MariaDB: Viable MySQL replacement

Colin Charles, Team MariaDB, SkySQL AB
colin@mariadb.org | byte@bytebot.net
http://mariadb.org/
http://bytebot.net/blog/ | @bytebot on Twitter
SCALE12x, Los Angeles, California, USA
22 February 2014
whoami

- Work on MariaDB at SkySQL Ab
- Merged with Monty Program Ab, makers of MariaDB
- Formerly MySQL AB (exit: Sun Microsystems)
- Past lives include Fedora Project (FESCO), OpenOffice.org
- MHA experience
  - since November 2011 (MHA 0.52, 0.53)
  - NRE work to make it run in a Solaris 10 environment... with no Internet access!
  - Continued deployment advice + work for data centre use
  - Much thanks to SkySQL for the experience

Sunday, 23 February 14
MySQL? Percona Server? MariaDB?
Agenda

- 4 years: major server releases (5.1, 5.2, 5.3, 5.5, 5.5+TokuDB, Galera Cluster) and 10.0 series
- Delving into history of previous releases
- MariaDB 10.0
- Client libraries, Galera Cluster
- Roadmap
What isn’t covered

- MariaDB Enterprise
- Galera Cluster + GUI + API
- mariadb.com
- SkySQL
- trademarks...
What is MariaDB?

- Community developed branch of MySQL
- Feature enhanced
- Fully compatible & feature complete with MySQL
Backed by MariaDB Foundation

- Driver of the MariaDB project
- Foundation not controlled by single entity/person; has a Board
- Ensure MariaDB is compatible with MySQL, maintain mariadb.org, keep community voice
- Major sponsors: SkySQL, Parallels, Booking.com, Automattic, OpenQuery, Percona, Webyog, more!
Aims of MariaDB

• Compatible, drop-in replacement to MySQL
• data on disk & on the wire the same
• same file names, sockets, port
• Stable (bug-free) releases with no regressions
• GPLv2
Why MariaDB 10.0?

• The 5.5 merge took about a year (!)
• In MariaDB 5.5, we have over 1.5 million lines of extra code ~61MB diff
• MySQL 5.6 refactored with huge losses in commit history
• We’re not a patch set against MySQL
  • MariaDB clearly does not depend on MySQL for future development
MariaDB 10.0 in a nutshell

• Built on MariaDB 5.5
• Backported features from MySQL 5.6
• Multiple new features
Where to get it

- Shipped as default in many Linux distributions or as an optional package
- 10.0 series under development/unstable branches in some distributions
- https://downloads.mariadb.org/
- apt, yum repositories for the latest 10.0 releases + repository configuration tool
Quick recap of MariaDB 5.5

- Built on MariaDB 5.3 + MySQL 5.5
- Non-blocking client library
- LIMIT ROWS EXAMINED
- Extended keys for XtraDB/InnoDB
- New SphinxSE
- New threadpool
Quick recap of MariaDB 5.3

- Built on MySQL 5.1, with MariaDB 5.1+5.2
- Storage engines (Aria, XtraDB, FederatedX, etc.), table elimination, virtual columns, extended user statistics, segmented MyISAM keycache, pool of threads, optimizer features, microsecond precision, faster HANDLER interface, dynamic columns, HandlerSocket, group commit in the binary log, checksum for binary log events, progress reporting for ALTER TABLE/LOAD DATA INFILE, GIS precise operations, kill by query ID (not just thread ID) or by USER
### MariaDB 10.0

<table>
<thead>
<tr>
<th>Date</th>
<th>Version</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>12 Nov 2012</td>
<td>10.0.0</td>
<td>Alpha</td>
</tr>
<tr>
<td>6 Feb 2013</td>
<td>10.0.1</td>
<td>Alpha</td>
</tr>
<tr>
<td>24 Apr 2013</td>
<td>10.0.2</td>
<td>Alpha</td>
</tr>
<tr>
<td>11 June 2013</td>
<td>10.0.3</td>
<td>Alpha</td>
</tr>
<tr>
<td>16 Aug 2013</td>
<td>10.0.4</td>
<td>Alpha</td>
</tr>
<tr>
<td>7 Nov 2013</td>
<td>10.0.5</td>
<td>Beta</td>
</tr>
<tr>
<td>18 Nov 2013</td>
<td>10.0.6</td>
<td>Beta</td>
</tr>
<tr>
<td>27 Dec 2013</td>
<td>10.0.7</td>
<td>Beta</td>
</tr>
<tr>
<td>10 Feb 2014</td>
<td>10.0.8</td>
<td>RC</td>
</tr>
</tbody>
</table>
Backported features

- InnoDB (from MySQL 5.6.14)
- PERFORMANCE_SCHEMA
- Online ALTER TABLE (10.0.4)
- Upstream collation tables (MDEV-4928)
- Privileges on temporary tables
- GET DIAGNOSTICS
- Optimizer
  - ORDER BY...LIMIT optimization (shows only few rows of a result set)
Re-implemented from MySQL 5.6

• Error messages (w/system error string + meanings)

• `CURRENT_TIMESTAMP` / `NOW()` as DEFAULT for DATETIME columns

• `EXPLAIN` for INSERT/UPDATE/DELETE

• Temporal literals - `TIME’12:05:00’`
MariaDB 10 replication

- Global Transaction ID
- have complex replication topologies; simple failover & slave promotion
- doesn’t require restarts!
- new slave provisioning: `SET GLOBAL GTID_SLAVE_POS = BINLOG_GTID_POS("master-bin.00045", 600); CHANGE MASTER TO master_host="192.168.2.4", master_use_gtid=slave_pos; START SLAVE;`
- turning on GTID for slaves: `STOP SLAVE CHANGE MASTER TO master_use_gtid=current_pos; START SLAVE;`
- change masters: `STOP SLAVE CHANGE MASTER TO master_host="10.2.3.5"; START SLAVE;`
- Crash-safe slaves - GTID position stored in InnoDB table
Only in MariaDB 10.0: Multi-source replication

- Work from Taobao/Lixun Peng
- Known to be in-use for production purposes at Tumblr, FunPlus Game, Taobao, Flipkart
- Many users partition data across many masters... now you can replicate many masters to a single slave
- Great for analytical queries, complete backups, etc.
Only in MariaDB 10.0

- SHOW EXPLAIN for <thread_id> gets the query plan for a running statement
- Per-thread memory usage (Taobao)
  - INFORMATION_SCHEMA.PROCESSLIST has MEMORY_USAGE & EXAMINED_ROWS now
- SHOW STATUS has memory usage too
Only in MariaDB 10.0: CassandraSE

- MariaDB as a “data platform”
- Integration with NoSQL/Big Data DB, Apache Cassandra cluster, seen as a storage engine to MariaDB
- Combine (join) data between Cassandra & MariaDB & Oracle (via CONNECT)
- Write to Cassandra from SQL (SELECT, INSERT, UPDATE, DELETE)
Only in MariaDB 10.0: Engine independent persistent statistics

- InnoDB has persistent statistics in MySQL 5.6; we have an engine-independent version
- These statistics aren’t limited by the SE API, and are used by query optimizer to choose best execution plan for each statement
- Statistics collected for non-indexed columns too (unlike InnoDB’s)
Dynamic columns

- Allows you to create virtual columns with dynamic content for each row in table
- Basically a blob with handling functions (GET, CREATE, ADD, DELETE, EXISTS, LIST, JSON)
- Store different attributes for each item (like a web store). Hard to do relationally
- In MariaDB 10: name support (instead of referring to columns by numbers, name it), convert all dynamic column content to JSON array, interface with Cassandra
- `INSERT INTO tbl SET dyncol_blob=COLUMN_CREATE("column_name", "value");`
Group commit in the binary log

- `sync_binlog=1, innodb_flush_log_at_trx_commit=1`

![Graph showing Throughput with group commit](image.png)
Threadpool 5.5 vs 5.1

![Graph showing performance comparison between pool-of-threads and thread-per-connection models. The graph illustrates throughput (tpm) against concurrent clients.](image)
PAM Authentication

- Authentication using /etc/shadow
- Authentication using LDAP, SSH pass phrases, password expiration, username mapping, logging every login attempt, etc.
- INSTALL PLUGIN pam SONAME ‘auth_pam.so’;
- CREATE USER foo@host IDENTIFIED via pam
- Remember to configure PAM (/etc/pam.d or /etc/pam.conf)
SphinxSE

- CREATE TABLE t1 (...) ENGINE=SPHINX
  CONNECTION="sphinx://localhost:9312/test";
- Engine connects to Sphinx searchd
- Let indexing, searching, sorting, filtering be performed by Sphinx
  - instead of WHERE, ORDER BY, LIMIT
- Sphinx is optimized/fast for these tasks
- Most of the Sphinx API is exposed to engine
- JOIN search table with other MySQL tables
GIS support!

- MySQL has OpenGIS SFS (Simple feature access, SQL access method)
- Now, SQL with full geometry types
- ST_ prefix
- MyISAM, Aria for SPATIAL & non-spatial indexes
- Use Osmosis, you can load all OpenStreetMap data into MariaDB now

MariaDB 10.0.2

- Support for atomic writes on FusionIO DirectFS
- Optimizer collects & can use histogram-based statistics for non-indexed columns
- Better table discovery, so FederatedX has assisted discovery, Sequence engine (creates ascending/descending sequences, useful in joins)
- `SHOW PLUGINS SONAME;`
MariaDB 10.0.4

- SPIDER storage engine for database sharding merged
- Audit plugin
- complete PERFORMANCE_SCHEMA
- INFORMATION_SCHEMA with upstream defaults too
- Online ALTER for InnoDB and thread information for in-place operations
MariaDB 10.0.5

- automatically detect independent transactions, parallel within same table, adapts to master load, and preserves commit ordering
MariaDB 10.0.6

• Serious incompatibility and data corruption of DATETIME and DATE types due to get_innodb_type_from_mysql_type refactor combined with InnoDB Online DDL

• https://mariadb.atlassian.net/browse/MDEV-5248

• Fixed upgrades from MySQL 5.1 -> MariaDB

• Parallel replication improvements
MariaDB 10.0.7

• Mostly bug fixes, to stabilise the code
• XtraDB 5.6 merged (InnoDB still default)
• OQGraph v3 - stores data on disk, persistent, larger graph support
• INFORMATION_SCHEMA.METADATA_LOCK_INFO plugin to see active metadata locks
MariaDB 10.0.8 (RC)

- Upgraded bundled PCRE library
- Upgraded InnoDB
TokuDB

- Opensource - separate MariaDB 5.5+TokuDB/ integrated in 10.0.5
- Improved insert (10-20x faster) & query speed, compression (up to 90% space reduction), replication performance and online schema flexibility
- Uses Fractal Tree Indexes instead of B-Tree
- Tests & builds of TokuDB on multiple platforms (think greater distribution)
CONNECT Storage Engine

- Made by Olivier Bertrand
- Read, write & update files in different storage formats:
  - .DBF (dBASE format)
  - .CSV
  - .INI
  - XML
  - ODBC
- Possible to join data from CassandraSE, XtraDB and an ODBC data source like Oracle
Roadmap

• MariaDB has already a superset of features in MySQL
• MariaDB 10.0 series will be mostly compatible with MySQL 5.6
  • see no EXPLAIN FORMAT=JSON, InnoDB memcached plugin, etc.
• 5.6 complete + 5.7 equivalent should be 10.1 (release aim: end-2014, early-2015)
  • 100% OpenGIS compliance, ANALYZE TABLE as online operation, IPv6 datatype, query logging, summary per query, audit for specific users, etc. -- see jira
MariaDB 10 has MySQL Enterprise features

- Threadpool
- PAM authentication plugin
- Audit plugin
MariaDB Galera Cluster

- MariaDB Galera Cluster is made for today’s cloud based environments. It is fully read-write scalable, comes with synchronous replication, allows multi-master topologies, and guarantees no lag or lost transactions.

- Currently 5.5-based

- 10.0.7 release as alpha, wsrep API v25, Galera 25.3.2
MariaDB LGPL Connectors

- C
- Java
- ODBC
- Works with MariaDB, Percona Server, MySQL
- Credit goes to: old LGPL client library for C, Drizzle driver for Java

Sunday, 23 February 14
Benchmarks

- “Lies, damned lies, and statistics” - Mark Twain
- http://blog.mariadb.org/sysbench-oltp-mysql-5-6-vs-mariadb-10-0/
- Yes, we’ve gotten Oracle to notice MariaDB :-)}
Continued commitments

- Security
  - Since about a year now, we’re the go-to people for security - good track record
- We don’t like regressions
- We care about backward compatibility & introduce features carefully
  - `XtraDB innodb_adaptive_checkpoint=none|reflex|estimate|keep_average` (no more reflex...)

Sunday, 23 February 14
Community involvement

- Many features since MariaDB 5.2 have come from the community or are sponsored features (5.3, 5.5)
- Knowledgebase has 3,000+ articles in English
- 6-figure monthly downloads (more users from mirrors + distributions)
- Active mailing lists (stats on KB)
- Google Summer of Code 2013
MariaDB is gaining popularity

- Wikipedia (English, German), Wikidata running MariaDB 5.5
- Google has public commitment to adapt, QA MariaDB 10
- Fedora, OpenSUSE shipping MariaDB as a default
- Slackware, Chakra Linux, ArchLinux have followed suit
- Many success stories at KB: Limelight Networks, Nimbuzz, Paybox, FictionPress, OLX, SlashGear, Web of Trust, SpamExperts, Cougarboard, etc.
Well supported

- Everyone supports MariaDB from a support standpoint with the exception of Oracle
- SkySQL, Percona, etc.
- Jelastic has it as PaaS too
- We support all GA releases for 5 years for security, etc.
User stats plugin

• Disabled by default, consider enabling it to show use! http://mariadb.org/feedback_plugin/

data from over 133 countries!
Compatibility with MySQL

- No NDB Cluster
- XtraDB enabled as default up to 5.5, InnoDB in 10.0 (for now)
  - XtraDB from Percona Server 5.6 merged, but we’re thinking of just InnoDB as default
- Optimizer: no worse than current plan
• Can I replicate from MySQL 5.6 to MariaDB 10.x?
  • Yes

• Can I replicate from MariaDB 10.x to MariaDB 5.5?
  • Yes

• Can I replicate from MariaDB 10.x to MySQL 5.5/5.6?
  • No
FAQ on tools

- SELECT VERSION() returns 10.0.5-MariaDB
- Version string in handshake packet will be 5.5.30-mysql-10.0.2-MariaDB (mysql#68187, MDEV-4088)
- Tools should start recognising MariaDB for additional feature-set (mytop, HeidiSQL, etc.)
Conclusion

• We’ve spent a lot of time adding features, some in parallel, some ahead
• MariaDB is binary compatible with MySQL
• Open bugs system, test suite, discussion lists
• Opensource, feature rich, no commercial extensions
Books

- MariaDB Crash Course, Ben Forta (September 2011)
- Getting Started with MariaDB, Daniel Bartholomew (October 2013)
- MariaDB Cookbook (coming soon - early 2014)
Resources

- bugs: mariadb.org/jira
- low traffic **announce** mailing list: http://lists.askmonty.org/cgi-bin/mailman/listinfo/announce
- maria-discuss@lists.launchpad.net
- maria-developers@lists.launchpad.net
- fb.com/MariaDB.dbms
- twitter: @mariadb
- google plus: +MariaDB
- #maria on irc.freenode.net
Thanks/Q&A

Colin Charles, colin@mariadb.org | byte@bytebot.net
http://bytebot.net/blog/ | @bytebot
slides: slideshare.net/bytebot/

planetmysql.org
planetmariadb.org