**EXPERIMENT 02**

CLASS: BE CMPN A ROLL NO. : 19

NAME: REBECCA DIAS PID: 182027

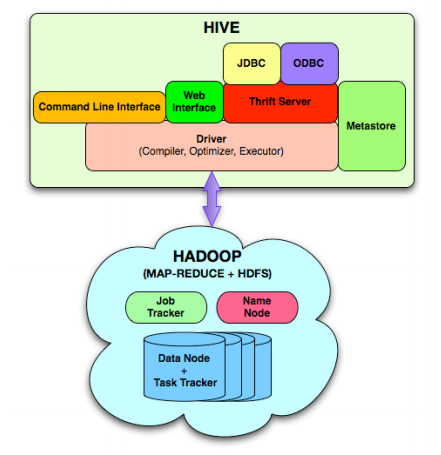
Aim:- To implement Hive commands in hadoop.

Theory :- Write about HIVE

Hive is a data warehouse infrastructure tool to process structured data in Hadoop. It resides on top of Hadoop to summarize Big Data, and makes querying and analyzing easy.

Apache Hive is an open source data warehouse software for reading, writing and managing large data set files that are stored directly in either the Apache Hadoop Distributed File System (HDFS) or other data storage systems such as Apache HBase. Hive enables SQL developers to write Hive Query Language (HQL) statements that are similar to standard SQL statements for data query and analysis. It is designed to make MapReduce programming easier because you don’t have to know and write lengthy Java code. Instead, you can write queries more simply in HQL, and Hive can then create the map and reduce the functions.

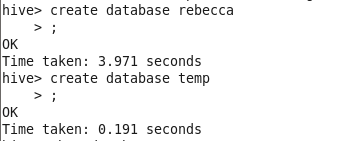
Included with the installation of Hive is the Hive metastore, which enables you to apply a table structure onto large amounts of unstructured data. Once you create a Hive table, defining the columns, rows, data types, etc., all of this information is stored in the metastore and becomes part of the Hive architecture. Other tools such as Apache Spark and Apache Pig can then access the data in the metastore.



Implementation:- Implement the following commands in HIVE

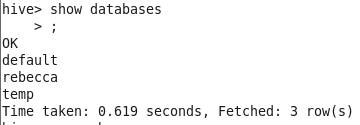
1. Creating a database in HIVE

Create database rebecca;



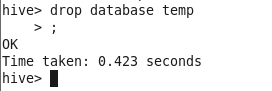
1. To display the existing databases

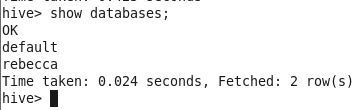
show databases;



1. To drop a database use

drop database temp;



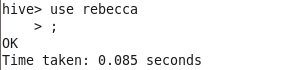


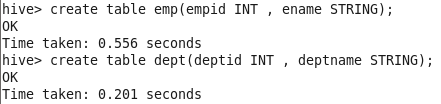
1. To create tables inside a database. First use that particular database and then use the same create table syntax for creating tables inside the database

use rebecca;

create table emp(empid INT, ename STRING);

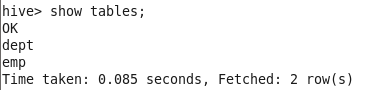
create table emp(empid INT, ename STRING) row format delimited fields terminated by ',';





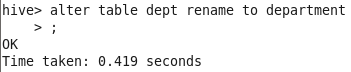
1. To display the tables in a particular database use the show tables syntax

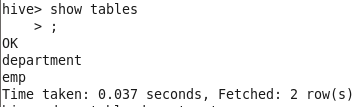
show tables;



1. In hive the table names can be altered. So the following command is used for altering the table name

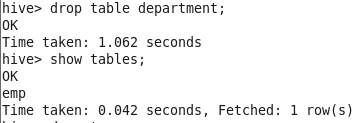
alter table dept rename to department;

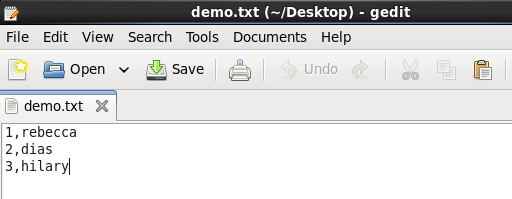




1. To drop tables use the following command. Drop table tablename

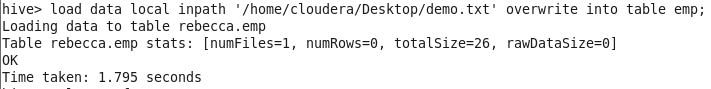
drop table department;



1. To create file in a local host and save it as demo.txt 
2. To load data from local path use the following command;

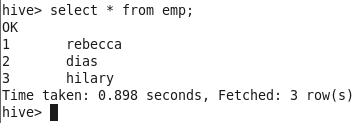
load data local inpath '/home/cloudera/Desktop/demo.txt' overwrite into table emp;





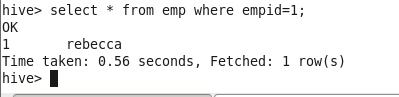
1. To view the contents of the table

Select \* from emp;



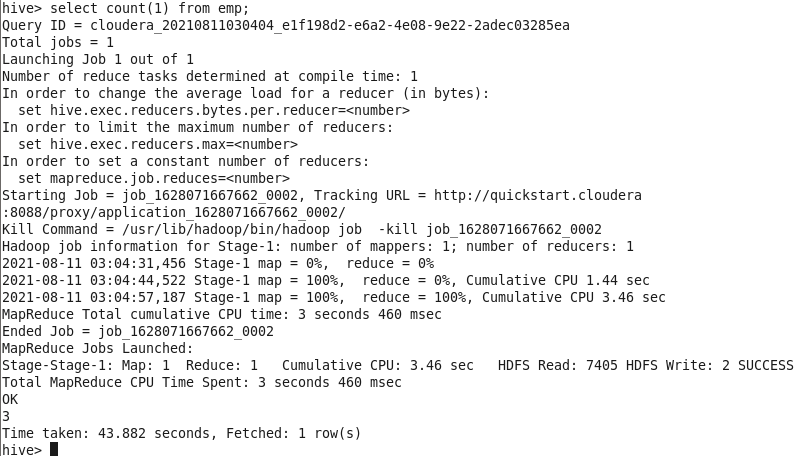
1. To give filter conditions

Select \* from emp where id=1;



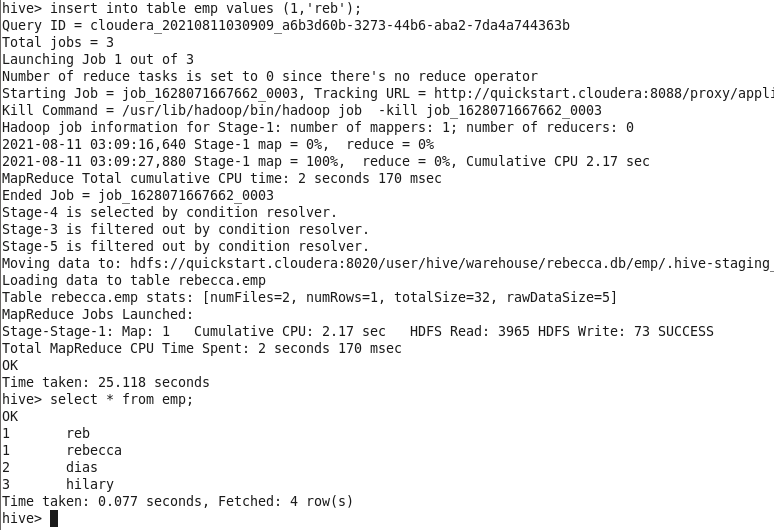
1. To calculate the number of rows in a table

select count(1) from emp;



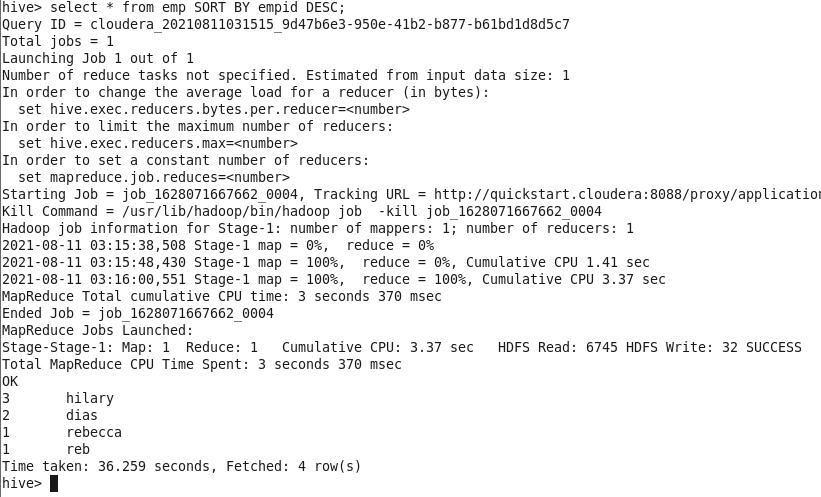
1. Insert into the tables

Insert into table\_name values (‘values’,1);



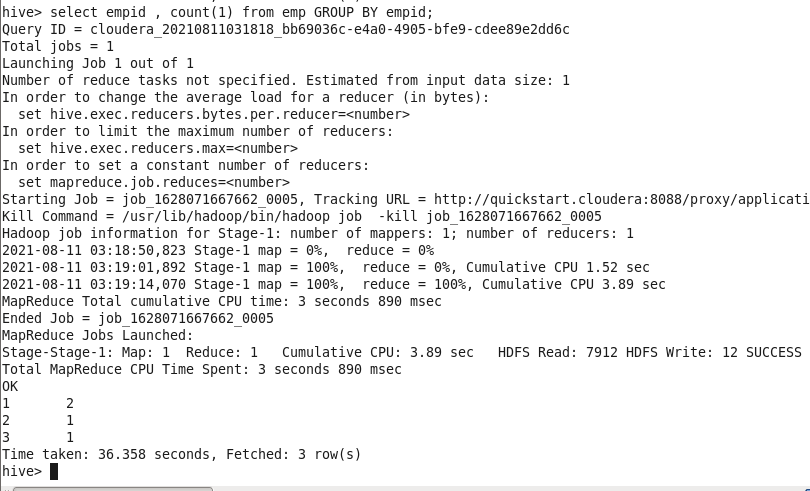
1. Sort by clause

Select \* from emp SORT BY empid DESC;



1. Group by Clause

Select empid , count(1) from emp GROUP BY empid;



Conclusion:-

In this experiment we learned the usage of Hive. We learnt the use of different commands for making databases, commands for reading and writing files, commands for making a directory and copying contents inside the created database.The usage of the databases used in hive was understood by us.