

**HARI PRASATH S**

**225229110**

## **Exercise 02 : File management tasks in Hadoop**

```
[cloudera@quickstart ~]$ cd
```

**1. Find the HDFS Command to check the health of the Hadoop file system.**

```
[cloudera@quickstart ~]$ hdfs fsck /
```

Connecting to namenode via http://quickstart.cloudera:50070/fsck?ugi=cloudera&path=%2F

FSCK started by cloudera (auth:SIMPLE) from /127.0.0.1 for path / at Thu Jul 13 21:09:53 PDT 2023

```
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
```

.....Status: HEALTHY

Total size: 858123172 B (Total open files size: 166 B)

Total dirs: 79

Total files: 931

Total symlinks: 0 (Files currently being written: 3)

Total blocks (validated): 929 (avg. block size 923706 B) (Total open file blocks (not validated): 2)

Minimally replicated blocks: 929 (100.0 %)

Over-replicated blocks: 0 (0.0 %)

Under-replicated blocks: 0 (0.0 %)

Mis-replicated blocks: 0 (0.0 %)

Default replication factor: 1

Average block replication: 1.0  
Corrupt blocks: 0  
Missing replicas: 0 (0.0 %)  
Number of data-nodes: 1  
Number of racks: 1  
FSCK ended at Thu Jul 13 21:09:53 PDT 2023 in 132 milliseconds

The filesystem under path '/' is HEALTHY

## 2. Attach the output screen shot for HDFS Command to display the list of Files and Directories in HDFS.

```
[cloudera@quickstart ~]$ hdfs dfs -ls /
```

Found 6 items

```
drwxrwxrwx - hdfs supergroup    0 2017-07-19 05:34 /benchmarks
drwxr-xr-x - hbase supergroup    0 2023-07-13 20:19 /hbase
drwxr-xr-x - solr solr          0 2017-07-19 05:37 /solr
drwxrwxrwt - hdfs supergroup    0 2023-07-13 20:19 /tmp
drwxr-xr-x - hdfs supergroup    0 2017-07-19 05:36 /user
drwxr-xr-x - hdfs supergroup    0 2017-07-19 05:36 /var
```

## 3. Create HDFS Directory with the name 'temp'

```
[cloudera@quickstart ~]$ hdfs dfs -mkdir /temp/
```

```
[cloudera@quickstart ~]$ hdfs dfs -ls /
```

Found 7 items

```
drwxrwxrwx - hdfs supergroup    0 2017-07-19 05:34 /benchmarks
drwxr-xr-x - hbase supergroup    0 2023-07-13 20:19 /hbase
drwxr-xr-x - solr solr          0 2017-07-19 05:37 /solr
drwxr-xr-x - cloudera supergroup 0 2023-07-13 21:10 /temp
drwxrwxrwt - hdfs supergroup    0 2023-07-13 20:19 /tmp
drwxr-xr-x - hdfs supergroup    0 2017-07-19 05:36 /user
```

```
drwxr-xr-x - hdfs supergroup 0 2017-07-19 05:36 /var
```

**4. Create file in HDFS with 'touchz' and show the size of the created file.**

```
[cloudera@quickstart ~]$
```

```
hdfs dfs -touchz /temp/sample00
```

**5. Find the HDFS command for show file size in HDFS folder**

```
[cloudera@quickstart ~]$ hdfs dfs -du /temp/sample00
```

```
0 0 /temp/sample00
```

```
[cloudera@quickstart ~]$ ^C
```

**6. Create file in local system using Linux command and display the content in that created file.**

```
[cloudera@quickstart ~]$ cat > te
```

```
welcome BHC
```

```
hello ds!^C
```

```
[cloudera@quickstart ~]$ cat te
```

```
welcome BHC
```

```
[cloudera@quickstart ~]$
```

```
[cloudera@quickstart ~]$ ls
```

```
cloudera-manager Documents enterprise-deployment.json lib Pictures Templates
```

```
cm_api.py Downloads express-deployment.json Music Public Videos
```

```
Desktop eclipse kerberos parcels te workspace
```

**7. Find HDFS command for copy locally created file into HDFS file system**

```
[cloudera@quickstart ~]$ hdfs dfs -ls /temp/
```

```
Found 1 items
```

```
-rw-r--r-- 1 cloudera supergroup 0 2023-07-13 21:11 /temp/sample00
```

```
[cloudera@quickstart ~]$ hdfs dfs -put temp /temp/
```

–put: Unknown command

**[cloudera@quickstart ~]\$ ls**

cloudera-manager Documents enterprise-deployment.json lib Pictures Templates  
cm\_api.py Downloads express-deployment.json Music Public Videos  
Desktop eclipse kerberos parcels te workspace

**[cloudera@quickstart ~]\$ hdfs dfs -ls /temp/**

Found 1 items

-rw-r--r-- 1 cloudera supergroup 0 2023-07-13 21:11 /temp/sample00

**[cloudera@quickstart ~]\$ hdfs dfs –put temp /temp/**

–put: Unknown command

**[cloudera@quickstart ~]\$ hadoop fs -ls**

**[cloudera@quickstart ~]\$ ls**

cloudera-manager Documents enterprise-deployment.json lib Pictures Templates  
cm\_api.py Downloads express-deployment.json Music Public Videos  
Desktop eclipse kerberos parcels te workspace

## **8. Copy locally created file into HDFS file system using ‘put’ command**

**[cloudera@quickstart ~]\$ ~]\$ hdfs dfs -put te**

bash: ~]\$: command not found

**[cloudera@quickstart ~]\$ ~]\$ hadoop fs -ls**

bash: ~]\$: command not found

**[cloudera@quickstart ~]\$ ls**

cloudera-manager Documents enterprise-deployment.json lib Pictures Templates  
cm\_api.py Downloads express-deployment.json Music Public Videos  
Desktop eclipse kerberos parcels te workspace

## **9. Remove Local file using Linux command**

**[cloudera@quickstart ~]\$ rm te**

rm: remove regular file `te'?

**[cloudera@quickstart ~]\$ ls**

```
cloudera-manager Documents enterprise-deployment.json lib Pictures Templates
cm_api.py Downloads express-deployment.json Music Public Videos
Desktop eclipse kerberos parcels te workspace
```

**[cloudera@quickstart ~]\$ rm te**

rm: remove regular file `te'? yes

**[cloudera@quickstart ~]\$ ls**

```
cloudera-manager Documents enterprise-deployment.json lib Pictures Videos
cm_api.py Downloads express-deployment.json Music Public workspace
Desktop eclipse kerberos parcels Templates
```

#### **10. Copy HDFS file to Local File System using copyToLocal command**

**[cloudera@quickstart ~]\$ hdfs dfs -copyToLocal /temp /home/cloudera/**

#### **11. Find another way to copy HDFS files into Local File System.**

**[cloudera@quickstart ~]\$ hadoop fs -get /temp /home/cloudera/**

get: `/home/cloudera/temp/sample00': File exists

#### **12. Count number of files in particular HDFS directory using 'count' command**

**[cloudera@quickstart ~]\$ hdfs dfs -count /temp**

```
1      1      0 /temp
```

#### **13. Remove the entire directory and all if its content from HDFS**

**[cloudera@quickstart ~]\$ hdfs dfs -rm -r temp**

rm: `temp': No such file or directory

**[cloudera@quickstart ~]\$ hdfs dfs -du /user**

```
12     12     /user/cloudera
```

```
0      0     /user/history
```

```
0    0    /user/hive
0    0    /user/hue
0    0    /user/jenkins
858119738 858119738 /user/oozie
0    0    /user/root
0    0    /user/spark
```

```
[cloudera@quickstart ~]$ hdfs dfs -du /user/hive
```

```
0 0 /user/hive/warehouse
```

**14. Copy HDFS file from one folder to another folder using 'cp' command**

```
[cloudera@quickstart ~]$ hdfs dfs -cp /user/hive/warehouse /user/hue
```

**15. Write the HDFS command to move files from source to destination.**

```
[cloudera@quickstart ~]$ hdfs dfs -du /user/hive
```

```
0 0 /user/hive/warehouse
```

```
[cloudera@quickstart ~]$ hdfs dfs -du /user/hue
```

```
0 0 /user/hue/warehouse
```

```
[cloudera@quickstart ~]$
```