1 syntax to create a hive table

a) CREATE EXTERNAL TABLE IF NOT EXISTS TestingTable1

(

BUYER\_ID BIGINT,

ITEM\_ID BIGINT,

CREATED\_TIME STRING

);

2 how to load data in to hive table

This will do in two ways

1) hive> load data inpath "/user/ssake/HiveExample/source\_file " INTO table table\_name;

2) hadoopfs -copyFromLocal "/user/ssake/HiveExample/source\_file" /tables path

Note: by default table path "/user/hive/warehouse/table\_name"

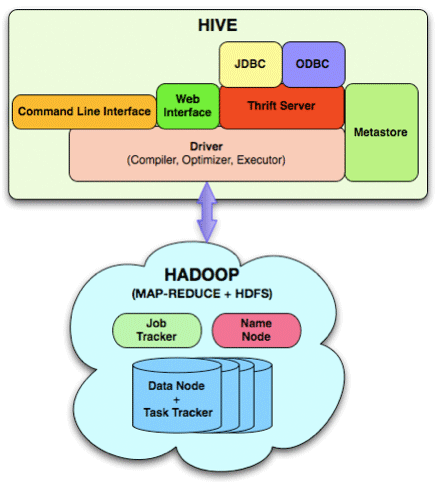
3 What is megastore in hive

a)meta data of the table information will store in SQL normally mysql .

4 what is default megastore database for hive

a)derby

5 hive architecture



6 what is thrift server

a) It is hive server need to start for java connectivity and pentaho report connectivity

7 how the table will store in hdfs (hadoop distributed file system)

a)folders

8 how the table partitions will store in hdfs

1. folders and sub folders in side that data will be placing like files csv files

**Joins**

9 what is inners join

Similar to SQL

Syntax example

SELECT a.\* FROM a JOIN b ON (a.id = b.id)

SELECT a.\* FROM a JOIN b ON (a.id = b.id AND a.department = b.department)

* SELECT a.val, b.val, c.val FROM a JOIN b ON (a.key = b.key1) JOIN c ON (c.key = b.key1)

10 what is outer join

1. similar to sql left outer join
2. SELECT a.val, b.val, c.val FROM a left outerJOIN b ON (a.key = b.key1) JOIN c ON (c.key = b.key1)
3. Right outer join

SELECT a.val, b.val, c.val FROM aright outer JOIN b ON (a.key = b.key1) JOIN c ON (c.key = b.key1)

11) Where condition in hive

SELECT a.key, a.value FROM a WHERE a.key in (SELECT b.key FROM );

Select employee\_no from employee where ( em\_dept=’it’ )and (emp\_sal>=300 and emp\_sal<500)

12)If condition in hive .

Select employee name from employee if(to\_date(emp\_creation date ) >= '2012-11-01' ,1,0))

13) What is udf in hive

a) User defined function custom logic has to implement in java addressed to hive to execute like map reduce program

14) How the hive Queries will be executing.

a) Query will convert in to mapreduceperogram and execute in parlal mode accross the cluster .

**Handling exceptions in hive**

15 There is webUi interface to track the error.

A)Log files and webui

16 port numnber for webui

1. http://localhost:50030--task tracker (data node)

http://localhost:50070 ---jobtacker (name node).

**Unix script knowledge required**

17 how to split the file

a)

18 how to capture the output of the hive Query to file

1. | sed -e 's/\t/,/g' >> /tmp/file\_name