HBase as a Data Warehouse

|  |  |
| --- | --- |
| **Objective:** | Use HBase as a data warehouse |
| **Location of Files:** | ~/materials/data/ |
| **Expected Outcome:** | You will build HBase queries and see results |
| **Before You Begin:** | Go to the Ambari URL for your lab instance. Verify that the HBase services are up and it is running. |

**Create a table in HBase**

**Confirm HBase via the GUI**

1. Access web UI for status:

In a browser go to http://[ip-address]:60010/master-status on your lab instance

You will see detailed information about your HBase instance on this page:

**Create a table in HBase**

1. Now go into the command terminal

2. Create a new directory under ~/

$ cd [space]

$ mkdir hbase\_lab

$ cd hbase\_lab

3. Enter HBase Shell

[[username]@localhost ~]$ hbase shell

You should see the prompt:

hbase(main):001:0>

4. Use the list command to display tables in HBase:

hbase(main):001:0> list

TABLE

SLF4J: Class path contains multiple SLF4J bindings.

SLF4J: Found binding in [jar:file:/usr/lib/hadoop/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]

SLF4J: Found binding in [jar:file:/usr/lib/zookeeper/lib/slf4j-log4j12-1.6.1.jar!/org/slf4j/impl/StaticLoggerBinder.class]

SLF4J: See [http://www.slf4j.org/codes.html#multiple\_bindings](http://www.slf4j.org/codes.html%23multiple_bindings) for an explanation.

ambarismoketest

1 row(s) in 8.4080 seconds

**Question: how is the command line different in HBase shell and Hive shell?**

<in hive, the command ends with a semicolon;>

**What do you get when use semicolon at the end, such as list;**

<nothing returns, no error either>

5. create a table in HBase

hbase(main):007:0> create 'table3','ch3'

You should see the table created:

=> Hbase::Table - table3

6. Now check the tables:

hbase(main):008:0\* list

TABLE

ambarismoketest

table3

7. Verify the new table

hbase(main):012:0> describe 'table3'

DESCRIPTION ENABLED

'table3', {NAME => 'ch3', DATA\_BLOCK\_ENCODING => 'NONE', BLOOMFILTER => 'ROW', REPLICATI true ON\_SCOPE => '0', VERSIONS => '1', COMPRESSION => 'NONE', MIN\_VERSIONS => '0', TTL => '2147483647', KEEP\_DELETED\_CELLS => 'false', BLOCKSIZE => '65536', IN\_MEMORY => 'false', BLOCKCACHE => 'true'}

1 row(s) in 0.0750 seconds

Review the output and notice the column ('ch3'), parameter properties such as COMPRESSION, DATA\_BLOCK\_ENCODING, whether is kept in memory...

Quiz: what is going to happen if you issue below command in HBase Shell:

create table 'tablea','ch1';

<nothing, table2 not created>

**Do you see a table 'tablea' using the list command?**

<no>

8. To exit the HBase Shell:

:0> exit

**END**