**HDFS COMMANDS**

1. **copyFromLocal(Copy a file or directory from Local to HDFS)**

hadoop fs –copyFromLocal /home/training/dvs/f1.txt dvs\_hdfs

**Note: Can be used for copying multiple files, similar pattern files, all the files, a directory**

1. **moveFromLocal(move a file or directory from Local to HDFS)**

hadoop fs –moveFromLocal /home/training/dvs/f1.txt dvs\_hdfs

1. **copyToLocal(Copy a file or directory from HDFS to Local)**

hadoop fs –copyToLocal dvs\_hdfs/f1.txt /home/training/dvs

1. **moveToLocal(Not yet implemented)**
2. **cp (copy a file from one location to another location inside HDFS)**

hadoop fs –cp dvs\_hdfs/file1 dvs\_hdfs1

1. **mv (move a file from one location to another location inside HDFS)**

hadoop fs –mv dvs\_hdfs/file1 dvs\_hdfs1

1. **put (Similar to copyFromLocal)**

hadoop fs –put /home/training/dvs/file1 dvs\_hdfs

1. **get (Similar to copyToLocal)**

hadoop fs –get dvs\_hdfs/file1 /home/training/dvs

1. **getmerge (writes multiple file contents in to a single file in Local File system)**

hadoop fs –getmerge dvs\_hdfs/file1 dvs\_hdfs/file2 /home/training/dvs/f3

**Note : This command is not supported in this version of Hadoop.**

1. **mkdir (Create a directory)**

hadoop fs –mkdir dvs\_hdfs

1. **touchz ( can create n no: of empty files in HDFS)**

hadoop fs –touchz dvs\_hdfs/file1

Hadoop fs –touchz f1 f2 f3 f4

1. **rm (Remove a file)**

hadoop fs –rm dvs\_hdfs/file1

Hadoop fs –rm \*.txt \*.csv

Hadoop fs –rm \*

Hadoop fs –skipTrash dvs\_hdfs/f1.txt (Shift delete)

1. **rmr (Can be used for removing a file or Directory recursively)**

hadoop fs –rmr dvs\_hdfs/file

hadoop fs –rmr dvs\_hdfs/Dir1

**Note: Can be used to remove similar pattern files(\*.sh, \*.txt etc), all the files(\*)**

1. **ls (Lists all the files & directories)**

hadoop fs –ls dvs\_hdfs

**ls|tail –n (Tail option with List)**

hadoop fs –ls dvs\_hdfs|tail -10

1. **ls|head –n (head option with List)**

hadoop fs –ls dvs\_hdfs|head -10

1. **cat (Displays the content of a file)**

hadoop fs -cat dvs\_hdfs/file

1. **text(Displays the content of zipped files)**

hadoop fs -text dvs\_hdfs/file.gz

1. **cat|tail –n (Display bottom n lines of a file)**

hadoop fs -cat dvs\_hdfs/file|tail 10

1. **cat|head –n (Display top n lines of a file)**

hadoop fs -cat dvs\_hdfs/file|head 10

1. **cat|wc –l (Counts the no:of lines in a file)**

hadoop fs -cat dvs\_hdfs/file1|wc –l

1. **cat|wc –w (Counts the no:of words in a file)**

hadoop fs –cat dvs\_hdfs/file1|wc –w

1. **cat|wc –c (Counts the no:of Characters in a file)**

hadoop fs -cat dvs\_hdfs/file1|wc –c

1. **du (Disk Usage of a file or directory)**

hadoop fs –du dvs\_hdfs/file1.txt

1. **du –h (formats & shows file or directory size in human readable format)**

hadoop fs –du -h dvs\_hdfs/f1.txt

1. **du –s(shows summary of the directories instead of each file)**

hadoop fs –du –s dvs\_hdfs

1. **df (Disk usage of the entire file system)**

hadoop fs –df

**O/P:**

Filesystem Size Used Available Use%

hdfs://nameservice1 328040332591104 102783556870823 210750795833344 31%

1. **df –h (Formats & shows in the human readable format)**

hadoop fs -df –h

**O/P:**

Filesystem Size Used Available Use%

hdfs://nameservice1 298.4 T 93.5 T 191.7 T 31%

1. **count(Counts all the Directories & Files in the given path)**

hadoop fs –count dvs\_hdfs

1. **fsck (To check file system health)**

hadoop fsck dvs\_hdfs

1. **fsck –files –blocks (Displays corresponding Files& their block level info)**

hadoop fsck dvs\_hdfs/file1.txt –files -blocks

1. **fsck –files –blocks –locations (Displays files& block level info including the block location)**
2. hadoop fsck dvs\_hdfs/f1.txt –files –blocks –locations -racks
3. **setrep(used to change the replication factor a file or a directory)**

hadoop fs –setrep 5 dvs\_hdfs/file1

Hadoop fs –setrep 5 –w dvs\_hdfs/ABC/f1.txt

-w It requests that the command waits for the replication to complete. This can potentially take a very long time.

Hadoop fs –setrep 5 dvs\_hdfs/sens\_data

Hadoop fs –setrep –R 5 dvs\_hdfs/sens\_data

1. **Controlling block size at file level without changing the block size in hdfs-site.xml**

Hadoop fs –D dfs.block.size=134217728 –put source\_path destination\_path

1. **Controlling replication at file level irrespective of the default replication set to 3**

Hadoop fs –D dfs.replication=2 –put source\_path destination\_path

1. **Setting replication factor for a directory in HDFS**

Hadoop fs –setrep 5 –R dvs\_hdfs/ABC

**Note:** All the files copied under this directory will be having a replication factor of 5 irrespective of the default replication set.

1. **Safe Mode**

Hadoop dfsadmin –safemode leave

Hadoop dfsadmin –safemode enter

Hadoop dfsadmin –safemode get

1. **Delete all the files in trash**

hadoop fs –expunge

1. **Copying a file from one cluster to another cluster**

hadoop fs -distcp hdfs://namenodeA/dvs\_hdfs/emp.csv hdfs://namenodeB/dvs\_hdfs

1. **Appending data to a existing file**

Hadoop fs –appendToFile f1.txt dvs\_hdfs/f2.txt

Note : Existing block will not be updated as we cant update data in HDFS. The last block of the data gets delated & new block gets created.