

1. First we need to login to EMR and then switch to root user and run MySQL connector code

sudo -i

wget <https://de-mysql-connector.s3.amazonaws.com/mysql-connector-java-8.0.25.tar.gz>

tar -xvf mysql-connector-java-8.0.25.tar.gz

cd mysql-connector-java-8.0.25/

sudo cp mysql-connector-java-8.0.25.jar /usr/lib/sqoop/lib/

```
root@ip-172-31-65-22:~/mysql-connector-java-8.0.25
mysql-connector-java-8.0.25/src/test/java/testsuite/simple/ServerControllerTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/simple/SimpleTransformer.java
mysql-connector-java-8.0.25/src/test/java/testsuite/simple/SplitIDotNameTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/simple/StatementsTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/simple/StringUtilsTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/simple/TestLifecycleInterceptor.java
mysql-connector-java-8.0.25/src/test/java/testsuite/simple/TransactionTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/simple/TraversalTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/simple/UpdatabilityTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/simple/UtilsTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/simple/XATest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/AsyncTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/BaseXDevAPITestCase.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/TestXDevAPIRequirements.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/AsyncQueryTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/BaseCollectionTestCase.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/BaseTableTestCase.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/BindTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/CollectionAddTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/CollectionFindTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/CollectionModifyTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/CollectionRemoveTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/CompressionTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/DevAPIBaseTestCase.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/IPv6SupportTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/MetadataTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/ResultTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/SchemaTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/SecureSessionTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/SessionFailoverTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/SessionTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/TableDeleteTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/TableInsertTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/TableSelectTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/TableTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/TableUpdateTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/TransactionTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/devapi/package-info.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/internal/InternalXBaseTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/internal/MySQLSessionTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/internal/XProtocolAsyncTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/internal/XProtocolAuthTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/internal/XProtocolTest.java
mysql-connector-java-8.0.25/src/test/java/testsuite/x/internal/package-info.java
[root@ip-172-31-65-22 ~]# cd mysql-connector-java-8.0.25/
[root@ip-172-31-65-22 mysql-connector-java-8.0.25]# sudo cp mysql-connector-java-8.0.25.jar /usr/lib/sqoop/lib/
[root@ip-172-31-65-22 mysql-connector-java-8.0.25]#
```

2. Command to ingest the data from RDS into the HBase Table

```
sqoop import --connect jdbc:mysql://database-using.cb0sqm06ygvf.us-east-1.rds.amazonaws.com:3306/taxi \
--username admin --password admin123 \
--table yellow_taxi --hbase-table yellow_taxi_hbase --column-family cf1 --hbase-create-table \
--hbase-row-key tpep_pickup_datetime,tpep_dropoff_datetime --hbase-bulkload --split-by payment_type
```

- This is a sqoop command to import data from MySQL to Hbase table 'yellow_taxi_hbase'.
- The '--connect' is for connection for MySQL DB.
- Then '--username' and '--password' are both used to connect MySQL DB. '--table' to specify the table name in MySQL table.
- '--hbase-table' specifies the name of the table to import to.
- '--column-family' specifies the name of column family
- '--hbase-create-table' create hbase table
- '--hbase-row-key' specifies the column that will be specified as row key in HBase

```
hadoop@ip-172-31-65-22:~$
[root@ip-172-31-65-22 mysql-connector-java-8.0.25]# sudo cp mysql-connector-java-8.0.25.jar /usr/lib/sqoop/lib/
[root@ip-172-31-65-22 mysql-connector-java-8.0.25]#
[root@ip-172-31-65-22 mysql-connector-java-8.0.25]# sqoop import --connect jdbc:mysql://database-using.cb0sqm06ygvf.us-east-1.rds.amazonaws.com:3306/taxi \
> --username admin --password admin123 \
> --table yellow_taxi --hbase-table yellow_taxi_hbase --column-family cf1 --hbase-create-table \
> --hbase-row-key tpep_pickup_datetime,tpep_dropoff_datetime --hbase-bulkload --split-by payment_type
[root@ip-172-31-65-22 mysql-connector-java-8.0.25]# exit
logout
hadoop@ip-172-31-65-22:~$ sqoop import --connect jdbc:mysql://database-using.cb0sqm06ygvf.us-east-1.rds.amazonaws.com:3306/taxi \
> --username admin --password admin123 \
> --table yellow_taxi --hbase-table yellow_taxi_hbase --column-family cf1 --hbase-create-table \
> --hbase-row-key tpep_pickup_datetime,tpep_dropoff_datetime --hbase-bulkload --split-by payment_type
Warning: /usr/lib/sqoop/./hcatalog does not exist! HCatalog jobs will fail.
Please set HCAT_HOME to the root of your HCatalog installation.
Warning: /usr/lib/sqoop/./accumulo does not exist! Accumulo imports will fail.
Please set SACCUMULO_HOME to the root of your Accumulo installation.
2024-05-07 05:31:20,418 INFO sqoop.Sqoop: Running Sqoop version: 1.4.7
2024-05-07 05:31:20,473 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.
2024-05-07 05:31:20,654 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
2024-05-07 05:31:20,654 INFO tool.CodeGenTool: Beginning code generation
Loading class 'com.mysql.jdbc.Driver'. This is deprecated. The new driver class is 'com.mysql.cj.jdbc.Driver'. The driver is automatically registered via the SPI and manual loading of the driver class is generally unnecessary.
2024-05-07 05:31:21,220 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM 'yellow_taxi' AS t LIMIT 1
2024-05-07 05:31:21,455 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM 'yellow_taxi' AS t LIMIT 1
2024-05-07 05:31:21,480 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /usr/lib/hadoop-mapreduce
2024-05-07 05:31:23,701 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-hadoop/compile/ce677ffd3afb3765012683c3e029108/yellow_taxi.jar
2024-05-07 05:31:23,726 WARN manager.MySQLManager: It looks like you are importing from mysql.
2024-05-07 05:31:23,726 WARN manager.MySQLManager: This transfer can be faster! Use the --direct
2024-05-07 05:31:23,727 WARN manager.MySQLManager: option to exercise a MySQL-specific fast path.
2024-05-07 05:31:23,727 INFO manager.MySQLManager: Setting zero DATETIME behavior to convertToNull (mysql)
2024-05-07 05:31:24,071 INFO mapreduce.ImportJobBase: Beginning import of yellow_taxi
2024-05-07 05:31:24,278 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar
2024-05-07 05:31:24,300 INFO mapreduce.HBaseImportJob: Creating missing HBase table yellow_taxi_hbase
2024-05-07 05:31:29,024 ERROR sqoop.Sqoop: Got exception running Sqoop: java.lang.NullPointerException
java.lang.NullPointerException
    at org.apache.sqoop.mapreduce.HBaseBulkImportJob.jobTeardown(HBaseBulkImportJob.java:148)
    at org.apache.sqoop.mapreduce.ImportJobBase.runImport(ImportJobBase.java:295)
    at org.apache.sqoop.manager.SqlManager.importTable(SqlManager.java:691)
    at org.apache.sqoop.manager.MySQLManager.importTable(MySQLManager.java:124)
    at org.apache.sqoop.tool.ImportTool.importTable(ImportTool.java:531)
    at org.apache.sqoop.tool.ImportTool.run(ImportTool.java:630)
    at org.apache.sqoop.Sqoop.run(Sqoop.java:145)
    at org.apache.hadoop.util.ToolRunner.run(ToolRunner.java:81)
    at org.apache.sqoop.Sqoop.runSqoop(Sqoop.java:181)
    at org.apache.sqoop.Sqoop.runTool(Sqoop.java:232)
    at org.apache.sqoop.Sqoop.runTool(Sqoop.java:241)
    at org.apache.sqoop.Sqoop.main(Sqoop.java:250)
hadoop@ip-172-31-65-22:~$
```

```
root@ip-172-31-65-22~#
EEEEEEEEEEEEEEEEEEEE MMMMMM RRRRRRRRRRRRRR
E:::|||||:|||||: M:::|||||: M:::|||||: R:::|||||:|:R
EE::|||:EEEEEEEE::|E M:::|||||: M:::|||||: M R:::|RRRRR::|:|:R
E:::|E EEEEE M:::|||||: M:::|||||: M RR::|:R R::|:R
E:::|E EEEEE M:::|:M::|: M::|:M::|: M R::|:R R::|:R
E:::|:EEEEEEEE M:::|M M::|M M::|M M::|M R::|RRRRR::|:|:R
E:::|:|||||:E M:::|M M::|M:M M::|M M::|M R:::|:|:|:RR
E:::|:EEEEEEEE M:::|M M:::|M M:::|M M:::|M R::|RRRRR::|:|:R
E:::|E M:::|M M::|M M::|M M::|M R::|:R R::|:R
E:::|E EEEEE M:::|M MMM M:::|M R::|:R R::|:R
EE::|||:EEEEEEEE::|E M:::|M M:::|M R::|:R R::|:R
E:::|||||:|||||:E M:::|M M:::|M RR::|:R R::|:R
EEEEEEEEEEEEEEEEEEEE MMMMMM RRRRRRR RRRRRR

[root@ip-172-31-65-22 ~]# hbase shell
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in [jar:file:/usr/lib/hadoop/lib/slf4j-reload4j-1.7.36.jar/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: Found binding in [jar:file:/usr/lib/hadoop/lib/client-facing-thirdparty/slf4j-reload4j-1.7.33.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.Reload4jLoggerFactory)
HBase Shell
Use "help" to get list of supported commands.
Use "exit" to quit this interactive shell.
For Reference, please visit: http://hbase.apache.org/2.0/book.html#shell
Version 2.4.15-amzn-0, Unknown, Tue Feb 21 18:23:24 UTC 2023
Took 0.0025 seconds
hbase:001> list
TABLE
yellow_taxi_hbase
1 row(s)
Took 0.5947 seconds
=> ["yellow_taxi_hbase"]
hbase:002> describe yellow_taxi_hbase
Traceback (most recent call last):
NameError (undefined local variable or method 'yellow_taxi_hbase' for #<HBaseReceiver:0x41582899>)
hbase:003> describe yellow_taxi_hbase:
Traceback (most recent call last):
NameError (undefined local variable or method 'yellow_taxi_hbase' for #<HBaseReceiver:0x41582899>)
hbase:004> describe 'yellow_taxi_hbase'
table yellow_taxi_hbase is ENABLED
yellow_taxi_hbase
COLUMN FAMILIES DESCRIPTION
(NAME => 'cf1', BLOOMFILTER => 'ROW', IN_MEMORY => 'false', VERSIONS => '1', KEEP_DELETED_CELLS => 'FALSE', DATA_BLOCK_ENCODING => 'NONE', COMPRESSION => 'FOREVER', MIN_VERSIONS => '0', BLOCKCACHE => 'true', TOLERANCE_KEY => '65536', REPLICATION_SCOPE => '0')

1 row(s)
Quota is disabled
Took 0.1934 seconds
hbase:005>
```