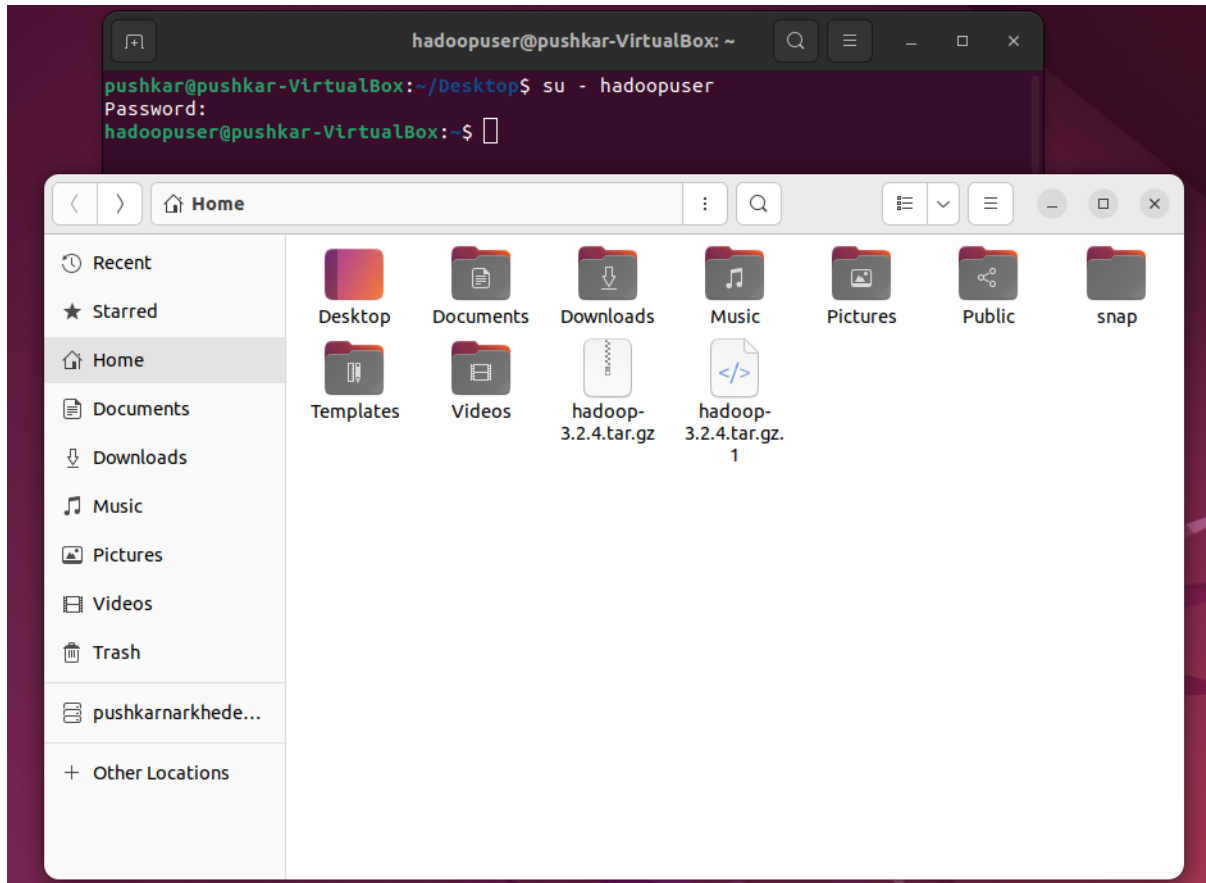


## ASSIGNMNT 02

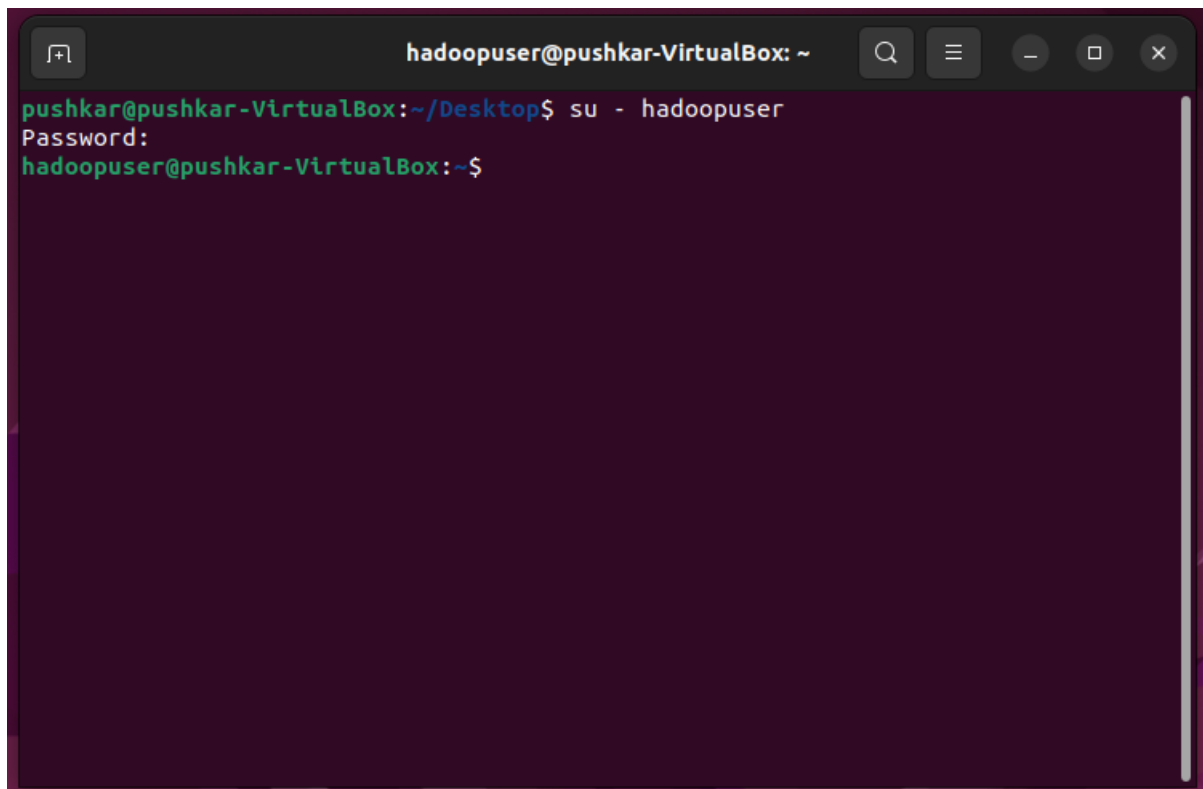
### File Management using Hadoop

#### 1. Files stored in Local storage



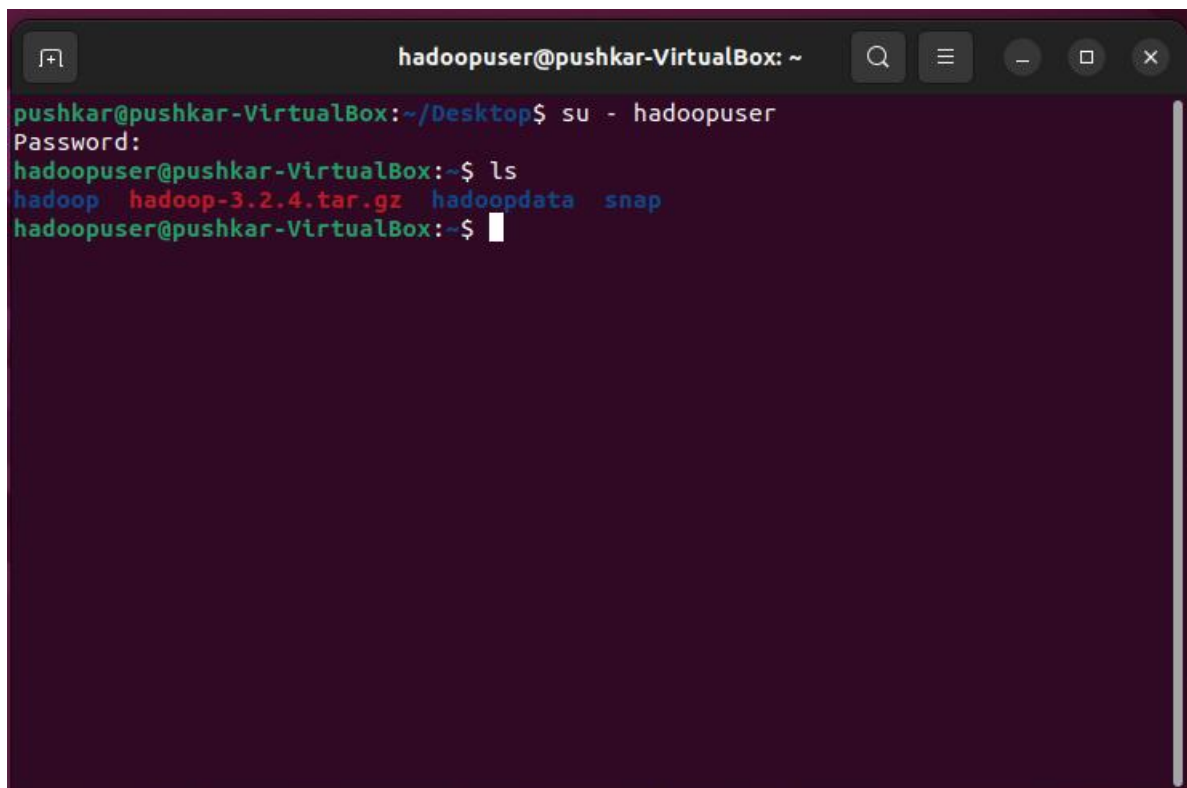
**Cluster type** – Single Node Cluster → means all files are saved into the single Hadoop cluster storage. To work with Hadoop file management commands we have to store some files from local storage to Hadoop cluster and perform operation

### Login into Hadoop

A terminal window titled 'hadoopuser@pushkar-VirtualBox: ~' with standard window controls. The prompt is 'pushkar@pushkar-VirtualBox:~/Desktop\$'. The user enters 'su - hadoopuser', followed by a 'Password:' prompt. After an invisible password entry, the prompt changes to 'hadoopuser@pushkar-VirtualBox:~\$'.

```
hadoopuser@pushkar-VirtualBox: ~
pushkar@pushkar-VirtualBox:~/Desktop$ su - hadoopuser
Password:
hadoopuser@pushkar-VirtualBox:~$
```

ls command ( Linux Local Storage Command )

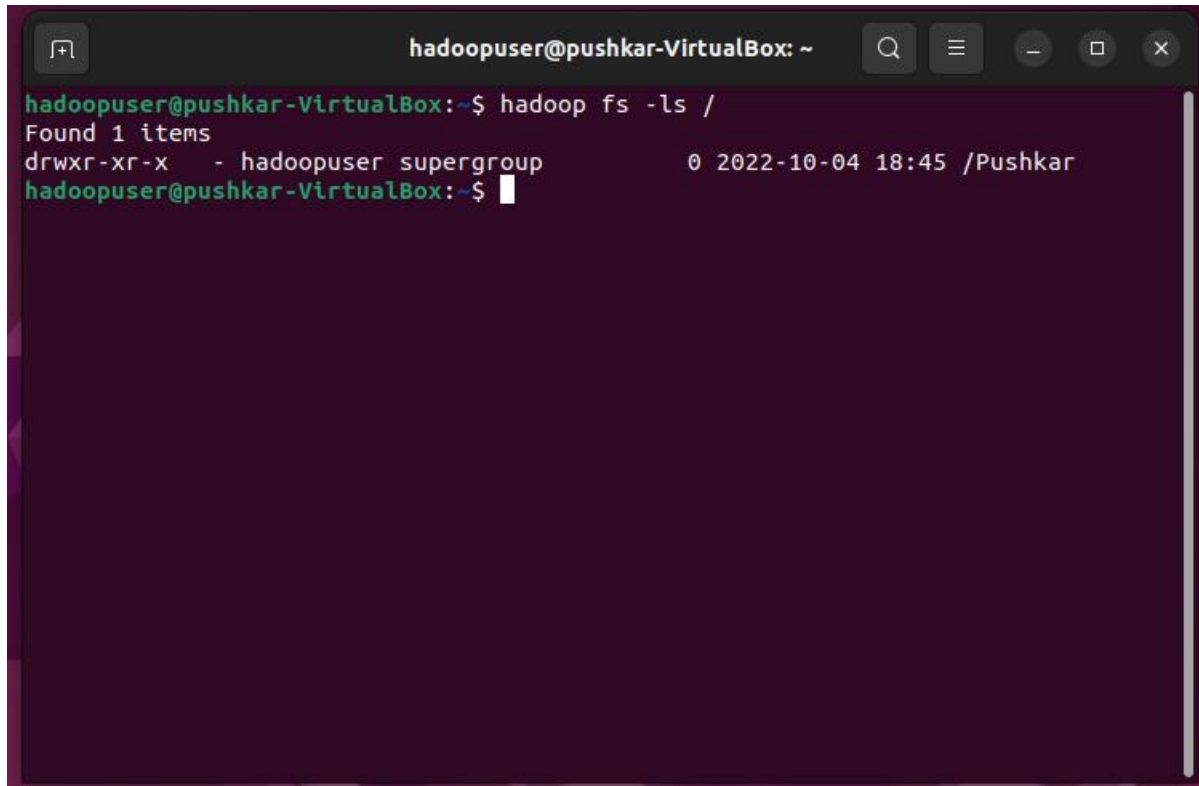
A terminal window titled 'hadoopuser@pushkar-VirtualBox: ~' with standard window controls. The prompt is 'pushkar@pushkar-VirtualBox:~/Desktop\$'. The user enters 'su - hadoopuser', followed by a 'Password:' prompt. After an invisible password entry, the prompt changes to 'hadoopuser@pushkar-VirtualBox:~\$'. The user then enters 'ls', and the output 'hadoop hadoop-3.2.4.tar.gz hadoopdata snap' is displayed. The prompt returns to 'hadoopuser@pushkar-VirtualBox:~\$' with a cursor.

```
hadoopuser@pushkar-VirtualBox: ~
pushkar@pushkar-VirtualBox:~/Desktop$ su - hadoopuser
Password:
hadoopuser@pushkar-VirtualBox:~$ ls
hadoop hadoop-3.2.4.tar.gz hadoopdata snap
hadoopuser@pushkar-VirtualBox:~$
```

### -ls List Directory

This command is used to list all the directory and files present in the cluster

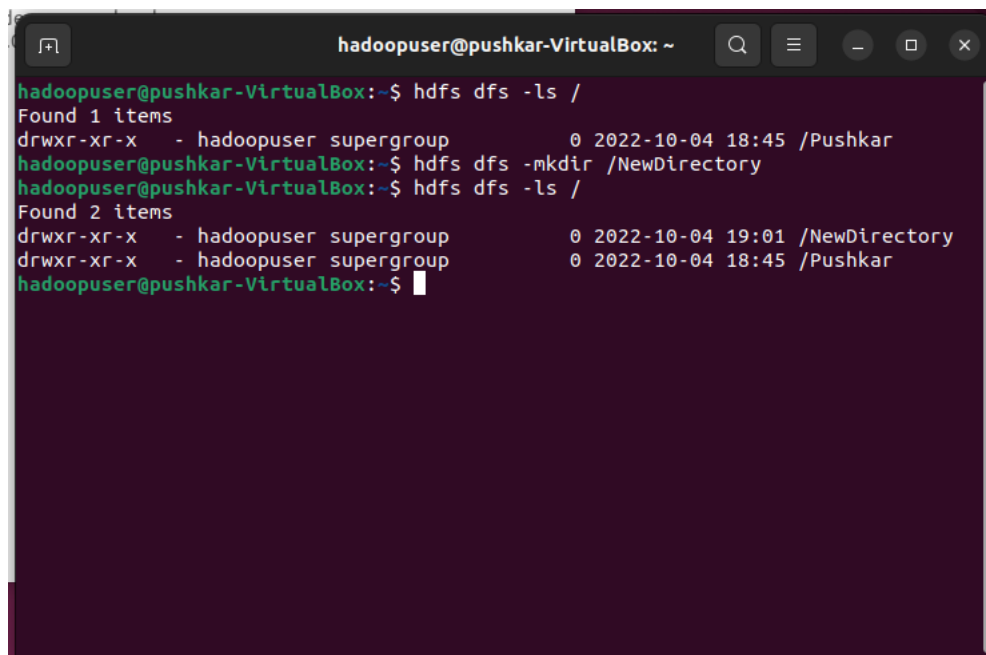
It also specifies the permission given to each directory and file

A terminal window titled 'hadoopuser@pushkar-VirtualBox: ~' showing the command 'hadoop fs -ls /' and its output. The output lists a single directory named '/Pushkar' with permissions 'drwxr-xr-x', owner 'hadoopuser', group 'supergroup', size '0', and timestamp '2022-10-04 18:45'.

```
hadoopuser@pushkar-VirtualBox:~$ hadoop fs -ls /
Found 1 items
drwxr-xr-x  - hadoopuser supergroup          0 2022-10-04 18:45 /Pushkar
hadoopuser@pushkar-VirtualBox:~$
```

### - mkdir ( Make Directory )

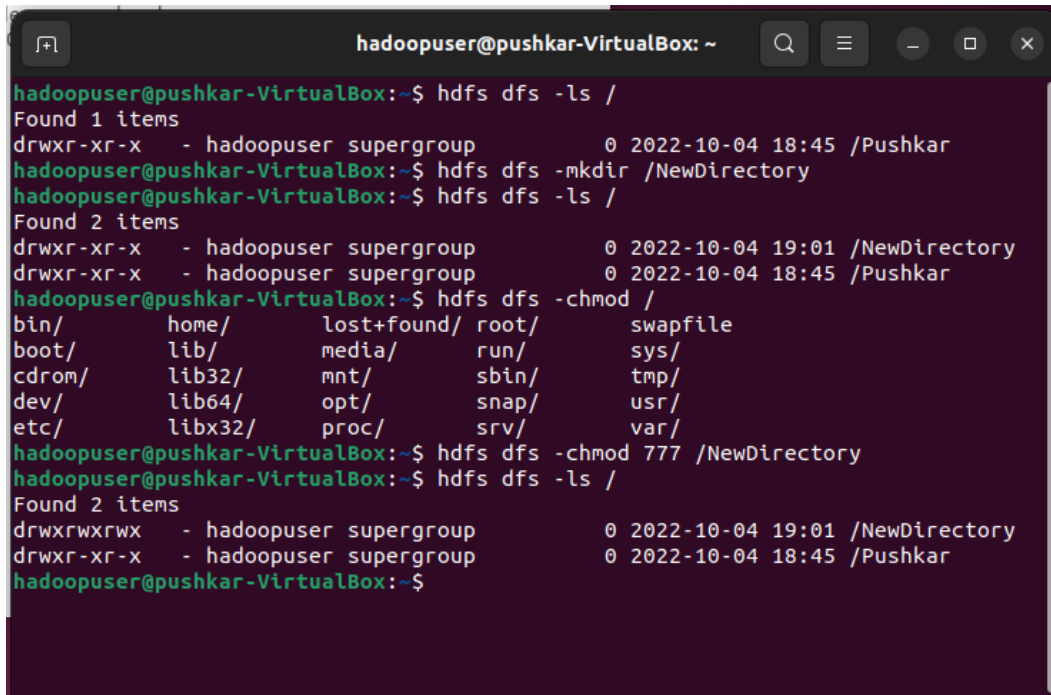
This command is used to make a new directory into the selected cluster

A terminal window titled 'hadoopuser@pushkar-VirtualBox: ~' showing two commands and their outputs. First, 'hdfs dfs -ls /' shows the existing '/Pushkar' directory. Then, 'hdfs dfs -mkdir /NewDirectory' is executed, and a second 'hdfs dfs -ls /' command shows two directories: '/NewDirectory' (created at 19:01) and '/Pushkar' (existing at 18:45).

```
hadoopuser@pushkar-VirtualBox:~$ hdfs dfs -ls /
Found 1 items
drwxr-xr-x  - hadoopuser supergroup          0 2022-10-04 18:45 /Pushkar
hadoopuser@pushkar-VirtualBox:~$ hdfs dfs -mkdir /NewDirectory
hadoopuser@pushkar-VirtualBox:~$ hdfs dfs -ls /
Found 2 items
drwxr-xr-x  - hadoopuser supergroup          0 2022-10-04 19:01 /NewDirectory
drwxr-xr-x  - hadoopuser supergroup          0 2022-10-04 18:45 /Pushkar
hadoopuser@pushkar-VirtualBox:~$
```

### - chmod

It is used to assign the read, write, and execute permission for any directory and files

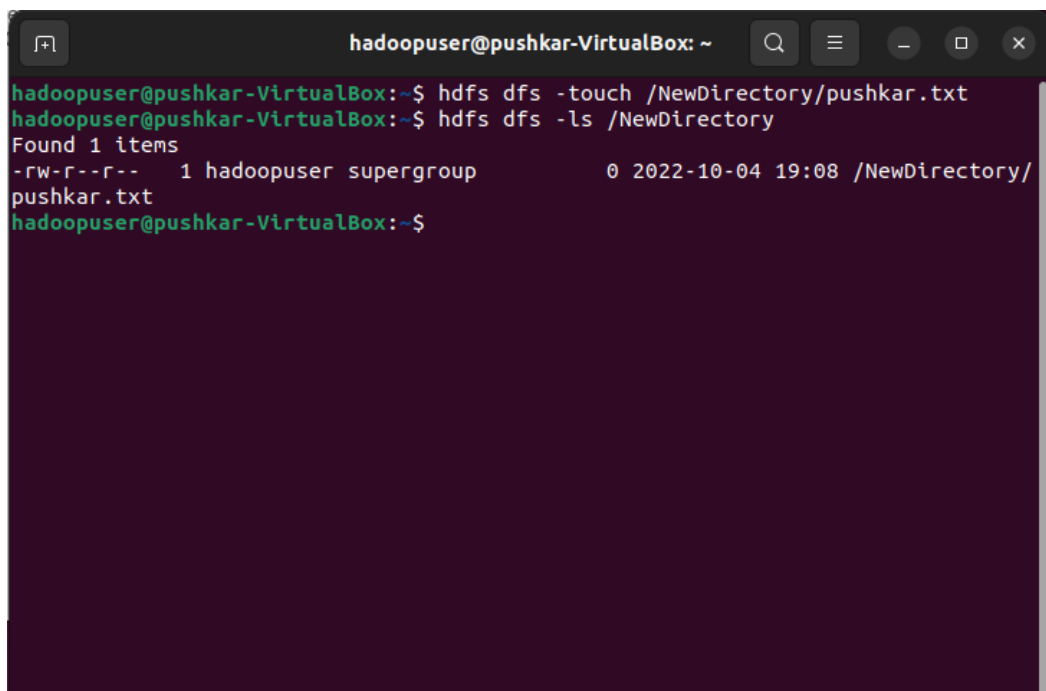
A terminal window titled 'hadoopuser@pushkar-VirtualBox: ~' showing the execution of HDFS commands. The user first lists the root directory, then creates a new directory, lists it, and finally changes its permissions to 777. The output shows the directory's creation and the successful permission change.

```
hadoopuser@pushkar-VirtualBox:~$ hdfs dfs -ls /
Found 1 items
drwxr-xr-x  - hadoopuser supergroup          0 2022-10-04 18:45 /Pushkar
hadoopuser@pushkar-VirtualBox:~$ hdfs dfs -mkdir /NewDirectory
hadoopuser@pushkar-VirtualBox:~$ hdfs dfs -ls /
Found 2 items
drwxr-xr-x  - hadoopuser supergroup          0 2022-10-04 19:01 /NewDirectory
drwxr-xr-x  - hadoopuser supergroup          0 2022-10-04 18:45 /Pushkar
hadoopuser@pushkar-VirtualBox:~$ hdfs dfs -chmod /
bin/      home/      lost+found/ root/      swapfile
boot/     lib/        media/     run/       sys/
cdrom/    lib32/     mnt/      sbin/      tmp/
dev/      lib64/     opt/      snap/     usr/
etc/      libx32/    proc/     srv/       var/
hadoopuser@pushkar-VirtualBox:~$ hdfs dfs -chmod 777 /NewDirectory
hadoopuser@pushkar-VirtualBox:~$ hdfs dfs -ls /
Found 2 items
drwxrwxrwx  - hadoopuser supergroup          0 2022-10-04 19:01 /NewDirectory
drwxr-xr-x  - hadoopuser supergroup          0 2022-10-04 18:45 /Pushkar
hadoopuser@pushkar-VirtualBox:~$
```

### -touch

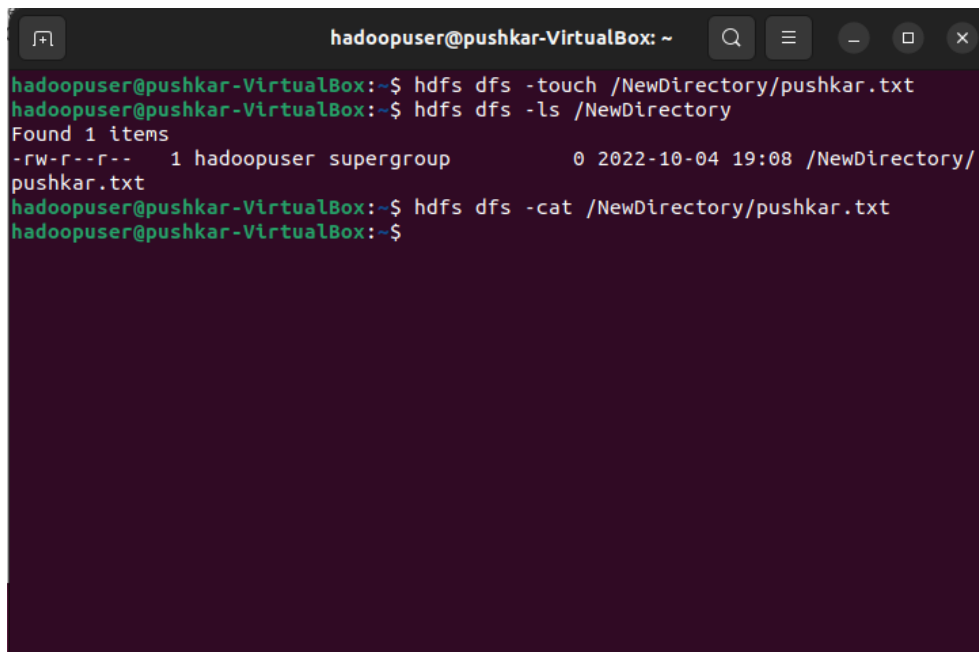
This command is used to create a new directory and file into the Hadoop cluster

It basically create a empty file and directory.

A terminal window titled 'hadoopuser@pushkar-VirtualBox: ~' showing the execution of HDFS commands. The user creates a new file in a directory and then lists the directory to confirm its creation. The output shows the file's creation with permissions -rw-r--r--.

```
hadoopuser@pushkar-VirtualBox:~$ hdfs dfs -touch /NewDirectory/pushkar.txt
hadoopuser@pushkar-VirtualBox:~$ hdfs dfs -ls /NewDirectory
Found 1 items
-rw-r--r--  1 hadoopuser supergroup          0 2022-10-04 19:08 /NewDirectory/
pushkar.txt
hadoopuser@pushkar-VirtualBox:~$
```

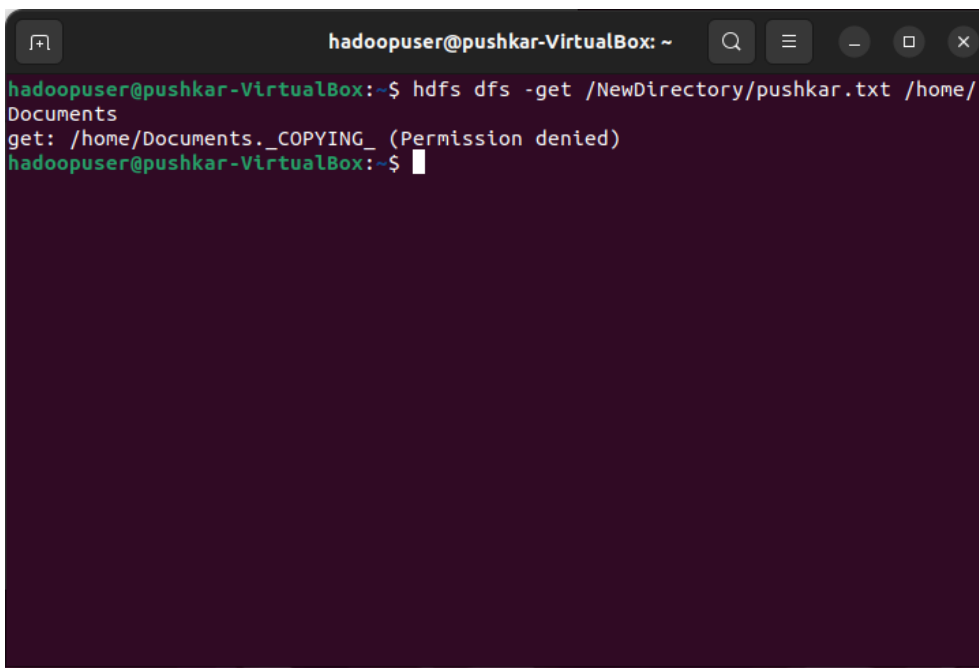
cat command for new created file

A terminal window titled 'hadoopuser@pushkar-VirtualBox: ~' with standard window controls. It shows the execution of 'hdfs dfs -touch /NewDirectory/pushkar.txt', followed by 'hdfs dfs -ls /NewDirectory' which lists the file with permissions '-rw-r--r--', owner 'hadoopuser', group 'supergroup', size '0', and timestamp '2022-10-04 19:08'. The final command 'hdfs dfs -cat /NewDirectory/pushkar.txt' is entered but produces no output.

```
hadoopuser@pushkar-VirtualBox:~$ hdfs dfs -touch /NewDirectory/pushkar.txt
hadoopuser@pushkar-VirtualBox:~$ hdfs dfs -ls /NewDirectory
Found 1 items
-rw-r--r--  1 hadoopuser supergroup          0 2022-10-04 19:08 /NewDirectory/
pushkar.txt
hadoopuser@pushkar-VirtualBox:~$ hdfs dfs -cat /NewDirectory/pushkar.txt
hadoopuser@pushkar-VirtualBox:~$
```

As the file is empty the content aren't shown

Get command for getting file from Hadoop cluster

A terminal window titled 'hadoopuser@pushkar-VirtualBox: ~' with standard window controls. It shows the execution of 'hdfs dfs -get /NewDirectory/pushkar.txt /home/Documents', which results in an error message: 'get: /home/Documents.\_COPYING\_ (Permission denied)'.

```
hadoopuser@pushkar-VirtualBox:~$ hdfs dfs -get /NewDirectory/pushkar.txt /home/
Documents
get: /home/Documents._COPYING_ (Permission denied)
hadoopuser@pushkar-VirtualBox:~$
```

**Creating a txt file using touch command and copying it into another directory**

```

hadoopuser@pushkar-VirtualBox: ~
hadoopuser@pushkar-VirtualBox:~$ hdfs dfs -cat /trial.txt
cat: `/trial.txt': No such file or directory
hadoopuser@pushkar-VirtualBox:~$ hdfs dfs -touch /trial.txt
hadoopuser@pushkar-VirtualBox:~$ hdfs dfs -ls /
Found 4 items
drwxrwxrwx   - hadoopuser supergroup          0 2022-10-04 19:24 /NewDir2
drwxrwxrwx   - hadoopuser supergroup          0 2022-10-04 19:23 /NewDirectory
drwxr-xr-x   - hadoopuser supergroup          0 2022-10-04 18:45 /Pushkar
-rw-r--r--   1 hadoopuser supergroup          0 2022-10-04 19:28 /trial.txt
hadoopuser@pushkar-VirtualBox:~$ hdfs dfs -cp /trial.txt /NewDir2/
hadoopuser@pushkar-VirtualBox:~$ hdfs dfs -ls /NewDir2
Found 1 items
-rw-r--r--   1 hadoopuser supergroup          0 2022-10-04 19:29 /NewDir2/trial
.txt
hadoopuser@pushkar-VirtualBox:~$

```

127.0.0.1:9870

How to Install and Config

Namenode information

+

127.0.0.1:9870/dfshealth.html#tab-overview

☆

🔍

🔔

☰

Hadoop

Overview

Datanodes

Datanode Volume Failures

Snapshot

Startup Progress

Utilities

## Overview 'hadoop.pushkar-VirtualBox.com:9000' (active)

|                       |  |
|-----------------------|--|
| <b>Started:</b>       | Tue Oct 04 18:07:37 +0530 2022                             |
| <b>Version:</b>       | 3.2.4, r7e5d9983b388e372fe640f21f048f2f2ae6e9eba           |
| <b>Compiled:</b>      | Tue Jul 12 17:28:00 +0530 2022 by ubuntu from branch-3.2.4 |
| <b>Cluster ID:</b>    | CID-6b881bfc-6d7a-4b3c-92ed-007f3766bdd4                   |
| <b>Block Pool ID:</b> | BP-1120510441-127.0.1.1-1664887034250                      |

## Summary

Security is off.

Safemode is off.

1 files and directories, 0 blocks (0 replicated blocks, 0 erasure coded block groups) = 1 total filesystem object(s).

Heap Memory used 34.71 MB of 87 MB Heap Memory. Max Heap Memory is 494 MB.

Non Heap Memory used 50.26 MB of 53.81 MB Committed Non Heap Memory. Max Non Heap Memory is <unbounded>.

|                                    |            |
|------------------------------------|------------|
| <b>Configured Capacity:</b>        | 0 B        |
| <b>Configured Remote Capacity:</b> | 0 B        |
| <b>DFS Used:</b>                   | 0 B (100%) |