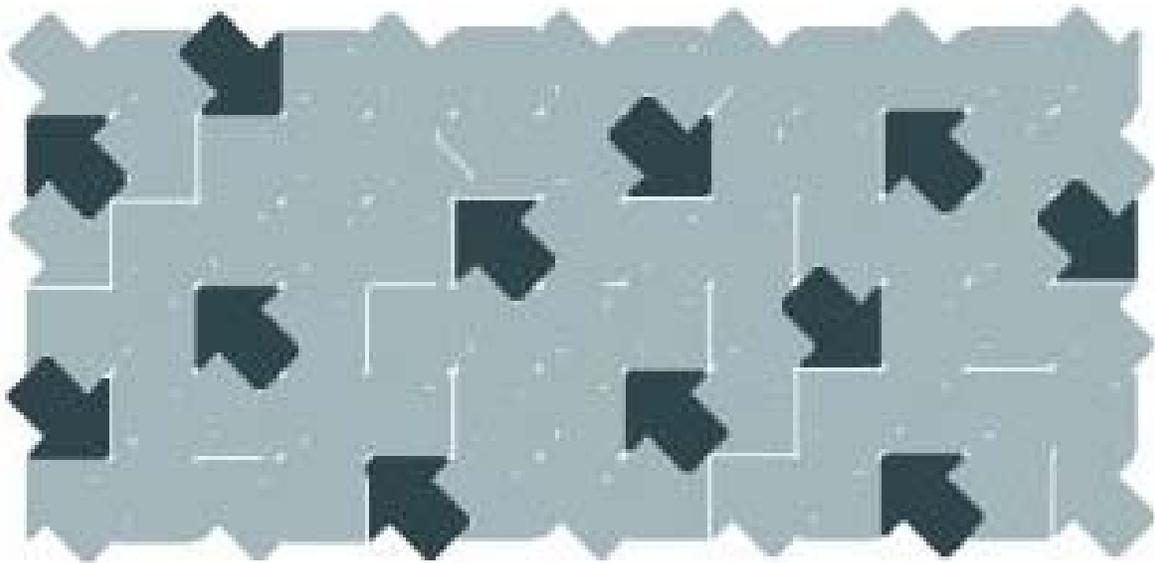# Layers of Virtual Storage

in VMware VI3:

Configuration without Confusion

A VMware Technical White Paper

Layers of Virtual Storage in VMware VI3: Configuration without Confusion
Paper Number: WP-023-INF-01-01

You can find the most up-to-date technical documentation on our Web site at
    http://www.vmware.com/support/
The VMware Web site also provides the latest product updates.
If you have comments about this documentation, submit your feedback to:
    docfeedback@vmware.com

VMware, Inc.
3401 Hillview Ave
Palo Alto, CA 94304
www.vmware.com

Contents

Layers of Virtual Storage in VMware VI3:

Configuration without Confusion

Introduction

Provisioning storage for a virtual environment requires navigating several layers of abstraction, making it desirable to view this as a stack in order to follow the complete path successfully. First, there are several options for configuring storage within the storage array. Next, several more choices are available at the ESX Server level for addressing and provisioning that storage. Then at the virtual machine level still more options determine how a virtual disk is partitioned and addressed. Because each of these three layers in the storage stack provides a set of configuration choices that may not be obvious from the other layers in the stack, it's easy to become confused about what virtual-to-physical mapping actually exists. The purpose of this paper is to define the options, choices and terms that can help trace the route through this complex set of dependencies.

Overview of the Storage Virtualization Stack

Because ESX Server addresses storage and presents virtual disks to a VM, there are already a few layers of virtualization in the mix before the introduction of a storage virtualization device. Even without actual storage virtualization devices, those layers of storage virtualization provide ease of management, better resource utilization rates (that is, less allocated but unused storage) and more flexible provisioning options. VMware's Virtual Machine File System, or VMFS, adds benefits that deliver significant value to the virtual environment.

Figure 1 shows a simplified view of the three layers of the storage technology stack for a virtual machine running on Virtual Infrastructure 3 environment. At the bottom is the storage array layer, which consists of physical disks that are presented as logical disks. Above the storage array layer is the operating system layer, which in the virtual environment is occupied by ESX Server. Storage from the VMFS volume is addressed in this layer as logical unit numbers (LUNs). The top of the stack is occupied by virtual machine layer, where storage is addressed by allocating a portion of the storage seen by the ESX Server to be set aside as virtual disks for the VM.
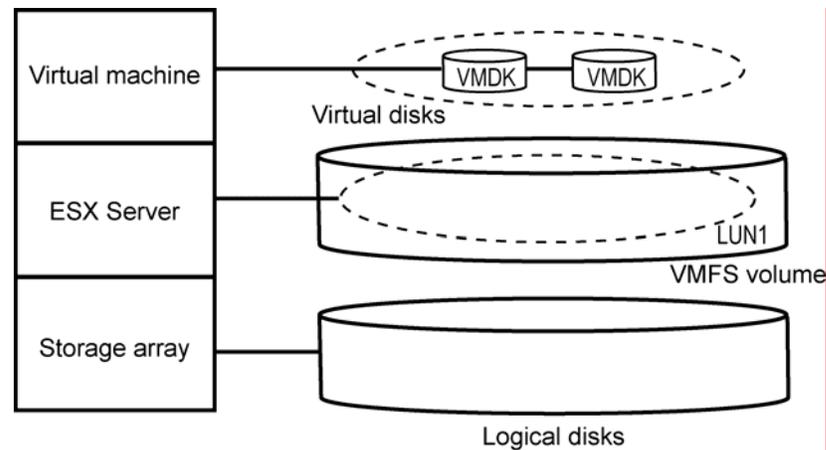
Figure 1. Three Layers of the Storage Stack for VI3

Configuring the Storage Array Layer

Within a storage array, there are several options to configure the storage resource. Typically the storage administration team will want to determine the desired LUN size and level of protection. In some cases they will also ask what class of performance is being requested.

RAID-1 and RAID-5 are common protection options offered. Another option is concatenating several disks to present a single LUN; that is, the LUN will often span several disks and is striped across the spindles. For the most part, the host OS (ESX Server in our case) will not know what level of RAID protection is being provided to the LUNs. The host OS also will not know what type of relationship exists between the logical LUN and its physical components
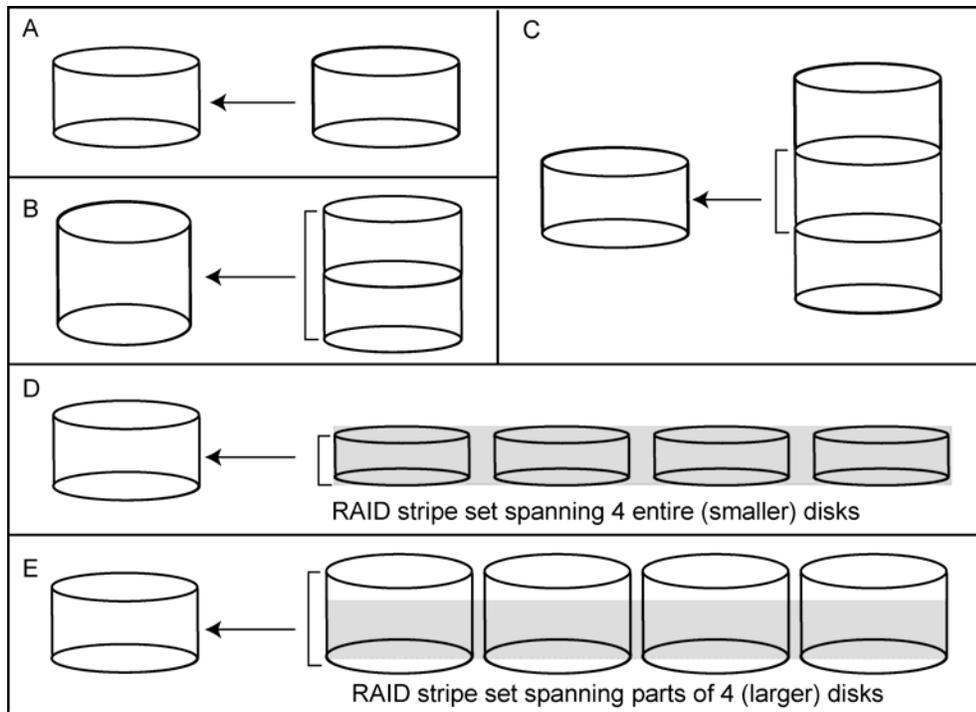
Figure 2. Configurations for a LUN at the ESX Server Level

Figure 2 illustrates the most common configurations for a LUN:

A. The entire physical disk

B. Several disks that are concatenated together

C. Part of a larger physical disk

D. A stripe set that covers the entire portion of several disks.

E. Made up of a stripe set across a portion of several disks

Note that the LUN configuration can have an effect on the I/O performance characteristic. The level of protection can affect performance as well. For example, if the LUN is being replicated to a remote site in a synchronous mode, the effect of latency could cause a significant degradation in overall performance. The limitation of the speed of light traveling through a vacuum results in latency of 1 millisecond for every 123 miles traveled. With synchronous mode replication, the result will be 2 milliseconds of delay for each 123 miles of separation, because the primary site must receive acknowledgement of the remote write from the secondary site before the write is considered to be completed.

Configuring the ESX Server Layer

Storage space is presented to the ESX Server system as volumes with logical unit numbers or, in the case of a network-attached storage, as NFS volumes. When the ESX Server discovers a logical unit number, the LUN is treated as a single storage target. The LUN can then be addressed as a raw disk for a raw disk map (RDM), or managed as a VMFS Volume or an extent of a multi-extent VMFS Volume.

As an RDM the entire LUN is typically addressed by single Virtual Machine (VM) and cannot be partitioned or addressed as a segment to separate VMs. To achieve an RDM of a size smaller than the size of the LUN, one option is to address only part of the LUN for a given VM and then lose or waste the remainder of that LUN. In this case, if you have a 100GB LUN available and the VM only requires a 50GB RDM, half the LUN is wasted. Allocating a single LUN to a single VM (which can be done with an RDM or with a VMFS volume dedicated to one VM) is known as "isolated storage."

When a LUN is part of a VMFS Volume it becomes a storage resource that can serve as a consolidated and clustered pool of storage for several VMs as well as several ESX Servers. This configuration ("consolidated storage") allows increased utilization and agility of that storage resource, as it is part of a larger pool of storage that can be centrally managed and increased in size as needed. Figure 3 illustrates the relative merits of isolated storage versus a consolidated or shared pool of storage.
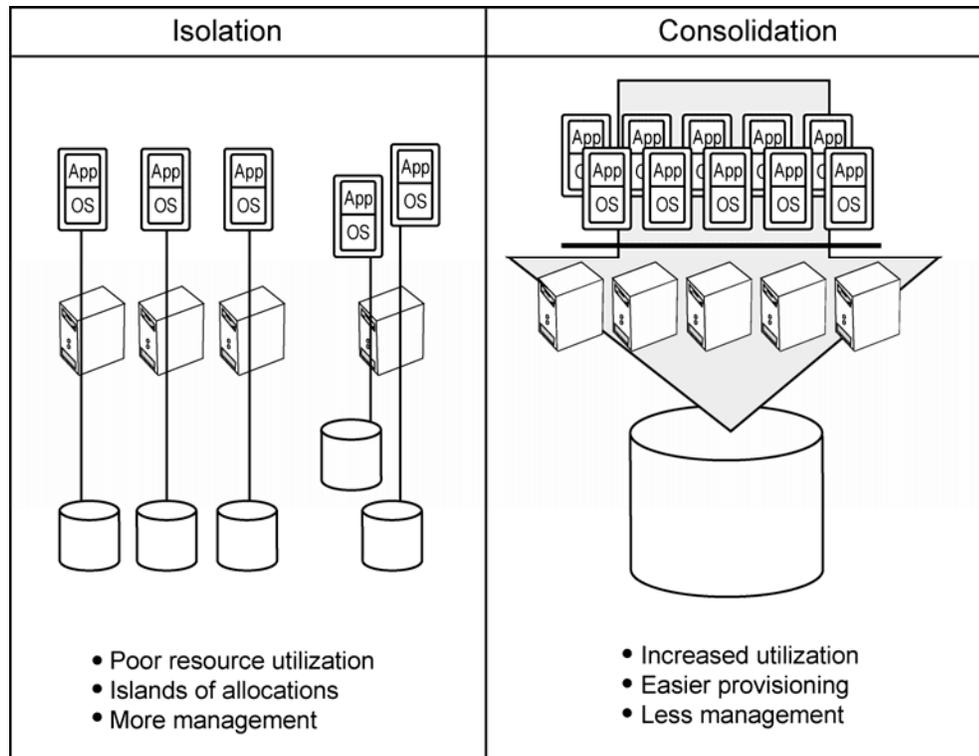
Figure 3. Comparing Islands of Storage and a Shared Pool of Storage

ESX Server virtual machines store their virtual disk files on specially formatted logical containers, or datastores, which can exist on different types of physical storage devices. The Virtual Infrastructure (VI) client is used to create datastores, and a datastore can use disk space on one physical device or several physical devices. Within an ESX 3.x Server datastores can be FC disks, iSCSI disks, NFS, or a disk local to that ESX Server.

Figure 4 shows the full set of choices available when configuring storage for an ESX server. When using local storage, you will not be able to leverage the VMotion capabilities that enable VMware Distributed Resource Schedule (DRS) and High Availability (HA). However, if you choose shared storage, you can choose from VMFS, RDM or network-attached storage (NAS) are options. In fact, a single ESX server can be configured to access datastores on all three of those options at the same time. Also, if using VMFS, you can opt for either FC SAN or iSCSI SAN storage.

Details about RDMs and iSCSI are a bit beyond the scope of this paper, and can be found in the SAN configuration guide located on the VMware.com website at VI3_SAN_Config_Guide.
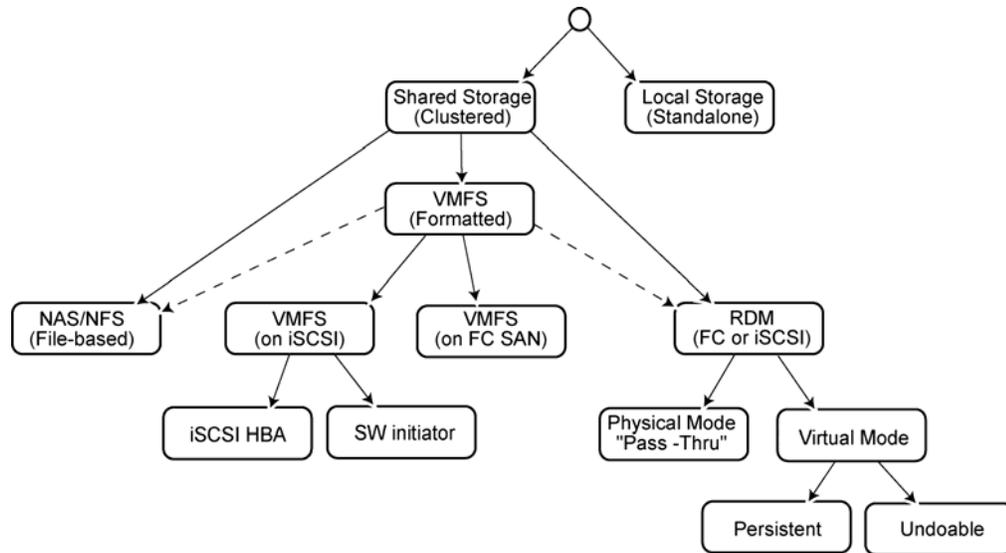
Figure 4. Storage Options for an ESX Server

Configuring the Virtual Machine Layer

Creating a Virtual Disk (VMDK) allows addressing storage in a Virtual Machine by allocating a portion of the storage seen by the ESX Server to be set aside for the VM. Editing the VM settings (Properties) lets you either create a new disk allocation or add to an existing one. You also have the option of pointing to an existing RDM, if that is the preferred access method. When creating a VMDK from an existing VMFS volume, you will need to select the datastore from which to allocate that VMDK. For a VM to be moved between two or more ESX Servers in cluster (in the case of Vmotion and HA, for example), the datastore must be seen by all nodes in that cluster.

The steps for recognizing the disk and formatting the disk for use within that environment depend on the OS. There are also different ways in which a disk that has been increased in size in the underlying layers can be recognized as a large disk at the Guest OS level. This configuration is beyond the scope of this paper and will not be covered here.

Conclusion

Configuring storage in the virtual environment requires navigating many interdependent elements, some physical and some abstract. Viewing the different levels of storage virtualization as a stack, with separate, distinct layers and interlocking dependencies that are not necessarily visible to each other, simplifies and clarifies the setup process. Proceeding step by step from the storage array to the ESX Server to the virtual machine, allows virtual storage configuration without confusion.