

Sqoop Setup

1. 簡介：

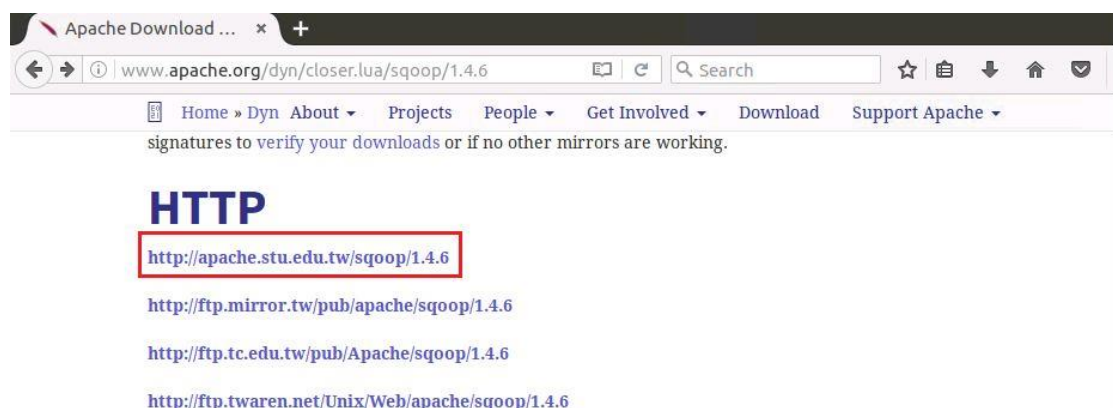
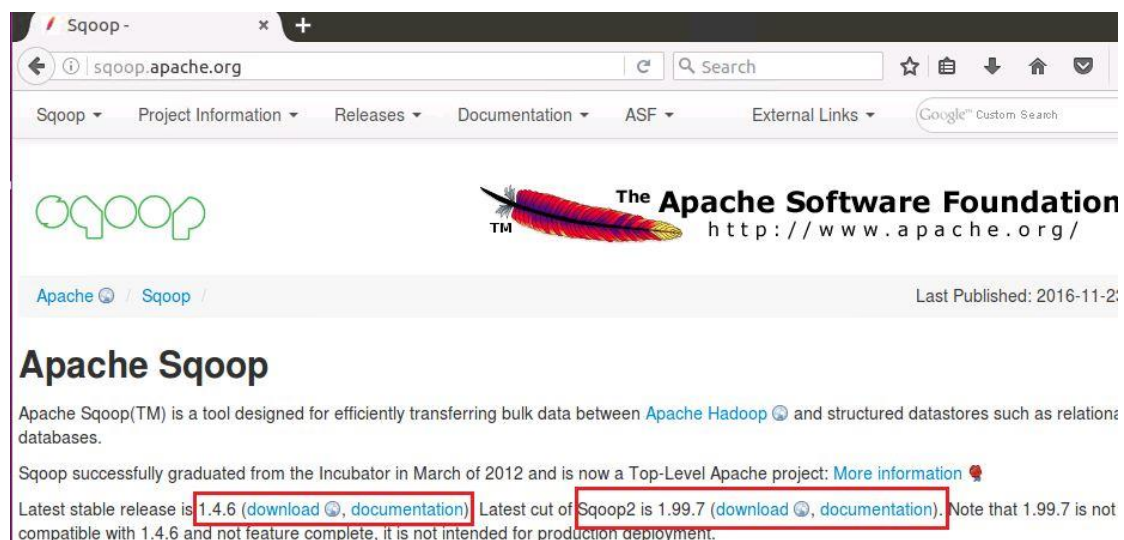
- Sqoop 是設計來讓結構化資料(如.MySQL)可以讓 HDFS 讀入和取出



- Sqoop 目前有 Sqoop1 和 Sqoop2 兩種版本。Sqoop1 較為簡化，本文件將進行的是 Sqoop1 的安裝和使用測試。

2. install Sqoop

- 訪問 Sqoop 網站，可以看到目前 Sqoop1 和 Sqoop2 的最新版本





- 請開啟 HadoopMaster 主機的終端
- 下載 Sqoop1，輸入 `wget`

http://apache.stu.edu.tw/squirrel/1.4.6/squirrel-1.4.6.bin_hadoop-2.0.4-alpha.tar.gz

```
hduser@hadoopmaster: ~
hduser@hadoopmaster:~$ wget http://apache.stu.edu.tw/squirrel/1.4.6/squirrel-1.4.6.bin_hadoop-2.0.4-alpha.tar.gz
--2016-12-25 11:19:26-- http://apache.stu.edu.tw/squirrel/1.4.6/squirrel-1.4.6.bin_hadoop-2.0.4-alpha.tar.gz
Resolving apache.stu.edu.tw (apache.stu.edu.tw)... 120.119.118.1, 2001:e10:c41:eee::1
Connecting to apache.stu.edu.tw (apache.stu.edu.tw)|120.119.118.1|:80... connected.
HTTP request sent, awaiting response... 200 OK
Length: 16870735 (16M) [application/x-gzip]
Saving to: 'squirrel-1.4.6.bin_hadoop-2.0.4-alpha.tar.gz'

squirrel-1.4.6.bin 1%[          ] 216.12K 87.9KB/s
```

- 執行解壓縮，輸入 `tar -zxvf squirrel-1.4.6.bin_hadoop-2.0.4-alpha.tar.gz`

```
hduser@hadoopmaster: ~
hduser@hadoopmaster:~$ tar -zxvf squirrel-1.4.6.bin_hadoop-2.0.4-alpha.tar.gz
squirrel-1.4.6.bin_hadoop-2.0.4-alpha/
squirrel-1.4.6.bin_hadoop-2.0.4-alpha/bin/
squirrel-1.4.6.bin_hadoop-2.0.4-alpha/conf/
squirrel-1.4.6.bin_hadoop-2.0.4-alpha/docs/
squirrel-1.4.6.bin_hadoop-2.0.4-alpha/docs/api/
squirrel-1.4.6.bin_hadoop-2.0.4-alpha/docs/api/com/
squirrel-1.4.6.bin_hadoop-2.0.4-alpha/docs/api/com/cloudera/
squirrel-1.4.6.bin_hadoop-2.0.4-alpha/docs/api/com/cloudera/squirrel/
squirrel-1.4.6.bin_hadoop-2.0.4-alpha/docs/api/com/cloudera/squirrel/lib/
```

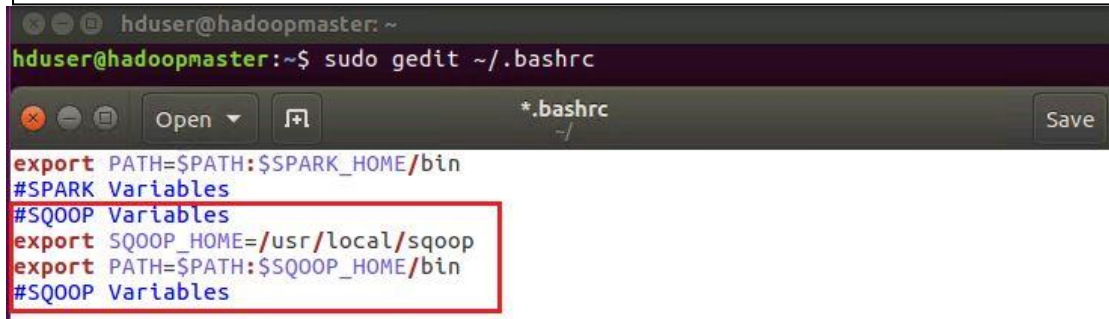
- 移動 Sqoop 資料夾，輸入 `sudo mv squirrel-1.4.6.bin_hadoop-2.0.4-alpha /usr/local/squirrel`

```
hduser@hadoopmaster: ~
hduser@hadoopmaster:~$ sudo mv squirrel-1.4.6.bin_hadoop-2.0.4-alpha /usr/local/squirrel
[sudo] password for hduser:
hduser@hadoopmaster:~$
```

- 修改 ~/.bashrc，輸入 `sudo gedit ~/.bashrc`

- 輸入下列指令

```
export SQOOP_HOME=/usr/local/sqoop
export PATH=$PATH:$SQOOP_HOME/bin
```



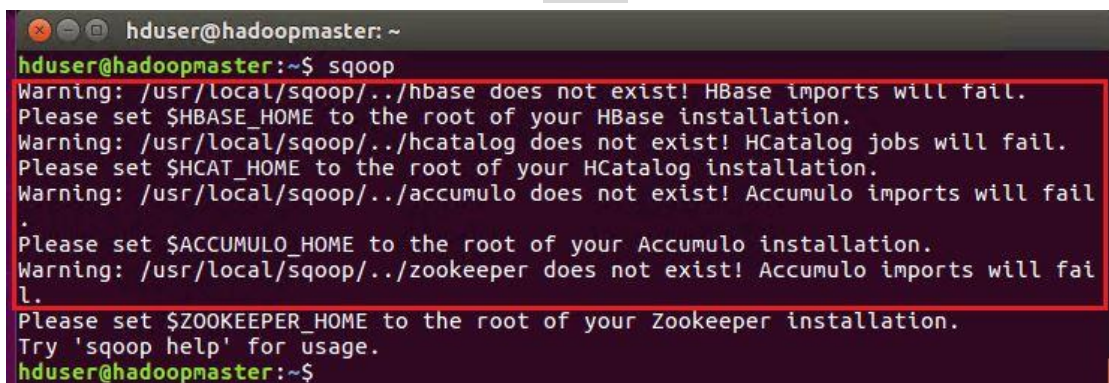
```
hduser@hadoopmaster: ~
hduser@hadoopmaster:~$ sudo gedit ~/.bashrc
*.bashrc
Save
export PATH=$PATH:$SPARK_HOME/bin
#SPARK Variables
#SQOOP Variables
export SQOOP_HOME=/usr/local/sqoop
export PATH=$PATH:$SQOOP_HOME/bin
#SQOOP Variables
```

- 執行更新過的~/.bashrc 檔，輸入 `source ~/.bashrc`



```
hduser@hadoopmaster:~$ source ~/.bashrc
hduser@hadoopmaster:~$
```

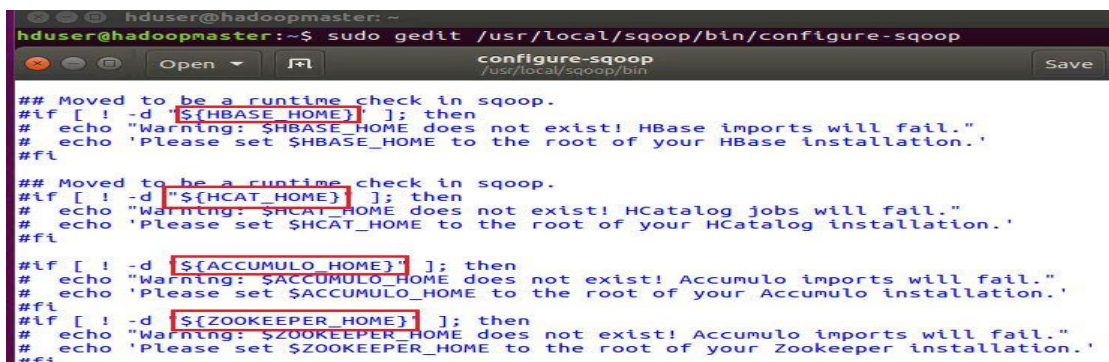
- 此時已可以執行 Sqoop，輸入 `Sqoop` 可看到許多 Warning



```
hduser@hadoopmaster: ~
hduser@hadoopmaster:~$ sqoop
Warning: /usr/local/sqoop/../../hbase does not exist! HBase imports will fail.
Please set $HBASE_HOME to the root of your HBase installation.
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 $ZOOKEEPER_HOME to the root of your Zookeeper installation.
Try 'sqoop help' for usage.
hduser@hadoopmaster:~$
```

- 修改 `configure-sqoop` 檔，輸入 `sudo gedit /usr/local/sqoop/bin/configure-sqoop`
- 將
 - ✓ `HBASE_HOME`
 - ✓ `HCAT_HOME`
 - ✓ `HIVE_CONF_DIR`
 - ✓ `ACCUMULO_HOME`
 - ✓ `ZOOKEEPER_HOME`

的 runtime check warning message 設定以 `#` 註解掉



```
hduser@hadoopmaster: ~
hduser@hadoopmaster:~$ sudo gedit /usr/local/sqoop/bin/configure-sqoop
configure-sqoop
Save
## Moved to be a runtime check in sqoop.
# if [ ! -d "${HBASE_HOME}" ]; then
#   echo "Warning: $HBASE_HOME does not exist! HBase imports will fail."
#   echo 'Please set $HBASE_HOME to the root of your HBase installation.'
# fi
## Moved to be a runtime check in sqoop.
# if [ ! -d "${HCAT_HOME}" ]; then
#   echo "Warning: $HCAT_HOME does not exist! HCatalog jobs will fail."
#   echo 'Please set $HCAT_HOME to the root of your HCatalog installation.'
# fi
# if [ ! -d "${ACCUMULO_HOME}" ]; then
#   echo "Warning: $ACCUMULO_HOME does not exist! Accumulo imports will fail."
#   echo 'Please set $ACCUMULO_HOME to the root of your Accumulo installation.'
# fi
# if [ ! -d "${ZOOKEEPER_HOME}" ]; then
#   echo "Warning: $ZOOKEEPER_HOME does not exist! Accumulo imports will fail."
#   echo 'Please set $ZOOKEEPER_HOME to the root of your Zookeeper installation.'
# fi
```


- 再輸入 `sqoop` 就沒有 warning 了

```
hduser@hadoopmaster: ~  
hduser@hadoopmaster:~$ sqoop  
Try 'sqoop help' for usage.  
hduser@hadoopmaster:~$
```

3. 資料測試—資料匯入 HDFS

- 請將附檔的 ratings.txt 檔放到 /usr/local/sqoop 中
 - ✓ 可以用 `sudo gedit /usr/local/sqoop/ratings.txt`，再將文字內容貼上
- 安裝 MySQL (結構化資料來源)，如已有安裝 MySQL 可跳到匯入資料步驟
- 安裝 MySQL，輸入 `sudo apt-get install mysql-server`，帳密設定為 `hduser/hduser`

```
hduser@hadoopmaster: ~  
hduser@hadoopmaster:~$ sudo apt-get install mysql-server  
Reading package lists... Done  
Building dependency tree  
Reading state information... Done  
The following packages were automatically installed and are no longer required:  
  linux-headers-4.4.0-31 linux-headers-4.4.0-31-generic  
  linux-image-4.4.0-31-generic linux-image-extra-4.4.0-31-generic  
Use 'sudo apt autoremove' to remove them.  
The following additional packages will be installed:  
  libaio1 libevent-core-2.0-5 libhtml-template-perl mysql-client-5.7  
  mysql-client-core-5.7 mysql-common mysql-server-5.7 mysql-server-core-5.7  
Suggested packages:  
  libipc-sharedcache-perl mailx tinyca  
The following NEW packages will be installed:  
  libaio1 libevent-core-2.0-5 libhtml-template-perl mysql-client-5.7  
  mysql-client-core-5.7 mysql-common mysql-server mysql-server-5.7  
  mysql-server-core-5.7  
0 upgraded, 9 newly installed, 0 to remove and 186 not upgraded.  
Need to get 18.3 MB of archives.  
After this operation, 160 MB of additional disk space will be used.  
Do you want to continue? [Y/n]
```

```
hduser@hadoopmaster: ~  
Package configuration
```

Configuring mysql-server-5.7

While not mandatory, it is highly recommended that you set a password for the MySQL administrative "root" user.

If this field is left blank, the password will not be changed.

New password for the MySQL "root" user:

<Ok>

- 安裝完成後，輸入 `mysql -u root -p`，進入 MySQL

```
hduser@hadoopmaster:~$ mysql -u root -p
Enter password:
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 6
Server version: 5.7.16-0ubuntu0.16.04.1 (Ubuntu)

Copyright (c) 2000, 2016, 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>
```

- 建立資料庫 yelp 並建立使用者 hduser，將 yelp 的使用權放給 hduser
 - ✓ `create database yelp;`
 - ✓ `grant all on yelp.* to 'hduser' identified by 'hduser';`
- 輸入 `show databases` 可以查看資料庫狀況，看到已建立 yelp

```
mysql> create database yelp;
Query OK, 1 row affected (0.00 sec)

mysql> grant all on yelp.* to 'hduser' identified by 'hduser';
Query OK, 0 rows affected, 1 warning (0.00 sec)

mysql> show databases;
+-----+
| Database |
+-----+
| information_schema |
| mysql |
| performance_schema |
| sys |
| yelp |
+-----+
5 rows in set (0.00 sec)

mysql>
```

- 切換至 yelp 資料庫，輸入 `use yelp;`
- 建立 table，輸入 `create table ratings(userid integer,itemid integer, ratings integer);`

```
mysql> use yelp;
Database changed
mysql> create table ratings(userid integer,itemid integer,ratings integer);
Query OK, 0 rows affected (0.04 sec)
```

- 接著將測試資料 ratings.txt 內容灌入 table ratings，輸入 `load data local infile '/usr/local/sqoop/ratings.txt' into table ratings fields terminated by ':' optionally enclosed by '\"' lines terminated by '\n' ignore 1 lines(userid,itemid,ratings)`

```
mysql> load data local infile '/usr/local/sqoop/ratings.txt' into table ratings
fields terminated by ':' optionally enclosed by '\"' ignore 1 lines(userid,item
id,ratings);
Query OK, 263778 rows affected, 3 warnings (3.14 sec)
Records: 263778 Deleted: 0 Skipped: 0 Warnings: 3
```

- 檢視匯入的資料，輸入 `select * from ratings limit 10`

```
mysql> select * from ratings limit 10;
+-----+-----+-----+
| userid | itemid | ratings |
+-----+-----+-----+
| 0      | 0      | 4      |
| 0      | 1      | 5      |
| 0      | 7495   | 3      |
| 0      | 7496   | 5      |
| 0      | 7497   | 5      |
| 0      | 7498   | 4      |
| 0      | 23688  | 3      |
| 0      | 23689  | 2      |
| 0      | 23690  | 3      |
| 0      | 23691  | 5      |
+-----+-----+-----+
10 rows in set (0.00 sec)
```

- 離開 MySQL, 輸入 `quit`
- 然後將 MySQL 資料轉入 hdfs 中
 - ✓ 開啟伺服器 HadoopMaster & HadoopSlave1 & HadoopSlave2
 - ✓ 開啟 hadoop, 於 HadoopMaster 的終端輸入 `start-all.sh`

```
hduser@hadoopmaster: ~
hduser@hadoopmaster:~$ start-all.sh
This script is Deprecated. Instead use start-dfs.sh and start-yarn.sh
Starting namenodes on [hadoopmaster]
hadoopmaster: starting namenode, logging to /usr/local/hadoop/logs/hadoop-hduser
-namenode-hadoopmaster.out
hadoopslave1: starting datanode, logging to /usr/local/hadoop/logs/hadoop-hduser
-datanode-hadoopslave1.out
hadoopslave2: starting datanode, logging to /usr/local/hadoop/logs/hadoop-hduser
-datanode-hadoopslave2.out
Starting secondary namenodes [0.0.0.0]
0.0.0.0: starting secondarynamenode, logging to /usr/local/hadoop/logs/hadoop-hd
user-secondarynamenode-hadoopmaster.out
starting yarn daemons
starting resourcemanager, logging to /usr/local/hadoop/logs/yarn-hduser-resource
manager-hadoopmaster.out
hadoopslave1: starting nodemanager, logging to /usr/local/hadoop/logs/yarn-hduse
r-nodemanager-hadoopslave1.out
hadoopslave2: starting nodemanager, logging to /usr/local/hadoop/logs/yarn-hduse
r-nodemanager-hadoopslave2.out
hduser@hadoopmaster:~$ jps
6771 NameNode
6968 SecondaryNameNode
7244 Jps
7166 ResourceManager
```

- 由 MySQL 讀取資料, 需有 JDBC Driver(網址), 執行下列指令來取得
 - ✓ `wget`
 - <http://dev.mysql.com/get/Downloads/Connector-J/mysql-connector-java-5.1.40.tar.gz>
 - ✓ `tar mysql-connector-java-5.1.40.tar.gz`
 - ✓ `sudo mv mysql-connector-java-5.1.40 /usr/local/mysql-connector`

MySQL :: Download ... x +

dev.mysql.com/downloads/connector, Search ☆ 📁 ⬇ » ≡

Generally Available (GA) Releases **Development Releases**

Connector/J 5.1.40

Select Platform: Platform Independent ▾ [Looking for previous GA versions?](#)

Platform Independent (Architecture Independent), Compressed TAR Archive (mysql-connector-java-5.1.40.tar.gz)	5.1.40	3.7M	Download
MD5: 415a375cf8a096ef0aa775a4ae36916d Signature			
Platform Independent (Architecture Independent), ZIP Archive (mysql-connector-java-5.1.40.zip)	5.1.40	4.1M	Download
MD5: ed67bcb617bc949537219d42b8407d5 Signature			

MySQL :: Begin You... x +

dev.mysql.com/downloads/file/?id=46564 x Search ☆ 📁 ⬇ » ≡

Comment in the MySQL Documentation

[Login »](#)
using my Oracle Web account

[Sign Up »](#)
for an Oracle Web account

MySQL.com is using Oracle SSO for authentication. If you already have an Oracle Web account, click the Login link. Otherwise, you can signup for a free account by clicking the Sign Up link and following the instructions.

[No thanks, just start my download.](#)

```
hduser@hadoopmaster: ~
hduser@hadoopmaster:~$ wget http://dev.mysql.com/get/Downloads/Connector-J/mysql-connector-java-5.1.40.tar.gz
--2016-12-25 16:05:12-- http://dev.mysql.com/get/Downloads/Connector-J/mysql-connector-java-5.1.40.tar.gz
Resolving dev.mysql.com (dev.mysql.com)... 137.254.60.11
Connecting to dev.mysql.com (dev.mysql.com)|137.254.60.11|:80... connected.
HTTP request sent, awaiting response... 302 Found
Location: http://cdn.mysql.com//Downloads/Connector-J/mysql-connector-java-5.1.40.tar.gz [following]
--2016-12-25 16:05:13-- http://cdn.mysql.com//Downloads/Connector-J/mysql-connector-java-5.1.40.tar.gz
Resolving cdn.mysql.com (cdn.mysql.com)... 104.115.235.134
Connecting to cdn.mysql.com (cdn.mysql.com)|104.115.235.134|:80... connected.
HTTP request sent, awaiting response... 200 OK
Length: 3911557 (3.7M) [application/x-tar-gz]
Saving to: 'mysql-connector-java-5.1.40.tar.gz'

mysql-connector-jav 30%[====>] 1.15M 209KB/s eta 25s
```

- 接下來將 jar 檔移到/sqoop/lib，執行 cp
`/usr/local/mysql-connector/mysql-connector-java-5.1.40-bin.jar`
`/usr/local/sqoop/lib`

```
hduser@hadoopmaster: ~
hduser@hadoopmaster:~$ cp /usr/local/mysql-connector/mysql-connector-java-5.1.40-bin.jar /usr/local/sqoop/lib
```

- 可以將 ratings.txt 資料灌入 MySQL 了，執行 `sqoop import --bindir /usr/local/sqoop/lib --connect jdbc:mysql://localhost:3306/yelp?useSSL=false --username hduser --password hduser --query "select * from ratings where 1=1 and \${CONDITIONS} limit 2000" -target-dir /sqoop/ratings -m1`
 - ✓ `--bindir`→指定 import 過程中，用到的 class 檔路徑，若未指定會發生 `ClassNotFoundException`
 - ✓ `useSSL=false`→localhost 指定不使用 SSL 金鑰
 - ✓ `\${CONDITIONS}`→sqoop 匯入時的指定字串，必加

```
hduser@hadoopmaster: ~
hduser@hadoopmaster:~$ sqoop import --connect jdbc:mysql://localhost:3306/yelp?useSSL=false --username hduser --password hduser --query "select * from ratings where 1=1 and \${CONDITIONS} limit 2000" -target-dir /sqoop/ratings -m1
16/12/25 16:42:28 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6
16/12/25 16:42:28 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.
16/12/25 16:42:28 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
16/12/25 16:42:28 INFO tool.CodeGenTool: Beginning code generation
16/12/25 16:42:29 INFO manager.SqlManager: Executing SQL statement: select * from ratings where 1=1 and (1 = 0) limit 2000
16/12/25 16:42:29 INFO manager.SqlManager: Executing SQL statement: select * from ratings where 1=1 and (1 = 0) limit 2000
16/12/25 16:42:29 INFO manager.SqlManager: Executing SQL statement: select * from ratings where 1=1 and (1 = 0) limit 2000
16/12/25 16:42:29 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /usr/local/hadoop
Note: /tmp/sqoop-hduser/compile/1534557e78cb74f4dcea656ae1bb49c3/QueryResult.java uses or overrides a deprecated API.
Note: Recompile with -Xlint:deprecation for details.
16/12/25 16:42:33 INFO orm.CompilationManager: Writing jar file: /tmp/sqoop-hduser/compile/1534557e78cb74f4dcea656ae1bb49c3/QueryResult.jar
```

```
16/12/25 16:42:46 INFO mapreduce.Job: Counters: 20
File System Counters
  FILE: Number of bytes read=18051210
  FILE: Number of bytes written=18496540
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
  HDFS: Number of bytes read=0
  HDFS: Number of bytes written=20882
  HDFS: Number of read operations=4
  HDFS: Number of large read operations=0
  HDFS: Number of write operations=3
Map-Reduce Framework
  Map input records=2000
  Map output records=2000
  Input split bytes=87
  Spilled Records=0
  Failed Shuffles=0
  Merged Map outputs=0
  GC time elapsed (ms)=0
  Total committed heap usage (bytes)=213385216
File Input Format Counters
  Bytes Read=0
File Output Format Counters
  Bytes Written=20882
16/12/25 16:42:46 INFO mapreduce.ImportJobBase: Transferred 20.3926 KB in 9.7872 seconds (2.0836 KB/sec)
16/12/25 16:42:46 INFO mapreduce.ImportJobBase: Retrieved 2000 records.
hduser@hadoopmaster:~$
```


- 匯入成功後，可執行下列指令查看 hdfs 中的資料狀況

- ✓ `hadoop fs -ls /sqoop`
- ✓ `hadoop fs -ls /sqoop/ratings`
- ✓ `hadoop fs -tail /sqoop/ratings/part-m-00000`

```
hduser@hadoopmaster: ~  
hduser@hadoopmaster:~$ hadoop fs -ls /sqoop  
Found 1 items  
drwxr-xr-x  - hduser supergroup          0 2016-12-25 16:42 /sqoop/ratings  
hduser@hadoopmaster:~$ hadoop fs -ls /sqoop/ratings  
Found 2 items  
-rw-r--r--  2 hduser supergroup          0 2016-12-25 16:42 /sqoop/ratings/_SUCCESS  
-rw-r--r--  2 hduser supergroup    20882 2016-12-25 16:42 /sqoop/ratings/part-m-00000  
hduser@hadoopmaster:~$ hadoop fs -tail /sqoop/ratings/part-m-00000  
22,2  
21,71323,1  
21,71324,4  
21,71325,1  
21,71326,4  
21,71327,4  
21,71328,4  
21,71329,2  
21,71330,3  
21,71331,4  
21,71332,4  
21,71333,1
```

- 若之前執行失敗，已於 hdfs 產生 `sqoop/ratings`，再次執行會產 `IOException`，先執行 `hadoop fs -rm -R /sqoop` 刪除，後再執行前兩步

```
hduser@hadoopmaster:~$ hadoop fs -rm -R /sqoop  
16/12/25 16:26:37 INFO fs.TrashPolicyDefault: Namenode trash configuration: Deletion interval = 0 minutes, Emptier interval = 0 minutes.  
Deleted /sqoop
```

4. 資料測試一由 HDFS 讀出資料

- 由 HDFS 讀出資料放置到 MySQL 中，先建立放置處
 - ✓ 進入 MySQL，執行 `mysql -u root -p`
 - ✓ 使用 yelp 資料庫，執行 `use yelp;`
 - ✓ 建立放置 table，執行 `create table ratings_tmp(userid integer,itemid integer,ratings integer);`
 - ✓ 離開，輸入 `quit`

```
hduser@hadoopmaster:~$ mysql -u root -p
Enter password:
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 18
Server version: 5.7.16-0ubuntu0.16.04.1 (Ubuntu)

Copyright (c) 2000, 2016, 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 yelp;
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> create table ratings_tmp(userid integer,itemid integer,ratings integer);
Query OK, 0 rows affected (0.04 sec)

mysql> quit
Bye
hduser@hadoopmaster:~$
```

- 接下來執行 `sqoop export --connect jdbc:mysql://localhost/yelp?useSSL=false --username hduser -password hduser --table ratings_tmp -m1 --bindir /usr/local/sqoop/lib --export-dir /sqoop/ratings --input-fields-terminated-by ','`

```
hduser@hadoopmaster:~$ sqoop export --connect jdbc:mysql://localhost/yelp?useSSL=false --username hduser --password hduser --table ratings_tmp -m1 --bindir /usr/local/sqoop/lib --export-dir /sqoop/ratings --input-fields-terminated-by ','
16/12/25 17:30:53 INFO sqoop.Sqoop: Running Sqoop version: 1.4.6
16/12/25 17:30:53 WARN tool.BaseSqoopTool: Setting your password on the command-line is insecure. Consider using -P instead.
16/12/25 17:30:54 INFO manager.MySQLManager: Preparing to use a MySQL streaming resultset.
16/12/25 17:30:54 INFO tool.CodeGenTool: Beginning code generation
16/12/25 17:30:55 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `ratings_tmp` AS t LIMIT 1
16/12/25 17:30:55 INFO manager.SqlManager: Executing SQL statement: SELECT t.* FROM `ratings_tmp` AS t LIMIT 1
16/12/25 17:30:55 INFO orm.CompilationManager: HADOOP_MAPRED_HOME is /usr/local/hadoop

16/12/25 17:31:05 INFO mapreduce.Job: Counters: 20
File System Counters
  FILE: Number of bytes read=18055072
  FILE: Number of bytes written=18500436
  FILE: Number of read operations=0
  FILE: Number of large read operations=0
  FILE: Number of write operations=0
  HDFS: Number of bytes read=20888
  HDFS: Number of bytes written=0
  HDFS: Number of read operations=12
  HDFS: Number of large read operations=0
  HDFS: Number of write operations=0
Map-Reduce Framework
  Map input records=2000
  Map output records=2000
  Input split bytes=132
  Spilled Records=0
  Failed Shuffles=0
  Merged Map outputs=0
  GC time elapsed (ms)=13
  Total committed heap usage (bytes)=223870976
File Input Format Counters
  Bytes Read=0
File Output Format Counters
  Bytes Written=0
16/12/25 17:31:05 INFO mapreduce.ExportJobBase: Transferred 20.3984 KB in 4.5784 seconds (4.4554 KB/sec)
16/12/25 17:31:05 INFO mapreduce.ExportJobBase: Exported 2000 records.
hduser@hadoopmaster:~$
```

- 執行完成可以進入 MySQL 查看

```
mysql> use yelp;
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 ratings_tmp limit 10;
+-----+-----+-----+
| userid | itemid | ratings |
+-----+-----+-----+
| 0      | 0      | 4      |
| 0      | 1      | 5      |
| 0      | 7495   | 3      |
| 0      | 7496   | 5      |
| 0      | 7497   | 5      |
| 0      | 7498   | 4      |
| 0      | 23688  | 3      |
| 0      | 23689  | 2      |
| 0      | 23690  | 3      |
| 0      | 23691  | 5      |
+-----+-----+-----+
10 rows in set (0.00 sec)
```