Hadoop HDFS Operations using CLI

Interacting with HDFS using command line interface.

All the important commands can be accessed using

1. hdfs dfs –help
2. To check summary of filesystem

hdfs fsck /

```
hadoop@ubuntu:~$ hdfs fsck /
Connecting to namenode via http://localhost:9870/fsck?ugi=hadoop&
FSCK started by hadoop (auth:SIMPLE) from /127.0.0.1 for path / a

Status: HEALTHY
 Number of data-nodes:   1
 Number of racks:                     1
 Total dirs:                          3
 Total symlinks:                      0

Replicated Blocks:
 Total size:      0 B
 Total files:     0
 Total blocks (validated):       0
 Minimally replicated blocks:    0
 Over-replicated blocks:         0
 Under-replicated blocks:        0
 Mis-replicated blocks:          0
 Default replication factor:     1
 Average block replication:      0.0
 Missing blocks:                 0
 Corrupt blocks:                 0
 Missing replicas:               0

Erasure Coded Block Groups:
 Total size:      0 B
 Total files:     0
 Total block groups (validated):        0
 Minimally erasure-coded block groups:  0
 Over-erasure-coded block groups:       0
 Under-erasure-coded block groups:      0
 Unsatisfactory placement block groups: 0
 Average block group size:       0.0
 Missing block groups:           0
 Corrupt block groups:           0
 Missing internal blocks:        0
```

3. To create and list directories

```
hadoop@ubuntu:~$ hdfs dfs -mkdir -p ~/test
hadoop@ubuntu:~$ hdfs dfs -ls ~/
Found 1 items
drwxr-xr-x   - hadoop supergroup          0 2020-12-08 16:54 /home/hadoop/test
hadoop@ubuntu:~$ 
```

4. To remove directory

```
hadoop@ubuntu:~$ hdfs dfs -rm -r ~/test/
Deleted /home/hadoop/test
hadoop@ubuntu:~$ 
```

5. To create empty file

```
hadoop@ubuntu:~$ clear

hadoop@ubuntu:~$ hdfs dfs -touchz ~/sample.txt
hadoop@ubuntu:~$ hdfs dfs -ls ~/
Found 1 items
-rw-r--r--   1 hadoop supergroup          0 2020-12-08 17:04 /home/hadoop/sample.txt
hadoop@ubuntu:~$ 
```

6. To append text into file

```
hadoop@ubuntu:~$ hdfs dfs -appendToFile - ~/sample.txt
I am writing to this file and excited to learn about
how hdfs file system works.hadoop@ubuntu:~$
```

7. To get size and read from file.

```
hadoop@ubuntu:~$ hdfs dfs -du -s ~/sample.txt
80  80  /home/hadoop/sample.txt
hadoop@ubuntu:~$ hdfs dfs -cat ~/sample.txt
I am writing to this file and excited to learn about
how hdfs file system works.hadoop@ubuntu:~$ 
```

8. Copying file from local system into hdfs directory

```
hadoop@ubuntu:~$ hdfs dfs -copyFromLocal host.txt ~/
hadoop@ubuntu:~$ hdfs dfs -ls ~/
Found 2 items
-rw-r--r--   1 hadoop supergroup         40 2020-12-08 18:40 /home/hadoop/host.txt
-rw-r--r--   1 hadoop supergroup         80 2020-12-08 18:33 /home/hadoop/sample.txt
hadoop@ubuntu:~$ vim readme.md
hadoop@ubuntu:~$ hdfs dfs -ls ~/
Found 2 items
-rw-r--r--   1 hadoop supergroup         40 2020-12-08 18:40 /home/hadoop/host.txt
-rw-r--r--   1 hadoop supergroup         80 2020-12-08 18:33 /home/hadoop/sample.txt
hadoop@ubuntu:~$ hdfs dfs -put readme.md ~/
hadoop@ubuntu:~$ hdfs dfs -ls ~/
Found 3 items
-rw-r--r--   1 hadoop supergroup         40 2020-12-08 18:40 /home/hadoop/host.txt
-rw-r--r--   1 hadoop supergroup         20 2020-12-08 18:43 /home/hadoop/readme.md
-rw-r--r--   1 hadoop supergroup         80 2020-12-08 18:33 /home/hadoop/sample.txt
```

9. Copy from hdfs to local directory

```
hadoop@ubuntu:~$ hdfs dfs -copyToLocal ~/readme.md ~/
copyToLocal: `/home/hadoop/readme.md': File exists
hadoop@ubuntu:~$ hdfs dfs -copyToLocal ~/readme.md ~/sample/
hadoop@ubuntu:~$ cd sample
hadoop@ubuntu:~/sample$ ls
readme.md
hadoop@ubuntu:~/sample$ hdfs dfs -get ~/readme.md ~/sample/
get: `/home/hadoop/sample/readme.md': File exists
hadoop@ubuntu:~/sample$ hdfs dfs -get ~/host.txt ~/sample/
hadoop@ubuntu:~/sample$ ls
host.txt   readme.md
```

10. To move file

```
hadoop@ubuntu:~/sample$ hdfs dfs -mv ~/host.txt ~/test/
hadoop@ubuntu:~/sample$ hdfs dfs -ls ~/test/
Found 1 items
-rw-r--r--   1 hadoop supergroup         40 2020-12-08 18:40 /home/hadoop/test/host.txt
```