Sqoop Cloudera

-----*** Install database to import data ***-----

$ sudo apt-get install mysql-server -y

-----*** Download sample database ***-----

$ wget http://www.mysqltutorial.org/wp-content/uploads/2018/03/mysqlsampledatabase.zip

$ sudo apt install zip -y

$ unzip mysqlsampledatabase.zip

-----*** Load sample data to database ***-----

$ sudo mysql -u root <mysqlsampledatabase.sql

$ sudo mysql -u root

$ mysql>show databases;

Change bind address of database

$ cd /etc/mysql/mysql.conf.d/

$ sudo nano mysqld.cnf
bind address = <private-ip of db>

$ sudo service mysql restart

-----*** Create a user for sqoop ***-----

mysql>create user 'sqoopuser'@'%' identified by 'password';

mysql>grant all privileges on *.* to 'sqoopuser'@'%';

-----*** Sqoop ***-----

Log in to one of the datanodes

$ sqoop help

Install driver

$ wget https://dev.mysql.com/get/Downloads/Connector-J/mysql-connector-j_8.0.32-1ubuntu20.04_all.deb

$ sudo dpkg -i mysql-connector-j_8.0.32-1ubuntu20.04_all.deb

Check connection to database

$ sqoop list-databases --connect jdbc:mysql://<database_dns>:3306/ --username sqoopuser -P

$ sqoop list-tables --connect jdbc:mysql://<database_dns>:3306/classicmodels --username sqoopuser -P

-----*** Sqoop Import ***-----

sqoop import --connect jdbc:mysql://<database_dns>:3306/classicmodels --username sqoopuser --password password --table employees -m 1

sqoop import --connect jdbc:mysql://<database_dns>:3306/classicmodels --username sqoopuser --password password --table employees -m 1 --target-dir /user/hdfs/new/

-----*** Hive Inport ***-----

>>> Switch to hive user

$ sqoop import --connect jdbc:mysql://<database_dns>:3306/classicmodels --username sqoopuser --password password --table employees --hive-import --hs2-url "jdbc:hive2://<hs2-dns>:10000"

-----*** Sqoop Export ***-----

NOTE : It is mandatory that the table to be exported is created manually and is present in the database from where it has to be exported.

Create table in destination database

mysql -u root -p

mysql>use classicmodels;

mysql>create table export (a INT, b VARCHAR(20), c CHAR(40));

###

$ nano abc.csv
1,sam,6th street
2,drake,turing road
3,tim, baker street

$ hdfs dfs -put abc.csv /user/hdfs/exportdir

$ sqoop export --connect jdbc:mysql://<database_dns>:3306/classicmodels --username sqoopuser --password password --table export --export-dir /user/hdfs/exportdir/

Check data in database

mysql -u root -p

mysql> select * from export;