Session 11: Sqoop Flume Assignment 2 Question

Perform incremental load in Hive

Read from MySQL Table and load it in Hive table.

Create hive table if it does not exist.

If it exists, perform the incremental load.

mysql> use db1;

Database changed

mysql> create table acad(emp_id INT NOT NULL AUTO_INCREMENT,emp_name VARCHAR(100),emp_sal INT,PRIMARY KEY(emp_id));

Query OK, 0 rows affected (0.01 sec)

mysql> insert into acad values(5,"Sanam",50000);

Query OK, 1 row affected (0.00 sec)

mysql> insert into acad values(6,"Sonam",54000);

Query OK, 1 row affected (0.00 sec)

mysql> GRANT ALL PRIVILEGES ON db1.* TO 'acadgild'@'localhost';

Query OK, 0 rows affected (0.00 sec)

mysql> flush privileges;

Query OK, 0 rows affected (0.00 sec)

mysql> commit;

Query OK, 0 rows affected (0.00 sec)

mysql> insert into acad values(7,"Suraj",58000);

Query OK, 1 row affected (0.00 sec)

```
mysql> use db1;
Database changed
mysql> create table acad(emp_id INT NOT NULL AUTO_INCREMENT,emp_name
VARCHAR(100),emp_sal INT,PRIMARY KEY(emp_id));
Query OK, 0 rows affected (0.01 sec)

mysql> insert into acad values(5,"Sanam",50000);
Query OK, 1 row affected (0.00 sec)

mysql> insert into acad values(6,"Sonam",54000);
Query OK, 1 row affected (0.00 sec)

mysql> insert into acad values(7,"Suraj",58000);
Query OK, 1 row affected (0.00 sec)

mysql> select * from acad;
+--------+----------+---------+
| emp_id | emp_name | emp_sal |
+--------+----------+---------+
|      5 | Sanam    |   50000 |
|      6 | Sonam    |   54000 |
|      7 | Suraj    |   58000 |
+--------+----------+---------+
```

mysql> GRANT ALL PRIVILEGES ON db1.* TO 'acadgild'@'localhost';

Query OK, 0 rows affected (0.00 sec)

mysql> flush privileges;

Query OK, 0 rows affected (0.00 sec)

mysql> commit;

Query OK, 0 rows affected (0.00 sec)

```
mysql> GRANT ALL PRIVILEGES ON db1.* TO 'acadgild'@'localhost';
Query OK, 0 rows affected (0.00 sec)

mysql> flush privileges;
Query OK, 0 rows affected (0.00 sec)

mysql> commit;
Query OK, 0 rows affected (0.00 sec)
```

[acadgild@localhost ~]$ sqoop import --connect jdbc:mysql://localhost/db1 --username root --password root --table acad --target-dir sqoop_increment -m 1

```
acadgild@192.168.1.11's password:
Last login: Wed Oct 25 18:18:57 2017 from 192.168.1.9
[acadgild@localhost ~]$ sqoop import --connect jdbc:mysql://localhost/db1 --user
name root --password root --table acad --target-dir sqoop_increment -m 1
Warning: /usr/local/sqoop/../hcatalog does not exist! HCatalog jobs will fail.
Please set $HCAT_HOME to the root of your HCatalog installation.
Warning: /usr/local/sqoop/../accumulo does not exist! Accumulo imports will fail
.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
Warning: /usr/local/sqoop/../zookeeper does not exist! Accumulo imports will fai
l.
Please set $ZOOKEEPER_HOME to the root of your Zookeeper installation.
2017-10-25 18:33:07,265 INFO  [main] sqoop.Sqoop: Running Sqoop version: 1.4.5
2017-10-25 18:33:07,279 WARN  [main] tool.BaseSqoopTool: Setting your password o
n the command-line is insecure. Consider using -P instead.
2017-10-25 18:33:07,482 INFO  [main] manager.MySQLManager: Preparing to use a My
SQL streaming resultset.
2017-10-25 18:33:07,482 INFO  [main] tool.CodeGenTool: Beginning code generation
2017-10-25 18:33:07,743 INFO  [main] manager.SqlManager: Executing SQL statement
: SELECT t.* FROM `acad` AS t LIMIT 1
2017-10-25 18:33:07,768 INFO  [main] manager.SqlManager: Executing SQL statement
: SELECT t.* FROM `acad` AS t LIMIT 1
2017-10-25 18:33:07,777 INFO  [main] orm.CompilationManager: HADOOP_MAPRED_HOME
is /usr/local/hadoop-2.6.0
Note: /tmp/sqoop-acadgild/compile/4721f956449ac263d8a66c1d4888aabf/acad.java use
s or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
2017-10-25 18:33:09,849 INFO  [main] orm.CompilationManager: Writing jar file: /
tmp/sqoop-acadgild/compile/4721f956449ac263d8a66c1d4888aabf/acad.jar
2017-10-25 18:33:09,868 WARN  [main] manager.MySQLManager: It looks like you are
 importing from mysql.
2017-10-25 18:33:09,869 WARN  [main] manager.MySQLManager: This transfer can be
faster! Use the --direct
2017-10-25 18:33:09,869 WARN  [main] manager.MySQLManager: option to exercise a
MySQL-specific fast path.
2017-10-25 18:33:09,869 INFO  [main] manager.MySQLManager: Setting zero DATETIME
```

[acadgild@localhost ~]$ hadoop fs -cat /user/acadgild/sqoop_increment/part-m-00000

5,Sanam,50000

6,Sonam,54000

7,Suraj,58000

mysql> insert into acad values(8,"Surya",49000);

Query OK, 1 row affected (0.00 sec)

mysql> insert into acad values(9,"Anand",79000);

Query OK, 1 row affected (0.00 sec)

mysql> select * from acad;

+--------+----------+---------+

| emp_id | emp_name | emp_sal |

```
+--------+----------+---------+
|    5 | Sanam   |  50000 |
|    6 | Sonam   |  54000 |
|    7 | Suraj   |  58000 |
|    8 | Surya   |  49000 |
|    9 | Anand   |  79000 |
+--------+----------+---------+
```

5 rows in set (0.00 sec)

mysql> SELECT MAX(`emp_id`) FROM acad;

```
+---------------+
| MAX(`emp_id`) |
+---------------+
|          9 |
```

```
+---------------+
```

1 row in set (0.00 sec)

```
mysql> insert into acad values(8,"Surya",49000);
Query OK, 1 row affected (0.00 sec)

mysql> insert into acad values(9,"Anand",79000);
Query OK, 1 row affected (0.00 sec)

mysql> select * from acad;
+--------+----------+---------+
| emp_id | emp_name | emp_sal |
+--------+----------+---------+
|      5 | Sanam    |   50000 |
|      6 | Sonam    |   54000 |
|      7 | Suraj    |   58000 |
|      8 | Surya    |   49000 |
|      9 | Anand    |   79000 |
+--------+----------+---------+
5 rows in set (0.00 sec)

mysql> SELECT MAX(`emp_id`) FROM acad;
+---------------+
| MAX(`emp_id`) |
+---------------+
|             9 |
+---------------+
1 row in set (0.00 sec)
```

[acadgild@localhost ~]$ sqoop import --connect jdbc:mysql://localhost/db1 --username root --password root --table acad --target-dir sqoop_increment --incremental append --check-column emp_id --last-value 7 -m 1

```
[acadgild@localhost ~]$ sqoop import --connect jdbc:mysql://localhost/db1 --username root --password root --table acad --target-dir sqo
op_increment --incremental append --check-column emp_id --last-value 7 -m 1
Warning: /usr/local/sqoop/../hcatalog does not exist! HCatalog jobs will fail.
Please set $HCAT_HOME to the root of your HCatalog installation.
Warning: /usr/local/sqoop/../accumulo does not exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
Warning: /usr/local/sqoop/../zookeeper does not exist! Accumulo imports will fail.
Please set $ZOOKEEPER_HOME to the root of your Zookeeper installation.
2017-10-25 18:55:48,733 INFO   [main] sqoop.Sqoop: Running Sqoop version: 1.4.5
2017-10-25 18:55:48,747 WARN   [main] tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P inste
ad.
2017-10-25 18:55:48,966 INFO   [main] manager.MySQLManager: Preparing to use a MySQL streaming resultset.
2017-10-25 18:55:48,966 INFO   [main] tool.CodeGenTool: Beginning code generation
2017-10-25 18:55:49,235 INFO   [main] manager.SqlManager: Executing SQL statement: SELECT t.* FROM `acad` AS t LIMIT 1
2017-10-25 18:55:49,263 INFO   [main] manager.SqlManager: Executing SQL statement: SELECT t.* FROM `acad` AS t LIMIT 1
2017-10-25 18:55:49,268 INFO   [main] orm.CompilationManager: HADOOP_MAPRED_HOME is /usr/local/hadoop-2.6.0
Note: /tmp/sqoop-acadgild/compile/862f64c012df2416d20b868b812628af/acad.java uses or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
2017-10-25 18:55:50,796 INFO   [main] orm.CompilationManager: Writing jar file: /tmp/sqoop-acadgild/compile/862f64c012df2416d20b868b8126
28af/acad.jar
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/local/hbase/lib/slf4j-log4j12-1.6.4.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/local/hadoop-2.6.0/share/hadoop/common/lib/slf4j-log4j12-1.7.5.jar!/org/slf4j/impl/StaticLoggerB
inder.class]
```

```
2017-10-25 18:55:51,545 INFO   [main] tool.ImportTool: Lower bound value: 7
2017-10-25 18:55:51,545 INFO   [main] tool.ImportTool: Upper bound value: 9
```

[acadgild@localhost ~]$ hadoop fs -cat /user/acadgild/sqoop_increment/part-m-00000

5,Sanam,50000

6,Sonam,54000

7,Suraj,58000

 8,Surya,49000

9,Anand,79000