



PBXT: Technology Trends That Affect Your Database

OpenSQL Camp 2009

Vladimir Kolesnikov

PrimeBase Technologies GmbH

www.primebase.org



Contents

- PBXT Overview
- The Trends
- Engine Design & Architecture
- In-Memory Tables
- PBXT Performance Monitoring
- Some Benchmarks
- Release Schedules



PBXT Engine Features

- A pluggable engine for MySQL 5.1.x, MySQL 5.4.x, Drizzle, MariaDB
- Open Source (GPL v.2)
- Transactional, ACID-compliant
- Row-level locks
- Referential integrity
- Fast rollback and recovery
- Available on Linux (32/64 bit), [Open] Solaris (x86 and SPARC), BSD, OSX, Win32

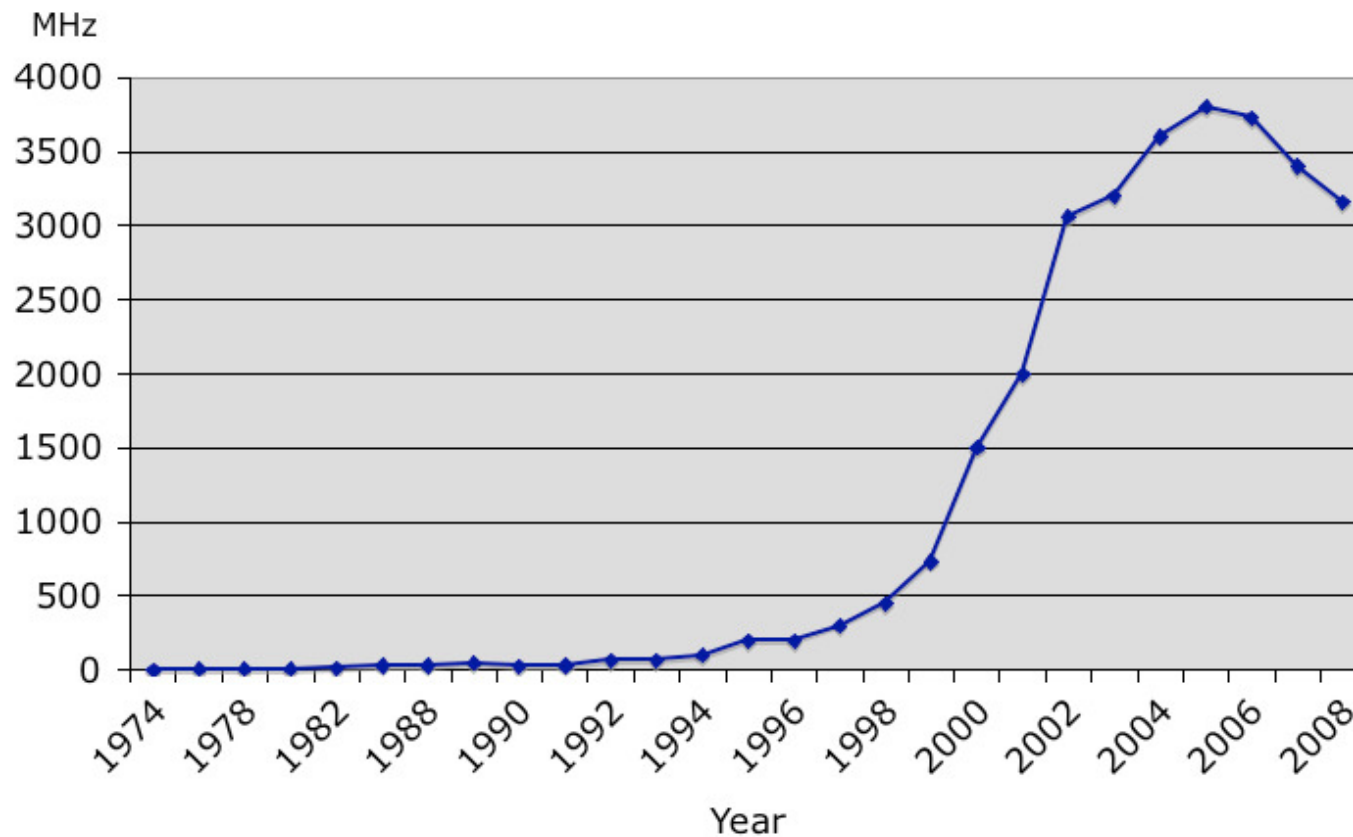


Technology Trends

- **Multi-core, multi-threaded CPUs**
minimize locked operation; split data
- **64-bit, loads of RAM**
cache smart; buffer data, process on background
- **Solid State Drives**
no random access penalty; no clustering needed;
write sequentially



Maximum CPU frequency, 1974-2008



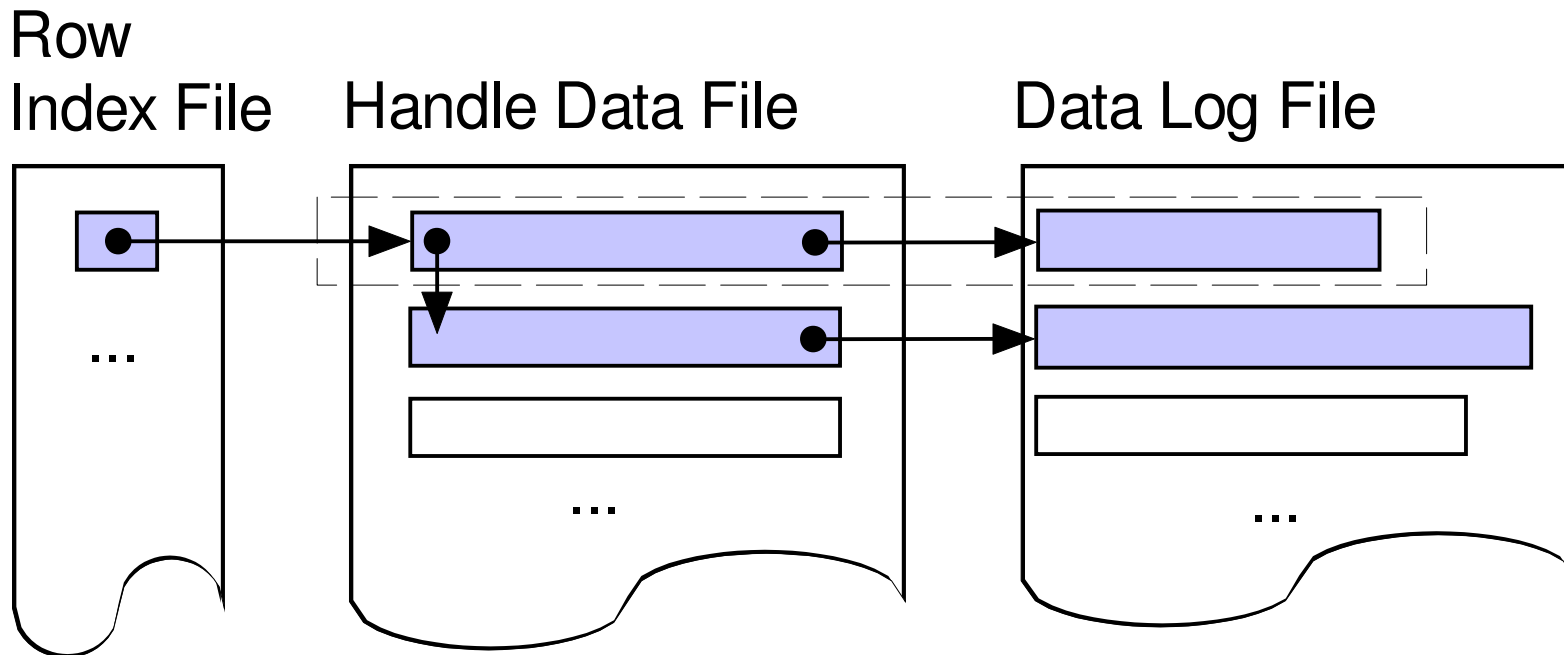


PBXT Engine Design

- Write-once, log-based
- Disk-based MVCC
- Fast commits
- No in-place updates
- No undo
- File-per-table



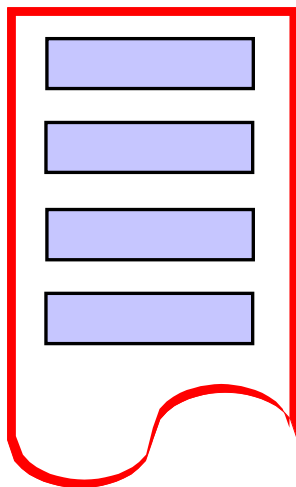
PBXT Table Data Layout





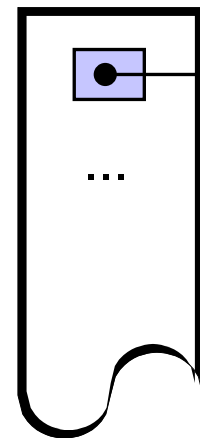
INSERT, UPDATE, DELETE

Transaction
Log File

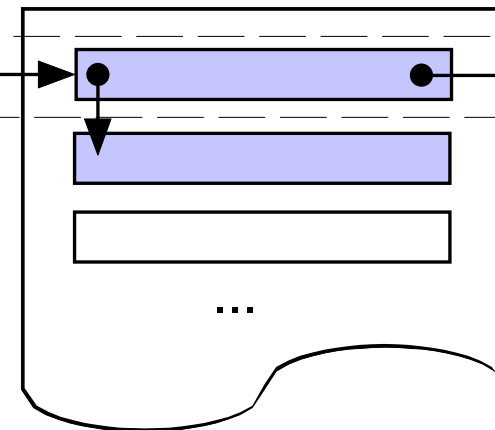


Row

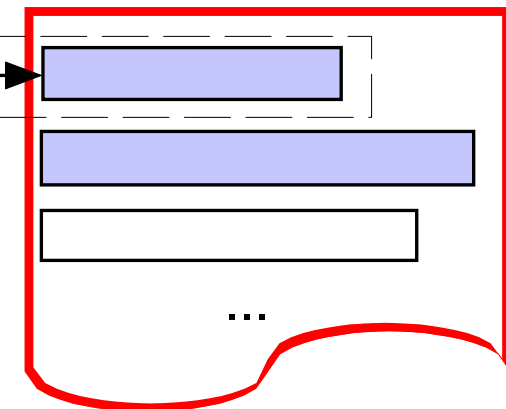
Index File



Handle Data File



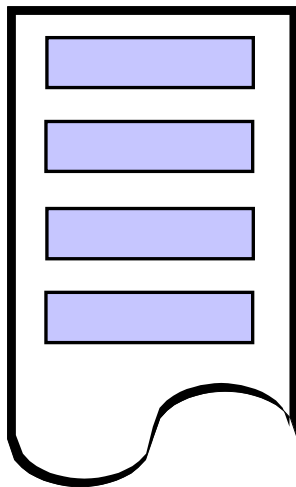
Data Log File





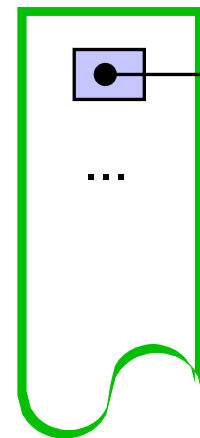
SELECT

Transaction
Log File

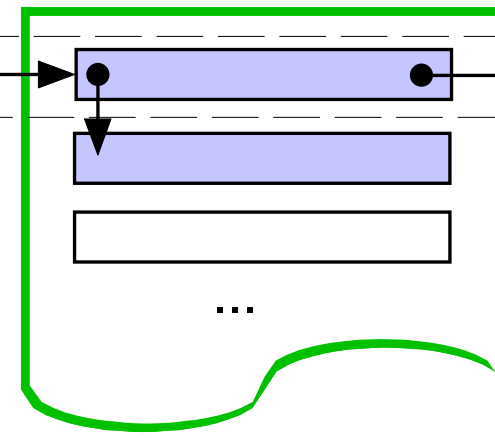


Row

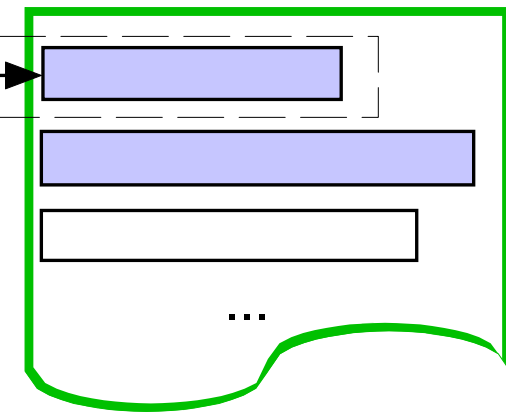
Index File



Handle Data File



Data Log File



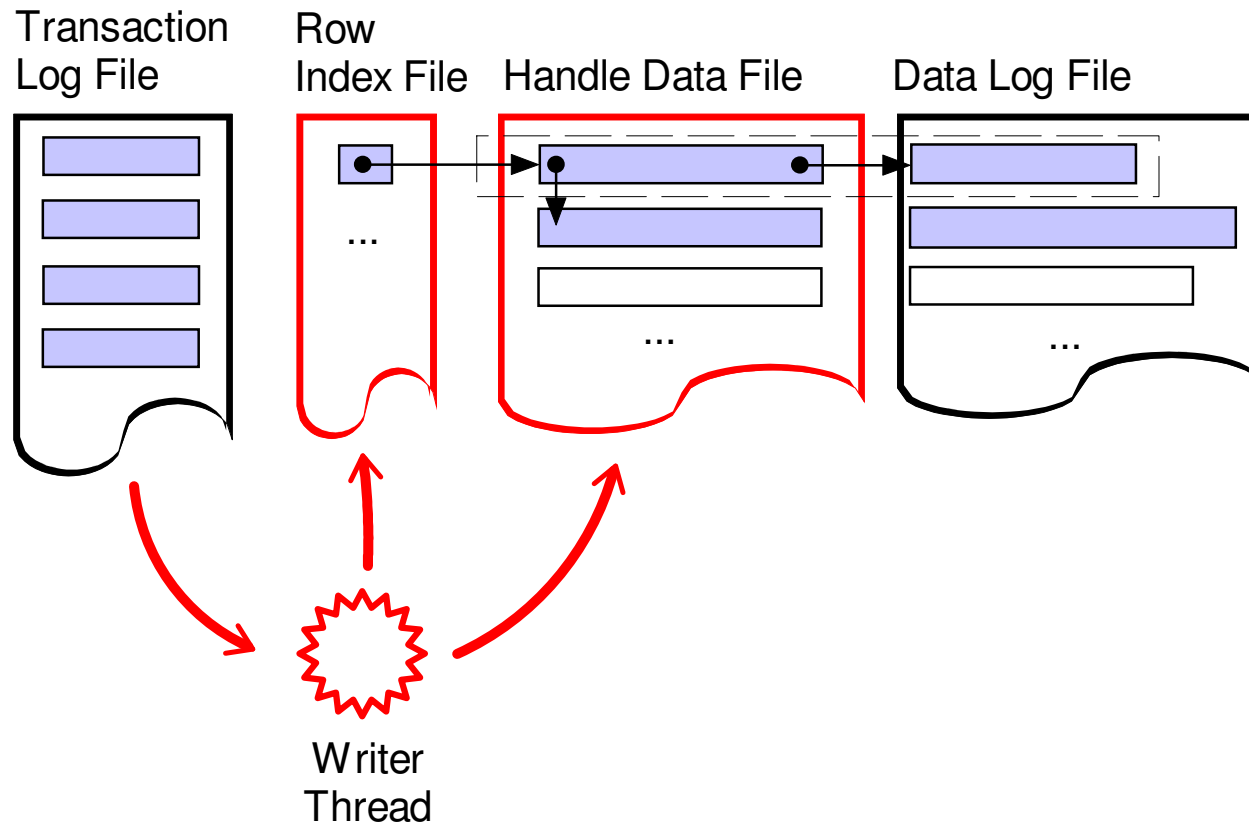


PBXT Background Threads

- Writer
- Sweeper
- Checkpointer
- Compactor

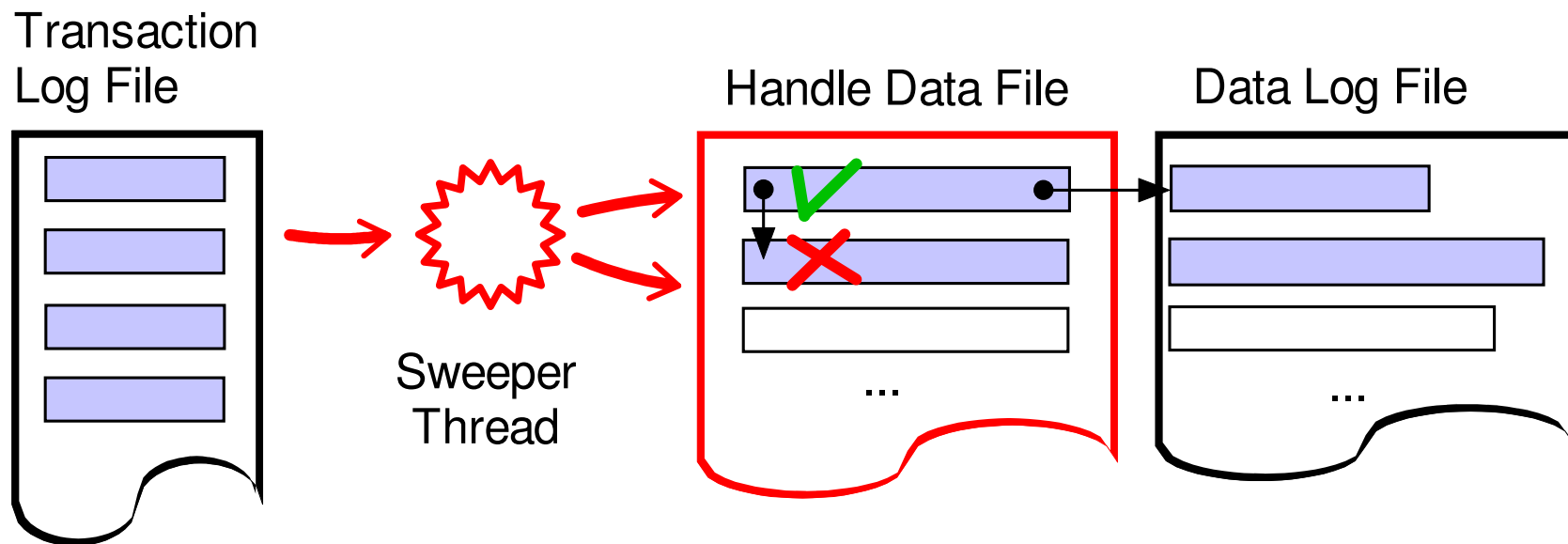


The Writer Thread



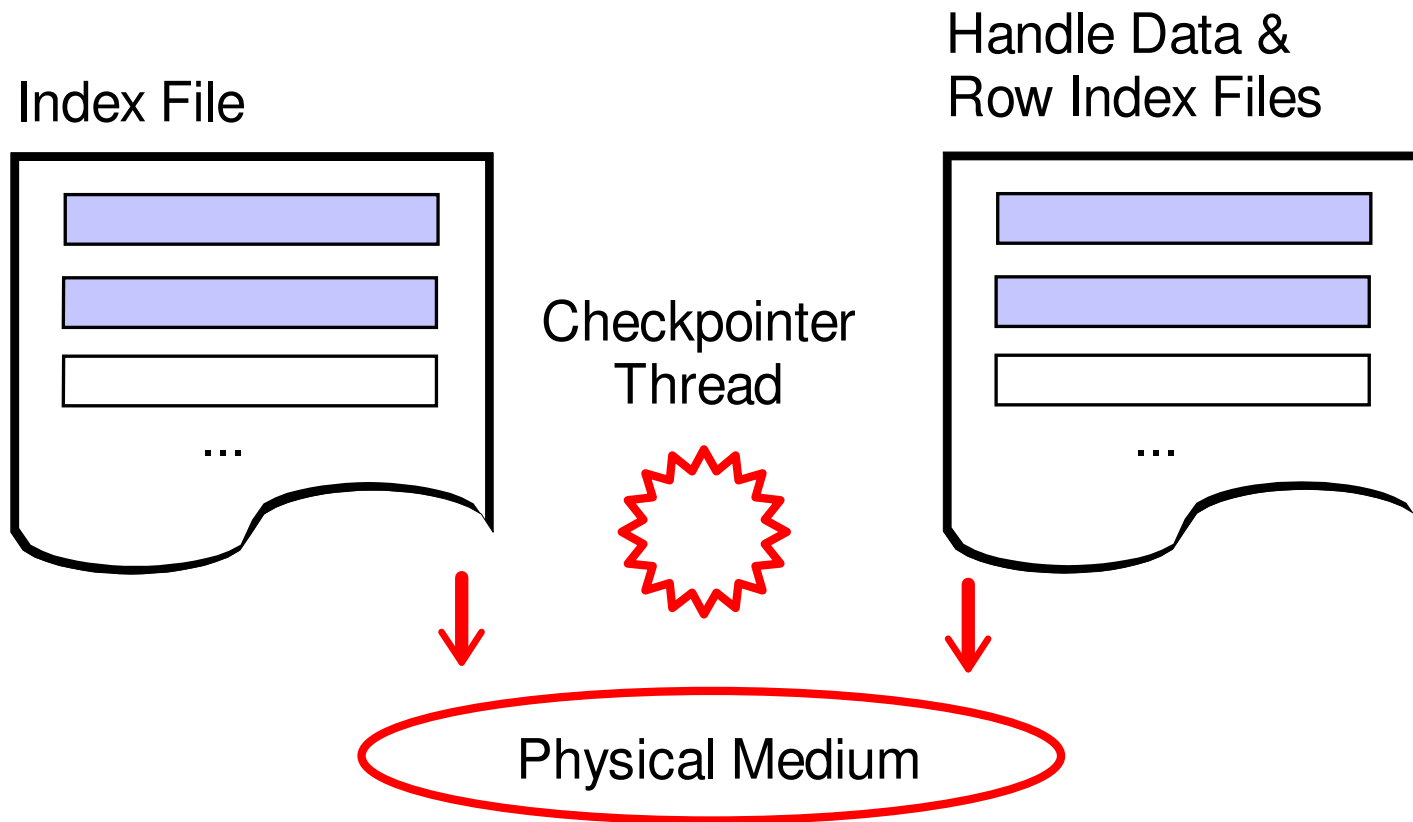


The Sweeper Thread



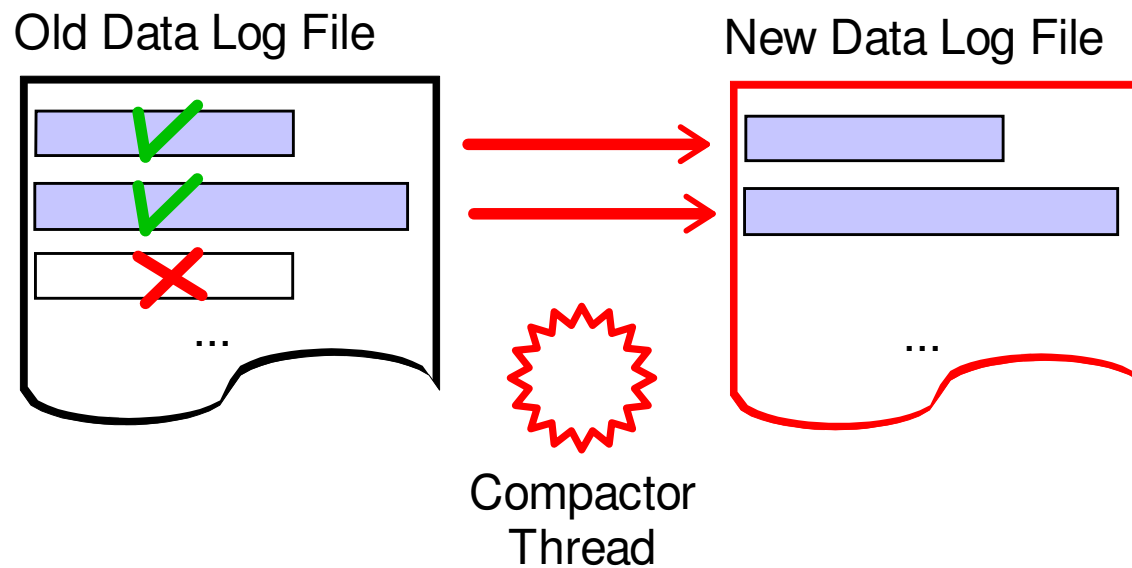


The Checkpointer Thread





The Compactor Thread





In-Memory Tables

- Fast in-memory storage
- Scalable (doesn't use table locking)
- Atomic, Consistent, Isolated
- Durability can be easily added (checkpoints or xlog)



In-Memory Tables

- Can be combined with disk-based PBXT tables
- Supports BLOB data
- `CREATE TABLE t1 (...) ENGINE=PBXT STORAGE MEMORY`



PBXT Performance Monitoring

- I_S.PBXT_STATISTICS
- XTSTAT
- SHOW ENGINE PBXT STATUS



PBXT_STATISTICS Table

- Transaction Statistics
- SQL Statistics
- Table and Index I/O Counters
- Table and Index Cache Status & Counters
- D-Log and X-Log Counters
- Total 48 counters



The XTSTAT Tool

- IOSTAT-like instant monitoring interface
- Local/remote monitoring
- Output filtering
- Suitable for scripting and logging



Sysbench r/w OLPT, 10m Rows, SSD

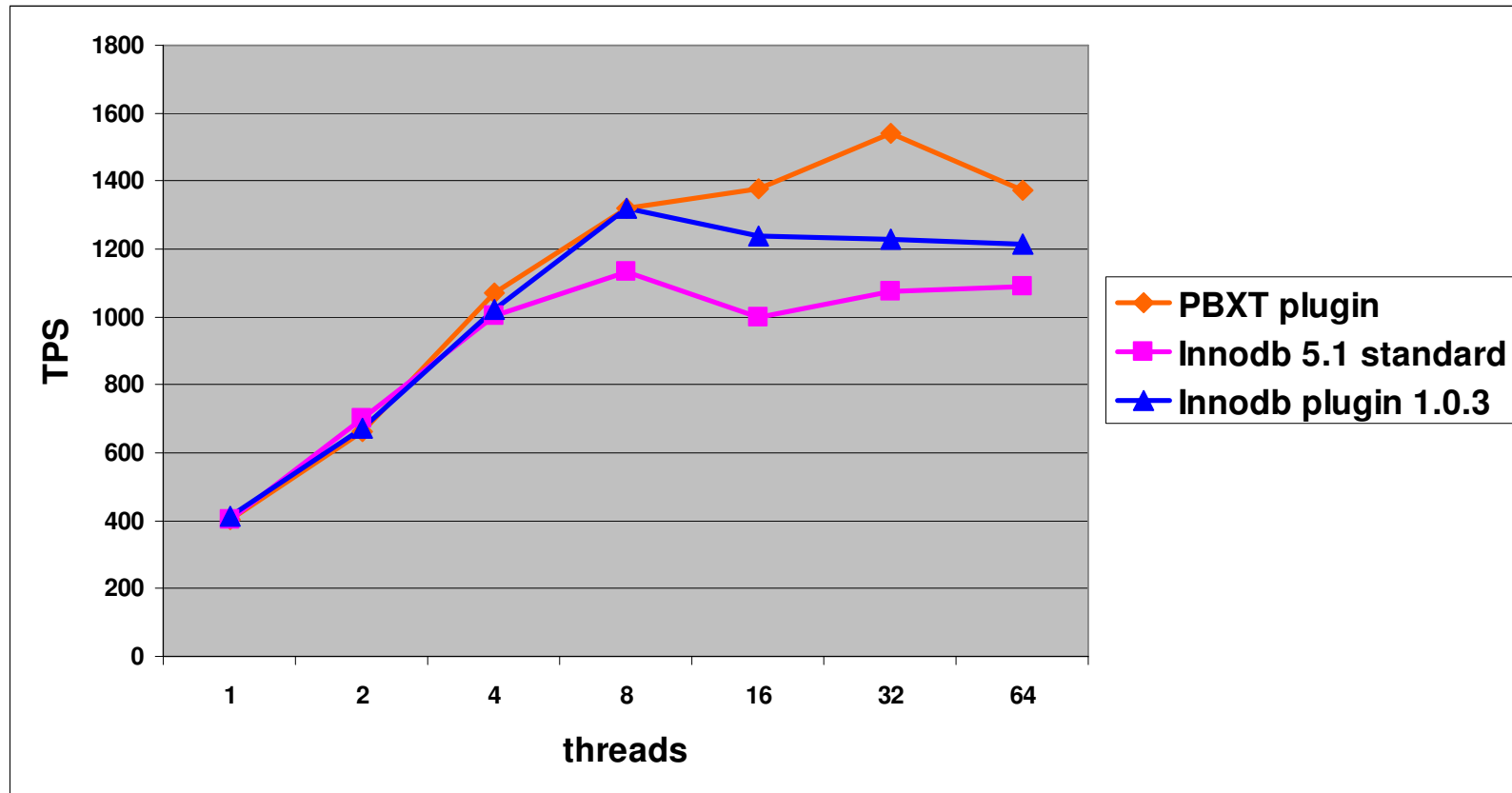




Table vs Index Scan (disk-bound)

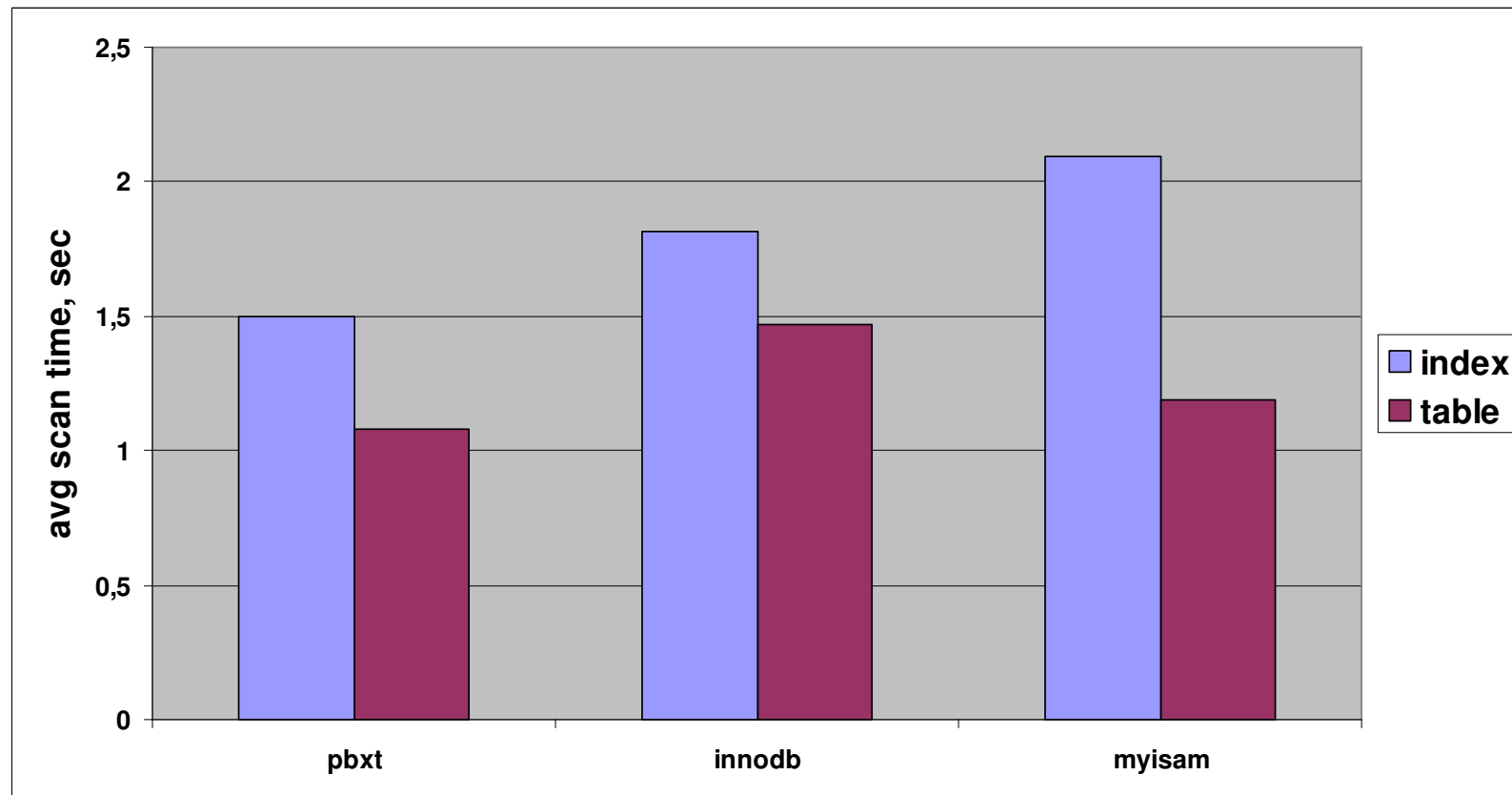
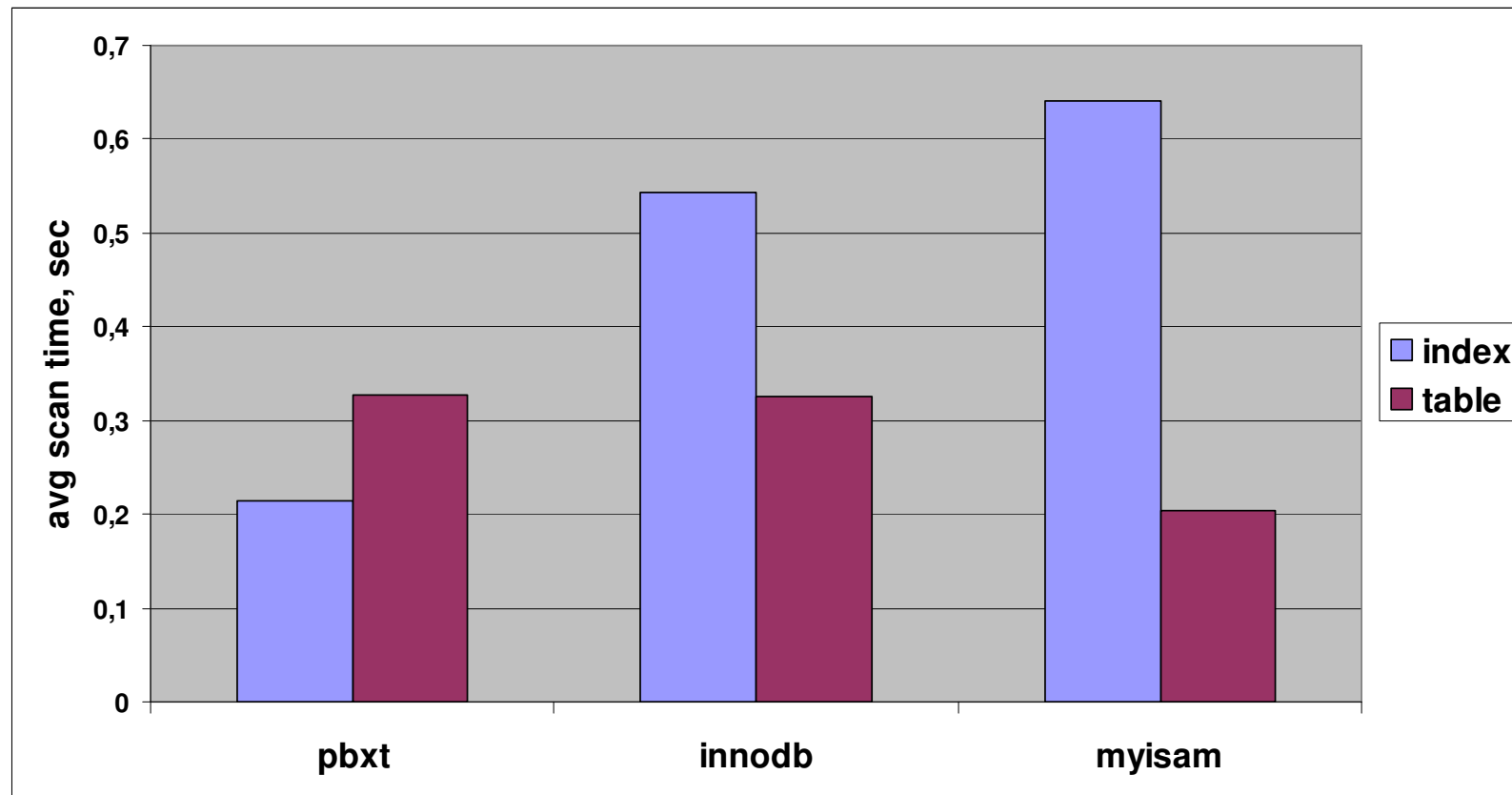




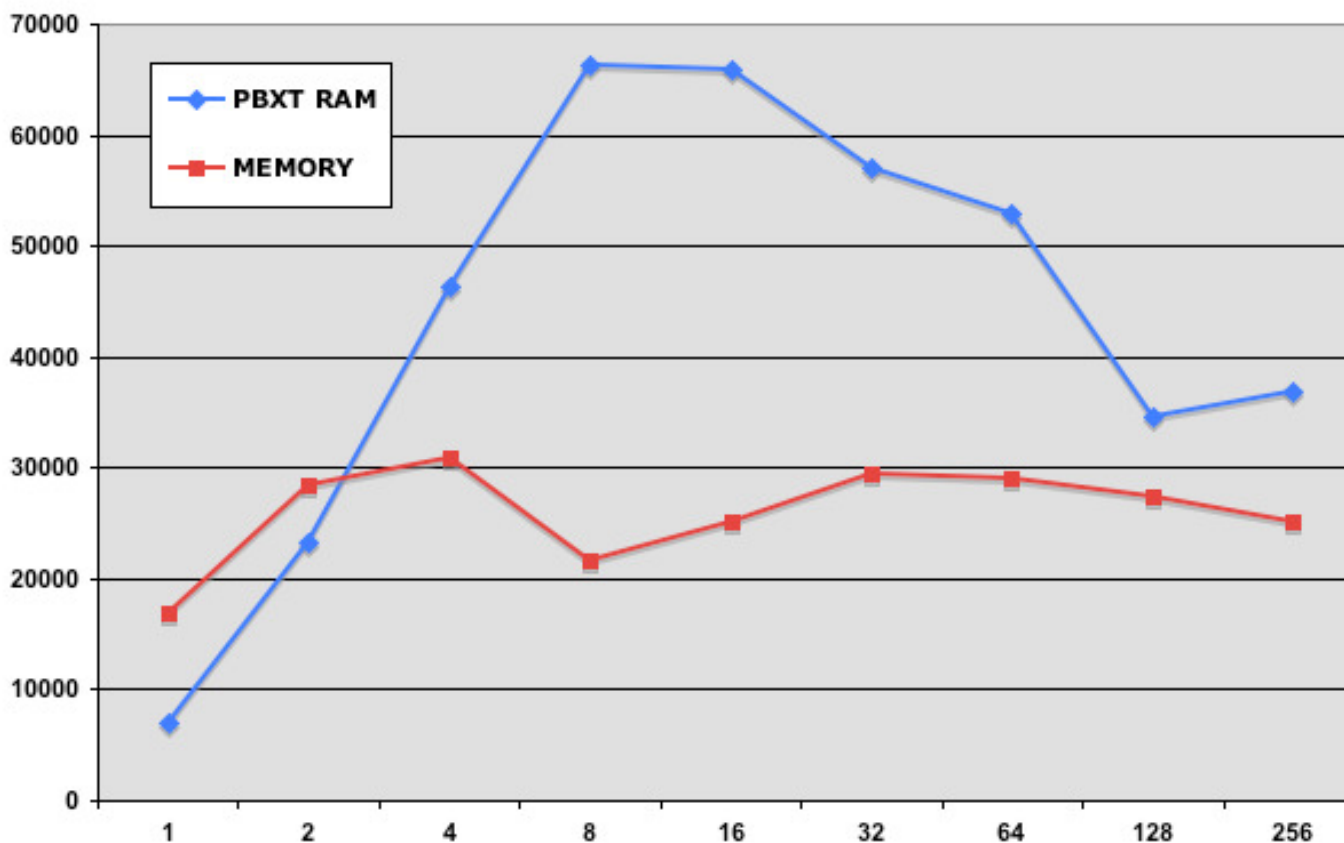
Table vs Index Scan (cache-bound)





In-Memory Tables: Mixed Workload

Mixed: INSERT/UPDATE/SELECTs per second





Releases Schedule

- September-October 2009 – RC3
- End of 2009 - GA



PBXT on Launchpad

- <https://code.launchpad.net/pbxt>
- <lp:pbxt> – the latest bugfix version
- <lp:pbxt/stable> – the latest stable version
- <lp:pbxt/rc2> – current RC
- <lp:~pbxt-core/pbxt/staging> – memory tables, all new features
- <https://bugs.launchpad.net/pbxt>
- <lp:~drizzle-pbxt/drizzle/pbxt>



Thank you!

Q&A

<http://www.primebase.org>

<https://launchpad.net/pbxt>

<http://pbxt.blogspot.com>