

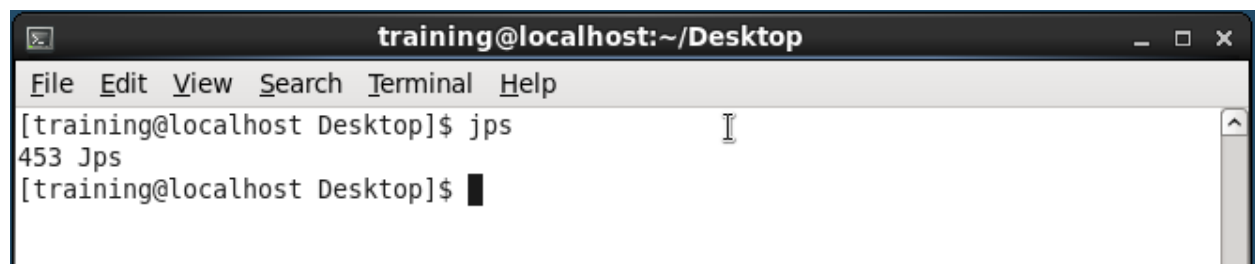
NAME=ADITYA ANAL

PROFILE=DATA ENGINEER

Hadoop commands tutorials

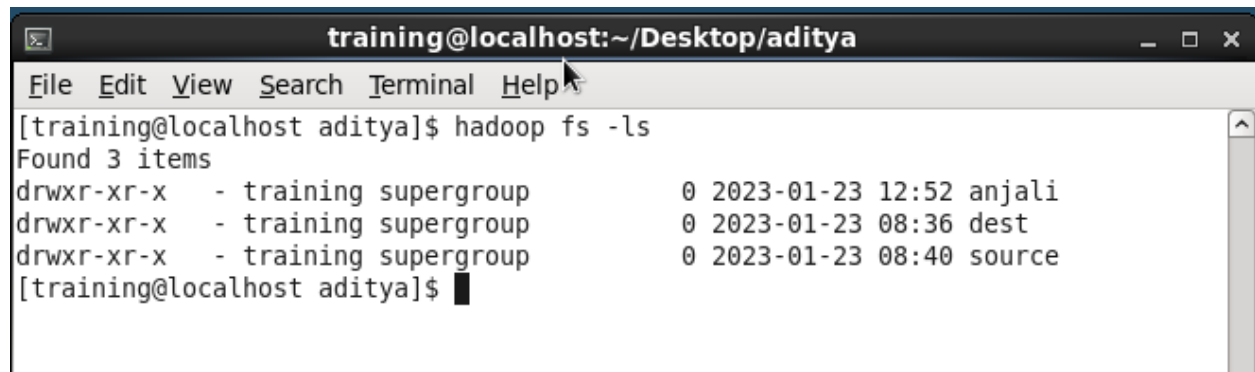
Hdfs: HDFS is the primary or major component of the Hadoop ecosystem which is responsible for storing large data sets of structured or unstructured data across various nodes and thereby maintaining the metadata in the form of log files.

1.Jps command: To check the Hadoop services are up and running use.

A terminal window titled 'training@localhost:~/Desktop' with a menu bar (File, Edit, View, Search, Terminal, Help). The command '[training@localhost Desktop]\$ jps' has been executed, resulting in the output '453 Jps'. The prompt '[training@localhost Desktop]\$' is visible on the next line.

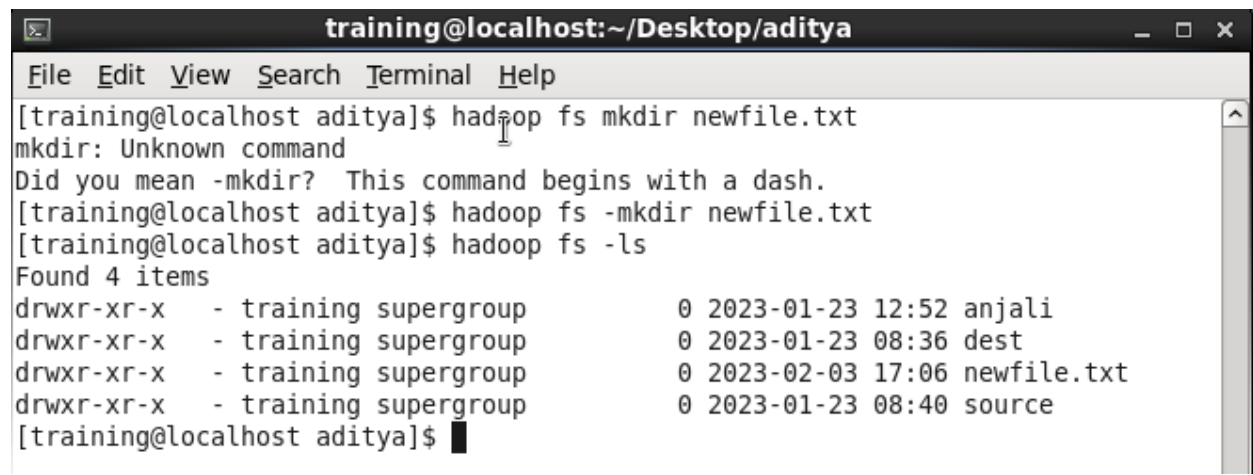
```
training@localhost:~/Desktop
File Edit View Search Terminal Help
[training@localhost Desktop]$ jps
453 Jps
[training@localhost Desktop]$
```

2.ls: This command is used to list all the files.

A terminal window titled 'training@localhost:~/Desktop/aditya' with a menu bar (File, Edit, View, Search, Terminal, Help). The command '[training@localhost aditya]\$ hadoop fs -ls' has been executed, resulting in the output 'Found 3 items' followed by a table of file details. The prompt '[training@localhost aditya]\$' is visible on the next line.

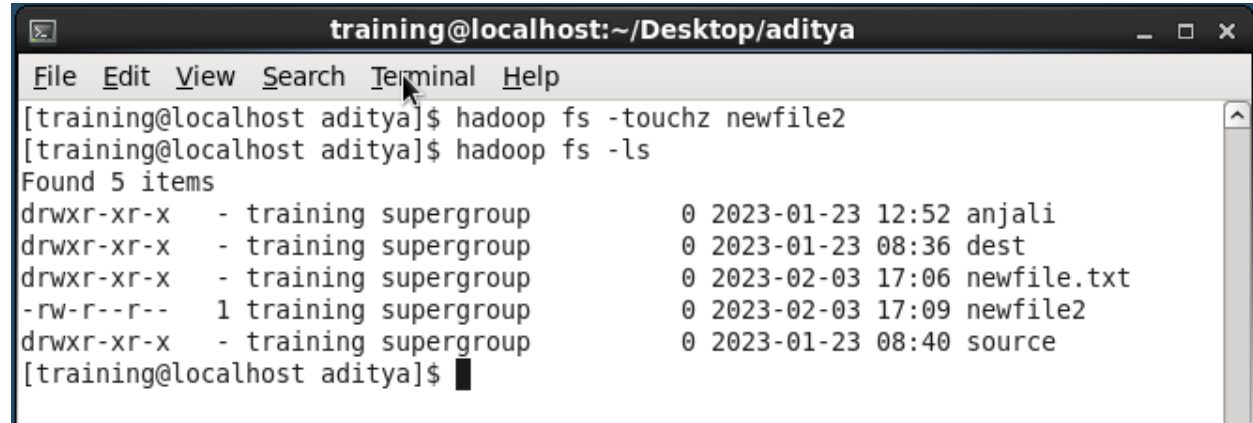
```
training@localhost:~/Desktop/aditya
File Edit View Search Terminal Help
[training@localhost aditya]$ hadoop fs -ls
Found 3 items
drwxr-xr-x - training supergroup      0 2023-01-23 12:52 anjali
drwxr-xr-x - training supergroup      0 2023-01-23 08:36 dest
drwxr-xr-x - training supergroup      0 2023-01-23 08:40 source
[training@localhost aditya]$
```

3.mkdir: To create a directory.

A terminal window titled 'training@localhost:~/Desktop/aditya' with a menu bar (File, Edit, View, Search, Terminal, Help). The terminal shows the following commands and output:

```
[training@localhost aditya]$ hadoop fs mkdir newfile.txt
mkdir: Unknown command
Did you mean -mkdir? This command begins with a dash.
[training@localhost aditya]$ hadoop fs -mkdir newfile.txt
[training@localhost aditya]$ hadoop fs -ls
Found 4 items
drwxr-xr-x - training supergroup          0 2023-01-23 12:52 anjali
drwxr-xr-x - training supergroup          0 2023-01-23 08:36 dest
drwxr-xr-x - training supergroup          0 2023-02-03 17:06 newfile.txt
drwxr-xr-x - training supergroup          0 2023-01-23 08:40 source
[training@localhost aditya]$
```

4.touchz: It creates an empty file.

A terminal window titled 'training@localhost:~/Desktop/aditya' with a menu bar (File, Edit, View, Search, Terminal, Help). The terminal shows the following commands and output:

```
[training@localhost aditya]$ hadoop fs -touchz newfile2
[training@localhost aditya]$ hadoop fs -ls
Found 5 items
drwxr-xr-x - training supergroup          0 2023-01-23 12:52 anjali
drwxr-xr-x - training supergroup          0 2023-01-23 08:36 dest
drwxr-xr-x - training supergroup          0 2023-02-03 17:06 newfile.txt
-rw-r--r-- 1 training supergroup          0 2023-02-03 17:09 newfile2
drwxr-xr-x - training supergroup          0 2023-01-23 08:40 source
[training@localhost aditya]$
```

5.rmr: This command deletes a file from HDFS *recursively*.

```
training@localhost:~/Desktop/aditya
File Edit View Search Terminal Help
[training@localhost aditya]$ hadoop fs -touchz newfile2
[training@localhost aditya]$ hadoop fs -ls
Found 5 items
drwxr-xr-x - training supergroup 0 2023-01-23 12:52 anjali
drwxr-xr-x - training supergroup 0 2023-01-23 08:36 dest
drwxr-xr-x - training supergroup 0 2023-02-03 17:06 newfile.txt
-rw-r--r-- 1 training supergroup 0 2023-02-03 17:09 newfile2
drwxr-xr-x - training supergroup 0 2023-01-23 08:40 source
[training@localhost aditya]$ hadoop fs -rmr anjali
\rmr: DEPRECATED: Please use 'rm -r' instead.
Deleted anjali
[training@localhost aditya]$
```

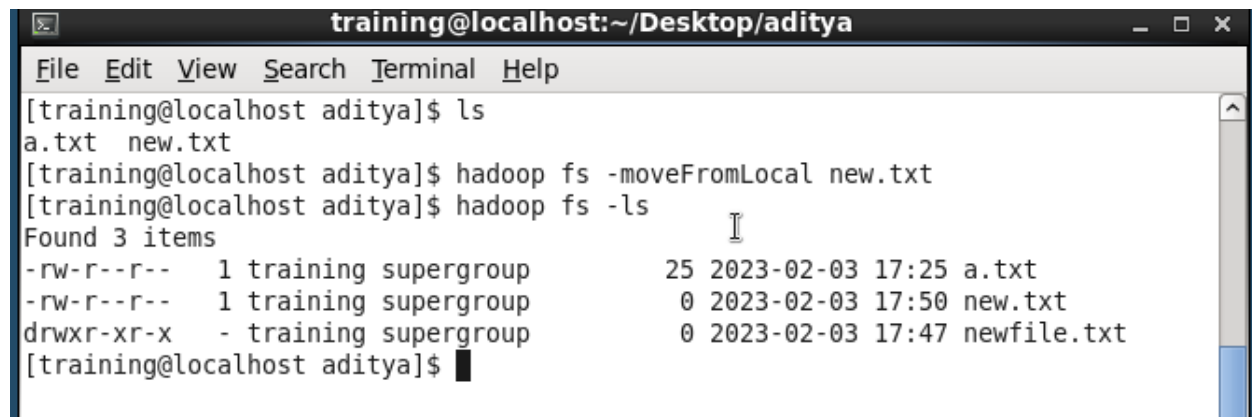
6.copyFromLocal (or) put: To copy files/folders from local file system to hdfs store.

```
training@localhost:~/Desktop/aditya
File Edit View Search Terminal Help
[training@localhost aditya]$ ls
a.txt
[training@localhost aditya]$ hadoop fs -copyFromLocal a.txt /
[training@localhost aditya]$ hadoop fs -ls /
Found 6 items
-rw-r--r-- 1 training supergroup 25 2023-02-03 17:27 /a.txt
drwxr-xr-x - training supergroup 0 2023-01-23 11:45 /dost
drwxr-xr-x - hbase supergroup 0 2023-02-03 11:50 /hbase
drwxrwxrwt - hdfs supergroup 0 2023-01-23 07:35 /tmp
drwxrwxrwx - hdfs supergroup 0 2023-01-23 07:35 /user
drwxr-xr-x - hdfs supergroup 0 2023-01-23 07:35 /var
[training@localhost aditya]$
```

7.copyToLocal (or) get: To copy files/folders from hdfs store to local file system.

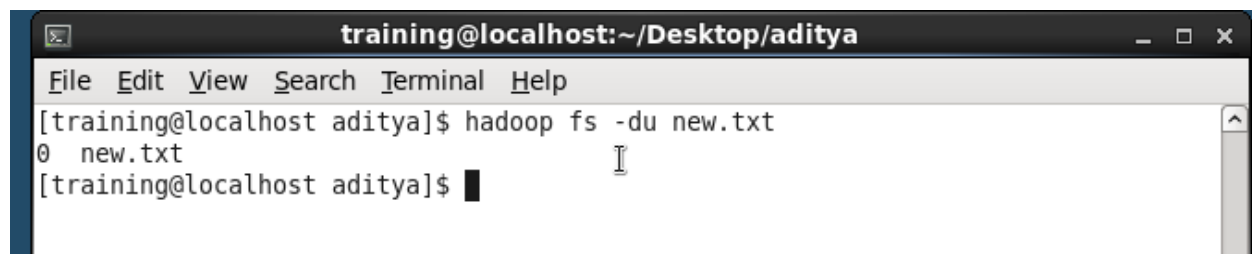
```
copyToLocal: must specify input/output path for  
[training@localhost aditya]$ pwd  
/home/training/Desktop/aditya  
[training@localhost aditya]$ hadoop fs -copyToLocal newfile.txt /home/training/  
Desktop/aditya  
[training@localhost aditya]$ ls  
a.txt  newfile.txt  
[training@localhost aditya]$
```

8.moveFromLocal: This command will move file from local to hdfs.

A terminal window titled 'training@localhost:~/Desktop/aditya' with a menu bar (File, Edit, View, Search, Terminal, Help). The terminal shows the following commands and output:

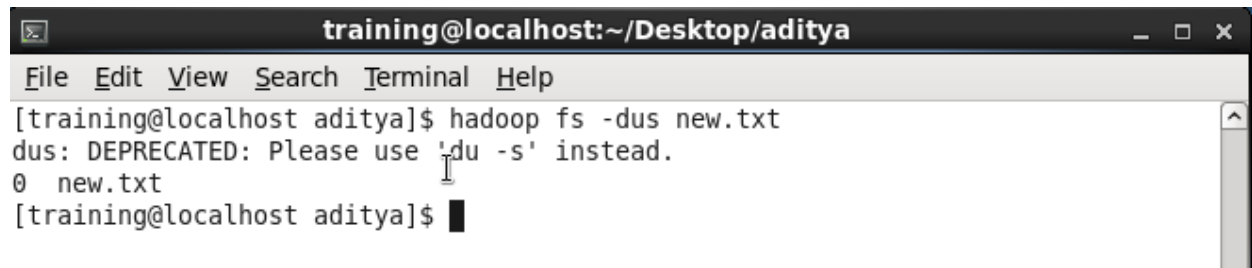
```
[training@localhost aditya]$ ls  
a.txt  new.txt  
[training@localhost aditya]$ hadoop fs -moveFromLocal new.txt  
[training@localhost aditya]$ hadoop fs -ls  
Found 3 items  
-rw-r--r--  1 training supergroup      25 2023-02-03 17:25 a.txt  
-rw-r--r--  1 training supergroup       0 2023-02-03 17:50 new.txt  
drwxr-xr-x  - training supergroup       0 2023-02-03 17:47 newfile.txt  
[training@localhost aditya]$
```

9.du: It will give the size of each file in directory.

A terminal window titled 'training@localhost:~/Desktop/aditya' with a menu bar (File, Edit, View, Search, Terminal, Help). The terminal shows the following commands and output:

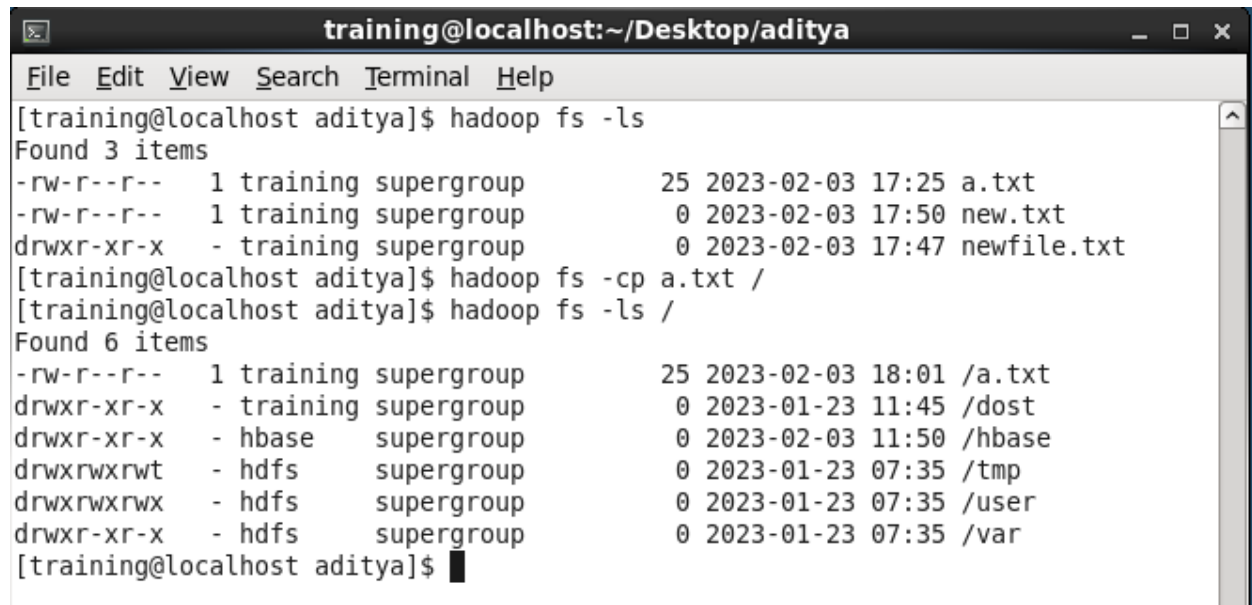
```
[training@localhost aditya]$ hadoop fs -du new.txt  
0 new.txt  
[training@localhost aditya]$
```

10.dus:: This command will give the total size of directory/file.



```
training@localhost:~/Desktop/aditya
File Edit View Search Terminal Help
[training@localhost aditya]$ hadoop fs -dus new.txt
dus: DEPRECATED: Please use 'du -s' instead.
0 new.txt
[training@localhost aditya]$
```

11.cp: This command is used to copy files within hdfs.



```
training@localhost:~/Desktop/aditya
File Edit View Search Terminal Help
[training@localhost aditya]$ hadoop fs -ls
Found 3 items
-rw-r--r-- 1 training supergroup 25 2023-02-03 17:25 a.txt
-rw-r--r-- 1 training supergroup 0 2023-02-03 17:50 new.txt
drwxr-xr-x - training supergroup 0 2023-02-03 17:47 newfile.txt
[training@localhost aditya]$ hadoop fs -cp a.txt /
[training@localhost aditya]$ hadoop fs -ls /
Found 6 items
-rw-r--r-- 1 training supergroup 25 2023-02-03 18:01 /a.txt
drwxr-xr-x - training supergroup 0 2023-01-23 11:45 /dost
drwxr-xr-x - hbase supergroup 0 2023-02-03 11:50 /hbase
drwxrwxrwt - hdfs supergroup 0 2023-01-23 07:35 /tmp
drwxrwxrwx - hdfs supergroup 0 2023-01-23 07:35 /user
drwxr-xr-x - hdfs supergroup 0 2023-01-23 07:35 /var
[training@localhost aditya]$
```

12.mv: This command is used to move files within hdfs.

```
training@localhost:~/Desktop/aditya
File Edit View Search Terminal Help
[training@localhost aditya]$ hadoop fs -ls
Found 3 items
-rw-r--r--  1 training supergroup      25 2023-02-03 17:25 a.txt
-rw-r--r--  1 training supergroup      0 2023-02-03 17:50 new.txt
drwxr-xr-x  - training supergroup      0 2023-02-03 17:47 newfile.txt
[training@localhost aditya]$ hadoop fs -mv new.txt /
[training@localhost aditya]$ hadoop fs -ls /
Found 6 items
drwxr-xr-x  - training supergroup      0 2023-01-23 11:45 /dost
drwxr-xr-x  - hbase supergroup          0 2023-02-03 11:50 /hbase
-rw-r--r--  1 training supergroup      0 2023-02-03 17:50 /new.txt
drwxrwxrwt  - hdfs supergroup          0 2023-01-23 07:35 /tmp
drwxrwxrwx  - hdfs supergroup          0 2023-01-23 07:35 /user
drwxr-xr-x  - hdfs supergroup          0 2023-01-23 07:35 /var
[training@localhost aditya]$ hadoop fs -ls
Found 2 items
-rw-r--r--  1 training supergroup      25 2023-02-03 17:25 a.txt
drwxr-xr-x  - training supergroup      0 2023-02-03 17:47 newfile.txt
[training@localhost aditya]$
```

13.stat: It will give the last modified time of directory or path.

```
training@localhost:~/Desktop/aditya
File Edit View Search Terminal Help
[training@localhost aditya]$ hadoop fs -stat a.txt
2023-02-03 22:25:40
[training@localhost aditya]$
```

14.hadoop fs: Give the list of all commands.

```
training@localhost:~/Desktop/aditya
File Edit View Search Terminal Help
[training@localhost aditya]$ hadoop fs
Usage: hadoop fs [generic options]
    [-cat [-ignoreCrc] <src> ...]
    [-chgrp [-R] GROUP PATH...]
    [-chmod [-R] <MODE[,MODE]... | OCTALMODE> PATH...]
    [-chown [-R] [OWNER][:[GROUP]] PATH...]
    [-copyFromLocal <localsrc> ... <dst>]
    [-copyToLocal [-ignoreCrc] [-crc] <src> ... <localdst>]
    [-count [-q] <path> ...]
    [-cp <src> ... <dst>]
    [-df [-h] [<path> ...]]
    [-du [-s] [-h] <path> ...]
    [-expunge]
    [-get [-ignoreCrc] [-crc] <src> ... <localdst>]
    [-getmerge [-nl] <src> <localdst>]
    [-help [cmd ...]]
    [-ls [-d] [-h] [-R] [<path> ...]]
    [-mkdir [-p] <path> ...]
    [-moveFromLocal <localsrc> ... <dst>]
    [-moveToLocal <src> <localdst>]
    [-mv <src> ... <dst>]
    [-put <localsrc> ... <dst>]
    [-rm [-f] [-r|-R] [-skipTrash] <src> ...]
    [-rmdir [--ignore-fail-on-non-empty] <dir> ...]
```

15.cat: To print file contents.

```
training@localhost:~/Desktop/aditya$ hadoop fs -cat source/new.txt
my name is aditya anal
i am coming form siwan
my barnch computer science
```

16.getmerge:to merge the two file in new file.

```
[training@localhost adityaWork]$ hadoop fs -getmerge source dest mergedfile.txt
[training@localhost adityaWork]$ cat mergedfile.txt
my name is aditya anal
i am coming form siwan
my barnch computer science
This is other test file for testing the getmerge command in haddop
we are learning HAdoop
my name is aditya anal
```
