



Sanjay Rao
Sr Software Engineer, Red Hat

Optimizing Database Performance

- System Performance
 - I/O performance
 - Memory
 - CPU

- Database Performance
 - Database layout
 - Database tuning
 - Application tuning

System Performance

■ I/O Performance

• Multiple HBAs

- Install device-mapper-multipath
- Provides multipathing capabilities and lun persistence
- Default settings in /usr/share/doc/device-mapper-multipath-0.4.7/multipath.conf.defaults

• Understand storage features / limitations

- Maximum random and sequential read and writes per port
- Maximum random and sequential read and writes for the controller

• Low level I/O numbers

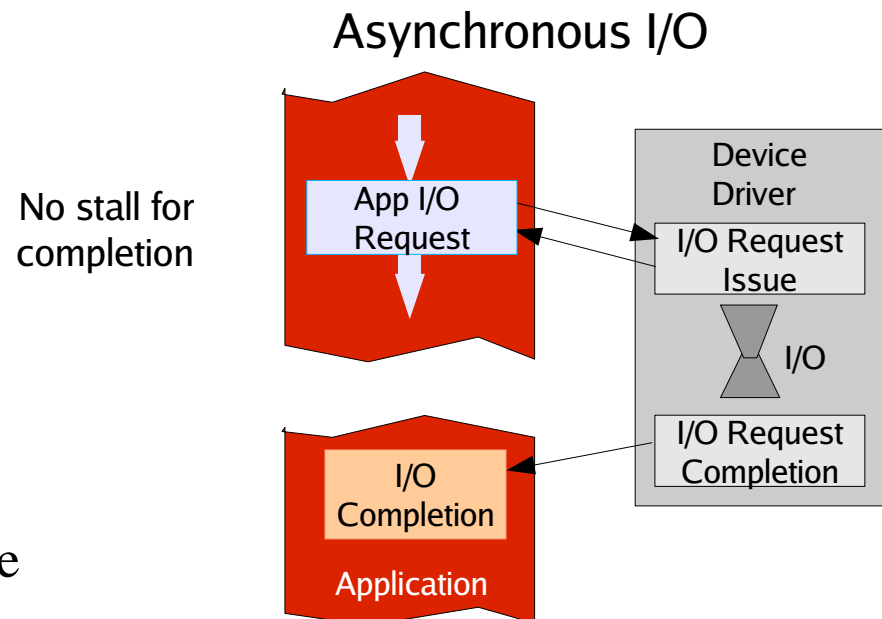
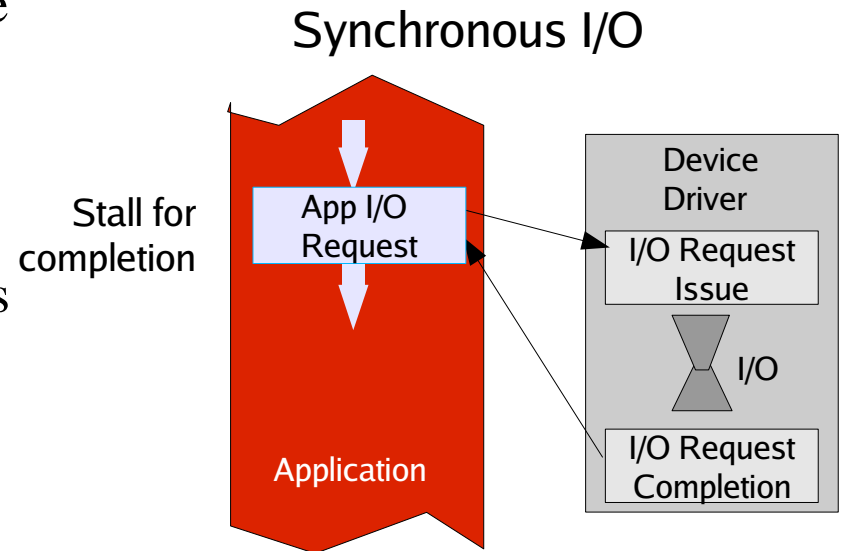
- Tools to use dd , aiocd , aio-stress
- Run I/O representative of the database implementation

• I/O Schedulers

- CFQ, Deadline, AS, Noop

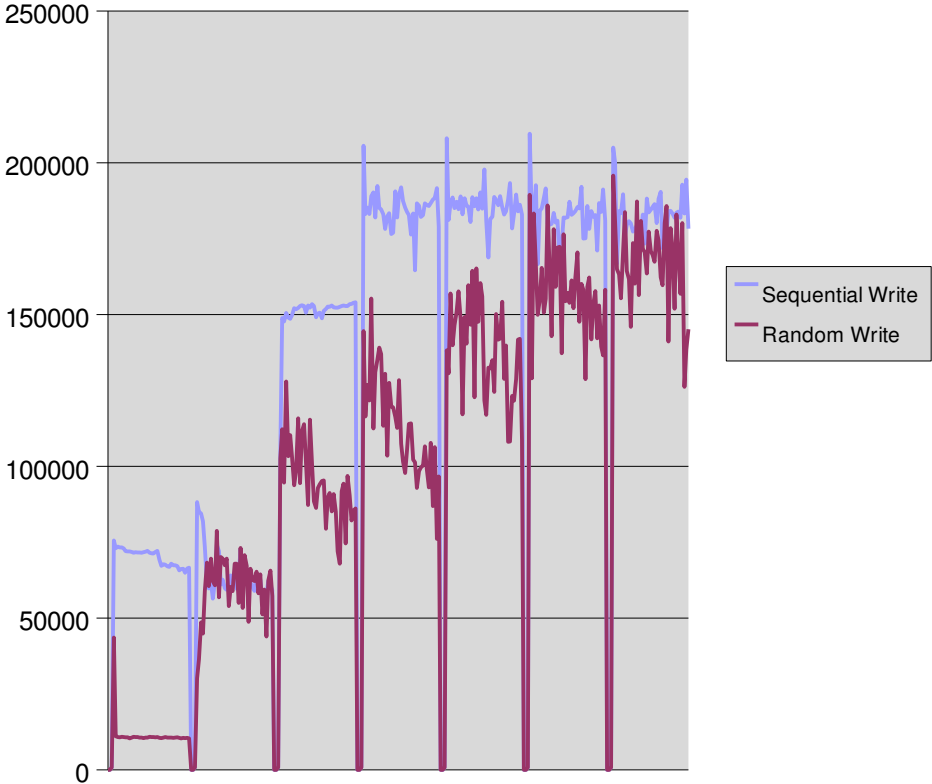
Asynchronous I/O to File Systems

- Allows application to continue processing while I/O is in progress
 - Eliminates Synchronous I/O stall
 - Critical for I/O intensive server applications
- Red Hat Enterprise Linux – since 2002
 - Support for RAW devices only
- With Red Hat Enterprise Linux 4, significant improvement:
 - Support for Ext3, NFS, GFS file system access
 - Supports Direct I/O (e.g. Database applications)
 - Makes benchmark results more appropriate for real-world comparisons

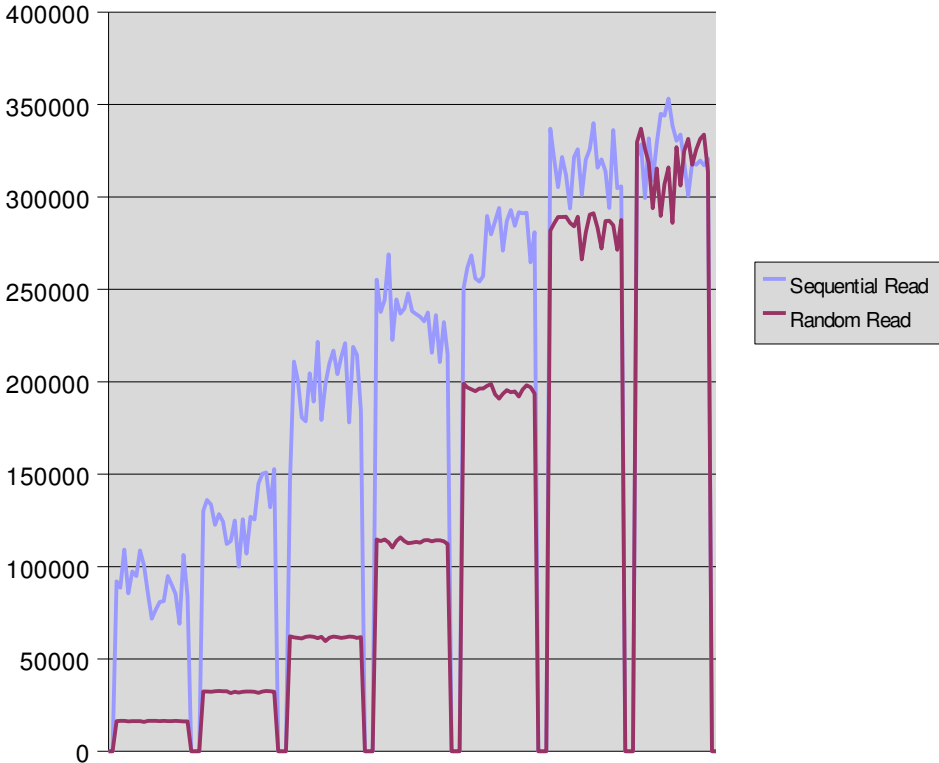


AIO - DIO Performance

AIO - DIO Writes

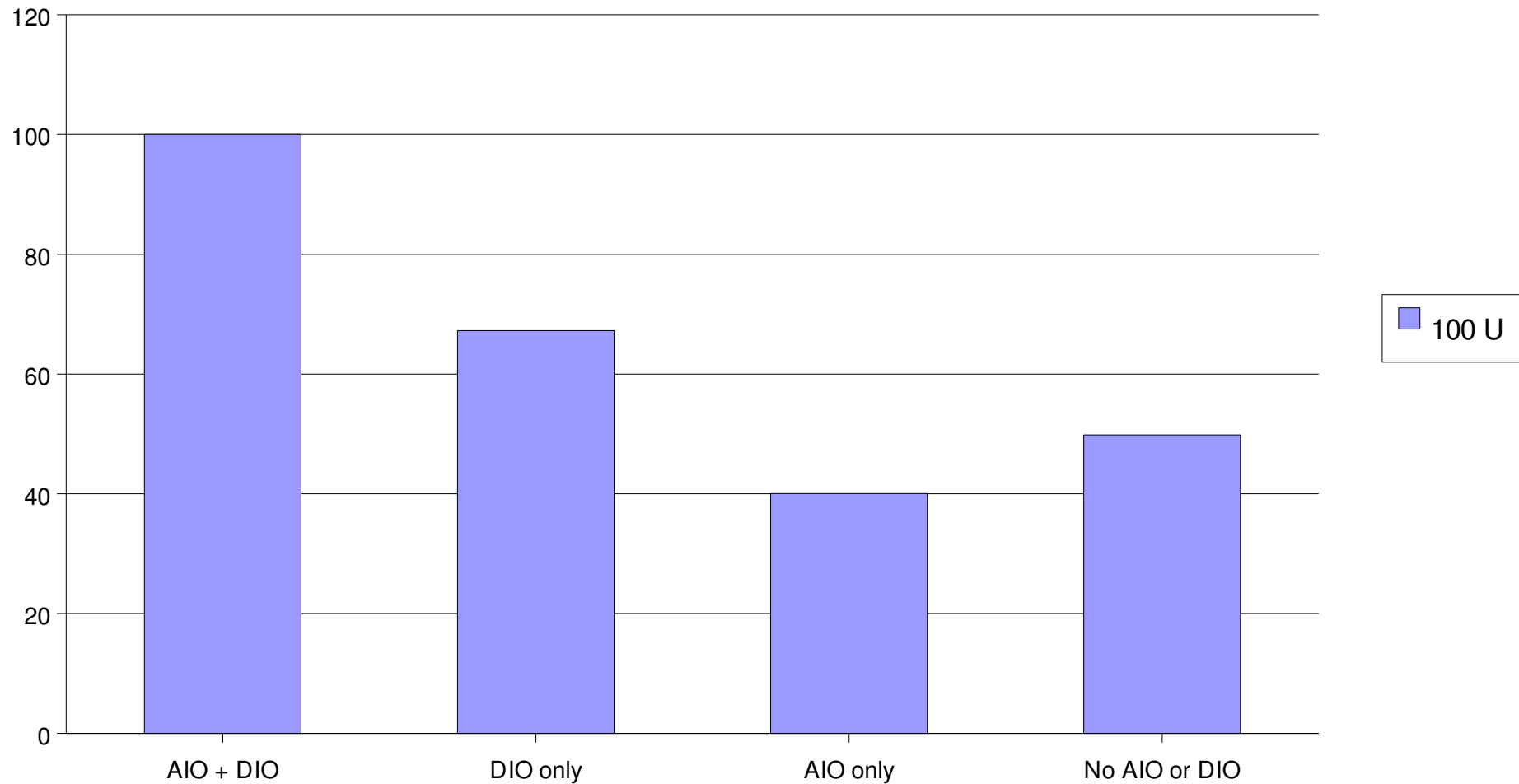


AIO - DIO Reads



RHEL5.1 Oracle 10.2 - I/O options

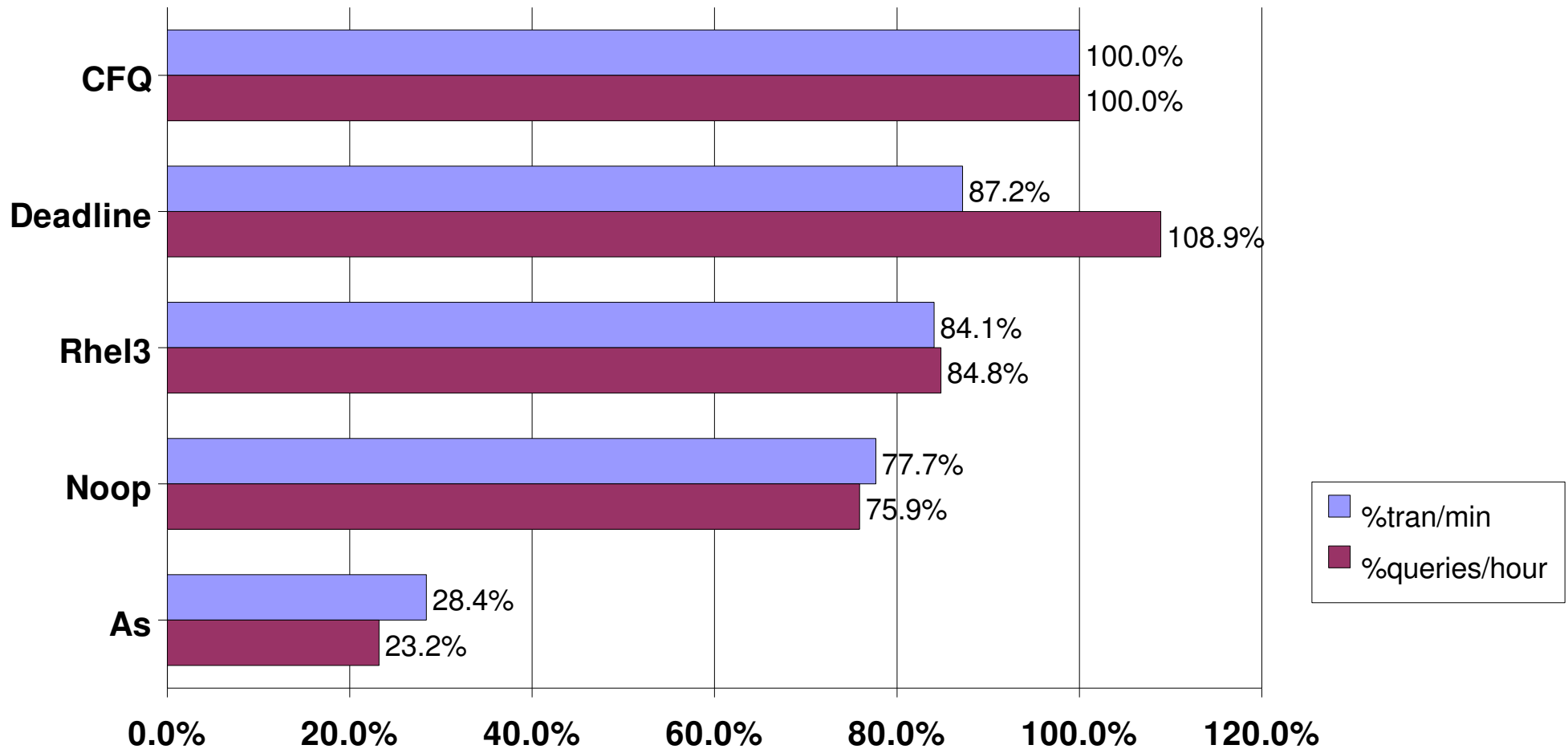
RHEL5.1 with Oracle 10.2 - I/O Options



Disk IO tuning - RHEL4/5

- RHEL4/5 – 4 tunable I/O Schedulers
 - CFQ – elevator=cfq. Completely Fair Queuing default, balanced, fair for multiple luns, adaptors, smp servers
 - NOOP – elevator=noop. No-operation in kernel, simple, low cpu overhead, leave opt to ramdisk, raid cntrl etc.
 - Deadline – elevator=deadline. Optimize for run-time-like behavior, low latency per IO, balance issues with large IO luns/controllers (NOTE: current best for FC5)
 - Anticipatory – elevator=as. Inserts delays to help stack aggregate IO, best on system w/ limited physical IO – SATA
- RHEL4 - Set at boot time on command line
- RHEL5 – Change on the fly
 - `echo deadline > /sys/block/<sdX>/queue/scheduler`

RHEL5 IO schedules vs RHEL3 for Database Oracle 10G oltp/dss (relative performance)



System Performance

■ Memory Performance

- Huge Pages

- 2MB huge pages
- Set value in `/etc/sysctl.conf` (`vm.nr_hugepages`)

- NUMA

- Localized memory access for certain workloads improves performance

- Swap

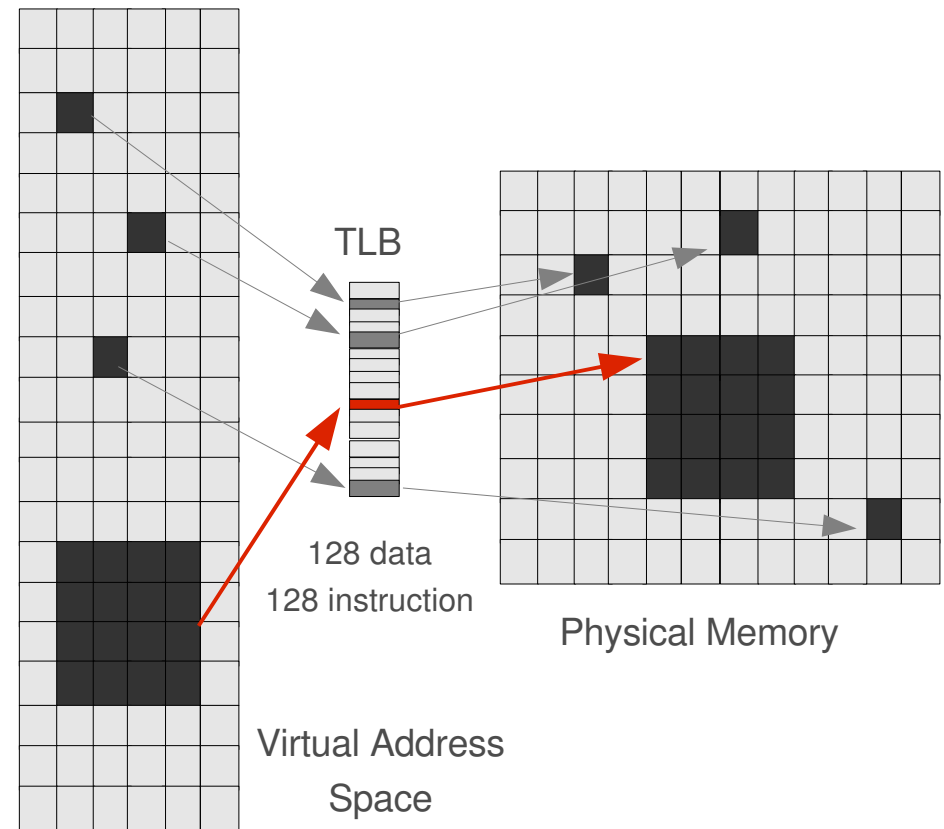
- Set value of `vm.swappiness` (Default 60)

- CPU Performance

- Multiple cores

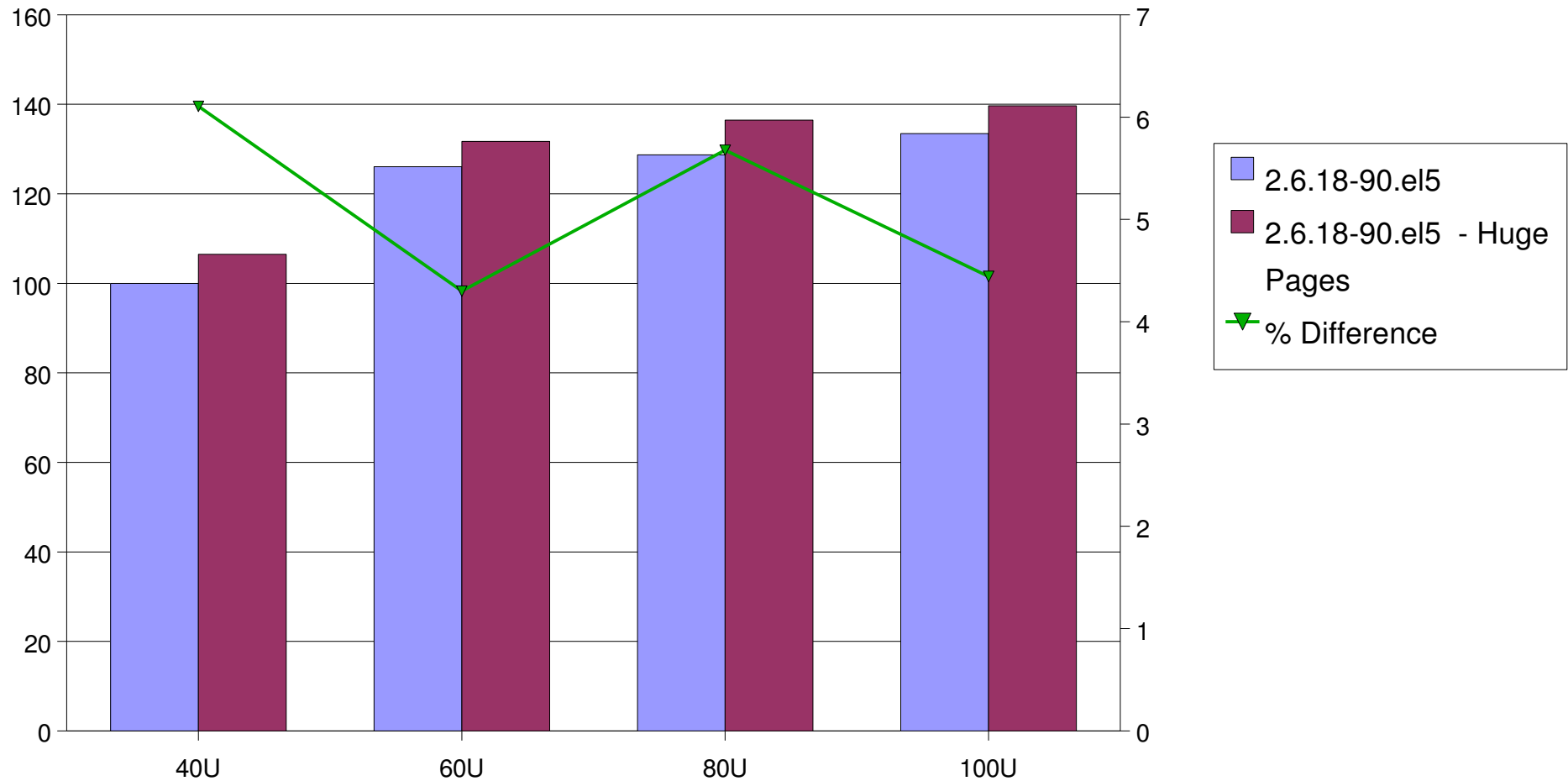
Huge Pages

- The Translation Lookaside Buffer (TLB) is a small CPU cache of recently used virtual to physical address mappings
- TLB misses are extremely expensive on today's very fast, pipelined CPUs
- Large memory applications can incur high TLB miss rates
- Huge pages permit memory to be managed in very large segments
 - Standard page: 4KB
 - Default huge page: 2MB
 - 500:1 difference
- File system mapping interface
- Ideal for databases running OLTP workloads



RHEL5.2 -Oracle 10.2 Huge pages

Relative performance



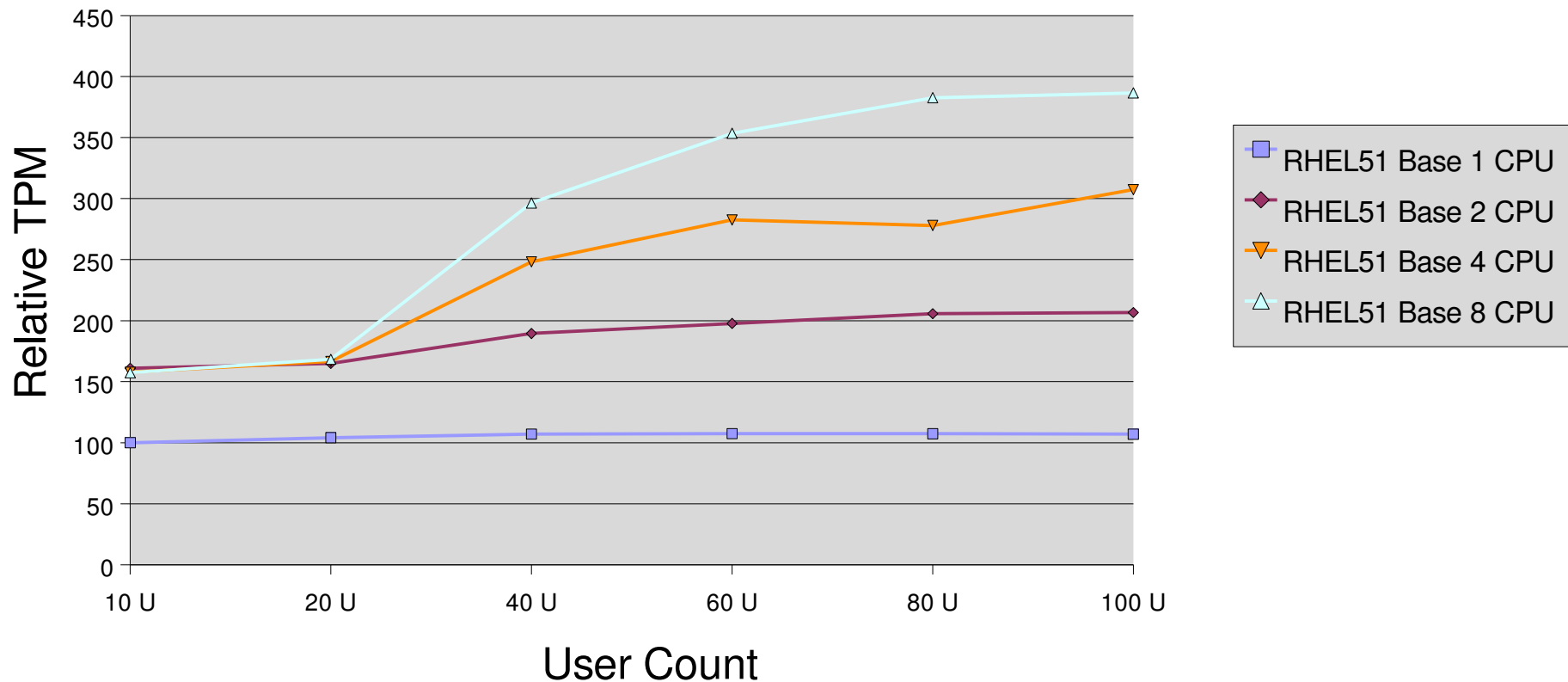
Database Performance

- Database layout
 - Separate logs from data files
 - Separate data files from Index files
 - Separate data file based on access patterns
- Database tuning
 - OLTP / DSS specific settings
- Application tuning
 - Design
 - Avoid exclusive locking / keep it to minimum
 - Take advantage of database features

Database Scaling - Intel

RHEL5.1 Scaling - Oracle 10.2 OLTP

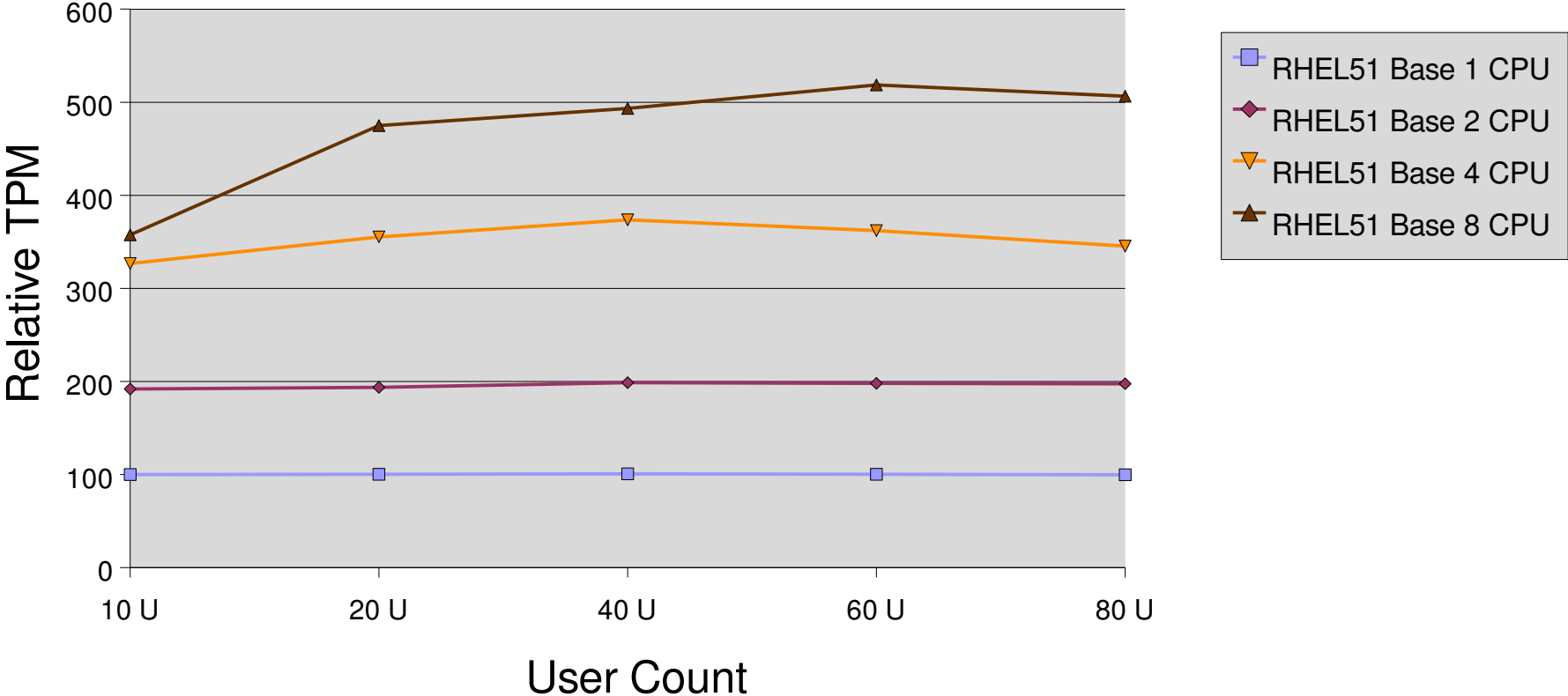
Intel Quad Core 2.33 GHz



Database Scaling - AMD

RHEL5.1 Scaling - Oracle 10.2 OLTP

AMD Dual Socket Quad Core 2.0 GHz



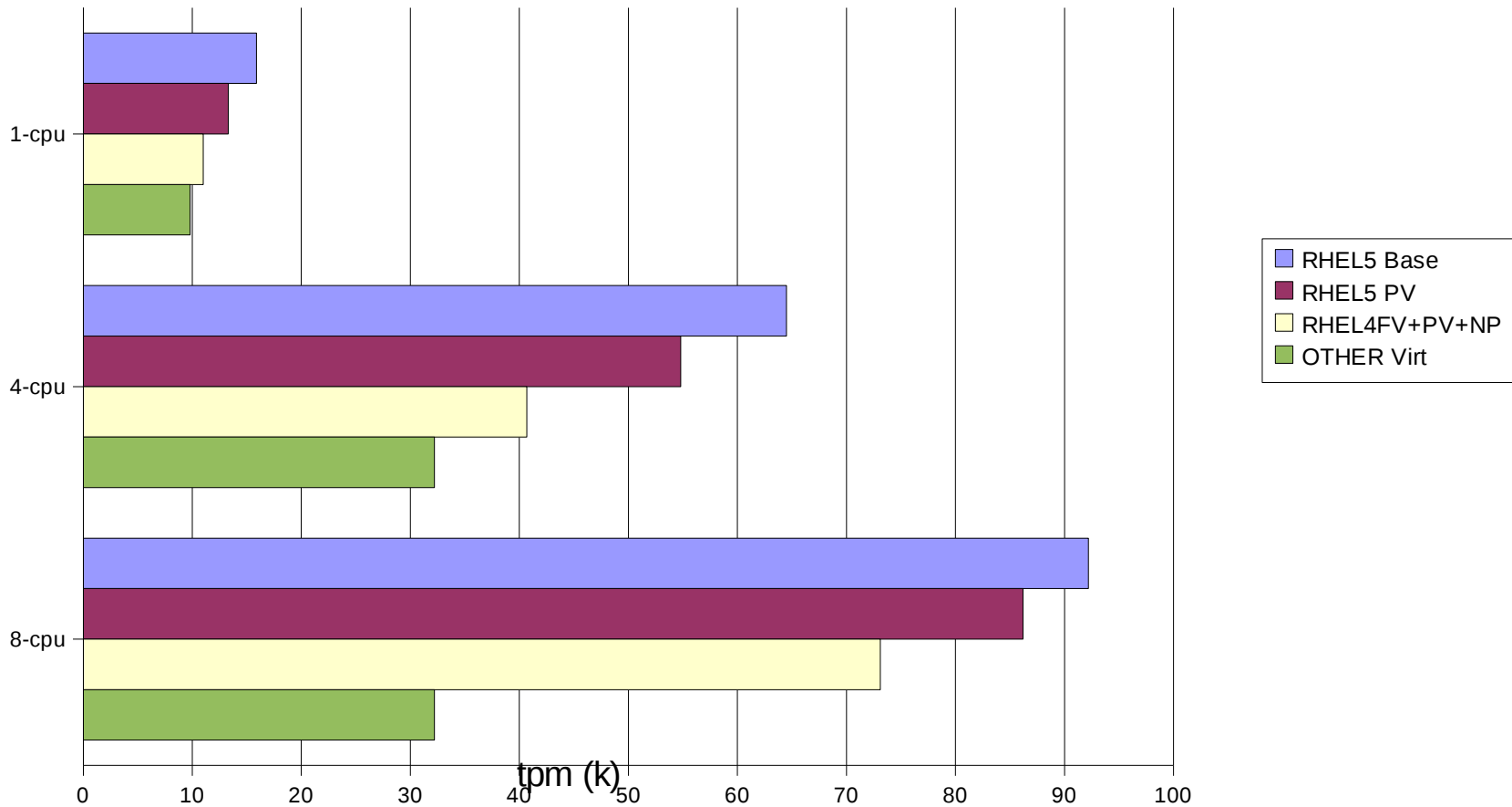
Red Hat Enterprise Linux 5 - Virtualization

*RHEL 5.1 provides the most **scalable** Virtualization platform*

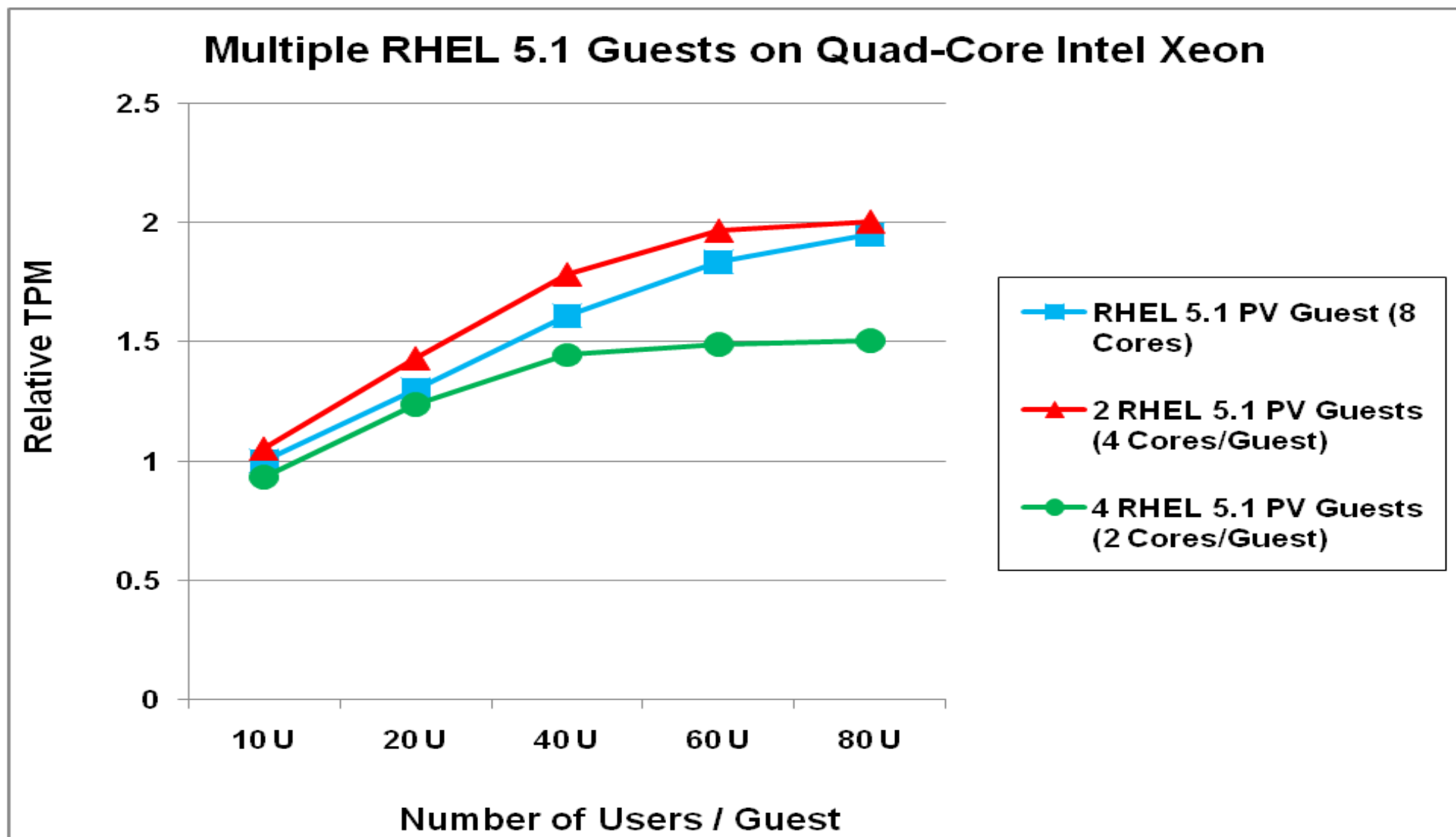
- Scales from single CPU servers to largest multi core machines
- The only virtualization platform that spans the entire enterprise
 - From x86 up to Itanium2
- Scales to meet the needs of enterprise class workloads
 - Supports up to 32 virtual CPUs per guest
 - Use Async I/O and Direct I/O within Guest (tap:aio)
 - Up to 16GB memory per guest
 - Up to 16TB of memory per host
 - *Dynamically allocate resources to respond to business needs*
 - *Add memory or CPU resources on-the-fly*
 - *Scale-up or scale-down resources*
 - *Live Migrate servers to take advantage of unused resources*

Red Hat 5.1 Virtualization Performance

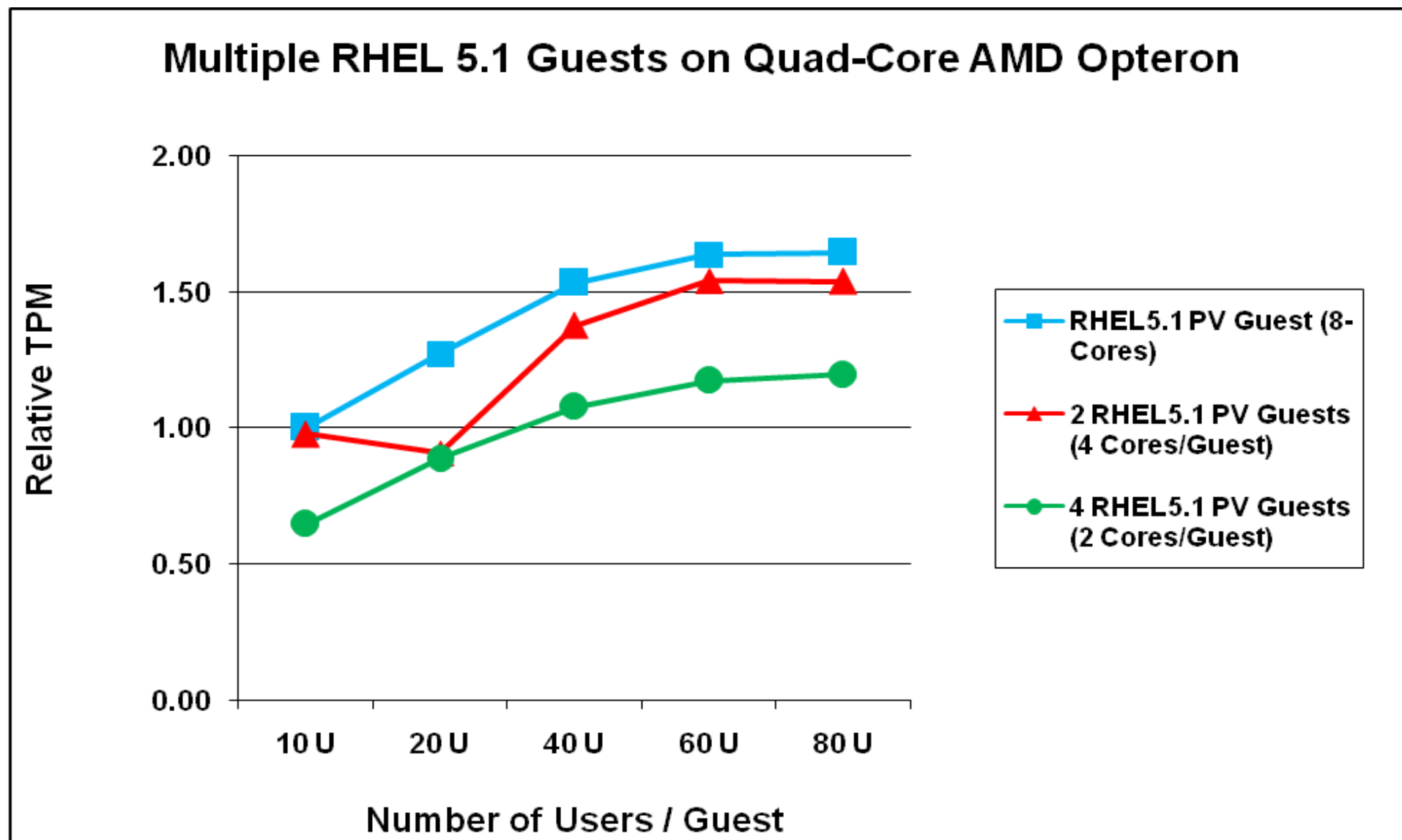
RHEL5.1 Virt Performance of Oracle 10G
vs Other Virt , AMD 1-4-8 cpu, 8Gb mem



Red Hat 5.1 Virt Multi-Guest Scalability

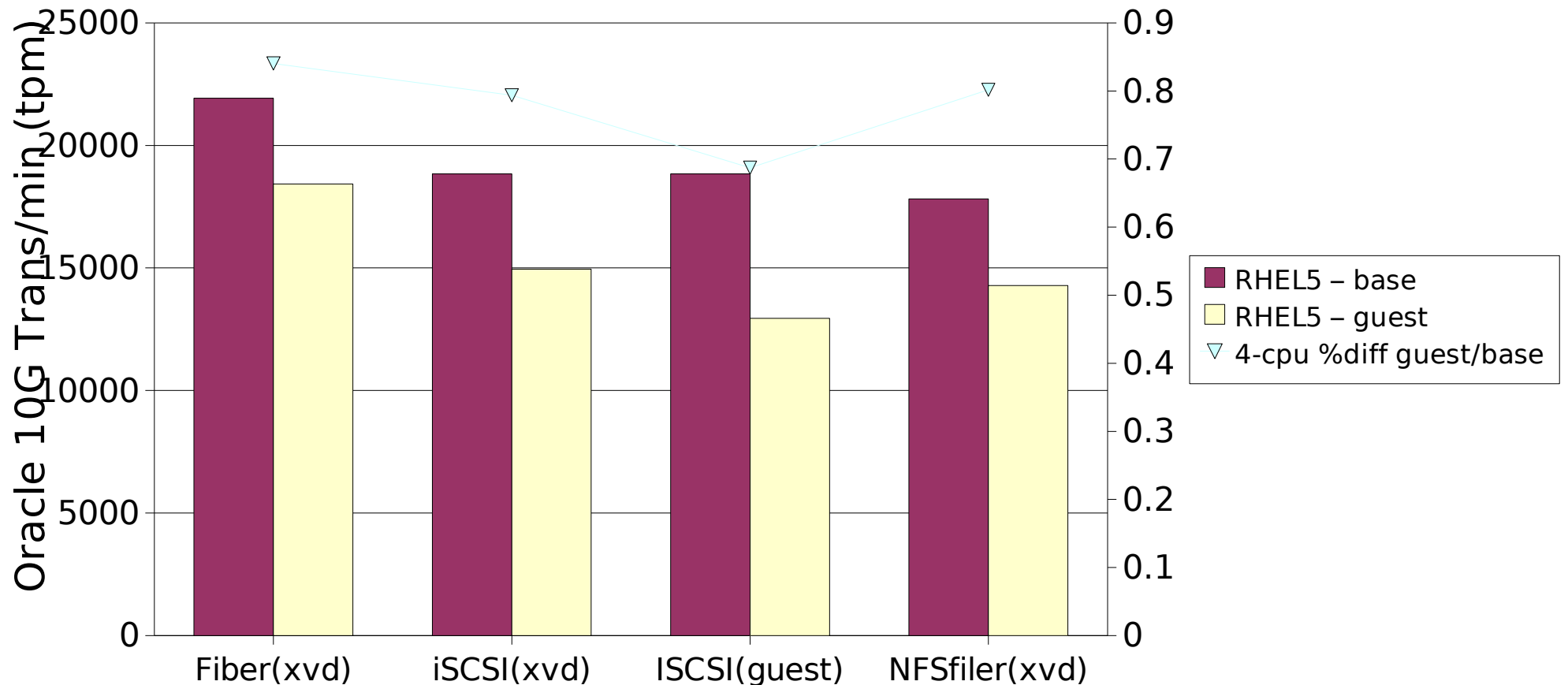


Red Hat 5.1 Virt Multi-Guest Scalability



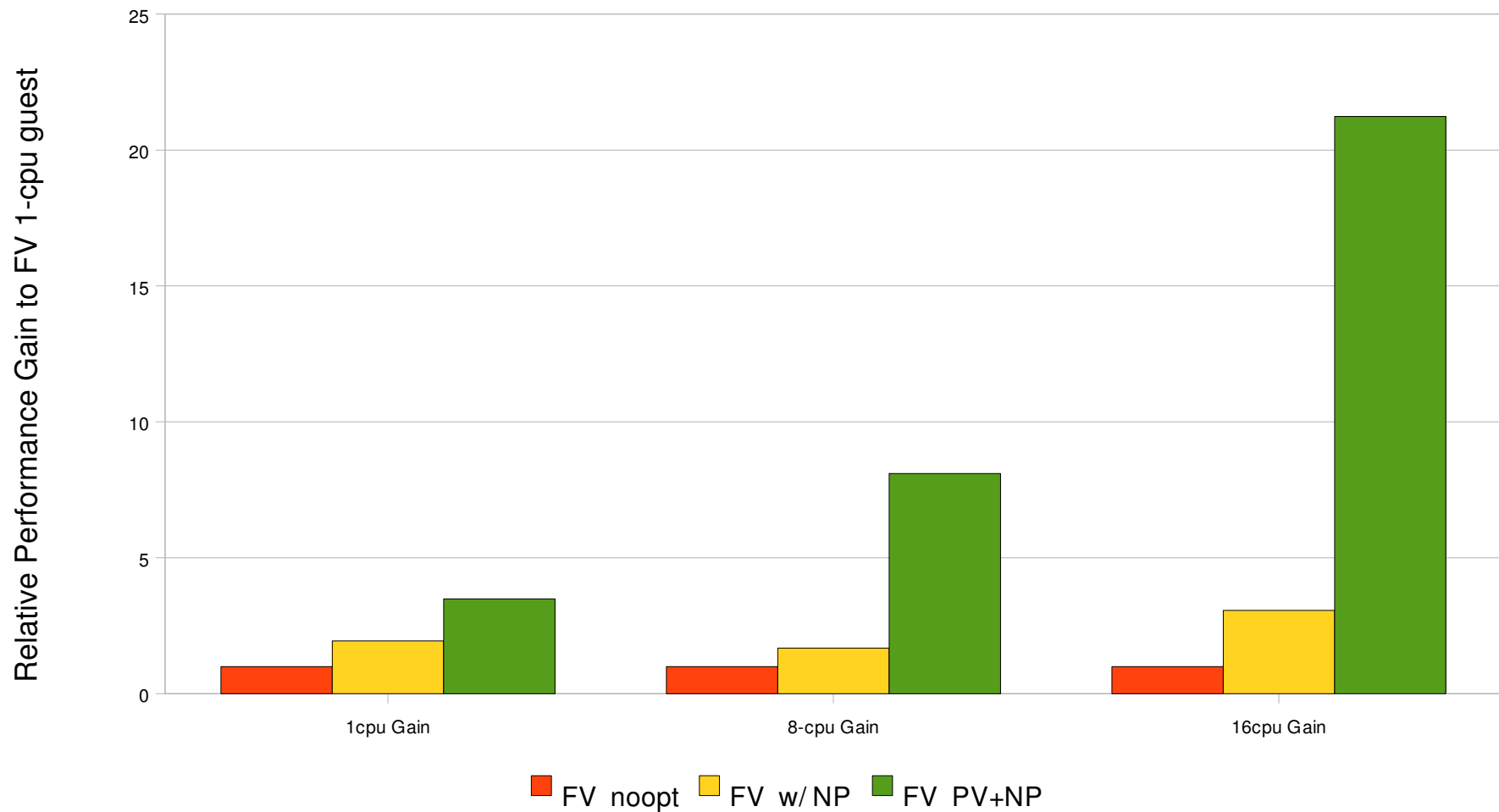
Red Hat Virt Storage Alternatives

RHEL5.1 Virt Application Performance
w/ various Storage



Red Hat 5 – AMD Rapid Virtualization Indexing

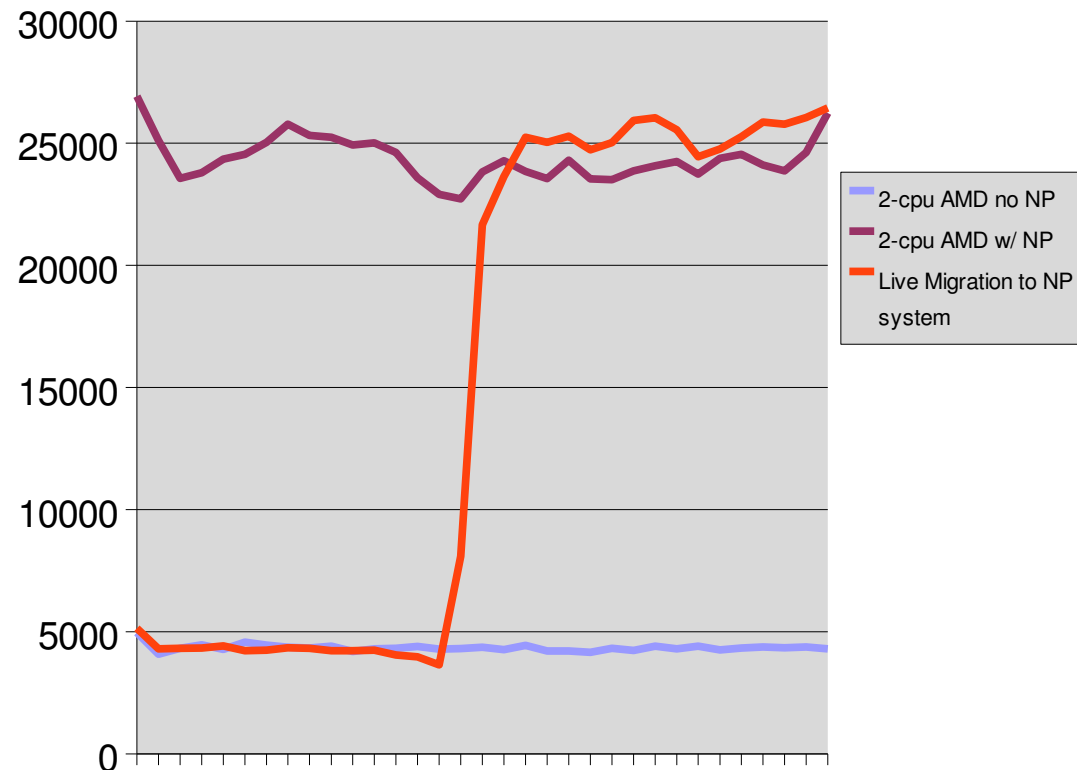
RHEL5.1 Virt HPDL585 16-cpu AMD Database Scaling



RHEL5.1 Virtualization “Live Migration” w/ Oracle 10G R2 on GFS cluster

- RHEL 5.1 Advance Platform
 - Virtualization
 - GFS cluster
 - Shared Fiber Channel SAN
- RHEL4u4 FV guest
 - OLTP running on older AMD sys (not using nested paging)
 - Live migration to newer system (faster w/ nested paging HW)
 - App migrated without shutdown
 - Performance impact immediate
- Migration examples
 - <http://et.redhat.com/~jmh/VirtDemo/>
 - http://people.redhat.com/dshaks/ora_amd_demo.html/

RHEL5.1 Virtualization Database Performance
Transactions/min during Migration from
older AMD to new AMD w/ Nested Paging



Resources

- Red Hat
 - <http://www.redhat.com/>
- Virtualization Infocenter
 - <http://www.openvirtualization.com/>
 - <http://www.redhat.com/promo/virtualization/>
- libvirt
 - <http://www.libvirt.org/>
- Virt-Manager
 - <http://virt-manager.et.redhat.com/>
- Red Hat Cluster Suite
 - <http://www.redhat.com/gfs/>
- View more information
 - <http://www.redhat.com/promo/webinars/>

Where to Look for Performance Results

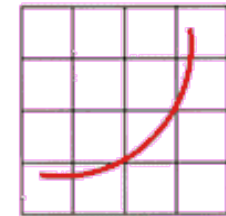
Transaction Processing Performance Council

- www.tpc.org



Standard Performance Evaluation Corporation

- www.spec.org



spec

Oracle Applications Standard Benchmark

- www.oracle.com/benchmarks

SAP Standard Application Benchmarks

- www.sap.com/benchmark



Questions?

