

Roll No: 5117060

Experiment 02

Aim: To create Table in my MYSQL and pass it on to the HDFS using Sqoop

Output:

Creating Table in MYSQL

```
[cloudera@quickstart ~]$ sudo su root
[root@quickstart cloudera]# mysql -u root -p
Enter password:
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 21
Server version: 5.1.66 Source distribution

Copyright (c) 2000, 2012, Oracle and/or its affiliates. All rights reserved.

Oracle is a registered trademark of Oracle Corporation and/or its
affiliates. Other names may be trademarks of their respective
owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

mysql> use shivam;
Reading table information for completion of table and column names
You can turn off this feature to get a quicker startup with -A

Database changed
mysql> select * from employee
-> ;
+-----+-----+
| id    | ename    |
+-----+-----+
| 1     | Shivam   |
| 2     | Abdul Aziz |
| 3     | Asma     |
| 4     | Abuzar   |
+-----+-----+
4 rows in set (0.01 sec)

mysql> █
```

Transfer From MYSQL to HDFS using sqoop

```

[cloudera@quickstart ~]$ sqoop import --connect jdbc:mysql://quickstart.cloudera:3306/shivam --username root --password cloudera --m 1 --table employee
Warning: /usr/lib/sqoop/./accumulo does not exist! Accumulo imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
20/11/17 23:08:20 INFO sqoop.Sqoop: Running Sqoop version: 1.4.5-cdh5.4.2
20/11/17 23:08:20 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.
20/11/17 23:08:20 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
20/11/17 23:08:20 INFO tool.CodeGenTool: Beginning code generation
20/11/17 23:08:21 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `employee` AS t LIMIT 1
20/11/17 23:08:21 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `employee` AS t LIMIT 1
20/11/17 23:08:21 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /usr/lib/hadoop-mapreduce
Note: /tmp/sqoop-cloudera/compile/85c99c87a622737554f919f47e29b949/employee.java uses or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
20/11/17 23:08:25 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-cloudera/compile/85c99c87a622737554f919f47e29b949/employee.jar
20/11/17 23:08:25 WARN manager.MySQLManager: It looks like you are importing from mysql.
20/11/17 23:08:25 WARN manager.MySQLManager: This transfer can be faster! Use the --direct
20/11/17 23:08:25 WARN manager.MySQLManager: option to exercise a MySQL-specific fast path.
20/11/17 23:08:25 INFO manager.MySQLManager: Setting zero DATETIME behavior to convertToNull (mysql)
20/11/17 23:08:25 INFO mapreduce.ImportJobBase: Beginning import of employee
20/11/17 23:08:25 INFO Configuration.deprecation: mapred.job.tracker is deprecated. Instead, use mapreduce.jobtracker.address
20/11/17 23:08:25 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar
20/11/17 23:08:26 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps
20/11/17 23:08:27 INFO client.RMProxy: Connecting to ResourceManager at /0.0.0.0:8032
20/11/17 23:08:30 INFO db.DBInputFormat: Using read committed transaction isolation
20/11/17 23:08:30 INFO mapreduce.JobSubmitter: number of splits:1
20/11/17 23:08:30 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1605682012347_0001

```

```
20/11/17 23:08:31 INFO impl.YarnClientImpl: Submitted application application
_1605682012347_0001
20/11/17 23:08:31 INFO mapreduce.Job: The url to track the job: http://quicks
tart.cloudera:8088/proxy/application_1605682012347_0001/
20/11/17 23:08:31 INFO mapreduce.Job: Running job: job_1605682012347_0001
20/11/17 23:08:49 INFO mapreduce.Job: Job job_1605682012347_0001 running in u
ber mode : false
20/11/17 23:08:49 INFO mapreduce.Job: map 0% reduce 0%
20/11/17 23:09:02 INFO mapreduce.Job: map 100% reduce 0%
20/11/17 23:09:04 INFO mapreduce.Job: Job job_1605682012347_0001 completed su
ccessfully
20/11/17 23:09:04 INFO mapreduce.Job: Counters: 30
  File System Counters
    FILE: Number of bytes read=0
    FILE: Number of bytes written=134813
    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=38
    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)=10339
    Total time spent by all reduces in occupied slots (ms)=0
    Total time spent by all map tasks (ms)=10339
    Total vcore-seconds taken by all map tasks=10339
    Total megabyte-seconds taken by all map tasks=10587136
  Map-Reduce Framework
    Map input records=4
    Map output records=4
    Input split bytes=87
    Spilled Records=0
    Failed Shuffles=0
    Merged Map outputs=0
    GC time elapsed (ms)=115
    CPU time spent (ms)=1700
    Physical memory (bytes) snapshot=118976512
    Virtual memory (bytes) snapshot=1505083392
    Total committed heap usage (bytes)=60751872
  File Input Format Counters
    Bytes Read=0
  File Output Format Counters
    Bytes Written=38
20/11/17 23:09:04 INFO mapreduce.ImportJobBase: Transferred 38 bytes in 37.50
```

```
34 seconds (1.0132 bytes/sec)
20/11/17 23:09:04 INFO mapreduce.ImportJobBase: Retrieved 4 records.
[cloudera@quickstart ~]$ hdfs dfs -ls
Found 8 items
drwxr-xr-x  - cloudera cloudera          0 2020-11-17 04:13 Shivam
-rw-r--r--  1 cloudera cloudera      174693 2020-10-17 12:03 a.txt
-rw-r--r--  1 cloudera cloudera          0 2020-11-17 01:55 b.txt
drwxr-xr-x  - cloudera cloudera          0 2020-11-17 23:09 employee
drwxr-xr-x  - cloudera cloudera          0 2020-10-17 11:56 out
drwxr-xr-x  - cloudera cloudera          0 2020-10-17 12:06 out1
drwxr-xr-x  - cloudera cloudera          0 2020-10-17 12:15 out2
-rw-r--r--  1 cloudera cloudera    5458199 2020-10-17 11:45 words.txt
[cloudera@quickstart ~]$ hdfs dfs -ls employee/
Found 2 items
-rw-r--r--  1 cloudera cloudera          0 2020-11-17 23:09 employee/_SUCCESS
-rw-r--r--  1 cloudera cloudera      38 2020-11-17 23:09 employee/part-m-000000
[cloudera@quickstart ~]$ hdfs dfs -cat employee/part-m-000000
1,Shivam
2,Abdul Aziz
3,Asma
4,Abuzar
[cloudera@quickstart ~]$
```