WQD7007 Big Data Management

Introduction to HBase

Introduction

- In this lab, we are going to practice how do we load data using HDFS and have simple practice using HBase.
- We will download a sample data and upload it to HDFS.
  - Download from here and extract. You should obtain **geolocation.csv** and **trucks.csv**
- For HBase, simple CRUD operation is carried out.
  - Hbase is not a relational database, therefore it does not support any structured query language like SQL.

Start HBase Shell

- Installation:
  - Refer to http://hbase.apache.org/book.html#quickstart
- Access the hbase shell: `hbase shell`

Create database

- Create table ‘test’ with column name ‘cf’:
  - `create 'test', 'cf'`
- Display table content
  - `list 'test'`
- Create (insert) data:
  - `put 'test', 'row1', 'cf:a', 'value1'`
  - `put 'test', 'row2', 'cf:b', 'value2'`
  - `put 'test', 'row3', 'cf:c', 'value3'`

Create and access “Contacts” Database

- `create 'Contacts', 'Personal', 'Office’`
- `put 'Contacts', '1000', 'Personal:Name', 'Taylor Swift'`
- `put 'Contacts', '1000', 'Personal:Phone', '603-3322 8883'`
- `put 'Contacts', '1000', 'Office:Phone', '1-425-000-0002'`
- `put 'Contacts', '1000', 'Office:Address', 'Centrepoint, Bandar Utama Malaysia’`
- `scan ‘contacts’`
- `put 'Contacts', '2000', 'Personal:Name', 'Ricky Martin'`
- `put 'Contacts', '2000', 'Personal:Phone', '603-640 7111'`
- `put 'Contacts', '2000', 'Office:Phone', '604-430 8288'`
- `put 'Contacts', '2000', 'Office:Address', '3730, Persiaran APEC, Cyberjaya'`
- `scan 'Contacts', {COLUMNS => ['Personal’]}`
- `scan 'Contacts', {COLUMNS => ['Personal:Name’]}`
- `get 'Contacts', '1000’`

- HOW TO ACCESS DATA FROM HDFS?