**Sqoop Import statement:**

sqoop import \

--connect jdbc:mysql://cxln2:3306/sqoopex \

--username 'sqoopuser' \

--password 'NHkkP876rp' \

--query 'select \* from sqoopex.ssv20230313 where $CONDITIONS' \

--target-dir hdfs://cxln1.c.thelab-240901.internal:8020/user/bigdatacloudxlab14968/ssv20230313/mysql\_sqoopex\_ssv20230314 \

--fields-terminated-by '\t' \

--lines-terminated-by '\n' \

--m 1

**Explanation:**

--connect -> to connect to mysql server

--username -> username of the mysql server used

--password -> password of the mysql server used

--query -> query to fetch the data from table in mysql server

--target-dir -> the path of the directory where you want to store the data in hdfs (we can fetch it from ambari in cloudx)

Steps to import data:

* Create a directory with your name like ssv20230313 in hdfs
* Go to mysql terminal
* Show databases;
* Use sqoopex;
* Create table ssv20230314(Firstname varchar(50), Lastname varchar(50));
* Insert into ssv20230314 Values((‘saisri’,’vemulapalli’),(‘ashok’,’bagam’));
* Select \* from ssv20230314;
* Go back to cloudx terminal and copy paste the sqoop import code there
* If it executes successfully then go to your directory and check if the data is imported successfully or not.
* Read the data from the file in your hdfs using the following command:

**hdfs dfs -ls /user/bigdatacloudxlab14968/ssv20230313/mysql\_sqoopex\_ssv20230314**

* Read the data from the file using the following command:

**hdfs dfs -cat /user/bigdatacloudxlab14968/ssv20230313/mysql\_sqoopex\_ssv20230314/part-m 00000/**

* If there is data in that file, then the data is displayed on the screen.

**SQOOP EXPORT:**

* We can do it in two ways. One is exporting data directly from Hadoop cluster to MySQL and the other way is to load the data from Hadoop to hive table and then export it to MySQL.
* Exporting data from Hadoop to hive and hive to MySQL.
* Copy the data from Hadoop cluster to local using copyToLocal command.

**hdfs dfs -copyToLocal /user/bigdatacloudxlab14968/ssv20230314/mysql\_sqoopex\_ssv14032023/** **part-m-00000**

* Go to hive.
* Create a new database with your name if there is no existing database.
* **Use database saisri;**
* **Create table ssv20230314\_test(Firstname varchar(50), Lastname varchar(50));**
* Load the data to hive table using the following command

**LOAD DATA LOCAL INPATH "part-m-00000" INTO TABLE ssv20230314\_test;**

* **Select \* from ssv20230314\_test;**
* It displays the data from hive table if the data is loaded successfully.
* Exit from hive and go back to CloudEx web console
* Export the data to MySQL command using the following command:

**sqoop export \**

**--connect jdbc:mysql://cxln2:3306/sqoopex \**

**--username 'sqoopuser' \**

**--password 'NHkkP876rp' \**

**--table ssv20230314\_test1 \**

**--export-dir /apps/hive/warehouse/saisri.db/ssv20230314\_test2 \**

**--input-fields-terminated-by '\t' \**

**--input-lines-terminated-by '\n' \**

**--m 1**

* If the export is successful, then it displays as follows:

**Exported 2 records.**

**Text

Description automatically generated**

* Go to MySQL and check if the data is exported to mysql table or not.

**Select \* from ssv20230314\_test1;**

**Text

Description automatically generated**

**To export data from Hadoop cluster to mysql:**

**sqoop export \**

**--connect jdbc:mysql://cxln2:3306/sqoopex \**

**--username 'sqoopuser' \**

**--password 'NHkkP876rp' \**

**--table ssv20230314\_test2 \**

**--export-dir hdfs://cxln1.c.thelab-240901.internal:8020/user/bigdatacloudxlab14968/ssv20230314/mysql\_sqoopex\_ssv14032023 \**

**--input-fields-terminated-by '\t' \**

**--input-lines-terminated-by '\n' \**

**--m 1**

**Text

Description automatically generated**

**Text

Description automatically generated**

**To fetch mysql hostname details:**

**SELECT SUBSTRING\_INDEX(USER(), '@', -1) AS ip, @@hostname as hostname, @@port as port, DATABASE() as current\_database;**

**To check disk space in cloudex:**

**hdfs dfs -du -h /user/bigdatacloudxlab14968/.Trash**

**To remove file or directory which is not in use to clear the disk space:**

**hdfs dfs -rm -r /user/bigdatacloudxlab14968/.Trash**