SQOOP ASSIGNMENT

Task 1: Use Sqoop tool to export data present in SQOOPOUT folder made while demo of Import table

Now 1st check where the sqoop output folder present and whether it contains data or not.

My SQOOPOUT folder present under

[acadgild@localhost ~]$ hadoop fs -cat /satya/sqoop/SQOOPOUT/p*

```
[acadgild@localhost ~]$ hadoop fs -cat /satya/sqoop/SQOOPOUT/p*
18/11/07 11:09:34 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
1,satya,Tavarekera,8989890000
2,priya,BTM,9898002323
3,muna,Agra,7878999900
4,kuna,Silk Board,7676889900
5,swapna,HSR,8787990012
6,prakash,HSR,9889007878
7,jhuna,HSR,9876500123
[acadgild@localhost ~]$
```

To export data, we have to use the bellow command

Sqoop export --connect jdbc:mysql://localhost:3306/emp --username root --password Root@123 --table person1 --export-dir /satya/sqoop/SQOOPOUT -m 1

NOTE:

The target table(person1) must be exist in MySQL.

```
mysql> describe person1;
+----------------+--------------+------+-----+---------+-------+
| Field          | Type         | Null | Key | Default | Extra |
+----------------+--------------+------+-----+---------+-------+
| person_id      | int(11)      | YES  |     | NULL    |       |
| person_name    | varchar(50)  | YES  |     | NULL    |       |
| person_address | varchar(200) | YES  |     | NULL    |       |
| person_mobile  | bigint(20)   | YES  |     | NULL    |       |
+----------------+--------------+------+-----+---------+-------+
4 rows in set (0.01 sec)

mysql> select * from person1;
Empty set (0.00 sec)
```

How it works:

Sqoop calls the JDBC driver written in the connect statement

For connect statement it uses the jdbc driver,server IP address and port where the my sql is running and for which schema we are running.

--connect jdbc:mysql://localhost:3306/emp

The --username and --password required to connect to the MySql database.

The --table argument tells us to which table we want to copy the data from the hdfs

The -m argument tells how many mapper we want to run, if we will not provide by default 4 map task will start

--export-dir : Here we need to provide from which location we need to read the file from HDFS

Now let us run the export command and check whether our data copied or not to mysql data base

```
[acadgild@localhost ~]$ sqoop export --connect jdbc:mysql://localhost:3306/emp --username root --password Root@123 --table person1 --export-dir /satya/sq
oop/SQOOPOUT -m  1
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 fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
18/11/07 11:46:30 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6
18/11/07 11:46:30 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.
18/11/07 11:46:30 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
18/11/07 11:46:30 INFO tool.CodeGenTool: Beginning code generation
Wed Nov 07 11:46:30 IST 2018 WARN: Establishing SSL connection without server's identity verification is not recommended. According 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 us
ing SSL the verifyServerCertificate property is set to 'false'. You need either to explicitly disable SSL by setting useSSL=false, or set useSSL=true and
 provide truststore for server certificate verification.
18/11/07 11:46:31 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `person1` AS t LIMIT 1
18/11/07 11:46:32 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `person1` AS t LIMIT 1
18/11/07 11:46:32 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /home/acadgild/install/hadoop/hadoop-2.6.5
Note: /tmp/sqoop-acadgild/compile/d6c850a7b74efa7b973c364cb91e718d/person1.java uses or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
18/11/07 11:46:35 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-acadgild/compile/d6c850a7b74efa7b973c364cb91e718d/person1.jar
18/11/07 11:46:35 INFO mapreduce.ExportJobBase: Beginning export of person1
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/StaticLogger
Binder.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/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an explanation.
SLF4J: Actual binding is of type [org.slf4j.impl.Log4jLoggerFactory]
18/11/07 11:46:35 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable
18/11/07 11:46:35 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar
18/11/07 11:46:37 INFO Configuration.deprecation: mapred.reduce.tasks.speculative.execution is deprecated. Instead, use mapreduce.reduce.speculative
18/11/07 11:46:37 INFO Configuration.deprecation: mapred.map.tasks.speculative.execution is deprecated. Instead, use mapreduce.map.speculative
18/11/07 11:46:37 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps
18/11/07 11:46:37 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032
18/11/07 11:46:40 INFO input.FileInputFormat: Total input paths to process : 1
18/11/07 11:46:40 INFO input.FileInputFormat: Total input paths to process : 1
18/11/07 11:46:41 INFO mapreduce.JobSubmitter: number of splits:1
18/11/07 11:46:41 INFO Configuration.deprecation: mapred.map.tasks.speculative.execution is deprecated. Instead, use mapreduce.map.speculative
18/11/07 11:46:41 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1541566516605_0002
18/11/07 11:46:42 INFO impl.YarnClientImpl: Submitted application application_1541566516605_0002
18/11/07 11:46:42 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1541566516605_0002/
18/11/07 11:46:42 INFO mapreduce.Job: Running job: job_1541566516605_0002
18/11/07 11:46:54 INFO mapreduce.Job: Job job_1541566516605_0002 running in uber mode : false
18/11/07 11:46:54 INFO mapreduce.Job:  map 0% reduce 0%
```

```
18/11/07 11:46:54 INFO mapreduce.Job:  map 0% reduce 0%
18/11/07 11:47:04 INFO mapreduce.Job:  map 100% reduce 0%
18/11/07 11:47:05 INFO mapreduce.Job: Job job_1541566516605_0002 completed successfully
18/11/07 11:47:05 INFO mapreduce.Job: Counters: 30
        File System Counters
                FILE: Number of bytes read=0
                FILE: Number of bytes written=127575
                FILE: Number of read operations=0
                FILE: Number of large read operations=0
                FILE: Number of write operations=0
                HDFS: Number of bytes read=316
                HDFS: Number of bytes written=0
                HDFS: Number of read operations=4
                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)=6887
                Total time spent by all reduces in occupied slots (ms)=0
                Total time spent by all map tasks (ms)=6887
                Total vcore-milliseconds taken by all map tasks=6887
                Total megabyte-milliseconds taken by all map tasks=7052288
        Map-Reduce Framework
                Map input records=7
                Map output records=7
                Input split bytes=136
                Spilled Records=0
                Failed Shuffles=0
                Merged Map outputs=0
                GC time elapsed (ms)=105
                CPU time spent (ms)=1790
                Physical memory (bytes) snapshot=107487232
                Virtual memory (bytes) snapshot=2061332480
                Total committed heap usage (bytes)=32571392
        File Input Format Counters
                Bytes Read=0
        File Output Format Counters
                Bytes Written=0
18/11/07 11:47:05 INFO mapreduce.ExportJobBase: Transferred 316 bytes in 28.5024 seconds (11.0868 bytes/sec)
18/11/07 11:47:05 INFO mapreduce.ExportJobBase: Exported 7 records.
You have new mail in /var/spool/mail/acadgild
```

Before export command

```
mysql> select * from person1;
Empty set (0.00 sec)
```

Now let us check the person1 table whether our data copied here or not

```
mysql> select * from person1;
+-----------+-------------+----------------+---------------+
| person_id | person_name | person_address | person_mobile |
+-----------+-------------+----------------+---------------+
|         1 | satya       | Tavarekera     |    8989890000 |
|         2 | priya       | BTM            |    9898002323 |
|         3 | muna        | Agra           |    7878999900 |
|         4 | kuna        | Silk Board     |    7676889900 |
|         5 | swapna      | HSR            |    8787990012 |
|         6 | prakash     | HSR            |    9889007878 |
|         7 | jhuna       | HSR            |    9876500123 |
+-----------+-------------+----------------+---------------+
7 rows in set (0.00 sec)
```