



VMware ESX Server

Providing LUN Security

VMware ESX Server provides strong security and performance isolation for virtual machine storage. Each virtual machine sees only the virtual disks that have been presented to its virtual SCSI adapters. Virtual machines cannot see the physical Fibre Channel HBAs on the ESX Server host on which they run. Nor, in typical use cases, do they see the LUNs on which their virtual disks reside. Emerging mechanisms for LUN security in a virtual environment from Fibre Channel HBA vendors provide an alternative for accomplishing the same goals.

In a physical Fibre Channel SAN environment, LUN security is typically accomplished through a combination of LUN masking and zoning. Using these approaches in a vendor-recommended way ensures that a given LUN can be accessed only by a single host, as identified by the world wide names (WWN) of its HBAs.

In a virtual environment, this situation changes slightly. It is now possible to have multiple virtual machines on a single physical host. Furthermore, to facilitate the use of advanced technologies such as VMotion, multiple ESX Server hosts may have their LUN masking and zoning set up to allow for broad access, with control being maintained by VMFS, the distributed file system that is included as part of ESX Server.

As this document explains in greater detail:

- Virtual machines can see and access only specific units of storage that the ESX Server administrator explicitly allows. This is true whether the virtual machine is using virtual disks on a VMFS file system or raw device mappings.
- The operating system within the virtual machine cannot change its own storage access nor interrogate a unit of storage in a way that allows it to discover any other storage units not defined by the ESX Server administrator.
- Special mechanisms are built in so concurrent access by multiple ESX Server systems to storage can take place safely. This protection enables advanced functionality such as VMotion.

This document addresses the following topics:

- [Storage Isolation for Virtual Machines on page 2](#)
- [VMFS and VMotion on page 2](#)
- [Raw Device Mappings on page 3](#)
- [N_Port ID Virtualization on page 3](#)
- [Continuing Innovation on page 4](#)



Storage Isolation for Virtual Machines

A host running ESX Server is attached to a Fibre Channel SAN in the same way that any other host is. It uses Fibre Channel HBAs, with the drivers for those HBAs installed in the software layer that interacts directly with the hardware. In environments that do not include virtualization software, the drivers are installed on the operating system. In the case of ESX Server, the drivers are installed in the ESX Server component known as VMkernel — a lightweight microkernel that controls the underlying hardware.

ESX Server also includes VMFS, a distributed file system and volume manager that creates and manages virtual volumes on top of the LUNs that are presented to the ESX Server host. Those virtual volumes, usually referred to as virtual disks, are allocated to specific virtual machines.

Virtual machines have no knowledge or understanding of Fibre Channel. The only storage available to virtual machines is SCSI devices. Put another way, a virtual machine does not have virtual Fibre Channel HBAs, but only has virtual SCSI adapters. Each virtual machine is able to see only the virtual disks that are presented to it on its virtual SCSI adapters.

This isolation is complete, with regard to both security and performance. A VMware virtual machine has no visibility into the WWN, the physical Fibre Channel HBAs, or even the target ID or other information about the LUNs upon which its virtual disks reside. The virtual machine is isolated to such a degree that software executing in the virtual machine cannot even detect that it is running on a SAN fabric. Even multipathing is handled in a way that is transparent to a virtual machine. Furthermore, virtual machines can be configured to limit the bandwidth they use to communicate with storage devices in order to assure desired quality of service levels.

Consider the example of running the Microsoft Windows operating system inside a VMware virtual machine. The virtual machine sees only the virtual disks chosen by the ESX Server administrator at the time the virtual machine is configured. This operation of configuring a virtual machine to see only certain virtual disks is effectively LUN masking in the virtualized environment. It has the same security benefits as LUN masking in the physical world; it is just done with a different set of tools. There is no security hole in this system. Software executing in the virtual machine — including the Windows operating system — is only aware of the virtual disks attached to the virtual machine. Even if the Windows operating system attempts to issue a SCSI command — Report LUNs — in an attempt to discover other targets, ESX Server prevents it from discovering any SCSI information that is not appropriate to its isolated and virtualized view of its storage environment.

VMFS and VMotion

Additional complexities in the storage environment arise when a cluster of ESX Server hosts is accessing common targets or LUNs. As a distributed file system, VMFS ensures that all of the hosts in the cluster cooperate to ensure correct permissions and safe access to the VMFS volumes. File locks are stored on disk as part of the volume metadata, and all ESX Server hosts utilizing the volumes are aware of the ownership. Ownership of files and various distributed file system activities are rendered exclusive and atomic by the use of standard SCSI reservation primitives.

Each virtual disk (sometimes referred to as a VMDK file) is exclusively owned by a single powered-on virtual machine. No other virtual machine on the same or another ESX Server host is allowed to access that virtual disk. This situation does not change fundamentally when there is a cluster of ESX Server hosts, with multiple virtual machines powered on and accessing virtual



disks on a single VMFS volume. Because of this fact, VMotion — which enables hot migration of a virtual machine from one ESX Server host to another — is a protected operation.

Conceptually, the VMotion process is similar to the following steps:

- Power down a virtual machine.
- Release its exclusive lock on the virtual disk files associated with it.
- Power on the same virtual machine on another server.

At the end of the process, the powered-on virtual machine again gains exclusive access to the previously released virtual disk or disks. In practice, this process is more complex because VMotion preserves live virtual machine state while it moves the virtual machine from one ESX Server host to the other. This is analogous to the way physical machines arbitrate access to shared storage in active or passive clusters such as those created with Microsoft Cluster Server.

Raw Device Mappings

VMware has received requests from vendors of certain SAN management, volume management, or clustering software to allow these applications (running in a virtual machine) to see the physical storage to a greater degree.

VMware's focus on protection and isolation of virtual machines allows these applications to control and manipulate only the virtual environment presented to them, thus preventing them from controlling certain aspects of the physical storage world— such as manipulating a snapshot or taking other similar actions on the SAN device.

In response to these requests, VMware has enabled an advanced storage option known as raw device mapping, which allows an ESX Server administrator to configure certain virtual machines so they can see the details of an underlying physical LUN.

Note that raw device mapping provides access to the LUN, not the fabric. While raw SCSI commands can now be passed through to the LUN, the virtual machine does not have the ability to see the SAN or scan the targets. As a result, virtual machines with raw device mappings behave exactly like virtual machines with virtual disks.

With this ability to see LUNs, certain applications, such as array-based replication, can now function as expected. The ESX Server administrator can grant these privileges only at the time the virtual machine is configured, and the feature is configured at the level of individual virtual machines. In other words, it is possible to have virtual machines using raw device mappings and virtual machines using standard virtual disks on the same ESX Server host. Because of the isolation guarantees that remain in place, these virtual machines have no knowledge of each others' storage.

N_Port ID Virtualization

N_Port ID Virtualization (NPIV) is a feature of newer Fibre Channel HBAs that parallels the software-based storage isolation functionality provided by ESX Server. With NPIV, each virtual machine can have its own WWN, which is aligned with the SAN administrator's traditional view of a 1:1 mapping between the server (in this case a virtual machine) and a WWN.

The primary difference between NPIV and the storage isolation in ESX Server is that a per-virtual-machine WWN is now exposed to the physical world. This feature primarily benefits SAN management applications, as they now have the ability to show per-virtual-machine WWNs, rather than per-host WWNs that may be associated with multiple virtual machines.



With NPIV, SAN administrators see the virtual machines and can identify their traffic and add a redundant layer of LUN masking and quality-of-service controls, but the virtual machines still are isolated as before and do not have any additional visibility into the SAN. They still do not see Fibre Channel HBAs and cannot see the Fibre Channel targets or scan for LUNs that are not explicitly assigned to them by the ESX Server administrator.

Because of these qualities, VMware believes that NPIV is valuable for customers with more complex installations, where SAN management is an entirely separate discipline and organization from server management. And because of this belief, VMware, working in conjunction with our Fibre Channel HBA partners, has prototyped versions of ESX Server utilizing NPIV and has demonstrated high-end features such as VMotion. Utilizing NPIV, a WWN can be uniquely associated with a virtual machine, and it remains associated with that virtual machine even as the virtual machine is dynamically transferred, using VMotion, across physical ESX Server hosts in a virtualized environment.

Continuing Innovation

VMware ESX Server provides strong security and performance isolation for virtual machine storage. This level of SAN security has given more than 20,000 customers the confidence to deploy ESX Server — in the majority of cases, using Fibre Channel SANs for storage.

VMware continues to innovate in the storage arena, to provide an ever-greater set of storage options. Raw device mappings are an addition to this portfolio of storage options that is attractive primarily to advanced users of ESX Server and Fibre Channel SANs. NPIV, which has been prototyped with ESX Server, is another potential addition to the list.

VMware, Inc. 3145 Porter Drive Palo Alto, CA 94304 www.vmware.com

Copyright © 1998-2006 VMware, Inc. All rights reserved. Protected by one or more of U.S. Patent Nos. 6,397,242, 6,496,847, 6,704,925, 6,711,672, 6,725,289, 6,735,601, 6,785,886, 6,789,156, 6,795,966, 6,880,022, 6,961,941, 6,961,806 and 6,944,699; patents pending. VMware, the VMware "boxes" logo and design, Virtual SMP and VMotion are registered trademarks or trademarks of VMware, Inc. in the United States and/or other jurisdictions. Microsoft, Windows and Windows NT are registered trademarks of Microsoft Corporation. Linux is a registered trademark of Linus Torvalds. All other marks and names mentioned herein may be trademarks of their respective companies. Revision 20060309 Item: ESX-ENG-Q206-199
