

Ex No: 2 File Management tasks in Hadoop

AIM:

To perform various file operation in HDFS

Step 1: Adding Files and Directories to HDFS

Before running Hadoop programs on data stored in HDFS, the data needs to be added to HDFS. Let's start by creating a directory and adding a file to it.

1. Create a directory in HDFS:

```
hadoop fs -mkdir /user/myfile
```

This command creates a new directory named myfile in the /user directory in

HDFS. 2. Add a file to HDFS:

```
hadoop fs -put a.txt
```

This command uploads the file a.txt from the local filesystem to the root directory of HDFS.

3. Add the file to the newly created directory:

```
hadoop fs -put a.txt /user/myfile
```

This command uploads the file a.txt from the local filesystem directly into the /user/myfile directory in HDFS.

Step 2: Retrieving Files from HDFS

To copy files from HDFS back to the local filesystem, use the get command. Here's how to retrieve a.txt:

```
hadoop fs -cat a.txt
```

This command displays the contents of the file a.txt directly to the console. To actually copy the file to the local filesystem, you would use:

```
hadoop fs -get a.txt /local/path
```

Replace /local/path with the desired path on your local filesystem.

Step 3: Deleting Files from HDFS

To delete a file from HDFS, use the rm command. Here's how to

```
delete a.txt: hadoop fs -rm a.txt
```

This command removes the file a.txt from HDFS.

Output

The successful execution of the above commands will result in the following:

- Creation of the /user/myfile directory in HDFS.
- Addition of a.txt to HDFS and then to /user/myfile.
- Retrieval of a.txt from HDFS to the local filesystem.
- Deletion of a.txt from HDFS.

Result

The program of file management tasks in Hadoop has been executed successfully, and the output has been verified