
# InnoDB: Status, Architecture, and Latest Enhancements

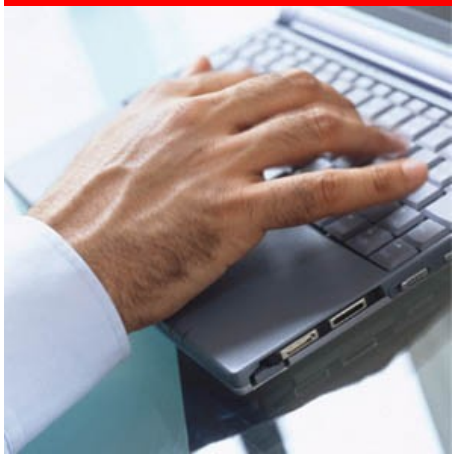
*COLLABORATE 11, April 13, 2011*

*Calvin Sun, Sr. Manager, Oracle*

MySQL™

The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions.
The development, release, and timing of any features or functionality described for Oracle's products remains at the sole discretion of Oracle.

# Agenda



- Introduction to InnoDB
- InnoDB Architecture
- What We Have Done?
- What's in the Plan?
- Q&A

MySQL™
INNODB®

# Introduction to InnoDB

# What is InnoDB?

- The most popular transactional storage engine for MySQL
- Architected and written by Dr. Heikki Tuuri
- Followed Gray & Reuter's "*Transactions Processing: Concepts & Techniques*"; also modeled on Oracle architecture
- Added unique subsystems/features
  - Doublewrite buffer
  - Change buffering
  - Adaptive hash index
- Many key features for performance and data integrity

# InnoDB Key Characteristics


- Row-level locking
- Multi-version concurrency control
- Efficient indexing (covering indexes)
- Fuzzy checkpoint
- Adaptive hash indexing
- Table compression
- Group commit
- Fast DDL operations
- Insert buffering

# InnoDB Key Characteristics

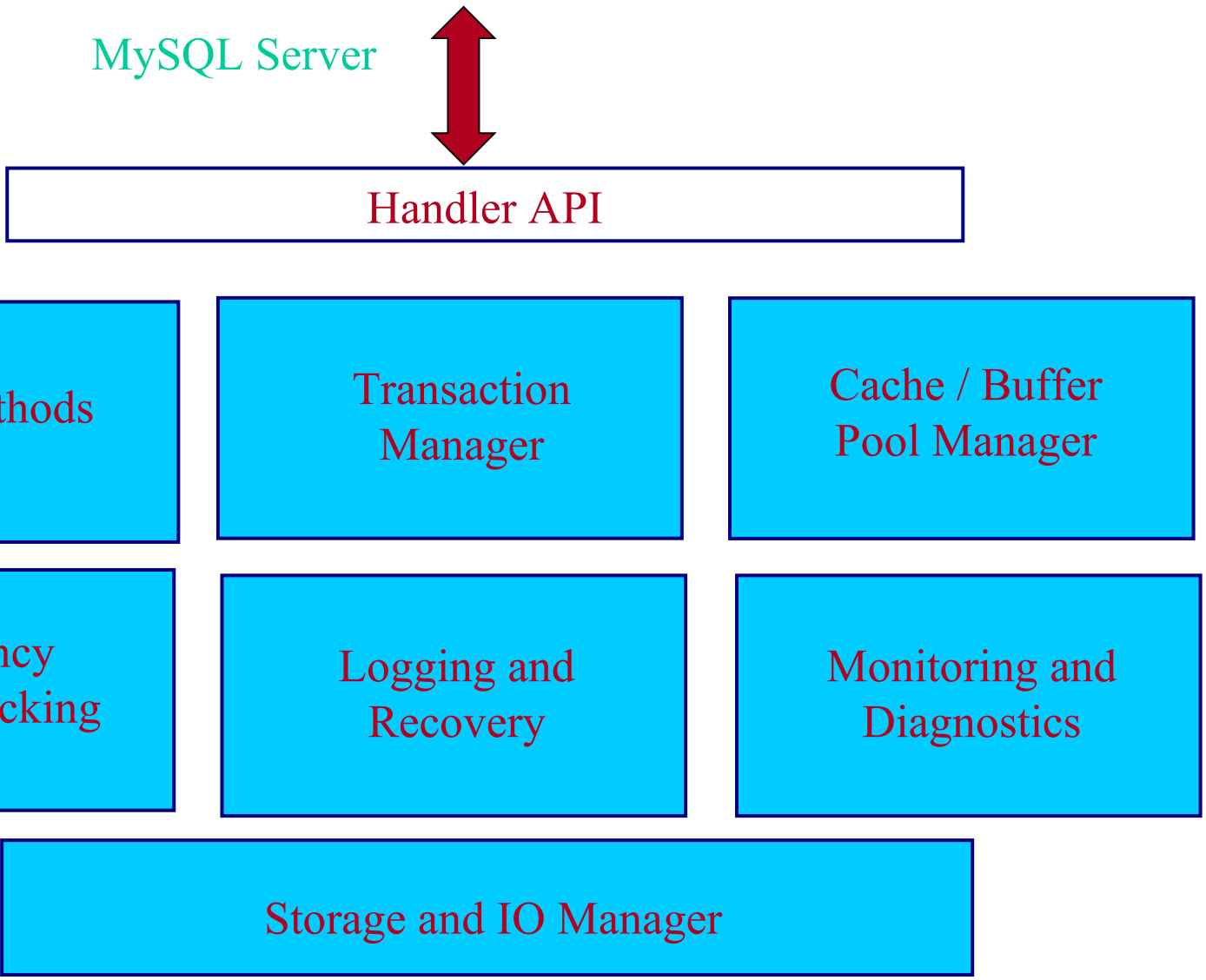

- ACID-compliant transactions
- Two phase commit
- Automatic crash recovery
- Doublewrite buffer
- Referential integrity
- Online backup with MySQL Enterprise Backup
- Well written, well tested code

# InnoDB Key Characteristics

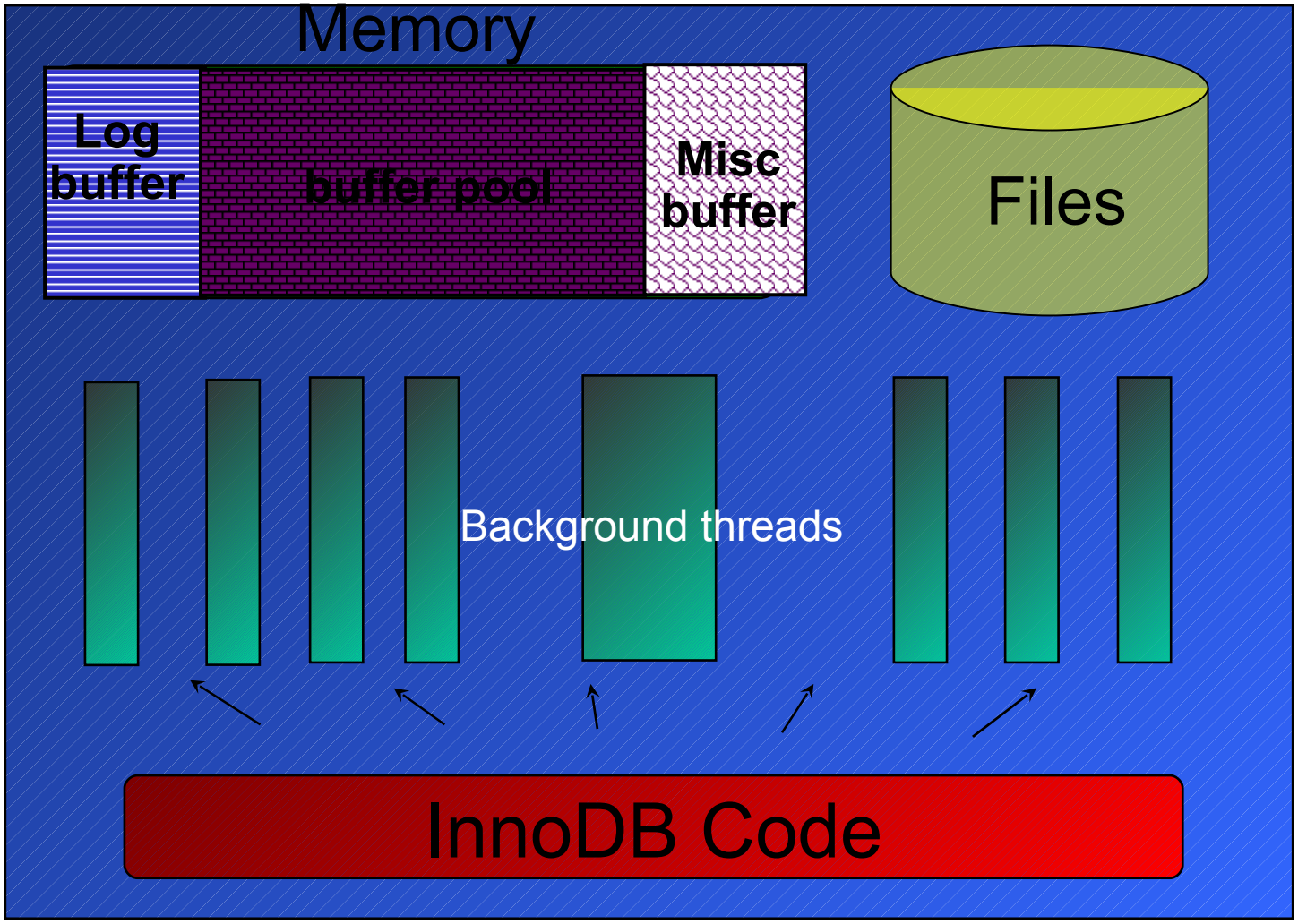

- Distributed by MySQL since 2001
- Deployed by millions of users
- Wide use in large-scale enterprise customers

# InnoDB Architecture

# InnoDB Architecture: Component Model



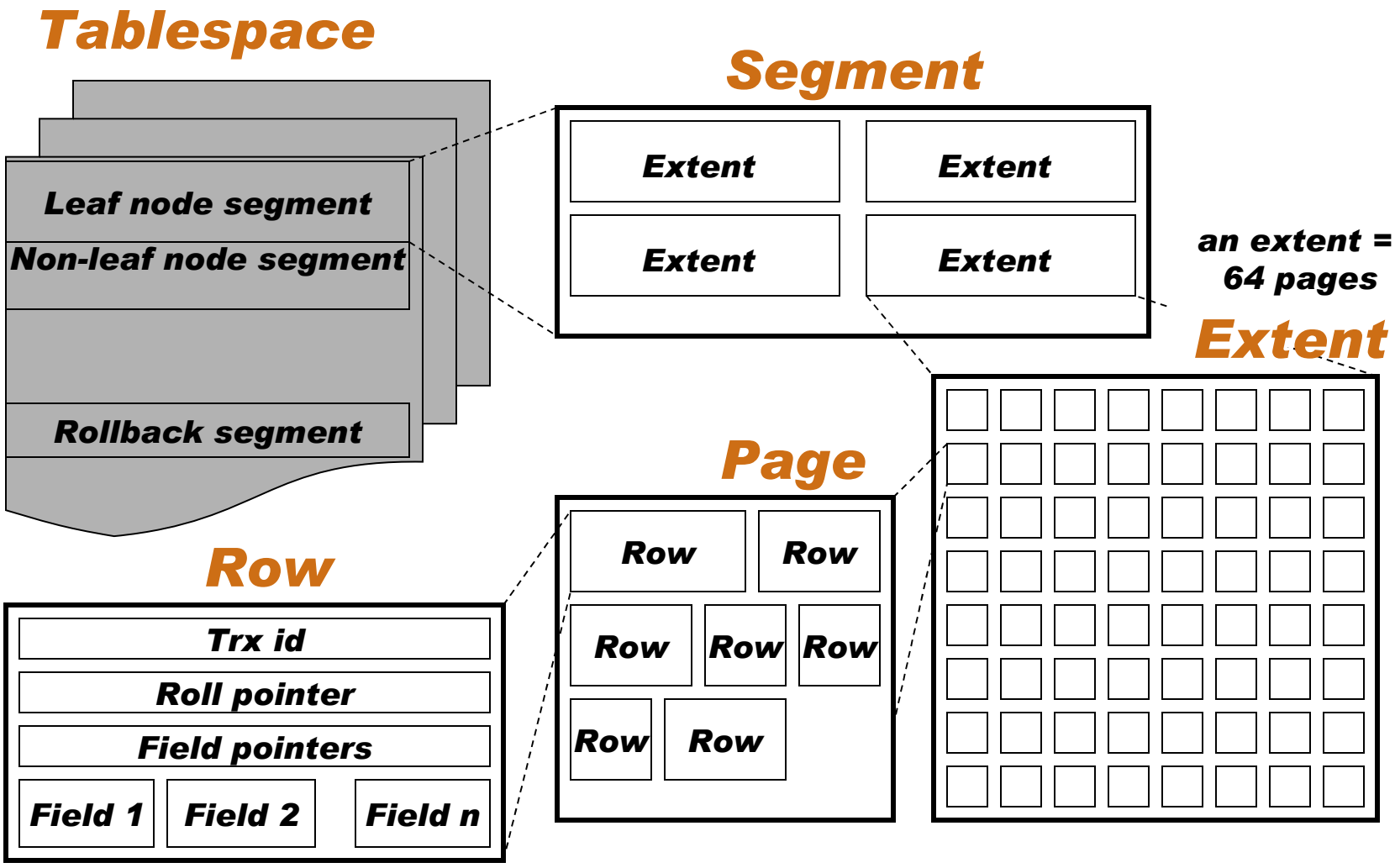
# InnoDB Architecture: Runtime Model




Threads:
-master
-read io
-write io
-ibuf io
-log io
-monitor
-and more

Buffer Pool:
-data
-index
-undo
-adaptive hash index

# InnoDB Architecture: Data Storage



# What We Have Done?

# InnoDB Plugin 1.0 for 5.1

- Performance and Scalability Enhancements
  - Faster locking for improved scalability
  - Using operating system memory allocators
  - Multiple background I/O threads
  - Group commit
  - Controlling master thread I/O rate
  - Controlling flushing rate of dirty pages
  - Changes in read ahead algorithm
  - Using a portable PAUSE to InnoDB spin loop
  - Making buffer cache scan resistant
  - Improvements to crash recovery performance

# InnoDB Plugin 1.0 for 5.1

- Page compression
- Fast index creation
- Information schema tables
- Other changes for flexibility, easy of use

# InnoDB 1.1 for MySQL 5.5

- Performance and scalability
  - Multiple buffer pool instances
  - Multiple rollback segments
  - Extended change buffering with delete buffering and purge buffering
  - Improved purge scheduling
  - Improved log sys mutex
  - Separate flush list mutex
  - Native async I/O support on Linux
  - Windows performance improvements

# InnoDB 1.1 for MySQL 5.5

- Monitoring & Diagnostics
  - Performance schema for InnoDB
  - Improved InnoDB transaction reporting
  - Log start and end of InnoDB buffer pool initialization to the error log
- UTF-32 support
- Significantly reduced the number of kernel objects (Windows)

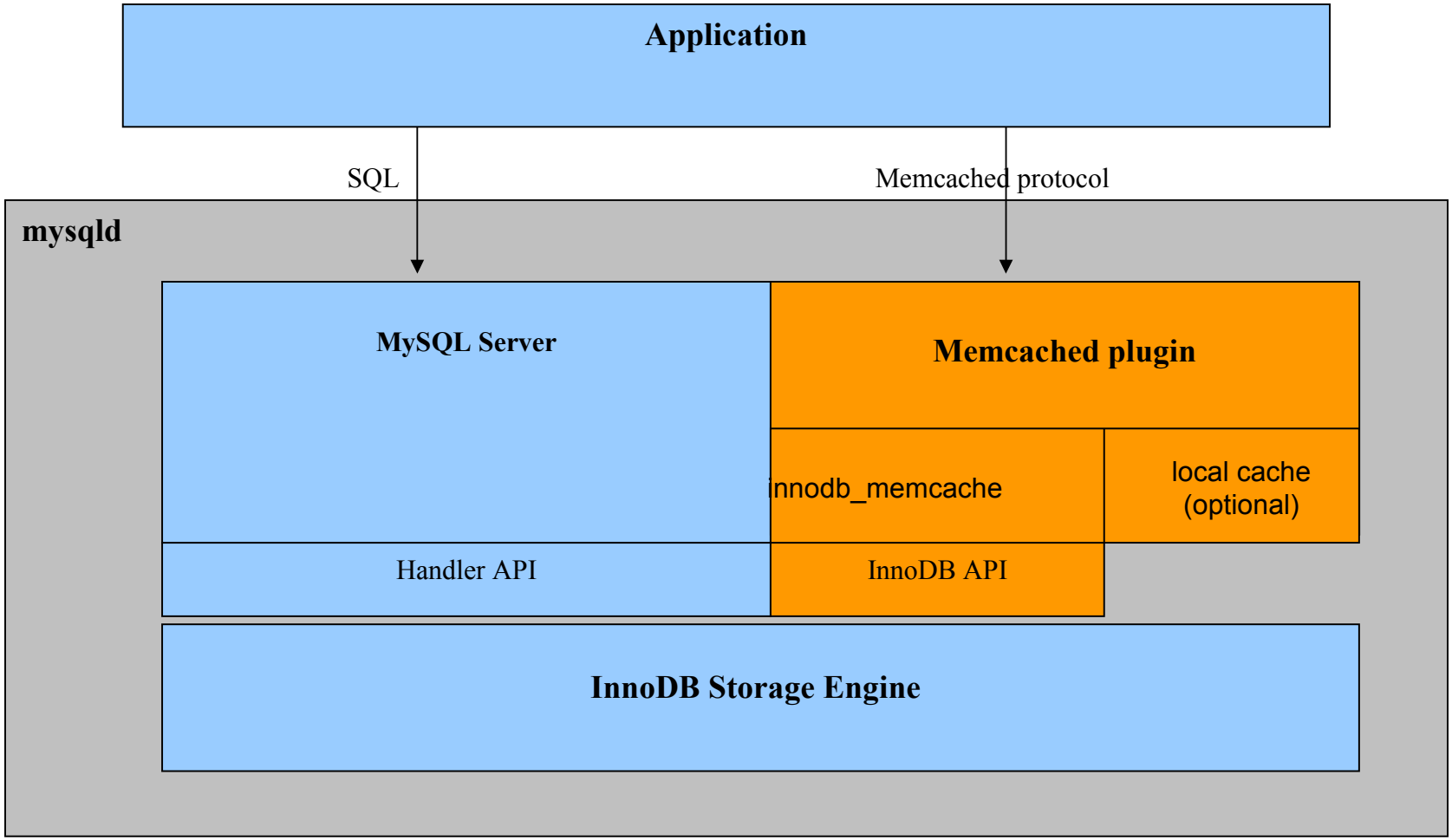
# InnoDB 1.2 for MySQL 5.6

- Performance and scalability
  - Split the kernel mutex
  - Multi threaded purge
  - Use rw_locks for page_hash
  - Add 'page_cleaner' thread to flush dirty pages
  - Ibuf merge rate improvement
  - Configurable data dictionary cache

# InnoDB 1.2 for MySQL 5.6

- Monitoring & diagnostics
  - InnoDB Information Schema Metrics Table
  - Information schema system tables for InnoDB
  - Information schema table for InnoDB buffer pool
  - InnoDB: report all deadlocks
- InnoDB persistent optimizer statistics
- MRR/ICP support for InnoDB

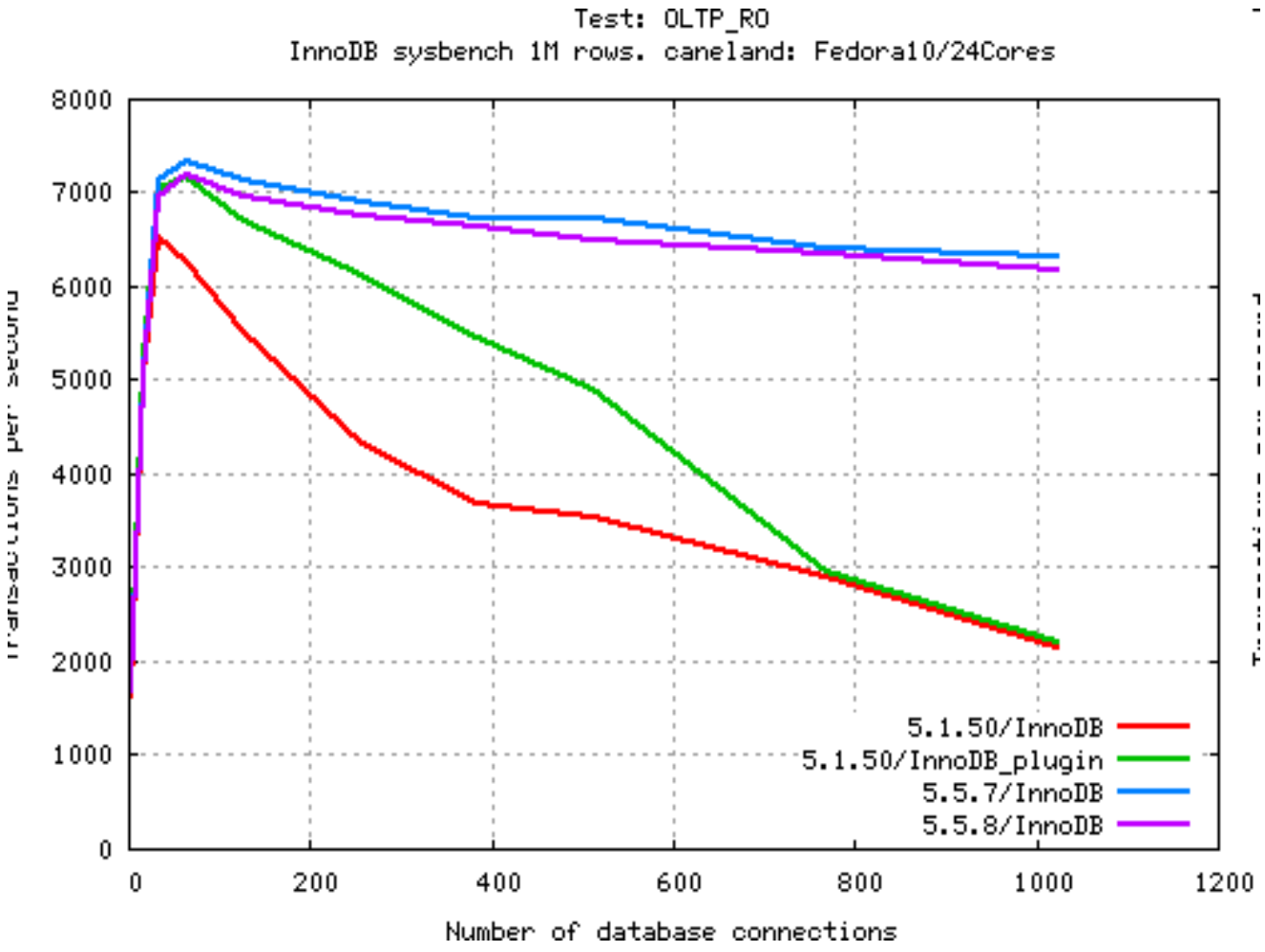
# NoSQL to InnoDB with memcached
# MySQL Labs



# NoSQL to InnoDB with memcached

- Features
  - Memcached as a daemon plugin of mysqld: both mysqld and memcached are running in the same process space, with very low latency access to data
  - Memcapable: support both memcached ascii protocol and binary protocol
  - Support multiple columns: users can map multiple columns into “value”
  - Optional local caching: three options – “innodb-only”, “cache-only”, and “caching”
  - Batch operations: batch read/write operations and commit together, resulting in better performance

# MySQL 5.5 Benchmarks – Linux



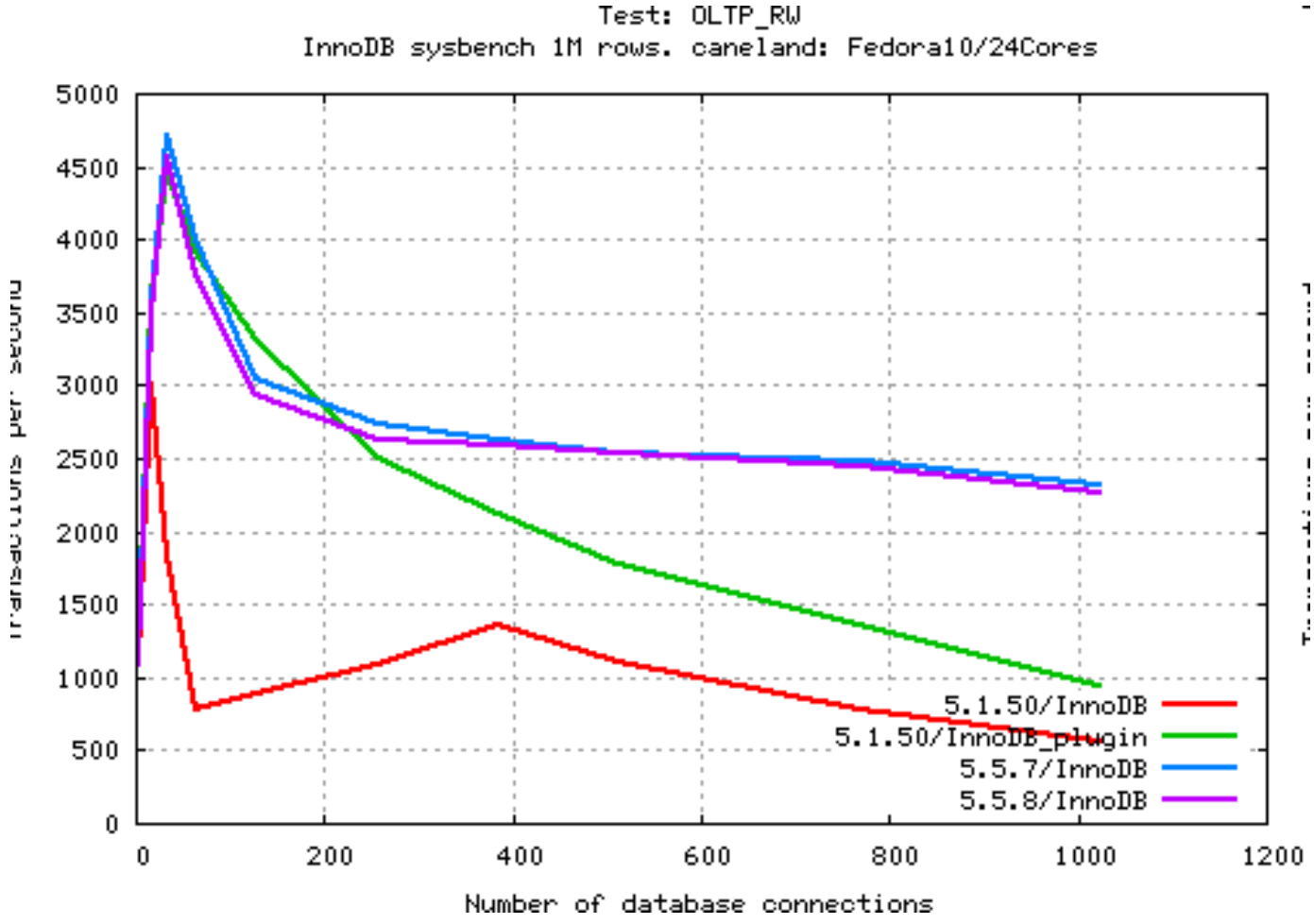

**MySQL 5.5.8**
(InnoDB 1.1)

**MySQL 5.1.50**
(InnoDB Plug-in)

**MySQL 5.1.50**
(InnoDB built-in)

Intel Xeon X7460 x86_64
4 CPU x 6 Cores/CPU
2.66 GHz, 32GB RAM
Fedora 10

# MySQL 5.5 Benchmarks – Linux



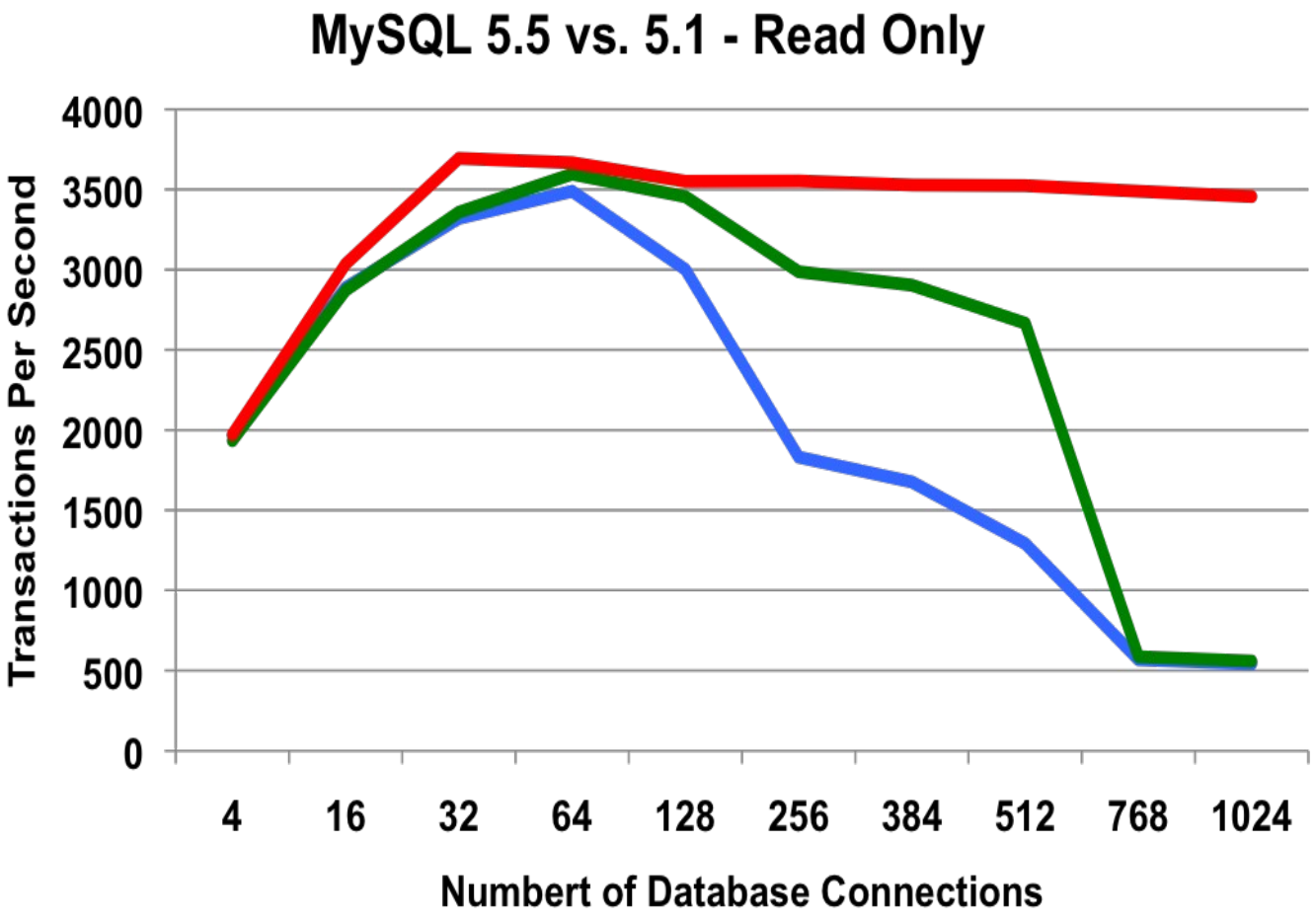

**MySQL 5.5.8**
(InnoDB 1.1)

**MySQL 5.1.50**
(InnoDB Plug-in)

**MySQL 5.1.50**
(InnoDB built-in)

Intel Xeon X7460 x86_64
4 CPU x 6 Cores/CPU
2.66 GHz, 32GB RAM
Fedora 10

# MySQL 5.5 Benchmarks – Windows



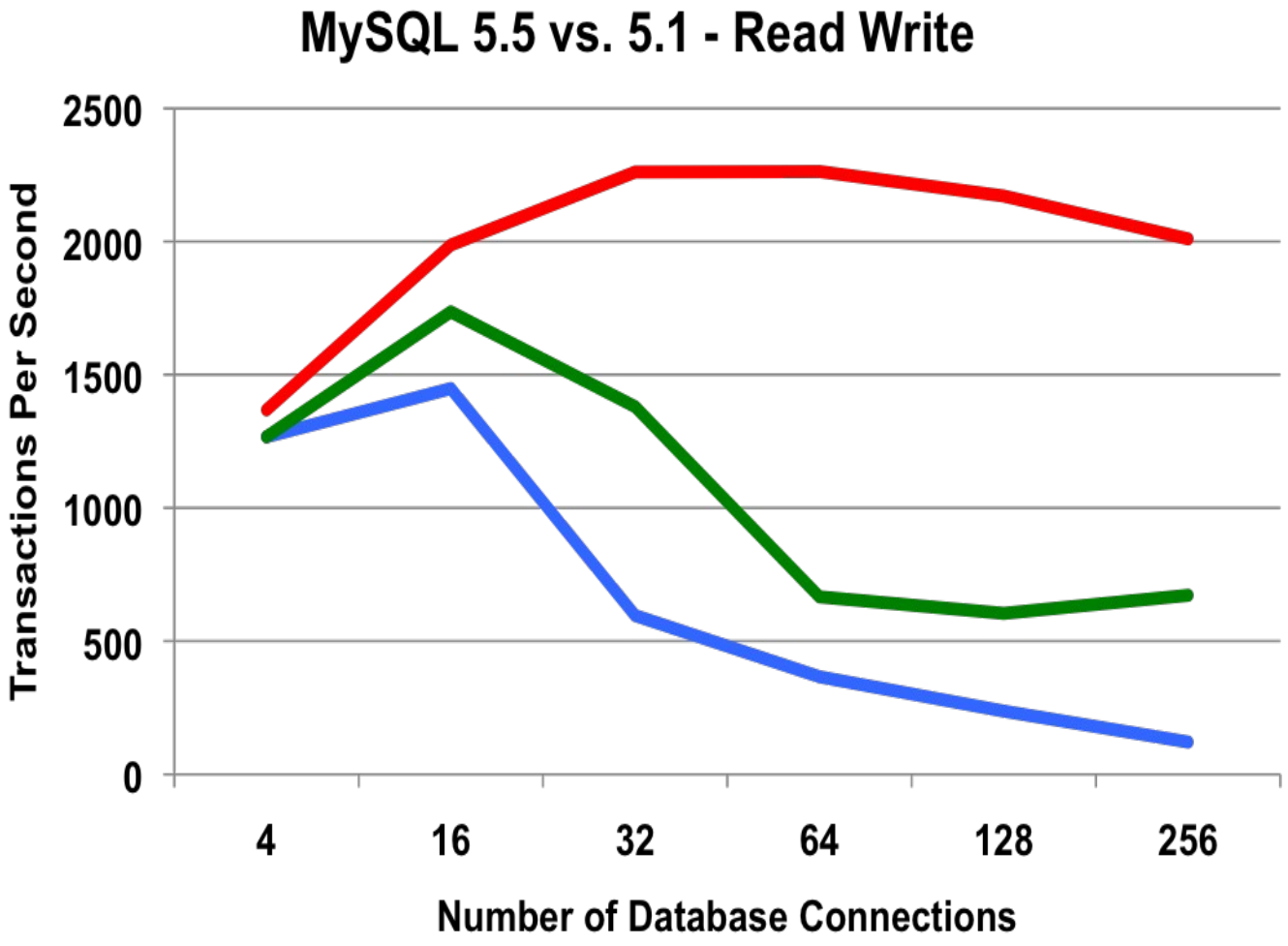

**MySQL 5.5.6**
(InnoDB 1.1)

**MySQL 5.1.50**
(InnoDB Plug-in)

**MySQL 5.1.50**
(InnoDB built-in)

Intel x86_64
4 CPU x 2 Cores/CPU
3.166 GHz, 8GB RAM
Windows Server 2008

# MySQL 5.5 Benchmarks – Windows




**MySQL 5.5.6**
(InnoDB 1.1)

**MySQL 5.1.50**
(InnoDB Plug-in)

**MySQL 5.1.50**
(InnoDB built-in)

Intel x86_64
4 CPU x 2 Cores/CPU
3.166 GHz, 8GB RAM
Windows Server 2008

# What's in the Plan?

# InnoDB Focus

- Performance and scalability
- Online operations
- Monitoring & diagnostics
- Long outstanding requested features

# InnoDB Roadmap

- Performance and Scalability
  - Auto-extension of files in background
  - Group commit with binlog
  - Improvements on InnoDB flushing
  - Increase the max size of redo log files
  - Improve thread scheduling
  - Preload buffer pool
  - Fast checksum

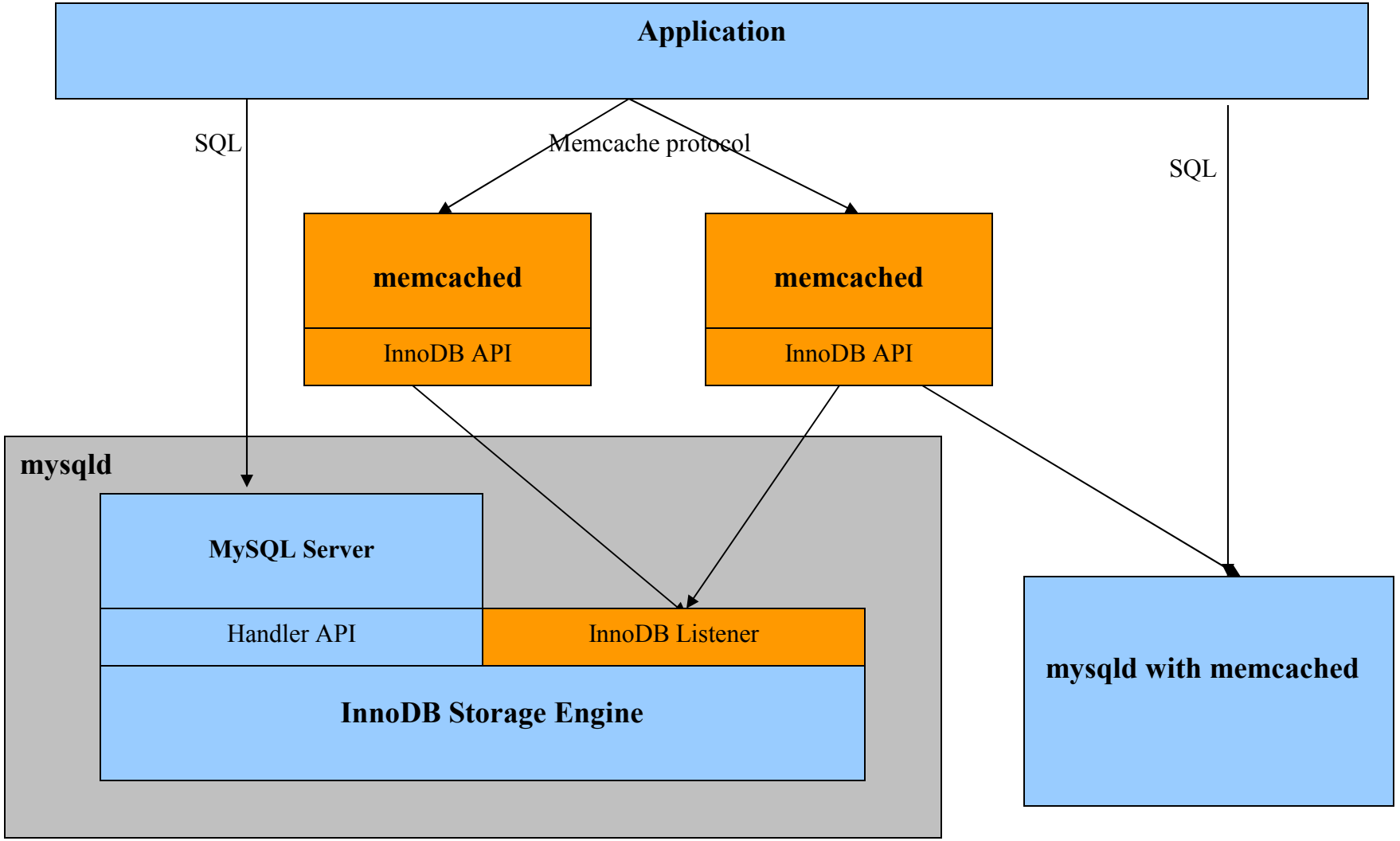
# InnoDB Roadmap

- Optimized for SSD
  - Various page sizes support
  - Separate tablespaces for undo log, doublewrite buffer, and change buffer
  - Flexible tablespaces management
  - Additional tuning
- Online DDLs
  - Online add index
  - Online drop index
  - Online index rebuild

# InnoDB Roadmap

- Monitoring and Diagnostics
  - More metrics counters
  - Additional information schema tables and performance statistics
  - DTrace probes for InnoDB
- Full-text search support for InnoDB
- Lift the limit of index key prefixes

# InnoDB with Remote Memcached



# Resources

**Transactions on InnoDB** (blogs.innodb.com)

- NoSQL to InnoDB with Memcached
- Get started with InnoDB Memcached Daemon plugin
- MySQL 5.6: InnoDB scalability fix – Kernel mutex removed
- MySQL 5.6: Multi threaded purge
- MySQL 5.6: Data dictionary LRU
- Information Schema for InnoDB System Tables
- Introducing page_cleaner thread in InnoDB
- InnoDB Persistent Statistics at last
- Tips and Tricks for Faster DDL

**MySQL Developer Zone** (http://dev.mysql.com/)

**DimitriK's Weblog** (http://dimitrik.free.fr/blog/index.html)


## San Francisco 2011

October 2–6

Register and Exhibit Now! oracle.com/openworld