



PERCONA
Performance Consulting Experts

State of Percona Server 5.5

Date, time, place:

Percona Live SF
16-Feb-2011

Reporter:

Vadim Tkachenko
Co-founder, CTO,
Percona Inc

Who is speaker

CTO and co-founder of Percona Inc.

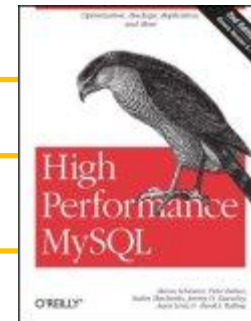
Lead of Percona Server/XtraDB and XtraBackup development

Co-author of “High Performance MySQL” 2nd edition

Author on MySQLPerformanceBlog.com and SSDPerformanceBlog.com

More reading on <http://www.percona.com/about-us/our-team/vadim-tkachenko/>

Email: vadim@percona.com

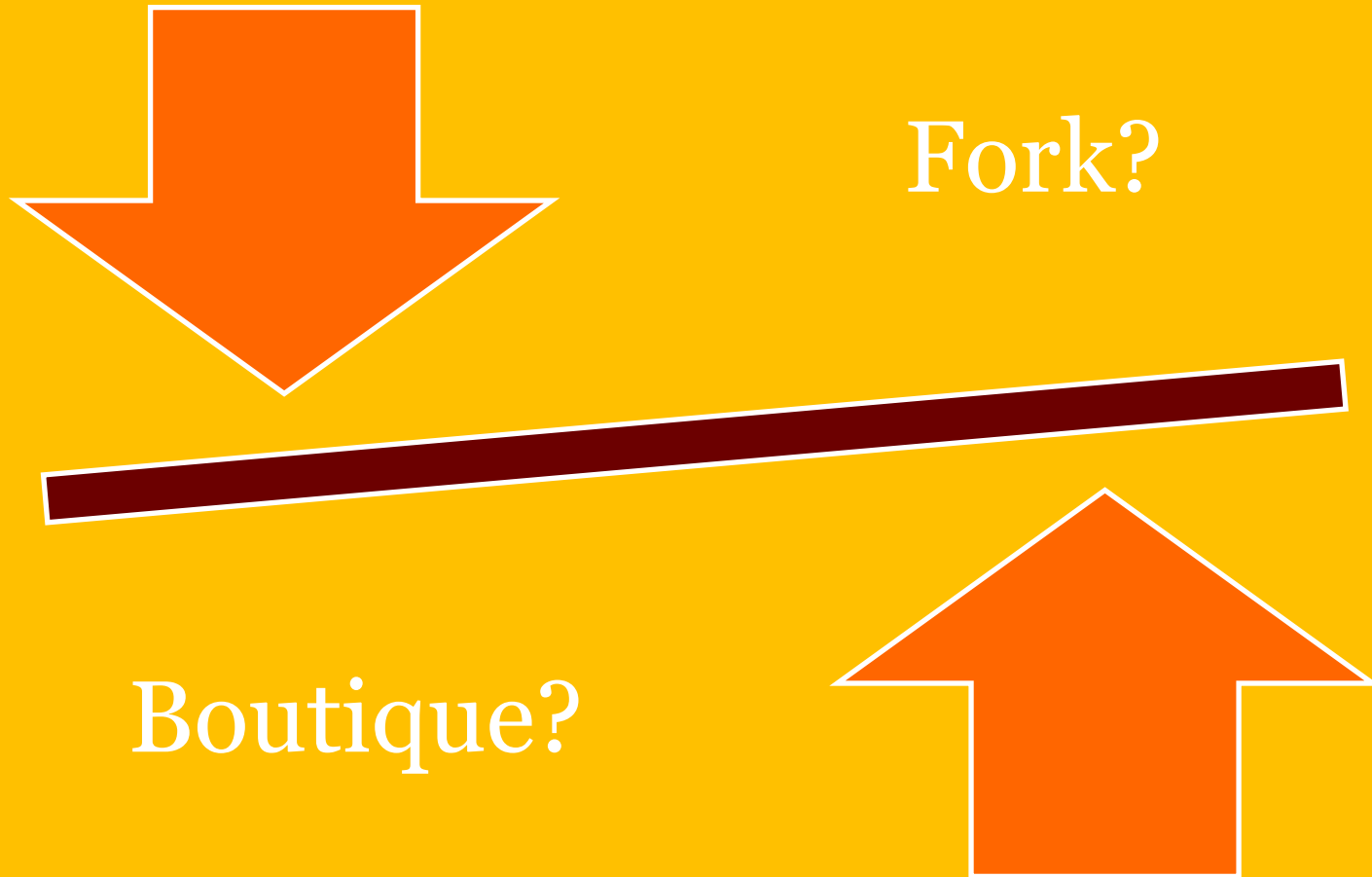


Percona Server 5.5

It is there

- Percona Server 5.5-20 Beta
- <http://percona.com/downloads/Percona-Server-5.5/>

So what is Percona Server ?



So what is Percona Server ?

Free + OpenSource

Like free
beer

Like
freedom

Still old good MySQL



... with high performance tuning



Percona Server

Backwards-compatible replacement for MySQL

Much faster and more scalable

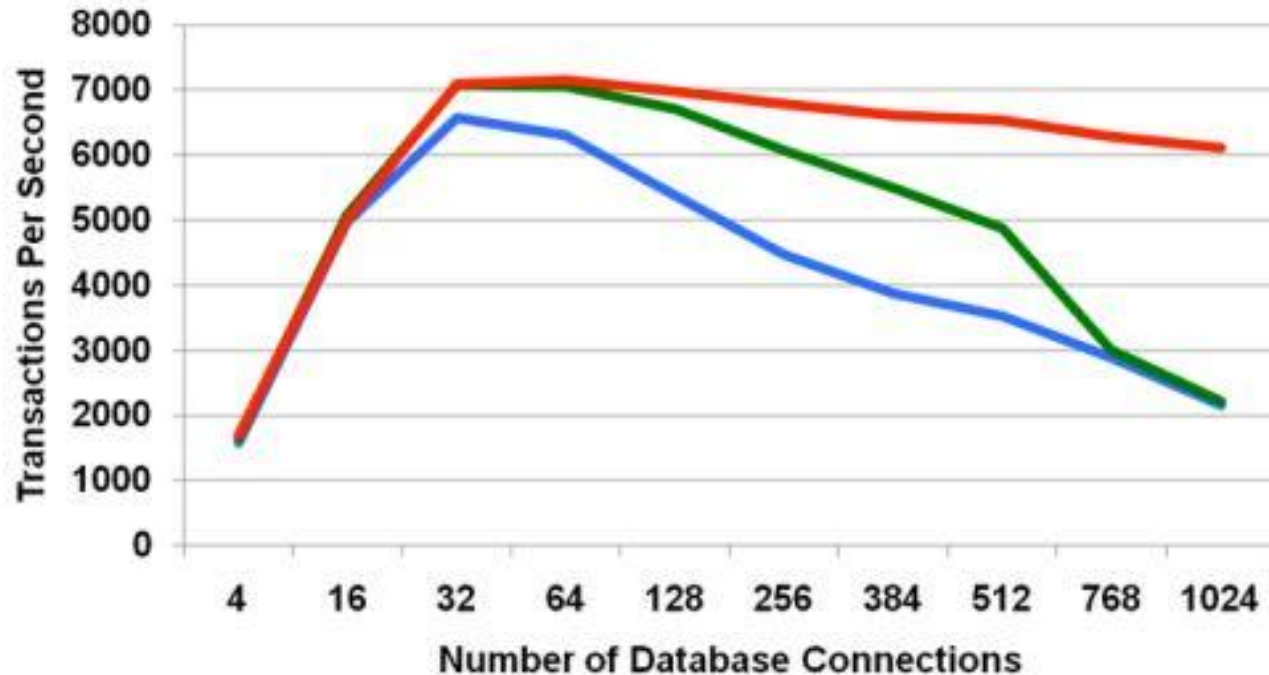
Easier to monitor and tune

Features to make operational tasks easier.

5.5

What's under the hood

All good stuff from MySQL 5.5



MySQL 5.5.6
(New InnoDB)

MySQL 5.1.50
(InnoDB Plug-in)

MySQL 5.1.50
(InnoDB built-in)

graph is taken from public Oracle / MySQL presentations

plus

IO/Flash/SSD

- Adaptive flushing
- Big log files
- Doublewrite configuration

plus

IO/Flash/SSD

- Adaptive flushing
- Big log files
- Doublewrite configuration

Multi
CPU/Cores/Big
memory

- Fix for adaptive hash read-write lock
- Buffer pool scalability

plus

IO/Flash/SSD

- Adaptive flushing
- Big log files
- Doublewrite configuration

Multi CPU/Cores/Big memory

- Fix for adaptive hash read-write lock
- Buffer pool scalability

Performance diagnostic

- Variables in SHOW STATUS
- Enhanced SLOW LOG

plus

IO/Flash/SSD

- Adaptive flushing
- Big log files
- Doublewrite configuration

Multi CPU/Cores/Big memory

- Fix for adaptive hash read-write lock
- Buffer pool scalability

Performance diagnostic

- Variables in SHOW STATUS
- Enhanced SLOW LOG

Cloud

- Fast warmup

plus

IO/Flash/SSD

- Adaptive flushing
- Big log files
- Doublewrite configuration

Multi CPU/Cores/Big memory

- Fix for adaptive hash read-write lock
- Buffer pool scalability

Performance diagnostic

- Variables in SHOW STATUS
- Enhanced SLOW LOG

Cloud

- Fast warmup

SaaS/Shards

- Export/Import InnoDB tables
- Partial backup / restore

plus

IO/Flash/SSD

- Adaptive flushing
- Big log files
- Doublewrite configuration

Multi CPU/Cores/Big memory

- Fix for adaptive hash read-write lock
- Buffer pool scalability

Performance diagnostic

- Variables in SHOW STATUS
- Enhanced SLOW LOG

Cloud

- Fast warmup

SaaS/Shards

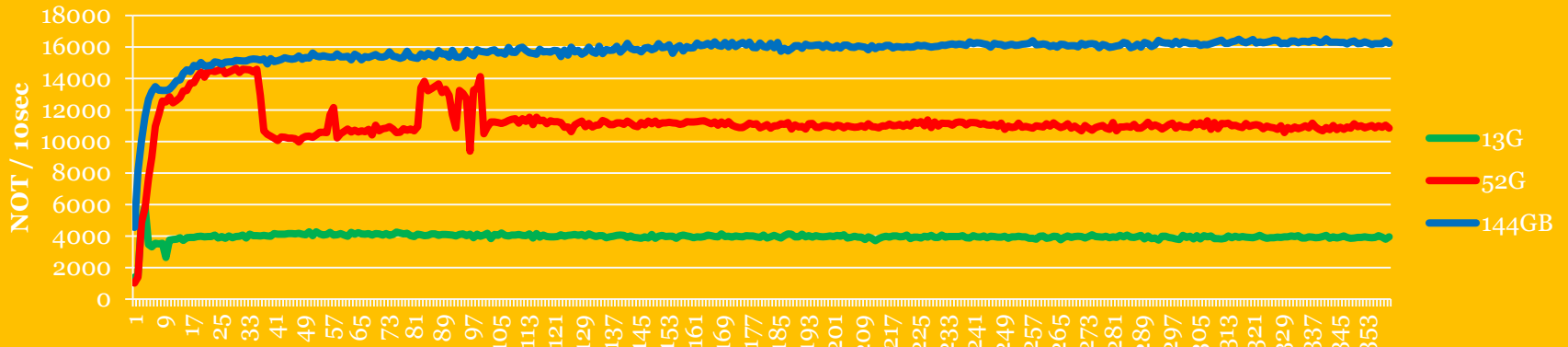
- Export/Import InnoDB tables
- Partial backup / restore

NOSQL

- HandlerSocket plugin

Details ?

tpcc-mysql 1000W, Virident tachION

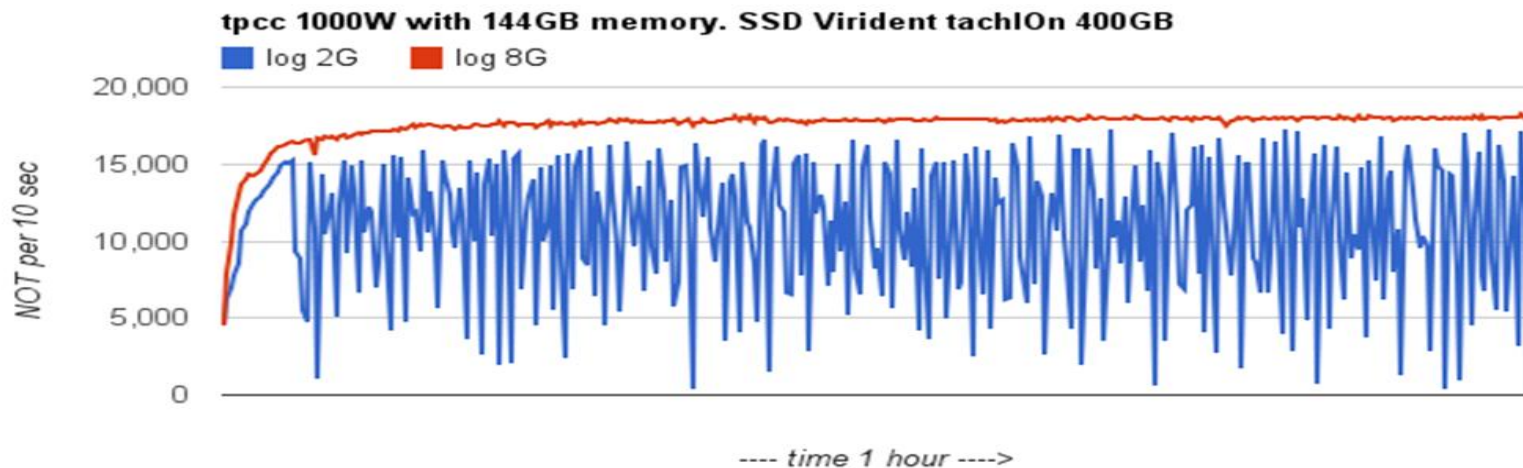
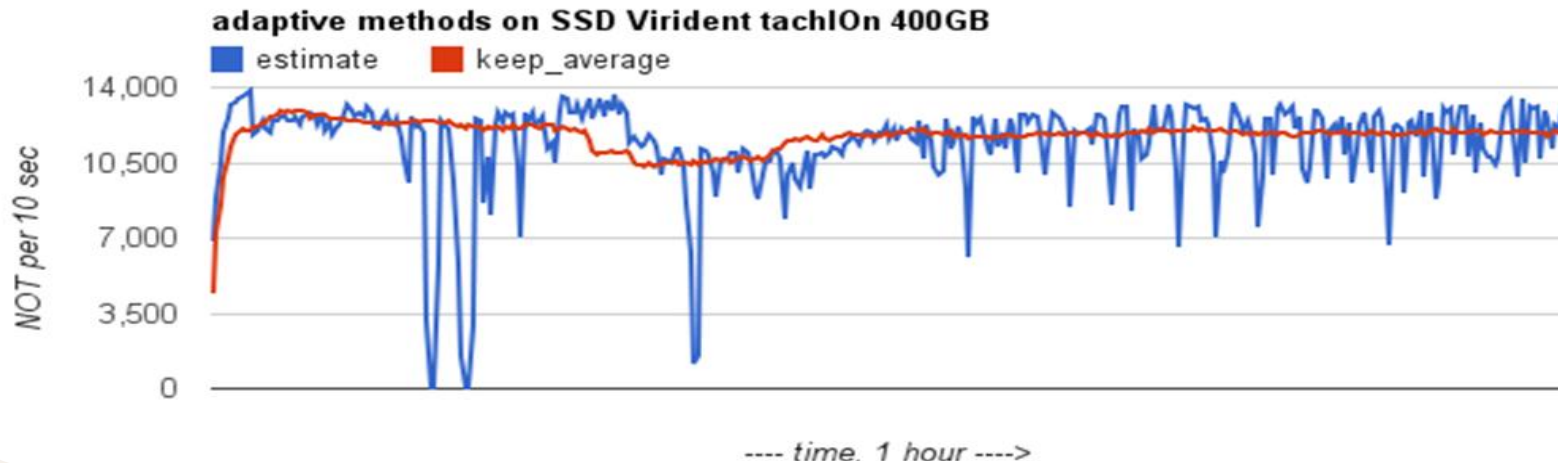


Adaptive_flushing_method
= keep_average

Innodb_log_file_size
>4G

innodb_log_block_size
=4096

More benchmarks



More SSD goodies

`innodb_read_ahead = none |
random | linear | both`

`innodb_flush_neighbor_pages
= 0 | 1`

Scalability

Problem:
Adaptive_hash

- Global rw_lock for all indexes

Solution

- Hash, one entry per index
- Innodb_adaptive_hash_partitions

Fast warmup

Problem

- It takes hours/days to get data into `buffer_pool` after restart

Solution

- Preload stored content of `buffer_pool` after restart

Export / Import

Problem

- Impossible to move single InnoDB database/server to another server

Solution

- XtraBackup export + Percona Server import

SHOW ENGINE INNODB STATUS

- Pending flushes (fsync) log: 0; buffer pool: 1
- 246400 OS file reads, 20256 OS file writes, 1261 OS fsyncs
- 9 pending preads, 0 pending pwrites
- 12514.15 reads/s, 16384 avg bytes/read, 698.83 writes/s, 69.79 fsyncs/s
- -----
- INSERT BUFFER AND ADAPTIVE HASH INDEX
- -----
- Ibuf: size 219, free list len 117, seg size 337, 4787 merges
- merged operations:
- insert 4923, delete mark 0, delete 0
- discarded operations:
- insert 0, delete mark 0, delete 0
- Hash table size 27624761, node heap has 360 buffer(s)
- 2780.72 hash searches/s, 27576.24 non-hash searches/s
- ..
- Pending reads 11
- Pending writes: LRU 0, flush list 0, single page 0
- Pages made young 170, not young 0
- 8.70 youngs/s, 0.00 non-youngs/s
- Pages read 246248, created 10418, written 658
- 12516.15 reads/s, 541.25 creates/s, 55.59 writes/s
- Buffer pool hit rate 921 / 1000, young-making rate 0 / 1000 not 0 / 1000
- Pages read ahead 0.00/s, evicted without access 0.00/s
- LRU len: 256677, unzip_LRU len: 0
- I/O sum[0]:cur[3871], unzip sum[0]:cur[0]
- Main thread process no. 13756, id 1288038720, state: sleeping
- Number of rows inserted 77848, updated 155479, deleted 5956, read 667829
- 4065.39 inserts/s, 8129.79 updates/s, 311.77 deletes/s, 35000.80 reads/s

SHOW STATUS

Innodb_adaptive_hash_cells	Innodb_checkpoint_max_age	Innodb_ibuf_merged_deletes	Innodb_master_thread_background_loops
Innodb_adaptive_hash_heap_buffers	Innodb_checkpoint_target_age	Innodb_ibuf_merged_inserts	Innodb_master_thread_main_flush_loops
Innodb_adaptive_hash_hash_searches	Innodb_deadlocks	Innodb_ibuf_merges	Innodb_master_thread_sleeps
Innodb_adaptive_hash_non_hash_searches	Innodb_dict_tables	Innodb_ibuf_segment_size	Innodb_max_trx_id
Innodb_background_log_sync	Innodb_history_list_length	Innodb_ibuf_size	Innodb_mem_adaptive_hash
Innodb_buffer_pool_pages_LRU_flushed	Innodb_ibuf_discarded_deleted_marks	Innodb_lsn_current	Innodb_mem_dictionary
Innodb_buffer_pool_pages_made_not_young	Innodb_ibuf_discarded_deletes	Innodb_lsn_flushed	Innodb_mem_total
Innodb_buffer_pool_pages_made_young	Innodb_ibuf_discarded_inserts	Innodb_lsn_last_checkpoint	Innodb_mutex_os_waits
Innodb_buffer_pool_pages_old	Innodb_ibuf_free_list	Innodb_master_thread_1_second_loops	Innodb_mutex_spin_rounds
Innodb_checkpoint_age	Innodb_ibuf_merged_delete_marks	Innodb_master_thread_10_second_loops	Innodb_mutex_spin_waits
Innodb_purge_undo_no	Innodb_s_lock_os_waits	Innodb_s_lock_spin_rounds	Innodb_oldest_view_low_limit_trx_id
Innodb_row_lock_numbers	Innodb_x_lock_os_waits	Innodb_s_lock_spin_waits	Innodb_purge_trx_id
Innodb_x_lock_spin_rounds	Innodb_x_lock_spin_waits		

INFORMATION_SCHEMA

Buffer_pool content

System tablespace content

InnoDB optimizer statistics

USER / THREAD / INDEX statistics

Query response time histogram

SHOW Temporary tables

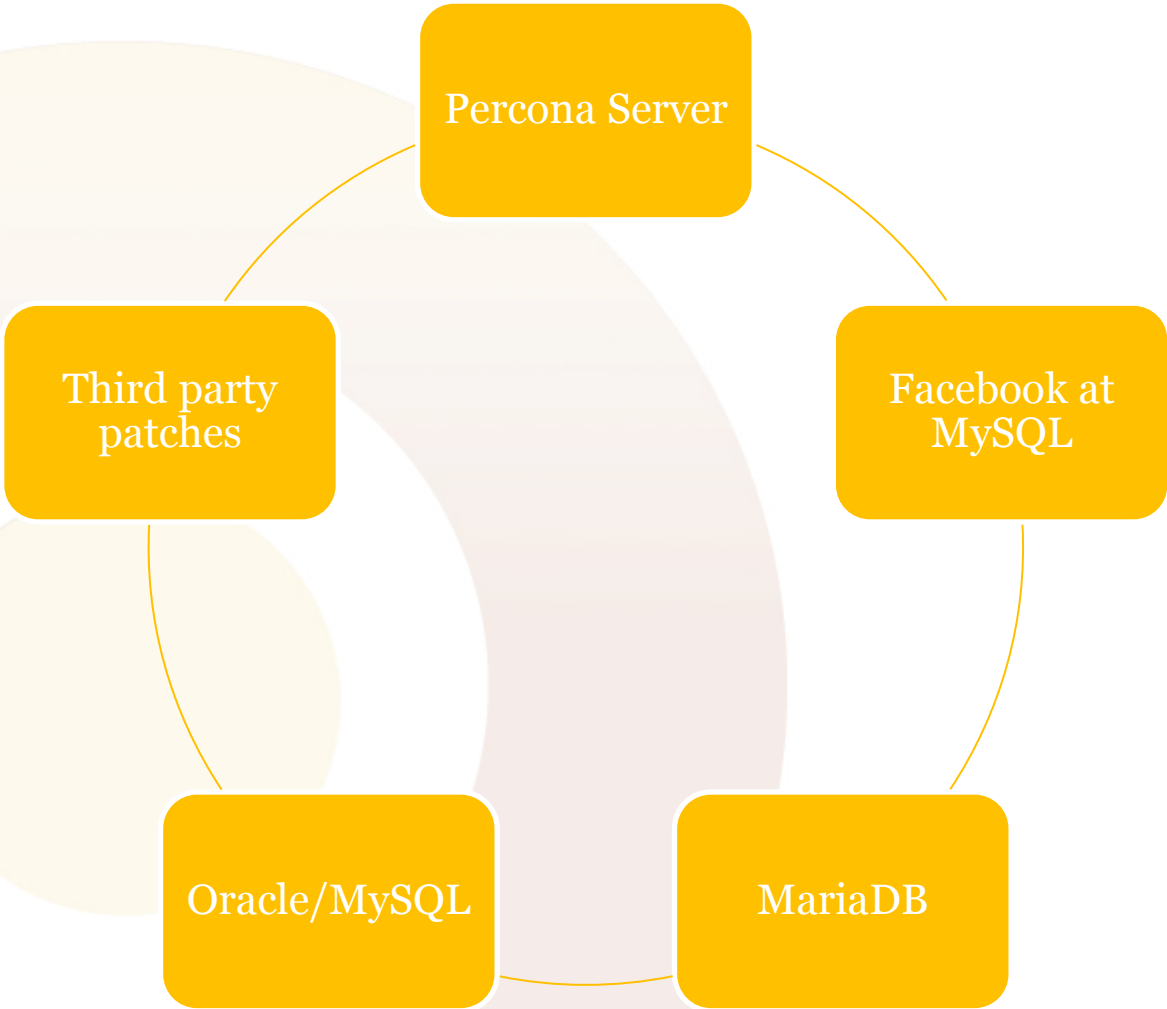
HandlerSocket

Direct access to data

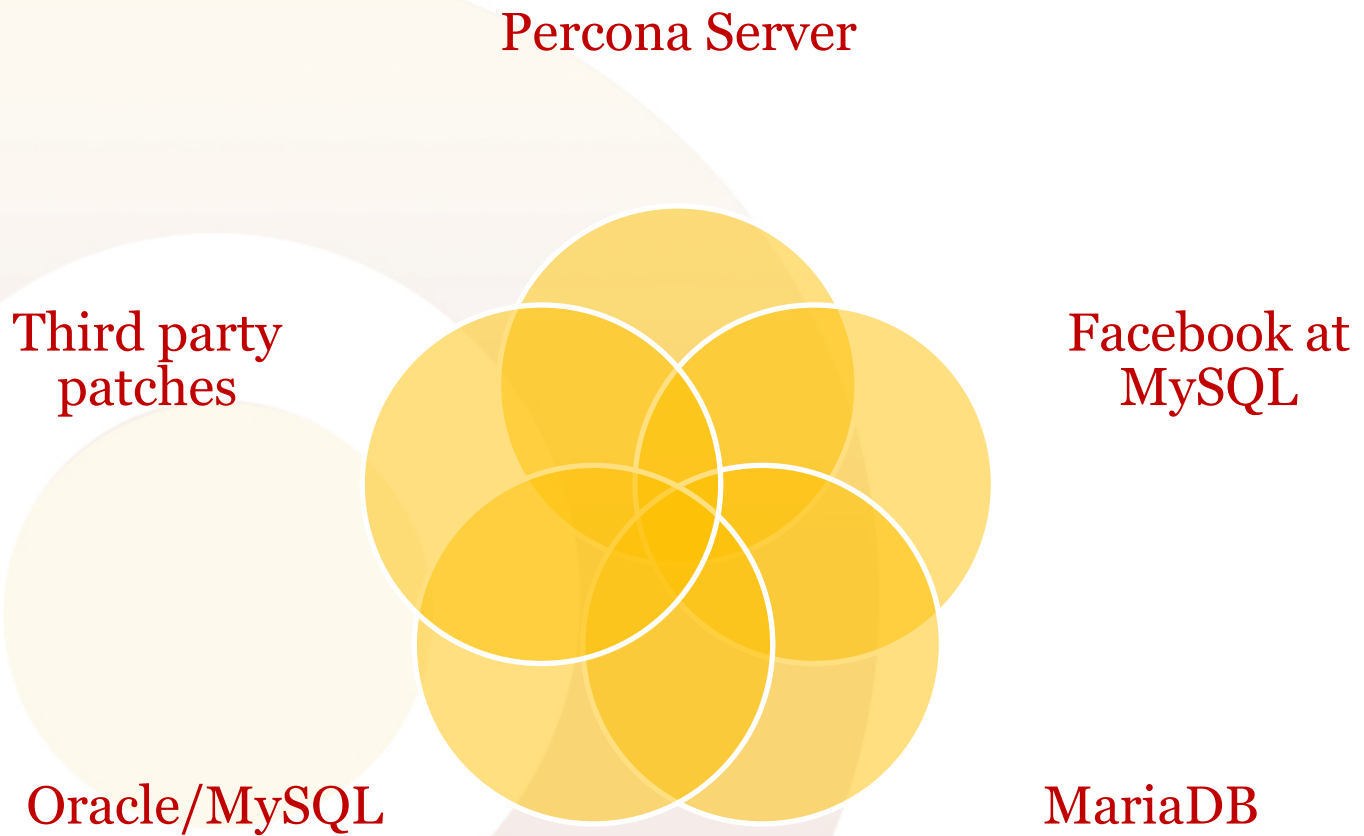
Avoid query parser

Avoid global mutexes

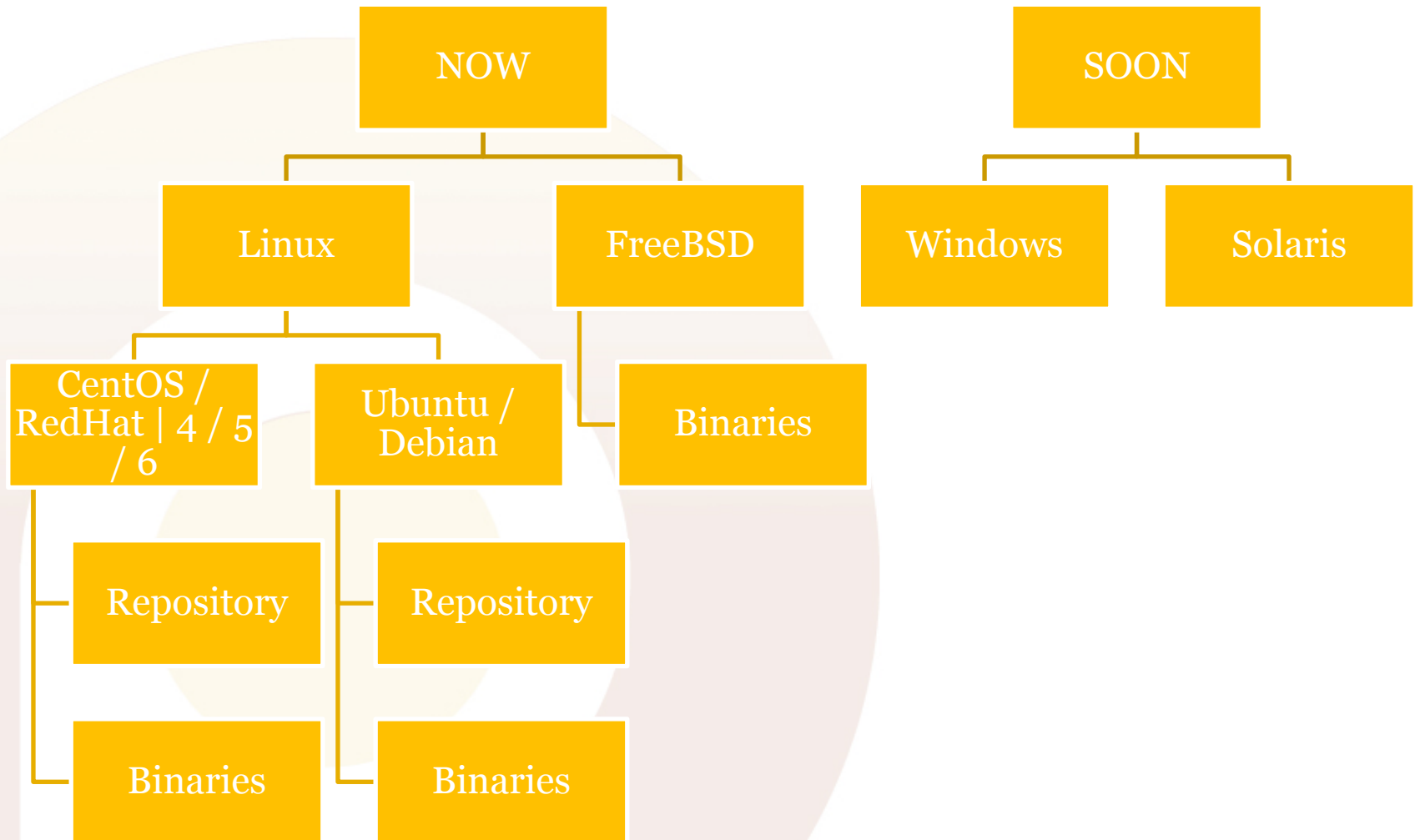
Relations



Relations



Supported platforms



When Percona Server 5.5 is stable

SOON

- Quality and testing takes time

But really when?

- O'Reilly MySQL Conference & Expo is good target

What is in future ?

Better performance and scalability

Better Cloud support

Better Flash/SSD support

Better Diagnostics

Support and maintenance?

Can I get help?

- Yes
- <http://www.percona.com/support/>

Can I get my bug fixed?

- Yes
- <http://www.percona.com/support/>

Can I get free answer?

- Yes
- <http://forum.percona.com/>
- only as time permits

Pictures

- <http://www.cartype.com/pages/1758/vespa>
- <http://www.flickr.com/photos/27587130@No2/2699140647/>

Thank you!