What's New in MySQL 5.7

A MySQL White Paper

October 2015

Introduction

MySQL is the most trusted and depended-on open source database platform in use today. 9 out of the top 10 most popular and highly-trafficked websites in the world rely on MySQL primarily due to its ubiquity across heterogeneous platforms and application stacks and for its proven performance, reliability, and ease of use. MySQL 5.7 builds on this momentum by delivering across the board improvements designed to enable innovative DBAs and Developers to create and deploy the next generation of web, embedded, and Cloud/SaaS/PaaS/DBaaS applications on the latest generation of development frameworks and hardware platforms.

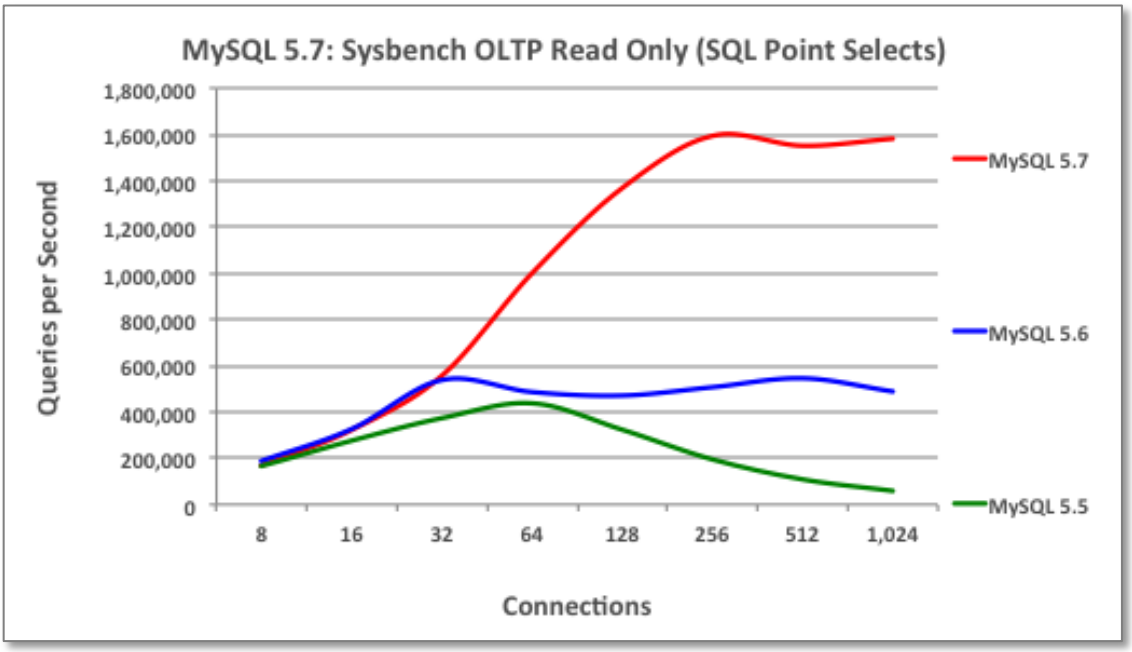
This paper serves as a DBA and Developer guide to MySQL 5.7 as it highlights the key new features. To conclude, there is a summary of how MySQL 5.7 can be implemented using best practices comprising production-ready support and services of the MySQL Enterprise Edition offering.

Performance Improvements

MySQL 5.7 delivers significant performance improvements on all workloads and sets new benchmark records.

3x Faster than MySQL 5.6: SQL Point Selects

MySQL 5.7 delivers 1,600,000 Queries/Second in the Point Select query benchmark, 3x faster than MySQL 5.6.



For context, the benchmark shown above was run on the following platform configuration:

- Intel(R) Xeon(R) CPU E7-8890 v3
- 4 sockets x 18 cores-HT (144 CPU threads)
- 2.5 Ghz, 512GB RAM
- Linux kernel 3.16

For more information on the MySQL 5.7 benchmarks, visit:
http://www.mysql.com/why-mysql/benchmarks/

The SysBench benchmark tool is freely available for application use-case specific benchmarks and can be downloaded here:
http://dev.mysql.com/downloads/benchmarks.html

Native JSON Support

In MySQL 5.7 a new data type for storing JSON data in MySQL tables has been added. Having a native data type for JSON provides some *major* benefits:

- **Document Validation -** Only valid JSON documents can be stored in a JSON column, so you get automatic validation of your data.
- **Efficient Access -** More importantly, when you store a JSON document in a JSON column, it is not stored as a plain text value. Instead, it is stored in an optimized binary format that allows for quicker access to object members and array elements.
- **Performance -** Improve your query performance by creating indexes on values within the JSON columns. This can be achieved with “functional indexes” on virtual columns.
- **Convenience** - The additional inline syntax for JSON columns makes it very natural to integrate Document queries within your SQL. For example (features.feature is a JSON column):

```
SELECT feature->"$.properties.STREET" AS property_street FROM features
WHERE id = 121254;
```

With MySQL 5.7, you can now seamlessly blend the best of the Relational and Document paradigms in a single tool, using whatever is best for each specific application and use case. This opens up a tremendous number of new use cases for MySQL users.

For more information:
- Native JSON Data Type and Binary Format
- JSON Functions, Part 1 - Manipulation JSON Data
- JSON Functions, Part 2 – Querying JSON Data
- Virtual Columns and Effective Functional Indexes in InnoDB
- JSON Manual

Performance Schema

Monitoring is important to our users and customers, and essential to any data management system. At the core of our monitoring strategy we have Performance Schema. The Performance Schema enhancements to MySQL 5.7 include numerous new monitoring capabilities, reduced footprint and overhead, and significantly improved ease of use with the new SYS Schema. On the monitoring side, we have instrumented:

- **Metadata Locking** - critical to understand metadata lock dependencies between sessions.
- **Stage Tracking** - follow the progress of long running operations (e.g. ALTER TABLE).
- **Transactions** - monitor all aspects of transaction processing within the server and storage engines.
- **Memory Usage** - aggregates memory usage statistics to understand and tune the memory consumption of a server.
- **Stored Programs –** instrumentation for stored procedures, stored functions, events scheduler events and table triggers.
- **Prepared Statements -** Provide aggregated statistics by prepared statements and expose prepared statements in use by the server.

We have exposed SHOW SLAVE STATUS information and user variables in Performance Schema. We also expose the GLOBAL and SESSION level STATUS and VARIABLES while preserving the corresponding SHOW STATUS and SHOW VARIABLES commands for backward compatibility.

For more information:
- Performance Schema Reference Manual
- Performance Schema and SYS Schema in MySQL 5.7
- Performance Schema Overview & New Features

SYS Schema

MySQL SYS Schema is a database schema with a set of objects (views, stored procedures, stored functions, tables, and triggers) that were implemented to give easy, readable, DBA and Developer based use case access to the wealth of instrumentation data available primarily within Performance Schema, but also within various INFORMATION_SCHEMA tables as well.

MySQL SYS Schema is included in the MySQL 5.7 Server by default and provides summary views to answers common questions, like:

- *"Who is taking up all the resources on my database server?"*
- *"Which hosts are hitting my database server those most?"*

- *"Where is all the memory going on my instance?*

For more information:
- The MySQL SYS Schema in MySQL 5.7
- Performance Schema and SYS Schema in MySQL 5.7
- MySQL SYS Schema Version 1.5.0 Released

InnoDB Improvements

InnoDB Online Operations

MySQL 5.7 supports additional online operations for developers and DBAs that need to manage MySQL in production environments. The biggest feature we have added in this space is the ability to resize your InnoDB Buffer Pool online. Now you can adjust your Buffer Pool size dynamically in order to adapt to changing needs without restarting MySQL. InnoDB can also now automatically truncate the InnoDB UNDO log and tablespaces as an online operation, thus eliminating one common cause of large shared tablespace files (ibdata1). Lastly, we have added support for online RENAME INDEX and enlarge VARCHAR column size operations.

For more information:
- Truncate Undo Tablespace
- Innodb Buffer Pool Online Resize
- Innodb Overview of Online DDL

InnoDB General Tablespaces

You can now group schemas and tables together into logical and physical tablespaces, thus improving resource utilization and providing an easy and efficient way to do data migrations in various cases (e.g. using a dedicated tablespace per customer/user).

For more information:
- InnoDB General Tablespaces

InnoDB Transparent Page Level Compression

InnoDB has supported table compression since MySQL 5.1. InnoDB Page Level Compression--the idea originally introduced by FusionIO (now part of SanDisk)--is a new feature in MySQL 5.7 and it complements the existing InnoDB table compression, allowing them to coexist in the same server instance. Users can now choose the compression scheme that makes the most sense for their use case, even on a table-by-table basis.

Supported compression algorithms include Zlib and LZ4. When a page is written, it is compressed using the specified compression algorithm. The compressed data is written to disk, where the hole punching mechanism then releases empty blocks from the end of the page. If compression fails, data is written out as-is.

InnoDB now also supports 32K and 64K page sizes, which is a nice complement to the page level compression. In general, larger page sizes will typically increase the amount of redundant data and thus improve the potential compression ratios.

All of the above work, when also combined with the new support in 5.7 for a user configurable fill factor and page merge thresholds, allow InnoDB to make the best use of your storage space.

For more information:
- InnoDB Transparent Page Compression
- InnoDB Supporting Page Sizes of 32k and 64k

InnoDB Native Partitioning

In MySQL 5.7 InnoDB now contains native support for partitioning. The InnoDB native partitioning removes overhead and reduces memory requirements by up to 90%. This work also paves the way for us to both remove existing limitations on partitioned tables, as well as add new features such as global indexes and parallel query execution.

For more information:
- InnoDB Native Partitioning

InnoDB Native Full Text Search

We introduced InnoDB Full Text Search in 5.6. We have now added greater flexibility and further optimizations in MySQL 5.7. For example, FullText indexes in InnoDB now support an external parser. The plugin can either replace the built-in parser or it can act as a front-end for it. We have also implemented Optimizer hints that are passed down to InnoDB about a query so that InnoDB may skip part of the Full-Text search processing, e.g. not to compute the ranking values if they are not needed.

As of MySQL 5.7.6, MySQL also provides a built-in full-text N-gram parser plugin that supports Chinese, Japanese, and Korean (CJK), as well as an installable MeCab full-text parser plugin for Japanese.

For more information:
- InnoDB Supports Full-Text Plugin Parser

- MySQL Full-Text Plugins
- InnoDB Full-Text N-gram Parser
- InnoDB MeCab Full-Text Parser Plugin

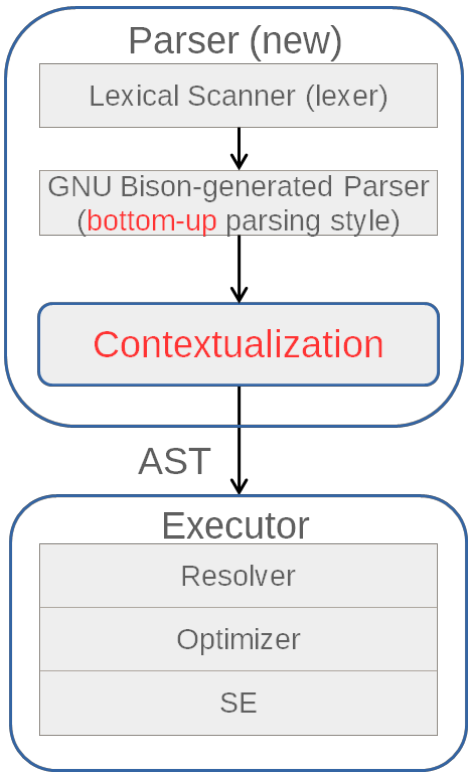
InnoDB Cache Retention

InnoDB now automatically retains the hottest 25% of your buffer pool across MySQL restarts. No longer do you need to use any manual process to preload or warm your data cache, nor do you suffer the associated performance penalty when a restart of MySQL is required.

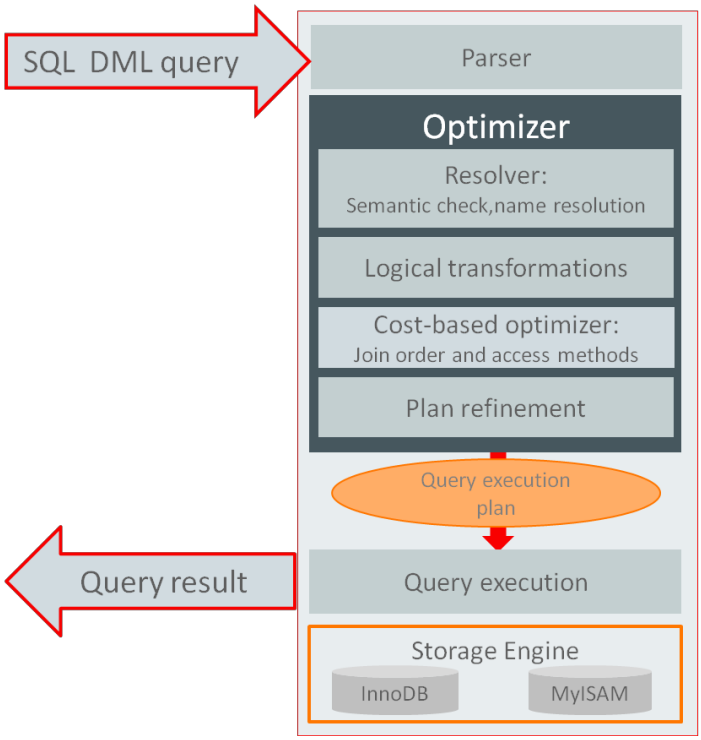
For more information:
- InnoDB Preload Buffer Pool
- MySQL Dumping and Reloading the InnoDB Buffer Pool

Optimizer Improvements

We have refactored the MySQL optimizer and as a result, large parts of the parser, optimizer, and cost model have been rewritten to improve maintainability, extendibility, and performance.



New MySQL Parser



New MySQL Optimizer

Optimizer Cost Model

To generate optimal execution plans, the optimizer uses a cost model that is based on estimates for the costs of various operations that occur during query execution. The optimizer has historically relied on a set of compiled-in default "cost constants" to make execution plan decisions.

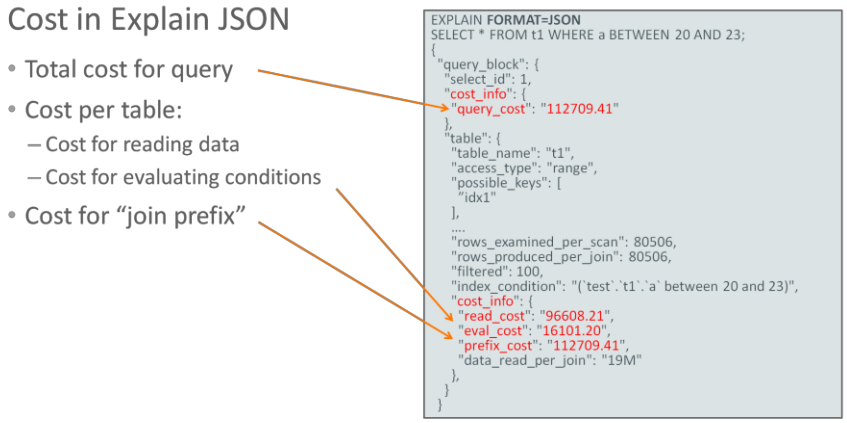
In MySQL 5.7, the cost based optimizer has been improved to make dynamic, intelligent, and ultimately better cost based decisions. For example, better cost estimates take into account new hardware architectures (larger buffers & caches, SSDs, etc.). Better cost estimates will lead to better decisions by the optimizer and thus better query performance.

The MySQL 5.7 optimizer includes a configurable database of cost estimates to use during execution plan construction. These estimates are stored in the server_cost and engine_cost tables within the mysql system database.

For more information:
- The MySQL Optimizer Cost Model Project
- Optimizer Cost Model Manual
- Optimizer Cost Model Presentation

JSON Explain

We have enhanced the JSON EXPLAIN output by printing the related cost estimates, and the estimated amount of data to be processed. This will make it easier for a DBA to see the critical differences between execution plans.

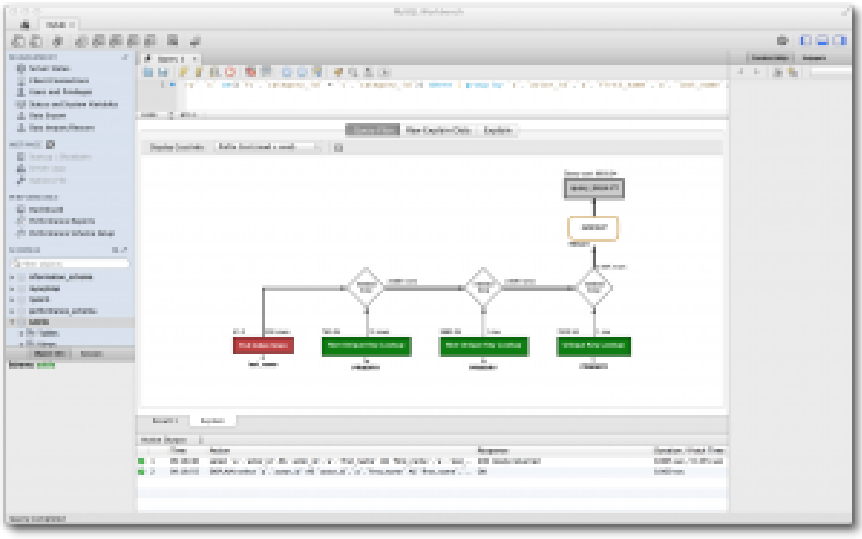


Explain Costs

For more information:
- Optimizer Cost Model Improvements in MySQL 5.7.5 DMR
- Explain Output Manual

Workbench Visual Explain

These new insights also allow for clients to provide additional visual representations of this data. The new Visual Explain functionality in MySQL Workbench is a great example.



For more information:
- MySQL EXPLAIN Explained

New and Improved Optimizer Hints

We have added parser rules to support a new hint syntax for optimizer hints. Hints can now be specified directly after the SELECT, INSERT, REPLACE, UPDATE, and DELETE keywords in an SQL statement, enclosed in /*+ */ style comments. Not only does this provide the end-user with a more consistent, powerful, and easy-to-use method of managing hints, but the internal refactoring done makes it far easier to add Server side support for additional hints moving forward. For example, we have added many new index and join related hints in MySQL 5.7.

For more information:
- New Optimizer Hints in MySQL
- Optimizer Hints Manual

Query Rewrite Plugin

MySQL 5.7 implements a Query Rewrite Plugin. The Query Rewrite Plugin allows you to specify how certain queries that arrive at the server should be rewritten before they are processed and executed by the server. For example, the Query Rewrite Plugin can enforce/prevent a certain query plan by adding hints to a query. The Query Rewrite Plugin is a superb tool for handling problematic queries when you cannot rewrite the query within the application itself, e.g. because it originates from an external tool (like an Object Relational Mapping tool such as Hibernate) or from some other 3rd party application. As part of this work, we've also created two APIs that can be used to write custom Query Rewrite Plugins.

For more information:
- Query Rewrite Plugins
- Write Yourself a Query Rewrite Plugin Part 1
- Write Yourself a Query Rewrite Plugin Part 2

Generated Columns

Generated or "Computed" Columns are virtual columns that are not physically stored in the table, unless the column is explicitly defined as *STORED*. This offers a convenient way to cache the results of commonly used expressions, and offer what's often referred to as "functional indexes" or "function based indexes".

It also makes it convenient and efficient to mix dynamic schema-less data within your relational tables and SQL, for example when using JSON Documents.

For more information:
- Generated Columns in MySQL 5.7.5
- Create Table Manual

Server Side Statement Timeouts

This brings users the ability to set query timeout limits at the global level, on a per-session basis, and for individual queries. For example:

```
SELECT /*+ MAX_EXECUTION_TIME(1000) */ * FROM my_table;
```

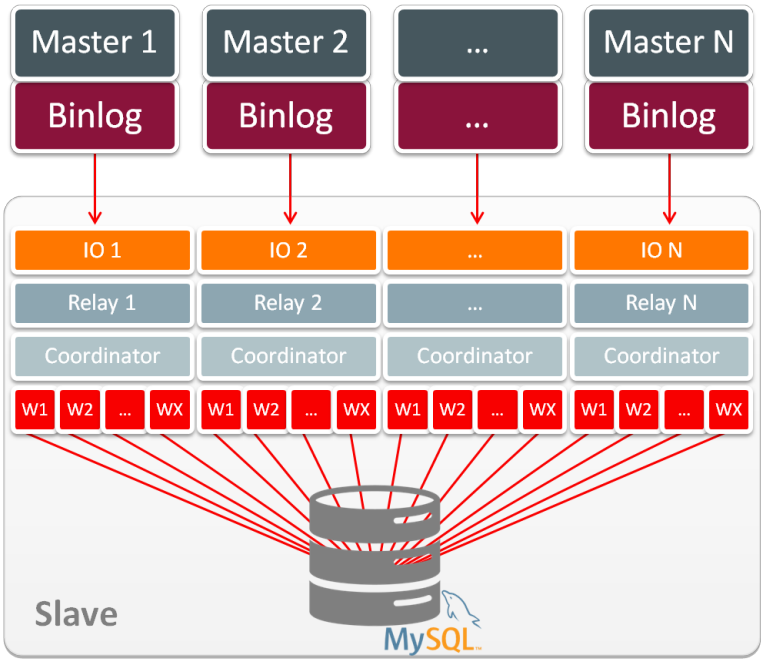
For more information:
- Server Side SELECT Statement Timeouts

Replication Improvements

Multi-Source Replication

MySQL Multi-Source Replication enables a replication slave to receive transactions from multiple sources simultaneously. Multi-source replication can be used to:

- Consolidate data from multiple servers to a single server
- Back up multiple servers to a single server
- Merge table shards



The monitoring interfaces now provide details for each replication channel, and we have also introduced new `performance_schema` tables to easily monitor your entire multi-source replication topology.

For more information:
- Introducing Multi-Source Replication
- Multi-Source Replication Manual

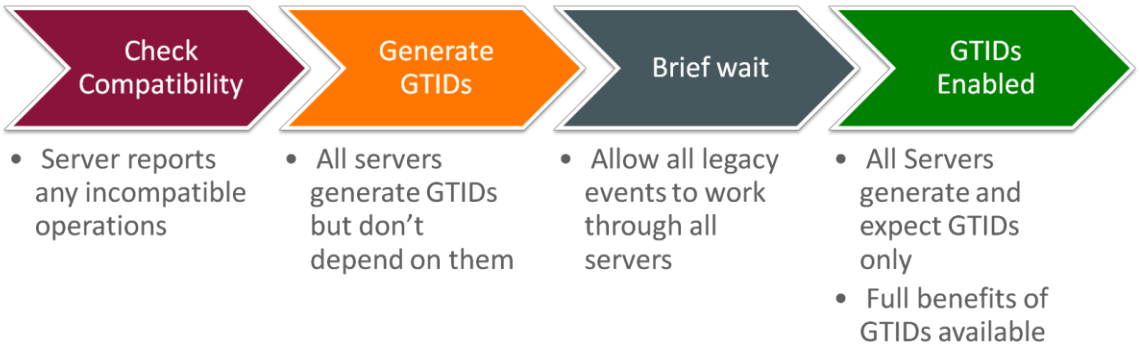
Transaction Based Parallel Replication

MySQL 5.7 adds intra-schema multi-threaded slaves. With this implementation (`slave-parallel-type=LOGICAL-CLOCK`) the slave will be able to apply transactions in parallel, even within a single database or schema, as long as they have a disjoint read and write set. This work allows the slave to keep up with the master, thus eliminating the most common cause of “slave lag”.

For more information:
- MySQL 5.7 Enhanced MTS: configuring slave for Intra-database parallelization

Online Replication Changes

You can now enable Global Transaction ID based replication as an online operation, allowing you to take advantage of next generation replication features without incurring downtime in your MySQL production environments.



You can also change replication filters online now, providing a variety of ways to configure data replication within your MySQL farm, and finally you can perform master failover operations (`CHANGE MASTER`) without stopping replication execution on slaves.

For more information:
- Replication Enable GTIDs Online
- Enabling GTIDs Without Downtime
- Change Replication Filter

Semi-Sync Replication Enhancements

The Semi-Sync replication plugin has improved semantics that provide better performance and reliability.

For more information:
- Loss-Less Semi-Sync Replication

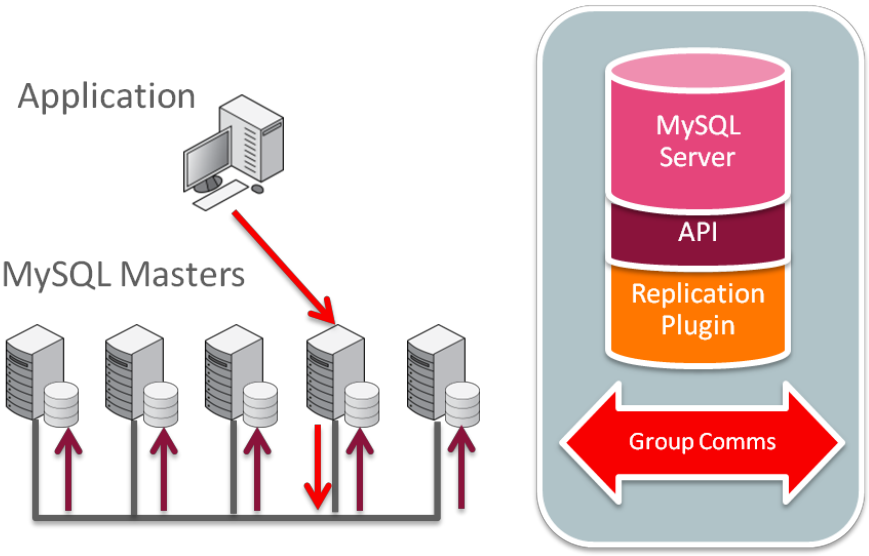
Enhanced Monitoring

In addition to the legacy `SHOW` commands, we have added a variety of new Performance Schema tables that offer unprecedented insights into what is happening and how things are performing. This allows you to easily ensure that your replication topology is healthy and performing well, while also providing the information needed to debug any issues that may occur.

For more information:
- Performance Schema Replication Tables

Group Replication

This work brings native active/active write anywhere replication clusters to MySQL. You can now choose a mix of appropriate replication models based on your specific needs: asynchronous, semi-synchronous, and (virtually) synchronous (Group Replication).



Group Replication takes care of membership, consistency, and other management related functions without the need for manual intervention or custom tooling. Easy High Availability for MySQL has arrived!

For more information:
- MySQL Group Replication: Hello World
- MySQL Goup Replication Blog Posts
- State of Group Replication Presentation

High Availability Improvements

Support for tracking session transaction state offers better support for load balancing across nodes in a cluster.

Server version tokens offer better support for caching in distributed systems.

A new data migration tool--mysqlpump--improves data migration and sharding operations between nodes (e.g. shard rebalancing).

Improved Replication options in HA groups
- Improved slave performance with logical clock based parallelization
- Loss-less Semi-Sync Replication plugin supporting multi-node acks
- Synchronous replication (Group Replication plugin)

For more information:
- MySQL Session Tracking
- MySQL Version Tokens
- Introducing mysqlpump

GIS Improvements

We have seen tremendous growth in mobile applications and the need for location based services. To address these needs, MySQL 5.7 includes significant GIS improvements.

New & Improved Spatial Extensions

MySQL 5.7 replaces the legacy code used for geometry computations with an external library, namely Boost.Geometry. Boost.Geometry is a powerful, reliable, open source geometric engine with an active and vibrant development community that we are actively engaged in. We have two of the lead Boost.Geometry developers in-house, and we happily contribute our code back to the upstream project. The end result of all this is that we have a better performing, more fully featured, reliable, and standards compliant MySQL GIS.

InnoDB Spatial Indexes

Furthermore, InnoDB now supports Spatial Indexes (implemented as R-Tree indexes). This makes any spatial search far more efficient, while adding the transactional support, MVCC, and ACID guarantees that Enterprises need today.

GeoJSON

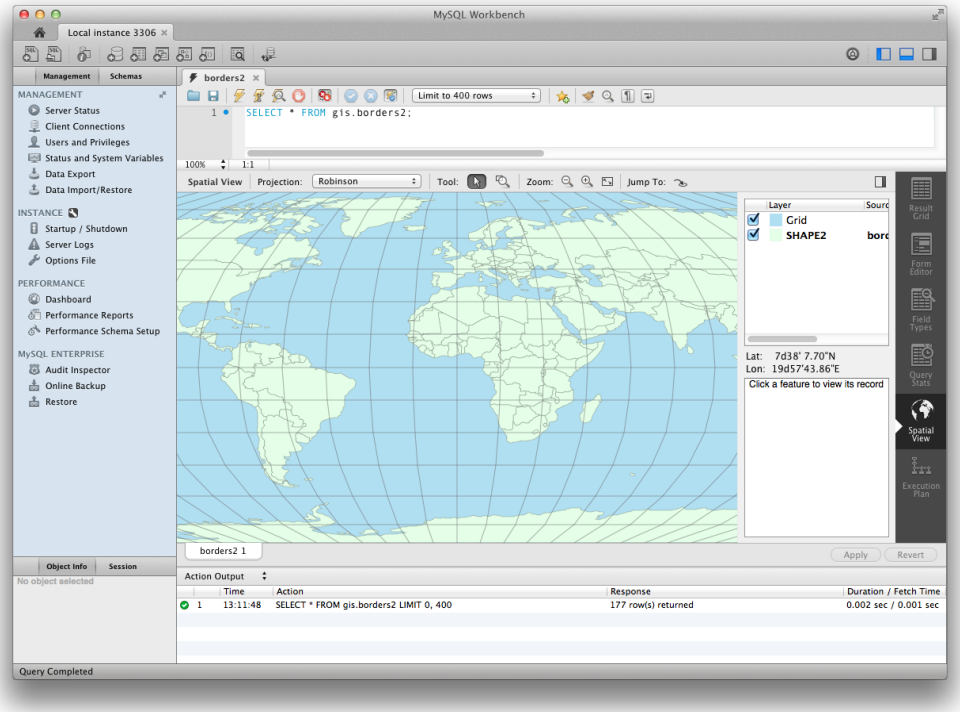
In recent years, GeoJSON has become a popular data format for exchanging GIS data due the fact that it's easy to read, it's simple, and it's lightweight. In MySQL 5.7 we have added support for parsing and generating GeoJSON documents via two new functions: ST_GeomFromGeoJson() and ST_AsGeoJson(). These functions make it easy to connect MySQL with other GeoJSON enabled software and services, such as the Google Maps Javascript API.

GeoHash

In MySQL 5.7 we introduced new functions for encoding and decoding Geohash data. Geohash is a well known system for encoding and decoding longitude and latitude coordinates (in the WGS 84 coordinate system) into a text string.

Spatial Aware Client Tools

Our popular graphical client tool, MySQL Workbench, now includes a new Spatial Viewer. This is automatically presented as a viewing option when `GEOMETRY` columns are detected in the results. We've also now incorporated valid spatial data in the Sakila sample database (the address.location column) to make it easy and convenient to test out the new GIS work.



MySQL Workbench provides visual tools for developing GIS applications

For more information:
- MySQL 5.7 and GIS, an Example
- GeoJSON Functions
- Geohash Functions
- Spatial Extensions Manual

Security Improvements

MySQL 5.7 adds “secure by default” installations, as well as many features that help users keep their database instances free from unintended use. This includes better data encryption, better password handling, better transport layer security, and more.

AES 256 Encryption

MySQL supports multiple AES Encryption modes. We have enhanced the security strength of our Advanced Encryption Standard (AES) encryption/decryption functions—AES_ENCRYPT() and AES_DECRYPT()—by adding support for larger key sizes and different block modes.

Password Rotation Policies

MySQL enables database administrators to expire account passwords manually, and to establish a policy for automatic password expiration. Any user who connects to the server using an account for which the password is past its permitted lifetime must change the password.

For more information:
- Password Expiration Policy in MySQL Server 5.7

MySQL Secure Installation

mysql_secure_installation enables you to improve the security of your MySQL installation in the following ways:

- You can set a password for root accounts.
- You can remove root accounts that are accessible from outside the local host.
- You can remove anonymous-user accounts.
- You can remove the test database (which by default can be accessed by all users, even anonymous users).

MySQL deployments installed using mysqld --initialize are secure by default.

- The installation process creates only a single root account, 'root'@'localhost', automatically generates a random password for this account, and marks the password expired. The MySQL administrator must connect as root using the random password and assign a new password.
- Installation creates no anonymous user accounts.
- Installation creates no test database.

MySQL Enterprise Edition

- **MySQL Database** – The most secure and up to date version of MySQL is used to power the most demanding online, web, cloud and OLTP applications and services. The MySQL commercial server is a fully integrated transaction-safe, ACID compliant database with full commit, rollback, crash- recovery and row level locking capabilities. MySQL is known for its performance, reliability and ease-of-use across all Linux, UNIX, Mac OS X, and Windows platforms.

- **MySQL Enterprise Backup** – MySQL Enterprise Backup performs online, non-blocking backups of your MySQL databases. Perform full, incremental, and partial backups for all InnoDB data while MySQL is fully available for transactional operations. All backup operations are executed

in parallel for quick results and also support compression options that reduce the size of backup images by up to 90%. Recovery options include backward compatible full recovery, precise one-click point-in-time recovery, and partial recovery for restoring a specific set of objects

- **MySQL Enterprise Firewall** - Blocks SQL Injection attacks that can result in loss of valuable personal and financial data. Whitelist creation, real-time threat monitoring, SQL statement blocking and alerting all enable DBAs to protect vital data assets. Acting as an intrusion detection system, MySQL Enterprise Firewall can notify administrators to SQL statement activity that does not match an approved whitelist.

- **MySQL Enterprise Audit** - MySQL Enterprise Edition provides an easy to use, policy-based auditing solution that helps developers implement stronger security controls and satisfy regulatory compliance without changes to existing applications.

- **MySQL Enterprise Scalability** - To meet the sustained performance and scalability of ever increasing users, query, and data loads--MySQL Enterprise Edition provides the MySQL Thread Pool. The Enterprise Thread Pool provides a highly scalable, queue-based thread-handling model designed to reduce overhead in managing client connections and statement execution threads.

- **MySQL Enterprise Authentication** - MySQL Enterprise Edition provides ready to use external authentication modules so DBAs and developers can easily integrate MySQL with existing security infrastructures, including Linux Pluggable Authentication Modules (PAM) and Windows Active Directory.

- **MySQL Enterprise Encryption** - To protect sensitive data throughout its lifecycle, MySQL Enterprise Encryption provides industry standard functionality for asymmetric encryption (Public Key Cryptography). MySQL Enterprise Encryption also provides encryption, key generation, digital signatures and other cryptographic features to help organizations protect confidential data and comply with regulatory requirements such as HIPAA, Sarbanes-Oxley, and the PCI Data Security Standard.

- **MySQL Enterprise High Availability** - MySQL Enterprise Edition offers a wide range of solutions for database high availability, to automatically detect and recover from failures--whether these occur at the network, host, OS or database layer--as well as minimize downtime resulting from scheduled maintenance activities.

- **MySQL Enterprise Monitor and Query Analyzer** – Continuously monitors MySQL databases and proactively alerts DBA to potential

problems, queries, and tuning opportunities before they impact key systems or applications. The Monitor provides a set of MySQL expert advisors that provide insight and detailed guidance on fixing and tuning MySQL configurations and variables for optimal security, performance and availability. The built-in, advanced Query Analyzer allows developers to visually find and tune expensive query code without the need for Slow Query Log, SHOW PROCESS LIST or other labor intensive methods.

- **Oracle Enterprise Manager for MySQL**
  The Oracle Enterprise Manager for MySQL provides Oracle developers and DBAs with real-time monitoring and delivers comprehensive performance, availability and configuration information for your MySQL databases. In addition to the default MySQL compliance rules, you can also create custom critical and warning thresholds can for each of the collected metrics.

- **MySQL Workbench** – Provides GUI-based data modeling, SQL development, deployment, database migration and comprehensive administrative tools (server configuration, user administration, object management) for database architects, developers, and DBAs.

- **Oracle Premier Lifetime Support for MySQL** – Provide direct access to expert MySQL Support engineers who are ready to assist users in the development, deployment, and management of MySQL applications. The MySQL Support team is composed of seasoned MySQL developers and database experts who understand the issues and challenges users face because they've overcome these same challenges themselves.

Conclusion

MySQL is the most popular and widely used open source database in the world because of its performance, reliability, and ease of use. MySQL 5.7 builds on this momentum by providing DBAs and developers with the added performance, scalability, and across the board improvements needed when building the next generation of web-based, cloud-based, and embedded applications and services. These improvements, further demonstrating that Oracle drives MySQL innovation, make MySQL 5.7 by far the most comprehensive and feature rich release in our history.

Additional Resources

MySQL 5.7 Downloads
http://www.mysql.com/downloads/

MySQL 5.7 Documentation
http://dev.mysql.com/doc/refman/5.7/en/

MySQL Enterprise Edition
http://mysql.com/products/enterprise/

MySQL Customers and Case Studies
http://www.mysql.com/customers

MySQL Professional Services and Consulting
http://mysql.com/consulting/

For more information on Oracle's MySQL products and services, please visit:
http://www.mysql.com/products/