

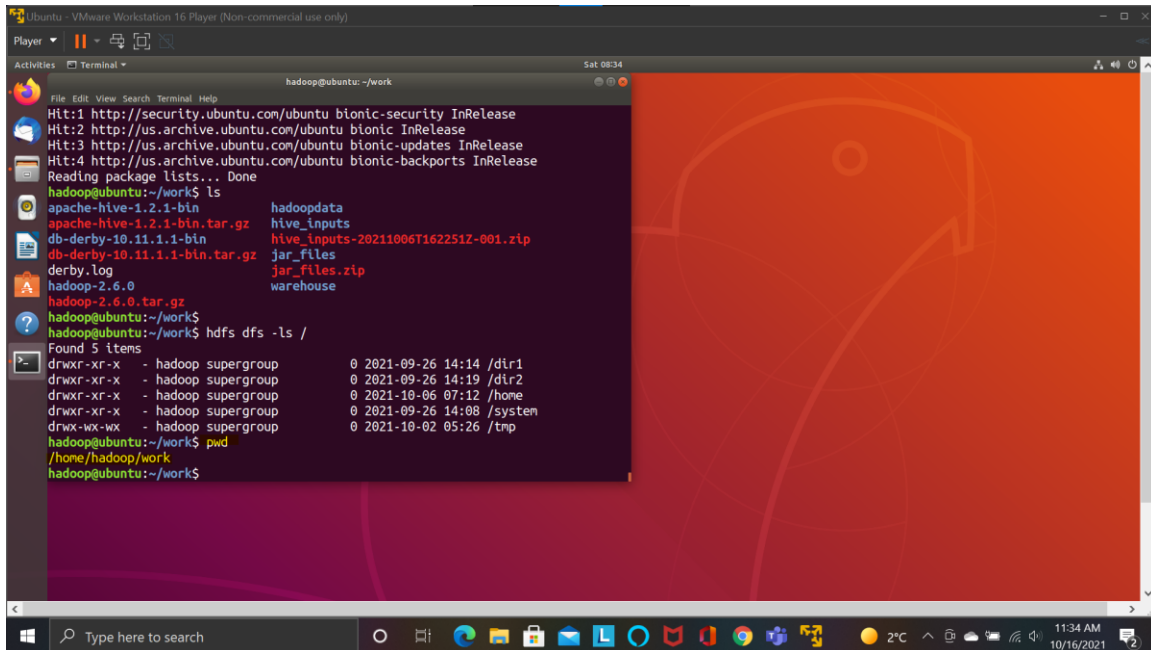
Lab Assignment 1

Hadoop Commands

Submitted by: Aadarsha Chapagain

1. Print current working directory

pwd



The screenshot shows a terminal window titled 'hadoop@ubuntu: ~/work'. The user has run the 'ls' command, which lists files in the current directory. The output shows several files including 'hadoopdata', 'hive_inputs', 'hive_inputs-20211006T162251Z-001.zip', 'jar_files', 'jar_files.zip', and 'warehouse'. The user then runs the 'hdfs dfs -ls /' command, which lists the root directory of the Hadoop file system. The output shows five items: 'dir1', 'dir2', 'home', 'system', and 'tmp'. Finally, the user runs the 'pwd' command, which outputs '/home/hadoop/work'.

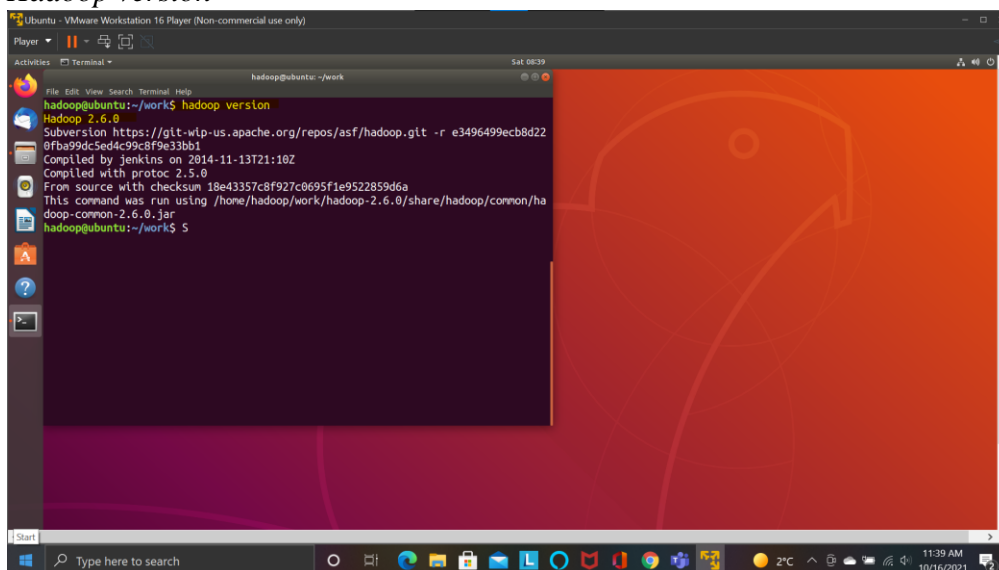
```
hadoop@ubuntu:~/work$ ls
apache-hive-1.2.1-bin      hadoopdata
apache-hive-1.2.1-bin.tar.gz  hive_inputs
db-derby-10.11.1.1-bin     hive_inputs-20211006T162251Z-001.zip
db-derby-10.11.1.1-bin.tar.gz jar_files
derby.log                 jar_files.zip
hadoop-2.6.0              warehouse
hadoop-2.6.0.tar.gz

hadoop@ubuntu:~/work$ hdfs dfs -ls /
Found 5 items
drwxr-xr-x - hadoop supergroup    0 2021-09-26 14:14 /dir1
drwxr-xr-x - hadoop supergroup    0 2021-09-26 14:19 /dir2
drwxr-xr-x - hadoop supergroup    0 2021-10-06 07:12 /home
drwxr-xr-x - hadoop supergroup    0 2021-09-26 14:06 /system
drwx-wx-wx - hadoop supergroup    0 2021-10-02 05:26 /tmp

hadoop@ubuntu:~/work$ pwd
/home/hadoop/work
hadoop@ubuntu:~/work$
```

2. Print the Hadoop version

Hadoop version

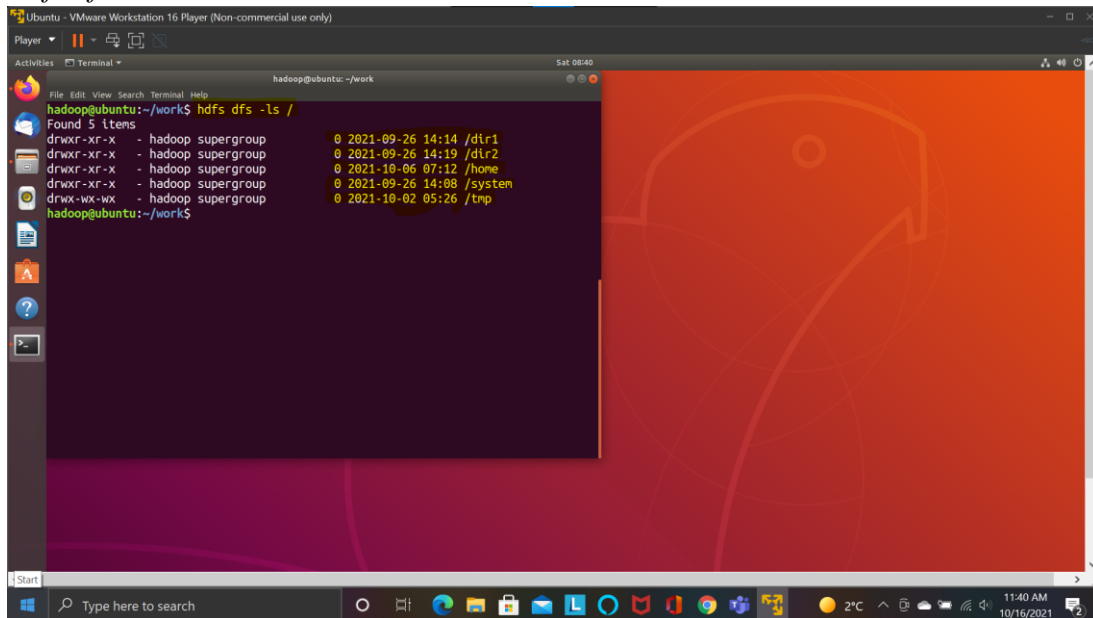


The screenshot shows a terminal window titled 'hadoop@ubuntu: ~/work'. The user has run the 'hadoop version' command, which outputs the Hadoop version and other details. The output shows 'Hadoop 2.6.0' and provides the Subversion URL, commit ID, and compilation date. It also shows the compilation protocol and the source checksum. Finally, the user runs the 'S' command, which outputs 'S'.

```
hadoop@ubuntu:~/work$ hadoop version
Hadoop 2.6.0
Subversion https://git-wip-us.apache.org/repos/asf/hadoop.git -r e3496499ecb8d22
0fba99dc5ed4c99c8f9e33bb1
Compiled by Jenkins on 2014-11-13T21:10Z
Compiled with protoc 2.5.0
From source with checksum 18e43357c8f927c0695f1e9522859d6a
This command was run using /home/hadoop/work/hadoop-2.6.0/share/hadoop/common/hadoop-common-2.6.0.jar
hadoop@ubuntu:~/work$ S
S
```

3. List the contents of the root directory in HDFS

Hdfs dfs -ls /

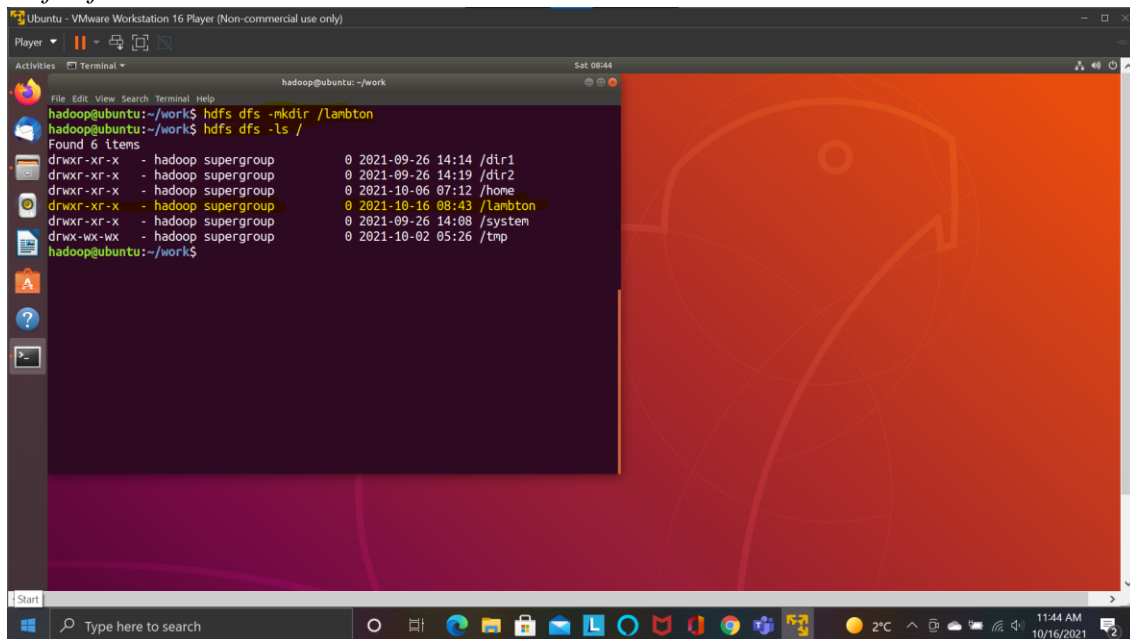


The screenshot shows a terminal window titled 'hadoop@ubuntu: ~/work' within a VMware Workstation 16 Player. The terminal displays the command `hadoop@ubuntu:~/work$ hdfs dfs -ls /` and its output, which lists 5 items in the root directory of HDFS. The output is as follows:

```
Found 5 items
drwxr-xr-x - hadoop supergroup 0 2021-09-26 14:14 /dir1
drwxr-xr-x - hadoop supergroup 0 2021-09-26 14:19 /dir2
drwxr-xr-x - hadoop supergroup 0 2021-10-06 07:12 /home
drwxr-xr-x - hadoop supergroup 0 2021-09-26 14:08 /system
drwx-wx-wx - hadoop supergroup 0 2021-10-02 05:26 /tmp
```

4. Create a new directory (lambton)

Hdfs dfs -mkdir /lambton

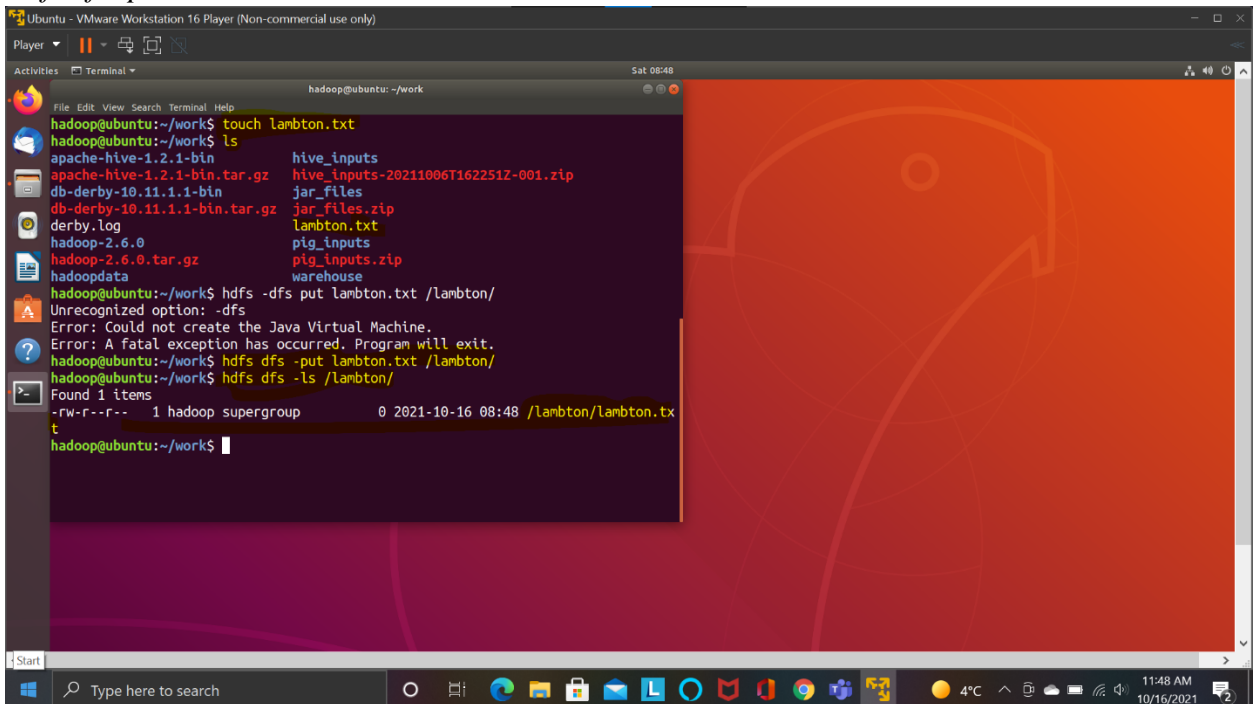


The screenshot shows a terminal window titled 'hadoop@ubuntu: ~/work' within a VMware Workstation 16 Player. The terminal displays two commands: `hadoop@ubuntu:~/work$ hdfs dfs -mkdir /lambton` and `hadoop@ubuntu:~/work$ hdfs dfs -ls /`. The output of the second command shows 6 items in the root directory of HDFS, including the newly created directory `/lambton`. The output is as follows:

```
Found 6 items
drwxr-xr-x - hadoop supergroup 0 2021-09-26 14:14 /dir1
drwxr-xr-x - hadoop supergroup 0 2021-09-26 14:19 /dir2
drwxr-xr-x - hadoop supergroup 0 2021-10-06 07:12 /home
drwxr-xr-x - hadoop supergroup 0 2021-10-16 08:43 /lambton
drwxr-xr-x - hadoop supergroup 0 2021-09-26 14:08 /system
drwx-wx-wx - hadoop supergroup 0 2021-10-02 05:26 /tmp
```

5. Add a text file (lambton.txt) from the local directory to HDFS

Hdfs dfs -put lambton.txt /Lambton/

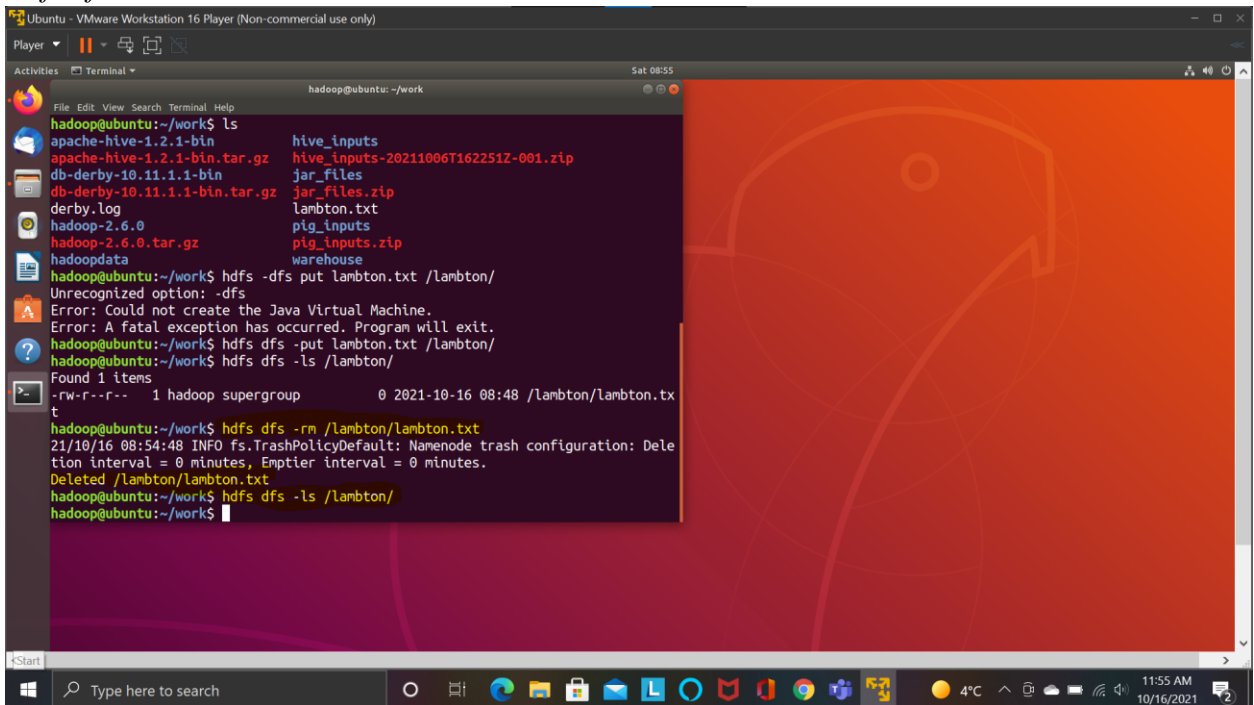


The screenshot shows a terminal window titled 'hadoop@ubuntu: ~/work'. The user has created a file 'lambton.txt' and listed the contents of the directory. They then attempted to upload the file to HDFS using 'hdfs -dfs put lambton.txt /lambton/'. This command resulted in an error: 'Unrecognized option: -dfs'. The user then corrected the command to 'hdfs dfs -put lambton.txt /lambton/' and successfully uploaded the file. Finally, they used 'hdfs dfs -ls /lambton/' to verify the upload, showing the file 'lambton.txt' with permissions '-rw-r--r--' and owner 'hadoop supergroup'.

```
hadoop@ubuntu:~/work$ touch lambton.txt
hadoop@ubuntu:~/work$ ls
apache-hive-1.2.1-bin          hive_inputs
apache-hive-1.2.1-bin.tar.gz  hive_inputs-20211006T162251Z-001.zip
db-derby-10.11.1.1-bin       jar_files
db-derby-10.11.1.1-bin.tar.gz jar_files.zip
derby.log                    lambton.txt
hadoop-2.6.0                 pig_inputs
hadoop-2.6.0.tar.gz          pig_inputs.zip
hadoopdata                   warehouse
hadoop@ubuntu:~/work$ hdfs -dfs put lambton.txt /lambton/
Unrecognized option: -dfs
Error: Could not create the Java Virtual Machine.
Error: A fatal exception has occurred. Program will exit.
hadoop@ubuntu:~/work$ hdfs dfs -put lambton.txt /lambton/
hadoop@ubuntu:~/work$ hdfs dfs -ls /lambton/
Found 1 items
-rw-r--r--  1 hadoop supergroup          0 2021-10-16 08:48 /lambton/lambton.txt
hadoop@ubuntu:~/work$
```

6. Delete a file (lambton.txt) from HDFS.

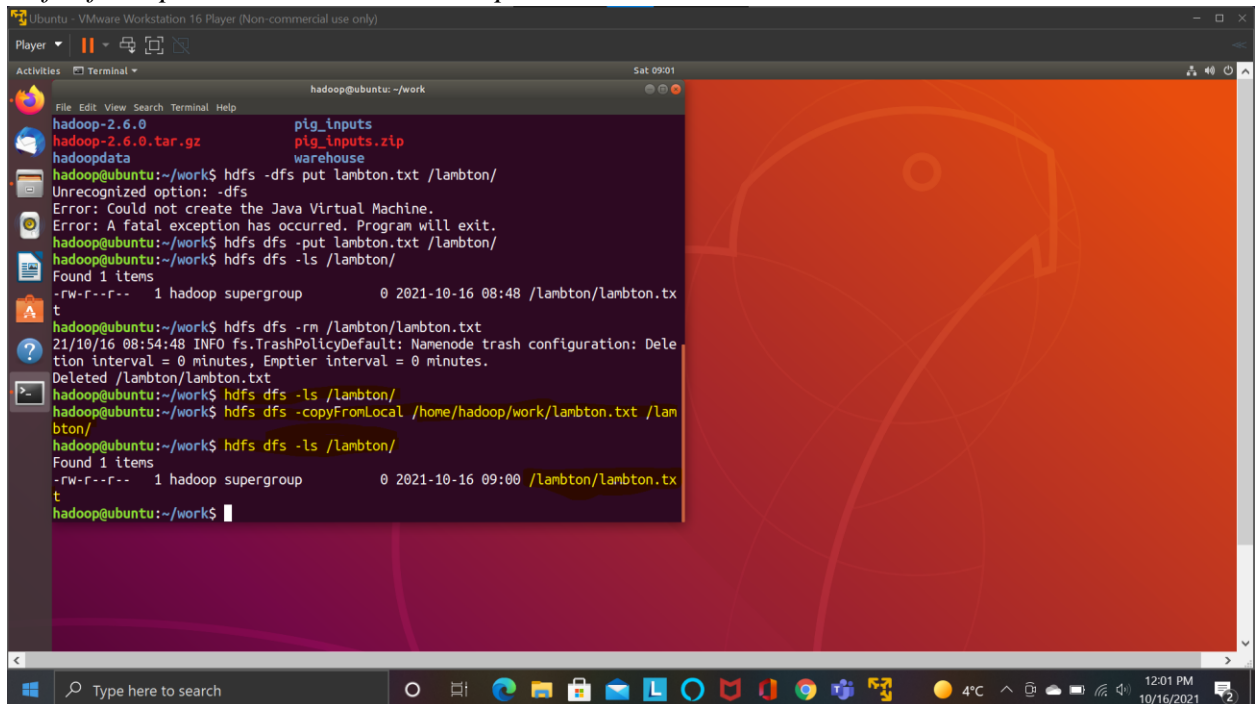
Hdfs dfs -rm /Lambton/Lambton.txt



This screenshot shows the same terminal window as the previous one, but now the user is deleting the file from HDFS. They run 'hdfs dfs -rm /lambton/lambton.txt', which successfully removes the file. The terminal output shows a confirmation message from the NameNode trash configuration.

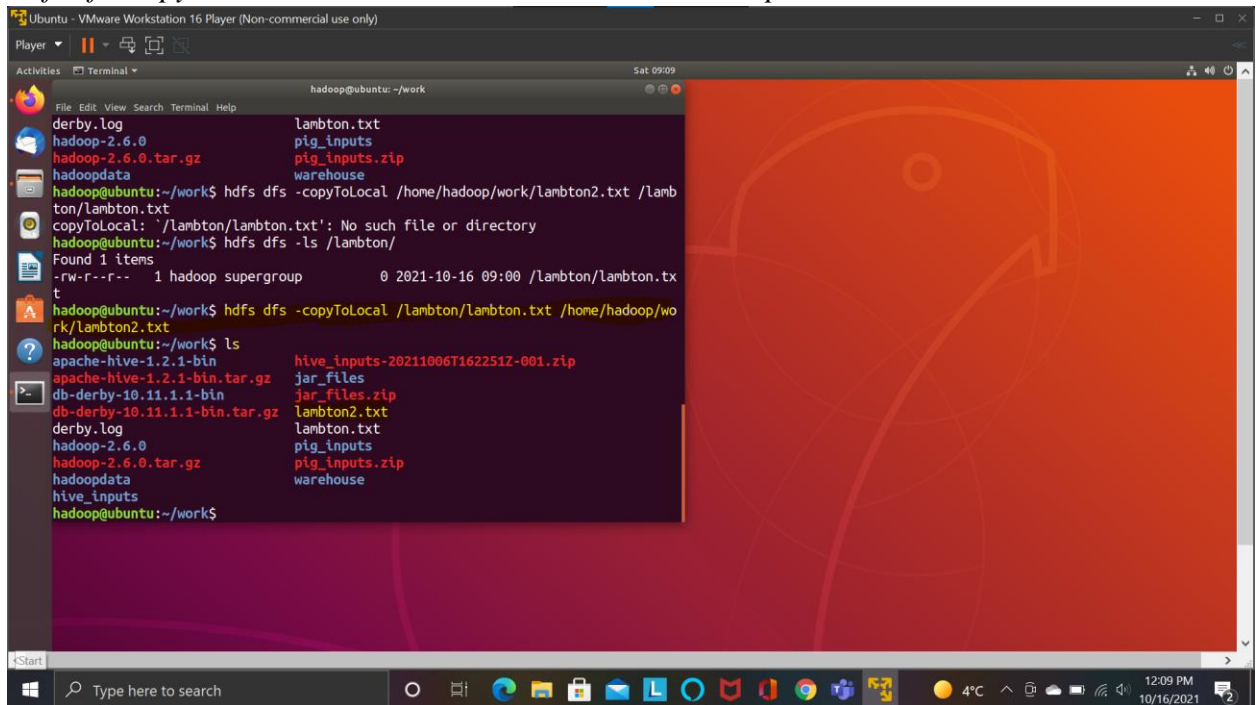
```
hadoop@ubuntu:~/work$ hdfs dfs -rm /lambton/lambton.txt
21/10/16 08:54:48 INFO fs.TrashPolicyDefault: Namenode trash configuration: Deletion interval = 0 minutes, Emptier interval = 0 minutes.
Deleted /lambton/lambton.txt
hadoop@ubuntu:~/work$ hdfs dfs -ls /lambton/
hadoop@ubuntu:~/work$
```

7. Add (lambton.txt) file from the local directory to HDFS using copyFromLocal utility
Hdfs dfs -copyFromLocal /home/hadoop/work/Lambton.txt /Lambton/



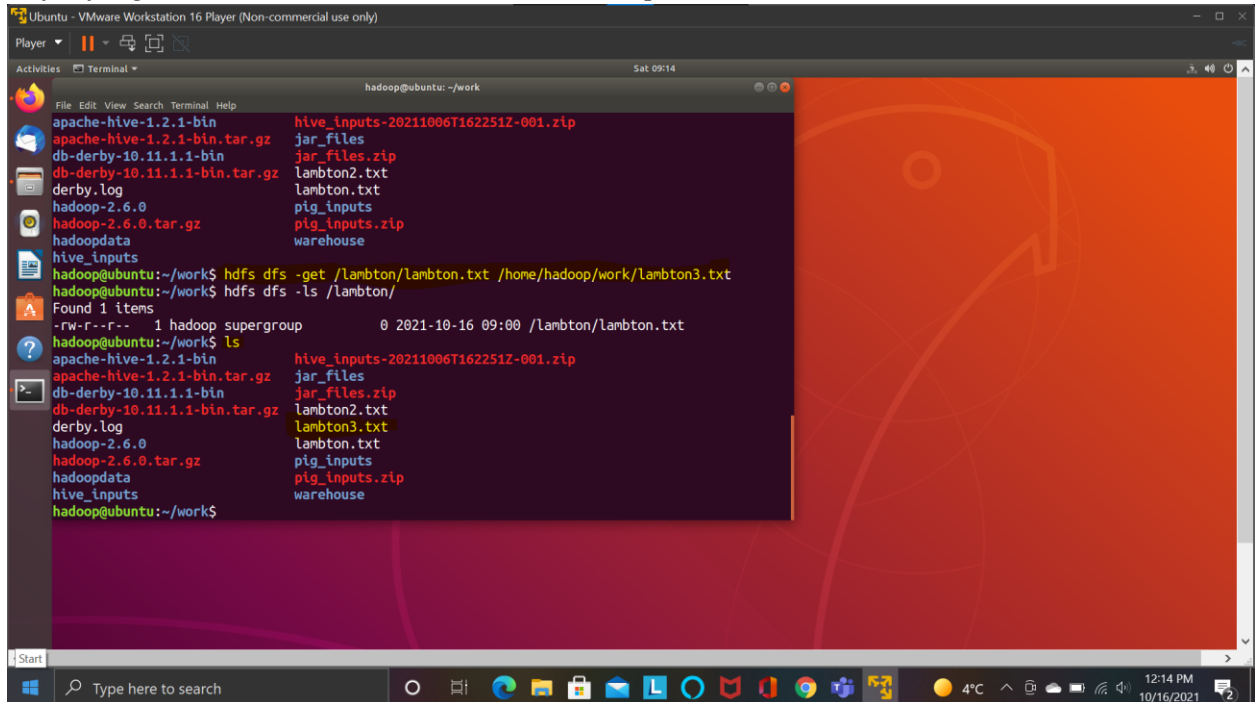
```
hadoop@ubuntu: ~/work
hadoop-2.6.0      pig_inputs
hadoop-2.6.0.tar.gz pig_inputs.zip
hadoopdata        warehouse
hadoop@ubuntu:~/work$ hdfs -dfs put lambton.txt /lambton/
Unrecognized option: -dfs
Error: Could not create the Java Virtual Machine.
Error: A fatal exception has occurred. Program will exit.
hadoop@ubuntu:~/work$ hdfs dfs -put lambton.txt /lambton/
hadoop@ubuntu:~/work$ hdfs dfs -ls /lambton/
Found 1 items
-rw-r--r--  1 hadoop supergroup          0 2021-10-16 08:48 /lambton/lambton.tx
t
hadoop@ubuntu:~/work$ hdfs dfs -rm /lambton/lambton.txt
21/10/16 08:54:48 INFO fs.TrashPolicyDefault: Namenode trash configuration: Dele
tion interval = 0 minutes, Emptier interval = 0 minutes.
Deleted /lambton/lambton.txt
hadoop@ubuntu:~/work$ hdfs dfs -ls /lambton/
hadoop@ubuntu:~/work$ hdfs dfs -copyFromLocal /home/hadoop/work/lambton.txt /Lam
bton/
hadoop@ubuntu:~/work$ hdfs dfs -ls /lambton/
Found 1 items
-rw-r--r--  1 hadoop supergroup          0 2021-10-16 09:00 /lambton/lambton.tx
t
hadoop@ubuntu:~/work$
```

8. Get (lambton.txt) file from HDFS directory to the local Directory using copyToLocal utility
Hdfs dfs -copyToLocal /Lambton/Lambton.txt /home/hadoop/work/lambton2.txt



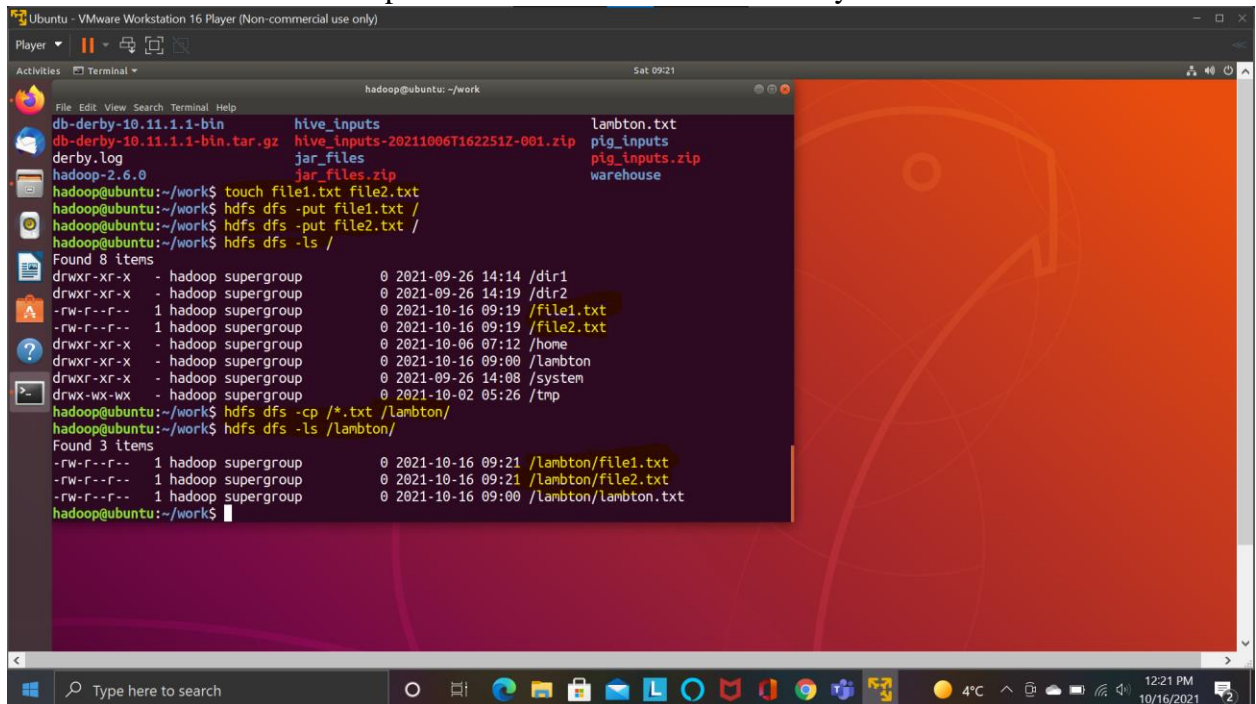
```
hadoop@ubuntu: ~/work
derby.log      lambton.txt
hadoop-2.6.0   pig_inputs
hadoop-2.6.0.tar.gz pig_inputs.zip
hadoopdata     warehouse
hadoop@ubuntu:~/work$ hdfs dfs -copyToLocal /home/hadoop/work/lambton2.txt /lamb
ton/lambton.txt
copyToLocal: '/lambton/lambton.txt': No such file or directory
hadoop@ubuntu:~/work$ hdfs dfs -ls /lambton/
Found 1 items
-rw-r--r--  1 hadoop supergroup          0 2021-10-16 09:00 /lambton/lambton.tx
t
hadoop@ubuntu:~/work$ hdfs dfs -copyToLocal /lambton/lambton.txt /home/hadoop/wo
rk/lambton2.txt
hadoop@ubuntu:~/work$ ls
apache-hive-1.2.1-bin      hive_inputs-20211006T162251Z-001.zip
apache-hive-1.2.1-bin.tar.gz jar_files
db-derby-10.11.1.1-bin     jar_files.zip
db-derby-10.11.1.1-bin.tar.gz lambton2.txt
derby.log                  lambton.txt
hadoop-2.6.0               pig_inputs
hadoop-2.6.0.tar.gz        pig_inputs.zip
hadoopdata                 warehouse
hive_inputs
hadoop@ubuntu:~/work$
```

9. Get (lambton.txt) file from HDFS directory to the local Directory using get utility
Hdfs dfs -get /Lambton/Lambton.txt /home/hadoop/work/lambton3.txt



The screenshot shows a terminal window titled 'hadoop@ubuntu: ~/work'. The user has listed files in the current directory, including 'lambton.txt'. They then execute the command 'hdfs dfs -get /lambton/lambton.txt /home/hadoop/work/lambton3.txt'. After this, they run 'hdfs dfs -ls /lambton/' and find one item: 'lambton.txt' with permissions '-rw-r--r--', owned by 'hadoop supergroup', dated '2021-10-16 09:00', and located at '/lambton/lambton.txt'.

10. Copy any 2 files from one to another directories present in HDFS using cp command in HDFS
Hdfs dfs -cp /.txt /Lambton/*
File1.txt and file2.txt were copied from root to Lambton directory.



The screenshot shows a terminal window titled 'hadoop@ubuntu: ~/work'. The user creates two files, 'file1.txt' and 'file2.txt', using 'touch'. They then upload them to the root of HDFS with 'hdfs dfs -put file1.txt /' and 'hdfs dfs -put file2.txt /'. A subsequent 'hdfs dfs -ls /' command shows the root directory listing, including 'file1.txt' and 'file2.txt'. Finally, they execute 'hdfs dfs -cp /*.txt /lambton/' to copy these files to the 'lambton' directory. A final 'hdfs dfs -ls /lambton/' command shows three items: 'file1.txt', 'file2.txt', and 'lambton.txt', all with permissions '-rw-r--r--' and owned by 'hadoop supergroup'.