



What's New in VMware vSphere™ 5.0 – Storage

TECHNICAL MARKETING DOCUMENTATION
V 1.0 / UPDATED MAY 2011

Table of Contents

Introduction	3
vSphere VMFS-5	3
vSphere Storage DRS.....	4
Datastore Clusters.....	4
Placement Recommendations.....	5
Affinity Rules and Maintenance Mode	6
vSphere Storage APIs – Storage Awareness.....	7
Profile-Driven Storage	7
Virtual Machine Compliance Status	8
Fibre Channel over Ethernet Software Initiator	9
iSCSI Initiator Enhancements	10
vSphere Storage I/O Control	11
vSphere Storage APIs – Array Integration.....	11
vSphere Thin Provisioning.....	11
Hardware Acceleration for NAS	13
T10 Compliancy	14
vSphere Storage vMotion	14
Conclusion	15

Introduction

VMware vSphere™ 5.0 (“vSphere”) brings many new capabilities to extend the benefits of vSphere 4.1. These new features and enhancements to core capabilities in vSphere provide more performance optimization and easier provisioning, monitoring and troubleshooting. This paper focuses on the storage-specific features and enhancements that are available in vSphere 5.0 and provides an overview of how they optimize storage utilization, ease monitoring, and increase operational efficiency. Wherever possible, we will also provide use cases and requirements that might apply to these new functions.

The topics to be covered in this paper are:

- vSphere® VMFS-5
- vSphere® Storage Distributed Resource Scheduler (Storage DRS)
- vSphere® Storage APIs - Storage Awareness
- Profile-Driven Storage
- Fibre Channel over Ethernet
- vSphere® Storage I/O Control (Storage I/O Control)
- vSphere® Storage APIs – Array Integration
- vSphere® Storage vMotion® (Storage vMotion)

This will provide a technical overview of new capabilities and enhancements as well as links to additional information about each of these new storage features.

vSphere VMFS-5

vSphere 5.0 introduces a new version of vSphere VMFS (VMFS), VMFS-5, which contains many important architectural changes that enable greater scalability and performance while reducing complexity. Although numerous fundamental changes have been made, the following enhancements are significant from operational and architectural aspects:

- 64TB device support
- Unified block size
- Improved subblock mechanism

vSphere 5.0 facilitates a nondisruptive upgrade from VMFS-3 to VMFS-5, ensuring consistency across virtual infrastructures. The unified block size, 1MB, enables easier deployments and reduced complexity from an architectural and operational aspect while maintaining the scalability and the flexibility that were previously found only with large block sizes. It should be noted that volumes that are upgraded from VMFS-3 to VMFS-5 will retain their original block size, because modifying the block size would require a reformat of the volume. To enable greater scalability and reduce storage overhead associated with small files, various enhancements have been made to VMFS-5. These enhancements include optimized subblock sizes and the allocation of these blocks. They have resulted in support for large volumes (64TB on a single extent) and higher virtual machine density while reducing the overhead associated with small files. VMFS-5 is capable of allocating 30,000 subblocks of 8KB for files such as log files and virtual machine metadata (.vmx files). For files with a size smaller than 1KB, “small file blocks” are used. The following table depicts the most significant architectural changes for VMFS-5 in comparison to VMFS-3.

FEATURE	VMFS-3	VMFS-5
Support for larger single-extent VMFS volumes up to 64TB	No	Yes
Support for pass-through RDMs larger than 2TB	No	Yes
Subblock for space efficiency	Yes (64KB, maximum of approx. 3K)	Yes (8KB maximum of approx. 30K)
Small-file support	No	Yes (1KB)
Unified block size	No	Yes

Table 1. VMFS-5 to VMFS-3 Comparison

vSphere Storage DRS

Virtual machine provisioning has always imposed operational challenges. Monitoring datastore capacity and I/O load has proven to be very difficult and as a result is often neglected. During the provisioning process of virtual machines and virtual disks, datastores are often randomly selected, leading to hot spots and over- or underutilized datastores. Storage DRS is a new feature introduced in vSphere 5.0. It provides smart virtual machine placement and load balancing mechanisms based on I/O and space capacity. Storage DRS will help decrease operational effort associated with the provisioning of virtual machines and the monitoring of the storage environment.

Datastore Clusters

With vSphere 5.0, a new VMware vCenter™ object called a datastore cluster is introduced. Datastore clusters form the basis of Storage DRS. A datastore cluster is a collection of datastores aggregated into a single unit of consumption from an administrator's perspective. When a datastore cluster is created, Storage DRS can manage the storage resources comparably to how Storage DRS manages compute resources in a cluster. As with a cluster of hosts, a datastore cluster is used to aggregate storage resources, enabling smart and rapid placement of new virtual machines and virtual disk drives as well as load balancing of existing workloads. Figure 1 depicts a datastore cluster of 12TB formed by four 3TB datastores.

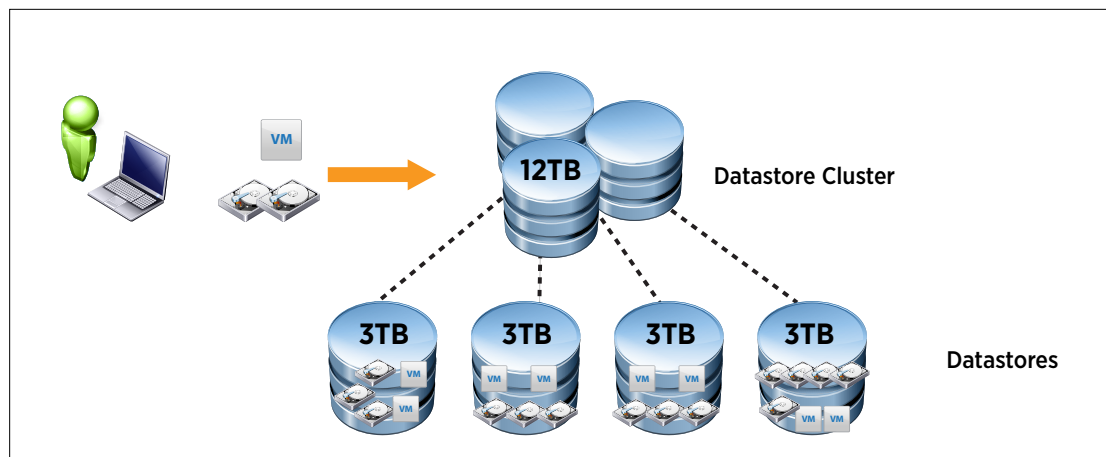


Figure 1. Datastore Cluster

Placement Recommendations

Storage DRS provides initial placement and ongoing balancing recommendations, helping vSphere administrators make placement decisions based on space and I/O capacity. During the provisioning of a virtual machine, a datastore cluster can be selected as the target destination for this virtual machine or virtual disk, after which a recommendation for initial placement is made based on space and I/O capacity. Initial placement in a manual provisioning process has proven to be very complex in most environments. As a result, crucial provisioning factors such as current space utilization and I/O load are often ignored. Storage DRS ensures that initial placement recommendations are made in accordance with space constraints and with respect to the goals of space and I/O load balancing. These goals aim to minimize the risk of storage I/O bottlenecks and the performance impact on virtual machines. Figure 2 depicts these definable thresholds.

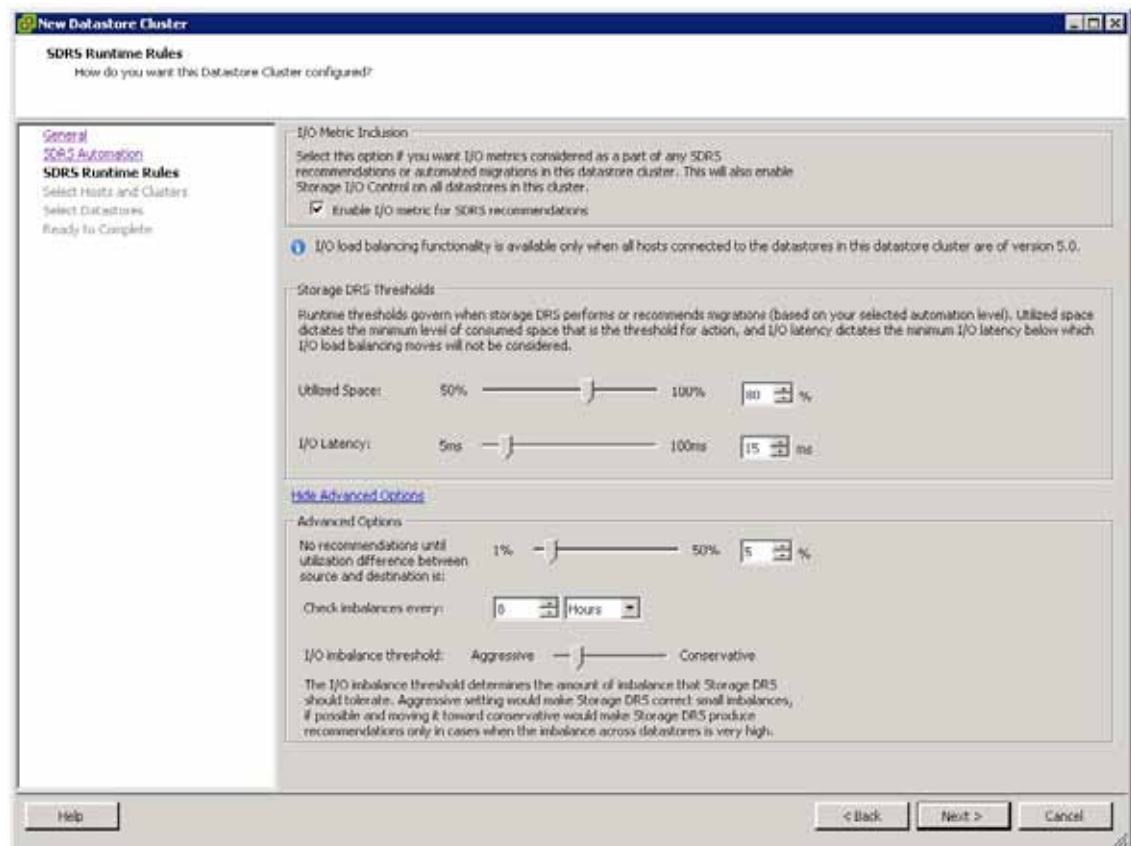


Figure 2. Defining Thresholds for Storage DRS

Ongoing balancing recommendations are made when one or more datastores in a datastore cluster exceed the user-configurable space utilization or I/O latency thresholds. These thresholds are typically defined during the configuration of the datastore cluster. Storage DRS applies the datastore utilization reporting mechanism of VMware vCenter Server, to make recommendations whenever the configured utilized space threshold is exceeded. I/O load is evaluated by default every 8 hours. When the configured maximum space utilization or the I/O latency threshold (15ms by default) is exceeded, Storage DRS will calculate all possible moves, to balance the load accordingly while considering the cost and the benefit of the migration.

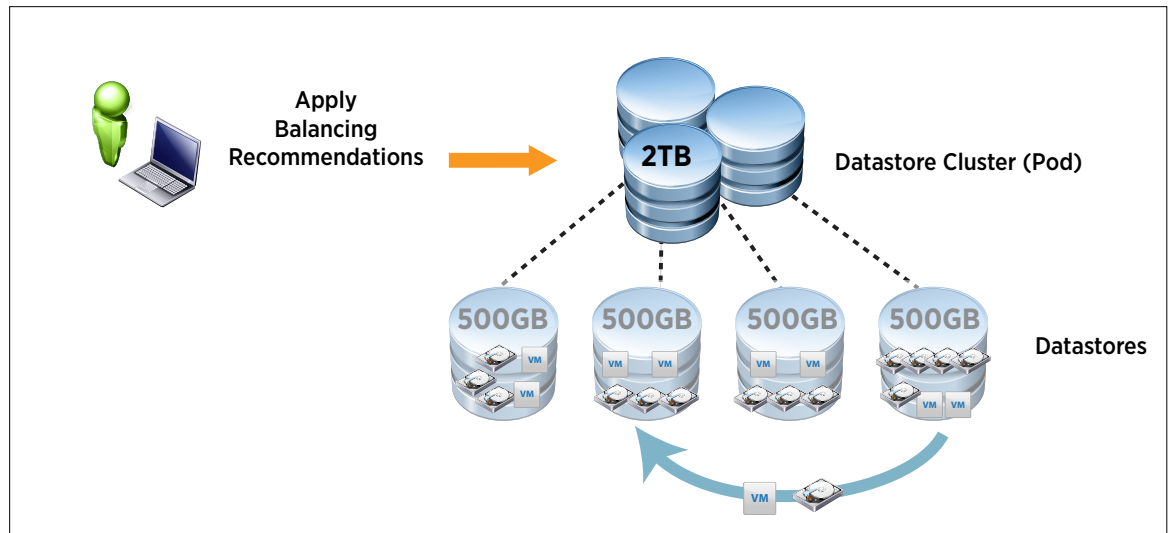


Figure 3. Defining Thresholds for Storage DRS

Affinity Rules and Maintenance Mode

Storage DRS affinity rules enable control over which virtual disks should or should not be placed on the same datastore within a datastore cluster. By default, a virtual machine's virtual disks are kept together on the same datastore. Storage DRS offers three types of affinity rules:

- **VMDK Anti-Affinity**
Virtual disks of a virtual machine with multiple virtual disks are placed on different datastores.
- **VMDK Affinity**
Virtual disks are kept together on the same datastore.
- **VM Anti-Affinity**
Two specified virtual machines, including associated disks, are placed on different datastores.

In addition, Storage DRS offers Datastore Maintenance Mode, which automatically evacuates all virtual machines and virtual disk drives from the selected datastore to the remaining datastores in the datastore cluster.

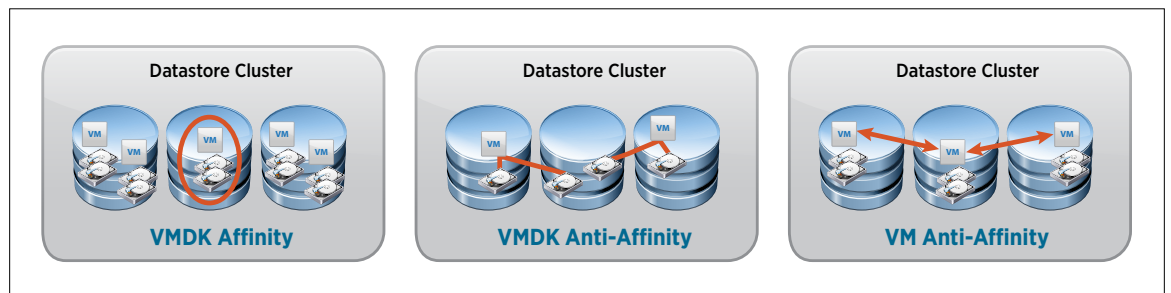


Figure 4. Affinity Rules

Storage DRS works with both VMFS- and NFS-based datastores. But mixing VMFS and NFS datastores in a single datastore cluster is currently not supported. More details about this feature, and implementation tips and design considerations, are described in the documentation: *Resource Management Guide*, chapter 11.

vSphere Storage APIs – Storage Awareness

vSphere® Storage APIs - Storage Awareness is a new set of APIs that will enable VMware vCenter Server to detect the capabilities of the storage array LUNs/datastores, making it much easier to select the appropriate disk for virtual machine placement or the creation of datastore clusters. vSphere Storage APIs – Storage Awareness can also provide help in facilitating, for example, the troubleshooting process or conversations between vSphere and storage administrators. Storage capabilities, such as RAID level, thin or thick provisioned, replication state and much more, can now be made visible within VMware vCenter Server either via “system-defined capabilities,” which are per-datastore descriptors, or via attributes that are exposed via Storage Views and SMS API. vSphere Storage APIs – Storage Awareness aims to eliminate the need for maintaining massive spreadsheets detailing the storage capabilities of each LUN that are needed to guarantee the correct SLA to virtual machines.

Profile-Driven Storage

Managing datastores and matching the SLA requirements of virtual machines with the appropriate datastore can be challenging and cumbersome tasks. vSphere 5.0 introduces Profile-Driven Storage, which will enable rapid and intelligent placement of virtual machines based on SLA, availability, performance or other requirements and provided storage capabilities.

Using Profile-Driven Storage, various storage characteristics, typically defined as a tier, can be requested in a virtual machine storage profile. These profiles are used during provisioning, cloning and Storage vMotion to ensure that only those datastores or datastore clusters that are compliant with the virtual machine storage profile are made available. The virtual machine storage profile can also help select a similar type of datastores when creating a Storage DRS datastore cluster. Profile-Driven Storage will reduce the amount of manual administration required for virtual machine placement while improving virtual machine SLA storage compliance.

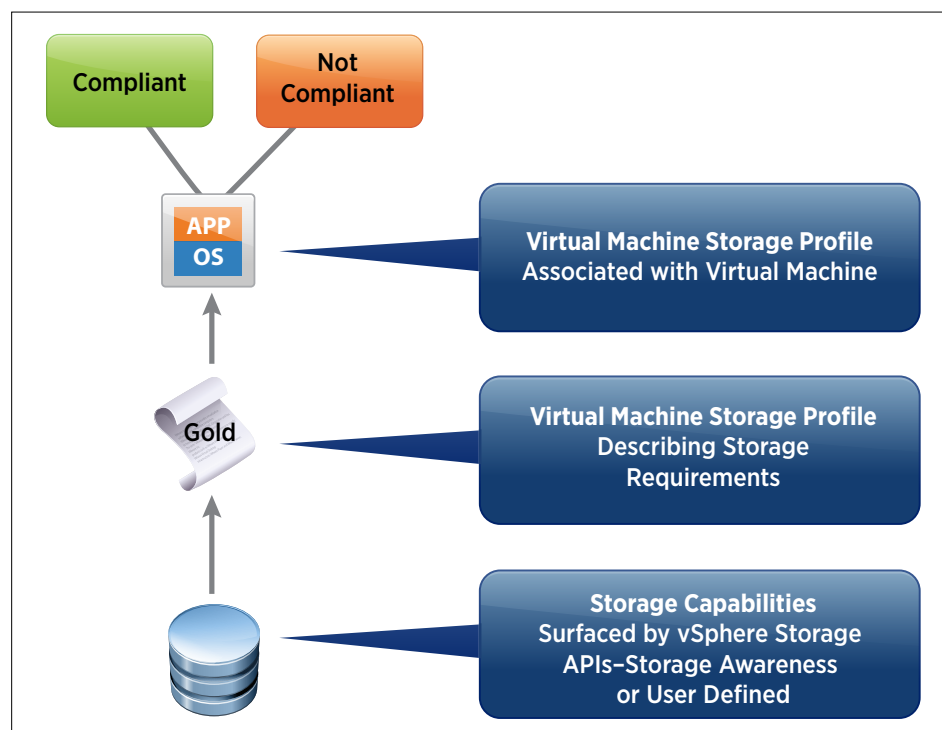


Figure 5. Virtual Machine Storage Profile

Profile-Driven Storage delivers these benefits by taking advantage of the following:

- Full integration with vSphere Storage APIs – Storage Awareness, enabling usage of storage characterization supplied by storage vendors.
- Support for NFS, iSCSI and Fibre Channel (FC) storage, and all storage arrays on the HCL.
- Enabling the vSphere administrator to tag storage based on customer- or business-specific descriptions.
- Using storage characterizations and/or administrator-defined descriptions to create virtual machine placement rules in the form of storage profiles.
- Providing an easy means to check a virtual machine's compliance with these rules. This ensures that the virtual machine is not deployed or migrated to an incorrect type of storage without the administrator's being informed about it.

Virtual Machine Compliance Status

As mentioned, virtual machine storage profiles enable you to check compliance of all virtual machines and the associated virtual disks in a single pane of glass (Figure 6). But they also enable you to see this on a per-virtual machine basis, ensuring that even “virtual machine administrators” who don't necessarily have access to all the different layers can validate compliance on the virtual machine summary's tab (Figure 7).

Name	Compliance Status	Last Checked	Compliance Failure
zimbra-srv-001			
vm home	Compliant	5/2/2011 8:32:04 AM	
Hard disk 1	Compliant	5/2/2011 8:32:04 AM	

Figure 6. Virtual Machine Storage Profiles Single Pane of Glass for Compliance

VM Storage Profiles	
VM Storage Profiles:	Zimbra-Servers
Profiles Compliance:	Compliant (5/2/2011 8:32:04 AM)

Figure 7. Per-Virtual Machine Profile Compliance

As a result, managing storage tiers, provisioning, migrating, cloning virtual machines and correct virtual machine placement in vSphere deployments have become more efficient and user friendly. Profile-Driven Storage removes the need for maintaining complex and tedious spreadsheets and validating compliance manually during every migration or creation of a virtual machine or virtual disk.

Fibre Channel over Ethernet Software Initiator

vSphere 4.0 introduced support for hardware Fibre Channel over Ethernet (FCoE) adaptors. vSphere 5.0 will take the adoption of FCoE to a new level by introducing a software FCoE adaptor. The software FCoE adaptor will require a network adaptor that can support partial FCoE offload capabilities before it can be enabled (Figure 8).

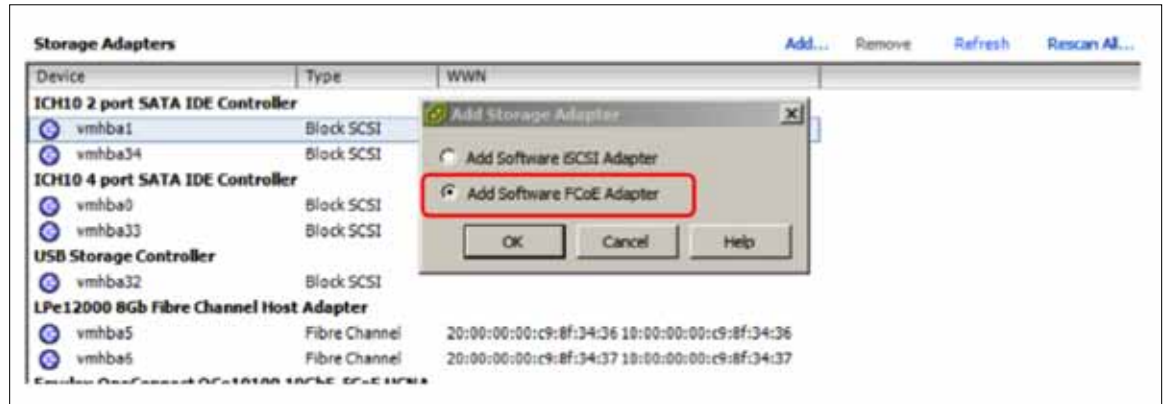


Figure 8. Adding a Software FCoE Adaptor

After the software FCoE adaptor has been enabled, a new adaptor is created. After the creation of the adaptor, devices that are presented to this particular adapter will be recognized (Figure 9) by vSphere, enabling the creation of a new datastore. The software FCoE adaptor will enable you to use FCoE capabilities without the need to resort to specific FCoE hardware adapters, creating a more agile and flexible infrastructure.

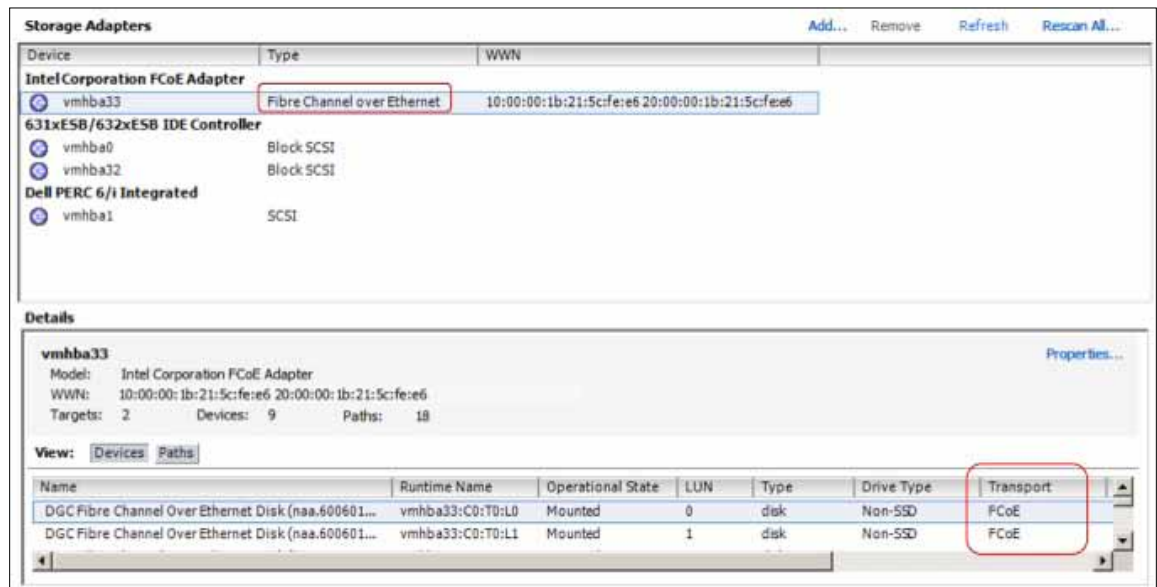


Figure 9. FCoE Adaptor

iSCSI Initiator Enhancements

With vSphere 4.0, a new iSCSI initiator was released that was more efficient and capable of driving more IOPS than ever before. The entire iSCSI software initiator stack was rewritten and optimized for performance. vSphere 4.0 enabled multiple VMkernel network adaptors to connect to the iSCSI array. Configuration was possible only through the command line and was considered to be a difficult task. vSphere 5.0 introduces the ability to fully configure iSCSI through the vSphere client, as depicted in Figure 10.

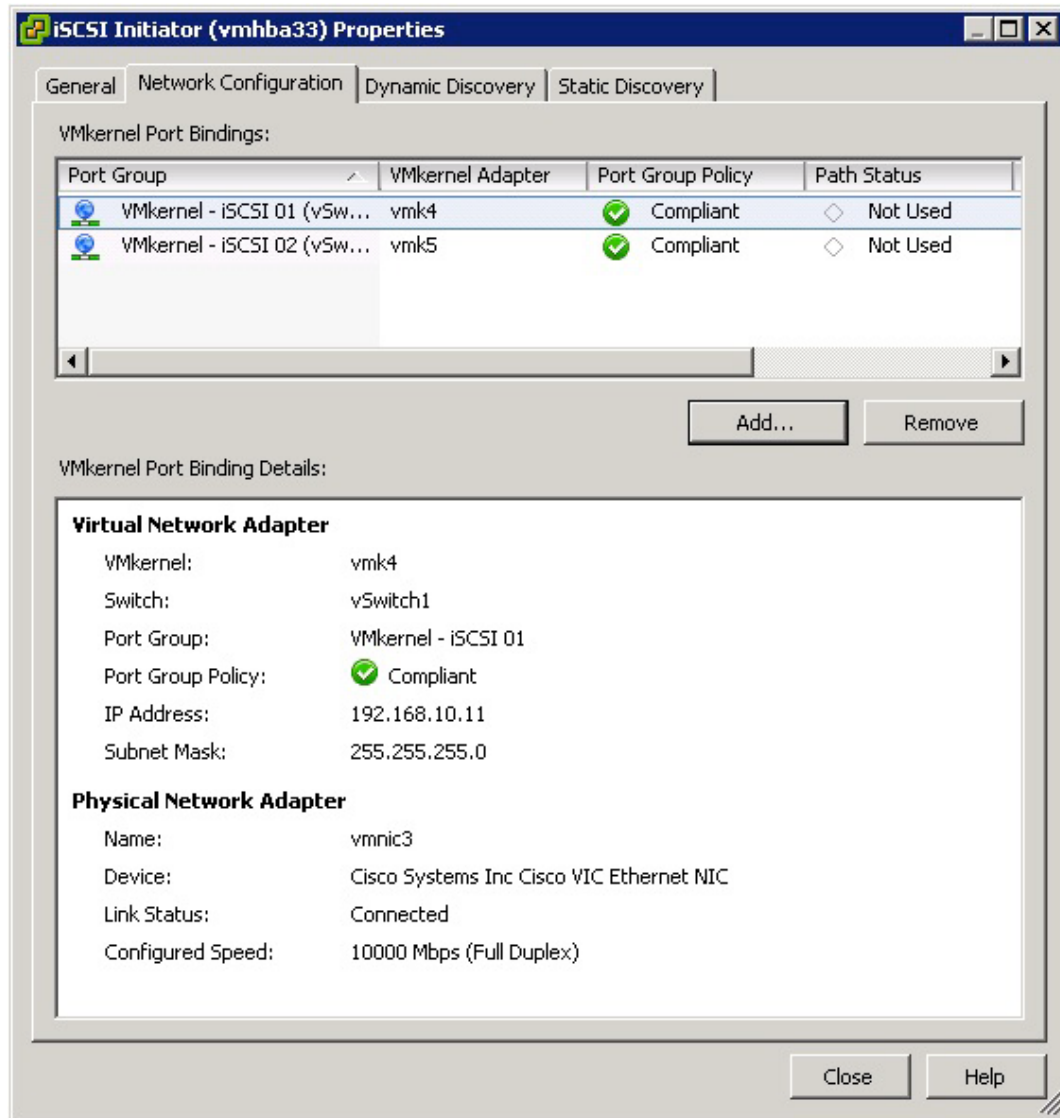


Figure 10. iSCSI Initiator

vSphere Storage I/O Control

vSphere Storage I/O Control was initially introduced in vSphere 4.1 to provide I/O prioritization of virtual machines running on a cluster of VMware® ESX® servers that had access to a shared, iSCSI or FC, storage pool. It extended the familiar constructs of shares and limits that existed for CPU and memory, to address storage utilization through a dynamic allocation of I/O queue slots across a cluster of VMware ESX servers.

vSphere 5.0 extends Storage I/O Control to provide clusterwide I/O shares and limits for NFS datastores. This means that no single virtual machine should be able to create a bottleneck in any environment regardless of the type of shared storage used. Storage I/O Control automatically throttles a virtual machine that is consuming a disparate amount of I/O bandwidth when the configured latency threshold has been exceeded. This enables other virtual machines using the same datastore to receive their fair share of I/O. Storage DRS and Storage I/O Control are the perfect partners for preventing deprecation of service-level agreements while providing long-term and short-term I/O distribution fairness.

vSphere Storage APIs – Array Integration

vSphere Storage APIs – Array Integration (VAAI) were first introduced with vSphere 4.1, enabling offload capabilities support for three primitives:

1. Full copy, enabling the storage array to make full copies of data within the array
2. Block zeroing, enabling the array to zero out large numbers of blocks
3. Hardware-assisted locking, providing an alternative mechanism to protect VMFS metadata

With vSphere 5.0, support for the VAAI primitives has been enhanced and additional primitives have been introduced:

- vSphere® Thin Provisioning (Thin Provisioning), enabling the reclamation of unused space and monitoring of space usage for thin-provisioned LUNs
- Hardware acceleration for NAS
- SCSI standardization by T10 compliancy for full copy, block zeroing and hardware-assisted locking

vSphere Thin Provisioning

vSphere 5.0 introduces multiple VAAI enhancements for environments using array-based thin-provisioning capabilities. Historically, the two major challenges of thin-provisioned LUNs have been the reclamation of dead space and the challenges surrounding the monitoring of space usage. VAAI Thin Provisioning introduces the following:

- Dead Space Reclamation informs the array about the datastore space that is freed when files are deleted or removed from the datastore by Storage vMotion. The array can then reclaim the freed blocks of space.
- Out-of-Space Conditions monitors the space usage on thin-provisioned LUNs to prevent running out of physical space. A new, advanced warning has been added to vSphere 5.0 for a thin-provisioned out-of-space condition.

Dead Space Reclamation offers the ability to reclaim blocks of a thin-provisioned LUN on the array when a virtual disk is deleted or migrated to a different datastore by, for example, Storage DRS (Figure 10). Historically, when virtual machines were migrated from a datastore, the blocks used by the virtual machine prior to the migration were still being reported as “in use” by the array. This meant that usage statistics from the storage array might have been misleading and expensive disk space possibly might have been wasted. With this new VAAI primitive, the storage device will be informed that the blocks are no longer used, resulting in better reporting of disk space consumption and allowing for the reclamation of these unused blocks.

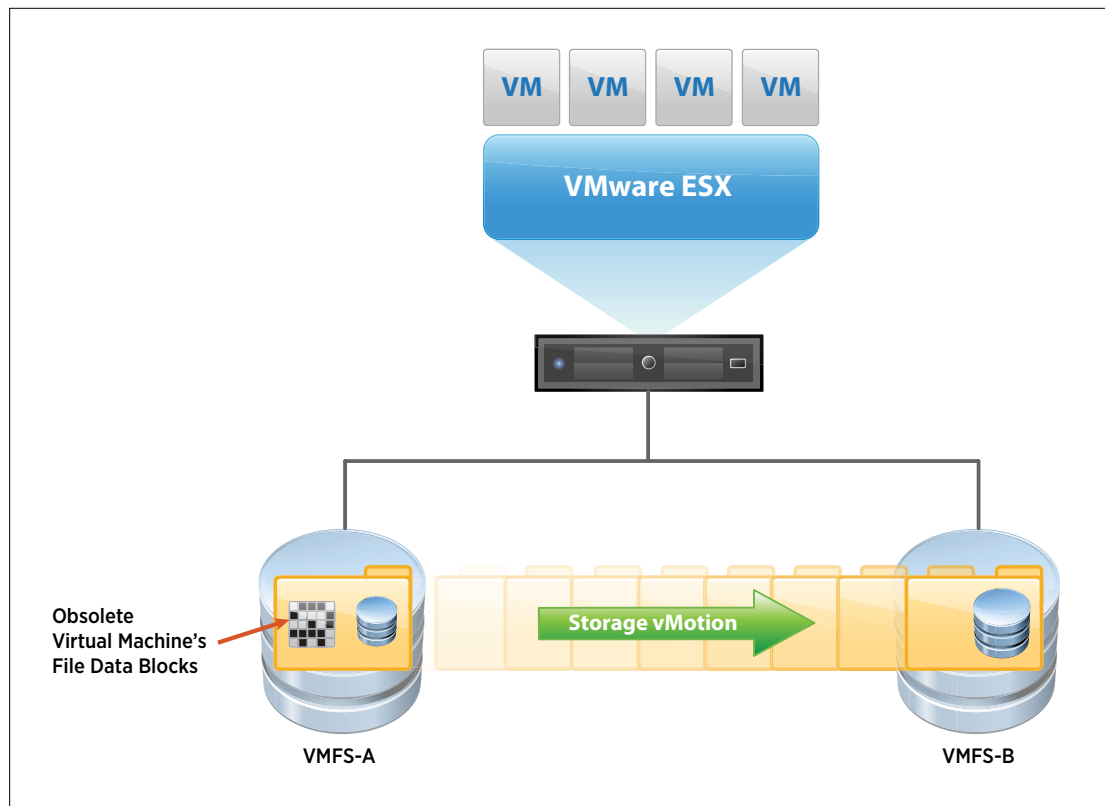


Figure 11. Reclamation of Dead Space

An “out-of-space condition” is the nightmare of every storage and virtualization administrator in array-based thin-provisioned environments. “Storage oversubscription” in thin-provisioned environments can lead to catastrophic scenarios when an out-of-space condition is encountered. vSphere 5.0 mitigates these problems and simplifies storage management through the addition of advanced warnings (Figure 12) and errors when thresholds are reached for thin-provisioned datastores. Enhancements include mechanisms that temporarily pause a virtual machine when disk space is exhausted, enabling the allocation of additional space to the datastore, or the migration of an existing virtual machine, without resulting in the failure of the virtual machine.

Summary Virtual Machines Resource Allocation Performance Configuration Local Users & Groups Events Permissions					
Show all entries ▾					
Description	Type	Date Time	Task	Description, Type or Target	
Space utilization on thin-provisioned device naa:6090a028b0c5d7c772cc64962f014006 exceeded configured threshold. Affected datastores (if any): cormac-eql-test-4mar2011.	warning	04/03/2011 09:12:52		cormac-eql-test-4mar2011	

Figure 12. Out-of-Space Condition

Hardware Acceleration for NAS

Hardware Acceleration for NAS will enable faster provisioning and the use of thick virtual disks through two newly introduced VAAI primitives:

- Full File Clone – Similar to Full Copy. Enables virtual disks to be cloned by the NAS device
- Reserve Space – Enables creation of thick virtual disk files on NAS

Full File Clone enables fast provisioning of virtual machines on NAS devices, similar to the experience “Full Copy” provided for iSCSI/FC arrays.

Prior to vSphere 5.0, a virtual disk was created as a thin-provisioned disk, not even enabling the creation of a thick disk. Starting with vSphere 5.0, VAAI NAS extensions enable NAS vendors to reserve space for an entire virtual disk (Figure 13). This enables the creation of thick disks on NFS datastores.

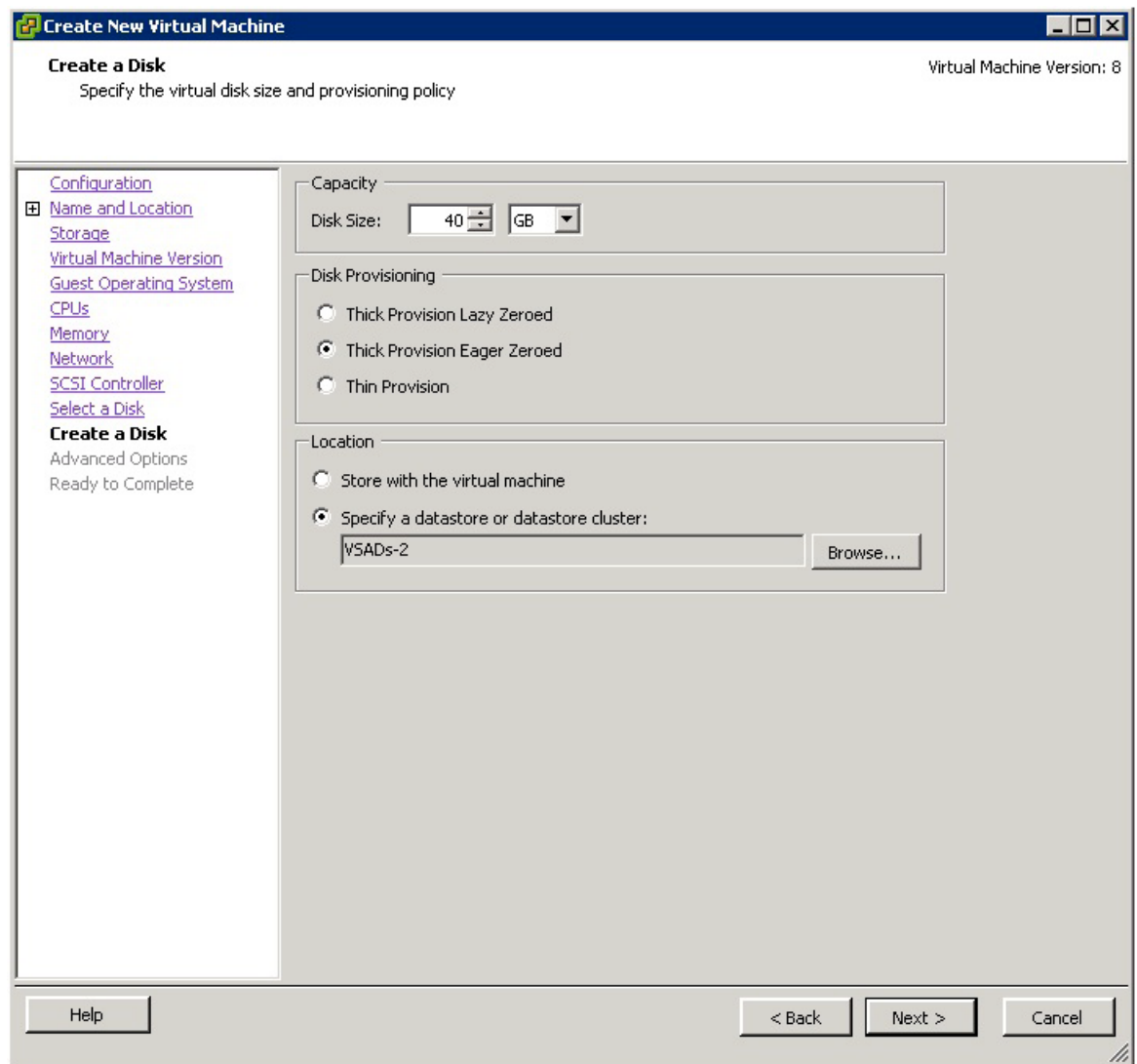


Figure 13. Creation of Thick Disks

NAS VAAI plug-ins are not shipped with vSphere 5.0. They are developed and distributed by storage vendors but are signed by the VMware certification program to guarantee quality.

T10 Compliance

vSphere 4.1 introduced T10 compliance for block zeroing, enabling vendors to utilize the T10 standards with the default, shipped plug-in. vSphere 5.0 introduces enhanced support for T10, enabling the usage of VAAI primitives without the need to install a plug-in, as well as enabling support for many additional storage devices.

vSphere Storage vMotion

vSphere® Storage vMotion® (Storage vMotion) enables live migration for running virtual machine disk files from one storage location to another with no downtime or service disruption. With vSphere 5.0, multiple enhancements have been introduced to increase efficiency of the Storage vMotion process, to improve overall performance, and for enhanced supportability. Storage vMotion in vSphere 5.0 now also supports the migration of virtual machines with a vSphere snapshot and the migration of linked clones.

The enhancements to Storage vMotion include a new and more efficient migration process through the use of a new feature called Mirror Mode. Mirror Mode enables a single-pass block copy of the source disk to the destination disk by mirroring I/Os of copied blocks. Not only has the efficiency of Storage vMotion increased, but also the migration time predictability, making it easier to plan migrations and reducing the elapsed time per migration.

The mirror driver is enabled on a per-virtual machine basis and resides within the VMkernel. When the guest OS of the virtual machine that is undergoing the process using Storage vMotion initiates a write to an already copied block, the mirror driver (Figure 14) will synchronously mirror this write and wait for both acknowledgements before communicating this to the guest OS.

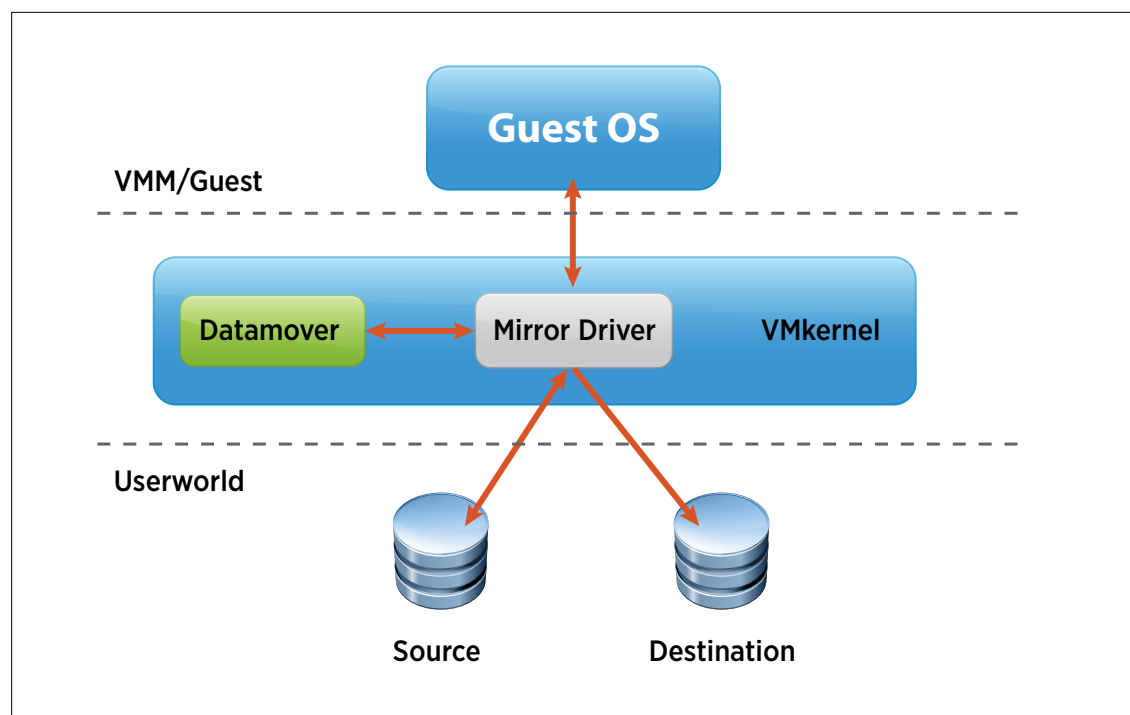


Figure 14. Data Flow with Storage vMotion Mirror Mode

Conclusion

vSphere 5.0 adds many new storage features to an already rich set of capabilities supported in vSphere 4.1. It reduces complexity while providing greater scalability. Virtual machine provisioning historically has imposed operational challenges. Monitoring and manually balancing workloads, or provisioning virtual machines based on I/O capacity and space utilization, has proven to be very difficult and has often been neglected, leading to hot spots and over- or underutilized datastores. Storage DRS is a new feature introduced in vSphere 5.0 that provides smart virtual machine placement and load balancing mechanisms based on I/O and space capacity. The result is a decrease in the operational effort associated with the provisioning of virtual machines and monitoring of the storage environment.

About the Author

Duncan Epping is principal architect in the Technical Marketing group at VMware and is focused on storage and VMware ESXi. Previously, he worked at Oracle and with several consultancy companies, where he had more than 10 years of experience in the design and development of infrastructures and deployment best practices. He was among the first VMware certified design experts (VCDX 007). He is the co-author of many books, including *VMware vSphere 4.1 HA and DRS Technical Deepdive*, the *VMware vSphere 4.0 Quick Start Guide*, *Foundation for Cloud Computing with VMware vSphere 4* and his most recent publication, *Foundation for Cloud Computing with VMware vCloud Director*.

