

mySQL class 5/9/05

Linux stuff

overview of file system

"everything is a file"--devices abstracted as files also

some basic linux commands

which

man

apropos

id

bash built-in features

history

command completion

file name completion

A little database history

extracted from: <http://math.hws.edu/vaughn/cpsc/343/2003/history.html>

<http://wwwdb.web.cern.ch/wwwdb/aboutdbs/history/industry.html>

extended, edited and modified by Steve Glasser

Back in the day:

Each program stored data in it's own files.

Problems:

each data set required it's own program to access data

redundant data was stored, leading to errors, etc.

First databases--hierarchical databases

all related data had parent-child relationship

(this is the same as your file system)

relationships were created using pointers to physical locations on disk

any guess to what problems this created?

1970-72: E.F. Codd proposed relational model for databases in a landmark paper on how to think about databases. He disconnects the schema (logical organization) of a database from the physical storage methods. This system has been standard ever since.

1970's: Two main prototypes for relational systems were developed during 1974-77.

Ingres: Developed at UCB. This ultimately led to Ingres Corp., Sybase, MS SQL Server, Britton-Lee, Wang's PACE. This system used QUEL as query language.

System R: Developed at IBM San Jose and led to IBM's SQL/DS & DB2, Oracle, others. The language created by the System R group, SQL (Structured Query Language) has become the industry standard for relational databases and is now an ISO standard.

The term Relational Database Management System (RDBMS) is coined during this period.

1976: P. Chen proposed the Entity-Relationship (ER) model for database design giving yet another important insight into conceptual data models. Such higher level modeling allows the designer to concentrate on the use of data instead of logical table structure.

Mid-1980's: Development of the IBM PC gives rise to many DB companies and products such as RIM, RBASE 5000, PARADOX, OS/2 Database Manager, Dbase III, IV (later Foxbase, even later Visual FoxPro), Watcom SQL.

Early 1990's: An industry shakeout begins with fewer surviving companies offering increasingly complex products at higher prices. Much development during this period centers on client tools for application development such as PowerBuilder (Sybase), Oracle Developer, VB (Microsoft), etc. Client-server model for computing becomes the norm for future business decisions. Development of personal productivity tools such as Excel/Access (MS) and ODBC. This also marks the beginning of Object Database Management Systems (ODBMS) prototypes.

Mid-1990's: Kaboom! The usable Internet/WWW appears. A mad scramble ensues to allow remote access to computer systems with legacy data.

Late-1990's through current: The large investment in Internet companies fuels market boom for Web/Internet/DB connectors. Active Server Pages, Front Page, Java Servlets, JDBC, Enterprise Java Beans, ColdFusion, Dream Weaver, Oracle Developer 2000, etc are examples of such offerings. Open source solution come online with widespread use of gcc, cgi, Apache, MySQL, etc. Online Transaction processing (OLTP) and online analytic processing (OLAP) comes of age with many merchants using point-of-sale (POS) technology on a daily basis.

Future trends: Huge (terabyte) systems are appearing and will require novel means of handling and analyzing data.

Large science databases such as genome project, geological, national security, and space exploration data.

Clickstream analysis.

Data mining.

Data warehousing,
data marts.

More of this in the future without a doubt. Smart/personalized shopping using purchase history, time of day, etc.

Successors to SQL (and perhaps RDBMS) will be emerging in the future. Most attempts to standardize SQL successors has not been successful. SQL92, SQL2, SQL3 are still underpowered and more extensions are hard to agree upon.

Most likely this will be overtaken by XML and other emerging techniques. XML with Java for databases is the current poster child of the "next great thing". Check in tomorrow to see what else is news.

Probably there will be a continuing shakeout in the RDBMS market. Linux with Apache supporting MySQL (or even Oracle) on relatively cheap hardware is a major threat to high cost legacy systems of Oracle and DB2. Object Oriented Everything, including databases, seems to be always on the verge to sweeping everything before it. Object Database Management Group (ODMG) standards are proposed and accepted and maybe something comes from that.

Ethical/security/use issues tend to be diminished at times but always come back. Should you be able to consult a database of the medical records/genetic makeup of a prospective employee? Should you be able to

screen a prospective partner/lover for genetic diseases? Should amazon.com keep track of your book purchasing? Should there be a national database of convicted sex offenders/violent criminals/drug traffickers? Who is allowed to do Web tracking? How many times in the last six months did you visit a particular sex chat room/porn site/political satire site? Who should be able to keep or view such data? Who makes these decisions?

mysql stuff

what is DDL?

Used to create new db objects and modify existing objects
such as user accounts, tables

What is DML?

Used to change the data values in the database; require explicit commit
(in transactional db)

What's a primary key?

A field or combination of fields that uniquely identifies each row
in each table.

Must not be null.

Usually auto-generated, auto-incremented.

Short vs long commands:

```
mysql(1)                mysql database                mysql(1)
```

NAME

mysql - text-based client for mysqld, a SQL-based relational
database daemon

USAGE

```
mysql [OPTIONS] [Database]
```

SYNOPSIS

```
mysql [-B|--batch] [-#|--debug= logfile] [-T|--debug-info]
[-e|--exec= command] [-f|--force] [-?|--help]
[-h|--host= hostname] [-n|--unbuffered] [-p[pwd]] [--password=
[pwd]] [-P|--port= pnum] [-q|--quick] [-r|--raw]
[-s|--silent] [-S|--socket= snum] [-u|--user= uname] [-v|--
verbose] [-V|--version] [-w|--wait]
```

DESCRIPTION

The mysql program provides a curses-based interface to the SQL-
based database server daemon, mysqld(1). Full

fuller documentation, refer to the HTML documents installed with
the package.

<snip><snip><snip><snip><snip><snip><snip><snip><snip><snip>

Note that almost all options are available in long and short form.

Some options take values, eg password, user, etc.

Note also there are some quirks-the long options which take input use the
'=' sign, the short options do not.

Some options have default values--

user: login name (ODBC for windows)

Options on the command line override options read from a file.

/etc/my.cnf

In this file you can use any of the long options that the program supports.

Options are written without the leading dashes.

Options are normally grouped by the name of the program to which the options apply.

Here is an example from an actual my.cnf file:

```
# The following options will be passed to all MySQL clients
```

```
[client]
```

```
#password      = your_password
```

```
port           = 3306
```

```
socket         = /var/run/mysql/mysql.sock
```

```
# Here follows entries for some specific programs
```

```
# The MySQL server
```

```
[mysqld]
```

```
port           = 3306
```

```
socket         = /var/run/mysql/mysql.sock
```

```
skip-locking
```

```
key_buffer    = 16M
```

```
max_allowed_packet = 1M
```

```
table_cache   = 64
```

```
sort_buffer_size = 512K
```

```
net_buffer_length = 8K
```

```
myisam_sort_buffer_size = 8M
```

There are three standard locations for an options file:

/etc (for global options)

the_data_dir (for server-specific options)

~/ (for per-user options)

You can force a non-standard options file to be used with

```
--defaults-file=
```

I note that per the MySQL book, any program can use this option. However the man pages show it only for mysqld_safe.

Selecting a database

from outside the program:

```
eg MySQL -u foo -p name_of_db
```

from inside the program-2 ways:

```
mysql> use name_of_db
```

```
mysql> select * from name_of_db.name_of_table
```

Using the MySQL client.

Two sources of input: script file or std in.

Interactive mode:

Write the command.

Terminate the command.

Terminators:

```
;
```

```
\g
```

```
\G
```

Try each and see what they do!

Dumb MySQL tricks:

Command line history and editing-very similar to the bash shell.

Commands can span lines.

More than one command on the same line (use terminators, eh?)

Use script file:

Inside: MySQL> source script_file;

Outside: MySQL [-u username -p] db_name<my_commands.sql

controlling MySQL output

2 modes:

 interactive: like a table of ascii art.

 Batch: tab-separated values.

Override batch output with command-line options:

(these do not work from within mysql, AFAIK)

 -B or -batch

 -T or -table any guess what these do?

 -H or -html

 -X or -xml try both long and short options; YMMV

MySQL client commands; or what you can type at the prompt:

 sql-- select, insert, etc.

 internal commands-- quit, source, status, help, etc.

 these have a short form: eg \q \. \s \h

what works when (internal commands): (demo this)

 long: at first MySQL prompt

 short: at any prompt w/in MySQL

 --safe-updates option

 update and delete: only with a 'where' clause or 'limit' clause

 select (single table) limited to 1000 rows unless 'where' clause

 select (multi-table) limited to 1,000,000 rows

Handy utility programs

1) MySQLimport

2) mysqldump

3) MySQLcheck

4) myisamchk

MySQLimport

 loads data into tables

 ~\$ MySQLimport [options] db_name input_file

 input_file_name=>table_name

 table must exist (not for moving a db to another host)

 read the options from the man pages

mysqldump

 dumps content of db into text files, along with instructions to

 rebuild the database on another server

 what does mysqldump not include? What option fixes this problem?

Undump

 use MySQL in batch mode, eh?

 Who can give the syntax?

mysqlcheck

client program-issues sql statements to the server
can optimize, repair myisam tables
safe to run with server up

myisamchk

stop the server!
Acts directly on the database files in /var/whatever.
Therefore you must be logged on to the machine where the files are.
User must have read/write on these files.

mySQL connectors flyover

install it on the client

odbc

mySQL connector/ODBC
mySQL's version of odbc connector
there are numerous others, both commercial and freeware
available for most platforms

What the flock is ODBC?

(pronounced as separate letters) Short for Open DataBase Connectivity, a standard database access method developed by the SQL Access group in 1992. The goal of ODBC is to make it possible to access any data from any application, regardless of which database management system (DBMS) is handling the data. ODBC manages this by inserting a middle layer, called a database driver, between an application and the DBMS. The purpose of this layer is to translate the application's data queries into commands that the DBMS understands. For this to work, both the application and the DBMS must be ODBC-compliant -- that is, the application must be capable of issuing ODBC commands and the DBMS must be capable of responding to them.

When programming to interact with ODBC you only need to talk the ODBC language (a combination of ODBC API function calls and the SQL language). The ODBC Manager will figure out how to contend with the type of database you are targeting. Regardless of the database type you are using, all of your calls will be to the ODBC API. All that you need to do is have installed an ODBC driver that is specific to the type of database you will be using.

jdbc

same drill, almost