MySQL: The Whys, Whats, and Watch-outs

Marc Nozell

GNHLUG
19-Jun-2008

- Marc Nozell (marc@nozell.com)
- Previously worked on the technical side of the MySQL AB / HP partner relationship
  - Presented at MySQL User Conferences, MySQL 5.0 trained
  - Technical White papers & benchmarks on Linux/ProLiant servers — http://h20338.www2.hp.com/ActiveAnswers/cache/81882-0-0-0-121.html (http://tinyurl.com/4dsqx5)
- Now doing Oracle Application benchmarking...

- What is MySQL
- SQL Examples
- MySQL AB, the company
- Advanced MySQL
- More resources

- http://www.mysql.com/
- Pronounced My-S-Q-L
- Multithreaded, multiuser, DBMS, ACID compliant
- MySQL AB copyright holder (acquired by Sun)
- The "M” in the LAMP stack

- Ask questions in a standard way:
  SELECT price,model FROM cars WHERE color = 'blue' ORDER BY status ASC LIMIT 0,10
- Less disk space
- data consistency – transactions, data types, etc
- centralized data store, let someone else backup
- Let database developers optimize code
- Scaling, Distributed, High Availablity, etc

- You like writing code to parse data, store/read data, multiple access/update, security, etc (hey that looks like a database)
  - Consider sqlite3
- Unstructured, large data (ie: photos, possible however)
- Stand alone application (could use MySQL Embedded)
- Doesn't run on your platform

- Yes, your platform is supported
  - Windows; Windows x64; Linux (non RPM packages); Linux (non RPM, Intel C/C++ compiled, glibc-2.3); Red Hat Enterprise Linux 3 RPM (x86); Red Hat Enterprise Linux 3 RPM (AMD64 / Intel EM64T); Red Hat Enterprise Linux 3 RPM (Intel IA64); Red Hat Enterprise Linux 4 RPM (x86); Red Hat Enterprise Linux 4 RPM (AMD64 / Intel EM64T); Red Hat Enterprise Linux 4 RPM (Intel IA64); Red Hat Enterprise Linux 5 RPM (x86); Red Hat Enterprise Linux 5 RPM (AMD64 / Intel EM64T); SuSE Linux Enterprise Server 9 RPM (x86); SuSE Linux Enterprise Server 9 RPM (AMD64 / Intel EM64T); SuSE Linux Enterprise Server 9 RPM (Intel IA64); SuSE Linux Enterprise Server 10 RPM (x86); SuSE Linux Enterprise Server 10 RPM (AMD64 / Intel EM64T); SuSE Linux Enterprise Server 10 RPM (Intel IA64); Ubuntu 6.06 LTS (Dapper Drake); Linux x86 generic RPM (statically linked against glibc 2.2.5); Linux x86 generic RPM (dynamically linked); Linux AMD64 / Intel EM64T generic RPM; Linux Intel IA64 generic RPM; Solaris (pkgadd packages); Solaris (TAR packages); FreeBSD (TAR packages); Mac OS X (package format); Mac OS X (TAR packages); HP-UX (depot packages); HP-UX (TAR packages); IBM AIX; IBM i5/OS (SAVF packages); IBM i5/OS (TAR packages); QNX; Novell NetWare; SCO OpenServer 6; Source;

- mysql
- mysqladmin
- mysqldump
- mysqlimport
- mysqlreport
- mysqlcheck
- Mysqlbug

- SQL – Structured Query Language – a standard but all vendors extend it, especially datatypes and stored procedures
- ACID – Atomic, Consistent, Isolated, Durable
- Transactions – all parts either complete successfully, or none are
  - BEGIN
  - UPDATE ...
  - INSERT ...
  - COMMIT

- Schema – table format
- Normalized – abstracted to reduce duplicated data.
  - business address in large company with multiple employees
- Foreign keys – references to other tables
  - don't want to add an employee to a business address that doesn't exist.

- Create database
- Create table
- Alter table
- Insert
- Select
- Update
- Delete
- Drop

- CREATE DATABASE mystore;
- USE mystore;
- SHOW DATABASES;

- CREATE TABLE books (
  id INT,
  title VARCHAR(50),
  author VARCHAR(60)
  );
- CREATE TABLE publishers (
  id int auto_increment primary key,
  name varchar(30));
- SHOW TABLES;
- SHOW COLUMNS FROM books;

- ALTER TABLE books
  ADD COLUMN pid INT;
- ALTER TABLE books \
  ADD genre ENUM('scifi', 'romance', 'poetry',
  'reference'),
  ADD isbn VARCHAR(30));

- INSERT INTO books (title, author, isbn, genre)
  \
  VALUES (”MySQL book”, ”Fred Smith”, ”0-123-345-789”, 'poetry');
- Id will automatically be incremented, publisher will not have a value.
- INSERT INTO publishers (name) VALUES (” O'Reilly”);

- UPDATE books SET genre = 'reference' WHERE title = ”MySQL book”;
- UPDATE books SET pid = 1 WHERE title = 'MySQL book';

- SELECT * FROM books;
- SELECT COUNT(*) FROM books;
- SELECT title, author FROM books;
- SELECT title, author FROM books LIMIT 0,10;
- SELECT title, author FROM books WHERE publisher = 1;

- SELECT title, isbn a FROM books WHERE author LIKE '%Smith%';
- SELECT title FROM books b, publishers p WHERE b.publisher = p.id AND p.name = ” O'Reilly” GROUP BY p.name DESC ORDER BY b.title;
- EXPLAIN SELECT title, isbn a FROM books WHERE author LIKE '%Smith%';

- Trade space for time (and some time on insert/update/delete)
- CREATE INDEX myidx ON books (author(3));
- SHOW KEYS ON books;

- DELETE FROM books; ## BAD
- DELETE FROM books WHERE id = 1;
- DELETE FROM books WHERE title NOT LIKE '%MySQL%';

- DROP TABLE books;
- DROP DATABASE mystore;

- Limited SQL profiling – enable slow query log
- May buffers, queues to change, monitor – See enterprise monitor
- Show status; show process_list, etc
- Provide hints to the SQL optimizer in comments. Select ... /*! SQL_NO_CACHE */ WHERE ...
- Fast cpus, more cores (to a point), fast disks, splinter your data...

- MySQL Administrator
- MySQL Query Browser
- MySQL Migration toolkit
- mytop

- Application Programming Interface
- All major languages – C, C++, perl, python, Ruby, .NET, Java, etc
- Also JDBC and ODBC interfaces (note licenses)

- use DBI;
  $handle = DBI->connect("dbi:mysql:database=mysql;host=db.host.name;
  user=muser;password=muserpwd")
  or die "Couldn't connect to database: $DBI::errstr\n";
- $sql = "SELECT firstname, lastname FROM usertable where lastname like \"foo\"
  and state = \"New %\"";
- $statement = $db_handle->prepare($sql) or die "Couldn't prepare query '$sql':
  $DBI::errstr\n";
- $statement->execute() or die "Couldn't execute query '$sql': $DBI::errstr\n";
- while ($row_ref = $statement->fetchrow_hashref()) {
- print "firstname: $row_ref->{firstname}\tlastname: $row_ref->{lastname}";
- }
- $db_handle->disconnect();

- Monty Widenius started as Swedish consulting
- MySQL released internally 1995
- MySQL Windows 1/1998
- MySQL 3.23 1/01-- still in production
- Marten Mickos, CEO
- MySQL V4.0 3/03 – unions
- MySQL V4.1 10/04 – R-Trees, B-Trees, subqueries, prep statements
- MySQL V5.0 10/05 – cursors, stored procs, triggers, views, XA Transactions
- MySQL V5.1 beta 11/05 – event scheduler, partitioning, plugin API, row-based replication, server log table
- Sun announced intent to aquire 1/08

- Enterprise Database Services and Support
  - Enterprise Server
    - fully patched, monthly, built by MySQL.com
  - Enterprise Monitor
    - Web based monitoring of all MySQL servers, proactive notification, lots of advisors (performance, admin, security, schema, memory util, etc), etc
  - 24x7 production support – 4 levels
  - MySQL migration toolkit – convert schema, move data to MySQL
- MySQL Certification training, testing

- MySQL DB, libraries, clients are GPLV2
- Note libraries are not LGPL
- Enterprise tools are not GPL
- Can negotiate one-off non-GPL licenses

- Stored procedures
  - put SQL statements in DB
  - tries to be standard, but plan on re-writing if migrating
- Triggers
  - do something when SQL is executed, eg: transfer money to another account when a threshold is reached

- Views
  - consistent view of results from changes, limits based on security
- Cursors
  - iterate through results from select

- Storage Engines
- Master Slave Replication / DRDB
- Approach to scaling

- How data is written to disk
- Each have own guarentees
- Transparent to the user/application (ALTER TABLE cars ENGINE=InnoDB)
- MySQL comes with many, some from ISVs for specialized uses (Data Warehousing, OLTP, etc)

- MyISAM – default, no transactions, fast, fulltext search
- InnoDB – transactions, now owned by Oracle but still GPL, faster recovery time, similar to Oracle's use of tablespaces
- MEMORY – In memory, no transactions

- BDB – transactions, mostly replaced by InnoDB
- FEDERATED – stores on remote database
- ARCHIVE – no indexes, compressed, write once, handy for SOX compliance
- CSV – text files(?)
- NDB Cluster (very specialized)
- More are in development – FALCON, etc
- Use either MyISAM or InnoDB

- Master DB for inserts/update/delete; Slaves for read
- Any number of slave servers, applies stream of updates from master
- Row based replication, SQL based in dev
- Each server has unique id to prevent loops
- Requires lots of storage – each slave has copy of master's data

- Scales well for read-mostly applications (flickr, youtube, etc)
- mysql> show slave status
  mysql> show master status
- For best results, needs application support
- Useful for scaling, backup, data mining, remote offices, SOX

- For high availablity
- Master/Master with slaves, etc

- MySQL V5.0 maxes at 4cores
- MySQL V5.1 supposed to address
- Use slaves
- Use memcache (!)
- Splinter your data; apps know where to query

- mysqldump > backup.sql
  - works with all storage engines
  - slow
  - dumps as SQL statements
  - in-consistent database
  - many options
- InnoDB – purchase mysqlhotbackup from InnoDB directly
- Use replication and backup a slave

- Many, many ISVs
- backup, data migration, db interop, management, special purpose (frequently closed source) storage engines (data warehousing, OLTP) etc

- MySQL docs http://dev.mysql.com/doc/refman/ (can leave comments and view other versions)
- Aggregated Developers & users blogs: http://planetmysql.org/
- CTO, Brian ”Krow” Aker: http://krow.net/
- MySQL users conference has videos, etc http://www.mysqluc.com/

- Boston MySQL Meetup http://mysql.meetup.com/137/
- Jeremy Zawodny blog http://jeremy.zawodny.com/blog/
- Jay Pipes blog http://jpipes.com/

- The Practical SQL Handbook (Addison Wesley) — Bowman, Emerson, Darnovsky
- MySQL in a Nutshell (O'Reilly) — Russell J.T. Dyer
- MySQL Stored Procedures (O'Reilly) — Guy Harrison w/ Steven Feuersten

- High Performance MySQL: Optimization, Backups, Repliction and Load Balancing (O'Reilly) — Jeremey D. Zawodny & Derek J. Balling
- MySQL 5.0 Certification Study Guide (MySQL Press) — Paul Dubois, Stefan Hinz, Carsten Pedersen

- http://www.mysqluc.com/
- O'Reilly show
- Most of MySQL AB devs, executives, etc attend
- very high quality ½ day tutorials
- ISV and focused presentations – youtube, flickr, smugmug, lucas films, etc
- Some talks are available on youtube