



Red Hat Reference Architecture Series

Scaling Oracle® 11g R2 in a Red Hat® Enterprise Virtualization (RHEV) 3.0 Environment

Brett Thurber, Principal Software Engineer
RHCA, RHCVA

Version 1.0

January 2012





1801 Varsity Drive™
Raleigh NC 27606-2072 USA
Phone: +1 919 754 3700
Phone: 888 733 4281
Fax: +1 919 754 3701
PO Box 13588
Research Triangle Park NC 27709 USA

Linux is a registered trademark of Linus Torvalds. Red Hat, Red Hat Enterprise Linux and the Red Hat "Shadowman" logo are registered trademarks of Red Hat, Inc. in the United States and other countries.

Oracle® is a U.S. registered trademark of Oracle Corporation.

UNIX is a registered trademark of The Open Group.

Intel, the Intel logo and Xeon are registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

All other trademarks referenced herein are the property of their respective owners.

© 2012 by Red Hat, Inc. This material may be distributed only subject to the terms and conditions set forth in the Open Publication License, V1.0 or later (the latest version is presently available at <http://www.opencontent.org/openpub/>).

The information contained herein is subject to change without notice. Red Hat, Inc. shall not be liable for technical or editorial errors or omissions contained herein.

Distribution of modified versions of this document is prohibited without the explicit permission of Red Hat Inc.

Distribution of this work or derivative of this work in any standard (paper) book form for commercial purposes is prohibited unless prior permission is obtained from Red Hat Inc.

The GPG fingerprint of the security@redhat.com key is:
CA 20 86 86 2B D6 9D FC 65 F6 EC C4 21 91 80 CD DB 42 A6 0E

Send feedback to refarch-feedback@redhat.com



Table of Contents

1 Executive Summary.....	1
2 Red Hat Enterprise Virtualization.....	2
2.1 RHEV Hypervisor.....	2
2.2 Red Hat Enterprise Virtualization.....	3
3 Reference Architecture Configuration.....	5
3.1 Environment.....	5
3.2 Software Configuration.....	6
3.2.1 Operating Systems.....	6
3.2.2 Applications, Tools and Packages.....	6
3.3 Hardware Configuration.....	7
3.3.1 Servers.....	7
3.3.2 Storage.....	8
4 Test Methodology.....	9
4.1 Workload.....	9
4.2 Configuration & Workload.....	9
4.3 Performance Test Plan.....	11
4.3.1 Scale-Out of Resources.....	11
4.3.2 Scale-Up of Resources.....	14
4.4 Tuning & Optimizations.....	15
4.4.1 Services.....	15
4.4.2 Pinning.....	16
4.4.3 Storage.....	16
5 Test Results.....	18
5.1 Scaling Multiple 8-vCPU Guests.....	19
5.2 Scaling Multiple 4-vCPU Guests.....	21
5.3 Scaling Multiple 2-vCPU Guests.....	23
5.4 Scaling Multiple 1-vCPU Guests.....	25
5.5 Scaling-Up Resources in a Single Guest.....	27
5.6 Consolidated Virtualization Efficiency.....	29
6 Conclusion.....	30



Appendix A: Revision History.....	31
Appendix B: Contributors.....	31



1 Executive Summary

This paper describes the performance and scaling of Oracle running in Red Hat Enterprise Linux 6 guests on a Red Hat Enterprise Virtualization 3.0 platform. The host was deployed on a Dell PowerEdge R810 server equipped with 128 GB of RAM comprised of dual sockets each with a 2.26GHz Intel® Xeon® 7560 Nehalem-EX processor with support for Hyper-Threading technology, totaling 16 cores and 32 hyper-threads. The workload used was a common Oracle Online Transaction Processing (OLTP) workload.

Scaling Out Virtual Machines

First, a series of tests involve scaling out multiple independent VMs, each comprised of one, two, four, and eight CPUs. The goal is to demonstrate scalability of OLTP workloads across multiple VMs while still maintaining good performance.

Scaling Up A Virtual Machine

Second, the performance of the Oracle OLTP workload was measured by comparing a single VM with increasing resources. Resources include the number of processors, amount of memory, size of the Oracle System Global Area (SGA), and user counts. The goal is to demonstrate OLTP scalability and performance within a single VM by increasing resources.

Bare Metal Comparison

Finally, a comparison of OLTP workloads across both virtual and physical machines shows the efficiency of running multiple workloads, across multiple virtual machines vs. a single physical machine running a single workload. The goal is to demonstrate the benefits of virtualization through performance and the scalability of multiple OLTP workloads.

The data presented in this paper establishes that Red Hat Enterprise Linux 6 virtual machines running within a Red Hat Enterprise Virtualization 3 environment provide an effective, production-ready, platform for hosting multiple virtualized Oracle OLTP workloads. The combination of the ability to both scale-up and scale-out contribute to the effectiveness of Red Hat Enterprise Virtualization for Oracle. The number of actual users and throughput supported in any specific customer situation will, of course, depend on the specifics of the customer application used and the intensity of user activity. However, the results demonstrate that in a heavily virtualized environment, good throughput was retained even as the number and size of guests/virtual machines was increased until the physical server and storage throughput were fully subscribed.

For this reference architecture, scaled resources include CPU, memory, Oracle SGA size, and number of Oracle users. A finite storage configuration was selected for use and it's associated resources were not scaled out as to better reflect a real world enterprise configuration.



2 Red Hat Enterprise Virtualization

2.1 RHEV Hypervisor

A hypervisor is a computer software platform that allows multiple “guest” operating systems to run concurrently on a host computer. The guest virtual machines interact with the hypervisor which translates guest I/O and memory requests into corresponding requests for resources on the host computer.

Running fully virtualized guests, i.e., guests with unmodified guest operating systems, used to require complex hypervisors and previously incurred a performance penalty for emulation and translation of I/O and memory requests.

Over the last few years chip vendors Intel and AMD have been steadily adding CPU features that offer hardware enhancements to support virtualization. Most notable are:

1. First-generation hardware assisted virtualization: Removes the requirement for hypervisor to scan and rewrite privileged kernel instructions using Intel VT (Virtualization Technology) and AMD's SVM (Secure Virtual Machine) technology.
2. Second-generation hardware assisted virtualization: Offloads virtual to physical memory address translation to CPU/chip-set using Intel EPT (Extended Page Tables) and AMD RVI (Rapid Virtualization Indexing) technology. This provides significant reduction in memory address translation overhead in virtualized environments.
3. Third-generation hardware assisted virtualization: Allows PCI I/O devices to be attached directly to virtual machines using Intel VT-d (Virtualization Technology for directed I/O) and AMD IOMMU along with SR-IOV (Single Root I/O Virtualization) which allows special PCI devices to be split into multiple virtual devices. This provides significant improvement in guest I/O performance.

The great interest in virtualization has led to the creation of several different hypervisors. However, many of these pre-date hardware-assisted virtualization, and are therefore somewhat complex pieces of software. With the advent of the above hardware extensions, writing a hypervisor has become significantly easier and it is now possible to enjoy the benefits of virtualization while leveraging existing open source achievements to date.

Red Hat Enterprise Virtualization uses the Kernel-based Virtual Machine (KVM),¹ which turns Linux into a hypervisor. Red Hat Enterprise Linux (RHEL) 5.4 provided the first commercial-strength implementation of KVM, which is developed as part of the upstream Linux community. RHEV 3.0 uses the RHEL 6 KVM hypervisor and inherits performance, scalability and hardware support enhancements from RHEL 6.



2.2 Red Hat Enterprise Virtualization

Virtualization offers tremendous benefits for enterprise IT organizations – server consolidation, hardware abstraction, and internal clouds deliver a high degree of operational efficiency.

Red Hat Enterprise Virtualization (RHEV) combines the KVM hypervisor (powered by the Red Hat Enterprise Linux kernel) with an enterprise grade, multi-hypervisor management platform that provides key virtualization features such as live migration, high availability, power management, and virtual machine life cycle management. Red Hat Enterprise Virtualization delivers a secure, robust virtualization platform with unmatched performance and scalability for Red Hat Enterprise Linux and Windows guests. Red Hat Enterprise Virtualization consists of the following two components:

- **RHEV Manager (RHEV-M):** A feature-rich virtualization management system that provides advanced capabilities for hosts and guests.
- **RHEV Hypervisor:** A modern, scalable, high performance hypervisor based on RHEL KVM. It can be deployed as RHEV-H, a small footprint secure hypervisor image included with the RHEV subscription, or as a RHEL server (purchased separately) managed by RHEV-M.

A **host** is a physical server which provides the CPU, memory, and connectivity to storage and networks that are used for the virtual machines (VM). The local storage of the standalone host is used for the RHEV-H executables along with logs and enough space for ISO uploads.

A **cluster** is a group of hosts of similar architecture. The requirement of similar architecture allows a virtual machine to be migrated from host to host in the cluster without having to shut down and restart the virtual machine. A cluster consists of one or more hosts, but a host can only be a member of one cluster.

A **data center** is a collection of one or more clusters that have resources in common. Resources that have been allocated to a data center can be used only by the hosts belonging to that data center. The resources relate to storage and networks.

A **storage domain** is a shared or local storage location for virtual machine image files, import/export or for ISO images. Storage domain types supported in RHEV 3.0 are NFS, iSCSI, Fibre Channel, and local disk storage.

The RHEV **network** architecture supports both virtual machine traffic as-well-as traffic among RHEV hypervisors and the RHEV-M server. All hosts have a network interface assigned to the logical network named *rhevm*. This network is used for the communications between the hypervisor and the manager. Additional logical networks are created on the data center and applied to one or more clusters. To become operational, the host attaches an interface to the local network. While the actual physical network can span across data centers, the logical network can only be used by the clusters and hosts of the creating data center.



Figure 2.2-1: Red Hat Enterprise Virtualization provides a graphical representation of a typical Red Hat Enterprise Virtualization environment with each component listed.

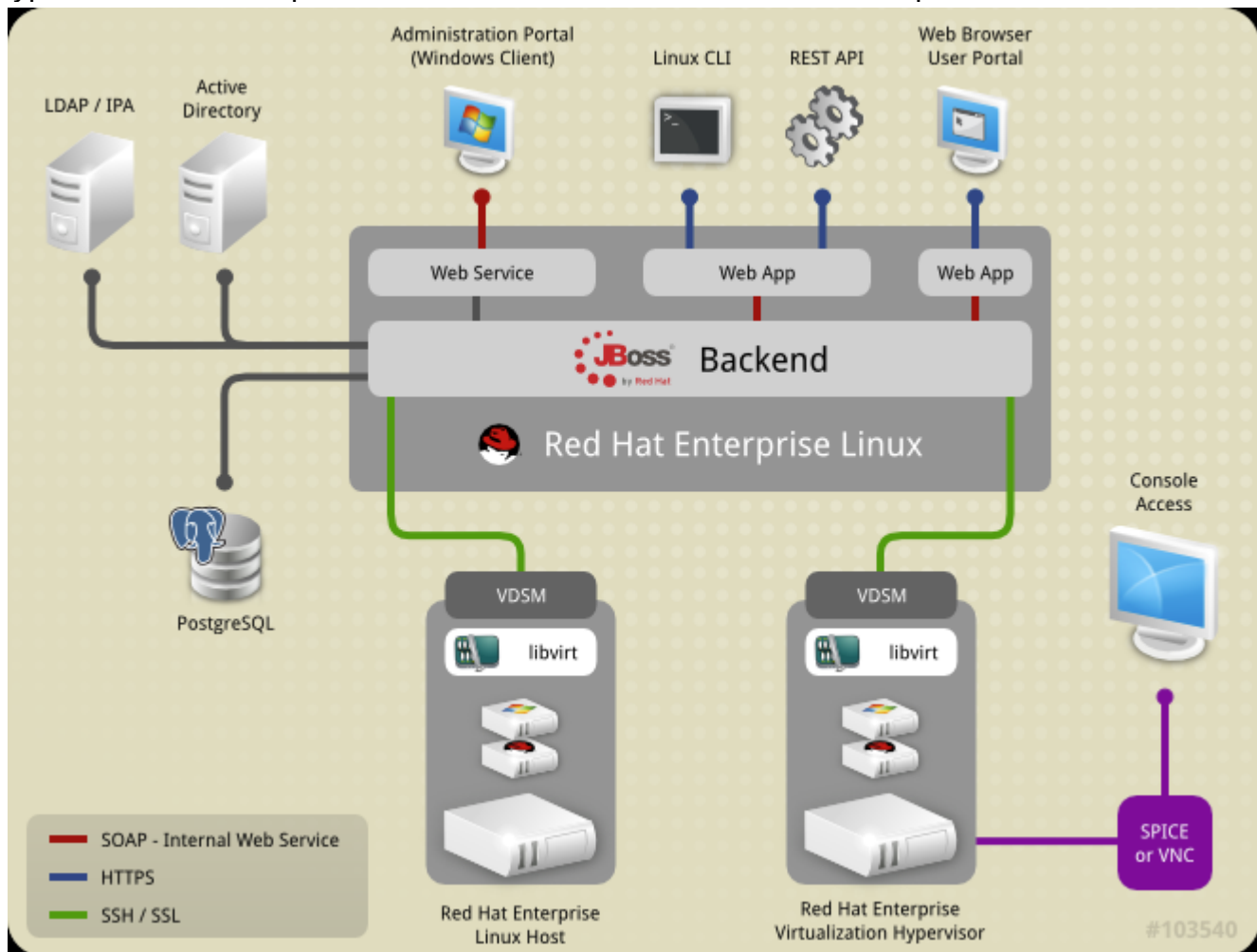


Figure 2.2-1: Red Hat Enterprise Virtualization



3 Reference Architecture Configuration

3.1 Environment

The following section details the reference architecture configuration used in this guide as depicted in **Figure 3.1-1: RHEV Scaling Environment**.

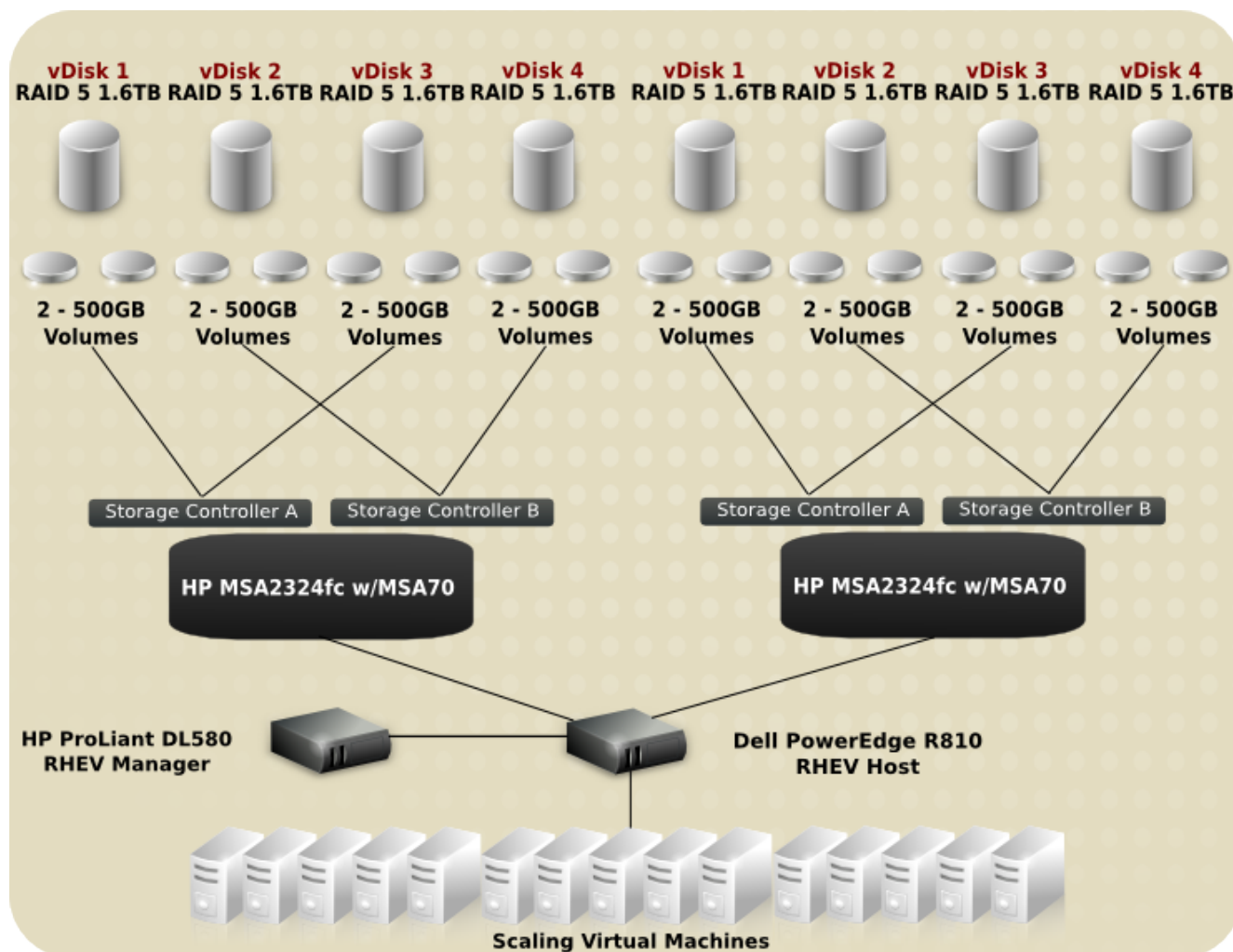


Figure 3.1-1: RHEV Scaling Environment



3.2 Software Configuration

3.2.1 Operating Systems

Operating systems with revisions used as referenced in **Table 3.2.1-1: Operating System Revisions**.

Software	Role	Version
Red Hat Enterprise Linux (RHEL)	RHEV Manager	6.1 2.6.32-131.0.15.el6
Red Hat Enterprise Linux w/KVM	RHEV Host	6.2 2.6.32-202.el6 vdsm-4.9-106.el6
Red Hat Enterprise Linux (RHEL)	Guest	6.1 2.6.32-131.0.15.el6

Table 3.2.1-1: Operating System Revisions

3.2.2 Applications, Tools and Packages

Applications, tools and package revisions used as referenced in **Table 3.2.2-1: Applications, Tools and Package Revisions**.

Software	Version
Red Hat Enterprise Virtualization Manager (RHEV-M)	3.0.0_0001-46.el6
Oracle	11.2.0.1.0

Table 3.2.2-1: Applications, Tools and Package Revisions



3.3 Hardware Configuration

3.3.1 Servers

Server hardware with configuration specifics used as referenced in **Table 3.3.1-1: Server Hardware**.

Hardware	Specifications
RHEV Host [1 x Dell PowerEdge R810]	Dual Socket, 8 Core, 16 total cores: Intel® Xeon® X7560 @2.26GHz, 128GB RAM
	4 x Hard Drive, 146G, SAS, 15K (RAID 5)
	2 x Qlogic HBA Fibre Channel Adapter QLE2562
	2 x Broadcom 1GbE BASE-T MC Server Adapter
	1 x Broadcom 10GbE Dual Port SFP+ Adapter
RHEV Manager [1 x HP ProLiant DL580 G5]	Quad Socket, 4 Core, 16 total cores: Intel® Xeon® X7350@2.93GHz, 64GB RAM
	4 x Hard Drive, 73G, SAS, 10K (RAID 5)
	2 x Broadcom Corporation NetXtreme II BCM5708 Gigabit Ethernet
	1 x Intel Corporation 82572EI Gigabit Ethernet Controller

Table 3.3.1-1: Server Hardware



3.3.2 Storage

Table 3.3.2-1: Storage Hardware displays the storage hardware used in this reference environment with firmware revision information.

Hardware	Specifications
2 x HP StorageWorks MSA2324fc Fibre Channel Storage Array + HP StorageWorks 70 Modular Smart Array with Dual Domain IO Module [24+25 x 146GB 10K RPM SAS disks]	Storage Controllers: Code Version: M112R14 Loader Code Version: 19.009
	Memory Controller: Code Version: F300R22
	Management Controller Code Version: W441R39 Loader Code Version: 12.015
	Expander Controller: Code Version: 1112
	CPLD Code Version: 8
	Hardware Version: 56
	Expansion Module: 2.28
1 x HP StorageWorks 8/24 SAN Switch	Firmware: v6.4.0a

Table 3.3.2-1: Storage Hardware

Table 3.3.2-2: Storage LUNs displays the LUN configuration and mappings for each host in the reference environment.

Volume(s)	Size	Count	Host Presentation	Purpose
vm1-vm16	500 GB (ea.)	16	rhev3-beta2	VM storage domain(s)

Table 3.3.2-2: Storage LUNs



4 Test Methodology

4.1 Workload

An Oracle OLTP workload was chosen as it represents a common database implementation exercising both the memory and I/O sub-systems of the virtual machines. Tests were performed on each guest configuration in 15 minute exercises on a 60 warehouse database.

4.2 Configuration & Workload

The RHEV host is configured with dual Intel X7560 processors, each being a 2.26 GHz eight-core processor supporting Hyper-Threading technology. **Table 4.2-1: Host BIOS Settings** displays the BIOS settings configured on the host to include CPU and Non-Uniform Memory Access performance enhancements.

Parameter	Setting	Function
Logical Processor	Disabled	Hyper-threading
Turbo Mode	Enabled	Allows for overclocking
C1E	Disabled	CPU Power Management
C States	Disabled	CPU Power Management
Node Interleaving	Disabled	NUMA behavior

Table 4.2-1: Host BIOS Settings



Each guest is configured with a vCPU that maps to each core available on the host for a total of 16 cores. For example, a guest sized for two vCPUs would be configured to use two sockets within RHEV-M as shown in **Figure 4.2-1: Guest CPU Configuration**.

Edit Server Virtual Machine

General

Console

Host

High Availability

Resource Allocation

Boot Options

Custom Properties

Data Center: Scale-Oracle

Host Cluster: Scale-Oracle

Name: VM2-Oracle

Description:

Based on Template: Blank

Memory Size: 15 GB

Total Cores: 2 (slider from 2 to 32)

CPU Sockets: 2 (slider from 1 to 16)

Operating System: Red Hat Enterprise Linux 6.x x64

OK Cancel

Figure 4.2-1: Guest CPU Configuration

Demonstrating the scaling of RHEV means several aspects of the workload (user count, SGA size) and guest configuration (vCPU count, memory) were scaled accordingly with the size of the guest. The database was held constant to demonstrate that results were the effect of scaling the guests and not the application. However, per guest factors such as the amount of system memory, the size of the Oracle SGA, and the number of Oracle users were increased with each vCPU. To that extent, an Oracle load of 20 users with a 6GB SGA was allocated per vCPU in each guest. For example, a 4-vCPU guest executed the OLTP workload with 30GB of system memory and 80 Oracle clients using a 24GB SGA.

The host system possessed a total 128GB of memory. Even distribution of this memory among the vCPUs would allow for 8GB per vCPU. However, 7.5GB was allocated to each vCPU in order to leave memory for the hypervisor as well as guests that may have oversubscribed the processing power of the hypervisor.



Table 4.2-2: Guest/Workload Configurations lists the totals used for each guest configuration.

VCPUs per Guest	Guest Memory	Oracle Users	Oracle SGA
1	7.5 GB	20	6 GB
2	15 GB	40	12 GB
4	30 GB	80	24 GB
8	60 GB	160	48 GB

Table 4.2-2: Guest/Workload Configurations

4.3 Performance Test Plan

Scale-out:

The scale-out data set highlights the results of scaling a number of concurrent 1-vCPU, 2-vCPU, 4-vCPU, or 8-vCPU guests executing the OLTP workload.

Scale-up:

The scale-up data set was collected by increasing the number of vCPUs, guest memory, Oracle SGA size and number of users running the workload on a single guest.

Virtualization Efficiency:

Efficiency is shown by comparing the data when all the physical CPUs are allocated to executing the workload using the bare metal host (no virtualization), sixteen 1-vCPU guests, eight 2-vCPU guests, four 4-vCPU guests, and two 8-vCPU guests.

4.3.1 Scale-Out of Resources

RHEV provides capabilities that allow system administrators to easily deploy and configure virtual machines through the use of templates². By using this feature, it becomes a simple process to scale out resources on an as-needed basis. Template deployment was used as the method to deploy a standard configuration used for testing. Each guest was customized to define identity and resource allocation in order to include memory assigned, number of vCPUs defined, custom properties for CPU and NUMA pinning, and database parameters used for testing each configuration.



Table 4.3.1-1: 1-vCPU, 16-Guest Configuration displays the storage, CPU, and memory configuration for each guest in the sixteen virtual machine test case.

Guest	CPU Pin	NUMA Zone Pin	Memory per Guest	RHEV Storage Domain	LUN Configuration	Storage Enclosure
VM1-Oracle	0	0	7,680 MB	VM1-Oracle	Single 12 disk, RAID 5, 500GB	1
VM2-Oracle	1	1	7,680 MB	VM2-Oracle	Single 12 disk, RAID 5, 500GB	1
VM3-Oracle	2	0	7,680 MB	VM3-Oracle	Single 12 disk, RAID 5, 500GB	1
VM4-Oracle	3	1	7,680 MB	VM4-Oracle	Single 12 disk, RAID 5, 500GB	1
VM5-Oracle	4	0	7,680 MB	VM5-Oracle	Single 12 disk, RAID 5, 500GB	1
VM6-Oracle	5	1	7,680 MB	VM6-Oracle	Single 12 disk, RAID 5, 500GB	1
VM7-Oracle	6	0	7,680 MB	VM7-Oracle	Single 12 disk, RAID 5, 500GB	1
VM8-Oracle	7	1	7,680 MB	VM8-Oracle	Single 12 disk, RAID 5, 500GB	1
VM9-Oracle	8	0	7,680 MB	VM9-Oracle	Single 12 disk, RAID 5, 500GB	2
VM10-Oracle	9	1	7,680 MB	VM10-Oracle	Single 12 disk, RAID 5, 500GB	2
VM11-Oracle	10	0	7,680 MB	VM11-Oracle	Single 12 disk, RAID 5, 500GB	2
VM12-Oracle	11	1	7,680 MB	VM12-Oracle	Single 12 disk, RAID 5, 500GB	2
VM13-Oracle	12	0	7,680 MB	VM13-Oracle	Single 12 disk, RAID 5, 500GB	2
VM14-Oracle	13	1	7,680 MB	VM14-Oracle	Single 12 disk, RAID 5, 500GB	2
VM15-Oracle	14	0	7,680 MB	VM15-Oracle	Single 12 disk, RAID 5, 500GB	2
VM16-Oracle	15	1	7,680 MB	VM16-Oracle	Single 12 disk, RAID 5, 500GB	2

Table 4.3.1-1: 1-vCPU, 16-Guest Configuration



Table 4.3.1-2: 2-vCPU, 8-Guest Configuration displays the storage, CPU, and memory configuration for each guest in the eight virtual machine test case.

Guest	CPU Pin	NUMA Zone Pin	Memory per Guest	RHEV Storage Domain	LUN Configuration	Storage Enclosure
VM1-Oracle	0,2	0	15,360 MB	VM1-Oracle	Two 12 disk, RAID 5, 1TB	1 and 2
VM2-Oracle	5,7	1	15,360 MB	VM2-Oracle	Two 12 disk, RAID 5, 1TB	1 and 2
VM3-Oracle	8,10	0	15,360 MB	VM3-Oracle	Two 12 disk, RAID 5, 1TB	1 and 2
VM4-Oracle	13,15	1	15,360 MB	VM4-Oracle	Two 12 disk, RAID 5, 1TB	1 and 2
VM5-Oracle	1,3	1	15,360 MB	VM5-Oracle	Two 12 disk, RAID 5, 1TB	1 and 2
VM6-Oracle	4,6	0	15,360 MB	VM6-Oracle	Two 12 disk, RAID 5, 1TB	1 and 2
VM7-Oracle	9,11	1	15,360 MB	VM7-Oracle	Two 12 disk, RAID 5, 1TB	1 and 2
VM8-Oracle	12,14	0	15,360 MB	VM8-Oracle	Two 12 disk, RAID 5, 1TB	1 and 2

Table 4.3.1-2: 2-vCPU, 8-Guest Configuration

Table 4.3.1-3: 4-vCPU, 4-Guest Configuration displays the storage, CPU, and memory configuration for each guest in the four virtual machine test case.

Guest	CPU Pin	NUMA Zone Pin	Memory per Guest	RHEV Storage Domain	LUN Configuration	Storage Enclosure
VM1-Oracle	0,2,4,6	0	30,720 MB	VM1-Oracle	Four 12 disk, RAID 5, 2TB	1 and 2
VM2-Oracle	1,3,5,7	1	30,720 MB	VM2-Oracle	Four 12 disk, RAID 5, 2TB	1 and 2
VM3-Oracle	8,10,12,14	0	30,720 MB	VM3-Oracle	Four 12 disk, RAID 5, 2TB	1 and 2
VM4-Oracle	9,11,13,15	1	30,720 MB	VM4-Oracle	Four 12 disk, RAID 5, 2TB	1 and 2

Table 4.3.1-3: 4-vCPU, 4-Guest Configuration



Table 4.3.1-4: 8-vCPU, 2 Guest Configuration displays the storage, CPU, and memory configuration for each guest in the two virtual machine test case.

Guest	CPU Pin	NUMA Zone Pin	Memory per Guest	RHEV Storage Domain	LUN Configuration	Storage Enclosure
VM1-Oracle	0,2,4,6,8,10,12,14	0	61,440 MB	VM1-Oracle	Eight 12 disk, RAID 5, 4TB	1 and 2
VM2-Oracle	1,3,5,7,9,11,13,15	1	61,440 MB	VM2-Oracle	Eight 12 disk, RAID 5, 4TB	1 and 2

Table 4.3.1-4: 8-vCPU, 2 Guest Configuration

4.3.2 Scale-Up of Resources

Through the use of the RHEV-M Portal, a system administrator has the ability to adjust guest properties easily to meet a specific workload requirement without having to spend time on-site re-configuring hardware. vCPUs, memory and database resources were increased with each OLTP run utilizing a single virtual machine.

Table 4.3.2-1: Scale-Up Configuration displays the configuration parameters defined for single VM resource scaling.

Guest	vCPUs	Memory	SGA	Users
VM1-Oracle	1	7,680 MB	6,144 MB	20
VM1-Oracle	2	15,360 MB	12,288 MB	40
VM1-Oracle	4	30,720 MB	24,576 MB	80
VM1-Oracle	8	61,440 MB	49,152 MB	160

Table 4.3.2-1: Scale-Up Configuration

NOTE – SGA is set to 80% of memory assigned to the guest across all test configurations.



4.4 Tuning & Optimizations

The following items were implemented to optimize performance and scalability within the guest and host operating systems to include implementation of new features within RHEV 3.0.

4.4.1 Services

Several processes deemed unnecessary for testing were disabled, using the `chkconfig` command, on each guest.

Services		
abrt-ccpp	cups	messagebus
abrt-oops	haldaemon	portreserve
abrt-d	irqbalance	postfix
acpid	iscsi	qpidd
atd	iscsid	rhnsd
auditd	kdump	rhsmcertd
autofs	ksm	rpcbind
avahi-daemon	ksmtuned	rpcgssd
bluetooth	libvirt-guests	rpcidmapd
cgconfig	libvirtd	spice-vdagentd
cpuspeed	mcelogd	virt-who
crond	mdmonitor	

Table 4.4.1-1: Disabled Services

Security Enhanced Linux (SELinux) is disabled across the guests and host.



4.4.2 Pinning

Each guest is pinned for CPU core(s) and NUMA zones on the host. This is accomplished via the guest *Custom Properties* value within the RHEV-M console as depicted in **Figure 4.4.2-1: Guest CPU and NUMA Pinning**.



Figure 4.4.2-1: Guest CPU and NUMA Pinning

NOTE: Custom pinning is performed using *hooks*³ within RHEV 3.0. In order to utilize *hooks*, the host must be installed using Red Hat Enterprise Linux 6.2 with KVM.

4.4.3 Storage

Within the RHEV Environment, each guest is configured to a dedicated storage domain. This provides the ability to explicitly assign a VM to a defined volume or storage location ensuring that only the defined VM has access to it. Other options include a single storage domain for all VMs for shared storage access or direct LUN mapping through the use of *hooks*. Direct LUN mappings provide the benefits of highly customized storage layouts per VM for configurations designed to meet a specific need such as dedicated disks for database operations or to meet a particular performance goal.



Figure 4.4.3-1: RHEV Storage Domain Configuration depicts an example RHEV Storage Domain configuration used for testing a sixteen VM, single vCPU configuration.

Data Centers		Clusters	Hosts	Storage	Virtual Machines	Pools	Templates
New Domain		Import Domain	Edit	Remove			
	Domain Name	Domain Type	Storage Type	Format	Cross Dat		
▲	Template-Oracle	Data (Master)	Fibre Channel	V2	Active		
▲	VM10-Oracle	Data	Fibre Channel	V2	Active		
▲	VM11-Oracle	Data	Fibre Channel	V2	Active		
▲	VM12-Oracle	Data	Fibre Channel	V2	Active		
▲	VM13-Oracle	Data	Fibre Channel	V2	Active		
▲	VM14-Oracle	Data	Fibre Channel	V2	Active		
▲	VM15-Oracle	Data	Fibre Channel	V2	Active		
▲	VM16-Oracle	Data	Fibre Channel	V2	Active		
▲	VM1-DB2	Data	Fibre Channel	V2	Active		
▲	VM1-Oracle	Data	Fibre Channel	V2	Active		
▲	VM2-Oracle	Data	Fibre Channel	V2	Active		
▲	VM3-Oracle	Data	Fibre Channel	V2	Active		
▲	VM4-Oracle	Data	Fibre Channel	V2	Active		
▲	VM5-Oracle	Data	Fibre Channel	V2	Active		
▲	VM6-Oracle	Data	Fibre Channel	V2	Active		
▲	VM7-Oracle	Data	Fibre Channel	V2	Active		
▲	VM8-Oracle	Data	Fibre Channel	V2	Active		
▲	VM9-Oracle	Data	Fibre Channel	V2	Active		

Figure 4.4.3-1: RHEV Storage Domain Configuration



5 Test Results

Multiple factors can affect scaling. Among those are hardware characteristics, application characteristics and virtualization overhead.

Hardware:

The most prominent hardware characteristics relevant to the tests in this guide include limited storage throughput and system architecture. The disk I/O requirements of a single database instance may not be extreme but this quickly compounds as multiple systems are executed in parallel against a limited I/O bandwidth on the hypervisor. The system architecture includes Hyper-Threading technology which provides a boost in performance beyond 16 cores. However, the performance of the two threads on any hyper threaded core is not expected to be equal that of two non-hyper threaded cores as Linux treats each processing thread as a separate CPU so therefore hyper-threading was disabled throughout the testing. The system architecture also includes NUMA, which allows faster access to nearby memory, albeit slower access to remote memory. This architecture has two NUMA nodes, one for each processor. Restricting a process or a guest to a specific NUMA node allows cache sharing and memory access performance boosts.

Application:

The specific scaling, up (increased amounts of resources) or out (multiple instances of similar sized systems), can effect various applications in different ways. The added memory and CPU power of scaling up will typically help applications that do not contend with a limited resource, where scaling out may provide a multiple of the limited resource. Conversely, scaling out may not be suited for applications requiring a high degree of coordination for the application, which could occur in memory for a scale-up configuration. Additionally, virtualization can be used to consolidate multiple independent homogenous or heterogeneous workloads onto a single server.

Virtualization:

As it is not entirely running directly on physical hardware and requires the hypervisor layer (which consumes processing cycles), some performance cost is associated with any virtualized environment. The amount of overhead can vary depending on the efficiency of the hypervisor and the assorted drivers used.



5.1 Scaling Multiple 8-vCPU Guests

This section presents the results obtained when running multiple 8-vCPU guests (each running an independent Oracle OLTP workload) on a four-socket, eight-core Dell PowerEdge R810 host. For the following tests, two sockets, eight cores were utilized for a total of 16 cores.

Figure 5.1-1: Scaling Multiple 8-vCPU Guests is an illustration depicting the scale out of workload and resources as multiple 8-vCPU guests are added.

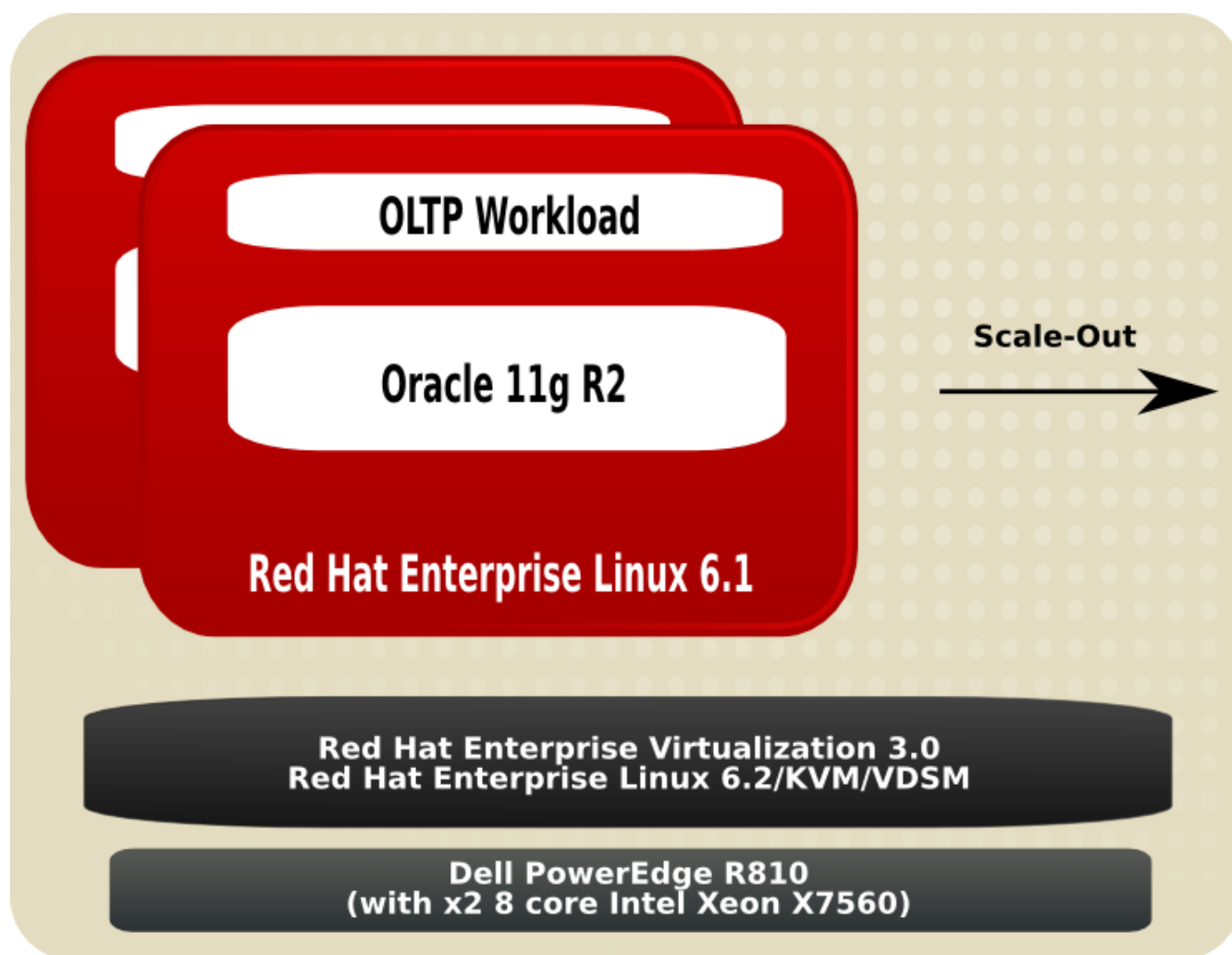


Figure 5.1-1: Scaling Multiple 8-vCPU Guests



Figure 5.1-2: Results of Scaling Multiple 8-vCPU Guests graphs the scalability achieved by increasing the number of 8-vCPU RHEL guests running independent OLTP workloads. The throughput demonstrates excellent (near-linear) scaling. As guests are added, the throughput per guest decreases slightly due to I/O contention and virtualization overhead.

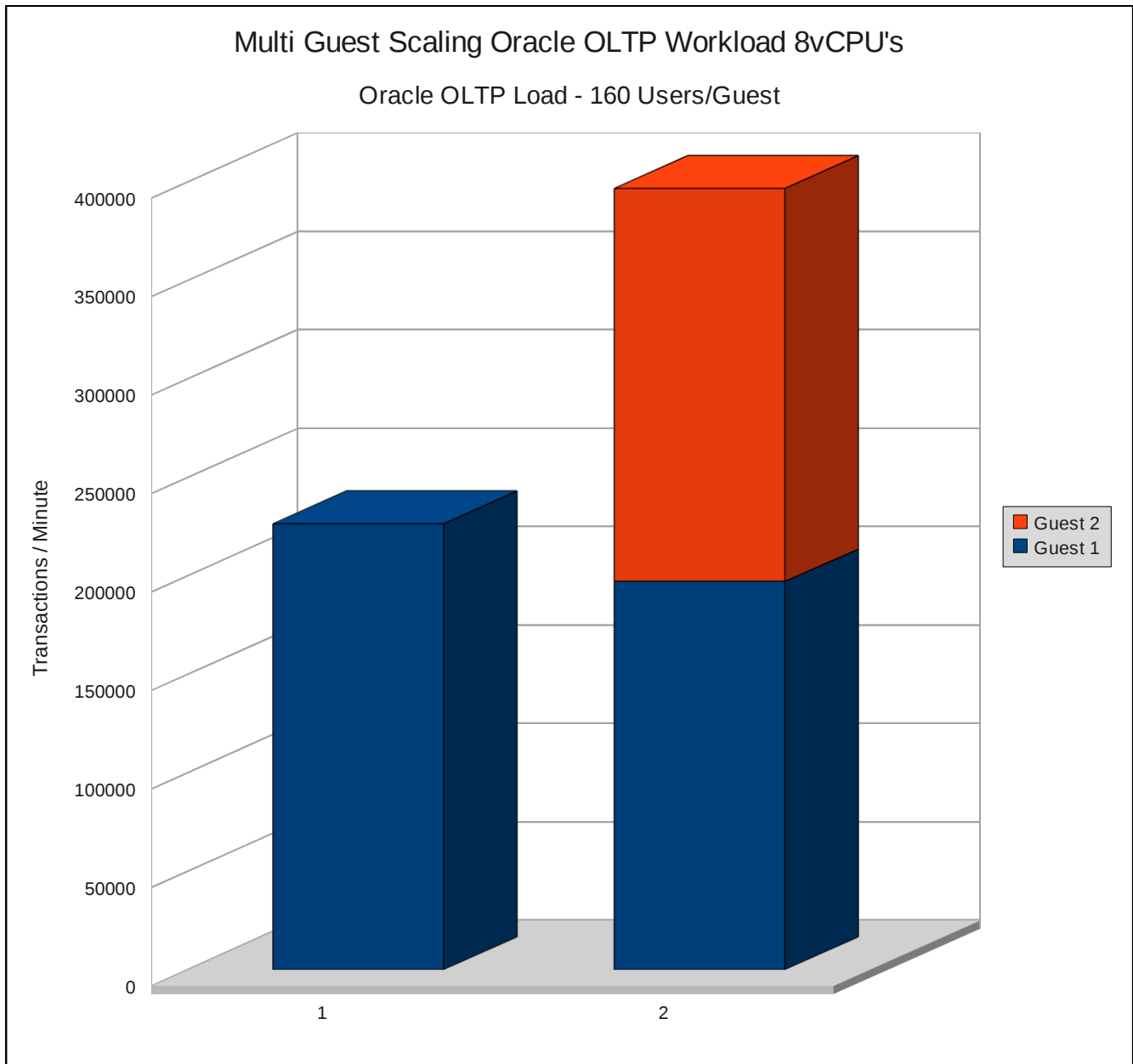


Figure 5.1-2: Results of Scaling Multiple 8-vCPU Guests



5.2 Scaling Multiple 4-vCPU Guests

This section presents the results obtained when running multiple 4-vCPU guests (each running an independent Oracle OLTP workload) on a four-socket, eight-core, Dell PowerEdge R810 host. For the following tests, two sockets, eight cores were utilized for a total of 16 cores.

Figure 5.2-1: Scaling Multiple 4-vCPU Guests is an illustration depicting the scale out of workload and resources as multiple 4-vCPU guests are added.

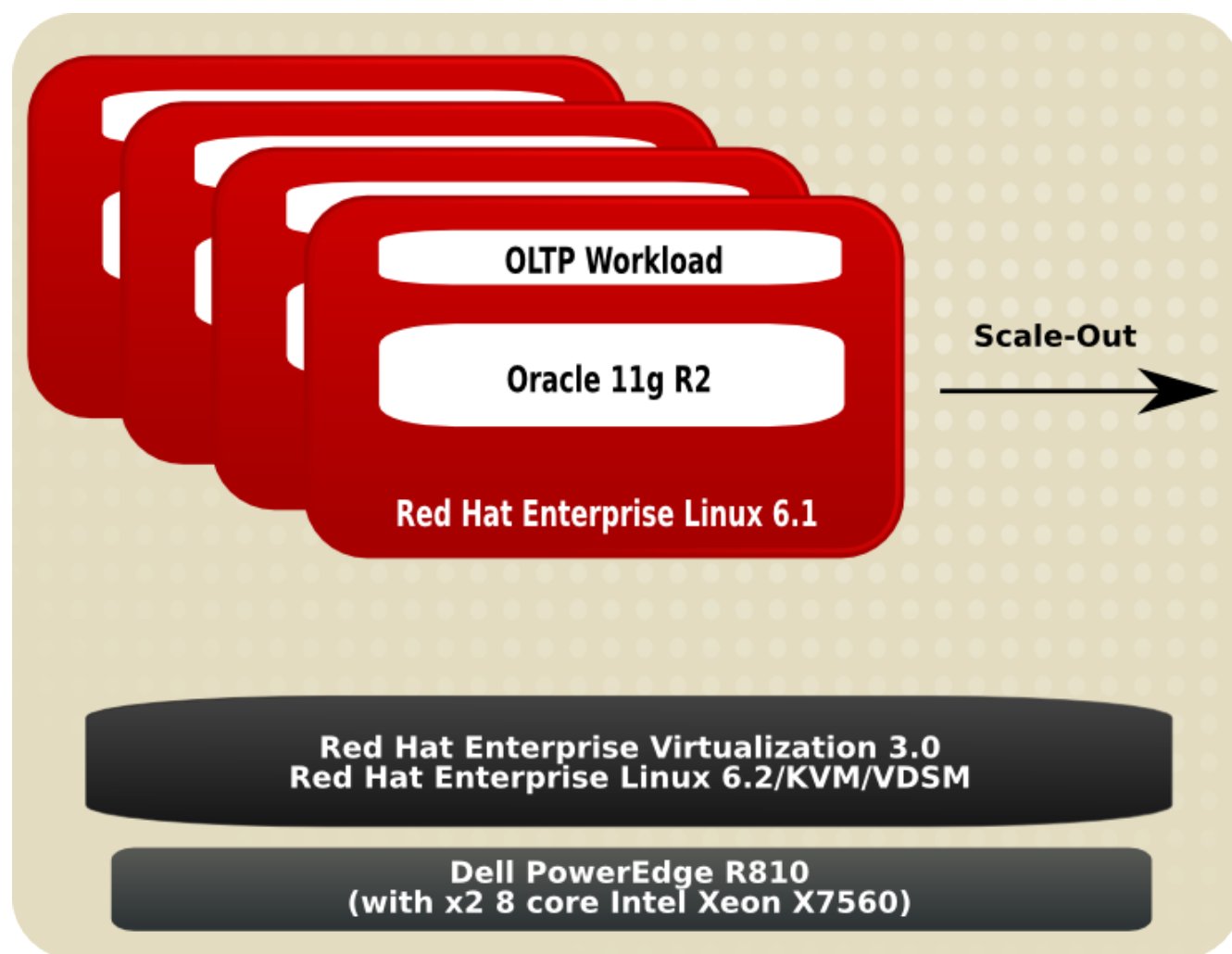


Figure 5.2-1: Scaling Multiple 4-vCPU Guests



Figure 5.2-2: Results of Scaling Multiple 4-vCPU Guests displays the scalability achieved by increasing the number of 4-vCPU Red Hat Enterprise Linux 6 guests from one to four, running independent OLTP workloads. The throughput demonstrates good scaling. As guests are added, the throughput per guest decreases slightly due to I/O contention and virtualization overhead.

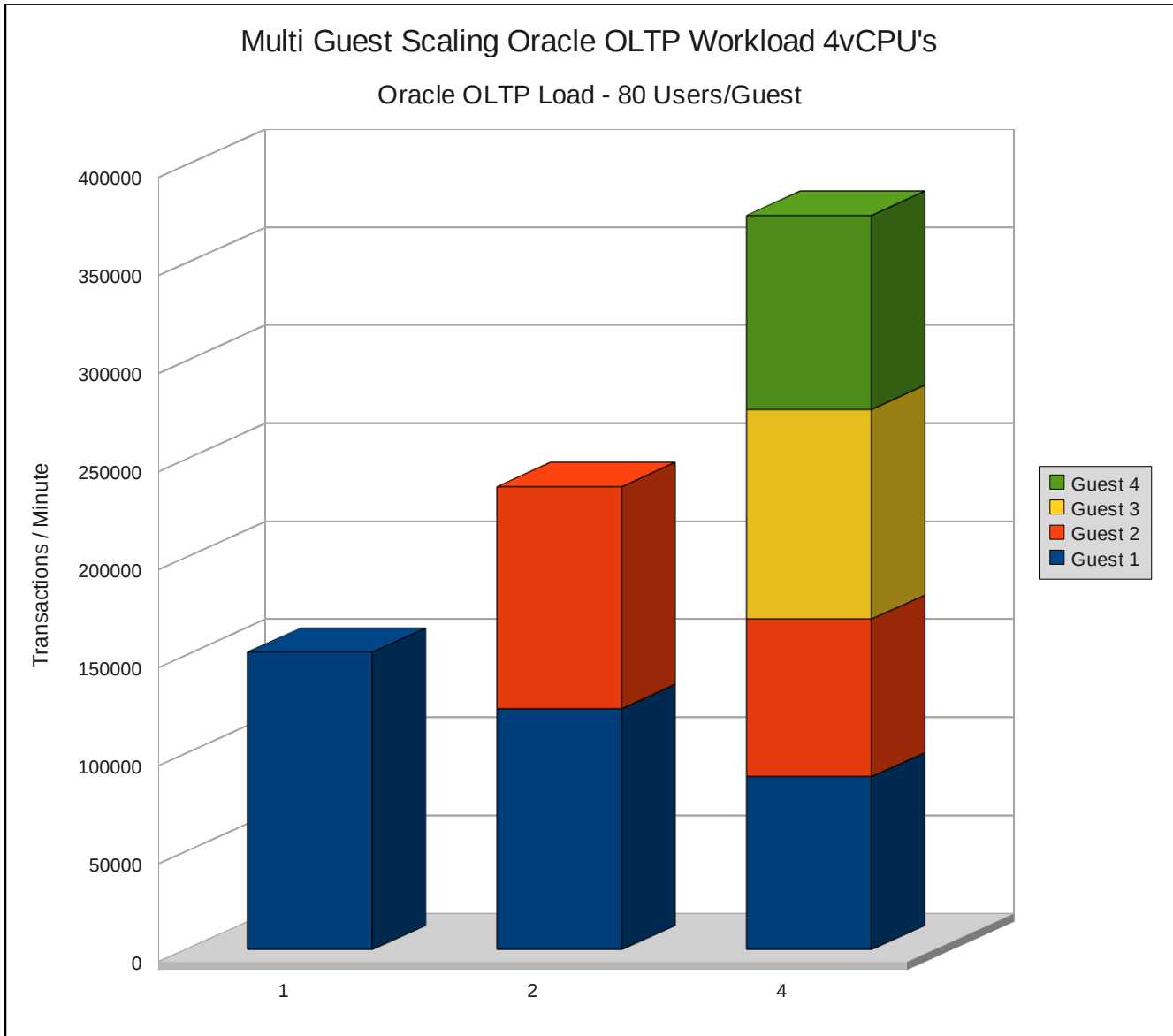


Figure 5.2-2: Results of Scaling Multiple 4-vCPU Guests



5.3 Scaling Multiple 2-vCPU Guests

This section presents the results obtained when running multiple 2-vCPU guests (each running an independent Oracle OLTP workload) on a four-socket, eight-core, Dell PowerEdge R810 host. For the following tests, two sockets, eight cores were utilized for a total of 16 cores.

Figure 5.3-1: Scaling Multiple 2-vCPU Guests is an illustration depicting the scale out of workload and resources as multiple 2-vCPU guests are added.

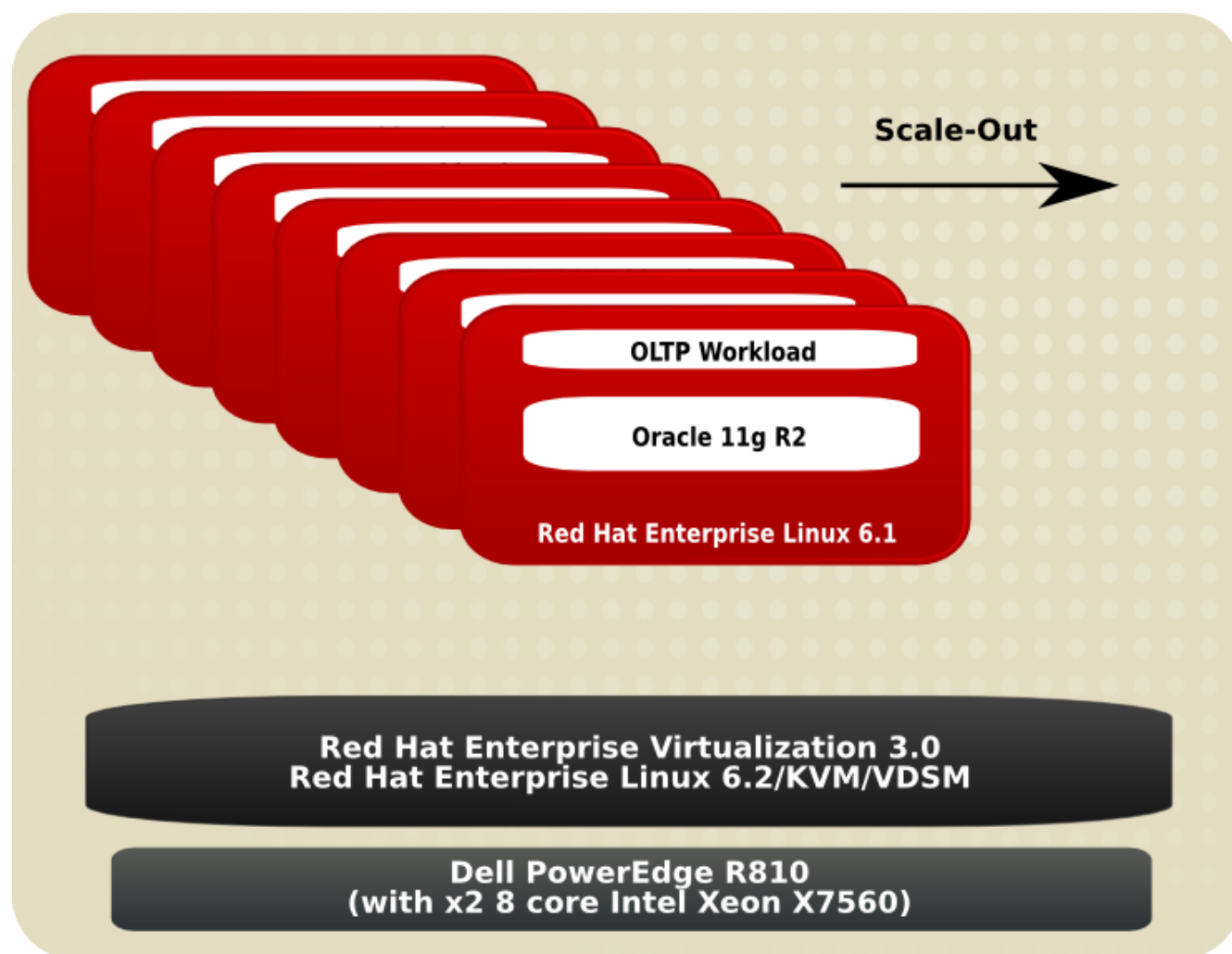


Figure 5.3-1: Scaling Multiple 2-vCPU Guests



Figure 5.3-2: Results of Scaling Multiple 2-vCPU Guests displays the scalability achieved by increasing the number of 2-vCPU Red Hat Enterprise Linux 6 guests from one to eight, running independent OLTP workloads. The throughput demonstrates good scaling. As guests are added, the throughput per guest decreases slightly due to I/O contention and virtualization overhead.

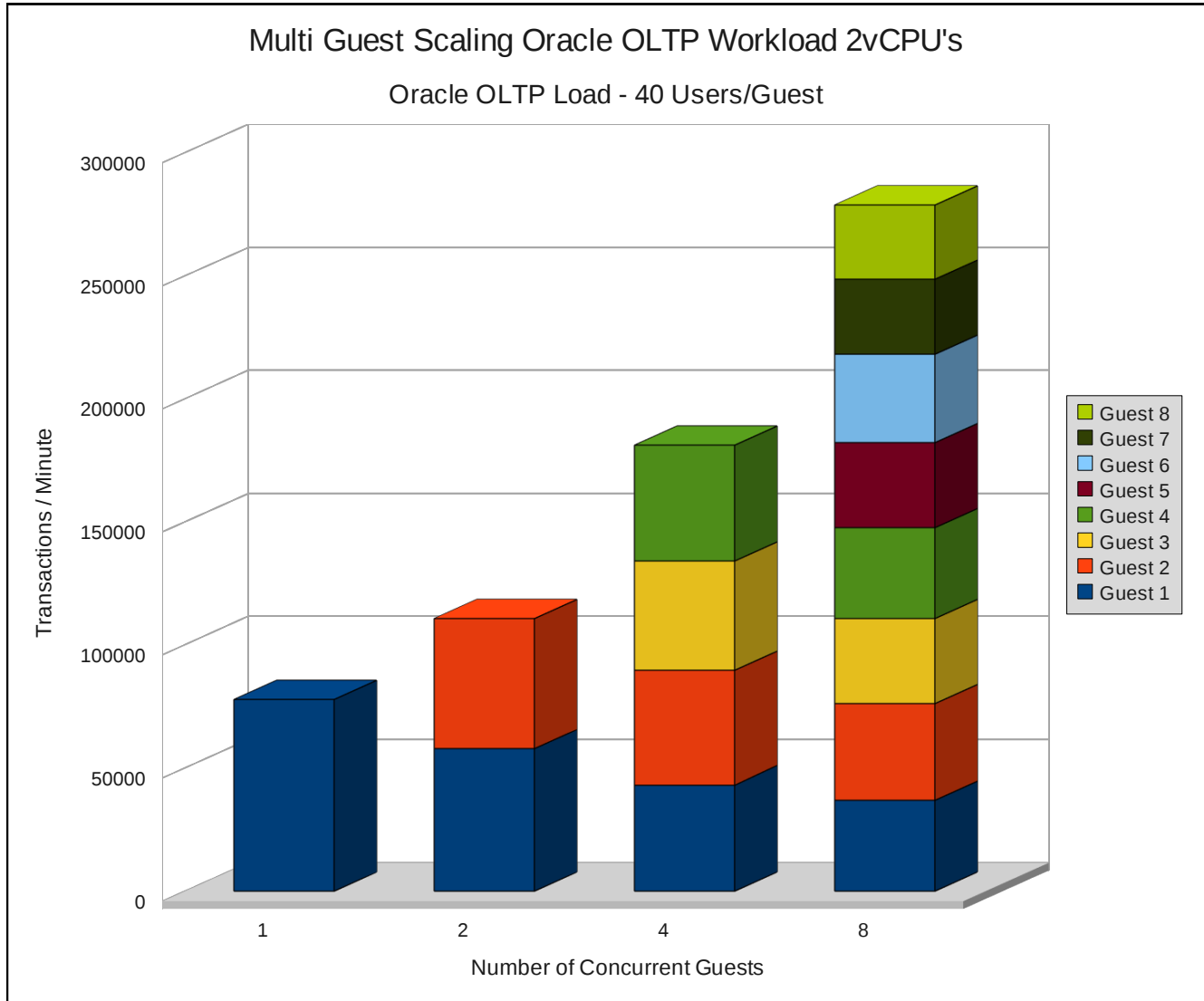


Figure 5.3-2: Results of Scaling Multiple 2-vCPU Guests



5.4 Scaling Multiple 1-vCPU Guests

This section presents the results obtained when running multiple 1-vCPU guests (each running an independent Oracle OLTP workload) on a four-socket, eight-core, Dell PowerEdge R810 host. For the following tests, two sockets, eight cores were utilized for a total of 16 cores.

Figure 5.4-1: Scaling Multiple 1-vCPU Guests is an illustration depicting the scale out of workload and resources as multiple 1-vCPU guests are added.

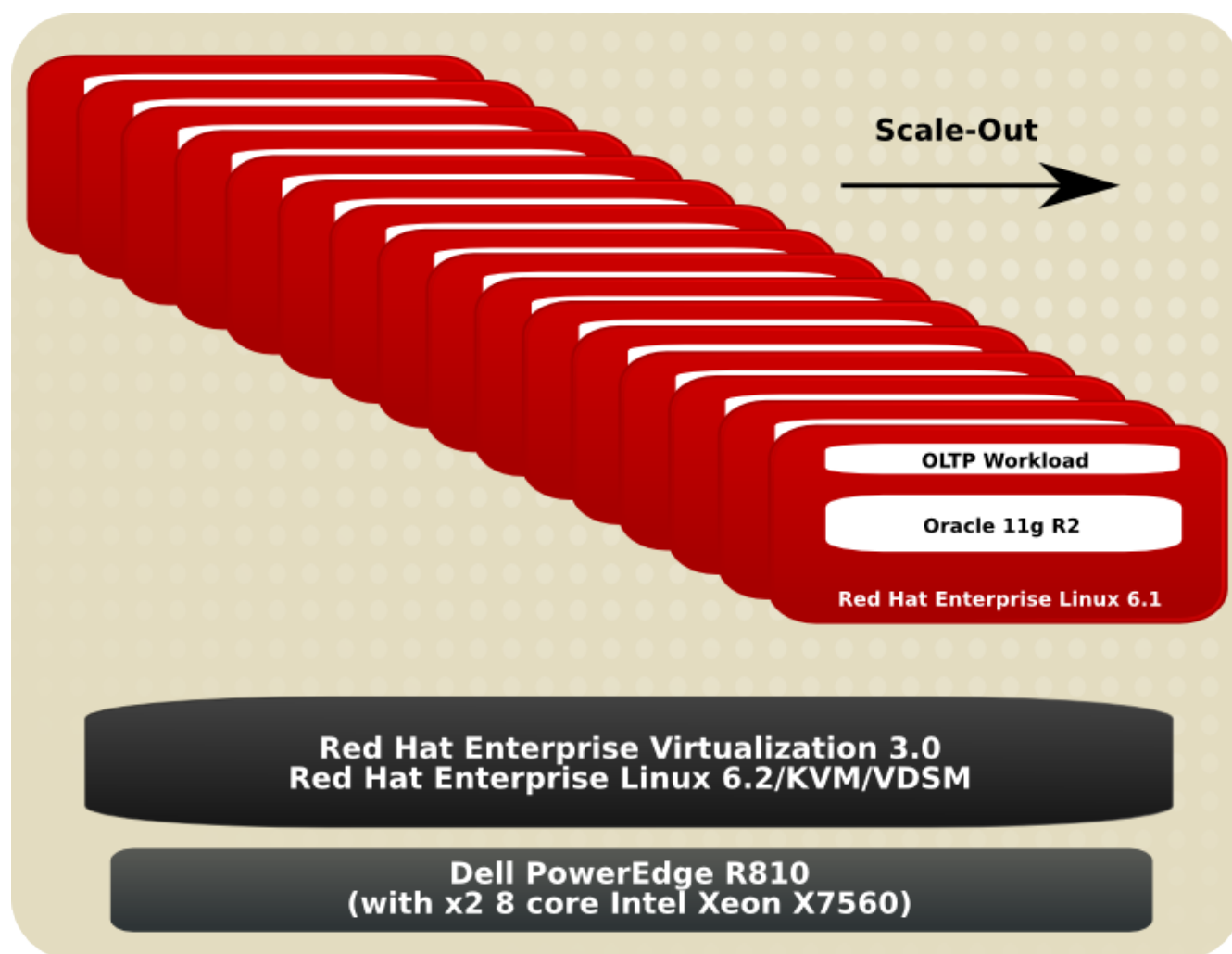


Figure 5.4-1: Scaling Multiple 1-vCPU Guests



Figure 5.4-2: Results of Scaling Multiple 1-vCPU Guests displays the scalability achieved by increasing the number of 1-vCPU Red Hat Enterprise Linux 6 guests from one to sixteen, each running independent OLTP workloads. The throughput demonstrates good scaling. As guests are added, the throughput per guest decreases due to I/O contention and virtualization overhead.

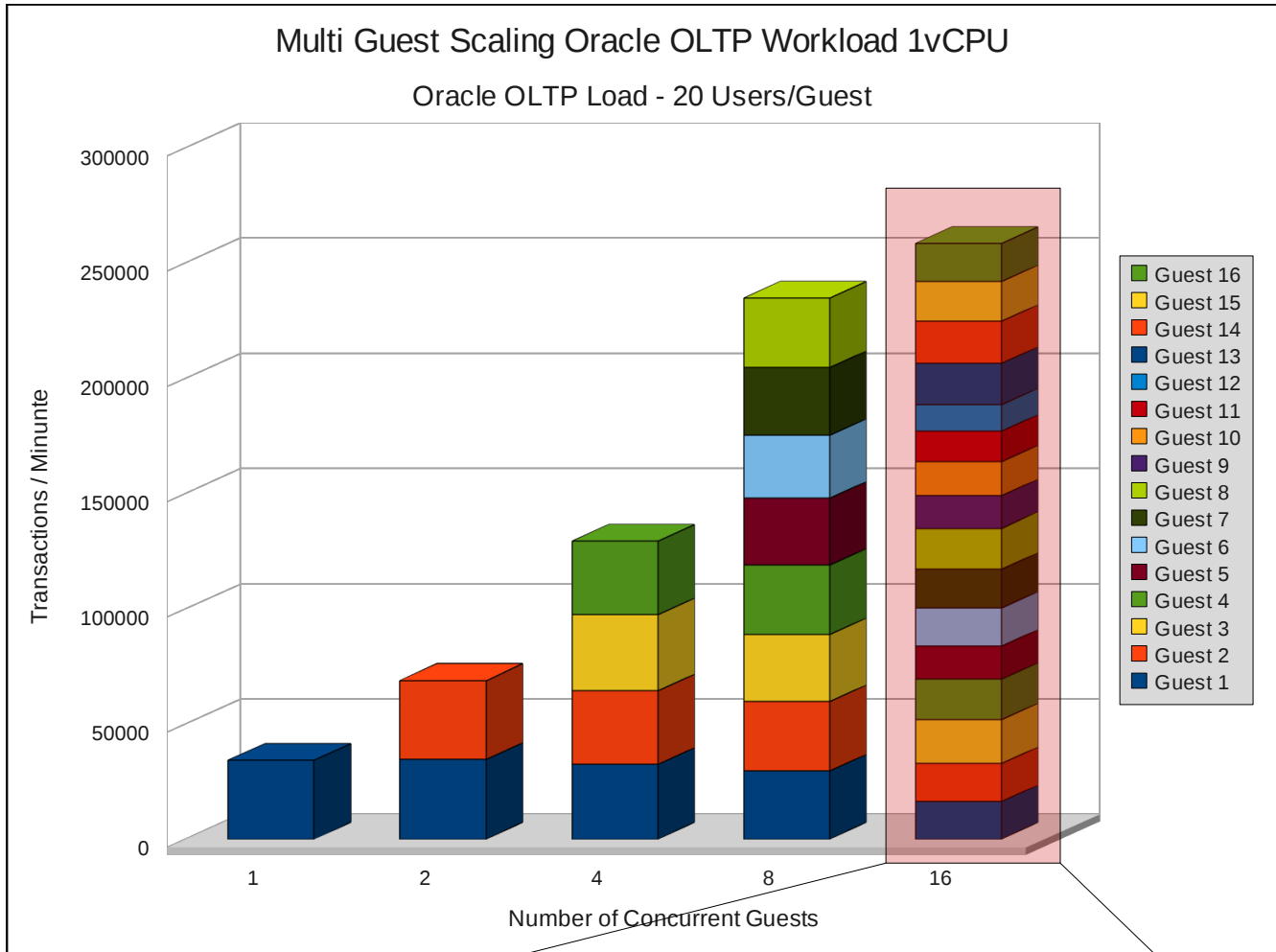


Figure 5.4-2: Results of Scaling Multiple 1-vCPU Guests

NOTE: Per guest performance begins to plateau with increasing guests due to storage hardware limitations. As additional guests are added, increased volume usage across the LUNs occurs. The storage environment used only allows eight virtual disks to be created. To accommodate sixteen guests, two volumes per virtual disk were used. The sharing of virtual disks amongst each guest increases storage I/O contention creating a reduction in efficiency. This simply indicates that more storage resources are needed in order to achieve optimal scaling and performance.



5.5 Scaling-Up Resources in a Single Guest

This section presents the results obtained when running an Oracle OLTP workload on a single guest with increasing amounts of memory and vCPUs.

Figure 5.5-1: Scaling the Memory and vCPUs in a Single Guest illustrates the configuration as resources are added.

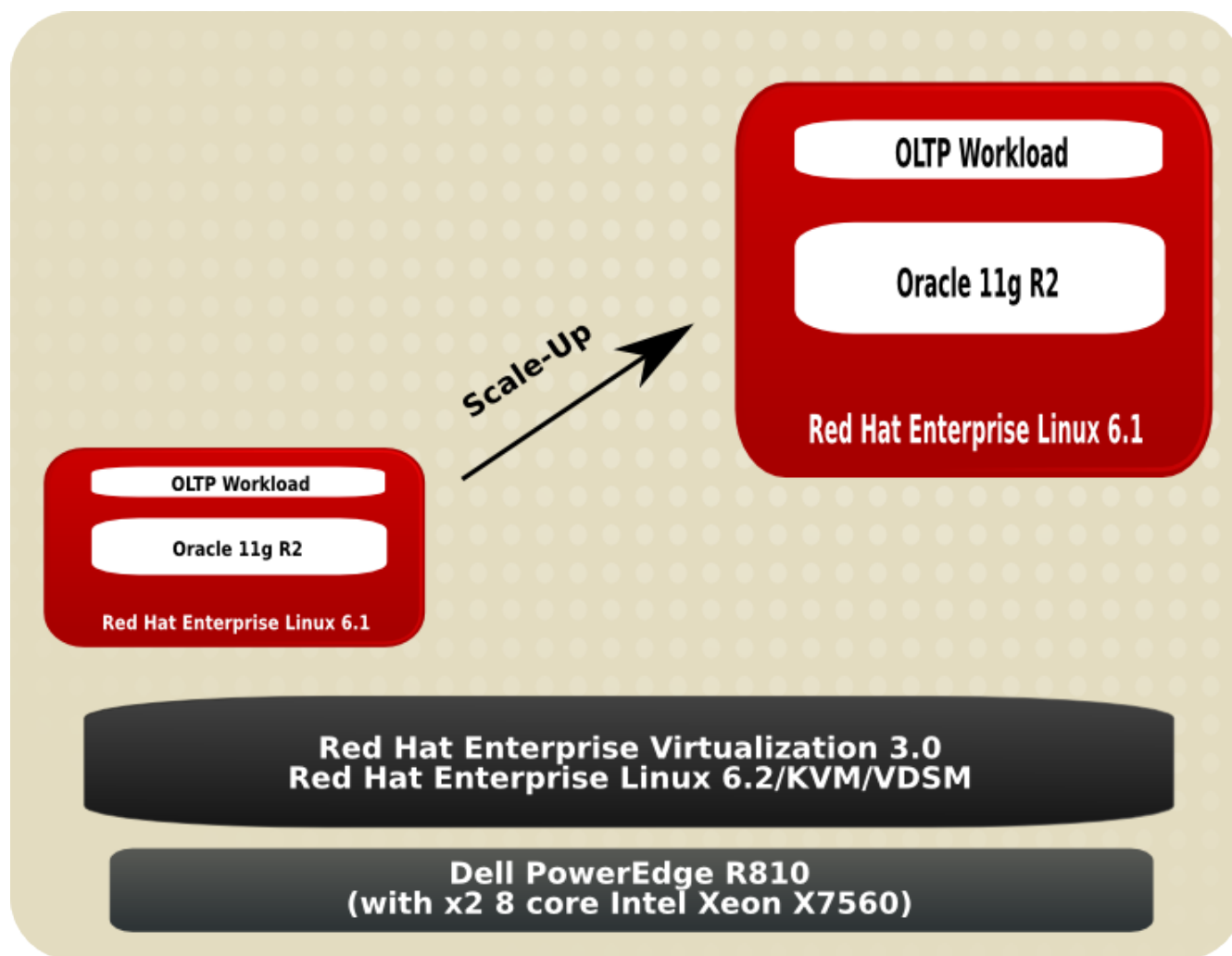


Figure 5.5-1: Scaling the Memory and vCPUs in a Single Guest



Figure 5.5-2: Scaling Up Resources in a Single Guest graphs the results when the OLTP workload was executed on a guest with one, two, four, and eight vCPUs with 7.5GB of memory for each vCPU. The throughput demonstrates good scaling. As vCPUs are added, the overall throughput per vCPU decreases due to I/O contention and virtualization overhead.

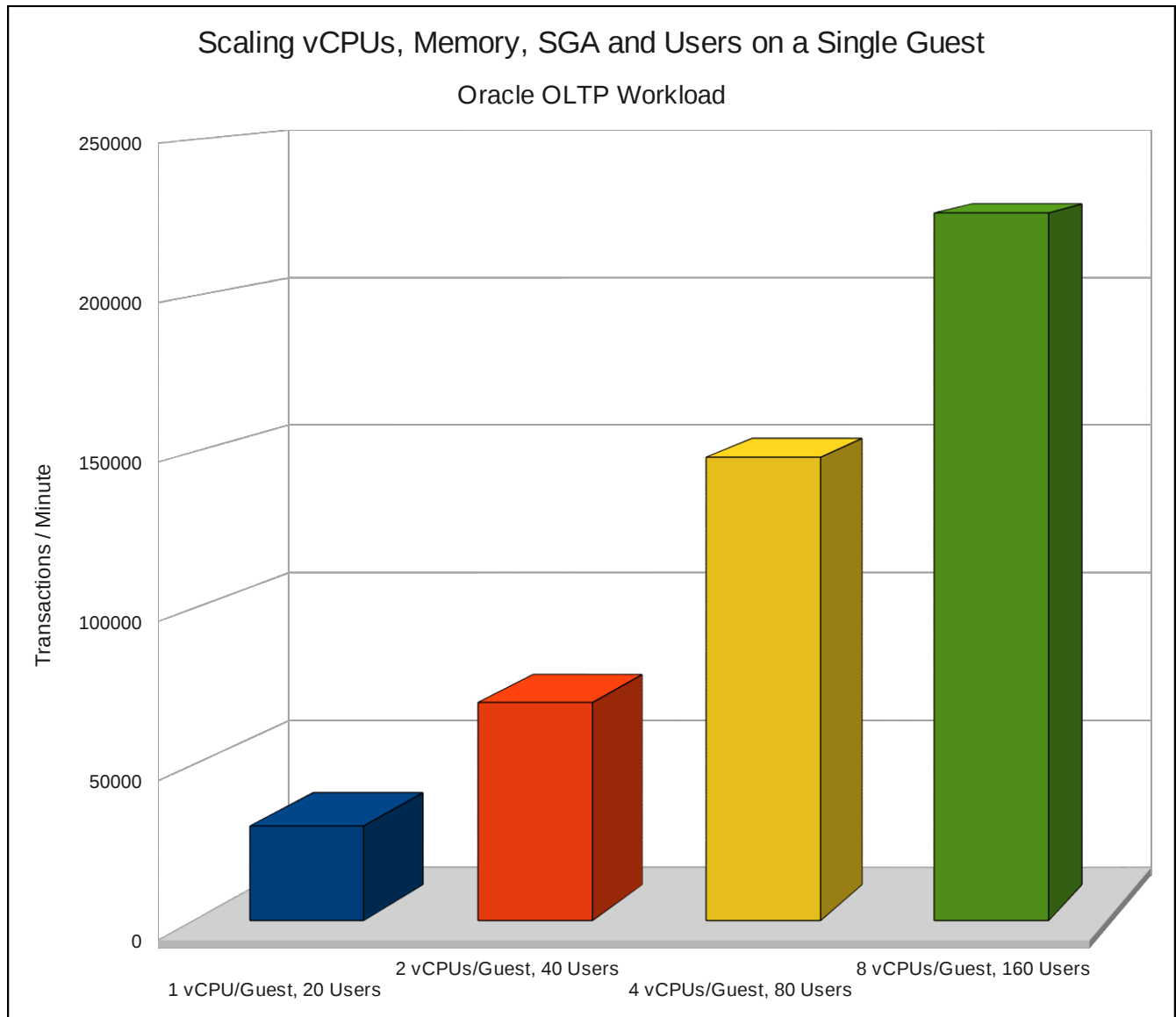


Figure 5.5-2: Scaling Up Resources in a Single Guest



5.6 Consolidated Virtualization Efficiency

Figure 5.6-1: Virtualization Efficiency compares the throughput performance of a 16 core bare-metal configuration to various virtual machine configurations totaling 16 vCPUs. In the virtual environment, this test was run with sixteen 1-vCPU guests, eight 2-vCPU guests, four 4-vCPU guests, and two 8-vCPU guests.

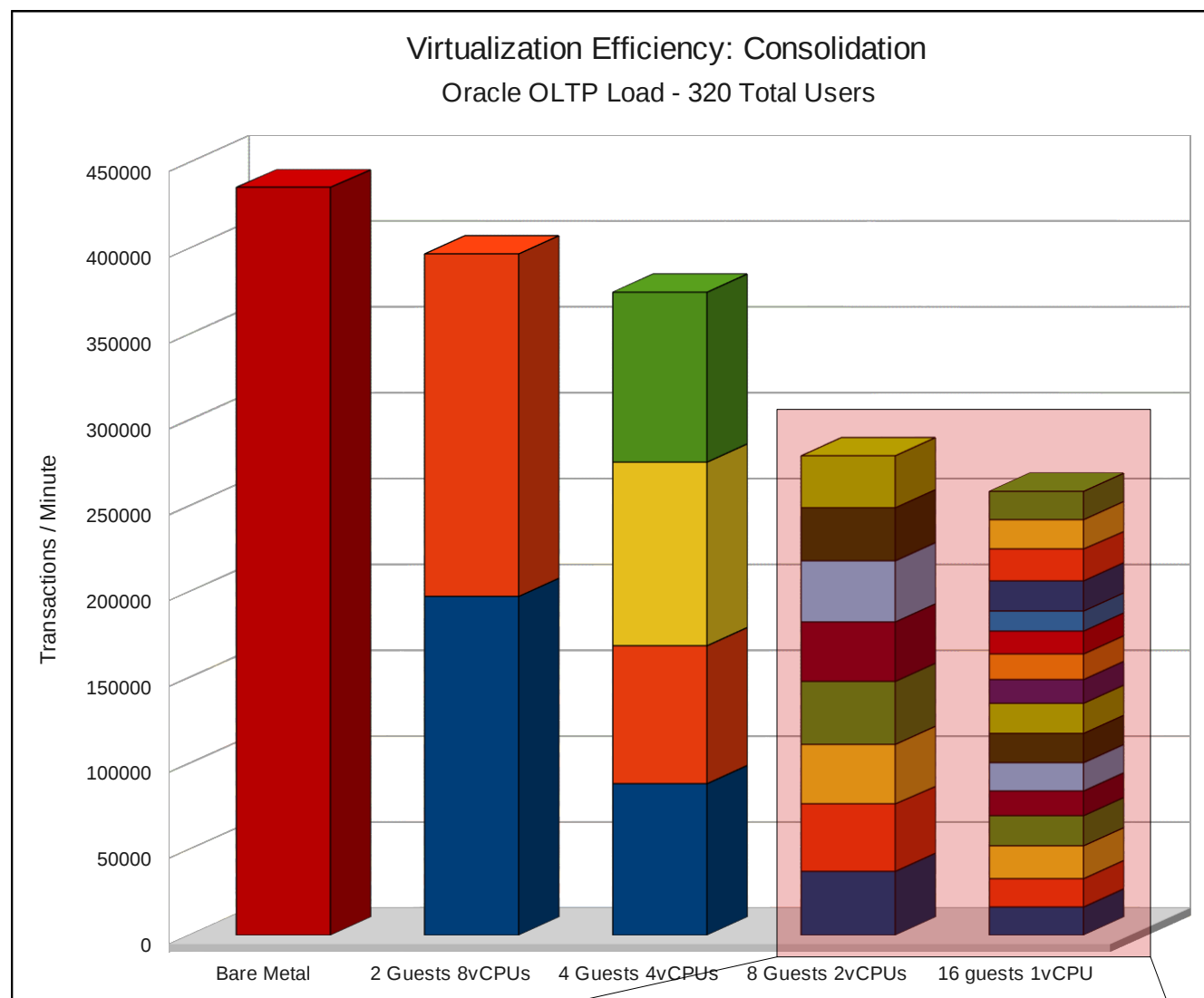


Figure 5.6-1: Virtualization Efficiency

NOTE: Lower throughput for the 8 Guest, 2vCPU and 16 Guest, 1vCPU configurations are a direct result of limited storage resources vs. RHEV scalability. This simply indicates that more storage resources are needed in order to achieve optimal scaling and performance.



6 Conclusion

This paper describes the performance and scaling of the Oracle OLTP workload running in Red Hat Enterprise Linux 6.1 guests on a Red Hat Enterprise Linux 6.2 host within a RHEV 3.0 environment. The host system was deployed on an Dell PowerEdge R810 server equipped with 128 GB of RAM and comprised of dual CPUs, each with a 2.26 GHz Intel Xeon X7560 Nehalem-EX processor totaling 16 non-hyperthreaded cores.

The data presented in this paper clearly establishes that RHEV virtual machines using Red Hat Enterprise Linux 6.2 with KVM as the hypervisor on a Dell PowerEdge R810, provide an effective production-ready platform for hosting multiple virtualized Oracle OLTP workloads. The combination of virtualization flexibility and the ability to both scale-up and scale-out contribute to the effectiveness of RHEV for Oracle. Examples include:

- ease of deployment
- ease of resource changes for each guest
- single management interface for multiple machines
- reporting capabilities
- reduced single points of failure inherent to many hardware only solutions
- ease of scaling multiple workloads across multiple guests

The number of actual users and throughput supported in any specific customer situation will ultimately depend on the specifics of the customer application used, the intensity of user activity, and the limitations of the hardware capacity involved.

In closing, the results demonstrate that in a virtualized environment, good throughput was retained even as the number and size of guests are increased until the physical server and storage throughput were fully subscribed.



Appendix A: Revision History

Revision 1.0

Sunday 01/16/12

Brett Thurber

- Initial document

Appendix B: Contributors

Contributor	Title	Contribution
Sanjay Rao	Principal Software Engineer, Red Hat	Content, Review
Mark Wagner	Principal Software Engineer, Red Hat	Content, Review
Steve Reichard	Sr. Principal Software Engineer, Red Hat	Content, Review
Tim Wilkinson	Sr. Software Engineer	Content, Review



- 1 <http://www.redhat.com/promo/qumranet/>
- 2 http://docs.redhat.com/docs/en-US/Red_Hat_Enterprise_Virtualization/3.0/html/Administration_Guide/Administration_Guide-Create_Objects-Create_Desktop.html#Administration_Guide-Create_Server-Existing_Template
- 3 http://docs.redhat.com/docs/en-US/Red_Hat_Enterprise_Virtualization/3.0/html/Administration_Guide/VDSM_Hooks.html