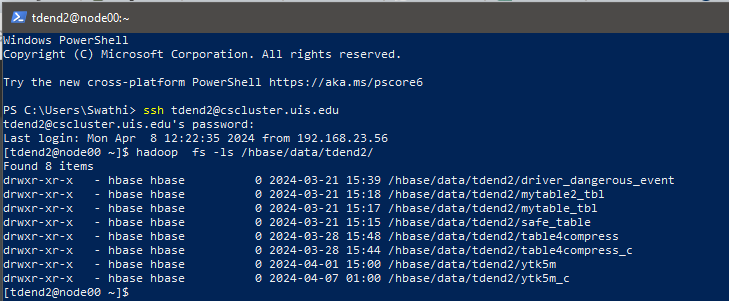
**NoSQL-CFDB-HBase-Hadoop**

**ASSIGNMENT#4:** 4. Table-level files

4.1. Table directories

Every table in HBase has its own directory, located under the namespaces under the HBase root directory in the HDFS. The directories of tables using the Hadoop command:

**hadoop fs -ls /hbase/data/tdend2/**



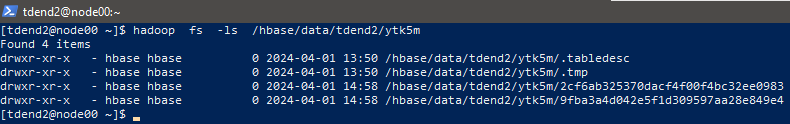
***>How many tables do your namespace have? I found 8 tables in namespace tdend2***

***------------------------------------------------------------------------------------------------------------------------------------------***

4.2. Table info and tmp directories

Each table directory contains a top-level file named ‘.tableinfo’ under the ‘.tabledesc’ directory, which stores the serialized HTableDescriptor for the table.

hadoop fs -ls /hbase/data/tdend2/ytk5m



5. Region-level files

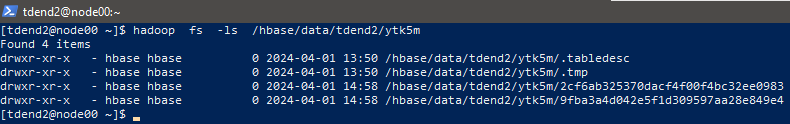
5.1. Region directories

Inside each table directory, there is a separate directory for every region comprising the table. The names of these directories are the MD5 hash portion of a region name, e.g. 0db1639806455f3c9503858cc79bf6fc

------------------------------------------------------------------------------------------------------------------------------------------

**ASSIGNMENT#3:** 5.2. Region name

hadoop fs -ls /hbase/data/tdend2/ytk5m



Found two regions as:

1)Region name : tdend2:ytk5m,125948559213021810688,1712001516129.2cf6ab325370dacf4f00f4bc32ee0983.

MD5 hash: 2cf6ab325370dacf4f00f4bc32ee0983

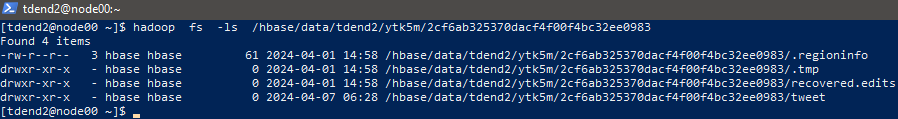
2)Region name : tdend2:ytk5m,,1712001516129.9fba3a4d042e5f1d309597aa28e849e4.

MD5 hash: **9fba3a4d042e5f1d309597aa28e849e4**

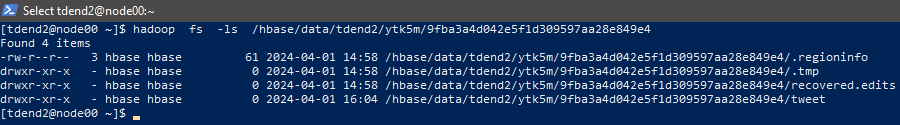
**-----------------------------------------------------------------------------------------------------------------------**

5.3. Inside Region

hadoop fs -ls /hbase/data/tdend2/ytk5m/**2cf6ab325370dacf4f00f4bc32ee0983**

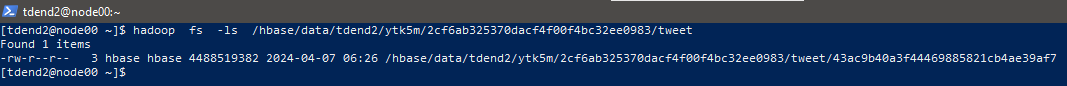


hadoop fs -ls /hbase/data/tdend2/ytk5m/ **9fba3a4d042e5f1d309597aa28e849e4**



**ASSIGNMENT#4:**  5.4. Column-family directory

hadoop fs -ls /hbase/data/tdend2/ytk5m/**2cf6ab325370dacf4f00f4bc32ee0983/tweet**



***The actual data file is 43ac9b40a3f44469885821cb4ae39af7 in the column family’tweet’ of the table ytk5m***

***------------------------------------------------------------------------------------------------------------------------------------------***

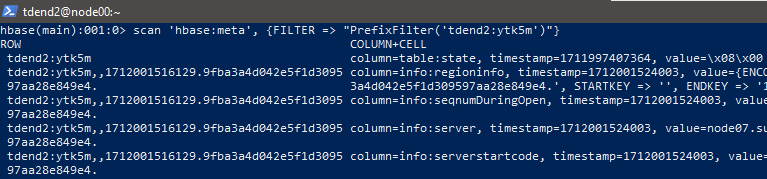
5.5. Region splits

Once the region needs to split because it has exceeded the maximum configured region size, a matching splits directory is created, which is used to stage the two new daughter regions. If this process is successful—usually this happens in a few seconds or less— they are moved up into the table directory to form the two new regions, each representing one-half of the original region.

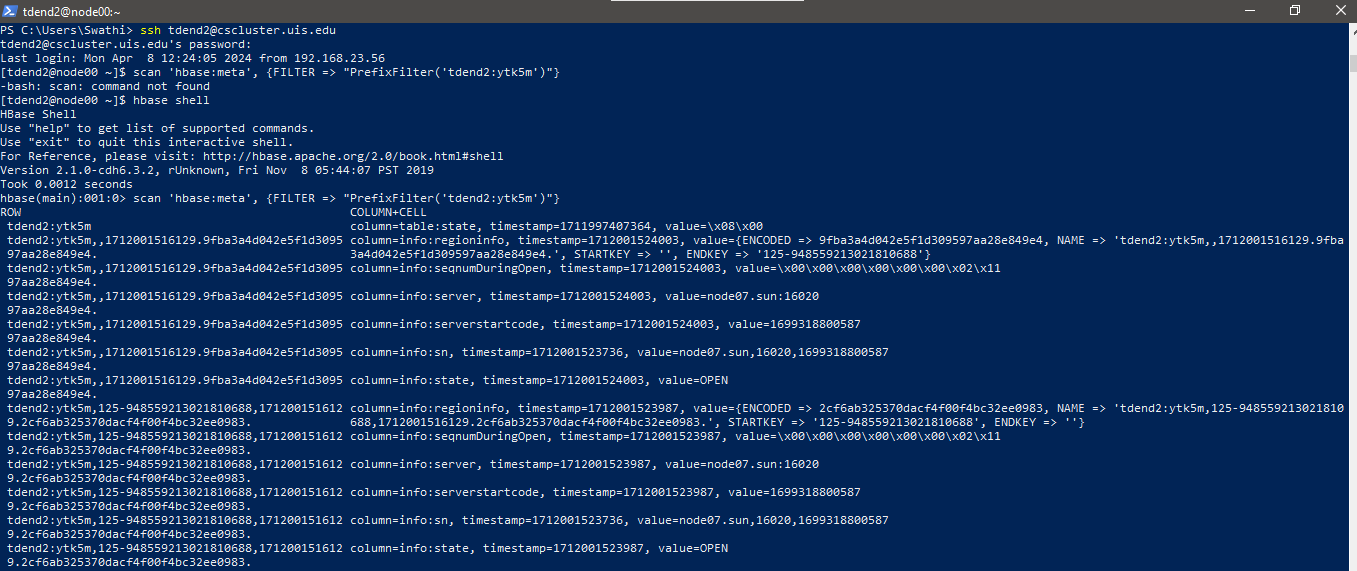
----------------------------------------------------------------------------------------------------------------------------------------

**Assignment#5:** 6. Region splits

**scan 'hbase:meta', {FILTER => "PrefixFilter('tdend2:ytk5m')"}**



OUTPUT:



I found 2 regions for table ytk5m as seen above and their information from 'hbase:meta' :

Region#1

REGION NAME => 'tdend2:ytk5m,,1712001516129.9fba3a4d042e5f1d309597aa28e849e4.',

STARTKEY => '', This is the first region of this table and so it is empty.

ENDKEY => '125-948559213021810688'

o MD5 hash portion of the name: 9fba3a4d042e5f1d309597aa28e849e4

Region server’s hostname is node07.sun

Region#2

REGION NAME => 'tdend2:ytk5m,125948559213021810688,1712001516129.2cf6ab325370dacf4f00f4bc32ee0983.',

STARTKEY => '125-948559213021810688',

ENDKEY => ' ' Note: This is the last region of this table and so it is empty and the start key of my next table ytk5m\_c is also ‘ ‘ and so it end key of this is ‘ ‘}

MD5 hash : 2cf6ab325370dacf4f00f4bc32ee0983

• Region server’s hostname: 'node05.sun'

======================================THE END=====================================