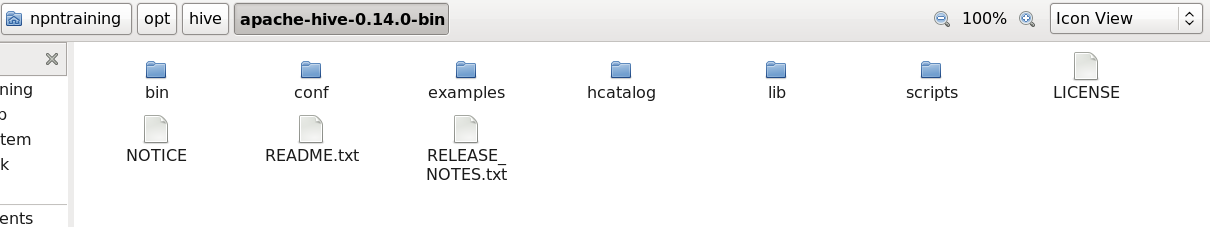
* **Hive**

**Hive home directory**

****

* + **Configuring Hive**

1. **Which is the default metastore in hive?**

Derby. It is nothing but a folder in file system

1. **How to Configure Hive metastore to derby?**

Audio path

I:\ProffessionalCareer\Technologies\BIGDATA\00\_NPNTraining\class5\20170722\_001-class5-1 (35:00 to 45:00)

Note: Goto Hive\_Home (npntraining/opt/hive/apache-hive-0.14.0-bin)/conf/hive-site.xml (if file is not there create it)

By default hive-site.xml file will not be there all default configurations will be there in hive-default.xml, If you want to override any property we have to create hive-site.xml.

1. Go to some file system location say desktop and create a folder by name HiveDB/ metastore\_db
2. Create hive-site.xml copy below 2 lines

<?xml version="1.0" encoding="UTF-8" standalone="no"?>

<?xml-stylesheet type="text/xsl" href="configuration.xsl"?>

1. Copy configuration tag

<configuration>

</configuration>

1. Go to hive-default.xml search for the property **javax.jdo.option.ConnectionURL** . Copy property tag to hive-site.xml and change database name to created folder as below

<value>jdbc:derby:;databaseName**=/home/npntraining/Desktop/HiveDB/metastore\_db**;create=true</value>

1. **What will happen if you do not configure metastore in hive-site.xml and keep the default?**

Where ever file system path you execute hive command derby database would have been created there.

1. **How to change hive-site.xml in cloudera/fusioninsight?**
   * **Commands/Configurations**
2. How to access hive command line interface(Hive CLI)?

**hive**

1. **How to come out of the Hive CLI?**

**quit;**

1. How will you execute other linux commads when you are in hive shell?

Using **!** for example: **!clear;**

* + - **HDFS**

1. **Where is default warehouse location of hive directory?**

Go to name node UI 🡪 Browse file system 🡪 usr/hive/warehouse

* + - **Databases**

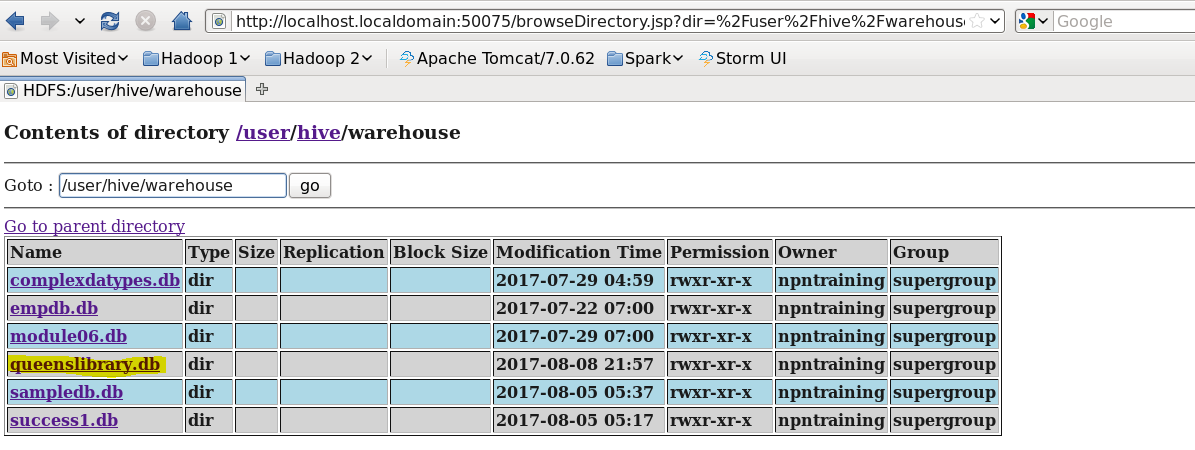
1. Command to list all databases?

**show databases;**

1. Command to create new database?

**create database queenslibrary;**

Once you create a database a folder will be created in hive warehouse location



1. Command to drop database?

**drop database if exists queenslibrary;**

1. To start using the database created ?

**use queenslibrary (database name);**

1. How to know which database user using?
2. **set hive.cli.print.current.db=true;** This is temporary solution.

Sets the property in current hive session only.

If one is out of the session the setting would be reset to default (false).

1. This is permanent solution add below property element in HIVE\_HOME/conf/hive-site.xml

**<property>**

**<name>hive.cli.print.current.db</name>**

**<value>true</value>**

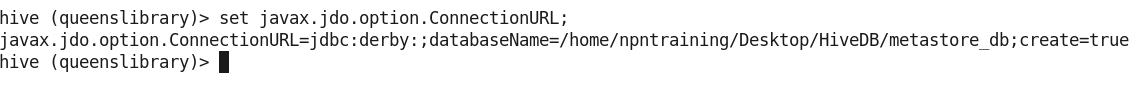
**</property>**

****

1. **How can you know the value of the property in hive?**

Using set command without giving a value for the property

Ex: set hive.cli.print.current.db;



* + - **Table**

1. command to create table?

create table employee(id int,name string,dept string,salary bigint)

row format delimited*Each and every row is formatted with a delimiter*

fields terminated by ','*Fields in the each row are terminated by comma, cannot have data separated by pipe or any other value*

lines terminated by '\n'

*Each line in the file is terminated by \n*

stored as TextFile;

*Files are stored as text file input format*

Once you created tables successfully a folder with the table name will be created inside database folder in hive warehouse folder.

1. Command to see list of tables created in databases?

**Show tables;**

1. How to list columns and columns types of the table?

**describe employee(Table name);**

1. How to see even more information about the table along with column and column type?

**describe extended employee;**

1. How to describe table in formatted way?

**describe formatted employee;**

* + - **Loading**
      * **Loading data from local file system to hive tables**

load data **local** inpath '/home/npntraining/Desktop/E-Learning/Hadoop/Modules/Module05/examples/01\_example/1.Employees.dat'overwrite into table employee;

* + - * **Loading the data from hdfs to hive tables**

First of all data should be loaded from local file system to hdfs using below command.

hadoop fs -put /home/npntraining/Desktop/E-Learning/Hadoop/Modules/Module05/examples/01\_example/3.Employees.dat /

Then loading data from HDFS to hive table

load data inpath '/3.Employees.dat' into table employee;

* + **Hive Query Language:**

It is similar to SQL

Select \* from employee;

For the above command there will be no map reduce

1. What happens when there is data type mismatch while querying data?

If there is data mismatch while querying null values will be returned. If there is empty line in file then all values will be null

Select count(\*) from employee;

As aggregation is involved in this Map reduce will happen here.

Note:

NPN training hive hands on examples:

/home/npntraining/Desktop/E-Learning/Hadoop/Modules/ Module05