**Sqoop import from MySQL to hdfs:**

sqoop import \

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

--username 'sqoopuser' \

--password 'NHkkP876rp' \

--query 'select \* from sqoopex.ssv20230316 where $CONDITIONS' \

--target-dir hdfs://cxln1.c.thelab-240901.internal:8020/user/bigdatacloudxlab14968/ssv20230316/mysql\_sqoopex\_ssv20230316 \

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

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

--m 1

Text

Description automatically generated

**Result in hdfs:**

**Text

Description automatically generated**

**Sqoop export from hdfs to MySQL:**

sqoop export \

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

--username 'sqoopuser' \

--password 'NHkkP876rp' \

--table ssv20230316\_1 \

--export-dir hdfs://cxln1.c.thelab-240901.internal:8020/user/bigdatacloudxlab14968/ssv20230316/mysql\_sqoopex\_ssv20230316 \

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

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

--m 1

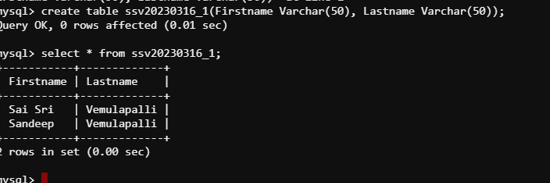
Text

Description automatically generated

Text

Description automatically generated

**Result in MySQL Table:**

****

**Sqoop export using null values in lastname:**

**hdfs dfs -text /user/bigdatacloudxlab14968/ssv20230316/mysql\_sqoopex\_ssv20230316/part-m-00000 | sed -e 's/Vemulapalli/NULL/g; s/Vemulapalli/NULL/g' | hdfs dfs -put - /user/bigdatacloudxlab14968/ssv20230316/mysql\_sqoopex\_ssv20230316/partout**

**hdfs dfs -cat /user/bigdatacloudxlab14968/ssv20230316/mysql\_sqoopex\_ssv20230316/partout**

**sqoop export \**

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

**--username 'sqoopuser' \**

**--password 'NHkkP876rp' \**

**--table ssv20230316\_test \**

**--export-dir hdfs://cxln1.c.thelab-240901.internal:8020/user/bigdatacloudxlab14968/ssv20230316/mysql\_sqoopex\_ssv20230316/partout \**

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

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

**--m 1**

***Results in MySQL Table:***

**Text

Description automatically generated**

**Importing dat from MySQL to hdfs using pyspark as csv file:**

**To import pyspark jar files:**

**pyspark --driver-class-path /usr/share/java/mysql-connector-java.jar**

**Graphical user interface, text

Description automatically generated**

**Pyspark Code:**

**from pyspark import SparkConf,SparkContext**

**from pyspark.sql import SQLContext**

**conf=SparkConf().setAppName("Dataimporting")**

**sc=SparkContext.getOrCreate(conf=conf)**

**sqlContext=SQLContext(sc)**

**url="jdbc:mysql://cxln2:3306/sqoopex"**

**table="ssv20230316"**

**user="sqoopuser"**

**password="NHkkP876rp"**

**target\_dir="/user/bigdatacloudxlab14968/ssv20230316/SSV\_Pyspark\_Mysql\_import"**

**jdbc\_properties={"user":user, "password":password, "driver":"com.mysql.jdbc.Driver"}**

**df=sqlContext.read.jdbc(url=url, table=table, properties=jdbc\_properties)**

**df.write.csv(target\_dir)**

**Text

Description automatically generated**

**Result in hdfs:**

**A picture containing text

Description automatically generated**

**Importing dat from MySQL to hdfs using pyspark as parquet file:**

**pyspark --driver-class-path /usr/share/java/mysql-connector-java.jar**

**from pyspark import SparkConf,SparkContext**

**from pyspark.sql import SQLContext**

**conf=SparkConf().setAppName("Dataimporting")**

**sc=SparkContext.getOrCreate(conf=conf)**

**sqlContext=SQLContext(sc)**

**url="jdbc:mysql://cxln2:3306/sqoopex"**

**table="ssv20230316"**

**user="sqoopuser"**

**password="NHkkP876rp"**

**target\_dir="/user/bigdatacloudxlab14968/ssv20230316/SSV\_Pyspark\_Mysql\_import/parquet"**

**jdbc\_properties={"user":user, "password":password, "driver":"com.mysql.jdbc.Driver"}**

**df=sqlContext.read.jdbc(url=url, table=table, properties=jdbc\_properties)**

**df.write.parquet(target\_dir)**

**Text

Description automatically generated**

**Result in hdfs as parquet file: A picture containing text, indoor, screenshot

Description automatically generated**

**Exporting data from hdfs to MySQL using PySpark:**

**from pyspark import SparkConf,SparkContext**

**from pyspark.sql import SQLContext**

**conf=SparkConf().setAppName("Dataexporting")**

**sc=SparkContext.getOrCreate(conf=conf)**

**sqlContext=SQLContext(sc)**

**url="jdbc:mysql://cxln2:3306/sqoopex"**

**table="ssv20230316\_test1"**

**user="sqoopuser"**

**password="NHkkP876rp"**

**source\_dir="/user/bigdatacloudxlab14968/ssv20230316/SSV\_Pyspark\_Mysql\_import/part-00000-181aae3c-3ce1-496b-9474-00f8e745e632.csv"**

**jdbc\_properties={"user":user, "password":password, "driver":"com.mysql.jdbc.Driver"}**

**df=sqlContext.read.format('csv').option('header', True).load(source\_dir)**

**df.write.jdbc(url=url, table=table, mode='overwrite', properties=jdbc\_properties)**

**df.show()**

**Text

Description automatically generated**

**Result in MySQL Table:**

**Text

Description automatically generated**