



PERCONA
Performance Consulting Experts

~~Still stuck with MySQL 5.0?~~
Afraid to upgrade? What's new in InnoDB-
plugin/5.1, InnoDB-plugin/5.5, XtraDB

Date, time, place:

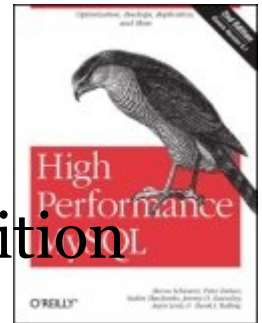
SF MySQL meetup, 14 Dec 2010

Reporter:

Vadim Tkachenko

Formal part

- Who am I?
 - 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-tkachev>
 - vadim@percona.com



What is this talk about ?

- ~~Promotion of Percona Server~~
- What is the InnoDB-plugin?
 - Advantages and benefits
- What is XtraDB?
 - Advantages and benefits
- What should you know about upgrading InnoDB?
- Does not cover general MySQL 5.1 / 5.5 features

Naming mess

- InnoDB
- InnoDB-plugin
- XtraDB
- InnoDB+

InnoDB

- Traditional transaction engine for MySQL
- Exists in MySQL 5.0 and MySQL 5.1
 - We name it InnoDB-builtin, -standard, -native, etc
- Very old and no new features for a long time

Let's agree on

- InnoDB is dead
 - No new features or performance fixes in InnoDB between MySQL 5.0 and MySQL 5.1
 - InnoDB-plugin is the default engine in MySQL 5.5

Reasons to use InnoDB in 5.1

- Not much
 - InnoDB-plugin is greatly superior
- The reasons
 - Server is less powerful than my laptop
 - Running “my homepage” website with 2 visitors per month
 - Your boss does not understand difference and does not allow to use InnoDB-plugin
- Exaggerating, but to show this point:
 - InnoDB is outdated

InnoDB-plugin

- Shipped with MySQL 5.1
 - Not enabled by default
- Default engine in MySQL 5.5
- To add more mess:
 - I will use name InnoDB when I mean InnoDB-plugin
 - Old-InnoDB will be “old”, “builtin”, “standard”
- InnoDB+
 - Rumors based closed source version of InnoDB

XtraDB

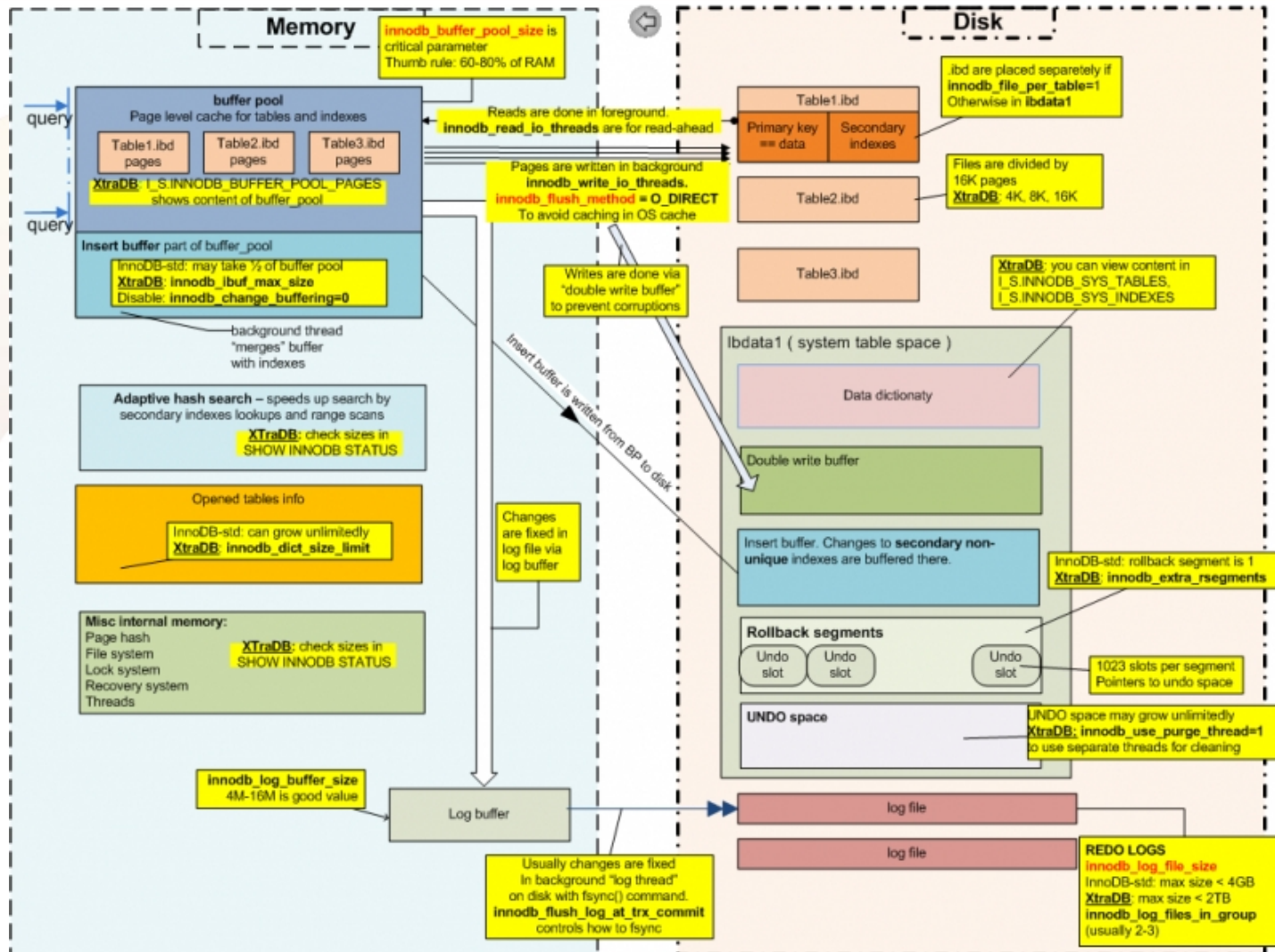
- InnoDB-plugin with enhancements from Percona
- Comes with Percona Server 5.1
 - No additional steps to enable it
 - Old-InnoDB is not available

Overview of new features

- I/O subsystem fixes
 - Most important improvements
 - InnoDB-plugin uses XtraDB implementation
- Hardware rw_locks
- Buffer_pool scan resistant for scalability
 - `SELECT * FROM big_table`, `mysqldump` are not dangerous
- Fast index creation
 - Good idea, but not fully implemented
- Compressed tables
 - Good idea, but questionable implementation
- Fast recovery

Quick look on InnoDB internals

<http://www.mysqlperformanceblog.com/2010/04/26/xtradb-innodb-internals-in-drawing/>



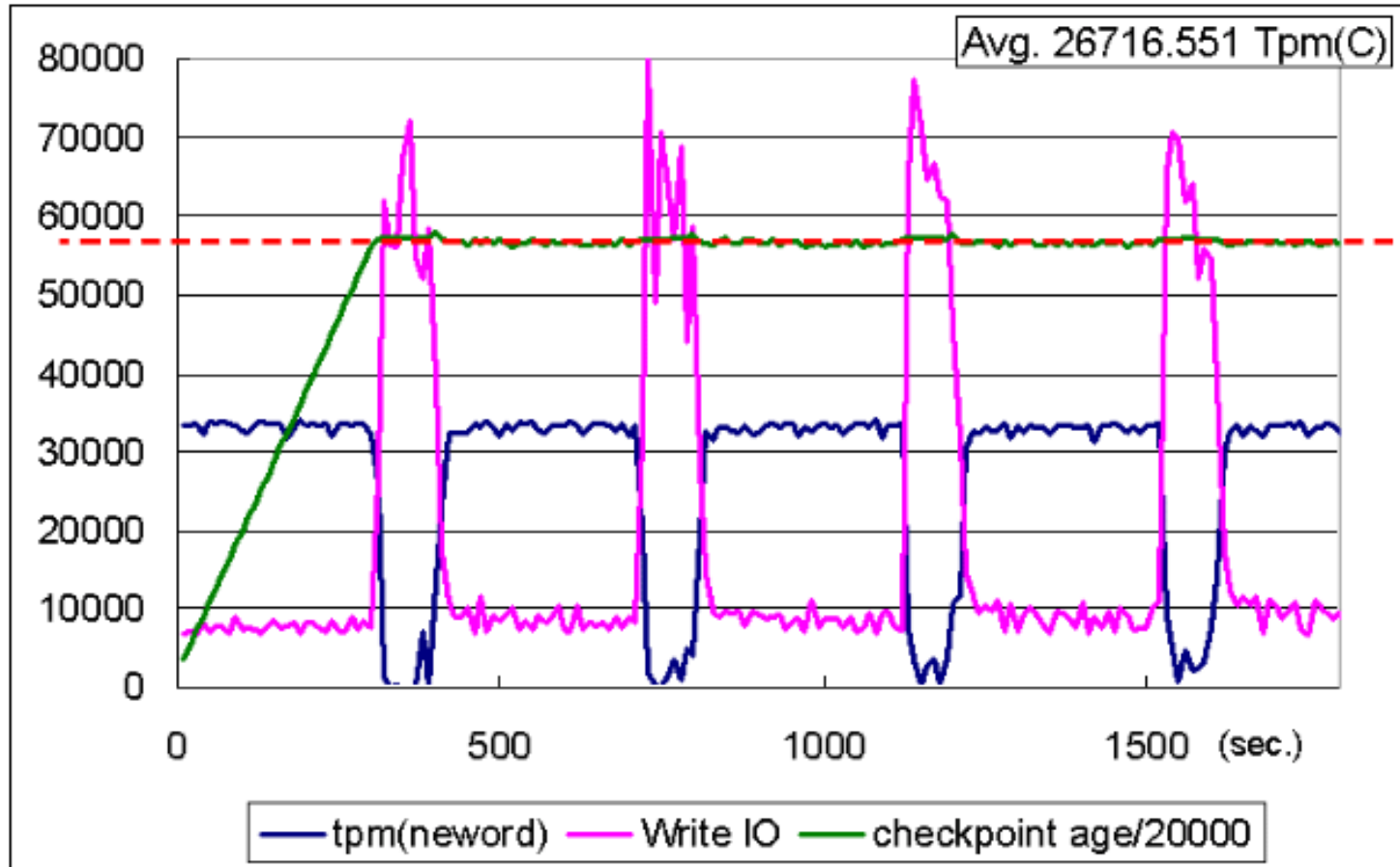
I/O subsystem

- Multiple read and write threads
 - innodb_read_io_threads
 - Affects only read-ahead requests, may keep default
 - innodb_write_io_threads
 - How many threads perform writes to disk
 - Good to understand what your I/O system allows
 - Single disk: 1
 - for RAID10 over 8 disks: 4 is good values
 - SSD PCI-E cards: 8-16 range

I/O subsystem

- `innodb_io_capacity`
 - Common understanding is that this is the amount of I/O's, but this assumption is wrong
 - Very easy to mistune
 - 200 is default, and is good for most
 - 400-500 may also work
 - High values may saturate your I/O system

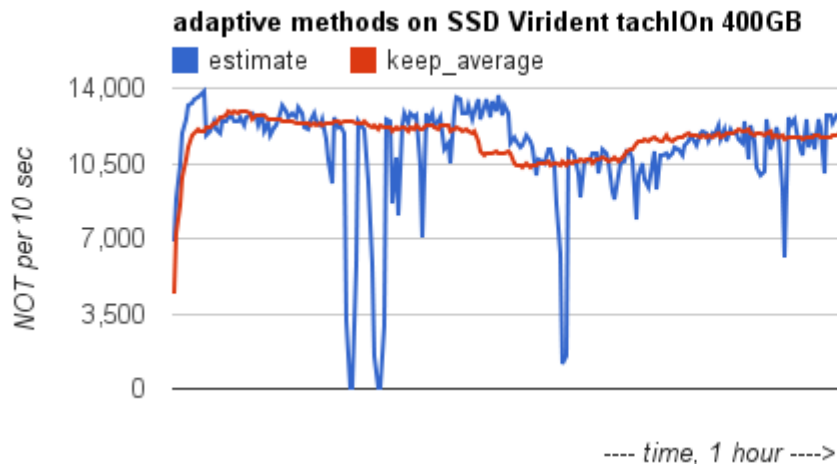
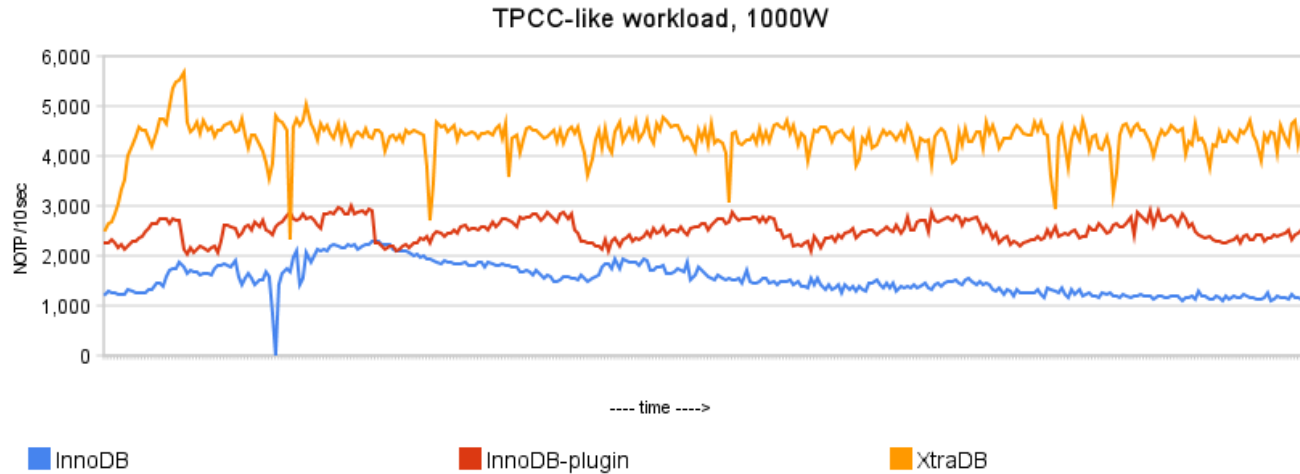
Adaptive stuff



Adaptive settings

- InnoDB (plugin)
 - `innodb_adaptive_flushing` = ON (default) | OFF
- XtraDB
 - `innodb_adaptive_checkpoint`=none|reflex|estimate|keep_average
 - Don't ask what this means
 - `keep_average` just added, not released

Adaptive cont

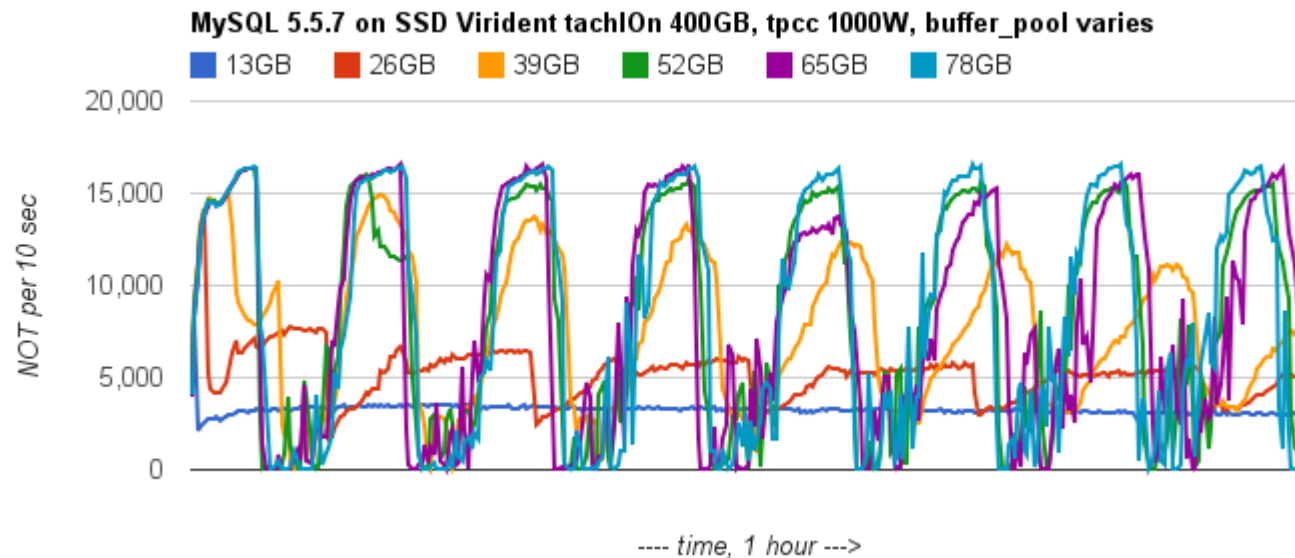


Adaptive – bottom line

- Whatever you use InnoDB-plugin or XtraDB for, you may not need to tune it 😊

But

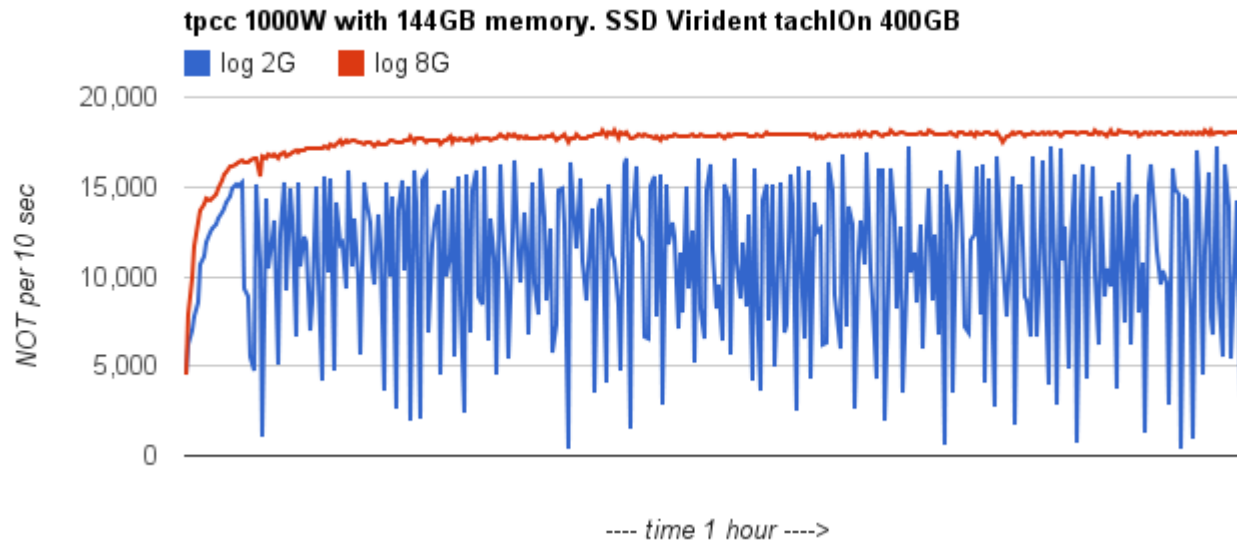
- MySQL 5.5 adaptive is really bad on SSD :



Misc. I/O Enhancements

- InnoDB-plugin
 - Read_ahead
 - “Random” removed
 - “linear” - innodb_read_ahead_threshold
 - Purge thread
 - Can be enabled
- XtraDB
 - Read_ahead
 - innodb_read_ahead = none | random | linear | both
 - none is recommended for SSD
 - innodb_flush_neighbor_pages = 0 | 1
 - 0 is recommended for SSD
 - innodb_log_file_size > 4GB

Log size effect



Still on I/O

- **insert_buffer**
 - Can take 1/2 of buffer_pool
- **InnoDB-plugin**
 - innodb_change_buffering = ON | OFF
- **XtraDB**
 - innodb_ibuf_max_size – maximum allowed size
 - innodb_ibuf_accel_rate, innodb_ibuf_active_contract
 - Don't ask. This is just to show how Percona is advanced

Hardware rw_locks

- Custom software implementation removed
- GCC atomic_builtin added
- Much more efficient on 8+ cores
- No config options
 - Can be disabled at compile time – there is no much sense to do it

Buffer_pool

- Scalability
 - XtraDB – just included, not tuning
 - InnoDB-plugin
 - innodb_buffer_pool_instances in MySQL 5.5
 - Tss... lazy way to ignore problem instead of real fix
 - Can be set 4 .. 8 .. 16 ... what is good setting is unknown – try and see
- Scan resistance
 - Now `SELECT * FROM big_table` will not kill your buffer_pool
 - Really it is just a bug fix to buffer pool logic, but announced as brand new feature ;-)

Buffer_pool settings

- Not much
- `innodb_buffer_pool_instances=X` in MySQL 5.5
 - X is from your experiments

FAST INDEX creation

- **ALTER TABLE t1 ADD KEY / CREATE INDEX idx ON t1**
 - Does not copy whole table as previously
 - Create index by sort and inserting in sorted order
 - Really much faster
 - Produces more compact indexes

FAST INDEX problems

- UTF-8 columns not handled (bug in MySQL)
- mysqldump does not support it
- OPTIMIZE TABLE / ALTER TABLE engine=InnoDB creates indexes in the old way
- Good for single index operation, pain if maintenance
- Not fully polished, reported bugs with .frm and .ibd out of sync
 - May be good idea to use on idle server
- Does not play well with Foreign Keys

COMPRESSED tables

- `ALTER TABLE t1 ENGINE=InnoDB
ROW_FORMAT=COMPRESSED
KEY_BLOCK_SIZE=4 | 8 | 16`
 - From MySQL docs: `KEY_BLOCK_SIZE=8` is usually a safe choice
- Good idea, but:
 - Doubles usage of `buffer_pool`
 - No scalability fixes for `buffer_pool` with compressed pages
- Use case:
 - Table with big BLOB/TEXT columns

Fast recovery

- Really just fix of annoying problem
- Recovery now is up to 10× faster

Misc. Enhancements

- Scalability of rollback segments
 - Automatic in MySQL 5.5
 - In MySQL 5.1/XtraDB in more limited form (requires reload of DB)
- `innodb_use_sys_malloc=ON` (default)
 - To use different memory allocator, built-in is bad one
 - Malloc by default, you can install tcmalloc and so on
- File formats
 - Antelope
 - Barracuda
 - Compression, dynamic row format
- Transactional Replication

Bottom line

- `innodb_buffer_pool_size=24G`
- `innodb_data_file_path=ibdata1:10M:autoextend`
- `innodb_file_per_table=1`
- `innodb_flush_log_at_trx_commit=2`
- `innodb_log_buffer_size=8M`
- `innodb_log_files_in_group=2`
- `innodb_log_file_size=2000M` # 4G for SSD like Virident , FusionIO
- `innodb_thread_concurrency=0`
- `innodb_flush_method = O_DIRECT`

- `#innodb_read_ahead = none` # for SSD
- `#innodb_flush_neighbor_pages = 0` # for SSD
- `#innodb_adaptive_checkpoint=keep_average` # for SSD

- `innodb_write_io_threads=8`
- `innodb_read_io_threads=8`
- `innodb_io_capacity=500`

Upgrade procedure to 5.1

- Quite simple
- Fully binary compatible
- Minimal actions for Percona Server/XtraDB
- For MySQL/InnoDB-plugin
 - ignore_builtin_innodb
 - plugin-load=innodb=ha_innodb_plugin.so;innodb_trx=ha_innodb_plugin.so;
 - innodb_locks=ha_innodb_plugin.so;innodb_lock_waits=ha_innodb_plugin.so;
 - innodb_cmp=ha_innodb_plugin.so;innodb_cmp_reset=ha_innodb_plugin.so;
 - innodb_cmpmem=ha_innodb_plugin.so;innodb_cmpmem_reset=ha_innodb_plugin.so
- Don't upgrade on Friday night or before Christmas

What version to use

- MySQL 5.5 should be really good
 - May be too risky to install earlier GA version
- MySQL 5.1.43+ are good, 5.1.52-53 looks pretty stable
 - 5.1.30 (14 November 2008 General Availability)
 - MySQL 5.1.43 (15 January 2010)
 - Took about 1 year to stabilize

Percona Server promotion

- **Percona Server 5.1.52-12.3**
 - HandlerSocket – NoSQL plugin to query InnoDB
 - Handles 200,000 req/sec from 2 clients
 - Only 5-7% busy, potential 1,000,000 req/sec
- **Parallel replication**
 - <http://www.percona.com/docs/wiki/percona-server:blueprint>
 - Looking for sponsors