

19BCE248

Big Data Analytics


Lab 3

D2

AIM: To run various commands of Hadoop

mkdir:

```
C:\hadoop-3.2.1\sbin>hadoop fs -mkdir \second_
```

<input type="checkbox"/>	drwxr-xr-x	Admin	supergroup	0 B	Sep 21 14:50	0	0 B	second	
--------------------------	------------	-------	------------	-----	--------------	---	-----	--------	---

rm:

```
C:\hadoop-3.2.1\sbin>hadoop fs -rm /first/submission.csv  
Deleted /first/submission.csv
```

rmdir:

```
C:\hadoop-3.2.1\sbin>hadoop fs -rmdir /first
```

put:

```
C:\hadoop-3.2.1\sbin>hadoop fs -put C:/Users/Admin/OneDrive/Desktop/commands.txt \second_
```

SHOW 20 entries

Search:

<input type="checkbox"/>	Permission	Owner	Group	Size	Last Modified	Replication	Block Size	Name	
<input type="checkbox"/>	-rw-r--r--	Admin	supergroup	141 B	Sep 21 14:50	3	128 MB	commands.txt	

Showing 1 to 1 of 1 entries

Previous 1 Next

cp:

```
C:\hadoop-3.2.1\sbin> hadoop fs -cp /second/commands.txt /third
2022-09-21 15:48:10,080 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2022-09-21 15:48:10,144 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
```

<input type="checkbox"/>	-rw-r--r--	Admin	supergroup	141 B	Sep 21 15:48	3	128 MB	third	
--------------------------	------------	-------	------------	-------	--------------	---	--------	-------	--

Executing wordcount:

```
C:\hadoop-3.2.1\sbin>hadoop jar C:\hadoop-3.2.1\share\hadoop\mapreduce\hadoop-mapreduce-examples-3.2.1.jar wordcount \second\commands.txt \output
2022-09-21 15:13:59,815 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032
2022-09-21 15:14:00,397 INFO mapreduce.JobResourceUploader: Disabling Erasure Coding for path: /tmp/hadoop-yarn/staging/Admin/.staging/job_1663749946279_0001
2022-09-21 15:14:00,500 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2022-09-21 15:14:00,610 INFO input.FileInputFormat: Total input files to process : 1
2022-09-21 15:14:00,663 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2022-09-21 15:14:00,693 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2022-09-21 15:14:00,709 INFO mapreduce.JobSubmitter: number of splits:1
2022-09-21 15:14:00,808 INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
2022-09-21 15:14:00,828 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1663749946279_0001
2022-09-21 15:14:00,829 INFO mapreduce.JobSubmitter: Executing with tokens: []
2022-09-21 15:14:01,012 INFO conf.Configuration: resource-types.xml not found
```

File for counting :

commands - Notepad

File Edit View

```
start-all.cmd
hdfs fs -put (path from local file) (path to hadoop)
hadoop jar (jar path) function (input from hadoop) (output on hadoop)
```

Output:

```
(input      1
(jar 1
(output    1
(path 2
-put 1
file) 1
from 2
fs    1
function 1
hadoop 1
hadoop) 3
hdfs 1
jar 1
local 1
on 1
path) 1
start-all.cmd 1
to 1
```