BIG DATA ANALYTICS

ASSIGNMENT 3: Manipulation of Data on HDFS

(Amitabh Saini, TY-E-4, 141541)

* To create a directory: mkdir

Eg. mkdir ARS

* To check present working directory: pwd
* For list of directories: ls
* To create a new file: touch

Eg. touch test.txt

* Editor : vi test.txt
* Quit from editor: Press Esc + Enter & then q or q! (to quit from file)
* 3 types of permissions:

1. Read 🡪 4
2. Write 🡪 2
3. Execute 🡪 1

* 3 people access file/directory:

1. User i.e. the creator (RWX)
2. Groups i.e. the set of people user belongs to (RWX)
3. Others i.e. the outside group (RWX)

* ls –l
* User: rw-

420 = 6 (total)

* Group: rw-

420 = 6 (total)

* Others: r--

400 = 4 (total)

Therefore, 664.

* To change mode (RWX)

chmod 777 test.txt

* rwxrwxrwx
* Vi editor modes:

1. Visual
2. Insert
3. Record

H 🡪 Up, J 🡪 Left, L 🡪 Right, K 🡪 Down, I 🡪 Write

* gedit test.sh (Shell Scripting)

In file 🡪 echo “Enter Name”

read name

echo “Hello $name”

* chmod 777 test.sh

./test.sh

o/p: Enter Name

Nidhi

Hello Nidhi

* 755 🡪 read-execute
* echo “Enter File Name”

read name

` touch $name

* Option: ` touch $dirname/$filename`
* echo “Enter Directory Name”

read name

` mkdir $name`

echo “Enter File Name”

cd $name

read name

` touch $name`

* ` 🡪 execute command
* 3 types of commands in Linux:

1. Terminal commands
2. Special commands
3. C codes

* Writing a line to a file in directory:

echo “DirName”

read name

` mkdir $name

echo “Filename”

read n

echo “Line”

read line

echo “$line” > $name/$n

CLOUDERA Commands:

[cloudera@quickstart ~]$ hadoop version

Hadoop 2.6.0-cdh5.8.0

Subversion http://github.com/cloudera/hadoop -r 57e7b8556919574d517e874abfb7ebe31a366c2b

Compiled by jenkins on 2016-06-16T19:38Z

Compiled with protoc 2.5.0

From source with checksum 9e99ecd28376acfd5f78c325dd939fed

This command was run using /usr/lib/hadoop/hadoop-common-2.6.0-cdh5.8.0.jar

[cloudera@quickstart ~]$ ls -l

total 196

drwxrwxr-x 2 cloudera cloudera 4096 Sep 19 23:29 amit

drwxrwxr-x 2 cloudera cloudera 4096 Aug 29 07:28 AS

-rwxrwxr-x 1 cloudera cloudera 5387 Aug 10 19:52 cloudera-manager

-rwxrwxr-x 1 cloudera cloudera 9964 Aug 10 19:52 cm\_api.py

drwxrwxr-x 2 cloudera cloudera 4096 Aug 10 19:52 Desktop

drwxrwxr-x 4 cloudera cloudera 4096 Aug 10 19:52 Documents

drwxr-xr-x 2 cloudera cloudera 4096 Aug 24 02:57 Downloads

drwxrwsr-x 9 cloudera cloudera 4096 Aug 24 03:06 eclipse

-rw-rw-r-- 1 cloudera cloudera 53653 Aug 10 19:52 enterprise-deployment.json

-rw-rw-r-- 1 cloudera cloudera 50513 Aug 10 19:52 express-deployment.json

-rwxrwxr-x 1 cloudera cloudera 5007 Aug 10 19:52 kerberos

drwxrwxr-x 2 cloudera cloudera 4096 Aug 10 19:52 lib

drwxr-xr-x 2 cloudera cloudera 4096 Aug 24 02:57 Music

-rwxrwxr-x 1 cloudera cloudera 4226 Aug 10 19:52 parcels

drwxr-xr-x 2 cloudera cloudera 4096 Aug 24 02:57 Pictures

drwxr-xr-x 2 cloudera cloudera 4096 Aug 24 02:57 Public

drwxr-xr-x 2 cloudera cloudera 4096 Aug 24 02:57 Templates

drwxr-xr-x 2 cloudera cloudera 4096 Aug 24 02:57 Videos

drwxrwxr-x 4 cloudera cloudera 4096 Aug 10 19:52 workspace

[cloudera@quickstart ~]$ mkdir ARS

[cloudera@quickstart ~]$ ls -l

total 200

drwxrwxr-x 2 cloudera cloudera 4096 Sep 19 23:29 amit

drwxrwxr-x 2 cloudera cloudera 4096 Sep 20 00:02 ARS

drwxrwxr-x 2 cloudera cloudera 4096 Aug 29 07:28 AS

-rwxrwxr-x 1 cloudera cloudera 5387 Aug 10 19:52 cloudera-manager

-rwxrwxr-x 1 cloudera cloudera 9964 Aug 10 19:52 cm\_api.py

drwxrwxr-x 2 cloudera cloudera 4096 Aug 10 19:52 Desktop

drwxrwxr-x 4 cloudera cloudera 4096 Aug 10 19:52 Documents

drwxr-xr-x 2 cloudera cloudera 4096 Aug 24 02:57 Downloads

drwxrwsr-x 9 cloudera cloudera 4096 Aug 24 03:06 eclipse

-rw-rw-r-- 1 cloudera cloudera 53653 Aug 10 19:52 enterprise-deployment.json

-rw-rw-r-- 1 cloudera cloudera 50513 Aug 10 19:52 express-deployment.json

-rwxrwxr-x 1 cloudera cloudera 5007 Aug 10 19:52 kerberos

drwxrwxr-x 2 cloudera cloudera 4096 Aug 10 19:52 lib

drwxr-xr-x 2 cloudera cloudera 4096 Aug 24 02:57 Music

-rwxrwxr-x 1 cloudera cloudera 4226 Aug 10 19:52 parcels

drwxr-xr-x 2 cloudera cloudera 4096 Aug 24 02:57 Pictures

drwxr-xr-x 2 cloudera cloudera 4096 Aug 24 02:57 Public

drwxr-xr-x 2 cloudera cloudera 4096 Aug 24 02:57 Templates

drwxr-xr-x 2 cloudera cloudera 4096 Aug 24 02:57 Videos

drwxrwxr-x 4 cloudera cloudera 4096 Aug 10 19:52 workspace

[cloudera@quickstart ~]$ cd ARS

[cloudera@quickstart ARS]$ touch test

[cloudera@quickstart ARS]$ ls -l

total 0

-rw-rw-r-- 1 cloudera cloudera 0 Sep 20 00:02 test

[cloudera@quickstart ARS]$ chmod 755 test

[cloudera@quickstart ARS]$ ls -l

total 0

-rwxr-xr-x 1 cloudera cloudera 0 Sep 20 00:02 test

[cloudera@quickstart ARS]$ gedit test

[cloudera@quickstart ARS]$ cat test

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

This is a file

[cloudera@quickstart ARS]$ hadoop fs -ls /user/cloudera

Found 1 items

-rw-r--r-- 1 cloudera cloudera 0 2016-09-19 23:33 /user/cloudera/test

[cloudera@quickstart ARS]$ hadoop fs -rm /user/cloudera/test

Deleted /user/cloudera/test

[cloudera@quickstart ARS]$ hadoop fs -copyFromLocal test /user/cloudera/

[cloudera@quickstart ARS]$ hadoop fs -ls /user/cloudera

Found 1 items

-rw-r--r-- 1 cloudera cloudera 811 2016-09-20 00:04 /user/cloudera/test

[cloudera@quickstart ARS]$ chmod 755 test

[cloudera@quickstart ARS]$ ls -l

total 8

-rwxr-xr-x 1 cloudera cloudera 811 Sep 20 00:03 test

-rw-r--r-- 1 cloudera cloudera 811 Sep 20 00:03 test~

[cloudera@quickstart ARS]$ hadoop fs -chmod 755 /user/cloudera/test

[cloudera@quickstart ARS]$ hadoop fs -ls /user/cloudera

Found 1 items

-rwxr-xr-x 1 cloudera cloudera 811 2016-09-20 00:04 /user/cloudera/test

[cloudera@quickstart ARS]$ rm test

rm: remove regular file `test'? y

[cloudera@quickstart ARS]$ ls -l

total 4

-rw-r--r-- 1 cloudera cloudera 811 Sep 20 00:03 test~

[cloudera@quickstart ARS]$ hadoop fs -copyToLocal /user/cloudera/test

[cloudera@quickstart ARS]$ ls -l

total 8

-rw-r--r-- 1 cloudera cloudera 811 Sep 20 00:06 test

-rw-r--r-- 1 cloudera cloudera 811 Sep 20 00:03 test~

[cloudera@quickstart ARS]$ hadoop fs -rm /user/cloudera/test

Deleted /user/cloudera/test

[cloudera@quickstart ARS]$ hadoop fs -ls /user/cloudera

[cloudera@quickstart ARS]$ hadoop fs -put test /user/cloudera/

[cloudera@quickstart ARS]$ hadoop fs -ls /user/cloudera

Found 1 items

-rw-r--r-- 1 cloudera cloudera 811 2016-09-20 00:07 /user/cloudera/test

[cloudera@quickstart ARS]$ rm test

rm: remove regular file `test'? y

[cloudera@quickstart ARS]$ hadoop fs -put /user/cloudera/test

put: `/user/cloudera/test': No such file or directory

[cloudera@quickstart ARS]$ hadoop fs -ls /user/cloudera

Found 1 items

-rw-r--r-- 1 cloudera cloudera 811 2016-09-20 00:07 /user/cloudera/test

[cloudera@quickstart ARS]$ hadoop fs -get /user/cloudera/test

[cloudera@quickstart ARS]$ ls -l

total 8

-rw-r--r-- 1 cloudera cloudera 811 Sep 20 00:13 test

-rw-r--r-- 1 cloudera cloudera 811 Sep 20 00:03 test~

[cloudera@quickstart ARS]$ hadoop fs -mkdir bda

[cloudera@quickstart ARS]$ hadoop fs -ls /user/cloudera

Found 2 items

drwxr-xr-x - cloudera cloudera 0 2016-09-20 00:13 /user/cloudera/bda

-rw-r--r-- 1 cloudera cloudera 811 2016-09-20 00:07 /user/cloudera/test

[cloudera@quickstart ARS]$ hadoop fs -rmdir /user/cloudera/bda

[cloudera@quickstart ARS]$ hadoop fs -ls /user/cloudera

Found 1 items

-rw-r--r-- 1 cloudera cloudera 811 2016-09-20 00:07 /user/cloudera/test

[cloudera@quickstart ARS]$ hadoop fs -chmod 755 /user/cloudera/test

[cloudera@quickstart ARS]$ hadoop fs -ls /user/cloudera

Found 1 items

-rwxr-xr-x 1 cloudera cloudera 811 2016-09-20 00:07 /user/cloudera/test

[cloudera@quickstart ARS]$

* We cannot edit file in cloudera because file is to be divided equally.
* Database Types:
* OLTP(Online Transmission Processing)
* OLAP(Online Analytic Processing) --Hadoop