Mysql Connection:

mysql -uroot -pcloudera

Sqoop version

Sqoop connect with mysql:

sqoop import --driver com.mysql.jdbc.Driver --connect jdbc:mysql://localhost/users --username root --password cloudera --table user

Importing the all tables:

sqoop import --driver com.mysql.jdbc.Driver --connect jdbc:mysql://localhost/users --username root --password cloudera --table user

UDEMY URL:

https://www.udemy.com/apache-sqoop-for-certifications-cca-and-hdpcd/learn/v4/t/lecture/8631410?start=765

Sqoop User Guide:

https://sqoop.apache.org/docs/1.4.6/SqoopUserGuide.html#\_connecting\_to\_a\_database\_server

Import statement:

sqoop import --connect jdbc:mysql://10.0.0.41:3306/users --username root --password cloudera --table user --target-dir = /user/root/data/

Still Running and no output

In ruuning state still accepted only:

The application might not be running yet or there is no Node Manager or Container available. This page will be automatically refreshed.

List the yarn cluster:

yarn node -list

Start the yarn cluster:

sudo service hadoop-yarn-resourcemanager start

sudo service hadoop-yarn-nodemanager start

sudo service hadoop-mapreduce-historyserver start

Delete the existing directory:

sqoop import --connect jdbc:mysql://localhost:3306/users --username root --P --table student --delete-target-dir --target-dir /user/root --m 2

M-using for giving the no of reducers in jobs.

Applying the where clause in Sqoop command:

sqoop import --connect jdbc:mysql://localhost:3306/users --username root --P --table student --where "s\_id = 10" --delete-target-dir --target-dir /user/root --m 2

Append the values in particular existing folder:

–incremental <mode>

–check-column <column name>

–last value <last check column value>

sqoop import --connect jdbc:mysql://localhost:3306/users --username root --P --table student --target-dir /user/root --incremental append --check-column s\_id --last-value 39 --m 2

Import all tables:

sqoop import-all-tables --connect jdbc:mysql://localhost:3306/users --username root --P

List the tables and databases:

sqoop list-tables --connect jdbc:mysql://localhost:3306/users --username root --P

sqoop list-databases --connect jdbc:mysql://localhost:3306/users --username root --P

Importing the selected column only:

sqoop import --connect jdbc:mysql://localhost:3306/users --username root --P --table student --columns "s\_id,s\_name" --where "s\_id > 10" --delete-target-dir --target-dir /user/root --m 2

AVRO file format:

sqoop import --connect jdbc:mysql://localhost:3306/users --username root --P --table student --columns "s\_id,s\_name" --where "s\_id > 10" --delete-target-dir --target-dir /user/root --as -avrodatafile

Parquet file format:

sqoop import --connect jdbc:mysql://localhost:3306/users --username root --P --table student --columns "s\_id,s\_name" --where "s\_id > 10" --delete-target-dir --target-dir /user/root --as-parquetfile

Delimeter example:

sqoop import --connect jdbc:mysql://localhost:3306/users --username root --P --table student --columns "s\_id,s\_name" --where "s\_id > 10" --delete-target-dir --target-dir /user/root --fields-terminated-by '\t' --lines-terminated-by '\n' \--optionally-enclosed-by '\"' -m1

Using data import from mysql to Hive:

sqoop import --connect jdbc:mysql://localhost:3306/users --username cloudera --P --table customer1 --m 2 --hive-import --create-hive-table --hive-table sqoop\_db.student1