EMC²
where information lives

EMC Solutions for Microsoft Office SharePoint Server

EMC® Celerra® Unified Storage Platforms

Applied Best Practices Guide

EMC NAS Product Validation
Corporate Headquarters
Hopkinton, MA 01748-9103
1-508-435-1000
www.EMC.com

Microsoft Office SharePoint Server EMC Celerra Unified Storage Platforms

Applied Best Practices Guide
P/N H4162.4

Contents

About this Document

This document provides source content that can be ported to a variety of EMC Solutions for Microsoft Office SharePoint Server (MOSS) documentation.

Purpose

Information in this document can be used as the basis for a solution build, white paper, best practices document, or training. Information in this document can also be used by other EMC organizations (for example, the technical services or sales organization) as the basis for producing documentation for a technical services or sales kit.

Audience

This document is intended for internal EMC personnel to use as content for finished documents to be available for internal EMC personnel, EMC partners, and customers.

Related documents

The following documents, located on Powerlink, provide additional, relevant information:

- *EMC Solutions for Microsoft Office SharePoint Server EMC Celerra Unified Storage Platforms – Reference Architecture*
- *EMC Solutions for Microsoft Office SharePoint Server Data Storage for EMC Celerra – Validation Test Report*
- *EMC Solutions for Microsoft Office SharePoint Server Data Backup for EMC Celerra – Validation Test Report*
- *EMC Solutions for Microsoft Office SharePoint Server on VMware ESX Server EMC Celerra iSCSI − Build Document*
- *Celerra Network Server 5.5 Best Practices for Performance − Best Practices Planning*
- *EMC Legato NetWorker Release 7.3 Microsoft Windows Version Installation Guide*
- *EMC Backup Manager for SharePoint 2.0 Release 2.0 Microsoft Windows Version Administration Guide*
- *Legato NetWorker Release 7.3 Multiplatform Version Administration Guide*

The following documents are available for download from the Microsoft website:

- *Microsoft iSCSI Software Initiator User's Guide*
- *Microsoft Storage Technologies: Deploying iSCSI SANs*
- *Microsoft SQL Server 2005 Operations Guide: Capacity and Storage Management*
- *Microsoft Office SharePoint Server 2007 Deployment for the Office SharePoint Server 2007*

Chapter 1 Microsoft Office SharePoint Server 2007 Best Practices

This chapter presents these topics:

Introduction

This section details recommendations for the configuration of a Microsoft Office SharePoint Server 2007 farm on Windows Server 2003.

Recommendation #1 Determine the type of topology to deploy

Determine the information architecture including logical structure, approximate sizing, reliability, and performance requirements for the environment in which to run the SharePoint Products and Technologies and SQL Server. Refer to the *Microsoft Office SharePoint Server 2007 Deployment for the Office SharePoint Server 2007* for more information.

Recommendation #2 Determine the capacity of the computer memory

For the best performance of SQL Server hardware, 4 GB is the minimum required memory for small size deployment, 8 GB is recommended for medium-size deployments, and 16 GB and above is recommended for large deployments.

Recommendation #3 Install MOSS 2007 SP1 on Windows Server 2003 SP2

Install Microsoft Office SharePoint Server 2007 x64 with SP1 onto a computer that has a new installation of Microsoft Windows Server 2003 x64 with SP2 (or later) and all critical updates.

MOSS 2007 SP1 has very important defect fixes and performance enhancements. The Microsoft Windows Server 2003 x64 platform offers far better memory management than other versions, which results in faster processing.

Recommendation #4 Dedicate a server to run SQL Server

Dedicate a server to run SQL Server. You should not run other farm roles on the server, unless you are deploying a single server.

For SQL Server databases hosting SharePoint Products and Technologies, it is recommended to install the SQL Server 64-bit version on a 64-bit operating system, unless you have a significant business reason not to, because it is anticipated that this is the last version of SharePoint Products and Technologies that will run on 32-bit operating systems and databases.

For optimal performance, host SharePoint Products and Technologies on the latest version of SQL Server with the latest service pack.

Recommendation #5 Limit content database size

To enhance the manageability and performance of SharePoint Products and Technologies the content database should be smaller than 100 GB. If the design requires a database larger than 100 GB, split the content from a site collection that is approaching 100 GB into a new site collection in a separate content database to avoid performance or manageability issues. This also helps ensure that the database can be backed up in a reasonable window.

Note: The limits recommended apply only to a server running SQL Server hosting SharePoint Products and Technologies, and are not general guidance for SQL Server.

Recommendation #6 Allocate storage for versions and the recycle bin

If the site will have versions or recycle bins, allocate additional storage to ensure the site quota is not exceeded. In libraries with version control enabled, the storage used for previous versions counts towards the site quota. In any site with the Recycle Bin enabled, stages will count towards the site quota. Therefore, you must plan enough additional storage space for the Recycle Bin. In addition, we recommend monitoring the number of days for retaining the deleted documents in each Recycle Bin stage.

Recommendation #7 Use quota templates to manage storage

Use quota templates to manage site collections with similar characteristics. A quota template sets storage limits for site collections, and also provides e-mail alerts when specified storage sizes are reached. Any change made to a quota template will not affect previously created sites, only new ones.

Recommendation #8 Set the Office SharePoint Server 2007 configuration order

Microsoft recommends configuring Office SharePoint Server 2007 in the order listed below. This order makes configuration easier and ensures that services and applications are in place before they are required by server types.

Install the Central Administration Web application on the following servers:

- An application server, such as a query server or a server that runs Excel Calculation Services, but not an index server (for performance reasons). If the farm will have an application server, install Office SharePoint Server 2007 on that server first. This also installs the Central Administration site.
- All the front-end Web servers.
- The index server (if using a separate server for search queries and indexing).
- The query servers, if separate from the index server.
- Other application servers (optional).

Note: Avoid configuring the index server as a query server.

Recommendation #9 Add servers to the farm

To achieve a minimal server farm configuration initially, install SQL Server on a single server followed by the installation and configuration of Office SharePoint Server 2007. This server will operate as the Web server and application server. You can add more servers at a later time if you need to expand the farm.

Note: SQL Server should be running on at least one server before installing Office SharePoint Server 2007.

Recommendation #10 Disable the Central Administration service on index servers

Since the Central Administration service is not required on index servers, disable the Central Administration service on all index servers in farms with multiple index servers. However, do not stop this service on the index server that hosts the Central Administration website. Consider the following items before stopping the service:

- Stopping this service on index servers can help avoid URL resolution problems with indexing.
- Do not stop this service for installations with only one index server in the farm.
- Make sure the service is running on another index server before stopping the service on a particular index server.

Note: Refer to the *Microsoft Office SharePoint Server 2007 Deployment for the Office SharePoint Server 2007* document for more information.

Recommendation #11 Disable the Windows SharePoint Service Web Application service

Disable all servers on the Windows SharePoint Service Web Application service that are not serving content, especially index servers. However, make sure that this service is enabled on the servers that are serving content.

To disable the service on the SharePoint Central Administration home page, click the Operations tab on the top link bar:

1. On the Operations page, in the Topology and Services section, click **Services** on the server.
2. On the Services on Server page, next to Window SharePoint Services Web Application, click **Stop**.

Note: Refer to the *Microsoft Office SharePoint Server 2007 Deployment for the Office SharePoint Server 2007* document for more information.

Recommendation #12 Plan your storage requirements for virtual machines

If you are implementing in a virtual environment, it is important to properly size your virtual disk files. These files are presented to the VM as physical disks, but actually reside as a file inside a VMFS or an NFS file system. When working with these files, a backup or replication utility will see them as a fully utilized space, even if the actual virtual disk is mostly unused. For this reason, it is important to plan your virtual disk sizes carefully.

Testing at the EMC RTP Solutions lab has shown that servers in a MOSS environment do not typically have a need for more than 20 GB to 30 GB of space in their system drive. (The requirements may be different in your environment.) It is tempting to overallocate storage space, to guard against future requirements; however, keep in mind that virtual machine disk files can be expanded in the future if additional space is required.

Recommendation #13 Create a VM template for reuse

Since many of the servers in the MOSS environment may require the same software installations, if you are using a virtual environment, we recommend creating a template of the virtual machine configured as Application Server using the “Clone to template…” option of the VC Server. Reuse this template to create the Index Server, Web Server, and any other additional Application Servers required in the farm. This will save time in the installation process, as well as provide a reference image for future expansion.

Recommendation #14 Set permissions for the VMkernel in the Celerra

If you are using a virtual environment, specify the IP address of the VMkernel in the Read/Write Hosts of the Celerra® NFS Export properties page to ensure that the Data Mover allows read/write access to the NFS export for only the VMkernels used by the MOSS environment. All other hosts with access to the NFS export should be limited to read-only access. This ensures that the file system used as a datastore by the MOSS virtual machines is not inadvertently modified by other clients.

The help file for the "NFS Exports" page in Celerra Manager provides an explanation on how the server handles export permissions based on the host access settings you specify. By default, if you do not set specific export permissions for an NFS export, all clients, referred to as hosts, that can access the Celerra Network Server have read/write access to the export.

Recommendation #15 Use the Microsoft SharePoint Capacity Planning tool for sizing

For sizing the MOSS environment, use the Microsoft SharePoint Capacity Planning tool available at:

http://technet.microsoft.com/en-us/library/bb961988.aspx

This tool takes numerous inputs, including: number of farms, number of users, user type (such as collaboration user or publishing user), expected content database size, and more. As output it provides an expected storage configuration. Based on this output, you can configure the MOSS environment.

Recommendation #16 Use the Microsoft Best Practices Analyzer for Sharepoint

Run the Microsoft Best Practices Analyzer tool called sharepointbpa.exe. This tool programmatically collects settings and values from data repositories such as Microsoft SQL, the Windows registry, metabase and performance monitor. Once collected, a set of comprehensive “best practice” rules are applied to the topology.

You can run this tool to get a detailed report listing the recommendations that can be made to the environment to achieve greater performance, scalability, and uptime.

Recommendation #17 Start the Office SharePoint Server Search and Windows SharePoint Services Help Search

Start the Office SharePoint Server Search and Windows SharePoint Services Help Search to facilitate the search feature in Microsoft Office SharePoint Server 2007.

Office SharePoint Server 2007 provides two search services: *Office SharePoint Server Search* and *Windows SharePoint Services Help Search*. Each of these services can be used to crawl, index, and query content. Each service uses a separate index.

The purpose of the Windows SharePoint Services Help Search service is to enable searching of the help system that is built into Office SharePoint Server 2007. If you do not want users to be able to search the help system, you do not need to start this service.

Use the Office SharePoint Server Search to crawl and index all content that you want users to search.

Recommendation #18 Crawl content in the site collections for search queries

Office SharePoint Server 2007 uses content sources to crawl content in your site collections. Before you use the enterprise search function in Office SharePoint Server 2007 to search for content, you must first crawl the content that you want to make available to search queries. Crawling content is the process by which the system accesses and parses content that is used to build a content index from which search queries can be serviced. Using the Shared Service Provider's administrator account, create content sources that specify the type of content and the URLs to be crawled, and the depth and time to crawl. After a content source is created, use the content source to crawl and index all the content in your Web application.

To prioritize and schedule crawls independently, a subset of content within the Shared Services Provider can be used as the source. This is particularly useful if the subset content changes frequently without having to revisit all content.

Avoid periods of peak usage and downtimes when scheduling a content crawl. Schedule full or incremental crawls based on the availability, performance, and bandwidth considerations of the servers running the search service and the servers hosting the crawled content.

Recommendation #19 Allocate storage for the SharePoint Content Index file

SharePoint uses the local file system by default to store data files created by the indexing process. Over time, the index files grow due to the indexing process. The growth of the SharePoint index files can reach 30 percent of the size of the content. A problem can result as the disk could run out of space causing the disablement of the indexing service. SharePoint's Search feature will not function properly when the disk is full. Allocate storage space sufficiently to accommodate the entire content index in a SharePoint environment. Provision storage for future growth of index files.

A healthy crawl schedule should include both incremental and full crawls. Only a full crawl will clean up index files created by the indexing process. Without full crawls, the disk used for indexes will eventually run out of space.

It may become necessary to move the index files to a larger disk. To move the indexes without re-crawling content, run the following command on the index server:

```
stsadm -o osearch -defaultindexlocation <path>
```

Recommendation #20 Disable IIS logging if security is not a requirement

IIS logging is an activity that can be disabled on a website, directory, or file. For example, it may not be required to know the rate at which a single .gif was downloaded during a particular month. Some sites group all incidental graphics in a folder and then disable logging for that folder. It is important to remember that logging provides important security-related information, so logging should always be used where content is a security risk; for example, folders that contain sensitive information. It should be noted that unnecessary levels of IIS logging can lead to performance degradation.

Recommendation #21 Set application pool recycling settings for better availability

Office SharePoint Server 2007 requires that application pools be recycled regularly. Follow these recommendations to keep sites up and running, even when you have to recycle processes for application pools:

1. Consider memory usage for recycling. When planning application pool recycling, consider the amount of memory used by each application pool and change the frequency based on the amount of memory used. Application pools that ordinarily use low memory resources will need fewer recycles than others that use lots of memory. We recommend the following settings, although your numbers will vary by how you use your installation and the features that you are using:

2. Turn off the recycle worker process (in minutes) parameter.

3. Configure a virtual memory-based recycle to occur at 1700 MB.

4. Use time-based recycles in environments with regular heavy loads at certain periods of the day. Set a scheduled recycle about 30 minutes before the peak traffic starts.

5. Recycle application pools at different times for different Web servers. If you have multiple Web servers in the farm, make sure that the application pools are set to recycle at different times on different Web servers.

6. Recycle application pools at different times for different IIS websites. Recycle different IIS websites at different times in order to avoid delay on the Web servers. If you have to recycle more than one application pool on a specific Web server at the same time, you should temporarily remove that Web server from the load balancer to avoid slow http service.

For more information, see the following resources:

- Recommendations for SharePoint Application Pool Settings (http://go.microsoft.com/fwlink/?LinkId=123977&clcid=0x409)
- Overlapped recycling and SharePoint memory-based recycling (http://go.microsoft.com/fwlink/?LinkId=125985)

- The section "Monitor and manage 32-bit worker process recycling" is available in the following book: *Planning and Deploying Service Pack 1 for Microsoft Office SharePoint Server 2007 in a Multi-server Environment* (http://go.microsoft.com/fwlink/?LinkId=125982)

Recommendation #22 Do not use Web gardens

The use of Web gardens (IIS application pools that are supported by multiple worker processes) is not recommended for enterprise content management sites because it has negative effects on page output caching. Generally, for SharePoint Products and Technologies, we have not found that Web gardens improve or detract from performance. Given the complexity of managing Web gardens, they are not recommended.

Recommendation #23 Use dedicated front-end Web servers for non-end-user services

A dedicated front-end Web server is a Web server that is not connected to the load balancer exposed to your end users. We recommend that you use dedicated front-end Web servers for running any non-end-user services that are expensive, such as:

- Search indexing
- Central Administration
- Profiles
- Excel Services

Recommendation #24 Enable only the features you need

Office SharePoint Server 2007 is a rich platform that offers many features. Your front-end resources will be better utilized if you only enable the features relevant to your user base. For details on disabling features, see the MSDN article *Working with Features* (http://go.microsoft.com/fwlink/?LinkID=105337&clcid=0x409).

Recommendation #25 Monitor SQL Server performance

MOSS depends on SQL Server for all of its data, so performance on the SQL Server will have a tremendous impact on Web front-end response times. We have found that it is best to monitor performance and capacity from the bottom of the stack to the top, because delays in the SQL Server will cause delays on front-end Web servers. For more information, see the section on SQL Server best practices.

To optimize search capabilities, it is extremely important to monitor SQL Server index fragmentation and follow the SQL Server defragmentation guidelines for SharePoint Products and Technologies provided in the Knowledge Base article *How to Defragment Windows SharePoint Services 3.0 databases and SharePoint Server 2007 databases* (http://go.microsoft.com/fwlink/?LinkID=105588&clcid=0x409). We have seen significantly improved search times in systems when this practice is followed.

Recommendation #26 Apply the ASP.NET # Induced GC counter hotfix

When you run a Microsoft ASP.NET 2.0 Web application that is built on the Microsoft .NET Framework 2.0, such as SharePoint Products and Technologies, the value of the # Induced GC performance counter increases very quickly. Additionally, CPU usage becomes high, and the computer performance decreases. To fix this issue, apply the hotfix available in this Knowledge base article, *FIX: The # Induced GC performance counter value increases quickly and CPU usage becomes high when you run an ASP.NET 2.0 Web application that is built on the .NET Framework 2.0* (http://go.microsoft.com/fwlink/?LinkId=105921&clcid=0x409).

Chapter 2 SQL Server Best Practices

This chapter presents these topics:

Introduction

This section details recommendations for the configuration of Microsoft SQL Server 2005 on Windows Server 2003.

General performance

This section details recommendations for improving general performance.

Recommendation #27 Plan for storage performance, not for capacity

The most common error made while planning the storage for Microsoft SQL Server is designing for storage capacity and not for performance or I/Os per second (IOPS). With advances in disk technology, the increase in storage capacity of a disk drive has outpaced the increase in IOPS by almost 1,000:1. With this effect it is rare to find a system that, when planned for performance, does not meet the storage capacity requirements for the workload. Hence, the IOPS capacity is the standard to be used while planning Microsoft SQL Server storage configurations. Only after considering the IOPS capacity of a configuration should the storage capacity (GB) be considered.

Recommendation #28 Use SQLIOSim.exe to validate storage configuration

When implementing any structural change to the storage subsystem used for SQL Server; we recommend that you validate that the system can support SQL Server loads. This includes the disk array, storage network, and any software that interacts with the I/O path. Microsoft publishes the SQLIOSim utility for this purpose.

Prior to SQL Server installation, or database deployment, download and install SQLIOSim.exe from Microsoft as documented in KB231619.

The SQLIOSim utility simulates the I/O patterns, and the problem identification methods used by SQL Server in order to expose potential data integrity issues. It should not be considered as a performance benchmarking tool.

Recommendation #29 Use clusters or VMHA for high availability

A cluster is a collection of servers known as nodes that together provide a single, highly available system for hosting applications such as Microsoft SQL Server 2005. Microsoft Windows Server 2003 Enterprise Edition 64-bit R2 supports clusters of up to eight nodes. The solution was validated using two nodes.

Microsoft clusters provide a highly available environment that can protect against Microsoft SQL Server 2005 server failures of hardware, operating systems, device drivers, or applications. If one of the nodes in a cluster is unavailable as a result of a failure, the service fails over to another node, which begins providing service in place of the failed node.

It is important to note that not all cluster configurations are supported in a virtual machine. At the time of publication, only 32-bit Microsoft clusters are supported by VMware. Support for 64-bit clusters is expected soon. Therefore, an alternative is to use VMware's built-in VMHA capabilities, which can achieve a similar level of high availability to Microsoft clustering.

Recommendation #30 Make SQL Server part of an Active Directory domain

The primary recommended method for security and account management in SQL Server is through Active Directory domain user accounts, using integrated security. This allows for greater security, at multiple levels, and makes user management easier.

The SQL Server should not be a domain controller, except in certain unusual circumstances. The added overhead of being a domain controller is likely to have a negative impact on the SQL Server performance.

Recommendation #31 Enable SQL Server to keep pages in memory

Microsoft SQL Server dynamically allocates and de-allocates memory based on the current state of the server, in an attempt to prevent memory pressure and swapping. However, if a process suddenly attempts to grab a substantial amount of memory, SQL Server may not be able to react quickly enough and the OS may swap some of SQL Server's memory to disk. Unfortunately, there is a good probability that the memory that was swapped to disk contains part of what SQL Server will soon be deallocating to decrease its memory use in response to the newly created memory pressure.

It is recommended that SQL Server be enabled to prevent its memory from being swapped. This is known as "Locking pages in ram". To do this, the account that the Microsoft SQL Server service is running under must be given the "Lock pages in memory" user right.

Recommendation #32 Enable Windows fast file initialization

When Microsoft SQL Server creates or expands a file, the file must be initialized. Previous versions of SQL Server had only one option and that was to initialize the space by writing all zeros to the space, which would cause a substantial performance impact if a file growth occurred. Microsoft SQL Server 2005 supports fast file initialization, which sets a file end pointer, and the process is then complete. This operation is nearly instantaneous and minimizes the impact of file growth on production systems. Fast file initialization is enabled at the OS level, by granting the user right "Perform volume maintenance tasks" to the account under which Microsoft SQL Server service is running. By default, this right is granted to administrators.

Recommendation #33 Do not allow database and log files to share physical spindles

It is highly recommended to ensure that the database data files and log files do not share the same physical spindles. This helps to prevent the loss of data due to loss of multiple drives, and improves performance.

Recommendation #34 Set your database file sizes and autogrow increments appropriately

Microsoft SQL Server 2005 supports the ability to automatically grow both data and log files as they fill. However, this should not be misconstrued as a method of database sizing. It is a best practice to set the file sizes appropriately and grow them manually at times of minimal system use, on a planned basis. Autogrowth should be used only as a safety net to prevent the files from becoming full and making the database read-only, at times when unpredicted substantial growth occurs.

When database files are expanded there is an impact to performance. This impact is minimized but not eliminated through fast file initialization.

Additionally, the file autogrowth increments should be set such that the time it takes for the growth to occur is short enough to minimize its impact to performance, but large enough to prevent many small allocations that invite file fragmentation. An adequate increase in file size that prevents fragmentation usually impacts the performance of the database. Hence, it is recommended that the file sizes be changed during periods of lowered activity on the database.

Log files have an additional issue, as there are virtual log files within a physical log file, and a virtual log file cannot span file growth increments. Thus, if the log file were set to grow at 1 MB increments, then the virtual log file would not be able to exceed 1 MB either. This will have a performance impact as discussed previously due to the impact of file expansion. This limit may also make certain transactions impossible to complete.

For all files, because of the impact to performance, it is recommended that an absolute growth increment (in MB or GB) be used instead of a percentage growth. It is also recommended that autoshrink should never be enabled. A file should not be shrunk, unless absolutely necessary, and only through a controlled manual action.

Recommendation #35 Plan your database filegroups based on your workload

SQL Server provides lots of options about how and where to lay out database tables and other structures on disk. The primary structure to control this behavior is a filegroup. Database structures are assigned to filegroups, which contain files on disk where that data can be stored. The placement of these data files is critical to the I/O performance of your database, and the recommendations for the best ways to set up filegroups vary based on the database workload. Please see “Appendix B” for a detailed discussion of how various database workloads impact your selection of a physical storage structure, and “Appendix C” for a discussion of how to plan your filegroups.

Recommendation #36 Consider table and index partitioning

Partitioning breaks up database objects by allowing subsets of the data to reside on separate filegroups. There are numerous ways that this ability can be helpful in a SQL environment.

- Improved manageability: Partitioning makes large tables or indexes more manageable. Maintenance operations can be performed on subsets of data and target only the data that is required instead of the whole table. This can be used for numerous improvements including shortened maintenance windows, improved backup coverage, and reduced backup storage requirements.

- Reduced costs: In some types of environments it is common to have very large tables that contain historical data. This data may be less valuable than current data, but still required for the application. Using partitioning allows sections of the table to be moved to lower-cost storage without impacting the whole table. For example, you could have the active part of a table using Fibre Channel (FC) disks, while the historical part of a table uses Serial ATA (SATA).

- Improved availability: Partitioning enables database partial availability, which can be used in some environments to reduce downtime due to planned, and unplanned, events. It also enables partial restores so that if a section of a table is damaged it can be restored from backup without impacting the other areas of the database.

Table and index partitioning can allow improved manageability, reduced costs, and improved availability, but they must be planned and implemented based on the needs of each individual environment. If they are implemented improperly, or without consideration for the environment they may actually have the opposite effect.

Recommendation #37 Plan the location, layout, and size of your tempdb

By default the tempdb database is rather small and gets its characteristics from the model database. Each time the Microsoft SQL Server service is started, tempdb is dropped and re-created with its initial parameters. Thus, if tempdb is initially 128 MB and during operations it autogrows to 4 GB, on restart it will be 128 MB again. Then it will have to go through the autogrow again, which will impact the performance of your database. To minimize this impact it is recommended that tempdb be sized appropriately for the environment. The easiest way to size the tempdb database is the following:

1. Start with a reasonable size tempdb for the size of databases that are in the same SQL Server instance. For example, a 1 GB tempdb database is a reasonable starting place for a sum total of instance databases between 10 GB and 100 GB, but not for 1 TB. A good starting place is to sum the total size of the databases in the instance and size tempdb between 1 percent and 10 percent of that size.

2. Set a valid autogrow increment that will allow the tempdb to grow without heavy fragmentation. The best way to do this is to set autogrow to 10 percent to 20 percent of the tempdb initial size. Do not use a percentage for the growth parameter; calculate the MB growth that corresponds to the percentage and set that as the autogrowth size. You should also make sure that fast file initialization is enabled.

3. Periodically verify the size and utilization of the tempdb database to see if it has grown significantly.

4. Reset the size of the tempdb database to something close to its size, before a shutdown. If our tempdb database from the previous example had grown from 1 GB to 5 GB, then resetting it to start at 5 GB would be advantageous, unless the new size is obviously excessive. For example, if the sum total of user databases was 10 GB and tempdb was 15 GB, this would seem excessive. It is possible that an odd set of scenarios came together to cause uncharacteristic tempdb growth. If you suspect that this may be the case, then the starting size should be set to something smaller than the current size. If the tempdb repeatedly grows to larger than is initially considered reasonable, then it is possible that this is simply the size of tempdb that is needed for your workload. From here, a DBA could diagnose what is causing the excessive growth and then determine if it is valid, or if anything needs tuning.

It is recommended that tempdb be placed on its own spindles, where tempdb and user database activity cannot cause physical disk contention with each other. The number of spindles will be determined on a case-by-case basis using the same principles that are applied to designing storage for user databases.

Recommendation #38 Use defaults for processors and memory

When Microsoft SQL Server is first installed, most of its tunable parameters are set to automatic and it is recommended that, on a server dedicated to SQL Server's use, these parameters should be left at their automatic defaults. The only time they should really be changed is if there are other workloads running on the same server, or if issues arise from the use of the defaults.

By default SQL Server will run at a standard priority and make all processors in the system available for use. Also by default, SQL Server will use as much memory as it needs until it

notices that memory pressure is starting to build. If other processes start consuming memory, SQL Server will begin decreasing its memory footprint appropriately to decrease the possibility of swapping occurring.

Recommendation #39 Use failover-aware applications

When a Microsoft SQL Server failover occurs, using MSCS clustering, Database Mirroring, or other technologies, all database connections are lost and any "in-flight" transactions are rolled back. To minimize data loss, it is recommended that all applications be failover-aware and have reconnect/retry logic. Thus, in case of a failover, the application will attempt to reconnect, and once it successfully reconnects, it will retry the transaction that was previously rolled back.

Recommendation #40 Disable hyperthreading on Microsoft SQL Server 2005 Servers

Intel hyperthreading technology allows multithreaded operating systems to view a single physical processor as if it were two logical processors. A processor that incorporates this technology shares CPU resources among multiple threads. In theory, this enables faster enterprise-server response times and provides additional CPU processing power to handle larger workloads. As a result, server performance will improve. However, testing has shown that hyperthreading can have a negative impact on many Microsoft SQL Server 2005-based processor loads. Unless it can be proven that hyperthreading helps the performance of a particular SQL Server 2005 workload, it is recommended that hyperthreading be disabled.

Hyperthreading must be disabled at the hardware (BIOS setting) level, not through the application of processor affinity or other software means.

Note: For more information, refer to http://blogs.msdn.com/slavao/archive/2005/11/12/492119.aspx.

Chapter 3 SQL Server Backup and Restore Best Practices

This chapter presents these topics:

Introduction

This section details recommendations for backing up and restoring SQL Server databases.

Recommendation #41 For point-in-time recovery use database log backups

A full database backup combined with a chain of log backups allows a database to be restored to a given point in time, at the granularity of a transaction. This is the highest level of granularity that it is possible to achieve with Microsoft SQL Server. The full backup may be taken with SQL Server native backup functionality, third-party tools, or a snapshot tool like EMC® Replication Manager. However, Replication Manager (RM) cannot perform transaction log (in SQL Server terms) backups. To achieve point-in-time recoverability, RM full backups would need to be combined with SQL Server log backups.

Recommendation #42 When possible, schedule backups for minimal disruption

When a backup is initiated, Microsoft SQL Server 2005 will do a checkpoint to flush all dirty pages to disk. When this is done on a machine that has a large amount of RAM (possibly most of which is dirty pages) that is also under a heavy I/O load, the backup may take substantially longer — sometimes as much as two to 20 times longer. Try to schedule your backups for times when the system is not under its heaviest load. It is also recommended that backup overhead be taken into account when designing RAID groups.

Recommendation #43 Do a full backup when you change the database recovery model

There are times when users might change their recovery model to simple or bulk logged, and then back to Full. The change to Full does not take complete effect until after a full database backup is performed. Therefore, a database that is changed from Simple to Full may lose data if the backup is taken before the recovery model change, instead of after the change.

A database does not start maintaining a log in Full recovery mode until a full backup is done and will maintain that recovery mode only as long as nothing is done to break the log chain. For example, if the command Backup Log With Truncate Only is issued, then the database log will no longer operate in Full recovery mode, because the log chain was broken. The only way to bring the log back into Full recovery mode is to then take another full database backup.

Chapter 4 SQL Server Database Mirroring Best Practices

This chapter presents these topics:

Introduction

This section details recommendations for SQL Server database mirroring.

Recommendation #44 If using MSCS or VMHA at the principle, do not use a witness server

When running in High Availability mode (Synchronous with a witness), in the case of a cluster or VMHA failover, it is likely that the database will fail over to the Mirror, before the cluster failover can complete. Therefore, when using MSCS or VMHA, it is recommended that only High Protection (Synchronous without a witness) or High Performance (Asynchronous) be used.

Recommendation #45 Plan for high I/O levels at the mirror site

The method that database mirroring uses to commit transactions at the mirror causes substantially more write I/O than occurs on the principle. Therefore, depending on the data load, more storage resources may be required at the mirror than at the principle. Some tests have indicated that the mirror site might need to handle as much as four times the level of I/O as the principal site. It is recommended that you monitor the mirror site for performance bottlenecks that may impact your data protection and recovery plan.

Recommendation #46 Ensure that interdatabase consistency is not needed

Database mirroring only maintains intradatabase consistency. There is no mechanism in database mirroring to maintain consistency between multiple databases. If your application requires interdatabase consistency then database mirroring as it is implemented in Microsoft SQL Server 2005 is not a recommended technology for data protection in that environment.

Recommendation #47 Use other methods to sync some SQL Server objects

Database mirroring operates only within the realm of a single database, and cannot be used for system databases like master. Therefore, in addition to database mirroring a separate mechanism must be used to keep other objects like user accounts, jobs, and security assignments above the database level, and the system databases in sync from the principle to the mirror system. This can be done with a SQL Server job that runs at a regular interval or with some third-party tools that help in this. Either way it is important to realize that dependent objects need to be synced as well.

Chapter 5 Microsoft Windows Server 2003 Best Practices

This chapter presents these topics:

Introduction

This section details recommendations for the configuration of Windows Server 2003 for use with a Microsoft SQL Server instance.

Recommendation #48 Only use hardware that is approved by Microsoft

Using hardware that is on the Windows Hardware Compatibility List (WHCL) decreases the possibility of compatibility problems and increases the level of support that Microsoft will provide, should a problem occur.

Note: In the case of using Windows on VMware ESX, use only hardware that is approved by VMware.

Recommendation #49 Use the latest verified NIC driver

For best performance and stability it is recommended to install the latest vendor NIC driver that has been validated for use with Windows 2003.

Note: In the case of using Windows on VMware ESX, use the latest driver that is approved by VMware.

Recommendation #50 When using MSCS, reboot the passive node occasionally

Many times configuration changes, especially disk configuration changes, may not be detected by the passive node until a reboot is completed. If the passive node has not detected these changes then a cluster failover may not succeed.

Recommendation #51 Use a dedicated VLAN for cluster heartbeat connectivity

If MSCS is used, it is recommended that the cluster heartbeat network be physically isolated from other networks. For example, in a two-way cluster it is common for a crossover cable to be used between the two machines as the heartbeat network.

Chapter 6 Virtual Infrastructure Best Practices

This chapter presents these topics:

Introduction

This section details recommendations when using VMware to virtualize the SQL Server environment.

Recommendation #52 Use VMware vCenter Server

The management of the virtual data center is made much easier and more featureful, through the use of VirtualCenter Server. VC Server is important for the proper functioning of several VM features such as VMHA, DRS, and VMotion. Also, VC Server adds ease of use to the following features and many others:

- VM Template creation and deployment throughout the data center
- Historical and realtime performance tracking of ESX hosts and VMs
- Tracking of VM-to-host mapping
- Initialization of new storage, including alignment of VMFS volumes

Recommendation #53 Make vCenter Server highly available

The most critical piece of any VMware ESX implementation is a combination of the VirtualCenter server and the license server (together referred to as the vCenter Server). Without these two components, which are usually installed on the same host, it is possible to lose much of ESX's functionality, including the ability to do things like DRS, VMHA, VMotion, and/or simply starting a VM. Therefore, it is important that the vCenter Server be highly available. The easiest and most reliable way to accomplish this is to run the vCenter Server on physical machines in an MSCS cluster. It is also critical that the Microsoft SQL Server database that the vCenter Server uses be implemented for high availability.

Recommendation #54 Be aware of virtual machine time considerations

Due to the way the Hypervisor-based virtualization solutions virtualize, the clock tick coming from the physical processor and some measurements of time can be skewed in the virtual machine. There are numerous ways that this can impact a system, and there are many ways to address the situation. As a best practice, refer to all the relevant documentation from your Hypervisor provider to understand how it can impact your environment. For example, VMware has a guide on timekeeping called *Timekeeping in VMware Virtual Machine*s at: http://www.vmware.com/pdf/vmware_timekeeping.pdf

Chapter 7 Network Best Practices

This chapter presents these topics:

Introduction

This section details recommendations for the configuration of your IP networks for use with Microsoft SQL Server 2005.

Recommendation #55 Use Gigabit Ethernet switches with VLAN capabilities

For best performance, use GbE switches that are capable of setting up virtual LANs (VLAN) to segment different kinds of traffic.

Recommendation #56 Use CAT6 cables for GbE connectivity

Use CAT6 cables for best performance and reliability. CAT6 cables show superior results when compared to CAT5E cables when used for 1000 Mb connectivity.

Recommendation #57 Set network speed and duplexing

Once the setup is completed and it has been verified that the infrastructure supports GbE properly, then the switch ports and NIC ports should be configured to 1 Gb/s and FULL duplex. During setup it may be necessary to use AUTO settings to ensure that everything works properly in a new environment, however, the proper speed and duplex settings should be set explicitly in production systems.

Recommendation #58 Plan for network high availability

One common oversight is to provide for high availability at the server level using clusters and the storage level using RAID, but forget about the network connecting them. To ensure uninterrupted communication between systems in your environment, plan your networks for high availability. This includes redundant switches and paths as well as redundant NIC ports and cards.

Chapter 8 Storage System Best Practices

This chapter presents these topics:

Introduction

This section details recommendations for the configuration of your EMC storage system for use with Microsoft SQL Server 2005.

Recommendation #59 Plan storage layouts for performance, not for capacity

The most common error made while planning storage for Microsoft SQL Server is designing for capacity and not for performance or I/Os per second (IOPS). To properly plan the disk layout there must be an estimate as to the number of IOPS that need to be supported on a sustained basis, the peak IOPS, and the duration of the peak.

Many customers gather data while the application is running, then use a 90th percentile to determine the level that should be planned for. There are three primary variables used for determining the number of spindles for database storage:

- IOPS (or sometimes MB/s, if a serial workload)
- RAID level — When planning for performance, striped RAID 1 (RAID 10) will require fewer spindles, than RAID for almost all read/write workloads. They are approximately equal in a read-only workload. Please see Appendix B on RAID group planning for additional information.
- Latency goals — The following are the Microsoft guidelines for good performance with SQL Server 2005:

Table 1 Latency goals

	Read	Write
Average latency	20 ms	10 ms
Max latency	50 ms	50 ms

With advances in disk technology, the increase in storage capacity of a disk drive has outpaced the increase in IOPS by almost 1,000:1. As a result, it is rare to find a system that does not meet the storage capacity requirements for the workload. Therefore, the IOPS capacity is the standard to be used while planning Microsoft SQL Server storage configurations. Only after considering the IOPS capacity of a configuration should the storage capacity (GB) be considered.

Recommendation #60 Use diskpart to align your LUNs for best performance

It is recommended to align your disk partition using diskpart. When a Windows partition is created, it is created starting at the 64th sector. This misaligns the partition with the physical disk, which can cause the I/O operation to straddle stripe element boundaries and result in a significant reduction in performance. Performance is improved as much as 40 percent when diskpart is used to align.

Note: In-depth discussion on this subject can be found in the *Using diskpar and diskpart to Align Partitions on Windows Basic and Dynamic Disks* white paper.

After the LUN creation process is complete on the production system, the active MSCS node should be able to see the LUN as a raw volume.

Partition the LUN using the Microsoft command line utility DISKPART, ensuring that the partition is created using the ALIGN=64 switch.

The following example uses diskpart against drive 4:

```
C:\>Diskpart

Microsoft DiskPart version 5.2.3790.1830

Copyright (C) 1999-2001 Microsoft Corporation.

On computer: JC27Q91X32


DISKPART> list disk

  Disk ###  Status      Size     Free     Dyn  Gpt

  --------  ----------  -------  -------  ---  ---

  Disk 1    Online       136 GB   112 GB

  Disk 2    Online       267 GB      0 B

  Disk 3    Online       267 GB      0 B

  Disk 4    Online       600 GB   600 GB

DISKPART> select disk 4

Disk 4 is now the selected disk.

DISKPART> create partition primary align=64

DISKPART succeeded in creating the specified partition.
```

Using the Microsoft Disk Manager select the drive letter or mount point to be associated with the corresponding LUN. After selecting this information, format the drive NTFS at 64k Allocation Unit Size for database data and log files.

Note: This is no longer required with Windows 2008.

Recommendation #61 Align new storage volumes on virtualized servers

When using a virtualized server, there are additional layers of abstraction that must be considered for alignment. All of the layers between the host application and the storage must be aligned, including NFS, NTFS, and VMware's VMFS. It is still necessary to align NFS and NTFS volumes by following the instructions in Recommendation #60 . If any layer in the stack is misaligned, performance will decrease.

Since VMFS is automatically aligned when the volume is created by the VMware VI client, that is the recommended method for creating a new VMware VMFS volume.

Note: See the VMware white paper entitled *Recommendations for Aligning VMFS Partitions* for more information.

Recommendation #62 Set the NTFS allocation unit to 64 KB

When formatting a new drive using Disk Administrator, the performance will be affected by the allocation or block size that is chosen. For Microsoft SQL Server 2005, Microsoft recommends using a 64k block size.

Note: For more information refer to http://www.microsoft.com/technet/prodtechnol/sql/bestpractice/pdpliobp.mspx.

Recommendation #63 Do not exceed 80% utilization of LUNs

For the best performance, the utilized drive (NTFS formatted) capacity must not exceed 80 percent. There will be performance bottlenecks if this threshold is exceeded. This is because NTFS needs additional space to work efficiently. If the space is not available, NTFS cannot function to its full potential and performance can degrade. This situation may spur additional performance degradation by creating excessive disk fragmentation.

If an iSCSI LUN that is servicing Microsoft SQL Server 2005 reaches 80 percent utilized drive capacity, do one or more of the following:

- Remove unnecessary data from the disk.
- Move some of the data to disks with more space.
- Add more disk space.

Following this recommendation will also provide you with some protection against application failure if there is unexpected growth in your database.

Recommendation #64 Use Virtual Provisioning for iSCSI LUNs

Virtual, or thin provisioning, gives the ability to present an application with more capacity than is physically allocated to it in the array. Virtual provisioning improves storage capacity utilization and simplifies storage management by presenting an application with sufficient capacity for an extended period of time. The physical storage is allocated from a shared pool of capacity to the application as it is needed. When additional physical storage space is required, disk drives can be nondisruptively added to the central storage pool. This reduces the time and effort required to provision additional storage, and avoids provisioning storage that may not be needed.

Note: It is possible to run out of storage space if steps are not taken to monitor utilization, and add more resources if needed. Virtual provisioning can be very helpful for ensuring high levels of storage utilization but must be managed so that it does not have unintended consequences such as failed writes and potential data corruption.

Recommendation #65 Use clone replicas for testing and development

When testing changes to an application where access to production data is required, but modification to that data is not, it is common to create a replica of the database so that changes can be evaluated with no impact on the production system. In such a case, it is possible to create a

logical point-in-time replica, which is called a snapshot, or a physical replica, which is called a clone.

Snapshot replicas track changed blocks of data but refer to the production data for blocks that are not changed. Clones, on the other hand, create a complete copy of the data. In a traditional testing and development scenario, the replica will be subjected to at least a sample of the production load in order to measure the impact of any changes. In a snapshot replica, this load will interact with, and could potentially negatively impact, the production workload. For this reason, it is recommended that testing and development operations where the replica will need to respond to a load should use clone replicas.

Recommendation #66 Plan storage operations for minimal disruption

Some operations on the storage array can consume resources and may cause an impact on the production system if executed during a heavy production load. For this reason, storage-based operations that may consume resources should be scheduled for non-peak hours to minimize the potential for such occurrences.

Chapter 9 Celerra Storage Best Practices

This chapter presents these topics:

Introduction

This section details recommendations for configuring EMC® Celerra® storage for use with Microsoft SQL Server 2005.

Recommendation #67 Use Gigabit Ethernet for storage connections

Maintaining optimal network performance is crucial to the deployment of Microsoft SQL Server 2005 because there is considerable network traffic generated by Microsoft SQL Server 2005. For optimal network performance, use Gigabit Ethernet cabling, switches, and network interface cards for network connections between Microsoft SQL Server 2005 servers and Celerra systems.

Use a dedicated iSCSI VLAN or a completely separate, private network for storage.

Recommendation #68 Use the most recently available Celerra software

Install the latest available Celerra release code or patch to take advantage of new features, functionality, and bug fixes. Refer to the most recent Celerra release notes for detailed information.

Note: The solution was validated using DART 5.6.37.

Recommendation #69 Use dedicated Celerra file systems for SQL Server storage

Use file systems created for Microsoft SQL Server 2005 operations only for those operations and not for any other I/O operations. This ensures more predictable performance from Microsoft SQL Server 2005.

Recommendation #70 Disable DNS on the storage network

The validated solution does not require DNS service on the storage network. Further, operational conflicts can occur if DNS service is not configured properly. For these reasons disabling DNS service on the storage network is recommended.

Recommendation #71 Use Celerra storage pools for volume management

The volume layout described in the reference architecture is designed to accommodate the I/O requirements of SQL Server database and log files while maintaining the physical separation of these elements consistent with accepted industry standards.

Note: The *Celerra Network Server 5.5 Best Practices for Performance - Best Practices Planning* white paper provides additional information.

Recommendation #72 Fail over Data Movers before rebooting

The primary Data Mover will automatically fail over to the standby Data Mover if the primary Data Mover panics or fails. However, the primary Data Mover will not automatically fail over if it is rebooted. Therefore, before rebooting the primary Data Mover, manually fail over to the standby Data Mover and make sure the database is operating properly, then reboot the primary

Data Mover and perform a failback operation. Similarly, fail over to the standby Data Mover prior to performing any maintenance on the primary Data Mover.

Recommendation #73 Use uncached mode for NFS datastores on Virtualized Servers

If you are using a virtual machine to host SQL Server then you have the option of using an NFS share on the Celerra instead of a traditional iSCSI LUN to hold the datastore. When mounting the NFS file system on the Celerra Data Mover you should add the following flag:

```
-option uncached
```

Testing has shown that under most virtual machine loads, the uncached option yields substantial performance benefits.

Note: Further information can be found on Powerlink® in the document *VMware ESX Server Optimization with EMC Celerra NFS Performance Study Technical Note.*

Recommendation #74 Verify that your TcpWindowSize setting is correct

The **TcpWindowSize** parameter of the Windows TCP/IP stack determines the amount of available buffer on the receiver side. **TcpWindowSize** should be set to 0x0000faf0 (64240) in the registry. The Windows Registry entry is as follows:

HKEY_LOCAL_MACHINE\System\CurrentControlSet\Services\Tcpip\Parameters\ TcpWindowSize (REG_DWORD)

Note: This recommendation is also cited as recommendation #105 (specifically, iSCSI recommendation #52) in the *Celerra Network Server 5.5 Best Practices for Performance - Best Practices Planning* white paper.

Recommendation #75 Configure the system for high availability

Configure the iSCSI target with multiple Network Portals. When creating an iSCSI target on the Celerra, enter more than one Data Mover interface IP address in the Network Portals text box.

Use two or more separate paths from the Microsoft SQL Server 2005 cluster to multiple Data Mover ports for high availability.

Set up the Data Movers in active/passive mode (with one Data Mover acting as a standby) for high availability.

Recommendation #76 Create persistent iSCSI target connections and bindings

If a service or application uses an iSCSI target volume or device, that volume or device must be bound in order for it to be available when the service or application is started by Windows.

To make target volumes and devices persistent, issue the **PersistentLoginTarget** command from the command line or select **Automatically restore this connection when the system boots** in the Log On to Target dialog box when configuring Microsoft iSCSI Initiator.

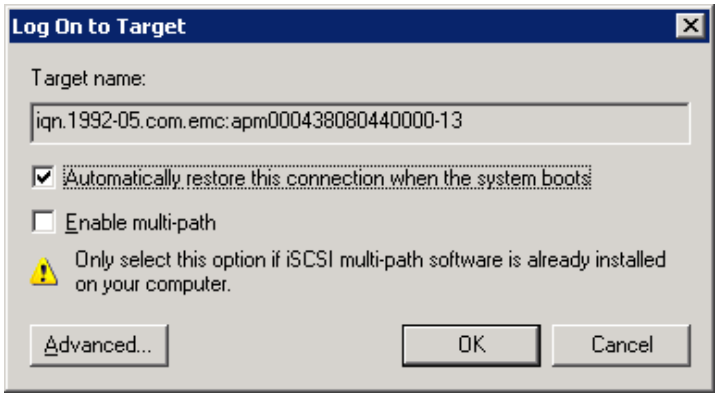

Figure 1 Log On to Target dialog box with the automatic restore option selected

To bind persistent target volumes and devices, issue the **BindPersistentVolumes** and **BindPersistentDevices** commands from the command line or use the **Bound Volumes/Devices** tab in the iSCSI Initiator Properties dialog box.

Recommendation #77 Increase your iSCSI time-out value

By default, the Microsoft iSCSI Initiator time-out value is set to 60 seconds. The time-out value determines how much time the initiator will hold a request before reporting an iSCSI connection error. The value can be increased to accommodate longer outages, such as a Data Mover or cluster failure event.

The recommendation for this setting will vary depending on how your environment is set up to respond to a failure. In most cases EMC recommends a 600-second timeout at the iSCSI initiator level.

To change the time-out value, search the Windows Registry for the MaxRequestHoldTime entry under HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet, and change the value to 600.

The following is an example of the Windows Registry entry:

HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\Class\ {4D36E97B-E325-11CE-BFC1-08002BE10318}\0001\Parameters

MaxRequestHoldTime = 0x00000258 (DWORD) (600)

Recommendation #78 Be aware of operation timings

Testing has shown that in some circumstances, operations that use Celerra replication can fail if other operations are running at the same time. For example, we have observed several times that if a differential backup is scheduled to run, and a replication update is already running, everything works as expected. However, if the differential backup is running when the replication update tries to start, an error is generated indicating that the operation has failed because a resource was busy. Manually starting the update after the backup is complete succeeds as expected.

If you are aware of the operational timings and interactions, troubleshooting the errors can be easier. Also, you can minimize the occurrence of these errors.

Recommendation #79 Use MC/S for iSCSI high performance and high availability

Microsoft recommends using Multiple Connections per Session (MC/S) instead of MPIO when the target supports MC/S, as Celerra does.

The solution was validated with Microsoft iSCSI Software Initiator 2.06 using MC/S in "round robin" mode. Please check the E-Lab™ Interoperability Navigator for the latest supported version of the Microsoft iSCSI Initiator.

Note: For more information, refer to the *Microsoft iSCSI Software Initiator User's Guide*.

Chapter 10 EMC Backup Manager for SharePoint 2.0 Best Practices

This chapter presents these topics:

Introduction

EMC Backup Manager for SharePoint provides online backup and recovery for Microsoft Office SharePoint Server and Microsoft Windows SharePoint Services (WSS). Backup Manager for SharePoint protects information to support continued operation while providing a safety net to ensure data availability.

This section details recommendations for the configuration of EMC Backup Manager for SharePoint 2.0 with MOSS 2007.

Recommendation #80 Use EBMS 2.0 for small-to-medium installations

EMC recommends using EMC Backup Manager for SharePoint (EBMS) 2.0 for small-to-medium SharePoint installations that are starting to leverage SharePoint as more of an Intranet solution for document storage and collaboration.

A small environment contains one or two servers with five to 20 users and up to 20 sites. This is characteristic of departments of medium size organizations or within small companies.

A medium environment contains two or three servers with 21-100 users, up to 500 sites. This is typically found in medium to larger organizations.

Recommendation #81 Separate production and backup-to-disk environments

To ensure backup data is kept separate from the production environment, install and run EBMS 2.0 on a separate server.

Recommendation #82 Use EMC NetWorker with EBMS 2.0 for full disaster recovery

EBMS 2.0 protects MOSS 2007 database content, websites, subsites, and individual items but does not protect SharePoint installations. To develop a SharePoint disaster recovery solution that leverages EBMS 2.0, Microsoft Windows, and SQL database backups, EMC NetWorker® for Microsoft SQL Server 5.0 and EMC NetWorker for Windows should be used with EBMS 2.0.

To understand and protect the main components of a working SharePoint environment, refer to *Microsoft Office SharePoint Server 2007 Central Administration Guide*. This document is only available as part of the Office SharePoint Server 2007 Central Administration online help.

Note: NDMP backup to disk does not require the EBMS or EMC NetWorker Module for Microsoft SQL Server for full disaster recovery; however if granular backup and restore of the SharePoint sites is required, then EBMS 2.0 should be used.

Recommendation #83 Restore MOSS 2007 and SQL Server first

In the event of full disaster, restore the MOSS 2007 installation and SQL Server before restoring SharePoint objects using EBMS. The MOSS 2007 environments operate on Windows 2003 Servers and SQL Server 2005 databases. If there is any data loss of the MOSS 2007 farm,

standalone environment, or MOSS 2007 databases then they must be restored before EBMS restores can be initiated.

Recommendation #84 Determine the backup option for EBMS

Non-optimized backup is recommended in the following scenarios:

- For SharePoint sites that require partial and item-level restores.
- For SharePoint sites that do not contain uploaded files in its document library. This is because EBMS 2.0 does not back up the uploaded files that reside in a site's document library.

Optimized backup is recommended in the following scenarios:

- When item-level restores are not required, use the Optimized backup option; this is the highest level of backup option and therefore involves shorter backup windows.
- With Optimized backups only full restores can be performed. This caters to a level of protection that meets the need for recovery scenarios.
- For SharePoint sites that contain uploaded files in its document library.

Note: When the Non-optimized back option is selected, EBMS 2.0 does not back up the uploaded files in the site's documentary library; therefore, they are not protected.

Recommendation #85 Perform partial restore from an alternate location

If granular restore is required after an Optimized backup and the EBMS administrator has taken only optimized backups of a website but now needs to do a partial restore:

1. Restore the full backup to an alternate location.
2. Perform a non-optimized backup of only the data needed (the partial data).
3. Restore to the item-level from the non-optimized backup to any other suitable location within the SharePoint environment.

Recommendation #86 Create backup cycles for all SharePoint objects

Create backup cycles for all SharePoint objects in the production environment. A backup cycle is comprised of regular full backups for all objects, probably on a weekly basis, combined with daily backups for all file changes in the object.

Recommendation #87 Check the integrity of a site before scheduling backups

Before creating a backup cycle in the Microsoft Task Scheduler, run a full backup of each live object through the normal EBMS interface to check and repair any site problems. Use the Microsoft Task Scheduler to set up and run backups in a cycle.

Recommendation #88 Settings for backing up sites

For backing up sites successfully; the following settings are recommended for the saved configuration:

1. Review the amount of disk space that is required for full and daily backups.
2. Create backup batch files and use Microsoft Task Scheduler to create the backup cycles.
3. Optionally, set up Event Log Monitors to monitor the Windows Event Log.

Chapter 11 EMC NetWorker Best Practices

This chapter presents these topics:

Introduction

This section details recommendations for the configuration of your EMC NetWorker for use with MOSS 2007.

Recommendation #89 Install the appropriate NetWorker software

1. Install EMC NetWorker version 7.4 or later and configure it per the recommendation in the *EMC NetWorker Release 7.4 Multiplatform Version Installation Guide* available on Powerlink®.

2. Install EMC NetWorker Module for Microsoft SQL Server version 5.0 or later and configure it per the recommendation in the *EMC NetWorker Module for Microsoft SQL Server Release 5.0 Administrator's Guide* located on Powerlink.

Note: NDMP backup to disk does not require the EMC NetWorker Module for Microsoft SQL Server.

Recommendation #90 Evaluate the backup-to-disk performance requirement

1. Avoid running disk-intensive applications, such as virus scanning, on the backup client when it is backing up or restoring files.

2. Resource allocation from advanced features such as Celerra Replicator™ must be taken into consideration when evaluating backup-to-disk performance requirements. Performance could be impacted if Celerra Replicator is running in the background while a backup is taking place.

Recommendation #91 Select "file" or "adv_file" type devices

1. If the destination for the backup is on the Celerra CIFS Share then select adv_file type devices in NetWorker.

2. File type devices were introduced to back up data to disk rather than to tape. With a file type device, NetWorker allows only one activity at a time. Concurrent requests are served sequentially as would be expected from a tape device.

3. In the advanced file type device, NetWorker allows concurrent access to the device. This allows backup and restore to occur concurrently.

Recommendation #92 Determine the disk capacity for data to be backed up

While deciding to perform backup to disk, determine the disk capacity to hold all the data to be backed up. The following are a few considerations:

- The amount of source data to be backed up.

- **Backup method**: Full or incremental. If full backups are performed regularly instead of incremental backups, more disk space is required to accommodate the volume of data.

- **Backup frequency**: More disk space is required if full backups are performed daily rather than weekly.
- **Backup retention**: Period of time the backup data needs to be kept on the disk before deleting or moving it to an alternate storage for longer retention. The longer the data is kept on disk, the more disk capacity is required.
- **Disk volume configuration**: When using RAID 5, approximately 20 percent of the total disk capacity is allocated to parity data. The parity information provides protection if a disk failure occurs.

Recommendation #93 Dedicate file systems for NDMP backup to disk

EMC recommends use of dedicated file systems as NDMP backup-to-disk storage. To enforce this guideline, only file systems with 99 percent or higher of their space available will show up in the list of file systems you can associate with the NDMP backup to disk.

Recommendation #94 Install the NetWorker Server on a separate machine

To ensure that the NetWorker Server is up and running while restoring the NFS file system from the backup, the datastore for the NetWorker Server should not reside on the NFS datastore of the production file system that contains the MOSS farm servers.

Recommendation #95 Configure an auto-detected NDMP SCSI jukebox

For NDMP backup to disk, use the jbconfig command to configure an auto-detected NDMP SCSI Jukebox from the command prompt of the NetWorker Server; for more information refer to *EMC NetWorker 7.4 Multiplatform Version Administration Guide* available on Powerlink.

Recommendation #96 NDMP variables for NetWorker with Celerra

Set the following options in the NetWorker software for the NDMP backup to disk.

SNAPSURE=Y: Should be set so that a crash-consistent checkpoint of the Celerra file system is captured every time the backup is started. This checkpoint will be used for the NDMP backup of the production file system. The SNAPSURE=Y option processes and automates the checkpoint creation, management, and deletion tasks upon initiation of the backups; it also facilitates the backup operation while the server is running.

UPDATE=Y: Should be set to force the timestamp information to be updated with the time, date, level, and file system that is being backed up. This value must be set to `y`. When a file system is backed up at the same level as its previous backup, the old timestamp information entry is overwritten with the new information.

DIRECT=Y: Should be set to invoke DAR functionality. This option optimizes the data recovery operation by allowing the NDMP client to access backed-up data directly anywhere in a tape set without having to traverse the tape set sequentially. This method can save significant time in the restore operation.

HIST=Y: Should be set to create a file history. Set this option before you perform a backup in order to later perform a Direct Access Restore operation.

Recommendation #97 NetWorker settings for performance

Network performance is key to achieving optimum performance. The network must be tuned to avoid congestion and to guarantee that the backup window is not limited by network bandwidth.

• Use a dedicated backup network by configuring a separate network.

• Alternatively, use VLANs, to segregate backup from production network traffic.

• Network interface cards (NICs) for servers and clients must be set to full-duplex.

• Use link aggregation on the Celerra to increase the link speed beyond the limits of a single Ethernet port.

There are many network options available on Celerra using link aggregation and virtual interfaces. Carefully evaluate your network requirements for all of the traffic that will be served by the Celerra, then design a configuration that meets those needs.

Chapter 12 NetWorker Module for Microsoft Applications Best Practices

This chapter presents these topics:

Introduction

The EMC NetWorker Module for Microsoft Applications (NMM) works with Microsoft Volume Shadow Copy Service (VSS) technology to provide snapshot backup and recovery services for file systems, application data and operating system data.

The NMM client allows for the creation of point-in-time snapshot data. A snapshot can be retained on storage volumes for quick access. One can also perform a rollover of a snapshot to a traditional backup medium such as a tape, file type device, or advanced file type device. Data can be recovered either from a snapshot or from the backup.

Recommendation #98 Different policies for application server data

Typically, application data is backed up several times a day while operating system data and volumes require less frequent backups. For application servers such as Microsoft SharePoint Server, back up the server application data under a different schedule than the host operating system data and volumes. To accomplish this, create separate backup groups for operating system data and application data. Assign the appropriate snapshot policy and client resource to each backup group.

Recommendation #99 Installation path for application server program

Do not install application server program files on the same volume as the application database files and log files.

Recommendation #100 NetWorker modules and the NetWorker for Microsoft Applications Client

If you attempt to use both the NMM Client and a NetWorker module to back up application data, the module backups will be promoted to Full backups.

Recommendation #101 Use EMC NMM 2.0 for SharePoint full disaster recovery

NMM provides recovery for the NMM Client itself, as well as the applications NMM protects, such as Microsoft Office SharePoint Services. NMM backs up the SharePoint configuration and content databases, and the query and index servers. Disaster recovery of the entire NMM Client machine requires completion of the following tasks:

1. Recovery of the NMM Client
2. Full recovery of the applications on that machine

SharePoint Disaster Recovery is more than restoring a full backup. In a disaster recovery, you need to restore the Internet Information Service (IIS), SQL, and SharePoint Servers that make up the farm, including file systems and registries of each machine.

To understand and protect the main components of a working SharePoint environment, refer to the *Microsoft Office SharePoint Server 2007 Central Administration Guide*. This document is only available as part of the Office SharePoint Server 2007 Central Administration online help.

Recommendation #102 Keep backups and recoveries in sync

For a complete backup strategy, you should set up a consistent schedule of full farm-level backups.

When scheduling a backup of individual content databases in between the full farm level backup of SharePoint, we recommend you to perform a full farm-level backup immediately after a significant change to the SharePoint farm (such as adding new SharePoint objects to the existing farm, creating a new web application). The existing NetWorker backup configuration should also be updated with the new SharePoint objects so that you have consistent backup data.

Recommendation #103 Multiple Client resources for the same NMM client host

In a NetWorker server you can create multiple client resources for the same NMM Client and these clients can be added to different backup groups that can be scheduled for backup. If these client resources are added to different backup groups that are scheduled to run at different time intervals, then you should ensure that the Start Time attribute for each backup group is spaced far enough apart so that the backups of the host's client resources should not overlap to avoid contention.

For example, in a NetWorker server we can have NMM client resources with the same save sets backing up the data to different target devices, and if these client resources are placed in a backup group scheduled to overlap at a particular interval of time, then this leads to the contention of data to be backed up between the backup groups. Ensure that the Start Time attributes are spaced far apart for the client resources.

Recommendation #104 Rollback of SharePoint SQL databases

NMM does not support a SharePoint SQL rollback. SharePoint content databases are backed up by NMM, but the master SQL database is not. The master database can be backed up separately and then used in a SharePoint SQL rollback. To back up SharePoint for rollback of SharePoint SQL databases, take a backup of the master database each time a NMM SharePoint backup is taken. To perform a full rollback, restore the backups created by the NMM SharePoint backup and then restore the SharePoint SQL master database.

Chapter 13 Backup with Data Protection Manager 2007 Best Practices

This chapter presents these topics:

Introduction

Microsoft System Center Data Protection Manager (DPM) expands the basic data protection capabilities included in SQL Server by adding the ability to provide backups for selected databases with more granular control over your recovery time objective (RTO) and recovery point objective (RPO). Using only the tools provided with Windows Server and SharePoint, it is possible to take periodic full backups, but the frequency of these backups will vary according to the speed of your backup system and the amount of data you need to back up. The frequency at which you can create backups will control both the RPO and the RTO available to you. For example, with nightly tape backup, your RPO or "potential data loss" will be one business day, meaning that any server outage will likely cost up to an entire business day of data (and productivity) that will be unrecoverable. Meanwhile, your RTO, indicating how long it will actually take to recover, will vary according to the amount of data that has to be restored.

Recommendation #105 Satisfy DPM 2007 Deployment Best Practices

Prior to System Center Data Protection Manager (DPM) 2007 installation, ensure that the DPM server and the servers it is going to back up meet network and security requirements. You must also ensure that they are running on supported operating systems and that they meet the minimum hardware and software requirements.

For information about DPM 2007 deployment, see *Deployment Best Practices* (http://technet.microsoft.com/en-us/library/bb808899.aspx).

Recommendation #106 Apply Hotfix 940349 for the protected computers

The DPM agent specifically requires the update rollup package 940349 to resolve issues with Volume Shadow Copy Service (VSS) snapshot in Windows Server 2003. The 940349 patch fixes approximately seven issues with VSS snapshot, for details, see (http://go.microsoft.com/fwlink/?LinkId=99034). Therefore, before you install protection agents on the computers you are going to protect, you must apply hotfix 940349. You must install the hotfix on your 64-bit and 32-bit servers. If you are installing a protection agent on Windows Vista, the 940349 hotfix is not required.

For more details, see Microsoft Knowledge Base article 940349, "Availability of a Volume Shadow Copy Service (VSS) update rollup package for Windows Server 2003 to resolve some VSS snapshot issues" (http://go.microsoft.com/fwlink/?LinkId=99034).

Note: After installing hotfix 940349 and restarting the DPM server and/or the protected server, we recommend that you refresh the protection agents in DPM Administration Console. To refresh the agents, in the Management task area, click the Agents tab, select the computer, and then in the Actions pane, click Refresh information. If you do not refresh the protection agents, Error ID: 31008 may appear because DPM only refreshes the protection agents every 30 minutes.

Recommendation #107 Apply DPM 2007 feature pack

Apply the DPM 2007 Update Rollup feature pack that includes previously released hotfixes alongside several new features aimed to enhance overall applicability and usability of DPM. The feature pack is available at http://support.microsoft.com/default.aspx/kb/949779/

Recommendation #108 Plan and deploy your recovery topology

For DPM that is protecting Office SharePoint Server, we recommend that for site and item recovery, you configure a recovery farm — a second farm — that is only used to restore data. The recovery farm is not intended to be a live farm. The recovery farm does not need to have the same hardware as your primary farm. We recommend that you use a single server installation or a virtual farm.

For information about deploying the recovery farm, see "How to Create a Recovery Farm" (http://technet.microsoft.com/en-us/library/dd180789.aspx).

Recommendation #109 Create a separate backup network

To ensure that DPM backups do not slow down your primary network, we recommend that you configure a separate backup network using DPM 2007 Management Shell (PowerShell) cmdlets (custom commands). The backup network address is created when you put separate network interface cards (NICs) on the DPM server and the protected servers and connect them through a separate LAN. As a result, backup data traffic does not impact the primary network.

For more information on setting up your backup network address, see "Using Backup Network Address" (http://technet.microsoft.com/en-us/library/cc964298.aspx).

Setting up your network

Before you can set up a backup network address, you need to:

1. Ensure that the name resolution of the protected server on the DPM server can resolve the backup address of the protected server and vice versa.

2. Configure the backup subnet and the corresponding subnet mask using Add-BackupNetworkAddress.

Note: The subnet should cover the entire range of network addresses for the DPM server and the servers you intend to protect.

3. Restart the DPM agent on the DPM server and the protected computers. It may cause ongoing tasks to fail. Post a restart, watch out for alerts, and perform the recommended actions, if needed.

Recommendation #110 Start and register the WSS Writer Service for the protected front-end Web server

Before you can start protecting server farms on servers running Windows SharePoint Services 3.0 or Microsoft Office SharePoint Server 2007, you must start and configure the Windows SharePoint Services VSS Writer service (WSS Writer service).

After you install the protection agent on the Windows SharePoint Services Web Front End (WFE) server, you must provide the protection agent with the credentials for the Windows SharePoint Services farm.

You perform the following procedure for a single WFE server. If your Windows SharePoint Services farm has multiple WFE servers, you must select only one WFE server when you configure protection in the Create New Protection Group Wizard.

To start and configure the WSS Writer service:

1. On the WFE server, at the command prompt, change the directory to <DPM installation location>\bin\.

2. Type ConfigureSharepoint.exe.

3. When prompted, enter your Windows SharePoint Services farm administrator credentials.

The administrator credentials you provide for the Windows SharePoint Services farm must be a local administrator on the WFE server.

Note: DPM uses a single front-end Web server to protect the server farm. When you add other front-end Web servers or remove front-end Web servers other than the server used by DPM, there is no impact on protection of the farm. If the front-end Web server that DPM uses to protect the farm is unavailable, or if you need to remove the front-end Web server that DPM is using while continuing protection of the server farm, see "Changing the Front-End Web Server" at http://technet.microsoft.com/en-us/library/bb809017.aspx.

Note: You must rerun ConfigureSharepoint.exe whenever the Windows SharePoint Services farm administrator password changes.

Recommendation #111 Plan your Initial Replica creation process

In DPM, a replica is a complete copy of the protected data on a single volume, database, or storage group. The DPM protection agent on the protected computer sends the data selected for protection to the DPM server. A replica of each member in the protection group is created. Replica creation is one of the more resource-intensive DPM operations, with its greatest impact being on network resources.

Typically, the performance of the replica creation will be limited by the speed of the network connection between the DPM server and the protected computers. That is, the amount of time that it takes to transfer a 1 GB volume from a protected computer to the DPM server will be determined by the amount of data per second that the network can transmit.

On an extremely fast network, such as a gigabit connection, the speed of replica creation will be determined by the disk speed of the DPM server or that of the protected computer, whichever is slower.

To avoid the network load of replica creation, we recommend that you create replicas manually from tape or other removable media when creating the initial replica, which can take from hours to days, depending on the amount of data to protect. For more information, see "Creating Replicas Manually" (http://technet.microsoft.com/en-us/library/bb808911.aspx).

Note: If the network goes down during synchronization, DPM will attempt to continue the synchronization from the point where it left off last. If the network goes down during consistency check, DPM will attempt to continue the check if the network comes back up in 5 minutes. However, if the network remains down for longer than 5 minutes the replica is marked as Inconsistent.

Recommendation #112 Use network bandwidth usage throttling

To better manage network traffic, we recommend that you configure network bandwidth usage throttling for each protected computer. In addition, you can specify different network bandwidth usage throttling rates for work hours, non-work hours, and weekends, and you can define the times for each of these categories.

Network bandwidth usage throttling limits the amount of network bandwidth that DPM can use to create and synchronize replicas. Throttling helps to ensure that network bandwidth is available to applications other than DPM.

The advantage of using network bandwidth usage throttling is that it enables you to limit the amount of network resources a synchronization job can consume. The disadvantage of network bandwidth usage throttling is that it can lengthen the amount of time each synchronization job takes to complete.

For more information, see "Network Bandwidth Usage Throttling" (http://technet.microsoft.com/en-us/library/bb808938.aspx).

Recommendation #113 Disable the protection agent while performing general maintenance on servers running Windows SharePoint Services

General maintenance includes tasks such as disk and file maintenance, updating operating systems and applications, and protecting data by using antivirus software and performing regular backups. Some special considerations apply when you are performing server maintenance on computers running Microsoft Windows SharePoint Services that are protected by System Center Data Protection Manager 2007 (DPM).

When you need to perform maintenance on a protected server and do not want protection jobs to continue for the duration of the maintenance, we recommend you disable the protection agent.

To disable a protection agent:

1. In the DPM Administrator Console, click **Management** on the navigation bar.

2. On the Agents tab, in the display pane, select the name of the computer with the protection agent you want to disable.

3. In the Actions pane, click **Disable protection agent**.

4. In the dialog box, click **OK** to confirm that you want to proceed.

Recommendation #114 Stop protection by retaining the replica when a database is removed

When a database is added to or removed from a Windows SharePoint Services farm, DPM will mark the replica as inconsistent and alert the administrator.

When a database is added, the alert includes a link to modify the protection group. After you complete the Modify Group Wizard, DPM performs a consistency check. Protection of the farm, including the newly added database, continues.

When a database is removed, you should stop protection of the server farm using the retain replica option, and then add the farm to the protection group again.

Appendix A Performance Monitoring and Tuning

This appendix presents these topics:

Introduction

Note: Some Windows counters may not be correct in a VM. Because of the nature of how VMs share processor resources, any counters that rely on timing based on processor ticks will be skewed and therefore unreliable

Two principle items can have tremendous impact on a SQL Server storage subsystem: server memory and I/O subsystem setup (number of physical disks, RAID level, paths to the storage, and so on).

Server memory serves as the SQL Server's primary cache. In general, the more memory available for database caching, the fewer I/O operations the storage subsystem will need to service. The number of physical disk drives and RAID level used (where the database data files and log files will be stored) determines the sustainable I/O rate the database can use, without exceeding acceptable latencies. Adding more drives will generally increase the I/O rate, provided the storage connection bandwidth is not exceeded. Using the proper RAID level will also benefit the number of sustained I/Os the system can handle. Also, additional storage connection paths or additional storage arrays can be added to lighten the load of saturated storage components. Monitoring and tuning SQL Server database installations should be a standard practice in all deployments.

There are several tools and methodologies that can be used to help monitor and tune performance, including SQLTrace (SQL Server Profiler is the GUI), Dynamic Management Views (DMVs), Perfmon, and others. The intent of this section is not to exhaustively discuss any of these tools or methodologies. This section is intended to provide more detailed information about some of the most commonly used Perfmon counters.

Windows Performance Monitor

The Windows 2003 Performance Monitor (Perfmon) allows administrators to view or collect realtime performance counter information on a wide variety of operating system, SQL Server, and hardware components. Table 2 provides a brief description about the counters.

Table 2 Windows Performance Monitor counters

Counter name	Performance object	Description
% Processor Time	Processor	An indication of how busy the system processors are; if the processors are very busy, the I/O system is not likely to impact the overall system performance. It is important to note that when using an Intel-based system with hyperthreading turned on, this counter is no longer accurate. In such a system, if this counter is hovering near 50%, it is probable that the machine is processor-bound (the processor is the bottleneck).
Pages / sec	Memory	This is one of the key indicators of operating system level memory pressure. Some level of occasional paging is normal for most systems; however, if there are sustained periods of substantial paging or the paging occurs during periods of poor performance, then this is a strong indication of a memory shortage.
% Idle time	PhysicalDisk or LogicalDisk	An indication of how much time a given disk/volume is not busy. If it is busy almost all the time (near 0%), the disk/volume may be a bottleneck.

Counter name	Performance object	Description
Avg. Disk Queue length	PhysicalDisk or LogicalDisk	Average number of read and write requests outstanding, over the sample interval, on the disk/volume. Large queues and high disk-busy times usually indicate performance bottlenecks. When using RAID arrays, this number can be much higher than for a single physical disk drive, because a RAID array is backed by many physical disks. A useful rule of thumb is to try to keep the queue length to less than two per physical disk in the RAID array.
Current Disk Queue length	PhysicalDisk or LogicalDisk	The instantaneous number of outstanding requests on the disk, at the exact moment of sampling. Useful for tracking down temporal hot spots. If the disk queue varied between 0 and 128 during a sample interval, then this counter could be anything from 0 to 128, even though the average over the sample period might be 16. For example, if during the sample interval the disk queue length samples were (0,0,0,128), the average disk queue length would be 32, but the current disk queue length would be 128. For most situations, this is not a useful counter.
Avg. Disk sec / transfer	PhysicalDisk or LogicalDisk	Average response times, in milliseconds (ms), for disk read and write operations. Typical values of fewer than 10 ms are very good. Consistent values greater than 20 ms may indicate a problem.
Avg. Disk sec / Read	PhysicalDisk or LogicalDisk	Average response time (ms) for read operations. Useful for further isolating general response time issues. An average of less than 20 ms is desirable.
Avg. Disk sec / Write	PhysicalDisk or LogicalDisk	Average response time (ms) for write operations. Useful for further isolating general response time issues. An average of less than 10 ms is desirable.
Disk Bytes / sec	PhysicalDisk or LogicalDisk	Number of bytes transferred to or from the disk/volume per second.
Disk Transfers/sec	PhysicalDisk or LogicalDisk	Number of transfers to or from the disk/volume, regardless of transfer size. Otherwise known as IOPS.
Avg. Disk Bytes / Transfer	PhysicalDisk or LogicalDisk	A measure of the relative I/O composition of the system. This is an average, but on disks/volumes that exclusively contain database data files, it will tend toward 8 KB for most random data workloads. If the value is significantly higher than 8 KB, the workload may be more sequential in nature and benefit from additional caching. For disks/volumes that contain database log files, this value can vary substantially depending upon the workload.

Counter name	Performance object	Description
Buffer Cache Hit Ratio	SQLServer : Buffer Manager	This object is useful for helping to determine whether SQL Server has sufficient memory available to it. Values of 98% or higher are excellent, 94% or higher are acceptable, and lower values are either an indication of insufficient memory or an extremely random data workload. Buffer Cache Hit Ratio is approximately (Page Lookups per sec – Page Reads per sec) divided by Page Lookups per sec. ReadAheads are not considered a cache miss, since they are not immediately being requested by a query processor; however they are disk reads and make the Buffer Cache Hit Ratio a bit misleading. For example, if all pages needed from the disk were prefetched by the read ahead manager, then the Buffer Cache Hit Ratio would be at or near 100%, which is only semi-true, as a Buffer Cache Hit Ratio of 100% would imply no physical reads would be occurring. Therefore, it is not uncommon to use an alternate calculation that takes ReadAheads into account, such as (Page Lookups per sec – (Page Reads per sec + ReadAhead Pages per sec)) divided by Page Lookups per sec.
Page Life Expectancy	SQLServer : Buffer Manager	This object is also useful for helping to determine whether SQL Server has sufficient memory available to it. Values of less than 300 (5 minutes) are usually an indication of insufficient memory.
Page Lookups / sec	SQLServer : Buffer Manager	Number of requests to find a page in the buffer pool. This counter includes pages that are requested from the buffer pool, not found, and read in from disk (Page Read).
Page Reads / sec	SQLServer : Buffer Manager	Number of physical database page reads issued. Pages reads in this counter are a direct result of needing the page to finish the execution of a query.
ReadAhead Pages / sec	SQLServer : Buffer Manager	Number of pages read in anticipation of use. This is done by the ReadAhead manager and is also termed a prefetch. Effectively, if SQL Server sees certain types of I/O activity, it will try to get the data from disk, before it is actually requested. For example, if a serial scan of a table is being done, rather than reading in individual pages (8 KB I/Os), the ReadAhead manager will begin prefetching entire extents (64 KB I/Os), or several contiguous extents (128, 256, 512 KB, and so on), as a single I/O request, thereby increasing the efficiency of the I/O subsystem and improving Query response time.

Appendix B RAID Group Planning

This appendix presents these topics:

Introduction

All RAID groups have two important quantities that can be consumed by an application workload: storage capacity and performance capacity. It is possible to have a RAID group whose capacity (size in gigabytes) is fully used, but the performance capacity of the RAID group is underutilized. However, it is far more common for a RAID group to be under a demanding performance load, while using only a portion of its storage capacity. Since the increased capacity of disk spindles has dramatically outpaced the increase in the performance of those spindles, one of the most common errors encountered in SQL Server deployments is to find a RAID group that has been designed for capacity instead of performance.

When designing a RAID group, the first thing to consider is the level of performance that is needed, then verify if the needed capacity is available. The methods of designing a RAID group discussed herein take a conservative view of RAID group design and do not rely on features such as caching or prefetching for sustained performance.

RAID level attributes

There are three primary RAID levels that are often discussed and offer fault tolerance: RAID 1, RAID 5, and RAID 1 with striping (RAID 10).

Note: RAID 0 is sometimes discussed but seldom recommended because it does not provide fault tolerance. Any spindle failure in a RAID 0 group will render the entire group unusable

The discussion of the various RAID levels is usually done in the rudimentary terms of capacity, whereas RAID 1 and 10 are thought of as 2n or needing twice the number of disks to store a given amount of information when compared with a non-RAID protected system; and RAID 5 is thought of as n+1 or needing one additional spindle than is required outside of a RAID implementation to store the data. These facts are true, but are seldom the most important part of a design discussion. As previously presented, the most common limitation of a disk array for SQL Server is its performance capacity, not storage capacity.

Each RAID level has a performance impact based on the type of fault tolerance that RAID level implements. This performance impact is normally seen only during writes. For example, RAID 1 and 10 both use mirroring, therefore everything that is written to a given spindle must also be written to its partner mirror. For this reason, each logical I/O issued by the host server actually turns into two physical I/Os inside the storage array and RAID 1 and 10 are thought of as having a 2x write penalty.

RAID 5 is more difficult to understand. One might think that RAID 5 would have some sort of x+1 write penalty, like its n+1 overhead in storage capacity. However, RAID 5 actually has a 4x write penalty. The specifics of this calculation are out of the scope for this document but easily obtainable from trusted sources. Effectively each logical I/O from the host is broken down into four physical I/Os:

- Read data disk being written
- Read parity disk for the stripe parity value
- Write new data to data disk
- Write new stripe parity to the parity disk

To summarize: RAID protection has an impact to the usable storage capacity, and performance capacity of an array of disks. The impact is determined by the RAID level that is selected for the array.

As the percentage of writes increases from 0 percent (read only) to 100 percent (write only), the gap in performance between RAID 1 or 10 and RAID 5 widens rapidly. The following graph shows the effects of RAID level on different % write / % read workloads.

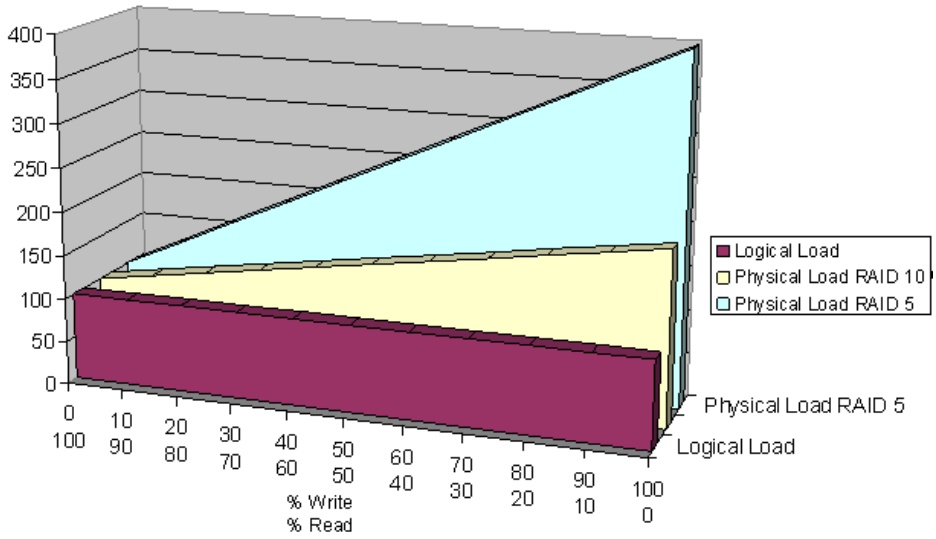

Figure 2 RAID overhead effect for random I/O

Table 3 presents some facts about the three RAID levels that are being discussed.

Table 3 RAID level performance characteristics

RAID level	Random	Serial	Read	Write
1[1]	Good[1]	Good[1]	Good[1]	Good[1]
5[2]	Moderate	Good	Excellent	Poor[3]
10	Excellent	Excellent	Excellent	Better[1]

Notes:

1. All RAID 1 groups are by definition limited to two drives. This places a distinct upper limit on their potential performance. RAID 10 is a method for striping data across multiple RAID 1 groups to avoid this limit.

2. RAID 5 takes a substantial performance impact during the failure of a drive and subsequent rebuild of its replacement. Therefore, this should be taken into account when planning.

3. Although RAID 5 writes perform poorly, because of the 4x penalty, there is a special case that is rarely found, where RAID 5 writes can actually outperform RAID 1 or RAID 10 called a "full stripe write." This occurs when the write is aligned with the stripe and is the exact same width as a full stripe. For example, if each disk held 32 KB per stripe and a 4+1 array was created, then a full stripe write would need to be 128 KB in size and be aligned so that it created one full stripe across the disks. Improved performance cannot be guaranteed unless testing on your workload proves this.

Estimating required performance

One of the often misunderstood facets of Microsoft SQL Server is that it is not an application. It is an environment that houses databases of various types and attributes. The performance characteristics of one database can vary substantially from another database. Therefore, it is not possible to discuss how Microsoft SQL Server will perform in general for all possible workloads, but it is possible to discuss the performance of a given database when the workload characteristics are defined. Databases are usually broken down into two general classes, OnLine Transaction Processing (OLTP) and OnLine Analytical Processing (OLAP).

The usual attributes of the I/O patterns involved with each type of database, as well as tempdb, are discussed in Table 4.

Table 4 Microsoft SQL Server file types and performance attributes

File type	Performance attributes
User Database Data File (OLTP)	The database data file for most OLTP-type applications usually has the following characteristics: • Smaller I/Os

File type	Performance attributes
	• Random I/Os • High percentage of writes compared to reads • Not usually a very large database (aged data is usually archived to a data warehouse) Based on this, RAID 10 will usually provide the best performance for a given number of spindles. To state this another way, the needed performance can usually be achieved with fewer spindles using RAID 10, rather than RAID 5.
User Database Data File (OLAP or Data Warehouse)	The database data file for most OLAP-type applications usually has the following characteristics: • Larger I/Os • Serial I/Os • Low percentage of writes compared to reads, sometimes read-only • Usually a very large database Based on this, RAID 5 will usually provide adequate performance and much more usable space for a given number of spindles.
Database Log File	The database log file(s) for all databases have the following characteristics: • Smaller I/Os (some multiple of 512 bytes). • Highly serialized I/Os. • Almost exclusively writes, with occasional reads during large rollbacks or log backups. • Size is dependent upon several factors and difficult to predict without more details about the database workload. • A log file is the single most important piece of information for database recovery from either a crash or database restore. • Every transaction that modifies data is limited by log write speed. Because of the critical nature of the log files both in terms of performance and recoverability, RAID 10 is the recommended standard for database logs. There are times when RAID 5 may provide adequate performance (because of full stripe writes), but upon drive failure, RAID 5 performance will likely drop below needed levels. Also, RAID 5 cannot survive a double drive fault, while it is possible that RAID 10 will survive such a failure.
tempdb Data File	The database data file(s) for tempdb usually has the following characteristics: • Smaller or larger I/Os, depending upon usage, but many times it is larger I/Os. • Serial or random I/Os, although a given workload might be somewhat serial, many workloads running simultaneously may give tempdb more of a random I/O appearance. • Usually a near 50/50 split of writes and reads. • Size can vary wildly. Based on the unpredictable nature of tempdb combined with its usually large percentage of writes, RAID 10 will usually provide the best performance for a given number of spindles.

Please remember that these are general characteristics and that a specific user database might generate an I/O workload that varies substantially from those presented. Therefore, the only real way to determine the I/O performance needs of a given database is to run tests with that database.

For example, in an OLTP-type database, it is critical to know what level of I/O performance will be needed from the data and log RAID groups in terms of IOPS. SQL Server has many inherent buffering algorithms that are used to try to decrease I/O levels and the efficiency of these algorithms is entirely database and workload dependent.

Therefore, to get accurate performance estimates it is best to run tests with as close to "real world" conditions as possible. During these tests, Performance Monitor Logs can be used to capture the characteristics (reads per second and writes per second) of the volumes used for storing database files. This information can then be used to do an initial RAID group design.

Note: IOPS counters averaged over time should not be used as the basis for a RAID group design. It is recommended to find the 90th percentile of the IOPS samples and design for that performance level. This will allow your system to respond well to spikes in demand.

Note: The IOPS requirements for a RAID group should be calculated independently for both reads and writes

Calculating disk spindle requirements

Once the read and write IOPS are known, they can be plugged in to the following formula.

$$\#ofSpindles = \frac{ReadsPerSecond + (WritesPerSecond * RAIDMultiplier)}{Recommended\ IOpsPerSpindle}$$

Note: Your spindle count may need to be adjusted to conform to the requirements for the RAID level you have selected. For example, you cannot build a seven-spindle RAID 10 set. In such a case you would need to build an eight-spindle RAID 10 set.

Now that we know the formula and two of the three variables needed (read and write IOPS), the only variable still needed is the number of IOPS that a given rotational speed spindle can support.

This number is even harder to compute than any of the others discussed so far and has a very broad range of possibilities. Primarily the number of IOPS that a spindle can support is derived either by observation of workloads or computed mathematically with certain assumptions made. The primary factor that influences the number of IOPS a spindle can support is the distance between where the current I/O is occurring and where the next I/O will occur. This is referred to as the "locality of reference."

Since it is impossible to know the locality of reference ahead of time in all non-serial workloads, certain assumptions are common to estimate how many IOPS a spindle is capable of maintaining. These assumptions usually involve taking an average between the minimum distance that would need to be moved and the maximum distance that would need to be moved and then averaging them.

Note: The exact math involved in making these assumption and estimates is beyond the scope of this document. It is generally accepted that a 15,000 rpm disk spindle can support approximately 180 IOPS under a variety of common workloads and conditions while maintaining an acceptable latency. We will use this estimate.

The following table shows some estimated spindle counts for a few example workloads, using the formula supplied previously. Notice that in almost every example the number of spindles required for RAID 10 to hit a certain performance level is lower than that of RAID 5. Therefore, it could be said that since fewer spindles are needed, when it comes to non-read-only workloads, RAID 10 is less expensive than RAID 5.

Table 5 shows the number of spindles that would be needed for various workloads using either RAID 5 or 10.

Table 5 Number of spindles required for a series of sample workloads

Total IOPS	%Read	%Write	Read IOPS	Write IOPS	RAID 5	RAID 10
1000	100	0	1000	0	7	6
1000	75	25	750	250	11	8
1000	50	50	500	500	15	10
1000	25	75	250	750	19	10
2000	100	0	2000	0	13	12
2000	75	25	1500	500	21	14
2000	50	50	1000	1000	29	18
2000	25	75	500	1500	37	20

After a RAID group initial configuration is determined, it is then recommended that it be tested under the specific workload that it will be performing to ensure that it meets the required performance level.

Another important point is that I/O levels are being discussed at the RAID group level. Therefore, if multiple LUNs exist on a given RAID group, then the aggregate sum of the needed IOPS performance from all LUNs on that RAID group would need to be used in the previous equation for its design.

Summary

All of the data in your database environment must at some point pass through the disk subsystem. At present, for most implementations, this will have some level of RAID protection. An understanding of the various RAID protection levels and their impact to both the storage capacity and performance capacity of your disks is critical when designing a database for your workload.

Appendix C Filegroup Planning

This appendix presents these topics:

Introduction

The proper number of data files per filegroup (or database, if a single filegroup is used) is dependent on a number of factors.

tempdb

In SQL Server 2000 the recommendation for tempdb was to have one file for every CPU core. Therefore, if you have a machine with four CPU sockets and use dual-core CPUs, tempdb should be broken into eight files. This is primarily because of an issue with contention in the GAM and SGAM areas of the files. Contention has been decreased substantially in SQL Server 2005, but it is still a possibility since tempdb will potentially be used much more than in previous versions, because of new features like row versioning. As usual, there is a cost associated with having multiple files per filegroup. This is because SQL Server will stripe data across all files in a given filegroup (proportional fill) and will likely cause all the files to be accessed simultaneously. If the files are located on LUNs that are on the same RAID group (set of spindles), this will induce head movement, which increases latency and decreases throughput. Therefore, the decreased contention of multiple files must be balanced against the increased I/O load. A good midway starting point might be to break tempdb into a number of files equal to half the number of CPU cores.

The following Microsoft KB articles discuss this in more detail.

FIX: Concurrency enhancements for the tempdb database

"Possible thread starvation detected" message may be reported in the SQL Server error log when all worker threads for a particular UMS scheduler are in a wait state in SQL Server

User databases

For user databases, a similar decision must be made and similar criteria should be used. However, the contention of user databases is usually much lower, and the user (administrator) should start with a smaller number of files (like one or two) and then increase, if necessary.

Log files

Increasing the number of files available to a SQL Server database does absolutely nothing for performance. If a database has two log files, Microsoft SQL Server will fill the first log file before beginning to use the second log file. Therefore, the only use of a second log file is to expand a database's logs onto a new volume.