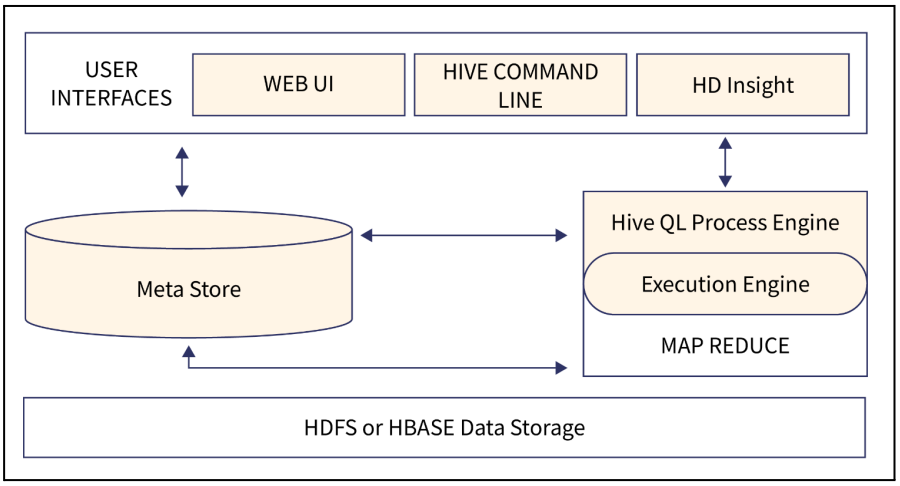
EXPERIMENT NO. 4

Aim : Execute HIVE commands to load, insert, retrieve, update, or delete table data.

Theory:

1. **Describe the Hive Architecture.**
2. **Hive Comparison with Traditional Database.**



Unit Name	Operation
User Interface	Hive is a data warehouse infrastructure software that can create interaction between user and HDFS. The user interfaces that Hive supports are Hive Web UI, Hive command line, and Hive HD Insight (In Windows server).
Meta Store	Hive chooses respective database servers to store the schema or Metadata of tables, databases, columns in a table, their data types, and HDFS mapping.
HiveQL Process Engine	HiveQL is similar to SQL for querying on schema info on the Metastore. It is one of the replacements of traditional approach for MapReduce program. Instead of writing MapReduce program in Java, we can write a query for MapReduce job and process it.
Execution Engine	The conjunction part of HiveQL process Engine and MapReduce is Hive Execution Engine. Execution engine processes the query and generates results as same as MapReduce results. It uses the flavor of MapReduce.

HDFS or HBASE	Hadoop distributed file system or HBASE are the data storage techniques to store data into file system.

Practical
Start CMD: [training@localhost ~]$ hive

1. Enter the command hive and will open the hive console.

```
hive> create database dj;
OK
Time taken: 3.038 seconds
```

2. All databases.

```
hive> show databases;
OK
```

3. Create the database dj using diff. query

```
hive> create database if not exists dj;
OK
Time taken: 0.0060 seconds
```

4.hive> show databases;

4.1 hive> describe database emp;

```
OK
emp             hdfs://localhost/user/hive/warehouse/emp.db
Time taken: 0.043 seconds
```

5. Want work under the same database.

```
hive> use dj;
OK
Time taken: 2.94 seconds
```

6. If you want to copy an external file(text/csv) use the following query to create a table.

Syntax To Create Table in Hive

```
CREATE TABLE [IF NOT EXISTS] <table-name> (
<column-name>   <data-type>,
<column-name>   <data-type>,
```

<column-name> <data-type>,
.
.
.
<column-name> <data-type>
)
COMMENT 'Add if you want'
LOCATION 'Location On HDFS'
ROW FORMAT DELIMITED
FIELDS TERMINATED BY ','

Note:

1. We can add a comment to the table as well as to each individual column.

2. ROW FORMAT DELIMITED shows that whenever a new line is encountered the new record entry will start.

3. FIELDS TERMINATED BY ‘,’ shows that we are using ‘ , ’ delimiter to separate each column.

4. We can also override the default database location with the *LOCATION* option.

7. Example

create table stud(rollno int,firstname string,lastname string,contactno float)
COMMENT 'djsce college student'
ROW FORMAT DELIMITED
FIELDS TERMINATED by ','
STORED AS TEXTFILE;

Describe stud;

8. find location of the table using the command (open http://localhost:50070/dfshealth.jsp on Browser)

Open new terminal and enter
hadoop dfs -ls /user/hive/warehouse

hadoop dfs -ls /user/hive/warehouse/dj.db

9.If you want to load data from a text /csv file into a table use the following command.

LOAD Command

```
hive> load data local inpath '/home/training/stud1.txt' into table stud1;
Copying data from file:/home/training/stud1.txt
Copying file: file:/home/training/stud1.txt
Loading data to table default.stud1
OK
Time taken: 0.635 seconds
```

9.1.hive> select * from tcs;

```
OK
123   Rathj  das     11192345
124   ram    lila    222123445
125   sham   sunder  33323412
```

10. Alter Table Statement

10.1 Rename To… Statement

The following query renames the table from **stud1** to **emp**.

```
hive> ALTER TABLE stud1  RENAME TO emp;
```

10.2 Change Statement

The following queries rename the column name and column data type using the above data:

```
hive> ALTER TABLE emp CHANGE name ename String;
hive> select * from emp;
OK
hive> describe aws;
hive> ALTER TABLE emp CHANGE salary salary Double;
```

10.3 Replace Statement

The following query deletes all the columns from the **emp** table and replaces it with **empid** and **name** columns:

```
hive> ALTER TABLE employee REPLACE COLUMNS ( eid INT empid Int, ename STRING name String);
```

11. Create table from an existing table

```
create table cc as select * from ab where rollno=122;
```

```
Select * from cc;
```

12. Drop the table structure

```
Drop table cc;
```

Conclusion: hence we study Hive Hadoop's ecosystem for SQL query processing.