



PERCONA
Performance Consulting Experts

Over 20,000 QPS, the xtradb performance show

Date, time, place:

FOSDEM 2011
Brussels, BE Feb 5-6

Reporter:

Istvan Podor,
Percona

Introduction

- Currently a Percona consultant
- Several alexa top500 websites
- Driven by Performance is everything principal

Why show?

- Can't publish exact numbers
- Don't even have the exact numbers
- Got 20 minutes
- More like a tale, less technical

Introduction - about the talk

- Alexa top500 (1m+ unique, few hundred million hit / day)
- I got hired as a LAMP performance engineer
- Really fast development
- (almost) impossible demands
- Large, resource demanding queries
- Extreme peak traffic, suddenly

The environment

- Really strong machines as masters
- Bottlenecks are the slaves, serving too few queries
- Going to talk about the slaves now
- Mysql I'm talking about was running almost on default factory settings

Not covering here

- Linux/network/kernel etc
- Schema changes
- Query optimization

The plan?

- Start working hard for months as I should to rewrite queries and optimize everything? :/
- Slow and stressful

The Plan!

- Install Percona Server! (aka: mysql with xtradb)
- Tune it
- As a fan of mysqlperformanceblog.com, obvious plan to come up with :)
- Measuring by LC balance algorithm
 - // part of the show

Installing xtradb

- Until the final result, I used only one server
- Things changed just by installing it
 - Higher CPU usage
 - Lower IO wait
- I got questioned, is that a good thing?

Installing xtradb

- Yes it is! Why?
- The server was spent much less time on waiting on disks, utilizing CPUs much more. How that could be bad? :)
- CPU IO wait dropped to 20-40% from 40-80%
- We were instantly gaining ~10% 'performance'
 - (remember the Least-Connection measure)

Some schema changes anyhow

- Hey, this thing works! :)
- All 'secondary' (logging/statistics) altered to myisam
- All 'primary' tables altered to InnoDB
- 'instabilize' and tune MyISAM
 - Disable query cache (we don't need that)
 - Concurrent inserts (no worries on secondary tables)
 - Delayed index write (we don't care about these)

Some schema changes anyhow

- Schema changes doesn't come with noticeable performance improvement
- But better memory utilization regarding to 'important' tables

'Force' InnoDB into memory

- Increase buffer pool (`innodb_buffer_pool_size`)
- Data dictionary to memory (`innodb_additional_mem_pool_size`)
- Delay flushing changes as long as possible
 - Log file size (`innodb_log_file_size`)
 - Log buffer size (`innodb_log_buffer_size`)

'Force' InnoDB into memory

- Comes with another significant drop in IO usage
- Also speeded up query execution time
- Another ~10-15% improvement

'get rid of' isolation

- **Flush to disk directly** (`innodb_flush_method = O_DIRECT`)
- **Flush logs once in a while** (`innodb_flush_log_at_trx_commit`)
- **Lower isolation level** (`innodb_trx_isolation`)
- **Disable XA** (`innodb_support_xa`)
- **Increase purge lag** (`innodb_max_purge_lag`)
- **And dirty pages** (`innodb_max_dirty_pages_pct`)

'get rid of' isolation

- Not that significant improvement after the previous changes (just an additional few percents, ~15%)
- Disk IO wait dropped down to 10-20%

Tune xtradb

- Increase concurrent thread works
 - Threads concurrency (innodb_thread_concurrency)
 - Commit concurrency (innodb_commit_concurrency)
- Increase concurrent IO works
 - File IO threads (innodb_file_io_threads)
 - Write threads (innodb_write_io_threads)
 - Read threads (innodb_read_io_threads)
- Increase IO capacity
 - Hardcoded by default, < 5.5 (innodb_io_capacity = $200 * \#$ of disks)

Tune xtradb

- These changes came with no improvement
- ... in normal workflow
- But under heavy load, the server performed much more better in compare to the others
- Disks wasn't loaded with unnecessary operations

What's happened?

- One server were able to replace 3 other
- High percent of IOwait in CPU utilization gone
- ~50% improvement in normal workload
- ~300%+ improvement in high workload
- Even fast response time when taking all the load of 3 servers
- New record of QPS/Server > ~21,600 (against ~8000)
- Linux load from 3,5-6 dropped to 0.8-1,6

What if a server failed?

- Rebuild from backup, no worries
- (or just simply take down another one, and SCP the whole data dir)

Thoughts

- Specific environment, numbers probably would be different for you
- Risking slaves integrity
- In case of a full power outage, we would be in huge trouble
- Forced to do this way by all the circumstances
- It's not just that easy, xtradb != miracle :)

Questions/Comments?

- Percona Inc. :: www.percona.com
- Performance blog :: www.mysqlperformanceblog.com
- Contact in case of questions or discussions
 - istvan(dot)podor at percona . com