

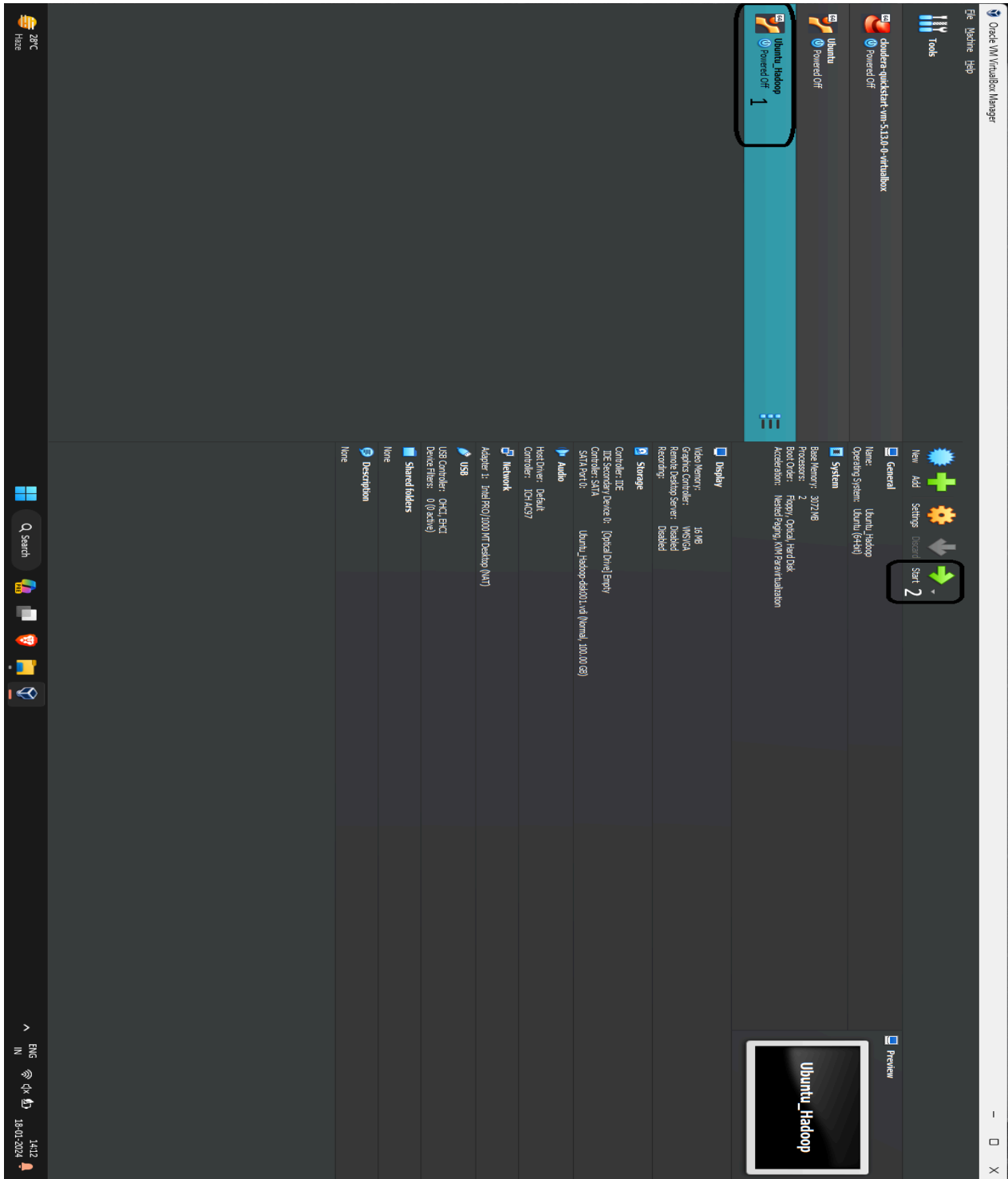
Follow the step to perform hadoop in ubuntu

1. Open the **oracle VM VirtualBox**

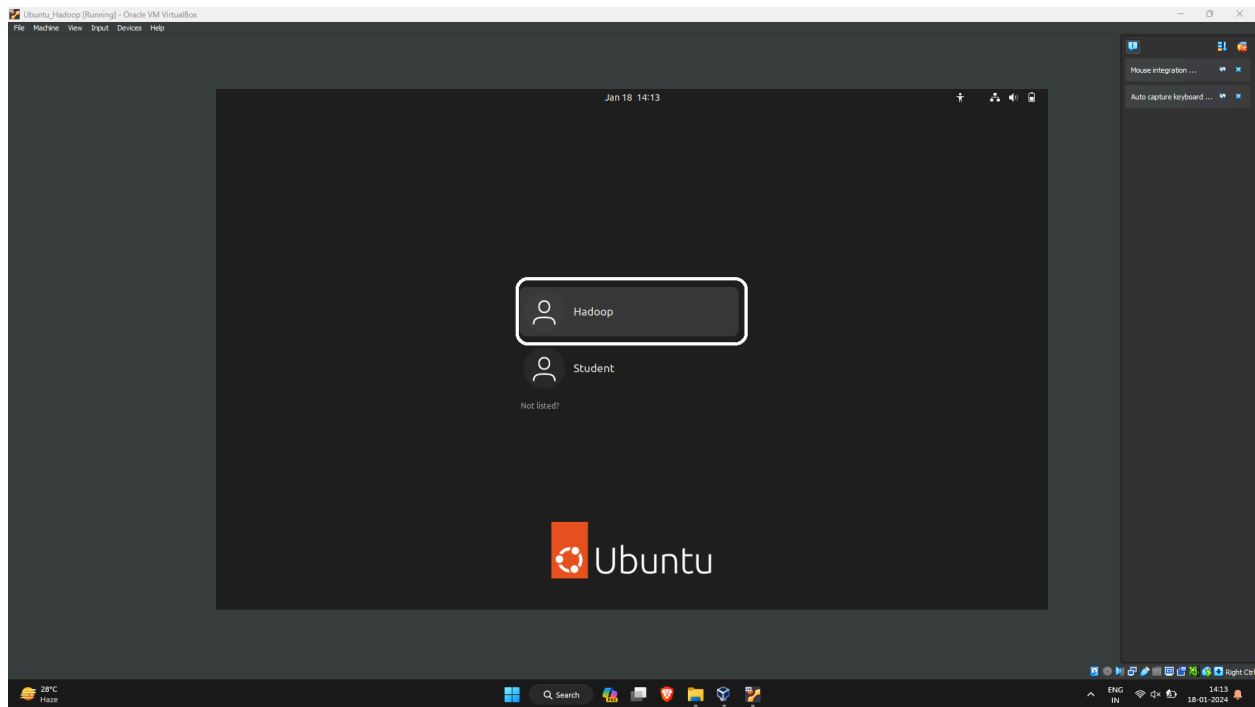


2. When **oracle VM Virtual Box** is opened the

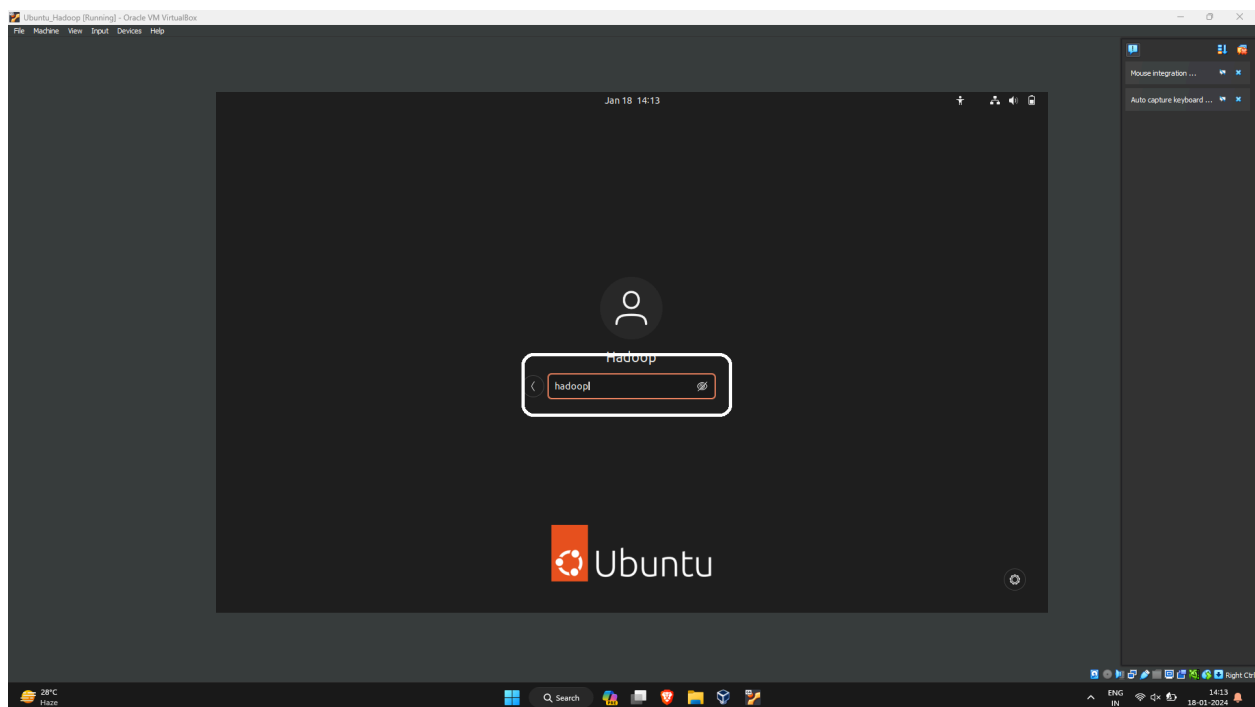
1. Click on “**Ubuntu_Hadoop**”
2. Click on “**Arrow ->** ” button”



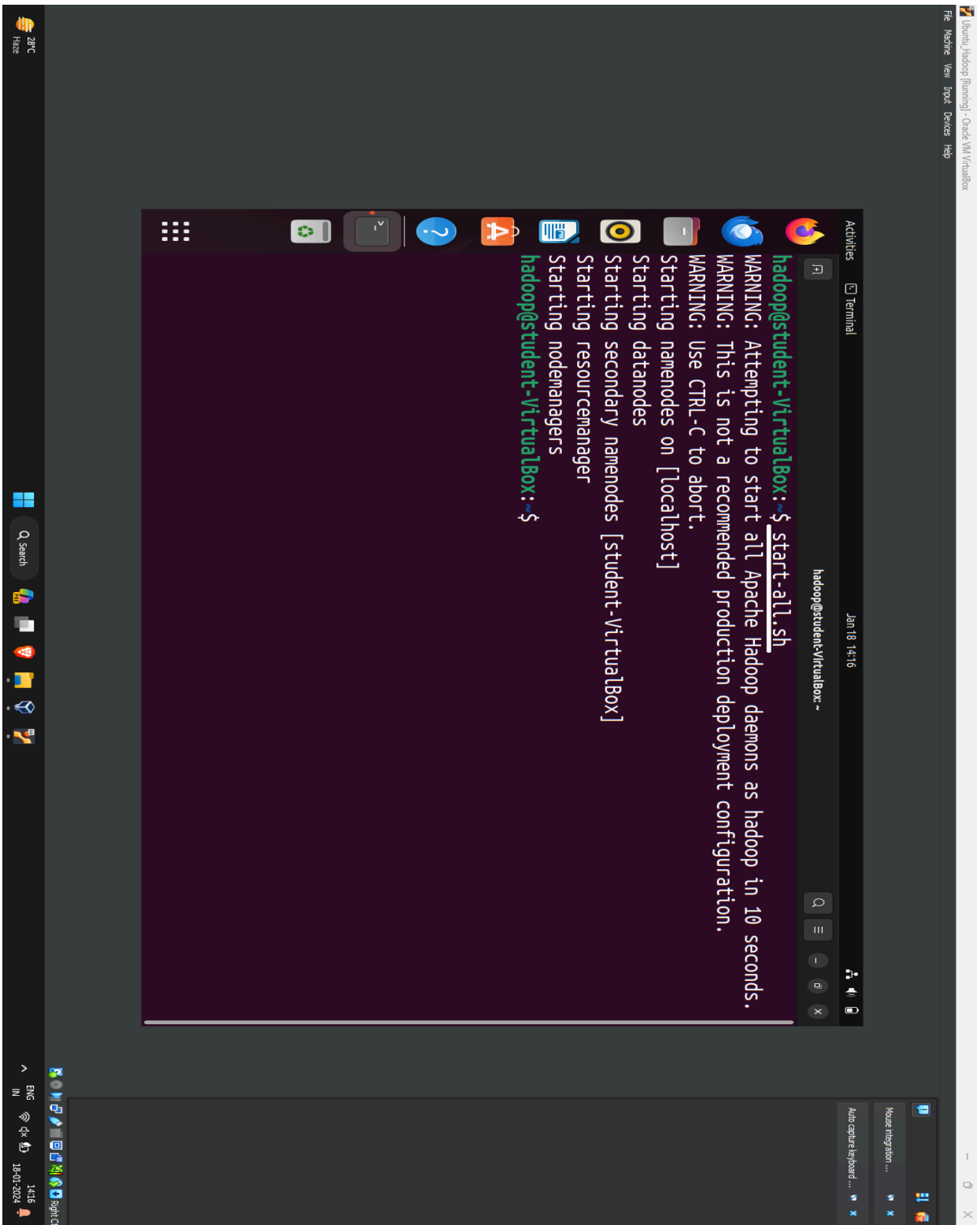
3. After opening the Ubuntu then Click on “**hadoop**” account



4. Select hadoop account and password is “**hadoop**”



5. Open **Terminal** and type **start-all.sh** to start all the hadoop related services

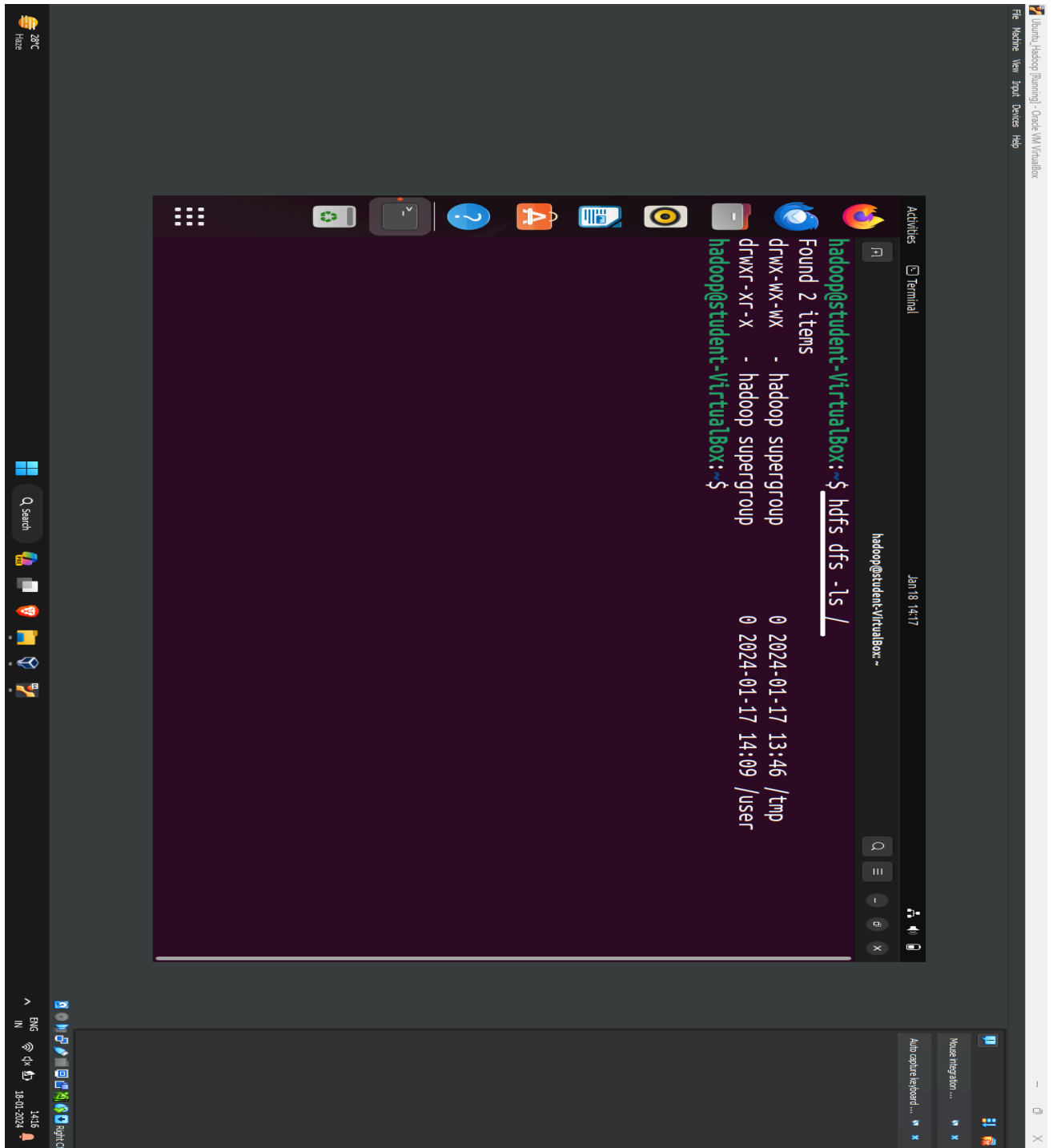


The screenshot shows a terminal window titled "Terminal" with the prompt "hadoop@student-VirtualBox: ~". The user has entered the command "start-all.sh". The terminal output shows several warning messages and the successful start of various Hadoop services. The background of the terminal window is a dark purple color with a grid of icons. The top of the window shows the Ubuntu logo, the temperature (28°C), and the date and time (Jan 18 14:16). The bottom of the window shows the system tray with icons for network, sound, and power, along with the date and time (14:16, 18-01-2024).

```
hadoop@student-VirtualBox:~$ start-all.sh
WARNING: Attempting to start all Apache Hadoop daemons as hadoop in 10 seconds.
WARNING: This is not a recommended production deployment configuration.
WARNING: Use CTRL-C to abort.
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [student-VirtualBox]
Starting resourcemanager
Starting nodemanagers
hadoop@student-VirtualBox:~$
```

6. To see what files are present in the hadoop type this command

hadoop@student-VirtualBox \$ hdfs dfs -ls /



7. **Task** : Create file on local and upload to hadoop file system

1. Creating file on local ubuntu machine for command is

hadoop@student-VirtualBox \$ touch file.txt

2. Check file is created locally or not for command is

hadoop@student-VirtualBox \$ ls

3. Upload file to hadoop file system from local ubuntu machine for command is

**hadoop@student-VirtualBox \$ hdfs dfs -put
'/home/hadoop/file.txt' '/'**

4. Check file is uploaded to hadoop file system or not for command is

hadoop@student-VirtualBox \$ hdfs dfs -ls /

```
Activities [Terminal] Jan 18 14:18
hadoop@student-VirtualBox:~$ touch file.txt
hadoop@student-VirtualBox:~$ ls
  apache-hive-3.1.2-bin
  apache-hive-3.1.2-bin.tar.gz
  derby.log
  Desktop
  dfsdata
  Documents
  Downloads
  file.txt
  hadoop-3.3.6
  hadoop-3.3.6.tar.gz
  metastore_db
  Music
hadoop@student-VirtualBox:~$ hdfs dfs -put '/home/hadoop/file.txt' '/'
hadoop@student-VirtualBox:~$ hdfs dfs -ls /
Found 3 items
-rw-r--r-- 1 hadoop supergroup 0 2024-01-18 14:18 /file.txt
drwx-wx-wx - hadoop supergroup 0 2024-01-17 13:46 /tmp
drwxr-xr-x - hadoop supergroup 0 2024-01-17 14:09 /user
hadoop@student-VirtualBox:~$
```

8.**Task** : Create file on hadoop file system and upload to local

1. Creating file on hadoop file system machine for command is

hadoop@student-VirtualBox \$ hdfs dfs -touchz /file1.txt

2.Check file is created on hadoop file system or not for command is

hadoop@student-VirtualBox \$ hdfs dfs -ls /

3.Upload file local ubuntu machine to from hadoop file system for command is

hadoop@student-VirtualBox \$ hdfs dfs -get '/file1.txt'
'/home/hadoop/'

4.Check file is uploaded to local ubuntu machine or not for command is

hadoop@student-VirtualBox \$ ls

Activities Terminal

Jan 18 14:20

hadoop@student-VirtualBox: ~

hadoop@student-VirtualBox:~\$ hdfs dfs -touchz /file1.txt

hadoop@student-VirtualBox:~\$ hdfs dfs -ls /

Found 4 items

-rw-r--r-- 1 hadoop supergroup 0 2024-01-18 14:18 /file.txt

-rw-r--r-- 1 hadoop supergroup 0 2024-01-18 14:19 /file1.txt

drwx-wx-wx - hadoop supergroup 0 2024-01-17 13:46 /tmp

drwxr-xr-x - hadoop supergroup 0 2024-01-17 14:09 /user

hadoop@student-VirtualBox:~\$ hdfs dfs -get '/file1.txt' '/home/hadoop/'

hadoop@student-VirtualBox:~\$ ls

apache-hive-3.1.2-bin file.txt pig_1705481238301.log

apache-hive-3.1.2-bin.tar.gz hadoop-3.3.6 hadoop-3.3.6 Public

derby.log hadoop-3.3.6.tar.gz snap

Desktop metastore_db spark-3.5.0-bin-hadoop3

dfsdata Music spark-3.5.0-bin-hadoop3.tgz

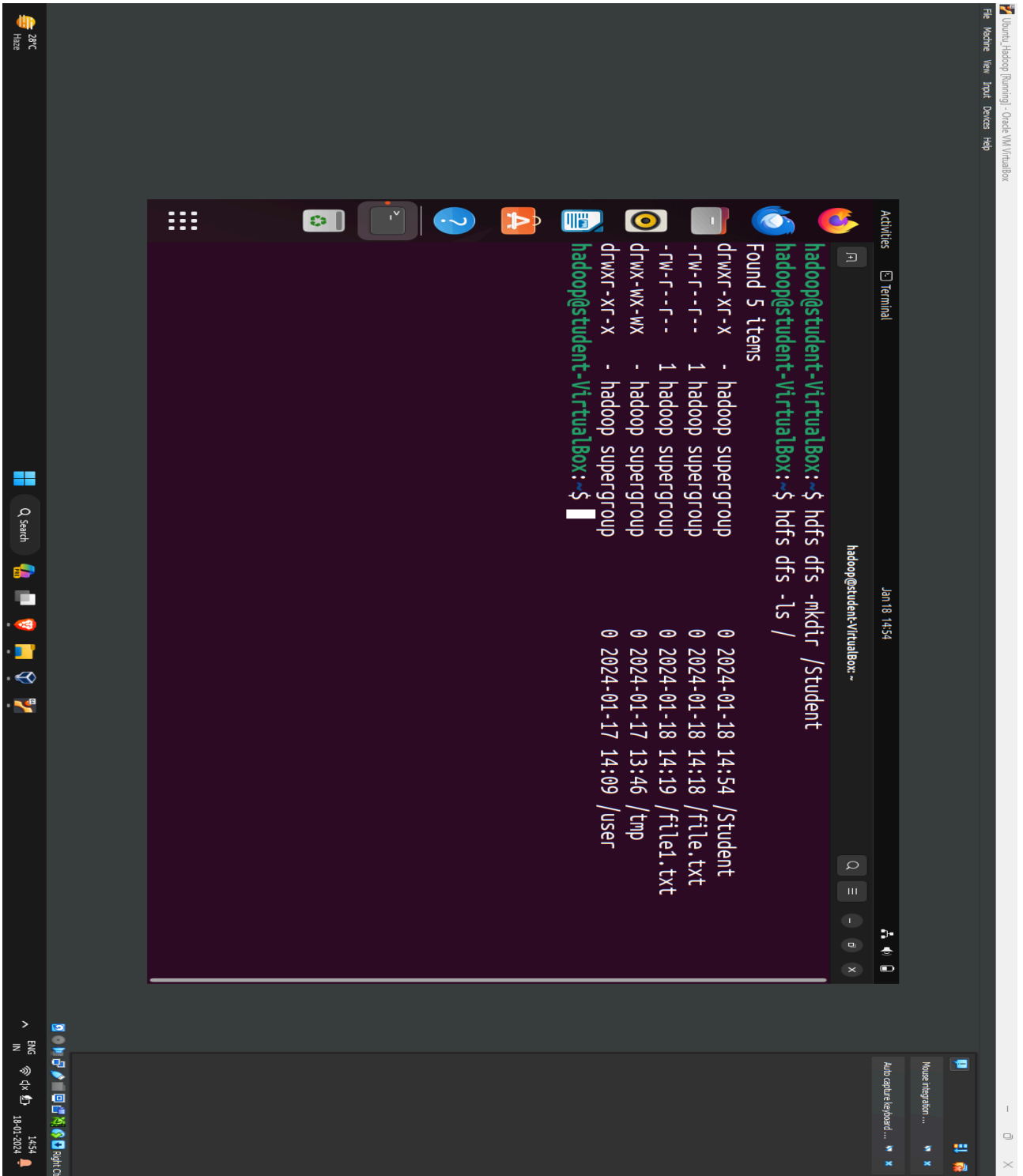
Documents Pictures Templates

Downloads pig-0.17.0 pig-0.17.0.tar.gz tmpdata

file1.txt Videos

hadoop@student-VirtualBox:~\$

9. Creating Directory on the hadoop file system named as **Student** and uploading **file.txt** from local ubuntu machine to Student folder which is created in hadoop file system.



The screenshot shows a terminal window within a virtual machine environment. The terminal displays the following commands and output:

```
hadoop@student-VirtualBox:~$ hdfs dfs -mkdir /Student
hadoop@student-VirtualBox:~$ hdfs dfs -ls /
Found 5 items
drwxr-xr-x - hadoop supergroup          0 2024-01-18 14:54 /Student
-rw-r--r-- 1 hadoop supergroup          0 2024-01-18 14:18 /file.txt
-rw-r--r-- 1 hadoop supergroup          0 2024-01-18 14:19 /file1.txt
drwx-wx-wx - hadoop supergroup          0 2024-01-17 13:46 /tmp
drwxr-xr-x - hadoop supergroup          0 2024-01-17 14:09 /user
hadoop@student-VirtualBox:~$
```

The terminal window is titled "Terminal" and shows the user "hadoop@student-VirtualBox". The system clock indicates "Jan 18 14:54". The virtual machine's desktop environment is visible in the background, including a search bar and various application icons.

```
Activities Terminal Jan 18 14:56  
hadoop@student-VirtualBox: ~  
hadoop@student-VirtualBox:~$ hdfs dfs -mkdir /Student  
hadoop@student-VirtualBox:~$ hdfs dfs -ls /  
Found 5 items  
drwxr-xr-x - hadoop supergroup 0 2024-01-18 14:54 /Student  
-rw-r--r-- 1 hadoop supergroup 0 2024-01-18 14:18 /file.txt  
-rw-r--r-- 1 hadoop supergroup 0 2024-01-18 14:19 /file1.txt  
drwx-wx-wx - hadoop supergroup 0 2024-01-17 13:46 /tmp  
drwxr-xr-x - hadoop supergroup 0 2024-01-17 14:09 /user  
hadoop@student-VirtualBox:~$ hdfs dfs -put '/home/hadoop/file.txt' '/Student/'  
hadoop@student-VirtualBox:~$ hdfs dfs -ls /Student/  
Found 1 items  
-rw-r--r-- 1 hadoop supergroup 0 2024-01-18 14:55 /Student/file.txt  
hadoop@student-VirtualBox:~$
```