

PRACTICAL NO 2

Aim: To implement file management tasks in Hadoop System (HDFS)

Description:

Hadoop HDFS is a distributed file system that provides redundant storage space for files having huge sizes. It is used for storing files that are in the range of terabytes to petabytes.

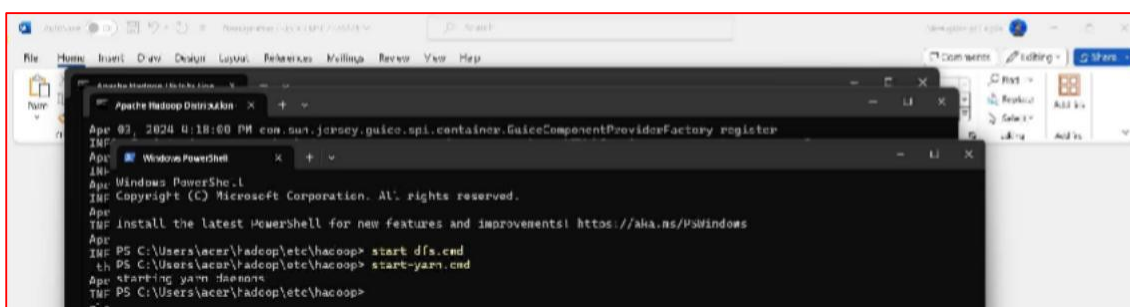
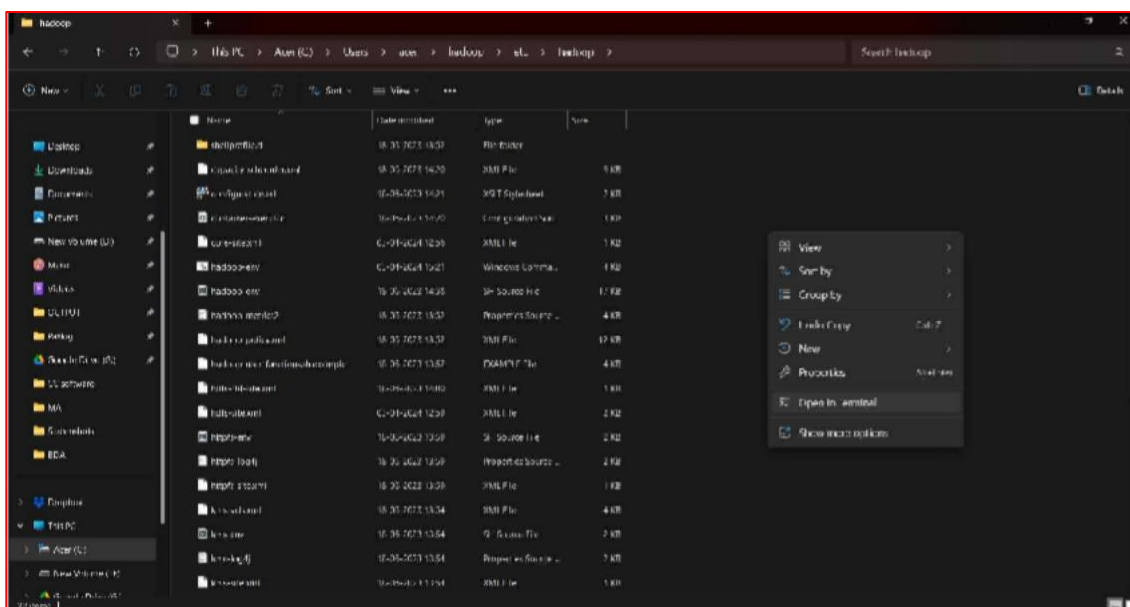
Go to hadoop inside that go to etc folder and then open hadoop directory and right click choose the option open in terminal and run below commands.

hadoop > etc > hadoop > right click > open in terminal (here it will open in powershell)

Commands:-

start-dfs.cmd

start-yarn.cmd



1) Open new command prompt and check the version

Command:

```
hadoop version
```

Output:

```
Command Prompt
Microsoft Windows [Version 10.0.22621.3374]
(c) Microsoft Corporation. All rights reserved.

C:\Users\acer>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 /C:/Users/acer/hadoop/share/hadoop/common/hadoop-common-3.3.6.jar

C:\Users\acer>
```

2) mkdir

Create a directory in Hadoop

Command:

```
hdfs dfs -mkdir /practical
```

Output:

```
PS C:\Users\acer\hadoop\etc\hadoop> hdfs dfs -mkdir /practical
PS C:\Users\acer\hadoop\etc\hadoop> |
```

3) ls

This command is used to enlist the files and directories present in HDFS.

Command:

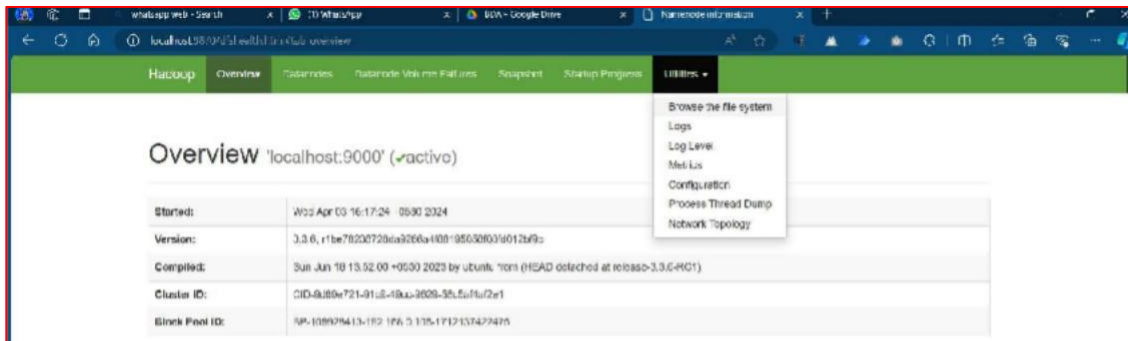
```
hdfs dfs -ls /
```

Output:

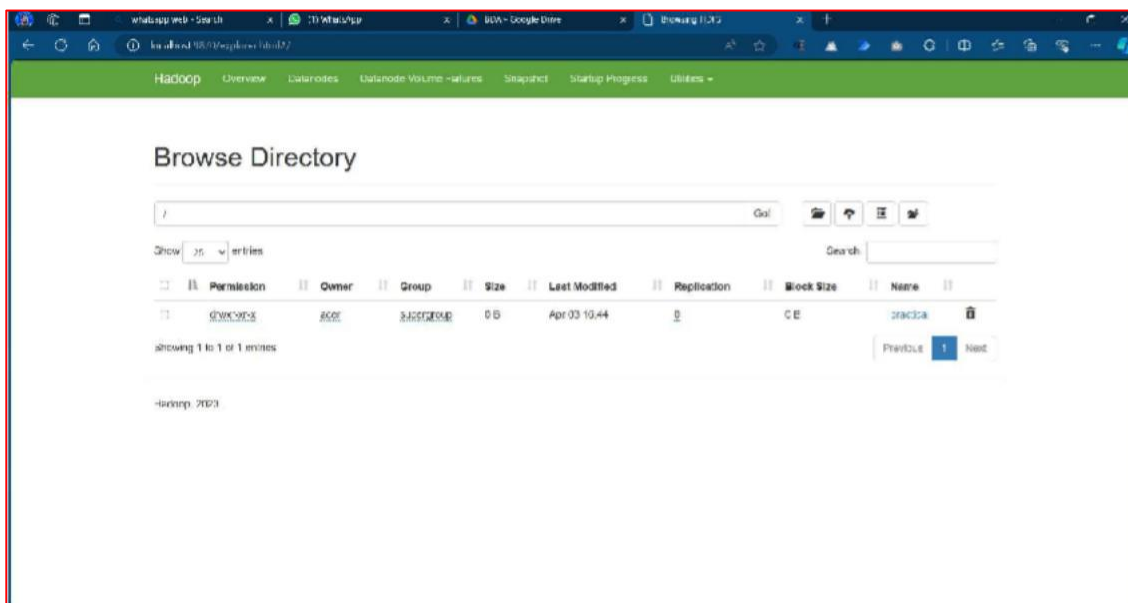
```
PS C:\Users\acer\hadoop\etc\hadoop> hdfs dfs -ls /
Found 1 items
drwxr-xr-x - acer supergroup 0 2024-04-03 16:44 /practical
PS C:\Users\acer\hadoop\etc\hadoop> |
```

Now check the website whether directory is created or not by going to localhost:9870 through your browser

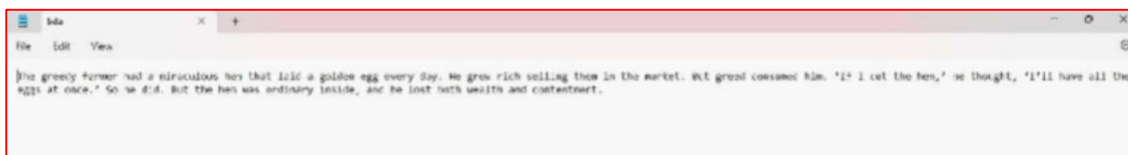
Once the page is open click on 'utilities' and select 'Browse the file system'

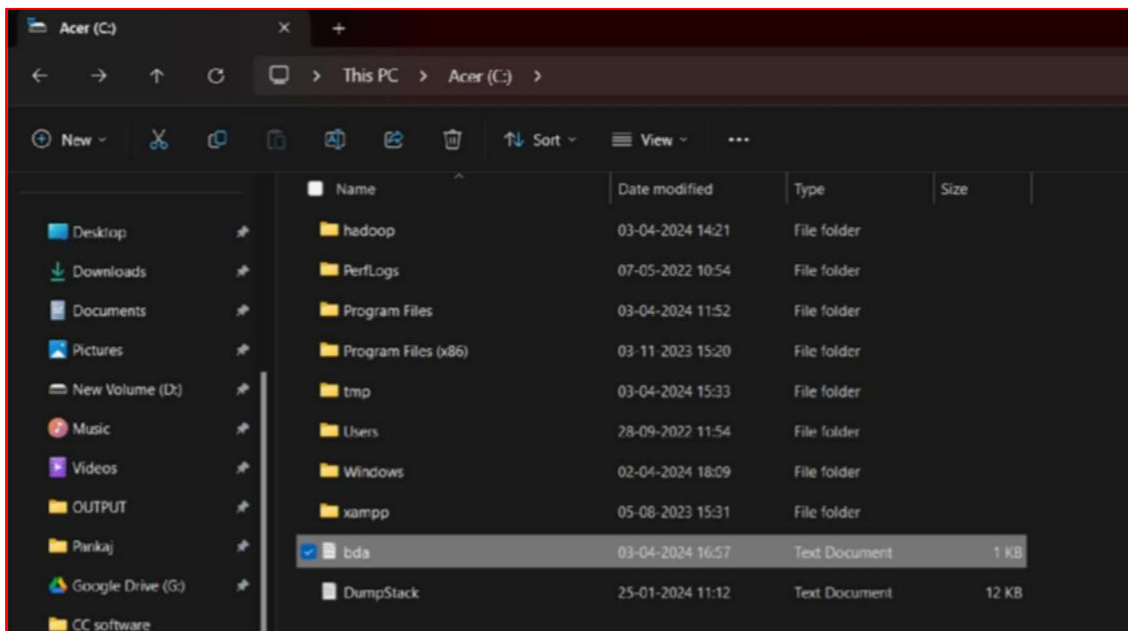


New window will appear named browse directory



4) Create a txt file in desktop or local drive and write 50 words (bda.txt)





5) put

This command is used to copy file from local system to Hadoop file system

Command:

```
hdfs dfs -put C:/bda.txt /practical
```

```
hdfs dfs -ls /practical
```

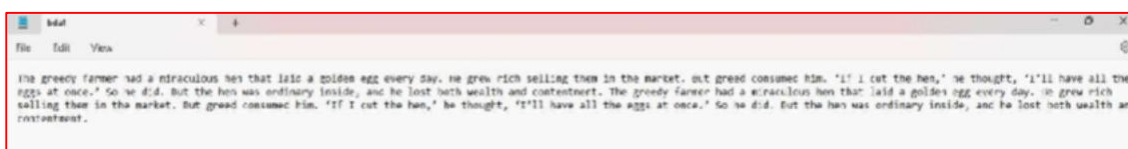
Output:

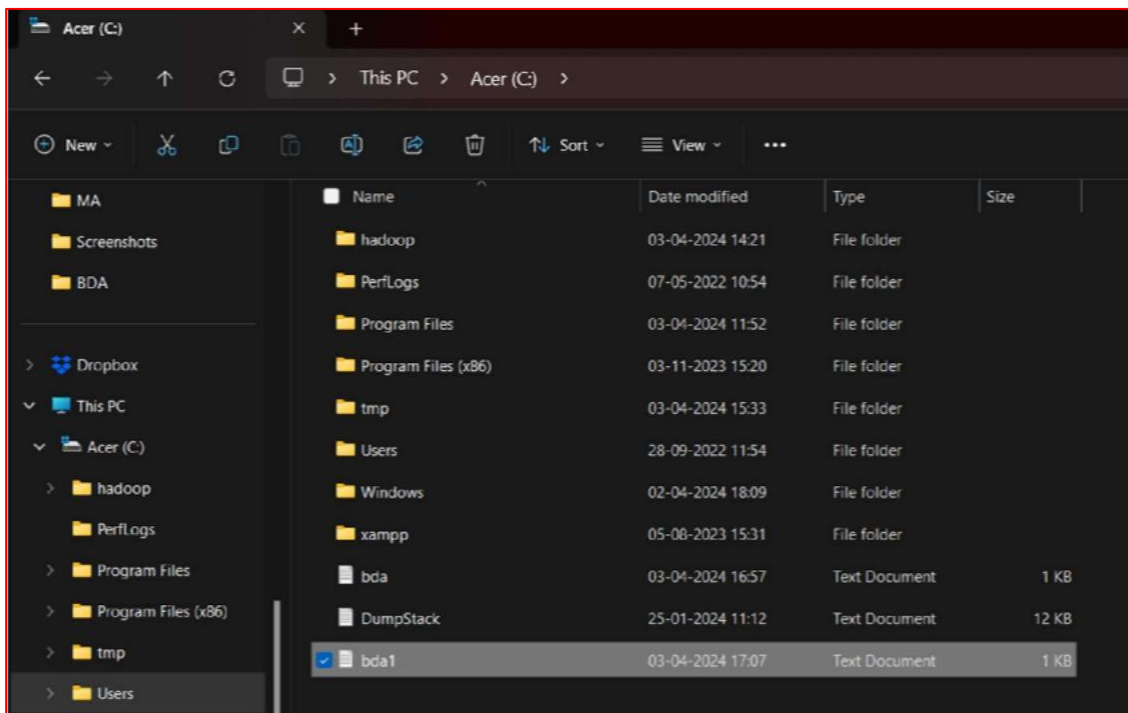
```
PS C:\Users\acer\hadoop\etc\hadoop> hdfs dfs -put C:/bda.txt /practical
PS C:\Users\acer\hadoop\etc\hadoop> hdfs dfs -ls /practical
Found 1 items
-rw-r--r-- 1 acer supergroup          298 2024-04-03 17:05 /practical/bda.txt
PS C:\Users\acer\hadoop\etc\hadoop> |
```

6) copyFromLocal

This command is similar to put command which is used to copy the file from local system to Hadoop system.

Create one more file in Desktop or local drive and save it (bda1.txt)





Command:

```
hdfs dfs -copyFromLocal C:/bda1.txt /practical
```

```
hdfs dfs -ls /practical
```

Output:

```
PS C:\Users\acer\hadoop\etc\hadoop> hdfs dfs -copyFromLocal C:/bda1.txt /practical
PS C:\Users\acer\hadoop\etc\hadoop> hdfs dfs -ls /practical
Found 2 items
-rw-r--r--  1 acer supergroup      298 2024-04-03 17:05 /practical/bda.txt
-rw-r--r--  1 acer supergroup      597 2024-04-03 17:11 /practical/bda1.txt
PS C:\Users\acer\hadoop\etc\hadoop> |
```

7) get

This command is used to copy the file from Hadoop System to local system

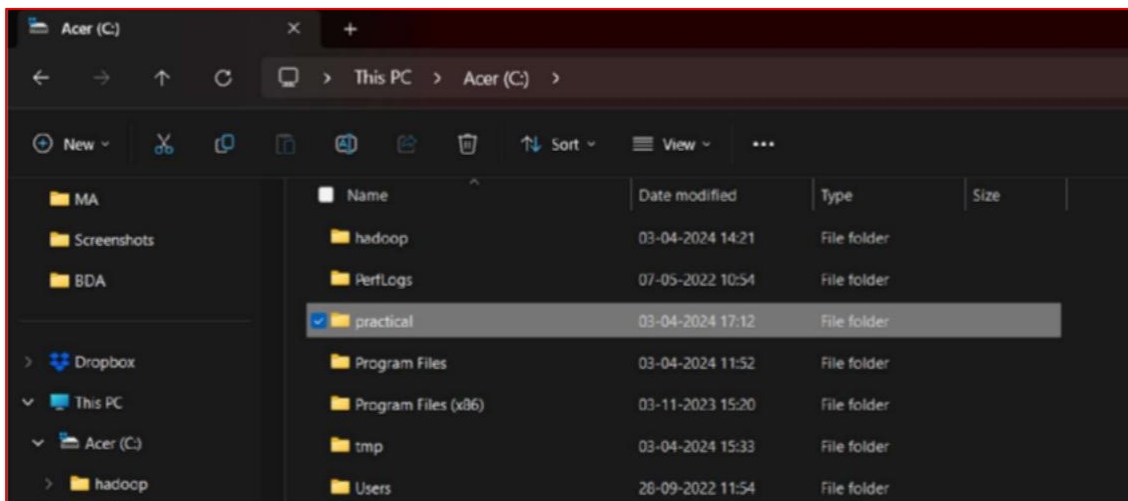
Command:

```
hdfs dfs -get /practical C:/
```

Output:

It will be automatically created on desktop or local drive (i.e. acer (C:))

```
PS C:\Users\acer\hadoop\etc\hadoop> hdfs dfs -get /practical C:/
PS C:\Users\acer\hadoop\etc\hadoop> |
```



8) cat

This command is used to display the content of the file.

Command:

```
hdfs dfs -cat /practical/bda.txt
```

Output:

```
PS C:\Users\acer\hadoop\etc\hadoop> hdfs dfs -cat /practical/bda.txt
The greedy farmer had a miraculous hen that laid a golden egg every day. He grew rich selling them in the market. But greed consumed him. If I cut the hen, I thought, I'll have all the eggs at once. So he did. But the hen was ordinary inside, and he lost both wealth and contentment.
PS C:\Users\acer\hadoop\etc\hadoop>
```

9) mkdir

Create a directory in Hadoop

Command:

```
hdfs dfs -mkdir /practical2
```

```
hdfs dfs -ls /
```

Output:

```
PS C:\Users\acer\hadoop\etc\hadoop> hdfs dfs -mkdir /practical2
PS C:\Users\acer\hadoop\etc\hadoop> hdfs dfs -ls /
Found 2 items
drwxr-xr-x - acer supergroup 0 2024-04-03 17:11 /practical
drwxr-xr-x - acer supergroup 0 2024-04-03 17:15 /practical2
PS C:\Users\acer\hadoop\etc\hadoop>
```

10) mv

This command is used to move the file from one directory to another directory within HDFS system.

Command:

```
hdfs dfs -mv /practical/bda.txt /practical2
```

```
hdfs dfs -ls /practical2
```

Output:

```
PS C:\Users\acer\hadoop\etc\hadoop> hdfs dfs -mv /practical/bda.txt /practical2
PS C:\Users\acer\hadoop\etc\hadoop> hdfs dfs -ls /practical2
Found 1 items
-rw-r--r--  1 acer supergroup      298 2024-04-03 17:05 /practical2/bda.txt
PS C:\Users\acer\hadoop\etc\hadoop> |
```

11) cp

This command is used to copy the file from one directory to another directory within HDFS system.

Command:

```
hdfs dfs -cp /practical/bda1.txt /practical2
```

```
hdfs dfs -ls /practical2
```

Output:

```
PS C:\Users\acer\hadoop\etc\hadoop> hdfs dfs -cp /practical/bda1.txt /practical2
PS C:\Users\acer\hadoop\etc\hadoop> hdfs dfs -ls /practical2
Found 2 items
-rw-r--r--  1 acer supergroup      298 2024-04-03 17:05 /practical2/bda.txt
-rw-r--r--  1 acer supergroup      597 2024-04-03 17:19 /practical2/bda1.txt
PS C:\Users\acer\hadoop\etc\hadoop> |
```

12) moveFromLocal

This command is used to move the file from the local filesystem to the Hadoop filesystem.

Command:

```
hdfs dfs -moveFromLocal C:/bda.txt /practical
```

```
hdfs dfs -ls /practical
```

Output:

```
PS C:\Users\acer\hadoop\etc\hadoop> hdfs dfs -moveFromLocal C:/bda.txt /practical
PS C:\Users\acer\hadoop\etc\hadoop> hdfs dfs -ls /practical
Found 2 items
-rw-r--r--  1 acer supergroup      298 2024-04-03 17:20 /practical/bda.txt
-rw-r--r--  1 acer supergroup      597 2024-04-03 17:11 /practical/bda1.txt
PS C:\Users\acer\hadoop\etc\hadoop> |
```

13) rm

This command is used to remove the files from the directory

Command:

```
hdfs dfs -rm /practical/bda.txt
```

```
hdfs dfs -ls /practical
```

Output:

```
PS C:\Users\acer\hadoop\etc\hadoop> hdfs dfs -rm /practical/bda.txt
Deleted /practical/bda.txt
PS C:\Users\acer\hadoop\etc\hadoop> hdfs dfs -ls /practical
Found 1 items
-rw-r--r--  1 acer supergroup      597 2024-04-03 17:11 /practical/bda1.txt
PS C:\Users\acer\hadoop\etc\hadoop> |
```