

Practical2

To implement Various Hadoop HDFS Commands

1) Hadoop Version

```
[cloudera@quickstart ~]$ hadoop version
Hadoop 2.6.0-cdh5.4.2
Subversion http://github.com/cloudera/hadoop -r
15b703c8725733b7b2813d2325659eb7d57e7a3f
Compiled by jenkins on 2015-05-20T00:03Z
Compiled with protoc 2.5.0
From source with checksum de74f1adb3744f8ee85d9a5b98f90d
This command was run using /usr/jars/hadoop-common-2.6.0-cdh5.4.2.jar
```

2) LS Command

->hdfs dfs -ls /

```
[cloudera@quickstart ~]$ hadoop fs -ls /
Found 5 items
drwxr-xr-x - hbase supergroup          0 2022-02-07 21:07 /hbase
drwxr-xr-x - solr solr                  0 2015-06-09 03:38 /solr
drwxrwxrwx - hdfs supergroup            0 2022-02-07 21:09 /tmp
drwxr-xr-x - hdfs supergroup            0 2015-06-09 03:38 /user
drwxr-xr-x - hdfs supergroup            0 2015-06-09 03:36 /var
```

2) MKDIR Command

```
[cloudera@quickstart ~]$ hdfs dfs -mkdir /rjc2122
[cloudera@quickstart Desktop]$ hadoop fs -ls /
Found 6 items
drwxr-xr-x - hbase supergroup          0 2022-02-07 21:07 /hbase
drwxr-xr-x - cloudera supergroup        0 2022-02-14 18:49 /rjc2122
drwxr-xr-x - solr solr                  0 2015-06-09 03:38 /solr
drwxrwxrwx - hdfs supergroup            0 2022-02-07 21:09 /tmp
drwxr-xr-x - hdfs supergroup            0 2015-06-09 03:38 /user
drwxr-xr-x - hdfs supergroup            0 2015-06-09 03:36 /var
[cloudera@quickstart Desktop]$ hdfs dfs -ls /
Found 6 items
drwxr-xr-x - hbase supergroup          0 2022-02-07 21:07 /hbase
drwxr-xr-x - cloudera supergroup        0 2022-02-14 18:49 /rjc2122
drwxr-xr-x - solr solr                  0 2015-06-09 03:38 /solr
drwxrwxrwx - hdfs supergroup            0 2022-02-07 21:09 /tmp
drwxr-xr-x - hdfs supergroup            0 2015-06-09 03:38 /user
drwxr-xr-x - hdfs supergroup            0 2015-06-09 03:36 /var
```

3) ->Hadoop dfsadmin -safemode leave

Use this command if you are getting "you are in safe mode "

4) copyFromLocal Command

```
[cloudera@quickstart Desktop]$ hdfs dfs -copyFromLocal
/home/cloudera/Desktop/file01 /rjc2122
[cloudera@quickstart Desktop]$ hdfs dfs -ls /rjc2122
Found 1 items
-rw-r--r--    1 cloudera supergroup          38 2022-02-14 19:05
/rjc2122/file01
```

5) If getting any error due to permissions

Use-> **export HADOOP_USER_NAME=hdfs**

6) Put Command

```
[cloudera@quickstart Desktop]$ hdfs dfs -put file02 /rjc2122
[cloudera@quickstart Desktop]$ hdfs dfs -ls /rjc2122
Found 2 items
-rw-r--r--    1 cloudera supergroup          38 2022-02-14 19:05
/rjc2122/file01
-rw-r--r--    1 cloudera supergroup          32 2022-02-14 19:09
/rjc2122/file02
```

7) copyToLocal Command

```
[cloudera@quickstart Desktop]$ hdfs dfs -copyToLocal /rjc2122/file01
/home/cloudera/Desktop/rjclocal
[cloudera@quickstart Desktop]$ ls /home/cloudera/Desktop/rjclocal
file01
```

8) CAT Command

```
[cloudera@quickstart Desktop]$ hdfs dfs -cat /rjc2122/file04
hello world
```

9) Cp Command

```
[cloudera@quickstart Desktop]$ hdfs dfs -mkdir /rjcnew
[cloudera@quickstart Desktop]$ hdfs dfs -cp /rjc2122/file01 /rjcnew
[cloudera@quickstart Desktop]$ hdfs dfs -ls /
Found 7 items
drwxr-xr-x    - hbase      supergroup          0 2022-02-07 21:07 /hbase
drwxr-xr-x    - cloudera   supergroup          0 2022-02-14 19:09 /rjc2122
drwxr-xr-x    - cloudera   supergroup          0 2022-02-14 19:37 /rjcnew
drwxr-xr-x    - solr       solr                0 2015-06-09 03:38 /solr
drwxrwxrwx    - hdfs       supergroup          0 2022-02-07 21:09 /tmp
drwxr-xr-x    - hdfs       supergroup          0 2015-06-09 03:38 /user
drwxr-xr-x    - hdfs       supergroup          0 2015-06-09 03:36 /var
[cloudera@quickstart Desktop]$ hdfs dfs -ls /rjcnew
Found 1 items
```

```
-rw-r--r-- 1 cloudera supergroup 38 2022-02-14 19:37
/rjcnew/file01
```

10) MV Command

```
[cloudera@quickstart Desktop]$ hdfs dfs -mv /rjc2122/file02 /rjcnew
[cloudera@quickstart Desktop]$ hdfs dfs -ls /rjcnew
Found 2 items
-rw-r--r-- 1 cloudera supergroup 38 2022-02-14 19:37
/rjcnew/file01
-rw-r--r-- 1 cloudera supergroup 32 2022-02-14 19:09
/rjcnew/file02
```

11) RM Command

```
[cloudera@quickstart Desktop]$ hdfs dfs -rm /rjcnew/file01
22/02/14 19:58:14 INFO fs.TrashPolicyDefault: Namenode trash
configuration: Deletion interval = 0 minutes, Emptier interval = 0
minutes.
Deleted /rjcnew/file01
[cloudera@quickstart Desktop]$ hdfs dfs -ls /rjcnew
Found 1 items
-rw-r--r-- 1 cloudera supergroup 32 2022-02-14 19:09
/rjcnew/file02
```

hdfs dfs -rm -r

```
[cloudera@quickstart Desktop]$ hdfs dfs -rm -r /rjcnew
22/02/14 20:00:37 INFO fs.TrashPolicyDefault: Namenode trash
configuration: Deletion interval = 0 minutes, Emptier interval = 0
minutes.
Deleted /rjcnew
[cloudera@quickstart Desktop]$ hdfs dfs -ls /
Found 6 items
drwxr-xr-x - hbase supergroup 0 2022-02-07 21:07 /hbase
drwxr-xr-x - cloudera supergroup 0 2022-02-14 19:57 /rjc2122
drwxr-xr-x - solr solr 0 2015-06-09 03:38 /solr
drwxrwxrwx - hdfs supergroup 0 2022-02-07 21:09 /tmp
drwxr-xr-x - hdfs supergroup 0 2015-06-09 03:38 /user
drwxr-xr-x - hdfs supergroup 0 2015-06-09 03:36 /var
```

12) MoveFromLocal Command

```
[cloudera@quickstart Desktop]$ hdfs dfs -moveFromLocal
/home/cloudera/Desktop/file03 /rjc2122
[cloudera@quickstart Desktop]$ hdfs dfs -ls /rjc2122
Found 2 items
-rw-r--r-- 1 cloudera supergroup 38 2022-02-14 19:05
/rjc2122/file01
-rw-r--r-- 1 cloudera supergroup 15 2022-02-14 20:05
/rjc2122/file03
```

13) MoveToLocal Command

```
[cloudera@quickstart Desktop]$ hdfs dfs -moveToLocal /rjc2122/file03
/home/cloudera/Desktop/rjclocal
moveToLocal: Option '-moveToLocal' is not implemented yet.
```

14) Tail Command

```
[cloudera@quickstart Desktop]$ hdfs dfs -tail /rjc2122/file01
i am a student of msdsai batch 21-22
```

```
[cloudera@quickstart Desktop]$ tail
/home/cloudera/Desktop/rjclocal/file01
i am a student of msdsai batch 21-22
```

15) Expunge Command

```
[cloudera@quickstart Desktop]$ hdfs dfs -expunge
22/02/14 20:23:09 INFO fs.TrashPolicyDefault: Namenode trash
configuration: Deletion interval = 0 minutes, Emptier interval = 0
minutes.
```

16) Replication Command

```
[cloudera@quickstart Desktop]$ hdfs dfs -setrep 3/rjc2122/file01
-setrep: Not enough arguments: expected 2 but got 1
Usage: hadoop fs [generic options] -setrep [-R] [-w] <rep> <path> ...
[cloudera@quickstart Desktop]$ hdfs dfs -setrep 3 /rjc2122/file01
Replication 3 set: /rjc2122/file01
```

Browsing HDFS - Mozilla Firefox

Cloudera Live : Welcom... * Index of file:/// * Browsing HDFS *

localhost:50070/explorer.html#/rjc2122

Cloudera Hue Hadoop HBase Impala Spark Solr Oozie Cloudera Manager Getting Started

Hadoop Overview Datanodes Snapshot Startup Progress Utilities

Browse Directory

/rjc2122 Go!

Permission	Owner	Group	Size	Replication	Block Size	Name
-rw-r--r--	cloudera	supergroup	38 B	3	128 MB	file01
-rw-r--r--	cloudera	supergroup	15 B	1	128 MB	file03

Hadoop, 2014.

17) DU Command

```
[cloudera@quickstart Desktop]$ hdfs dfs -du /rjc2122
38 114 /rjc2122/file01
15 15 /rjc2122/file03
[cloudera@quickstart Desktop]$ hdfs dfs -du /rjc2122/file01
38 114 /rjc2122/file01
```

18) Df command

```
[cloudera@quickstart Desktop]$ hdfs dfs -df
Filesystem                Size      Used    Available
Use%
hdfs://quickstart.cloudera:8020  58665738240  394661888  48055541760
1%
```

->hdfs dfs -df -h

With h parameter the information is human readable

```
[cloudera@quickstart Desktop]$ hdfs dfs -df -h
Filesystem                Size      Used    Available  Use%
hdfs://quickstart.cloudera:8020  54.6 G  376.4 M    44.8 G      1%
[cloudera@quickstart Desktop]$ hdfs dfs -fsck /rjc2122
-fsck: Unknown command
[cloudera@quickstart Desktop]$ hdfs dfs fsck /rjc2122
fsck: Unknown command
```

19) Fsck command

```
[cloudera@quickstart Desktop]$ hdfs fsck /rjc2122
Connecting to namenode via http://quickstart.cloudera:50070
FSCK started by cloudera (auth:SIMPLE) from /127.0.0.1 for path /rjc2122
at Mon Feb 14 20:48:04 PST 2022
.
/rjc2122/file01: Under replicated BP-989008105-127.0.0.1-
1433846136903:blk_1073742209_1387. Target Replicas is 3 but found 1
replica(s).
.Status: HEALTHY
Total size:      53 B
Total dirs:      1
Total files:      2
Total symlinks:   0
Total blocks (validated):  2 (avg. block size 26 B)
Minimally replicated blocks:  2 (100.0 %)
Over-replicated blocks:      0 (0.0 %)
Under-replicated blocks:      1 (50.0 %)
Mis-replicated blocks:        0 (0.0 %)
Default replication factor: 1
Average block replication: 1.0
Corrupt blocks: 0
Missing replicas: 2 (50.0 %)
Number of data-nodes: 1
Number of racks: 1
FSCK ended at Mon Feb 14 20:48:04 PST 2022 in 4 milliseconds
```

```
The filesystem under path '/rjc2122' is HEALTHY
[cloudera@quickstart Desktop]$ hdfs fsck /rjc2122-files
Connecting to namenode via http://quickstart.cloudera:50070
FSCK started by cloudera (auth:SIMPLE) from /127.0.0.1 for path /rjc2122-
files at Mon Feb 14 20:50:42 PST 2022
```

```
Path '/rjc2122-files' does not exist
[cloudera@quickstart Desktop]$ hdfs fsck /rjc2122 -files
Connecting to namenode via http://quickstart.cloudera:50070
FSCK started by cloudera (auth:SIMPLE) from /127.0.0.1 for path /rjc2122
at Mon Feb 14 20:51:09 PST 2022
/rjc2122 <dir>
/rjc2122/file01 38 bytes, 1 block(s): Under replicated BP-989008105-
127.0.0.1-1433846136903:blk_1073742209_1387. Target Replicas is 3 but
found 1 replica(s).
/rjc2122/file03 15 bytes, 1 block(s): OK
Status: HEALTHY
Total size:      53 B
Total dirs:      1
Total files:     2
Total symlinks:  0
Total blocks (validated): 2 (avg. block size 26 B)
Minimally replicated blocks: 2 (100.0 %)
Over-replicated blocks: 0 (0.0 %)
Under-replicated blocks: 1 (50.0 %)
Mis-replicated blocks: 0 (0.0 %)
Default replication factor: 1
Average block replication: 1.0
Corrupt blocks: 0
Missing replicas: 2 (50.0 %)
Number of data-nodes: 1
Number of racks: 1
FSCK ended at Mon Feb 14 20:51:09 PST 2022 in 3 milliseconds
```

The filesystem under path '/rjc2122' is HEALTHY

20) Touchz Command

```
[cloudera@quickstart Desktop]$ hdfs dfs -touchz /rjc2122/file04
[cloudera@quickstart Desktop]$ hdfs dfs -ls /rjc2122
Found 3 items
-rw-r--r-- 3 cloudera supergroup 38 2022-02-14 19:05
/rjc2122/file01
-rw-r--r-- 1 cloudera supergroup 15 2022-02-14 20:05
/rjc2122/file03
-rw-r--r-- 1 cloudera supergroup 0 2022-02-14 20:57
/rjc2122/file04
```

```
[cloudera@quickstart Desktop]$ echo "hello world" | hdfs dfs -
appendToFile - /rjc2122/file04
[cloudera@quickstart Desktop]$ hdfs dfs -cat /rjc2122/file04
hello world
```

21) Stat Command

```
[cloudera@quickstart Desktop]$ hdfs dfs -stat /rjc2122
2022-02-15 04:05:16
```

22) ->hdfs dfs -stat %b /rjc/file1

```
[cloudera@quickstart Desktop]$ hdfs dfs -stat %b /rjc2122/file01
38
```

23) ->hdfs dfs -stat %o /output_abc

```
[cloudera@quickstart Desktop]$ hdfs dfs -stat %o /rjc2122/file01
134217728
```

24) ->hdfs dfs -stat %r /output_abc

```
[cloudera@quickstart Desktop]$ hdfs dfs -stat %r /rjc2122/file01
3
```

25) ->hdfs dfs -stat %y /output_abc

```
[cloudera@quickstart Desktop]$ hdfs dfs -stat %y /rjc2122/file01
2022-02-15 03:05:40
```

26) Checksum Command

```
[cloudera@quickstart Desktop]$ hdfs dfs -checksum /rjc2122/file04
/rjc2122/file04 MD5-of-0MD5-of-512CRC32C
0000020000000000000000000000cb719ad85249ddab138233798a828f23
```

27) Help command

```
[cloudera@quickstart Desktop]$ hdfs dfs -help ls
-ls [-d] [-h] [-R] [<path> ...] :
  List the contents that match the specified file pattern. If path is not
  specified, the contents of /user/<currentUser> will be listed.
Directory entries
  are of the form:
    permissions - userId groupId sizeOfDirectory(in bytes)
    modificationDate(yyyy-MM-dd HH:mm) directoryName

  and file entries are of the form:
    permissions numberOfReplicas userId groupId sizeOfFile(in bytes)
    modificationDate(yyyy-MM-dd HH:mm) fileName

  -d Directories are listed as plain files.
  -h Formats the sizes of files in a human-readable fashion rather than
  a number
    of bytes.
  -R Recursively list the contents of directories.
```

28) Get command

```
[cloudera@quickstart Desktop]$ hdfs dfs -get /rjc2122/file04  
/home/cloudera/Desktop
```

```
[cloudera@quickstart ~]$ ls /  
bin    dev    home  lib64      media  opt    root  selinux  sys  usr  
boot  etc    lib   lost+found mnt    proc  sbin  srv      tmp  var  
[cloudera@quickstart ~]$ ls /home/cloudera/Desktop  
Eclipse.desktop    file01    file02~  hdfs commands  rjclocal  x1  
Enterprise.desktop file01~   file03~  newfile        sid  
Express.desktop    file02    file04   pract          x
```