

Assignment no.2 (File Management tasks in Hadoop)

1. Create a directory in HDFS at given path(s).

Usage: `hadoop fs -mkdir <paths>`

Example: `hadoop fs -mkdir /user/saurzcode/dir1 /user/saurzcode/dir2`

```
king_of_mz@PranavPC:~$ /home/hadoop_dev/hadoop2/bin/hadoop fs -mkdir /home/king_of_mz/dir1
king_of_mz@PranavPC:~$ ls
abc.c  a.out  Desktop  dir1  Documents  Downloads  file.c  hadoop-3.3.4.tar.gz  hadoop-3.3.4.tar.gz.sha512
```

2. List the contents of a directory.

Usage : `hadoop fs -ls <args>` Example: `hadoop fs -ls /user/saurzcode`

```
king_of_mz@PranavPC:~$ /home/hadoop_dev/hadoop2/bin/hadoop fs -ls /home/king_of_mz/
Found 38 items
-rw-r--r-- 1 king_of_mz king_of_mz 11691 2022-10-06 11:38 /home/king_of_mz/.bash_history
-rw-r--r-- 1 king_of_mz king_of_mz 220 2021-06-19 18:29 /home/king_of_mz/.bash_logout
-rw-r--r-- 1 king_of_mz king_of_mz 3771 2021-06-19 18:29 /home/king_of_mz/.bashrc
drwxr-xr-x - king_of_mz king_of_mz 4096 2022-06-07 00:33 /home/king_of_mz/.cache
drwxr-xr-x - king_of_mz king_of_mz 4096 2022-06-07 00:33 /home/king_of_mz/.config
drwxr-xr-x - king_of_mz king_of_mz 4096 2022-10-06 09:42 /home/king_of_mz/.gnupg
drwxr-xr-x - king_of_mz king_of_mz 4096 2021-06-19 18:37 /home/king_of_mz/.local
drwxr-xr-x - king_of_mz king_of_mz 4096 2021-06-19 18:46 /home/king_of_mz/.mozilla
-rw-r--r-- 1 king_of_mz king_of_mz 807 2021-06-19 18:29 /home/king_of_mz/.profile
-rw-r--r-- 1 king_of_mz king_of_mz 253 2021-08-31 14:27 /home/king_of_mz/.python_history
drwxr-xr-x - king_of_mz king_of_mz 4096 2021-06-19 18:39 /home/king_of_mz/.ssh
-rw-r--r-- 1 king_of_mz king_of_mz 0 2021-06-19 18:42 /home/king_of_mz/.sudo_as_admin_successful
drwxr-xr-x - king_of_mz king_of_mz 4096 2021-06-20 12:23 /home/king_of_mz/.thunderbird
-rw-r--r-- 1 king_of_mz king_of_mz 5 2022-10-06 09:07 /home/king_of_mz/.vboxclient-clipboard.pid
-rw-r--r-- 1 king_of_mz king_of_mz 5 2022-10-06 09:07 /home/king_of_mz/.vboxclient-display-svgx11.pid
-rw-r--r-- 1 king_of_mz king_of_mz 5 2022-10-06 09:07 /home/king_of_mz/.vboxclient-draganddrop.pid
-rw-r--r-- 1 king_of_mz king_of_mz 5 2022-10-06 09:07 /home/king_of_mz/.vboxclient-seamless.pid
-rw-r--r-- 1 king_of_mz king_of_mz 6073 2022-10-06 10:18 /home/king_of_mz/.viminfo
drwxr-xr-x - king_of_mz king_of_mz 4096 2022-10-06 10:00 /home/king_of_mz/Desktop
drwxr-xr-x - king_of_mz king_of_mz 4096 2021-06-19 18:37 /home/king_of_mz/Documents
drwxr-xr-x - king_of_mz king_of_mz 4096 2021-08-18 15:40 /home/king_of_mz/Downloads
drwxr-xr-x - king_of_mz king_of_mz 4096 2021-06-19 18:37 /home/king_of_mz/Music
drwxr-xr-x - king_of_mz king_of_mz 4096 2021-06-19 18:37 /home/king_of_mz/Pictures
drwxr-xr-x - king_of_mz king_of_mz 4096 2021-06-19 18:37 /home/king_of_mz/Public
drwxr-xr-x - king_of_mz king_of_mz 4096 2021-06-19 18:37 /home/king_of_mz/Templates
drwxr-xr-x - king_of_mz king_of_mz 4096 2021-06-19 18:37 /home/king_of_mz/Videos
-rwxrwxr-x 1 king_of_mz king_of_mz 16696 2022-09-07 20:16 /home/king_of_mz/a.out
-rw-rw-r-- 1 king_of_mz king_of_mz 78 2022-09-07 20:11 /home/king_of_mz/abc.c
drwxr-xr-x - king_of_mz king_of_mz 4096 2022-10-06 11:47 /home/king_of_mz/dir1
drwxrwxr-x - king_of_mz king_of_mz 4096 2022-09-15 22:39 /home/king_of_mz/file.c
-rw-rw-r-- 1 king_of_mz king_of_mz 695457782 2022-07-29 23:41 /home/king_of_mz/hadoop-3.3.4.tar.gz
-rw-rw-r-- 1 king_of_mz king_of_mz 160 2022-07-29 23:41 /home/king_of_mz/hadoop-3.3.4.tar.gz.sha512
-rw-rw-r-- 1 king_of_mz king_of_mz 0 2021-06-20 08:56 /home/king_of_mz/help
drwxrwxr-x - king_of_mz king_of_mz 4096 2022-10-06 11:12 /home/king_of_mz/input
-rw-rw-r-- 1 king_of_mz king_of_mz 71 2022-09-07 20:16 /home/king_of_mz/pqr.c
drwxr-xr-x - king_of_mz king_of_mz 4096 2021-08-18 15:38 /home/king_of_mz/snap
drwxrwxr-x - king_of_mz king_of_mz 4096 2022-10-05 21:01 /home/king_of_mz/vdxt
-rw-rw-r-- 1 king_of_mz king_of_mz 68 2022-09-07 20:13 /home/king_of_mz/xyz.c
```

3. Upload and download a file in HDFS.

a) `hadoop fs -put`:

Copy single src file, or multiple src files from local file system to the Hadoop data file system

Usage: `hadoop fs -put <localsrc> ... <HDFS_dest_Path>`

Example: `hadoop fs -put /home/saurzcode/Samplefile.txt /user/ saurzcode/dir3/`

```
king_of_mz@PranavPC:~$ /home/hadoop_dev/hadoop2/bin/hadoop fs -put /home/king_of_mz/abc.c /home/king_of_mz/dir1
king_of_mz@PranavPC:~$ cd dir1
king_of_mz@PranavPC:~/dir1$ ls
abc.c
```

b) `hadoop fs -get`:

Copies/Downloads files to the local file system

Usage: `hadoop fs -get <hdfs_src> <localdst>`

Example: `hadoop fs -get /user/saurzcode/dir3/Samplefile.txt /home/`

```
king_of_mz@PranavPC:~/dir1$ /home/hadoop_dev/hadoop2/bin/hadoop fs -get /home/king_of_mz/dir1/temp.txt /home/king_of_mz/
king_of_mz@PranavPC:~/dir1$ ls
abc.c  pqr.c  temp.txt
king_of_mz@PranavPC:~/dir1$ cd
king_of_mz@PranavPC:~$ ls
abc.c  a.out  Desktop  dir1  Documents  Downloads  file.c  hadoop-3.3.4.tar.gz  hadoop-3.3.4.tar.gz.sha512  help  input  Music  Pictures  pqr.c  Public  snap  Templates  temp.txt
king_of_mz@PranavPC:~$
```

4. See contents of a file

Same as unix cat command:

Usage: `hadoop fs -cat <path[filename]>`

Example: `hadoop fs -cat /user/saurzcode/dir1/abc.txt`

```
king_of_mz@PranavPC:~$ /home/hadoop_dev/hadoop2/bin/hadoop fs -cat /home/king_of_mz/dir1/abc.c
#include <stdio.h>
void main()
{
printf("the world is full of adventures");
}
```

5. Copy a file from source to destination

This command allows multiple sources as well in which case the destination must be a directory.

Usage: `hadoop fs -cp <source> <dest>`

Example: `hadoop fs -cp /user/saurzcode/dir1/abc.txt /user/saurzcode/dir2`

```
king_of_mz@PranavPC:~$ /home/hadoop_dev/hadoop2/bin/hadoop fs -cp /home/king_of_mz/dir1/temp.txt /home/king_of_mz/snap
king_of_mz@PranavPC:~$ cd snap
king_of_mz@PranavPC:~/snap$ ls
snap-store  temp.txt
```

6. Copy a file from/To Local file system to HDFS

a) copyFromLocal

Usage: `hadoop fs -copyFromLocal <localsrc> URI`

Example: `hadoop fs -copyFromLocal /home/saurzcode/abc.txt /user/saurzcode/abc.txt`

```
king_of_mz@PranavPC:~/snap$ /home/hadoop_dev/hadoop2/bin/hadoop fs -copyFromLocal /home/king_of_mz/xyz.c /home/king_of_mz/dir1/xyz.c
king_of_mz@PranavPC:~/snap$ cd
king_of_mz@PranavPC:~$ cd dir1
king_of_mz@PranavPC:~/dir1$ ls
abc.c  pqr.c  temp.txt  xyz.c
king_of_mz@PranavPC:~/dir1$
```

7. Move file from source to destination.

Note:- Moving files across filesystem is not permitted.

Usage : `hadoop fs -mv <src> <dest>`

Example: `hadoop fs -mv /user/saurzcode/dir1/abc.txt /user/saurzcode/dir2`

```
king_of_mz@PranavPC:~$ /home/hadoop_dev/hadoop2/bin/hadoop fs -mv /home/king_of_mz/dir1/temp.txt /home/king_of_mz/dir2
king_of_mz@PranavPC:~$ cd dir1
king_of_mz@PranavPC:~/dir1$ ls
abc.c  pqr.c  xyz.c
king_of_mz@PranavPC:~/dir1$ cd
king_of_mz@PranavPC:~$ cd dir2
king_of_mz@PranavPC:~/dir2$ ls
temp.txt
```

8. Remove a file or directory in HDFS.

Remove files specified as argument. Deletes directory only when it is empty

Usage : `hadoop fs -rm <arg>`

Example: `hadoop fs -rm /user/saurzcode/dir1/abc.txt`

```
king_of_mz@PranavPC:~/dir1$ ls
abc.c  pqr.c  xyz.c
king_of_mz@PranavPC:~/dir1$ /home/hadoop_dev/hadoop2/bin/hadoop fs -rm /home/king_of_mz/dir1/abc.c
2022-10-06 12:26:02,911 INFO Configuration.deprecation: io.bytes.per.checksum is deprecated. Instead use
Deleted /home/king_of_mz/dir1/abc.c
king_of_mz@PranavPC:~/dir1$ ls
pqr.c  xyz.c
```

9. Display last few lines of a file.

Similar to tail command in Unix.

Usage : `hadoop fs -tail <path[filename]>`

Example: `hadoop fs -tail /user/saurzcode/dir1/abc.txt`

```
king_of_mz@PranavPC:~/dir1$ /home/hadoop_dev/hadoop2/bin/hadoop fs -tail /home/king_of_mz/dir1/convo.c
```

```
scanf("%c",address);
printf("\nAshish:ohh I am from %s",address);

printf("\nVishal:Btw you are from which div?");
printf("\nAshish:I am from div ");
scanf("%c",div);
printf("Vishal:Great we are in same division");
}
```

10. Display the aggregate length of a file.

Usage : `hadoop fs -du <path>`

Example: `hadoop fs -du /user/saurzcode/dir1/abc.txt`

```
king_of_mz@PranavPC:~/dir1$ /home/hadoop_dev/hadoop2/bin/hadoop fs -du /home/king_of_mz/dir1/convo.c
526 526 /home/king_of_mz/dir1/convo.c
```