

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

--I have created another table which is different from demo session but performed same operations as per Assignments.

mysql> show databases;

```
+-----+
| Database      |
+-----+
| information_schema |
| metastore      |
| mysql          |
| oozie          |
| performance_schema |
| simplidb      |
| sys            |
```

7 rows in set (1.37 sec)

---Using simplidb db to create table employee.

mysql> use simplidb;

mysql> create table employee(

emp_id INT NOT NULL,

emp_name varchar(20) NULL,

emp_dep varchar(20) NULL,

salary INT NULL,

loc varchar(20) NULL ,

PRIMARY KEY (emp_id)

);

Query OK, 0 rows affected (1.96 sec)

--Inserted values in to the table employee using insert command.

```
Mysql> insert into employee values(101,'soundarya','IT',25000,'Hyd');
```

Query OK, 1 row affected (0.24 sec)

```
mysql> insert into employee values(102,'janaki','IT',30000,'Ban');
```

Query OK, 1 row affected (0.24 sec)

```
mysql> insert into employee values(103,'sai','CLAIMS',15000,'Hyd');
```

Query OK, 1 row affected (0.01 sec)

```
mysql> insert into employee values(104,'Jinna','FIN',55000,'Ban');
```

Query OK, 1 row affected (0.20 sec)

```
mysql> insert into employee values(105,'Chetan','FIN',65000,'Hyd');
```

Query OK, 1 row affected (0.06 sec)

```
mysql> select * from employee;
```

```
+-----+-----+-----+-----+-----+
| emp_id | emp_name | emp_dep | salary | loc |
+-----+-----+-----+-----+-----+
| 101 | soundarya | IT | 25000 | Hyd |
| 102 | janaki | IT | 30000 | Ban |
| 103 | sai | CLAIMS | 15000 | Hyd |
| 104 | Jinna | FIN | 55000 | Ban |
| 105 | Chetan | FIN | 65000 | Hyd |
+-----+-----+-----+-----+-----+
```

5 rows in set (0.00 sec)

--Imported the data from mysql to hdfs ---using Import command.

```
[acadgild@localhost ~]$ sqoop import --connect jdbc:mysql://localhost/simplidb --table employee --
username root -P --target-dir /hadoopdata/hdfs/simp -m 1
```

Total megabyte-milliseconds taken by all map tasks=58985472

Map-Reduce Framework

Map input records=5

Map output records=5

Input split bytes=87

Spilled Records=0

Failed Shuffles=0

Merged Map outputs=0

GC time elapsed (ms)=585

CPU time spent (ms)=10410

Physical memory (bytes) snapshot=94982144

Virtual memory (bytes) snapshot=2063585280

Total committed heap usage (bytes)=32571392

File Input Format Counters

Bytes Read=0

File Output Format Counters

Bytes Written=125

18/05/13 19:16:16 INFO mapreduce.ImportJobBase: Transferred 125 bytes in 160.0371 seconds (0.7811 bytes/sec)

18/05/13 19:16:16 INFO mapreduce.ImportJobBase: Retrieved 5 records.

[acadgild@localhost ~]\$ hadoop fs -ls /hadoopdata/hdfs/simp

18/05/13 19:22:13 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable

Found 2 items

-rw-r--r-- 1 acadgild supergroup 0 2018-05-13 19:16 /hadoopdata/hdfs/simp/_SUCCESS

-rw-r--r-- 1 acadgild supergroup 125 2018-05-13 19:16 /hadoopdata/hdfs/simp/part-m-00000

```
[acadgild@localhost ~]$ hadoop fs -cat /hadoopdata/hdfs/simp/part-m-00000
```

18/05/13 19:22:47 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable

IT,101,soundarya,Hyd,25000

IT,102,janaki,Ban,30000

CLAIMS,103,sai,Hyd,15000

FIN,104,Jinna,Ban,55000

FIN,105,Chetan,Hyd,65000

---Now as per task 1 need to export the data from HDFS to MYSQL. Created table employee1 as follows to export the data from /hadoopdata/hdfs/simp/part-m-00000 to employee1.

```
mysql> create table employee1(
```

```
-> emp_dep varchar(20) NULL,
```

```
-> emp_id INT NOT NULL,
```

```
-> emp_name varchar(20),
```

```
-> loc varchar(20
```

```
-> ) NULL,
```

```
-> sal INT NULL,
```

```
-> PRIMARY KEY (emp_id)
```

```
-> );
```

Query OK, 0 rows affected (1.14 sec)

```
mysql> describe employee1;
```

Field	Type	Null	Key	Default	Extra
emp_dep	varchar(20)	YES		NULL	

emp_id	int(11)	NO	PRI	NULL		
emp_name	varchar(20)	YES		NULL		
loc	varchar(20)	YES		NULL		
sal	int(11)	YES		NULL		

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

5 rows in set (0.16 sec)

---EXPORT DATA from HDFS to employee1-----

```
[acadgild@localhost ~]$ sqoop export -m 1 --connect jdbc:mysql://localhost/simplidb --username root
--password Root@123 --table employee1 --export-dir /hadoopdata/hdfs/simp/part-m-00000
```

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.

18/05/13 19:28:50 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6

18/05/13 19:28:51 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.

18/05/13 19:28:52 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.

18/05/13 19:28:52 INFO tool.CodeGenTool: Beginning code generation

Sun May 13 19:28:54 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 using 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.

18/05/13 19:29:00 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `employee1` AS t LIMIT 1

18/05/13 19:29:00 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `employee1` AS t LIMIT 1

18/05/13 19:29:00 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /home/acadgild/install/hadoop/hadoop-2.6.5

Note: /tmp/sqoop-acadgild/compile/cc4fc98041ad28d4c37f142370bed6b4/employee1.java uses or overrides a deprecated API.

Note: Recompile with -Xlint:deprecation for details.

18/05/13 19:29:18 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-acadgild/compile/cc4fc98041ad28d4c37f142370bed6b4/employee1.jar

18/05/13 19:29:18 INFO mapreduce.ExportJobBase: Beginning export of employee1

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/StaticLoggerBinder.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]

18/05/13 19:29:20 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable

18/05/13 19:29:21 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar

18/05/13 19:29:29 INFO Configuration.deprecation: mapred.reduce.tasks.speculative.execution is deprecated. Instead, use mapreduce.reduce.speculative

18/05/13 19:29:29 INFO Configuration.deprecation: mapred.map.tasks.speculative.execution is deprecated. Instead, use mapreduce.map.speculative

18/05/13 19:29:29 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps

18/05/13 19:29:30 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032

18/05/13 19:29:46 INFO input.FileInputFormat: Total input paths to process : 1

18/05/13 19:29:46 INFO input.FileInputFormat: Total input paths to process : 1

18/05/13 19:29:47 INFO mapreduce.JobSubmitter: number of splits:1

18/05/13 19:29:47 INFO Configuration.deprecation: mapred.map.tasks.speculative.execution is deprecated. Instead, use mapreduce.map.speculative

18/05/13 19:29:49 INFO mapreduce.JobSubmitter: Submitting tokens for job:
job_1526185864223_0010

18/05/13 19:29:54 INFO impl.YarnClientImpl: Submitted application application_1526185864223_0010

18/05/13 19:29:56 INFO mapreduce.Job: The url to track the job:
http://localhost:8088/proxy/application_1526185864223_0010/

18/05/13 19:29:56 INFO mapreduce.Job: Running job: job_1526185864223_0010

18/05/13 19:30:54 INFO mapreduce.Job: Job job_1526185864223_0010 running in uber mode : false

18/05/13 19:30:54 INFO mapreduce.Job: map 0% reduce 0%

18/05/13 19:31:47 INFO mapreduce.Job: map 100% reduce 0%

18/05/13 19:31:49 INFO mapreduce.Job: Job job_1526185864223_0010 completed successfully

18/05/13 19:31:50 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=264

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)=44614

Total time spent by all reduces in occupied slots (ms)=0

Total time spent by all map tasks (ms)=44614

Total vcore-milliseconds taken by all map tasks=44614

Total megabyte-milliseconds taken by all map tasks=45684736

Map-Reduce Framework

Map input records=5

Map output records=5

Input split bytes=136

Spilled Records=0

Failed Shuffles=0

Merged Map outputs=0

GC time elapsed (ms)=395

CPU time spent (ms)=7550

Physical memory (bytes) snapshot=95412224

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

18/05/13 19:31:50 INFO mapreduce.ExportJobBase: Transferred 264 bytes in 140.7353 seconds (1.8759 bytes/sec)

18/05/13 19:31:50 INFO mapreduce.ExportJobBase: Exported 5 records.


```
mysql> select * from employee1;
```

```
+-----+-----+-----+-----+
| emp_dep | emp_id | emp_name | loc | sal |
+-----+-----+-----+-----+
| IT      | 101    | soundarya | Hyd | 25000 |
| IT      | 102    | janaki    | Ban | 30000 |
| CLAIMS  | 103    | sai       | Hyd | 15000 |
| FIN     | 104    | Jinna     | Ban | 55000 |
| FIN     | 105    | Chetan    | Hyd | 65000 |
```

```
5 rows in set (0.01 sec)
```