

Assisted Practice 3: HDFS Commands

Problem Scenario: Write the commands to work with HDFS.

Objective: In this demonstration, you will explore the basic commands of HDFS.

Tasks to Perform:

1. Create a text file named "Sample.txt" in the vi editor
2. Create a directory named "Simplilearn" in HDFS
3. Transfer a sample text file from your local filesystem to the HDFS directory
4. Show the content of the HDFS file uploaded on the command prompt.
5. Get details of all the replicated blocks of the sample file uploaded
6. Remove the text file from the HDFS directory

Steps to Perform:

Step 1: Log in to your LMS account

Step 2: Open the course "**Big Data Hadoop and Spark Developer**"

Step 3: On the left side, click on the "**PRACTICE LABS**" tab and click on the "**LAUNCH LAB**" button

Big Data Hadoop and Spark Developer
52% Self-Learning Videos Watched | 0/3 Projects Done

BDH

IMP: Dear learner,
Please note: This lab is configured based on the curriculum covered during the live virtual classes.
All details pertaining to the exercises in this lab are provided in the e-books available in your LMS account.

Instructions:

- When you go to the practice Lab page click on the LAUNCH LAB button and it would redirect you to the login credential page.
- To start a Cloudera manager, click on the Auth URL and enter the username and password as given.
- Similarly, for the web console, navigate to the Auth URL, and enter the credentials.
- For FTP, click on the URL and enter the credentials.

Tools:

Cluster Setup

FTP ,Cloudera Infrastructure ,Hadoop 3.x ,HDFS ,YARN ,Kafka ,Flume ,Sqoop ,Pig ,Hive ,Hbase ,Scala ,Spark ,Spark - Core ,Spark SQL ,Spark GraphX ,Spark .ML ,Cloudera Manager

[You can download the Lab Guides from HERE](#)

Your Labs are ready. **LAUNCH LAB**

Step 4: Again, click on the **“LAUNCH LAB”** button

Big Data Hadoop and Spark Developer
52% Self-Learning Videos Watched | 0/3 Projects Done

BDH

IMP: Dear learner,
Please note: This lab is configured based on the curriculum covered during the live virtual classes.
All details pertaining to the exercises in this lab are provided in the e-books available in your LMS account.

Instructions:

- When you go to the practice Lab page click on the LAUNCH LAB button and it would redirect you to the login credential page.
- To start a Cloudera manager, click on the Auth URL and enter the username and password as given.
- Similarly, for the web console, navigate to the Auth URL, and enter the credentials.
- For FTP, click on the URL and enter the credentials.

Tools:

Cluster Setup

FTP ,Cloudera Infrastructure ,Hadoop 3.x ,HDFS ,YARN ,Kafka ,Flume ,Sqoop ,Pig ,Hive ,Hbase ,Scala ,Spark ,Spark - Core ,Spark SQL ,Spark GraphX ,Spark .ML ,Cloudera Manager

[You can download the Lab Guides from HERE](#)

Your Labs are ready. **LAUNCH LAB**

Your Lab Setup is Ready!

This Lab will get reset on
30th September 2022, 2:33 PM

To ensure continued practice, the reset date is automatically set to 120 days from the last lab start date. You'll lose all the lab progress on reset activity. Take a backup of your code if you are going to be away for more than 120 days.

LAUNCH LAB

Note: Practice labs are disabled on course expiry date by default.

Step 5: Click on **“Webconsole”** and click on the **“Auth Url”**

Step 6: Copy the “**Username**” and the “**Password**” provided to log in to the Web console

Step 7: Paste the “**Username**” and the “**Password**” on the console and click on enter

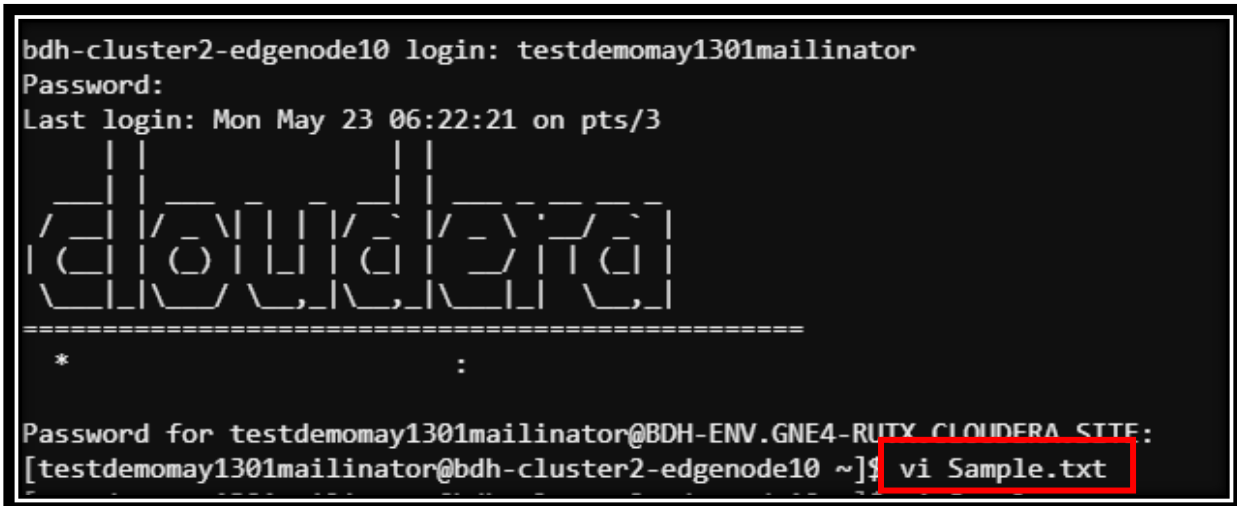
Note: The password will not be visible when pasted on the console

```
bdh-cluster2-edgenode10 login: testdemomay1301mailinator
Password:
Last login: Wed May 18 05:09:33 on pts/1

doug@bdh-cluster2-edgenode10 ~$
*                               :
Password for testdemomay1301mailinator@BDH-ENV.GNE4-RUTX.CLOUDERA.SITE:
[testdemomay1301mailinator@bdh-cluster2-edgenode10 ~]$
```

Step 8: Create a Sample.txt file using the vi editor using the below command

Command: vi Sample.txt

A terminal window showing a login process. The prompt is 'bdh-cluster2-edgenode10 login:'. The user enters 'testdemomay1301mailinator'. The prompt changes to 'Password:'. The user enters a password. The terminal shows 'Last login: Mon May 23 06:22:21 on pts/3'. Below this is a large ASCII art logo for 'DOUGERA'. A separator line of dashes follows. Then, the prompt is 'Password for testdemomay1301mailinator@BDH-ENV.GNE4-RUTX.CLOUDERA.SITE:'. The user enters a password. The terminal shows the prompt '[testdemomay1301mailinator@bdh-cluster2-edgenode10 ~]\$'. The command 'vi Sample.txt' is entered and highlighted with a red box.

```
bdh-cluster2-edgenode10 login: testdemomay1301mailinator
Password:
Last login: Mon May 23 06:22:21 on pts/3

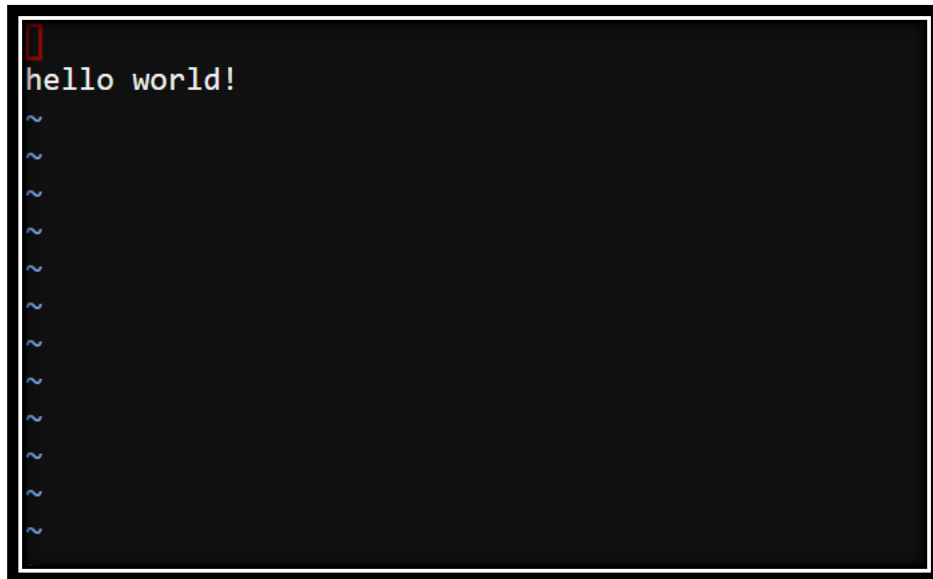
DOUGERA

=====
*                               :

Password for testdemomay1301mailinator@BDH-ENV.GNE4-RUTX.CLOUDERA.SITE:
[testdemomay1301mailinator@bdh-cluster2-edgenode10 ~]$ vi Sample.txt
```

Step 9: Add some text inside the Sample.txt file after the below screen appears:



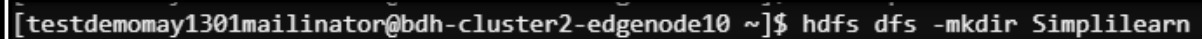
A screenshot of a terminal window with a black background and a white border. The text 'hello world!' is written in white at the top left. Below it, there are several lines of tilde symbols (~) in white, indicating a file listing or directory structure.

Step 10: To come out of the vi editor, click on esc:wq

Step 11: Create a directory named “**Simplilearn**” on HDFS using the below command:

Command:

```
hdfs dfs -mkdir Simplilearn
```

A terminal window screenshot showing the command 'hdfs dfs -mkdir Simplilearn' being executed. The prompt is '[testdemomay1301mailinator@bdh-cluster2-edgenode10 ~]\$'.

Step 12: Copy the Sample.txt file from the local system to the HDFS Simplilearn directory using the below command:

Command:

```
hdfs dfs -put Sample.txt Simplilearn/
```

```
[testdemomay1301mailinator@bdh-cluster2-edgenode10 ~]$ hdfs dfs -put Sample.txt Simplilearn/
```

Step 13: Using the command prompt, display the contents of the Sample.txt file that was uploaded to the HDFS

Command:

```
hdfs dfs -cat Simplilearn/Sample.txt
```

```
[testdemomay1301mailinator@bdh-cluster2-edgenode10 ~]$ hdfs dfs -cat Simplilearn/Sample.txt
hello world
```

Step 14: Get details of all the replicated blocks of the Sample.txt file uploaded on the HDFS

Command:

```
hdfs fsck Simplilearn/Sample.txt -files -locations -blocks
```

```
[testdemomay1301mailinator@bdh-cluster2-edgenode10 ~]$ hdfs fsck Simplilearn/Sample.txt -files -locations -blocks
Connecting to namenode via https://bdh-cluster2-master10.bdh-env.gne4-rutx.cloudera.site:9871/fsck?ugi=testdemomay1301%2Fuser%2Ftestdemomay1301mailinator%2FSimplilearn%2FSample.txt
FSCK started by testdemomay1301mailinator (auth:KERBEROS_SSL) from /10.0.2.134 for path /user/testdemomay1301mailinator
59 UTC 2022

/user/testdemomay1301mailinator/Simplilearn/Sample.txt 12 bytes, replicated: replication=2, 1 block(s): OK
0. BP-1117704909-10.0.2.99-1646027839873:blk_1073859012_118214 len=12 Live_repl=2 [DatanodeInfoWithStorage[10.0.2.182,DISK], DatanodeInfoWithStorage[10.0.2.120:1004,DS-41ea42fc-a235-4944-a435-7727565aee41,DISK]]

Status: HEALTHY
Number of data-nodes: 3
Number of racks: 1
Total dirs: 0
Total symlinks: 0
```

Step 15: Remove the Sample.txt file from the HDFS directory

Command:

hdfs dfs -rm -r Simplilearn/Sample.txt

```
[testdemomay1301mailinator@bdh-cluster2-edgenode10 ~]$ hdfs dfs -rm -r Simplilearn/Sample.txt
Deleted Simplilearn/Sample.txt
[testdemomay1301mailinator@bdh-cluster2-edgenode10 ~]$
```