

Reem Almijmaj

ID: 3217747723

DSCI 551

Lab2

Give examples for each of the following HDFS commands:

– ls

To list the contents of the root

```
[ec2-user@ip-172-31-22-162 ~]$ hdfs dfs -ls /  
Found 3 items  
drwxr-xr-x - ec2-user supergroup          0 2022-02-10 08:13 /DSCI_LAB2  
drwxr-xr-x - ec2-user supergroup          0 2022-02-10 04:32 /ReemAlmijmaj  
drwxr-xr-x - ec2-user supergroup          0 2022-02-10 08:04 /dsci551  
[ec2-user@ip-172-31-22-162 ~]$ |
```

To list the contents of the directory dsci551

```
[ec2-user@ip-172-31-22-162 current]$ hdfs dfs -ls /dsci551  
Found 2 items  
drwxr-xr-x - ec2-user supergroup          0 2022-02-09 01:08 /dsci551/input  
-rw-r--r-- 1 ec2-user supergroup      8347 2022-02-10 06:22 /dsci551/reem.xml
```

– mkdir

```
[ec2-user@ip-172-31-22-162 ~]$ hdfs dfs -mkdir /dsci551Lab2  
[ec2-user@ip-172-31-22-162 ~]$ |
```

Just to make sure that dsci551Lab2 is exists, please check the last line:

```
[ec2-user@ip-172-31-22-162 ~]$ hdfs dfs -ls /  
Found 3 items  
drwxr-xr-x - ec2-user supergroup          0 2022-02-10 04:32 /ReemAlmijmaj  
drwxr-xr-x - ec2-user supergroup          0 2022-02-10 06:22 /dsci551  
drwxr-xr-x - ec2-user supergroup          0 2022-02-10 06:32 /dsci551Lab2
```

– rmdir

```
[ec2-user@ip-172-31-22-162 ~]$ hdfs dfs -rmdir /dsci551Lab2
```

Just to prove that this command works, I included a screenshot for all the directories and as you can see dsci551Lab2 is not there.

```
[ec2-user@ip-172-31-22-162 ~]$ hdfs dfs -ls /  
Found 2 items  
drwxr-xr-x - ec2-user supergroup          0 2022-02-10 04:32 /ReemAlmijmaj  
drwxr-xr-x - ec2-user supergroup          0 2022-02-10 06:22 /dsci551  
[ec2-user@ip-172-31-22-162 ~]$ |
```

- rm

As you can see the file reem.xml was deleted.

```
[ec2-user@ip-172-31-22-162 ~]$ hdfs dfs -rm /dsci551/reem.xml
Deleted /dsci551/reem.xml
[ec2-user@ip-172-31-22-162 ~]$
```

- put

```
[ec2-user@ip-172-31-22-162 current]$ ls
edits_000000000000000000000001-000000000000000000000002  edits_000000000000000000000070-000000000000000000000071  edits_000000000000000000000075-000000000000000000000076  fsimage_000000000000000000000086  reem.xml
edits_000000000000000000000003-000000000000000000000066  edits_000000000000000000000072-000000000000000000000072  edits_000000000000000000000077-000000000000000000000086  fsimage_000000000000000000000086  seen_txid
edits_000000000000000000000067-000000000000000000000067  edits_000000000000000000000073-000000000000000000000073  edits_000000000000000000000087-000000000000000000000088  fsimage_000000000000000000000088  VERSION
edits_000000000000000000000068-000000000000000000000069  edits_000000000000000000000074-000000000000000000000074  edits_inprogress_000000000000000000000089  fsimage_000000000000000000000089  Yasmin.xml
[ec2-user@ip-172-31-22-162 current]$ hdfs dfs -put reem.xml /dsci551
```

I put reem.xml in the directory dsci551

```
[ec2-user@ip-172-31-22-162 ~]$ hdfs dfs -ls /dsci551
Found 3 items
drwxr-xr-x - ec2-user supergroup 0 2022-02-09 01:08 /dsci551/input
-rw-r--r-- 1 ec2-user supergroup 31 2022-02-10 08:04 /dsci551/lab2.txt
-rw-r--r-- 1 ec2-user supergroup 0 2022-02-11 20:18 /dsci551/reem.xml
```

- get

```
[ec2-user@ip-172-31-22-162 lab2]$ ls
[ec2-user@ip-172-31-22-162 lab2]$ hdfs dfs -get /DSCI_LAB2/lab2.txt /home/ec2-user/lab2
```

To make sure that I downloaded the file lab2.txt to lab2 directory:

```
[ec2-user@ip-172-31-22-162 lab2]$ ls
lab2.txt
```

- cp

```
[ec2-user@ip-172-31-22-162 lab2]$ hdfs dfs -cp /dsci551/lab2.txt /DSCI_LAB2
```

To make sure that I have copy the file to DSCI_LAB2, please check the snapshot:

```
[ec2-user@ip-172-31-22-162 lab2]$ hdfs dfs -ls /DSCI_LAB2
Found 1 items
-rw-r--r-- 1 ec2-user supergroup 31 2022-02-10 08:13 /DSCI_LAB2/lab2.txt
```

- cat

I used cat to show the content of the reem.xml file

```
ec2-user@ip-172-31-22-162:~
[ec2-user@ip-172-31-22-162 ~]$ hdfs dfs -cat /dsci551/reem.xml
<?xml version="1.0"?>
<fsimage><version><layoutVersion>-66</layoutVersion><onDiskVersion>1</onDiskVersion><oivRevision>a3b9c37a397ad4188041dd80621bdeefc46885f2</oivRevision></version>
<NameSection><namespaceId>1933799804</namespaceId><genStampV1>1000</genStampV1><genStampV2>1011</genStampV2><genStampV1Limit>0</genStampV1Limit><lastAlllocatedBlockId>1073741835
</NameSection>
<ErasureCodingSection>
<erasureCodingPolicy>
<policyId>1</policyId><policyName>RS-6-3-1024k</policyName><cellSize>1048576</cellSize><policyState>DISABLED</policyState><ecSchema>
<codecName>rs</codecName><dataUnits>6</dataUnits><parityUnits>3</parityUnits></ecSchema>
</erasureCodingPolicy>
<erasureCodingPolicy>
<policyId>2</policyId><policyName>RS-3-2-1024k</policyName><cellSize>1048576</cellSize><policyState>DISABLED</policyState><ecSchema>
<codecName>rs</codecName><dataUnits>3</dataUnits><parityUnits>2</parityUnits></ecSchema>
```