Software to Simplify and Share SAN Storage

sanbolic

Enabling High Availability for Citrix® XenDesktop™ and XenApp™ - Which Option is Right for You?

White Paper

By Andrew Melmed, Director of Enterprise Solutions, Sanbolic, Inc.

June 2010

www.sanbolic.com

Introduction

With Citrix ***XenDesktop*** virtual desktop and ***XenApp*** application delivery infrastructures, the engine behind the dynamic provisioning of virtual machines (VMs) hosting virtual desktops and XenApp servers is Citrix ***Provisioning Services***™ (aka ***PVS***™). To ensure continuous uptime, PVS includes a high availability (HA) feature that allows the VMs to maintain access to their vDisk image files (contains OS, drivers, applications) when one or more PVS Servers fails unexpectedly. The prerequisite for PVS HA: the vDisks streamed to the virtual machines *must* be accessible by multiple PVS Servers at all times, allowing the VMs to re-establish connectivity to their vDisks through one of the remaining online PVS Servers.

Because numerous options exist that allow multiple PVS Servers to access the same vDisks, organizations planning XenDesktop and XenApp deployments often find it difficult to determine which option will not only meet their initial needs, but also affords them the ability to quickly and easily expand these infrastructures so they can realize the greatest return on their investments in these solutions.

The following list identifies the options most frequently considered by organizations looking to enable PVS HA to ensure maximum uptime for virtual machines hosting XenDesktop virtual desktops and XenApp servers:

- Replication
- Read-Only LUNs
- Shared Storage – Windows Shares or NAS
- Shared Storage – SAN storage with Sanbolic® Melio 2010™

To help organizations gain a better understanding of these options and which might prove most appropriate for their environment, this paper describes each of the options outlined above, how they work and the pros and cons associated with their use.

Options for enabling PVS HA

Replication

Description: Replication refers to provisioning multiple LUNs on SAN storage and assigning one LUN to each PVS Server, ensuring that each PVS Server has sole access to its own respective LUN.

How it works: After a vDisk is created in a PVS data store located on a LUN, the vDisk gets copied to every other LUN so that every PVS Server can stream the vDisk to target devices (VMs) simultaneously.

Pros: Good performance can be achieved as each vDisk is streamed to target devices from multiple PVS Servers in tandem via separate data paths (fibre channel or iSCSI) to individual LUNs (one LUN per server).

By its very nature, vDisks are protected since the latest version of each vDisk is stored in multiple locations, with each LUN basically serving as a backup for all other LUNs. This ensures the vDisks are available in the event that one or more LUNs become inaccessible or unusable.

Cons: Although this option meets the requirement of allowing multiple PVS Servers to access the same vDisk image files so PVS HA can be enabled, it also introduces additional management overhead that can be both time-consuming and error prone as every time a vDisk is modified (i.e., contents and/or properties), the vDisk has to be copied to every other LUN used by the other PVS Servers in order to maintain vDisk version consistency among all the PVS Servers in the PVS Farm. Not only can this introduce the potential for vDisk versions being out of sync, it also introduces additional network traffic that can adversely affect application performance each time a modified vDisk is copied over the LAN.

Multiple LUNs must be provisioned and maintained and multiple PVS data stores must be created in the PVS Administrator console (one for each LUN), each of which can only be accessed by a single PVS Server. This makes scaling the solution more difficult as a new LUN must be provisioned and maintained every time another PVS Server is added to the PVS Farm.

Although LUNs may be expandable (if the storage array supports LUN extension), they cannot be expanded without first stopping all I/O activity against the volume. This requires that the Stream service on the PVS Server connected to the LUN be stopped whenever additional storage capacity is needed to ensure that all target devices accessing their vDisks on the LUN fail over to another PVS Server to avoid downtime.

While replication provides high availability of vDisks, it does not offer high availability for the PVS database as there is no inherent mechanism that allows a PVS database to be hosted by another server if the initial server hosting the database were to fail unexpectedly. To obtain this type of functionality, another technology, such as Microsoft® MSCS™ or Windows Failover Clustering™ for SQL Server™, would need to be implemented, introducing additional configuration and management to the overall solution.

It should also be noted that although all vDisk Access modes are supported, the write cache files for target devices *must* be stored on a device's local disk or in device RAM. Since shared storage cannot be used, consolidation of write cache files cannot be achieved.

Lastly, using separate LUNs to store multiple copies of the same vDisk image files leads to poor utilization of storage resources.

Read-Only LUNs

Description: With the release of PVS 5.1, Citrix introduced the concept of "Read-Only" LUNs for PVS HA, which allowed PVS vDisks (VHDs) to be stored on SAN storage for high availability without the use of mandatory locking mechanisms such as NAS gateways or cluster file systems to maintain the integrity of the vDisks as multiple Provisioning Servers shared access to a LUN containing the PVS data store.

In the latest release, PVS 5.6, the "Read-Only" LUN feature has been enhanced to make it easier for PVS administrators to maintain (update) vDisks stored on "Read-Only" LUNs. To accomplish this, data stores can now be configured to operate in "Managed" mode, allowing administrators to update vDisk image files by toggling between two modes of operation using a separate management utility.

How it works: In order for "Managed Read-Only" stores to be used, several prerequisites must be met:

1. PVS Servers must be running Windows Server 2008 R1 or R2.
2. PVS Servers must connect to the "Read-Only" LUN containing the "Managed Read-Only" store via iSCSI.
3. vDisks placed in the "Managed Read-Only" store must be created and stored in another location that supports both read and write access.

4. The LUN containing the “Managed Read-Only” store must be configured to operate in shared read-only mode or shared read/write mode without requiring the use of a separate locking mechanism to ensure the integrity of the data on the LUN is not comprised as multiple PVS Servers read and write to it at the same time.

Note: *Citrix recommends that* ***separate*** *“Read-Only” LUNs be used to store each Standard Image vDisk, allowing vDisks to be taken offline for maintenance independently of one another. While this approach may suffice for a small number of vDisks, for implementations using a larger number of vDisks, following this recommendation introduces additional management overhead as multiple LUNs would have to be created, managed and maintained. This requires PVS administrators spend more time managing storage resources.*

After configuring a “Managed Read-Only” store, a new management wizard called the “Store Management Wizard” is used to 1) designate a “Maintenance Server” and 2) toggle the data store between “Active” (read-only) mode and “Maintenance” (read-write) mode. In “Active” mode, all PVS Servers in the PVS Farm have read-only access to the managed store, allowing vDisks to be streamed to the target devices by multiple PVS Servers simultaneously. In “Maintenance” mode, the store transitions to read/write mode and only one PVS Server, the Maintenance Server, has access to the store. The Maintenance Server is used to perform updates on the vDisk image files (i.e., modify the contents and/or properties of the vDisks). Once the changes have been made, the data store is set back to “Active” mode so that all PVS Servers regain access to the store to stream the updated vDisks to target devices.

Note: *As PVS administrators toggle back and forth between “Active” and “Maintenance” mode whenever a change to a vDisk is necessary, it is imperative they ensure “Managed Read-Only” stores are never left in read/write mode when multiple PVS Servers start accessing the vDisks as corruption of the vDisks could occur.*

Pros: Since all PVS Servers in a PVS Farm are able to access a single LUN that contains the vDisks, “Read-Only” LUNs address the PVS HA requirement of allowing multiple PVS Servers to access the same vDisk image files.

vDisk maintenance is simplified – changes can be made to the contents and/or properties of a vDisk without having to copy the vDisk to multiple LUNs in order for all PVS Servers to stream the most recent version of the vDisk to target devices.

Good performance can be achieved with “Read-Only” LUNs as each vDisk is streamed to target devices from multiple PVS Servers in tandem via separate data paths (fibre channel or iSCSI) to a single LUN.

Additional PVS Servers can be introduced into the PVS Farm without the need to provision additional LUNs, providing a degree of scalability to the solution.

To protect vDisks stored in “Managed Read-Only” stores, snapshots of the LUN containing the vDisks can be taken (if the storage array supports snapshots) or vDisks can be copied over the LAN to another storage location, ensuring the vDisks are available for backup purposes.

Cons: Although “Read-Only” LUNs provide high availability of vDisks, they offer no high availability for PVS databases as there is no inherent mechanism that allows a PVS database to be hosted by another server if the initial server hosting the database were to fail unexpectedly. To obtain this type of functionality, another technology, such as Microsoft MSCS or Windows Failover Clustering for SQL Server, would need to be implemented, introducing additional configuration and management to the overall solution.

If the storage array supports LUN extension, LUNs may be expandable. However, a LUN cannot be expanded without first stopping all I/O activity against the volume. This requires that the Stream service on the PVS Server connected to the LUN be stopped whenever additional storage capacity is needed to ensure that all target devices accessing their vDisks on the LUN fail over to another PVS Server to avoid downtime.

Access modes for vDisks stored in “Managed Read-Only” stores are restricted to Standard Image mode or Difference Disk mode with device-side write caching (on disk or in RAM) or server-side write caching. With server-side write caching, the write cache files for the target devices *must* be stored on a separate LUN that supports both read and write access. vDisks operating in Private Image mode cannot be stored in “Managed Read-Only” stores.

While “Read-Only” LUNs can simplify vDisk maintenance by allowing highly available vDisks to be stored in a single LUN, the use of a single read-only LUN that must be continuously switched from read-only mode to read-write mode whenever vDisk changes need to be made presents additional complexities that can make the overall management of a XenDesktop or XenApp solution awkward and somewhat cumbersome.

Shared Storage – Windows Shares or NAS

Description: File Shares hosted by Windows Servers or NAS devices are used to store PVS vDisks.

How it works: A File Share is created on a Windows Server or NAS device and configured as a PVS data store within the PVS Administrator console. vDisk image files are created and stored on the File Share, allowing all PVS Servers in a PVS Farm to access the vDisks concurrently using a file-sharing protocol such as CIFS or NFS, thus meeting the requirement for enabling PVS HA.

Pros: vDisk maintenance is simplified – changes can be made to the contents and/or properties of a vDisk without having to copy the vDisk to multiple locations in order for all PVS Servers to stream the most recent version of the vDisk to target devices.

All vDisk Access modes are supported and consolidation of target device write cache files can be achieved as all device write cache files can be stored on a single File Share.

Cons: Although vDisk management is simplified by storing all vDisk image files in a single location, when it comes to other major aspects of the solution, such as performance, scalability and availability, File Shares hosted by Windows Servers or NAS devices introduce limitations that cause them to fall short of effectively addressing these aspects.

For example, as a result of the locking contentions issues that are often encountered with file sharing protocols in mid- to large-size environments, the performance capabilities of File Shares can be constrained, resulting in sub-par application performance that can have an adverse affect on the user experience. This can often be seen once I/O demands imposed by multiple applications running concurrently on multiple virtual desktops begin to overwhelm the locking mechanisms employed by file sharing protocols.

Adding more PVS Servers to support the build-out of XenDesktop and/or XenApp infrastructures further increases the potential for locking contention issues to occur, making it difficult for organizations to extent the benefits of either solution to a larger percentage of their users.

Expanding storage volumes to store additional vDisks is also not a task that can be completed quickly or easily as most file sharing technologies that support volume expansion require all I/O operations be stopped prior to commencing with the expansion of a volume, affecting user productivity.

Providing high availability of vDisks to ensure maximum uptime for virtual machines hosting virtual desktops or XenApp servers can be difficult to achieve. If a server hosting a file share was to encounter an unexpected outage (i.e., loss of network connectivity, hard drive failure, loss of power, etc.) PVS HA would be rendered useless as none of the PVS Servers would be able to access the vDisks located on the File Share. While NAS devices offer some protection in the form dual power supplies, fans and network interfaces, neither File Servers nor NAS devices provide any means for ensuring high availability of a PVS database in the event that a PVS Server (or another SQL server hosting the PVS database) encounters an unexpected failure that results in a system crash. To obtain this type of functionality, another technology, such as Microsoft® MSCS™ or Windows Failover Clustering™ for SQL Server™, would need to be implemented, introducing additional configuration and management to the overall solution.

To provide data protection for vDisks, the vDisk image files must be copied to another folder on the server hosting the File Share or to another folder on a different server. If a NAS device is used, the vDisks could be copied to another Share on the same NAS device or to another Share on a different NAS device. Copying vDisks between File Shares located on different Windows Servers or NAS devices introduces additional network traffic that could impact application performance and affect the user experience.

Shared Storage – SAN storage with Sanbolic Melio 2010

Description: Using a combination of SAN storage and SAN software provided by **Sanbolic**, this option greatly simplifies vDisk maintenance as all vDisk image files are stored on a single LUN. The big difference between this option and "Read-Only" LUNs… LUNs formatted with Sanbolic's cluster file system "**Melio FS**™" operate in read/write mode, allowing multiple PVS Servers to share concurrent read-and-write access to vDisks stored on Melio volumes at all times.

How it works: Sanbolic ***Melio 2010*** is a product suite comprised of an all-purpose, 64-bit cluster file system that employs an advanced, multi-layer locking mechanism to ensure data integrity is maintained as multiple servers read and write to the Melio volume at the same time; and a cluster volume manager that simplifies the management of Melio shared volumes while offering advanced features such as the creation of volume sets or stripe sets to improve I/O performance, as well as dynamic volume expansion.

Pros: Instead of requiring PVS administrators use a separate management utility to modify the properties of a data store to allow a designated PVS Server to modify the contents and/or properties of a vDisk, with Melio 2010, *any* PVS Server can be used to update a vDisk at any time as all PVS Servers have continuous read AND write access to the LUN containing the data store for the vDisk image files, addressing the requirement for enabling PVS HA.

All vDisk Access modes are supported and consolidation of target device write cache files can be achieved as all write cache files can be stored on a single LUN formatted with Melio FS, introducing an additional level of flexibility and versatility to XenDesktop and XenApp solutions.

Using Melio 2010 with SAN storage, high levels of performance and scalability can also be achieved. For example, using the cluster volume manager included in Melio 2010, PVS administrators can create stripe sets comprised of multiple LUNs located on different storage arrays. This configuration enhances I/O performance by leveraging multiple storage processors, additional caching and more spindles. To allow XenDesktop and XenApp infrastructures to scale quickly and seamlessly, additional PVS Servers can be implemented into a PVS Farm dynamically and Melio volumes containing the vDisks and write cache files can be expanded on the fly, with little management effort and no interruption to user productivity.

In addition to providing high availability of vDisks, Melio 2010 also includes support for SQL Server consolidation and availability via active/active application clustering, adding another layer of fault-tolerance to XenDesktop and XenApp solutions. For example, if a server hosting a PVS database fails unexpectedly, another server could begin hosting the database automatically, minimizing system downtime to ensure maximum user productivity.

To protect vDisks, Melio 2010 includes a VSS software provider that can be invoked by Sanbolic's Snapper™ utility, SILM™ or a third-party VSS agent, allowing PVS administrators to protect all vDisks through VSS-based snapshots of entire Melio volumes or individual vDisks at various points in time. The vDisks can then be mounted on any PVS Server for backup and recovery purposes. vDisks can also be automatically copied or moved to another backup volume based on user-defined policies; ensuring vDisks are available at all times.

Cons: It takes about 20 minutes to set up Melio shared storage for all PVS Servers in a PVS Farm.

Conclusion

Depending on the size, scope, nature and requirements of a Citrix XenDesktop virtual desktop or XenApp application delivery infrastructure, there is the potential for any one of the options described in this paper to prove effective for enabling PVS HA to ensure maximum uptime for virtual machines hosting virtual desktops and/or XenApp Servers.

For larger XenDesktop and/or XenApp infrastructures deployed in enterprise data centers however, only the combination of highly scalable, highly available SAN Storage and **Sanbolic *Melio 2010*** offers the levels of performance, scalability, flexibility and manageability necessary to enhance XenDesktop virtual desktop and XenApp application delivery infrastructures powered by Provisioning Services, while providing high availability of PVS vDisks, write cache files and databases, as well as a seamless growth path to support future expansions as demands continue to grow and organizations look to achieve the greatest return on their investments in these leading technologies.

Sanbolic Inc.
304 Pleasant Street, 2nd Floor
Watertown, MA 02472
Phone: 617-833-4249
Fax: 617-926-2808
Email: sales@sanbolic.com

CiTRIX ready™

Microsoft® GOLD CERTIFIED Partner