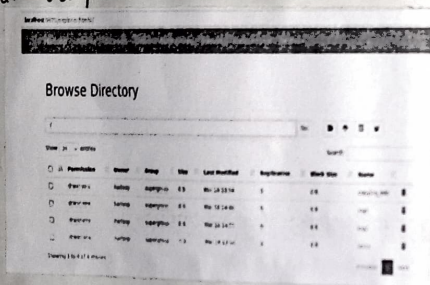


a) 1. lists of files & directories output

```
hadoop@cmrct-virtual-machine:~$ hadoop fs -ls /
Found 2 items
drwxr-xr-x   - hadoop supergroup          0 2023-03-18 13:54 /copying_into
drwxr-xr-x   - hadoop supergroup          0 2023-03-18 12:34 /test3
```

2) -mkdir output



Ex. No. Week-4

Date

Page No.

## Hadoop commands

a) Introducing Hadoop command

1. hdfs dfs -ls: Lists the files and directories in the Hadoop Distributed File System (HDFS)  
command:

```
hadoop@cmrct-virtual-machine$ hadoop fs -ls /
```

2. hdfs dfs -mkdir: creates a new directory in the HDFS  
command:

```
hadoop@cmrct-virtual-machine$ hadoop dfs -mkdir /log2
```

```
hadoop@cmrct-virtual-machine:~$ hadoop fs -mkdir /log2
hadoop@cmrct-virtual-machine:~$
```

3. hdfs dfs -put: copies a file from the local file system to the HDFS

command:

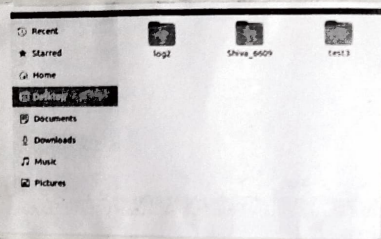
```
$ hadoop fs -put /home/hadoop/Desktop/shiva_6609/test3 /log1
```

```
hadoop@cmrct-virtual-machine:~$ hadoop fs -put /home/hadoop/Desktop/shiva_6609/test3 /log1
hadoop@cmrct-virtual-machine:~$
```

4) -cat output

```
hadoop@cnrcet-virtual-machine:~$ hadoop fs -cat /log1
cat: '/log1': Is a directory
```

5) -get output



6) -rm output

```
hadoop@cnrcet-virtual-machine:~$ hadoop fs -rm -r /log2
Deleted /log2
```

7) start-stop command output

```
hadoop@cnrcet-virtual-machine:~$ start-yarn.sh
Starting resourcemanager
Starting nodemanagers
hadoop@cnrcet-virtual-machine:~$ start-dfs.sh
Starting namenodes on [localhost]
Starting datanodes
Starting secondary namenodes [cnrcet-virtual-machine]
hadoop@cnrcet-virtual-machine:~$ stop-yarn.sh
Stopping nodemanagers
Stopping resourcemanager
hadoop@cnrcet-virtual-machine:~$ stop-dfs.sh
Stopping namenodes on [localhost]
Stopping datanodes
Stopping secondary namenodes [cnrcet-virtual-machine]
hadoop@cnrcet-virtual-machine:~$
```

Ex. No.

Date

Page No.

4. hdfs dfs -cat: Displays the content of a file stored in the HDFS.

command:

```
$hadoop fs -cat /log1
```

5. hdfs dfs -get: copies a file from the HDFS to the local file system.

command:

```
$hadoop fs -get /log2 /home/hadoop/Desktop/
```

```
hadoop@cnrcet-virtual-machine:~$ hadoop fs -get /log2 /home/hadoop/Desktop/
hadoop@cnrcet-virtual-machine:~$
```

6. hdfs dfs -rm: Deletes a file or directory from the HDFS.

command:

```
$hadoop fs -rm -r /log2
```

7. start & stop commands.

command

```
$ start-yarn.sh
```

```
$ start-dfs.sh
```

```
$ stop-yarn.sh
```

```
$ stop-dfs.sh
```



b) output

localhost:5003 (-active)

## Overview localhost:5003 (-active)

**Master:** Sat Mar 24 13:49:42 +0530 2023  
**Version:** 3.3.4-rc556a7a7e4d1e0c4220a134643e3e76095a36d0  
**Compiled:** Fri Jul 29 16:02:51 +0530 2022 by shiva from branch 3.3.4  
**Cluster ID:** C0-3c4779a-e0e1-4566-90ed-1975a9938c  
**Block Pool ID:** BP-34661c3895-127-C-1.3.387612.472844

## Summary

c) output

```
hadoop@crccet-virtual-machine:~$ hdfs fsck /
Connecting to namenode via http://localhost:8025/fsc?op=hadoopoppath=/
fsck started by hadoop (auth:SIMPLE) from [127.0.0.1 for path / at Sat Mar 24 14:01:37 2023]
```

```
STATUS: Healthy
Number of data-nodes: 0
Number of racks: 0
Total disks: 0
Total symlinks: 0
```

```
Replicated blocks:
Total size: 0 B
Total files: 0
Total blocks (validated): 0
Nonfully replicated blocks: 0
Over-replicated blocks: 0
Under-replicated blocks: 0
Mis-replicated blocks: 0
Orphaned replication factor: 1
Average block replication: 0.0
Missing blocks: 0
Corrupt blocks: 0
Missing replicas: 0
Blocks queued for replication: 0
```

```
Erasure-coded block groups:
Total size: 0 B
Total files: 0
Total block groups (validated): 0
Nonfully erasure-coded block groups: 0
Over-erasure-coded block groups: 0
Under-erasure-coded block groups: 0
Mis-erasure-coded block groups: 0
Average block group size: 0.0
Missing block groups: 0
Corrupt block groups: 0
Missing internal blocks: 0
Blocks queued for replication: 0
fsck ended at Sat Mar 24 14:01:37 2023 in 5 milliseconds
```

The filesystem under path '/' is HEALTHY

Ex. No.

Date

Page No.

b) Navigating the location where the Datanodes store data  
 To navigate the location, we just have to type  
 "localhost:9870/" in the browser.

c) checking the current status of HDFS by using the  
 "fsck" command.

command:

```
$ hdfs fsck /
```

d) Loading the small size data by using the "copyFromLocal"  
 To load the data through there are few steps to be  
 followed.

1. create a directory where you want to copy files.  
 \$ hdfs dfs -mkdir /copying-into

2. Check the file where created or not.

i) go to the browser type "localhost:9870"

ii) click the utilities

iii) Browse the file system.

3. check the files you want to copy

```
$ cd /home/hadoop/Desktop/shiva-6609/
home/hadoop/Desktop/shiva-6609 $ ls
```

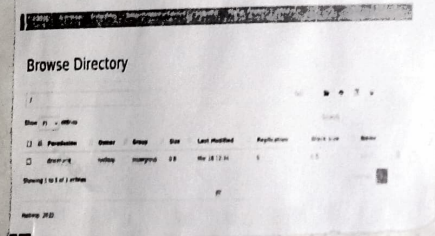
```
hadoop@crccet-virtual-machine:~$ cd /home/hadoop/Desktop/shiva_6609/
hadoop@crccet-virtual-machine:~/Desktop/shiva_6609$ ls
*.png *.txt
hadoop@crccet-virtual-machine:~/Desktop/shiva_6609$
```



d) output

```
hadoop@cmrct-virtual-machine: ~/Desktop/Shiva_6609$ hdfs dfs -copyFromLocal /home/hadoop/Desktop/Shiva_6609/test3 /copying_into
hadoop@cmrct-virtual-machine: ~/Desktop/Shiva_6609$
```

e) output



f) output

```
hadoop@cmrct-virtual-machine: $ hdfs dfs -ls /
Found 1 items
drwxr-xr-x 1 hadoop supergroup 0 2023-03-18 12:34 /test3
hadoop@cmrct-virtual-machine: $ hdfs dfs -get /test3 /home/hadoop/Desktop/Shiva_6609
hadoop@cmrct-virtual-machine: $ cd /home/hadoop/Desktop/Shiva_6609/
hadoop@cmrct-virtual-machine: /Desktop/Shiva_6609$ ls
test3
hadoop@cmrct-virtual-machine: /Desktop/Shiva_6609$
```

Ex. No.

Date

Page No.

4. Here is the main command to be executed.

```
$ hdfs dfs -copyFromLocal /home/hadoop/Desktop/Shiva_6609/test3 /copying-into
```

Path1: File Location

path2: Where to be loaded

2) Show the output of the copyFromLocal command.

check the localhost whether the file/data is copied or not

1. go to browser 9870/

2. Utilities

3. go to the browser the file system.

There you can check whether copied or not.

4) Reading the data from HDFS

The data can be read by -get command.

1. Check the data you want to read

```
$ hdfs dfs -ls /
```

2. use -get command and give the path where you want to read

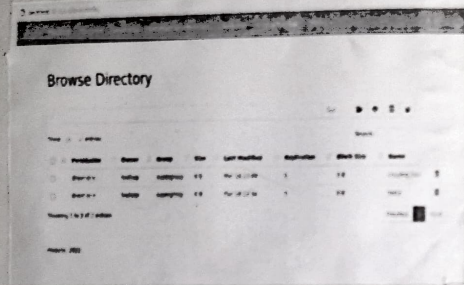
```
$ hdfs dfs -get /test3 /home/hadoop/Desktop/Shiva_6609
```

3. check whether data is read or not!

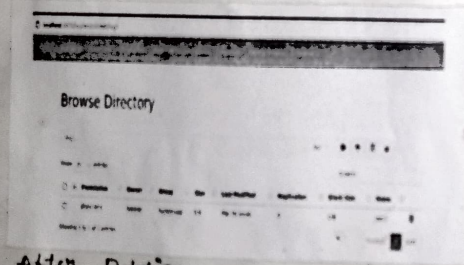
```
$ cd /home/hadoop/Desktop/Shiva_6609/
```

```
:~/Desktop/Shiva_6609$ ls
```

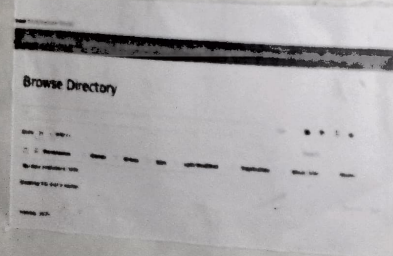
g) output



h) before Deletion



After Deletion



Ex. No.

Date

Page No.

g) creating a Directory in HDFS

we can create the directory using -mkdir

\$hdfs dfs -mkdir /test3

check whether the directory created or not.

\$hdfs dfs -ls /

```
hadoop@cwccet-virtual-machine1:~$ hdfs dfs -mkdir /test3
hadoop@cwccet-virtual-machine1:~$ hdfs dfs -ls /
Found 2 items
drwxr-xr-x 2 hadoop supergroup 4096 2023-03-18 12:11 /test3
```

or in localhost:9870/

h) Removing files from HDFS

1. firstly, check the files you want to delete from the localhost 9870 → Utilities → go to "Browse the file system".

2. To delete we use -rm -r command:

\$hadoop fs -rm -r /test3

```
hadoop@cwccet-virtual-machine1:~$ hadoop fs -rm -r /test3
Deleted /test3
hadoop@cwccet-virtual-machine1:~$
```