



# INNODB<sup>®</sup>

Transactional Storage for MySQL  
FAST. RELIABLE. PROVEN.

## Innovative Technologies for Performance and Data Protection

Ken Jacobs, VP Product Strategy, Oracle

**INNOBASE**

# What is InnoDB?

- A pluggable storage engine for MySQL
- Provides high performance and reliability for mission-critical applications
  - efficient, effective storage management
  - transactions and multi-user concurrency
  - referential integrity
- A standard part of MySQL Enterprise Server and MySQL Community Server editions
- Distributed & supported by MySQL since 2001
- Most popular & proven transactional engine

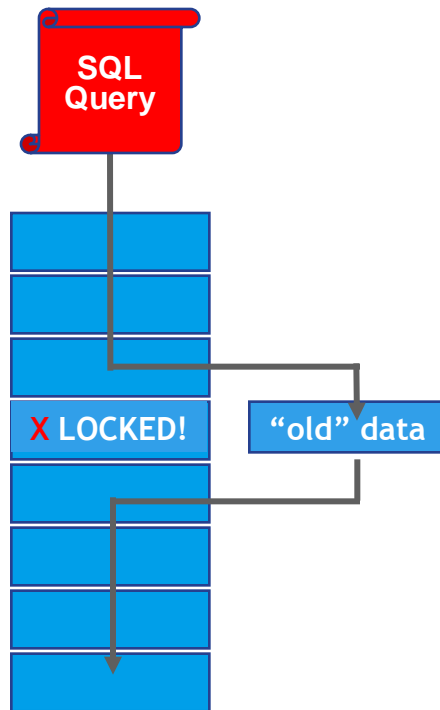
# History of InnoDB & Innobase

- Dr. Heikki Tuuri, was professor at Helsinki University
  - simultaneously with Linus Torvalds
- started work on InnoDB in 1994
- founded Innobase Oy in 1995
- Signed contract with MySQL AB in 2000
  - First InnoDB shipped with MySQL 4.x in 2001
- Acquired by Oracle in October, 2005

# Key Basic Features of InnoDB for Performance and Data Integrity

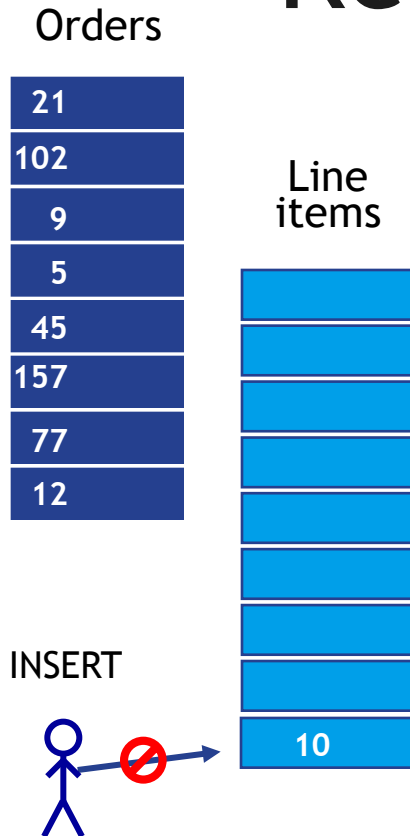
- Based on sound database computer science
  - Gray & Reuters definitive text on database design
  - Internal subsystems isolated from one another
- Architected and written by one person
- Borrowed some ideas from Oracle
- Key characteristics for performance
  - Row-level locking, multi-version concurrency control
  - Multi-threaded, direct i/o, buffer pool, efficient indexing ...
- Key features for data protection, integrity, reliability
  - Transactions / ACIDity
  - Automatic crash recovery
  - Referential integrity

# InnoDB Transactions & Locking



- Full transaction support
  - atomicity, consistency, isolation, durability
- SQL-standard isolation levels
  - SERIALIZABLE
  - REPEATABLE READ
  - READ COMMITTED
  - RED UNCOMMITTED
- Unlimited row-level locking
- Multi-version read-consistency
- Automatic deadlock detection

# InnoDB and Referential Integrity

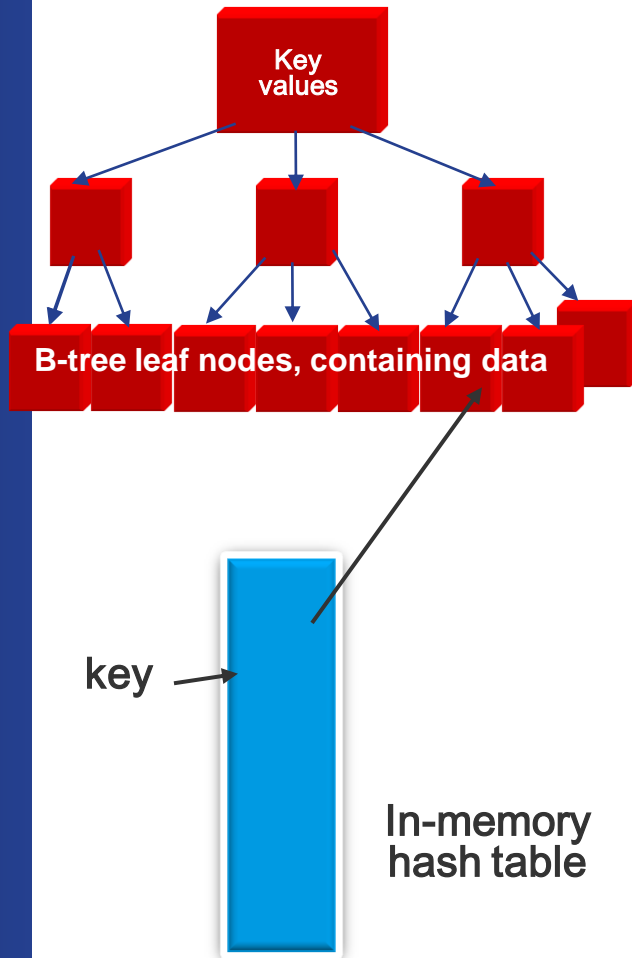


- Use SQL to define relationships between tables (foreign keys)
- InnoDB prevents changes that violate those relationships
- Supports SQL clauses in CREATE TABLE
  - ON UPDATE  
RESTRICT | CASCADE | SET NULL
  - ON DELETE  
RESTRICT | CASCADE | SET NULL
- Protects data integrity

# Innovative InnoDB Features for Performance & Reliability

- Adaptive Hash Indexes
- Insert Buffer
- Doublewrite buffer
- The newest “stuff”
  - InnoDB Plugin
  - MySQL 5.4

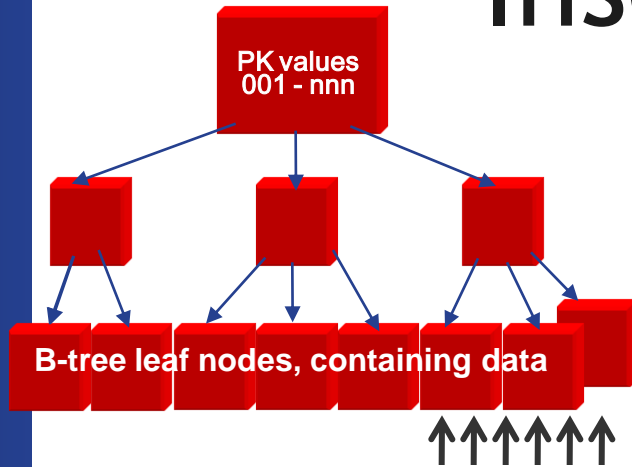
# Adaptive Hash Indexing



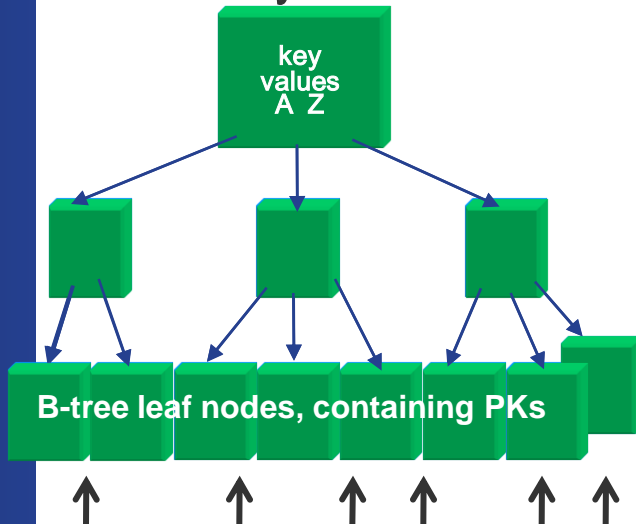
- Automatically creates hash index on prefix of key for frequent queries
- Unique to InnoDB, approximates in-memory dbs
- Performance benefit of adaptive hash index:
  - Modified read/write sysbench: approx 2X
  - Secondary index joins: up to 5X

# Insert Buffering

Clustered index



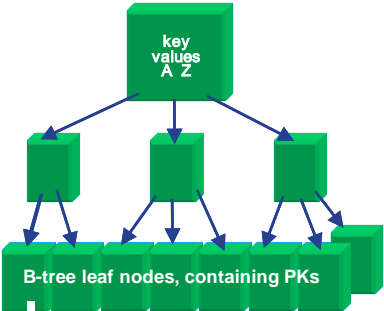
Secondary index



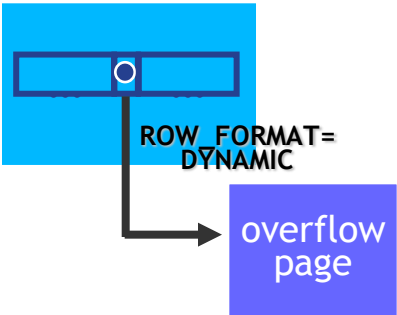
**INNOBASE**

- Defers writes to secondary indexes on INSERTs
- Unique to InnoDB, saves random writes, improves insert speed
- Performance benefit of insert buffering:
  - [mysqlha.blogspot.com/2008/12/innodb-insert-performance.html](http://mysqlha.blogspot.com/2008/12/innodb-insert-performance.html)
  - as much as 7.2x faster than the theoretical rate of inserts in a "normal" DBMS

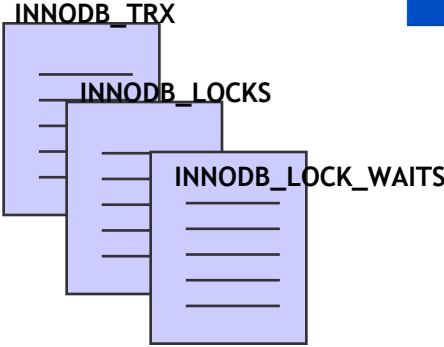
# InnoDB Plugin Features



fast index creation



new row storage format



Info schema tables

... and MORE!

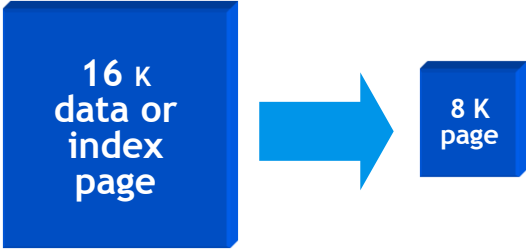
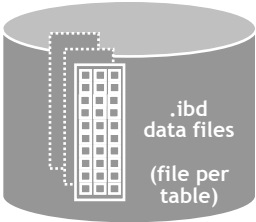


table compression

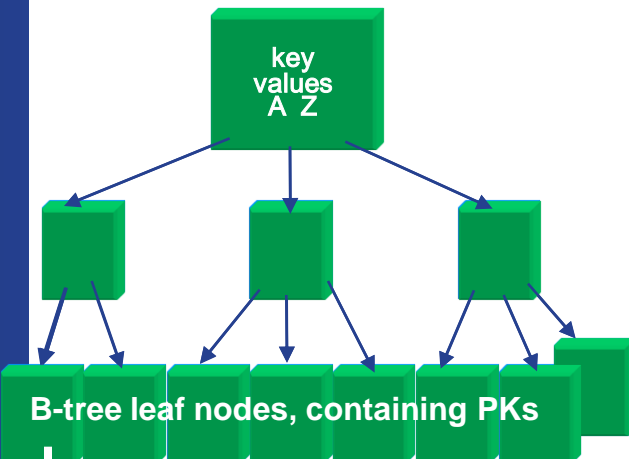
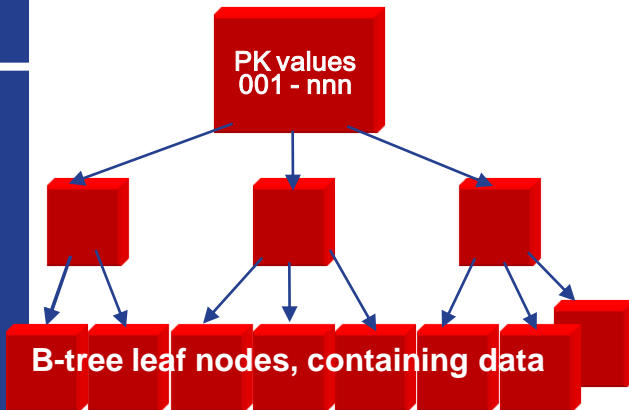


file format management

# InnoDB Plugin Features



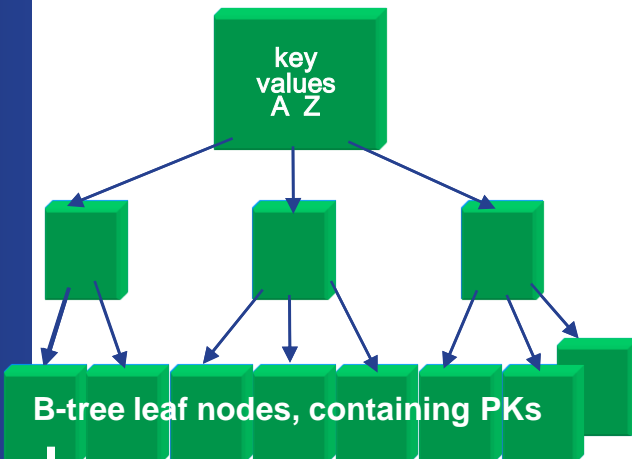
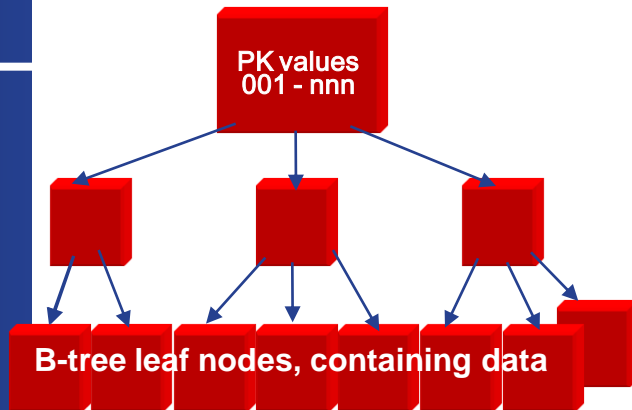
# InnoDB Fast Index Create



**INNOBASE**

- MySQL 5.1 rebuilds the entire table, row-by-row, to create a new secondary index
- The InnoDB Plugin builds just the new indexes, not the entire table
  - Sorts data on secondary key
  - Inserts the rows into the index
- MUCH faster, since the table is not re-created and because the data are inserted in order
- DROP INDEX for secondary index is even FASTER; data dictionary change only

# Fast Index Create - Performance




**INNOBASE**

**Creation  
Time**

**File size**    **Approx 3 GB**

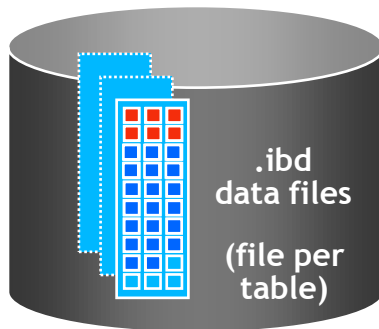
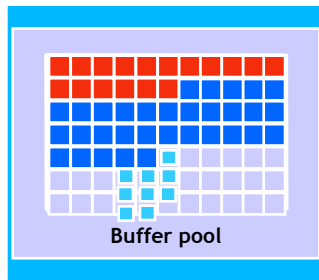
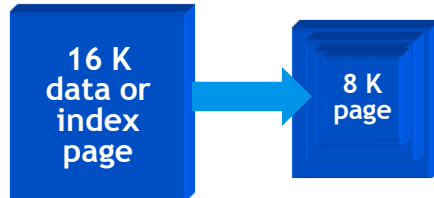
**MySQL 5.1**    **88 minutes**

**InnoDB Plugin**    **8 minutes** 😊

- Fast index creation reduced size of secondary indexes by 30%
-  Faster index scans!
- Even faster speedup for larger tables!

Test performed by  
Peter Zaitsev of Percona

# InnoDB Table Compression



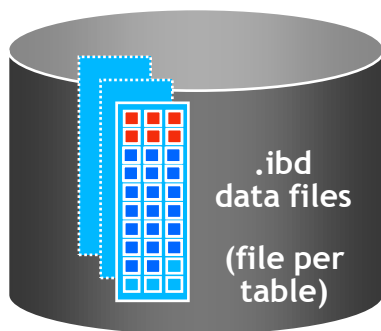
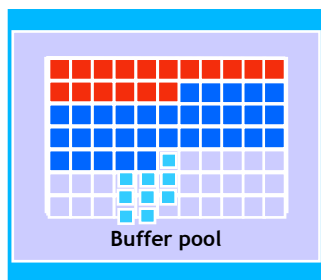
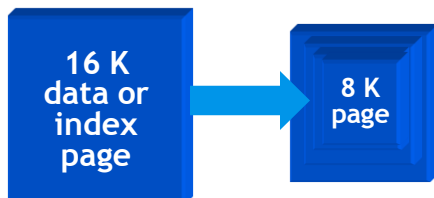
**INNOBASE**

- Specify compressed page size per table

```
CREATE TABLE t ...  
KEY_BLOCK_SIZE=8
```

- InnoDB tries to compress data & index pages from normal 16 kB size to specified compressed page size
- Typical values for the compressed page size are 8 kB and 4 kB
- InnoDB keeps some uncompressed pages in the buffer pool with compressed copy
- Requires `innodb_file_per_table=1` in `my.cnf` and new **“Barracuda”** `innodb_file_format`

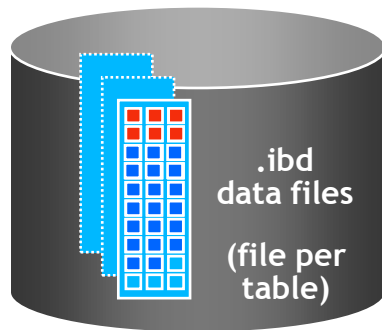
# Performance of Compression



**INNOBASE**

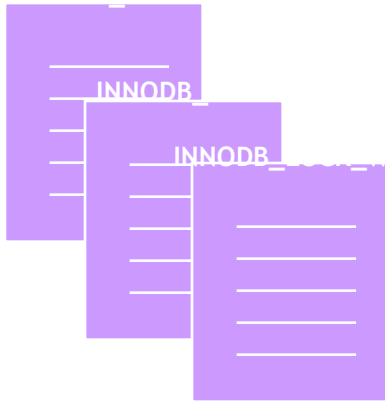
- Sysbench transactional workload may run up to 50% faster compressed vs. non-compressed
  - sysbench data compresses well
- But, DBT2 runs 30 % slower with compressed tables
  - DBT2 data does not compress well
- Percona has run benchmarks on the InnoDB plugin
  - Typical MySQL workloads: Forum, Social Network, Clickstream
- File size 50% smaller but load time increased 2x (not using fast index creation!)
- Query speed improved by 2x

# File Format Management



- Pre-plugin format: “Antelope”
- Enable with `innodb_file_format`
- New “Barracuda” format enables compression, `ROW_FORMAT=DYNAMIC`
  - Fast index creation, other features do not require Barracuda file format
  - “Antelope” precludes new features
- Built-in InnoDB can access “Antelope” databases, but not “Barracuda” databases
- Preserves ability to downgrade easily

# Information Schema Tables



- Query InnoDB locks, other status
  - Previously available only in cryptic output of `SHOW INNODB STATUS`
- In Plugin, Info Schema tables must be installed:  
`INSTALL PLUGIN INNODB_LOCKS  
SONAME 'ha_innodb.so', etc.`
- Info Schema tables reside in MySQL's “`INFORMATION_SCHEMA`” pseudo-database
- Example:  

```
mysql -uroot -p...  
mysql> USE INFORMATION_SCHEMA;  
mysql> SELECT * FROM INNODB_TRX;
```

# Future of InnoDB & Innobase

- Better, faster, more scalable
- More functional, even more reliable
- “Enhanced” great partnership with Sun/MySQL



and  
ANNOUNCING ....

# Embedded InnoDB

- For developers, device makers, ISVs, OEMs
- For “edge” applications, devices, robotics, ...
- The high-performance, reliability and rich functionality of InnoDB
- Flexible, powerful, programmatic API
- Operational characteristics for embedded applications

# Embedded InnoDB: Performance and Reliability

- All the capabilities of InnoDB
  - Multi-core scalability, multi-version concurrency with row-level locking, deadlock detection, automatic crash recovery, clustered & secondary indexes, adaptive hashing, insert buffering, fast index creation, compression, etc., etc. ...
- Based on the InnoDB Plugin

# Embedded InnoDB

## Flexible, Powerful API

- Low-level C/C++ non-SQL, ISAM-like API
- Functional interfaces for all data access & management, configuration operations
  - create/drop tables & indexes
  - ins/upd/del/query via primary, secondary indexes
  - cursor management, transaction management
  - start, stop and configure
- Relational db features: strong data types, schema management
  - Differs from key-value pair embedded dbs

# Embedded InnoDB

## Operational Flexibility

- Designed for environments requiring automatic operations w/o DBA
  - small footprint (838KB)
  - automatic configuration
  - API calls to set configuration parameters
  - fully-multi-threaded
  - restartable without process reload

# Embedded InnoDB

- Source and binary libraries
- Platforms: Windows, Linux
  - ... others TBD
- Open source / GPL
- “Early Adopter” release NOW available
- Visit

[www.innodb.com/products  
/embedded-innodb](http://www.innodb.com/products/embedded-innodb)

# What else you need to know

- Oracle/Innobase commitment
  - Continued development
  - Hiring/recruiting
  - Coming new relationship!!
- InnoDB Hot Backup
- Where to get info
  - Innodb.com, Forums, Blogs
  - BoF

# InnoDB Forums

forums.innodb.com

**INNODB®**

 [Announcing the InnoDB Plugin 1.0 for MySQL 5.1](#)

**Goto:** [Search](#) · [My Control Center](#) · [Private Messages](#) · [Log Out \(KenJacobs\)](#)

## Forums

### Announcements

News and announcements about InnoDB, the InnoDB Plugin and InnoDB Hot Backup. You may also find here announcements and news about Innobase or the [www.innodb.com](http://www.innodb.com) website.

Options: [Mark Forum Read](#) ▪  [RSS](#)

### InnoDB Storage Engine

Get support for using the standard built-in InnoDB storage engine distributed by MySQL from other users and from InnoDB developers themselves.

Options: [Mark Forum Read](#) ▪  [RSS](#)

### InnoDB Plugin

Discussions about installation and use of the plugin version of InnoDB

Options: [Mark Forum Read](#) ▪  [RSS](#)

### InnoDB Hot Backup

Discuss and get help for the InnoDB Hot Backup utility

Options: [Mark Forum Read](#) ▪  [RSS](#)

### Suggestion Box

This is the place to make suggestions and request new features for any of the InnoDB products: the InnoDB storage engine (built-in or the plugin) and for InnoDB Hot Backup.

Options: [Mark Forum Read](#) ▪  [RSS](#)

**INNOBASE**

# New InnoDB Blog

blogs.innodb.com

## Transactions on InnoDB

“The Word” about InnoDB Products and Technology

Home

### Introducing the InnoDB Blog

Well, here we are ... the first post to the InnoDB blog. Now there is a blog dedicated solely to InnoDB products and technology. The Innobase team will be posting here regularly on all manner of topics regarding the InnoDB storage engine. We plan to provide timely updates and important technical information about InnoDB-related products including the built-in InnoDB distributed by MySQL, the InnoDB Plugin and InnoDB Hot Backup. We invite you to visit regularly and post your comments.

We've borrowed the name “Transactions on” from the computer-science journal *Transactions on Database Systems*, published by the ACM society for computing professionals. Like that journal, this blog will cover a wide range of database topics, specifically as they relate to InnoDB.

Users of InnoDB know a transaction is an atomic all-or-nothing set of changes made to a collection of data. But according to [Webster's Dictionary](#), a transaction also is “a communicative action or activity involving two parties or things that reciprocally affect or influence each other”. So, a transaction is also an exchange of ideas.

Welcome to this place to transact in ideas about InnoDB!

Posted by [Ken Jacobs](#) in [Uncategorized](#) | [Edit](#)

### InnoDB sites

- [Documentation](#)
- [Downloads](#)
- [Forums](#)
- [www.innodb.com](http://www.innodb.com)

### MySQL sites

- [MySQL BugDB](#)
- [MySQL Documentation](#)
- [MySQL InnoDB Forum](#)

April 2009

M	T	W	T	F	S	S
		1	2	3	4	5
..	-	-	-	..	..	..

**INNOBASE**

# Summary:

## Fast. Reliable. Proven. Improved!

- InnoDB is the leading transactional storage engine for MySQL
- InnoDB delivers high performance and protects data integrity for MySQL applications
  - transactions, row-level locking, multi-version concurrency control, referential integrity and crash recovery
- New features in the **InnoDB Plugin** add more performance, reliability, flexibility, ease of use
- InnoDB Hot Backup permits on-line backup of running MySQL databases
- The InnoDB and MySQL relationship is strong, and the future of InnoDB is sound

# Summary:

## Fast. Reliable. Proven. Improved!

- InnoDB is the leading transactional storage engine for MySQL
- InnoDB de... protects data integrity
  - transac... concurrency control... ery
- New feat... more performance... of use
- InnoDB H... p of running M...
- The Innobase and MySQL relationship is strong, and the future of InnoDB is sound

Even more  
Good News!

Oracle / Innobase +  
Sun / MySQL

# For More Information ...

## 2009 MySQL User Conference Sessions

- Ballroom C: InnoDB Birds of a Feather, Wed 7:30pm
- Matthew Yonkovit & Yves Trudeau:  
Distributed InnoDB Caching w/memcached, Tue 2pm
- Heikki Tuuri: Crash Recovery and Media Recovery in InnoDB, Wed, 2pm
- Peter Zaitsev: InnoDB database recovery techniques, Wed 3:05pm
- Mark Callaghan: MySQL Performance on EC2 Wed, 3:05pm
- Ben Handy & Justin Tolmer:  
High Availability & Scalability Patches from Google, Wed 4:25pm
- Calvin Sun & Heikki Tuuri:  
InnoDB File Formats & Source Code Structure Wed, 5:15pm
- Mark Callaghan: The evolution of a MySQL deployment at Google Tues, 9am
- Vadim Tkachenko: InnoDB Performance & Usability Patches , Thu 2pm
- Heikki Tuuri: Concurrency Control: How it Really Works Thurs, 2:50pm

Please visit [www.innodb.com](http://www.innodb.com),  
[blogs.innodb.com](http://blogs.innodb.com) and [forums.innodb.com](http://forums.innodb.com)



Q & A  
QUESTIONS  
ANSWERS

**INNOBASE**

# INNOBASE

an **ORACLE**® company

## INNODB

### INNODB Plugin

### INNODB Hot Backup

### Embedded INNODB

