Bigdata Assignment – 4.3

Perform and explain the code flow and the associated result for the below tasks. Candidates should

create and use their own employee dataset for the same. Share the screenshot of the commands used

and its associated result.

● Transfer data between Mysql and HDFS (Import and Export) using Sqoop.

● Transfer data between Mysql and Hive (Import and Export only selected columns) using Sqoop.

Task 1 -

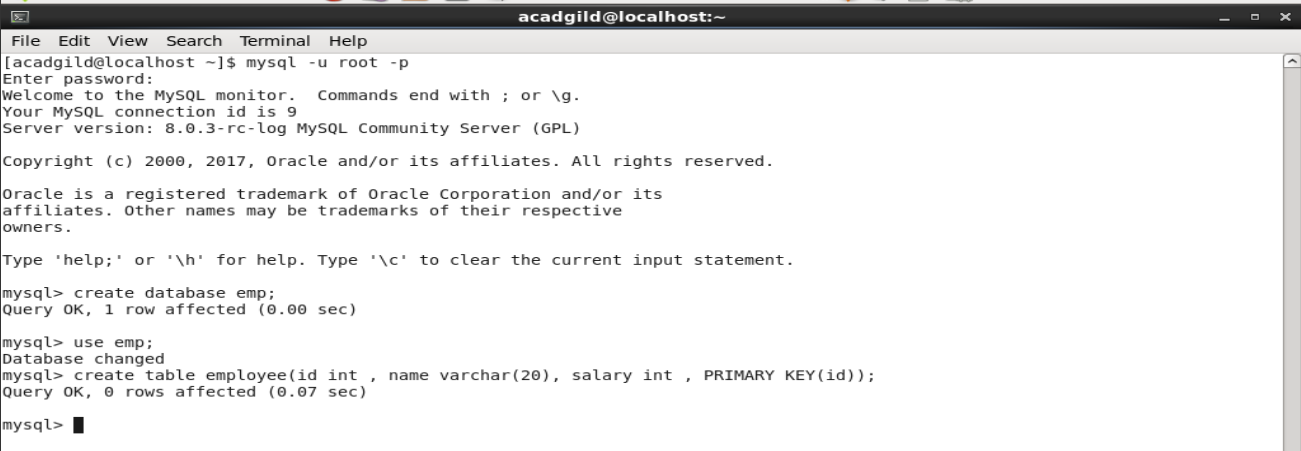
For importing data into hdfs from mysql

* Created a database and used it and it created a table.

**create database emp;**

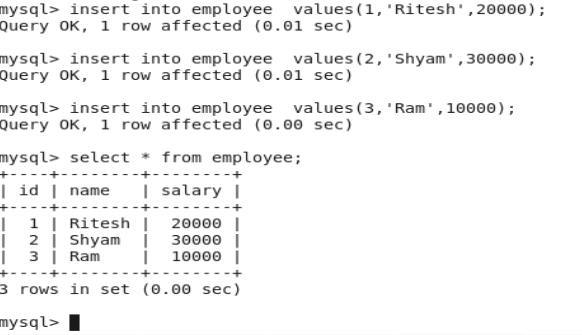
**use emp;**

**create table employee(id int, name varchar(20) , salary int , PRIMARY KEY(id));**



* Inserted data into the table.

**Insert into employee values(1,'Ritesh',20000);**



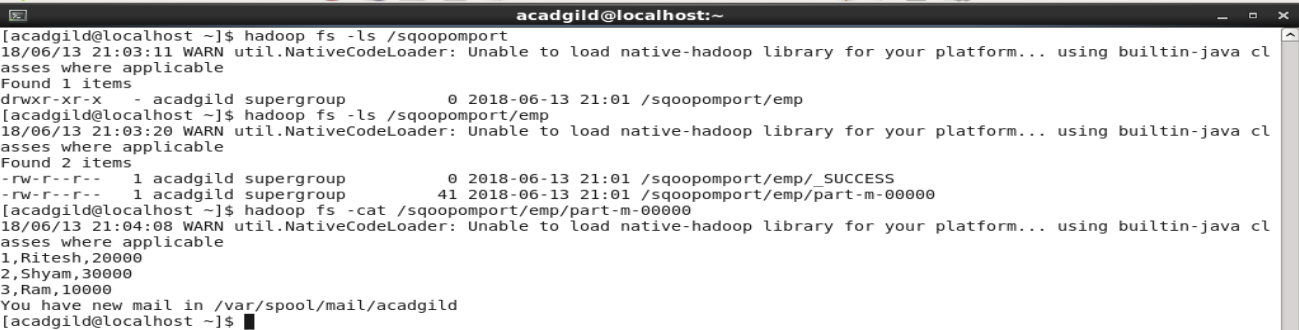
* Imported the mysql data into hdfs by using sqoop

**sqoop import –connect jdbc:mysql://localhost/emp –username root –password** [**Root@123**](mailto:Root@123) **–table employee –m 1 –target -dir /sqoopomport/emp**



* Then we checked whether the file is imported into hdfs or not, then its contents.

**hadoop fs -cat /sqoopomport/emp/part-m-00000**



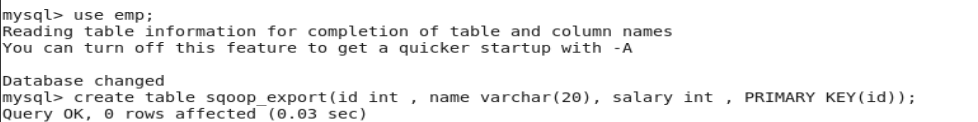
In the above screenshot it is visible that the same data we got in the hdfs

For exporting data into mysql from hdfs:

* Used the created database and created a table in it with same columns.

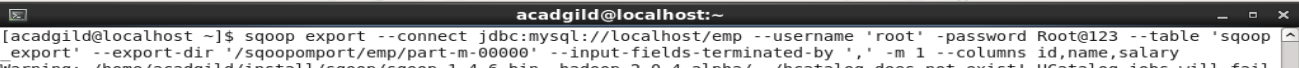
**use emp;**

**create table sqoop\_export(id int, name varchar(20) , salary int , PRIMARY KEY(id));**



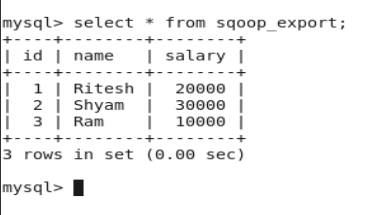
* Transfered the data from hdfs into mysql using sqoop

**sqoop export –connect jdbd:mysql://localhost/emp –username 'root' –password** [**Root@123**](mailto:Root@123) **–table 'sqoop\_export' –export-dir '/sqoopomport/emp/part-m-00000' –input-fields-terminated-by ',' -m 1 –columns id,name,salary**



* Checked the data of the table

**select \* from sqoop\_export ;**



**Task 2**

Transfer of data from mysql into hive:

* Tranfer of data using sqoop from mysql table which was created earlier into hive.

**sqoop import –connect jdbc:mysql://localhost/emp –username 'root' –pasword** [**Root@123**](mailto:Root@123) **–split-by id –columns id,name,salary –table employee –target-dir 'sqoopimport/emphive' –hive-import –create-hivetable –hive-table default.nysqlemployee -m 1**



* In hive , in the default database we checked the table is present or not , then the content of the table.

**show databases;**

**use default;**

**show tables;**

**select \* from mysqlemployee;**



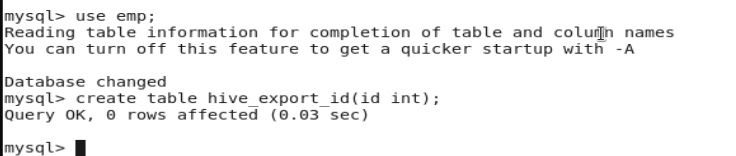
In the above screenshot , we can see the contents of the table same as in mysql table.

Transfer of data from hive into mysql:

* In mysql , in the emp database we created a table with id as only columns so tha we can export only id from hive.

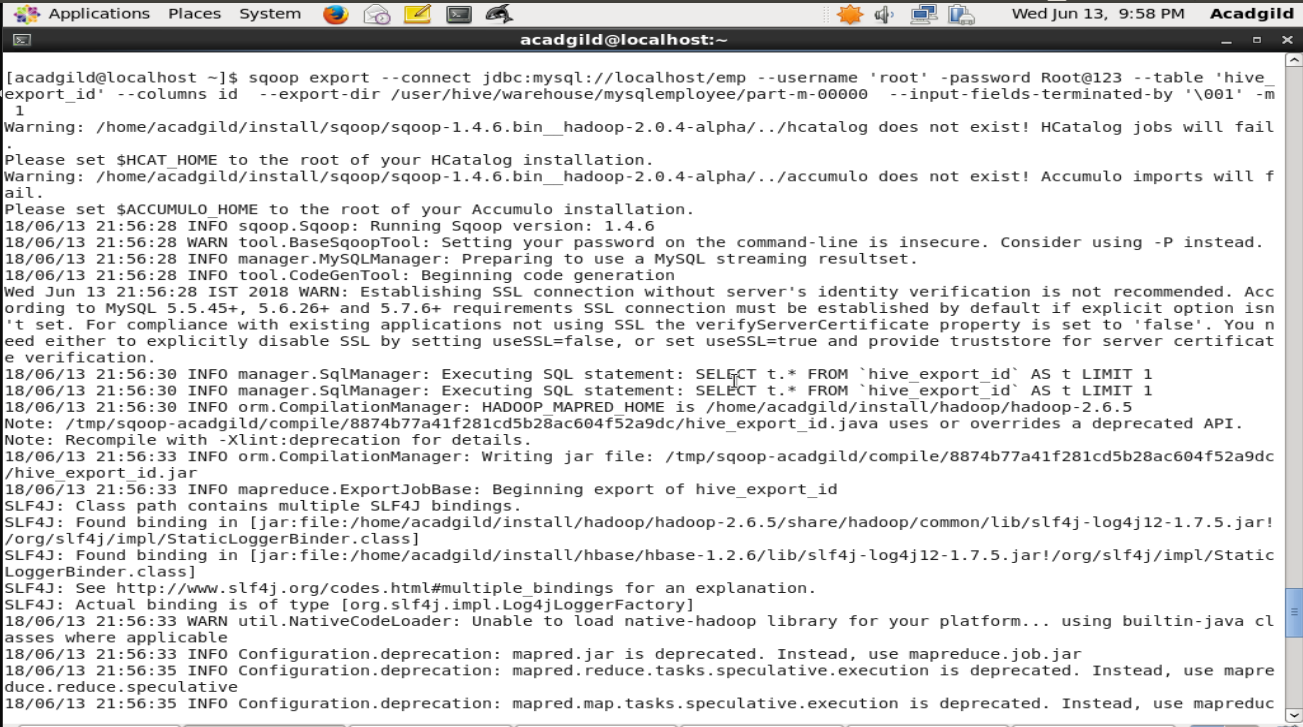
**use emp;**

**create table hive\_export\_id(id int);**



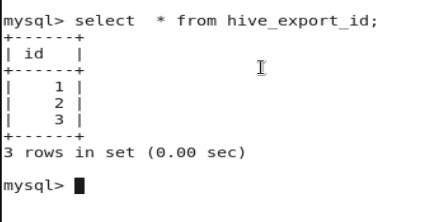
* We transfered the data from hive to mysql using sqoop.

**sqoop export –connect jdbc:mysql://localhost/emp –username 'root' –password** [**Root@123**](mailto:Root@123) **–table 'hive\_export' –columns id –export-dir /user/hive/warehouse/mysqlemployee/part-m-00000 –input-fields-terminated-by '\001' -m 1**



* In mysql , we checked the contents of the table where data is exported.

**select \* from hive\_export\_id;**



We can see we exported only id's from the hive table into mysql table.It is same as in the hive table.