

Ejercicio Sqoop: importar una tabla en MySQL a Hive

Creación de la tabla en MYSQL

1. Probamos la conexión con mysql
 - a. `Mysql -u root -p`
 - b. `Password: cloudera`
2. Vemos las bbdd que contiene
 - a. `Show databases;`
3. Creamos en MYSQL la table que queremos importar en hive
 - a. `create database pruebaadb;`
4. Creamos una tabla con datos que luego importaremos a hive mediante sqoop
 - a. `use pruebaadb;`
 - b. `create table tabla_prueba (nombre varchar(30), edad int);`
5. comprobamos que se ha creado
 - a. `show tables;`
6. Importamos algunas filas
 - a.

```
INSERT INTO tabla_prueba VALUES
("Alberto",22);

INSERT INTO tabla_prueba VALUES ("Luis",
23);

INSERT INTO tabla_prueba VALUES
("Pablo", 24);

INSERT INTO tabla_prueba VALUES
```

7. Comprobamos que los datos se han insertado en la tabla
 - a. `Select * from tabla_prueba;`
 - b. `Describe tabla_prueba;`

Creación de la tabla en HIVE

Creamos la tabla en hive donde se importarán los datos que acabamos de crear

1. Accedemos a hive
 - a. `hive`
2. Creamos una base de datos para esta prueba y accedemos a ella
 - a. `create database prueba_sqoop_hive;`
 - b. `use prueba_sqoop_hive;`
3. Comprobamos que está en el warehouse de hive
 - a. `hadoop fs -ls /user/hive/warehouse`
4. creamos la estructura de la table que contendrá los datos importados desde mysql con sqoop

a.

```
Drop table tabla_prueba_hive;  
CREATE TABLE tabla_prueba_hive  
(nombre string,  
edad int  
)  
ROW FORMAT DELIMITED  
STORED AS TEXTFILE;
```

5. Comprobamos que se ha creado con éxito

a. `Show tables;`

Importamos la tabla con SQOOP

1. Dado que la “bbdd” Accumulo no está configurada, abrimos un Shell y ejecutamos los siguientes comandos para evitar warnings molestos.

- a. `sudo mkdir /var/lib/accumulo`
- b. `ACCUMULO_HOME='/var/lib/accumulo'`
- c. `export ACCUMULO_HOME`

2. En un Shell escribimos lo siguiente para ver que sqoop está conectado con nuestro mysql:

- a. `sqoop list-databases --connect jdbc:mysql://localhost
--username root --password cloudera`

3. Ahora listamos la tabla “table_prueba” de la bdd “pruebadb” que hemos creado en MySQL

- a. `sqoop list-tables --connect
jdbc:mysql://localhost/pruebadb --username root --
password cloudera`

4. Usando los argumentos de importación hive mostrados en las slides del curso, importar la tabla creada en Mysql en la estructura creada en hive. Usar como conector (jdbc:mysql://localhost/bbddMysql) y un solo mapper.

- a. `sqoop import --connect
jdbc:mysql://localhost/pruebadb --table tabla_prueba
--username root --password cloudera -m 1 --hive-
import --hive-overwrite --hive-table
prueba_sqoop_hive.tabla_prueba_hive`