What's New in VMware® vSphere™ 4.1 — Performance

VMware vSphere 4.1

WHITE PAPER

vmware®

Table of Contents

Introduction

VMware® vSphere™ 4.1 ("vSphere") delivers dramatic performance and scale for the platform, introducing new performance capabilities for the most efficient aggregation of all datacenter resources into an elastic pool of computing power that applications utilize on an as needed basis, to ensure SLAs are met. Improvements include:

- Faster vCenter server startups
- Faster vSphere client connects
- Faster adding host operation

vSphere 4.1 scales to an unmatched number of virtual machines and virtualized hosts to support the buildout of private and public clouds at even lower operational costs than before:

- Better vSphere client responsiveness and snappier user interaction
- Support for larger number of concurrent operations with lowered latencies and increased throughput
- Better load balancing due to improved VMware Distributed Resource Scheduler (VMware DRS) algorithm
- Lower resource usage by vCenter agents

vSphere 4.1 enables higher consolidation ratios with unequaled performance by providing ground breaking new memory-management technology and expanding its resource pooling capabilities with new granular controls for storage and the network. Some of the highlights are:

Compute/Performance

- Memory compression—reclaim application performance by up to 30 percent by reducing memory contention as a bottleneck

Storage

- Storage I/O Control—set storage quality of service priorities per virtual machine for guarantee millisecond access to storage resources.
- Performance reporting—deliver key storage performance statistics regardless of storage protocol

Network

- Transmit Performance—increase Virtual-machine-to-virtual-machine throughput communication by 2x

Scalability

- vSphere 4.1—fully virtualize the datacenter and scale at least 2x more than ever before

Scalability Enhancements

New resource limits

In vSphere 4.1, a number of scalability enhancements have been incorporated into vCenter.

VCENTER	VSPHERE 4	VSPHERE 4.1	IMPROVEMENT
Virtual machines per cluster	1,280	3,000	3x
Hosts per vCenter Server	300	1,000	3x
Powered-on virtual machines per vCenter Server	3,000	10,000	3x
Registered virtual machines per vCenter Server	4,500	15,000	>3x
Concurrent vSphere clients	30	100	3x

CPU Enhancements

Wide VM NUMA (new feature)

VMware ESX®/ESXi™ use a sophisticated NUMA scheduler to dynamically balance processor load and memory locality of virtual machines when running on NUMA host. This is especially important with large workloads. A virtual machine that has more virtual processors than the number of physical processor cores available on a single socket, must span NUMA nodes. This improves the performance of certain memory-intensive workloads for wide VMS.

How does it improve performance?
Wide VM NUMA support can improve performance with the following conditions:

- The virtual machine is wider than the number of cores per NUMA node; in other words, 8vCPU virtual machine on a 4 physical CPU system or 4vCPU virtual machine on a 2 physical CPU system
- The system has many NUMA nodes; that is, 8vcpu virtual machine on a four-socket quad-core system, 4vcpu virtual machine on a four-socket dual-core system
- Better memory locality leads to optimal memory access latencies, improving application performance

Performance benefit, of course, depends on workload and configuration. For example, internal studies based on SPECjbb, a memory-intensive Java workload, show up to 7 percent performance improvement.

Memory Enhancements

Memory Compression (new feature)

Memory compression in vSphere 4.1 is a new hierarchy in VMware's memory overcommit technology, a VMware key differentiator. This is a new level of the memory overcommit positioned between the use of ballooning and disk swapping. The new feature optimizes the use of memory by compressing a virtual machine's memory on the fly, freeing up space as needed. This is yet another way for this technology to reclaim some performance as a last resort when memory is under contention before disk swapping.

- Transparent page sharing
- Ballooning
- Memory compression
- Disk swapping

How does it work?

Accessing compressed memory is faster than accessing memory swapped to disk. When a virtual page has exhausted transparent page sharing and ballooning it must be swapped to disk. In vSphere 4.1, there is an attempt first to compress the page and store it in the virtual machine's compression cache. Using the "Advanced Settings" dialog box in the vSphere client, the maximum compression cache size can be set.

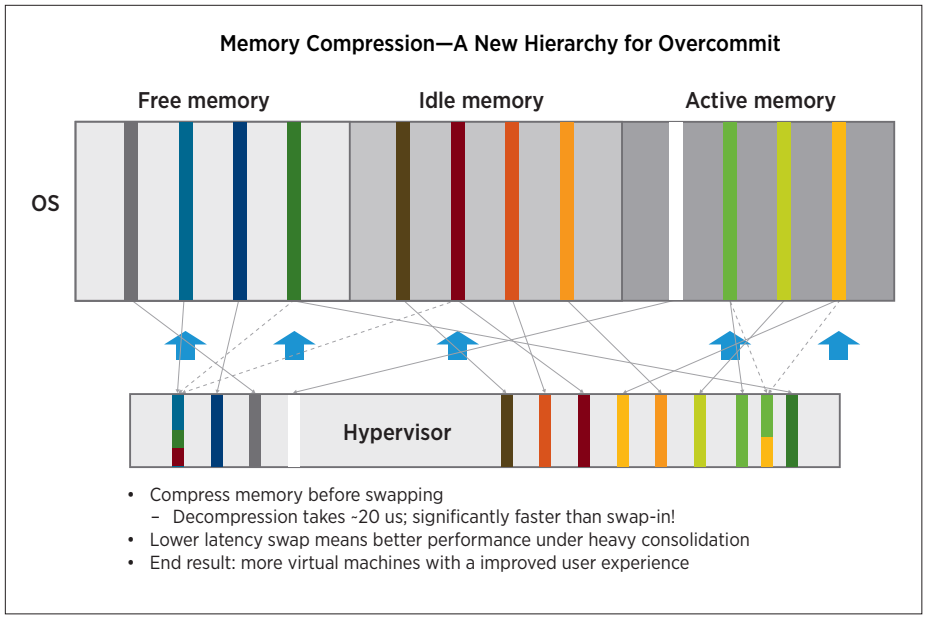

Figure 1.

Performance Improvements

The following graph depicts the relative cumulative performance (16 virtual machines) of the Oracle Swingbench workload with and without memory compression enabled. The workloads fit into memory until about 70GB. As the memory size decreases, this forces the workload to incur swapping, which is to be expected. The cost of disk swapping is very expensive, lowering workload performance.

Memory compression recovers much of the performance by avoiding swapping to disk as much as possible. It is not until transparent page sharing, ballooning and all compression options are exhausted that swapping to disk is used. Sustainability of performance is longer with memory compression.

As a host becomes memory starved there is:

- Up to 15 percent performance improvement, with some memory overcommitment
- Up to 25 percent performance improvement, with heavy memory overcommitment

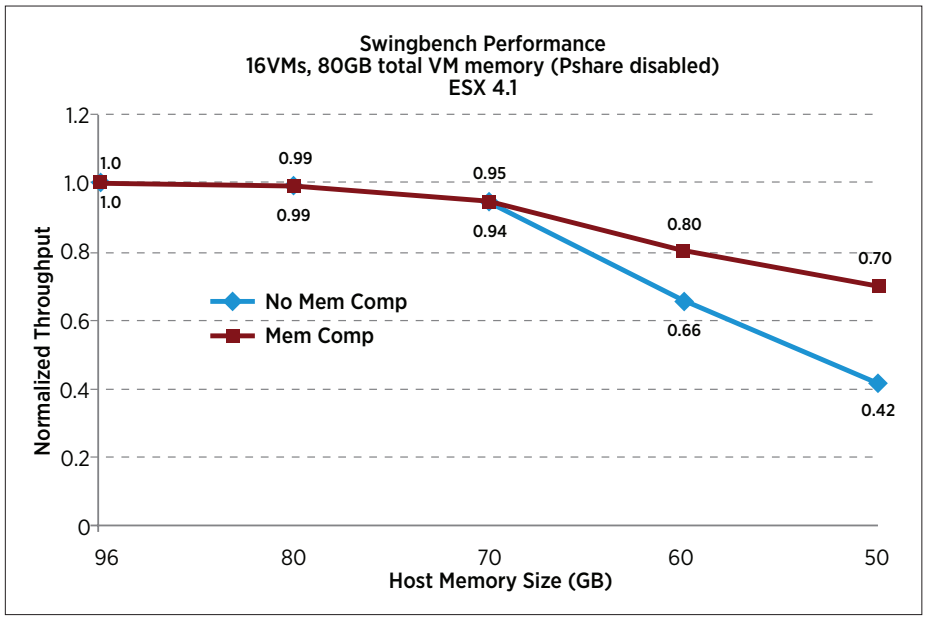

Figure 2.

Storage Enhancements

Storage I/O Control (new feature)

Using Storage I/O control (SIOC), vSphere administrators can ensure that the most important virtual machines get adequate I/O resources in times of congestion. Storage I/O control enables cluster-wide virtual machine storage I/O prioritization enabling sustainable performance for the most mission critical applications. SIOC extends the concepts of CPU and memory shares and limits to handle storage I/O resources as well. SIOC now enables vSphere administrators to set these I/O shares and limits to virtual machines, and prioritize I/O bandwidth. This feature delivers traffic-management controls, allowing flexible partitioning of physical NIC bandwidth between different traffic types, including virtual machine, vMotion, FT, and IP storage traffic (enabled with Distributed Switch only).

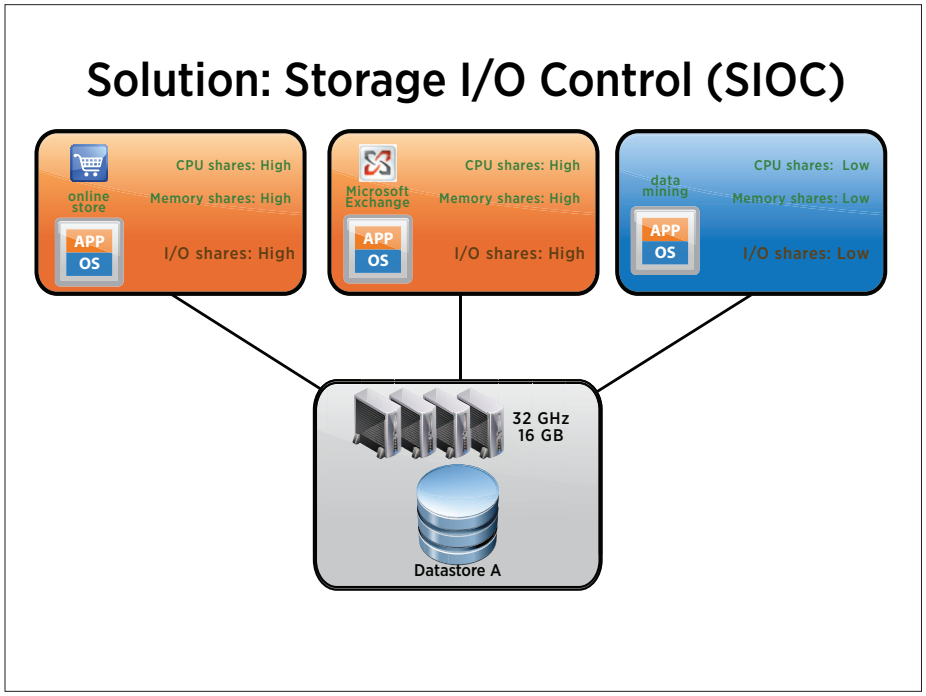

Figure 3.

How does it work?

When SIOC on a datastore is enabled (FC or iSCSI supported), ESX/ESXi begins to monitor the device latency that hosts detect when communicating with that datastore. When device latency exceeds a "congestion" threshold of 30ms sustained for at least 4 seconds, the datastore is considered to be congested, and each virtual machine that accesses that datastore is allocated I/O resources in proportion to their I/O shares as configured by the administrators per virtual machine. SIOC will use the I/O shares for each virtual machine in that datastore to balance the use of I/O slots on each ESX server to align priority with throughput. One can adjust shares for each virtual machine based on need, ensuring that storage allocations are done in line with virtual machine priorities.

This feature enables per-datastore priorities/shares for virtual machines to improve total throughput and has cluster-level enforcement for shares for all virtual machines accessing a datastore. Configuring SIOC is a three-step process for a datastore:

1. Enable SIOC.
2. Set a congestion threshold.
3. Configure the number of storage I/O shares and upper limit of I/O operations per second (IOPS) allowed for each virtual machine. By default, all virtual machine shares are set to "Normal" (1,000), with unlimited IOPS.

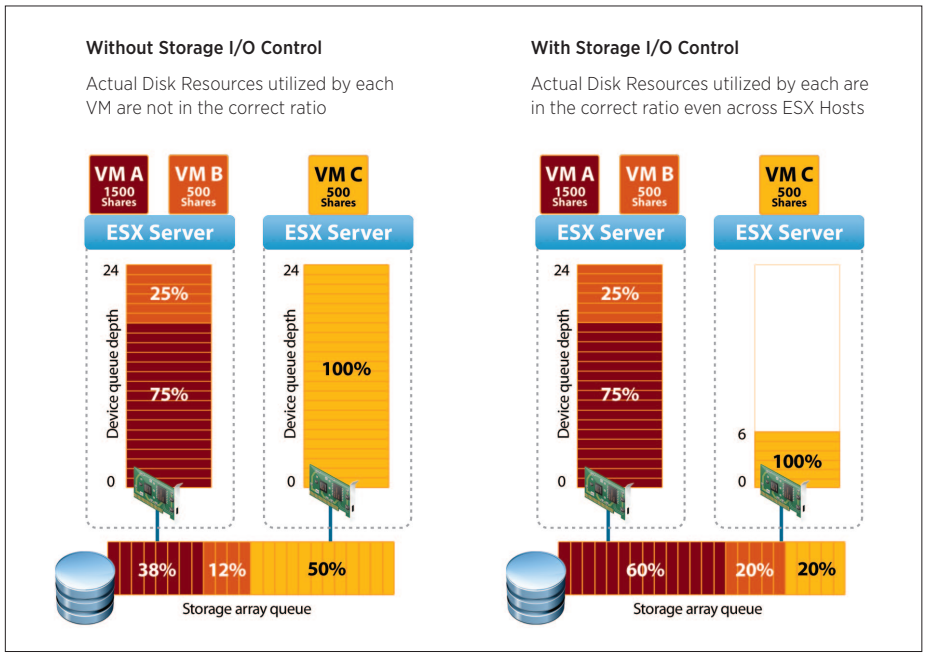

Figure 4.

Performance Improvements

The SIOC feature enforces proportional I/O share of virtual machines across hosts in an ESX cluster and supports various FC and iSCSI SAN storage. Figure 5 compares the performance of the critical workload as the portion of the host's I/O queue allocated to the virtual machine was varied.

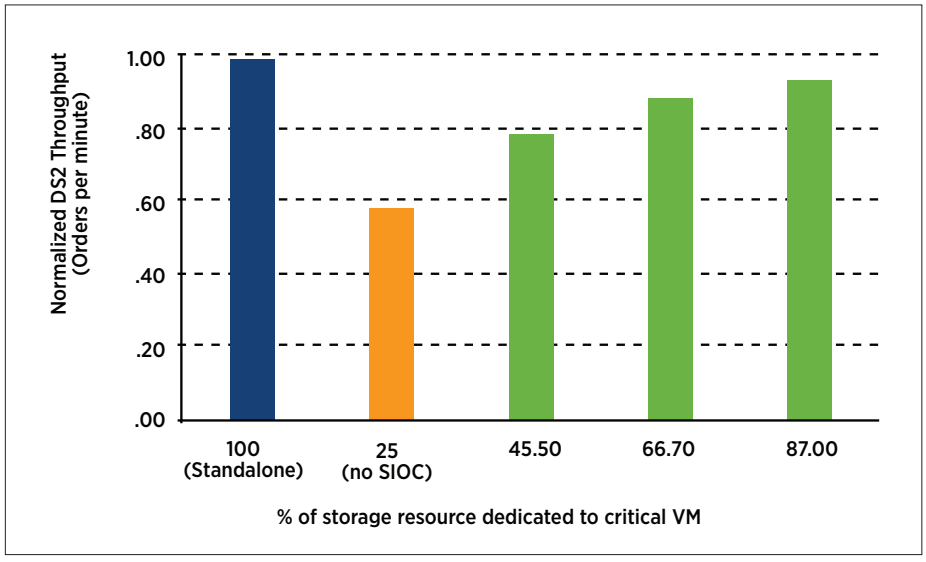

Figure 5. Impact of Storage I/O Control on performance of DVD Store workload

The improvements were measured with the DVD Store Version 2 (DS2) workload, which is a complete, online e-commerce test application, with a back-end database component, a Web application layer, and driver programs. In Figure 5, the blue bar represents performance of the workload in isolation. The orange bar represents the workload while 4 other identical loads are running simultaneously. Performance drops by 41 percent when the workload has to equally share resources with the other workloads running. With SIOC enabled and the workload prioritized as high, very little change in performance occurs and the high-priority workload is performing at 94 percent of the blue bar.

Storage Protocol Improvements

Optimized networked storage performance is crucial for optimal virtualized datacenter performance and scalability. vSphere 4.1 has a number of key improvements.

8Gb FC storage support

vSphere 4.1 now supports 8Gb FC storage arrays. ESX/ESXi can be deployed with 8Gb end-to-end FC SANs. With 8Gb support, ESX is able to effectively double the measured throughput over 4KB with transfer size greater than 8K, as shown in the following graph.

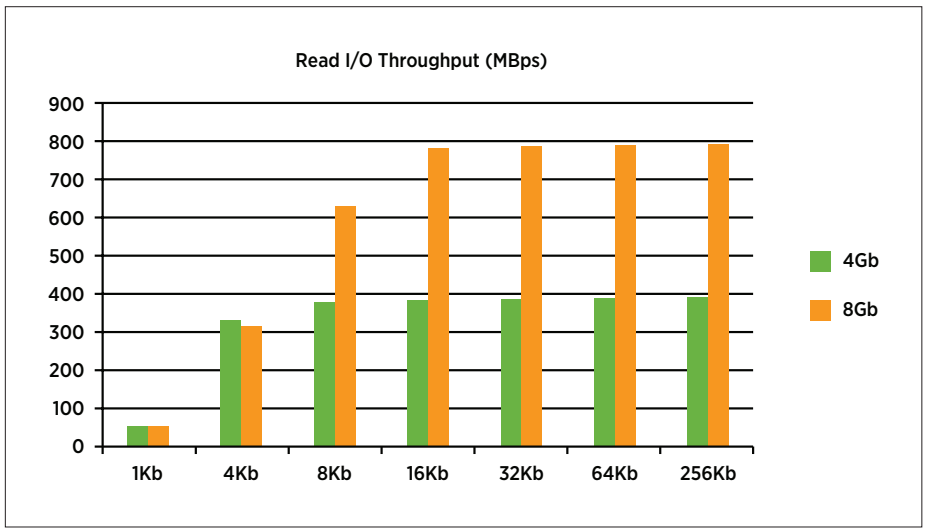

Figure 6.

NFS improvements

In vSphere 4.1, NFS throughput over 10GbE shows improvements of 12–40 percent for Reads, and 32–124 percent for Writes. These results were measured with storage micro-benchmarks.

For most workloads, improvements would vary depending on the workload characteristics.

Software iSCSI improvements

In vSphere 4.1, software iSCSI throughput over 10GbE shows improvements of 6–23 percent for Reads, and 8–19 percent for Writes. These results were measured with storage micro-benchmarks.

For most workloads, improvements would of course vary depending on the workload characteristics.

Network Enhancements

Transmit Performance

In the earlier vSphere releases, networking packet transmissions were synchronous. This resulted in the execution of the guest within a virtual machine being blocked during the duration of the transmission, and in vSphere 4.1 this has been made asynchronous, with the ability of the async transmission thread to run on any CPU core. This frees up more CPU cycles for virtual machines, making more CPU resources available. This has resulted in improving:

- Virtual-machine-to-virtual-machine throughput by 2x, to up to 19 Gb/sec
- Virtual-machine–to-native throughput by 10 percent

Software and Hardware Large Receive Offload

ESX now supports hardware LRO for Linux. This allows ESX to aggregate packets before sending them to a virtual machine, which reduces per packet processing overhead in the virtual machine (does not have to be interrupted each time, networking stack processing cost amortized across multiple packets). Receive networking tests indicate 5–30% improvement in throughput and 40–60% decrease in CPU cost, depending on the workload.

Fault-Tolerant Virtual Machine Throughput

There is a significant 3.6x increase in fault-tolerant throughput in vSphere 4.1 (9Gb/sec) over vSphere 4.0 (at 2.5 Gb/sec) using 10GbE. Fault-tolerant virtual machines generating large amounts of logging traffic now can fully utilize the bandwidth of a 10GbE network.

VDI Enhancements

A number of improvements in the platform have enhanced VDI performance:

1. Transparent memory compression improves boot-storm time (up to 40 percent) under moderate memory overcommit.
2. Critical VDI operations such as "Create Virtual Machine," "PowerOn/Off," "Un/Register" and "Reconfigure," improved significantly compared to vSphere 4.0.
   a. "Create Virtual Machine" is 60 percent faster; "PowerOn" is 3.4x faster.
   b. Achieve significantly higher limits in vSphere 4.1.
3. "Async Transmit" improves VDI performance by as much as 8 percent.

vMotion™ Enhancements

Scalable vMotion

The number of concurrent live migrations has increased from a default of two in vCenter 4.0 to as many as eight (on a 10GbE network) in vCenter 4.1, with the same advantages of zero downtime, continuous service availability and complete transaction integrity. vMotion always leveraged performance of a 10GbE network but now scalable vMotion leverages even more to support the simultaneous migration of up to eight virtual machines.

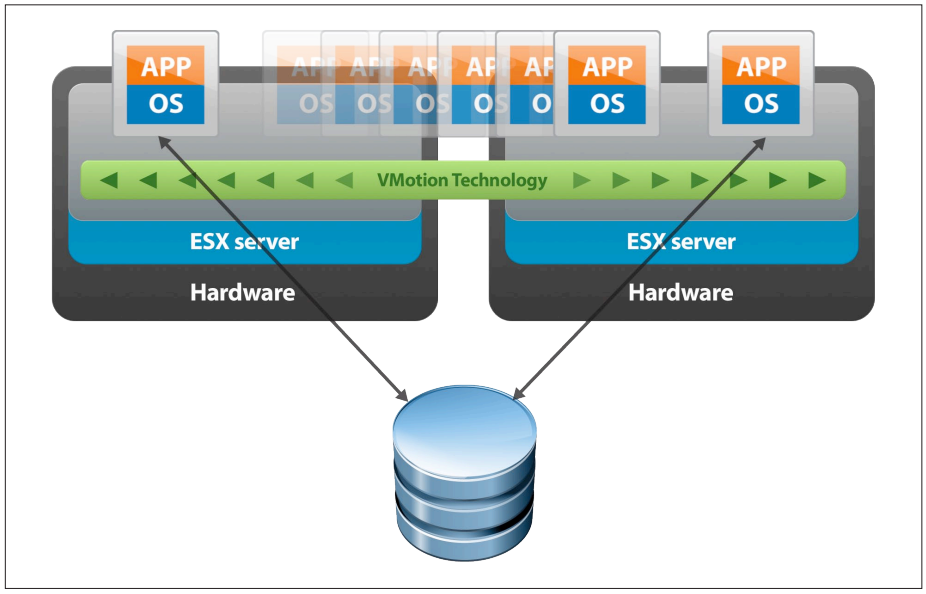

Figure 7.

Performance Improvements

In vSphere 4.1, scalable vMotion improves throughput by greater than 3x (8Gb/sec) on 10GbE links, thereby improving vMotion performance and scalability. Specific optimizations in ESX 4.1 include:

- Reorganized for a new network I/O model (STREAMS — vMotion has been optimized to maximize its transmit and receive rates to dynamically available host resources, allowing hosts with available processors and memory resources to execute with higher concurrency, lower latency, and increasing bandwidth over a network faster than 1Gbps)
- Optimized virtual machine memory handling
- Optimized vMotion convergence logic (how and when vMotion exits)
- "Quick resume" (RDPI allows the guest to progress concurrently with the memory updates, thereby reducing any latency period for many workloads, especially those with high page-dirty rates — memory intensive workloads, some database workloads)

Storage vMotion

VMware Storage vMotion is a component of VMware vSphere that provides an interface for live migration of virtual machine disk files within and across storage arrays, with no downtime or disruption in service. Storage vMotion relocates virtual machine disk files from one shared storage location to another, with zero downtime, continuous service availability and complete transaction integrity. Customers use Storage vMotion to

- Simplify array migrations and storage upgrades
- Dynamically optimize storage I/O performance
- Efficiently manage storage capacity

Storage vMotion is fully integrated with VMware vCenter Server to provide easy migration and monitoring.

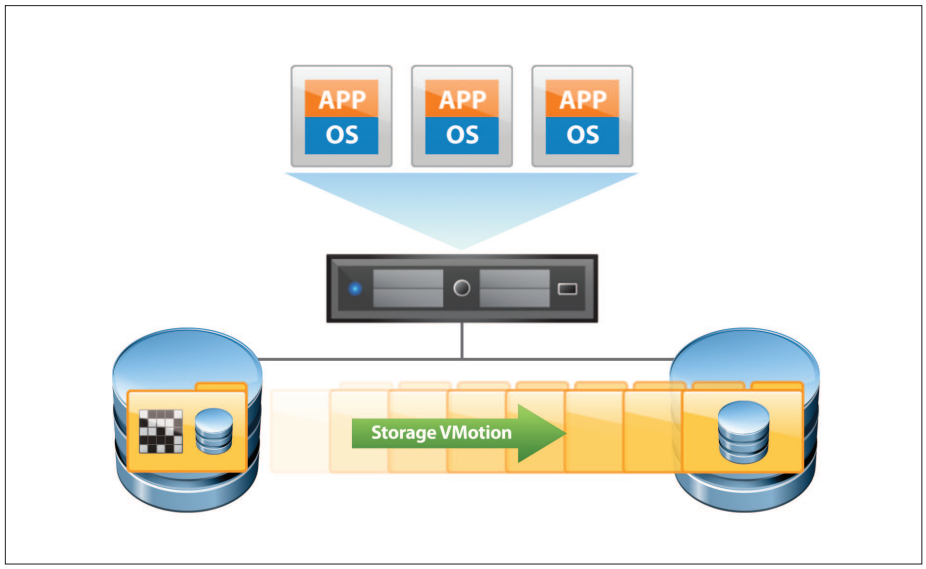

Figure 8.

Performance Improvements

Depending on the application workload characteristics and the size of the disk, time spent for updating the destination after cloning is reduced by up to 25 percent.

VMware vCenter and Performance Management Enhancements

A number of vCenter optimizations have resulted in significantly reduced latencies of common provisioning operations. Some operations, such as registering virtual machines and reconfiguring virtual machines, have improved by as much as 3x since vSphere 4.0.

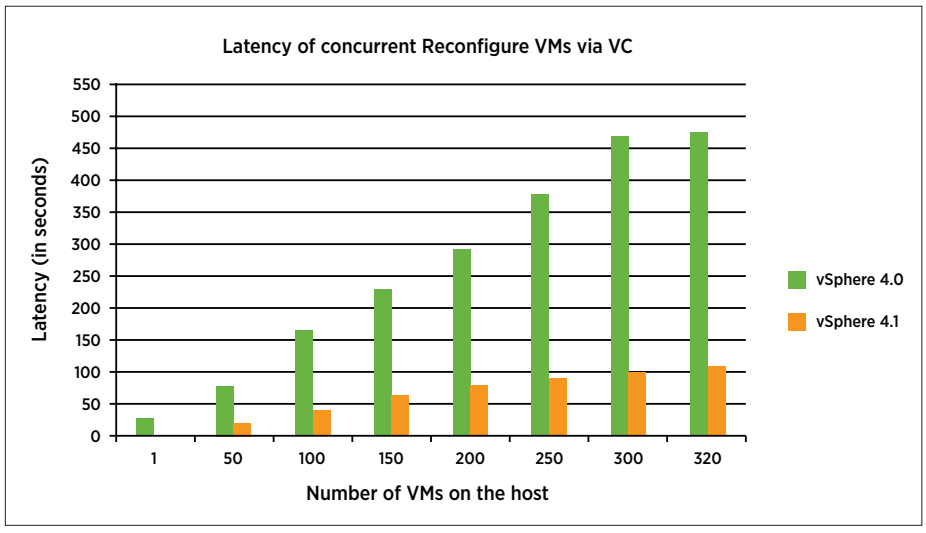

Figure 9.

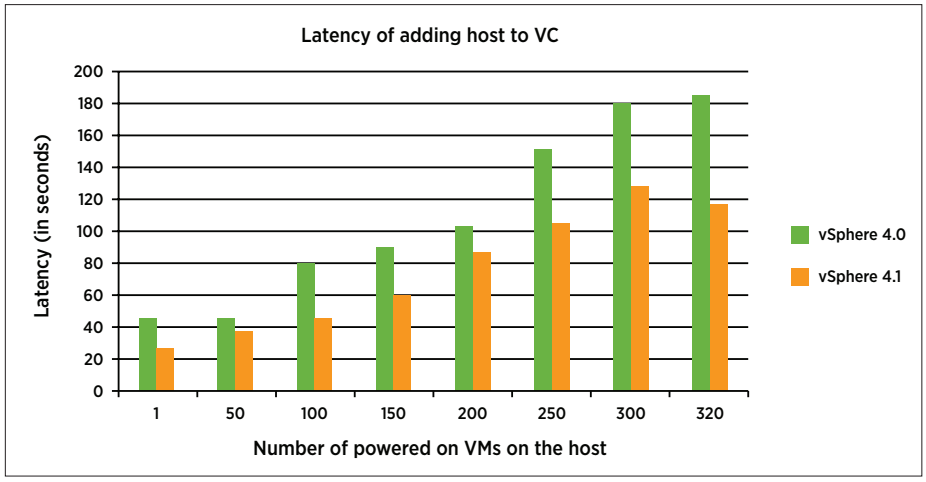

Figure 10.

Storage Performance Monitoring

New storage performance monitoring for vSphere 4.1 has comprehensive host and virtual machine storage performance statistics, enabling proactive monitoring to simplify troubleshooting, heterogeneous customer storage environments supported (FC, iSCSI, NFS), real-time and historical trending (vCenter), and esxtop (for ESX) and resxtop (ESXi) support. Tools support varied usage scenarios:

- GUI for trending and user-friendly comparative analysis
- Command line for scripting/drill-down at host

The new additions support throughput and latency statistics for:

- Datastore per host
- Storage adapter and path per host
- Datastore per virtual machine
- VMDK per virtual machine

INVENTORY OBJECT	PER COMPONENT	PROTOCOL SUPPORT	VCENTER	ESXTOP
Host	Datastore	FC/NFS/iSCSI	Yes	Yes
	Storage adapter	FC	Yes	Yes
	Storage Path	FC	Yes	Yes
	LUN	FC/iSCSI	Yes	Yes
Virtual machine	Datastore	FC/NFS/iSCSI	Yes	Yes
	VMDK	FC/NFS/iSCSI	Yes	Yes

New Throughput and Latency Matrics

Summary

VMware innovations continue to ensure that VMware vSphere 4.1 pushes the envelope of performance and scalability. The numerous performance enhancements in vSphere 4.1 enable organizations to get even more out of their virtual infrastructure and further reinforce the role of VMware as industry leader in virtualization.

vSphere 4.1 represents advances in performance, to ensure that even the most resource-intensive and scale-out applications, such as large databases and Microsoft Exchange email systems, can run on private clouds powered by vSphere.

References

VMware Storage vMotion
www.vmware.com/files/pdf/VMware-Storage-vMotion-DS-EN.pdf

VMware vMotion
www.vmware.com/files/pdf/VMware-vMotion-DS-EN.pdf

vSphere Resource Management Guide
www.vmware.com/files/pdf/vsp_41_resource_mgmt.pdf

Managing Performance Variance of Applications using Storage IO Control
www.vmware.com/go/tp-managing-performance-variance

Configuration Maximums VMware vSphere 4.1
www.vmware.com/files/pdf/vphere4/r41/vsp_41_config_max.pdf

VMware, Inc. 3401 Hillview Avenue Palo Alto CA 94304 USA Tel 877-486-9273 Fax 650-427-5001 www.vmware.com