

Meerza Ahmed | DSCI 551

*Due 2/02/24 @ 11:59 PM

Q: [10 pts] For each of the following HDFS dfs commands, submit a screenshot showing an example of the command and also explain what the command does.

1. 'ls'

- The 'ls' command is used to list all files in the current directory, normally the command excludes hidden files but you can use -a as an argument to display hidden files such as 'bash'.

```
ubuntu@ip-172-31-46-240:~/hadoop-3.3.6$ bin/hdfs dfs -ls /
Found 1 items
drwxr-xr-x - ubuntu supergroup          0 2024-01-28 01:38 /home
ubuntu@ip-172-31-46-240:~/hadoop-3.3.6$
```

2. 'mkdir'

- The 'mkdir' command creates a directory in the current directory, normally you do mkdir + "/dir name"

```
ubuntu@ip-172-31-46-240:~/hadoop-3.3.6$ bin/hdfs dfs -mkdir /example.txt
ubuntu@ip-172-31-46-240:~/hadoop-3.3.6$ bin/hdfs dfs -ls /
Found 2 items
drwxr-xr-x - ubuntu supergroup          0 2024-01-28 01:39 /example.txt
drwxr-xr-x - ubuntu supergroup          0 2024-01-28 01:38 /home
ubuntu@ip-172-31-46-240:~/hadoop-3.3.6$
```

3. 'put'

- The 'put' command is used to copy files from the local machine to Hadoop

```
ubuntu@ip-172-31-46-240:~/hadoop-3.3.6$ bin/hdfs dfs -put README.txt /home/readme
ubuntu@ip-172-31-46-240:~/hadoop-3.3.6$ bin/hdfs dfs -ls /home
Found 1 items
-rw-r--r-- 1 ubuntu supergroup          175 2024-01-28 01:49 /home/readme
```

4. 'get'

- The 'get' command is used to copy a file from Hadoop to the local machine

```
ubuntu@ip-172-31-46-240:~/hadoop-3.3.6$ bin/hdfs dfs -get /home testREADME.txt
ubuntu@ip-172-31-46-240:~/hadoop-3.3.6$ ls
LICENSE-binary LICENSE.txt NOTICE-binary NOTICE.txt README.txt bin etc include lib libexec licenses-binary logs output sbin share testREADME.txt
ubuntu@ip-172-31-46-240:~/hadoop-3.3.6$
```

5. 'rm'

- The 'rm' command is used to delete a file

```
ubuntu@ip-172-31-46-240:~/hadoop-3.3.6$ bin/hdfs dfs -ls /home
Found 2 items
-rw-r--r-- 1 ubuntu supergroup          175 2024-01-28 01:49 /home/readme
-rw-r--r-- 1 ubuntu supergroup          175 2024-01-28 01:54 /home/test
ubuntu@ip-172-31-46-240:~/hadoop-3.3.6$ bin/hdfs dfs -rm /home/test
Deleted /home/test
ubuntu@ip-172-31-46-240:~/hadoop-3.3.6$ bin/hdfs dfs -ls /home
Found 1 items
-rw-r--r-- 1 ubuntu supergroup          175 2024-01-28 01:49 /home/readme
ubuntu@ip-172-31-46-240:~/hadoop-3.3.6$
```

6. 'rmdir'

- The 'rmdir' command this is similar to the rm command but is used to delete directories

```
ubuntu@ip-172-31-46-240:~/hadoop-3.3.6$ bin/hdfs dfs -rmdir /home
ubuntu@ip-172-31-46-240:~/hadoop-3.3.6$ bin/hdfs dfs -ls /
ubuntu@ip-172-31-46-240:~/hadoop-3.3.6$
```

7. 'cp'

- The 'cp' command is used to copy a file from a source and into a destination

```
ubuntu@ip-172-31-46-240:~/hadoop-3.3.6$ bin/hdfs dfs -cp /home/readme /home/test
ubuntu@ip-172-31-46-240:~/hadoop-3.3.6$ bin/hdfs dfs -ls /
Found 1 items
drwxr-xr-x  - ubuntu supergroup          0 2024-01-28 01:54 /home
ubuntu@ip-172-31-46-240:~/hadoop-3.3.6$ bin/hdfs dfs -ls /home
Found 2 items
-rw-r--r--  1 ubuntu supergroup        175 2024-01-28 01:49 /home/readme
-rw-r--r--  1 ubuntu supergroup        175 2024-01-28 01:54 /home/test
ubuntu@ip-172-31-46-240:~/hadoop-3.3.6$
```

8. 'cat'

- The 'cat' command prints the contents of a file it means concatenate, you do cat + file path to call it

```
ubuntu@ip-172-31-46-240:~/hadoop-3.3.6$ bin/hdfs dfs -cat /home/readme/README.txt
cat: "/home/readme/README.txt": No such file or directory
ubuntu@ip-172-31-46-240:~/hadoop-3.3.6$ bin/hdfs dfs -cat /home/readme
For the latest information about Hadoop, please visit our website at:

    http://hadoop.apache.org/

and our wiki, at:

    https://wiki.apache.org/confluence/display/HADOOP/
ubuntu@ip-172-31-46-240:~/hadoop-3.3.6$
```