DELL Reference Architecture Guide

Deploying Microsoft® SQL Server® with Dell™ EqualLogic™ PS5000 Series Storage Arrays

A Dell Technical White Paper

Database Solutions Engineering

By Anthony Fernandez and Kevin Guinn
Dell Product Group
September 2008

CONTENTS

Abstract

This white paper provides a reference design, architectural overview, and configuration guidelines for deploying Microsoft® SQL Server® 2005 Standard Edition with SP2 with Microsoft Windows Server® 2003 R2 with SP2. This architecture is presented as an integrated solution that uses Dell PowerEdge™ R805 servers and Dell™ EqualLogic™ PS5000-series iSCSI storage arrays.

This technical report summarizes best practices and makes recommendations that provide a robust starting point to build a database solution that offers high performance, scalability, and reliability.

Introduction

This white paper focuses on providing a scalable design for enterprise level database solutions running on Microsoft SQL Server 2005 and Dell EqualLogic PS iSCSI Storage.

Dell PowerEdge servers are proven to be the best choice to deliver cost-effective performance for highly reliable databases[1]. Using a building-block approach, customers can build and scale their applications on Dell hardware at a fraction of the cost as their need grows. Dell EqualLogic arrays provide an ideal storage solution to implement databases that will provide seamless scalability and performance.

The Dell EqualLogic PS5000-series arrays offer the performance, features, ease of use, and scalability to simplify and consolidate complex environments. In addition, by using familiar IP technologies for the storage interconnection, the arrays offer an affordable easy-to-use solution.

This Reference Architecture Guide is intended to help IT professionals design and configure Microsoft SQL Server 2005 database solutions using Dell servers and EqualLogic storage that apply “best practices” derived from laboratory and real-world experiences.

The purpose of this whitepaper is to provide a detailed view of the benefits of a SQL Server deployment on Dell EqualLogic arrays. It provides a summary and references to all the best practices to configure Dell server and EqualLogic storage hardware and software for an optimal solution deployment.

This Dell Deployment Guidance whitepaper documents the results from engineering testing, captures details of the hardware and software tested, and provides tips and hints for deploying and operating the solution. Additional Dell Deployment Guidance whitepapers for SQL Server are delivered online at www.dell.com/sql.

Executive Summary

This paper provides a high-level overview of iSCSI technology, and clarifies its role within an enterprise IT infrastructure. The paper then examines the unique features of the Dell EqualLogic array and describes a platform for robust and scalable SQL Server deployments. This architecture is complemented by design considerations and best practices that improve the performance of the database solution. Specific guidance, tuning tips, and recommendations are provided for the storage, disk, and network subsystems. Similarly, it offers configuration guidelines and tuning tips for Windows Server and SQL Server.

[1] As of July 1, 2008, Dell holds the number one position and six of the top ten TPC-C Price/Performance results listed on the TPC website at http://tpc.org/tpcc/results/tpcc_price_perf_results.asp.

iSCSI Technology

Internet SCSI (iSCSI) is an industry standard developed to enable the transmission of SCSI block storage commands and data through IP data networks. Figure 1 illustrates how SCSI commands and data are encapsulated within a TCP/IP packet.

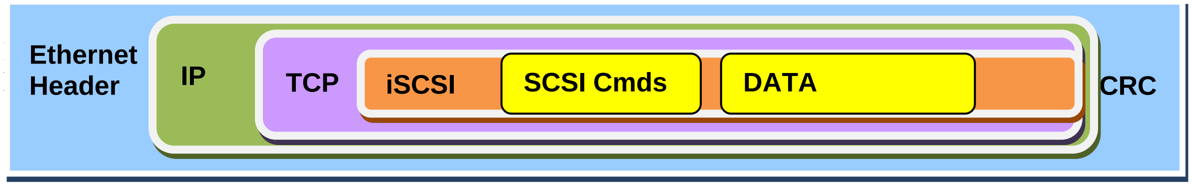

Figure 1: Encapsulation of Commands and Data in a Network Packet

Storage arrays based on iSCSI offer some key benefits over competing enterprise network storage devices. iSCSI makes use of Ethernet, which is a widely-deployed and well-understood infrastructure component for most IT organizations. Compared to Fibre Channel, the acquisition cost of adapters, switches, and other components is greatly reduced. As a result, iSCSI offers a reduced total cost of ownership compared to traditional FC storage area network (SAN) technologies.

Because it builds upon the SCSI protocol suite, iSCSI presents block devices to the operating system. This is in contrast to traditional network-attached storage (NAS) devices, which generally present file-level access to data shares. While it is possible to use a data share to house SQL Server data files, Microsoft recommends the use of block devices[2]. Storage provisioning operations, such as creating partitions and formatting volumes on block storage devices are performed using familiar OS tools. Finally, block devices are often faster and more efficient than file shares.

In an iSCSI SAN, devices are categorized as initiators or targets. An iSCSI initiator is generally configured on a host server, and employs one or more Ethernet ports that are used for iSCSI communication. The iSCSI initiator function can be implemented in hardware, as in the case of an iSCSI host bus adapter (HBA), but can also be implemented in software thereby further reducing the cost of an iSCSI solution. An iSCSI target is provided by a storage array and uses one or more Ethernet ports to present block storage devices to the initiator.

[2] “Description of support for network database files in SQL Server”: http://support.microsoft.com/kb/304261

Benefits of Dell EqualLogic Arrays in an Enterprise iSCSI SAN

Figure 2 illustrates an iSCSI SAN that is comprised of several different host servers accessing volumes provided by a set of EqualLogic storage arrays. Functionality that is associated with Fibre Channel SANs is available in this type of enterprise iSCSI deployment. The following figure demonstrates a storage consolidation deployment, where multiple hosts are granted access to separate volumes. The storage for each host server is centrally provisioned and managed from the pool of storage arrays. The balance of this paper will focus on deploying a database server in an iSCSI SAN environment.

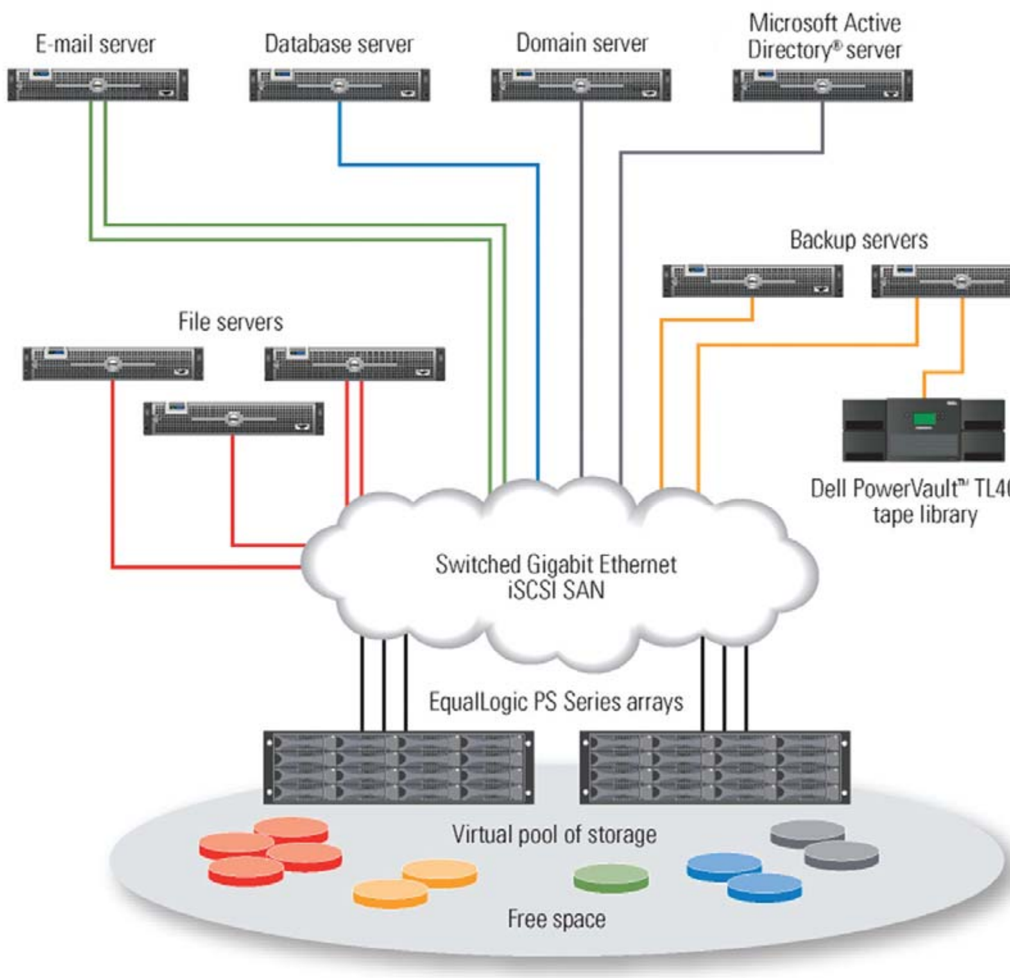

Figure 2: Enterprise iSCSI SAN with Storage Consolidation

Historically, one of the greatest challenges in database deployment has been provisioning storage in a manner that optimizes the performance characteristics for various critical functions, such as SQL Server data files, log files, and tempdb files. Regardless of the type of storage used, administrators had to be careful in determining RAID levels and creating volumes that balance the performance and capacity needs associated with each of these roles.

Consider the following example:

> A log volume is desired for a database and has an estimated current size requirement of 10 GB. To accommodate this need, storage administrators are asked to provision a 50-GB volume, which will allow for both expected growth and a factor of safety for any unforeseen circumstances. The storage system administrators are aware that performance is important, and want to use four dedicated disks, or spindles, in a RAID 1/0 configuration to meet the expected I/O demands of this volume. However, the company has begun to standardize on 146 GB disks, meaning that the total capacity of the RAID set would be significantly larger than the 50 GB required by the volume. In fact, the volume occupies about one fifth of the total capacity of the four-spindle RAID set, but the four disks were required to satisfy the latency, data protection, and other characteristics required by the database.

In a SAN environment, scenarios such as this are common, and the challenge of efficiently and effectively assigning storage that strikes a balance between the performance and capacity demands of several disparate application servers becomes even more difficult.

EqualLogic PS5000-series storage arrays include storage virtualization technology that has been designed to reduce this type of complexity. To better understand how this is possible, it is helpful to be familiar with some of the terminology used to describe these arrays and their functions:

- **Member** – a single PS5000-series array is known as a member
- **Group** – a set of one or more members that can be centrally managed
- **Pool** – a RAID that can consist of the disks from one or more members
- **Volume** – a LUN or virtual disk that represents a subset of the capacity of a pool
- **Non-Tiered Configuration** – a group that contains a single pool
- **Tiered Configuration** – a group that contains multiple pools

When a member is initialized, it can be configured with RAID 1/0, RAID 5, or RAID 5/0. The EqualLogic PS5000-series arrays provide automatic load balancing of volumes among members that participate in the same pool; this can improve the aggregate performance of all of the configured volumes without administrative input.

Solution Overview

This section describes the components required to deploy a highly scalable database solution with the Dell EqualLogic PS5000-series storage arrays. Based on the results of engineering experimentation, the reference architecture in Figure 3 has been developed. This Dell Solution for Windows SQL Server 2005 is comprised of the following components:

- Dell PowerEdge R805 server running Windows Server 2003 and SQL Server 2005
- Two Dell PowerConnect™ 5400-series Ethernet switches for iSCSI traffic
- Dell EqualLogic PS5000-series iSCSI storage arrays

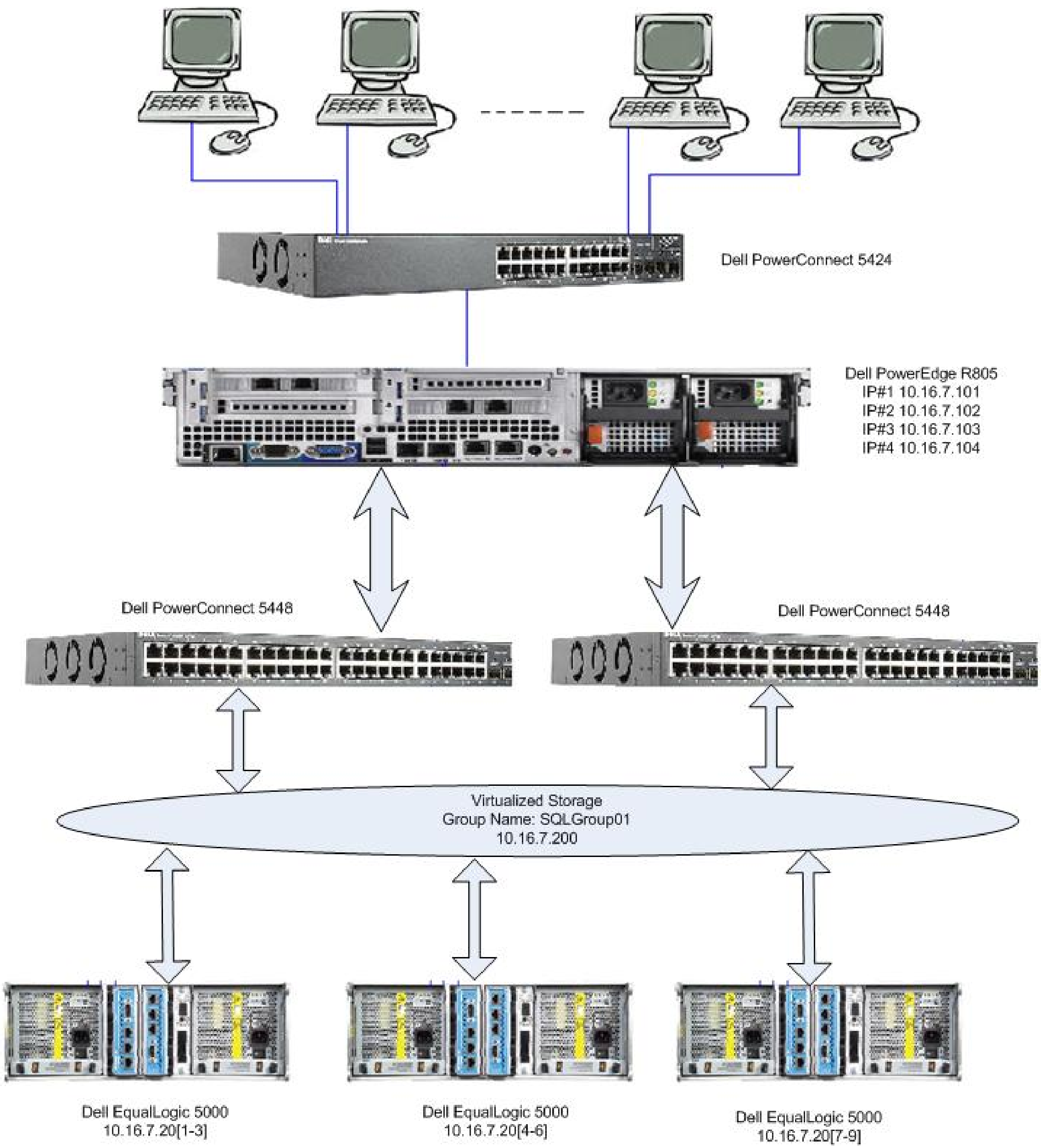

Figure 3: Key Components of this Reference Architecture

Dell PowerEdge R805 Server with Windows Server 2003 and Windows SQL Server 2005

Compared to servers with two processor sockets that occupy two rack units (2U), this server offers more RAM, more integrated network connectivity, and more expansion slots. These traits, which make this server well suited for typical database workloads, are listed in Table 1.

Table 1: PowerEdge R805 Features

<table>
<tr><th>Feature</th><th>PowerEdge R805 as Tested</th><th>PowerEdge R805 Maximum Configuration</th></tr>
<tr><td>Processor</td><td>Two Quad-Core AMD® Opteron™ 2346HE, 55W, 1.8 GHz</td><td>Two Quad-Core AMD Opteron 2360SE, 105W, 2.5 GHz</td></tr>
<tr><td>Memory</td><td>16 GB</td><td>16 total RAM slots (8 per socket), offering up to 128 GB of RAM</td></tr>
<tr><td>Front Side Bus</td><td colspan="2">HyperTransport at 1000MHz</td></tr>
<tr><td>I/O Slots</td><td colspan="2">2 PCIe Risers
3 – PCIe x8 (4GB/s) slots
1 - PCIe x4 (2GB/s) slot

Details:
Riser 1: slot 1 is PCIe x8, slot 2 is PCIe x4
Riser 2: slot 3 and slot 4 are both PCIe x8

Adding expansion NICs to PCIe x8 slots on separate buses is a best practice. For example, use slots 1 and 3.</td></tr>
<tr><td>Internal Storage</td><td>Two 2.5" SAS drives</td><td>Two 2.5” drives, either SATA or SAS</td></tr>
<tr><td>Internal RAID Controller</td><td>PERC6i, providing RAID 1 for the OS with cache and battery backup</td><td>PERC6i offers RAID 0 or RAID 1 for internal hard drives. SAS6iR also available.</td></tr>
<tr><td>Network Adapters</td><td>Integrated dual-port Broadcom 5708 Gigabit Ethernet with optional TOE license key - PCIe x4 (2GB/s)

Additional Integrated dual-port Broadcom 5708 Gigabit Ethernet - PCIe x4 (2GB/s)

Expansion PCIe GigE NICs for additional connectivity</td><td>TOE is available for all four integrated network ports. Additional expansion GigE and 10 GigE PCIe adapters may be added.</td></tr>
</table>

During testing, up to six GigE ports were used to provide iSCSI connectivity, depending on the target workload. With OLTP workloads, the network port usage remained low, even with only two iSCSI network ports on the host. With OLAP workloads, the bandwidth provided by two host iSCSI ports was insufficient. Adding additional ports helped alleviate the bottleneck. When multiple members were configured, increasing the number of host ports available helped balance the demands of the OLAP workload.

Additional network ports were also needed to provide connections between the SQL Server host and client systems or application servers that access the database. Thus, having the four on-board ports, and the ability to add additional port via PCIe expansion cards proved to be valuable in this Dell Solution for Windows SQL Server.

Dell PowerConnect 5400-series Ethernet Switches

The Dell PowerConnect 5424 and 5448 are manageable switches that provide 24 and 48 GigE ports, respectively. These switches support jumbo frames, link aggregation, VLANs, and other features that are desirable for use in an iSCSI SAN.

By default, these switches are configured to use Quality of Service (QoS) to prioritize iSCSI traffic. However, when used with the EqualLogic PS5000-series arrays, this should be disabled because the traffic between members that enables the automatic load balancing functionality does not use the iSCSI protocol, and can therefore inadvertently be delayed or blocked. Table 2 details the configuration settings that are optimal for use in this Dell Solution for SQL Server.

Table 2: PowerConnect 5400-series Configuration Options for EqualLogic iSCSI Networks

Feature	Setting
iSCSI Optimization (QoS)	Disabled
Jumbo Frames	Enabled[3]
Flow Control	Enabled for all ports
Storm Control	Enabled for Broadcast and Multicast on all ports
Spanning Tree Protocol	Global Setting: Rapid STP Enabled Port Settings: • Disabled for any iSCSI initiator and target ports • Enabled for inter-switch (trunk) ports
Link Aggregation	Enabled for inter-switch (trunk) ports in multi-switch networks
VLANs	Use a VLAN to isolate iSCSI traffic if a physically separate iSCSI network is not feasible.

To provide redundant connection paths and fault resilience in the event of a switch failure, a two-switch iSCSI network topology is recommended. In such a configuration, there should be multiple ports with link aggregation providing the inter-switch, or trunk, connection. Further, each EqualLogic controller should have two connections to one switch and one connection to the other switch. The active and stand-by controller in the same member should have dual connections to alternate switches. This topology is illustrated for a single PS5000 array in Figure 4. In addition, at least two network adapter ports on each host server should be configured for iSCSI, with one of these ports connected to each of the two switches.

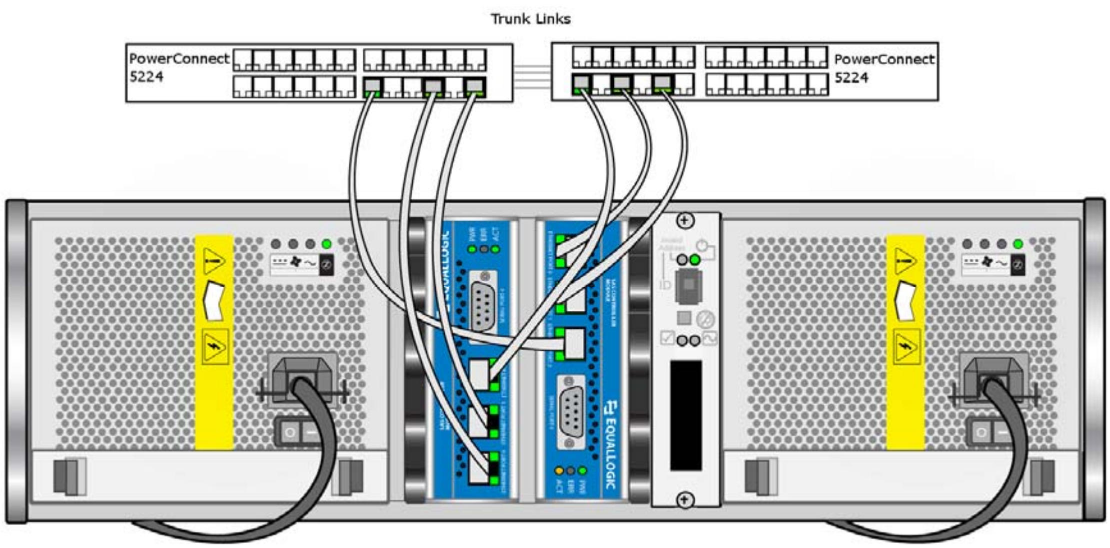

Figure 4: PS5000-series Array in a Two-Switch Redundant iSCSI Topology

[3] To benefit from this setting, also configure initiator ports for a maximum data size of 9000 bytes (or maximum frame size of 9018 bytes, including the TCP header).

Dell EqualLogic PS5000-series iSCSI Storage Arrays

Table 3 highlights the key specifications and features of the Dell EqualLogic PS5000-series storage arrays. Additional information about how these features are applied to this Dell Solution for SQL Server is presented in the remainder of this paper.

Table 3: EqualLogic PS5000-series Features and Specifications

Feature or Specification	Description
Peer Storage Architecture	♦ Additional members can be added online, increasing: o Overall storage capacity. o Number of iSCSI connections. ♦ Central management of a group. ♦ Flexible configurations: pools and tiers. ♦ Automatic volume management and load balancing among members. ♦ Virtualized storage
Storage Management	♦ Simplified deployment and provisioning o Volume management o Group management ♦ Role-based administration ♦ E-mail Home capability
Industry Standards	♦ Gigabit Ethernet – three Copper ports per controller. ♦ iSCSI protocol – supports a wide range of initiators. ♦ Serial Attached SCSI or Serial ATA hard-disk drives.
iSCSI Host Connectivity	Up to 512 hosts can be connected using standards-compliant initiators.
Disk Capacity and Scalability	♦ Sixteen hard-disk drives per member, providing between 2.3 TB and 16 TB, depending on the model in the PS5000-series that has been selected. ♦ Online addition of members to expand the capacity of the system. ♦ Multi-Path I/O (MPIO) ♦ Thin Provisioning for planned growth and non-disruptive online volume expansion.
Additional Included Features	♦ Volume Consistency Groups ♦ Volume Clones and Volume Snapshots (Writable) ♦ Auto-Snapshot Manager and Smart Copy, including a Volume ShadowCopy Service provider and intelligent integration with SQL Server. ♦ Auto-Replication for disaster recovery deployments

Architectural Design Considerations

This reference architecture provides recommendations and guidelines for implementing a robust and scalable database solution. These guidelines implement best practices based on advice from Microsoft and EqualLogic, and on experiments performed in the Dell engineering labs.

The following sections explain these recommendations in the context of the various subsystems that are integral to the Dell Solution for SQL Server.

Storage Considerations

With every new technology, customers need to have a clear view of their current infrastructure limitations in order to develop a suitable solution that meets their specific needs. For example, the key questions for a storage subsystem may be based on considerations such as storage consolidation, planning for future growth, or designing to alleviate existing performance limitations. Table 4 outlines several key considerations and explains how this reference architecture addresses the requirements.

Table 4: Storage Subsystem Design Considerations

Storage Subsystem Objectives	
Reliability and Availability (Redundancy)	♦ Dual controllers: one active and one stand-by for PS5000-series arrays. ♦ 2GB battery-backed cache for up to 72 hours of data protection. ♦ Redundant power supplies and cooling fans. ♦ Hot Spare disks.
Performance	♦ 16 hard-disk drives per array -- 10k rpm SAS (PS5000X) or 15k SAS (PS5000XV) are recommended for databases. ♦ Three GigE Ports per Controller with Multi-Path Input/Output (MPIO) support. ♦ Automatic Load Balancing among members.
Scalability	♦ Additional members can be added to a pool to provide additional storage capacity and connectivity. ♦ Tiered configurations allow the RAID level to be matched with intended volume workloads.
Data Protection	♦ RAID 5, 1/0, or 5/0 configured on each member. ♦ Volume Snapshots. ♦ Volume Cloning. ♦ Auto-Snapshot Manager (ASM) for easy integration with SQL Server, Volume ShadowCopy Service (VSS), and enterprise backup strategies[4]. ♦ Automatic Replication for Disaster Recovery.
Ease of Deployment and Management	♦ Simplified initial configuration of a member. ♦ Centralized management of a group of members. ♦ Easy to add members to a pool or a group. ♦ Automatic Load Balancing of volumes.
Service Level Commitments	♦ Multiple pools in tiered configurations allow volumes to be provisioned with appropriate RAID levels based on their intended usage. ♦ Additional members can be added to increase the capacity of a pool. ♦ Thin Provisioning enables online volume expansion.

[4] For more information about ASM and VSS integration, refer to "SQL Server Advanced Protection and Fast Recovery with EqualLogic Auto-Snapshot Manager" at https://www.equallogic.com/uploadedFiles/Resources/White_Papers/WP911_SQL-Server_Advanced-Protection.pdf

Disk and RAID Considerations

Choosing hard-disk drives of the correct type, interface, rotational speed, and capacity and determining the appropriate RAID level and number of disks for volumes created on those spindles are challenging tasks for IT administrators. Several key factors that influence these decisions are outlined in Table 5.

Table 5: Disk and RAID Design Considerations

Disk and RAID Level Considerations	
Disk Type / Interface	♦ The EqualLogic PS5000-series arrays offer models with either SATA or SAS drives. ♦ SAS drives are recommended for typical database workloads. ♦ SATA drives are appropriate and offer a cost-effective option for volume clones, snapshots, backup enablement, and near-line storage in a tiered configuration.
Disk Rotational Speed	♦ Higher rpm disks offer faster access times and better transfer times than disks with slower rotational speeds. ♦ Lower rpm disks generally provide additional per-spindle capacity vs. disks with faster rotational speeds. ♦ 15k disks are recommended for Transaction Processing (OLTP) environments where small random I/Os make their performance characteristics desirable. ♦ Higher-capacity drives may be well-suited for workloads with large sequential I/O patterns, such as Analytical Processing (OLAP), Data Warehouses, or Decision Support Systems (DSS).
RAID Levels for Database Volumes	Non-tiered configuration (single member or multi-member pool): ♦ OLTP: RAID 1/0 ♦ OLAP (higher capacity): RAID 5 ♦ OLAP (higher performance): RAID 5/0 Tiered configuration (multiple pools): ♦ OLAP data volumes: RAID 5 or 5/0 ♦ OLAP log volumes: RAID 1/0 ♦ SQL Server tempdb volumes: RAID 1/0 ♦ OLTP data or log volumes: RAID 1/0
Hot Spare Strategy	♦ Hot spare disks allow automatic re-builds of data to reestablish RAID data protection in the event of a disk failure. ♦ Spares are recommended for all storage arrays that provide production enterprise applications. ♦ In test and development environments, the extra capacity made available may outweigh the risk of data loss that is associated with a second failure occurring within a degraded RAID set.
Database Workload Performance and Scalability	♦ Add members to a pool to reduce latency for OLTP databases. ♦ Employ tiered configurations for large OLAP databases, or for SQL Server hosts that handle mixed workloads.

Client and iSCSI Network Considerations

A SQL Server host in an iSCSI SAN is connected to several different networks. To simplify deployment and management, it is desirable to separate iSCSI storage networks from client-access networks. Having at least two network interface ports for each of these functions, and dividing ports that serve the same function between separate adapters can greatly reduce the risk that a single-component failure is disruptive to the operation of this Dell Solution for SQL Server. Several key factors for designing iSCSI and client networks are provided in Table 6.

Table 6: Client and Storage Network Design Considerations

Client and Storage Network Considerations	
Isolated Client and Storage Networks	♦ Where possible, use a separate physically-isolated network for iSCSI traffic that is independent from the network on which the SQL Server host handles client or application requests. ♦ Otherwise, employ VLANs to isolate the network traffic based on these functional roles.
Redundant Storage Network Topology	Use two interconnected switches (as illustrated in Figure 4) to provide network resilience in the event of a switch or network port failure.
MPIO for iSCSI Networks	♦ Configure at least two host network ports for iSCSI connectivity. ♦ Use MPIO to provide load balancing and fault resilience among multiple host iSCSI ports. ♦ Install the EqualLogic MPIO Device Specific Module (DSM) to improve load balancing and routing of iSCSI traffic between the host and the members.
Host iSCSI Network Configuration	♦ Configure iSCSI host network interfaces to complement the switch configuration settings listed in Table 2. o Enable Jumbo Frames, set maximum data size to 9000 bytes (or maximum frame size to 9018 bytes). o Enable Flow Control, set to generate and respond to messages. ♦ Consider increasing the number of Receive Descriptors and/or Transmit Descriptors to 2048, particularly for OLTP workloads. ♦ Configure the iSCSI link down (35s) and Windows disk (60s) timeout values by running the `EqlSetupUtil.exe`[5] utility that is installed by the EqualLogic Host Integration Toolkit (HIT). ♦ Select "Maximize data throughput for network applications" from the Windows Server 2003 Network Connection Control Panel. ♦ Place expansion network adapters on separate PCIe busses to maximize throughput (see Table 1).
Redundant Client Network Connectivity	Use more than one adapter port and configure NIC Teaming (or Bonding) to provide redundancy for the client/application network.
Identify Potential I/O Bottlenecks based on Workload	♦ For OLAP workloads, add additional iSCSI host ports to scale the host I/O capability; these jobs require a lot of bandwidth, and can rapidly saturate iSCSI GigE links. ♦ For OLTP workloads, it is unlikely that the iSCSI GigE networks will become saturated; the number of spindles available to handle the required throughput is generally the principal concern.

[5] After the HIT is installed, this utility will be located in the `Program Files\EqualLogic\bin` directory. Refer to the release notes for more information about the syntax for the utility.

Server and Operating System Considerations

There are several hardware and software configuration guidelines that can improve the performance of a SQL Server solution. Table 7 describes the principal design factors that impact the server hardware and operating system.

Table 7: Server Hardware and OS Configuration Design Considerations

Server Hardware and Operating System Configuration Considerations	
OS Volume	♦ Use a RAID volume to offer protection for the OS. Example: RAID 1 configured on the two internal hard-disk drives. ♦ Where practical, create a separate volume for the OS page file. ♦ If you are using Boot-to-SAN (iSCSI Boot), consider using the internal drives for the page file.
iSCSI Network Ports	♦ Use at least two host network ports for iSCSI traffic. ♦ Configure MPIO for all iSCSI initiators on the host. **NOTE:** Do **not** configure NIC Teaming or Bonding for iSCSI. ♦ For redundancy and availability, use a multi-switch topology (see Figure 4). ♦ Be aware of potential bandwidth bottlenecks: o OLAP workloads feature large sequential I/Os which can saturate a small number of GigE links. o Plan for approximately 100 MB/s of iSCSI bandwidth per GigE port – for example, if your workload generates 600 MB/s, you should configure six iSCSI ports on the SQL Server. ♦ The number of active storage ports should meet or exceed the number of required host ports – in the example above, two member arrays would provide six active storage ports, and therefore would be able to handle the workload.
Server RAM	♦ Use the SQL Server Advisor on www.dell.com/sql to help estimate memory requirements based on your workload. ♦ If more than 4GB of RAM is required, consider 64-bit extended (x64) versions of Windows Server 2003 and SQL Server 2005. ♦ Monitor server available memory using the “Memory: Available Mbytes” counter in perfmon. o Goal: maintain at least 500MB of available memory to reduce the potential for excessive OS page swapping and other factors that can degrade performance. o If actual numbers are below this threshold, reduce the SQL Server Max Memory (default is “all available”) and consider adding additional RAM.
Server Processors	♦ Use the SQL Server Advisor on www.dell.com/sql to help estimate the CPU requirements based on your workload. ♦ Consider your workload: o OLTP rule-of-thumb: invest in I/O throughput first, engineering testing has confirmed that CPU is rarely the bottleneck for transactional workloads with small I/Os and basic query structures.
SQL Server Volumes	♦ Maintain separate volumes for data, logs, and tempdb. ♦ Partition alignment: set the offset to 64 KB for SQL data, log, and tempdb volumes on PS5000-series arrays. ♦ Format volumes with NTFS, and set the allocation unit size to 64 KB.
Configure OS Parameters	♦ Processor Scheduling: “Best Performance for Background Services” SQL Server 2005 runs as a service. ♦ Memory Usage: “Best Performance for Programs” SQL Server 2005 controls its own memory pool and maintains its own cache.

SQL Server Considerations

After configuring the operating system and the various solution sub-systems, there are some SQL Server configuration parameters and design guidelines that can help improve the performance of this Dell Solution for SQL Server. Table 8 outlines some of the key considerations.

Table 8: SQL Server Design Considerations

SQL Server Considerations	
Data, Log, and TempDB Volumes	♦ Create separate volumes for SQL Server data files, log files, and TempDB data and log files. ♦ Plan for storage scalability based on your workload: o OLTP tends to be throughput-bound. Adding an additional member allows additional spindles to be used by the volume, which will relieve this bottleneck. See Figure 6 for an example. o OLAP tends to be bandwidth-bound. Adding an additional member provides more iSCSI ports on the storage array. When combined with additional iSCSI ports on the host, this can mitigate the bottleneck.
File Groups and Files	Apply best practices for the number of volumes, file groups, and files as documented in the SQL Server 2005 Deployment Guidance and Design Consideration white papers on www.dell.com/sql.
SQL Server Configuration Options	♦ Apply best practices for SQL Server Performance optimization as documented in the SQL Server 2005 Deployment Guidance and Design Consideration white papers on www.dell.com/sql. ♦ If your configuration includes SQL Server 2005 Enterprise Edition, configure SQL Server to "keep pages in memory." To configure this, the account under which the Microsoft SQL Server service runs must be granted rights to "Lock pages in memory". See http://technet.microsoft.com/en-us/library/ms190730.aspx for detailed instructions. This privilege is not automatically granted to local administrators.

Real-world SQL Server deployments often serve mixed workloads. In such cases, a tiered configuration on the EqualLogic PS5000-series arrays, as illustrated in FIGX, can meet the needs of the various components of the composite workload. In this example, a separate SQL Server file group has been configured to house specific databases that are often used for reporting and analysis. These particular databases experience large sequential I/O and are therefore well suited for RAID 5 or RAID 5/0 volumes; the balance of the SQL Server system primarily serves high-volume OLTP transactions which are better candidates for RAID 1/0 volumes.

The tiered configuration allows multiple RAID types to be used within the solution, and administrators can create volumes in whichever pool is the best fit. Further, more than one volume can be configured in the pool, and all of the volumes will benefit from the automatic load balancing across the members in that pool.

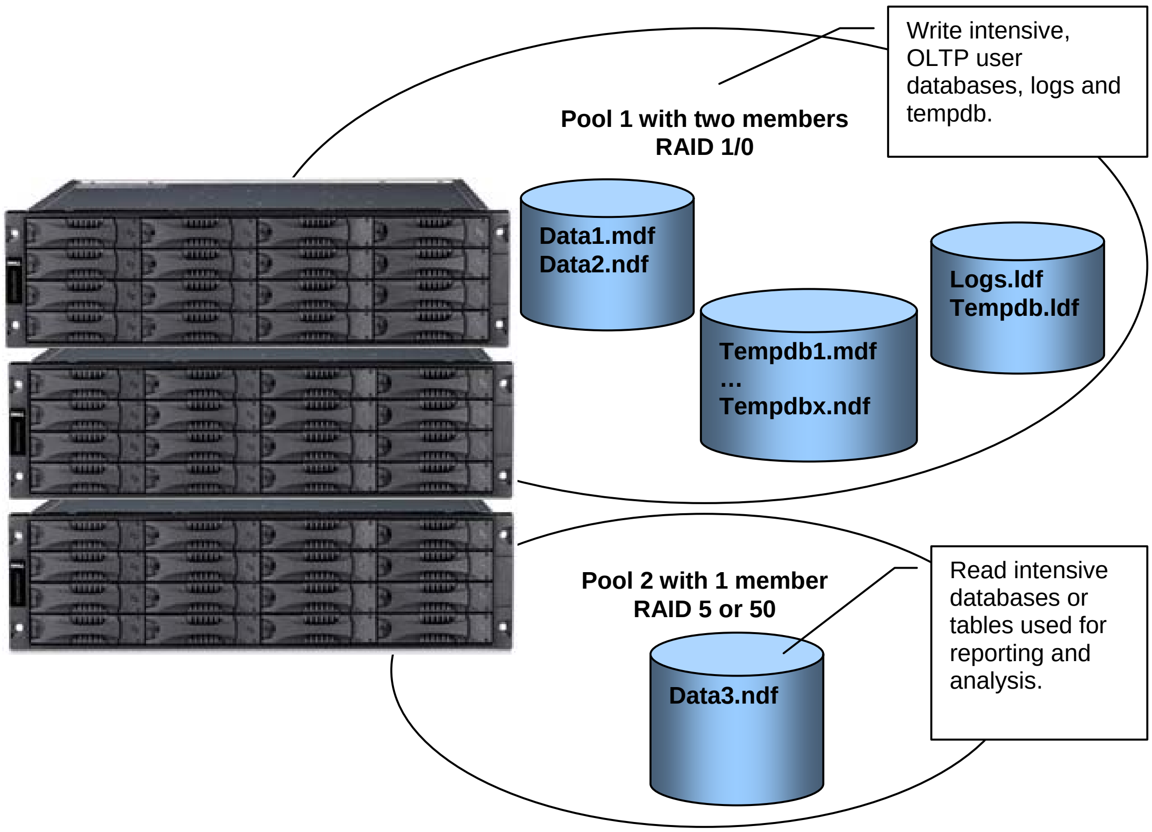

Figure 5: Three-Member Tiered Configuration of EqualLogic PS5000-series Arrays

EqualLogic PS5000-series OLTP Scalability Test

To verify the scalability of the EqualLogic PS5000-series storage arrays, lab experiments were run with a 100 GB database that follows the TPC-C schema for OLTP. The server configuration was held constant (see Table 1), with four iSCSI network ports throughout the course of the experiments. Tests were conducted in which a varying number of simulated users (from 100 to 4000) performed tasks and queries consistent with a TPC-C benchmark. For each test, the average query response time was measured, and then the number of users was increased by 100 in an iterative fashion.

After completing the iterations with a single member, a second member was added. This member was configured with the same RAID level, and was added to the same pool to allow the volumes to be distributed across additional spindles. The test iterations were then repeated with the two-member configuration. Finally, a third member was added to the pool, and the test iterations were completed again.

The engineering hypothesis was that with a single member hosting the data, log, and tempdb files, it was unlikely to provide the throughput necessary to handle large user loads; the average query response time was expected to rapidly exceed the threshold. This assumption was based on the fact that with a single member, only fourteen spindles were being used for the SQL Server data, log, and tempdb volumes. Further, there was an expectation that adding additional members – therefore increasing the number of spindles available for the SQL Server volumes – should alleviate this bottleneck. The average query response time for test runs with larger numbers of users were expected to improve as additional members were added.

Figure 6 shows the results of this lab experiment. With a single member, the threshold value for average query response times was exceeded with approximately 700 users. After adding a second member, approximately 1800 users could be added while keeping the average query response time within the acceptable range. With three members configured in the storage pool, the response time remained in this range beyond 4000 users.

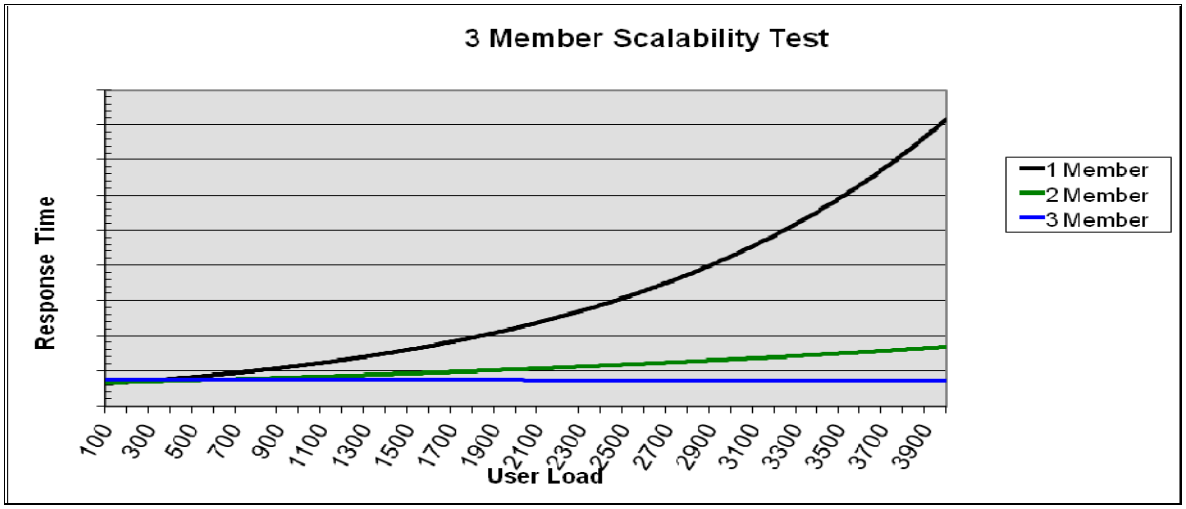

Figure 6: Improved OLTP Response Time with Additional Members

This OLTP scalability experiment validated the hypothesis, and demonstrated that the automatic load balancing capability of the EqualLogic arrays was able to distribute the OLTP workload among multiple members efficiently. As more spindles were made available, a greater number of users were able to access the database while maintaining desirable performance (as measured by average query response times). This functionality makes it possible to scale the workload without the need to reconfigure the SQL Server volumes. As additional member arrays were added to the storage pool, the EqualLogic Peer Storage architecture recognized the constraints and automatically distributed the volumes across a greater number of spindles, thereby allowing the performance of the system to improve without requiring additional administrative intervention.

Conclusion

The storage subsystem of a SQL Server solution experiences different potential bottlenecks based on the how the database server is used. With minimal configuration changes, the reference architecture that has been presented in this paper can be easily tuned to support common database workloads.

With 64-bit extended (x64) editions of the Windows Server operating system and Windows SQL Server 2005, database servers can efficiently make use of more than 4 GB of system RAM. The PowerEdge R805 server provides two processor sockets, and offers support for up to 128 GB of RAM. Because of the available expansion slots, this server offers many GigE network ports to be configured; these ports are used either for iSCSI host connectivity, or for communication with clients and/or application servers. Regardless of the workload, MPIO allows load balancing and fault tolerance for the iSCSI connections, while NIC Bonding or Teaming provides these functions for the client network. For OLTP workloads, two iSCSI host networks will often provide sufficient throughput. For OLAP workloads, it is beneficial to allocate additional iSCSI host ports to increase the aggregate available bandwidth of these connections.

The Dell EqualLogic PS5000-series storage arrays provide a unique Peer Storage Architecture that enables multiple members to work together and offer good scalability for a database solution. For high-volume OLTP workloads, additional members can be added to a storage pool; this automatically allows the throughput of the overall solution to improve, as the automatic load balancing function of the arrays can distribute the configured volumes so that they benefit from the added spindles. For OLAP workloads, additional members increase the number of iSCSI ports on the storage subsystem, which similarly increases the overall bandwidth of the system.

When these components are combined and the best practices listed in this paper are applied, the resulting Dell Solution for SQL Server offers good performance and excellent scalability across a variety of common database workloads. Careful considerations of design criteria, backed by practical laboratory evaluations, helps ensure that this reference architecture effectively meets the needs of a variety of enterprise-scale database deployments.

Solution Components

The following table summarizes the components, drivers and firmware used at the time of testing for this reference configuration. For the latest drivers please visit http://support.dell.com

Table 9: Solution Components

Solution Components			
PowerEdge Servers	**Model**	**BIOS**	**ESM/BMC Firmware**
	R805	1.0.2	2.00
Internal Disks RAID			
PERC 6/i	Firmware version = 6.0.3-0002; Driver version = 2.14.0.64		
SAS Backplane Firmware	Firmware version = 1.0.6		
Additional Network Adapters			
Intel® vPRO™ 1000PT Single/Dual Port	Driver version = 9.9.12.0		
Intel vPRO 1000PT Quad Port	Driver version = 10.0.25.0		
Integrated Network Adapters			
Broadcom NetXtreme II 5708 Gigabit Ethernet Driver	Driver version = 12.1.0 (A01), VBD = 4.1.5 LOM Firmware = 4.4.1 (A01)		
Ethernet Teaming Driver[6]	Driver version = 6.2.33		
iSCSI SAN Switches			
Dell PowerConnect 5424 / 5448 Gigabit Ethernet Switches	Firmware = v1.0.0.35		
iSCSI Storage			
Dell EqualLogic iSCSI Storage	PS5000XV; Firmware = V4.0.1		
Software			
RDBMS	Microsoft Windows SQL 2005 x64 Standard Edition SP2		
Operating systems	Microsoft Windows Server 2003 R2 with SP2 Standard x64 Edition		
iSCSI Initiator	Microsoft iSCSI Initiator 2.0.7		
EqualLogic Host Integration Tool Kit	3.1.1		

[6] NIC Teamming is not recommended for iSCSI network. Use Teaming for high availability only on the public LAN interfaces.

References

"Deploying Microsoft® SQL Server® 2005 Standard x64 Edition SP2 with Microsoft Windows Server™ 2003 Standard x64 Edition with SP2 on Dell™ PowerEdge™ Servers and Dell™ PowerVault™ Storage"
http://www.dell.com/downloads/global/solutions/public/white_papers/sql_server_2005_se_x64_sp2.pdf

"Deploying Microsoft SQL Server in an iSCSI SAN"
http://www.equallogic.com/uploadedfiles/Resources/Tech_Reports/tr1013-sql-server.pdf

"Deploying Microsoft Windows Server 2003 in an iSCSI SAN"
http://www.equallogic.com/uploadedfiles/Resources/Tech_Reports/tr1015-ms-windows-2003.pdf

"Deploying Microsoft Multipath I/O in an iSCSI SAN"
http://www.equallogic.com/uploadedfiles/Resources/Tech_Reports/tr-ms-windows-mpio.pdf

"Choosing a Member RAID Policy"
http://www.equallogic.com/uploadedfiles/Resources/Tech_Reports/tr-raid-opt_TR1020.pdf

"Deploying Pools and Tiered Storage in a PS Series SAN"
http://www.equallogic.com/uploadedFiles/Resources/Tech_Reports/tr1027-tiered-storage.pdf

"SQL Server Data Protection with Auto-Snapshot Manager"
http://www.equallogic.com/uploadedFiles/Resources/Tech_Reports/TR1032_ASM_SQL.pdf

"Aligning Disk Sectors for Optimal Performance"
https://www.equallogic.com/uploadedFiles/Resources/Tech_Reports/tr-ms-sector-align_TR1012_v2-0.pdf

"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

"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

"Microsoft SQL Server I/O subsystem requirements for the tempdb database"
http://support.microsoft.com/kb/917047/en-us

"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

Figures

Tables