

Hadoop HDFS Shell Commands

Introduction

Hadoop Distributed File System (HDFS) provides a command-line interface for managing files, directories, and storage operations across the Hadoop cluster. These commands are similar to UNIX commands (ls, cat, du, etc.) but include HDFS-specific functionalities such as replication control, balancing, and distributed copy. Below is a complete list of Hadoop shell commands with concise yet clear descriptions for quick reference.

HDFS Commands

hadoop version

Displays the installed Hadoop version.

```
hadoop@ubuntu-mate:~$ hadoop version
Hadoop 3.3.6
Source code repository https://github.com/apache/hadoop.git -r 1be78238728da9266a4f88195058f08fd012bf9c
Compiled by ubuntu on 2023-06-18T08:22Z
Compiled on platform linux-x86_64
Compiled with protoc 3.7.1
From source with checksum 5652179ad55f76cb287d9c633bb53bbd
This command was run using /home/hadoop/hadoop-3.3.6/share/hadoop/common/hadoop-common-3.3.6.jar
```

hadoop fs -ls /

Lists all files and directories in the root directory of HDFS.

```
hadoop@ubuntu-mate:~$ hadoop fs -ls /
Found 2 items
drwxr-xr-x - hadoop supergroup 0 2025-09-01 10:34 /system
drwxr-xr-x - hadoop supergroup 0 2025-09-01 10:34 /user
```

hadoop fs -du hdfs:/

Reports disk usage and available space of the HDFS file system.

```
hadoop@ubuntu-mate:~$ hadoop fs -du hdfs:/
0 0 hdfs:///system
0 0 hdfs:///user
```

hadoop fs -count hdfs:/

Counts number of directories, files, and total size under the specified path.

```
hadoop@ubuntu-mate:~$ hadoop fs -count hdfs:/
4          0          0 hdfs:///
```

hadoop fsck - /

Checks health and consistency of the HDFS file system. This command is deprecated. New command is '*hdfs fsck - /*'

```
hadoop@ubuntu-mate:~$ hdfs fsck - /
Connecting to namenode via http://localhost:9870/fsck?ugi=hadoop&path=%2F
FSCK started by hadoop (auth:SIMPLE) from /127.0.0.1 for path / at Sat Sep 06 14:08:53 UTC 2025

Status: HEALTHY
  Number of data-nodes: 0
  Number of racks:      0
  Total dirs:           4
  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
  Blocks queued for replication: 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
  Blocks queued for replication: 0
FSCK ended at Sat Sep 06 14:08:53 UTC 2025 in 5 milliseconds

The filesystem under path '/' is HEALTHY
```

```
hadoop@ubuntu-mate:~$ hadoop fsck - /
WARNING: Use of this script to execute fsck is deprecated.
WARNING: Attempting to execute replacement "hdfs fsck" instead.
```

hadoop balancer

Balances data across DataNodes to optimize storage distribution. This command is deprecated, new command is '*hdfs balancer*'.

```
hadoop@ubuntu-mate:~$ hdfs balancer
2025-09-06 14:18:27,816 INFO balancer.Balancer: namenodes = [hdfs://localhost:9000]
2025-09-06 14:18:27,818 INFO balancer.Balancer: parameters = Balancer.BalancerParameters [BalancingPolicy.Node, threshold = 10.0, max idle iteration = 5, #excluded nodes = 0, #included nodes = 0, #source nodes = 0, #blockpools = 0, run during upgrade = false]
2025-09-06 14:18:27,818 INFO balancer.Balancer: included nodes = []
2025-09-06 14:18:27,819 INFO balancer.Balancer: excluded nodes = []
2025-09-06 14:18:27,819 INFO balancer.Balancer: source nodes = []
Time Stamp      Iteration#  Bytes Already Moved  Bytes Left To Move  Bytes Being Moved  NameNode
2025-09-06 14:18:27,822 INFO balancer.NameNodeConnector: getBlocks calls for hdfs://localhost:9000 will be rate-limited to 20 per second
2025-09-06 14:18:28,481 WARN hdfs.DataStreamer: DataStreamer Exception
org.apache.hadoop.ipc.RemoteException(java.io.IOException): File /system/balancer.id could only be written to 0 of the 1 minReplication nodes. There are 0 datanode(s) running and 0 node(s) are excluded in this operation.
```

hadoop fs -mkdir /user/soham

Creates a new directory named 'soham' inside /user directory in HDFS.

```
hadoop@ubuntu-mate:~$ hadoop fs -mkdir /user/soham
hadoop@ubuntu-mate:~$ hadoop fs -ls /user
Found 2 items
drwxr-xr-x - hadoop supergroup      0 2025-09-01 10:39 /user/hadoop
drwxr-xr-x - hadoop supergroup      0 2025-09-06 14:14 /user/soham
```

hdfs dfs -put /home/hadoop/Downloads/image.jpg /user/soham

Uploads 'image.jpg' from local directory to HDFS.

```
hadoop@ubuntu-mate:~$ hdfs dfs -put /home/hadoop/Downloads/image.jpg /user/soham
hdfs dfs -ls -h /user/soham
Found 1 items
-rw-r--r--  1 hadoop soham      1.4 M 2025-09-06 15:48 /user/soham/image.jpg
```

hadoop fs -ls /user/soham

Lists files and directories inside /user/NIM/hadoop in HDFS.

```
hadoop@ubuntu-mate:~$ hadoop fs -ls /user/soham
Found 1 items
-rw-r--r--  1 hadoop soham    1460259 2025-09-06 15:48 /user/soham/image.jpg
```

hadoop fs -put /home/hadoop/Downloads /user/soham

Uploads the entire local 'retail' directory to /user/NIM/hadoop in HDFS.

```
hadoop@ubuntu-mate:~$ hadoop fs -put /home/hadoop/Downloads /user/soham
hadoop@ubuntu-mate:~$ hadoop fs -ls /user/soham
Found 2 items
drwxr-xr-x   - hadoop soham          0 2025-09-06 15:55 /user/soham/Downloads
-rw-r--r--   1 hadoop soham    1460259 2025-09-06 15:48 /user/soham/image.jpg
```

`hadoop fs -du -s -h /user/soham`

Displays total space used by /user/NIM/hadoop in human-readable format.

```
hadoop@ubuntu-mate:~$ hadoop fs -du -s -h /user/soham
2.9 M   2.9 M   /user/soham
```

`hadoop fs -rm /user/soham/image.jpg`

Deletes the 'image.jpg' file from the user/soham directory in HDFS.

```
hadoop@ubuntu-mate:~$ hadoop fs -rm /user/soham/image.jpg
Deleted /user/soham/image.jpg
```

`hadoop fs -ls /user/soham`

Lists files in the /user/soham directory to confirm deletion.

```
hadoop@ubuntu-mate:~$ hadoop fs -ls /user/soham
Found 1 items
drwxr-xr-x   - hadoop soham          0 2025-09-06 15:55 /user/soham/Downloads
```

`hadoop fs -rm /user/soham/Downloads*`

Deletes all files inside the /user/soham/Downloads directory.

```
hadoop@ubuntu-mate:~$ hadoop fs -rm /user/soham/Downloads/*
Deleted /user/soham/Downloads/.~lock.Hadoop_HDFS_Shell_Commands_Final.docx#
Deleted /user/soham/Downloads/Hadoop_HDFS_Shell_Commands.docx
Deleted /user/soham/Downloads/Hadoop_HDFS_Shell_Commands_Final.docx
Deleted /user/soham/Downloads/SmartCity_Crime_Violence_BigData_Guide.docx
Deleted /user/soham/Downloads/image.jpg
```

`hadoop fs -expunge`

Empties the trash in HDFS permanently.

```
hadoop@ubuntu-mate:~$ hadoop fs -expunge
```

hadoop fs -rm -r /user/soham

Deletes the entire /user/soham directory recursively.

```
hadoop@ubuntu-mate:~$ hadoop fs -rmr /user/soham
rmr: DEPRECATED: Please use '-rm -r' instead.
Deleted /user/soham
```

hadoop fs -ls user/soham

Lists contents of the /user/soham directory again.

```
hadoop@ubuntu-mate:~$ hadoop fs -ls user/soham
ls: `user/soham': No such file or directory
```

hadoop fs -copyFromLocal /home/hadoop/Downloads/image.jpg /user/new

Copies image.jpg from local machine to HDFS directory.

```
hadoop@ubuntu-mate:~$ hadoop fs -copyFromLocal /home/hadoop/Downloads/image.jpg /user/new
copyFromLocal: `/user/new/image.jpg': File exists
```

hadoop fs -cat /user/new/lorem_ipsum.txt

Displays contents of /user/new/lorem_ipsum.txt stored in HDFS.

```
hadoop@ubuntu-mate:~$ hadoop fs -cat /user/new/lorem_ipsum.txt
Lorem ipsum dolor sit amet, consectetur adipiscing elit.
Sed do eiusmod tempor incididunt ut labore et dolore magna aliqua.
Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat.
Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur.
Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.hadoop@ubuntu-mate:~$
```

hadoop fs -copyToLocal /user/new/lorem_ipsum.txt /home/hadoop/Downloads

Downloads lorem_ipsum.txt from HDFS to the local 'Downloads' directory.

```
hadoop@ubuntu-mate:~$ ls /home/hadoop/Downloads
Hadoop_HDFS_Shell_Commands.docx  Hadoop_HDFS_Shell_Commands_Final.docx  image.jpg  lorem_ipsum.txt  SmartCity_Crime_Violence_BigData_Guide.docx
```

hadoop fs -cp /user/new/*.txt /user/new/soham

Copies all .txt files within HDFS between directories.

```
hadoop@ubuntu-mate:~$ hadoop fs -cp /user/new/*.txt /user/new/soham
```

hadoop fs -get /user/new/lorem_ipsum.txt /home/hadoop/Documents

Alternative to copyToLocal, retrieves a file from HDFS to local system.

```
hadoop@ubuntu-mate:~$ hadoop fs -get /user/new/lorem_ipsum.txt /home/hadoop/Documents
get: `/home/hadoop/Documents/lorem_ipsum.txt': File exists
```

hadoop fs -tail /user/new/lorem_ipsum.txt

Displays the last 1 KB of lorem_ipsum.txt from HDFS.

```
hadoop@ubuntu-mate:~$ hadoop fs -tail /user/new/lorem_ipsum.txt
Lorem ipsum dolor sit amet, consectetur adipiscing elit.
Sed do eiusmod tempor incididunt ut labore et dolore magna aliqua.
Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat.
Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur.
Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.hadoop@ubuntu-mate:~$ ^C
```

hadoop fs -chmod 600 /user/new/lorem_ipsum.txt

Changes file permissions to owner-read/write only.

```
hadoop@ubuntu-mate:~$ hadoop fs -ls /user/new
Found 3 items
-rw-r--r-- 1 hadoop soham 1460259 2025-09-06 16:07 /user/new/image.jpg
-rw----- 1 hadoop soham 449 2025-09-06 16:13 /user/new/lorem_ipsum.txt
drwxr-xr-x - hadoop soham 0 2025-09-06 16:21 /user/new/soham
```

hadoop fs -chown root:root /user/new/lorem_ipsum.txt

Changes file owner and group to root.

```
hadoop@ubuntu-mate:~$ hadoop fs -chown root:root /user/new/lorem_ipsum.txt
hadoop@ubuntu-mate:~$ hadoop fs -ls /user/new
Found 3 items
-rw-r--r-- 1 hadoop soham 1460259 2025-09-06 16:07 /user/new/image.jpg
-rw----- 1 root root 449 2025-09-06 16:13 /user/new/lorem_ipsum.txt
drwxr-xr-x - hadoop soham 0 2025-09-06 16:21 /user/new/soham
```

hadoop fs -chgrp training /user/new/lorem_ipsum.txt

Changes group ownership of lorem_ipsum.txt to 'training'.

```
hadoop@ubuntu-mate:~$ hadoop fs -ls /user/new
Found 3 items
-rw-r--r-- 1 hadoop soham 1460259 2025-09-06 16:07 /user/new/image.jpg
-rw----- 1 root training 449 2025-09-06 16:13 /user/new/lorem_ipsum.txt
drwxr-xr-x - hadoop soham 0 2025-09-06 16:21 /user/new/soham
```

hadoop fs -mv /user/new/soham /user/new/bds

Moves the hadoop directory to a new path in HDFS.

```
hadoop@ubuntu-mate:~$ hadoop fs -mv /user/new/soham /user/new/bds
```

```
hadoop@ubuntu-mate:~$ hadoop fs -ls /user/new/bds
Found 1 items
drwxr-xr-x  - hadoop soham          0 2025-09-06 16:21 /user/new/bds/soham
```

hadoop fs -setrep -w 2 /user/new/lorem_ipsum.txt

Changes replication factor of lorem_ipsum.txt to 2.

```
hadoop@ubuntu-mate:~$ hadoop fs -setrep -w 2 /user/new/lorem_ipsum.txt
Replication 2 set: /user/new/lorem_ipsum.txt
```

hadoop fs -distcp hdfs://namenodeA/apache_hadoop

hdfs://namenodeB/hadoop

Performs distributed copy between HDFS clusters.

```
hadoop@ubuntu-mate:~$ hadoop fs -distcp hdfs://namenodeA/apache_hadoop hdfs://namenodeB/hadoop
```

hadoop dfsadmin -safemode leave

Takes NameNode out of safe mode.

```
hadoop@ubuntu-mate:~$ hadoop dfsadmin -safemode leave
WARNING: Use of this script to execute dfsadmin is deprecated.
WARNING: Attempting to execute replacement "hdfs dfsadmin" instead.

Safe mode is OFF
```

hadoop fs

Displays a list of available Hadoop shell commands.

```

hadoop@ubuntu-mate:~$ hadoop fs
Usage: hadoop fs [generic options]
[-appendToFile [-n] <localsrc> ... <dst>]
[-cat [-ignoreCrc] <src> ...]
[-checksum [-v] <src> ...]
[-chgrp [-R] GROUP PATH...]
[-chmod [-R] <MODE[,MODE]... | OCTALMODE> PATH...]
[-chown [-R] [OWNER][:[GROUP]] PATH...]
[-concat <target path> <src path> <src path> ...]
[-copyFromLocal [-f] [-p] [-l] [-d] [-t <thread count>] [-q <thread pool queue size>] <localsrc> ... <dst>]
[-copyToLocal [-f] [-p] [-crc] [-ignoreCrc] [-t <thread count>] [-q <thread pool queue size>] <src> ... <localdst>]
[-count [-q] [-h] [-v] [-t <storage type>]] [-u] [-x] [-e] [-s] <path> ...]
[-cp [-f] [-p | -p[topax]] [-d] [-t <thread count>] [-q <thread pool queue size>] <src> ... <dst>]
[-createSnapshot <snapshotDir> [<snapshotName>]]
[-deleteSnapshot <snapshotDir> <snapshotName>]
[-df [-h] [<path> ...]]
[-du [-s] [-h] [-v] [-x] <path> ...]
[-expunge [-immediate] [-fs <path>]]
[-find <path> ... <expression> ...]
[-get [-f] [-p] [-crc] [-ignoreCrc] [-t <thread count>] [-q <thread pool queue size>] <src> ... <localdst>]
[-getfacl [-R] <path>]
[-getfattr [-R] {-n name | -d} [-e en] <path>]
[-getmerge [-nl] [-skip-empty-file] <src> <localdst>]
[-head <file>]
[-help [cmd ...]]
[-ls [-C] [-d] [-h] [-q] [-R] [-t] [-S] [-r] [-u] [-e] [<path> ...]]
[-mkdir [-p] <path> ...]
[-moveFromLocal [-f] [-p] [-l] [-d] <localsrc> ... <dst>]
[-moveToLocal <src> <localdst>]
[-mv <src> ... <dst>]
[-put [-f] [-p] [-l] [-d] [-t <thread count>] [-q <thread pool queue size>] <localsrc> ... <dst>]
[-renameSnapshot <snapshotDir> <oldName> <newName>]
[-rm [-f] [-r] [-R] [-skipTrash] [-safely] <src> ...]
[-rmdir [--ignore-fail-on-non-empty] <dir> ...]
[-setfacl [-R] [{-b|-k} {-m|-x <acl_spec>} <path>][--set <acl_spec> <path>]]
[-setfattr {-n name [-v value] | -x name} <path>]
[-setrep [-R] [-w] <rep> <path> ...]
[-stat [format] <path> ...]
[-tail [-f] [-s <sleep interval>] <file>]
[-test [-defswrz] <path>]
[-text [-ignoreCrc] <src> ...]
[-touch [-a] [-m] [-t TIMESTAMP (yyyyMMdd:HHmmss) ] [-c] <path> ...]
[-touchz <path> ...]
[-truncate [-w] <length> <path> ...]
[-usage [cmd ...]]

```

hadoop fs -help

Shows detailed help for all or a specific command.

```
hadoop@ubuntu-mate:~$ hadoop fs -help
```



```
-appendToFile [-n] <localsrc> ... <dst> :
  Appends the contents of all the given local files to the given dst file. The dst
  file will be created if it does not exist. If <localSrc> is -, then the input is
  read from stdin. Option -n represents that use NEW_BLOCK create flag to append
  file.

-cat [-ignoreCrc] <src> ... :
  Fetch all files that match the file pattern <src> and display their content on
  stdout.

-checksum [-v] <src> ... :
  Dump checksum information for files that match the file pattern <src> to stdout.
  Note that this requires a round-trip to a datanode storing each block of the
  file, and thus is not efficient to run on a large number of files. The checksum
  of a file depends on its content, block size and the checksum algorithm and
  parameters used for creating the file.

-chgrp [-R] GROUP PATH... :
  This is equivalent to -chown ... :GROUP ...

-chmod [-R] <MODE[,MODE]... | OCTALMODE> PATH... :
  Changes permissions of a file. This works similar to the shell's chmod command
  with a few exceptions.

  -R          modifies the files recursively. This is the only option currently
               supported.

  <MODE>       Mode is the same as mode used for the shell's command. The only
               letters recognized are 'rwxXt', e.g. +t,a+r,g-w,+rwx,o=r.

  <OCTALMODE>  Mode specified in 3 or 4 digits. If 4 digits, the first may be 1 or
               0 to turn the sticky bit on or off, respectively. Unlike the
               shell command, it is not possible to specify only part of the
               mode, e.g. 754 is same as u=rwx,g=rx,o=r.

  If none of 'augo' is specified, 'a' is assumed and unlike the shell command, no
  umask is applied.

-chown [-R] [OWNER][:[GROUP]] PATH... :
  Changes owner and group of a file. This is similar to the shell's chown command
  with a few exceptions.

  -R          modifies the files recursively. This is the only option currently
               supported.
```

```
Generic options supported are:
-conf <configuration file>      specify an application configuration file
-D <property=value>             define a value for a given property
-fs <file:///hdfs://namenode:port> specify default filesystem URL to use, overrides 'fs.defaultFS' property from configurations.
-jt <local|resourceManager:port> specify a ResourceManager
-files <file1,...>              specify a comma-separated list of files to be copied to the map reduce cluster
-libjars <jar1,...>            specify a comma-separated list of jar files to be included in the classpath
-archives <archive1,...>       specify a comma-separated list of archives to be unarchived on the compute machines
```