ASSIGNMENT11.2

Perform incremental load in Hive
Read from MySQL Table and load it in Hive table.
Create hive table if it does not exist.
If it exists, perform the incremental load.

mysql> create table employee1
    -> (id int,name varchar(20),salary int);
Query OK, 0 rows affected (0.07 sec)

mysql> insert into employee1 values(1,'vishnu',500);
Query OK, 1 row affected (0.04 sec)

mysql> insert into employee1 values(2,'vipin',300);
Query OK, 1 row affected (0.00 sec)


```
[acadgild@localhost ~]$ sqoop import --connect
jdbc:mysql://localhost/db1 --username 'root' -P --table 'employee1'
--target-dir '/display' --hive-import --create-hive-table
--hive-table 'employeehive' -m 1
Warning: /usr/local/sqoop/../hcatalog does not exist! HCatalog jobs
will fail.
Please set $HCAT_HOME to the root of your HCatalog installation.
Warning: /usr/local/sqoop/../accumulo does not exist! Accumulo
imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
Warning: /usr/local/sqoop/../zookeeper does not exist! Accumulo
imports will fail.
Please set $ZOOKEEPER_HOME to the root of your Zookeeper
installation.
2017-11-03 23:32:07,098 INFO  [main] sqoop.Sqoop: Running Sqoop
version: 1.4.5
Enter password:
2017-11-03 23:32:08,944 INFO  [main] tool.BaseSqoopTool: Using
Hive-specific delimiters for output. You can override
2017-11-03 23:32:08,944 INFO  [main] tool.BaseSqoopTool: delimiters
with --fields-terminated-by, etc.
2017-11-03 23:32:09,463 INFO  [main] manager.MySQLManager: Preparing
to use a MySQL streaming resultset.
2017-11-03 23:32:09,491 INFO  [main] tool.CodeGenTool: Beginning code
generation
2017-11-03 23:32:10,162 INFO  [main] manager.SqlManager: Executing
SQL statement: SELECT t.* FROM `employee` AS t LIMIT 1
2017-11-03 23:32:10,271 INFO  [main] manager.SqlManager: Executing
SQL statement: SELECT t.* FROM `employee` AS t LIMIT 1
2017-11-03 23:32:10,294 INFO  [main] orm.CompilationManager:
HADOOP_MAPRED_HOME is /usr/local/hadoop-2.6.0
```

Note:
/tmp/sqoop-acadgild/compile/77e0717fd2a983e81d99806d7310dfea/employee
.java uses or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
2017-11-03 23:32:14,419 INFO [main] orm.CompilationManager: Writing
jar file:
/tmp/sqoop-acadgild/compile/77e0717fd2a983e81d99806d7310dfea/employee
.jar
2017-11-03 23:32:14,468 WARN [main] manager.MySQLManager: It looks
like you are importing from mysql.
2017-11-03 23:32:14,468 WARN [main] manager.MySQLManager: This
transfer can be faster! Use the --direct
2017-11-03 23:32:14,468 WARN [main] manager.MySQLManager: option to
exercise a MySQL-specific fast path.
2017-11-03 23:32:14,468 INFO [main] manager.MySQLManager: Setting
zero DATETIME behavior to convertToNull (mysql)
2017-11-03 23:32:14,483 INFO [main] mapreduce.ImportJobBase:
Beginning import of employee
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in
[jar:file:/usr/local/hbase/lib/slf4j-log4j12-1.6.4.jar!/org/slf4j/imp
l/StaticLoggerBinder.class]
SLF4J: Found binding in
[jar:file:/usr/local/hadoop-2.6.0/share/hadoop/common/lib/slf4j-log4j
12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an
explanation.
2017-11-03 23:32:15,175 WARN [main] util.NativeCodeLoader: Unable to
load native-hadoop library for your platform... using builtin-java
classes where applicable
2017-11-03 23:32:15,198 INFO [main] Configuration.deprecation:
mapred.jar is deprecated. Instead, use mapreduce.job.jar
2017-11-03 23:32:16,450 INFO [main] Configuration.deprecation:
mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps
2017-11-03 23:32:16,844 INFO [main] client.RMProxy: Connecting to
ResourceManager at /0.0.0.0:8032
2017-11-03 23:32:21,565 INFO [main] db.DBInputFormat: Using read
commited transaction isolation
2017-11-03 23:32:21,766 INFO [main] mapreduce.JobSubmitter: number
of splits:1
2017-11-03 23:32:21,844 INFO [main] Configuration.deprecation:
mapred.job.name is deprecated. Instead, use mapreduce.job.name
2017-11-03 23:32:21,845 INFO [main] Configuration.deprecation:
mapred.cache.files.timestamps is deprecated. Instead, use
mapreduce.job.cache.files.timestamps
2017-11-03 23:32:21,845 INFO [main] Configuration.deprecation:
mapreduce.map.class is deprecated. Instead, use
mapreduce.job.map.class
2017-11-03 23:32:21,846 INFO [main] Configuration.deprecation:

mapreduce.inputformat.class is deprecated. Instead, use
mapreduce.job.inputformat.class
2017-11-03 23:32:21,848 INFO [main] Configuration.deprecation:
mapreduce.outputformat.class is deprecated. Instead, use
mapreduce.job.outputformat.class
2017-11-03 23:32:21,848 INFO [main] Configuration.deprecation:
mapred.output.value.class is deprecated. Instead, use
mapreduce.job.output.value.class
2017-11-03 23:32:21,849 INFO [main] Configuration.deprecation:
mapred.output.dir is deprecated. Instead, use
mapreduce.output.fileoutputformat.outputdir
2017-11-03 23:32:21,849 INFO [main] Configuration.deprecation:
mapred.cache.files is deprecated. Instead, use
mapreduce.job.cache.files
2017-11-03 23:32:21,849 INFO [main] Configuration.deprecation:
mapred.working.dir is deprecated. Instead, use
mapreduce.job.working.dir
2017-11-03 23:32:21,849 INFO [main] Configuration.deprecation:
mapred.job.classpath.files is deprecated. Instead, use
mapreduce.job.classpath.files
2017-11-03 23:32:21,849 INFO [main] Configuration.deprecation:
user.name is deprecated. Instead, use mapreduce.job.user.name
2017-11-03 23:32:21,849 INFO [main] Configuration.deprecation:
mapred.reduce.tasks is deprecated. Instead, use mapreduce.job.reduces
2017-11-03 23:32:21,849 INFO [main] Configuration.deprecation:
mapred.cache.files.filesizes is deprecated. Instead, use
mapreduce.job.cache.files.filesizes
2017-11-03 23:32:21,849 INFO [main] Configuration.deprecation:
mapred.output.key.class is deprecated. Instead, use
mapreduce.job.output.key.class
2017-11-03 23:32:22,368 INFO [main] mapreduce.JobSubmitter:
Submitting tokens for job: job_1509729939048_0001
2017-11-03 23:32:23,521 INFO [main] impl.YarnClientImpl: Submitted
application application_1509729939048_0001 to ResourceManager at /
0.0.0.0:8032
2017-11-03 23:32:23,690 INFO [main] mapreduce.Job: The url to track
the job:
http://http://localhost:8088/proxy/application_1509729939048_0001/
2017-11-03 23:32:23,691 INFO [main] mapreduce.Job: Running job:
job_1509729939048_0001
2017-11-03 23:32:46,419 INFO [main] mapreduce.Job: Job
job_1509729939048_0001 running in uber mode : false
2017-11-03 23:32:46,429 INFO [main] mapreduce.Job: map 0% reduce 0%
2017-11-03 23:32:58,765 INFO [main] mapreduce.Job: map 100% reduce
0%
2017-11-03 23:32:59,834 INFO [main] mapreduce.Job: Job
job_1509729939048_0001 completed successfully
2017-11-03 23:33:00,215 ERROR [main] tool.ImportTool: Imported
Failed: No enum constant

```
org.apache.hadoop.mapreduce.JobCounter.MB_MILLIS_MAPS


[acadgild@localhost ~]$ hadoop fs -cat /display/part-m-00000
17/11/03 23:35:21 WARN util.NativeCodeLoader: Unable to load
native-hadoop library for your platform... using builtin-java classes
where applicable
1#vishnu#500
2#vipin#300
```

Incremental importing

mysql> insert into employee1 values(3,'vishnuvijay',500);
Query OK, 1 row affected (0.00 sec)

mysql> insert into employee1 values(4,'vijay',500);
Query OK, 1 row affected (0.00 sec)

```
[acadgild@localhost ~]$ sqoop import --connect
jdbc:mysql://localhost/db1 --username 'root' -P --table 'employee1'
--target-dir '/display' --hive-import --create-hive-table
--hive-table 'employeehive1' --incremental append --check-column id
--last-value 3 -m 1
Warning: /usr/local/sqoop/../hcatalog does not exist! HCatalog jobs
will fail.
Please set $HCAT_HOME to the root of your HCatalog installation.
Warning: /usr/local/sqoop/../accumulo does not exist! Accumulo
imports will fail.
Please set $ACCUMULO_HOME to the root of your Accumulo installation.
Warning: /usr/local/sqoop/../zookeeper does not exist! Accumulo
imports will fail.
Please set $ZOOKEEPER_HOME to the root of your Zookeeper
installation.
2017-11-04 01:09:52,322 INFO  [main] sqoop.Sqoop: Running Sqoop
version: 1.4.5
Enter password:
2017-11-04 01:09:54,197 INFO  [main] tool.BaseSqoopTool: Using
Hive-specific delimiters for output. You can override
2017-11-04 01:09:54,197 INFO  [main] tool.BaseSqoopTool: delimiters
with --fields-terminated-by, etc.
2017-11-04 01:09:54,791 INFO  [main] manager.MySQLManager: Preparing
to use a MySQL streaming resultset.
2017-11-04 01:09:54,813 INFO  [main] tool.CodeGenTool: Beginning code
generation
2017-11-04 01:09:55,534 INFO  [main] manager.SqlManager: Executing
SQL statement: SELECT t.* FROM `employee1` AS t LIMIT 1
2017-11-04 01:09:55,669 INFO  [main] manager.SqlManager: Executing
SQL statement: SELECT t.* FROM `employee1` AS t LIMIT 1
2017-11-04 01:09:55,714 INFO  [main] orm.CompilationManager:
```

HADOOP_MAPRED_HOME is /usr/local/hadoop-2.6.0
Note:
/tmp/sqoop-acadgild/compile/ce8dd9bd0a6f4247b7a683c483db3490/employee
1.java uses or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
2017-11-04 01:10:01,941 INFO [main] orm.CompilationManager: Writing
jar file:
/tmp/sqoop-acadgild/compile/ce8dd9bd0a6f4247b7a683c483db3490/employee
1.jar
SLF4J: Class path contains multiple SLF4J bindings.
SLF4J: Found binding in
[jar:file:/usr/local/hbase/lib/slf4j-log4j12-1.6.4.jar!/org/slf4j/imp
l/StaticLoggerBinder.class]
SLF4J: Found binding in
[jar:file:/usr/local/hadoop-2.6.0/share/hadoop/common/lib/slf4j-log4j
12-1.7.5.jar!/org/slf4j/impl/StaticLoggerBinder.class]
SLF4J: See http://www.slf4j.org/codes.html#multiple_bindings for an
explanation.
2017-11-04 01:10:03,211 WARN [main] util.NativeCodeLoader: Unable to
load native-hadoop library for your platform... using builtin-java
classes where applicable
2017-11-04 01:10:04,687 INFO [main] tool.ImportTool: Maximal id
query for free form incremental import: SELECT MAX(`id`) FROM
employee1
2017-11-04 01:10:04,689 INFO [main] tool.ImportTool: Incremental
import based on column `id`
2017-11-04 01:10:04,689 INFO [main] tool.ImportTool: Lower bound
value: 3
2017-11-04 01:10:04,689 INFO [main] tool.ImportTool: Upper bound
value: 4
2017-11-04 01:10:04,690 WARN [main] manager.MySQLManager: It looks
like you are importing from mysql.
2017-11-04 01:10:04,690 WARN [main] manager.MySQLManager: This
transfer can be faster! Use the --direct
2017-11-04 01:10:04,690 WARN [main] manager.MySQLManager: option to
exercise a MySQL-specific fast path.
2017-11-04 01:10:04,690 INFO [main] manager.MySQLManager: Setting
zero DATETIME behavior to convertToNull (mysql)
2017-11-04 01:10:04,698 INFO [main] mapreduce.ImportJobBase:
Beginning import of employee1
2017-11-04 01:10:04,753 INFO [main] Configuration.deprecation:
mapred.jar is deprecated. Instead, use mapreduce.job.jar
2017-11-04 01:10:04,847 INFO [main] Configuration.deprecation:
mapred.map.tasks is deprecated. Instead, use mapreduce.job.maps
2017-11-04 01:10:05,393 INFO [main] client.RMProxy: Connecting to
ResourceManager at /0.0.0.0:8032
2017-11-04 01:10:08,491 INFO [main] db.DBInputFormat: Using read
commited transaction isolation
2017-11-04 01:10:08,622 INFO [main] mapreduce.JobSubmitter: number

of splits:1
2017-11-04 01:10:08,662 INFO [main] Configuration.deprecation:
mapred.job.name is deprecated. Instead, use mapreduce.job.name
2017-11-04 01:10:08,664 INFO [main] Configuration.deprecation:
mapred.cache.files.timestamps is deprecated. Instead, use
mapreduce.job.cache.files.timestamps
2017-11-04 01:10:08,664 INFO [main] Configuration.deprecation:
mapreduce.map.class is deprecated. Instead, use
mapreduce.job.map.class
2017-11-04 01:10:08,667 INFO [main] Configuration.deprecation:
mapreduce.inputformat.class is deprecated. Instead, use
mapreduce.job.inputformat.class
2017-11-04 01:10:08,672 INFO [main] Configuration.deprecation:
mapreduce.outputformat.class is deprecated. Instead, use
mapreduce.job.outputformat.class
2017-11-04 01:10:08,672 INFO [main] Configuration.deprecation:
mapred.output.value.class is deprecated. Instead, use
mapreduce.job.output.value.class
2017-11-04 01:10:08,673 INFO [main] Configuration.deprecation:
mapred.output.dir is deprecated. Instead, use
mapreduce.output.fileoutputformat.outputdir
2017-11-04 01:10:08,674 INFO [main] Configuration.deprecation:
mapred.cache.files is deprecated. Instead, use
mapreduce.job.cache.files
2017-11-04 01:10:08,674 INFO [main] Configuration.deprecation:
mapred.working.dir is deprecated. Instead, use
mapreduce.job.working.dir
2017-11-04 01:10:08,674 INFO [main] Configuration.deprecation:
mapred.job.classpath.files is deprecated. Instead, use
mapreduce.job.classpath.files
2017-11-04 01:10:08,675 INFO [main] Configuration.deprecation:
user.name is deprecated. Instead, use mapreduce.job.user.name
2017-11-04 01:10:08,675 INFO [main] Configuration.deprecation:
mapred.reduce.tasks is deprecated. Instead, use mapreduce.job.reduces
2017-11-04 01:10:08,675 INFO [main] Configuration.deprecation:
mapred.cache.files.filesizes is deprecated. Instead, use
mapreduce.job.cache.files.filesizes
2017-11-04 01:10:08,675 INFO [main] Configuration.deprecation:
mapred.output.key.class is deprecated. Instead, use
mapreduce.job.output.key.class
2017-11-04 01:10:09,038 INFO [main] mapreduce.JobSubmitter:
Submitting tokens for job: job_1509729939048_0014
2017-11-04 01:10:09,978 INFO [main] impl.YarnClientImpl: Submitted
application application_1509729939048_0014 to ResourceManager at /
0.0.0.0:8032
2017-11-04 01:10:10,211 INFO [main] mapreduce.Job: The url to track
the job:
http://http://localhost:8088/proxy/application_1509729939048_0014/
2017-11-04 01:10:10,233 INFO [main] mapreduce.Job: Running job:

```
job_1509729939048_0014
2017-11-04 01:10:27,005 INFO  [main] mapreduce.Job: Job
job_1509729939048_0014 running in uber mode : false
2017-11-04 01:10:27,026 INFO  [main] mapreduce.Job:  map 0% reduce 0%
2017-11-04 01:10:43,703 INFO  [main] mapreduce.Job:  map 100% reduce
0%
2017-11-04 01:10:43,750 INFO  [main] mapreduce.Job: Job
job_1509729939048_0014 completed successfully
2017-11-04 01:10:44,204 ERROR [main] tool.ImportTool: Imported
Failed: No enum constant
org.apache.hadoop.mapreduce.JobCounter.MB_MILLIS_MAPS


[acadgild@localhost ~]$ hadoop fs -cat /display/part-m-0000017/11/04
01:17:46 WARN util.NativeCodeLoader: Unable to load native-hadoop
library for your platform... using builtin-java classes where
applicable
3#vishnuvijay#500
4#vijay#500
```