

SQOOP COMMAND

```
sqoop import \  
--connect jdbc:mysql://upgraddetest.cyaieic9bmnf.us-east-1.rds.amazonaws.com/testdatabase \  
--username student \  
--password STUDENT123 \  
--table patients_information \  
--direct \  
-m 1 \  
--hive-import \  
--create-hive-table \  
--hive-table patient_contacts \  
--fields-terminated-by ';
```

SNAPSHOT OF DATA COLLECTED

```
hive> select * from patient_contacts;  
OK  
1      Alex S   XDC test Address      8982739282      1      23      NULL  
2      Sammy A New Building Address  2382739282      2      45      NULL  
3      Karan C  Aws Address      8923739282      3      56      NULL  
4      Dara M  India Address   2182739282      4      67      NULL  
5      Pam     ABC test Address  4982739282      5      72      NULL  
Time taken: 1.39 seconds, Fetched: 5 row(s)  
hive>
```

```
ec2-user@ip-10-0-0-188 ~]$ sqoop import \
--connect jdbc:mysql://upgradtest.cya1elc9bmnf.us-east-1.rds.amazonaws.com/testdatabase \
--username student \
--password STUDENT123 \
--table patients information \
--direct \
--m 1 \
--hive-import \
--create-hive-table \
--hive-table patient_contacts \
--fields-terminated-by ';'

21/11/07 10:06:00 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6-cdh5.15.1
21/11/07 10:06:00 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.
21/11/07 10:06:00 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
21/11/07 10:06:00 INFO tool.CodeGenTool: Beginning code generation
21/11/07 10:06:01 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `patients_information` AS t LIMIT 1
21/11/07 10:06:01 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `patients_information` AS t LIMIT 1
21/11/07 10:06:01 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /opt/cloudera/parcels/CDH/lib/hadoop-mapreduce
Note: /tmp/sqoop-ec2-user/compile/f4340fc9b10a686a133d831ff1f6eefb/patients_information.java uses or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
21/11/07 10:06:02 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-ec2-user/compile/f4340fc9b10a686a133d831ff1f6eefb/patients_informati
on.jar
21/11/07 10:06:02 INFO manager.DirectMySQLManager: Beginning mysqldump fast path import
21/11/07 10:06:02 INFO mapreduce.ImportJobBase: Beginning import of patients_information
21/11/07 10:06:03 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar
21/11/07 10:06:04 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps
21/11/07 10:06:04 INFO client.RMProxy: Connecting to ResourceManager at ip-10-0-0-188.ec2.internal/10.0.0.188:8032
21/11/07 10:06:12 INFO db.DBInputFormat: Using read committed transaction isolation
21/11/07 10:06:14 INFO mapreduce.JobSubmitter: number of splits:1
21/11/07 10:06:14 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1636275300252_0010
21/11/07 10:06:15 INFO impl.YarnClientImpl: Submitted application application_1636275300252_0010
21/11/07 10:06:15 INFO mapreduce.Job: The url to track the job: http://ip-10-0-0-188.ec2.internal:8088/proxy/application_1636275300252_0010/
21/11/07 10:06:15 INFO mapreduce.Job: Running job: job_1636275300252_0010
21/11/07 10:06:23 INFO mapreduce.Job: Job job_1636275300252_0010 running in uber mode : false
21/11/07 10:06:23 INFO mapreduce.Job: map 0% reduce 0%
21/11/07 10:06:29 INFO mapreduce.Job: map 100% reduce 0%
ec2-user@ip-10-0-0-188 ~]$
21/11/07 10:06:30 INFO mapreduce.Job: Counters: 30
File System Counters
FILE: Number of bytes read=0
FILE: Number of bytes written=177017
FILE: Number of read operations=0
FILE: Number of large read operations=0
FILE: Number of write operations=0
HDFS: Number of bytes read=87
HDFS: Number of bytes written=230
HDFS: Number of read operations=4
HDFS: Number of large read operations=0
HDFS: Number of write operations=2
Job Counters
Launched map tasks=1
Other local map tasks=1
Total time spent by all maps in occupied slots (ms)=3674
Total time spent by all reduces in occupied slots (ms)=0
Total time spent by all map tasks (ms)=3674
Total vcore-milliseonds taken by all map tasks=3674
Total megabyte-milliseonds taken by all map tasks=3762176
Map-Reduce Framework
Map input records=1
Map output records=5
Input split bytes=87
Spilled Records=0
Failed Shuffles=0
Merged Map outputs=0
GC time elapsed (ms)=48
CPU time spent (ms)=1370
Physical memory (bytes) snapshot=237748224
Virtual memory (bytes) snapshot=2798776320
Total committed heap usage (bytes)=240648192
File Input Format Counters
Bytes Read=0
File Output Format Counters
Bytes Written=230
21/11/07 10:06:30 INFO mapreduce.ImportJobBase: Transferred 230 bytes in 26.6177 seconds (8.6409 bytes/sec)
21/11/07 10:06:30 INFO mapreduce.ImportJobBase: Retrieved 5 records.
21/11/07 10:06:30 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `patients_information` AS t LIMIT 1
21/11/07 10:06:30 INFO hive.HiveImport: Loading uploaded data into Hive
Logging initialized using configuration in jar:file:/opt/cloudera/parcels/CDH-5.15.1-1.cdh5.15.1.p0.4/jars/hive-common-1.1.0-cdh5.15.1.jar!/hive
-log4j.properties
OK
Time taken: 3.212 seconds
Loading data to table default.patient_contacts
Table default.patient_contacts stats: [numFiles=1, totalSize=230]
OK
Time taken: 0.473 seconds
```