How to Stop Hating MySQL: Fixing Common Mistakes and Myths

Sheeri K. Cabral
Database Team Lead
The Pythian Group, www.pythian.com

cabral@pythian.com
LISA 2008
Who I Am

- MySQL DBA
- MySQL User Group
- MySQL Podcast (OurSQL, on hiatus), videos (technocation.org)
- Lots of community stuff
Myths about MySQL

- Uses too much memory
- Slow
- I need more features!
More Myths about MySQL

- Don't use ENUM
- Schema changes take forever
- You have to restart to log
- No partitioning
MySQL uses too much memory!

- INFORMATION_SCHEMA.GLOBAL_VARIABLES (5.1)
- SHOW GLOBAL VARIABLES (5.0)
- LIKE
  - '%cache%'
  - '%buffer%'
Common Mistake: Wasting Memory

- bdb_cache_size
- bdb_log_buffer_size
- 32-bit operating system
Myth: Query Cache Is Faster!

- Use memcached for caching common queries

- `query_cache_type`
  - 1 or ON
  - 2 or DEMAND – SQL_CACHE
  - SQL_NO_CACHE

- Default query_cache_size=0
MyISAM Index Cache

- `key_buffer_size` – globally allocated on startup

- How much is being used?
  - STATUS variables
  - `Key_blocks_unused` vs. `Key_blocks_used`
  - \((\text{Key\_blocks\_\%} \times \text{key\_cache\_block\_size}) / \text{key\_buffer\_size}\)
InnoDB Buffer Pool

• `innodb_buffer_pool_size` – globally allocated on startup

• How much is being used?
  – STATUS variables
  – `Innodb_buffer_pool_pages_free`
  – `Innodb_buffer_pool_pages_total`
Allocated Per-thread

- `binlog_cache_size = 32 Kb`
- `max_binlog_cache_size = 4 Gb`
- `net_buffer_length = 16 Kb`
  - 1 Mb max
  - `max_allowed_packet`
Per-Thread, Allocated As Needed

- `join_buffer_size` – memory buffer for joins not using indexes
- `read_buffer_size` – memory buffer for sequential table scans
- `read_rnd_buffer_size` – memory buffer for random table seeks
Per-Thread, Allocated As Needed

- preload_buffer_size – when pre-loading indexes

- sort_buffer_size – for sorting

- myisam_sort_buffer_size
  - Index sorting only – REPAIR, OPTIMIZE, creating indexes
Temporary Tables

- `tmp_table_size`
- `max_heap_table_size`
- Large rows, BLOBs written to disk
Performance Caches

- Global Variables
  - table_cache
    - STATUS variables Opened_tables
  - thread_cache_size
    - STATUS variables Connections and Threads_created
Size Matters

- Larger/fragmented data/indexes use more memory
- Larger/fragmented data/indexes take more time to search
- Clustered indexes
- innodb_file_per_table
Size Matters – Data Diet

- OPTIMIZE – how often?
- Purge/archive regularly
- Follow large deletes with OPTIMIZE
Size Matters – Schema Diet

- IP addresses
- TIMESTAMP vs DATETIME
- Strings vs. Numbers vs. ENUM/SET
Myth: MySQL is Slow

- Memory/disk tradeoffs
- RAID
- Temporary tables
  - Created_tmp_tables
  - Created_tmp_disk_tables
Disk I/O

- Many data changes
  - Binary logs
  - Data
  - Index

- INSERT DELAYED

- Batch update/deletes
Network

- Large queries vs. small
  - CPU compute time
  - Network traffic
  - Large data sets

- INSERT....ON DUPLICATE KEY UPDATE
Size Matters, So Does Performance

- Choosing Correct Table Types:
  - MyISAM
  - InnoDB
  - BLACKHOLE
  - MERGE
  - MEMORY
  - ARCHIVE
Query Optimization

- Use EXPLAIN


Know What You're Doing

- Subqueries
- VIEW
- TRIGGER
Know What You're Doing

- LIMIT

- Stored Procedures
  - compiled per thread

- Indexes
  - Selectivity
  - Functions
  - Overhead vs. utility
More Features: Do Not Want!

- Look into Drizzle, a MySQL fork
- http://www.drizzleproject.org
Feedback

- Questions?
- Comments?
- Why else do you hate MySQL?

cabral@pythian.com
How to Stop Hating MySQL:
Fixing Common Mistakes and Myths

Sheeri K. Cabral
Database Team Lead
The Pythian Group, www.pythian.com

cabral@pythian.com
LISA 2008
Who I Am

- MySQL DBA
- MySQL User Group
- MySQL Podcast (OurSQL, on hiatus), videos (technocation.org)
- Lots of community stuff
Myths about MySQL

- Uses too much memory
- Slow
- I need more features!
More Myths about MySQL

- Don't use ENUM
- Schema changes take forever
- You have to restart to log
- No partitioning
MySQL uses too much memory!

- INFORMATION_SCHEMA.GLOBAL_VARIABLES (5.1)

- SHOW GLOBAL VARIABLES (5.0)

- LIKE
  - '%cache%'
  - '%buffer%'
Common Mistake: Wasting Memory

- bdb_cache_size
- bdb_log_buffer_size
- 32-bit operating system
Myth: Query Cache Is Faster!

- Use memcached for caching common queries

- `query_cache_type`
  - 1 or ON
  - 2 or DEMAND – SQL_CACHE
  - SQL_NO_CACHE

- Default `query_cache_size`=0
MyISAM Index Cache

- `key_buffer_size` – globally allocated on startup

- How much is being used?
  - STATUS variables
  - `Key_blocks_unused` vs. `Key_blocks_used`
  - `(Key_blocks_% * key_cache_block_size) / key_buffer_size)`
InnoDB Buffer Pool

- `innodb_buffer_pool_size` – globally allocated on startup

- How much is being used?
  - STATUS variables
  - `Innodb_buffer_pool_pages_free`
  - `Innodb_buffer_pool_pages_total`
Allocated Per-thread

- `binlog_cache_size = 32 Kb`
- `max_binlog_cache_size = 4 Gb`
- `net_buffer_length = 16 Kb`
  - 1 Mb max
  - `max_allowed_packet`
Per-Thread, Allocated As Needed

- `join_buffer_size` – memory buffer for joins not using indexes
- `read_buffer_size` – memory buffer for sequential table scans
- `read_rnd_buffer_size` – memory buffer for random table seeks
Per-Thread, Allocated As Needed

- `preload_buffer_size` – when pre-loading indexes
- `sort_buffer_size` – for sorting
- `myisam_sort_buffer_size`
  - Index sorting only – REPAIR, OPTIMIZE, creating indexes
Temporary Tables

- `tmp_table_size`

- `max_heap_table_size`

- Large rows, BLOBs written to disk
Performance Caches

- Global Variables
  - table_cache
    - STATUS variables Opened_tables
  - thread_cache_size
    - STATUS variables Connections and Threads_created
Size Matters

- Larger/fragmented data/indexes use more memory
- Larger/fragmented data/indexes take more time to search
- Clustered indexes
- innodb_file_per_table
Size Matters – Data Diet

- OPTIMIZE – how often?
- Purge/archive regularly
- Follow large deletes with OPTIMIZE
Size Matters – Schema Diet

- IP addresses
- TIMESTAMP vs DATETIME
- Strings vs. Numbers vs. ENUM/SET
Myth: MySQL is Slow

- Memory/disk tradeoffs
- RAID
- Temporary tables
  - Created_tmp_tables
  - Created_tmp_disk_tables
Disk I/O

• Many data changes
  – Binary logs
  – Data
  – Index

• INSERT DELAYED
• Batch update/deletes
Network

- Large queries vs. small
  - CPU compute time
  - Network traffic
  - Large data sets

- INSERT....ON DUPLICATE KEY UPDATE
Size Matters, So Does Performance

- Choosing Correct Table Types:
  - MyISAM
  - InnoDB
  - BLACKHOLE
  - MERGE
  - MEMORY
  - ARCHIVE
Query Optimization

- Use EXPLAIN


Know What You're Doing

- Subqueries
- VIEW
- TRIGGER
Know What You're Doing

- LIMIT

- Stored Procedures
  - compiled per thread

- Indexes
  - Selectivity
  - Functions
  - Overhead vs. utility
More Features: Do Not Want!

- Look into Drizzle, a MySQL fork
- http://www.drizzleproject.org
Feedback

- Questions?
- Comments?
- Why else do you hate MySQL?

cabral@pythian.com