Percona Server for MySQL: What it is and How it is Done

Satya Bodapati, Yura Sorokin

Percona

OPEN SOURCE DATABASE CONFERENCE

PERCONA

LIVE EUROPE

AMSTERDAM

Part 1: Agenda

- What is Percona Server for MySQL
- Percona Server for MySQL Features / Enhancements
- Deprecations and Removals

What Is Percona Server for MySQL

...a free, fully compatible, enhanced and open source drop-in replacement

for any MySQL database...

What Is Percona Server for MySQL

- Based on open-source MySQL community edition
- With enhancements
- With bug fixes
- Care is taken to maintain drop-in quality unless some specific features are enabled

Percona Server for MySQL Features

- Write-optimized storage engines
- Data Encryption
- Enterprise Features
- Percona Specific Features

Write-Optimized Storage Engines

Write-Optimized Storage Engines

- InnoDB is a battle-tested B-tree-based read-optimized storage engine
- LSM-tree-based RocksDB key-value store
- Efficient writes, efficient compression, reads not that bad
- RocksDB-based MyRocks storage engine
- New in 8.0: native partitioning
- New in 8.0: cross-engine consistent physical backups
- TokuDB fully supported but scheduled to EOL at the end of 8.0 cycle

Data Encryption

Data Encryption

- Do not write any unencrypted data to disk in an InnoDB-based server
- Support several key storage options
- Support key rotation
- Implemented through a combination of existing MySQL features, MariaDB security feature porting, and custom Percona development

Data Encryption: Percona Enhancements

- InnoDB Temporary tablespace encryption (ibtmp*). Undo from temporary tables still goes there
  - How? innodb_temp_tablespace_encryption=ON
- Session temporary tablespaces encryption (IBT files)
  - How? innodb_temp_tablespace_encryption=ON or default_table_encryption=ON
- System tablespace encryption (ibdata*)
  - How? (at bootstrap) innodb_sys_tablespace_encrypt=ON

Data Encryption: Percona Enhancements

- Parallel Doublewrite buffer encryption
  - How? innodb_parallel_dblwr_encrypt=ON
- InnoDB temp file (DDL log) encryption
  - How? innodb_encrypt_online_alter_logs=ON
- Server temp file encryption
  - How? --encrypt-tmp-files=ON

Data Encryption: Percona Enhancements

- Key storage in Hashicorp Keyring Vault
  - How? keyring_valut plugin
- Key versioning and rotation (pre-GA)
  - How? innodb_encryption_threads, default_table_encryption=ONLINE_TO_KEYRING, innodb-encryption-rotate-key-age

	MySQL 8.0	MariaDB	Percona Server 5.7	Percona Server 8.0
File-per-table tablespace encryption	✓	✓	✓	✓
General tablespace encryption	✓	✓	✓	✓
Temporary tablespace encryption		? may be fixed	✓	✓
System tablespace encryption		✓	✓	✓
Parallel Doublewrite buffer encryption	N/A	N/A	✓	✓
mysql.ibd encryption	✓	✓	N/A	✓
InnoDB redo log encryption	✓	✓	✓	✓
InnoDB undo log encryption	✓	✓	✓	✓
InnoDB temp file (DDL log) encryption		✓(?)	✓	✓
Server temp file encryption		✓(?)	✓	✓
Binary log encryption	✓	✓	✓	✓
Key storage in a local file	✓	✓	✓	✓
Key storage in Hashicorp Keyring Vault			✓	✓
Key storage in Amazon KMS	✓(EE)	✓		
Key storage in Oracle Vault	✓(EE)			
Key versioning and rotation		✓(?)	pre-GA	pre-GA

Percona Server for MySQL 8.0: Enterprise Features

Percona Server for MySQL 8.0: Enterprise Features

- Threadpool
- Audit Plugin
- PAM authentication plugin

Percona Server for MySQL 8.0: Enterprise Features

Threadpool:

- Have to increase *max_connections*? Just increasing the number may result in decrease in performance
- Why? MySQL uses one OS thread per connection
- Use threadpool plugin to reduce the number of OS threads used which will then reduce the context switching and hot locks contentions
- Allows you to use higher *max_connections* value while maintaining performance
- An opensource alternative to the MySQL Enterprise Threadpool Plugin

Percona Server for MySQL 8.0: Enterprise Features

Audit Plugin:

- Provides monitoring and logging of connections and queries
- Logs can be stored in multiple formats like CSV, JSON, XML, etc.
- This implementation is an alternative to the MySQL Enterprise Audit Log Plugin

Percona Server for MySQL 8.0: Enterprise Features

PAM authentication plugin:

- PAM stands for Pluggable Authentication Modules (available on Linux)
- Percona PAM Authentication Plugin is a free and Open Source implementation of MySQL's authentication plugin (Enterprise only)
- Acts a mediator between MySQL server, MySQL client, and PAM stack

Percona Specific Features

Percona Server for MySQL 8.0: Memory Engine VARCHAR/BLOB

	Percona Server Memory SE	MySQL Memory SE	MySQL TempTable SE
Can store VARCHAR	✓		✓
Can store BLOB (JSON, TEXT, ...)	✓		✓
Users can create tables	✓	✓	
Query Optimiser can use if VARCHAR			✓

Percona Server for MySQL 8.0: InnoDB Enhancements

- Dedicated LRU Flusher per buffer pool instance
- Solved the famous “single page flushing” issue on IO bound workloads
  - See *innodb_empty_free_list_algorithm*
- Parallel doublewrite buffer

Percona Server for MySQL 8.0: Column Compression with Dictionaries

- A problem to solve: store lots of small JSON documents efficiently
- InnoDB table compression: limited compression ratio
- InnoDB page compression: just don’t
- Compression at application level: limited compression ratio, needs application changes
- MyRocks: maybe?

Percona Server for MySQL 8.0: Column Compression with Dictionaries

- CREATE TABLE t (foo JSON COLUMN_FORMAT COMPRESSED) ENGINE=InnoDB;
- Now “foo” gets compressed and decompressed transparently on each access
- Create a dictionary with shared words between different rows for compression ratio improvement

Percona Server for MySQL 8.0: Column Compression with Dictionaries

- CREATE COMPRESSION DICTIONARY address_parts (‘country’ ‘state’ ‘city’ ...)
  - CREATE TABLE ... address JSON COLUMN_FORMAT COMPRESSED WITH COMPRESSION_DICTIONARY address_parts ...

Percona Server for MySQL 8.0: Backup Support

- MySQL 8.0: LOCK INSTANCE FOR BACKUP
- MySQL 8.0: performance_schema.log_status
- Percona Server for MySQL 8.0: LOCK TABLES FOR BACKUP blocks less
- Percona Server for MySQL 8.0: performance_schema.log_status extended for MyRocks
- Percona Server for MySQL 8.0: START TRANSACTION WITH CONSISTENT SNAPSHOT consistent across storage engines and binlog

Partitioned TokuDB and MyRocks Tables

- MySQL dropped support for legacy partitioning handler, all storage engines must implement partitioning natively
- Implemented for TokuDB and MyRocks in 8.0
- That is not enough for upgrades: 8.0 server cannot read 5.7-format partitioned tables
- Hence native partitioning implemented in 5.7 too for the sole purpose of upgrade
    - ALTER TABLE ... UPGRADE PARTITIONING

SET STATEMENT FOR → /*+ SET_VAR */

- Percona Server 5.7 (and MariaDB) per-statement variable assignments:
- SET STATEMENT sort_buffer_size = 100000 FOR SELECT name, id ...
- MySQL 8.0 query optimizer hint for setting variables:
- SELECT /*+ SET_VAR(sort_buffer_size=100000) */ name, id ...
- MySQL 8.0 implementation is restricted compared to Percona Server 5.7 for MySQL
- Percona Server for MySQL 8.0 addresses most restrictions

“Userstat” Duration Columns Now Floating-Point

- INFORMATION_SCHEMA.CLIENT_STATISTICS, THREAD_STATISTICS, USER_STATISTICS tables column CONNECTED_TIME, BUSY_TIME, CPU_TIME types changed:
- Percona Server for MySQL 5.7: integers, providing 1 second resolution only
- Percona Server for MySQL 8.0: floating-point doubles

SHOW [EFFECTIVE] GRANTS

- MySQL 5.7 SHOW GRANTS
  - Shows assigned but not effective grants
- Percona Server for MySQL 5.7 SHOW GRANTS
  - The above was considered a bug
  - Changed to show effective but not assigned grants
- Percona Server for MySQL 8.0 SHOW [EFFECTIVE] GRANTS
  - The above was also considered a bug
  - SHOWS GRANTS: shows assigned grants
  - SHOW EFFECTIVE GRANTS: shows effective grants

[innodb_]kill_idle_transactions

- Percona Server for MySQL 5.7: kill_idle_transactions and innodb_kill_idle_transactions
  - Identically-behaving aliases
  - innodb_kill_idle_transactions deprecated
- Percona Server for MySQL 8.0: kill_idle_transactions only

Deprecated / Removed Features

Deprecated / Removed Features

- Scalability metrics plugin
  - Subtly broken architecture with no easy fix
  - No users we are aware of
- INFORMATION_SCHEMA.THREAD_STATISTICS CONCURRENT_CONNECTIONS column
- All of our query cache patches

Removed Features due to Upstream

- Query Response Time plugin: replaced by Performance Schema execution time histograms
- innodb_flush_method=ALL_O_DIRECT
  - Avoided keeping redo logs in the kernel filesystem cache
  - But MySQL 8.0 redo logging is dependant on logs being cached

Removed Features due to Lack of Uptake

- INFORMATION_SCHEMA.XTRADB_RSEG
- Expanded program option modifiers
- pseudo_server_id
- CSV_MODE
- max_slowlog_files and max_slowlog_size
- innodb_show_verbose_locks

Let us know if you use any of those!

Part II

How it is done

Source Code

GitHub repositories, branches, GCA

Source Code

GitHub

https://github.com/percona/percona-server/

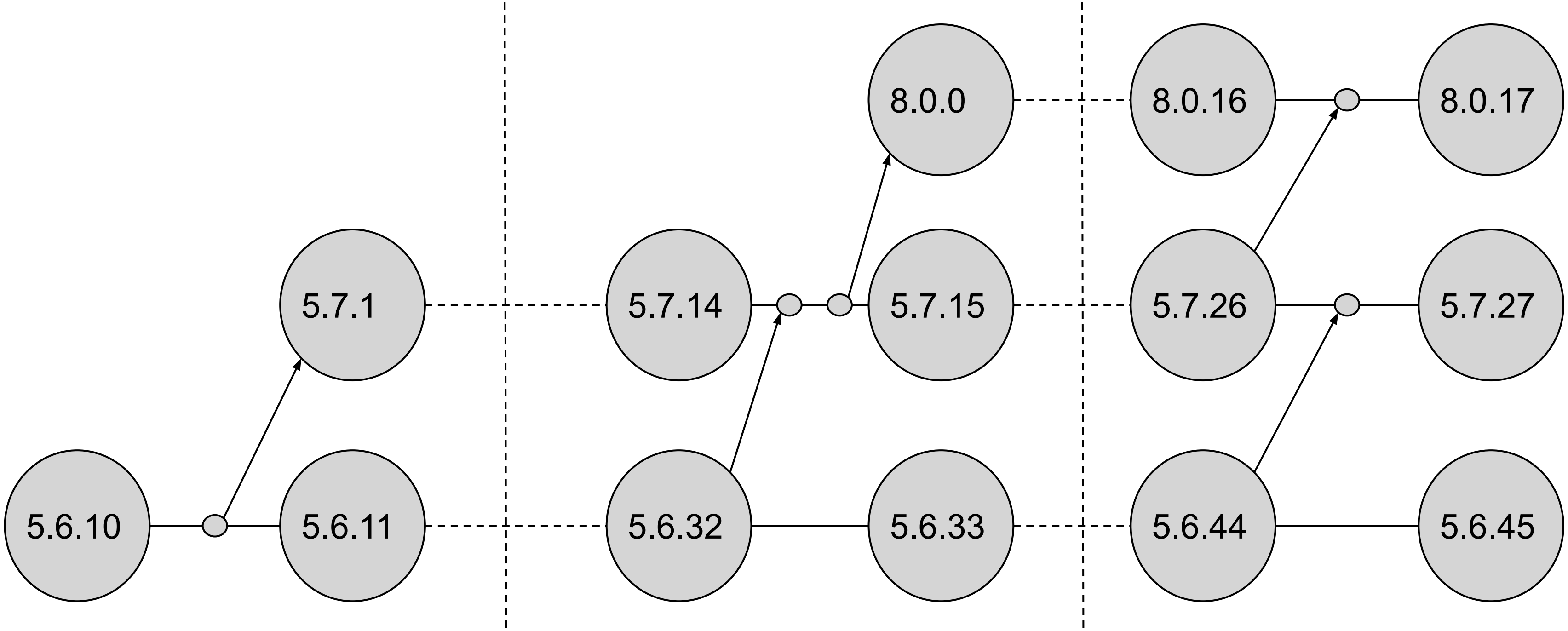
3 Active branches: 5.6, 5.7 and 8.0

```
git clone -b 5.6 https://github.com/percona/percona-server/
git clone -b 5.7 https://github.com/percona/percona-server/
git clone -b 8.0 https://github.com/percona/percona-server/
```

Git worktrees

https://git-scm.com/docs/git-worktree

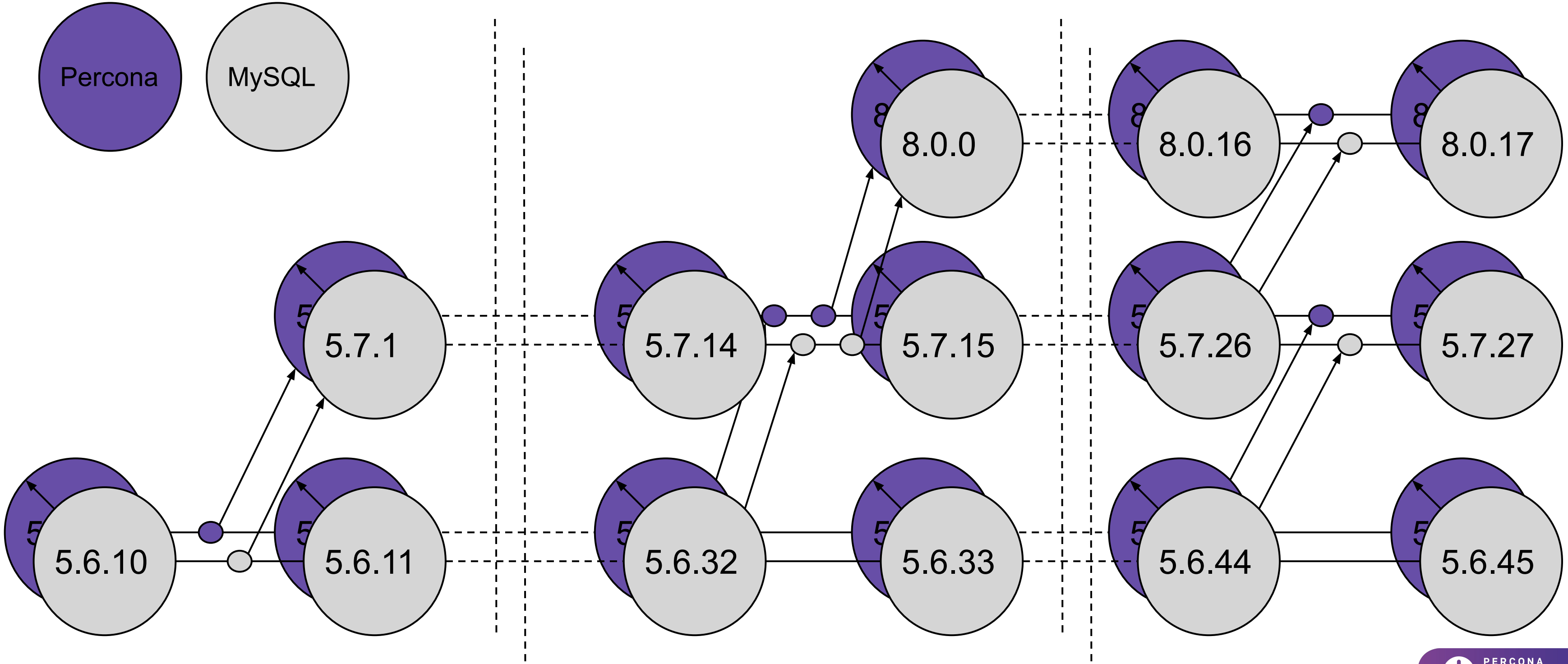
MySQL Server Source Code Diagram



Percona Server Source Code

- MySQL Server 5.7 has branched from 5.6 at some point.
- MySQL 5.6 is periodically merged into 5.7.
- MySQL Server 8.0 has branched from 5.7 at some point and therefore has also a common point with both 5.7 and 5.6.
- MySQL 5.7 is periodically merged into 8.0 and therefore almost everything from 5.7 and 5.6 is in 8.0.
- Percona Server 5.7 has branched from both MySQL Server 5.7 and Percona Server 5.6.
- Percona Server 8.0 has branched from both MySQL Server 8.0 and Percona Server 5.7.

Percona Server Source Code Diagram



Building Percona Server

Platforms, Compilers, CMake options, etc.

Platforms

RHEL / CentOS

- CentOS 6 64-bit
- CentOS 6 32-bit
- CentOS 7
- RHEL 8

Debian

- Debian 8.x Jessie
- Debian 9.x Stretch
- Debian 10.x Buster

Ubuntu

- Ubuntu 16.04 Xenial
- Ubuntu 18.04 Bionic
- Ubuntu 19.04 Disco

Checking beta releases earlier: CentOS 8 Beta, Ubuntu 19.10 Eoan Beta

Compilers

- GCC 4.8
- GCC 5.5
- GCC 6.5
- GCC 7.4
- GCC 8.3
- GCC 9.2

- Clang 4.0.1
- Clang 5.0.2
- Clang 6.0.1
- Clang 7.1.0
- Clang 8.0.1
- Apple LLVM 10.0.1 (clang-1001.0.46.4) (partially supported)

Checking compiler releases earlier before they are included in the next OS release.

CMake Options

https://dev.mysql.com/doc/refman/5.6/en/source-configuration-options.html
https://dev.mysql.com/doc/refman/5.7/en/source-configuration-options.html
https://dev.mysql.com/doc/refman/8.0/en/source-configuration-options.html

- -DBUILD_CONFIG: **mysql_release**
- -DFEATURE_SET: **community**
- -DCMAKE_BUILD_TYPE: **RelWithDebInfo** / **Debug**
- -DWITH_<THIRD_PARTY_LIBRARY>: **bundled** / **system** (-DWITH_SSL, -DWITH_ZLIB)

The most important one

- -DMYSQL_MAINTAINER_MODE=**ON**

Always using out-of-source builds

Fixing Bugs

Percona-specific, upstream, escalations, etc.

Bugs

Oracle MySQL Bug tracker
https://bugs.mysql.com/

Percona Server bug tracker
https://jira.percona.com/projects/PS/
upstream label
Upstream Bug URL field

Escalations have priority

Percona always submits our upstream fixes to Oracle under OCA

A number of Percona developers nominated MySQL Community Contributor Award:
https://blogs.oracle.com/mysql/mysql-community-contributor-award-program-2019-v2

Propagating the Fix Through Branches

- git chekcout -b ps-5.6-feature_branch <5_6_and_5_7_comon_commit>
- fix 5.6 code
- git commit

- git chekcout -b ps-5.7-feature_branch <5_7_and_8_0_comon_commit>
- git merge ps-5.6-beature_branch --no-commit
- apply 5.7-specific changes
- git commit

- git chekcout -b ps-8.0-feature_branch 8.0
- git merge ps-5.7-beature_branch --no-commit
- apply 8.0-specific changes
- git commit

Upstream Merges

Merging new code, reverting/combining Percona fixes, null merges, etc.

New Wave of Upstream Releases

5.6.45, 5.7.27 and 8.0.17 released on 2019-07-22

Merge blueprints (https://jira.percona.com/browse/PS-5363)

- List of Percona fixes to be reverted as the same issues were fixed by Oracle
- List of the MTR test cases for “security” issues Oracle fixed with code-only changes without proper test coverage.

- git fetch origin --tags
- git remote add upstream https://github.com/mysql/mysql-server.git
- git fetch upstream --tags
- git checkout -b ps-8.0.17-merge 8.0
- git merge mysql-8.0.17 --no-commit

3-step Merge

Make sure you have 'merge.conflictstyle = diff3' in your '.gitconfig'.

This will allow to see not only ours ('<<<') and theirs ('>>>') changes in files with merge conflicts but baseline ('|||') as well.

1. **“Documentation update”** commit
   Updating man/* from the upstream release tarball
2. **“Merge with conflicts”** commit
   git merge mysql-8.0.17
   git add . (also includes files with conflict markers)
3. **“Conflicts resolved”** commit
   git checkout --conflict=merge <list of conflict files from the previous step>
   Resolve actual conflicts

All 3 commits are parts of a single Pull Request

Code Checkers

Sanitizers, Valgrind, etc.

Code Checkers

GCC **Address Sanitizer** https://gcc.gnu.org/onlinedocs/gcc/Instrumentation-Options.html
Clang **Address Sanitizer** https://clang.llvm.org/docs/AddressSanitizer.html
GCC **Undefined Behavior Sanitizer** https://gcc.gnu.org/onlinedocs/gcc/Instrumentation-Options.html
Clang **Undefined Behavior Sanitizer** https://clang.llvm.org/docs/UndefinedBehaviorSanitizer.html
Clang **Memory Sanitizer** https://clang.llvm.org/docs/MemorySanitizer.html

Valgrind http://www.valgrind.org

In Percona Server Jira:

- **asan** label - 63 issues
- **valgrind** label - 62 issues

Always use the latest versions: Sanitizers from GCC 9.2 / Clang 8.0.1, Valgrind 3.15.0

Continuous Integration

TravisCI, CircleCI, Jenkins, MTR runs, pquery, etc.

Continuous Integration

CircleCI

- makes sure our code is properly **clang-formatted**

TravisCI

- Builds our code with different compilers
- Builds our code in Debug / Release mode
- Checks various cmake options (system / bundled)

Jenkins (AWS EC2 spot instances)

- Builds and run MTR test cases on different platforms (currently 8 in 2 modes)

pquery (https://github.com/Percona-QA/pquery)

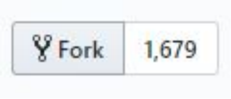
- multi-threaded test program created to stress test the MySQL server (in any flavor), either randomly or sequentially, for QA purposes

Summary

3 steps to build a successful software product

All You Need is Fork

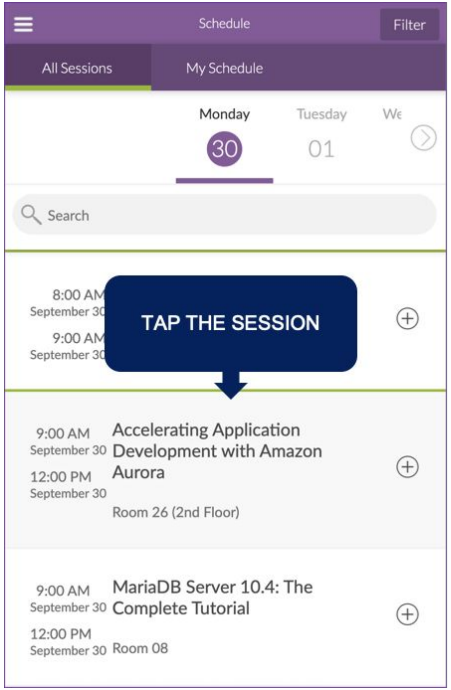
mysql / mysql-server



1. Fork mysql-server repository
2. Invest 13 years of development effort
3. Profit :)

Questions?

Rate My Session

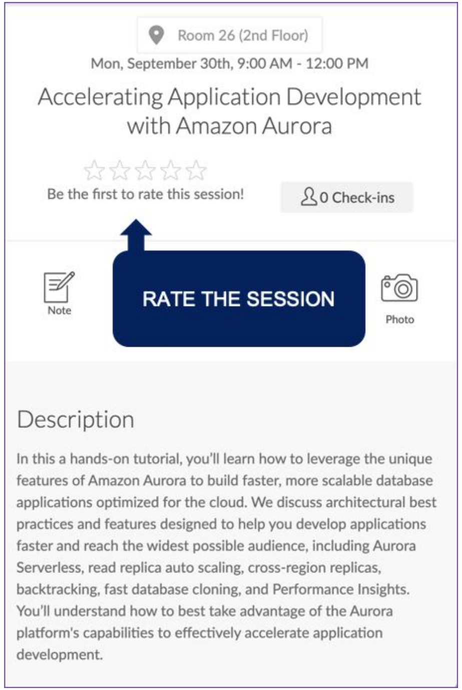


TAP THE SESSION

8:00 AM
September 30
9:00 AM
September 30

9:00 AM
September 30
12:00 PM
September 30
Accelerating Application Development with Amazon Aurora
Room 26 (2nd Floor)

9:00 AM
September 30
12:00 PM
September 30
MariaDB Server 10.4: The Complete Tutorial
Room 08



Room 26 (2nd Floor)

Mon, September 30th, 9:00 AM - 12:00 PM

Accelerating Application Development with Amazon Aurora

☆☆☆☆☆

Be the first to rate this session!

RATE THE SESSION

Description

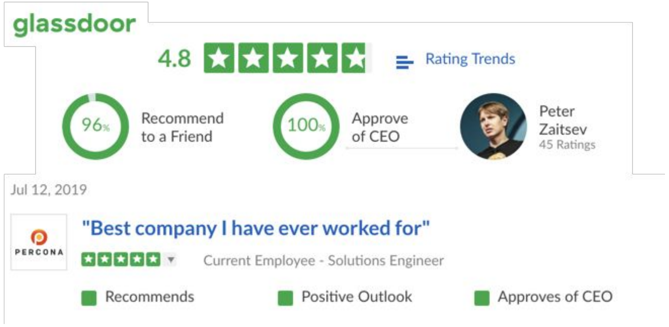
In this a hands-on tutorial, you'll learn how to leverage the unique features of Amazon Aurora to build faster, more scalable database applications optimized for the cloud. We discuss architectural best practices and features designed to help you develop applications faster and reach the widest possible audience, including Aurora Serverless, read replica auto scaling, cross-region replicas, backtracking, fast database cloning, and Performance Insights. You'll understand how to best take advantage of the Aurora platform's capabilities to effectively accelerate application development.

We’re Hiring!

Percona’s open source database experts are true superheroes, improving database performance for customers across the globe.

Our staff live in nearly 30 different countries around the world, and most work remotely from home.

Discover what it means to have a Percona career with the smartest people in the database performance industries, solving the most challenging problems our customers come across.



glassdoor

4.8 ★★★★★ Rating Trends

96% Recommend to a Friend

100% Approve of CEO

Peter Zaitsev
45 Ratings

Jul 12, 2019

PERCONA

"Best company I have ever worked for"

★★★★★ Current Employee - Solutions Engineer

🟩 Recommends 🟩 Positive Outlook 🟩 Approves of CEO

Thank You