UNIVERSITY INSTITUTE of COMPUTING
Asia's Fastest Growing University

NAAC GRADE A+ ACCREDITED UNIVERSITY

Experiment No. 9th

Student Name: Rishav Kumar **UID: 22MCC20039**
Branch: MCA - CCD **Section/Group: 22MCD-1/ Grp A**
Semester: IV **Date of Performance: 3rd April 24**
Subject Name: Big Data & Analytics Lab **Subject Code: 22CAH-782**

1. **Aim/Overview of the practical:**

   Creation and alternation of table in the database in Hive.

2. **Code/Steps & Output for practical:**

In Apache Hive we can create tables to store structured data so that later on we can process it. The table in the hive is consists of multiple columns and records. The table we create in any database will be stored in the sub-directory of that database. The default location where the database is stored on HDFS is */user/hive/warehouse*. The way of creating tables in the hive is very much similar to the way we create tables in SQL. We can perform the various operations with these tables like Joins, Filtering, etc.

Below are the steps to launch a hive on your local system.

Step 1: Start all your Hadoop Daemon

```
start-dfs.sh                    # this will start namenode, datanode and secondary
namenode

start-yarn.sh                   # this will start node manager and resource manager

jps                             # To check running daemons
```

```
dikshant@dikshant:~$ hive
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/dikshant/apache-hive-3.1.2-bin/lib/log4j
-slf4j-impl-2.10.0.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/home/dikshant/hadoop/share/hadoop/common/lib/
slf4j-log4j12-1.7.25.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.apache.logging.slf4j.Log4jLoggerFactory]
Hive Session ID = 56c93e61-b0a2-4d3c-8dac-7a654477eff2

Logging initialized using configuration in jar:file:/home/dikshant/apache-hive-3
.1.2-bin/lib/hive-common-3.1.2.jar!/hive-log4j2.properties Async: true
Hive Session ID = 6114c8f6-6862-44be-a6bc-a8eaeba6cd67
Hive-on-MR is deprecated in Hive 2 and may not be available in the future versio
ns. Consider using a different execution engine (i.e. spark, tez) or using Hive
1.X releases.
hive>
```

Step 2: Launch hive from terminal

Creating Table in Hive

Let's create a database first so that we can create tables inside it. The command for creating a database is shown below.

Syntax To Make Database:

```
hive> CREATE DATABASE student_detail;
OK
Time taken: 0.341 seconds
hive> show databases;
OK
default
student_detail
Time taken: 0.037 seconds, Fetched: 2 row(s)
hive>
```

```
CREATE DATABASE <database-name>;
```

Syntax To Create Table in Hive

```
CREATE TABLE [IF NOT EXISTS] <table-name> (
<column-name>    <data-type>,
<column-name>    <data-type> COMMENT 'Your Comment',
.
)
COMMENT 'Add if you want'
LOCATION 'Location On HDFS'
ROW FORMAT DELIMITED
FIELDS TERMINATED BY ',';
```

Example:

```
CREATE TABLE IF NOT EXISTS student_data(

Student_Name STRING COMMENT 'This col. Store the name of student',
Student_Rollno INT COMMENT 'This col. Stores the rollno of student',
Student_Marks FLOAT)
ROW FORMAT DELIMITED
FIELDS TERMINATED BY ',';
```

We have successfully created the **table student_data** in our **student_detail database** with 3 different fields Student_Name, Student_Rollno, Student_Marks as STRING, INT, FLOAT respectively.

We can list down the table available in our database with the help of the command explained below.

Syntax & Command:

```
SHOW TABLES IN student_detail;
```

```
hive> CREATE TABLE IF NOT EXISTS student_data (
    > Student_Name STRING COMMENT 'This col. Store the name of student',
    > Student_Rollno INT COMMENT 'This col. Stores the rollno of student',
    > Student_Marks FLOAT)
    > ROW FORMAT DELIMITED
    > FIELDS TERMINATED BY ',';
OK
Time taken: 1.358 seconds
hive>
```

Hive provides us the functionality to perform Alteration on the Tables and Databases. *ALTER TABLE* command can be used to perform alterations on the tables.

1. Renaming Table Name

ALTER TABLE with RENAME is used to change the name of an already existing table in the hive.

Syntax:

```
ALTER TABLE <current_table_name> RENAME TO <new_table_name>;
```

```
hive> ALTER TABLE demo RENAME TO customer;
OK
Time taken: 0.135 seconds
hive> show tables;
OK
customer
Time taken: 0.021 seconds, Fetched: 1 row(s)
hive>
```

2. ADD Columns

Syntax:

ALTER TABLE <table_name> ADD COLUMNS (<col-name> <data-type> COMMENT ", <col-name> <data-type> COMMENT ", …..)

```
hive> ALTER TABLE customer ADD COLUMNS ( contact BIGINT COMMENT 'Store the custo
mer contact number');
OK
Time taken: 0.093 seconds
```

3. CHANGE Column

CHANGE in ALTER TABLE is used to change the name or data type of an existing column or attribute.

Syntax:

```
ALTER TABLE <table_name> CHANGE <column_name> <new_column_name> <new_data_type>;
```

```
hive> ALTER TABLE customer CHANGE demo_name customer_name STRING;
OK
Time taken: 0.131 seconds
hive> describe customer;
OK
customer_name           string
contact                 bigint                  Store the customer contact numbe
r
Time taken: 0.05 seconds, Fetched: 2 row(s)
hive>
```

```
hive> ALTER TABLE customer REPLACE COLUMNS (
    > customer_name STRING
    > );
OK
Time taken: 0.143 seconds
hive> describe customer;
OK
customer_name           string
Time taken: 0.068 seconds, Fetched: 1 row(s)
hive>
```

Learning outcomes (What I have learned):

a) Learned about Creation and alteration of tables in a database.