



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

VMware vSphere 4.1

TECHNICAL WHITE PAPER

Table of Contents

Scalability enhancements 4
 New resource limits 4
 CPU enhancements 4
 Wide VM NUMA 4
 Memory enhancements 5
 Memory compression 5
 Storage enhancements 7
 Storage I/O control (SIOC) 7
 8Gb Fibre Channel (FC) support 9
 NFS improvements 9
 Software iSCSI improvements 9
 Network enhancements 10
 Transmit performance 10
 Software and hardware large receive offload (LRO) 10
 Fault Tolerance (FT) throughput 10
 VDI enhancements 10
 VMware® vMotion™ (vMotion) enhancements 10
 Scalable vMotion 10
 VMware Storage vMotion 11
 VMware vCenter™ and performance management enhancements 12
 Storage performance monitoring 13
 Summary 14
 References 14

Introduction

VMware® vSphere™ 4.1 (“vSphere”), the industry’s first cloud operating system, pushes further ahead the performance and scalability of the VMware virtualization platform. vSphere 4.1 enables higher consolidation ratios with unequaled performance by providing groundbreaking new memory-management technology and expanding its resource-pooling capabilities with new granular controls for storage and the network. 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.

Some of the highlights are:

Compute/Performance

- Memory compression — reclaim application performance by reducing memory contention as a bottleneck

Storage

- Storage I/O Control— set storage quality of service priorities per virtual machine for better access to storage resources for high priority applications.
- Performance reporting — deliver new key storage performance statistics for NFS

vMotion™

- Support larger number of concurrent operations with lowered latencies and increased throughput

Network

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

VMware vCenter™ Scalability

- vSphere 4.1 — fully virtualize the data center and scale at least 3x more than ever before
- Faster vCenter server startups
- Faster vSphere client connects
- Faster adding host operation
- Better vSphere client responsiveness and snappier user interaction

Scalability Enhancements

New resource limits

In vSphere 4.1, a number of scalability enhancements have been incorporated into VMware 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)

Modern processors from both Intel and AMD have memory local to a processor socket. Access to this memory is faster from the socket associated with the memory than from the other sockets on the system. The speed of memory access is thus nonuniform across sockets and such systems are said to have a non-uniform memory access (NUMA) architecture. In such systems a set of processors and the memory local to it is referred to as a NUMA node.

VMware ESX®/ESXi™ uses a sophisticated NUMA-aware scheduler to ensure that virtual machines with vCPUs running on a NUMA node access memory that is local to that node. A virtual machine that has more virtual processors than the number of processors available on a NUMA node must span NUMA nodes. Such a virtual machine, with this release, will be intelligently NUMA-managed for optimal performance. It will now benefit from the memory locality in a manner similar to a virtual machine with a smaller vCPU count that does not span across NUMA nodes.

Performance Improvement

The new Wide VM NUMA feature ensures that a virtual machine with a vCPU count spanning NUMA nodes is managed by the NUMA scheduler, resulting in better memory locality for the applications within the virtual machines. This better memory locality leads to improved performance for the applications.

Performance benefit depends on the workload and configuration. Internal studies based on SPECjbb, a server Java workload, show up to 7% 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 new level of memory overcommit is positioned between the use of ballooning and disk swapping. In periods of memory contention on the host, right before swapping a page to the disk, we try to compress it first. This is yet another way for this technology to reclaim some performance when memory is under contention before disk swapping as a last resort.

- 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. The maximum compression cache size can be set using the "Advanced Settings" dialog box in the vSphere client.

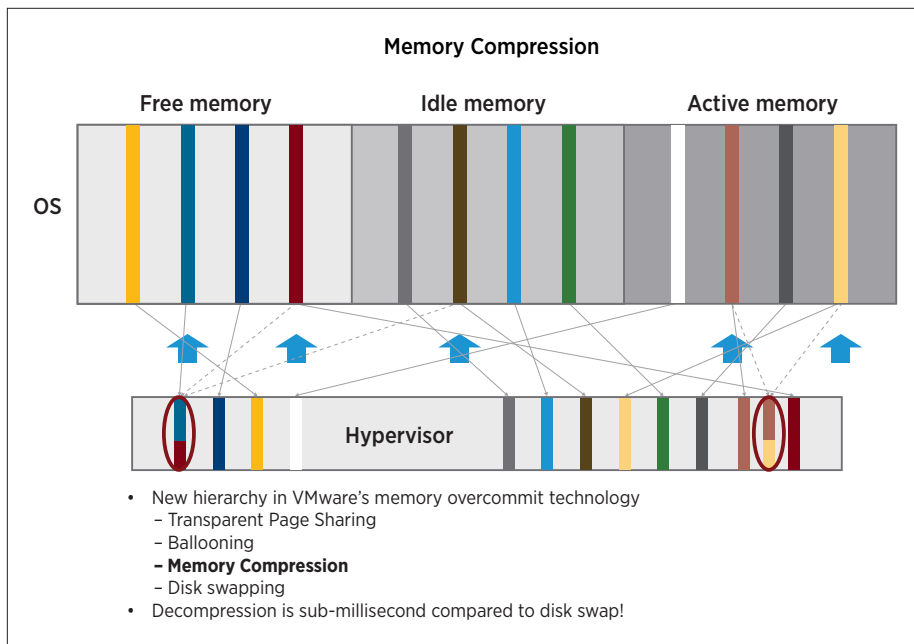


Figure 1.

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 workload fits into memory until about 70GB. As the memory size decreases, this forces the workload to incur swapping, which is to be expected. Disk swapping is 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. As a host becomes memory starved there is:

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

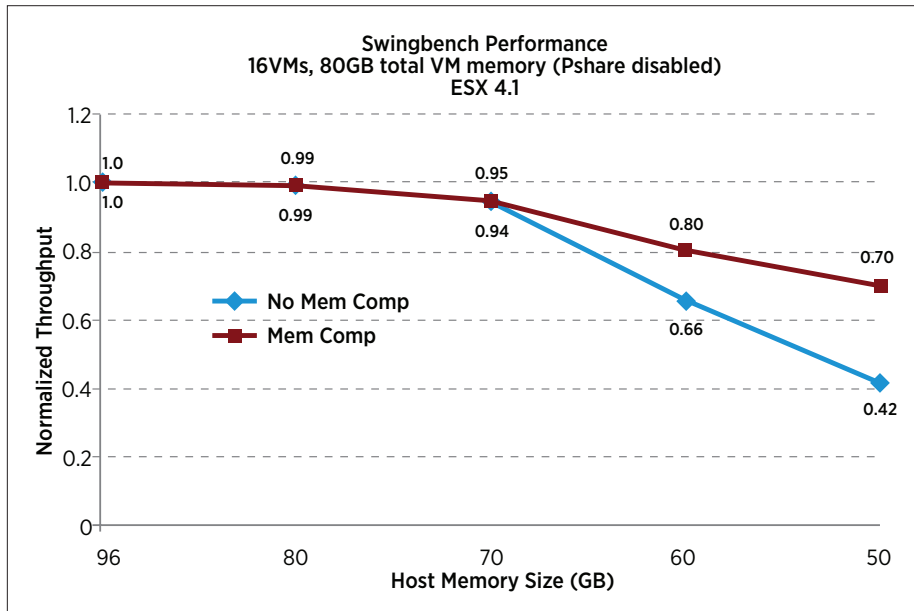


Figure 2.

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. SIOC enables cluster-wide virtual machine storage I/O prioritization. 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 to prioritize I/O bandwidth. This would enable sustained performance for the most mission-critical applications.

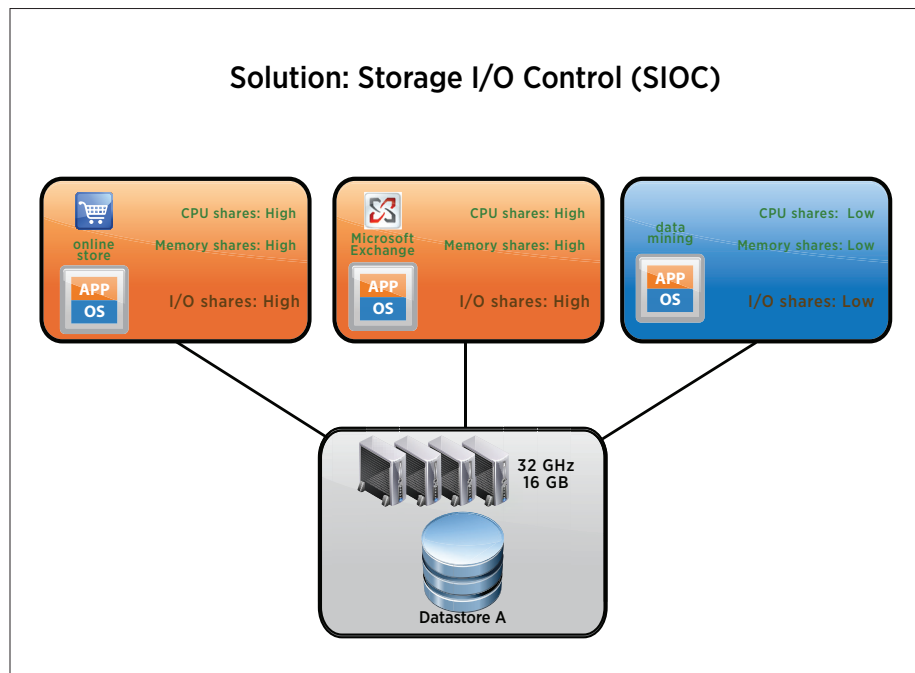


Figure 3.

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 default “congestion” threshold of 30ms sustained for at least 4 seconds, the datastore is considered to be congested. 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.

Storage I/O shares can be adjusted for each virtual machine based on need, ensuring that virtual machine 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 provides an example of how the assignments of I/O shares using SIOC results in configured prioritized access to the storage array.

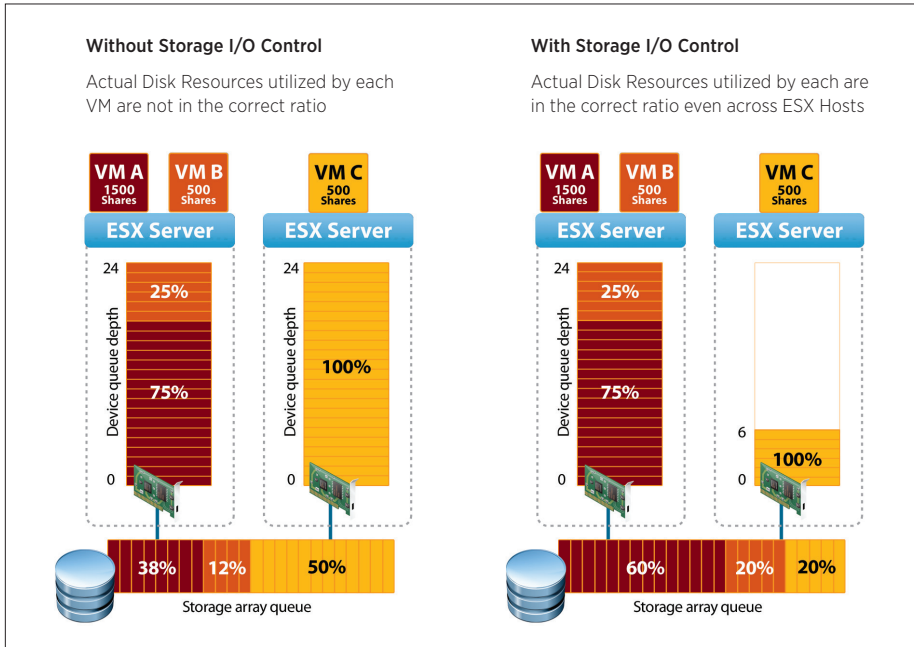


Figure 4.

Performance Improvements

The SIOC feature enforces proportional I/O share of virtual machines across hosts in an ESX cluster and supports both 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.

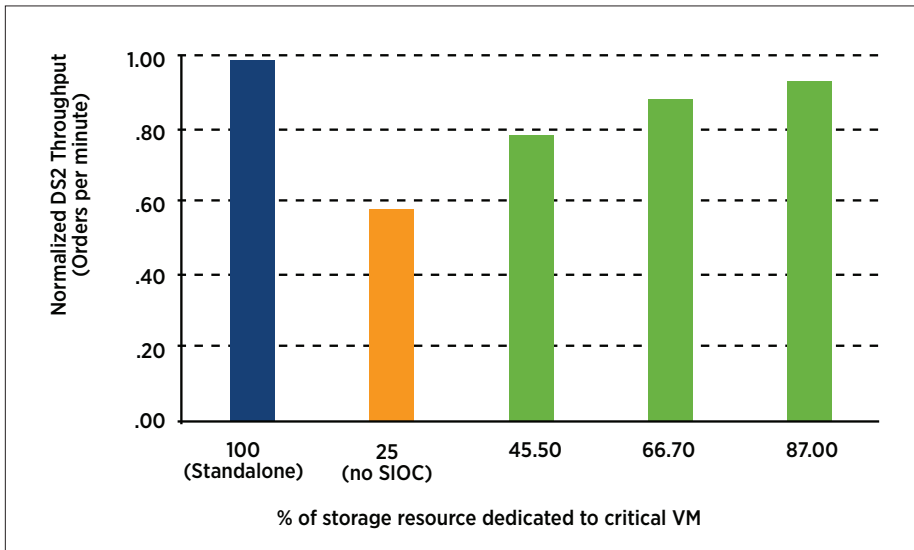


Figure 5. Impact of Storage I/O Control on performance of DVD Store workload

In Figure 5, the blue bar represents performance of the workload in isolation. The orange bar represent the workload while four other identical loads are running simultaneously. Performance drops by 41% 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% 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 4Gb with transfer size greater than 8KB, as shown in the following graph.

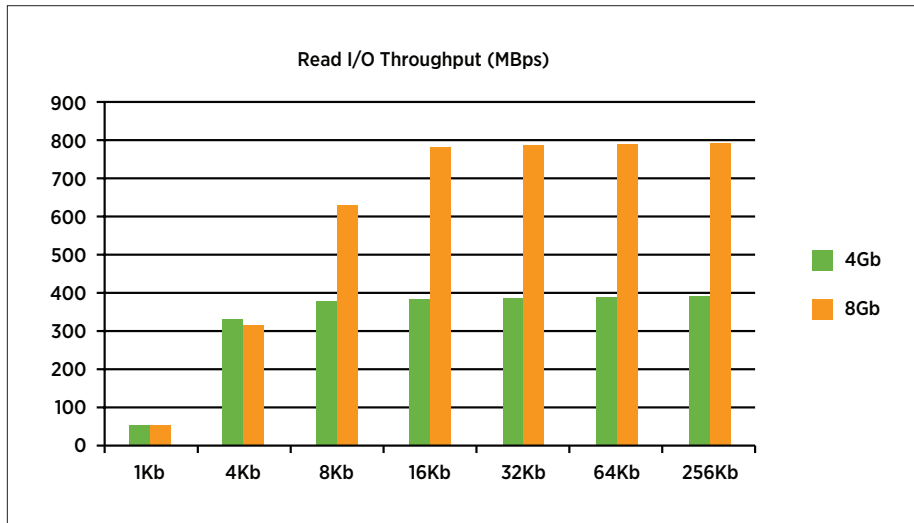


Figure 6.

NFS improvements

Using storage microbenchmarks, we observe that vSphere 4.1 NFS shows improvements in the range of 12–40% for Reads, and improvements in the range of 32–124% for Writes, over 10GbE.

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

Software iSCSI improvements

Using storage microbenchmarks, we observe that vSphere 4.1 Software iSCSI shows improvements in the range of 6–23% for Reads, and improvements in the range of 8–19% for Writes, over 10GbE.

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

Network Enhancements

Transmit Performance

In vSphere 4.1 virtual machine packet transmissions have been made asynchronous, with the ability for the asynchronous (async) transmit thread to run on any CPU core. Asynchronous transmissions enable the virtual machine to continue to make progress during the packet transmissions. In addition, allowing the async thread to run on any available CPU results in better scalability.

- Virtual-machine-to-virtual-machine throughput improves by as much as 2x, to up to 19 Gb/sec
- Virtual-machine-to-native throughput improves by 10%

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

Improvements in the VMware Fault Tolerance (VMware FT) module as well as the networking stack has resulted in significant improvement in fault-tolerant networking throughput using 10GbE. Fault-tolerant virtual machines generating large amounts of logging traffic now can fully utilize (~9Gb/sec) the bandwidth of a 10GbE network.

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. Memory compression improves boot-storm time (up to 40%) 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% faster; “PowerOn” is 3.4x faster.
3. Achieves significantly higher limits in vSphere 4.1 (see *Configuration Maximums VMware vSphere 4.1*).
4. Asynchronous network transmissions (discussed in “Network Enhancements”) improves VDI performance by as much as 8% .

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.

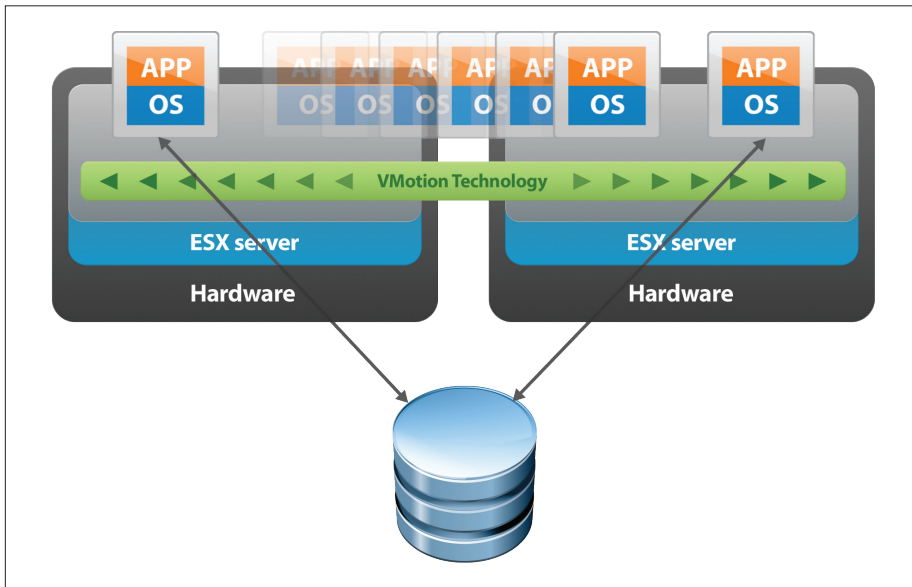


Figure 7.

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 leverage 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.

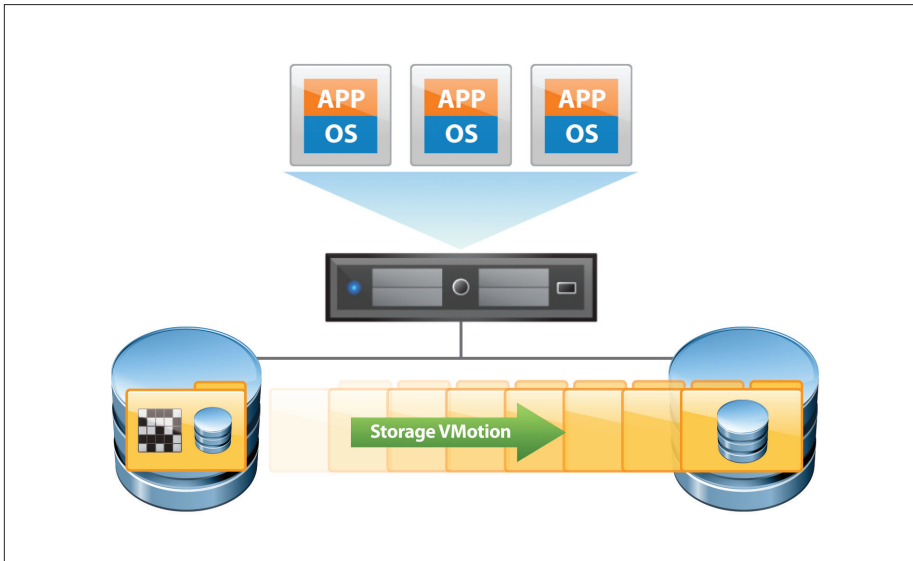


Figure 8.

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%.

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 as compared to vSphere 4.0.

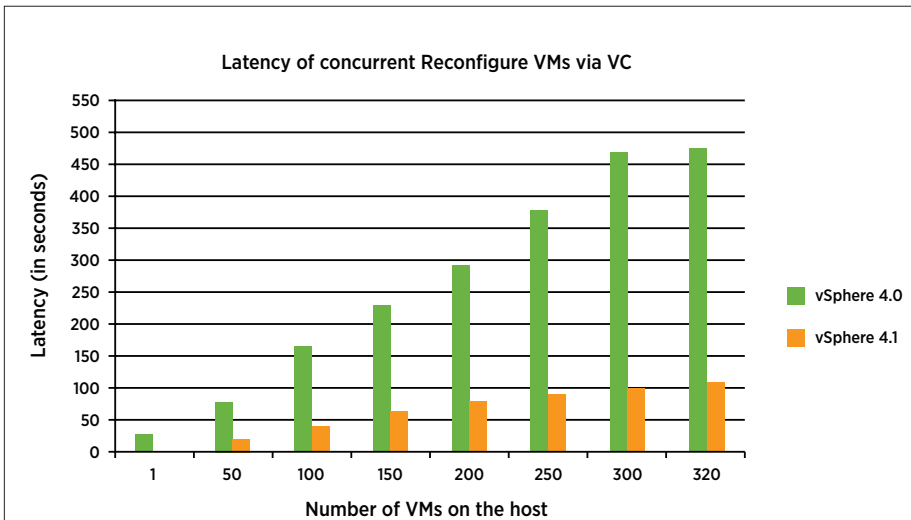


Figure 9. Latency of concurrent reconfigure VMs via vCenter

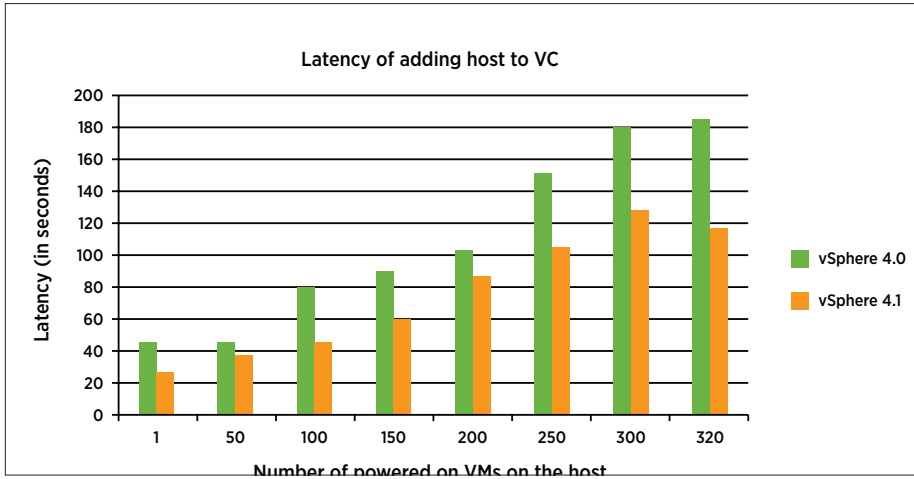


Figure 10. Latency of adding host to vCenter

Storage Performance Monitoring

New storage performance monitoring statistics for NFS are now available in vSphere 4.1. These 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 Metrics

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 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/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/pdf/vsphere4/r41/vsp_41_config_max.pdf

