MySQL performance in a cloud

Mark Callaghan

Special thanks

Eric Hammond (http://www.anvilon.com) provided documentation that made all of my work much easier.

What is this thing called a cloud?

- Deployment trends
- Technology
- Public versus private

Deploying MySQL in a cloud

- New problems
- New benefits
- Differences from traditional deployment
- Performance can be good, but ...
  - Virtualization techniques matter
  - May need InnoDB patches to tolerate IO latency

Impact from requirements

Database in direct attached storage:

- backups and binlogs archived in the cloud
- use MySQL replication to maintain a failover target
- less can go wrong

Database in network attached storage

- another MySQL server can takeover on failure

Focus on InnoDB performance

- Network attached storage
- Direct attached storage
- Multi-core servers
- Virtualization overhead
- Patches that improve performance

Benchmarks

- Start with simple benchmarks
- iibench
  - IO bound workload
  - great for finding bottlenecks in storage engines
  - started by Tokutek
- sysbench
  - OLTP workload
- wisconsin
  - query processing workload

What is different?

- Not much, MySQL runs great here
- Multi-core scalability matters because 8-cores costs more
- May need ability to tolerate IO latency

Make InnoDB faster

- link with tcmalloc
- use XFS
- reduce mutex contention for multi-core servers
- IO performance
    - multiple background IO threads
    - increase IO rate on busy servers

Factors for IO latency

- O_DIRECT versus buffered IO
- SATA writeback cache
- Flash erase cycles
- Network versus direct attached storage
- IO scheduler
- Excessive prefetching from the OS
- Hardware RAID write cache
- File system limits on concurrent reads/writes per file
- Ability of storage engine to issue concurrent IO requests

Tuning for IO bound loads

innodb_read_io_threads

- In Percona and Google patches
- Helps when there is a lot of prefetching for full table scans

innodb_write_io_threads

- In Percona and Google patches
- Helps when writes have a lot of latency
- Writes have a lot of latency when:
    - using O_DIRECT without SATA writeback cache
    - using O_DIRECT without HW RAID write cache
    - using network attached storage

Tuning for IO bound loads (2)

innodb_io_capacity
- In Google and Percona patches
- Helps when there are many writes to issue

faster IO
- Increases rate at which background IO is done
- Increase size of IO request arrays
- Google and Percona patches have changes for this

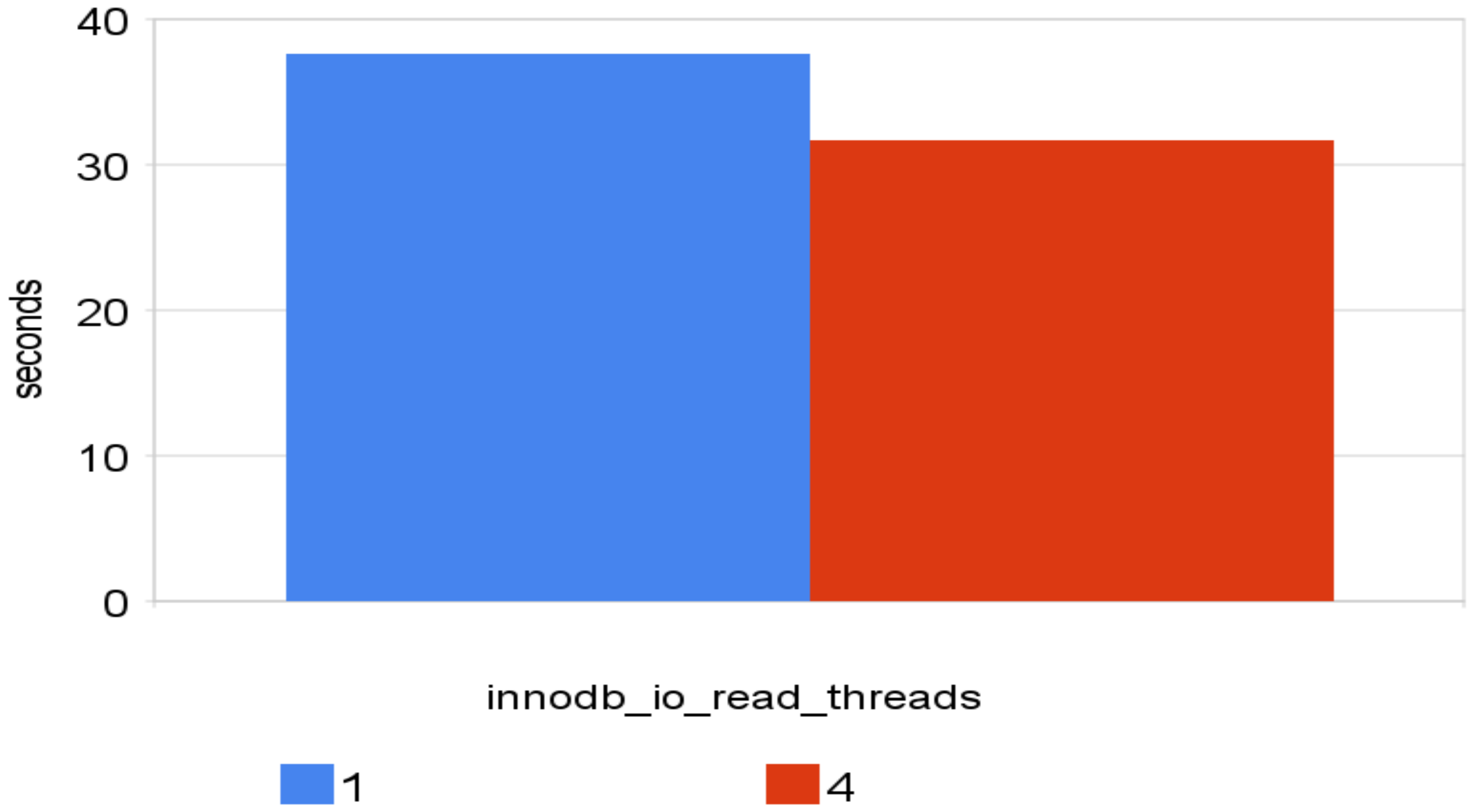
SHOW INNODB STATUS
- Google and Percona added more output
- Google patch includes average IO time for reads and writes

Network attached storage tests

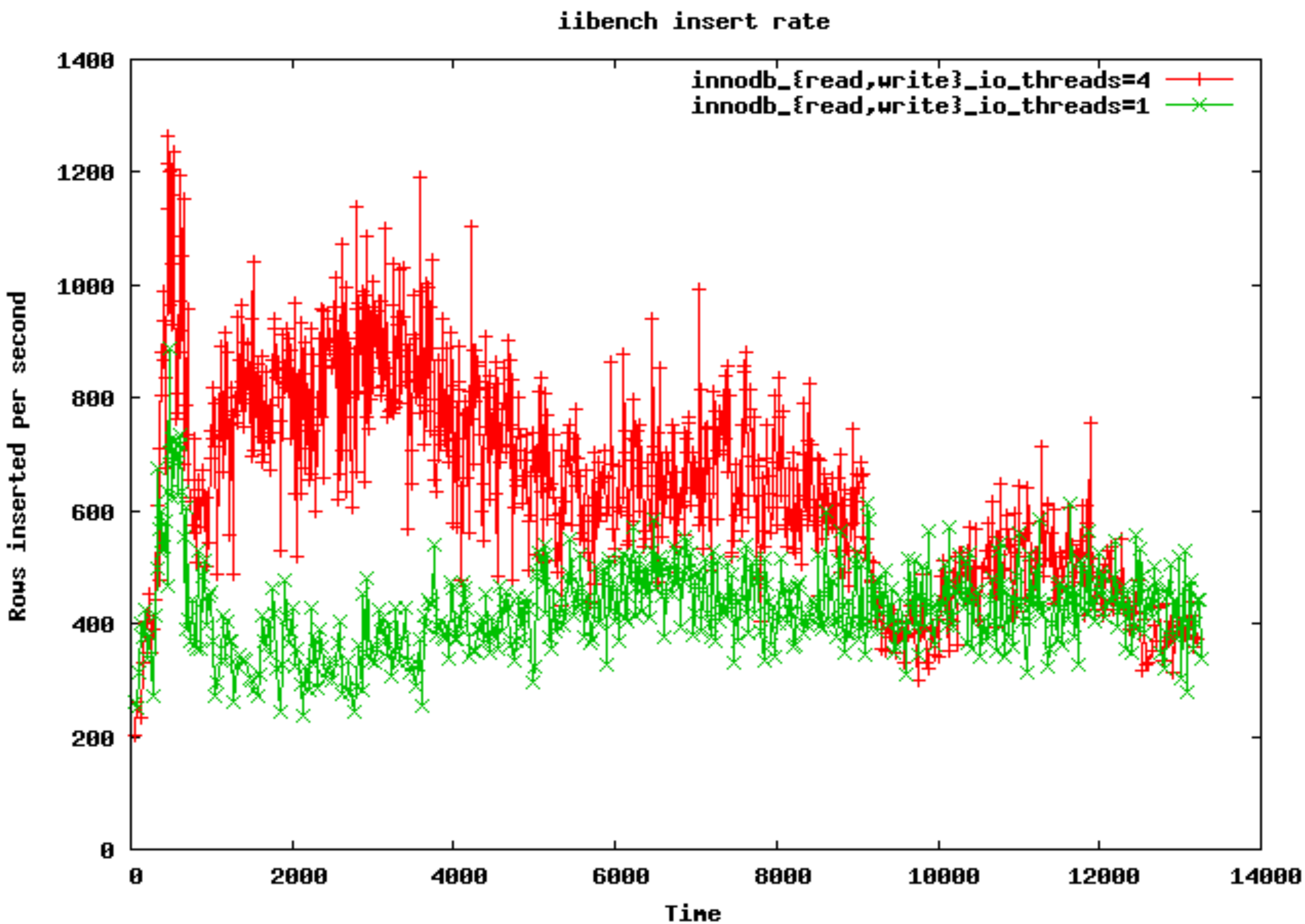
Server:
- 2 CPU cores, 4G or 8G RAM
- SW RAID 0 striped over 4 network volumes
- 1M RAID stripe size
- XFS
- MySQL 5.0.37 + v3 Google patch + tcmalloc
- Innodb with 1G buffer pool, O_DIRECT, innodb_flush_log_at_trx_commit=2

Concurrent query performance with network attached storage:

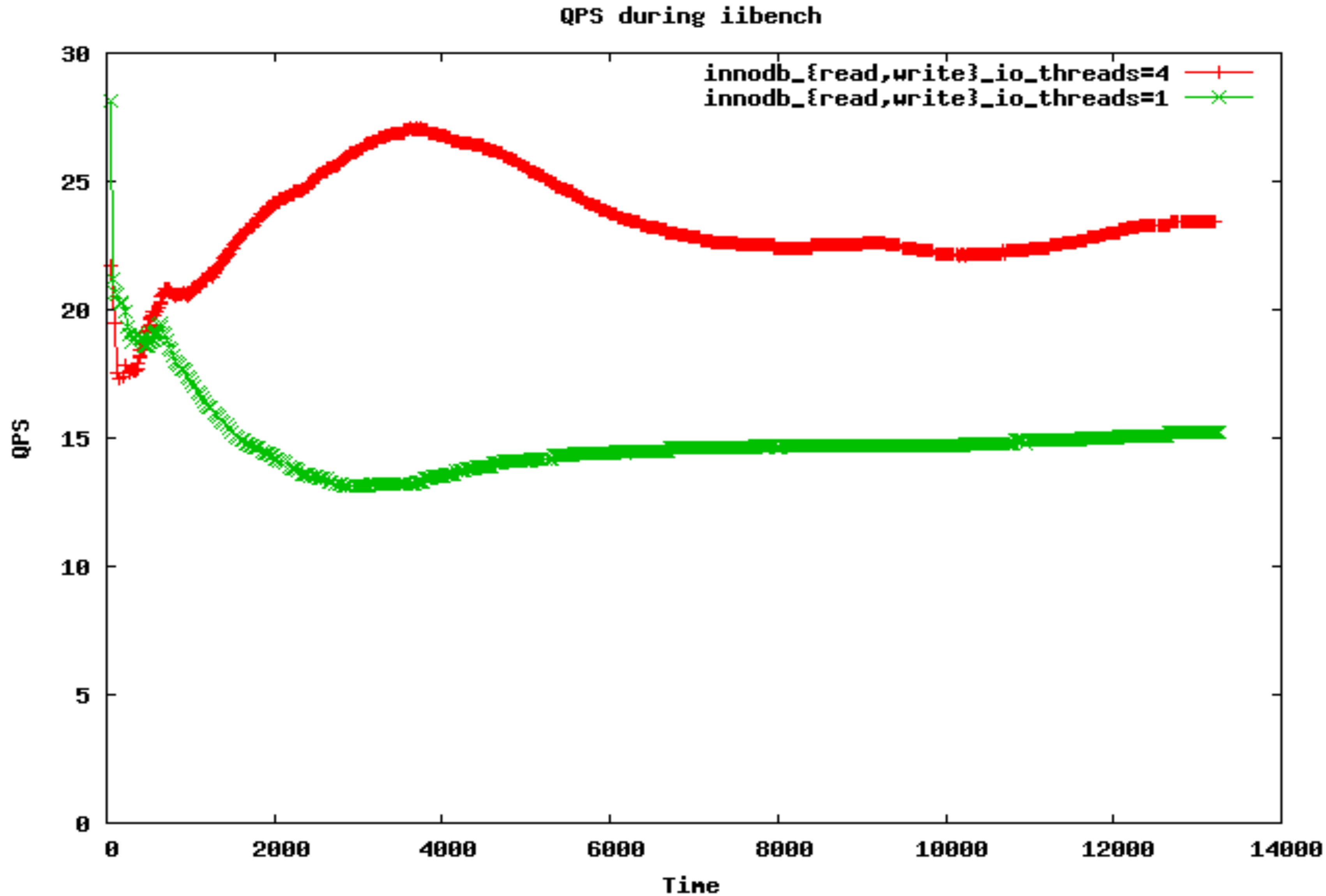
- 4 concurrent queries, IO bound



iibench insert rate



iibench QPS rate from 4 threads concurrent with inserts



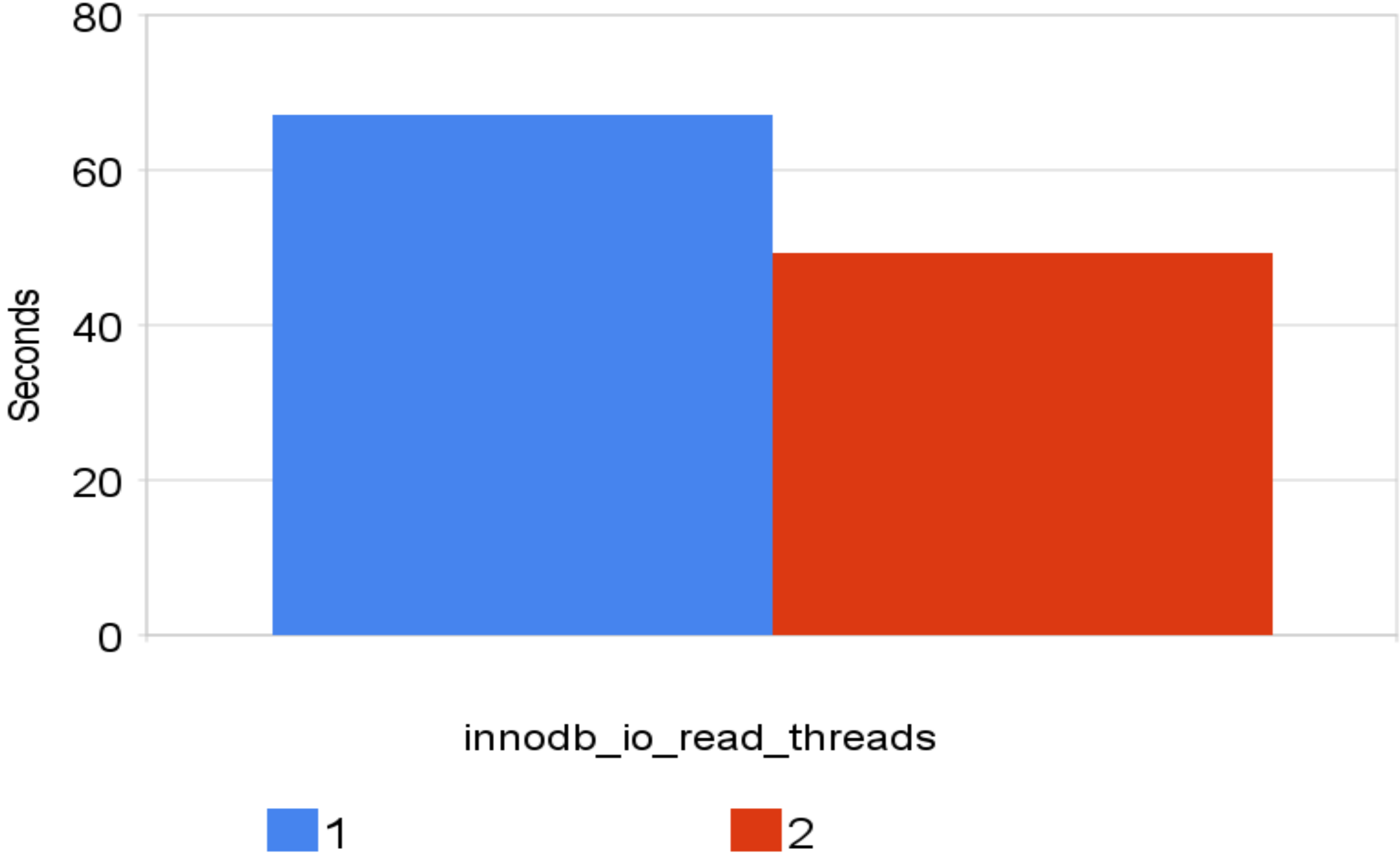
Direct attached storage tests

Server:

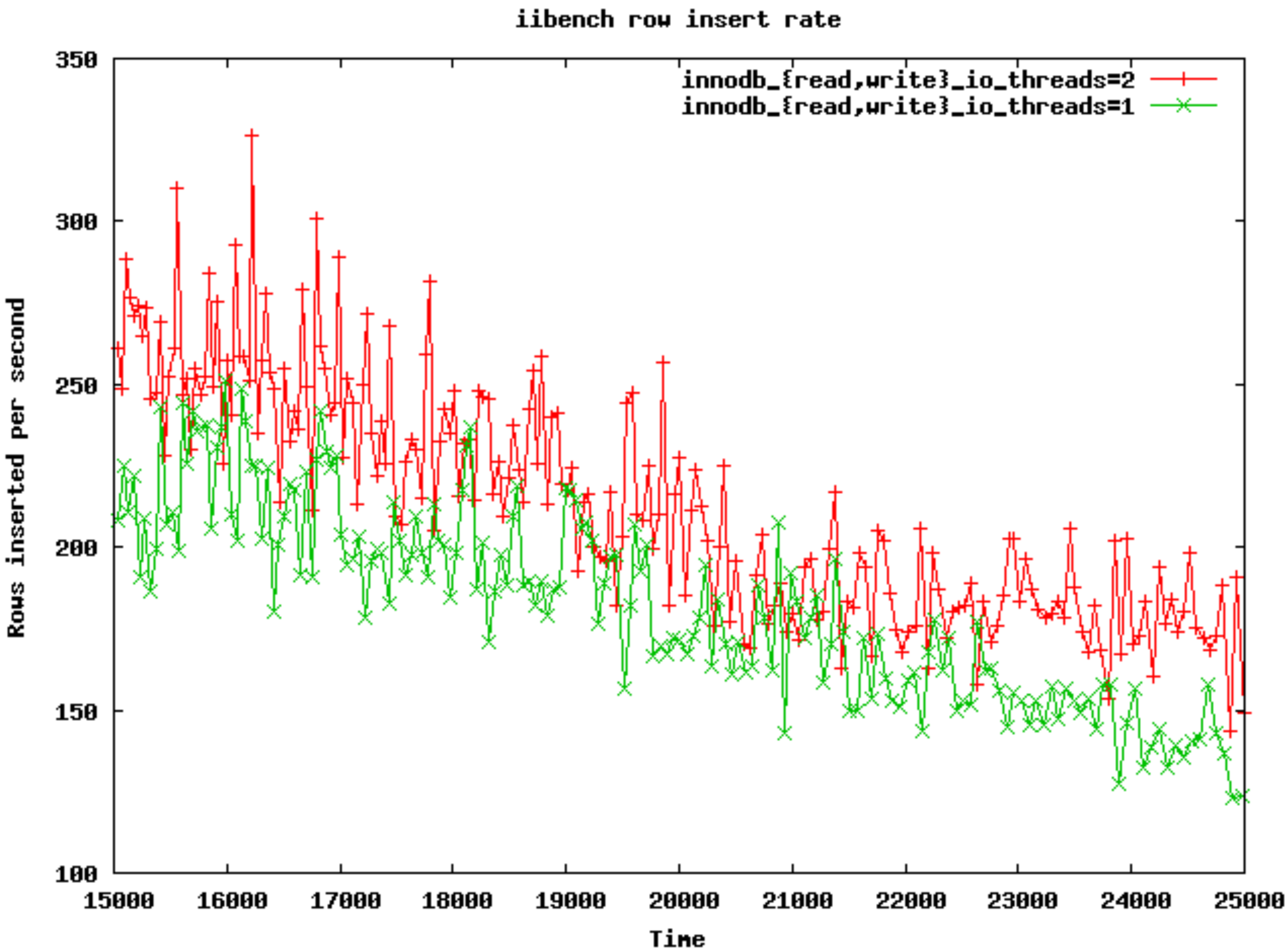
- 2 CPU cores, 4G or 8G RAM
- SW RAID 0 striped over 2 disks
- 1M RAID stripe size
- XFS
- Innodb with 1G buffer pool, O_DIRECT, innodb_flush_log_at_trx_commit=2
- MySQL 5.0.37 + v3 Google patch + tcmalloc

Concurrent query performance with direct attached storage:

- 2 concurrent queries, IO bound



iibench insert rate

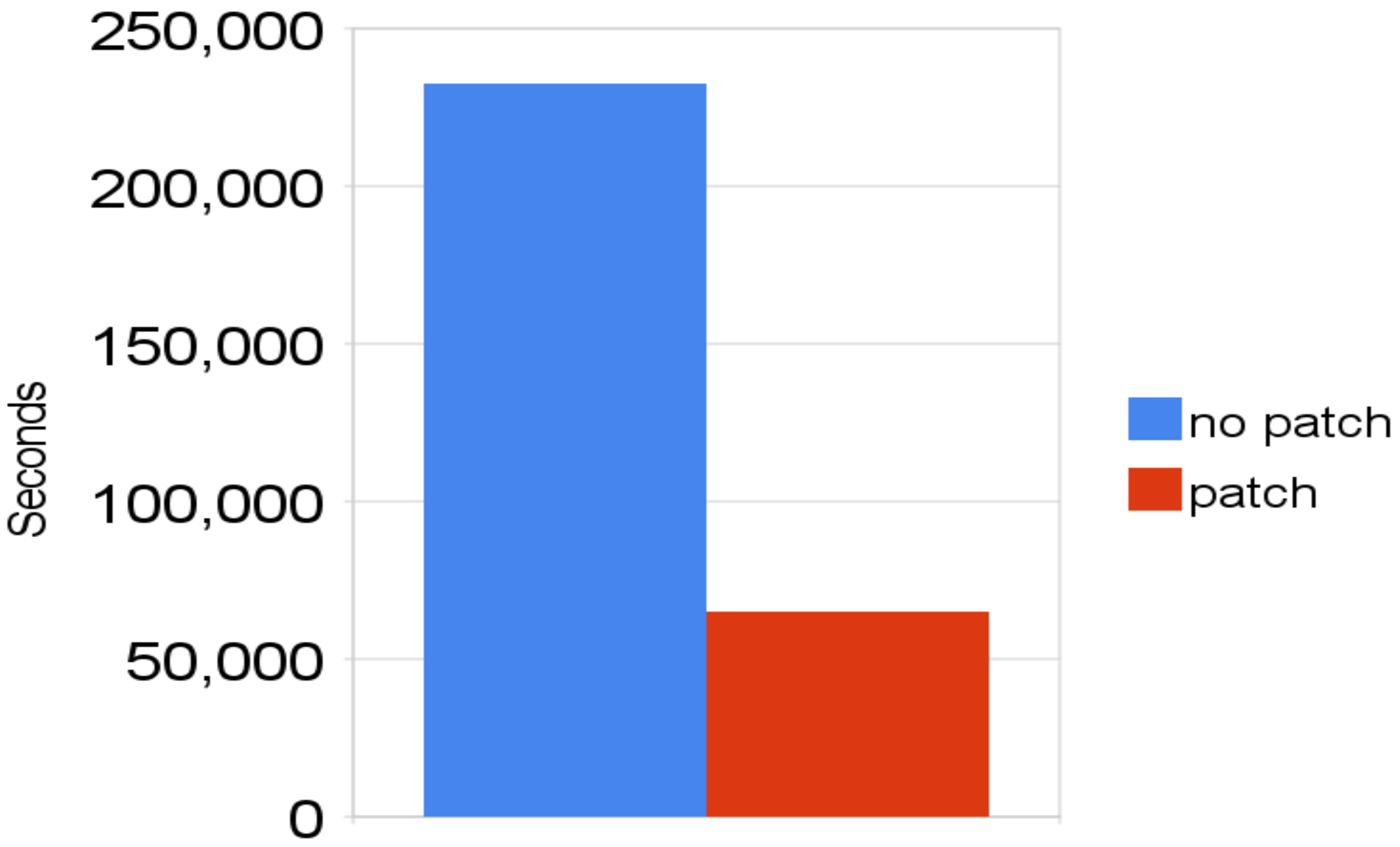


Direct attached storage tests (2)

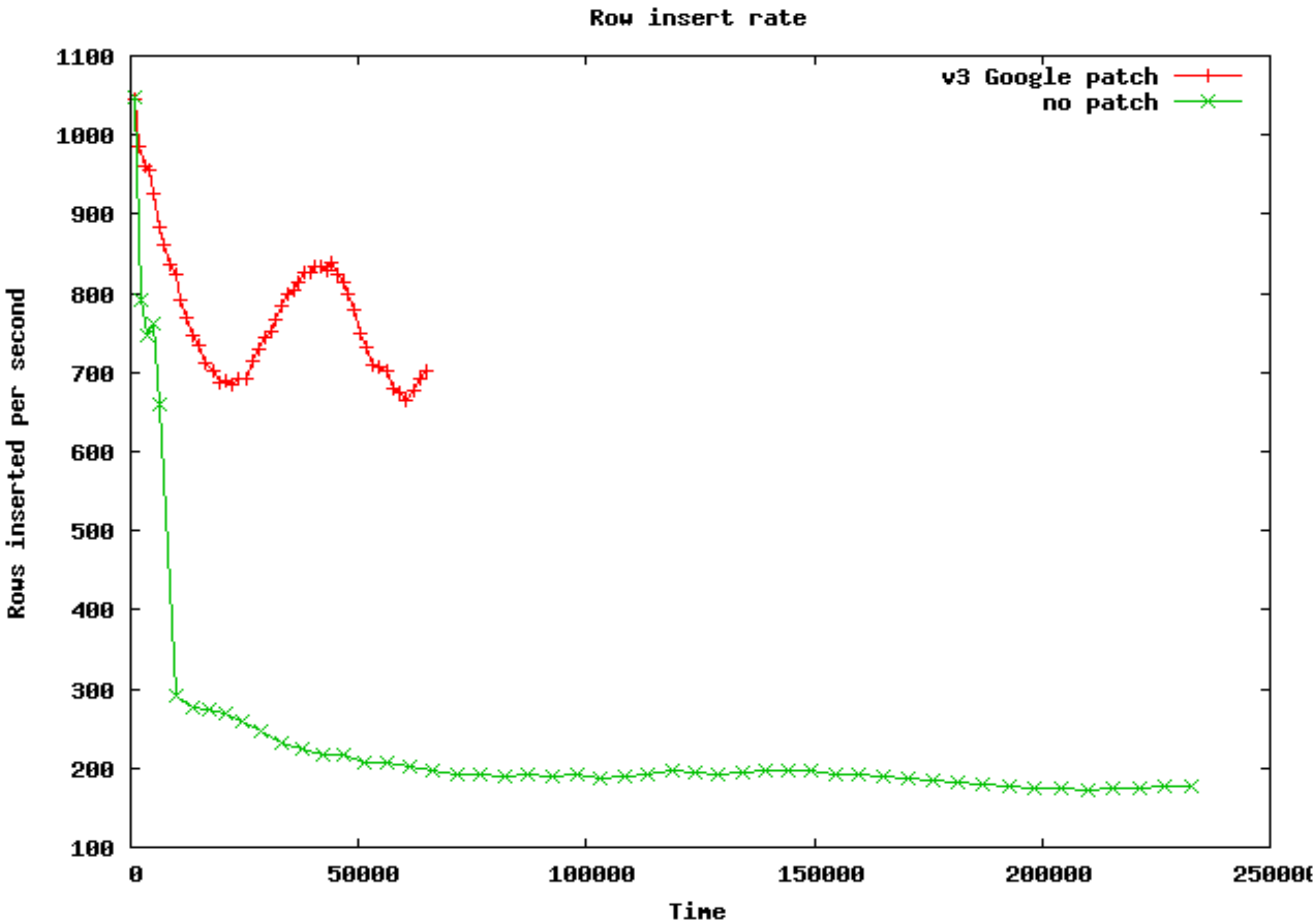
Server:

- 8 CPU cores, 4G or 8G RAM
- SW RAID 0 striped over 10 disks
- 1M RAID stripe size
- ext-2
- Innodb with 1G buffer pool, O_DIRECT, innodb_flush_log_at_trx_commit=2
- MySQL 5.0.37 + v3 Google patch + tcmalloc

Time to load 50M rows in iibench



Row insert rate while loading 50M rows in iibench

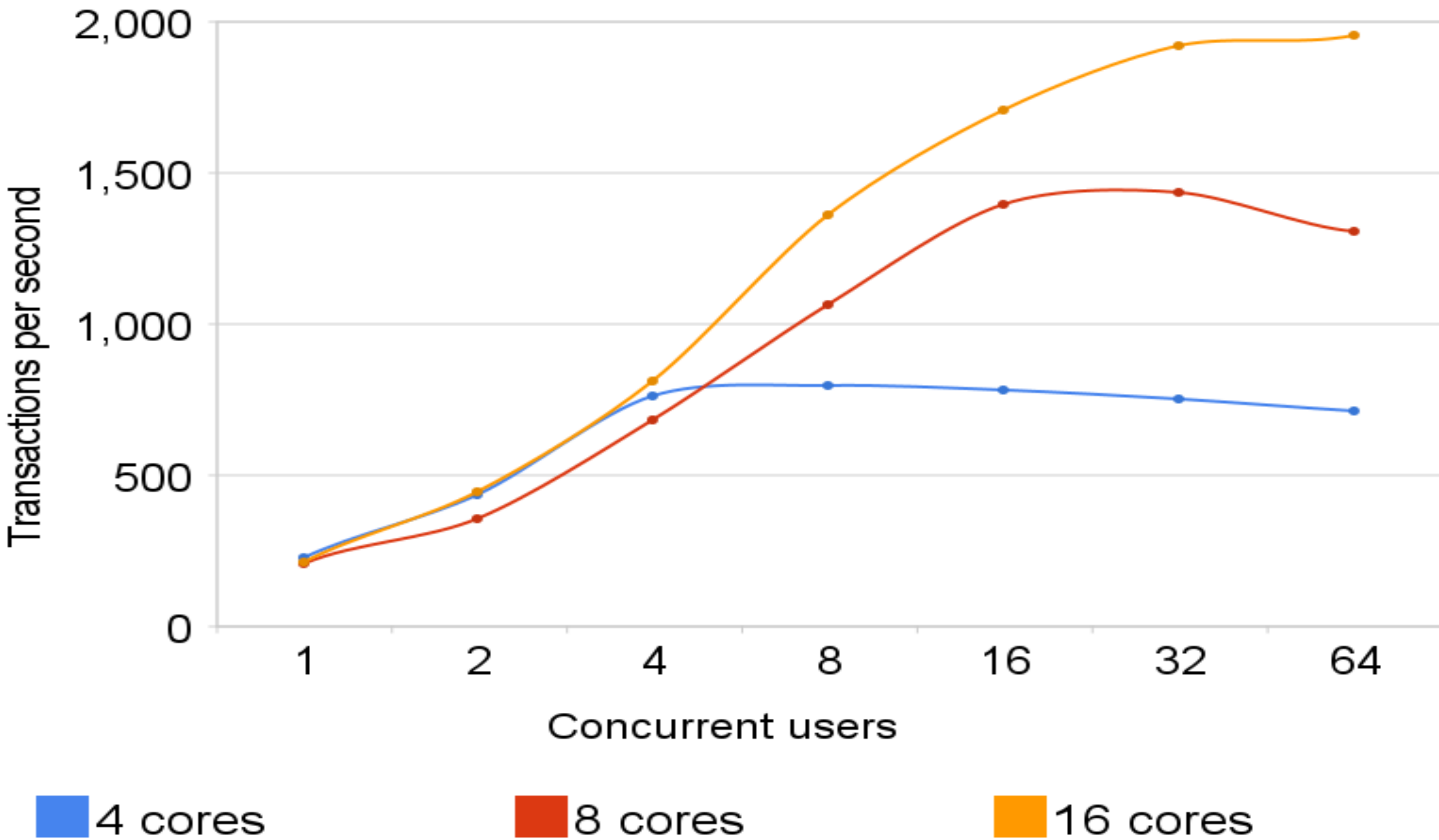


Multi-core servers

- How do MySQL and InnoDB scale on SMP?
- Test configuration:
  - CPU bound workload
  - MySQL 5.0.37 with v3 Google patch
  - 4, 8 and 16 core servers
  - mysqld linked with tcmalloc

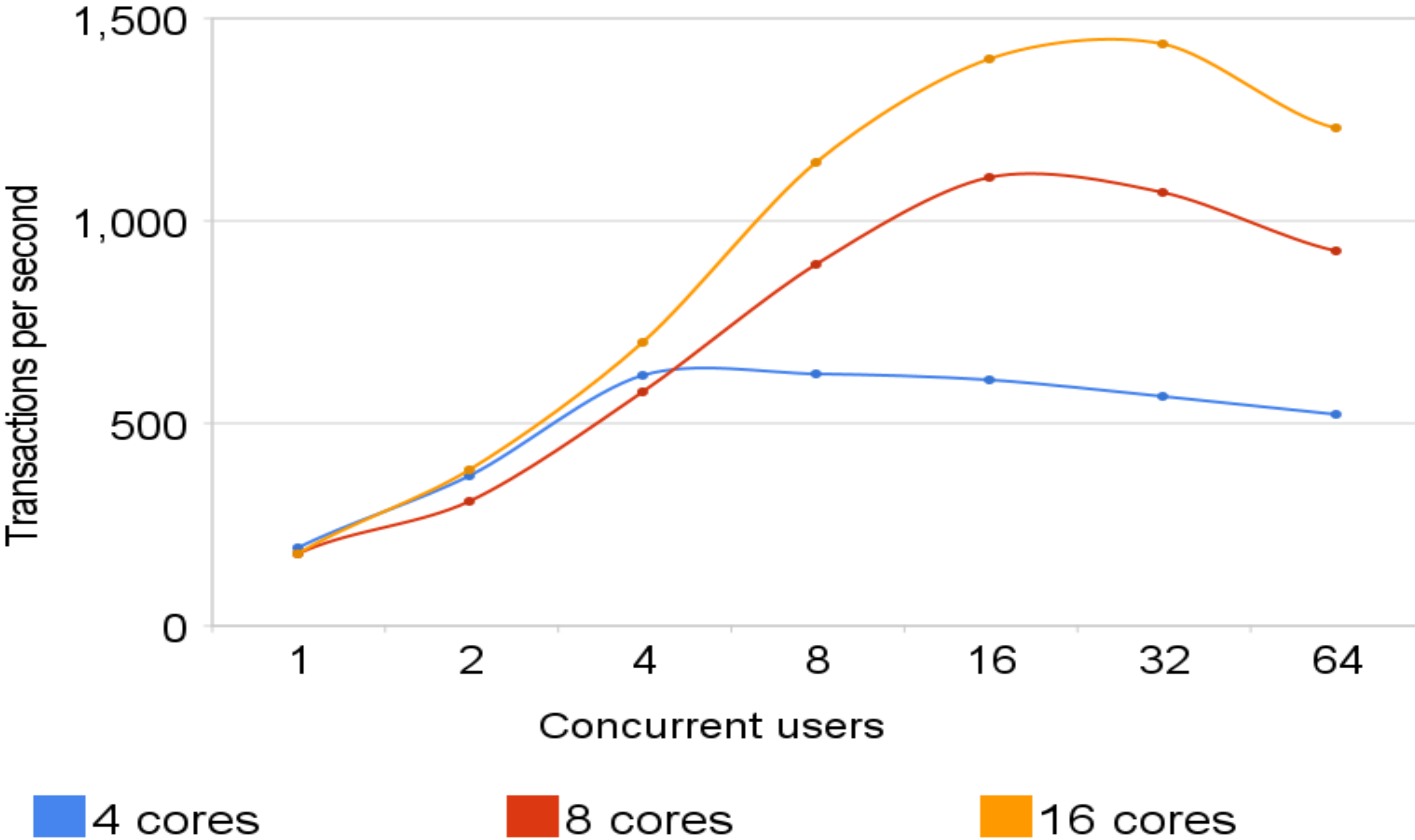
CPU speedup without virtualization:

- modified sysbench readonly, CPU bound
- measure transactions per second



CPU speedup without virtualization:

- modified sysbench readwrite, CPU bound
- measure transactions per second



Virtualization overhead

KVM tests

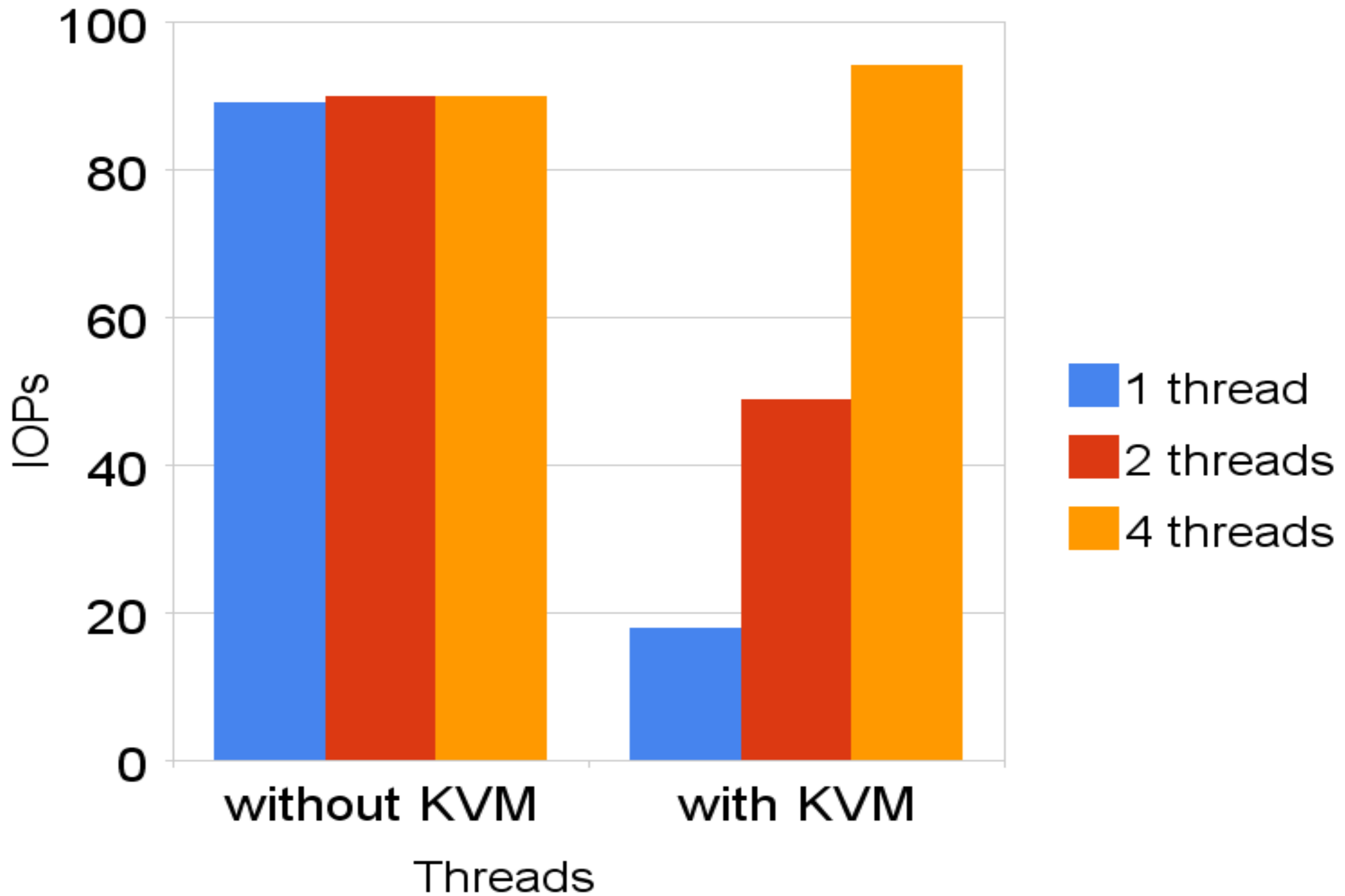
- Ubuntu 8.04
- 4 core server, 1 disk, 4G RAM, supports AMD-V
- MySQL 5.0.77 with tcmalloc
- MySQL 5.0.37 with v3 Google patch and tcmalloc
- Note that KVM is much improved since this version

Xen tests

- Linux 2.6
- 8 CPU cores, enough RAM to cache database
- hardware on server with Xen faster than non-Xen server
- Xen server has 4 disks in SW-RAID 0 using XFS, 16G RAM
- MySQL 5.0.37 with tcmalloc and v3 Google patch

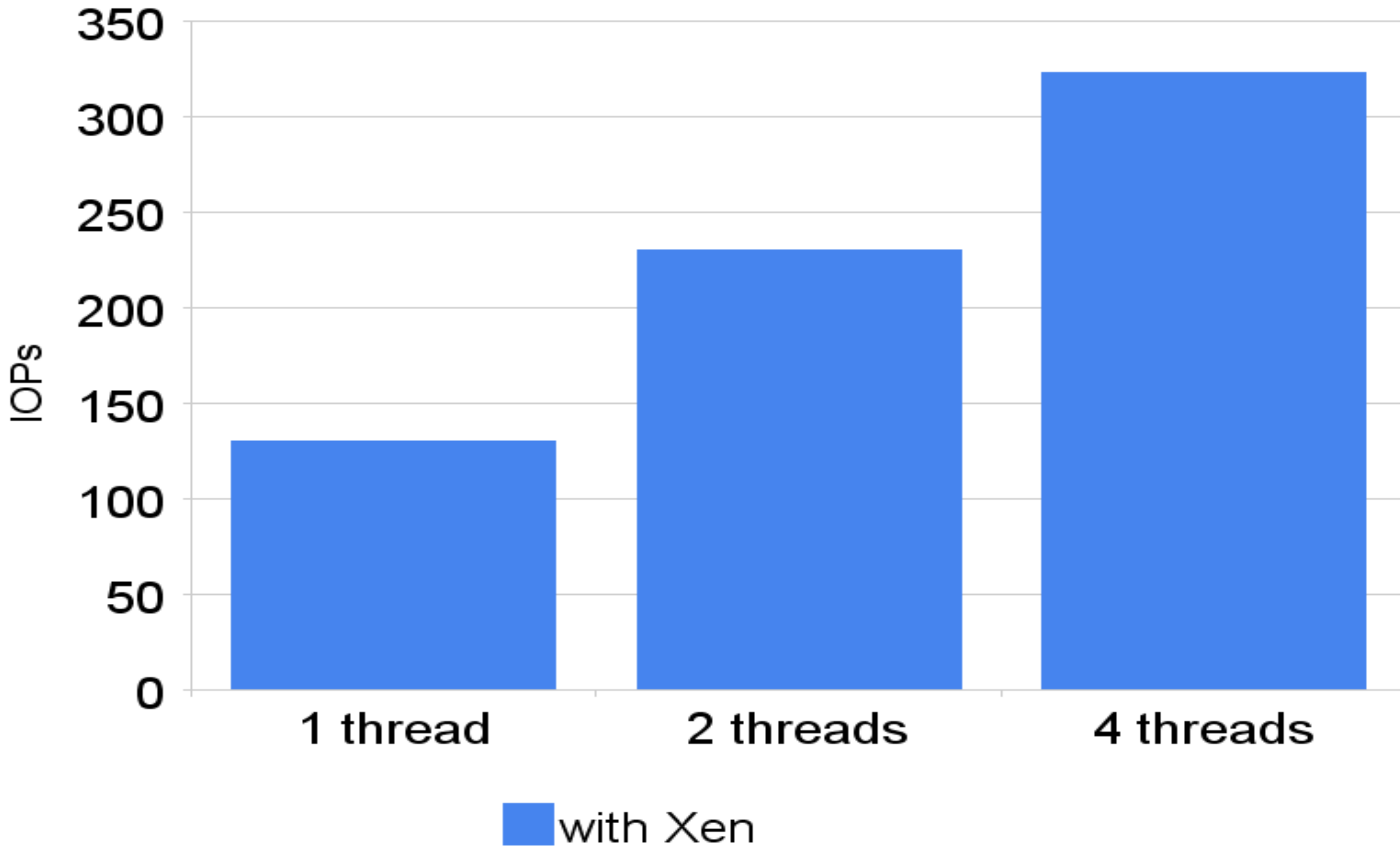
KVM random IO performance:

- sysbench fileio rndrd, 8G file



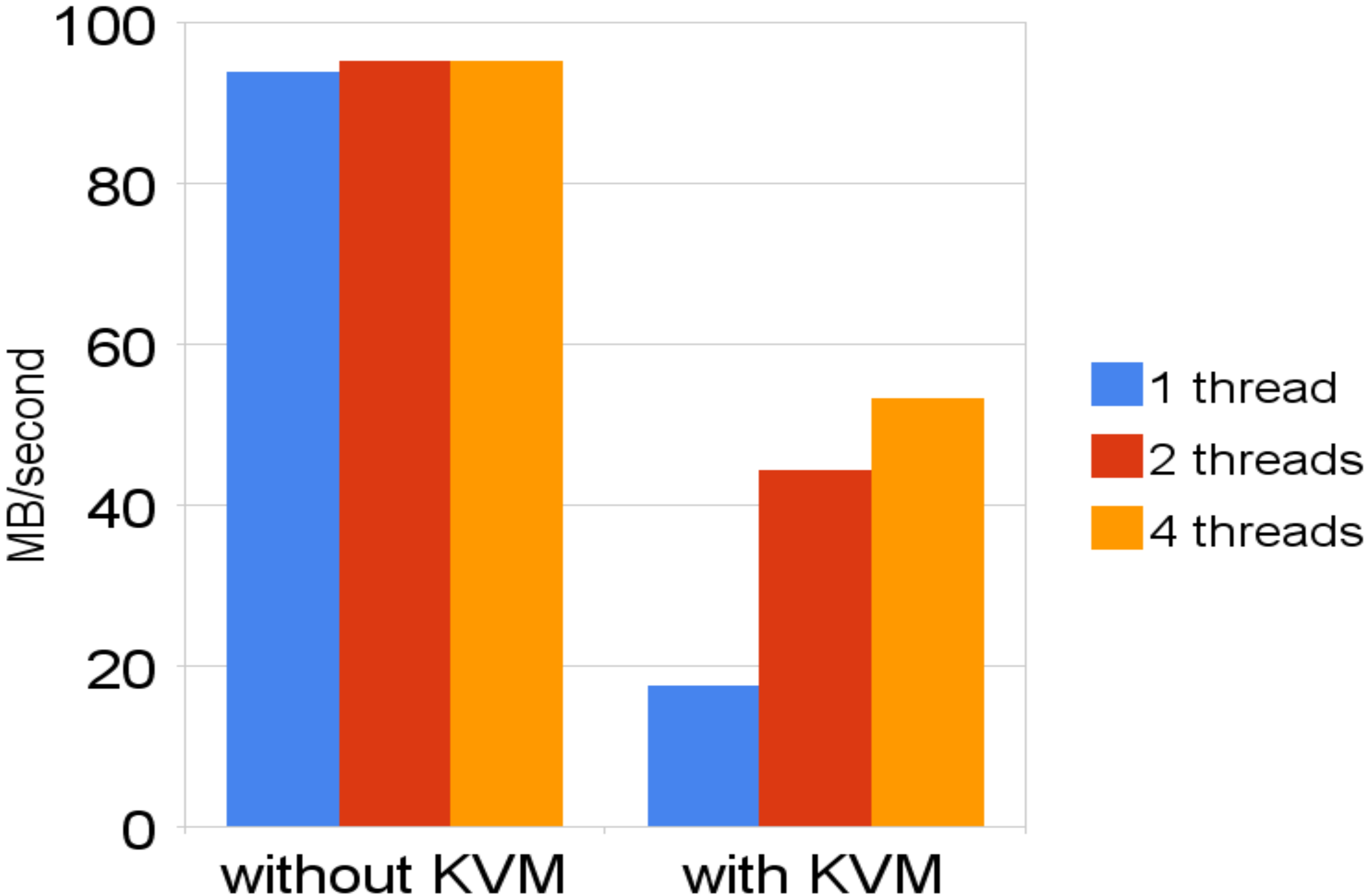
Xen random IO performance:

- sysbench fileio rndrd, 16G file



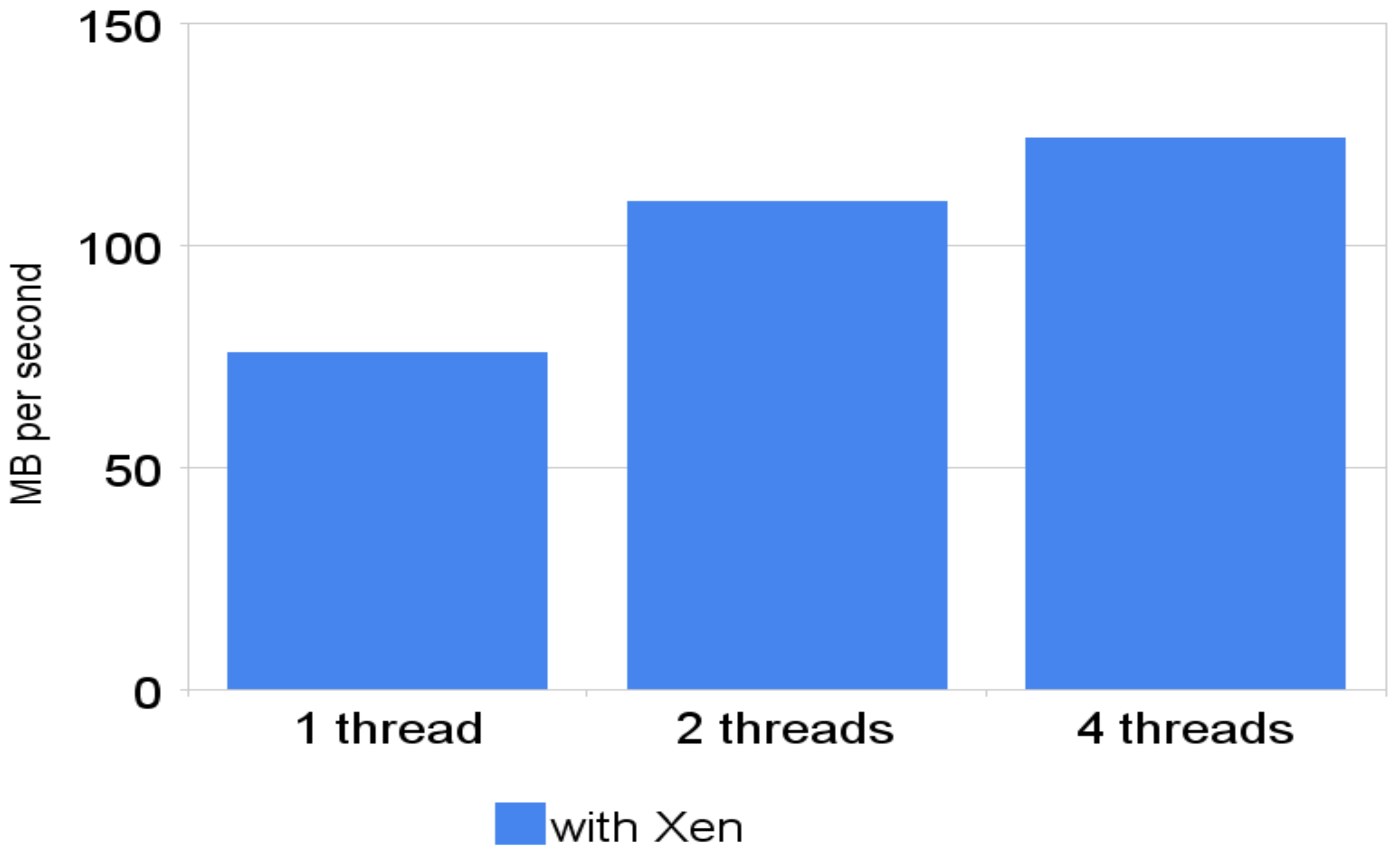
KVM sequential IO performance:

- sysbench fileio seqrd, 8G file



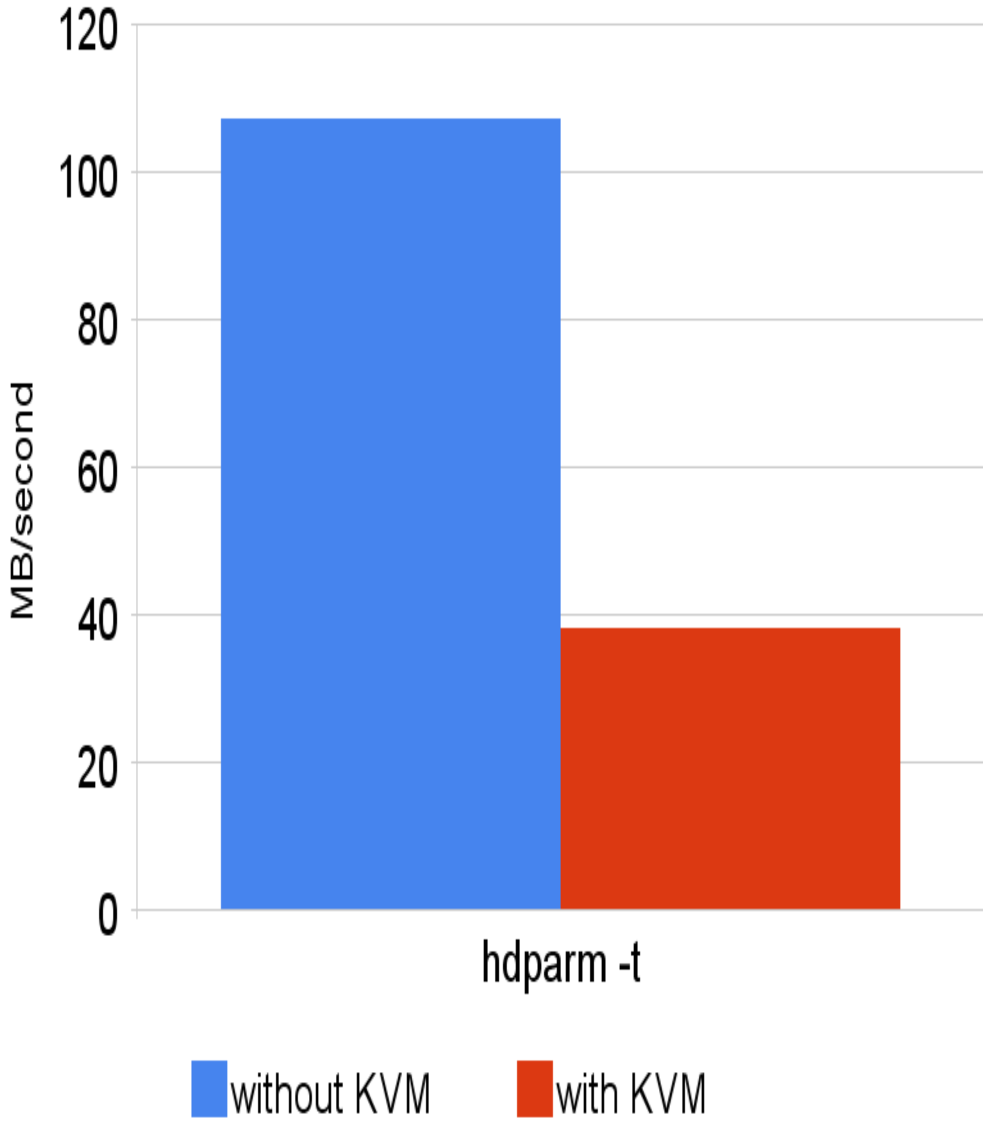
Xen sequential IO performance:

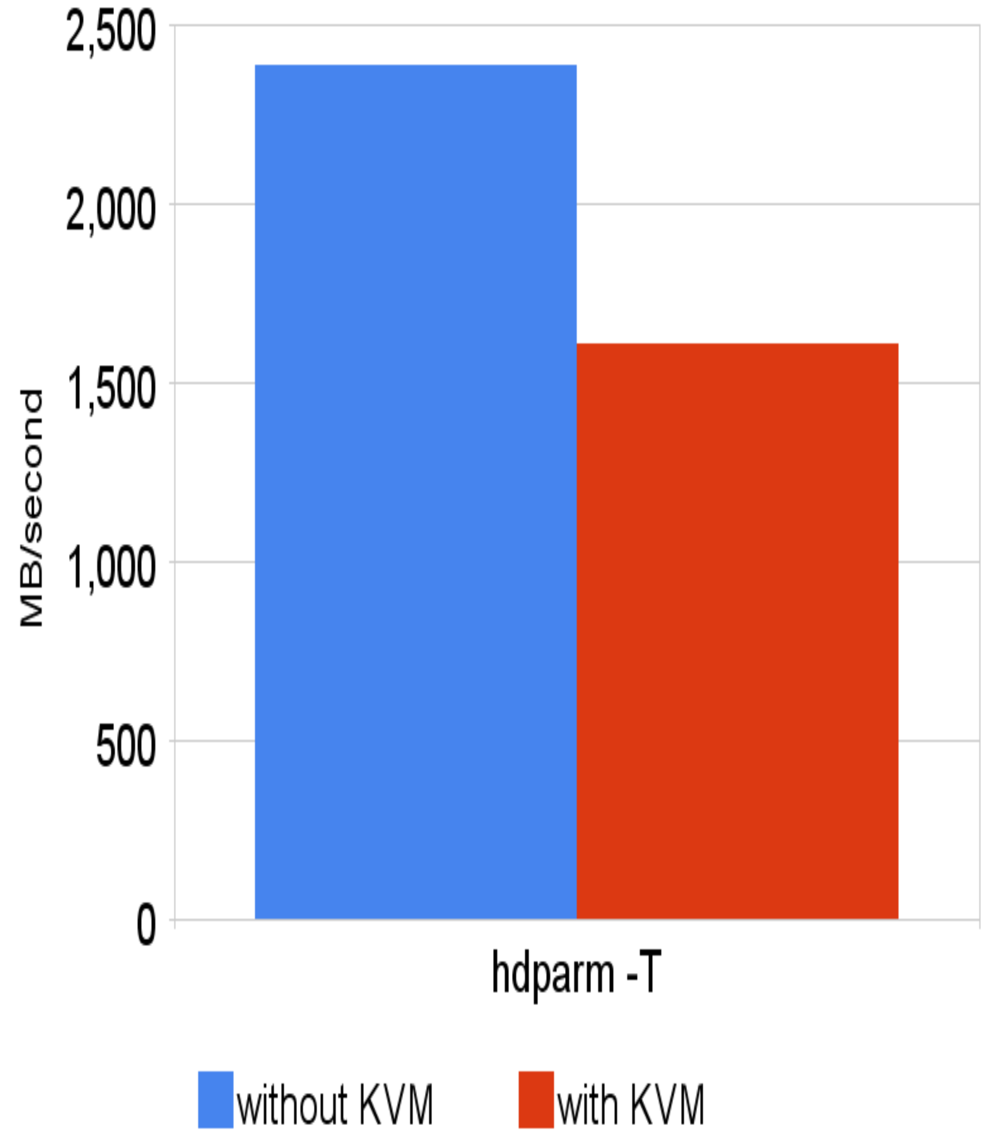
- sysbench fileio seqrd, 16G file



KVM sequential IO performance:

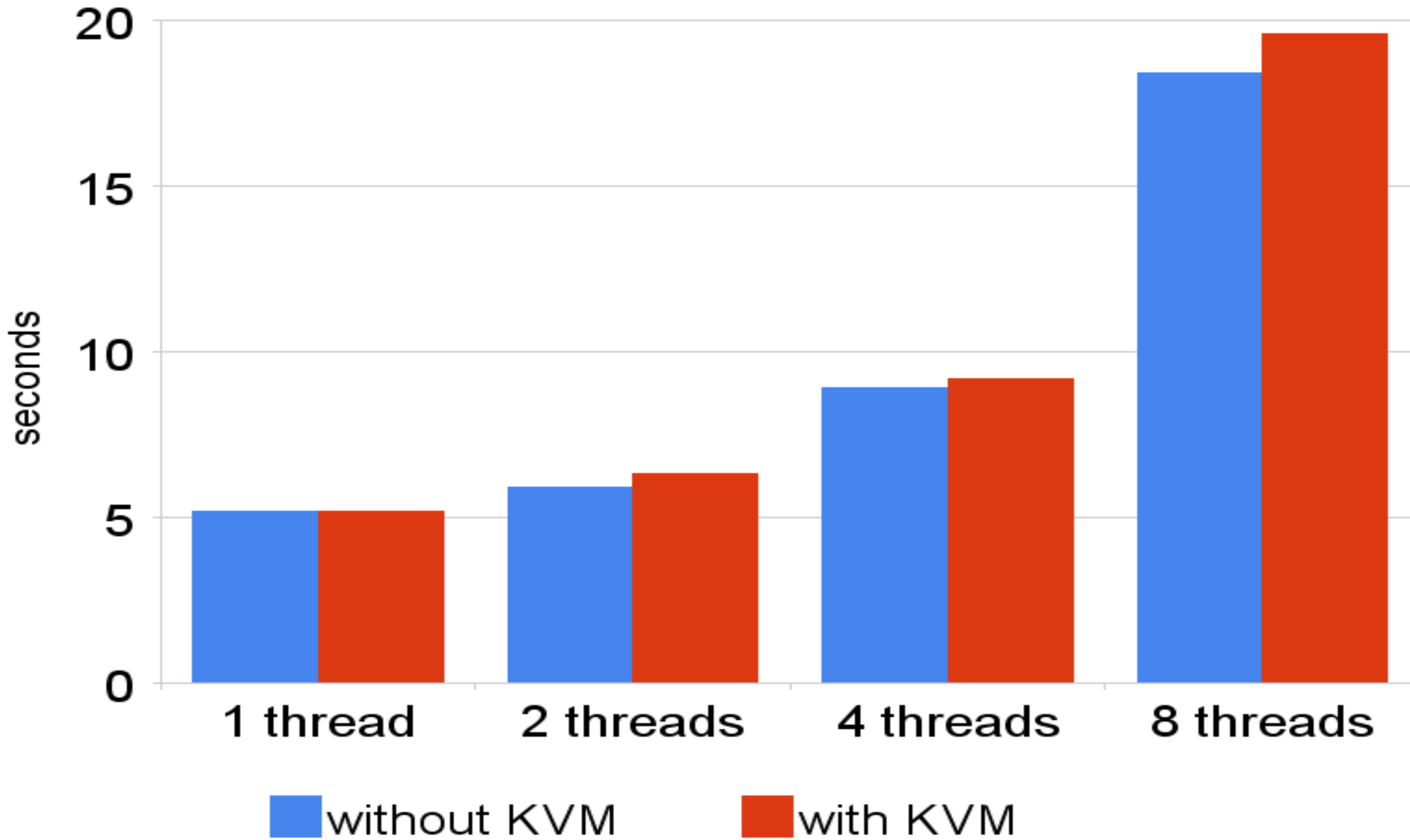
- hdparm -t, hdparm -T





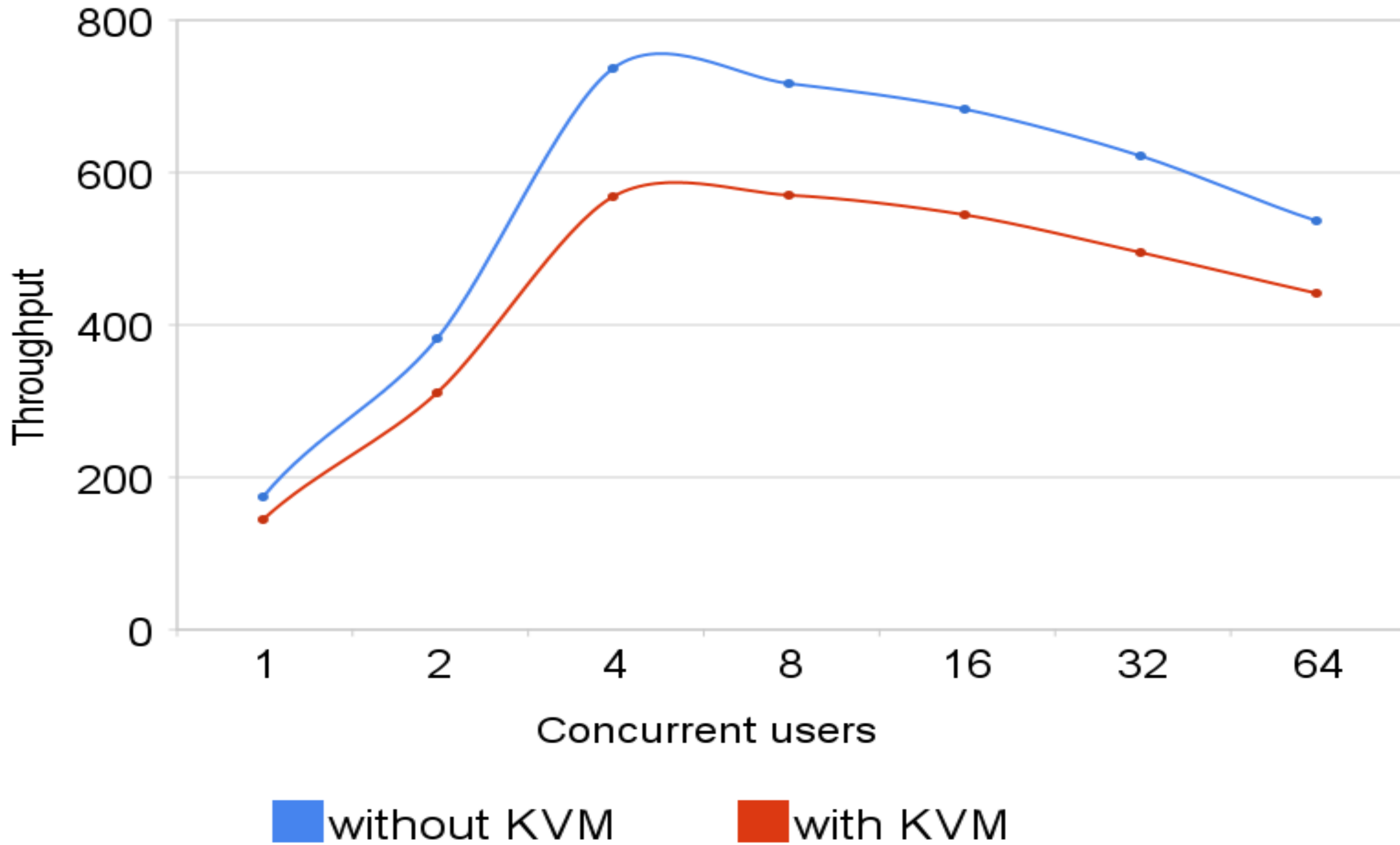
KVM CPU performance:

- modified wisconsin benchmark, CPU bound
- measure time to run all queries



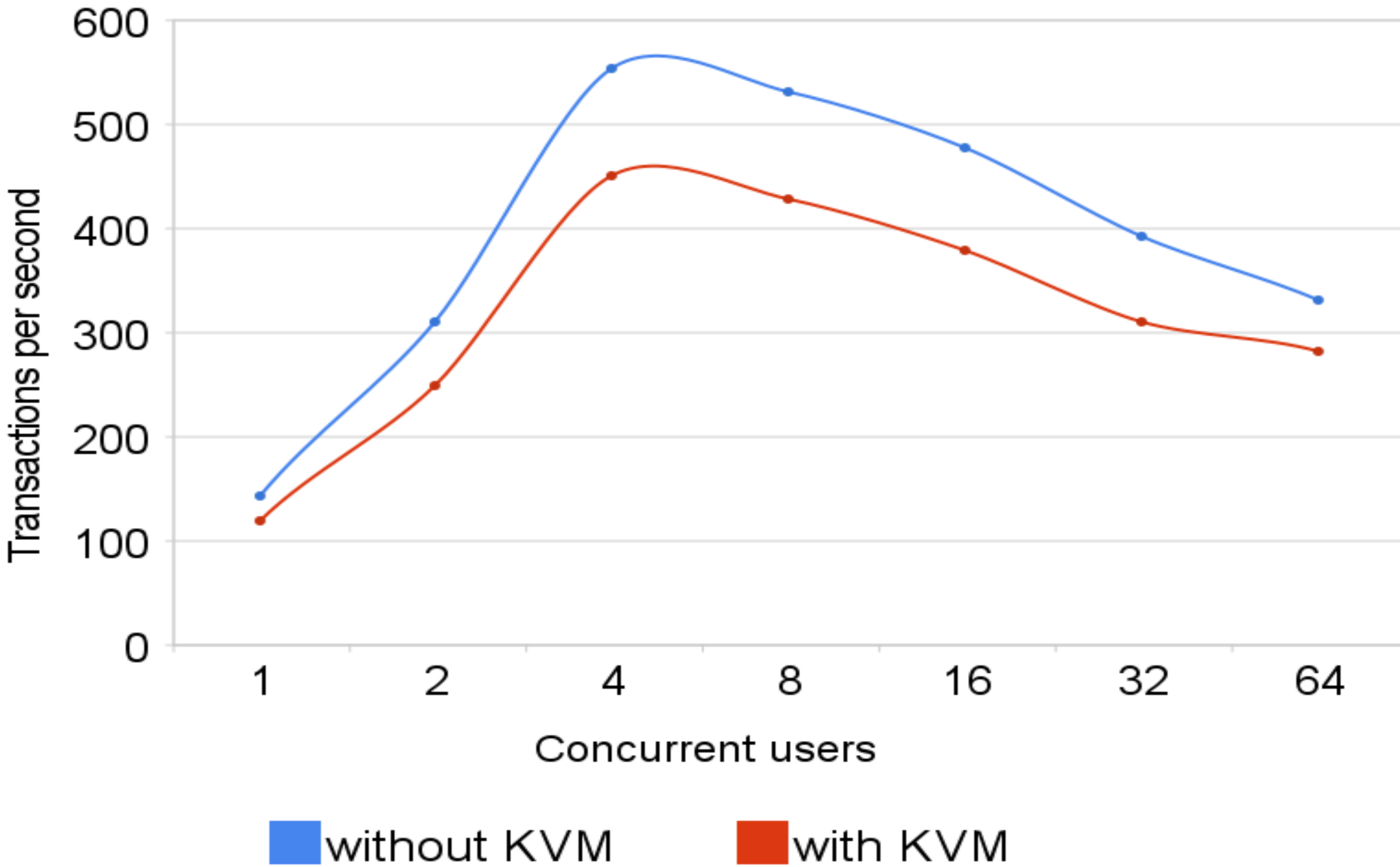
KVM CPU performance:

- modified sysbench readonly, CPU bound
- measure transactions per second



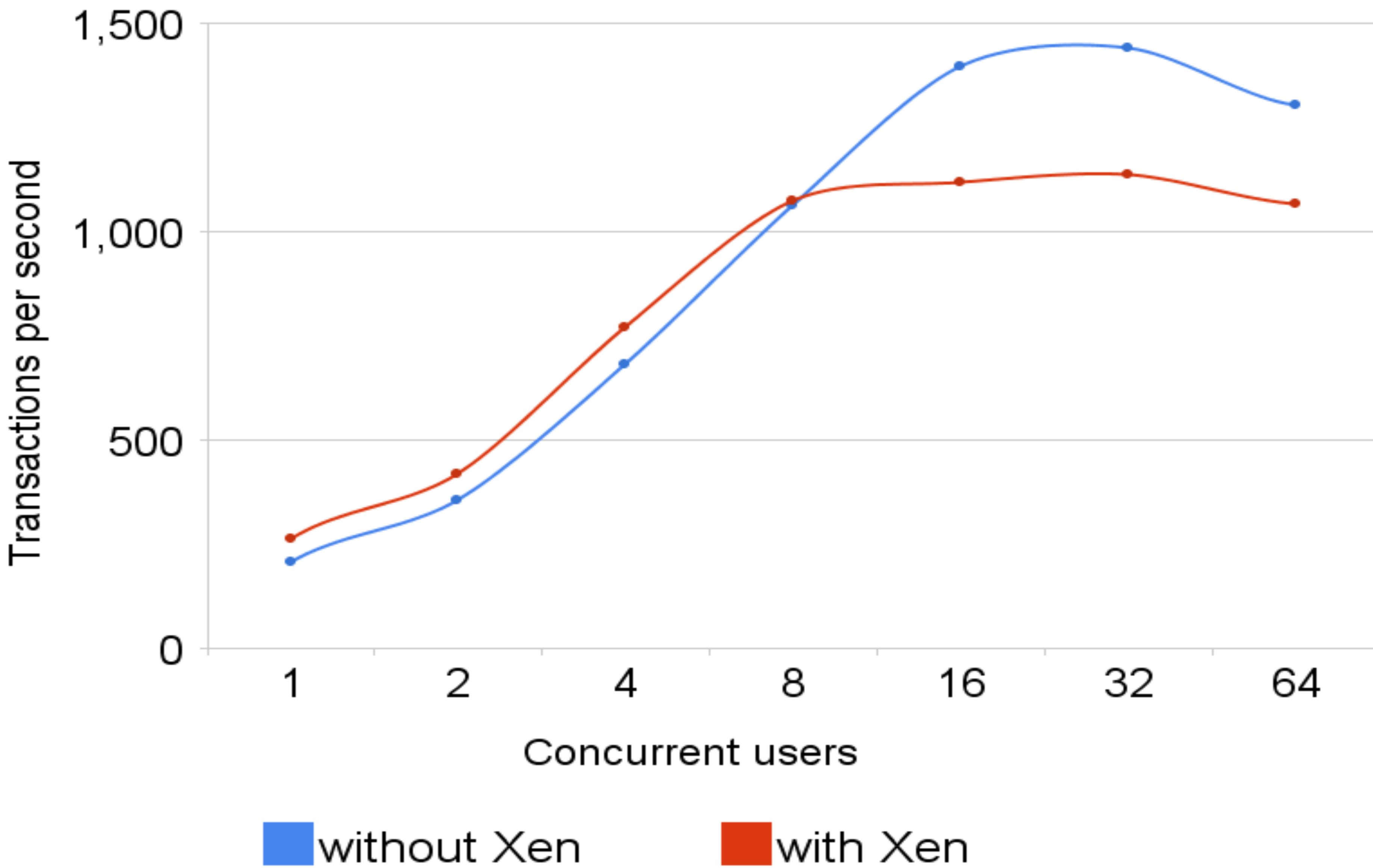
KVM CPU performance:

- modified sysbench readwrite, CPU bound
- measure transactions per second



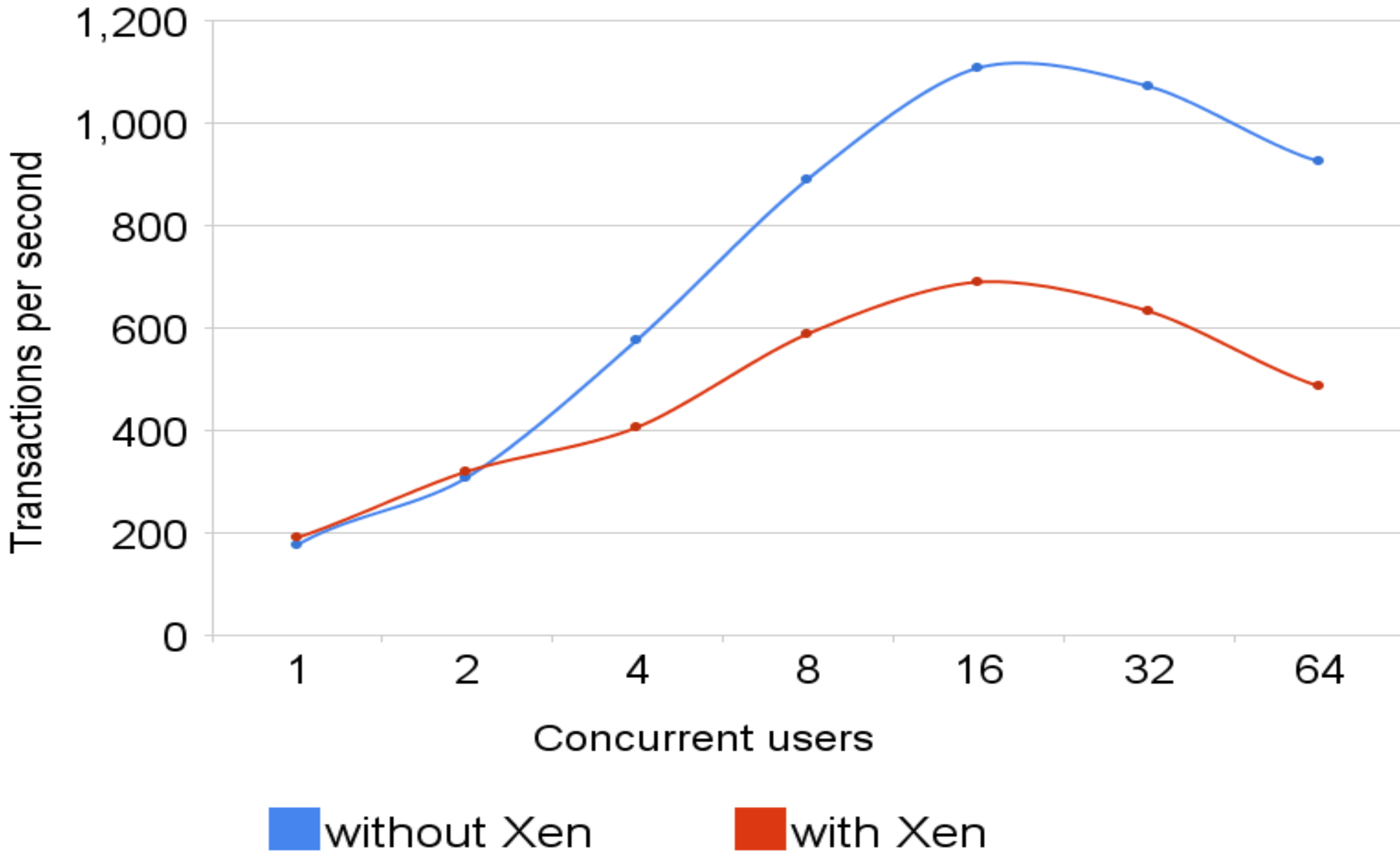
Xen CPU performance:

- modified sysbench OLTP readonly, CPU bound

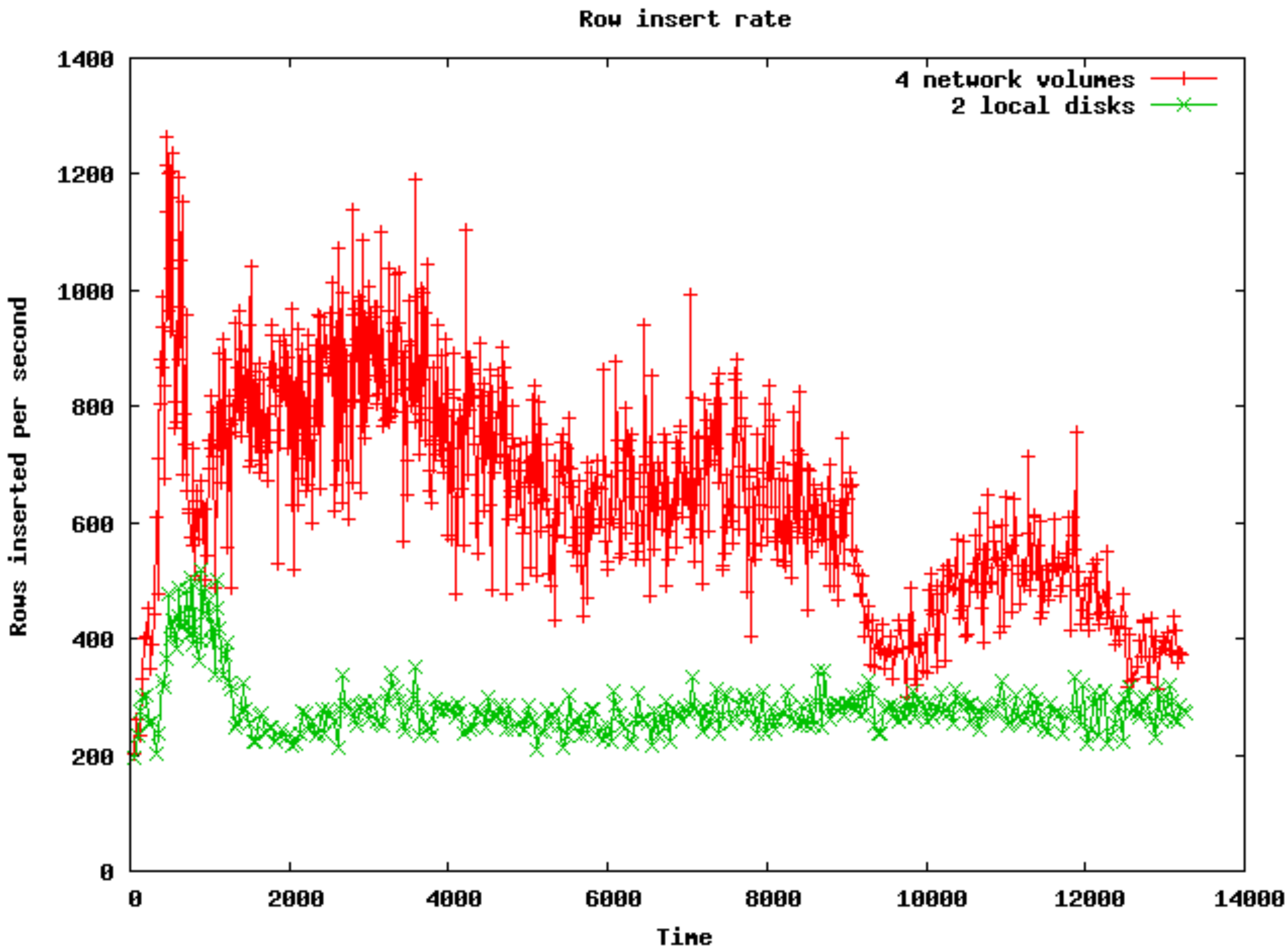


Xen CPU performance:

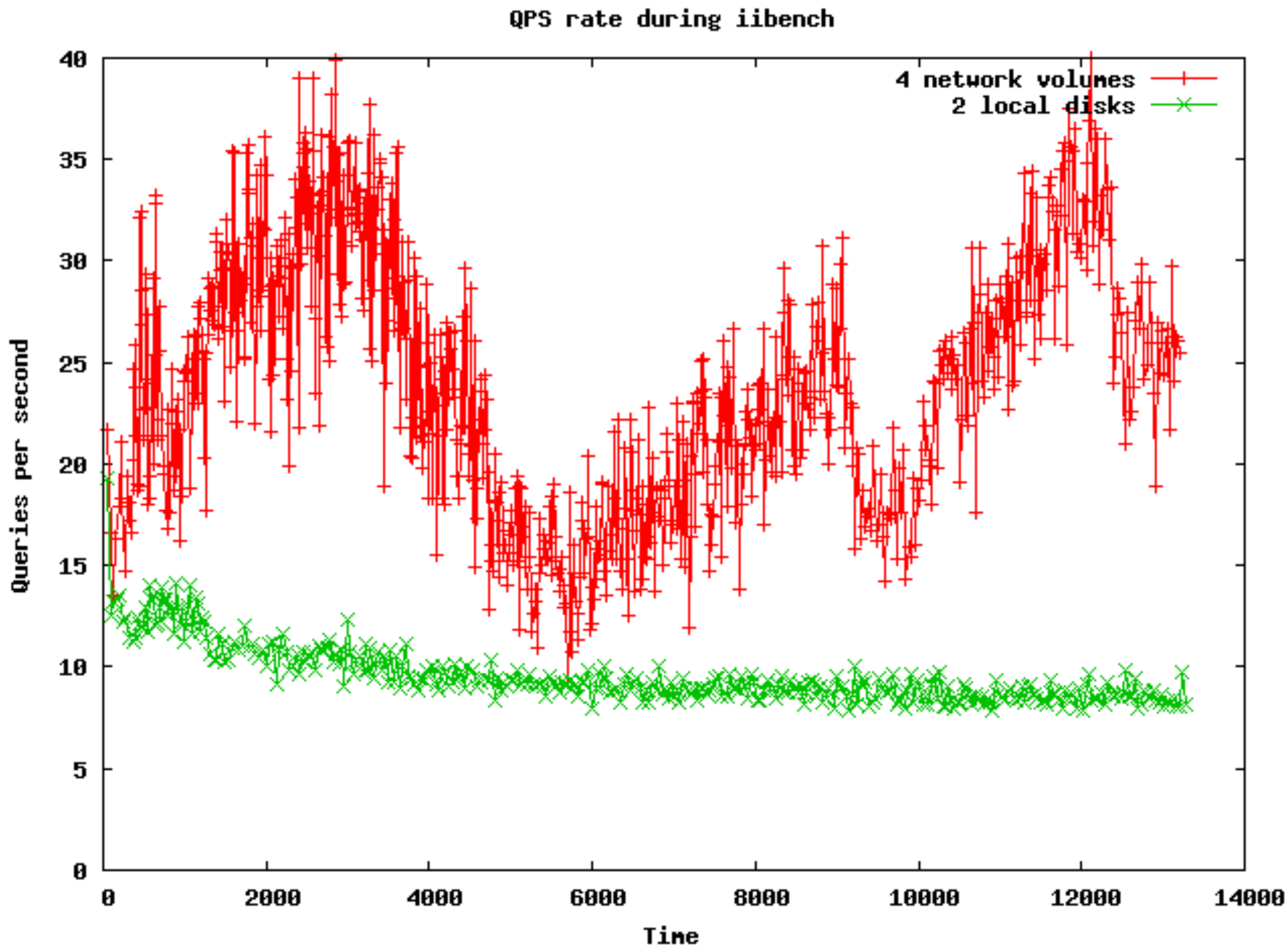
- modified sysbench OLTP readwrite, CPU bound



iibench insert rate comparing 2 local disks versus 4 network volumes



iibench QPS rate comparing 2 local disks versus 4 network volumes



Patches

All of these changes are available in some combination of the v3 Google patch, Percona builds and now ....

MySQL 5.4!

Make appropriate choices

- remote versus direct attached storage
- configuration
- storage engine
- IO scheduler
- file system
- patches