FreeBSD 7.0 and beyond

Kris Kennaway
The FreeBSD Project
kris@FreeBSD.org

February 23, 2008

Introducing FreeBSD 7.0

- FreeBSD 7.0 will be the next release of FreeBSD, and the first major release in 2 years.
- Due for release *within hours!*
- FreeBSD 7.0 brings major changes to the BSD and open source operating system landscape.

Outline:

I. The SMPng project: a 7 year journey
   - "Symmetric Multi-Processor, next generation"

II. Some of the major new features appearing in FreeBSD 7.0

III. What the future holds for FreeBSD

Multiprocessor support, old and new: FreeBSD 4.x

FreeBSD 4.x is a single-threaded kernel with limited multiprocessor support.

- Only one process at a time can execute in the kernel ("Giant lock" around entire kernel)

The historical BSD kernel architecture worked very well for single-processor systems. It fundamentally does not scale to multi-processor systems, which are now almost ubiquitous.

Multiprocessor support, old and new: The SMPng project

Goal: Re-design the FreeBSD kernel as a multi-threaded system, for "next generation" SMP support (June 2000)

- Multiple CPUs must be able to execute kernel code in parallel
- Balance the performance needs of Uni-Processor (UP) and SMP systems (not always different needs)
- A major challenge...
- ...now complete

SMPng and the Universal Development Model

The SMPng project followed a simple 3 step process:

SMPng, step 1: First, make it work; FreeBSD 5.x

FreeBSD 5.0-5.2.1 (2003-01-17 - 2004-02-22)

- Debut of the new architectural model for symmetric multiprocessor support in FreeBSD.

FreeBSD 5.3 (2004-11-06), 5.4 (2005-05-09)

- The fundamental architectural changes were largely in place
- Some initial progress with kernel parallelism by 5.3 and 5.4 (network stack, virtual memory, ...)

SMPng, step 2: Then make it work *well*; FreeBSD 6.x

FreeBSD 6.0 (2005-11-01), 6.1 (2006-05-08), 6.2 (2007-01-15), 6.3 (2008-01-18)

- Stabilized the work of the 5.x branch
- Performance benefits from subsequent development work
  - e.g. Virtual File System (VFS) and Unix File System (UFS) now allow parallel access
- Large parts of the kernel may now operate in parallel, with significant performance gains on many common workloads

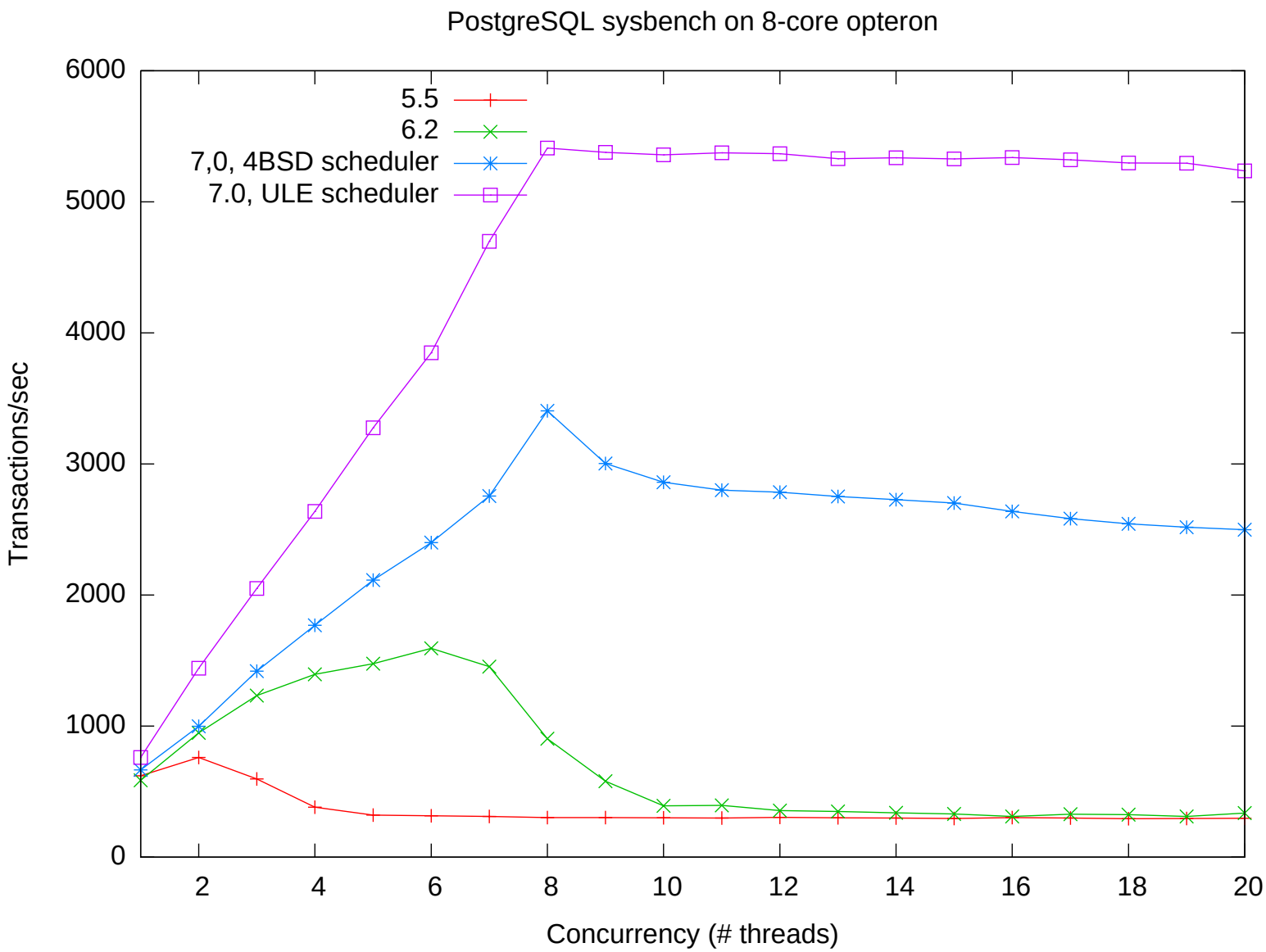
SMPng, step 3: Then make it **fast; FreeBSD 7.0**

- The goals of the SMPng project have been achieved
- The FreeBSD 7 kernel is a fully parallel system
  - The "Giant lock" is no longer present on almost all possible workloads
- Major shift of focus from correctness to optimization, with impressive results

A case study: SQL database performance

- "Online transaction processing" benchmark; `/usr/ports/benchmarks/sysbench`
- Transaction-based queries
- Read-only: no disk access to avoid benchmarking disk performance
- Clients and servers on the same system
- PostgreSQL 8.2.4 (process-based + System 5 Inter-Process Communication (IPC)
- MySQL 5.0.45 (thread-based)
- Test hardware:
  1. 4 * 2-core Opteron (amd64 mode)
     - 2.2GHz CPUs, 4 GB RAM
  2. 2 * 4-core Xeon E5320 (i386 mode)
     - 1.8GHz CPUs, 3.5GB RAM

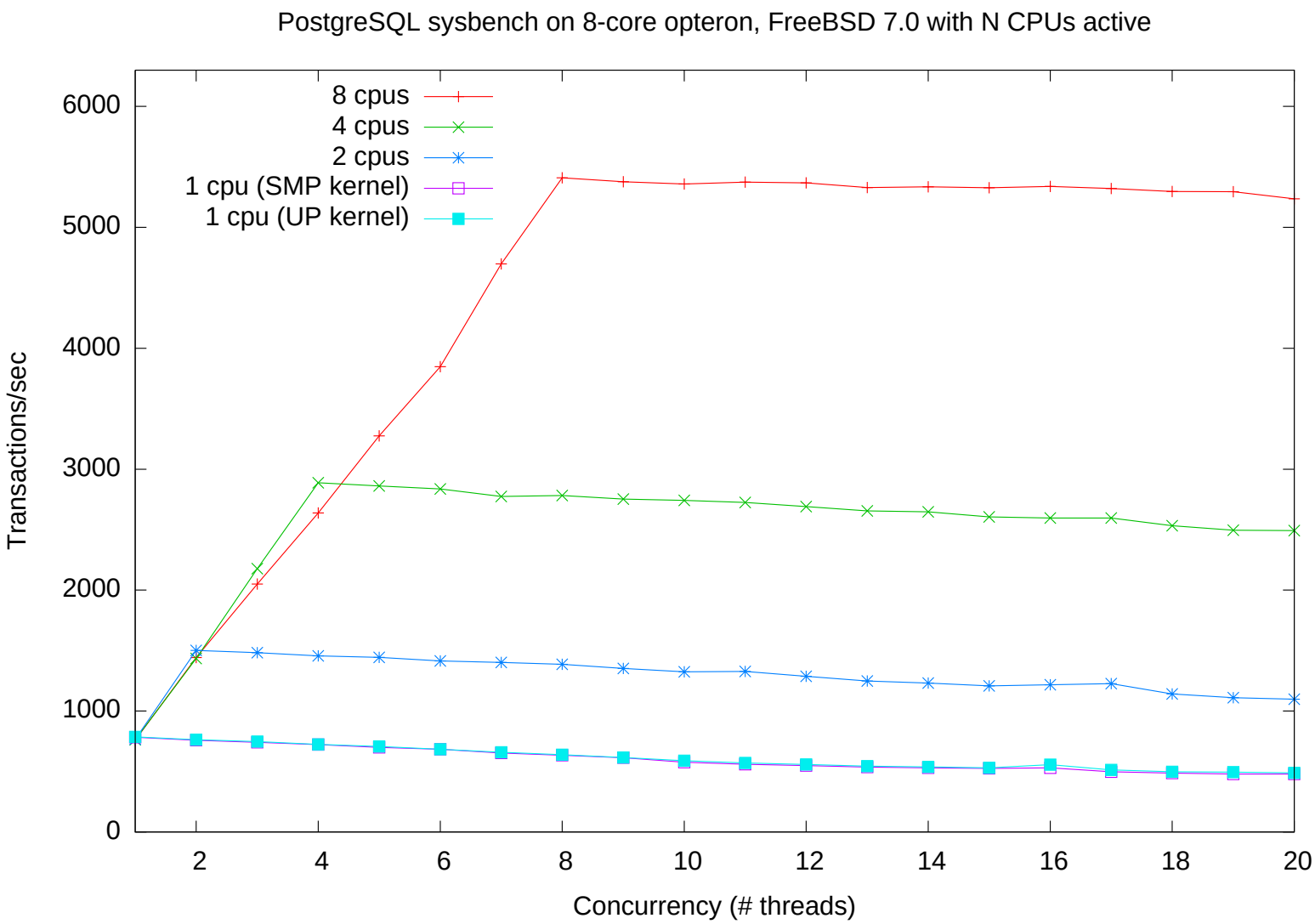
FreeBSD PostgreSQL performance: 5.5, 6.2 and 7.0



Performance of PostgreSQL

- The ULE scheduler has significantly better performance than 4BSD (historical BSD scheduler)
  - Better interactivity for desktop users also
  - 4BSD will remain the default in 7.0, changing in 7.1
  - You can easily switch to ULE by recompiling your kernel
- PostgreSQL with ULE has linear scaling to 8 CPUs and minimal degradation at higher loads; close to ideal performance from the hardware.
- No significant performance problems in the FreeBSD 7 kernel on this workload

FreeBSD 7.0: Scaling with varying number of CPUs

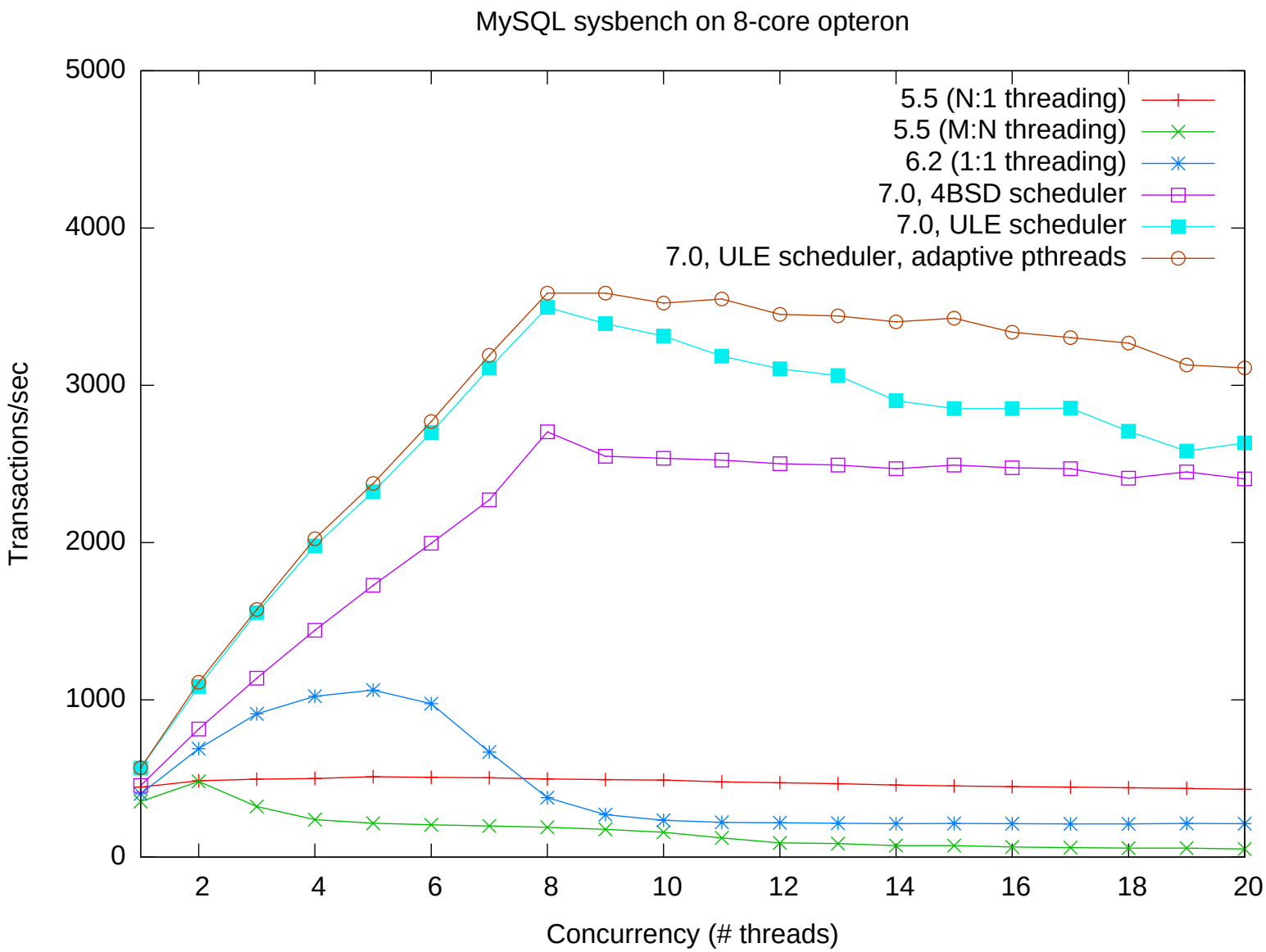


FreeBSD 7.0: Scaling with varying number of CPUs (2)

Notes:

- Performance from 1 $\longrightarrow$ 2 $\longrightarrow$ 4 $\longrightarrow$ 8 CPUs scales linearly
- Consistently stable performance at high loads
- No significant overhead from SMP kernel on UP system

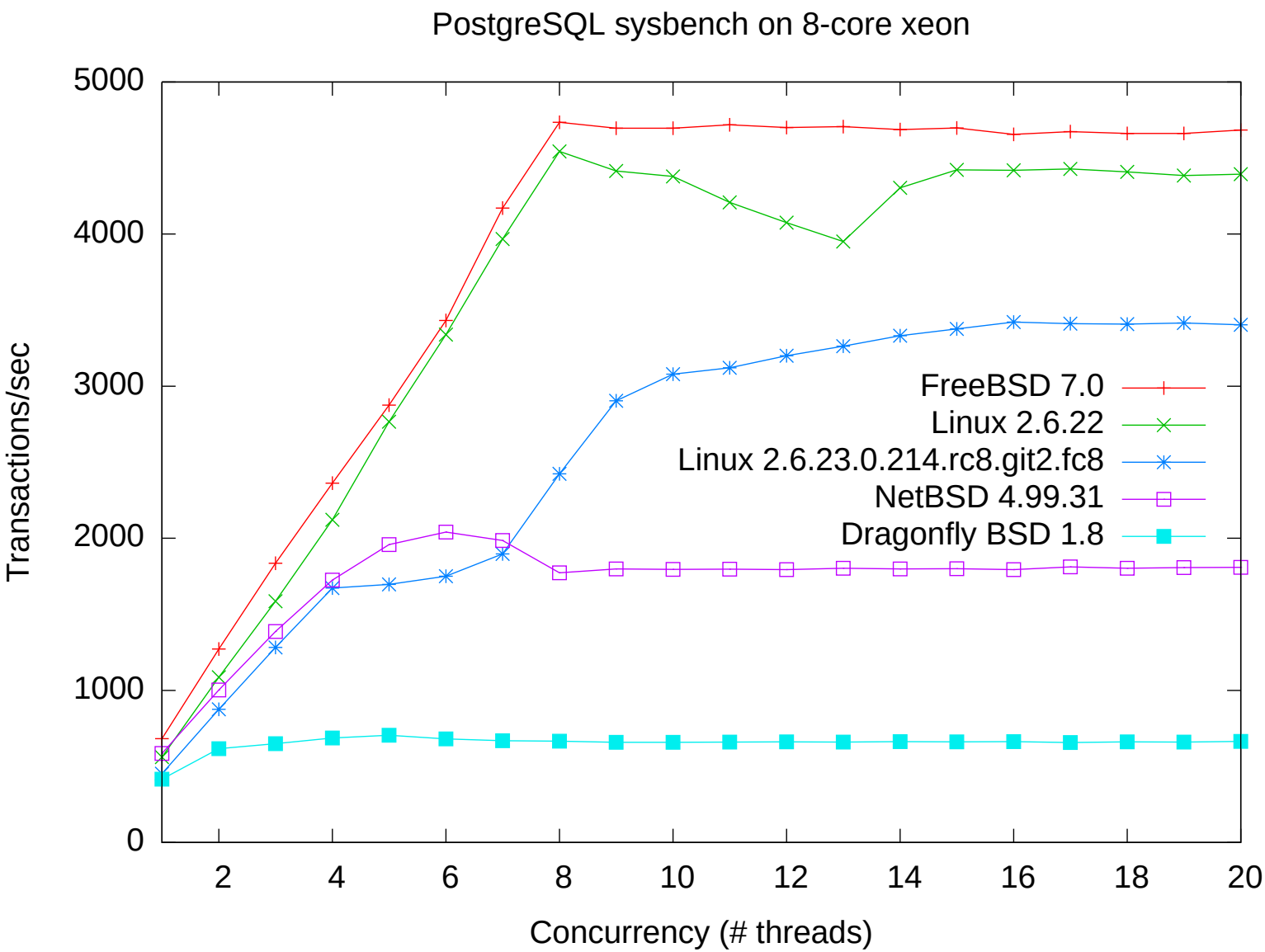
FreeBSD MySQL performance: 5.5, 6.2 and 7.0



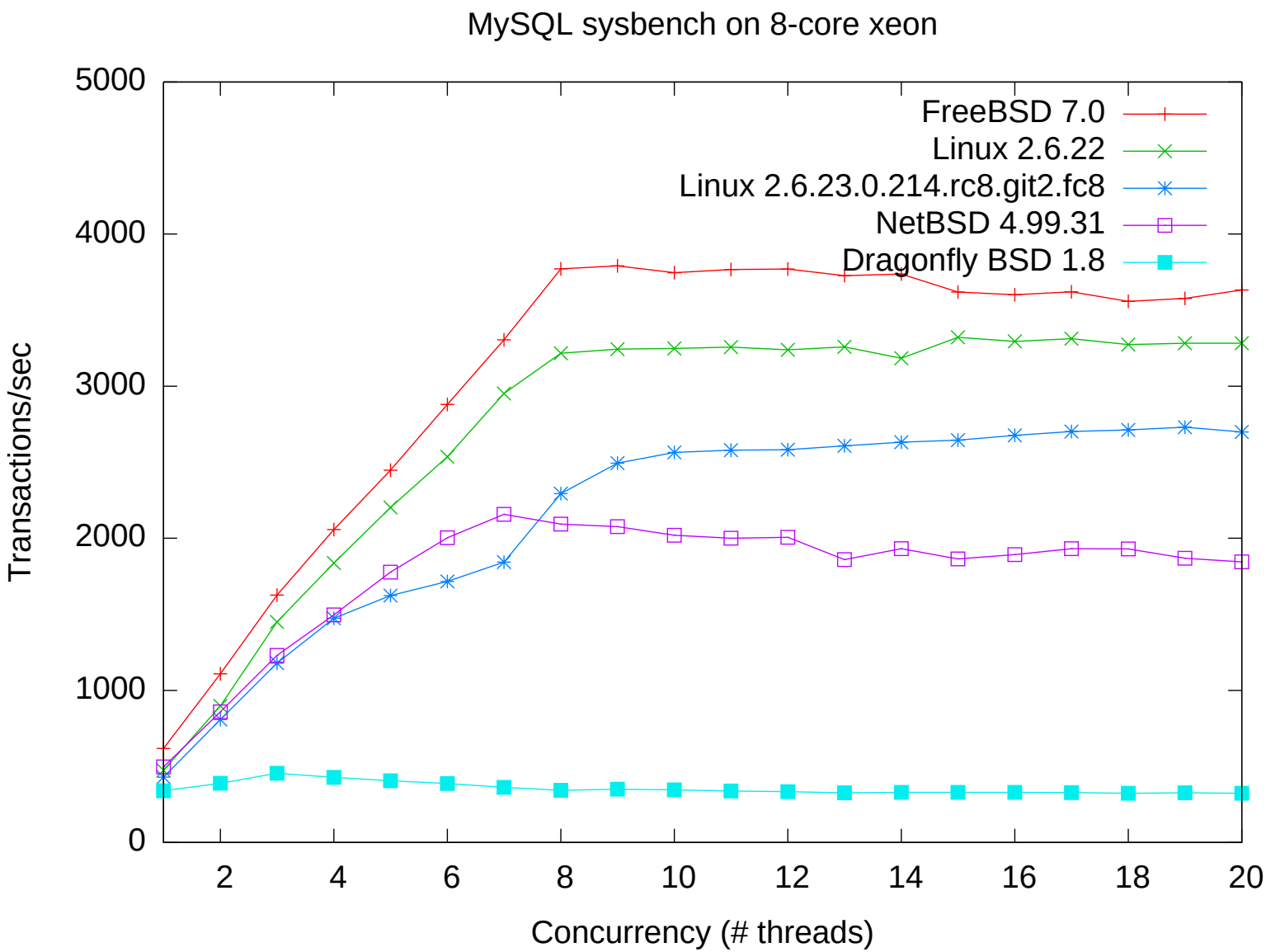
Understanding MySQL performance

- Again, linear scaling up to 8 client threads (= # CPUs)
- The degradation above 8 threads is due to scaling problems within MySQL (not a FreeBSD kernel issue)
- Heavy contention on pthread mutexes within the application
- Recent change to libpthread to reduce the performance loss from heavily contended pthread mutexes
  - Non-portable "adaptive" mutex type defined by glibc, used by MySQL
- Ultimately a MySQL architectural problem
- NB: On this benchmark PostgreSQL is 35% - 45% faster than MySQL at all loads

FreeBSD vs other operating systems: PostgreSQL



FreeBSD vs other operating systems: MySQL



Comments on other operating systems

Linux

- Major improvement since we initially publicized our benchmarks in February 2007
  - The 2.6.20.1 kernel performs very badly
- Publication of FreeBSD 7.0 performance comparisons motivated improvements in Linux
- 2.6.22 is still 15% slower than FreeBSD 7.0
- The new CFS scheduler in 2.6.23 is "Completely Fair" ...to FreeBSD

NetBSD

- Good initial progress on SMP support

Part II: New features debuting in FreeBSD 7.0

FreeBSD 7.0 brings updates to almost every part of the operating system (more than 18000 code changes), as well as several major new features.

1. Filesystem/storage
2. Networking
3. New CPU architectures
4. Security systems

Filesystem and storage subsystem changes

New filesystems

- ZFS
  - Sun's amazing new filesystem moves the goalposts.
- unionfs: overlay multiple filesystem hierarchies into one. Broken for many years but now usable again.
- XFS support (read-only)
- CODA distributed filesystem support fixed
- Performance improvements for various filesystems (including NFS)

Storage subsystem changes:

- SCSI layer (CAM) is now parallelized, including many drivers. Performance benefits for SCSI device access.
- Various new GEOM (pluggable storage layer) modules

Network stack changes

- Complete elimination of giant lock from network stack
- On-going cleanup and performance work
- SCTP (Stream Control Transmission Protocol)
- Migration from KAME IPSec to Fast IPSec
  - Improved performance
  - Hardware acceleration with cryptographic accelerators
  - Both IPv4 and IPv6

Network drivers; security subsystems

- Support for commonly encountered 10 gigabit ethernet drivers: Chelsio (cxgb), Intel (ixgbe), Myricom (mxge), Neterion (nxge)
- Transmit Segmentation Off-load (TSO)/Large Receive Off-load (LRO); off-load send/receive into the ethernet driver
- New devices supported
- Much improved wireless (802.11) layer

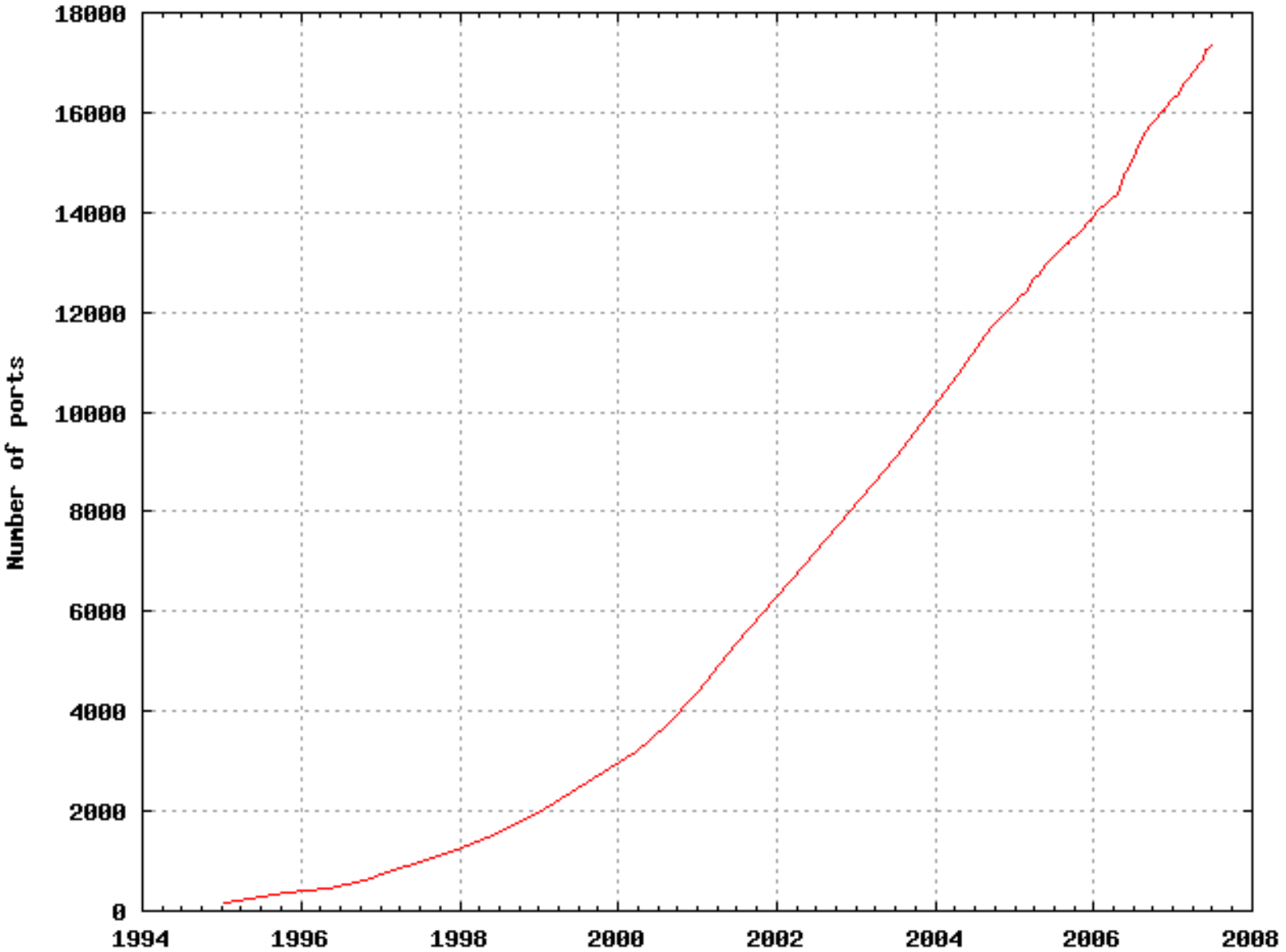
Security: Audit subsystem

- fine-grained, configurable logging of security-relevant events
  - System calls, application and user space activities
- Builds on the other advanced security features developed by the TrustedBSD project for FreeBSD

New CPU architectures

- Improved support for ARM architecture
  - FreeBSD/ARM used as the basis for growing number of embedded devices
- Sun Ultrasparc T1 (preliminary)
  - 8 cores, 4 threads per core = 32 logical CPUs per package
  - A very interesting new CPU architecture, and one to watch in the future
  - T2: 8 threads * 8 cores = 64 logical CPUs per package!
- X-box!

Growth of FreeBSD Ports Collection

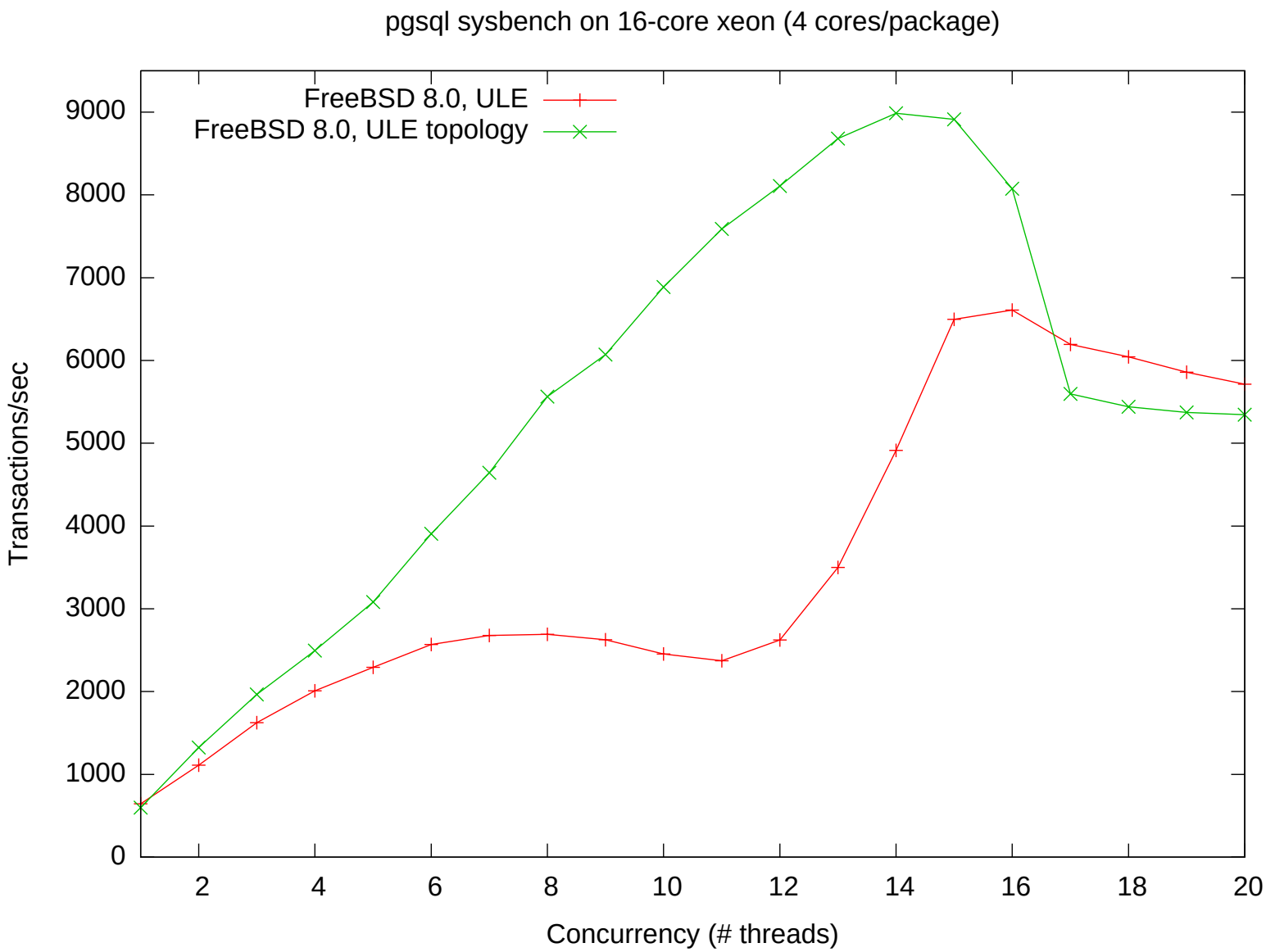


Part III: What the future holds for FreeBSD

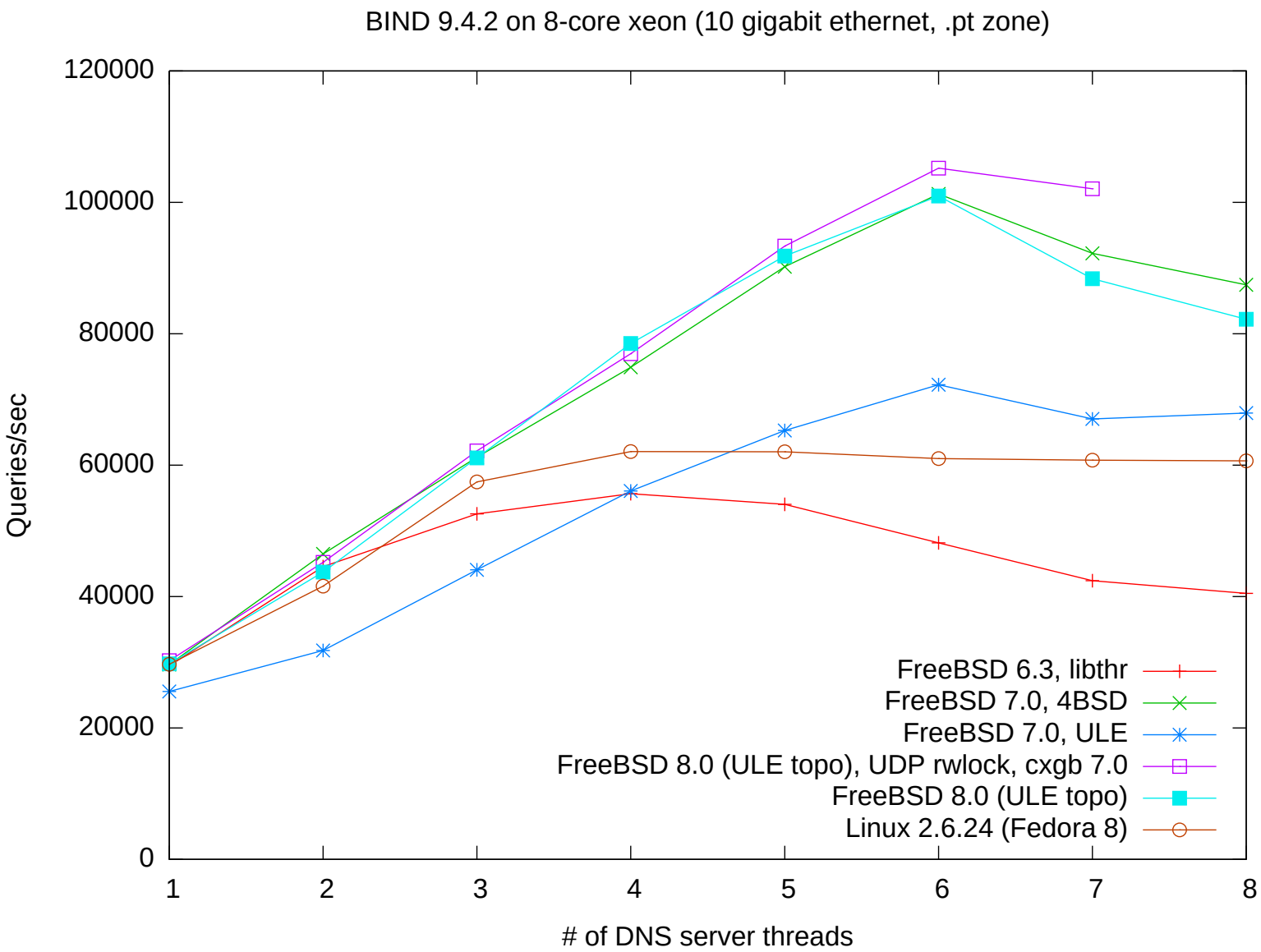
As we finalize the launch of the FreeBSD 7.x branch, development is also beginning on FreeBSD 8.0-CURRENT, due some time in 2009 (maybe).
Some ongoing performance work:

- 16-core systems (AMD/Intel)
- Improved parallel network performance
- Filesystem performance
- **If you find a workload that FreeBSD 7.0 performs poorly on, we want to hear about it!**

Topology-aware ULE scheduler



Improved parallel network performance



Further work in progress

More work in progress:

- Virtualization support: xen, network stack virtualization, ...
- Porting FreeBSD (again) to MIPS architecture
- DTrace support from Sun; powerful and extensible debugging and system analysis framework
- BLUFFS: BSD Logging Updated Fast File System. UFS with filesystem-level journalling.
- Stuff we haven't even thought of yet!

Summary

- FreeBSD 7.0 brings FreeBSD back to the forefront of OS performance on modern hardware (it's good to be back).
- Provides advanced features not available in other open source operating systems
- An attractive platform for both high end and embedded hardware.
- An excellent new foundation for the years ahead.

Where can I get FreeBSD 7?

- Download the CD images:

  `ftp://ftp.freebsd.org/pub/FreeBSD/releases/`