VMware vSphere® VMFS

Technical Overview and Best Practices

A VMWARE TECHNICAL WHITE PAPER UPDATED FOR VMWARE vSPHERE 5.1
VERSION 3.0

vmware®

Table of Contents

Introduction

VMware vSphere® VMFS is a high-performance cluster file system (CFS) that enables virtualization to scale beyond the boundaries of a single system. Designed, constructed and optimized for the virtual infrastructure, VMFS increases resource utilization by providing multiple virtual machines with shared access to a consolidated pool of clustered storage. And it offers the foundation for virtualization spanning multiple servers, enabling services such as virtual machine snapshots, VMware vSphere Thin Provisioning, VMware vSphere vMotion®, VMware vSphere Distributed Resource Scheduler™ (vSphere DRS), VMware vSphere High Availability (vSphere HA), VMware vSphere Storage DRS™ and VMware vSphere Storage vMotion®.

VMFS reduces management overhead by providing a highly effective virtualization management layer that is especially suitable for large-scale enterprise datacenters. Administrators employing VMFS find it easy and straightforward to use, and they benefit from the greater efficiency and increased storage utilization offered by the use of shared resources.

This paper provides a technical overview of VMFS, including a discussion of features and their benefits. It highlights how VMFS capabilities enable greater scalability and decreased management overhead and it offers best practices and architectural considerations for deployment of VMFS.

Background

In today's IT environment, systems administrators must balance competing goals: finding ways to scale and consolidate their environment while decreasing the management overhead required to provision and monitor resources. Virtualization provides the answer to this challenge. VMware vSphere 5.1 enables administrators to run more workloads on a single server, and it facilitates virtual machine mobility without downtime.

A key feature of vSphere is the ability for all machines to dynamically access shared resources such as a pool of storage. VMware® vCenter™ provides a management interface that can easily provision, monitor and leverage the shared disk resources. Without such an intelligent interface, the operational costs of scaling virtual machine workloads and their storage resources might affect the benefits of virtualization.

VMware has addressed these needs by developing VMFS to increase the benefits gained from sharing storage resources in a virtual environment. VMFS plays a key role in making the virtual environment easy to provision and manage. It provides the foundation for storage access to virtual machines by making available an automated CFS along with cluster volume management capabilities for the virtual environment.

VMFS Technical Overview

VMFS is a high-performance CFS that provides storage virtualization that is optimized for virtual machines. Each virtual machine is encapsulated in a small set of files; VMFS is the default storage management interface for these files on physical SCSI disks and partitions.

VMFS empowers IT organizations to greatly simplify virtual machine provisioning by efficiently storing the entire machine state in a central location. It enables multiple instances of vSphere hosts to access shared virtual machine storage concurrently. It also enables virtualization-based distributed infrastructure services such as vSphere DRS, vSphere HA, vMotion, vSphere Storage DRS and Storage vMotion to operate across a cluster of vSphere hosts. In short, VMFS provides the foundation that enables the scaling of virtualization beyond the boundaries of a single system.

Figure 1 shows how multiple vSphere hosts with several virtual machines running on them can use VMFS to share a common clustered pool of storage.

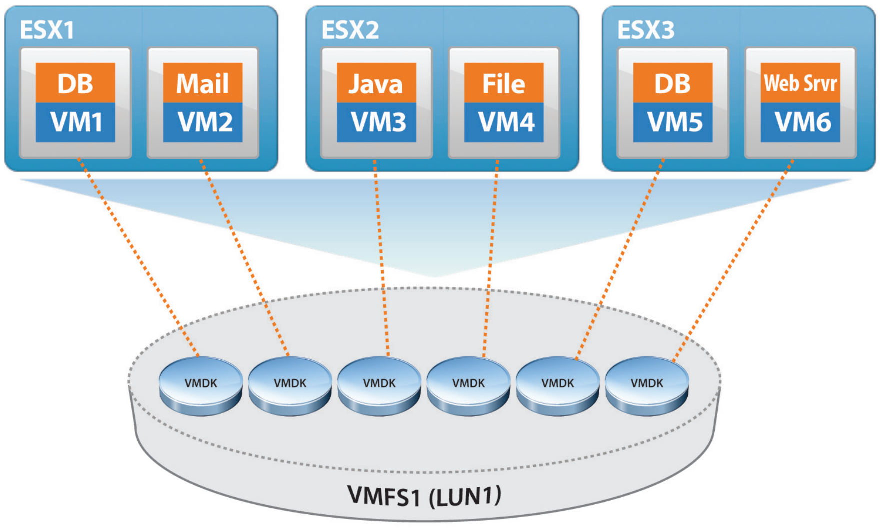

Figure 1. VMFS as a Common Pool of Storage

Each of the three vSphere hosts has two virtual machines running on it. The lines connecting them to the disk icons for the virtual machine disks (VMDKs) are logical representations of the association between and allocation of the larger VMFS volume, which is made up of one large logical unit number (LUN). A virtual machine detects the VMDK as a local SCSI target. The virtual disks are really just files on the VMFS volume, shown in the illustration as a dashed oval.

Each vSphere host stores its virtual machine files in a specific subdirectory on VMFS. When a virtual machine is operating, VMFS has a lock on those files so other vSphere hosts cannot update them. VMFS ensures that the virtual machine cannot be opened by more than one vSphere host in the cluster unless explicitly so instructed in case of clustering applications running in the virtual machine, fault tolerance or linked clones.

Each of the three vSphere hosts detects the entire LUN. The LUN is a clustered volume, and VMFS provides the distributed lock management that arbitrates access, enabling vSphere hosts to share the clustered pool of storage. The point of control moves from the SAN to the VMkernel, with no loss of security.

Features of VMFS

The following technical features of VMFS are among those that make it suitable for use in a virtual environment:

- Automated file system with hierarchical directory structure
- Optimization for virtual machines in a clustered environment
- Lock management and distributed logical volume management
- Dynamic datastore expansion by spanning multiple storage extents
- CFS with journal logging for fast recovery
- Thin-provisioned virtual disk format for space optimization
- Virtual machine–level point-in-time snapshot copy management
- Encapsulation of the entire virtual machine state in a single directory
- Support for VMware vSphere Storage APIs – Array Integration (VAAI)

Benefits of VMFS

As an intelligent and automated storage interface for virtual machine environments, VMFS provides both an automated CFS capability and intelligent cluster volume management functions. It has a number of benefits that make it particularly well suited as a CFS for the virtual environment. It is included with vSphere at no additional cost and is tailored to virtual machine performance patterns.

Enables Automated CFS Capability

VMFS is automated and optimized for virtual machines. It enables multiple vSphere hosts to access the same virtual machine storage. Virtual machines can be dynamically and automatically migrated between vSphere hosts.

Optimizes Virtual Machine Access

VMFS provides the SCSI access layer for virtual machines to efficiently read and write data on the underlying disk. VMFS5, the most recent release, uses a unified 1MB file block allocation for large files, and subblock allocation for small files and directories. VMFS is rigorously tested and certified for a wide range of Fibre Channel, Fibre Channel over Ethernet (FCoE) and iSCSI storage systems, and it is optimized to support large files while also performing many small concurrent writes.

Encapsulates the Entire Virtual Machine State in a Single Directory

VMFS stores all of the files that make up the virtual machine in a single directory and automatically creates a new subdirectory for each new virtual machine. This location is often referred to as the "virtual machine home."

Simplifies Provisioning and Administration of Virtual Machines

VMFS reduces the number of steps required to provision storage for a virtual machine. It also reduces the number of interactions required between virtualization administration (vSphere administrators) and the storage administration team to allocate storage to a new virtual machine. vSphere administrators appreciate the automated file naming and directory creation as well as the user-friendly hierarchical file system structure that eases navigation through the files that form the virtual machine environment.

Provides Distributed Infrastructure Services for Multiple vSphere Hosts

VMFS provides on-disk locking that enables concurrent sharing of virtual machine disk files across many vSphere hosts. In fact, VMFS enables virtual disk files to be shared by as many as 32 vSphere hosts. Furthermore, it manages storage access for multiple vSphere hosts and enables them to read and write to the same storage pool at the same time. It also provides the means by which vSphere DRS and vMotion can dynamically move an entire virtual machine from one vSphere host to another in the same cluster without having to restart the virtual machine. vSphere Storage DRS and Storage vMotion offer the capability to move a virtual machine home from one datastore to another without downtime, which enables migration of virtual machines off an overcrowded datastore or to a datastore with a different performance capacity.

Facilitates Dynamic Growth

Through the use of a volume management layer, VMFS enables an interface to storage resources so that several types of storage (Fibre Channel, iSCSI and FCoE) can be presented as datastores on which virtual machines can reside. Enabling dynamic growth of those datastores through aggregation of storage resources provides the ability to increase a shared storage resource pool without incurring downtime. With the addition of the VMFS Volume Grow feature, a datastore on block-based storage now can be expanded on an underlying LUN that has been expanded within the storage array. And VMFS also enables dynamic growth of the virtual disk for guest operating systems (OS) that support this capability.

Provides Intelligent Cluster Volume Management

VMFS simplifies administration with an intelligent interface that makes it easy to manage allocation and access of virtual disk resources, providing the ability to recognize and mount snapshot copies at the datastore or LUN level. VMFS has a volume signature that can be resignatured to manage additional but convergent copies of a given datastore on block-based storage.

Optimizes Storage Utilization

VMFS enables virtual disk thin provisioning as a means to dramatically increase storage utilization. With dynamic allocation and intelligent provisioning of available storage capacity in a datastore, Thin Provisioning reduces the amount of space that is allocated but not used in a datastore.

Enables High Availability with Lower Management Overhead

VMFS enables portability of virtual machines across vSphere hosts to provide high availability while lowering management overhead. As a CFS and cluster volume manager (CVM), VMFS enables unique virtualization services that leverage live migration of running virtual machines from one vSphere host to another. VMFS also facilitates automatic restart of a failed virtual machine on a separate vSphere host, and it supports clustering virtual machines across different vSphere hosts. File-level lock management provides the foundation needed for the multiserver virtualization that enables vSphere HA, vSphere DRS, vMotion, vSphere Storage DRS, Storage vMotion and VMware vSphere Fault Tolerance (FT), causing less downtime and faster recovery.

Simplifies Disaster Recovery

Because VMFS stores virtual machine files in a single subdirectory, disaster recovery, testing and cloning are greatly simplified. The entire state of the virtual machine can be remotely mirrored and easily recovered in the event of a disaster.

And with automated handling of virtual machine files, VMFS provides encapsulation of the entire virtual machine so that it easily can become part of a disaster recovery solution. The following VMFS features are among those that are especially useful in disaster recovery:

- Hardware independence between primary and secondary sites
- Encapsulation—all files for a virtual machine in one place

• Robust journal file system capability for CFS metadata
• Integration of raw disk maps (RDMs) in the VMFS structure
• Resignature option to handle storage array–based snapshots

VMware vCenter Site Recovery Manager™ Server (SRM Server) and VMware vSphere Replication leverage many of these features in the replication and disaster recovery of virtual environments.

Comparing VMFS to Conventional File Systems

Conventional file systems allow only one server to have read/write access to a specific file at a given time. In contrast, VMFS is a CFS that leverages shared storage to enable multiple vSphere hosts to have concurrent read and write access to the same storage resources. VMFS also has distributed journaling of changes to the VMFS metadata to enable fast and resilient recovery across multiple vSphere clusters.

On-disk locking in VMFS ensures that a virtual machine is not powered on by multiple installations of a vSphere host at the same time. With vSphere HA enabled, if a server fails, the on-disk lock for each virtual machine is released, enabling the virtual machine to be restarted on other vSphere hosts. Moreover, VMFS provides virtual machine–level snapshot capability, enabling fast point-in-time recovery. VMware vSphere Data Protection (VDP), along with backup products from many VMware partners, leverages this feature to provide consistent backup of virtual environments.

VMFS does not have every feature found today in other CFS and CVM systems. However, there is no other CFS or CVM that provides the capabilities of VMFS. Its distributed locking methods forge the link between the virtual machine and the underlying storage resources in a manner that no other CFS or CVM can equal. The unique capabilities of VMFS enable virtual machines to join a VMware cluster seamlessly, with no management overhead.

Best Practices for Deployment and Use of VMFS

This section offers some best practices, insight and experience in addressing several questions that often arise when deploying and using VMFS volumes. It is not intended to provide the definitive answer for every question, because often there is no single right answer. The intent here is to discuss what the trade-offs and considerations are, as well as to offer some insights in choosing the answer that best fits a specific configuration.

The following topics are addressed:

• How large should LUNs be created for a given VMFS volume?
• Should we isolate storage for virtual machines or share a consolidated pool of storage?
• Should we use RDMs or VMFS volumes?
• Should we use disk spanning? If so, are there any concerns or suggestions?

How Large a LUN?

The best way to configure a LUN for a given VMFS volume is to size for throughput first and capacity second. That is, you should aggregate the total I/O throughput for all applications or virtual machines that might run on a given shared pool of storage; then make sure you have provisioned enough back-end disk spindles (disk array cache) and appropriate storage service to meet the requirements.

This is actually no different from what most system administrators do in a physical environment. It just requires an extra step, to consider when to consolidate a number of workloads onto a single vSphere host or onto a collection of vSphere hosts that are addressing a shared pool of storage.

Each storage vendor likely has its own recommendation for the size of a provisioned LUN, so it is best to check with the vendor. However, if the vendor's stated optimal LUN capacity is backed with a single disk that has little or no storage array write cache, the configuration might result in low performance in a virtual environment. In this case, a better solution might be a smaller LUN striped within the storage array across many physical disks, with some write cache in the array. The RAID protection level also factors into the I/O throughput performance.

Because there is no single correct answer to the question of how large your LUNs should be for a VMFS volume, the more important question to ask is, "How long would it take one to restore the virtual machines on this datastore if it were to fail?" The recovery time objective (RTO) is now the major consideration when deciding how large to make a VMFS datastore. This equates to how long it would take an administrator to restore all of the virtual machines residing on a single VMFS volume if there were a failure that caused data loss. With the advent of very powerful storage arrays, including Flash storage arrays, the storage performance has become less of a concern. The main concern now is how long it would take to recover from a catastrophic storage failure. Another important question to ask is, "How does one determine whether a certain datastore is overprovisioned or underprovisioned?" There are many performance screens and metrics that can be investigated within vCenter to monitor datastore I/O rates and latency. Monitoring these metrics is the best way to determine whether a LUN is properly sized and loaded. Because workload can vary over time, periodic tracking is an important consideration. vSphere Storage DRS, introduced in vSphere 5.0, can also be a useful feature to leverage for load balancing virtual machines across multiple datastores, from both a capacity and a performance perspective.

Isolation or Consolidation?

The decision whether to "isolate" or "consolidate" storage resources for virtual environments is a topic of some debate. The basic answer depends on the nature of the I/O access patterns of that virtual machine. If you have a very heavy I/O-generating application, in many cases VMware vSphere Storage I/O Control can assist in managing fairness of I/O resources among virtual machines. Another consideration in addressing the "noisy neighbor" problem is that it might be worth the potentially inefficient use of resources to allocate a single LUN to a single virtual machine. This can be accomplished using either an RDM or a VMFS volume that is dedicated to a single virtual machine. These two types of volumes perform similarly (within 5 percent of each other), with varying read and write sizes and I/O access patterns. Figure 2 illustrates the differences between isolation and consolidation.

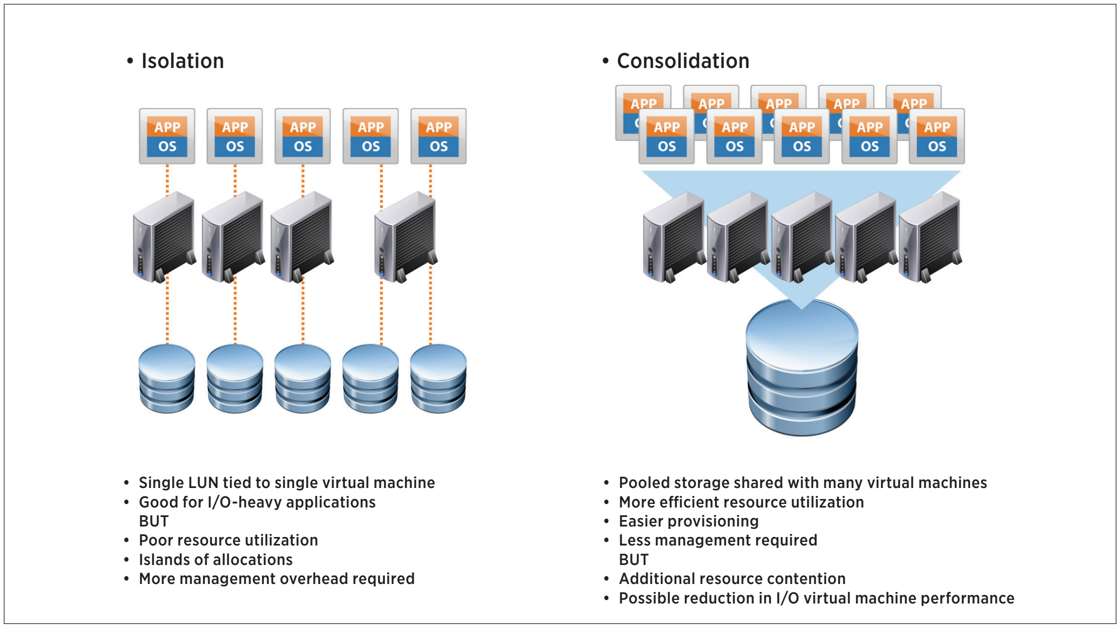

Figure 2. Differences Between Consolidation and Isolation

The following are detailed arguments regarding isolation and consolidation:

Isolated Storage Resources

One school of thought suggests limiting the access of a single LUN to a single virtual machine. In the physical world, this is quite common. When using RDMs, such isolation is implicit, because each RDM volume is mapped to a single virtual machine.

The downside to this approach is that as you scale the virtual environment, you soon reach the upper limit of 256 LUNs per host. You also must provision an additional disk/LUN each time you want to increase storage capacity for the virtual machine. This can lead to significant management overhead; in some environments, it might take days for a request to provision a disk/LUN to be serviced by the storage administration team.

Another consideration is that every time you must expand the capacity of a virtual machine, your minimum commit size is the allocation of a LUN. Although many arrays allow LUNs of any size, the storage administration team might balk at carving up lots of small LUNs because this configuration makes it more difficult for them to manage the array. Most storage teams prefer to allocate LUNs that are fairly large; they like the system administration or applications teams to divide those LUNs into smaller chunks higher up in the stack. VMFS suits this allocation scheme perfectly and is one of the reasons VMFS is so effective in the virtualization storage management layer.

Consolidated Pools of Storage

The consolidation school wants to gain additional management productivity and resource utilization by pooling the storage resource and sharing it, with many virtual machines running on several vSphere hosts. Dividing this shared resource among many virtual machines enables better flexibility as well as easier provisioning and ongoing management of the storage resources for the virtual environment.

Compared to strict isolation, consolidation normally offers better utilization of storage resources. The cost is additional resource contention, which under some circumstances can lead to reduction in virtual machine I/O performance. However, vSphere offers Storage I/O Control and vSphere Storage DRS to mitigate these risks.

At this time, there are no clear rules of thumb regarding the limits of scalability. For most environments, the ease of storage resource management and incremental provisioning offers gains that outweigh any performance impact. As you will see later in this paper, however, there are limits to the extent of consolidation.

Best Practice: Mix Consolidation with Some Isolation

In general, use vSphere Storage DRS to detect and mitigate storage latency and capacity bottlenecks by load balancing virtual machines across multiple VMFS volumes. Additionally, vSphere Storage I/O Control can be leveraged to ensure fairness of I/O resource distribution among many virtual machines sharing the same VMFS datastore. However, the vSphere Storage DRS and vSphere Storage I/O Control features might not always be available. Another option is to separate heavy I/O workloads from the shared pool of storage. This optimizes the performance of those high-transactional throughput applications—an approach best characterized as "consolidation with some level of isolation."

Because workloads can vary significantly, there is no exact formula that determines the limits of performance and scalability regarding the number of virtual machines per LUN. These limits also depend on the number of vSphere hosts sharing concurrent access to a given VMFS volume. The key is to remember the upper limit of 256 LUNs per vSphere host and consider that this number can diminish the consolidation ratio if you take the concept of "one LUN per virtual machine" too far.

Many different applications can easily and effectively share a clustered pool of storage. And the increase in disk utilization and improvements in management efficiency clearly can compensate for the minor performance reductions caused by the additional contention.

Use of RDMs or VMFS?

Another question is when to use VMFS and when to use RDMs. This section explains the trade-offs.

About RDMs

First, a little more detail about RDMs. As illustrated in Figure 3, an RDM file is a special file in a VMFS volume that manages metadata for its mapped device. The mapping file is presented to the management software as an ordinary disk file that is available for the usual file system operations. To the virtual machine, the storage virtualization layer presents the mapped device as a virtual SCSI device. Key contents of the metadata in the mapping file include the location (name resolution) and locking state of the mapped device.

RDM files contain metadata used to manage and redirect disk accesses to the physical device. Employing RDMs provides the advantages of direct access to a physical device while keeping some advantages of a virtual disk in the VMFS file system. In effect, the RDM merges VMFS manageability with raw device access.

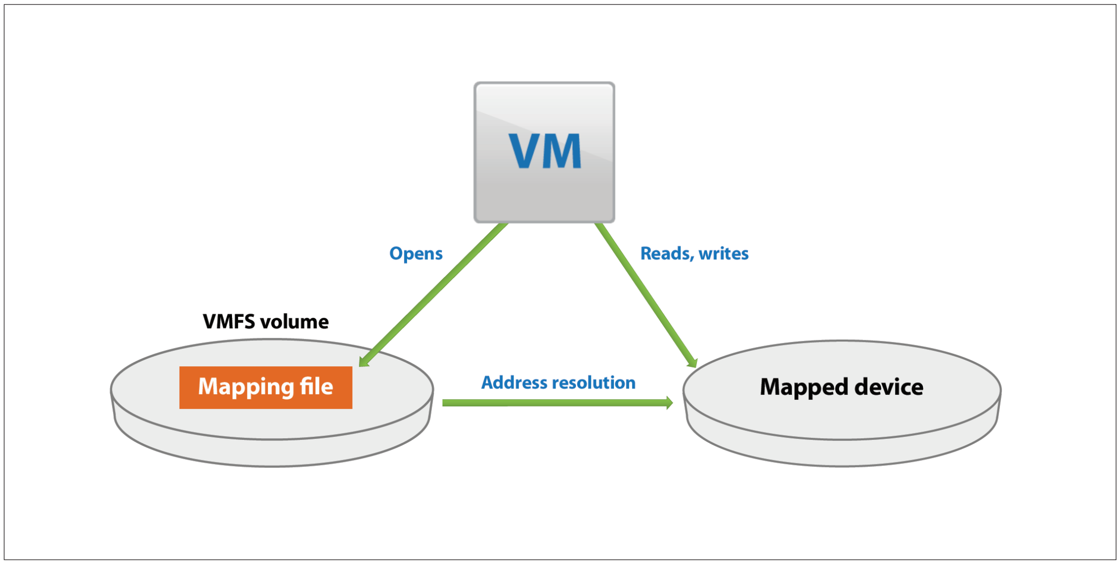

Figure 3. Raw Disk Mapping

An RDM is a symbolic link from a VMFS volume to a raw volume. The mapping makes volumes appear as files in a VMFS volume. The mapping file—not the raw volume—is referenced in the virtual machine configuration. The mapping file, in turn, contains a reference to the raw volume.

Using RDMs, you can do the following:

- Use vMotion to migrate virtual machines using raw volumes.
- Add raw volumes to virtual machines using the VI client.
- Use file system features such as distributed file locking, permissions and naming.

RDMs have the following two compatibility modes:

- Virtual compatibility mode enables a mapping to act exactly like a virtual disk file, including the use of storage array snapshots.
- Physical compatibility mode enables direct access of the SCSI device, for those applications needing lower level control.

vMotion, vSphere DRS and vSphere HA are all supported for RDMs that are in both physical and virtual compatibility modes.

Why Use VMFS?

For most applications, VMFS is the clear choice. It provides the automated file system capabilities that make it easy to provision and manage storage for virtual machines running on a cluster of vSphere hosts. VMFS has an automated hierarchical file system structure with user-friendly file-naming access. It automates the subdirectory naming process to make administration more efficient in managing RDMs. It enables a higher disk utilization rate by facilitating the process of provisioning the virtual disks from a shared pool of clustered storage.

As you scale the number of vSphere hosts and the total capacity of shared storage, VMFS greatly simplifies the process. It also enables a larger pool of storage than might be addressed via RDMs. Because the number of LUNs that a given cluster of vSphere hosts can discover is currently capped at 256, you can reach this number rather quickly if mapping a set of LUNs to every virtual machine running on the vSphere host cluster.

Using RDMs usually requires more frequent and varied dependence on the storage administration team, because each LUN must be sized for the needs of each specific virtual machine to which it is mapped.

With VMFS, however, you can carve out many smaller VMDKs for virtual machines from a single VMFS volume. This enables the partitioning of a larger VMFS volume—or a single LUN—into several smaller virtual disks, which facilitates a centralized management utility (vCenter) to be used as a control point. The control point resides at the vSphere host level, between the storage array and the virtual machine.

With RDMs, there is no way to break up the LUN and address it as anything more than a single disk for a given virtual machine. One example of this limitation is a case where a user provisioned several 50GB LUNs and wanted to test relative performance on a few virtual machines. The plan called for testing with 10GB virtual disks. With an RDM, the only choice was to address the entire 50GB RDM to the virtual machine and use only the first 10GB. This wasted the other 40GB of that LUN. Using VMFS with a 50GB volume, on the other hand, enabled the creation of five directly addressable virtual disks of 10GB each on the shared VMFS volume.

Even if performance is the main consideration, you can employ a single VMFS volume for each virtual machine, in much the same way as an RDM volume is isolated to a single virtual machine. (Used this way, the VMFS and the RDM volumes provide similar performance.) The bigger question is whether to isolate or consolidate, and that is not limited to use of RDMs for isolation and VMFS for consolidation.

Why Use RDMs?

Even with all the advantages of VMFS, there still are some cases where it makes more sense to use RDM storage access. The following scenarios call for raw disk mapping:

- Migrating an existing application from a physical environment to virtualization
- Using Microsoft Cluster Service (MSCS) for clustering in a virtual environment
- Implementing N-Port ID Virtualization (NPIV)
- Separating heavy I/O workloads from the shared pool of storage

RDM Scenario 1: Migrating an Existing Application to a Virtual Server

Figure 4 shows a typical migration from a physical server to a virtual one. Before migration, the application running on the physical server has two disks (LUNs) associated with it. One disk is for the OS and application files; a second disk is for the application data.

To begin, use the VMware vCenter Converter™ to build the virtual machine and to load the OS and application data into the new virtual machine.

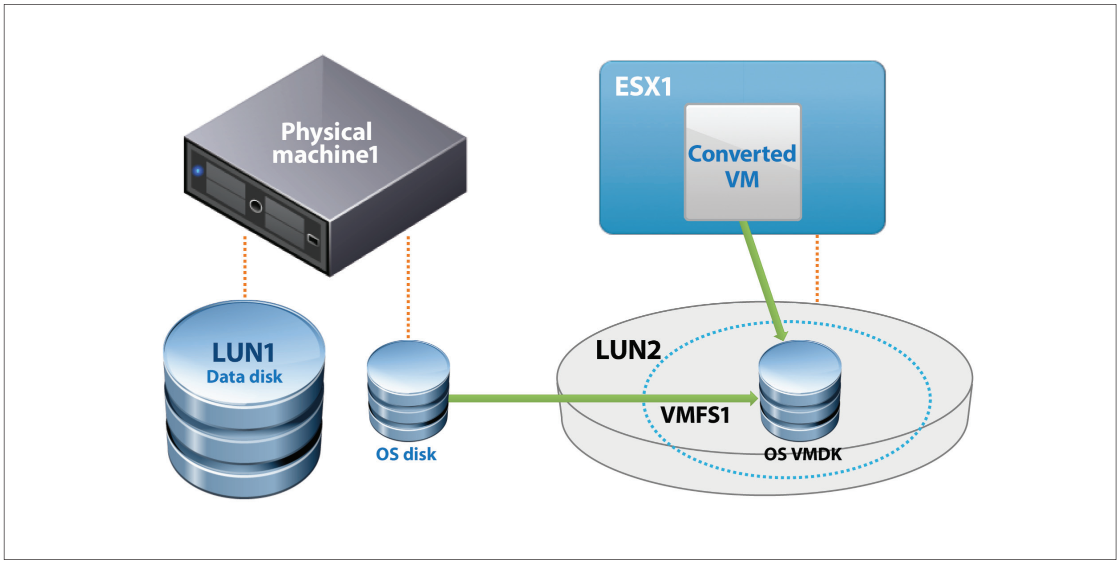

Figure 4. VMware vCenter Converter

Next, remove access to the data disk from the physical machine and make sure the disk is properly zoned and accessible from the vSphere host. Then create an RDM for the new virtual machine pointing to the data disk. This enables the contents of the existing data disk to be accessed just as they are, without the need to copy them to a new location.

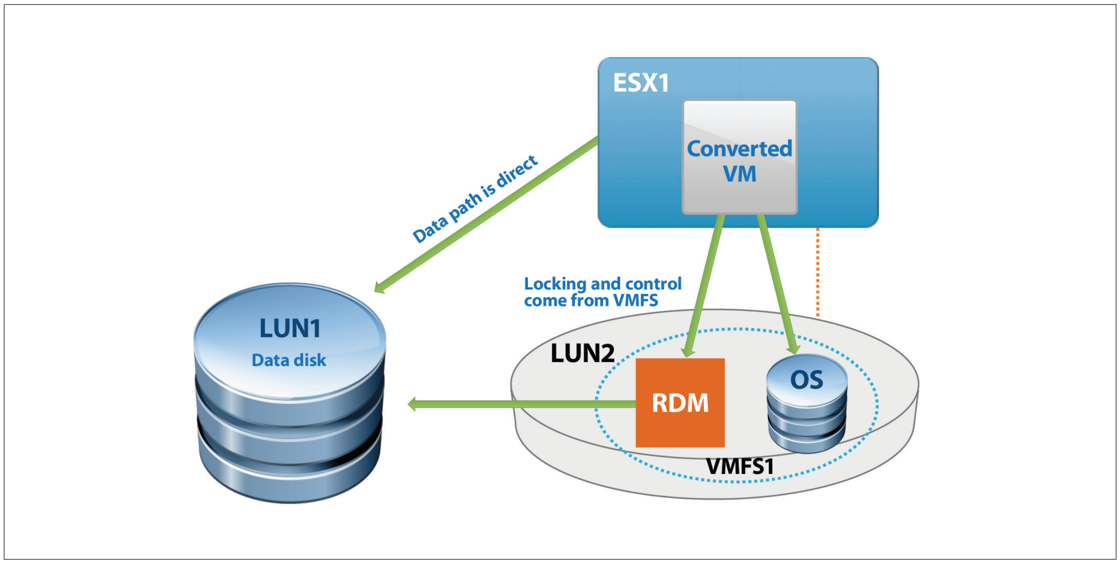

Figure 5. Raw Disk Map Creation

The path to the data disk located on the RDM is stored in the VMFS. Although VMFS provides security access control and locking, the data path to the RDM is direct access. As with virtual disks, VMFS controls access to make sure there is no simultaneous access to the data disk from other virtual machines. Because RDMs enable vMotion, the VMFS can transfer control of the RDM to the destination vSphere host when a virtual machine migrates.

RDM Scenario 2: Using Microsoft Cluster Service in a Virtual Environment

Another common use of RDMs is for MSCS configurations. These can use clusters in a single vSphere host (cluster in a box); clusters between virtual machines on separate vSphere hosts (cluster across boxes); or clusters across both physical and virtual machines. Each of these scenarios has different requirements for shared storage, which are summarized in the following table:

	CLUSTER IN A BOX	CLUSTER ACROSS BOXES	N+1 CLUSTERING
Virtual disks	Yes	No	No
Passthrough RDM (physical compatibility mode)	No	Yes	Yes
Non-passthrough RDM (virtual compatibility mode)	Yes	Yes	No

Table 1. Summary of MSCS Shared Storage

Using MSCS in a virtual environment requires a disk for each node in which files specific to that node are stored. Other files require shared access for the quorum disk. Those disks must support native file system access, which requires the use of RDMs in physical compatibility mode. This is another example where RDMs provide a more flexible solution for the use of virtualization technology.

For more information on MSCS configurations supported with vSphere hosts, refer to the VMware *Setup for Microsoft Cluster Service* documentation at:

- http://pubs.vmware.com/vsphere-50/topic/com.vmware.ICbase/PDF/vsphere-esxi-vcenter-server-50-mscs-guide.pdf
- http://www.vmware.com/pdf/vsphere4/r41/vsp_41_mscs.pdf
- http://www.vmware.com/pdf/vSphere4/r40/vsp_40_mscs.pdf
- http://www.vmware.com/pdf/vi3_301_201_mscs.pdf

More details can be found in VMware knowledge base article 1004617.

When and How to Use Disk Spanning

It is generally best to begin with a single LUN in a VMFS volume. To increase the size of that resource pool, you can provide additional capacity by either 1) adding a new VMFS extent to the VMFS volume or 2) increasing the size of the VMFS volume on an underlying LUN that has been expanded in the array (via a dynamic expansion within the storage array). Adding a new extent to the existing VMFS volume will result in the existing VMFS volume's spanning across more than one LUN. However, until the initial capacity is filled, that additional allocation of capacity is not yet put to use. The VMFS does not stripe I/O across LUNs when more than one LUN is allocated to a given VMFS volume.

Expanding the VMFS volume on an existing, larger LUN will also increase the size of the VMFS volume, but it should not be confused with spanning. However, the VMFS Volume Grow feature can be used to expand a VMFS volume that spans a few VMFS extents as well as one that spans multiple LUNs, provided there is space on those LUNs to expand the VMFS extents onto.

From a management perspective, it is preferable that a single large LUN with a single extent host your VMFS. Using multiple LUNs to back multiple extents of a VMFS volume entails presenting every LUN to each of the vSphere hosts sharing the datastore. Although multiple extents might have been required prior to the release of vSphere 5 and VMFS5 to produce VMFS volumes larger than 2TB, VMFS5 now supports single-extent volumes up to 64TB.

Gaining Additional Throughput and Storage Capacity

Additional capacity with disk spanning does not necessarily increase I/O throughput capacity for that VMFS volume. It does, however, result in increased storage capacity. If properly provisioned on the underlying storage array, the additional capacity can be allocated on LUNs other than the first LUN and will result in additional throughput capacity as well. It is very important to be certain you are adding LUNs of similar performance capability (RAID level and I/O density) when adding to an existing VMFS volume.

The current size limit for a VMFS5 extent is 64TB. In VMFS3, the extent size was 2TB. For large VMFS3 volumes, spanning was required to concatenate multiple 2TB extents. There is a limit of 32 extents in a VMFS volume; the size limit of any VMFS volume is 64TB. Spanning of multiple volumes (LUNs) was required to reach that upper limit and is needed for any VMFS3 volume greater than 2TB in size. This is not necessary for VMFS5.

Suggestions for Rescanning

In prior versions of vSphere, it was recommended that before adding a new VMFS extent to a VMFS volume, you make sure a rescan of the SAN is executed for all nodes in the cluster that share the common pool of storage. However, in more recent versions of vSphere, there is an automatic rescan that is triggered when the target detects a new LUN, so that each vSphere host updates its shared storage information when a change is made on that shared storage resource. This auto rescan is the default setting in vSphere and is configured to occur every 300 seconds.

Conclusion

VMware vSphere VMFS provides the foundation for virtualization to span multiple systems. It enables optimized I/O handling of multiple virtual machines sharing a common clustered pool of storage. It also provides more efficient storage management for both the initial provisioning and the ongoing management of the shared storage resources. Leveraging VMware vSphere Thin Provisioning can dramatically increase the utilization rates of your storage resources in a virtual environment. VMFS insight and monitoring are built into VMware vCenter as an intelligent interface that enables efficiency and easy management.

VMFS is the leading cluster file system for virtualization. It is unique in the market today and provides the capabilities that empower virtual machines to scale beyond the limits of a single server, without downtime. VMFS enables VMware vSphere Storage vMotion to migrate virtual machines from one storage resource to another without downtime. It also improves resource utilization and lowers the management overhead for storage administration in a virtual environment. VMFS is a key reason why VMware vSphere is today's leading virtualization solution, one that is more scalable and reliable than any other offering currently available.

About the Author

Cormac Hogan is a senior technical marketing architect within the Cloud Infrastructure Product Marketing group at VMware. He is responsible for storage in general, with a focus on core VMware vSphere storage technologies and virtual storage, including the VMware vSphere Storage Appliance. Cormac has written a number of storage-related white papers and has given numerous presentations on storage best practices and new features. He has been in technical marketing since 2011.

Additional Reading

For those customers who are considering upgrading from VMFS3 to VMFS5, the following white paper mentions a number of considerations that should be taken into account: http://www.vmware.com/resources/techresources/10242.

Glossary

CFS – Cluster file system.

CVM – Cluster volume manager.

Datastore – A formatted file system that a vSphere host mounts and uses as a shared storage pool. It is built upon either a VMFS volume for block-based storage or a mount point for NFS shared storage.

DRS – VMware vSphere Distributed Resource Scheduler.

vSphere host – The hardware on which the vSphere software is loaded and running.

Extent – Part of a larger allocation of a resource, or an appendage to an existing resource.

Fibre Channel (FC) – An ANSI-standard, gigabit-speed network technology used to build storage area networks and transmit data. FC components include HBAs, switches and cabling.

HA – High availability.

iSCSI – Internet small computer serial interface.

LUN – Logical unit number—what the server detects as a single disk resource.

MSCS – Microsoft Cluster Service.

NFS – Network file system, provided by network attached storage (NAS).

RAID – Redundant array of independent disks.

RDM – Raw disk map.

SAN – Storage area network.

SCSI – Small computer serial interface.

Storage vMotion – A live migration of a virtual machine on disk files from one datastore to another. The virtual machine home changes location, but the virtual machine remains running on the same vSphere host.

VCB – VMware consolidated backup.

vCenter – Centralized management utility.

VMware vSphere Client™ – Virtual infrastructure client, a management access point.

VMDK – Virtual machine disk, a file in the VMFS that backs the storage object for the disk for a virtual machine.

VMFS3 – Version 3 of the Virtual Machine File System.

VMFS5 – Version 5 of the Virtual Machine File System.

VMFS extent – The portion of a VMFS volume that resides on a single LUN.

VMFS volume – The aggregation of block-based storage that backs a datastore. Normally a single LUN, it can be up to 32 LUNs when a VMFS volume spans across multiple VMFS extents.

VMkernel – Core management component of the vSphere host system.

vMotion – A means by which the virtual machine can be moved from one vSphere host to another without any downtime (the virtual machine remains on the same datastore but runs on a new vSphere host).

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