Assignment 6 -SQOOP

Task 1

Use Sqoop tool to export data present in SQOOPOUT folder made while demo of Import table

Solution:

Here we need to export file which has data in table format in HDFS to a table in MySQL.

First checked for the content of the file placed in HDFS Location as shown below:

```
File  Edit  View  Search  Terminal  Help
[acadgild@localhost ~]$ hadoop fs -ls /sqoopout/queryoutput/
18/05/16 01:20:22 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
Found 5 items
-rw-r--r--   1 acadgild supergroup          0 2018-05-16 00:52 /sqoopout/queryoutput/_SUCCESS
-rw-r--r--   1 acadgild supergroup         24 2018-05-16 00:52 /sqoopout/queryoutput/part-m-00000
-rw-r--r--   1 acadgild supergroup         35 2018-05-16 00:52 /sqoopout/queryoutput/part-m-00001
-rw-r--r--   1 acadgild supergroup         37 2018-05-16 00:52 /sqoopout/queryoutput/part-m-00002
-rw-r--r--   1 acadgild supergroup         42 2018-05-16 00:52 /sqoopout/queryoutput/part-m-00003
[acadgild@localhost ~]$ hadoop fs -cat /sqoopout/queryoutput/part*
18/05/16 01:21:20 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
1,Shyam,Ram,Patna,Bihar
2,Tanya,Priya,Whitefiled,Bangalore
3,James,Brown,New York,United States
4,Jhon,Miller, Los Angeles ,United States
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$ █
```

Then Created a table Person1 in MYSQL Simplidb database with column in same order in hdfs output shown in screen shot below :

```
mysql> use simplidb
Database changed
mysql> Create table Person1 ( person_id INT NOT NULL, lname varchar(20), fname v
archar(20), area varchar(20), PRIMARY KEY (person_id) );
Query OK, 0 rows affected (0.97 sec)

mysql> select * from Person1;
Empty set (0.13 sec)

mysql> describe Person1;
+-----------+-------------+------+-----+---------+-------+
| Field     | Type        | Null | Key | Default | Extra |
+-----------+-------------+------+-----+---------+-------+
| person_id | int(11)     | NO   | PRI | NULL    |       |
| lname     | varchar(20) | YES  |     | NULL    |       |
| fname     | varchar(20) | YES  |     | NULL    |       |
| area      | varchar(20) | YES  |     | NULL    |       |
+-----------+-------------+------+-----+---------+-------+
4 rows in set (0.11 sec)

mysql> █
```
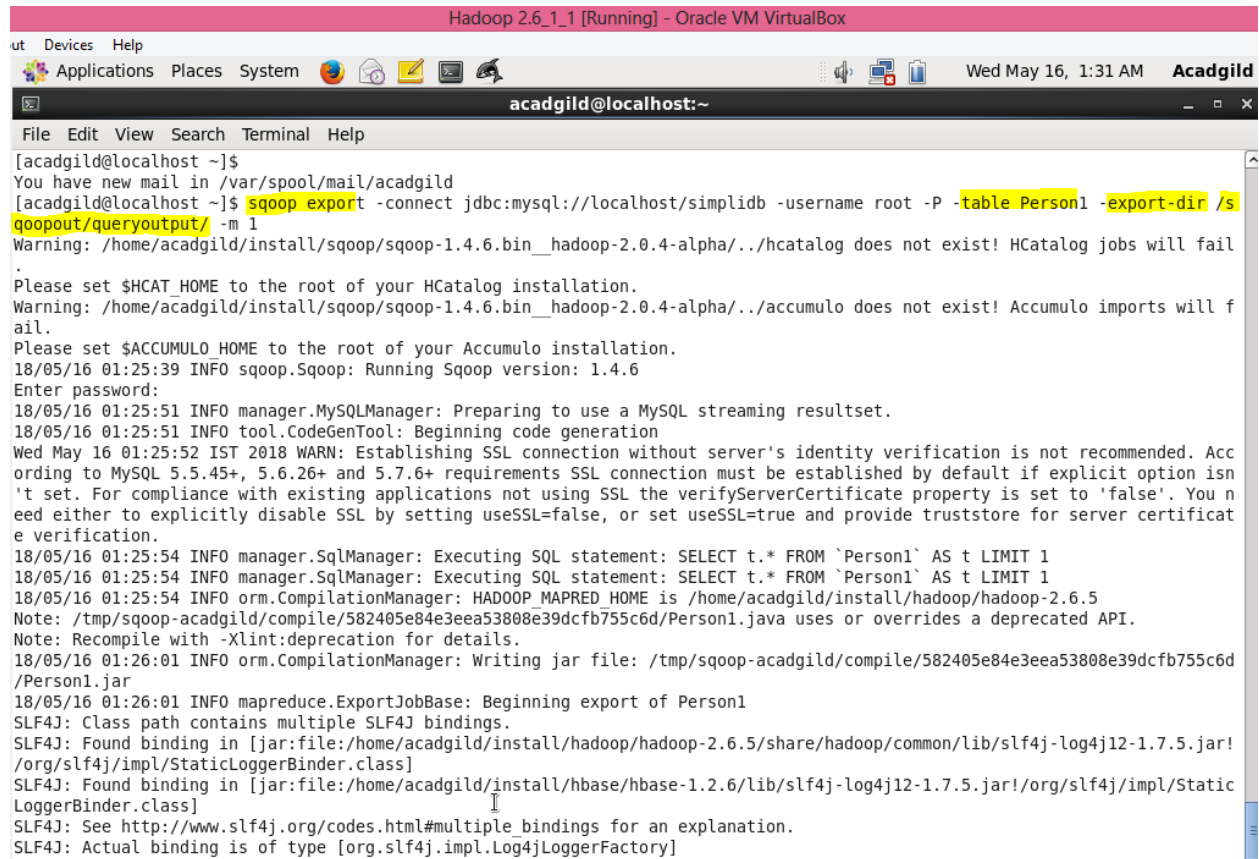
Then Export the data from HDFS file to MYSQL Person2 table using the below export command :

] Sqoop export –connect jdbc:mysql://localhost/simplidb –username root –P –table Person1 – export-dir /sqoopout/queryoutput/ -m 1

Below is the screenshot:

ut   Devices   Help

```
                FILE: Number of bytes written=127551
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=502
                HDFS: Number of bytes written=0
                HDFS: Number of read operations=13
                HDFS: Number of large read operations=0
                HDFS: Number of write operations=0
        Job Counters
                Launched map tasks=1
                Data-local map tasks=1
                Total time spent by all maps in occupied slots (ms)=23734
                Total time spent by all reduces in occupied slots (ms)=0
                Total time spent by all map tasks (ms)=23734
                Total vcore-milliseconds taken by all map tasks=23734
                Total megabyte-milliseconds taken by all map tasks=24303616
        Map-Reduce Framework
                Map input records=4
                Map output records=4
                Input split bytes=352
                Spilled Records=0
                Failed Shuffles=0
                Merged Map outputs=0
                GC time elapsed (ms)=145
                CPU time spent (ms)=2380
                Physical memory (bytes) snapshot=106487808
                Virtual memory (bytes) snapshot=2061336576
                Total committed heap usage (bytes)=32571392
        File Input Format Counters
                Bytes Read=0
        File Output Format Counters
                Bytes Written=0
18/05/16 01:27:02 INFO mapreduce.ExportJobBase: Transferred 502 bytes in 57.3852 seconds (8.7479 bytes/sec)
18/05/16 01:27:02 INFO mapreduce.ExportJobBase: Exported 4 records.
You have new mail in /var/spool/mail/acadgild
[acadgild@localhost ~]$
```

Now see the results in MYSQL DB Table Person2. The data should be exported to Person1 . Four rows should be reflected in table as shown below:

```
mysql> use simplidb
Database changed
mysql> Create table Person1 ( person_id INT NOT NULL, lname varchar(20), fname v
archar(20), area varchar(20), PRIMARY KEY (person_id) );
Query OK, 0 rows affected (0.97 sec)

mysql> select * from Person1;
Empty set (0.13 sec)

mysql> describe Person1;
+-----------+-------------+------+-----+---------+-------+
| Field     | Type        | Null | Key | Default | Extra |
+-----------+-------------+------+-----+---------+-------+
| person_id | int(11)     | NO   | PRI | NULL    |       |
| lname     | varchar(20) | YES  |     | NULL    |       |
| fname     | varchar(20) | YES  |     | NULL    |       |
| area      | varchar(20) | YES  |     | NULL    |       |
+-----------+-------------+------+-----+---------+-------+
4 rows in set (0.11 sec)

mysql> select * from Person1;
+-----------+-------+--------+-------------+
| person_id | lname | fname  | area        |
+-----------+-------+--------+-------------+
|         1 | Shyam | Ram    | Patna       |
|         2 | Tanya | Priya  | Whitefiled  |
|         3 | James | Brown  | New York    |
|         4 | Jhon  | Miller | Los Angeles |
+-----------+-------+--------+-------------+
4 rows in set (0.01 sec)

mysql>
```

Task2

Use Sqoop tool to export data present in SQOOPOUT folder made while demo of Import table

with parameter person_id =3.

Created a table Person2 in MYSQL Simplidb database with column in same order in hdfs output
shown in screen shot below :

```
                              acadgild@localhost:~                          _  □  x
File  Edit  View  Search  Terminal  Help
mysql> use simplidb
Database changed
mysql> Create table Person2 ( person_id INT NOT NULL, lname varchar(20), fname varchar(20), area varchar(20), PRIMARY KEY (pe
rson_id) );
Query OK, 0 rows affected (0.34 sec)

mysql> select * from Person2;
Empty set (0.06 sec)

mysql> describe Person2;
+-----------+-------------+------+-----+---------+-------+
| Field     | Type        | Null | Key | Default | Extra |
+-----------+-------------+------+-----+---------+-------+
| person_id | int(11)     | NO   | PRI | NULL    |       |
| lname     | varchar(20) | YES  |     | NULL    |       |
| fname     | varchar(20) | YES  |     | NULL    |       |
| area      | varchar(20) | YES  |     | NULL    |       |
+-----------+-------------+------+-----+---------+-------+
4 rows in set (0.11 sec)

mysql>
```

Then Exececute the below command in terminal as shown below in the screen shot to insert row
where Person_id= 3

```
[acadgild@localhost ~]$ sqoop export -connect jdbc:mysql://localhost/simplidb -username root -P -table Person2 -export-dir /s
qoopout/queryoutput/part-m-00002
Warning: /home/acadgild/install/sqoop/sqoop-1.4.6.bin__hadoop-2.0.4-alpha/../hcatalog does not exist! HCatalog jobs will fail
.
Please set $HCAT_HOME to the root of your HCatalog installation.
Warning: /home/acadgild/install/sqoop/sqoop-1.4.6.bin__hadoop-2.0.4-alpha/../accumulo does not exist! Accumulo imports will f
ail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
18/05/19 22:47:35 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6
Enter password:
18/05/19 22:47:43 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
18/05/19 22:47:43 INFO tool.CodeGenTool: Beginning code generation
Sat May 19 22:47:43 IST 2018 WARN: Establishing SSL connection without server's identity verification is not recommended. Acc
ording to MySQL 5.5.45+, 5.6.26+ and 5.7.6+ requirements SSL connection must be established by default if explicit option isn
't set. For compliance with existing applications not using SSL the verifyServerCertificate property is set to 'false'. You n
eed either to explicitly disable SSL by setting useSSL=false, or set useSSL=true and provide truststore for server certificat
e verification.
18/05/19 22:47:45 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `Person2` AS t LIMIT 1
18/05/19 22:47:45 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `Person2` AS t LIMIT 1
18/05/19 22:47:45 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /home/acadgild/install/hadoop/hadoop-2.6.5
Note: /tmp/sqoop-acadgild/compile/ccaa9223a32f76dfb5efba4b01e80481/Person2.java uses or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
18/05/19 22:47:50 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-acadgild/compile/ccaa9223a32f76dfb5efba4b01e80481
/Person2.jar
18/05/19 22:47:50 INFO mapreduce.ExportJobBase: Beginning export of Person2
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/home/acadgild/install/hadoop/hadoop-2.6.5/share/hadoop/common/lib/slf4j-log4j12-1.7.5.jar!
/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/home/acadgild/install/hbase/hbase-1.2.6/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/Static
LoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
18/05/19 22:47:50 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java cl
asses where applicable
18/05/19 22:47:50 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar
18/05/19 22:47:53 INFO Configuration.deprecation: mapred.reduce.tasks.speculative.execution is deprecated. Instead, use mapre
duce.reduce.speculative
```

Then Query in Mysql to see the result in Person2 Table as shown below:

```
File  Edit  View  Search  Terminal  Help
mysql> use simplidb
Database changed
mysql> Select * from Person2;
+-----------+-------+-------+----------+
| person_id | lname | fname | area     |
+-----------+-------+-------+----------+
|         3 | James | Brown | New York |
+-----------+-------+-------+----------+
1 row in set (0.02 sec)

mysql>
```