TASK 2: Use Sqoop tool to export data present in SQOOPOUT folder made while demo of Import table

with parameter person_id =3.

----I have inserted one more row with emp_id =106 as follows and then imported the data to hdfs using where clause in import statement.

mysql> insert into employee values(106,'Shamshad','IT',60000,'Ban');

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 |

| 106 | Shamshad | IT | 60000 | Ban |

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

6 rows in set (0.00 sec)

acadgild@localhost ~]$ sqoop import --connect jdbc:mysql://localhost/simplidb --table employee --username root -P --target-dir /hadoopdata/hdfs/simp/append -m 1 --where emp_id=106

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 19:45:22 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6

Enter password:

18/05/13 19:45:28 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.

18/05/13 19:45:28 INFO tool.CodeGenTool: Beginning code generation

Sun May 13 19:45:29 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 19:45:35 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `employee` AS t LIMIT 1

18/05/13 19:45:35 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `employee` AS t LIMIT 1

18/05/13 19:45:35 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /home/acadgild/install/hadoop/hadoop-2.6.5

Note: /tmp/sqoop-acadgild/compile/fc71e0b87d6b0ec2e431440f376636b6/employee.java uses or overrides a deprecated API.

Note: Recompile with -Xlint:deprecation for details.

18/05/13 19:45:52 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-acadgild/compile/fc71e0b87d6b0ec2e431440f376636b6/employee.jar

18/05/13 19:45:53 WARN manager.MySQLManager: It looks like you are importing from mysql.

18/05/13 19:45:53 WARN manager.MySQLManager: This transfer can be faster! Use the --direct

18/05/13 19:45:53 WARN manager.MySQLManager: option to exercise a MySQL-specific fast path.

18/05/13 19:45:53 INFO manager.MySQLManager: Setting zero DATETIME behavior to convertToNull (mysql)

18/05/13 19:45:53 INFO mapreduce.ImportJobBase: Beginning import of employee

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 19:45:55 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable

18/05/13 19:45:55 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar

18/05/13 19:46:00 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps

18/05/13 19:46:01 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032

Sun May 13 19:46:14 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 19:46:16 INFO db.DBInputFormat: Using read commited transaction isolation

18/05/13 19:46:16 INFO mapreduce.JobSubmitter: number of splits:1

18/05/13 19:46:19 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1526185864223_0011

18/05/13 19:46:22 INFO impl.YarnClientImpl: Submitted application application_1526185864223_0011

18/05/13 19:46:22 INFO mapreduce.Job: The url to track the job: http://localhost:8088/proxy/application_1526185864223_0011/

18/05/13 19:46:22 INFO mapreduce.Job: Running job: job_1526185864223_0011

18/05/13 19:47:14 INFO mapreduce.Job: Job job_1526185864223_0011 running in uber mode : false

18/05/13 19:47:14 INFO mapreduce.Job: map 0% reduce 0%

18/05/13 19:48:03 INFO mapreduce.Job: map 100% reduce 0%

18/05/13 19:48:05 INFO mapreduce.Job: Job job_1526185864223_0011 completed successfully

18/05/13 19:48:06 INFO mapreduce.Job: Counters: 30

        File System Counters

                FILE: Number of bytes read=0

FILE: Number of bytes written=128039

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=26

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)=40988

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

Total time spent by all map tasks (ms)=40988

Total vcore-milliseconds taken by all map tasks=40988

Total megabyte-milliseconds taken by all map tasks=41971712

Map-Reduce Framework

Map input records=1

Map output records=1

Input split bytes=87

Spilled Records=0

Failed Shuffles=0

Merged Map outputs=0

GC time elapsed (ms)=395

CPU time spent (ms)=9000

Physical memory (bytes) snapshot=104955904

Virtual memory (bytes) snapshot=2063593472

Total committed heap usage (bytes)=32571392

File Input Format Counters

Bytes Read=0

File Output Format Counters

Bytes Written=26

18/05/13 19:48:06 INFO mapreduce.ImportJobBase: Transferred 26 bytes in 125.7399 seconds (0.2068 bytes/sec)

18/05/13 19:48:06 INFO mapreduce.ImportJobBase: Retrieved 1 records.

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

18/05/13 19:49:10 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 19:48 /hadoopdata/hdfs/simp/append/_SUCCESS

-rw-r--r--   1 acadgild supergroup   26 2018-05-13 19:47 /hadoopdata/hdfs/simp/append/part-m-00000

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

18/05/13 19:50:27 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable

IT,106,Shamshad,Ban,60000

----Now using command --update-mode allowinsert which updates the record or if record does not exist then inserts that particular record----

[acadgild@localhost **~]$ sqoop export -m 1 --connect jdbc:mysql://localhost/simplidb --username root --password Root@123 --table employee1 --update-key "emp_id" --update-mode allowinsert --export-dir** /hadoopdata/hdfs/simp/append/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 20:18:40 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6

18/05/13 20:18:40 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.

18/05/13 20:18:42 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.

18/05/13 20:18:42 INFO tool.CodeGenTool: Beginning code generation

Sun May 13 20:18:43 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 20:18:49 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `employee1` AS t LIMIT 1

18/05/13 20:18:49 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `employee1` AS t LIMIT 1

18/05/13 20:18:49 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /home/acadgild/install/hadoop/hadoop-2.6.5

Note: /tmp/sqoop-acadgild/compile/9e76b075dbe6269b710f06609ed13f09/employee1.java uses or overrides a deprecated API.

Note: Recompile with -Xlint:deprecation for details.

18/05/13 20:19:07 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-acadgild/compile/9e76b075dbe6269b710f06609ed13f09/employee1.jar

18/05/13 20:19:07 WARN manager.MySQLManager: MySQL Connector upsert functionality is using INSERT ON

18/05/13 20:19:07 WARN manager.MySQLManager: DUPLICATE KEY UPDATE clause that relies on table's unique key.

18/05/13 20:19:07 WARN manager.MySQLManager: Insert/update distinction is therefore independent on column

18/05/13 20:19:07 WARN manager.MySQLManager: names specified in --update-key parameter. Please see MySQL

18/05/13 20:19:07 WARN manager.MySQLManager: documentation for additional limitations.

18/05/13 20:19:07 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 20:19:09 WARN util.NativeCodeLoader: Unable to load native-hadoop library for your platform... using builtin-java classes where applicable

18/05/13 20:19:09 INFO Configuration.deprecation: mapred.jar is deprecated. Instead, use mapreduce.job.jar

18/05/13 20:19:15 INFO Configuration.deprecation: mapred.reduce.tasks.speculative.execution is deprecated. Instead, use mapreduce.reduce.speculative

18/05/13 20:19:16 INFO Configuration.deprecation: mapred.map.tasks.speculative.execution is deprecated. Instead, use mapreduce.map.speculative

18/05/13 20:19:16 INFO Configuration.deprecation: mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps

18/05/13 20:19:17 INFO client.RMProxy: Connecting to ResourceManager at localhost/127.0.0.1:8032

18/05/13 20:19:27 INFO input.FileInputFormat: Total input paths to process : 1

18/05/13 20:19:27 INFO input.FileInputFormat: Total input paths to process : 1

18/05/13 20:19:28 INFO mapreduce.JobSubmitter: number of splits:1

18/05/13 20:19:28 INFO Configuration.deprecation: mapred.map.tasks.speculative.execution is deprecated. Instead, use mapreduce.map.speculative

18/05/13 20:19:30 INFO mapreduce.JobSubmitter: Submitting tokens for job: job_1526185864223_0012

18/05/13 20:19:33 INFO impl.YarnClientImpl: Submitted application application_1526185864223_0012

18/05/13 20:19:34 INFO mapreduce.Job: The url to track the job:
http://localhost:8088/proxy/application_1526185864223_0012/

18/05/13 20:19:34 INFO mapreduce.Job: Running job: job_1526185864223_0012

18/05/13 20:20:31 INFO mapreduce.Job: Job job_1526185864223_0012 running in uber mode : false

18/05/13 20:20:31 INFO mapreduce.Job: map 0% reduce 0%

18/05/13 20:21:22 INFO mapreduce.Job: map 100% reduce 0%

18/05/13 20:21:24 INFO mapreduce.Job: Job job_1526185864223_0012 completed successfully

18/05/13 20:21:25 INFO mapreduce.Job: Counters: 30

      File System Counters

            FILE: Number of bytes read=0

            FILE: Number of bytes written=127740

            FILE: Number of read operations=0

            FILE: Number of large read operations=0

            FILE: Number of write operations=0

            HDFS: Number of bytes read=172

            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)=43166

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

            Total time spent by all map tasks (ms)=43166

Total vcore-milliseconds taken by all map tasks=43166

Total megabyte-milliseconds taken by all map tasks=44201984

Map-Reduce Framework

Map input records=1

Map output records=1

Input split bytes=143

Spilled Records=0

Failed Shuffles=0

Merged Map outputs=0

GC time elapsed (ms)=448

CPU time spent (ms)=7630

Physical memory (bytes) snapshot=103227392

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 20:21:25 INFO mapreduce.ExportJobBase: Transferred 172 bytes in 128.5367 seconds (1.3381 bytes/sec)

18/05/13 20:21:25 INFO mapreduce.ExportJobBase: Exported 1 records.

--Checking whether the last updated column exported to MYSQL or not----

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 |
| IT      |   106 | Shamshad  | Ban  | 60000 |
+---------+--------+-----------+------+-------+
```

6 rows in set (0.01 sec)