CSP554—Big Data Technologies

Assignment #12 – Cassandra

Readings

Read Chapters 9 and 13 from our next book: Pramod J. Sadalage and Martin Fowler. 2012. *NoSQL Distilled: A Brief Guide to the Emerging World of Polyglot Persistence*. Addison-Wesley.(PS)

Worth: 12 points
Due by the start of the next class period

Assignments should be uploaded via the Blackboard portal

Exercise 1) (4 points)
Read the article "A Big Data Modeling Methodology for Apache Cassandra" available on the blackboard in the 'Articles' section. Provide a ½ page summary including your comments and impressions.

Exercise 2) (2 points)
a) Follow the instructions "Installing and Starting Cassandra" included on the blackboard with this assignment.

b) You might want to open two terminal windows and connect them (via ssh) to the Cassandra VM you set up in step "a". This will allow you to edit files in one window and start cqlsh in the other to execute them.

c) Create a file in your working directory called init.cql and enter the following commands. Use your IIT id as the name of your keyspace… For example, if your id is A1234567, then replace <IIT id> below with that value:

      CREATE KEYSPACE <IIT id> WITH REPLICATION = { 'class' : 'SimpleStrategy', 'replication_factor' : 1 };

d) Then execute this file in the CQL shell as follows…

      source './init.cql';

e) At this point you have created a keyspace unique to you. So make that keyspace the default by entering:

      USE <IIT id>;

Now create a file in your working directory called ex2.cql. In this file write the command to create a table named 'Music' with the following characteristics:

Attribute Name	Attribute Type	Primary Key / Cluster Key
artistName	text	Primary Key
albumName	text	Cluster Key
numberSold	int	Non Key Column
cost	int	Non Key Column

Execute ex2.cql. Then execute the shell command 'DESCRIBE TABLE Music' and include the output as the result of this exercise.

Exercise 3) (2 points)

Now create a file in your working directory called ex3.cql. In this file write the commands to insert the following records into table 'Music'...

artistName	albumName	numberSold	cost
Mozart	Greatest Hits	100000	10
Taylor Swift	Fearless	2300000	15
Black Sabbath	Paranoid	534000	12
Katy Perry	Prism	800000	16
Katy Perry	Teenage Dream	750000	14

a) Execute ex3.cql. Provide the content of this file as the result of this exercise.
b) Execute the command 'SELECT * FROM Music;' and provide the output of this command as another result of the exercise.

Exercise 4) (2 points)
Now create a file in your working directory called ex4.cql. In this file write the commands to query only Katy Perry songs. Execute ex4.cql. Provide the content of this file and result of executing this file as the result of this exercise.

Exercise 5) (2 points)
Now create a file in your working directory called ex5.cql. In this file write the commands to query only albums that have sold 700000 copies or more. Execute ex5.cql. Provide the content of this file and the result of executing this file as the result of this exercise.

Remember to terminate your Cassandra VM when you complete this assignment.