A brief introduction to MySQL

Ch. Morisset
IA-UNAM

History

MySQL is a relational database management system (RDBMS). It was created in 1995.

It is the second widely used (after SQLite, which is included in any Android and iOS device...)

SQL stands for “Structured Query Language”.

The “My” comes from the name of the co-founder daughter, My.

In 2008, Sun Microsystems bought MySQL for $1 billion. In 2009 Oracle entered into an agreement to purchase Sun and to continue to enhance MySQL.

In January 2009, prior to Oracle's acquisition of Sun and MySQL, Monty Widenius started a GPL-only fork, MariaDB.

Usage

- MySQL is part of the LAMP package (Linux, Apache, MySQL, PHP) which is widely used for web server creation. WAMP and MAMP are for Windows and Mac.
- It is used by a lot of big companies (Wikipedia, Google, Youtube, Reuters, ...)
- It supports various database engines (the underlying software component that a DBMS uses to create, read, update and delete (CRUD) data from a database), between others MyISAM (default until v5.5) and InnoDB (default since v5.6).

SQL

- The request language is SQL. Other RDBMS using SQL are Oracle, PostgreSQL, SQLite, Microsoft SQL server, Microsoft Access, between many others.
- Some small differences in syntax can exist between the RDBMS.
- ADQL is the Astronomical Database Query Language, used in Virtual Observatory, see: http://www.ivoa.net/documents/latest/ADQL.html

Client and server

- Most of the users will only need to have a **client** access to a database, but not to manage themselves a database.
- It's similar to have an access to the web using a browser. Everybody does it, it's easy. Another story is having a server managing its own web pages.
- Almost every Linux distribution comes with MySQL ot MariaDB installed.
- OSX: have to install from MySQL web page (http://dev.mysql.com/downloads/mysql/). Need to register to Oracle. Notice: for OSX 10.7, use the 10.6 version.

Connect to a db

To connect to the 3MdB database (as an example):

mysql -h 132.248.1.102 3MdB -u OVN_user -p

Enter password: **********

mysql>help

Basic commands

```
mysql> show tables; # list the tables available in the current base
+---------------+
| Tables_in_3MdB |
+---------------+
| abion          |
| lines          |
| tab            |
| teion          |
| temis          |
+---------------+
5 rows in set (0.00 sec)
```

Basic commands

```
mysql> describe `lines`; # `` are need as lines is a MySQL keyword
+--------+-------------+------+-----+---------+----------------+
| Field  | Type        | Null | Key | Default | Extra          |
+--------+-------------+------+-----+---------+----------------+
| Nl     | bigint(20)  | NO   | PRI | NULL    | auto_increment |
| label  | varchar(15) | YES  |     | NULL    |                |
| id     | varchar(20) | YES  |     | NULL    |                |
| lambda | double      | YES  |     | NULL    |                |
| name   | varchar(40) | NO   |     | NULL    |                |
| used   | int(2)      | YES  |     | 1       |                |
+--------+-------------+------+-----+---------+----------------+
6 rows in set (0.00 sec)
```

Basic commands

SELECT is the command to obtain a result from a query.

```
mysql> select count(*) from tab; # number of elements
+----------+
| count(*) |
+----------+
|   665491 |
+----------+

mysql> select min(N) from tab; # arithmetic operations are available
+---------+
| min(N)  |
+---------+
| 1743958 |
+---------+
```

Aliases and limit

```
mysql> SELECT min(N) AS MIN, max(N) AS MAX FROM tab;
+---------+---------+
| MIN     | MAX     |
+---------+---------+
| 1743958 | 2500069 |
+---------+---------+

mysql> SELECT id, lambda, name FROM `lines` LIMIT 10;
+------+--------+------------------+
| id   | lambda | name             |
+------+--------+------------------+
| Bac  |   3646 | BalmHead         |
| cout |   3646 | OutwardBalmPeak  |
| cref |   3646 | ReflectedBalmPeak|
| H  1 |   4861 | H I 4861         |
| TOTL |   4861 | H I 4861         |
| H  1 |   6563 | H I 6563         |
| H  1 |   4340 | H I 4340         |
| H  1 |   4102 | H I 4102         |
| H  1 |   3970 | H I 3970         |
| H  1 |   3835 | H I 3835         |
+------+--------+------------------+
```

Where

```
mysql> SELECT name FROM `lines` WHERE lambda > 5000 AND lambda < 6000;
+-----------------+
| name            |
+-----------------+
| He I 5876       |
| He I 5876 Bcase |
| [N I] 5198      |
| [N I] 5200      |
| [N II] 5755     |
| N II 5755 rec   |
| N II 5679 totl  |
| [O I] 5577      |
| [O III] 5007    |
| [Cl III] 5538   |
| [Cl III] 5518   |
| [Ar III] 5192   |
| [Fe III] 5271   |
| [Fe VI] 5177    |
| [Fe VII] 5721   |
| [Fe VII] 5277   |
+-----------------+
```

Where and order

```
mysql> SELECT count(*) from `lines` WHERE lambda > 5000 AND lambda < 6000;
+----------+
| count(*) |
+----------+
|       16 |
+----------+

mysql> SELECT name from `lines` WHERE lambda > 5000 AND lambda < 6000 ORDER BY lambda;
+-----------------+
| name            |
+-----------------+
| [O III] 5007    |
| [Fe VI] 5177    |
| [Ar III] 5192   |
| [N I] 5198      |
| [N I] 5200      |
| [Fe III] 5271   |
| [Fe VII] 5277   |
| [Cl III] 5518   |
| [Cl III] 5538   |
| [O I] 5577      |
| N II 5679 totl  |
| [Fe VII] 5721   |
| N II 5755 rec   |
| [N II] 5755     |
| He I 5876 Bcase |
| He I 5876       |
+-----------------+
```

Count and group

```
mysql> SELECT ref, count(*) AS number FROM tab GROUP BY
ref ORDER by number;
+--------------+--------+
| ref          | number |
+--------------+--------+
| HII_CHIm     |   7854 |
| CALIFA       |  28080 |
| DIG_HR       |  41327 |
| PNe_2014_c13 |  45280 |
| PNe_2014     | 542950 |
+--------------+--------+
```

Join tables

In some databases, the data are disseminated in multiple tables.

Keys are used to associate entries from one table with entries from another table.

Ex: N in `tab` and N in `teion` are referring to the same model.

Join tables

```
mysql> SELECT
    O__3__5007A/H__1__4861A as O3,
    N__2__6584A/H__1__4861A as N2,
    T_OXYGEN_VOL_2 as T_O3
FROM
    tab, teion
WHERE
    tab.N = teion.N
      AND
    tab.ref = 'PNe_2014' # need the tab.ref, as ref is also in teion
LIMIT
    10;
```

O3	N2	T_O3
0.00075005260466636 05	4.066774515375886	7329.55042055
13.943986612997021	0.10823972778663463	13623.7705353
10.07287308055993	0.7346007859302494	14425.8946549
9.76004806591307	0.1620998406576782	13404.7666047
4.0190422636700 53	0.090125301557351 14	8062.68788933
0.6178622832763104	4.310980124863453	8401.1907834
0.00002081717553084 7115	0.8547542220355914	5531.31483827
6.498702975397857	0.171329331937283 84	8608.60687521
3.43898063798645 74	11.445522984383846	11931.6793072
0.501324852257867	12.311852742730316	10782.7816288

Format the output

Using ROUND or FORMAT functions:

```
mysql> SELECT
    ROUND(O__3__5007A/H__1__4861A, 4) as O3,
    ROUND(N__2__6584A/H__1__4861A, 4) as N2,
    ROUND(T_OXYGEN_VOL_2,1) as T_O3
FROM
    tab,teion
WHERE
    tab.N=teion.N
      AND
    tab.ref = 'PNe_2014'
LIMIT
    10;
```

O3	N2	T_O3
0.0008	4.0668	7329.6
13.9440	0.1082	13623.8
10.0729	0.7346	14425.9
9.7600	0.1621	13404.8
4.0190	0.0901	8062.7
0.6179	4.3110	8401.2
0.0000	0.8548	5531.3
6.4987	0.1713	8608.6
3.4390	11.4455	11931.7
0.5013	12.3119	10782.8

Running requests

Put the request into a file (e.g. req1.sql) containing for example:

```
SELECT ROUND(O__3__5007A/H__1__4861A, 4) as O3,
ROUND(N__2__6584A/H__1__4861A, 4) as N2, ROUND(T_OXYGEN_VOL_2,1) as
T_O3 FROM tab,teion WHERE tab.N=teion.N AND tab.ref = 'HII_CHIm'
```

Notice that there is no LIMIT anymore (a lot of results). Run it from a terminal:

```
mysql -h 132.248.1.102 3MdB -u OVN_user -p < req1.sql >
req1.res
```

The result is obtained in less than a second and store in req1.res. It contains 7855 lines, easy to read from python or IDL, starting like:

```
O3      N2      T_O3
0.4108  1.1447  5409.4
2.5814  0.1484  18626.0
0.0021  0.9481  3460.5
4.1895  0.0581  11935.3
0.0251  0.1854  5271.2
4.7034  0.0064  16001.7
2.6313  0.0036  18857.0
1.6673  0.6845  14112.8
1.6881  0.1417  6489.9
6.5569  0.3675  13707.0
...
```

Special functions

- The reference manual is 3500 pages big!...
- I'll here describe a few number of functions, have a look at the online doc for more.

Strings

- Substring, reverse, ltrim, etc...
- like and %:

```
mysql> SELECT count(*), ref FROM tab WHERE
ref LIKE 'PNe_2014%' GROUP BY ref;
+----------+---------------+
| count(*) | ref           |
+----------+---------------+
|   542950 | PNe_2014      |
|    45280 | PNe_2014_c13  |
+----------+---------------+
```

http://www.tutorialspoint.com/mysql/mysql-string-functions.htm

Numeric functions

- Log10, avg, sqrt, pow, abs, sin, …
- http://www.tutorialspoint.com/mysql/mysql-numeric-functions.htm

User defined variables

We define 2 variables (@maxo and @mino):

```
mysql> select @maxo:=log10(max(O__3__5007A/H__1__4861A)),
@mino:=log10(min(O__3__5007A/H__1__4861A)) from tab where
ref = 'HII_CHIm';
```

We can use the variables in any request:

```
mysql> select count(*) from tab where ref='HII_CHIm' and
log10(O__3__5007A/H__1__4861A) > (@mino+1) and
log10(O__3__5007A/H__1__4861A) < (@maxo-1);
+----------+
| count(*) |
+----------+
|     3509 |
+----------+
```

Optimization

- EXPLAIN before SELECT.

```
mysql> EXPLAIN SELECT ROUND(O__3__5007A/H__1__4861A,
4) as O3, ROUND(N__2__6584A/H__1__4861A, 4) as N2,
ROUND(T_OXYGEN_VOL_2,1) as T_O3 FROM tab,teion WHERE
tab.N=teion.N AND tab.ref = 'HII_CHIm' LIMIT 10;

+----+-------------+-------+--------+---------------+---------+---------+------------+------+-------------+
| id | select_type | table | type   | possible_keys | key     | key_len | ref        | rows | Extra       |
+----+-------------+-------+--------+---------------+---------+---------+------------+------+-------------+
|  1 | SIMPLE      | tab   | ref    | PRIMARY,ref   | ref     | 122     | const      | 7802 | Using where |
|  1 | SIMPLE      | teion | eq_ref | PRIMARY       | PRIMARY | 8       | 3MdB.tab.N |    1 |             |
+----+-------------+-------+--------+---------------+---------+---------+------------+------+-------------+
```

Manage a database

- Have to install and run the server.
- PhpMyAdmin
- Easy import data from a file
- Next chapter...

Links

- http://www.mysqltutorial.org/
- http://www.astro.rug.nl/~belikov/VO2016/
- http://cdn.oreillystatic.com/en/assets/1/event/2/Top%2020%20DB%20Design%20Tips%20Every%20Architect%20Needs%20to%20Know%20%20Presentation.pdf
- http://cas.sdss.org/dr7/en/help/howto/search/