White Paper

Dell Reference Architecture Guide

Deploying Microsoft® SQL Server™ 2005 with Microsoft Windows Server® 2008 on Dell™ PowerEdge™ Servers and Dell Storage

Abstract

This white paper provides an architectural overview and configuration guidelines for deploying Microsoft SQL Server 2005 x64 SP2 with Microsoft Windows Server® 2008x64 Enterprise or Standard Edition on Dell PowerEdge servers with Dell PowerVault and Dell/EMC CX3 storage. Using the knowledge gained through joint development, testing and support with Microsoft and EMC, the Dell Reference Architecture Guide for SQL Server documents "best practices" that can help speed SQL Server 2005 solution implementation and help simplify operations, improve performance and availability.

Table of Contents

Introduction

Dell™ PowerEdge™ servers, Dell/EMC CX3 family of storage and Dell PowerVault™ storage systems can be used to deploy highly reliable and sustainable Microsoft® SQL Server™ 2005 databases. The purpose of this whitepaper is to provide a detailed view of the Dell SQL Server 2005 Reference Architecture deployed on Windows Server 2008. It will outline best practices for configuring Dell server and storage hardware and provide Dell recommendations for configuring Windows Server 2008 and SQL Server 2005 for an optimal solution deployment.

This Reference Architecture Guide is intended to help IT professionals design and configure Microsoft SQL Server 2005 database solutions using Dell servers and storage that apply "best practices" derived from laboratory and real-world experiences. This white paper provides a recommended approach for implementing SQL Server 2005 x64 SP2 database on Dell PowerEdge 9th and 10th generation servers, Dell PowerVault MD3000 SAS storage system, Dell/EMC CX3 Series of storage and Microsoft Windows Server 2008x64 Enterprise or Standard Edition.

Dell Deployment Guidance whitepapers provide a fast time-to-market by performing point in time testing of a specific configuration or use case, which enables Dell to rapidly test a new or emerging solution, document this testing and provide customers with guidance on selecting an appropriate hardware / software combination.

Dell Deployment Guidance whitepapers document the testing results in the form of a technical white paper that captures details of the hardware and software tested and provides tips and hints for deploying and operating the solution. Dell Deployment Guidance whitepapers for SQL Server are delivered online at www.dell.com/sql .

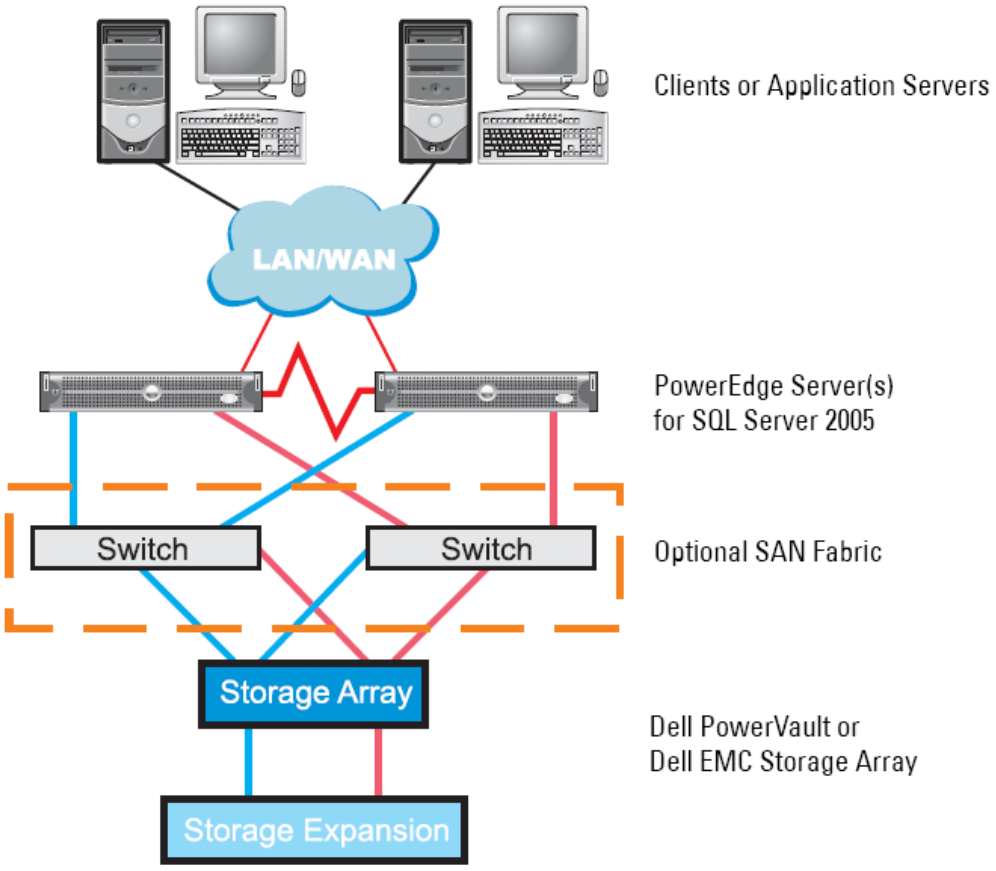
Overview of this White Paper

The balance of this white paper will provide the reader with a detailed view of the Dell Reference Architecture for SQL Server 2005, best practices for configuring the hardware and software components and pointers for obtaining more information.

Architecture Overview

An architectural overview of the Dell Solution for SQL Server 2005 is shown in Figure 1 below, comprised of the following components:

- Client systems that access data stored within the SQL Server database
- Client-server network made up of network controllers, cables and switches
- Dell PowerEdge servers running Windows Server 2008 and SQL Server 2005
- Server-storage interconnect using Fibre Channel or Serial Attached SCSI (SAS)
- Dell/EMC CX3 Series of storage or Dell PowerVault MD3000 storage



Use Hardware and Software that are Tested and Validated by Dell

Using hardware that is on the Solution Deliverable List (SDL) at the end of this paper, can help ensure the successful deployment of your SQL Server 2005 solution by using components that have been tested and validated by Dell. All of the components on the SDL are also listed on the Windows Server Catalog. Using hardware that has been validated by both Dell and Microsoft will minimize the risk of encountering compatibility problems, and helps provides a solid foundation for your SQL Server database.

Similarly, to help assure the performance and stability of your Dell Solution for SQL Server, it is recommended to install the latest tested and validated software, firmware, and driver versions for HBAs, NICs, storage arrays, and other components as listed in the SDL and Server Catalog .

NOTE: In SQL Server 2005 failover clusters, it is recommended that all nodes be configured with identical hardware, and employ the same software, driver, and firmware revisions.

Storage Configuration

Configure Storage Host Bus Adapters

Depending on the configuration of your Dell Solution for SQL Server 2005, either the Dell PowerVault MD3000 or the Dell/EMC CX3 Series storage arrays may be used. For configurations with the MD3000 it is a minimum requirement that at least one dual-port SAS 5/E controller be installed on the PowerEdge server hosting SQL Server 2005 to provide redundant links to the PowerVault MD3000 storage array.

Connecting both ports of the dual-port SAS 5/E controllers protects against potential data loss or corruption if either a SAS 5/E controller port or MD3000 storage controller fails. Similarly, when using a CX3 storage array, it is recommended that at least two 4Gb Fibre Channel ports are available on the PowerEdge server. This can be accomplished using two single-port Fibre Channel host bus adapters (HBAs), or using a dual-port HBA.

To configure dual storage array controllers for a Power Edge Server hosting SQL Server 2005, follow the procedure described in the appropriate Dell SQL Server 2005 Deployment Guide.

For a list of validated HBAs, consult the Solution Deliverable List at the end of this paper.

For connection and configuration details for the Dell PowerVault MD3000 or the Dell/EMC CX3-series**, see** Appendix A**.**

Configure Disks for SQL Server 2005

The physical disks in the MD3000 or Dell/EMC CX3 storage array provide the physical storage capacity for the SQL Server database. Before data can be stored, the physical storage capacity must be configured on the storage array. Depending on the storage array in the solution, different terms may be used to refer to similar groupings of physical disks into volumes that are presented by the storage array to the Windows Server operating systems, and by the OS to SQL Server 2005. Throughout this paper, the terms in Table 1 may be used interchangeably. A disk group (MD3000) or RAID Group (CX3) is a set of physical disks that are logically grouped and assigned a particular RAID level. Each group of physical disks provides the overall capacity needed to create one or more virtual disks (MD3000) or LUNs (CX3), which are logical entities that the server uses to store data.

Table 1 - Storage Logical Entities and Terminology

Logical Entity	MD3000 Terminology	Dell/EMC Terminology
Set of physical disks that contains one or more logical disks of the same RAID level	Disk Group	RAID Group
Logical disk or RAID volume that is presented to the operating system	Virtual Disk	LUN
Set of one or more hosts and one or more logical disks that are accessible by the host(s)	Host Group	Storage Group

SQL Server 2005 uses three specific storage areas: TempDB, Transaction Log File, and Data files for user defined databases. TempDB is a system database which is automatically created during the installation of SQL Server 2005. TempDB is a shared working area for all databases on the server, for various activities, including temporary tables, sorting, processing sub queries, building aggregates to support GROUP BY or ORDER BY clauses, queries using DISTINCT, cursors, and hash joins. Therefore, the required size of the TempDB area can vary depending on the type and complexity of database queries. Complex queries that aggregate data from many tables, such as those required for business intelligence workloads, will utilize the TempDB more extensively and will require additional storage space for this purpose.

In addition to TempDB, SQL Server 2005 also maintains a Transaction Log File for each database. The Transaction Log File stores the details of all the modifications made to a SQL Server database and the details of the transactions that performed each modification. This information is critical for maintaining database consistency and for aiding database recovery. In SQL Server 2005, data for user defined databases are stored in a separate data storage area. This data file contains user defined database objects such as tables, indexes, and stored procedures. All databases have at least one data file and at least one

transaction log file. Data from a given database is physically stored in one of the data files allocated to that database.

It is a best practice to separate the TempDB and Transaction Log files onto their own virtual disks on separate disk groups. The separation of TempDB and log files can enable better I/O performance by ensuring that these files do not share the same physical disks. Figure 2 illustrates a sample disk group and virtual disk configuration which separates TempDB, transaction logs, and data onto distinct sets of physical disks.

When designing the storage layout, various considerations need to be evaluated depending on the business requirements for the database system. These considerations include the following:

- Cost
- Performance
- Reliability/Availability
- Manageability
- Scalability and capacity management
- Type of Workload: OLTP vs. OLAP

The following figure is a sample storage layout on a MD3000 storage array with separate virtual disks for TempDB, transaction logs, and data files. Because of the RAID5 data set, this example is better suited read intensive applications like, Data Warehousing. RAID5 provides good performance and offers higher capacity than RAID 1/0 and is best suited for a read-intensive environment. However, RAID5 is less suitable for write-intensive OLTP workloads due to the overhead associated with calculating and writing parity information, RAID 1/0 is recommended for transaction processing workloads. The maximum MD3000 storage configuration can be expanded to include up to two MD1000 storage enclosures, for a total of 45 spindles.

Figure 2 - Data Warehousing on MD3000 storage

Microsoft recommends that the number of virtual disks/LUNs required for data should be equal to the number of CPU socket. Table 2 and Table 3 illustrate recommended virtual disk configuration and sizes for single-socket host servers and for dual-socket host servers respectively.

Table 2 - Single-Socket Server - Storage Configuration

Virtual Disks	Minimum Size	Disk Group RAID	Used for
Data1	Database size	10, or 5 for read-only	data
TempDB	Depending on required TempDB size	10, or 1	TempDB
Logs	10% of database size	10, or 1	logs

Table 3 - Dual-Socket Server - Storage Configuration

Virtual Disks	Minimum Size	Disk Group RAID	Used for
Data1	Database size /2	10, or 5 for read-only	data
Data2	Database size /2	10, or 5 for read-only	data
TempDB	Depending on required TempDB size	10, or 1	TempDB
Logs	10% of database size	10, or 1	logs

RAID 1/0 is generally considered the optimal choice for SQL Server 2005 virtual disk implementations because it offers balanced read and write performance and high fault tolerance by combining the best attributes of mirroring and striping[1]. When possible, the disk groups on which the transaction logs, TempDB and data virtual disks reside should be configured with RAID 1/0.

Because additional drives are required to implement RAID 1/0, it may not be the preferred choice for certain databases. In these cases, RAID 1 can be used as an alternative for the disk group of the TempDB virtual disk. For the disk group containing the transaction logs, RAID 1 is an alternate recommendation as RAID 1 provides protection from drive hardware failure. For the disk group containing the data files, RAID 5 provides a cost effective alternative especially for predominantly read-only workloads such as a data warehouse database. However, RAID 5 is not suitable for write-intensive workloads, such as in an OLTP database, due to the additional overhead associated with re-reading the data, calculating parity, and writing parity.

Each virtual disk created in storage will be mapped to one partition at the OS level. Depending on your server configuration, several data files may be created on this partition. For details on operating system partitions, please refer to the "Configuring File System" section below.

Network Configuration

This section provides best practice recommendations for setting Network speed, configuring Network Interface cards (NICs) and using IPv6, which is installed and enabled by default with Windows Server 2008.

The server requires a unique public IP address. The best practice is to have the public IP address registered with DNS. If a DNS server is not available, then the IP address must be registered in the hosts file in any client system or application server that accesses the SQL Server.

For a single server running SQL Server 2005, it is recommended that at least two network interfaces are used, and that the interfaces be teamed (or bonded) to provide additional redundancy for the database server.

For a SQL Server 2005 cluster, each server must contain at least two network interfaces: one for the cluster Interconnect (heartbeat), and one for client connections (public). It is recommended that a third network interface be added, which allows network interface teaming to be employed on the public network.

Set Network Speed and Duplex

Although automatic detection of network adapter and link properties (such as link speed and duplex) is useful during the initial configuration of your Dell Solution for SQL Server, these automatic settings introduce the possibility of negotiating a connection with diminished network performance. Therefore, it is recommended that production SQL servers be manually configured to connect at 1 Gb/s with full duplex

[1] "Physical Database Storage Design", Kathy Lu, Lewis Bruck, Microsoft TechNet, February 12, 2007. http://www.microsoft.com/technet/prodtechnol/sql/2005/physdbstor.mspx

communication. This configuration can generally be configured in the NIC driver properties, and may also be configured on managed or web-managed GbE switches.

Configure NIC Teaming

To help guard against network access failures, PowerEdge servers can be configured to provide redundant links to the client network. Using NIC teaming software, two NIC interfaces on the PowerEdge server can be bonded together to operate under a single IP address. Provided the NICs are from the same vendor, NIC teaming software provides load balancing and failover functionality, balancing the workload and routing network traffic between the two NIC interfaces. If a failure occurs, affecting one of the NIC interfaces – examples include switch port failure, cable disconnection, or failure of the NIC interface itself -- network traffic is routed to the remaining operable NIC interface. Failover occurs transparently to the SQL Server 2005 database with no network communication interruption or changes to the server IP address.

NOTE: In a SQL Server 2005 failover cluster, NIC teaming can only be used on the public network. NIC Teaming is not supported on the heartbeat network.

Configure Failover Cluster Networks

When deploying a SQL Server 2005 Instance on a Windows Server 2008 Failover Cluster, ensure that you are following best practices for configuring the cluster, including the following

Use Static IP Addresses for Cluster Nodes and Clustered SQL Server Instances
All cluster networks should be configured using static IP addressing. This reduces the chance that SQL Server 2005 resources will become unavailable due to failures of DHCP servers or other network infrastructure components. Each node should use a static IP address for its public and private network interfaces. The cluster management and SQL server instance IP addresses should also be statically assigned.

Use a Dedicated Network or Virtual LAN (VLAN) for the Cluster Heartbeat
The cluster public and private networks must be configured to use distinct IP subnets. Further, it is recommended that the cluster heartbeat network be physically isolated from other networks. In a two-node cluster, a simple cable between the two machines can be used for heartbeat network. The gigabit Ethernet adapters in Dell PowerEdge servers generally do not require the use of a crossover cable when they are directly connected in this manner. Alternately, a dedicated switch or VLAN can be configured for the node-to-node communications within the cluster.

Use IPv6 to Improve Network Security in Windows Server 2008

The size of an address in IPv6 is 128 bits, which allows for large address spaces. In IPv6, IPSec support is a protocol requirement. This requirement provides a standards–based solution for network security needs for devices, applications and services, and promotes interoperability between different IPv6 implementations. Failover clusters fully support IPv6 for both node-to-node and node-to-client communication.
For more information on IPv6, please refer the link below:

http://technet.microsoft.com/en-us/library/bb726954.aspx

Operating System Configuration

Set the NTFS File System Allocation Unit Size

All SQL Server volumes, including those used for Data, Logs and TempDB, should be formatted using the NTFS file system, and it is recommended that the allocation unit size for these volumes be set to 64 KB[2].

2 Predeployment I/O Best Practices http://www.microsoft.com/technet/prodtechnol/sql/bestpractice/pdpliobp.mspx

For additional recommendations regarding SQL data files and volumes, refer to *Configuring the File System* in this document.

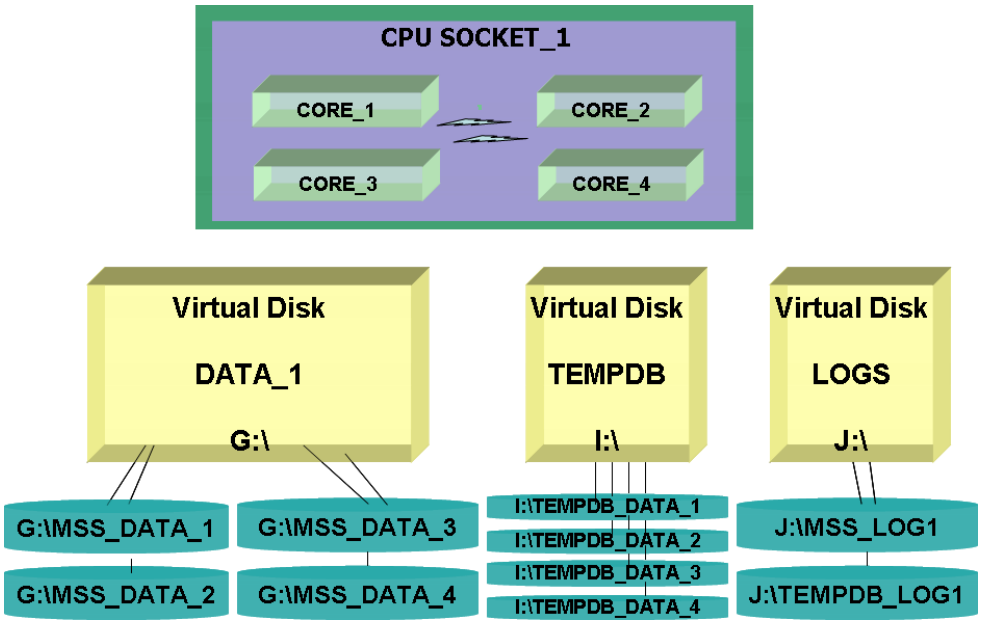
Disk Partition Alignment

The performance of an I/O-intensive application, such as SQL Server 2005, can be greatly influenced by how the partition tables, the master boot record (MBR), and other metadata are laid out on the actual LUNs on which the SQL Server data resides. Historically, manual steps were required to ensure that this metadata did not cause a partition misalignment which could inadvertently require extra I/O operations for every read or write performed by SQL Server 2005. In Windows Server 2008, this is no longer a concern, because the first partition is automatically aligned to 1024KB – which can accommodate virtually any common RAID configuration and thereby eliminates the partition misalignment issue. For detailed information, please refer to Appendix C

Configure the File System

In SQL Server 2005, each database consists of at least one data file and at least one transaction log file. As mentioned in the *Configure Disks for SQL Server 2005* section above, the recommended number of virtual disks for data should equal the number of physical CPU cores, and each virtual disk created in the storage system is mapped to an operating system file partition. To achieve optimal performance that scales with heavy workloads, Microsoft recommends that the number of data files configured for a SQL Server 2005 database equal the total number of CPU cores installed on the server[3]. Therefore, within each of the partitions used for data, the recommended number of data files should equal the number of cores per CPU socket. However there is a cost associated with having multiple files in a file group. If these files share spindles then this can lead to contention when they are accessed simultaneously. Therefore a decision to balance the decreased contention of multiple files against the increased I/O load has to be made. Based on our lab results, the recommendation is to break data files into a number of files equal to half the number of cores.

For example, a PowerEdge server with a single quad core processor should have four data files for each data volume as illustrated in Figure 3 and Table 4



[3] Lu and Bruck, op. cit.

Figure 3 – File System Layout for a PowerEdge server with a Single Quad-core Processor

Table 4 – Data, Log, and TempDB Files for a PowerEdge Server with a Single Quad-core Processor

Drive Letter	Folder Name	Format	Used for
G:	MSS_DATA_1, MSS_DATA_2, MSS_DATA_3, MSS_DATA_4	NTFS 64 KB	data
I:	TEMPDB_DATA_1, TEMPDB_DATA_2, TEMPDB_DATA_3, TEMPDB_DATA_4	NTFS 64 KB	TempDB
J:	MSS_LOG1, TEMPDB_LOG1	NTFS 64 KB	Logs

The following table illustrates a sample file system configuration for a quad-core, dual-socket host server; details on configuring the TempDB files are covered below.

Table 5 – Data, Log, and TempDB Files for a PowerEdge Server with Two Quad-core Processors

Drive Letter	Folder Name	Format	Used for
G:	MSS_DATA_1, MSS_DATA_2, MSS_DATA_3, MSS_DATA_4	NTFS 64 KB	Data
H:	MSS_DATA_5, MSS_DATA_6, MSS_DATA_7, MSS_DATA_8	NTFS 64 KB	Data
I:	TEMPDB_DATA_1, TEMPDB_DATA_2, TEMPDB_DATA_3, TEMPDB_DATA_4, TEMPDB_DATA_5, TEMPDB_DATA_6, TEMPDB_DATA_7, TEMPDB_DATA_8	NTFS 64 KB	TempDB
J:	MSS_LOG1, TEMPDB_LOG1	NTFS 64 KB	Logs

The number of TempDB data files should equal the number of CPU cores. Refer to section "Configuring TempDB files" below for more details.

Check Total Memory Allocation

When you allocate memory for multiple instances of SQL Server (clustered or non-clustered), make sure that the summed value of the maximum server memory settings for all the instances of SQL Server, plus any other cluster resource and local application requirements, is less than the physical RAM available.

Deploy SQL Server in 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. Using Windows authentication provides ease of integration with application servers within the domain, and provides a centralized facility for managing user rights and roles.

NOTE: Cluster nodes running SQL Server 2005 must belong to an Active Directory domain. It is recommended that these nodes be configured as member servers, and not as domain controllers.

Provide High Availability for SQL Server

SQL Server 2005 includes support for Database Mirroring, which can be used to create a remote copy, or mirror, from a database on one SQL Server instance to a distinct SQL Server instance that resides on a separate server. Database mirroring can be used among single server instances, and can also be used in conjunction with instances configured within a failover cluster. Where failover clusters use a single, shared copy of the data, database mirroring provides a mechanism to replicate the data on the remote node. The

replication engine relies on the transaction logs to keep the remote copy of the database updated. Mirroring can be configured in either high-performance (asynchronous), or high-safety (synchronous) modes – depending on the distance between the nodes, the quality of the network transport, and the needs of the application tier. Database mirroring provides one mechanism to enhance the availability of your database. Another way to provide availability for your database is to deploy a SQL Server 2005 instance in a failover cluster. A failover cluster combines two or more server nodes with a shared storage device and provides a means for applications, such as SQL Server 2005, to resume operation in the event that one of the server nodes experiences a failure. In this context, a named instance of SQL Server 2005 can be installed within a "cluster resource group," which should contain the following:

- A network name and IP address, via which the SQL Server instance can be accessed by client systems or application servers
  At least two physical disk resources, to house the SQL Server data and log files
- The specific SQL Server resources that are created when a named instance of SQL 2005 is installed on set of clustered server nodes

For general information about configuring Microsoft Windows failover clusters with your Dell PowerEdge servers and Dell storage arrays, visit www.dell.com/ha. Throughout this paper, there have been extensions and adaptations of several best practices that explain differences that may apply when using a SQL Server 2005 instance within a failover cluster. Windows Server 2008 has simplified the deployment of failover clusters, and provides new tools that can be used to aid in deployment and troubleshooting of clustered applications. You should consider the following additional best practices when using SQL Server with Windows Server 2008 Failover Clusters:

Using the Cluster Validation Wizard
This new validation wizard in failover clusters, will allow you to perform tests to determine whether your system, storage, and network configuration is suitable for a cluster. These tests include specific simulations of cluster actions, for the following categories:

- **System Configuration tests** analyze whether the selected servers meet specific requirements for cluster membership, for example, the requirement that the servers must run the same operating system version and software updates.
- **Network tests** analyze whether the planned cluster networks meet specific requirements such as network redundancy.
- **Storage tests**. Tests to analyze whether the storage meets specific requirements, for example, whether the storage correctly supports the necessary SCSI commands and handles simulated cluster actions correctly.

Select a Quorum Type to Maximize Availability
Select a configuration option such that the quorum is not a single point of failure. With improvements in failover clusters, you can use the two cluster models that previously existed—the quorum resource model and the majority node set model—or a "hybrid" of the two. For example, in a two-node cluster, you can specify that if the quorum disk (now called a "witness disk") becomes unavailable, the cluster continues running as long as the copies of the cluster configuration database on the two nodes remain available.

Reboot any Passive Cluster Node When Altering Storage
Because of the way Windows Server detects and manages disk volumes, when the underlying disk configuration for a failover cluster changes, the passive node may not detect these changes until it has been rebooted. If the passive node has not detected these changes, then it may not be able to bring disk resources online and a cluster failover may not succeed. Therefore, it is recommended that the passive node be rebooted after any changes to the LUNs being presented to the cluster nodes.

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

Configure Non-Critical Services Within the Cluster
To prevent the failure of specific services from causing the SQL Server group to fail over, configure those services using Cluster Administrator in Windows, as follows:
Clear the **Affect the Group** check box on the **Advanced** tab of the **Full Text Properties** dialog box. However, if SQL Server causes a failover, the full-text search service restarts.[4]

If installing SQL Server 2005 components such as Integration Services, Notification Services, or Workstation Components, Microsoft Distributed Transaction Coordinator (MSDTC) should be installed and network DTC access should be enabled. To help ensure availability between multiple clustered applications, it is highly recommended that the MSDTC have its own resource group. If MSDTC cannot be configured to have its own resource group, the recommended alternate choice is to use the Cluster group and Quorum drive.

Follow Cluster Security Best Practices
In a SQL Server 2005 failover cluster, different service accounts have distinct privileges and management roles. The following best practices help provide separation of duties between these roles:

- The cluster service and SQL Server service domain accounts need to be a member of the Local Administrators group of each node, but they should not be a member of the Domain Administrators group.
- Use the same cluster service and SQL Server service accounts for all clusters in the same domain.
- The cluster service and SQL Server service accounts should only be used for their own specific purposes.
- By default, all local administrators of cluster nodes can use Cluster Administrator. This privilege includes the ability to take resource offline including SQL Server.
- Do not use local accounts on clustered nodes. Always use domain accounts for clusters. This is because during a failover, local account information is not failed over.
- If the BUILTIN/Administrators account is removed ensure that the account that the Cluster Service is running under can log into SQL Server for the IsAlive check. If it cannot, the IsAlive check fails.
- Auditing—Administrators can use auditing to capture information about who accessed a cluster and when it was accessed. Enable login auditing at the Operating System and SQL Server level. Examine the audit for login failure events and look for trends to detect any possible intrusion

For more information on security features refer this link

http://technet2.microsoft.com/windowsserver2008/en/library/13c0a922-6097-4f34-ac64-18820094128b1033.mspx?mfr=true

SQL Server 2005 Configuration

Separate Physical Spindles for Database Data Files, Log Files and TempDB

It is highly recommended that the database data files and log files and TempDB do not share the same physical spindles. This helps in preventing the loss of data in case of multiple drives failure, and in improving the performance. Refer to *Configure the File System* in this document for more information.

Failover Cluster Troubleshooting http://msdn2.microsoft.com/en-us/library/ms189117.aspx
Dell Reference Architecture for Microsoft SQL Server 2005 on Windows 2008 Server

Enable SQL Server to Keep Pages in Memory

Based on the current state of the server, Microsoft SQL Server dynamically allocates and de-allocates memory in an attempt to prevent memory pressure (condition when limited amount of memory is available) and swapping. However, if a process suddenly attempts to grab a substantial amount of memory, for example, a complex query that aggregate data from multiple tables, then 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 de-allocating 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 perform this, the account that the Microsoft SQL Server service is running under must be given the "Lock pages in memory" user right.

Enable Windows Instant File Initialization

By default, when a new data file is created or a data file is expanded as a result of AUTO GROW, the pages on the file are initialized by writing zeros before the file can be used. The file zeroing process can be time consuming and can lead to application timeout, especially during the AUTO GROW activity. One of the major improvements of SQL Server 2005 is the instant file initialization feature. This feature skips the file zeroing process when a data file is created or a file size is increased. Instead, the pages on the data file get overwritten with new data when pages are actually allocated. Microsoft's testing for creating and growing files shows a significant performance improvement when the instant data file initialization is used[5].

To enable instant file initialization, you must run the SQL Server 2005 service account under a Windows account and assign the Windows SE_MANAGE_VOLUME_NAME permission to that Windows account. Follow the procedure described in the appropriate Dell SQL Server2005 Deployment Guide to Configure a Windows Account to enable Instant File Initialization.

Set the Database File Sizes and Autogrow Increments

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 only be used as a safety net to prevent the files from becoming full and making the database read-only, when unpredicted substantial growth occurs. When a database's files are expanded there is an impact on performance. This impact is minimized but not nullified through fast file initialization.

Additionally, the file auto growth increments should be set such that the time it takes for the growth to occur is short enough to minimize its impact on 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 files are manually grown 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 an impact on the performance as discussed above 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. This way the file growth will be predictable as file sizes change. It is also recommended that Auto Shrink should never be enabled. A file should not be shrunk, unless necessary, and only through a controlled manual action.

[5] Lu and Bruck, op. cit.

Plan Database Filegroups Based on Workload

SQL Server provides many options for the data architect 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. The recommendations for the best ways to set up filegroups vary based on the database workload. Refer to *Configure Disks for SQL Server 2005* in this document for details.

Configure TempDB Files

When the SQL Server 2005 database is initially created, the TempDB files are placed on server internal drives. Follow the procedure described in Appendix B, to move the TempDB files to external disk drives dedicated for TempDB files.

By default, a single TempDB data file is created during installation of SQL Server 2005. Adding additional TempDB data files can help avoid latch contention on allocation pages and help solve I/O performance issues. The recommended number of TempDB data files should match the number of CPU cores of the server. To better utilize the allocation mechanism, all TempDB data files should be equal in size. But, there is a cost associated with having multiple files in a filegroup. SQL Server stripes data across all files in a filegroup and there is likelihood that all files may be accessed simultaneously. If these files are located on LUNs that share same physical disks (on same RAID group), this will induce head movement which in turn increases latency and decreases throughput. Therefore a decision to balance the decreased contention of multiple files against the increased I/O load has to be made. The recommendation is to separate tempdb into a number of files equal to half the number of cores. For more information, please refer to the following Microsoft KB article: "Microsoft SQL Server I/O subsystem requirements for the tempdb database" - http://support.microsoft.com/kb/917047/en-us

By default, the AUTO GROW option is set to "on" for TempDB files. However, expanding TempDB too frequently can lead to performance degradation. To avoid this issue, it is recommended to pre-allocate the TempDB space with a size large enough to accommodate the expected workload and set the file growth increment large enough to minimize TempDB expansions. In addition, Microsoft recommends setting the TempDB files FILEGROWTH increment to 10%[6].

Set Processor and Memory Parameters in a SQL Server

Default settings for processor and memory parameters on a SQL Server run at a standard priority which allow the use of all the processors in the system and also make available as much RAM as needed by SQL Server. But if you are running other applications on the SQL Server that are accessing the system resources such as memory and processor , then these settings can be changed depending on the application requirement to achieve desired performance

.

Use Microsoft SQL Server Best Practices Analyzer Tool

The Microsoft SQL Server Best Practices Analyzer is a diagnostic tool that gathers configuration information from an instance of Microsoft SQL Server, performs specific tests on the instance of SQL Server , proactively verifies that the configuration is set according to recommended best practices, reports all settings that differ from the default settings, reports recent changes in the instance of SQL Server.
For every issue, the SQL Server Best Practices Analyzer provides three kinds of data:

- Issues that were detected in the scanned SQL Server instance.
- Recommended configuration.
- Links to more detailed information about the issue and related topics.

---

[6] "Working with TempDB in SQL Server 2005", Wei Xiao, Matt Hink, Mirek, Sunil Agarwal, Microsoft TechNet, June 26, 2006, http://www.microsoft.com/technet/prodtechnol/sql/2005/workingwithtempdb.msp

Solution Deliverables List(SDL) for Dell Solution for SQL Server 2005 x64 on Microsoft Windows Server 2008 x64

This section contains detailed listing of server and storage hardware configurations, firmware, driver, OS and database versions for this white paper.

Minimum Required Component List			
	Validated Component(s)	**Minimum Single Node DB Configuration**	**Minimum HA DB Configuration**
PowerEdge™ Nodes	R805, R900, R905 and M1000	1 Only	2
Memory	All valid Dell Server memory configurations	1 GB	1GB (per node)
Dell PowerVault Storage	MD3000 + MD1000 Expansion Module [1] [2]	N/A (For Internal Storage) 1 (For External Storage)	1
Dell /EMC FC Storage Array	CX Family of storage	N/A (For Internal Storage) 1 (For External Storage)	1
HBAs	Dell SAS 5/E (for attaching to PowerVault MD3000)	N/A (For Internal Storage) 2 ports (For External Storage)	2 ports (Per Node)
	LP 1150e LP10000(except PE 6950) LPe12000, LPe12002 LPe1105-M4 (only M600, M605)		
	QLE 2460, QLE2462 QLA2362(only on PE6850) QLE 2560, QLE2562 QME2472 (only M600, M605)		
Fibre Channel Switch	Brocade SW200E, SW4100, SW5000, SW4424[4] McData M4400, M4700	N/A	N/A (For Direct Attached) 2 (8 port switches for 2-4 nodes, and 16 port switches for 5 or above nodes) [1]
Ethernet Ports	Intel® or Broadcom® Gigabit NICs	1	2ports(Per Node)[3]
Ethernet Switches (For Private Interconnect)	Gigabit-only Switches		

RAID Controllers (Used for internal storage only)	PERC 5/i, PERC 6i SAS 6i/R	1	1(Per Node)
Internal Drive	All valid Dell Server internal storage configurations	73 GB	73GB (Per Node)
SQL Server Software & Licenses	Microsoft® SQL Server™ 2005 x64 SP2		
Operating System	Microsoft® Windows Server™ 2008 x64 Standard /Enterprise Edition		
Recommended Support Contract	Dell ProSupport		

[1] Note: MD1000 is not validated in this configuration as a stand-alone unit

Notes:

1. This assumes you don't need any other ports for other functions on the switch. Fibre Channel HBAs are recommended for other servers.
    a. Qlogic QME 2472 is only supported on PowerEdge M600 and M605.
    b. Emulex LP1105-M4 is only supported on PE M600 and M605.
    c. SAS (PowerVault MD3000 storage) configuration requires a minimum of two (single or dual channel) SAS 5/E HBAs for all supported servers.

2. PowerVault MD3000 with 2 PowerVault MD1000 expansion (max 45 disks supported). MD1000 is not supported as a standalone unit.

3. This assumes one NIC for public network and two teamed NICs for private network for both Fibre Channel and SAS configuration .Only one NIC is used for private network in a SAS (PowerVault MD3000 storage) configuration. It is recommended that two teamed NIC s are homogenous.

4. Brocade SW4424 is only supported on M1000e

Validated Servers

Dell PowerEdge™ Servers	Model	BIOS[7]	ESM - BMC Firmware[1]	NOTES
	PE R900	1.1.4	v1.79	
	PE M605	1.1.4	A04 (iDRAC)	Requires PE M1000e blade enclosure
	PE M600	1.2.2	A05 (iDRAC)	Requires PE M1000e blade enclosure
	PE R805	1.0.1	X09	
	PE R905	1.0.2	A00	

Internal Disks RAID	
PERC 5/i	Firmware version = 5.2.1-0067; Driver version = 2.14.0.64
PERC 6/i	Firmware version = 6.0.1-0080; Driver version = 2.14.0.64
Network Interconnect	
Intel Proset Gigabit Family Of Adapters (Single/Dual/Quad Ports)	Driver version = 10.0
Broadcom NetXtreme II	NDIS Driver version = 3.5.8 VBD Driver version = 3.5.10.1
Broadcom NetXtreme	Driver version = v10.43d
NIC Teaming (Not available yet for TOE NICs)[2]	
Intel NIC teaming	Driver version = 8.4.2.0
Intel Proset for Windows Device Manager	Version = 12.2.40.13
Broadcom	Broadcom Advanced Control Suite Version = 10.5.11.0 BASP Teaming driver version = 6.2.30.0
Host Bus Adapter (HBA)	
Dell SAS HBA	
Dell SAS 5/E	Firmware version = 00.10.51.00.06.12.05.00 (A04); Storport MiniPort Driver version 1.24.04.40
Qlogic HBAs	
QLA 2362 (PE 6850 only) QLE 2460; QLE2462	Bios=v1.24; Storport MiniPort Driver version = v9.1.4.15L (x64)
QME2462	Bios=v1.24; Storport Miniport Driver version = v9.1.4.15L (x64)
QME2472	Bios=v1.29; Storport MiniPort Driver version = v9.1.4.17L (x64)
QLE2560, QLE2562	Bios=v2.02; Storport MiniPort Driver version = v9.1.7.16 A4 (x64)
Emulex HBAs	
LP10000-E (Not supported in PE6950)	Firmware=v1.91.a5; Storport MiniPort Driver version = 1.30.9 (x64)
LPe1150-E	Firmware=v2.72a2; Storport MiniPort Driver version = 2.01a4 (x64)
LPe11002	Firmware=v2.72a2; Storport MiniPort Driver version = 2.01a4 (x64)
LPe1105-M4	Firmware=v2.70a5; Storport MiniPort Driver version = 2.01a4 (x64)
LPe12000, LPe12002	Firmware= ud100a9; Storport MiniPort Driver version = 2.01a4 (x64)
Microsoft Storport Hotfix for Qlogic, Emulex and Dell SAS HBAs	
KB 940467	storport.sys driver version = 5.2.3790.4133

Continued on Next Page

Fibre Channel Switches		
Brocade Fibre Channel Switches		
M4400, M4700	Firmware version = 09.02.01 1 or higher	
Brocade Fibre Channel Switches		
SW 200E, SW4100, SW5000	Firmware version = v5.3.0 or higher	
SW 4424	Firmware version = v5.3.0 or higher	
Direct Attached SAS Storage		
Dell PowerVault Storage	MD3000	Firmware version = A04
	MD1000	Firmware version = A04
Fibre Channel Storage		
Dell / EMC	CX3-20 (c/f); CX3-80 (Release 26 or higher) C	
Database Software		
SQLServer2005	Microsoft® SQL Server™ 2005 x64 SP2	
Operating system	Windows Server 2008 x64 Standard or Enterprise Edition	
Power Path for Windows x64	PowerPath 5.1.2(www.emc.com)	

[1] Minimum BIOS and ESM/BMC versions. For latest updates to BIOS go to http://support.dell.com

[2] Do not remove 'Dell 2 Port TOE Key (Part number: FG027)'. Use the Broadcom Advanced Control Suite software (BACS) to disable TOE on the Broadcom LOM. To do this open BACS and navigate to the Resource Configuration screen from the Resource Allocations tab. Then, uncheck the box for TOE.

Conclusion

Dell Solutions for SQL Server 2005 are designed to simplify operations, improve utilization and cost-effectively scale as your needs grow over time. This reference configuration white paper provides a blueprint for setting up a standalone or clustered SQL Server 2005 SP2 database as implemented on Windows Server 2008 on Dell PowerEdge servers and Dell PowerVault or Dell/EMC storage arrays.

The best practices described here are intended to help achieve optimal performance of SQL Server 2005 on Windows Server 2008. To learn more about deploying SQL Server 2005 on PowerEdge server and PowerVault and Dell/EMC storage, please visit www.dell.com/sql or contact your Dell representative for up to date information on Dell servers, storage and services for SQL Server 2005 solutions.

Index of Tables

Index of Figures

References

"Advanced Configuration Options in Microsoft SQL Server 2005", Ananda Sankaran and Shabana M., Dell PowerSolutions August 2006.
http://www.dell.com/downloads/global/power/ps3q06-20060270-Sankaran.pdf

"Storage Top 10 Best Practices", Microsoft TechNet, October 17, 2006.
http://www.microsoft.com/technet/prodtechnol/sql/bestpractice/storage-top-10.mspx

"Physical Database Storage Design", Kathy Lu, Lewis Bruck, Microsoft TechNet, February 12, 2007.
http://www.microsoft.com/technet/prodtechnol/sql/2005/physdbstor.mspx

"Clustered SQL Server do's, don'ts, and basic warnings", Revision 13.1, September 14, 2007.
http://support.microsoft.com/kb/254321

"Working with TempDB in SQL Server 2005", Wei Xiao, Matt Hink, Mirek, Sunil Agarwal, Microsoft TechNet, June 26, 2006.
http://www.microsoft.com/technet/prodtechnol/sql/2005/workingwithtempdb.mspx

"Using diskpar and diskpart to Align Partitions on Windows Basic and Dynamic Disks", EMC PowerLink September, 2005.
https://powerlink.emc.com/nsepn/webapps/btg548664833igtcuup4826/km/live1/en_US/Offering_Technical/White_Paper/H1445_Using_Diskpar_Align_Partition_ldv.pdf

"Disk performance may be slower than expected when you use multiple disks in Windows Server 2003, in Windows XP, and in Windows 2000", Revision 1.1, October 11, 2007.
http://support.microsoft.com/default.aspx?scid=kb;EN-US;929491

"SQL Server 2000 I/O Basics", Bob Dorr, TechNet, January 21, 2005.
http://www.microsoft.com/technet/prodtechnol/sql/2000/maintain/sqliobasics.mspx#EYCAE

"SQL Server I/O Basics, Chapter 2", July 19, 2006.
http://www.microsoft.com/technet/prodtechnol/sql/2005/iobasics.mspx

"Predeployment I/O Best Practices", Mike Ruthruff, June 5, 2007.
http://www.microsoft.com/technet/prodtechnol/sql/bestpractice/pdpliobp.mspx?pf=true

"Vista's New Partitioning Rules."
http://www.multibooters.co.uk/partitions.html

"Diskpart Alignment"http://sqlblogcasts.com/blogs/christian/archive/2007/10/09/disk-sector-alignment.aspx

"New in Failover Clusters in Windows Serevr 2008"
http://www.microsoft.com/windowsserver2008/en/us/high-availability.aspx

"IPv6 Benefits over Existing Technology "
http://technet.microsoft.com/en-us/library/bb726954.aspx

APPENDIX A: Configure Storage Connections

Figure 2 illustrates the interconnection of a PowerEdge server hosting SQL Server 2005 and an MD3000 storage system where the database resides. Using two dual-port SAS 5/E controllers in an Active/Active configuration, commands and data flow over both SAS links between the server and storage system. If either of the twoSAS 5/E controller or an MD3000 storage controller fails, operations continue using the remaining SAS 5/E – MD3000 storage controller pair.

One Server with Two HBAs - Dual Redundant Path

The following image shows a single server attached to an MD3000 with two SAS5/E HBAs installed. Each HBA is cabled to each individual controller.

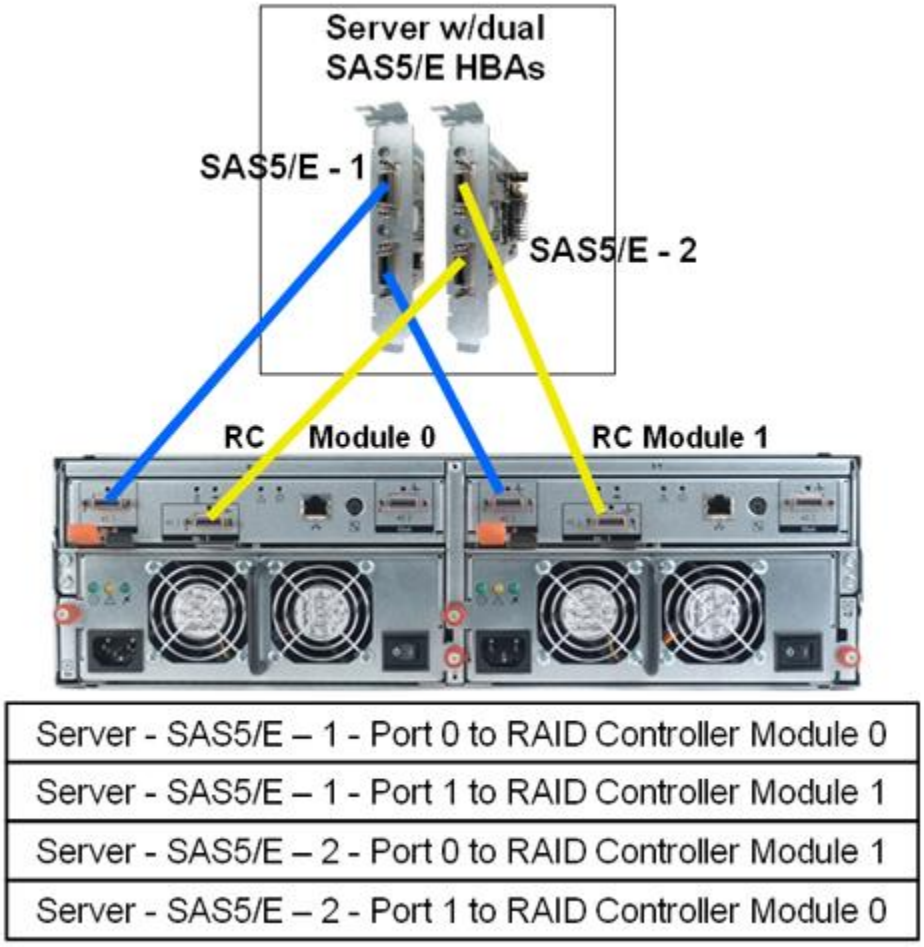


Figure 4 - Cabling a Direct Attached MD3000 Storage Array

Two Servers with Dual HBAs Each - Redundant Path

This configuration is supported for high availability clustering with the standard Edition of SQL Server 2005. The following image shows two servers with dual HBAs in each server. Each server has a redundant pathway to the controllers and the virtual disks, and is further protected against the loss of an individual HBA.

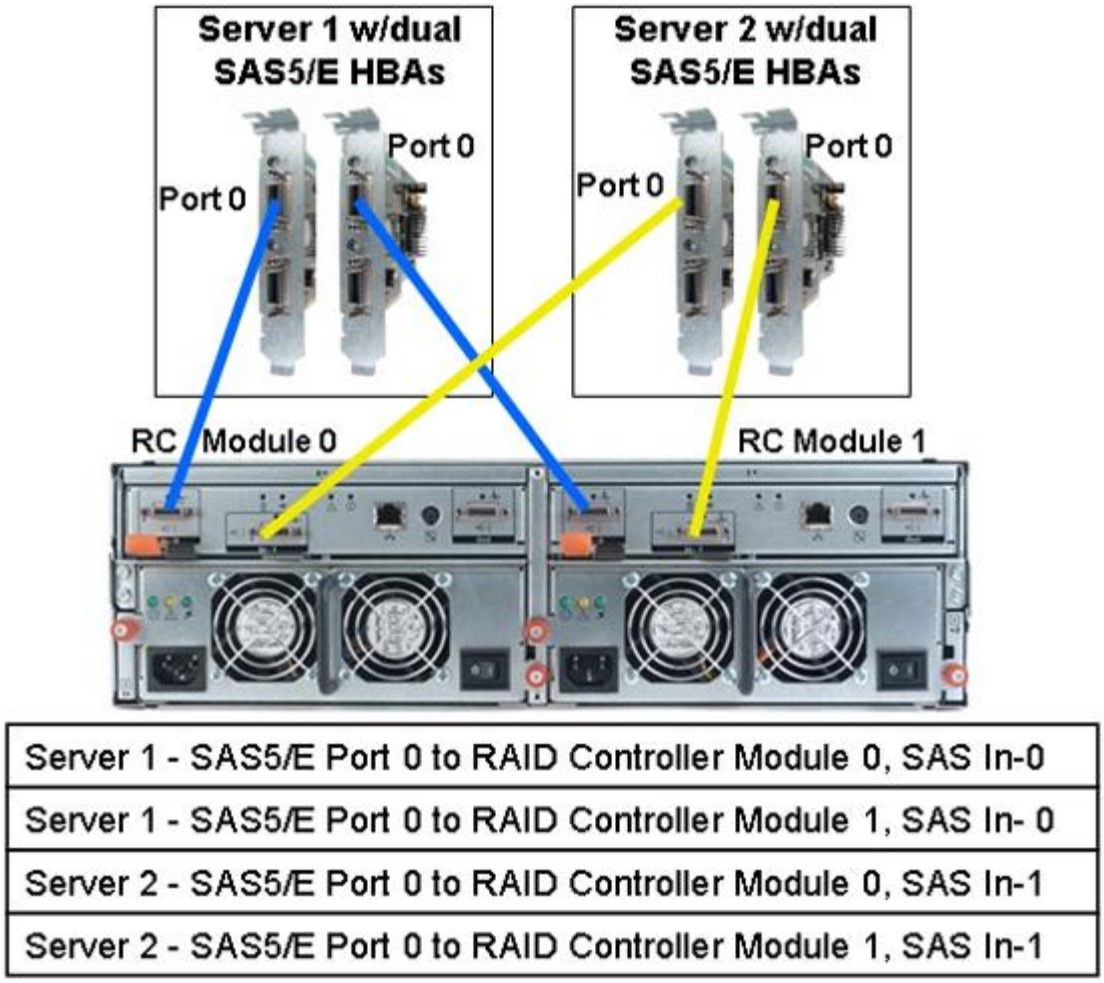


Figure 5 - Cabling two PowerEdge in cluster mode on MD3000 Storage Array

The MD3000 enclosure can support up to two cascaded MD1000 enclosures (see Figure 4), provided that the MDSM software, the EMM BIOS and firmware, and the cabling methodology are appropriate. The enclosures must be powered down for cascading.

WARNING:
Adding an MD1000 unit to an MD3000 storage chain is a data-destructive operation for any data on the MD1000s that was created with the MD1000 connected to a PERC5/E. The metadata format (DACstore) of configured drives in the MD3000 differs from the Disk Data Format used on configured drives attached to a PERC5/E. Data cannot be migrated by moving the MD1000 enclosure from a PERC5/E to an MD3000. The data must be backed up, confirmed, and restored to newly formed disk groups and virtual disks on the MD3000. No in-place migration is possible. This operation is not harmful to existing data on the MD3000.

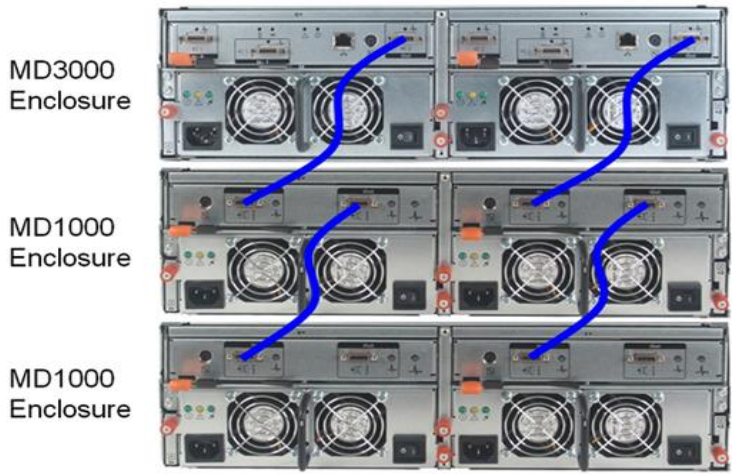


Figure 6 - MD3000 to MD1000 Enclosure Cascading

Configure Fibre Channel Storage Connections

Figure 5 illustrates the interconnection of a PowerEdge server hosting SQL Server 2005 and a CX3 storage system where the database resides. Using two FC HBAs in an Active/Active configuration, commands and data flow over both FC links between the server and storage system. If either a FC HBA or a CX3 storage processor fails, operations continue using the remaining FC HBA port – CX3 storage processor pair.

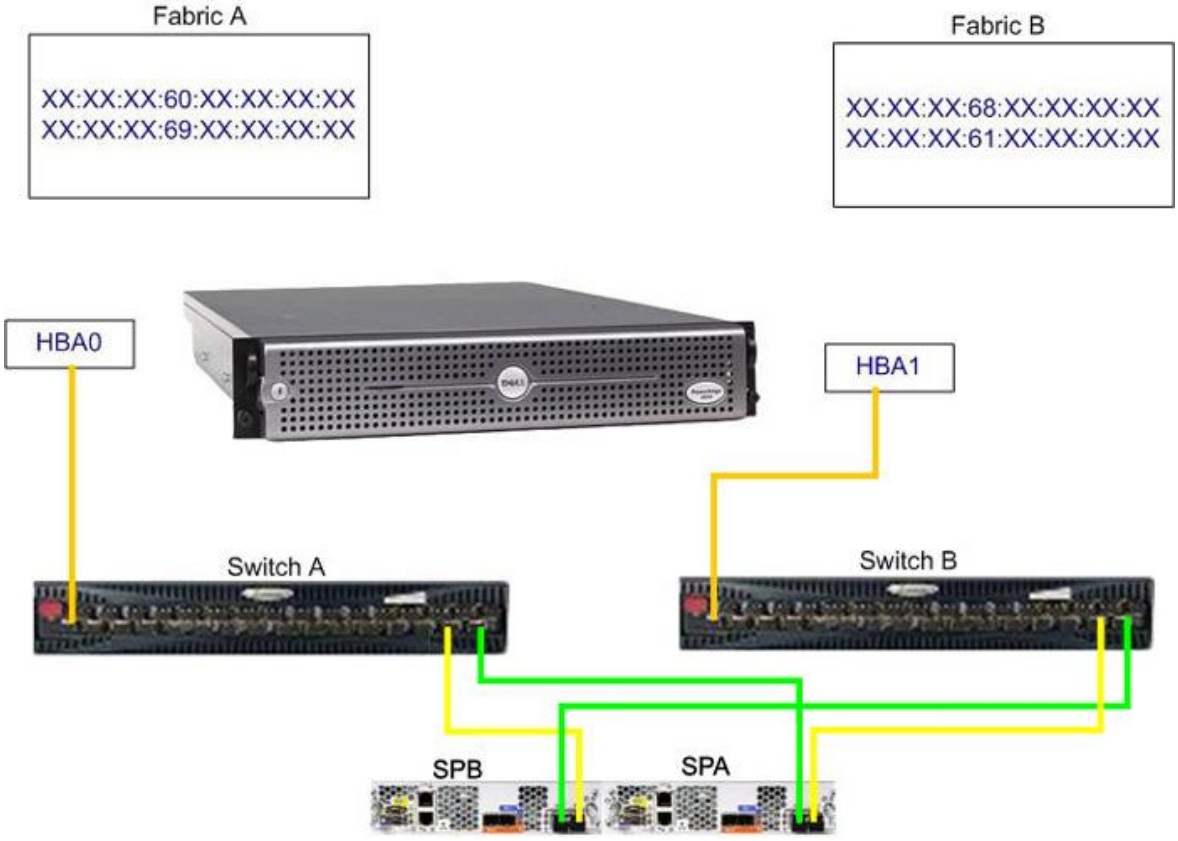


Figure 7- Cabling a Direct Attached CX3 Storage Array

NOTE: For two node clusters, follow the storage documentation to configure a storage group with multiple hosts and zoning information.

APPENDIX B: Moving TempDB Files to External Disks

1. Click Start and select Programs → Microsoft SQL Server 2005 and then click SQL Server Management Studio.

2. In the Connect to Server window:
   a. Select Database Engine from the Server type drop-down menu.
   b. Select the name of the server from the Server name drop-down menu.
   c. Select Windows Authentication from the Authentication drop-down menu.
   d. Click Connect to connect to the database.

3. In the Microsoft SQL Server Management Studio, click on the New Query in the tool bar.

4. In the new query window (right-hand side of the panel), type the following:

   USE Master
   GO
   ALTER DATABASE tempdb
   MODIFY FILE (NAME=tempdev, FILENAME='<Drive letter>:\TEMPDB_DATA_1\tempdb.mdf');
   GO
   ALTER DATABASE tempdb
   MODIFY FILE (name=templog, FILENAME='<Drive letter>:\TEMPDB_LOG1\templog.ldf');
   where <Drive letter> indicates the drive letters assigned to the external disks which were allocated for the temp database and log files .

5. Click Execute in the tool bar.

6. To add new files to tempdb, type the following in a new query window:
   USE Master
   GO
   ALTER DATABASE tempdb
   ADD FILE (NAME=TEMPDEV1, FILENAME='<location of the new temp file>', SIZE = <initial size of the file>, FILEGROWTH=10%);

   NOTE: To add additional files, repeat the previous entries starting with ADD FILE.

7. Click Execute in the tool bar.

APPENDIX C: Disk Partition Alignments

Unlike Windows2003, Windows Server 2008 automatically aligns the first partition to 1024KB to accommodate pretty much any RAID configuration. That means no need to align for any stripe/cache size up to 1024KB. Windows2008 by default leaves a 1MB offset at the start of a drive, which results in 2048 sectors not used by the partition, (on the new drives 1MB will be 256 sectors).
Extended partitions will start with these rules but end on a cylinder boundary. Logical partitions also adhere to these rules with the first logical having the 2048 sector offset from the beginning of the extended partition, the same as a first primary has to the start of the drive. This adds 2048 to the count so the first logical starts xxxxx4096 and ends xxxxx4095. Then each successive logical also has the 2048 offset from its own partition table, so the count adds again to xxxxx6144 to xxxxx6143, and etc for each further logical.

More Hidden Sectors Available
Hidden Sectors are just sectors on a drive that are not part of any partition. Under the old scheme of things, the first track on the drive is reserved for the MBR, which will give 63 hidden sectors. This causes subsequent data structures to become misaligned with respect to the track boundary. If a single I/O operation crosses a track boundary, it can consume added resources or cause additional work by the storage array leading to performance degradation. There are no gaps between primary partitions or between primary and extended, so no hidden sectors there. Each logical partition in an extended partition will have a reserved track immediately before it for its partition table, so there is another 63 hidden sectors before each logical partition. Now, the hidden sector figures become 2048 for the MBR and also 2048 before each logical hence no need for aligning partitions any more.