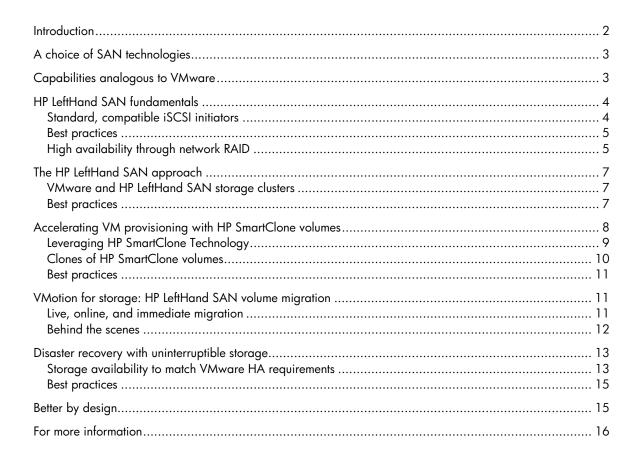


HP LeftHand SAN technology: enabling effective VMware deployments

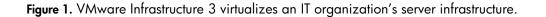
White paper

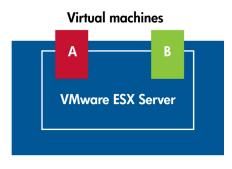


Introduction

This white paper describes how HP LeftHand SANs can help you deploy VMware Infrastructure 3 (VI3) software more effectively through volume snapshots, SmartClone volumes, volume migration, and high-availability features. With HP LeftHand SAN and VI3, organizations can reclaim valuable data center space and reduce their power and cooling requirements.

IT organizations are recognizing the benefits of virtualizing their data centers, and many of them are enabling server virtualization with VMware Infrastructure 3. VI3 allows IT organizations to manage their server infrastructure as a single pool of resources on which applications can be deployed and then moved dynamically from server to server as workload and business requirements dictate. (See Figure 1.) Some VI3 features, including VMware High Availability (HA), VMotion, and the Distributed Resource Scheduler (DRS), require shared storage.





With applications and their host operating systems encapsulated into individual virtual machines, they are no longer bound to a specific physical server. IT organizations can finally break free of the legacy of silo-oriented applications that limit sharing and require each application to run on its own dedicated set of underutilized hardware. By virtualizing and consolidating their applications onto a smaller number of more highly utilized physical servers, organizations can reclaim valuable data center space and reduce their power and cooling requirements. When servers are decoupled from applications, they are free to allocate resources to applications as their workloads change and as their business priorities dictate. Rolling out new applications does not always require the deployment of new servers because they can draw on the power of a shared infrastructure.

VMware Infrastructure 3 offers compelling features for IT organizations concerned with resource management, high availability, and disaster recovery. VMware VMotion allows virtual machines to be moved from server to server while running. Thus administrators can balance workloads and performance to match user demands—and even vacate a server in order to take it offline for maintenance. VMware HA monitors the state of running virtual machines and, in the event of a server failure, restarts the virtual machines on alternate servers.

A choice of SAN technologies

Features such as VMotion and VMware HA require continuous access to the underlying storage from multiple servers, which in turn requires the implementation of a shared storage infrastructure. VMware allows customers to choose between traditional Fibre Channel-based SANs, network attached storage, and iSCSI SANs. There are several advantages to using HP LeftHand iSCSI SANs.

HP LeftHand SANs use simple and cost-effective technology that helps IT organizations avoid the cost of implementing storage-specific networks, with their specialty host bus adapters (HBAs), switches, and optical interconnects. In addition, HP LeftHand SANs provide specific features that virtualize storage in ways that are similar to the way VMware virtualizes server resources. (See Figure 2.)

Figure 2. HP SAN/iQ[®] Software virtualizes the storage infrastructure with logical volumes supported by storage clusters.



HP LeftHand SANs are based on intuitive HP SAN/iQ Software and use standard IP networking, a technology that every IT organization understands and has existing infrastructure to support. Ethernet networks are everywhere, so equipment to support them is priced low by a highly competitive marketplace. This significantly reduces the cost of iSCSI SANs. The platform allows IT organizations to add storage as needed, incorporating new modules in its clusters dynamically and transparently.

Capabilities analogous to VMware

HP LeftHand SANs are certified to work with VMware, and are listed on VMware's Storage/SAN Compatibility Guide for ESX Server 3.x. In addition, several HP LeftHand SAN features complement VMware's virtualization capabilities by performing functions in storage that are analogous to those performed by VMware in servers:

- Organizations can clone storage volumes just as VMware clones virtual machines, but creating SmartClone volumes at the storage level allows duplicate volumes to be created and mounted in an instant for read/write activity—with minimal use of storage.
- Just as VMotion can move running virtual machines from server to server, HP SAN/iQ Software can move virtual disks from storage cluster to storage cluster without taking down or reconfiguring the virtual machine. This feature allows logical volumes to be moved to different storage clusters to balance storage capacity and workloads. You can even offload a cluster completely so that it can be taken out of service for upgrade or maintenance.

- The HP LeftHand SAN allows your storage infrastructure to be treated as a single, continuously available virtual pool of storage.
- You even can distribute the pool of storage among multiple sites, with the HP LeftHand SAN's builtin replication function maintaining identical copies of your data in two locations. VMware can restart failed virtual machines in alternate locations even though the failure of a data center has taken the primary site's servers and storage offline.

HP LeftHand SAN fundamentals

The HP LeftHand SAN uses standard iSCSI protocols and standard IP networking techniques to deliver clustered storage that is compatible with VMware. This section discusses some of the technology behind the HP LeftHand SAN and how it can best be leveraged with VMware products.

Standard, compatible iSCSI initiators

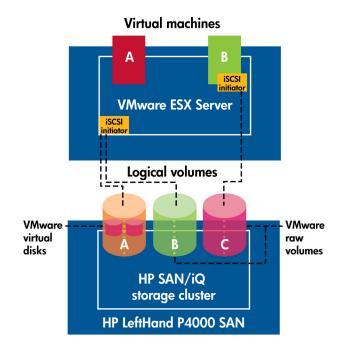
The HP LeftHand SAN supports standard iSCSI protocols, providing compatibility with virtually any operating system having an iSCSI initiator. The use of standard, open-source iSCSI protocols allows servers to utilize storage without the need to install any third-party server software.

The operating system's iSCSI initiator acts as if it is controlling a physical SCSI device, which makes the underlying IP network invisible to applications. When the iSCSI initiator is initialized, it is configured with a set of iSCSI qualified names (IQNs). These IQNs map to a set of IP addresses, where the initiator can discover which logical unit numbers (LUNs), or logical volumes, are available to it. This mapping function is important because it provides flexibility in storing logical volumes on the network.

Best practices

The use of standard, compatible iSCSI initiators allows servers running VMware to access HP LeftHand SANs through initiators in the host operating system, in one or more guest operating systems, or both. (See Figure 3.)

Figure 3. Virtual disks must be accessed through the ESX Server iSCSI initiator; raw volumes can be accessed by ESX Server or its guest operating systems.



If your guest operating systems boot from virtual disks stored in iSCSI storage, they must be accessed through the iSCSI initiator in VMware ESX Server or through an iSCSI HBA. These disks appear to virtual machines as local SCSI drives.

In addition, data that you wish to manage independently of VMware can be accessed from iSCSI initiators in the guest operating systems. This allows you to overcome the limit of 64 iSCSI targets that VMware ESX Server currently imposes, and it also allows you to manage large amounts of application-specific data (such as a database) directly through the SAN's management interfaces.

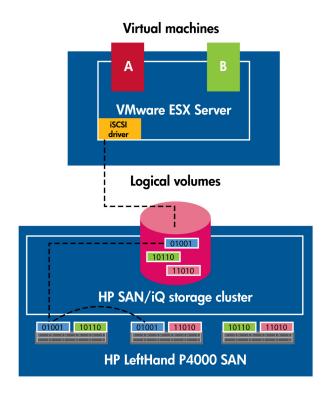
High availability through network RAID

HP LeftHand SANs implement network RAID across a cluster of storage modules. The cluster stripes data across storage modules in the cluster, and replicates blocks as dictated by each volume's replication level. You can scale a storage cluster by simply adding a new network storage module. Internal to the cluster, HP SAN/iQ Software then re-stripes the data to incorporate the new device. This is analogous to adding a new server to a resource pool in VMware Infrastructure 3. You can add a server, and VMware Dynamic Resource Scheduler (DRS) will then rebalance its managed set of virtual machines across the newly enlarged VI3 cluster.

Storage clusters are available to the server's iSCSI initiator through a virtual IP address (VIP). The use of a VIP virtualizes the cluster so that it can balance both the workload and the storage of specific blocks within the cluster while keeping this balance transparent to the servers that access it. If the storage module answering the VIP fails, the cluster will continue to respond to requests from one or more alternate servers in the cluster.

As data is written to the cluster, the single storage module receiving a data block from the server replicates the block to additional storage modules, in real time, according to the logical volume's network RAID level. Regardless of network RAID level, blocks are striped across the cluster for high performance. (See Figure 4.) The result is that the larger the cluster, the higher its performance. Each storage module's disks are typically combined as a RAID 5 volume internally, and the high availability of the cluster increases with the replication level assigned to each logical volume.

Figure 4. Blocks are striped across the storage modules in a storage cluster for performance, and they are replicated as dictated by the logical volume's properties (a replication level of 2 in this example).



The advantage of RAID is that it allows highly available storage to be created from a set of inexpensive disk drives. Even if one drive in a RAID array fails, the data will be continue to be available. Similarly, data will continue to be available even if any one storage module fails—or even if part of the cluster in a different location stops working.

The HP LeftHand SAN approach

What makes the HP LeftHand SAN superior to traditional SAN implementations is that it doesn't require expensive, purpose-built components, but instead uses hardware whose cost has been minimized by the forces of a competitive marketplace. Reliability is increased over that of a single server because data is replicated across storage modules. Unlike traditional SAN hardware, which typically has a refresh cycle of 24 to 36 months, HP LeftHand SANs benefit from the faster server innovation cycle of nine to 12 months, bringing your storage in line with your server technology.

The solution is also more flexible than traditional approaches because it allows your IT organization to support applications with storage policies that can be configured on a highly granular basis. The network RAID level is a property of the logical volume, not of the storage cluster. So network RAID, including block replication levels, can be assigned to each logical volume based on the requirements of the data stored there. You can change replication levels, and the cluster will automatically accommodate your changing requirements.

VMware and HP LeftHand SAN storage clusters

Storage is managed on any device through LUNs, or logical volumes. VMware provides access to virtual disks through a virtual SCSI device that is provided by each virtual machine to its guest operating system. This simplifies the relationship between applications and storage. It also allows guest operating systems to access their virtual disks through the same interface regardless of the underlying hardware platform. Whether a virtual disk is stored on direct attached, network attached, or SAN attached storage, the guest OS views a virtual disk through a SCSI device initiator.

Virtual disks can be stored as flat files, in which each block of the flat file simulates a block on a real disk drive. Many of these flat files are typically stored in a single logical volume. Virtual disks can also be stored as raw device-mapping volumes, where a virtual disk is assigned its own logical volume and the guest operating system manipulates blocks directly.

Best practices

The value-added features that the HP LeftHand SAN provides to VMware ESX Server are available on a logical volume basis. Replication, for example, is performed on logical volumes, and snapshots used for backups are taken of logical volumes. Even features like volume cloning are provided on a logical volume basis.

You should therefore take care to store related virtual disks together on the same logical volume. If, for example, you expect to replicate a Web server environment, consider putting all of its related virtual disks into a single logical volume that can be replicated as a unit. If you want to back up an entire application stack, it's best to place all of the related virtual disks into a single logical volume so that a single snapshot can capture all of the application OS, software, and data at once.

Large amounts of data—for example an enterprise database—are best stored in raw device-mapping volumes, where each block can be accessed through only one layer of indirection. This is preferable to the two layers of indirection imposed by storing a virtual disk as a flat file in a file system. Storing such data in a raw volume allows more flexibility, because the volume can be attached to a virtual machine as a virtual disk—or accessed directly from a guest operating system through its own iSCSI initiator.

Accelerating VM provisioning with HP SmartClone volumes

VMware Infrastructure 3 makes it easy to create and provision new virtual machines based on existing ones, a process referred to as cloning. Cloning adds to an IT organization's flexibility, responsiveness, and efficiency by allowing whole new environments to be created without being encumbered by physical hardware.

Organizations can use cloning in a number of ways. They can create and deploy a set of temporary Web servers to manage performance and capacity in response to spikes in workloads. They can reduce risk by cloning existing mission-critical applications and using the copies to test new patches or configurations before deploying them on production systems. Developers can create virtual laboratories with as many virtual machines as they need to develop and test their software. IT organizations can create "golden master" versions of their desktop environments and deploy new virtual PCs for users based on consistent, tested configurations that help to reduce the cost of maintaining a unique environment for each user.

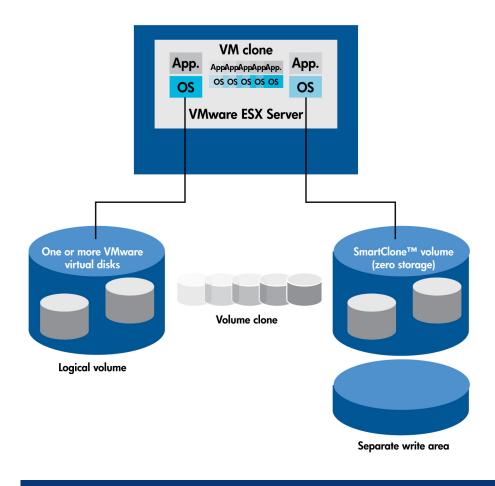
Cloning virtual machines through VMware VirtualCenter is a straightforward process that copies an existing VM's virtual disks, creates a new virtual machine based on the existing one's characteristics, and attaches the copied virtual disks to the new VM. This is an easy way to create new environments for long-term use. But there are two problems using VMware virtual machine cloning for short-lived virtual machines:

- The cloning process creates a new copy of each virtual disk, which takes a considerable amount of time when each virtual disk is tens of gigabytes in size.
- Each new virtual disk uses physical storage, depleting storage resources for data that, for the most part, won't vary from its original source.

Leveraging HP SmartClone Technology

When virtual machines have large amounts of data or are to be copied numerous times, a better solution is to use volume cloning in the storage cluster itself. The HP LeftHand SAN includes HP SmartClone[™] Technology, a block-level cloning mechanism that supports multiple clones of the same logical volume, each of which can be mounted for read/write operation. (See Figure 5.)

Figure 5. SmartClone capability allows VMware to use volume clones that can be modified without affecting the original source.



Using the HP SmartClone feature when creating a new virtual machine, you can create new virtual disks instantaneously, yet consume physical storage only for the disk blocks that are modified (plus a small amount of metadata). Using this approach, you can reduce staff time, increase responsiveness, and make more efficient use of your storage resources.

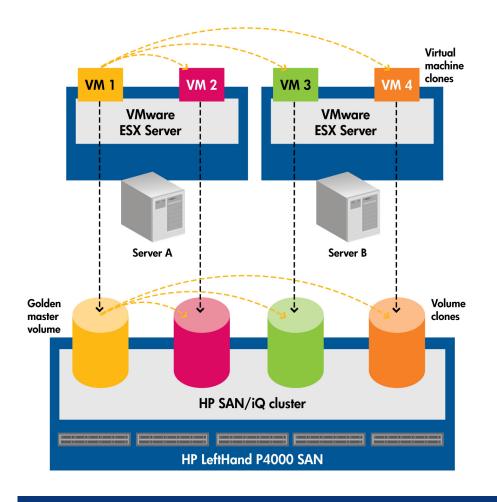
The workflow for creating a clone of a virtual machine using HP SAN/iQ Software is straightforward:

- 1. Use the HP LeftHand SAN interface to create a clone of the logical volume containing the virtual disk files you need to clone.
- 2. In VMware VirtualCenter, click **Rescan** to locate the new storage. The resignaturing option must be turned on in VMware ESX Server.
- 3. Add the new virtual disks to Virtual Infrastructure 3 software's inventory.

Clones of HP SmartClone volumes

The HP LeftHand SAN allows you to create clones of HP SmartClone volumes based on a single original volume, where each clone is independent of the other clones. (See Figure 6.) When you create a SmartClone volume, the HP LeftHand SAN also creates an independent write area holding the collection of blocks that are modified from the original, and the clone's metadata points back to the original unmodified blocks.

Figure 6. HP SmartClone Technology instantly makes space-efficient volume copies for use by virtual machines and physical servers.



The ability to have several SmartClone replicas of the same logical volume, each capable of being independently read and written to, allows you to have many virtual machines that start with the same "golden master" image, but consume only a minimal amount of physical storage space.

Best practices

The best way to use SmartClone volumes is to create them using the HP LeftHand SAN user interface, adding a comment for each clone that indicates its use. The comment allows administrators to evaluate a clone by tracing it back to the virtual machine that uses it. To assist in SAN administration, the HP LeftHand SAN also includes a map view detailing the relationship between volumes, their snapshots, and their clones.

VMotion for storage: HP LeftHand SAN volume migration

VMware VMotion allows you to move running virtual machines (VMs) from server to server without interrupting the services those VMs provide, which means you can dynamically change the mapping of virtual machines to servers. VMotion is the technology on which the VMware Dynamic Resource Scheduler (DRS) is based. It supports:

- Managing performance by moving virtual machines with high resource demands to lessutilized servers
- Re-balancing workloads after adding one or more servers to a resource pool
- Matching virtual machines to server capabilities as workloads change
- Moving volumes between physically different clusters, such as from SAS to SATA disk types, or between storage modules configured with RAID 0 to RAID 5
- Vacating a server for maintenance or upgrades

For all the same reasons, your IT organization must also to be able to move logical volumes between storage systems. Unfortunately, traditional Fibre Channel SANs fall short in the capabilities they provide. Moving a logical volume typically requires you to quiesce the data by shutting down the corresponding virtual machines; perform the copy using remote replication or remote copying software (often an additional-cost feature); reconfigure the virtual machines to access the relocated logical volume by its new identifier; and, finally, restart the virtual machines. This is a tricky process, prone to error, and it can result in significant downtime depending on the amount of data copied.

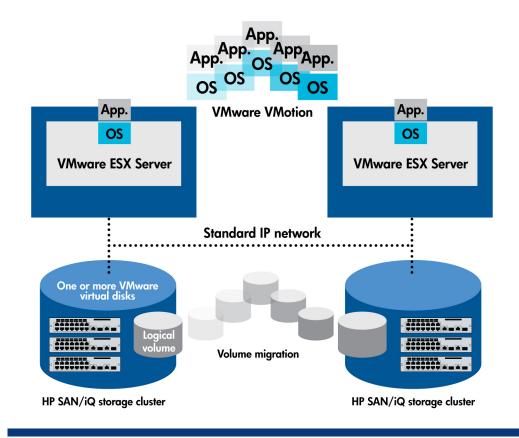
Live, online, and immediate migration

The HP LeftHand SAN allows IT organizations to move live logical volumes between storage clusters without interrupting the virtual machines using them—and without the VMs even being aware that their storage has been moved. IT organizations can thus enjoy the same benefits in storage as those they enjoy when using VMotion with their servers. HP LeftHand SAN volume migration allows you to:

- Manage performance by moving logical volumes with high I/O demands to less-utilized clusters
- Re-balance storage workloads after adding new storage clusters
- Match logical volumes to the performance characteristics of different storage clusters
- Move logical volumes between geographically separated clusters
- Vacate a cluster for maintenance or upgrades

VMotion and volume migration are analogous: The former moves virtual machines between servers; the latter moves logical volumes between storage clusters. These two features are independent of and transparent to each other. A storage cluster does not know when a virtual machine has been moved, and a virtual machine does not know when its virtual disks have been moved. (See Figure 7.) The IP address of the virtual machine remains the same when it moves, and when a logical volume is moved, the logical unit number remains the same as well.

Figure 7. Just as VMotion moves running virtual machines between servers, the HP LeftHand SAN migrates live volumes between clusters.



Behind the scenes

When two or more clusters are in the same HP LeftHand SAN management group, moving logical volumes between them is as simple as editing a volume's properties so that it has a different cluster than its owner. The volume's metadata is moved first, and then the volume's data blocks are moved asynchronously.

The servers or virtual machines accessing the logical volume continue to make requests to the original storage cluster. The original cluster acts as a proxy, passing requests on to the new cluster depending on which cluster currently owns the block being requested. If the original cluster is taken out of service or if the server accessing the storage is rebooted, the server's iSCSI initiator will discover the migrated volume by finding the same IQN on a new cluster.

Disaster recovery with uninterruptible storage

VMware HA provides cost-effective high availability for applications that do not have their own built-in HA mechanisms, providing organizations with a base level of disaster recovery capability regardless of how simple or complex the application. VMware HA continuously monitors server hardware, and automatically restarts any affected virtual machines on secondary servers in the event of a failure. By locating parts of a VMware server cluster in different data centers on the same campus—or even in different closets in the same building—IT organizations can withstand a disaster that makes one of the two sites unavailable.

For VMware HA to work with no human intervention, it requires continuous access to every virtual machine's virtual disks. In general, traditional Fibre Channel–based SANs are built with redundant components to achieve high availability levels. To survive the loss of an entire data center, however, they typically use remote replication software that adds to the cost of the SAN.

An issue with some with remote replication solutions is that a replicated logical volume has a different unique logical unit number (LUN) than the original. If an entire data center fails, VMware HA can restart virtual machines in the alternate location, but the server will fail to find the replicated storage because it has a different LUN. At this point, human intervention is required to reconfigure each server and virtual machine so that it accesses the appropriate LUN. Furthermore, when the original data center comes back online, a slow, manual process of re-synchronizing the replicated volumes with their original versions must be completed before virtual machines can be moved back to their original servers.

Storage availability to match VMware HA requirements

To deliver the continuous access to storage required by VMware HA, the HP LeftHand SAN simply splits a storage cluster between the primary and backup sites, and provides a high-speed link between them.

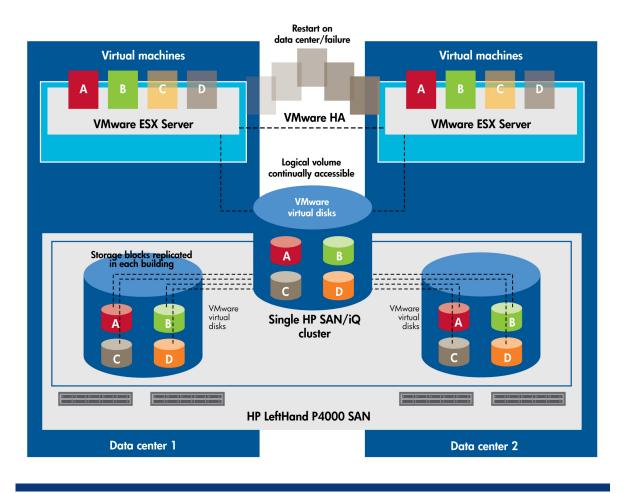
The cluster provides continuous availability as long as the following configuration conditions are met:

- An even number of storage modules is configured in the cluster.
- The logical volume data is replicated.
- Even-numbered modules are in one location and odd-numbered modules are in the other location.

This configuration helps ensure that every block stored in a logical volume is replicated between the two locations. The cluster automatically responds to the failure of any of its servers with the remaining servers responding to requests, regardless of whether the cluster is located entirely in one location or split between two locations.

Consider a simplified example where two data centers each host one physical server, and both are members of the same VMware HA cluster. (See Figure 8.) Data center 1 runs virtual machines A and B, and data center 2 runs virtual machines C and D. The virtual disks for each VM are stored in a single logical volume. As long as the logical volume's properties specify a network RAID replication level of 2, both locations will have an identical set of blocks because every block will be replicated across even and odd storage modules.

Figure 8. A storage cluster with an even number of storage modules can provide continuous availability to VMware HA, even with the loss of an entire data center.



If the power in data center 1 fails, VMware HA will restart virtual machines A and B on the server in data center 2, booting from virtual disks A and B. Unknown to VMware ESX Server, the HP LeftHand SAN recognized the failure of the modules located in data center 1, and it may have moved the ownership of the cluster's VIP address from a storage module in data center 1 to a module in data center 2. The failure of an entire data center is handled just as the failure of one or more servers in a cluster is handled. Access to storage is never interrupted, and the virtual machines restarted by VMware HA simply access their virtual volumes just as they did in data center 1. The cluster's IP address and the volume's LUN are unchanged.

Using a distributed storage cluster allows failover to execute seamlessly. Unlike failover using SANs and remote replication, failback is automatic and seamless. When the power is restored to data center 1, VMware can automatically rebalance the workload, using VMotion to move virtual machines A and B back to the server in data center 1. The modules in data center 2 continue to respond to storage requests. When the modules in data center 1 come back online, the cluster recognizes that they have out-of-date data blocks, and the cluster—internally and transparently—reconciles the differences by copying the changed blocks back from data center 2 to data center 1, restoring the desired replication level of 2.

The advantage of this solution is that both failover and failback are automatic, requiring no human intervention. Combined with VMware HA, this solution provides a disaster-recovery strategy for even the simplest of applications that encompasses both servers and storage. Moreover, it's a strategy that is suitable for organizations of all sizes. Small businesses can distribute their storage clusters across multiple closets in the same building; medium and large businesses can disperse their storage clusters across multiple data centers on a campus.

Best practices

This solution works when the following requirements are met:

- The storage cluster is split between two locations connected with a minimum of 1 Gb/second bandwidth and less than 2 milliseconds of latency. Storage blocks are replicated between even and odd storage modules synchronously, requiring a high-speed network connection between all modules. Likewise, requests coming into one half of the cluster may be load-balanced to the other half of the cluster, so network performance between cluster halves is important.
- Either a Virtual Manager or another storage module in a neutral location is required. The Virtual Manager allows for manual failover control, while using another storage module provides failover automation.
- The storage cluster must operate on the same IP subnet, as it maintains one and only one VIP address for incoming requests.
- An even number of storage modules must be deployed, with even-numbered storage modules in one location and odd-numbered storage modules in another.
- Logical volumes participating in the disaster-recovery plan must have their replication level set to 2 so that each block is stored in both locations.

Large companies can extend this solution to multiple data centers. For example, a cluster can be distributed across three locations, with storage modules deployed in threes; that is, storage modules numbered in multiples of 1, 2, and 3 are stored in their respective locations.

Similarly, smaller companies can achieve some disaster-recovery capability by locating cluster halves in different data centers, offices, or even closets in the same building. The value of the redundancy can be increased by powering and cooling the two halves of the cluster from different circuits, since power and cooling failure is a cause of data center downtime that is often overlooked.

Better by design

This white paper has focused on three specific capabilities of VMware Infrastructure 3 (VI3) that are complemented in storage by the HP LeftHand SAN:

- 1. You can use the HP LeftHand SAN to speed virtual machine cloning by replicating logical volumes using SmartClone functionality.
- 2. You can move live storage between storage clusters to optimize performance or capacity, or for administrative or business reasons, without interrupting the virtual machines that access the storage—indeed without them even knowing. HP SAN/iQ Software technology lets you create distributed storage clusters that automatically handle failover and failback so that VMware HA can operate with no human intervention.
- 3. Finally, using the HP LeftHand SAN through your operating system's built-in iSCSI initiators allows you to connect logical volumes to VMware ESX Server or to individual virtual machines, giving you control over how storage is configured and accessed.

With all of the features that directly benefit IT organizations running VI3, it may be easy to overlook the fundamental features of the HP LeftHand SAN that make it a superior solution for any data center environment. The HP LeftHand SAN delivers a network-centric, distributed, and clustered storage solution that offers inherent benefits. Because it is highly scalable, IT organizations can add storage modules as they are needed, knowing that the HP LeftHand SAN can incorporate the new capacity automatically, seamlessly, and with no downtime. In addition, the HP LeftHand SAN is a highly available solution that uses network RAID to achieve high availability levels with cost-effective, off-theshelf, state-of-the-art x86 server technology. The HP LeftHand SAN delivers a high level of reliability, allowing IT organizations to "dial in" the desired level of protection, most often on a per-volume basis.

HP SAN/iQ Software provides multiple layers of protection, including RAID on the individual storage modules, network RAID, snapshots, and remote copy. The HP LeftHand SAN delivers high performance without high price by using available resources from multiple spindles in each storage module and delivering them to multiple, aggregated network links.

Finally, the HP LeftHand SAN is easy to use. Thanks to its intuitive graphical user interface, customers can configure and manage their storage while allowing the storage clusters themselves to implement specific directives—from creating SmartClone volumes that support new virtual machines to incorporating new storage modules in a cluster. The HP LeftHand SAN's support for VMware, coupled with the overall benefits of using a network-centric, distributed, and clustered storage solution, makes it a natural choice for any IT organization running VMware Infrastructure 3 software.

For more information

For more information on HP LeftHand SANs, visit <u>www.hp.com/go/p4000</u>

Technology for better business outcomes

© Copyright 2009 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

