



Dell EqualLogic Best Practices Series

Deploying Microsoft SQL Server 2008 based Decision Support Systems using Dell EqualLogic 10GbE iSCSI Storage

A Dell Technical Whitepaper

Ananda Sankaran

Storage Interoperability Engineering

Dell Product Group
March 2010



THIS WHITE PAPER IS FOR INFORMATIONAL PURPOSES ONLY, AND MAY CONTAIN TYPOGRAPHICAL ERRORS AND TECHNICAL INACCURACIES. THE CONTENT IS PROVIDED AS IS, WITHOUT EXPRESS OR IMPLIED WARRANTIES OF ANY KIND.

© 2010 Dell Inc. All rights reserved. Reproduction of this material in any manner whatsoever without the express written permission of Dell Inc. is strictly forbidden. For more information, contact Dell.

Dell, the *DELL* logo, and the *DELL* badge, *PowerConnect™*, *Dell™ EqualLogic™*, *PowerEdge™* and *PowerVault™* are trademarks of Dell Inc. *Broadcom™* is a registered trademark of Broadcom Corporation. *Intel™* is a registered trademark of Intel Corporation in the U.S. and other countries. *Microsoft™*, *Windows™*, *Windows Server™*, and *Active Directory™* are either trademarks or registered trademarks of Microsoft Corporation in the United States and/or other countries.

Table of Contents

1	Introduction.....	1
1.1	Audience.....	1
1.2	The Rest of This Paper.....	1
2	Understanding DSS Storage Requirements	2
2.1	Strategy for Supporting DSS I/O Requirements.....	2
3	DSS Workload Test.....	4
3.1	Test Methodology	7
3.1.1	Workload Simulation Tool.....	7
3.1.2	Monitoring	8
4	Best Practices	8
4.1	Dell EqualLogic Storage Array Configuration	8
4.1.1	Choose Appropriate Disk Drive Type and Size	8
4.1.2	Setting an Appropriate RAID Policy	8
4.1.3	Storage Pool Configuration	9
4.2	SAN Design	9
4.2.1	Use Redundant Switch Paths Between Servers and Storage Arrays	9
4.2.2	Use Flow Control and Jumbo Frames.....	10
4.2.3	Pay Attention to Spanning Tree	11
4.2.4	Select Appropriate Physical Interconnects.....	11
4.3	Server Configuration.....	11
4.3.1	Hardware Configuration to Meet Performance and Capacity Requirements is Critical.	11
4.3.2	Think beyond Just CPU and Memory.....	12
4.3.3	Properly Size Your SAN Connection Paths.....	12
4.4	Operating System Configuration	13
4.4.1	Disk Partitions Should be Aligned With Storage Stripe Boundaries	13
4.4.2	NIC Driver Settings.....	13
4.4.3	Optimize MPIO Configuration With EqualLogic DSM.....	15
4.5	Microsoft SQL Server 2008	17
4.5.1	Properly Size the Storage Volumes for Each Database I/O Component.	17
4.5.2	Match RAID Levels to DSS I/O Components	18

4.5.3	Use SQL Server Table Partitioning to Improve Query Processing Performance	18
5	Test Results	19
Appendix A	Dell PowerEdge M1000e high-speed fabric architecture	22
Appendix B	System Specifications.....	23
Appendix C	SFP+ Connectivity Options	24
	Related Publications.....	25

1 Introduction

The goal of this paper is to describe best practices for deploying Decision Support System (DSS) or Data Warehousing (DW) solutions based on Microsoft® SQL Server® 2008 using Dell™ EqualLogic™ storage. DSS solutions require scalable storage platforms that offer high levels of performance and capacity. Deploying the storage platform using validated best practices and recommended settings for various system components (storage, server, switch, operating system, and application) will ensure optimal operation.

The Dell Storage Interoperability Engineering Lab conducted a series of end-to-end system integration tests using Dell PowerEdge servers, Dell PowerConnect network switches, and Dell Equallogic storage arrays. From those tests we validated optimal configuration settings for maximizing system performance and for designing storage architectures that are best suited to support the storage demands of DSS applications running on Microsoft® SQL Server® 2008.

Note: Decision Support System (DSS) is a term that defines a broad class of information processing applications which are used to support data analysis and decision making activities. Business Intelligence (BI), Online Analytical Processing (OLAP) and Data Warehousing (DW) systems provide decision support functionality and require similar system designs. Thus, the design and configuration best practices presented in this paper could apply to those scenarios as well.

1.1 Audience

This paper is intended for Solution Architects, Storage Network Engineers, System Administrators and IT Managers who need to understand how to design and deploy Microsoft® SQL Server® 2008 based DSS or DW solutions on Dell® PowerEdge® servers and Dell EqualLogic storage.

1.2 The Rest of This Paper

This rest of this paper contains the following sections:

- Section 2, Understanding DSS Storage Requirements, on page 2
- Section 3, DSS Workload Test, on page 4
- Section 4, Best Practices, on page 8
- Section 5, Test Results, on page 19

2 Understanding DSS Storage Requirements

DSS applications are typically designed to support complex analytical query activities using very large data sets. The queries executed on a DSS database typically take a long time to complete and usually require processing large amounts of data. A DSS query may fetch millions of records from the database for processing. To support these queries the server reads table data from the storage devices. The data I/O pattern from storage to server is predominantly sequential and usually consists of large blocks – typically 512KB or 1MB in size.

It is an important best practice to run DSS workloads on a physically separate system that is optimized for DSS query I/O activity. This will prevent DSS workloads from interfering with the operation of other workloads.

2.1 Strategy for Supporting DSS I/O Requirements

The large sequential I/O patterns in DSS queries require careful system design to insure that the servers, interconnect fabric and storage arrays are able to meet performance and capacity requirements. The overall performance of the DSS solution is determined by the performance characteristics of each component in the system: from the SQL based DSS application driving the I/O workload down to the individual disk drives in the storage area network (SAN) that store the data. The best practices presented in this paper focus on all of the component layers presented in Figure 1.

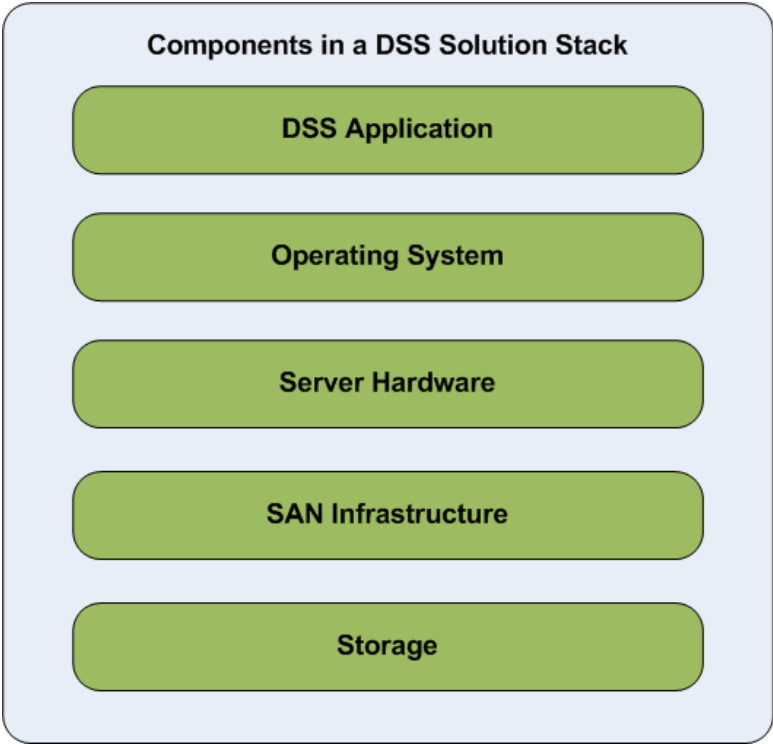


Figure 1: Components in a DSS Solution stack

Each of the component layers in the solution stack requires careful design and configuration to ensure optimal system performance. The considerations at each layer of the solution stack are described as follows:

- Storage** The storage arrays must be configured to support capacity and performance requirements. Considerations include: array controller capabilities, I/O port specifications, RAID policies, hard drive types and spindle speeds. The data volume layout within the storage arrays must also be optimized. Best practice details for storage array configuration are presented in Section 4.1.
- SAN Infrastructure** The SAN infrastructure must be configured to optimize performance (throughput) and availability. The large data I/O characteristics of DSS workloads require special device and protocol level settings. Best practice details for SAN design are presented in Section 4.2.
- Server Hardware** The server configuration must include enough processing capability, memory, I/O bus bandwidth and installed Network Interface Controllers (NICs) to support the DSS workload. Best practice details for server configuration are presented in Section 4.3.
- Operating System** There are multiple operating system settings that must be optimized to support device level capabilities, I/O performance, Multi-Path I/O (MPIO) and most efficient utilization of server resources. Best practice details for operating system configuration are presented in Section 4.4.
- DSS Application** The workload and I/O characteristics of the DSS application must be thoroughly understood before you can most effectively configure the other component layers in the system. In this paper we focus on how these characteristics affect storage array configuration. Details are presented in Section 4.5.

Furthermore, within each layer of the stack you must consider each of the following design goals:

- Availability** Ensuring high availability is a critical design guideline for each component in the solution. Redundant sub-components and connection paths must be maintained to avoid single points of failure.
- Scalability** Scalability is also critical for a DSS solution. Our solution architecture and best practices address both scale out and scale up considerations.
- Performance** End-to-end performance is a key consideration for optimal operation of the solution. Performance of each component in the stack must be balanced such that bottlenecks in one layer do not waste resources in other layers.

Note: To set up the iSCSI SAN in our test configuration we created a 10 gigabit Ethernet connection fabric using Dell PowerConnect 8024F 10GbE switches and Dell EqualLogic PS6010XV storage arrays. The arrays were configured with redundant multi-port 10GbE controllers.

For more information about the Dell PowerConnect 8024F switch:

<http://www.dell.com/downloads/global/products/pwcnt/en/switch-powerconnect-8024f-spec.pdf>

For more information about the Dell EqualLogic PS6010XV storage array:

<http://www.equallogic.com/products/default.aspx?id=8973>
<http://www.dell.com/downloads/global/products/pvaul/en/equallogic-ps6010xv-specsheet.pdf>

3 DSS Workload Test

To support development of this paper we deployed and tested a full-scale DSS Solution configuration at Dell labs. The following were the design goals for the test:

- Generate high I/O load on the SAN using a DSS workload running on Dell PowerEdge servers, Dell EqualLogic PS series 10GbE storage arrays and Dell PowerConnect 10GbE switches.
- Maximize end-to-end solution performance to confirm best practices for system configuration in each component layer.

Various storage configurations were built and tested using simulated DSS workloads. System behavior was observed during each simulation run while hardware and software settings were modified to maximize end-to-end solution performance. The following sections describe the test configuration and test methodology in more detail.

The first test configuration is shown in Figure 2. This test configuration includes one PS6010XV storage array for SQL Server® data and a second PS6010XV storage array for SQL Server® logs. It also includes one Dell PowerEdge M710 blade server connected to the two PS6010XV storage arrays via a 10GbE iSCSI storage network. We chose the PowerEdge M710 server to take advantage of its expanded I/O capability (it can support up to four PCIe x8 Mezzanine Cards). To create the 10GbE iSCSI storage network we installed two Broadcom® NetXtreme II 57710 10GbE Mezzanine Network Interface Cards (NICs) on the server in Fabric B and connected the Broadcom NICs to external PowerConnect 8024F switches via pass through modules installed on the PowerEdge M1000e blade enclosure. One port on each NIC was utilized in this configuration. Fabric A in the blade chassis was used for connecting to other LAN servers.

Note: More information about the fabric architecture on the Dell PowerEdge M1000e is included in Appendix A , Dell PowerEdge M1000e high-speed fabric architecture. System configuration version levels are provided in Appendix B , System Specifications.

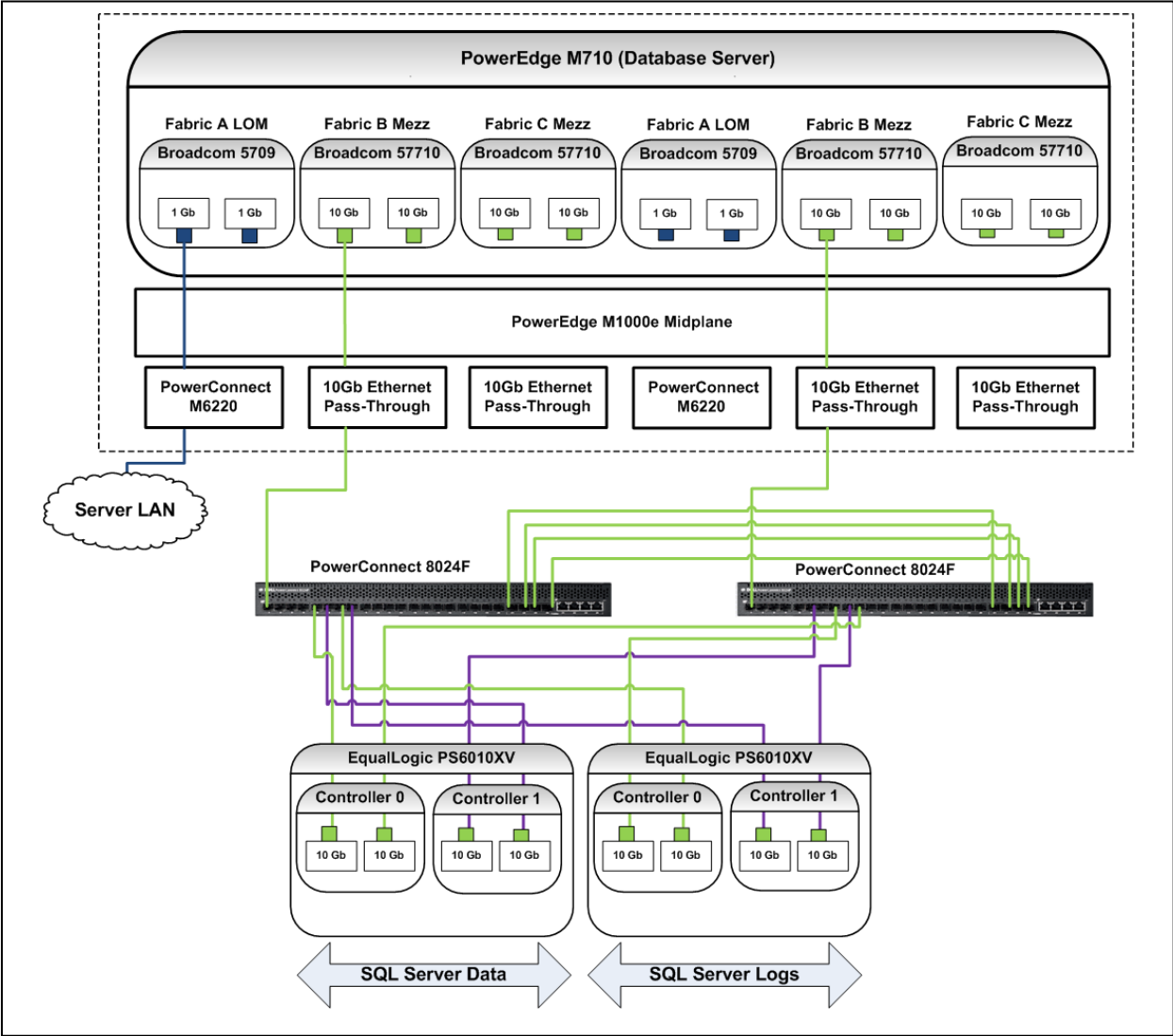


Figure 2: Test Configuration #1

The second test configuration, as shown in Figure 3, includes three PS6010XV arrays for SQL Server® 2008 data and one PS6010XV array for SQL Server® logs. The Storage arrays are connected to a Dell PowerEdge M710 blade server.

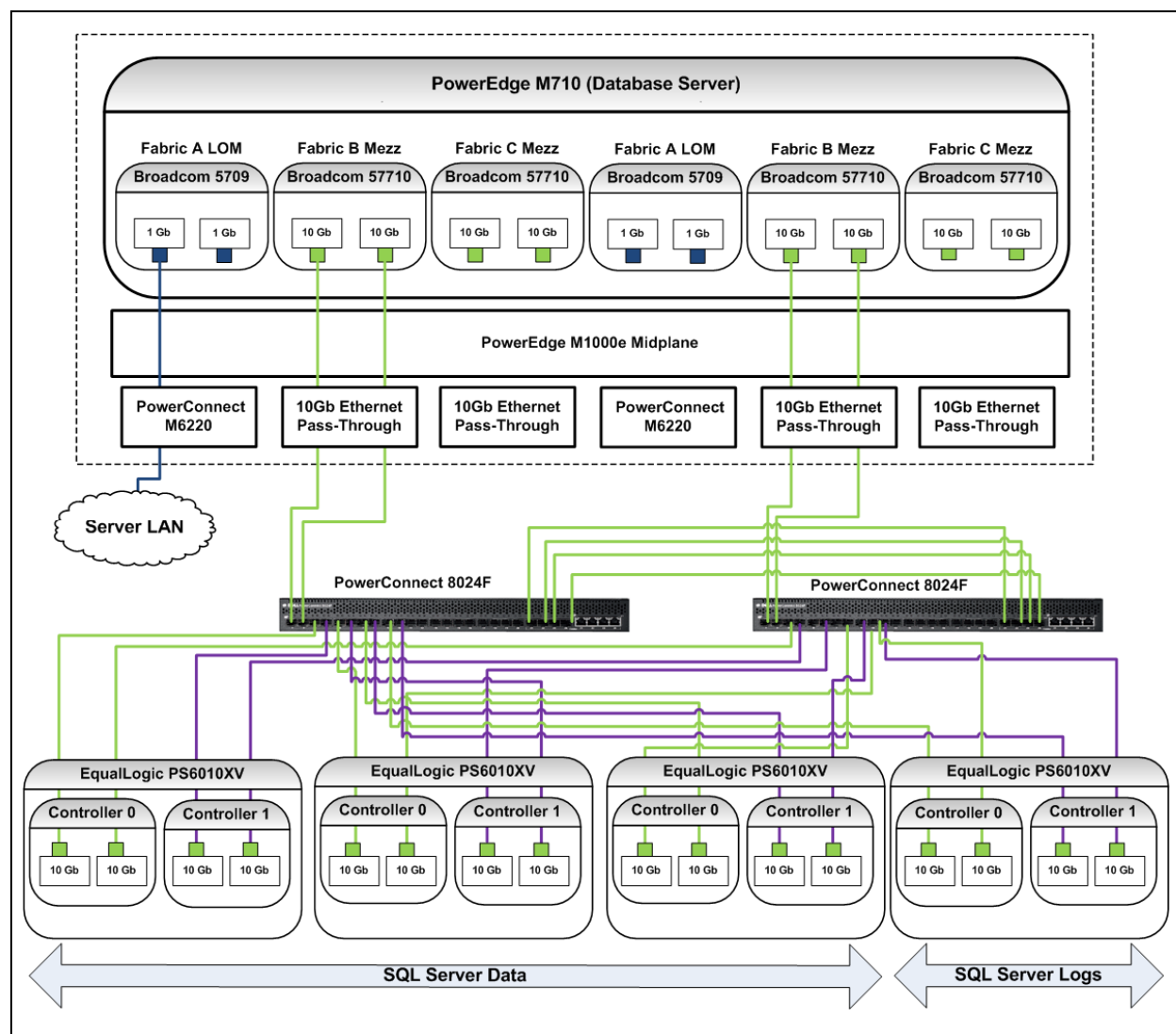


Figure 3: Test Configuration #2

Once again, we chose the PowerEdge M710 server to take advantage of its expanded I/O capability. To create the 10GbE iSCSI storage network we installed two Broadcom NetXtreme II 57710 10GbE Mezzanine Network Interface Cards (NICs) on the server in Fabric B and connected the Broadcom NICs to external PowerConnect 8024F switches via the pass through modules installed on the PowerEdge M1000e blade enclosure. In this configuration we utilized both ports on each NIC. (In Test Configuration #1 only one port on each NIC was utilized.) Similar to Configuration 1, Fabric A in this configuration was used for connecting to other LAN servers.

Please note that in both test configurations we implemented hardware redundancy at various levels to ensure high availability (HA). The HA configuration uses dual network cards on the servers, dual network switches and dual storage controllers to create redundant connection paths, thus eliminating single points of failure across the SAN fabric.

Note: For additional information about network configuration and high availability see the following documents:

- *Dell EqualLogic PS Series Network Performance Guidelines*, available at: <http://www.equallogic.com/resourcecenter/assetview.aspx?id=5229>
- *10 Steps to Deploying a High-Performance Fault-Tolerant SAN Using Dell EqualLogic PS Series Storage and Cisco Catalyst Network Infrastructure*, available at: <http://www.equallogic.com/resourcecenter/assetview.aspx?id=8299>

In both test configurations we configured the Dell PowerEdge M710 as the database server running Microsoft® SQL Server® 2008 Enterprise Edition. We used a second Dell PowerEdge M610 in each test configuration to execute the DSS query simulation tool (Quest BenchMark Factory). The M610 also hosted the EqualLogic SAN Headquarters (SAN HQ) monitoring application. Both the M710 and M610 servers were installed in the same Dell PowerEdge M1000e blade system chassis.

3.1 Test Methodology

We used a modified TPC-H¹ benchmark to model a DSS workload. Our DSS workload included a stream of 22 database SQL queries that read and then processed large amounts of data from the database. In each run we executed five query streams (simulating five end users) on the database server. These queries exercised large volumes of data to simulate reporting and analytical activity in a typical DSS workload. The test data warehouse was modeled after a sales activity database implemented using SQL Server® 2008. It included data that would typically be found in databases used for managing, selling and distributing sample products. The database size, including the data tables, table indices and all other auxiliary objects was approximately 220GB. The largest table in the database included ~600 million rows. In each test run the stream of DSS queries executed on the database server required reading large volumes of data from the attached Dell EqualLogic Storage arrays hosting the data warehouse.

3.1.1 Workload Simulation Tool

We used Quest Benchmark Factory® for Databases to simulate the DSS workload and to create and populate the DSS database under test.² We executed Benchmark Factory for Databases on a client system (the Dell PowerEdge M610). The test client was connected to the database server under test (the Dell PowerEdge M710) using Fabric A. Before beginning the first test run we created, populated and backed up the test database to unutilized storage space in the SAN. The database was restored as necessary before beginning subsequent test runs. This allowed us to start each test run from exactly the same state.

¹ TPC-H is an industry standard benchmark for simulating DSS workloads on database systems. See <http://www.tpc.org/tpch/default.asp>

² <http://www.quest.com/benchmark-factory/>

3.1.2 Monitoring

We configured Windows Server® 2008 Performance Monitor (PerfMon) to monitor the utilization of server resources – processor, memory, disk I/O and network. We also used the Dell EqualLogic SAN HeadQuarters (“SAN HQ”) application to monitor the utilization and operation of the storage resources: physical disks, logical volumes, network ports, member arrays and storage pool activity.

Note: SAN HeadQuarters (SAN HQ) can be used for performance and event monitoring across multiple EqualLogic SAN groups. It can help storage administrators monitor and analyze configuration, performance, capacity utilization and connection status. The information captured and provided by SAN HQ can be analyzed at a point-in-time or over a specified time range. SAN HQ is a powerful tool for gathering array performance data. For more details on using SAN HQ, please download the following Dell Technical Report: *“Monitoring your PS Series SAN with SAN HeadQuarters”*, available at:

<http://www.equallogic.com/resourcecenter/assetview.aspx?id=8749>

The SAN HeadQuarters application is available for download at:

<https://www.equallogic.com/support/>

4 Best Practices

In Section 2 we discussed the need to take an end-to-end approach when designing optimal configurations for SQL Server based DSS solutions. In this section of the paper we focus on best practices for creating and deploying optimized configurations.

4.1 Dell EqualLogic Storage Array Configuration

You should consider the following best practices when configuring Dell EqualLogic storage arrays for a DSS solution.

4.1.1 Choose Appropriate Disk Drive Type and Size

In both test configurations we used Dell EqualLogic PS6010XV storage arrays configured with 16 450GB SAS 15K RPM drives. We used the SAS 15K drives because they offer the highest level of disk I/O throughput.

SAS 10K drives (Dell EqualLogic PS6010X) or SATA drives (Dell EqualLogic PS6010E) can also be considered for DSS environments requiring lower levels of I/O throughput. Choice of drive size depends on the capacity requirements of the DSS database or the data warehouse.

4.1.2 Setting an Appropriate RAID Policy

In both test configurations we configured RAID 50 on each array using 14 disk drives, with the remaining 2 disk drives in each array configured as hot spares. We used RAID 50 because it offers a good balance between capacity utilization and read performance levels.

Since DSS environments are highly read intensive and require large storage capacities, it is a best practice to use RAID50 with EqualLogic PS Series storage arrays. RAID 10 is also recommended for environments that include a mix of both read and write operations. For more information on setting RAID policies see Table 1: Matching RAID levels to DSS I/O componentsTable 1, in Section 4.5.2.

4.1.3 Storage Pool Configuration

In each test configuration we included all member arrays in a single EqualLogic group. The storage arrays under test were managed using a single instance of the Group Manager user interface.

We created two storage pools in each test configuration. One pool was used for hosting the SQL Server® data volumes, including the tempdb volumes. We used a second pool for hosting the SQL Server® transaction logs. The transaction logs were physically separated from the DSS database disk drives for high levels of data protection and recoverability.

DSS workloads are primarily read intensive and they typically generate minimal transaction log activity. Thus, in smaller DSS database environments using Dell EqualLogic storage arrays the transaction log volumes can be hosted on the same pool and disk drives as the data volumes without sacrificing performance. However, if high levels of data protection or recoverability are required SQL Server mirroring or replication solutions can be considered.

For DSS solutions that require high levels of data protection and recoverability, it is a best practice to host data and log volumes on separate disk drives and pools, or consider application data mirroring solutions such as SQL Server mirroring or replication.

4.2 SAN Design

You should consider the following best practices when configuring the SAN for a DSS solution.

4.2.1 Use Redundant Switch Paths Between Servers and Storage Arrays

In our test configurations we used redundant switch paths in the SAN to ensure uninterrupted access to storage controllers in the event of a switch or interconnect cable failure or a scheduled switch downtime for maintenance. Figure 4 shows how we used two switches to create redundant connection paths between the server NICs and storage array controllers. This example illustrates the following best practices to follow while designing an iSCSI SAN for a DSS solution:

- Equally distribute the NIC ports installed on the server for iSCSI SAN connectivity across the redundant switches.
- Equally distribute the Ethernet ports from each controller of the storage array across the redundant switches.
- Interconnect the two switches with sufficient number of inter-switch links (ISLs) to support the workload under test.

In our testing we found that a link aggregation group (LAG) consisting of four 10Gbps ISLs between the Dell PowerConnect 8024F switches was required for the workload under test. It is a best practice to provision the LAG bandwidth based on the throughput requirement between the servers and arrays plus I/O workload characteristics. The required number of switches in the SAN design depends on the number of servers and the number of storage arrays required to support the workload.

Figure 4 does not show the separate switching infrastructure used for the Server LAN. It is a best practice to always isolate the SAN so that the iSCSI switching infrastructure is dedicated to storage traffic only.

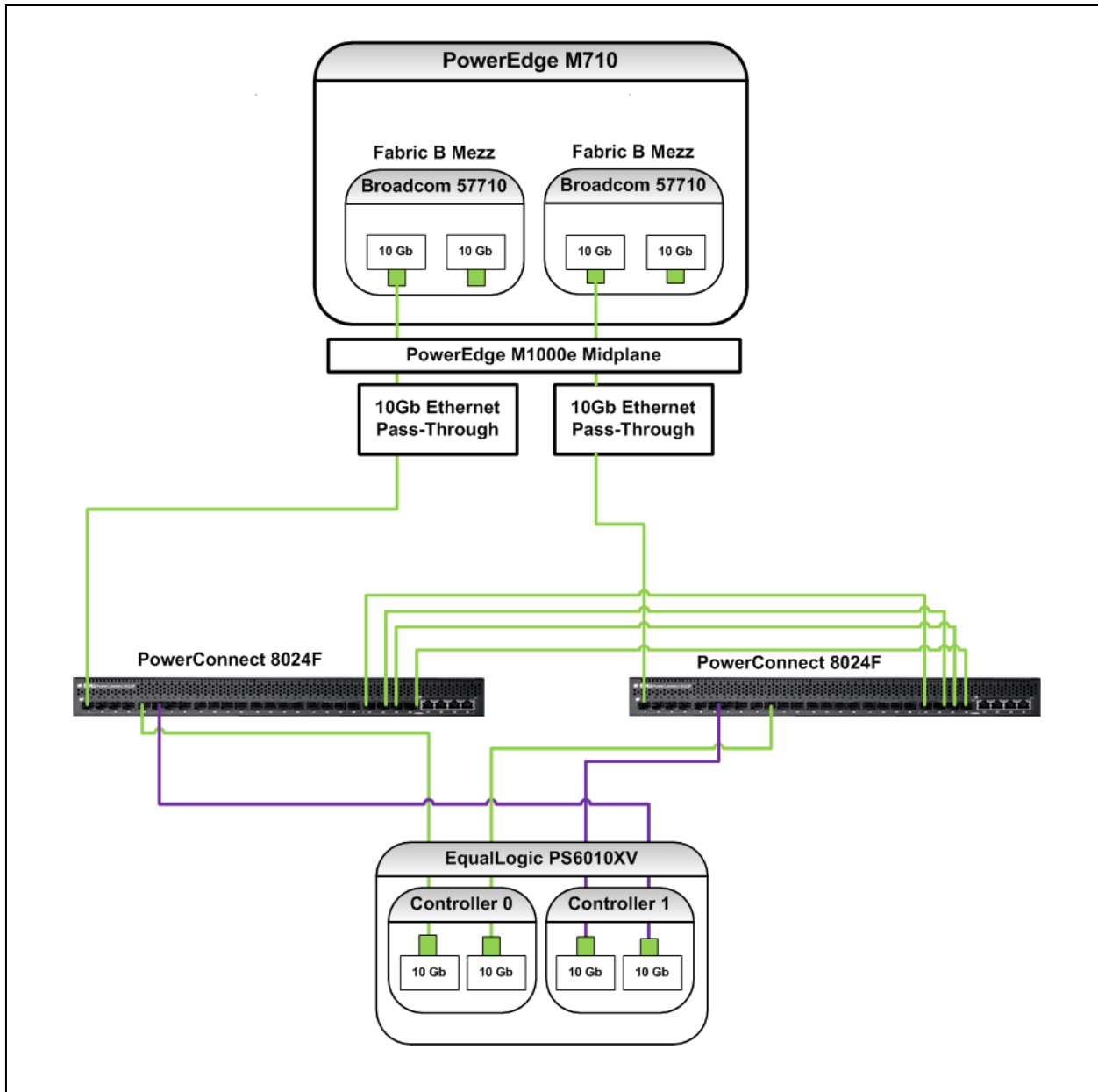


Figure 4: Redundant Connection Paths for High Availability

4.2.2 Use Flow Control and Jumbo Frames

Flow control increases the ability of the SAN to regulate packet flow between nodes (server, switches, and storage) during high I/O traffic conditions.

Jumbo frames are Ethernet frames that have a MTU larger than 1518 bytes. Jumbo frames can increase SAN throughput by allowing more data to be transferred in each Ethernet frame. This will reduce the time it takes to transfer large amounts of data across the SAN.

In our testing we enabled flow control (TX and RX) and used jumbo frames with a MTU setting of 9216 bytes on the switch ports. The jumbo frame size on the switch ports must match or exceed the settings at the end device ports (the server and storage NICs). Generally end-devices conform to a jumbo frame setting of at least 9000 bytes.

It is a best practice to enable flow control (both TX and RX) and jumbo frames on all ports, including LAG ports. In both test configurations we obtained maximum end-to-end throughput when we enabled flow control and jumbo frames on all iSCSI SAN switch ports (including LAG virtual ports).

4.2.3 Pay Attention to Spanning Tree

In our test configurations we disabled Spanning Tree Protocol (STP) on all switch ports connecting to end devices (the server and storage controller NICs). We also enabled the "portfast" setting on those switch ports.

Since NIC and storage controller ports are end points they cannot create loops. Thus you should enable the 'portfast' setting on those switch ports instead. With port fast enabled you eliminate the STP forwarding state delay time that occurs when a port becomes active. This will accelerate initialization of new iSCSI sessions if ports on an existing iSCSI connection path change status. This becomes very important in larger SAN fabric configurations with multiple hops.

4.2.4 Select Appropriate Physical Interconnects

The Dell PowerConnect 8024F switches support SFP+ ports for 10Gb Ethernet connectivity. 10GBASE SFP+ support is typically provided with 10GbE switches. Either SFP+ transceivers along with optical cables or SFP+ copper cables can be used for inter-connecting switches and for connecting end devices to the switches. SFP+ transceivers along with optical cables will support longer distances than SFP+ copper cables. Connection distance and cost will determine the type of the physical connection you use. The various physical connection options for SFP+ 10Gb Ethernet are provided in Appendix C , SFP+ Connectivity Options, on page 24.

4.3 Server Configuration

You should consider the following best practices when configuring servers for a DSS solution.

4.3.1 Hardware Configuration to Meet Performance and Capacity Requirements is Critical.

The Dell PowerEdge M710 database server used in this test was configured with two Intel Xeon X5570 quad-core processors and 32 GB of RAM. Our testing shows that this server configuration can adequately execute the DSS queries to drive the heavy I/O load on the storage arrays as required by our testing.

It is a best practice to use the Dell SQL Server Advisor Tool³ as a starting point for sizing server configurations to support SQL Server DSS workloads.

4.3.2 Think beyond Just CPU and Memory

In our test configurations we installed two Broadcom mezzanine NICs on the PowerEdge M710 database server for connecting to the external iSCSI storage via Fabric B. Each NIC is connected to the two blade CPU sockets internally through an eight-lane (x8) PCIe interface as shown in Figure 5. This 8 lane interconnect path provides up to 16 Gbps of PCIe 1.1 compliant bandwidth per mezzanine NIC.⁴ In our testing we found that this mezzanine NIC and PCIe lane configuration provided the required bandwidth for the workload under test.

It is a best practice to choose servers with appropriate PCIe technology, sufficient PCIe slots and the required number of PCIe lanes necessary to provide the data transfer bandwidth required by the DSS workload.

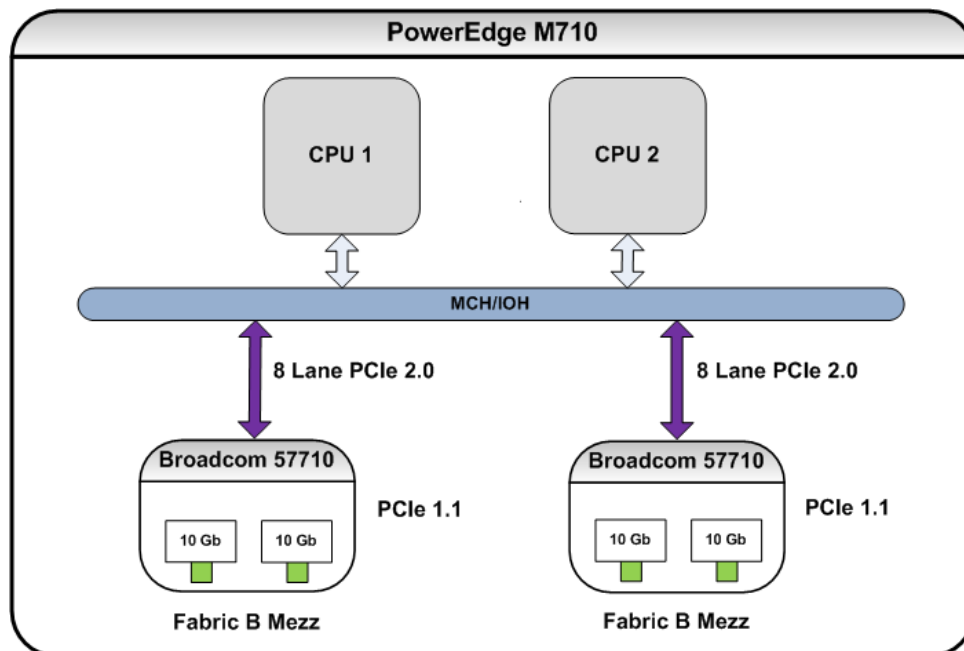


Figure 5: PCIe Paths

4.3.3 Properly Size Your SAN Connection Paths

In Test Configuration #1 we used two 10 GbE ports to connect to the SAN and in Test Configuration #2 we used four 10 GbE ports (see Figure 2 on page 5 and Figure 3 on page 6). We used two physical NICs per server to create redundant paths for high availability. In our testing we found that this configuration of SAN NIC ports provided the required bandwidth for the workload under test. We also installed separate NICs for the server LAN connections. We connected the workload simulation client

³ The Dell SQL Server Advisor Tool is available at: <http://www.dell.com/sql>

⁴ The Dell PowerEdge R710 supports PCIe 2.0. Use of PCIe 2.0 compliant NICs would increase this bandwidth further.

running Quest Benchmark Factory for Databases and the monitoring system running SAN HQ to the database server via the server LAN.

It is a best practice to provision the quantity and connection speed of the SAN NIC ports on your server based on the throughput requirements of the workload.

4.4 Operating System Configuration

You should consider the following best practices for operating system settings in a DSS solution.

4.4.1 Disk Partitions Should be Aligned With Storage Stripe Boundaries

You can enhance performance by ensuring alignment of file system partitions with the external storage striping scheme. In an optimal configuration your partition boundaries will align with the RAID striping scheme at block size multiples to minimize I/O operations. If partitions do not align with the RAID striping scheme (i.e. block size) then the system will generate more I/O operations than in an optimal configuration.

Windows Server® 2008 by default aligns partitions with an offset of 1024KB and thus accommodates all RAID stripe sizes up to 1024KB. This default alignment is sufficient for creating partitions on volumes hosted on EqualLogic Storage. Prior to Windows® 2008 the best practice is to manually align the data partitions with an offset suitable to the RAID stripe size. This can be done using a utility such as diskpart.exe.

Note: We recommend that you review the information provided by Microsoft in following paper: *“Microsoft SQL Server 2008: Disk Partition Alignment Best Practices for SQL Server”*, available at the following location:

<http://msdn.microsoft.com/en-us/library/dd758814.aspx>.

4.4.2 NIC Driver Settings

Jumbo frames provide a more efficient transport for the large data transfers typically generated by DSS workloads. Flow control increases the ability of the NIC to regulate packet flow between the NIC and the switch ports during high I/O traffic conditions.

It is a best practice to enable flow control (TX and RX) and jumbo frames on the Server NIC ports used for connecting to the SAN and match these settings with the switch and array controller port settings.

Utilize Offload Engine Capabilities in Your NICs

We configured the BroadCom 57710 NIC ports on the database server to operate in the TCP offload Engine (ToE) mode. For optimal performance with SQL DSS workloads you can enable ToE mode on the NIC to work in conjunction with the Microsoft® iSCSI software initiator, Windows® MPIO, and the Dell EqualLogic Multipath Device Specific Module (DSM) for MPIO.

The ToE mode can be enabled for the Broadcom NICs using the Broadcom Advanced Control Suite (BACS) application. Ensure that ToE and the NDIS (Network Driver Interface Specification) options are enabled on the iSCSI NIC ports. Disable iSCSI offload mode if ToE and NDIS are enabled on the same port. ToE mode will need to be enabled in the operating system before you enable it in BACS.

TCP offload is disabled by default in Windows Server® 2008 and on the NIC. You will need to make sure it is properly enabled in both locations. To configure ToE enablement in the Operating system, run the following commands in the administrator mode via a Windows® command prompt.

To configure OS TCP offload:

```
C:\>netsh int tcp set global chimney=enabled
```

To verify OS TCP offload:

```
C:\>netsh int tcp show global
```

Note: More information on enabling TCP offload in the operating system can be found in the following Microsoft Knowledgebase Articles:

- *Information about the TCP Chimney Offload, Receive Side Scaling, and Network Direct Memory Access features in Windows Server 2008:*
<http://support.microsoft.com/kb/951037>
- *The Microsoft Windows Server 2003 Scalable Networking Pack release:*
<http://support.microsoft.com/kb/912222>

As an alternative to ToE mode, the network card can also be setup in the iSCSI Offload Engine (iSOE) mode. The EqualLogic MPIO DSM is designed to work with ports operating in iSOE mode too. We did not validate iSOE mode in our test configuration.

Disable NIC Teaming and Unused NICs

Since we enabled MPIO in our test configurations, we disabled NIC teaming on all ports used for iSCSI connections⁵. We also disabled any unused NIC ports on the system to simplify system management and reduce the chances of making errors during configuration.

It is a best practice to disable NIC teaming on all ports used for iSCSI connections. You should also disable unused NIC ports to simplify system management and reduce the chance for configuration errors.

4.4.3 Optimize MPIO Configuration With EqualLogic DSM

In both test configurations we enabled the Dell EqualLogic Device Specific Module (DSM) for Multi-Path I/O (MPIO). The DSM is provided by the Dell EqualLogic Host Integration Toolkit ("HIT Kit") for the server operating system. The HIT Kit can be obtained from Dell EqualLogic support. Figure 6 shows the MPIO configuration for a volume hosted on 3 arrays. We used the default value (2) for the number of connections per member. This created a total of 6 connections to the volume since the volume spans three pool members. Only logical connections to the active storage controller are shown in Figure 6.

⁵ See *Microsoft iSCSI Initiator Version 2.x Users Guide*:
<http://download.microsoft.com/download/a/e/9/ae91dea1-66d9-417c-ade4-92d824b871af/uguide.doc>

Note: Configuring EqualLogic DSM for MPIO provides enhanced performance benefits for bandwidth intensive applications like SQL Server® DSS. The DSM also enables easier configuration by automating setup of the iSCSI sessions (connections) to the target volumes. MPIO settings can be specified using the "Remote Setup Wizard" program that is installed by the HIT Kit. The wizard offers settings for load-balancing policies such as "Least Queue Depth", "Round Robin" and "Fail-Over".

More information about the Dell EqualLogic Host Integration Toolkit can be found at the following link: <https://www.equallogic.com/support/download.aspx?id=9049>

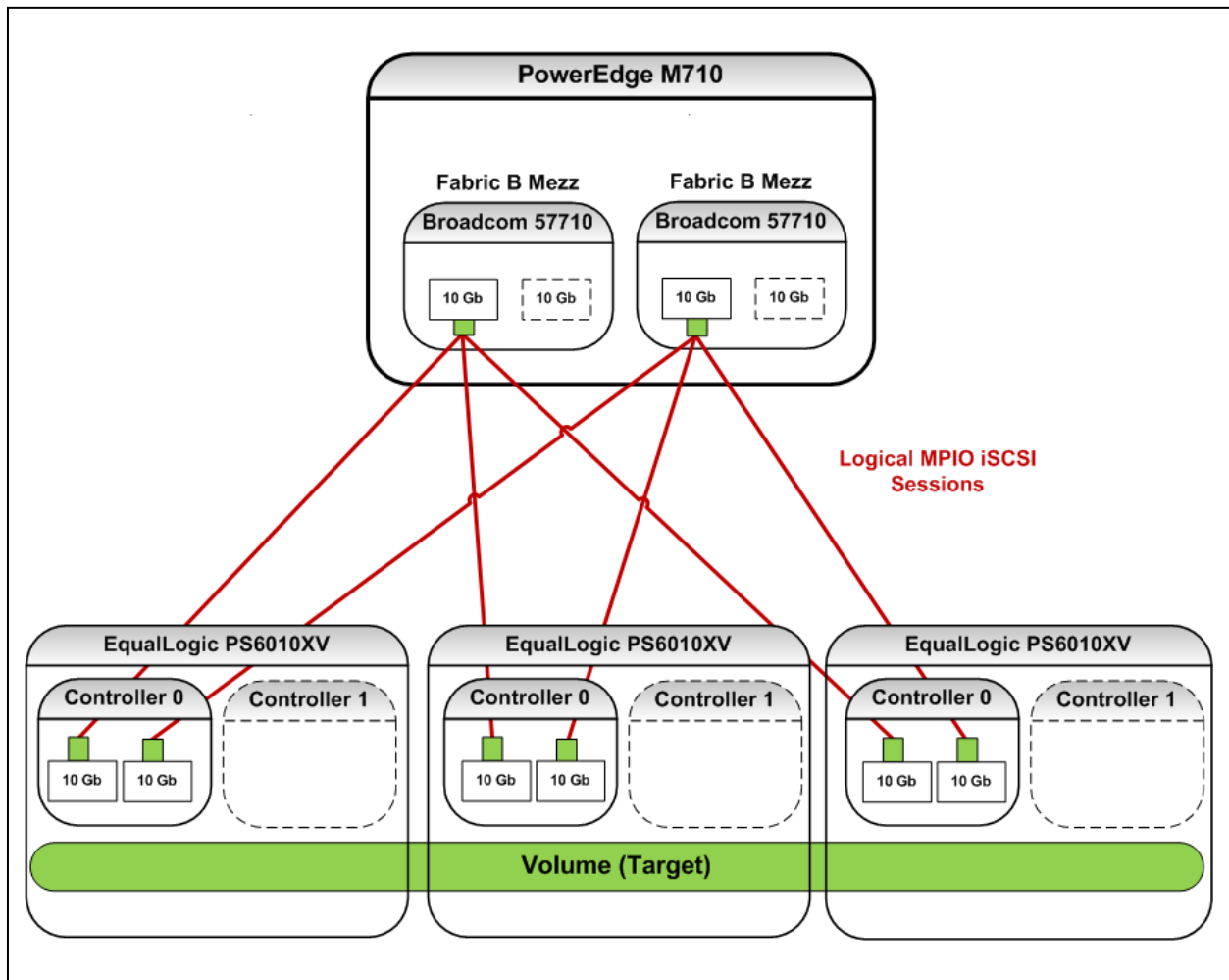


Figure 6: Logical MPIO iSCSI Sessions

During configuration of the MPIO DSM we chose the **Least Queue Depth** option as the default load balancing policy, as shown in Figure 7. Using the Least Queue Depth option will cause the best path to

be chosen for I/O (instead of the “Round Robin” approach, which would choose the next path in the sequence regardless of queue depth).

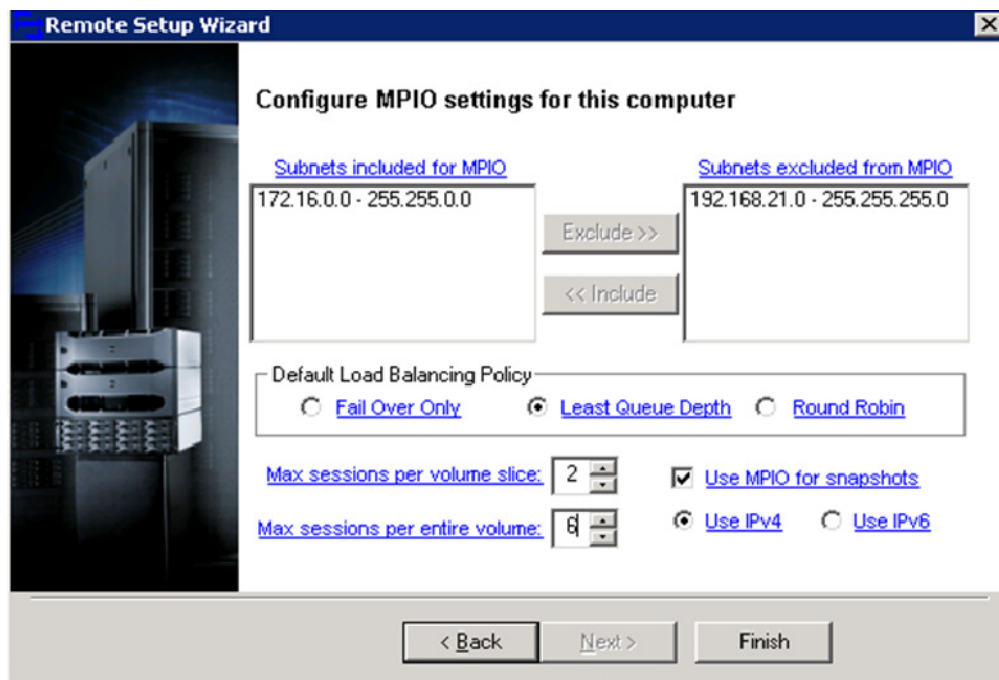


Figure 7: MPIO Settings

4.5 Microsoft SQL Server 2008

You should consider the following best practice guidelines for configuring SQL Server®.

4.5.1 Properly Size the Storage Volumes for Each Database I/O Component.

SQL Server® data layout on storage volumes is an important factor for optimal operation of DSS workloads. Storage volumes need to be appropriately sized for capacity and performance to host the database data, transaction logs and ‘tempdb’ databases. In the test configurations discussed in this paper, we created five volumes for hosting the database data. We created two additional volumes for hosting the ‘tempdb’ database and for the transaction logs. The data and ‘tempdb’ volumes were hosted in one storage pool. The log volumes were hosted in a separate storage pool.

DSS workloads often create large temporary data objects during query processing (see Section 4.1.3, Storage Pool Configuration). Therefore, it is critical to size the ‘tempdb’ appropriately. An improperly sized tempdb can significantly impact performance.

Note: For more information on tempdb sizing and best practices, see the following Microsoft Knowledgebase articles:

- *Capacity Planning for tempdb:* <http://msdn.microsoft.com/en-us/library/ms345368.aspx>
- *Optimizing tempdb Performance:* <http://msdn.microsoft.com/en-us/library/ms175527.aspx>
- *Troubleshooting Insufficient Disk Space in tempdb:* <http://msdn.microsoft.com/en-us/library/ms176029.aspx>

4.5.2 Match RAID Levels to DSS I/O Components

Different database types can generate different I/O patterns in a DSS environment. I/O performance can be optimized based on which RAID level is used. See Table 1 below for recommendations.

Volume Type:	Recommendation:
Database Volumes	DSS workloads are read intensive, therefore RAID 50 is recommended ⁶ to optimize for both capacity and performance.
TempDB Volumes	RAID 50 is recommended to optimize for both capacity and performance if write activity is low. If write activity is high, RAID 10 is recommended.
Log Volumes	DSS workloads have relatively low log activity, except during database refreshes. Therefore, RAID 50 is recommended to optimize for both capacity and performance. If DSS DB refreshes include very large bulk updates then RAID 10 is recommended.

Table 1: Matching RAID levels to DSS I/O components

4.5.3 Use SQL Server Table Partitioning to Improve Query Processing Performance

We used table partitioning in our test configurations so that SQL Server could most efficiently fetch data. We partitioned the largest table in our test into four subsets. The subset data was spread across four different volumes residing in their own file groups. The rest of the database tables were hosted on the fifth volume. We based this partitioning scheme on a particular table column that contained the most common field in the user queries.

Table partitioning is a recommended best practice to improve query processing so that SQL Server can efficiently fetch relevant data⁷. SQL Query efficiency and I/O throughput improvements depend on the partitioning scheme chosen, which in turn is heavily dependent on the data and workload characteristics. Typically, partitioning will offer benefits when implemented on frequently accessed

⁶ For more information, see "How to Select the Correct RAID for an EqualLogic SAN", available at <http://www.equallogic.com/resourcecenter/assetview.aspx?id=8071>

⁷ See: MSDN SQL Server Development Center, "Partitioned Table and Index Strategies Using SQL Server 2008": <http://msdn.microsoft.com/en-us/library/dd578580.aspx>

large tables and when based on the most commonly queried table field. We also recommend as a best practice that you align the indices with the same partitioning scheme used for table data.

5 Test Results

A PS6010XV array can deliver up to 2.5 times the throughput (megabytes per second) of a PS6000XV array⁸. However, every DSS implementation is different. Actual throughput may vary depending on database size and workload characteristics.

We conducted two SQL DSS tests to understand PS6010XV array scalability using the test configurations described earlier in this paper. The first configuration used a single storage array for SQL Server data and the second configuration used three storage arrays for SQL Server data. Transaction logs were hosted on a separate array in both configurations. Using Quest Benchmark Factory for Databases we executed five database query streams on the server.

The test workload and database (including same data and volume size) was deployed on two different test configurations to understand how performance scaled when increasing the number of storage array members in the pool. Server and Storage resource utilization metrics were monitored using the EqualLogic SAN HQ and Windows PerfMon tools. In particular the Server NIC throughput (data received from storage) was monitored. The intent was to capture the amount of data, including protocol overhead that the storage arrays in both configurations delivered to the server. Figure 8 shows relative instantaneous peak throughput at the Server NICs (receive) for both configurations, as measured using Windows Performance Monitor.

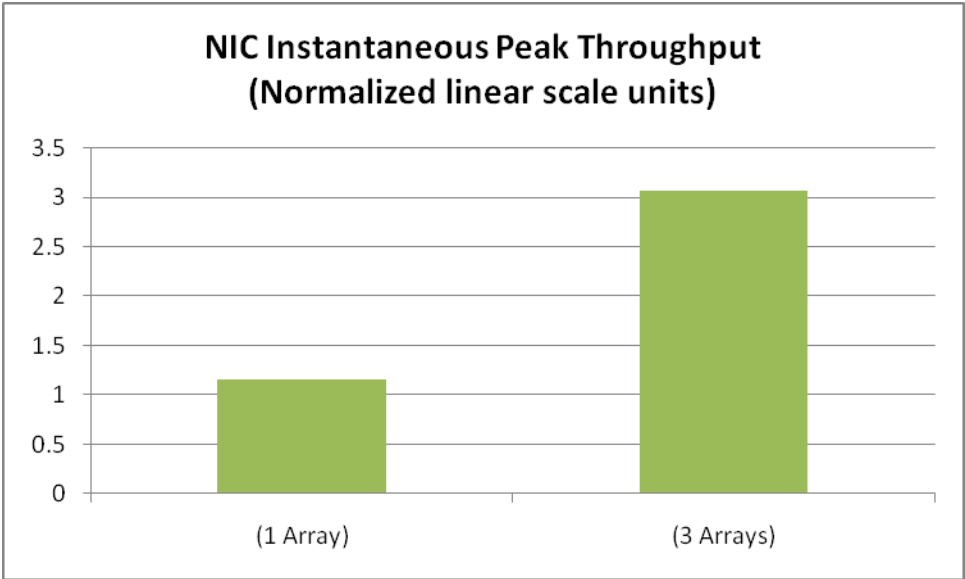


Figure 8: Instantaneous Peak Throughput

⁸ Dell EqualLogic PS6010XV product details:
<http://www.equallogic.com/products/default.aspx?id=8973>

The maximum throughput observed at the server NICs scaled almost linearly from the single array configuration to the three array configuration. This was due to the increased storage bandwidth available (three times the number of storage array 10GbE controller ports) and the increase in available storage resources (three times the number of storage controllers and disk drives). The peak throughput rates occurred when multiple queries were fetching very large data sets simultaneously.

The average processor utilization on the database server during both tests is shown in Figure 9. The response times measured at the test client for completion of all 5 query streams are shown in Figure 10. These results illustrate that as more I/O bandwidth is available (due to increased storage resources), the database server is able to reduce query response times while increasing processor utilization during query execution. In the 1 array configuration the processors received data at a slower rate due to limited storage throughput and were relatively idle. Queries took longer to execute because of limited storage throughput, even though server processing power was available. In the 3 array configuration the processors received data at a faster rate. Higher storage throughput enabled full utilization of server processing capability, with shorter query processing times.

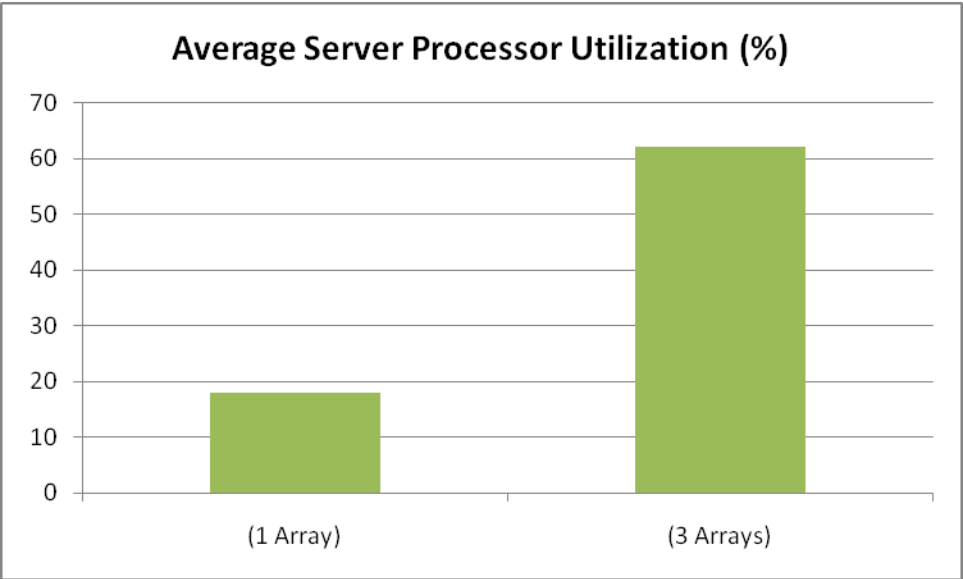


Figure 9: Average Processor Utilization

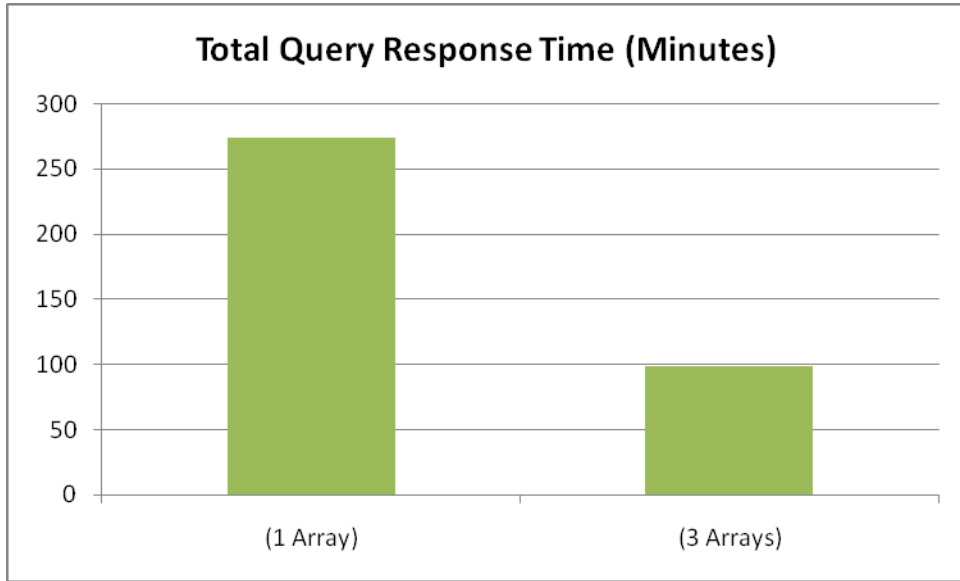


Figure 10: Total Query Response Time

Appendix A Dell PowerEdge M1000e high-speed fabric architecture

We used a Dell PowerEdge M1000e blade server enclosure for our test configurations. Figure 11 shows the interconnect architecture. For more information, see *Exploring the Dell PowerEdge M1000e Network Fabric Architecture*, available at:

<http://content.dell.com/us/en/enterprise/d/business~solutions~whitepapers~en/Documents~exploring-the-dell-poweredge.pdf.aspx>

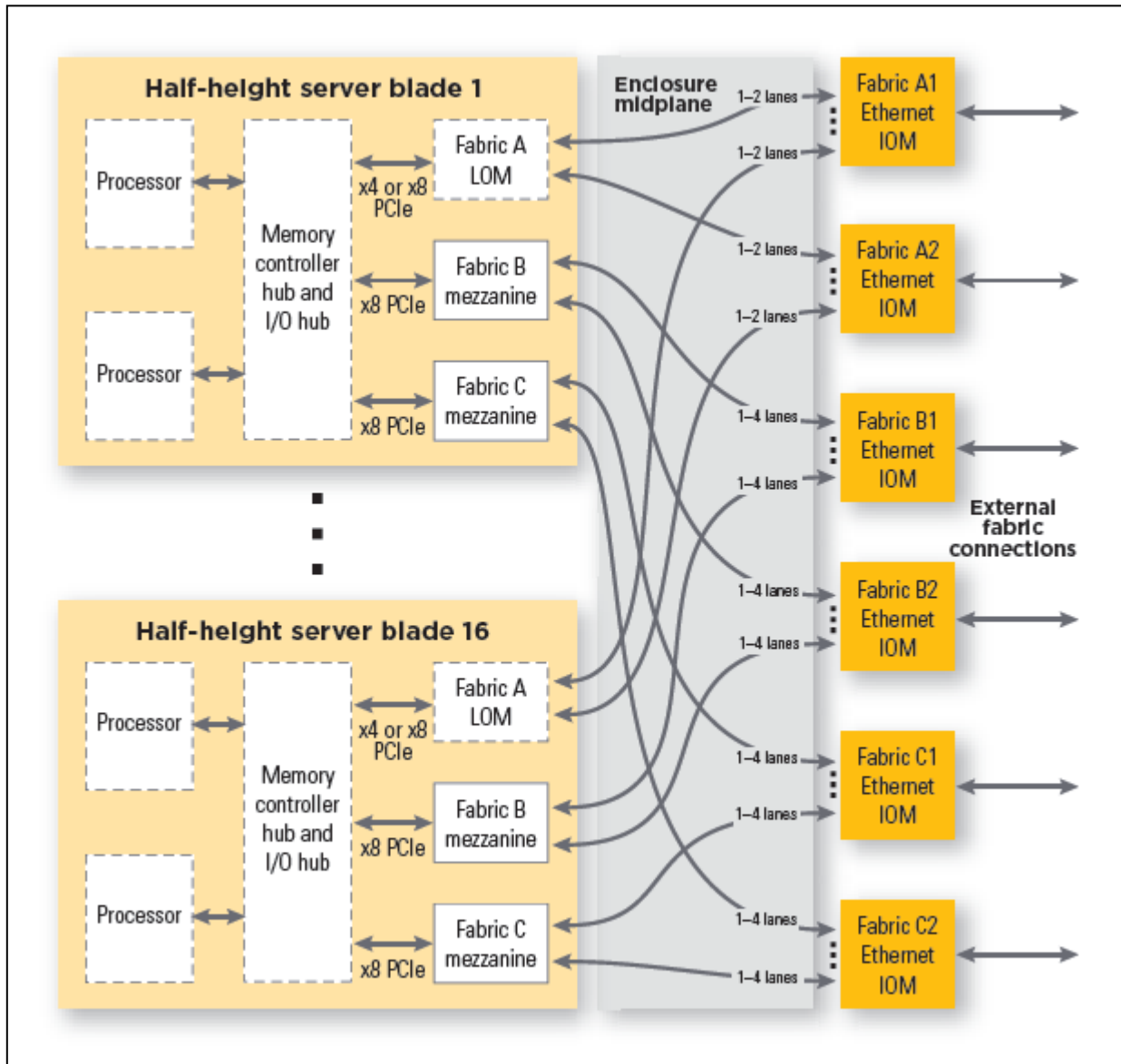


Figure 11: Dell PowerEdge M1000e Fabric Architecture

Appendix B System Specifications

Table 2 shows the configuration of the DSS server used in our test configurations.

Dell PowerEdge M710 Configuration ⁹	
Component	Description
Processors	2 x Intel Xeon X5570, 2.93Ghz, 8M Cache, 6.40 GT/s QPI, Turbo, HT
Memory	32GB Memory (8x4GB), 1066MHz Dual Ranked RDIMMs for 2 Processors
BIOS	1.2.7
Onboard NIC (Fabric A)	Broadcom 5709 Quad Port 1GbE NIC with TOE
iSCSI test network server initiator NIC (Fabric B)	Broadcom 57710 NetXtreme II 10GbE I/O Card ¹⁰ for M-Series Blades (NDIS mode) NIC Firmware: v5.0.9
Embedded Management	iDRAC6 Enterprise
Hard Drives	2 x 73GB 15K RPM SAS 2.5 inch hot-pluggable; Hard Drive Controller: SAS6/IR supporting 2HDs; configured as RAID 1 Mirror
Operating System	Windows Server 2008, Enterprise x64 with SP2 iSCSI NIC driver: 5.0.13.0 (Broadcom 14.1.1 A03 LAN Kit) MPIO enabled EqualLogic HIT Kit version 3.2
Application	SQL Server 2008 Enterprise x64 Edition with SP1

Table 2: Test System Configuration

⁹ Specification Guide: <http://www.dell.com/downloads/global/products/pedge/en/server-poweredge-m710-tech-guidebook.pdf>

¹⁰ Specification Guide: <http://www.dell.com/us/en/enterprise/networking/nic-broadcom-57710-blade/pd.aspx?refid=nic-broadcom-57710-blade&cs=555&s=biz>

Table 3 shows the configuration of the storage array and switch used in our test configurations.

Dell EqualLogic PS6010XV Configuration	
Storage Controllers	Dual 10GbE controllers with 4GB battery-backed memory
Controller Firmware	4.3.2
Hard Disk Drives	16 x SAS hot-pluggable disk drives; 15,000 RPM
Hard Drive Capacities	450GB each
RAID	RAID50
SAN Headquarters	V 2.0
Dell PowerConnect 8024F SAN Switch	
Firmware	3.1.1.4

Table 3: Array and Switch Configuration

Appendix C SFP+ Connectivity Options

We expect most initial implementations of 10GbE switches, iSCSI target adapters and host NICs will utilize SFP+ form factor physical connections. The SFP+ based Direct Access Copper solution (DAC or 10GSFP+Cu) is a fairly short-haul solution restricted to cable lengths of less than 15 meters. The standard copper SFP+ cable includes the SFP+ interface as part of the cable. Fiber SFP+ solutions will require the purchase of separate interface modules that match the cable standard being used. Various SFP+ cable options are listed in Table 4 below. The best practices presented in this paper do not differentiate between SFP+ physical implementations. They should apply to any of the various SFP+ cabling solutions available.

Connector	Cable	Distance	Notes
10GBASE-CX1	Twin Ax Copper	< 15m	SFP+ interface built into cable ends.
10GBASE-SR	MM Fiber (850nm)	26-300m	26m w/ standard FDDI grade MMF 300m w/ OM3 MMF
10GBASE-LR	SM Fiber (1310nm)	10km	Standard Single Mode Fiber
10GBASE-LRM	MM Fiber (1310nm)	220m	Standard Multi-Mode Fiber

Table 4: SFP+ cable options

Related Publications

The following Dell publications are referenced in this document or are recommended sources for additional information.

- *Dell EqualLogic PS Series Network Performance Guidelines*
<http://www.equallogic.com/resourcecenter/assetview.aspx?id=5229>
- *10 Steps to Deploying a High-Performance Fault-Tolerant SAN Using Dell EqualLogic PS Series Storage and Cisco Catalyst Network Infrastructure:*
<http://www.equallogic.com/resourcecenter/assetview.aspx?id=8299>
- *Monitoring your PS Series SAN with SAN HeadQuarters:*
<http://www.equallogic.com/resourcecenter/assetview.aspx?id=8749>

SAN Headquarters software download and documentation:

- <https://www.equallogic.com/support/download.aspx?id=8923> (Dell EqualLogic Support ID required)
- Dell EqualLogic Host Integration Toolkit (HIT Kit):
<https://www.equallogic.com/support/download.aspx?id=9049> (Dell EqualLogic Support ID required)
- Dell SQL Server Advisor tool:
<http://www.dell.com/sql>
- *Exploring the Dell PowerEdge M1000e Network Fabric Architecture*
<http://content.dell.com/us/en/enterprise/d/business~solutions~whitepapers~en/Documents~exploring-the-dell-poweredge.pdf.aspx>
- *Dell PowerEdge M710 Technical Guidebook:*
<http://www.dell.com/downloads/global/products/pedge/en/server-poweredge-m710-tech-guidebook.pdf>
- *Broadcom NetXtreme II 57710 Dual Port 10Gb Ethernet Mezzanine Card with TOE and iSCSI Offload for M-Series Blades:*
<http://www.dell.com/us/en/enterprise/networking/nic-broadcom-57710-blade/pd.aspx?refid=nic-broadcom-57710-blade&cs=555&s=biz>

The following Microsoft® publications are referenced in this document or are recommended sources for additional information.

- *Microsoft SQL Server 2008: Disk Partition Alignment Best Practices for SQL Server:*
<http://msdn.microsoft.com/en-us/library/dd758814.aspx>
- *Information about the TCP Chimney Offload, Receive Side Scaling, and Network Direct Memory Access features in Windows Server 2008:*

<http://support.microsoft.com/kb/951037>

- *The Microsoft Windows Server 2003 Scalable Networking Pack release:*
<http://support.microsoft.com/kb/912222>
- *Microsoft iSCSI Initiator Version 2.x Users Guide:*
<http://download.microsoft.com/download/a/e/9/ae91dea1-66d9-417c-ade4-92d824b871af/uguide.doc>
- *Capacity Planning for tempdb:*
<http://msdn.microsoft.com/en-us/library/ms345368.aspx>
- *Optimizing tempdb Performance:*
<http://msdn.microsoft.com/en-us/library/ms175527.aspx>
- *Troubleshooting Insufficient Disk Space in tempdb:*
<http://msdn.microsoft.com/en-us/library/ms176029.aspx>
- *Partitioned Table and Index Strategies Using SQL Server 2008:*
<http://msdn.microsoft.com/en-us/library/dd578580.aspx>



THIS WHITE PAPER IS FOR INFORMATIONAL PURPOSES ONLY, AND MAY CONTAIN TYPOGRAPHICAL ERRORS AND TECHNICAL INACCURACIES. THE CONTENT IS PROVIDED AS IS, WITHOUT EXPRESS OR IMPLIED WARRANTIES OF ANY KIND.

