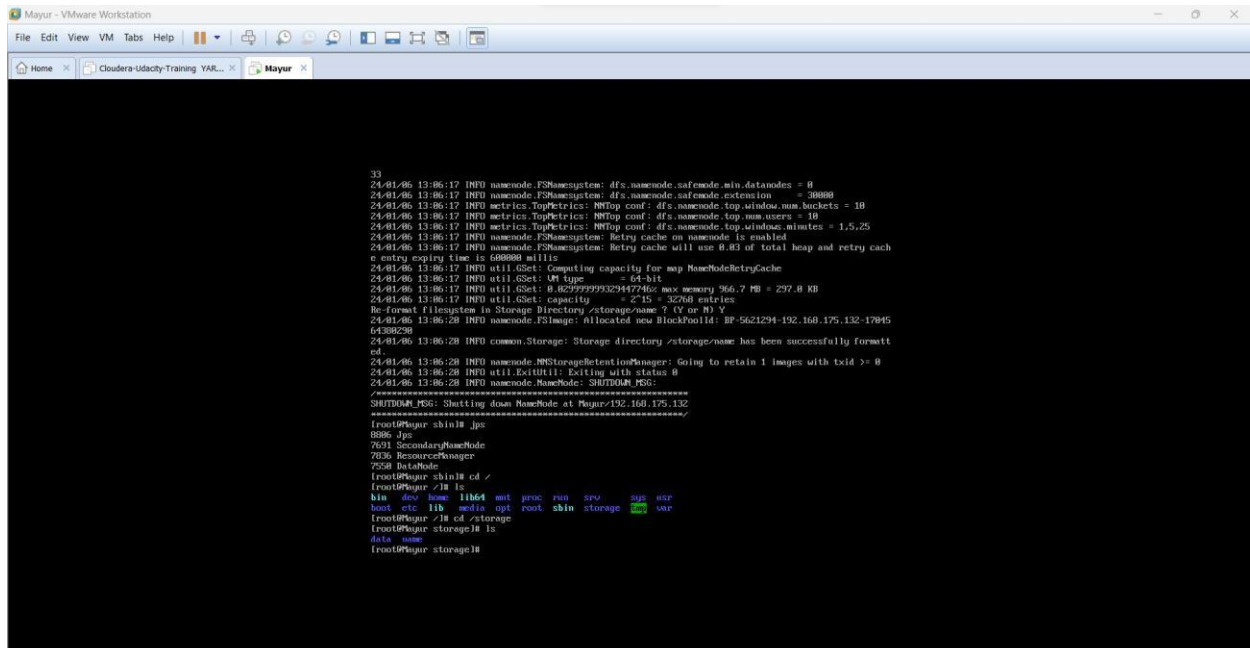
commands:

`sh start-dfs.sh`

`sh start-yarn.sh`

Step 11: Check all tools of Hadoop are working or not using `Jps` command. The tools are working means hadoop setup in linux is completed. In that point hadoop single node cluster setup is completed.

## Outcomes



To direct input to this VM, click inside or press Ctrl+G.

