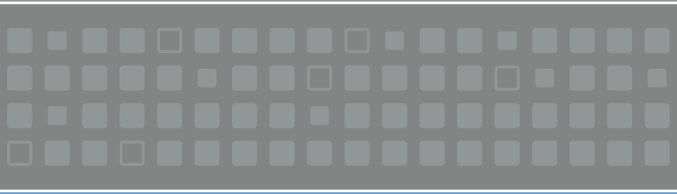
WHITE PAPER

What Is New in VMware vSphere™ 4: Storage



vmware®

What Is New in VMware vSphere™ 4: Storage

With the launch of vSphere 4, VMware introduces a new version of the hypervisor, ESX™ 4.0 and the management interface, vCenter. Some key benefits offered by this release are greater resource efficiency, management control and freedom of choice for virtual datacenter storage resources and connectivity options. VMware has added many new storage management capabilities that increase disk utilization rate, provide greater control of these storage resources and enable increased options for datastore protocol support and format of virtual disk format. These storage features provide new levels of flexibility and improved management agility for virtualization environments that run vSphere.

This document provides information about the following storage enhancements in vSphere sorted under the following categories:

Efficiency
- Virtual Disk Thin Provisioning
- Improved iSCSI Software Initiator Efficiency

Control
- New vCenter Storage Capabilities
- Dynamic Expansion of VMFS Volumes

Choice
- Enhanced Storage VMotion
- Pluggable Storage Architecture
- PV SCSI and Direct Path I/O

Increased Storage Efficiency

One big challenge facing many IT managers is to increase the level of resource utilization needed to run the datacenter. Specifically, the amount of wasted storage resources can be greatly reduced with the use of thin provisioning, wherein storage waste measures the amount of allocated storage that is not used. This ratio is also referred to as the storage disk utilization rate. In most datacenters this number is often well below 50 percent utilization. To address this issue, many storage vendors have released support for thin provisioning in their storage arrays. Another means to increase storage efficiency is to reduce the CPU overhead it takes to drive an I/O between a server and the storage. Reductions in this overhead can greatly increase the throughput of a given system.

Virtual Disk Thin Provisioning

VMware vSphere introduces full support for the option to create thin provisioned virtual disks when deploying or migrating Virtual Machines. vCenter has also been updated with new management screens and capabilities such as raising alerts, alarms and improved datastore utilization reports, to enable management of over provisioned datastores.

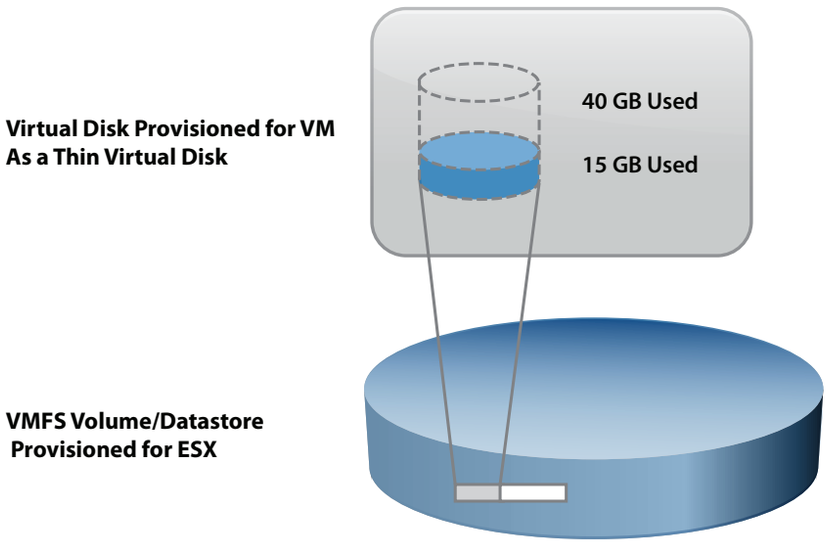
Virtual disk thin provisioning increases the efficiency of storage utilization for virtualization environments by using only the amount of underlying storage resources needed for that virtual disk. Upon initial allocation of the virtual disk, the storage requirements will be given 1 MB of space in the datastore. As that space is filled, additional 1 MB chunks of storage will be allocated for the virtual disk so that the underlying storage demand will grow as its size increases.

This capability is best understood when compared to a common practice in the airline industry: overbooking of flights. Airlines found that many people who made reservations did not show up for their flights. Therefore the airlines began to book more passengers than they had seats available on the flight. In the storage industry, a request for disks larger than what is needed for the application creates a similar problem. Thin provisioning of storage allows an administrator to over allocate the amount of storage resources currently available. Similar to how airlines carefully optimize the use of seats on their flights, storage administrators can also increase the utilization rates, and reduce the cost of their storage resources by over allocating them and adding more when the demand usage grows to meet the allocation.

vCenter now fully supports virtual disk thin provisioning within the virtual environment— such that a thin virtual disk is assigned only the amount of space in the datastore needed for that virtual disk. In the past, thin provisioning has been the default format for virtual disks created on NFS datastores in Virtual Infrastructure 3, and has been available through the command line for block-based datastores. However. VMware has integrated management of virtual disk thin provisioning and now fully supports this format for all virtual disks with the release of vSphere.

It is important to note that virtual disk thin provisioning should not be confused as the same thin provisioning that an array vendor might offer. In fact, with vSphere you now have the capability of doing thin provisioning at the datastore level in addition to thin provisioning at the storage array. (See Figure 1 for more details)

Figure 1. Virtual disk thin provisioning at the datastore level



In addition to using thin provisioning to save space for new disk allocations, many existing virtualization environments can benefit from the use of Storage VMotion™ as a means to reduce wasted (allocated yet unused) storage space. By changing virtual disk formats from thick to thin in conjunction with a migration of the VM home from one datastore to another while the VM is up and running, some customers have found they are able to regain upwards of 40 percent of their storage space.

Improved iSCSI Software Initiator Efficiency

With vSphere, VMware has released a new iSCSI software initiator that is far more efficient in its use of the ESX CPU cycles to drive storage I/O. The entire iSCSI software initiator stack was rewritten and tuned to optimize cache affinity, enhance the VMkernel TCP/IP stack, and make better use of internal locks. The new vSphere iSCSI SW initiator has been optimized for virtualization I/O and provides significant improvements in CPU efficiency and throughput when compared to the earlier 3.5 version.

The configuration steps for an iSCSI datastore have also been improved to make it easier to provision iSCSI storage for vSphere. The new iSCSI stack no longer requires a Service Console connection to communicate with an iSCSI target. The security protection has also been expanded with support for bi-directional CHAP authentication. This adds another level of security by allowing the initiator to authenticate the target. Global configuration settings made in the general tab will also propagate down to all targets. However, these can be overridden by settings on a per target basis, so users can configure unique parameters for each array discovered by the initiator.

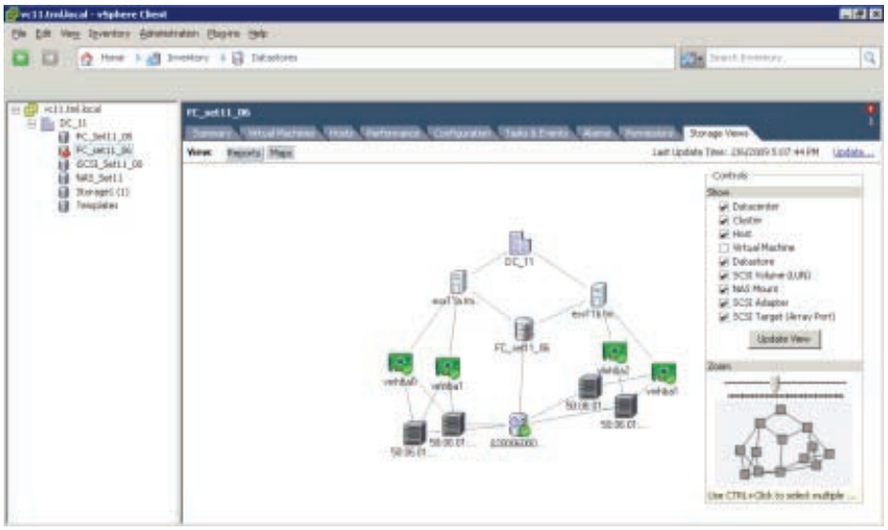
Improved Storage Resources Control

As the scope of storage resources have increased significantly with large deployments of virtualization environments, so has the need for greater automation and control of these resources. In the vSphere release, vCenter has been enhanced with several new storage specific capabilities to help the virtual administrator manage these environments with a higher degree of control. These enhancements provide administrators with proactive alerts and alarms to address issues before they interrupt the availability of applications running on those resources. vCenter allows setting permissions and quota limits on datastores, as well as per VM.

New vCenter Storage Capabilities

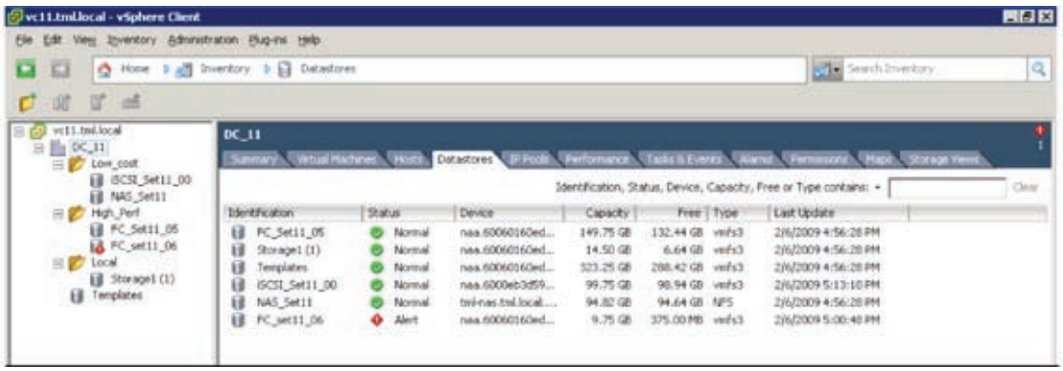
To help manage storage space in vSphere, each VM and ESX in the inventory now has its own storage tab that shows information about storage resources for those objects. The vCenter inventory also has a new view for listing datastore details. The datastore is now managed as an object within vCenter in a manner that allows the vSphere administrator to view, group, and set permissions for each datastore. vSphere also provides a detailed view of all the components in the storage layout (See Figure 2). These topology maps provide key information to administrators about which paths are available, as well as the grouping of objects sharing storage resources.

Figure 2. View of all components in the storage layout



In the highest level of visibility, users can arrange the list of datastores higher-archaically; grouping datastores by department (as shown in Figure 2), or by class of services. Details for each datastore can be investigated to reveal which ESX servers are accessing it and what level of utilization is being achieved. A new topology map is available for these datastores as well.

Figure 3. vSphere provides increased visibility and control of storage resources



As each VM and ESX now has its own storage tab, a user can more quickly determine what storage resources are associated with those objects. New storage performance charts also give users the ability to set alerts and alarms for a VM, ESX, or datastore level.

New alerts and alarms are a key capability for to effectively managing datastores in which thin provisioned disks might create an over allocation of storage resources. The vSphere administrator can be alerted when thin provisioned disk growth creates a need to either increase the size of a datastore or create more free space in that datastore before the lack of space causes a problem. These alerts and alarms have a set of default values that can be customized to increase the awareness of actions needing to be taken. For example an administrator can set an email alert to be sent once a datastore fills to a certain percentage.

To augment the introduction of thin provisioning, there are a few additional capabilities that might best be grouped under the heading of dynamic storage growth. One is the ability to grow the size of the datastore if the over allocation of virtual disks or the need for new virtual disks is required. The other is the ability to free up space in a datastore without having to shut down any virtual machines, in the event the datastore size cannot grow. Storage VMotion is the solution that addresses the later of these two, and VMFS Volume Grow, a new feature introduced in vSphere, includes an option to address the first.

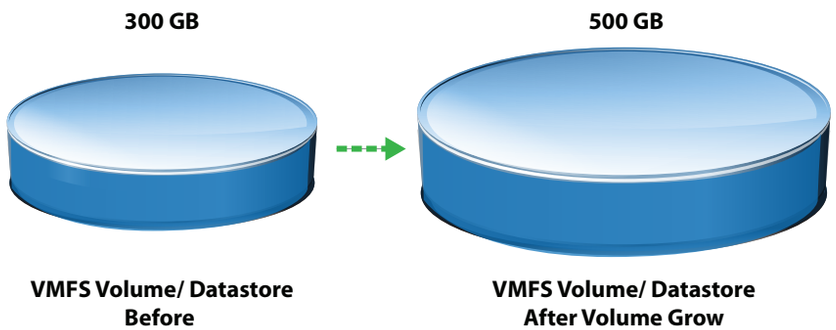
Dynamic Expansion of VMFS Volumes

VMFS Volume Grow offers a new way to increase the size of a datastore that resides on a VMFS volume. It complements the dynamic LUN expansion capability that exists in many storage array offerings today. If a LUN is increased in size, then VMFS Volume Grow enables the VMFS volume extent to dynamically increase in size as well.

Prior to vSphere, the only option for increasing the size of an existing VMFS volume was to add a new LUN via a process called spanning. Even if the LUN upon which the VMFS volume extent resided were to increase, the only option was to create a separate disk partition through the command line on that additional space made available and then add the new partition as you might add a second LUN to the VMFS Volume.

With VMFS Volume Grow, the process of increasing the size of the VMFS volume is integrated through the vCenter GUI as a choice within the VMFS volume properties screen. Provided that additional capacity on the existing extent is there, or has been recently been increased in capacity, the VMFS volume can now be expanded dynamically up to the 2 TB limit per LUN. For VMFS volumes that might already span multiple extents, the VMFS volume Grow can be used to grow each of those extents up to 2 TB as well.

Figure 4. VMFS Volume Grow enables the VMFS volume to dynamically increase in size along with the datastore



More Storage Resources Choices

As virtualization is deployed across an increasing set of applications in many datacenters, there is an increasing need to enable multiple storage protocols, storage from many vendors, and many different types of hardware adaptors. Several new features released in vSphere 4 provide additional options and choices to exist in these deployments. The following features enable increased options for choice in vSphere environments:

Enhanced Storage Vmotion

vSphere introduces several new capabilities to Storage VMotion. When Storage VMotion was introduced in VI3 release 3.5, it had a few limitations which vSphere 4 addresses. Storage VMotion is now fully integrated into vCenter and offers full support for migration across datastores of several protocol choices. Hence the enhanced Storage VMotion capabilities fall squarely in the realm of how vSphere provides an increased set of choices. First the ability to move a VM home from one datastore to another while changing the storage protocol in the process. The source datastore might be FC, iSCSI, or NFS, and the target datastore any of those three. Users can also change from a thick to thin virtual disk format or vice versa when a VM home is migrated with Storage VMotion.

The enhancements to Storage VMotion also include a more efficient migration process, as well as the removal of the 3.5 requirement for having 2 times the CPU and memory needed for the VM that is being migrated. Storage VMotion in vSphere leverages a change block tracking method that moves the VM home over to the new datastore in a few iterations instead of using the virtual disk snapshot method used in the 3.5 release.

Storage VMotion in vSphere provides more choices to support migrations from one storage vendor to another, choice of protocol, and choice of virtual disk format—all without having to take down the VM to relocate where it resides on the disk.

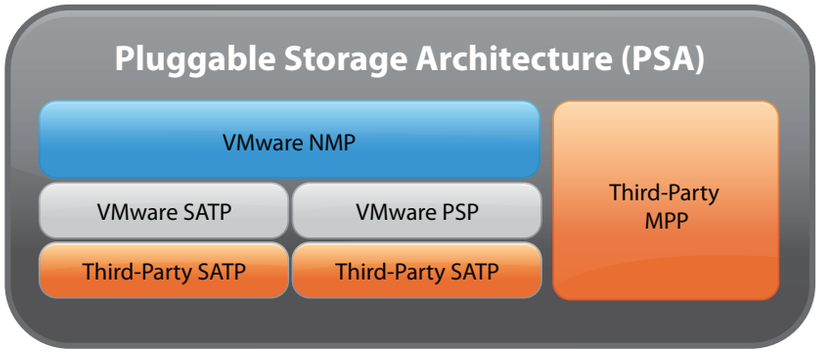
Using Storage VMotion to migrate VM from one datastore to another while also leveraging thin provisioning provides a means to reclaim what might be best described as wasted storage space. Many guest OS virtual disks are often allocated for more storage than they truly need for their current storage requirements. However, moving those disks from one datastore to another while also changing the virtual disk format from thick to thin in the same migration process enables that allocated but unused storage space to be reclaimed (as mentioned in the Thin Disk Virtual Provisioning section on page 3 of this white paper). Figure 5 shows how vCenter has the agility to leverage Storage VMotion with a change in the disk format in the process of moving the VM home. This is a key enabler of efficiency, control, and choice that comes from combining two new vSphere features.

Figure 5. vCenter leverages Storage VMotion to change disk format in the process of moving the VM home

Pluggable Storage Architecture (PSA)

In vSphere, leveraging third party storage vendor multi-path software capabilities has been introduced through a modular storage architecture that allows storage partners to write a plug-in for their specific capabilities. These modules can communicate with the intelligence running in the storage array to determine the best path selection, as well as leveraging parallel paths to increase performance and reliability of the I/O from the ESX to the storage array. By default, the native multipath driver (NMP) supplied by VMware will be used. It can be configured to support round-robin multipath as well. However, if storage vendor module is available, it can be configured to manage the connections between the ESX and the storage.

Figure 6. This pluggable storage architecture is enabled by the vStorage API for multi-Pathing



The basic version of those pluggable storage modules can provide Path Selection Plug-in (PSP) optimal path selection for high availability and reliability. The more advanced option for storage partners to integrate with PSA is for them to write a module that is Storage Array Type Plug-in (SATP) that can aggregate I/O across more than one channel for optimal I/O load balancing as well as high availability. This pluggable storage architecture is enabled by the vStorage API for multi-Pathing. It is one of several vStorage APIs that has been published to enable storage partners to integrate their vendor-specific capabilities to provide more optimal performance and to give joint customers more choice in how they configure the storage resources for virtualization. Another vStorage API also released in vSphere is the vStorage API for Data Protection.

vStorage APIs for Data Protection

This API takes the benefits of Consolidated Backup and makes them significantly easier to deploy, while adding several new features that deliver efficient, scalable backup, and restore of virtual machines.

Like Consolidated Backup, this API make it possible to offload backup processing from ESX servers, ensuring that you deliver the best consolidation ratios without disrupting applications and users. The vStorage API for Data Protection enables backup tools to directly connect the ESX servers and the virtual machines running on them without any additional software installation. They add the ability to enable backup tools to do efficient incremental, differential, and full-image backup and restore of virtual machines.

Paravirtualized SCSI
VMware Paravirtualized SCSI (PVSCSI) is a special purpose driver for high-performance storage adapters that offer greater throughput and lower CPU utilization for virtual machines. They are best suited for environments in which guest applications are very I/O intensive. VMware requires that you create a primary adapter for use with a disk that will host the system software (boot disk) and a separate PVSCSI adapter for the disk that will store user data, such as a database. The primary adapter will be the default for the guest operating system on the virtual machine. For example, a virtual machine with Microsoft Windows 2008 guest operating systems, LSI Logic is the default primary adapter. The PVSCSI driver is similar to vmxnet in that it is an enhanced and optimized special purpose driver for VM traffic and works with only certain Guest OS verision that currently include Windows Server 2003, 2008 and RHEl 5. It can also be shared by multiple VMs running on a single ESX, unlike the VMDirectPath I/O which will dedicate a single adaptor to a single VM.

VMDirectPath I/O for Storage
VMDirectPath I/O device access enhances CPU efficiency in handling workloads that require constant and frequent access to I/O devices. It enables virtual machines to directly access underlying hardware devices. This will map a single HBA to a single VM and not allow sharing of the HBA by more than a single Virtual Machine. However, other virtualization features, such as VMotion, hardware independence and sharing of physical I/O devices will not be available to the virtual machines using VMDirectPath I/O.

VMDirectPath I/O is experimentally supported for the following storage I/O devices:

- QLogic QLA25xx 8 Gb Fibre Channel adapters
- Emulex LPe12000 8 Gb Fibre Channel adapters
- LSI 3442e-R and 3801e (1068 chip based) 3 Gb SAS adapters

Conclusion

vSphere enables many new features for the management of virtualization storage resources. These capabilities provide vSphere administrators with new ways to increase efficiency for disk utilization and CPU cycles. With automated proactive alerts and alarms, vCenter provides more visibility and control of storage resource allocations, deployment, and monitoring. Also, new levels of reporting, topology maps, and datastore status and details screens provide an easy way to manage access and limit storage sprawl in VMware deployments. Storage VMotion has been enhanced to enable more choices for migrating VM homes from one type of storage to another. And finally, new choices exist for optimizing I/O traffic for I/O intensive Virtual Machines.

In short, VMware is now providing more storage options and features that provide greater control, choice, and efficiency to be leveraged in a vSphere environment. vSphere 4 enables administrators to do more with less—extending the benefits provided by virtualization to reach new levels of storage efficiency, management control, and additional choices in storage connectivity.

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

vmware®