

HADOOP - COMMAND REFERENCE

There are many more commands in "\$HADOOP_HOME/bin/hadoop fs" than are demonstrated here, although these basic operations will get you started. Running ./bin/hadoop dfs with no additional arguments will list all the commands that can be run with the FsShell system. Furthermore, \$HADOOP_HOME/bin/hadoop fs -help commandName will display a short usage summary for the operation in question, if you are stuck.

A table of all the operations is shown below. The following conventions are used for parameters:

```
"<path>" means any file or directory name.  
"<path>..." means one or more file or directory names.  
"<file>" means any filename.  
"<src>" and "<dest>" are path names in a directed operation.  
"<localSrc>" and "<localDest>" are paths as above, but on the local file system.
```

All other files and path names refer to the objects inside HDFS.

1. **ls <path>**
Lists the contents of the directory specified by path, showing the names, permissions, owner, size and modification date for each entry.
2. **lsr <path>**
Behaves like -ls, but recursively displays entries in all subdirectories of path.
3. **du <path>**
Shows disk usage, in bytes, for all the files which match path; filenames are reported with the full HDFS protocol prefix.
4. **dus <path>**
Like -du, but prints a summary of disk usage of all files/directories in the path.
5. **mv <src><dest>**
Moves the file or directory indicated by src to dest, within HDFS.
6. **cp <src> <dest>**
Copies the file or directory identified by src to dest, within HDFS.
7. **rm <path>**
Removes the file or empty directory identified by path.
8. **rmr <path>**

Removes the file or directory identified by path. Recursively deletes any child entries
i. e. , files or subdirectories of path.

9.

put <localSrc> <dest>

Copies the file or directory from the local file system identified by localSrc to dest within the DFS.

10.

copyFromLocal <localSrc> <dest>

Identical to -put

11.

moveFromLocal <localSrc> <dest>

Copies the file or directory from the local file system identified by localSrc to dest within HDFS, and then deletes the local copy on success.

12.

get [-crc] <src> <localDest>

Copies the file or directory in HDFS identified by src to the local file system path identified by localDest.

13.

getmerge <src> <localDest>

Retrieves all files that match the path src in HDFS, and copies them to a single, merged file in the local file system identified by localDest.

14.

cat <file-name>

Displays the contents of filename on stdout.

15.

copyToLocal <src> <localDest>

Identical to -get

16.

moveToLocal <src> <localDest>

Works like -get, but deletes the HDFS copy on success.

17.

mkdir <path>

Creates a directory named path in HDFS.

Creates any parent directories in path that are missing *e. g. , mkdir -p in Linux.*

18.

setrep [-R] [-w] rep <path>

Sets the target replication factor for files identified by path to rep.

The actual replication factor will move toward the target over time

19. **touchz <path>**
Creates a file at path containing the current time as a timestamp. Fails if a file already exists at path, unless the file is already size 0.
20. **test -[ezd] <path>**
Returns 1 if path exists; has zero length; or is a directory or 0 otherwise.
21. **stat [format] <path>**
Prints information about path. Format is a string which accepts file size in blocks , filename , block size , replication , and modification date .
22. **tail [-f] <file2name>**
Shows the last 1KB of file on stdout.
23. **chmod [-R] mode,mode,... <path>...**
Changes the file permissions associated with one or more objects identified by path.... Performs changes recursively with R. mode is a 3-digit octal mode, or {augo}+/-{rwxX}. Assumes if no scope is specified and does not apply an umask.
24. **chown [-R] [owner][:[group]] <path>...**
Sets the owning user and/or group for files or directories identified by path..... Sets owner recursively if -R is specified.
25. **chgrp [-R] group <path>...**
Sets the owning group for files or directories identified by path..... Sets group recursively if -R is specified.
26. **help <cmd-name>**
Returns usage information for one of the commands listed above. You must omit the leading '-' character in cmd.