**Experiment 11: Week 12:**

11. Install and Run Hive then use Hive to create, alter, and drop databases, tables, views, functions, and indexes

**ALGORITHM:**

**Apache HIVE INSTALLATION STEPS**

1. Install MySQL-Server

Sudo apt-get install mysql-server

1. Configuring MySQL UserName and Password
2. Creating User and granting all Privileges

Mysql –uroot –proot

Create user <USER\_NAME> identified by <PASSWORD>

1. Extract and Configure Apache Hive

tar xvfz apache-hive-1.0.1.bin.tar.gz

1. Move Apache Hive from Local directory to Home directory
2. Set CLASSPATH in bashrc

Export HIVE\_HOME = /home/apache-hive

Export PATH = $PATH:$HIVE\_HOME/bin

1. Configuring hive-default.xml by adding My SQL Server Credentials

<property>

<name>javax.jdo.option.ConnectionURL</name>

<value>

jdbc:mysql://localhost:3306/hive?createDatabaseIfNotExist=true

</value>

</property>

<property>

<name>javax.jdo.option.ConnectionDriverName</name>

<value>com.mysql.jdbc.Driver</value>

</property>

<property>

<name>javax.jdo.option.ConnectionUserName</name>

<value>hadoop</value>

</property>

<property>

<name>javax.jdo.option.ConnectionPassword</name>

<value>hadoop</value>

</property>

8) Copying mysql-java-connector.jar to hive/lib directory.

**SYNTAX for HIVE Database Operations**

**DATABASE Creation**

CREATE DATABASE|SCHEMA [IF NOT EXISTS] <database name>

**DropDatabaseStatement**

DROP DATABASE StatementDROP (DATABASE|SCHEMA) [IF EXISTS]

database\_name [RESTRICT|CASCADE];

**Creating and Dropping Table in HIVE**

CREATE [TEMPORARY] [EXTERNAL] TABLE [IF NOT EXISTS] [db\_name.] table\_name

[(col\_name data\_type [COMMENT col\_comment], ...)]

[COMMENT table\_comment] [ROW FORMAT row\_format] [STORED AS file\_format]

**Loading Data into table log\_data**

Syntax:

**LOAD DATA LOCAL INPATH '<path>/u.data' OVERWRITE INTO TABLE u\_data;**

**Alter Table in HIVE**

Syntax:

ALTER TABLE name RENAME TO new\_name

ALTER TABLE name ADD COLUMNS (col\_spec[, col\_spec ...])

ALTER TABLE name DROP [COLUMN] column\_name

ALTER TABLE name CHANGE column\_name new\_name new\_type

ALTER TABLE name REPLACE COLUMNS (col\_spec[, col\_spec ...])

**Creating and Dropping View**

CREATE VIEW [IF NOT EXISTS] view\_name [(column\_name [COMMENT

column\_comment], ...) ] [COMMENT table\_comment] AS SELECT …

**Dropping View**

Syntax:

DROP VIEW view\_name

**Functions in HIVE**

String Functions:- round(), ceil(), substr(), upper(), reg\_exp() etc

Date and Time Functions:- year(), month(), day(), to\_date() etc

Aggregate Functions :- sum(), min(), max(), count(), avg() etc

**INDEXES**

CREATE INDEX index\_name ON TABLE base\_table\_name (col\_name, ...)

AS 'index.handler.class.name'

[WITH DEFERRED REBUILD]

[IDXPROPERTIES (property\_name=property\_value, ...)]

[IN TABLE index\_table\_name]

[PARTITIONED BY (col\_name, ...)]

[

[ ROW FORMAT ...] STORED AS ...

| STORED BY ...

]

[LOCATION hdfs\_path]

[TBLPROPERTIES (...)]

**Creating Index**

CREATE INDEX index\_ip ON TABLE log\_data(ip\_address) AS

'org.apache.hadoop.hive.ql.index.compact.CompactIndexHandler' WITH DEFERRED

REBUILD;

**Altering and Inserting Index**

ALTER INDEX index\_ip\_address ON log\_data REBUILD;

**Storing Index Data in Metastore**

**SET**

hive.index.compact.file=/home/administrator/Desktop/big/metastore\_db/tmp/index\_ipaddress\_result;

**SET**

hive.input.format=org.apache.hadoop.hive.ql.index.compact.HiveCompactIndexInputFormat;

**Dropping Index**

DROP INDEX INDEX\_NAME on TABLE\_NAME;

**INPUT**

Input as Web Server Log Data







