

EXP NO: 2	Hadoop Implementation of file management tasks, such as Adding files and directories, retrieving files and deleting files
DATE:	

AIM: -

BACKGROUND THEORY: -

PROCEDURE: -

- Switch to superuser mode using `sudo su`.
- Adding Files to HDFS
 - To upload a local file to HDFS, use the `hdfs dfs -put`
- Listing Files and Directories
 - To list files and directories in a specific HDFS directory, use the `hdfs dfs -ls` command
- Creating Directories in HDFS
 - To create a new directory in HDFS, use the `hdfs dfs -mkdir` command
- Retrieving Files from HDFS
 - To download a file from HDFS to your local filesystem, use the `hdfs dfs -get` command
- Deleting Files from HDFS
 - To delete a file from HDFS, use the `hdfs dfs -rm` command
- Deleting Directories from HDFS
 - To delete a directory and its contents from HDFS, use the `hdfs dfs -rm -r` command
- Viewing File Contents in HDFS
 - To view the contents of a file in HDFS, use the `hdfs dfs -cat` command
- Checking Disk Usage of HDFS Directory
 - To check disk usage of files and directories in HDFS, use the `hdfs dfs -du -h` command
- Copying Files within HDFS
 - To copy a file or directory within HDFS, use the `hdfs dfs -cp` command
- Moving Files within HDFS
 - To move a file or directory within HDFS, use the `hdfs dfs -mv` command

CODING: -

- `sudo su`
- `hdfs dfs -put /local/path/to/file /hdfs/path/`
- `hdfs dfs -ls /hdfs/path/`
- `hdfs dfs -mkdir /hdfs/path/`
- `hdfs dfs -get /hdfs/path/to/file /local/path/`
- `hdfs dfs -rm /hdfs/path/to/file`
- `hdfs dfs -rm -r /hdfs/path/to/directory`
- `hdfs dfs -cat /hdfs/path/to/file`
- `hdfs dfs -du -h /hdfs/path/`

- `hdfs dfs -cp /hdfs/source/path /hdfs/destination/path`
- `hdfs dfs -mv /hdfs/source/path /hdfs/destination/path`

OUTPUT: -

```
aws Services Q Search [Alt+S]
hadoop-3.3.6/share/hadoop/hdfs/jdiff/Apache_Hadoop_HDFS_3.0.0.xml
hadoop-3.3.6/share/hadoop/hdfs/jdiff/Apache_Hadoop_HDFS_3.1.1.xml
hadoop-3.3.6/share/hadoop/hdfs/jdiff/Apache_Hadoop_HDFS_3.2.2.xml
hadoop-3.3.6/share/hadoop/hdfs/jdiff/Apache_Hadoop_HDFS_3.2.3.xml
hadoop-3.3.6/share/hadoop/hdfs/jdiff/Apache_Hadoop_HDFS_2.8.2.xml
hadoop-3.3.6/share/hadoop/hdfs/jdiff/Apache_Hadoop_HDFS_3.3.3.xml
hadoop-3.3.6/share/hadoop/hdfs/jdiff/Null.java
hadoop-3.3.6/share/hadoop/hdfs/jdiff/Apache_Hadoop_HDFS_2.8.3.xml
hadoop-3.3.6/share/hadoop/hdfs/jdiff/Apache_Hadoop_HDFS_3.3.5.xml
hadoop-3.3.6/share/hadoop/hdfs/jdiff/Apache_Hadoop_HDFS_2.8.0.xml
hadoop-3.3.6/share/hadoop/hdfs/jdiff/Apache_Hadoop_HDFS_3.0.3.xml
hadoop-3.3.6/share/hadoop/hdfs/jdiff/hadoop-hdfs-0.22.0.xml
hadoop-3.3.6/share/hadoop/hdfs/jdiff/Apache_Hadoop_HDFS_2.9.1.xml
hadoop-3.3.6/share/hadoop/hdfs/jdiff/Apache_Hadoop_HDFS_3.1.1.xml
hadoop-3.3.6/share/hadoop/hdfs/jdiff/hadoop-hdfs-0.20.0.xml
hadoop-3.3.6/share/hadoop/hdfs/jdiff/Apache_Hadoop_HDFS_3.0.0-alpha4.xml
hadoop-3.3.6/share/hadoop/hdfs/jdiff/Apache_Hadoop_HDFS_3.2.0.xml
hadoop-3.3.6/share/hadoop/hdfs/jdiff/Apache_Hadoop_HDFS_2.9.2.xml
hadoop-3.3.6/share/hadoop/hdfs/jdiff/Apache_Hadoop_HDFS_3.0.0-alpha2.xml
hadoop-3.3.6/share/hadoop/hdfs/jdiff/Apache_Hadoop_HDFS_3.0.2.xml
hadoop-3.3.6/share/hadoop/hdfs/jdiff/Apache_Hadoop_HDFS_2.10.0.xml
hadoop-3.3.6/share/hadoop/hdfs/jdiff/Apache_Hadoop_HDFS_3.1.0.xml
hadoop-3.3.6/share/hadoop/hdfs/jdiff/Apache_Hadoop_HDFS_3.0.1.xml
hadoop-3.3.6/share/hadoop/hdfs/jdiff/Apache_Hadoop_HDFS_3.2.1.xml
hadoop-3.3.6/share/hadoop/hdfs/jdiff/Apache_Hadoop_HDFS_3.2.4.xml
hadoop-3.3.6/share/hadoop/hdfs/jdiff/hadoop-hdfs-0.21.0.xml
hadoop-3.3.6/share/hadoop/hdfs/jdiff/Apache_Hadoop_HDFS_3.1.3.xml
hadoop-3.3.6/share/hadoop/hdfs/hadoop-hdfs-client-3.3.6-tests.jar
hadoop-3.3.6/share/hadoop/hdfs/hadoop-hdfs-httplibs-3.3.6.jar
[root@ip-172-31-40-166 local]# sudo mv hadoop-3.3.6 hadoop
[root@ip-172-31-40-166 local]# sudo nano ~/.bashrc
[root@ip-172-31-40-166 local]# sudo nano ~/.bashrc
[root@ip-172-31-40-166 local]# source ~/.bashrc
[root@ip-172-31-40-166 local]# hadoop version
Hadoop 3.3.6
Source code repository https://github.com/apache/hadoop.git -r 1be78238728da9266a4f8b195058f08fd012bf9c
Compiled by ubuntu on 2023-06-18T08:22Z
Compiled on platform linux-x86_64
Compiled with protoc 3.7.1
From source with checksum 5652179ad5576cb287d9c633bb53bbd
This command was run using /usr/local/hadoop/share/hadoop/common/hadoop-common-3.3.6.jar
[root@ip-172-31-40-166 local]#
```

```
Bytes Read=173
File Output Format Counters
Bytes Written=200
[root@ip-172-31-40-166 local]# cat /usr/local/hadoop/output/*
For 1
Hadoop, 1
about 1
and 1
at: 2
http://hadoop.apache.org/ 1
https://cwiki.apache.org/confluence/display/HADOOP/ 1
information 1
latest 1
our 2
please 1
the 1
visit 1
website 1
wiki, 1
[root@ip-172-31-40-166 local]#
```