

Lab: Managing Files On HDFS

Lab: Managing Files On HDFS

1. Table of Contents

1. Table of Contents.....	2
2. HDFS File Management.....	3

Lab: Managing Files On HDFS

2. HDFS File Management

The goal of this lab is to demonstrate HDFS file management through HDFS CLI.

Lets upload a file into HDFS

1. Launch a terminal on Jupyter
2. Create a file, call it 'hello.txt', with some text content in it

```
echo "Hello World" > /tmp/hello.txt
```

3. Lets create a directory in HDFS

```
hdfs dfs -mkdir /user/student/demo
```

4. Lets put it to HDFS

```
hdfs dfs -put /tmp/hello.txt /user/student/demo/
```

Lets run some read activities

1. See the contents of /user/student and /user/student/demo

```
hdfs dfs -ls /user/student  
hdfs dfs -ls /user/student/demo
```

2. See the contents of hello.txt

```
hdfs dfs -cat /user/student/demo/hello.txt
```

3. Lets do a checksum of hello.txt

```
hdfs dfs -checksum /user/admin/demo/hello.txt
```

4. Lets see the size usage of hello.txt

```
hdfs dfs -du /user/student/demo/hello.txt
```

5. lets download hello.txt to local

```
hdfs dfs -get /user/student/demo/hello.txt  
cat hello.txt
```

Lab: Managing Files On HDFS

6. Lets delete hello.txt and its directory

```
hdfs dfs -rm /user/student/demo/hello.txt  
hdfs dfs -rmdir /user/student/demo/
```