CSP554—Big Data Technologies

Assignment #11

Worth: 5 points (1 point for each problem)

Due by the start of the next class period

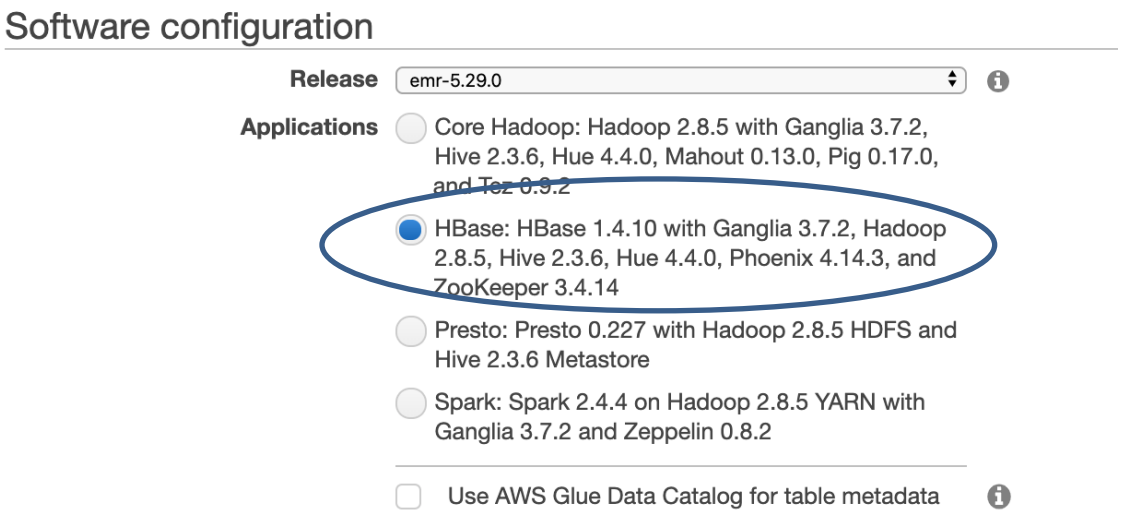
Assignments should be uploaded via the Blackboard portal.

Readings

NoSQL Distilled: Chapters 8 and 10

Starting HBase

Start up a Hadoop cluster as previously, but instead of choosing the "Core Hadoop" configuration chose the "HBase" configuration (see below), otherwise proceed as before:



Log on to the master Hadoop EC2 VM as per previous assignments and enter 'hbase shell' to start the HBase shell.

Note, when you start the HBase shell, your will see a long list of error messages. Ignore these.

There are HBase Cheat Sheets and HBase Shell Command documents on the blackboard in the "Free Books and Chapters" section that will help with the assignment.

Exercises

Exercise 1)

Create an HBase table with the following characteristics

Table Name: csp554Tbl

First column family: cf1

Second column family: cf2

Then execute the DESCRIBE command on the table and return command you wrote and the output as the results of this exercise.

Exercise 2)

Put the following data into the table created in exercise 1:

Row Key	Column Family	Column (Qualifier)	Value
Row1	cf1	name	Sam
Row2	cf1	name	Ahmed
Row1	cf2	job	Pilot
Row2	cf2	job	Doctor
Row1	cf2	level	LZ3
Row2	cf2	level	AR7

Execute the SCAN command on this table returning all rows, column families and columns. Provide the command and its result as the output of this exercise.

Exercise 3)

Using the above table write a command that will get the value associated with row (Row1), column family (cf2) and column/qualifier (level). Provide the command and its result as the output of this exercise.

Exercise 4)

Using the above table write command that will get the value associated with row (Row2), column family (cf1) and column/qualifier (name). Provide the command and its result as the output of this exercise.

Exercise 5)

Using the above table write a SCAN command that will return information about only two rows using the LIMIT modifier. Provide the command and its result as the output of this exercise.