Integrating EMC MirrorView with VMware ESX Server for

Business Continuity and Disaster Recovery

Dell™ PowerEdge™ servers running VMware® ESX Server virtualization software provide multiple host-level data protection features. However, none of these features can prevent catastrophic site-level failure resulting from serious accidents or natural disasters. Dell engineers integrated ESX Server with EMC® MirrorView™ software in a Dell/EMC storage area network to demonstrate site-level disaster recovery.

BY JACOB LIBERMAN AND DAVID ULBRICH

Related Categories:

Backup, recovery, and archiving (BURA)

Dell/EMC storage

Disaster recovery

Storage management

Storage software

Systems management

Virtualization

VMware

Visit www.dell.com/powersolutions for the complete category index.

Data security on mission-critical servers is a top priority for scalable enterprises. Traditionally, enterprises have achieved acceptable levels of business continuity by using costly reliability enhancements such as RAID configurations, uninterruptible power supplies, independent utility grids, backup generators, and redundant wide area network and Internet links, all of which can help dramatically improve individual server and single-site data center availability.

Dell PowerEdge servers running VMware ESX Server can provide building blocks for a cost-effective virtual infrastructure while enhancing the fault-tolerant capabilities enterprises demand. These enhancements include encapsulated virtual machines (VMs) that can be rapidly redeployed across different underlying server hardware architectures, along with the ability to cluster physical servers using VMs, which can help improve availability without additional hardware investment. Additionally, VMware VMotion™ technology allows administrators to seamlessly migrate running VMs between physical servers, which can help eliminate the need to schedule downtime for hardware maintenance.

However, none of these high-availability features can ensure business continuity following a catastrophic failure of the entire data center or hosting facility. Accidents, terrorist attacks, and natural disasters pose serious threats to data security and could take days or weeks to offset using traditional recovery methods. EMC MirrorView software, deployed in conjunction with an ESX Server virtual infrastructure on a Dell/EMC storage area network (SAN), can help provide data security and business continuance following a catastrophic site failure.

In May 2006, a team of Dell Enterprise Services engineers conducted business continuance tests to demonstrate the benefits of MirrorView and ESX Server integration. This article describes these tests and how enterprises can carry out such an integration.

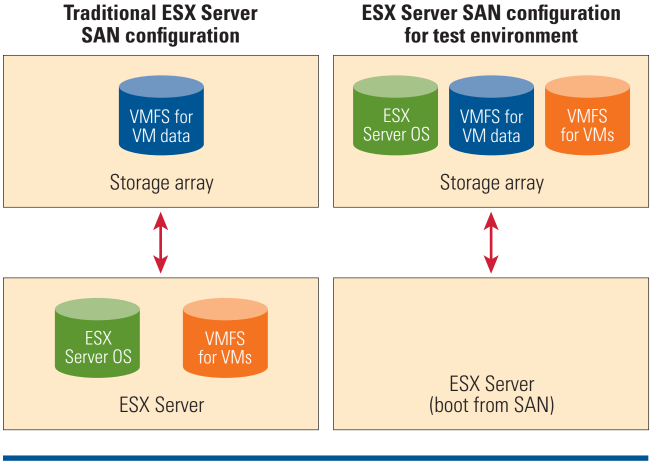

Figure 1. Traditional VMware ESX Server SAN configuration compared with the test configuration

Preparing the environment

The Dell team began by implementing a SAN typical of enterprise environments. It comprised two Dell/EMC storage arrays connected with two Brocade SilkWorm 4100 switches. The team designated a Dell/EMC CX600 as the primary production array and a remote CX400 as the secondary business continuance array. Both arrays were updated to base code release 02.19.x00.5.007.

Next, the team fabric-attached one Dell PowerEdge 1850 server to each array using dual QLogic QLE2360 host bus adapters (HBAs). Although ESX Server is typically installed on the local server hard disks, the team chose to use the ESX Server boot-from-SAN capability by deploying the primary server's OS directly to a 30 GB logical unit (LUN). Deploying the server OS directly to the SAN has two benefits. First, a single central installation is easier to deploy and maintain than multiple installations. Second, installing the OS on a LUN makes it compatible with MirrorView remote replication—a crucial consideration for disaster recovery.

After creating and presenting the LUN to the primary server, the team identified it as the boot target in both the server and HBA BIOS. Next, the team booted the server from the ESX Server 2.5.3 CD and entered "Boot from SAN" at the installation prompt.[1] The only modifications made to the default disk partitioning scheme were to increase the size of the root partition to 10 GB to allow for log growth, and to create a 2 GB swap partition to allow proper functionality of the ESX Server service console.

Two important considerations are worth noting when configuring boot from SAN. First, when configuring the server's storage group through the EMC Navisphere® Manager user interface, administrators should always present the target boot disk as host LUN 0. Second, in ESX Server 2.5.*x*, administrators must place one of the server's HBAs in shared mode so the service console can use it. The second HBA should be dedicated to the VMs.

Once the host installation had completed, the team presented two additional LUNs to the server, both configured as additional VMware File System 2 (VMFS2) file systems intended to store the VMs and their data. The team installed Red Hat® Enterprise Linux® AS 4 as the guest OS on one of the file systems to simulate a production VM typical of an enterprise environment.

Implementing EMC MirrorView/Synchronous for data protection

To achieve business continuance across the two test sites, the team installed and enabled EMC MirrorView/Synchronous (MirrorView/S) software on both storage arrays. MirrorView/S is a software application that maintains a copy image of a LUN in two separate locations. The production image on the primary array is mirrored to a backup image on the secondary array, and any changes made to the primary image are written to the secondary image before they are committed to the primary array's disks.[2] Therefore, MirrorView/S is well suited for Class A site-level business continuance in which data loss and extended recovery time are not feasible.[3]

The team felt that previous MirrorView disaster recovery guidelines for ESX Server did not allow for true Class A business continuance. These guidelines assumed the ESX Server OS would be installed on the servers' local disks, and that only the disks' data would be mirrored between the primary and secondary arrays. If a primary site failure occurred, the guidelines recommended promoting the secondary image, re-creating the VMs on the ESX Server system attached to the secondary array, and then reconnecting them to their data on the mirrored LUNs. The Dell team felt that this scenario would lead to unsatisfactory recovery times. Therefore, rather than implement the traditional ESX Server architecture—in which the VMs and OS reside on local disks while the data resides on the mirrored LUNs—the team deployed the entire virtual infrastructure to the SAN disks. Figure 1 illustrates the differences between the traditional ESX Server SAN configuration and the test configuration.

After installing and enabling MirrorView/S on both storage arrays, the team connected the arrays physically and logically. The highest-numbered front-end port of each storage processor on the

[1] The ESX Server 2.5.*x* raw disk mapping (RDM) feature is not compatible with HBAs set in shared mode for boot from SAN, so all VMs in this virtualized environment must reside on servers using the VMware File System (VMFS). Although RDM is typically recommended for ESX Server interoperability with MirrorView, the May 2006 EMC Support Matrix indicates that MirrorView now fully supports VMFS provided the storage array is running base software release 19 or later. ESX Server 3 allows RDM and boot-from-SAN interoperability. For details, visit www.vmware.com/support/kb/enduser/std_adp.php?p_faqid=1458.

[2] Supported distances vary across connection topologies. MirrorView/S, when used in conjunction with long-wave Cisco Inter-Switch Links or dense wavelength division multiplexing, synchronously mirrors images between arrays over distances of up to 200 km. MirrorView/Asynchronous, when deployed in conjunction with Fibre Channel to IP converters, can support distances of hundreds to thousands of miles.

[3] Class A recovery goals include zero data loss, internal recovery times of less than 30 minutes, and external recovery times of less than 24 hours. For more details on these guidelines, see "Data Recovery Outside the Core Data Center," by Michael Kimble, *Dell Power Solutions,* May 2006, www.dell.com/downloads/global/power/ps2q06-20050307-Kimble.pdf.

CX400 was zoned to the highest-numbered front-end port of each storage processor on the CX600. Next, the team joined both storage systems to the same Navisphere security domain so they could be managed through the same management interface. The team allocated write intent logs—which track changes to the mirror source LUN that have not yet been written to the mirror's images—on both storage arrays to help provide fast recovery if the primary storage system failed. Finally, the team enabled MirrorView connections between the storage arrays using the Manage MirrorView Connections dialog box in Navisphere Manager.

The production ESX Server system at the primary site was provisioned with three LUNs: one for the OS image, one to store VMs, and one to store the VM data. The team created three LUNs of equal block sizes on the secondary array to act as remote images. Next, the team created secondary images for the primary LUNs on the equal-size LUNs on the remote storage array, which began the initial synchronization process in which all data on the primary LUNs was copied to the corresponding secondary images.[4]

Once the secondary images were fully synchronized, they were organized into a single consistency group, which is a set of mirrored images that must remain consistent with one another. Binding the three LUNs that make up the ESX Server system into a single consistency group helped prevent individual mirrors from being promoted, fractured, or destroyed without their appropriate counterparts.

Finally, the team configured the host at the business continuance site to boot from the secondary image of the OS LUN by adding the LUN and the server to the same Navisphere storage group and making the necessary changes in both the server and HBA BIOS. *Note:* This step did not entail installing a second instance of ESX Server. The OS was already "installed" on the mirrored LUN, although it would not be available until the secondary image was promoted.

Testing business continuance

The first test simulated a graceful transition to the business continuance site—the same type of test enterprises should perform before releasing MirrorView to production. To start the test, the first ESX Server host was shut down cleanly and powered down. Next, the team fractured the consistency group using Navisphere Manager, which in turn fractured all the mirrors in that consistency group simultaneously. The team then promoted the remote consistency group to primary image status and powered up the remote PowerEdge 1850 server using the embedded remote access controller. The PowerEdge 1850 booted to the ESX Server recovery image without incident. The first test was successful.

The second test was intended to simulate the type of sudden business disruption that would occur during a true site-level catastrophe. The team configured a VM running Red Hat Enterprise Linux on the ESX Server system. Next, the team wrote a short shell script to generate I/O to the data LUN by continuously writing 1 MB time-stamped files. Once the script was running, the team simulated a catastrophic failure by cutting the power to the host server and storage array.

After fracturing the mirror and promoting the secondary image, the standby PowerEdge 1850 booted successfully without incident. When the team powered up the guest VM on the new server, ESX Server recognized that it was in a different location and asked whether it should keep the old configuration file or create a new one.[5] The team chose to keep the existing configuration file because the VM was, in a sense, copied to a new location. The VM successfully booted to a crash-consistent state in the same way that a physical server would detect that it had been shut down uncleanly. However, the team encountered no problems with the boot, and the VM mirrored data LUN contained all files written by the script before the production server and storage array lost power.

Providing site-level disaster recovery

These tests demonstrate the benefits of integrating EMC MirrorView/S with VMware ESX Server for site-level disaster recovery. To provision a highly available virtualized environment, the recommended procedure is to configure Dell PowerEdge servers running ESX Server to boot from a fully redundant Dell/EMC SAN, duplicate the business-critical primary-site hardware environment at the remote disaster recovery site, and mirror all server LUNs between the primary and secondary storage arrays. The VMware VirtualCenter management host can also be configured to boot from the SAN so that the entire virtual infrastructure can be replicated to the remote environment.

Infrastructure virtualization with VMware ESX Server has emerged as a crucial building block in the scalable enterprise. As such, the technologies surrounding and supporting virtualization must meet the same data security standards that enterprises demand of those supporting physical infrastructures. MirrorView/S, deployed in conjunction with ESX Server in a Dell/EMC SAN, can help provide site-level disaster recovery and business continuance that meet Class A data security guidelines.

Jacob Liberman is a training consultant in the Dell Enterprise Training Services Group and is a certified EMC Implementation Engineer. He has a B.A. from the University of Wisconsin and is currently pursuing an M.Ed. in Instructional Technology from the University of Texas at Austin.

David Ulbrich is a server support adviser in the Dell Enterprise Expert Center and is a VMware Certified Professional. He has a B.A. in Plan II Honors from the University of Texas at Austin and is currently pursuing an M.B.A. from the McCombs School of Business at the University of Texas at Austin.

[4] An additional benefit of using MirrorView/S for remote data replication is that the replication I/O is handled solely by the storage arrays; therefore, the synchronization process can have a minimal performance impact on the attached hosts.

[5] This message is normal and to be expected because of the way ESX Server tracks VM universally unique identifiers. For details, visit www.vmware.com/support/kb/enduser/std_adp.php?p_sid=&p_faqid=1541.