Big Data And Hadoop

Session 18 - Assignment 1

Problem Statement:

- Create a HBase table 'customer' with one column family 'details'.
- Write a shells script that loads the content of the file customers.dat in the HBase table.
- The structure of file looks like: id, name, location, age

Solution:

A. Place input file in HDFS

Step 1: The input file is present on local file system at **/home/acadgild/Abhilasha/hbase** as shown below.

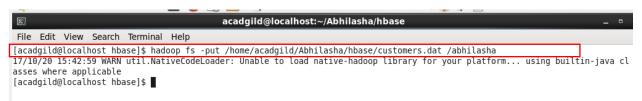
```
File Edit View Search Terminal Help

[acadgild@localhost hbase]$ pwd
/home/acadgild/Abhilasha/hbase
[acadgild@localhost hbase]$ ls -l
total 12
-rw-rw-r--. 1 acadgild acadgild 833 Sep 24 18:27 commands
-rw-rw-r--. 1 acadgild acadgild 257 Sep 24 17:55 commands~
-rw-rw-r--. 1 acadgild acadgild 60 Oct 17 22:09 customers.dat

[acadgild@localhost hbase]$ [acadgild@localhost hbase]$
```

Name of the file is customers.dat

Step 2: We now place the file on HDFS using the put command at location /abhilasha as follows



Step 3: On listing the directory /abhilasha, we find the file with name customers.dat placed.

```
acadgild@localhost:~/Abhilasha/hbase
 File Edit View Search Terminal Help
[acadgild@localhost hbase]$ hadoop fs -ls /abhilasha
17/10/20 15:43:44 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
Found 5 items
drwxr-xr-x
                acadgild supergroup
                                              0 2017-08-29 21:58 /abhilasha/FlumeTitanicData
             1 acadgild supergroup
-rw-r--r--
                                             60 2017-10-20 15:43 /abhilasha/customers.dat
drwxr-xr-x

    acadgild supergroup

                                              0 2017-09-05 16:52 /abhilasha/flume

    acadgild supergroup
    acadgild supergroup

drwxr-xr-x
                                              0 2017-09-17 15:18 /abhilasha/hive
drwxr-xr-x
                                              0 2017-08-29 22:53 /abhilasha/project12
[acadgild@localhost hbase]$
```

Step 4: We can view the content of this file using **cat** command as follows:

```
File Edit View Search Terminal Help

[acadqild@localhost hbase]$ hadoop fs -cat /abhilasha/customers.dat

17/10/20 15:44:21 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable

1,Amit,IND,18
2,Sumit,PAK,20
3,Rohit,AUS,26
4,Namit,UK,24[acadqild@localhost hbase]$
```

B. Start HBase

Step 1: We first start HBase using the command start-hbase.sh as follows:



Step 2: Next we start the hbase shell using the command **hbase shell** as follows:



C. Create table

Step 1: Now we create the table named **customer** using the command **create 'customer', 'details'** as follows. Here, **details** is the name of the column family of the table **customer**.



Step 2: Using list command, we can see the created table appear in the list.

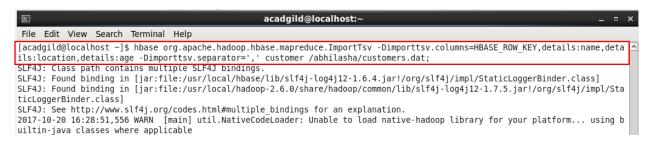


D. Load data from file into HBase table

We use the below command to read data from file **customers.dat** and load it in hbase table name **customer**

```
hbase org.apache.hadoop.hbase.mapreduce.ImportTsv -
Dimporttsv.columns=HBASE_ROW_KEY,details:name, details:location, details:age -
Dimporttsv.separator=','
customer /abhilasha/customers.dat
```

Here, we specify the delimiter to be ',' in the file. This is the field separator for the records in the input file.



The data in the table after execution of the above command can be viewed using the **scan** command as follows:

