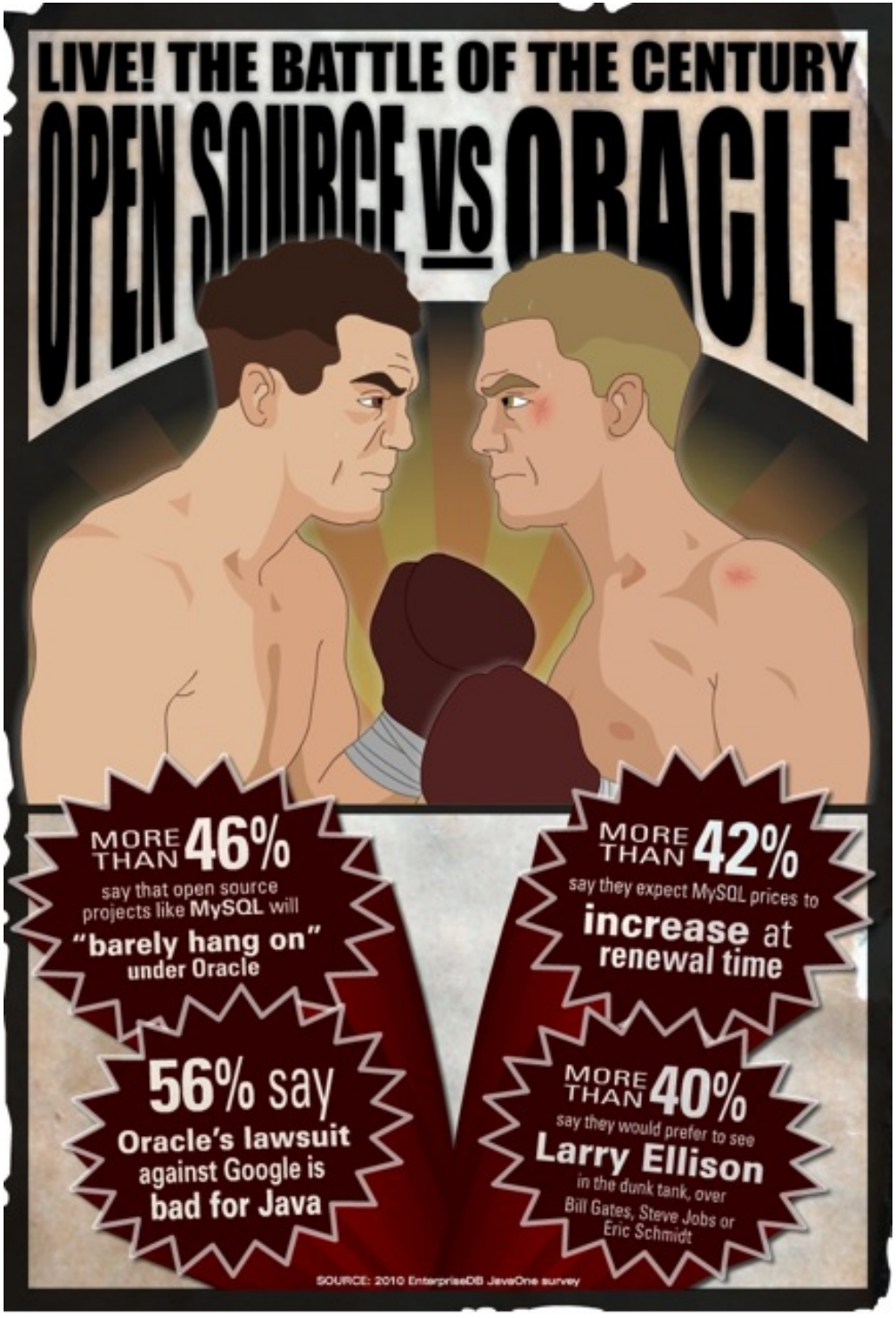
# Latest news from the MariaDB (and MySQL) community

Kristian Nielsen
knielsen@knielsen-hq.org

MariaDB Developer
Monty Program AB

Open Source Days 2011




EnterpriseDB survey
March 3, 2011

# Outline



# Outline



# The development community

There is much more to MySQL server development than the MySQL group @ Oracle!

- MariaDB
- Percona Server
- MySQLAtFacebook
- Google MySQL patches
- Continuent/Tungsten replication
- Galera synchronous replication
- Third-party storage engines
- Sphinx full-text search
- FederatedX, OQGraph, Spider
- ...

# MariaDB

- Fork of MySQL 5.1+, started 2009
- Initiated by Michael “Monty” Widenius, MySQL founder and original programmer.
- More open and community-oriented development than possible inside MySQL/Sun.
- Monty Program AB
  - Around 10 full-time developers
  - Similar number of sysadmin, web, documentation, administration, . . .
- Funding: NRE, L3 support partnership
- Some outside participation also (we want more)

# What does MariaDB offer?

MariaDB is a **drop-in replacement** for MySQL

- It is just another version of the MySQL server
  - Latest MySQL 5.1.55 + enhancements
- Fully backwards and forwards compatible
  - No dump/restore needed
- Fully SQL-level compatible
  - No need to change existing applications
- Full protocol and client library binary compatibility
  - No need to recompile existing applications
  - Can use MySQL client library against MariaDB (or vice versa).

# Easy migration

- Migration is easy, eg. Debian/Ubuntu:
  - Point `/etc/apt/sources.list` towards the MariaDB repository
  - `sudo apt-get install mariadb-server`
- Automatic upgrade (even from MySQL 5.0).
- Upgrade is tested automatically in Buildbot
  - After every commit
  - All distro versions

# MariaDB enhancements

- Bug fixes
- Performance
  - Percona XtraDB
  - MyISAM segmented key cache
  - Table elimination
  - ...
- Diagnostics
  - Extended user statistics
  - Slow query log extensions
  - `INFORMATION_SCHEMA` extensions
  - Microsecond precision in SHOW PROCESSLIST
  - ...
- Features
  - Virtual Columns
  - Storage engines: PBXT, Sphinx, OQGraph, ...

# Continous merging of new MySQL changes

- MariaDB continously merges changes to MySQL
  - Usually released within 1-2 months
- So *not* diverging
- “MariaDB is a *branch*, not a fork”
- Merge of MySQL 5.5 with MariaDB is in progress
- Merging is a lot of work
  - Especially since MySQL/Oracle does not cooperate
  - MariaDB has the needed skillset and momentum

# Open development

MariaDB is a full Open Source project, with an open development model:

- Public bug tracker (Launchpad)
- Public mailing lists
- Public IRC channels (FreeNode)
- All development branches public (Launchpad)
  - Commit rights to outsiders based on merit, not employment
- Public Worklog
- Open developer meetings twice a year
  - Lissabon next week still open :)

Please come and participate!

# Save the people, save the project

Two acquisitions in a row does not go unnoticed

- (Too) many key developers leaving
  - By my subjective opinion, 5 of the ten best/most important developers have left MySQL@Oracle
  - Fortunately, 4 of them are working full-time on MariaDB
- “Save the people, save the project”

Healty competition

- Keep MySQL management on their toes
- Having an alternative available reduces incentive for Oracle to stall development of MySQL

# Track record

Two year track record of releases

| | |
|---|---|
| 5.1.55 | 2011-03-01 |
| 5.1.53 | 2010-12-06 |
| 5.1.51 | 2010-11-19 |
| 5.1.50 | 2010-09-09 |
| 5.1.49 | 2010-08-09 |
| 5.1.47 | 2010-06-01 |
| 5.1.44 | 2010-05-10 |
| 5.1.42 | 2010-02-01 |
| 5.1.41 | 2010-01-13 |

| | |
|---|---|
| 5.2.5 | 2011-03-03 |
| 5.2.4 | 2010-12-06 |
| 5.2.3 | 2010-11-10 |
| 5.2.2 | 2010-09-28 |
| 5.2.1 | 2010-06-18 |
| 5.2.0 | 2010-04-10 |

# MariaDB is an independent alternative to MySQL

- MariaDB is the primary fork of MySQL
- A lot has been achieved in the first two years
- MariaDB has gained a lot of momentum
  - Likely to become the main focus of development efforts outside MySQL
- MariaDB is a viable, independent alternative to MySQL
  - No risk of loosing the “M” in LAMP.

(And remember, MySQL is still here)

# But there is more!

Percona

- High profile MySQL consultants
- Percona Server
  - Key enhancements based on their experience of what their customers need
  - More conservative about what to include
  - Individual patches, not a full fork
- Most patches included in MariaDB
  - (We will get all eventually)

# But there is (even) more!

MySQLAtFacebook

- Big MySQL installation
  - “X” 1000 servers
  - 13M queries per second
  - 450M/3.5M rows read/written per second
- Team of 3-4 working on the server code
- Plans for including their enhancements in MariaDB

Google

- Also run their own version of the server
- Several Google patches already in MariaDB/Percona Server, or planned to be included.

# Outline



# Upcoming feature: group commit

- Last year I worked on this feature
- Good example of the kind of development that is possible with MariaDB
- Will be in the next release, and also available in a feature preview
  - `http://kb.askmonty.org/v/mariadb-52-replication-feature-preview`

# Improving binlog commit performance

This is a feature to improve commit performance when the binlog is enabled.

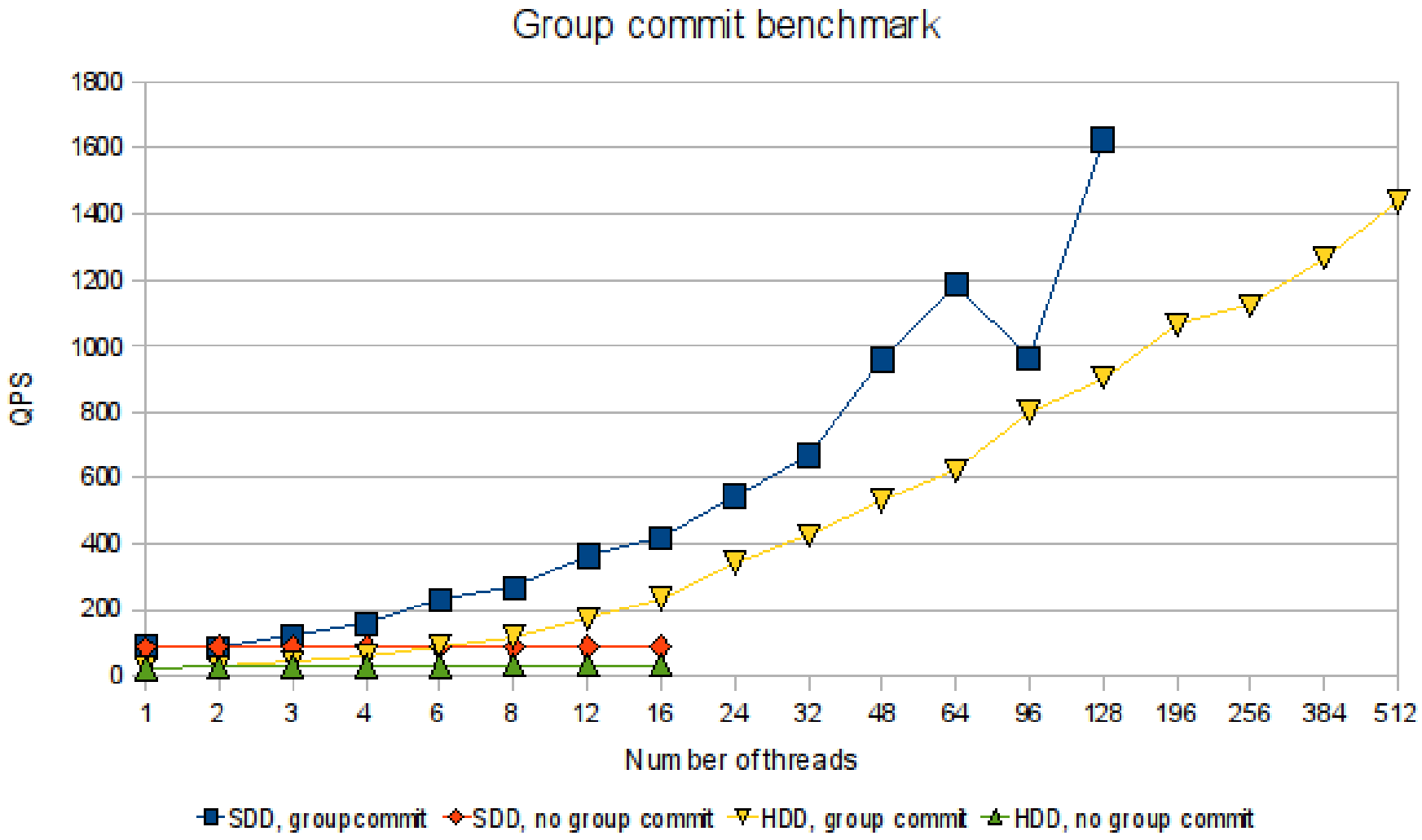
- binlog is what is used to provide MySQL replication

Big improvement when `sync_binlog=1`

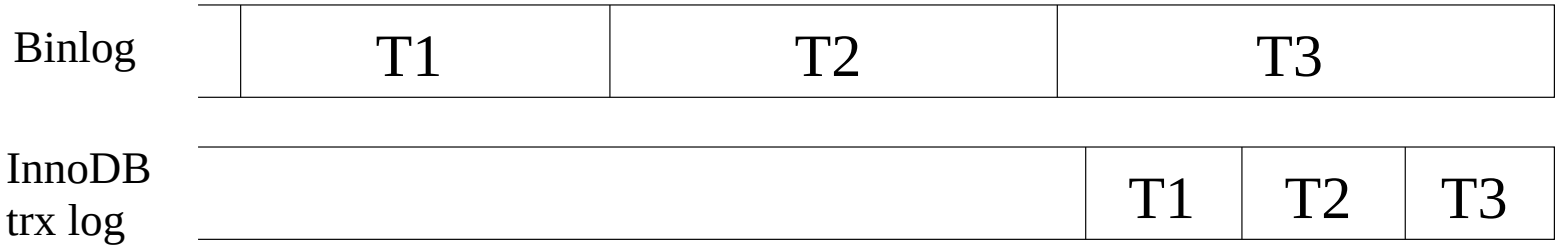
The `sync_binlog=1` option is needed to

- Ensure durability of commit (“D” in ACID)
- Recover replication into a consistent state after crash

# `sync_binlog=1` performs badly

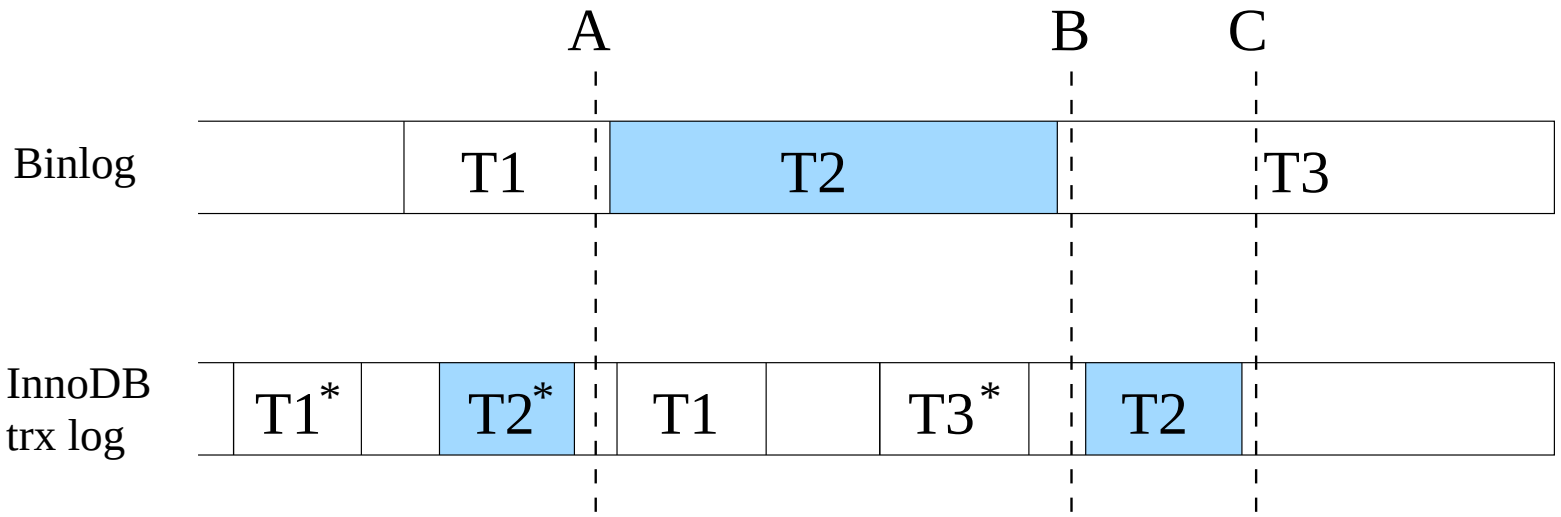



# Multiple transaction logs



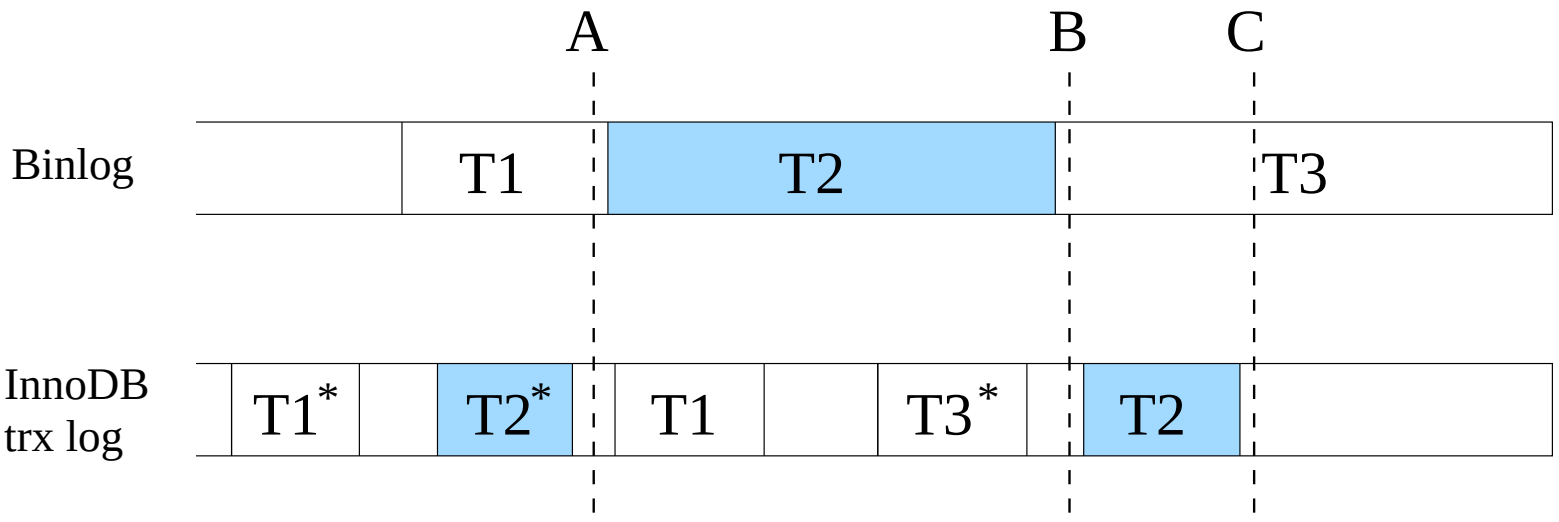

- Pluggable storage engines
  - Multiple transaction logs
- Both a blessing and a curse
  - Flexibility
  - Extra cost in complexity and performance
- After crash, must ensure that all logs have the same transactions committed
- Done using 2-phase commit

# 2-phase commit



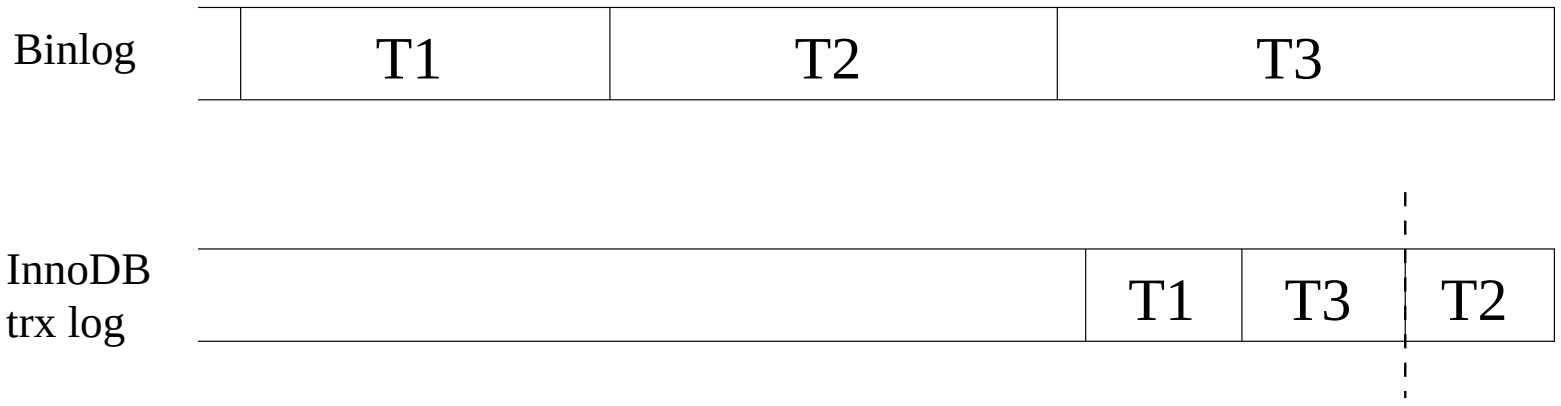

- First we **prepare** (∗) T2 in InnoDB
- Then we **commit** T2 in binlog
- Finally we **commit** T2 in InnoDB
- After crash we will rollback (A), commit (B), or do nothing (C) in InnoDB

# Problem: `fsync()` is expensive




- Problem: need **3** `fsync()` to disk per commit
- `fsync()` is expensive
  - Especially on traditional commodity hard disks
  - Also SSD or even with battery-backed-up RAID
- Solution: group commit
  - Write and `fsync()` many parallel transactions at once
  - Amortise the cost of `fsync()` over many commits.

# Consistent commit order

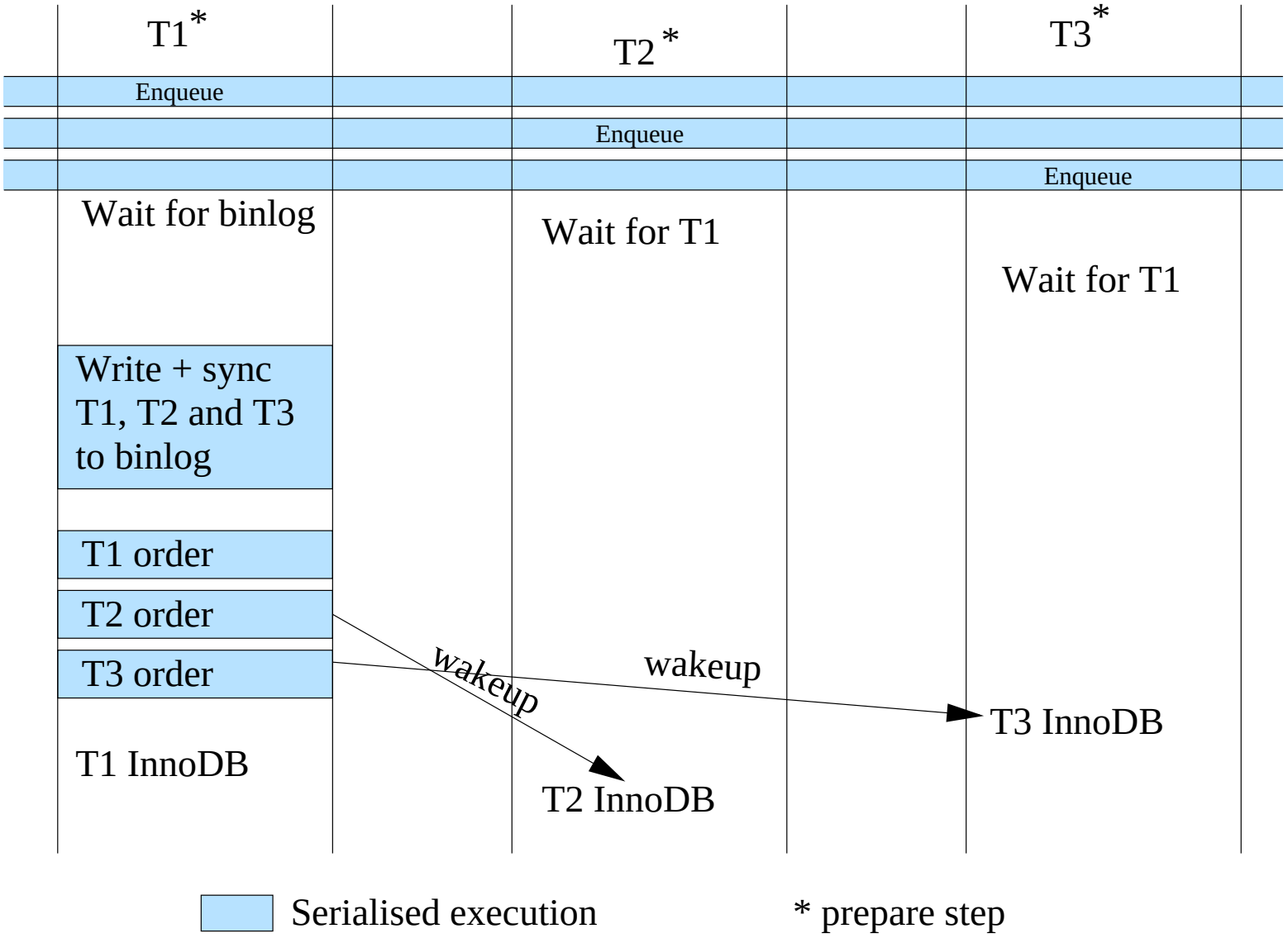



Need same commit order in diffent engines and in binlog

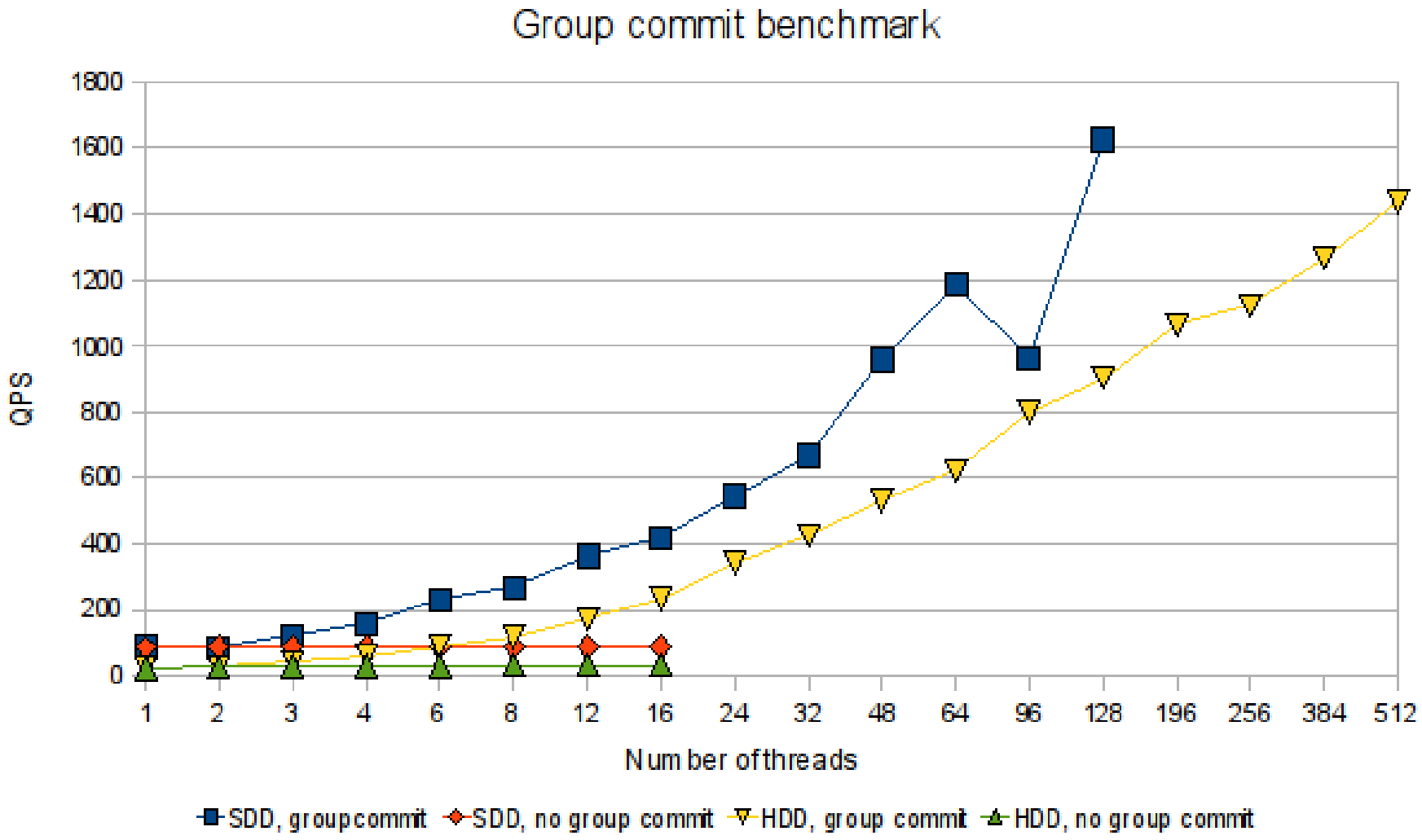
- Online backup takes snapshot of engine
- Could end with engine state that does not exist in binlog

Must coordinate commit across all engines and binlog

# The new commit algorithm



# Group commit scales well




# Extension to storage engine API

Extend the storage engine API

- `prepare()`
  - Write prepared trx in parallel, with group commit
- `prepare_ordered()`
  - Called serially, in commit order
- `commit_ordered()`
  - Called serially, in commit order
  - Fast commit to memory only
- `commit()`
  - Commit to disk in parallel, with group commit

# Implementation

- Implement the enhanced commit algorithm in the server core
- Extend the storage engine API
- Implement the group commit in binlog
- Extend InnoDB/XtraDB and PBXT to support the new API

Additional goodies

- `START TRANSACTION WITH CONSISTENT SNAPSHOT`
- Non-blocking `mysqldump --master-data --single-transaction`

# Full scale development possible in MariaDB

- With MariaDB we have the framework to develop on the server code at all levels
- Group commit is a fairly intrusive feature
  - Touches core commit algorithm, binlog, storage engines
  - Not just a “fringe feature” patch
- MariaDB collects sufficient skillset and procedures to facilitate such development
  - Core developers for architecture review
  - Code review
  - Testing/QA
  - Code maintenance/merging

# Other upcoming features

- Subqueries
  - Extension of the query optimiser
  - Greatly improved execution of big class of subqueries
- Microsecond support
- SQL query annotations in row-based replication
  - For debugging replication problems
- ...

# Outline



# Why MariaDB?

“Why is it called **Maria**DB?”

# Why MariaDB?

“Why is it called **Maria**DB?”

- Monty has two daughters
  - One is called My. . .

# Why MariaDB?

"Why is it called **Maria**DB?"

- Monty has two daughters
  - One is called My...
  - ...and the other is called Maria
- (And the MySQL trademark is owned by Oracle now)
- (And Monty btw. also has a son called Max
  - and yes, there used to be a "MaxDB" as well)

# Has/will Oracle killed MySQL?

“Has/will Oracle killed MySQL?”

# Has/will Oracle killed MySQL?

“Has/will Oracle killed MySQL?”

- No

# Has/will Oracle killed MySQL?

“Has/will Oracle killed MySQL?”

- No
- In the first year of MySQL@Oracle, things have been pretty much the same as always
- I think it is highly unlikely that they will, at least in the short to mid term
- Long-term, the question remains if Oracle will be motivated to improve MySQL towards eating into its other lucrative database offering
  - One important role of MariaDB is to provide such motivation

# Will Oracle embrace Open Source

“Has Oracle embraced and improved the Open Source nature of MySQL??”

# Will Oracle embrace Open Source

“Has Oracle embraced and improved the Open Source nature of MySQL??”

- No

# Will Oracle embrace Open Source

"Has Oracle embraced and improved the Open Source nature of MySQL??"

- No
- In Sun, there were plans and progress for opening up development, even if they were slow
- In Oracle, such plans and progress are just gone
  - and no signs that this will change
- Things are still much the same as one year ago, but everyone seems to be just waiting for Oracle to start closing up

My personal guess is that Oracle understands the need to thread carefully to avoid alienating the community (too much)

# Should I switch to PostgreSQL?

“Should I switch to PostgreSQL?”

# Should I switch to PostgreSQL?

“Should I switch to PostgreSQL?”

- No

# Should I switch to PostgreSQL?

“Should I switch to PostgreSQL?”

- No . . . , not unless you want to
- The MySQL server, under whatever name, remains as viable a long-term solution as it always was
- Independent support and development communities are available
- Still, PostgreSQL seems to be getting some extra interest, something that I personally welcome

# Should I switch to MariaDB?

“Should I switch to MariaDB (or Percona Server)?”

# Should I switch to MariaDB?

"Should I switch to MariaDB (or Percona Server)?"

- You can, though most likely there is no pressing need just yet
  - MySQL@Oracle is still under active development
  - Most distros and hosting providers etc. probably still default to MySQL
- For heavy users, it makes sense to evaluate the extra features of MariaDB (and Percona Server) for the next server upgrade
- I you want to switch, we encourage it and welcome feedback!

# Did Oracle do anything good?

“Did Oracle do anything good for MySQL?”

# Did Oracle do anything good?

“Did Oracle do anything good for MySQL?”

- Yes

# Did Oracle do anything good?

“Did Oracle do anything good for MySQL?”

- Yes
- Many good improvements to InnoDB released in MySQL 5.5
  - (Oracle owned InnoDB since 2005)
- Focus on multi-core scalability improvements in the server

# Did Oracle do anything bad?

“Did Oracle do anything bad for MySQL?”

# Did Oracle do anything bad?

“Did Oracle do anything bad for MySQL?”
There are at least some bad signs

- Backup project killed
  - Proprietary InnoDB backup rebranded as MySQL Enterprise Backup
- Foreign key project killed
- Client exception removed from source tree
  - But re-appeared on web site
  - Can it be retro-actively revoked?
- “Rumors”:
  - Moving to non-public bugtracker?
  - Less trees publicly mirrored on Launchpad?
  - Remove commit mails?
- Very little communication with community
  - Feel we are seen as enemies rather than potential partners

# I do not like Oracle!

“I do not like Oracle!”

# I do not like Oracle!

“I do not like Oracle!”

Then go with one of the alternatives:

- MariaDB (or Percona Server) for binaries and source code
- SkySQL (or lots of others) for L1 and L2 support
  - Backed by core developers from Monty Program for L3 support
- Percona for consulting

# How can I participate?

“How can I participate?”

- `maria-developers@lists.launchpad.net`
  `https://launchpad.net/~maria-developers`
- FreeNode IRC: `#maria`

# Outline



# Conclusion

- There is a lot more to MySQL than just MySQL
- MariaDB is available and mature, try it!
- The MySQL server is as alive and well as ever
- Will be interesting to see what the future brings

MariaDB:

`http://mariadb.org/`

Slides:

`http://knielsen-hq.org/maria/osd2011.pdf`

Contact:

`knielsen@knielsen-hq.org`