MySQL and The Architecture of Participation

foss.in 2007
December, 2007

... or learning how to contribute to MySQL

Colin Charles, Community Relations Manager, APAC
colin@mysql.com | http://bytebot.net/blog/

Who am I?

- Community Relations Manager, APAC
  - Distribution Ombudsman
  - Community Engineering
  - Summer of Code
  - Forge Dude
  - Generalised Dolphin Wrangler
- Previously:
  - Fedora Project FESCO and PowerPC hacker
  - OpenOffice.org contributor

What are my aims, today?

- Harness more contributors to MySQL
- Show you what we've done, to help you contribute
- Answer any of your queries, concerns, worries
  - I like questions, and will *attempt* to answer them interactively within these 90 minutes
- Generalise open source contributions, so even if you don't contribute to MySQL, you contribute to the FLOSS world
- ... evangelise MySQL all week :)

What we won't be covering today

- MySQL Enterprise
  - And the whole Enterprise/Community split
  - Community source+binary policy
- MySQL Enterprise Monitor
  - affectionately known as Merlin, this is built with open source tools/libraries, but itself, isn't open source
- MySQL, the company, business model, etc.
- We can however talk about this (and more), over a cup of *chai*

Have you ever...

- Submitted a bug report?
- Extended a Wiki page documenting something
- Hopped on IRC, at #project ?
- Used some open source software?
- Evangelised your favourite software?

You're already part of the Architecture of Participation

(now, pat yourselves on the back)

What is this buzz-phrase?

- Tim O'Reilly coined
- Describes the **nature of systems that are designed for user contribution**
- Surprisingly, not a term to sell more books ;-)
- Read more:
  - http://tinyurl.com/yw3fna

The MySQL back story

- A commercial entity, since its inception
- Open source runs deeply within the veins
  - GPLv2 since early 2000
- Development has been relatively closed
  - Developers, usually all from the community, are immediately hired
  - Code reviews, performed in secrecy
  - Hurdles, including legal issues, for patch submission
  - Commits only from internal developers

A glitch in the sea was detected

- We sought change
- Complement the existing “openness”
  - bugs.mysql.com bugs database
  - mailing lists
  - forums.mysql.com
  - dev.mysql.com developer zone
- Notice how the above are *minimum requirements* for any open source project?

Steps taken... to open up

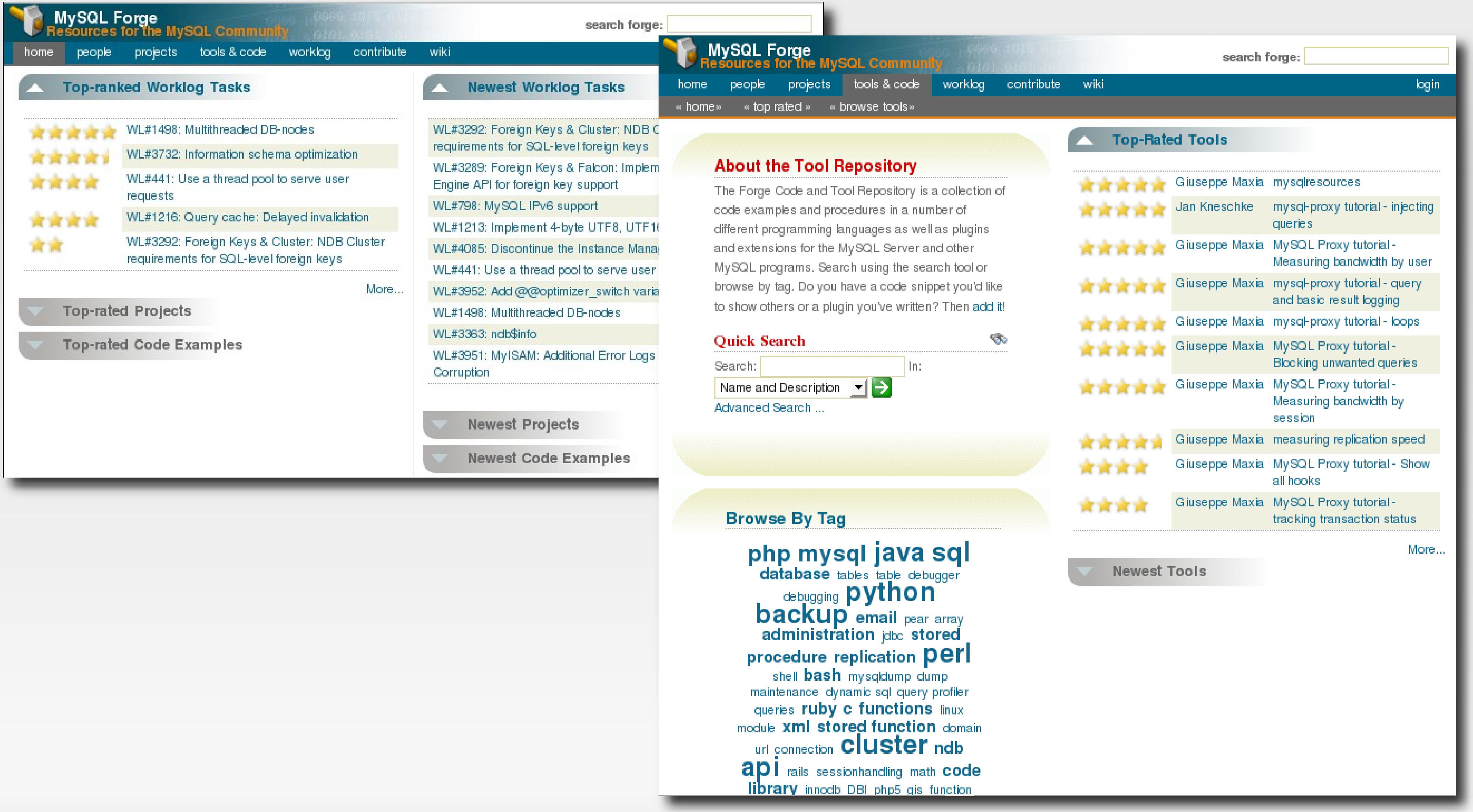
MySQL DevZone

- Developer Zone, is most visible, all over www.mysql.com
- Get your downloads
- Useful documentation (really!)
- Has some substantial article content, getting you using MySQL, more and more
- Problem? Its a mixture of Marketing and Community, so it doesn't *really* serve developers as much as they should be served

MySQL Forge

- http://forge.mysql.com/
- Sourceforge/Freshmeat equivalent, tailored to MySQL only
- Projects: allows listing of projects that make use of MySQL, with full tagging support
- Snippets: allows sharing of SQL, stored procedures, UDFs, etc.
- Lots of UI and functional requirements are required (Forge 2.0 will be out any day)

MySQL Forge 2.0



MySQL Forge
Resources for the MySQL Community

search forge:

Top-ranked Worklog Tasks

★★★★★	WL#1498: Multithreaded DB-nodes
★★★★½	WL#3732: Information schema optimization
★★★★	WL#441: Use a thread pool to serve user requests
★★★★	WL#1216: Query cache: Delayed invalidation
★★	WL#3292: Foreign Keys & Cluster: NDB Cluster requirements for SQL-level foreign keys

More...

Top-rated Projects

Top-rated Code Examples

Newest Worklog Tasks

- WL#3292: Foreign Keys & Cluster: NDB C
  requirements for SQL-level foreign keys
- WL#3289: Foreign Keys & Falcon: Implem
  Engine API for foreign key support
- WL#798: MySQL IPv6 support
- WL#1213: Implement 4-byte UTF8, UTF1
- WL#4085: Discontinue the Instance Mana
- WL#441: Use a thread pool to serve user
- WL#3952: Add @@optimizer_switch varia
- WL#1498: Multithreaded DB-nodes
- WL#3363: ndb$info
- WL#3951: MyISAM: Additional Error Logs
  Corruption

Newest Projects

Newest Code Examples

MySQL Forge
Resources for the MySQL Community

search forge:

About the Tool Repository

The Forge Code and Tool Repository is a collection of code examples and procedures in a number of different programming languages as well as plugins and extensions for the MySQL Server and other MySQL programs. Search using the search tool or browse by tag. Do you have a code snippet you'd like to show others or a plugin you've written? Then add it!

Quick Search

Search: In:
Name and Description

Advanced Search ...

Browse By Tag

php mysql java sql database tables table debugger debugging python backup email pear array administration jdbc stored procedure replication perl shell bash mysqldump dump maintenance dynamic sql query profiler queries ruby c functions linux module xml stored function domain url connection cluster ndb api rails sessionhandling math code library innodb DBI php5 gis function

Top-Rated Tools

★★★★★	Giuseppe Maxia	mysqlresources
★★★★★	Jan Kneschke	mysql-proxy tutorial - injecting queries
★★★★★	Giuseppe Maxia	MySQL Proxy tutorial - Measuring bandwidth by user
★★★★★	Giuseppe Maxia	mysql-proxy tutorial - query and basic result logging
★★★★★	Giuseppe Maxia	mysql-proxy tutorial - loops
★★★★★	Giuseppe Maxia	MySQL Proxy tutorial - Blocking unwanted queries
★★★★★	Giuseppe Maxia	MySQL Proxy tutorial - Measuring bandwidth by session
★★★★½	Giuseppe Maxia	measuring replication speed
★★★★	Giuseppe Maxia	MySQL Proxy tutorial - Show all hooks
★★★★	Giuseppe Maxia	MySQL Proxy tutorial - tracking transaction status

More...

Newest Tools

MySQL Forge Wiki

- This seems to a requirement for all *modern* open source projects
- Internals documentation, moved to the Wiki
  - http://forge.mysql.com/wiki/MySQL_Internals
  - If you plan on writing *any* substantial code (even a tiny patch), this is what you want to read
- Localised documentation
  - e.g. http://forge.mysql.com/wiki/Tamil
- Project pages, get Wiki pages too

Contributing to the Forge

- Add your projects
- Share your code snippets
  - http://forge.mysql.com/snippets/
- Improve documentation
- Localise documentation
- Fight spam on the Wiki
- Wiki Editorial, ala, what Wikipedia does

Quality Contribution Program (QCP)

- http://dev.mysql.com/qualitycontribution.html
- Our aim is to improve the quality of MySQL products, with active community participation
- Ego++: visibly acknowledge participants in program
- Rewards++: Access to MySQL Enterprise
- Participation... is easy!
  - Reported 3 verifiable bugs? Have 3 accepted test cases? One minor patch? Start today!

Participating in the QCP

- You're rewarded for:
  - bug reports
  - test cases
  - patches
- Counts your activity in the last 12-months
- Rewards, and ego-boosting therefore, aren't for life, its a continuing process
- Rules, in public:
  - http://forge.mysql.com/wiki/QCP_Project_Rules

Contributor License Agreement (CLA)

- Legal requirements, since MySQL is dual-licensed (commercial reasons)
- Copyright transfer+assignment to MySQL
- Very much modelled after the OOo JCA
- Required for more than 15-lines of code
- Rewards: book, t-shirt, USD$100 rebate for training/conference or $100 donated to FSF
- **No code is accepted without a CLA**
- http://forge.mysql.com/contribute/

Hanging out with developers

- Internal IRC, we will always need
  - support, *secret* projects (though none exist now!)
- However, developers in general, are available via public IRC
  - irc.freenode.net
  - #mysql-dev
  - Good signal-to-noise ratio, its generally developer-speak
    - ... with the occasional amount of developer chatter
  - Generally newbie free (they're at #mysql)

Subscribing to ~~lkml~~internals

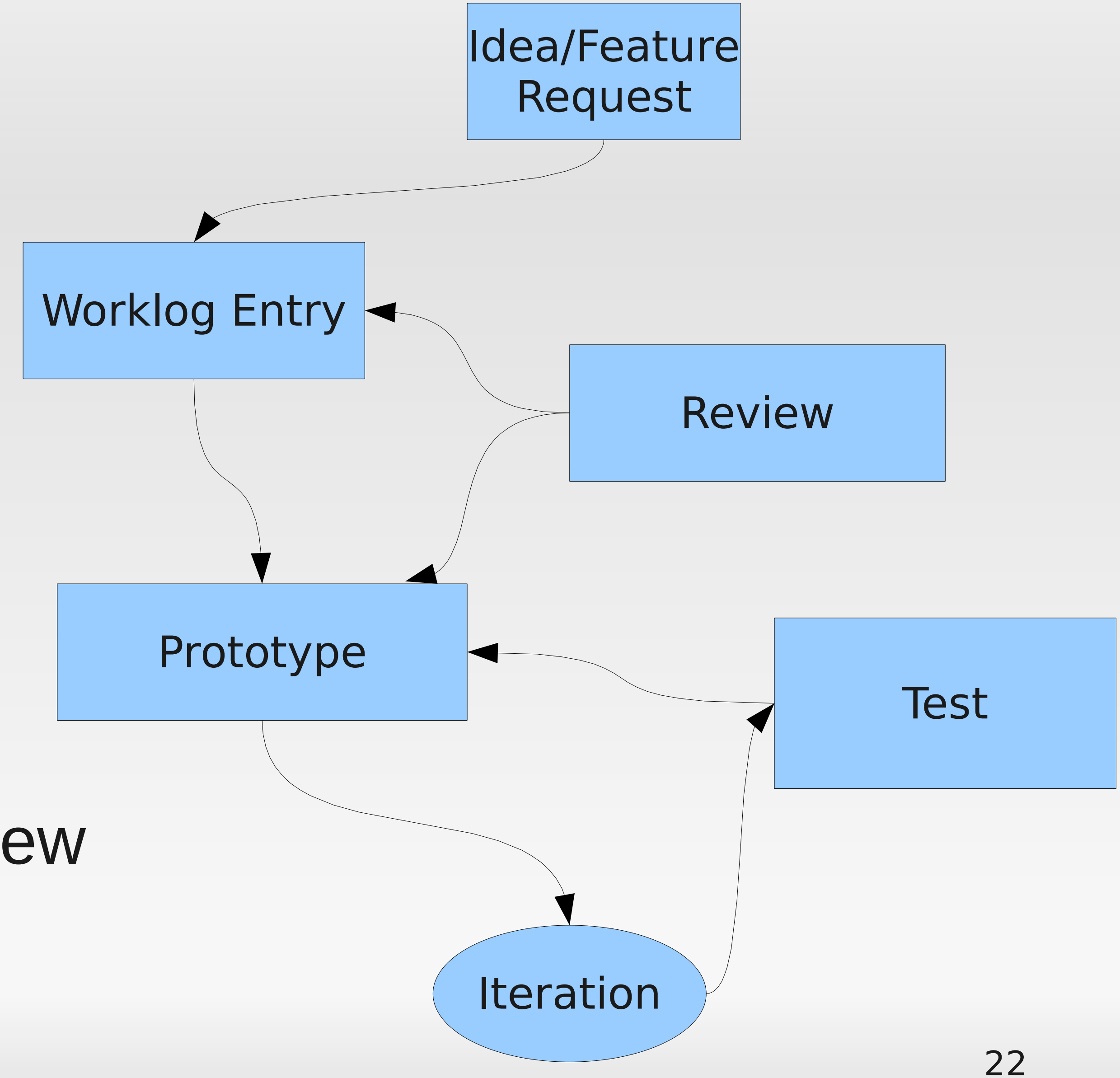
- internals@lists.mysql.com
- Subscribe to it, really. Its the LKML equivalent for MySQL
- All code reviews, are now public
- So if it doesn't happen attached to a bug, it will happen on list (usually Worklog items/new features)
- Commits? High (automated) traffic, not much discussion
    - http://lists.mysql.com/commits online forum

MySQL University

- Educating the masses, to code for MySQL
- Initially targeted at new hires, but opened up to anyone
  - Forge Wiki for documenting, #mysql-university on Freenode for chat, & ICECast streaming server for audio (with recorded OGGs for playback later)
- Over 40+ sessions in 2007
  - http://forge.mysql.com/wiki/Category:MySQLUniversity
  - Takes a summer (Northern Hemisphere) break
  - 6.30pm (summer+DST) or 7.30pm (winter) in India

How the development process works

- Somewhat agile
- Iterative
- Lots of refactoring
- Similar even for patches from externals
  - minus the usual worklog entry/review stages



Worklog

- Development and roadmap tasks for MySQL
  - This is a *roadmap* if anything
  - Describes features being worked on
  - References to WL#ids?
    - *Most* WLs are public
      - Older WL entries had customer details, so are marked private
- Has a user commenting interface!
  - Another avenue to discuss, besides bugs database, or internals list
- http://forge.mysql.com/worklog/

Getting the source

- We *still* use BitKeeper
- It is non-free software, and *very* expensive
- However, BitMover provides `bkf`, a free tool that allows cloning, and pulling updates
    - It doesn't allow committing code
- Our trees are public!
    - ... as long as the synchronisation doesn't break, they're also very up-to-date
- http://mysql.bkbits.net/

bkf 101

- `bkf clone bk://mysql.bkbits.net/mysql-5.0-community mysql-5.0-community`
  - clones the tree, to a local directory
- `bkf pull`
  - Updates the tree with the latest changes
- `bkf clone -rTAG bk://mysql.bkbits.net/mysql-5.0-community mysql-5.0-community-TAG`
  - replace TAG with mysql-5.0.45 or something, to get actual releases

Building MySQL 101

- Before making changes, build MySQL and ensure tests pass
- `BUILD/compile-dist`
  - builds mysql, as it would be built upstream
- `make test`
- `make dist`
  - source tarball generation
  - make dist --ignore ndb
- `scripts/make_binary_distribution`

Testing MySQL

- Use the **MySQL Sandbox**
- http://sourceforge.net/projects/mysql-sandbox
- Its really, MySQL in a one-click install
- `./express_install.pl mysql-5.0.45-linux-powerpc64.tar.gz`
- Check `~/msb_5.0.45` and run `./use`
- Linux/OSX only, sorry Windows folk
- Does not require root privileges, so can be run remotely on shell accounts, etc.

Coding Guidelines

- Read MySQL Internals Coding Guidelines
  - http://tinyurl.com/3cjkyz
- Its C/C++ skewed, but so is the mysql code
- Spaces (not tabs)... 2 spaces for indentation
  - sample .emacs and .vimrc online
  - common issue on internals list
- Commenting is crucial
  - Doxygen, Javadoc based (@tag)
  - http://forge.mysql.com/wiki/CommunityDoxygenProject

Code/Language coverage

- Server is mostly written in C
  - Falcon, NDB (Cluster) are written in C++
- Scripts, tests, utilities are largely written in shell (sh) and Perl
- Lua embedded in Proxy & GUI tools
- Connectors: There's (*at least*) one for every popular language, it would seem: Connector/J, Connector/.NET, Connector/Ruby, Connector/PHP, etc.

Code submission basics...

- Our Code Review Guidelines are ***public***
  - http://forge.mysql.com/wiki/Code_Review_Process
- Use `diff`, against a pristine source, please!
- `diff -purN ~/code/mysql/foo.c ~/code/mysql-working/foo.c`
- `diff -pu` or `diff -urN` works just as well

So... how to submit a patch

- Decide:
  - fix a bug? Patch is needed
  - new feature? It *may* be needed (justify yourself)
    - Bug report, S4
    - Discuss on internals
- Submit:
  - documentation (what does your patch do)
  - test case, if necessary
  - bug#id or WL#id it fixes
  - Attach to bug report, and ping internals

Things to think about

- MySQL runs on many operating systems, and platforms
- Your code must be easy to port (be portable)
  - `fork()`, `mmap()`, etc.
- Thread-safety, is locking sorted?
- You aren't using `printf`, or `gets`, right?
  - `snprintf`, `fgets`, you know, secure programming practices
- Run against valgrind?

Does it affect backwards compatibility?

- Column order changes == bad
- Breaks monitoring scripts that are out there
- Put it to the end (even though patch makes output look nicer, though)

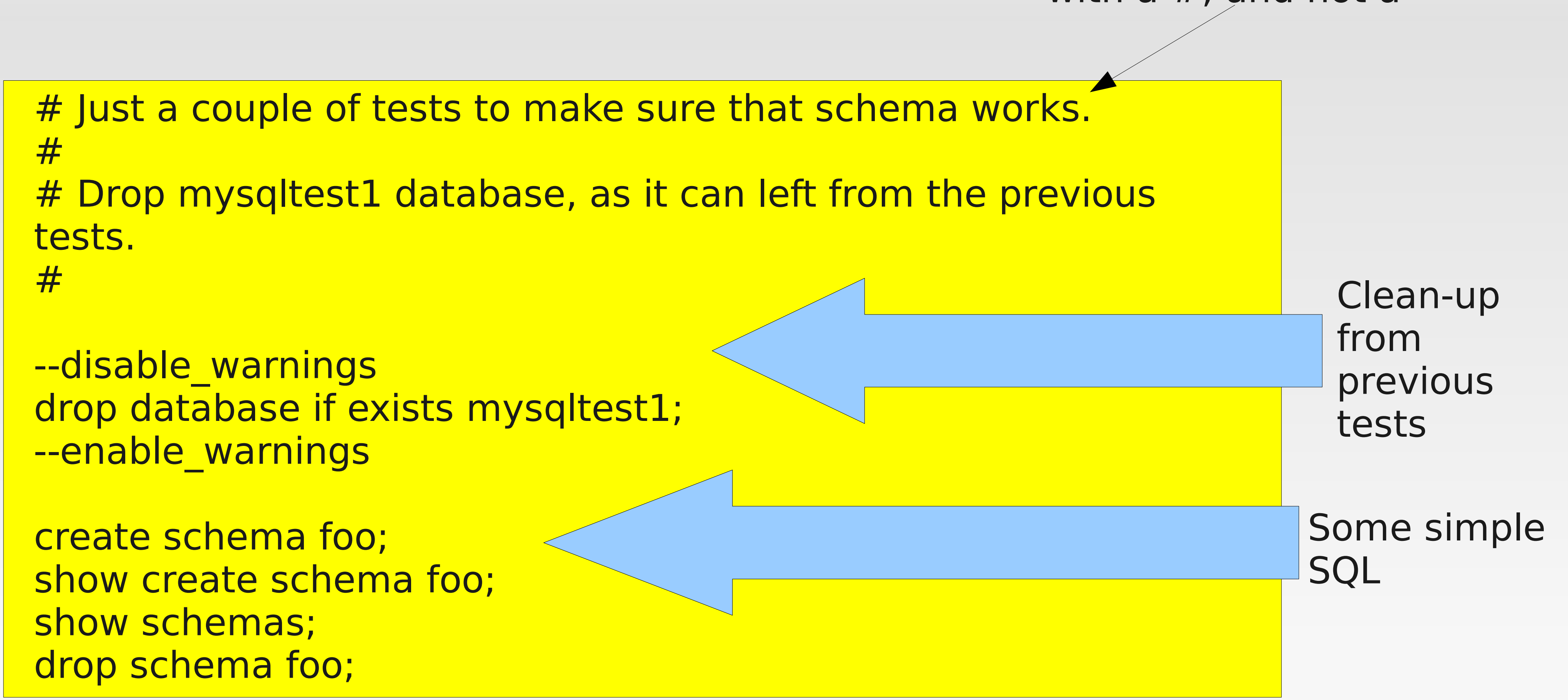
```
--- telco-new.orig/sql/slave.cc    2007-05-31 22:28:07.972811000 +1000
+++ telco-new/sql/slave.cc     2007-05-31 22:28:16.765360500 +1000
@@ -1210,6 +1210,8 @@ bool show_master_info(THD* thd, MASTER_I
                                         sizeof(mi->user)));
   field_list.push_back(new Item_return_int("Master_Port", 7,
                                     MYSQL_TYPE_LONG));
+  field_list.push_back(new Item_empty_string("Master_Bind",
+                                        sizeof(mi->bind_addr)));
   field_list.push_back(new Item_return_int("Connect_Retry", 10,
                                     MYSQL_TYPE_LONG));
   field_list.push_back(new Item_empty_string("Master_Log_File",
```

Writing good test cases

- http://forge.mysql.com/wiki/How_to_Create_Good_Tests
- Style is important: lines shorter than 80 characters, SQL keywords in UPPERCASE
- Headers are important: change/create log, what it tests against
- Tests, stored in `mysql-test/t/`
  - in `mysql-test/`, do `./mysql-test-run xa` for example
  - failure? watch `mysql-test/r/` for .reject
  - `./mysql-test-run --record xa` – records output to file

An example of a test case

- ./mysql_test_run schema



Comments, note they begin with a #, and not a --

```
# Just a couple of tests to make sure that schema works.
#
# Drop mysqltest1 database, as it can left from the previous
tests.
#

--disable_warnings
drop database if exists mysqltest1;
--enable_warnings

create schema foo;
show create schema foo;
show schemas;
drop schema foo;
```

Clean-up from previous tests

Some simple SQL

Storage Engines

- MySQL has a pluggable storage engine architecture (5.1 and above)
- Write your own storage engine, easily w/API
- Community engines
  - Amazon S3
  - PBXT
- Memcached, HTTP engine
- Engines for local uses? Baidu (China) created one against 5.0.18 for more efficient CJK functionality

Trees along the ocean

Current as of
October 2007

- Where do new features go?
  - http://forge.mysql.com/wiki/ContributionsPipeline
- SHOW PROFILE went into 5.0
  - New feature in a GA release?
- No new features in GA, since May 2007 (good)
- Community Contributions not in Enterprise, continues living on, via #ifdef
- 5.1 is in bug fix priority mode (no new features)
- 5.2 is the development tree
  - But its being merged with 6.0...

Feature Previews

- Long-term schedules of MySQL keep new features out of GA for *too long*
- Snapshots, feature previews
  - http://forge.mysql.com/wiki/Falcon_Feature_Preview
    - x86_64 and Win32-32bit binaries available
    - Not Big Endian capable, *yet*
  - http://forge.mysql.com/wiki/GIS_Functions
- While the Cluster “telco” edition isn't in binary form, it is available from mysql.bkbits.net

Summer of Code 2007

- We participated for the first time
- 8 projects that garnered us code
  - 4 extending the server
  - 1 extending the test suite
  - 3 creating completely new “3rd party” tools
- Thanks to the SoC, we now have IPv6 support in MySQL
  - http://mysql.bkbits.net:8080/mysql-5.2-ipv6/

MySQL Proxy

- Sits between client and MySQL server(s), that can monitor, analyse, or transform communication
- Load balancers, failover (without heartbeat), query filtering, query modification, etc.
- GPL licensed core (written in C), some scripts are commercial only (written in Lua)
  - Good, you can write your own, in Lua
- `mysql  --host=127.0.0.1 --port=4040 --user=... --pass=...`

MySQL Proxy Example

- A simple script, that displays all queries executed by the mysql server

```
function read_query(packet)
  if string.byte(packet) == proxy.COM_QUERY then
    print("we got a normal query: " .. string.sub(packet,2))
  end
end
```

- `mysql-proxy --proxy-lua-script=intercept.lua`
- Outuput: just that you got a normal query
  - `we got a normal query: select user, host, length(password) from mysql.user`

MySQL Proxy Scripts

- Scripts are in LUA (5.1 currently)
- Encouraged sharing at: http://forge.mysql.com/wiki/MySQL_Proxy
- Replication in Proxy scripts? Situations are endless
- Good snippets:
  - Measure bandwidth consumed by user - http://forge.mysql.com/snippets/view.php?id=111
  - Block unwanted queries - http://forge.mysql.com/snippets/view.php?id=108

Contributing isn't only about coding

- Its beyond patches, bugs, test cases, et al
- Write a blog, talk about MySQL
  - Syndicate to http://planetmysql.org/ (PR=8)
- Answer questions on forums, mailing lists, IRC
- Destroy the FUD (MySQL *Mythbusters*)
- Go to conferences, events, & evangelise
- Create podcasts, videoblogs
- Make artwork
- http://forge.mysql.com/wiki/Contributing

How we've fared so far...

- Still hard to not feel the urge to hire contributors
  - Growing company, we've still hired a majority of our community contributors
  - Only 5% from top Quality Contributors
- Internal-external rift, i.e. old habits die hard
  - Internal IRC is still a lot more populated, than freenode #mysql-dev
- BitKeeper, really hampers externals actually committing into the code-base

How we've fared so far... II

- Coding guidelines haven't taken into consideration extending Connectors or even the Proxy or GUI tools
- Not really set up to accept code for the above
- “We can rewrite that in 5-hours mentality”
  - SHOW PROFILE in 5.0 was community contributed, but re-basing it for 5.1 was such a nightmare, that it made sense to rewrite it
- Recently, invested in community infrastructure

One more thing

- SoC 2007 thought us about creating target lists
  - http://forge.mysql.com/wiki/SummerOfCode2007Ideas
- Continuing with the idea of lists, of Worklog items
- We will provide a “crash course” on the MySQL development process
  - i.e. More in-depth code/testing coverage, etc.
- You get *assigned* a MySQL “mentor” from the start
  - Assistance, guidance, reviews happen, etc.

A preview of the list

- Bugs
  - 5299, 9541, 12917, 16565, 17078, 17079, 17124, 19027, 19612, 20263, 20415, 20454, 20455, 23529, 29304, 29497, 29622, 29716, 29730, 29738, 30025, 30139, 30140, 30261, 30394, 31106, 31354, 31486, 32318, 32457, 32738, 20764, 21099, 23028, 23222, 22054, 25748, 26266, 14594, 17738, 19538, 22891, 23606, 25641, 26243, 26704, 26892, 1382, 25734, 30414, 17034
- Worklog
  - 2929, 881, 2814, 3262, 3129, 605, 2825

MySQL Developer Meeting

- Yearly. First time we allowed community contributors to show up (Heidelberg, 2007)

Thanks! Questions?

E-mail me:
byte@mysql.com
Catch me on IRC, at #mysql-dev:
ccharles / bytee