Beyond MySQL GA: patches, storage engines, forks, and pre-releases

Kristian Nielsen

`knielsen@knielsen-hq.org`

MariaDB Developer
Monty Program AB

FOSDEM 2010

There is more to MySQL than standard MySQL 5.1 GA!

In the last 3–4 years, a lot of exciting development has taken place on the MySQL server outside of MySQL AB/Sun/Oracle. In this talk, I want to present two main points:

- Give an overview of some of the different extensions and enhancements to MySQL that have been developed by the community and are now available
- Show what the options are for users to use and benefit from this development

Patches, storage engines, extensions, branches

Google patches

Google patches (mostly for 5.0, some 5.1):

- Better statistics/monitoring:
  - Lock contention
  - Binlog events
  - malloc() statistics
  - Filesorts
  - InnoDB I/O and wait statistics
  - `SHOW {USER,TABLE,INDEX,CLIENT}_STATISTICS`
- Performance improvements (lots of it for InnoDB):
  - SMP scalability (lock split, atomic operations, tcmalloc)
  - I/O scalability

Google patches (cont.)

- Semisync replication
- binlog event checksums and global transaction id
- Replication state in InnoDB table (crash recovery)
- Incremental table checksums (for consistency check)
- `SET GLOBAL innodb_disallow_writes=(0|1)`
- `KILL IF_IDLE <id>`
- `MAKE (USER|CLIENT) <who> DELAYED 100`
- Mapped users
- Audit log
- Lots of other small tweaks

Percona patches

Implemented or integrated by Percona, for MySQL 5.0:

- Percona-authored and external
- Focused on statistics/monitoring and performance/scalability
- Some of this ported to 5.1 in XtraDB (InnoDB parts)
- `innodb_check_fragmentation`
- `INFORMATION_SCHEMA.INNODB_BUFFER_POOL_CONTENT`
- `innodb_split_buffer_pool_mutex`
- `innodb_dict_size_limit`
- `microsec_process`
- `microslow_innodb`
- ...

Miscellaneous patches

Ebay patches:

- Variable-length rows in memory storage engine
- Pool-of-threads

Virtual columns:

- ```
  CREATE TABLE t
       (a VARCHAR(32),
        b VIRTUAL VARCHAR(5) AS (LEFT(a,5)))
  ```
- Patch by Andrey Zhakov

XtraDB storage engine

- Percona-developed improved InnoDB (for MySQL 5.1)).
- Improved statistics/monitoring
- Better CPU/IO scalability (performance)
- XtraBackup
- Includes a number of other patches (eg. Google patches)
    - Also some Percona-developed things
- Includes Oracle/Innobase InnoDB plugin

PBXT storage engine

- Transactional MVCC storage engine
- Uses novel write-once, log-based storage rather than traditional transaction log + write-back buffer cache
- Promising benchmarks
- Good to see new approaches, hardware is much different now than when traditional algorithms were designed
- Vladimir Kolesnikov will say more in the afternoon.

FederatedX storage engine

- Replacement for MySQL Federated storage engine, which is no longer maintained.
- A FederatedX table acts as a link to a table in another database server
- Can access tables in remote server with local SQL
- Possibility to do cross-server joins, etc.
- Contributed and maintained by Patrick Galbraith

Sphinx Storage Engine

- Interface to Sphinx Full-text search
- Spinx is an external full-text search engine
- Spinx storage engine allows to query the (external) full-text search using SQL.

Pinba storage engine

- Collect PHP statistics via UDP and present as MySQL tables
- Used to profile PHP servers to find bottlenecks (Pinba: PHP is not a bottleneck anymore).
- Collects information like
  - Host
  - Request count
  - Document size
  - Memory usage
  - CPU time
  - Wallclock time
  - Custom timers

OQGraph and Spider storage engines

OQGraph:

- Special-purpose storage engine for graph-type relations
- Optimised for queries such as needed for social network type applications (facebook, twitter, linkedin, ...)

Spider:

- Storage Engine for database sharding and clustering

Galera: Synchronous Multi-Master replication

- Developed by Codership
- Synchronous replication
- Changes made durable on *all* servers before `COMMIT` returns
- Optimisations to utilise query parallelism to get decent performance despite round-trip cost
- InnoDB specific

Drizzle

- Forked from MySQL 6.0 in 2008
- Radical rewrite/cleanup of MySQL code base
- No plans to converge Drizzle and MySQL code base
- David Axmark will say more

MySQL pre-releases

- MySQL Community Server 5.5.1
- Basis for next version of MySQL
- New release model for post-5.1 releases
- Similar to alpha/beta release

MySQL Previews

- `http://forge.mysql.com/wiki/`
  `Category:Software_Preview`
- Publishes selected feature branches
- Replication:
    - Delayed replication
    - Parallel SQL slave threads
    - Scriptable replication
    - Semi-synchronous replication
- `SHOW PROFILES`
- Online Backup
- Falcon
- Batched key access

So how can one use all of this stuff?

Alternate packages

- Combination of stock MySQL with various community contributions, maintained by 3rd parties
- Combine official MySQL releases with patches and storage engines
- Release source and binary packages
- Only practial method to use large number of external community contributions
- Problem if your favorite patch is not included
- Will lag behind official MySQL releases (typically 1 month, sometimes more)

Alternate packages for 5.0

MySQL 5.0:

- OurDelta
- Percona binaries
- Maintain series of patch files separately
- Need to update patch series for every new MySQL release, and resolve/rewrite code changes that conflict
- Easy to drop a patch

Alternate packages for 5.1

MySQL 5.1:

- MariaDB, a branch of MySQL 5.1
- Sergey Petrunya will say more 15:00 today
- Community patches and storage engines, plus own development
- Source and binary packages
- Maintained as a parallel bzr branch to MySQL on Launchpad
- Tighter integration between trees, more help from revision control software to control merging and conflicts
- Easier to handle major changes
- Still a lot of work to handle conflicts

Patch MySQL source tree and compile

- `patch -p1 -s < my-patch-file.diff`
- Resolve conflicts
- Custom storage engine instructions
- Loadable .so plugin (need identical build options and version)
- `quilt` (advanced merging of multiple patches)
- Too much effort for everyone to do this individually
    - need to integrate once in a way that many can reuse

Conclusion

- Last few years has seen an explosion in community-driven MySQL development outside of MySQL/Sun/Oracle
  - Good to be aware of what is available
- Alternate packages makes it easier to use these community-developed extensions
  - More work in currently being done to make this even easier in the future

Slides:

```
https://knielsen-hq.org/maria/fosdem2010.pdf
```

Contact:

```
knielsen@askmonty.org
```