Realtime
publishers

The Shortcut Guide™ To

Storage Considerations for Microsoft SharePoint

sponsored by

hp®

Wendy Henry

Copyright Statement

If you have any questions about these terms, or if you would like information about licensing materials from Realtime Publishers, please contact us via e-mail at info@realtimepublishers.com.

Chapter 4: High Availability, Disaster Recovery, and Maintenance

Regardless of how fast or efficient your storage solution for SharePoint is, the entire SharePoint infrastructure is vulnerable if you do not engage in regular maintenance and employ some type of data protection strategy. SharePoint is an organic collaboration system that will likely host volatile mission-critical data. Sustaining data availability and delivery speeds the SharePoint users have come to expect should be a top priority for any SharePoint administrator. SharePoint usage will evolve over time, so any high-availability (HA) or disaster recovery (DR) plan should be routinely reviewed, tested, and potentially modified to accommodate the system.

The terms *high availability* and *disaster recovery* have become almost synonymous in today's IT industry. Understandably, the need to provide persistent and quick access to critical data almost certainly entails the need to reproduce that same critical data should it go missing. But in truth these terms express two different phases of an overall goal. HA SharePoint solutions are strategies that provide redundancy and workload balancing to deliver data quickly to users without detrimental interruption. Depending on the HA solution employed, a noticeable interruption of data delivery can be virtually eliminated or at least mitigated by setting reasonable expectations. HA solution performance is often measured by the *uptime* of the system (the percentage of a time slice that the system is online and delivering to the users). For example, the ultimate goal of an IT system may be 99.999% uptime or less than 6 minutes downtime per year.

DR strategies, however, define methods of restoring critical data that has been rendered extinct to the SharePoint infrastructure. Think natural disaster, irreparable disk failure, or irreversible data corruption. And though most DR strategies are built around a redeemable backup operation for valid reasons, leaning on an alternative HA redundant data source can also provide some measure of recoverability. In fact, many of the new iSCSI SAN data bit replication offerings in the market are making a name as DR solutions as well as HA stratagems. Choosing an appropriate DR scheme will depend on your Service Level Agreement (SLA) and user demands.

Facilitating HA and DR strategies will require monitoring as discussed in previous chapters and routine maintenance tasks. Data may occasionally need to be relocated via volume migration to support online, near-line, and offline availability. Load balancing may involve automating the rollover of workloads in response to dynamic user activity on the SharePoint system. This chapter will discuss hardware and software availability challenges, common DR strategies, and routine maintenance operations that enhance a SharePoint environment stored on a SAN infrastructure.

Hardware Reliability Solutions

The first component of a SharePoint environment that should be examined for reliability is the hardware. After all, what good is software if you don't have functioning hardware to run it on? When SharePoint is built upon a SAN infrastructure, it is important to consider not only the disks but also the network paths to those disks and the server controlling them. From the HA and DR standpoints, SAN hardware is uniquely qualified to store duplicate data sets. Of course, exploiting such capabilities makes protecting the SAN hardware and network paths even more imperative.

RAID

The configuration of two or more relatively small and cheap disks employed simultaneously to enhance performance and reliability was originally coined as a Redundant Array of Inexpensive Disks (RAID) in 1987 by some early adopters of the practice at UC Berkley in California. However, the idea of stretching data across multiple disks to enhance performance while implementing parity calculation to provide fault tolerance actually dates back to a 1978 US Patent by an IBM employee. In fact, at that very early date in computer technology history, disk duplexing, which is now known as *disk mirroring*, was already thought to be a mature art! But it wasn't until 1992 that controller, disk, and adapter board manufacturers realized an interoperability need for standardization and formed the RAID Advisory Board. This group then changed the acronym to translate as a *Redundant Array of Independent Disks (RAID)*, the practice's currently accepted title.

In essence, a RAID configuration employs portions or entirety of two of more disks to represent a single logical area of potential data storage space (see Figure 4.1). The benefits can be twofold: by utilizing multiple mechanical disks simultaneously, more data is read or written within a finite time slice (performance) and by isolating only a portion of the data to potential disk failure induced loss, that portion can be regained mathematically via parity by processing the remaining portions (fault tolerance). There are several levels of accepted RAID implementation, as denoted in Table 4.1, but suffice it to say that using RAID for mechanical drives is a good idea.

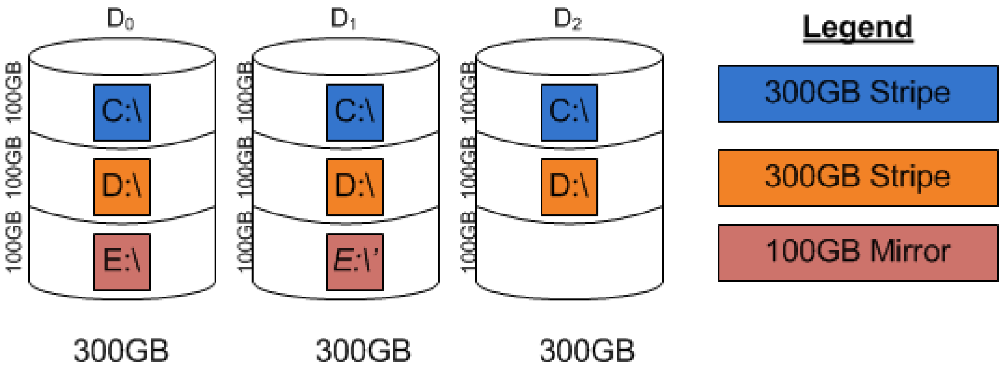
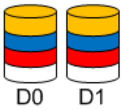
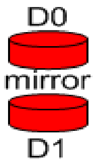
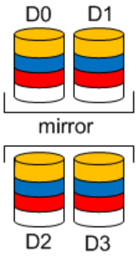
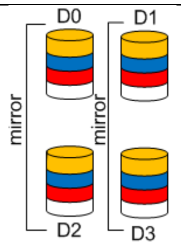
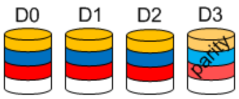


Figure 4.1: Simple example of RAID use across three independent disks.

Level	Schematic	Description
RAID 0 Stripe Set		Two or more equal areas of space on separate disks; data is fragmented per number of disks and fragments are written to the same sector on respective disks simultaneously without error checking ; performance enhancement only, NO FAULT TOLERANCE
RAID 1 Mirror Set		Two equal areas of space on separate disks; complete data stream is written to same sector on respective disks simultaneously generating an entirely redundant data set; connecting each disk in a mirror configuration to a separate disk controller is informally known as *disk duplexing*
RAID 0+1 Mirrored Stripe Set		Two or more equal areas of space on separate disks striped with RAID 0, then the stripe set is RAID 1 mirrored to equal areas on as many separate disks as in the stripe set
RAID 10 Striped Mirror Set		Two equal areas of space on separate disks mirrored with RAID 1, then the mirror is RAID 0 striped to equal areas on two or more separate disks ; essentially a stripe set made up of independently mirrored disks
RAID 2	n/a	Theoretical stripe set of 3+ disks rarely employed in practice; Hamming code error correction calculated bits are determined for each small data stripe then stored on no particular disk
RAID 3 Stripe Set with dedicated parity		Three or more equal areas of space on separate disks; data is fragmented per number of disks and fragments are written to same sector on respective disks simultaneously; parity bits are calculated at the bit or byte level then stored on a dedicated disk

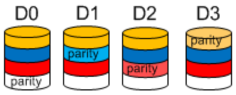
Level	Schematic	Description
RAID 4	Same as RAID 3	Same as RAID 3 except parity bits are calculated at the block level then stored on a dedicated disk
RAID 5 Stripe Set with distributed parity	 D0 D1 D2 D3 parity parity parity parity	Three or more equal areas of space on separate disks; data is fragmented per number of disks and fragments are written to same sector on respective disks simultaneously; parity bits are calculated at the block level then bit or byte interleave stored on no particular disk

Table 4.1: Standard RAID levels.

In addition to the RAID levels outlined in Table 4.1, some hardware manufacturers are implementing nested RAID beyond RAID 10, such as RAID 5+0 or RAID 5+1 (sometimes referred to as RAID 53) to further enhance the performance benefits of data striping with the fault tolerance of parity and mirrors. Be sure to gather transfer rate and disk access speed information from your vendor before settling on a highly nested RAID level. Sometimes too much of a good thing can be bad for you.

RAID 6

A grave limitation to RAID 3, 4, and 5 is that these configurations can recover successfully only from a single disk failing in the array. If a second disk or more fails in the array either before or during the rebuild of the first failed disk, fault tolerance is null and void and the entire data set on the array is lost. Just as concerning, if the single and only disk in a RAID 3 or 4 array that fails happens to be the dedicated parity disk, the entire array becomes unrecoverable and (depending on the firmware) may even become inaccessible!

An additional RAID 6 level is recognized by the Storage Networking Industry Association (SNIA) as providing recoverability from up to two disks failing in the array. Basically, RAID 6 is an enhanced version of RAID 5 that also uses block-level stripes and stores parity across distributed disks. However, RAID 6 calculates a second parity on the byte level and stores it separately across distributed disks. Like the first block parity, the second byte parity is also interleaved and stored on a different disk for each stripe.

RAID 6 protects data from loss during the vulnerable rebuild of the first disk failure. But you pay for this additional level of protection in diminished write performance and reduced data storage space:

- RAID 5 Useable space = s(n-1)/n
- RAID 6 Useable space = s(n-2)/n

Where s=sum of capacities from n disks and n=number of disks

Redundant Channels to External Storage

Protecting data from a disk failure is only half the battle. To be certain that the SharePoint server can access its SQL Server databases stored on a SAN solution, you must also ensure a persistent network path. The previous chapter outlined the implementation of two LAN switched paths between the SharePoint server and its SAN storage device. However, switching equipment and patch cables are not the only concern. Network Interface Cards (NICs) on the servers should also be redundant to provide failover in the event of a NIC device failure. Consider purchasing OEM teaming software that provides load balancing and fault tolerance of the NICs or choose a storage provider that supports Microsoft's MPIO drivers for Windows Server 200x OSs. Complete duplication of the entire network channel is imperative to providing uninterrupted access and must be installed correctly to provide automatic failover.

> **Warning**
>
> Duplicating network paths between a SharePoint server and SAN is often overlooked and may be restricted by budget constraints. However, failing to produce a redundant path to data can result in unacceptable downtime. In the event of failure, lengthy troubleshooting to determine the failed network component could violate SLA downtime limits for SharePoint.

First, you must isolate the network path between SharePoint and the SAN from the network path between SharePoint and its clients. Doing so might require extensive restructuring of your switching infrastructure. When it comes to hardware redundancy, simply implementing Virtual LANs on a Layer-3 switch is not sufficient. Though VLANs will help isolate the traffic, they do not protect your SharePoint environment from the failure of the switch equipment itself. To adequately separate client traffic from SAN communication, purchase two additional switches for the SAN communication paths.

Second, in addition to NIC failover, the network switch equipment must provide fault tolerance with minimal hands-on administration (see Figure 4.2). Many switch manufacturers offer fault-tolerance configurations that will automatically redirect TCP/IP traffic from failed ports to functioning ports and even between switches. Configure the most fault tolerance you can afford to automatically handle path interruptions between SharePoint and the SAN.

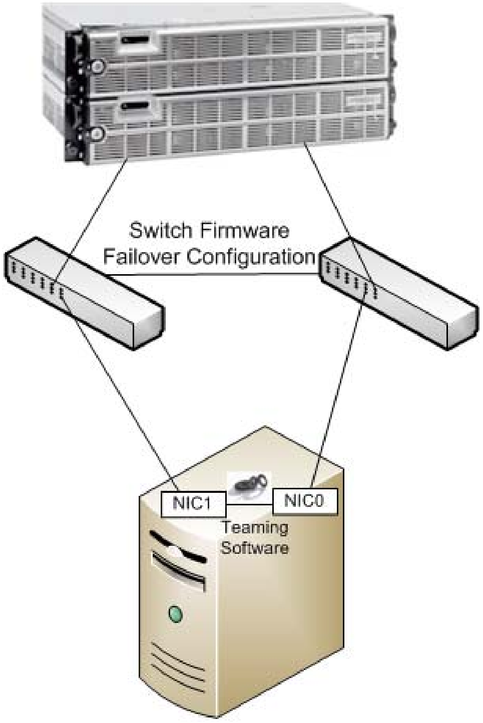


Figure 4.2: Optimal redundant network path configurations.

Establishing fault-tolerant disk arrays and redundant network paths protects SharePoint only from certain hardware failures. For example, RAID up to level 5 recovers only from a single disk failing. Should multiple disks in a RAID array fail, the SharePoint data set is lost. And even the most redundant of network paths will not mean a thing if the data is missing. Therefore, to assure fault tolerance of an entire SharePoint data set, you should invoke replication.

Replication: Establishing Duplicate Data Sets

In the past, SharePoint data replication was usually performed by an application or service running on the OS. And although Windows Server 2003 and 2008 still offer means of automatically copying data between volumes, there is a more efficient way to produce a redundant set of SharePoint data. Most SAN software offers data replication solutions to store data sent to the SAN disks in two separate locations. In fact, if you combine data replication with storage virtualization, the SAN platform can provide fault tolerance to the SharePoint server transparently (see Figure 4.3).

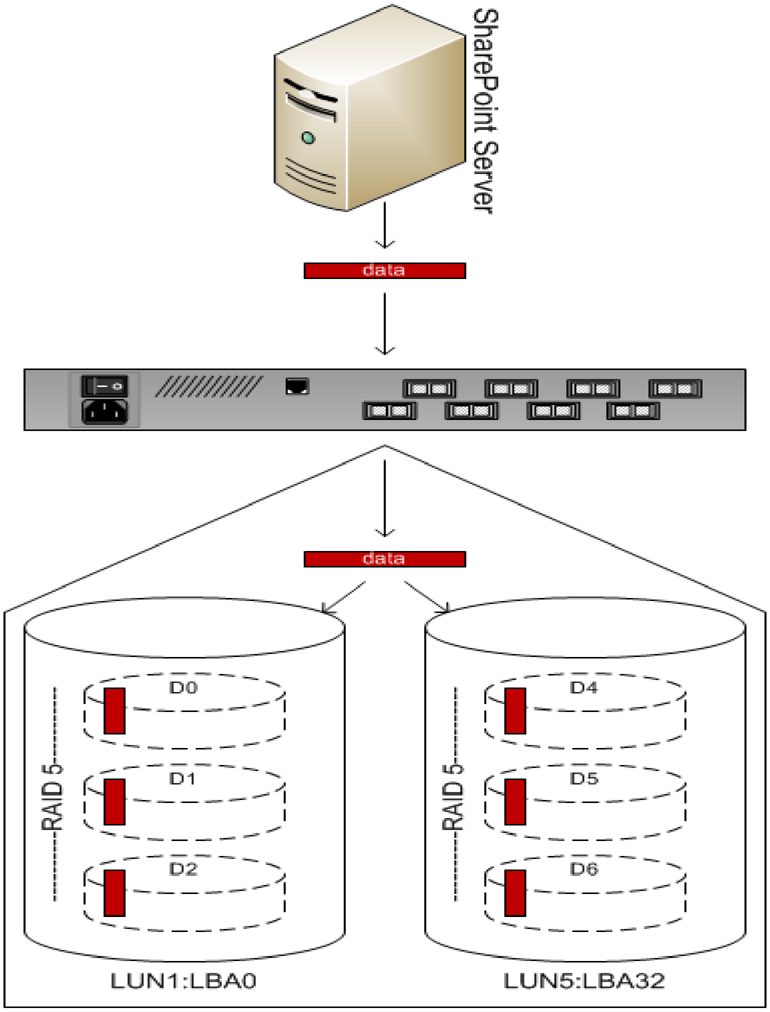


Figure 4.3: Data replication across two LUNs of a SAN: original data on LUN1, duplicate data on LUN5.

In essence, data saved to a SAN RAID stripe set can be optionally written again to one or more alternative RAID destinations to automatically establish duplicate sets of the data. Should the primary volume become unavailable, failover to duplicate data on a secondary volume is automatic within the SAN. Better yet, when the primary volume is again operational, the SAN can be configured to failback automatically as well. SharePoint benefits from such local synchronous data replication of its SQL Server databases as the services are always able to connect to their configuration, search, and content databases. In fact, large SharePoint Search environments can also benefit from replicating the index because rebuilding a large full-text index in the event of failure could take several hours (see Figure 4.4).

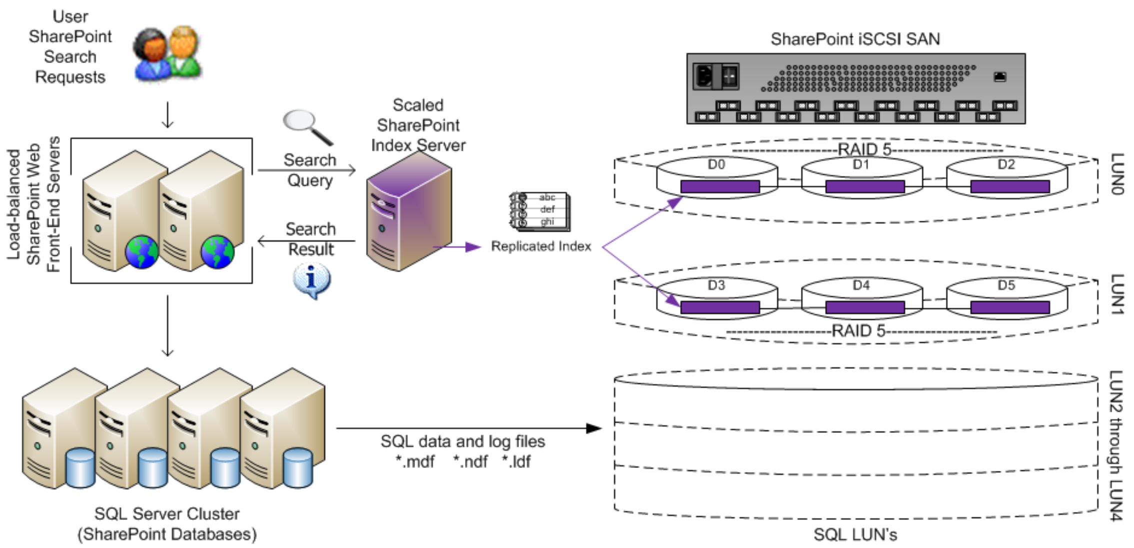


Figure 4.4: Storage of a scaled SharePoint index onto a replicated SAN facility.

The degree to which any SAN data is replicated (once, twice, and so on) depends on the reliability of the storage devices as well as the critical importance of the data. Often regulatory compliance will also dictate data replication requirements as well as data location. Local synchronous replication occurs within the same system and data is written to replication destinations synchronously so that data value latency is avoided. Replication can also occur remotely across systems or sites, but more on that later in this chapter.

Software Reliability Solutions

Establishing hardware, network, and data fault tolerance protects SharePoint from obvious tangible failures. Perhaps more subtle are software issues that present intermittently and can be more difficult to accurately diagnose. Protecting the Windows Server OS and SharePoint services is paramount. Microsoft SQL Server and the SAN controller OS must be highly available as well. Corrupting any one of these systems can render the SharePoint enterprise unstable, or worse, unavailable. Maintain synergy among all the SharePoint systems to keep SharePoint online.

Delivering SharePoint on a Virtual Machine

Virtual servers have gained significant popularity in the IT industry as a means to capitalize on sophisticated hardware investments. When it comes to setting up a scaled SharePoint enterprise on virtual servers, the WSS Web front-end servers and database servers as well as the MOSS application servers (Excel Services Server, Forms Server) are all viable candidates. By creating each of these systems as a virtual server and storing the virtual hard drive files on a SAN storage solution, all the SAN benefits such as storage virtualization, thin provisioning, snapshots, and data replication are available to support the SharePoint software.

To install SharePoint onto a virtual server, first the virtual server must be created on a host server. Select a host server that uses SAN disk destinations in its file system volume configuration. Once launched, a virtual server consumes significant memory and performs frequent I/O instructions, so invest wisely in the chosen host server's hardware. Load the virtualization application of your choice onto the host server and create a new virtual machine to support SharePoint. The choices of host OS platform and virtualization application are of little consequence except to note that the virtualization application must support Windows Server 2003 and Windows Server 2008 guest OSs and should be 64-bit compatible to get the most performance out of the virtual machine.

During creation of the virtual machine that will eventually house SharePoint or SQL Server, the virtualization application will undoubtedly prompt you to select file system destinations for the virtual machine files. Be sure to select a logical volume on the host OS that paths to your SAN (Figure 4.5). You must also configure the virtual machine to utilize the NIC of the host server to facilitate network connectivity to the LAN clients that will be requesting SharePoint resources.

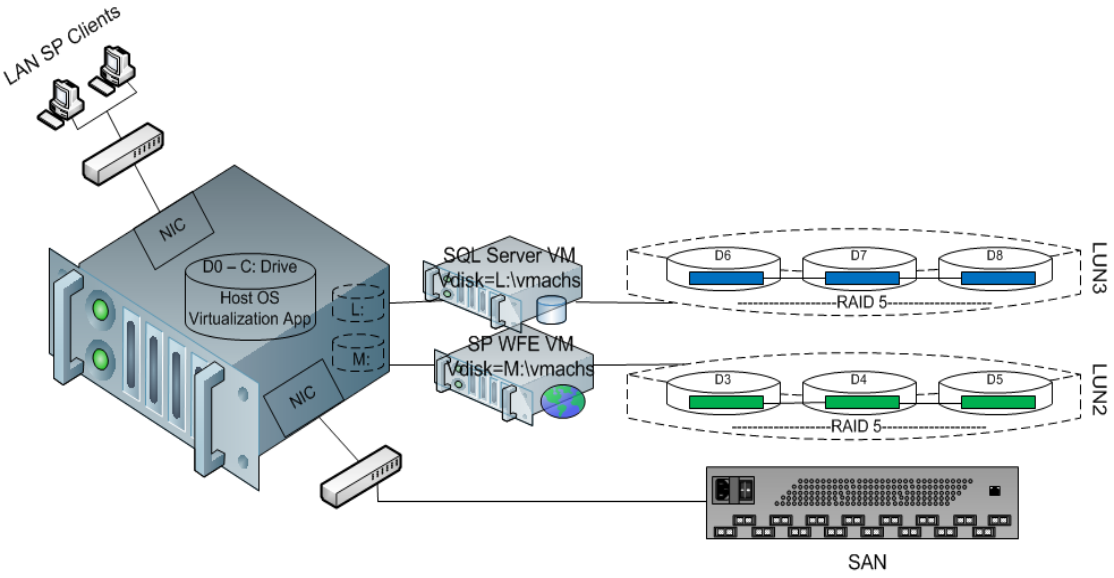


Figure 4.5: Example of two virtual machines hosted by one host server and stored on two separate SAN LUNs.

Configure the virtual machine file destinations carefully so as not to place codependent virtual machines on the same disks. In the illustrated example in Figure 4.5, a SharePoint WFE server frequently queries the SQL Server housing the SharePoint databases. Placing both the WFE and SQL virtual server files on the same LUN could inadvertently cause undue disk array contention. Distribute disk access by placing such dependent virtual servers on separate disks. You should also reduce disk contention within a SQL Server virtual machine by placing the virtual machine files on a separate disk from the SQL database files.

Note

Most virtualization applications offer the ability to extend a virtual machine onto multiple virtual disk files to be stored across multiple file system destinations. But such application-level logical striping is unnecessary if you use a well-designed SAN destination that already exploits multiple disk arrays.

By utilizing a well-designed virtual server solution stored on a SAN system, you can lean on the SAN data replication methods to provide highly available, synchronized versions of the virtual machine disk files. Make actual virtual machine disk file locations transparent to the virtualization application by employing SAN storage virtualization to avoid time-consuming reconfiguration of the virtual machine's settings in the event of disk file movement or failover. Employ SAN thin provisioning for flexible initial virtual machine disk file sizing. When you store your SharePoint virtual servers on SAN facilities, the SAN system's data protection offerings make your entire SharePoint software suite (OS, IIS, WSS, MOSS, and SQL) fault tolerant!

A Brief Note About Virtualization and Windows

Though detailing the steps necessary to create a virtual Windows server falls beyond the scope of this book, it is worth mentioning that not every virtual machine setup is the same. Different virtualization applications are available in the market, from Microsoft's own Virtual Server application to the industry leader in production virtual machine software—VMware. In fact, the Windows Server 2008 OS contains its own virtualization application module called Hyper-V that can be added as a server role to the OS.

Regardless of your virtualization application choice, the software will be interacted with via the host OS on an actual server. Within the host server's file system, the configuration settings of the virtual server machine are stored in a separate file from the virtual disk file containing the virtual server's stored data files. These files have different filename extensions depending on the manufacturer:

- Microsoft = .vmc(config)/.vhd(disk)
- VMware = .vmx(config)/.vmdk(disk)/supporting files

Microsoft now offers System Center 2008 Virtual Machine Manager in its enterprise-class management suite to migrate, manage, and monitor virtual machines hosted by a Windows Server OS. If you purchase MSSCVMM, you will need to configure it properly to support SAN storage of the virtual machine files (see the following link http://technet.microsoft.com/en-us/library/cc764269.aspx).

However, you would be better off using your SAN utilities to manage your virtual server files. You should also employ your SAN software's transparent redirection capabilities, such as storage virtualization and data replication failover, to avoid the need to reconfigure the virtualization application every time a virtual server file is relocated. Using the SAN abstracts virtual machine management from the Windows platform, allows more interoperability with non-Microsoft platforms (host OS and/or virtual application), and saves on Microsoft System Center licensing costs.

Providing Continuous Data Protection

Protecting SharePoint data, both configuration and content, is vital to uninterrupted delivery of mission-critical business user information. This chapter has thus far discussed protecting both hardware and the SharePoint suite of OS and application software. By storing SharePoint on a SAN storage facility, SharePoint data protection is provided by the SAN system itself. However, to ensure continuous data protection, the SAN system must be fault tolerant as well.

Different SAN manufacturers offer different methods of protecting their systems. Many mature SAN implementations use redundant controllers, power supplies, and other components to eliminate single points of failure within the SAN hardware. Alternatively, consider purchasing a SAN storage facility that implements logical storage volumes on clustered SAN nodes so that if any one SAN node fails, the volume continues uninterrupted data delivery from an alternative node in the SAN cluster. There are also SAN-aware smart network devices such as Cisco SAN switches running SAN-OS that can be implemented to enhance QoS and routing between multiple storage nodes of a SAN solution.

Note

Quality of Service (QoS), specifically relating to TCP/IP networking devices, is often assured via software policies that regulate packet drop thresholds or bit rates to relieve congestion on a busy network.

All storage systems should employ an uninterrupted power supply (UPS) unit to condition voltage and protect from power fluctuations or outages. Also, plug redundant power supplies into redundant power distribution units or circuits if possible. If the SAN operates in a single site, you might even consider investing in an alternative power supply solution (that is, generators) to maintain business data during long-term power outages.

Database Mirroring in Microsoft SQL Server

The most critical participant of the SharePoint enterprise is the Microsoft SQL Server. In a large-scale SharePoint farm, SQL Server is often installed onto a failover cluster using a SAN as the storage. New to SQL Server 2005 and 2008, within the SQL Server product ships the ability to duplicate and synchronize a chosen database between two separate SQL Server installations (called *instances*), whether they be running on the same or different servers. The technology, dubbed *database mirroring*, employs your choice of synchronous or asynchronous data writing to two separate databases, producing a complete replica of a given database. From a testing perspective, the ability to mirror a database between two SQL Server instances installed on a single server allows for analyzing updates and procedures against a mirror copy of the production database without the expense of a second lab server piece of hardware. However, from a production standpoint, it is best to mirror a critical database between two SQL Server instances installed on two separate servers to protect the database from OS corruption or server failure.

In Microsoft vocabulary, the server on which the original database resides is referred to as the *Principal Server* and the server housing the replica database is referred to as the *Mirror Server*. The premise is simply this: the SQL Server instance on the Principal communicates with the instance on the Mirror through TCP ports selected solely for database mirroring packets. Once established, all transactions written to the original production database are also written to the replica database. To set this up, a knowledgeable SQL Server DBA must:

- Create TCP Endpoint objects in both the Principal Server’s SQL Server instance and in the targeted SQL Server instance on the Mirror Server
- Set the original production database’s recovery model option to Full
- Perform a Full backup of the original production database and restore the backup to the targeted instance on the Mirror Server to create the initial copy of the replica database

Note

Database mirroring became available in SQL Server 2005 SP1. Only the Standard and Enterprise editions of SQL Server are capable of performing the Principal Server and Mirror Server roles (since Developer and Enterprise 180-day Evaluation editions support all functionalities of Enterprise edition they too will perform these roles).

Caution

Remember that allowing SharePoint’s setup program to automatically install SQL Server as well results in SQL Server Embedded edition being loaded (you may notice SQL Server Management Studio displays “Express” as the edition property of the instance for Embedded edition). Embedded edition does not support database mirroring.

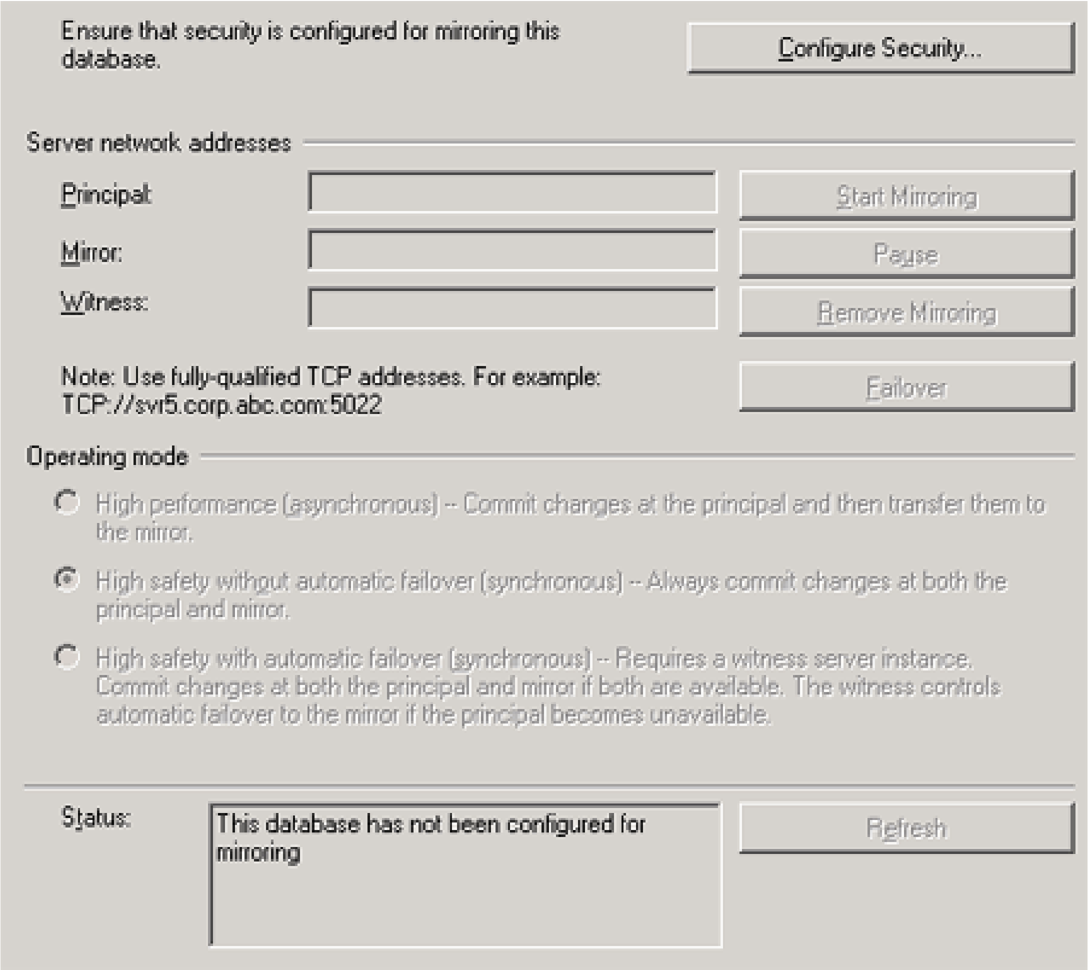
Unlike other SQL Server HA offerings, in database mirroring, the replica database residing on the Mirror Server remains unavailable to the SharePoint services until it is needed for failover reasons. There is no workload balancing. Failover to the replica can take place automatically if the database mirroring design employs a third SQL Server system in the role of Witness Server. The Witness Server will watch the database synchronization traffic between the Principal Server and the Mirror Server and will promote the Mirror Server if the transaction flow from the Principal Server ceases. If a Witness Server is not employed, failover to the Mirror Server in the event of a failure on the Principal Server must be implemented manually.

Why Implement SQL Server Centric Database Mirroring?

Though this book is not about Microsoft SQL Server per se, it is important to understand the HA offerings of this required data repository for SharePoint. Each SQL Server instance is autonomous as far as instance-level configuration settings and security. Each instance employs its own core MSSQLSERVER service to run the database engine and its own SQLSERVERAGENT service to handle automation. Each SQL Server instance will contain its own unique set of objects, including resident databases.

Despite the database protections outlined earlier in this guide, a SharePoint content database can still potentially become unavailable due to corruption within its instance. Recovering the entire OS of a SQL Server via server virtualization, data replication, and snapshots may be overkill when a simple SQL instance or database adjustment will do. In fact, over-correction caused by restoring an entire OS could potentially revert other services that should not have been set back! By establishing database mirroring to separate SQL Server instances, the SharePoint content database will continue to perform while the DBA distinguishes whether issues are at the database level, the SQL Server instance level, the SharePoint services level, or the OS level. Database mirroring buys troubleshooting time without invoking an entire OS failover.

Microsoft best practices for SharePoint encourage database mirroring for content databases only. Recall that configuration and search database failures may interrupt administration or navigation, but only content database failure will negatively impact the business users' ability to access critical information. Setting up mirroring on the content database can be accomplished by the DBA in SQL Server Management Studio GUI on the database mirroring property page of the original content database (see Figure 4.6). By clicking Configure Security, SSMS will launch the Configure Database Mirroring Security Wizard to walk through identifying mirroring instances and setting TCP port numbers, Endpoint names, and encryption status. After which, the mirroring Operating Mode can be changed.



Ensure that security is configured for mirroring this database.

Server network addresses

Principal:

Mirror:

Witness:

Note: Use fully-qualified TCP addresses. For example: TCP://svr5.corp.abc.com:5022

Operating mode

○ High performance (asynchronous) -- Commit changes at the principal and then transfer them to the mirror.

● High safety without automatic failover (synchronous) -- Always commit changes at both the principal and mirror.

○ High safety with automatic failover (synchronous) -- Requires a witness server instance. Commit changes at both the principal and mirror if both are available. The witness controls automatic failover to the mirror if the principal becomes unavailable.

Status: This database has not been configured for mirroring

Figure 4.6: Database Mirroring page of WSS_Content database properties.

Disaster Recovery

Designing HA solutions to meet user access needs and SLA promises is an admirable endeavor. But a SharePoint environment, like any critical system, also requires a workable DR plan to be complete. DR plans should address backup routines, standby strategies, and site outages. And although some HA solutions can be used to continue SharePoint data delivery in the event of a disaster, Microsoft best practices stipulate that both HA and DR designs are necessary for full protection of a SharePoint enterprise.

Backup and Restore Strategies

In recent years, the tried-and-true practice of creating backups as compressed files for the purpose of data recovery or offline storage has come under fire. There are two schools of thought on whether traditional backup procedures are necessary any longer in the wake of immediate data recovery options such as snapshots and bit/byte/block-level replication offered by many sophisticated SAN manufacturers. On one side, traditionalists argue the benefits of creating backup files for mobility, to meet regulatory compliance, and to recover from data corruption due to user error. In opposition, mavericks point out that today's SAN replication architectures allow for remote storage of replicated data (thus meeting regulatory compliance and mobility needs) while requiring less downtime to recover lost data. Microsoft best practices recognize both arguments and stress that both strategies have a place in SharePoint administration.

One of the advantages of newer data replication and snapshot technologies is the speed at which data versions can be captured. In fact, modern SLAs require ever-tighter data salvage tolerances with minimal value lineage loss. For example, creating a snapshot of data values every 5 minutes is not unheard of. In the event of corruption or failure, reverting to the last snapshot would lose only 5-minutes worth of work. In fact, data replication can be designed to be so synchronous that no work is lost at all! However, the speed of these technologies can also be their weakness when it comes to recovering from data corruption. If the corruption goes unnoticed for a time, the replicas and snapshots will indeed contain the corrupted values. And if these data replication and snapshot schemes are frequent, they are probably also set to overwrite or at least limit the replicas stored to avoid spanning excessive disk space. Thus, by the time the data value corruption is discovered, the previous pristine data value may no longer be available from snapshot or replica.

All of a sudden the traditional backup media gathering dust on the corner shelf doesn't look so obsolete anymore, does it? By combining HA tools with a backup and restore solution, you can provide SharePoint with the best of both schools. In the event of true data loss, redirecting requests to a replica or reverting to a snapshot provides immediate or near-immediate recovery. However, in the event of slowly emerging data corruption, a restore procedure of a pre-corruption backup file can restore your data to stable values. Table 4.2 outlines various free Microsoft utilities bound with the product that can be used to generate SharePoint backups.

Utility	Interface	Scope
STSADM.EXE	CLI	Entire SharePoint Farm ; Specific SharePoint Web Applications ; any or all SharePoint content databases ; Specific SharePoint Site Collections
SharePoint Designer 2007	GUI application	Individual SharePoint Sites
SP Central Administration	GUI web-based	Entire SharePoint Farm ; Specific SharePoint Web Applications ; any or all SharePoint content databases
SQL Server Management Studio	GUI application	Specific SharePoint databases

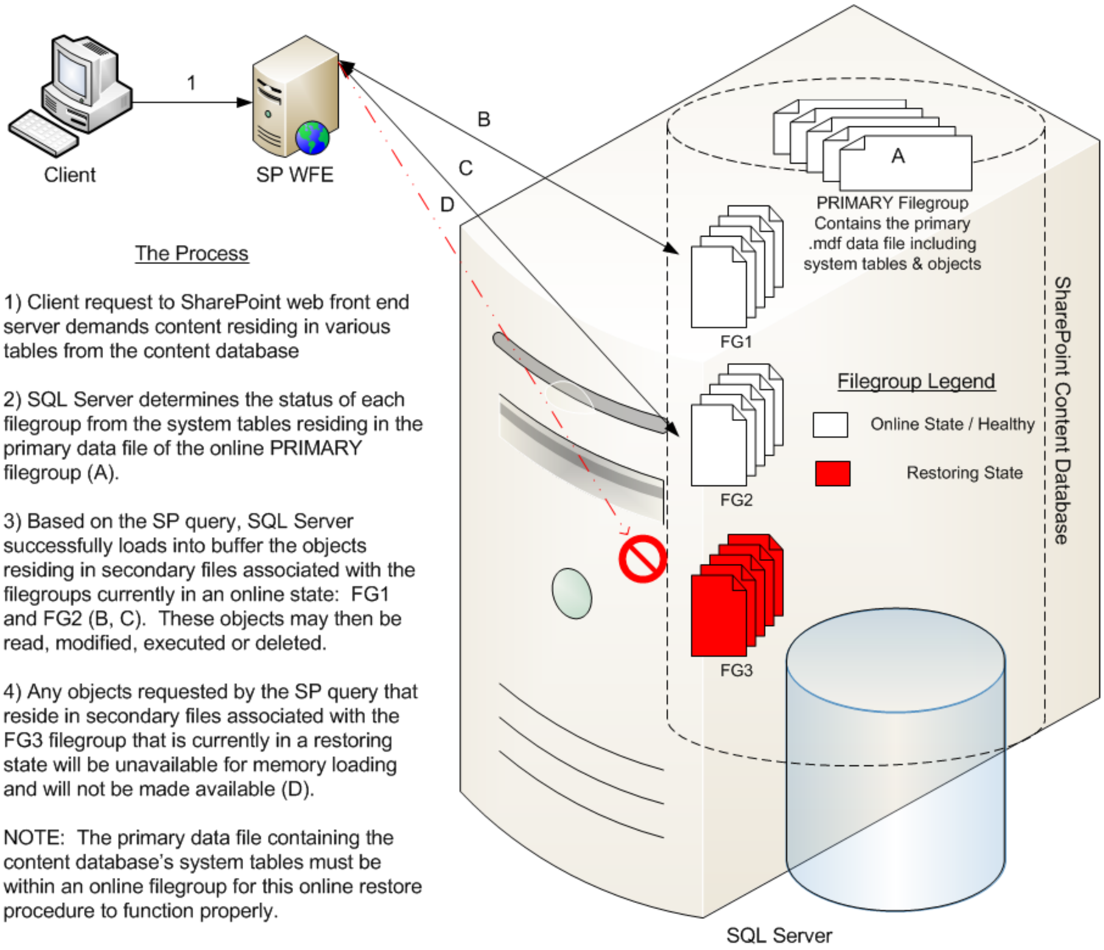
Table 4.2: SharePoint backup utilities.

When using STSADM.EXE from the command prompt, a simple parameter choice on the backup operation will pivot the scope of the operation from backing up only a site collection (–Url parameter) to offering the choice of farm, Web application, or content database (–backupmethod parameter using a –item designation).

Cross-Reference

For more information about using the backup operation of stsadm.exe, see the http://technet.microsoft.com/en-us/library/cc263441.aspx Web page at Microsoft TechNet. If you need granular backup of SharePoint content objects such as individual lists, libraries, pages, or Web parts, consider purchasing a third-party SharePoint backup application. Many manufacturers will show up on an Internet search, such as AvePoint, CommVault, Idera, and Quest Software.

SQL Server 2005 and 2008 now offer online restore procedures on very large databases (VLD) that have been structured into multiple data files across multiple filegroups. Essentially, the database is backed up at the file or filegroup level and in the event of a restore, only the currently restoring file or filegroup is held offline in a restoring state. The Primary filegroup and any other filegroups already restored or not requiring a restore can remain online to satisfy queries and transactions (see Figure 4.7). Consider restructuring large SharePoint content databases into multiple filegroups to take advantage of this functionality.



The Process

1) Client request to SharePoint web front end server demands content residing in various tables from the content database

2) SQL Server determines the status of each filegroup from the system tables residing in the primary data file of the online PRIMARY filegroup (A).

3) Based on the SP query, SQL Server successfully loads into buffer the objects residing in secondary files associated with the filegroups currently in an online state: FG1 and FG2 (B, C). These objects may then be read, modified, executed or deleted.

4) Any objects requested by the SP query that reside in secondary files associated with the FG3 filegroup that is currently in a restoring state will be unavailable for memory loading and will not be made available (D).

NOTE: The primary data file containing the content database's system tables must be within an online filegroup for this online restore procedure to function properly.

Figure 4.7: Online restore of a SQL Server 2005/2008 SharePoint content database.

Warm Standby Solutions

Maintaining redundant IT systems with the intent of bringing user requests to the replica system either immediately (as in load balancing) or eventually (as in a failover scenario) constitutes a valid standby solution. The class of a standby solution is determined by the degree to which the systems are duplicated and the latency of data synchronization between them. In terms of data redundancy:

- Cold standby employs highly latent asynchronous methods of data duplication to a replica system to be engaged for failover only
- Warm standby employs latent asynchronous data replication for load balancing and failover
- Hot standby virtually eliminates data latency with synchronous data manipulation on both the production and the secondary system and allows data to be serviced by either system for load balancing and failover

If you think of standby classes in terms of *site* redundancy, the picture becomes even clearer. A cold standby site is merely an empty building into which human and physical resources could be moved if the primary site becomes inhospitable. If the primary site suffers a disaster that debilitates IT hardware and data, business will be interrupted while resources are moved to the cold site. A warm standby site is a separate building housing duplicate physical resources to which data is replicated but at which no personnel qualified to work with or administer the systems reside. In the event of a disaster at the primary site, human resources could easily step into the warm standby site and begin using the replica systems until the primary site is recovered. Hot standby sites are those bearing duplicate physical and human resources to which all business requests can be quickly and easily redirected should the primary site become unreachable

Imagine a company that maintains two customer service call centers, one in Boston that usually answers all East and Central customer calls and one in San Diego that usually answers all Mountain and Pacific customer calls. The entire set of customer service data is maintained at each city. Should a natural disaster strike the San Diego call center, all incoming Mountain and Pacific customer phone calls would be immediately and transparently rerouted to Boston. In this scenario, Boston is a hot standby site for San Diego.

Of the three standby classes, obviously, the hot standby site or system is the most costly to implement. In the interest of balancing budget and downtime, the warm standby solutions are most frequently employed. Simple risk assessment is

$$\text{Risk} = \text{Vulnerability} \times \text{Probability}$$

Companies must invest reasonably in DR standby solutions that address not only the severity of a disastrous event but also the likelihood that it will happen.

Remote Replication—Multi-Site and Geo-Cluster

To create alternative standby sites, identical SharePoint data must reside in two separate locations. By employing a SAN solution that offers remote replication, you can lean on the storage facility to duplicate critical SharePoint databases to an alternative warm standby location. For more extensive DR, the Windows server accessing the SAN data should be able to failover to a fellow Windows server at the warm standby site to protect the system from server failure.

Recall that SQL Server is one of the most critical components of a SharePoint enterprise and is often installed onto a cluster HA design. In the past, one of the limitations of clustering the Windows Server OS via Microsoft Cluster Service (MSCS) was that the cluster nodes had to reside in the same physical location, or at least believe they did. Microsoft introduced the idea of geo-clustering in Windows Server 2000 MSCS to accommodate placing nodes of a single cluster in different physical sites then using VLAN switch technology to fool the cluster nodes' network interfaces into believing they were on the same IP subnet. By stretching a Windows Server cluster across the WAN, failover to alternate sites can occur within the cluster without the need for manual intervention.

In essence, creating a fully functioning warm standby site will use both the SAN remote replication technology and MSCS geographical disbursement technology to provide failover to an alternate site in the event of either storage facility interruptions or cluster node downtime. First, network paths across the WAN must be established between MSCS cluster node servers. One of the easiest ways to accomplish this is to install at least two NIC adapters in each node, then configure one with the internal IP address structure of the physical site (Microsoft calls this the *private interface*) and configure the other with an IP address of a separate VLAN reserved for cross-WAN cluster traffic only (Microsoft calls this the *public interface*) as seen in Figure 4.8.

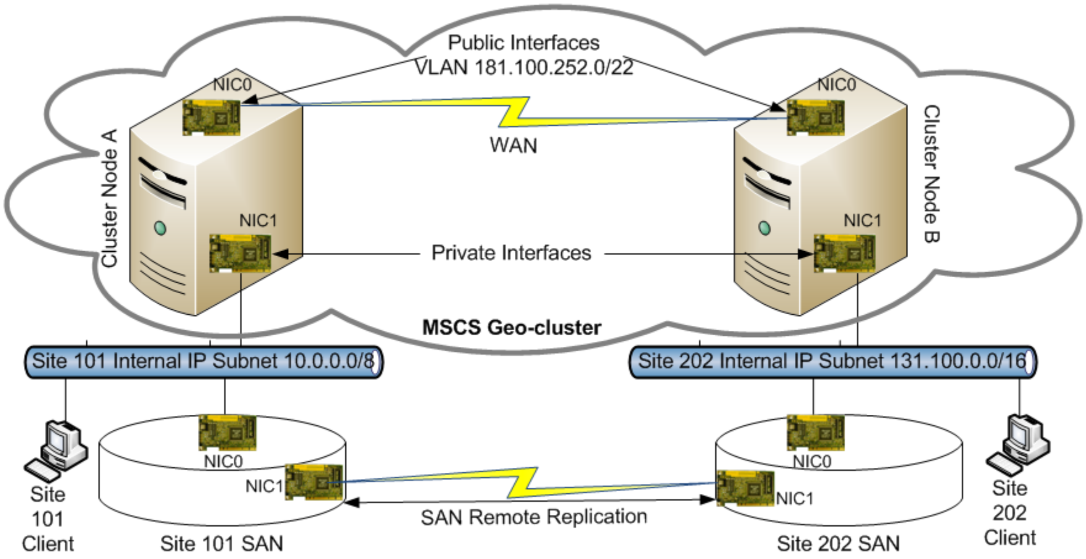


Figure 4.8: Typical Microsoft geo-cluster topology with SAN remote replication.

Similarly, the SAN nodes should be configured with at least two network interfaces as well; one on the internal IP address scheme that can communicate with the private interface of the resident Windows Server cluster node. And the other on a separate VLAN or public IP address scheme that can communicate with the data replication partner SAN in the alternate location.

A Few Gotchas About Microsoft Geo-Clusters

The idea of stretching a Windows Server MSCS cluster across physical locations is attractive. But before you begin configuring nodes, be aware of the following limitations and recommendations.

Microsoft recommends three NICs per cluster node server:

- 0=Internal network IP scheme for communication with clients
- 1=Separate internal network IP scheme for communication with other local cluster nodes and the local shared storage SAN
- 2=Public interface on WAN-capable IP scheme for communication with remote cluster nodes and the remote shared storage SAN replica

Also, the heartbeat of the cluster on a Windows Server 2003 MSCS design is hard coded at 500 milliseconds. Therefore, the WAN topology between geographically disbursed cluster nodes must be capable of handling heartbeat traffic within this time limit or unnecessary failovers may occur. Limit the competing network traffic on the WAN used for geo-clusters.

New to Windows Server 2008 CS, the heartbeat interval is now configurable and can be set to a reasonable time limit according to WAN workload. Also new to Windows Server 2008 CS geo-clusters, the public interfaces of the nodes can be on different subnets. It is no longer necessary to design identical VLANs at each site to accommodate WAN communication between disbursed nodes.

Maintenance

Throughout the life span of a SharePoint enterprise, certain data and system maintenance tasks must be addressed to keep the system running smoothly. Data may need to be relocated or archived, user base growth may necessitate new load-balancing schemes, and the software of the various participant systems will undoubtedly require patching, updating, or perhaps even upgrades. SharePoint is one of those platforms that can always be improved upon. Be sure to document all maintenance and test all procedures, as with any IT system, to avoid unnecessary failover traffic or downtime. Remember that SharePoint is unique in that it is an organic collaborative system and should be examined routinely to identify potential improvements.

Maintaining Volume Distribution: Online vs. Offline

Not all information within SharePoint is needed at all times. Over the course of use, certain content will by nature be sought less frequently. This is not to say that the content will be ready for archiving onto cheaper storage, but rather the data will be accessed infrequently enough to beg the question of whether it remain online or taken offline. To determine a SharePoint site's usage, you can invoke Usage Analysis from the Central Administration Operations page and watch the results. Over time, usage of certain content will wane, making its resident volume a candidate for being taken offline.

If you store related SharePoint content on the same SAN volume, when that related content is no longer frequently accessed, the volume can be taken offline to save on resources and prevent long searches. Most SAN manufacturers offer configuration settings that will allow individual volumes to be stopped and can even provide seek histories to determine which volumes are busy and which are not. In fact, some software allows for automating volume status based on configurable parameters that allows the SAN to react to SharePoint user activity.

Load Balancing: Automatic vs. Manual

Many discussions about improving SharePoint performance are centered on providing multiple servers for various roles or jobs within a SharePoint farm. For instance, answering client user HTTP requests is the job of the Web front end server role. Employing multiple WFE servers in a scaled SharePoint farm allows more concurrent user traffic and balances the workload of presenting SharePoint Web pages. Similarly, clustering SQL Server allows multiple cluster node servers to respond simultaneously to concurrent requests for SharePoint data by WFE servers and improves response times.

When choosing a load-balancing solution, determine whether the concurrent system response is to happen automatically or manually. Automatic load balancing is provided by the system software, such as a SAN configured with storage virtualization and data replication automatically balancing the bit retrieval from multiple disks. Manual load balancing requires human intervention to redirect a request to an alternate resource. Despite the best of automated load-balancing features, do not assume that the software is infallible. Monitor all automatic load-balancing designs and make sure the workload is evenly distributed.

Moving Data Sets by Migrating Volumes

Eventually, SharePoint content that is dated will become historical. To lower the cost of storing such data in accordance with regulations or for audit purposes, move the data to alternative, cheaper storage and release the space on your production SharePoint storage for more current information. Many SAN solutions ship with migration utilities to move entire volumes between disks or between heterogeneous systems. You might also consider backup strategies that compress the antique SharePoint data to removable media such as magnetic tape or DVD.

Maintaining Software and Upgrades

As with any critical IT system, all participant software in a SharePoint enterprise should be well maintained. Let us begin with the Windows Server OS on the WFE servers, MOSS application servers, SharePoint index servers, and SQL Server systems. Keep each Windows OS updated with critical patches, security updates, and service packs. In a large enterprise, this task can most easily be managed by invoking the Automatic Update Client service of the OS. To streamline Internet traffic, consider employing Microsoft Windows Server Update Service (WSUS) architecture to test, approve, and deploy OS updates to all your SharePoint servers.

Windows SharePoint Services and MOSS also require updating. However, the process of updating SharePoint in a scaled farm is unique and challenging. First, all users must be removed from SharePoint and kept out during update installation. This is most easily accomplished by stopping the WWW services on the WFE servers. Then, the SharePoint software update must be downloaded from Microsoft and installed individually on all servers in the SharePoint farm. At each server, upon installing the update, the SharePoint Products and Technologies Configuration Wizard will automatically launch. All wizards on all servers must then be advanced only to the prompt specifying binary files can only be installed via setup. Once all farm servers' wizards are at this prompt, return to the one WFE server in the farm that runs Central Administration and complete the wizard on it first.

> **Cross-Reference**
>
> For more information about updating Windows SharePoint Services, see the Microsoft TechNet article on deploying software updates at http://technet.microsoft.com/en-us/library/cc288269.aspx.

In addition to the Windows OS and SharePoint, keep SQL Server updated as well to ensure optimal performance of the SharePoint databases. And don't forget to keep your SAN solution's software current to avoid driver or network connectivity issues and keep storage virtualization, data replication, and thin provisioning utilities running at peak performance.

Summary

In this chapter, we examined implementing redundant hardware and network paths and RAID storage configurations, and setting up data replication to keep SharePoint data highly available. This chapter exposed server virtualization and database mirroring to protect SharePoint from OS and SQL Server instance failures. This chapter also compared two prominent viewpoints on traditional backup routines and discussed designing warm standby SharePoint sites via remote replication and geo-clusters for DR. Lastly, we examined routine maintenance concerns that should be on the priority list of all SharePoint administrators.

Hopefully, you have enjoyed the SharePoint storage opportunities introduced and detailed in this guide and can now make an informed choice of platforms.

Download Additional eBooks from Realtime Nexus!