



Best Practices and Sizing Guidelines for Transaction Processing Applications with Microsoft SQL Server 2012 using EqualLogic PS Series Storage Arrays

A Dell EqualLogic Best Practices Technical White Paper

Dell Storage Engineering
June 2014

Revisions

Date	Description
June 2014	Initial release

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.

© 2014 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.

PRODUCT WARRANTIES APPLICABLE TO THE DELL PRODUCTS DESCRIBED IN THIS DOCUMENT MAY BE FOUND AT: <http://www.dell.com/learn/us/en/19/terms-of-sale-commercial-and-public-sector> Performance of network reference architectures discussed in this document may vary with differing deployment conditions, network loads, and the like. Third party products may be included in reference architectures for the convenience of the reader. Inclusion of such third party products does not necessarily constitute Dell's recommendation of those products. Please consult your Dell representative for additional information.

Trademarks used in this text:

Dell™, the Dell logo, Dell Boom!™, Dell Precision™, OptiPlex™, Latitude™, PowerEdge™, PowerVault™, PowerConnect™, OpenManage™, EqualLogic™, Compellent™, KACE™, FlexAddress™, Force10™ and Vostro™ are trademarks of Dell Inc. Other Dell trademarks may be used in this document. Cisco Nexus®, Cisco MDS®, Cisco NX-OS®, and other Cisco Catalyst® are registered trademarks of Cisco System Inc. EMC VNX®, and EMC Unisphere® are registered trademarks of EMC Corporation. Intel®, Pentium®, Xeon®, Core® and Celeron® are registered trademarks of Intel Corporation in the U.S. and other countries. AMD® is a registered trademark and AMD Opteron™, AMD Phenom™ and AMD Sempron™ are trademarks of Advanced Micro Devices, Inc. Microsoft®, Windows®, Windows Server®, Internet Explorer®, MS-DOS®, Windows Vista® and Active Directory® are either trademarks or registered trademarks of Microsoft Corporation in the United States and/or other countries. Red Hat® and Red Hat® Enterprise Linux® are registered trademarks of Red Hat, Inc. in the United States and/or other countries. Novell® and SUSE® are registered trademarks of Novell Inc. in the United States and other countries. Oracle® is a registered trademark of Oracle Corporation and/or its affiliates. Citrix®, Xen®, XenServer® and XenMotion® are either registered trademarks or trademarks of Citrix Systems, Inc. in the United States and/or other countries. VMware®, Virtual SMP®, vMotion®, vCenter® and vSphere® are registered trademarks or trademarks of VMware, Inc. in the United States or other countries. IBM® is a registered trademark of International Business Machines Corporation. Broadcom® and NetXtreme® are registered trademarks of Broadcom Corporation. Qlogic is a registered trademark of QLogic Corporation. Other trademarks and trade names may be used in this document to refer to either the entities claiming the marks and/or names or their products and are the property of their respective owners. Dell disclaims proprietary interest in the marks and names of others.



Table of contents

Revisions.....	2
Executive summary	5
1 Introduction.....	6
1.1 Objective.....	6
1.2 Audience.....	6
1.3 Terminology.....	6
2 Database application workloads	9
2.1 OLTP.....	9
2.2 SQL Server I/O	9
3 Test configuration	11
3.1 Physical system configuration.....	11
3.2 Storage configuration	12
4 SQL Server 2012 OLTP I/O performance study	14
4.1 I/O profiling study with IOMeter	14
4.1.1 Test criteria.....	14
4.1.2 IOMeter test configuration.....	14
4.1.3 IOMeter test results	15
4.2 SQL Server 2012 OLTP Workload.....	16
4.2.1 Test criteria.....	16
4.2.2 OLTP test configuration.....	17
4.2.3 OLTP test results with one database	18
4.2.4 OLTP test results with two databases	19
5 Best practice recommendations.....	21
5.1 Storage.....	21
5.2 Network infrastructure.....	21
5.3 VMware vSphere ESXi Server/VM	21
5.4 SQL Server 2012 best practices	22
5.4.1 Database volumes.....	22
5.4.2 Buffer cache size.....	22
5.4.3 Table partitioning	22
5.4.4 Files and file groups.....	23



5.4.5	Data file growth	23
5.4.6	Transaction log file growth	24
5.4.7	Tempdb file growth	24
6	Conclusion	26
A	Solution components	27
B	Additional resources	29



Executive summary

Microsoft SQL Server database applications, ranging from web-based e-commerce sites to accounting systems to customer support programs, are at the heart of business functions. The data generated by these database applications can grow exponentially over time. Thus, designing a solution which can accommodate strict performance requirements and future capacity growth of databases are the key challenges faced by customers.

Unique characteristics of the Online transaction processing (OLTP) workload are a higher percentage of read I/O than write I/O and a high percentage of total I/O being concentrated on a relatively small percentage of the total data set. Dell EqualLogic PS Series hybrid storage arrays that combine solid state drives (SSDs) and hard disk drives (HDDs) within the same chassis with automated, dynamic data tiering are well suited for these OLTP applications. EqualLogic PS Series firmware automatically moves more frequently accessed data to high-performing SSDs and less frequently accessed data to HDDs, thereby helping optimize the balance between performance and capacity needs of a demanding SQL OLTP environment.

The following white paper presents the results of targeted storage I/O performance tests using EqualLogic PS series hybrid arrays and illustrates how the unique automated tiering capability of PS6210XS arrays can help address the needs and challenges of an OLTP database application.

In addition, this white paper shows the best practices to achieve the optimal desired performance for OLTP applications. The design must ensure that the entire ecosystem including server, storage and networking resources are sized and configured appropriately to meet the workload performance requirements.



1 Introduction

OLTP applications often need to deliver extremely fast response times and rapid access to data that is in high demand. Moreover, significant data growth over time in a typical OLTP system leads to high capacity requirements in addition to the performance requirements. While the OLTP data set size can grow exponentially over time, the active, frequently-accessed data set size remains relatively small in comparison to the overall OLTP system size.

The Dell EqualLogic PS6210 family arrays streamline storage management and improve performance. The PS6210XS hybrid flash array is ideal for supporting SQL OLTP deployments, because the automatic tiering function of the PS6210XS hybrid array optimizes the balance between performance and capacity needs.

EqualLogic PS6210XS hybrid arrays have 24 drives, seven high performance solid state disk drives and seventeen high capacity 10K RPM SAS hard disk drives, within a single chassis. The scale-out, peer storage architecture of the EqualLogic PS series allows organizations to start small and grown non-disruptively as the performance and capacity needs increase.

1.1 Objective

The objective of this paper is to present the results of SQL Server I/O performance tests conducted using EqualLogic PS6210XS hybrid arrays. It also provides sizing guidelines and best practices for running SQL OLTP workloads on these arrays.

The following two use cases were analyzed during the test studies for this paper:

- I/O profiling tests using IOMeter were used to establish baseline I/O performance characteristics for the EqualLogic PS6210XS array when running OLTP-like I/O patterns before deploying databases.
- Performance characterization tests were run using Dell Quest Benchmark Factory (QBMF), an industry standard benchmark tool, that simulated SQL transactions by running an OLTP like workload.

1.2 Audience

This paper is intended for system administrators, database administrators and storage architects interested in deploying Microsoft SQL Server 2012 virtualized database solutions using Dell EqualLogic PS Series hybrid SANS. Readers should have familiarity with Microsoft SQL Server 2012 system administration, EqualLogic storage array configuration and VMware ESXi administration tasks.

1.3 Terminology

The following terms are used throughout this document.

Autogrowth: Autogrowth properties control how the database file expands when its maximum file size is reached. This can be adjusted in the Files page for the database properties in SQL Server Management Studio.



Buffer cache: Buffer cache minimizes physical disk I/O by holding (buffering) copies of requested data blocks in memory. Data pages that are buffered in the cache can be served at memory access speed, much faster than going to disk for it. By default, SQL Server dynamically grows and shrinks the size of its buffer pool (cache), depending on the physical memory load that the operating system reports. However, the buffer cache size can be adjusted using the max server memory setting in SQL Server Management Studio.

Group: One or more EqualLogic PS Series arrays connected to an IP network that work together to provide SAN resources to host servers.

Hypervisor: The software layer that manages the access to hardware resources on a server. It sits between the hardware and the guest operating systems.

Member: A single physical EqualLogic array.

OLTP I/O Pattern: OLTP workloads tend to select a small number of rows at a time. These transfers happen anywhere in the data and are each fairly small in size – typically between 8K and 64K. This causes the I/O pattern to be random in nature. The key metric in measuring performance of OLTP workloads is the number of I/Os per second (IOPS) that can be achieved while maintaining a healthy response time.

Perfmon: perfmon.exe is a process associated with the Microsoft Windows Operating System. Performance Monitor or perfmon, measures performance statistics on a regular interval and saves those statistics to a file. The administrator picks the time interval, file format and the counters to monitor.

Pool: A logical collection that each array is assigned to after being added to a group and the array contributes its storage space.

Primary data file (.mdf): The primary data file contains the startup information for the database and points to the other files in the database. User data and objects can be stored in this file or in secondary data files. Every database has one primary data file.

Primary Filegroup: In SQL Server, the primary file and all system tables are allocated to the primary filegroup.

SAN HQ: EqualLogic SAN Headquarters (SAN HQ) monitors one or more PS Series groups and collects data over time for later analysis.

Secondary data file (.ndf): Secondary data files are optional, user defined, and store user data. Secondary files can be used to spread data across multiple disks by putting each file on a different disk drive. Additionally, if a database exceeds the maximum size for a single Microsoft Windows file, the secondary data files can be used so the database can continue to grow.

Transaction log file (.ldf): The transaction log files hold the log information used to recover the database. There must be at least one log file for each database.



User defined Filegroup: User-defined filegroups are specifically created by the user when the database is created or later modified. It can be created to group data files together for administrative, data allocation, and placement purposes.

Virtual Machine: An operating system implemented on a software representation of hardware resources (for example, processor, memory, storage, network). Virtual machines are usually identified as guests in relation to the hypervisor that executes the processes to allow them to run directly on the hardware.



2 Database application workloads

Different types of database applications have varying needs, and understanding the models for the most common database application workloads can be useful in predicting the possible application behavior. The most common database application workload models are Online Transaction Processing (OLTP) and Data Warehouse (DW). This paper focuses on the OLTP database workloads.

2.1 OLTP

OLTP database applications are optimal for managing rapidly changing data. These applications typically have many users who are performing transactions and changing real-time data. Although individual data requests by users usually reference only a few records, many of these requests are being made at the same time. Examples of different OLTP systems include airline ticketing systems, banking transaction systems, and web ordering systems.

Optimizing an OLTP database system running on a SQL server allows the maximum number of transactions through the system in the least amount of time. The key metric in measuring performance of OLTP workloads is the number of I/Os per second (IOPS), which reflects the number of transactions that can be achieved while maintaining a healthy response time. For OLTP transactions to take place, SQL relies on an efficient I/O subsystem. According to a *Microsoft SQL Server best practices article* (<http://technet.microsoft.com/en-us/library/cc966412.aspx>), an OLTP transaction profile is composed of the following pattern:

- OLTP processing is generally random in nature for both reads and writes issued against data files.
- Read activity, in most cases, is constant in nature.
- Write activity to the data files occurs during checkpoint operations whose frequency is determined by the recovery interval settings.
- Log writes are sequential with a varying size, which is dependent on the nature of the workload.
- Log reads are also sequential in nature.

2.2 SQL Server I/O

The first step in being able to determine the requirements for a storage system is to understand the application I/O pattern. An understanding of their behavior and frequency is needed in order to properly understand the requirements of that system.

The following terms describe the frequency and size of reads and writes sent by the SQL Server application that are received and processed by the storage system.

SQL Server reads: SQL Server performs two types of reads; logical and physical. Logical reads occur when the Database Engine requests a page from the buffer cache. Physical reads, on the other hand, occur when the page is not currently in the buffer cache and the database engine retrieves the data from the I/O storage subsystem. These copies are taken from the disk into the cache.



SQL Server writes: There are also logical and physical writes. Logical writes occur when data is modified in a page of the buffer cache and physical writes occur when the page is written to the disk from the buffer cache. Both reading from and writing to a page happen at the buffer cache. Each time a page is modified in the buffer cache, it is marked as dirty and is not immediately written back to the disk. A record of the changes is made in the log cache for every logical write. To avoid any loss of data, SQL Server makes sure that the log records are written first to a disk. The associated dirty page is removed from the buffer cache and written to a disk later. SQL Server uses a technique known as write-ahead logging that prevents writing a dirty page before the associated log record is written to a disk.

Transaction log: The SQL Server transaction log is a sequential write-intensive operation and is used to provide recoverability of data in case of database or instance failure.

Tempdb: Tempdb is a system database used by SQL Server as a temporary workspace. Access patterns for tempdb may vary but are generally more like OLTP data patterns.

Read-ahead: SQL Server 2012 supports a performance optimization mechanism called "Read-Ahead". Read-ahead identifies the data and index pages needed to perform a query in advance and tries to bring all those pages into the buffer cache before they are actually needed. This results in significant reduction in I/O to the backend SAN and also lowers application response time

The white paper *OLTP I/O Profile Study with Microsoft SQL 2012 Using EqualLogic PS Series Storage* at <http://en.community.dell.com/dell-groups/dtcmedia/m/mediagallery/20308518/download.aspx> provides more information about the impact of Buffer Cache Size on SQL Server 2012.



3 Test configuration

Section 3.1 illustrates the logical test configuration used to evaluate the solution, and Section 3.2 describes the volume layout configuration on an EqualLogic SAN.

3.1 Physical system configuration

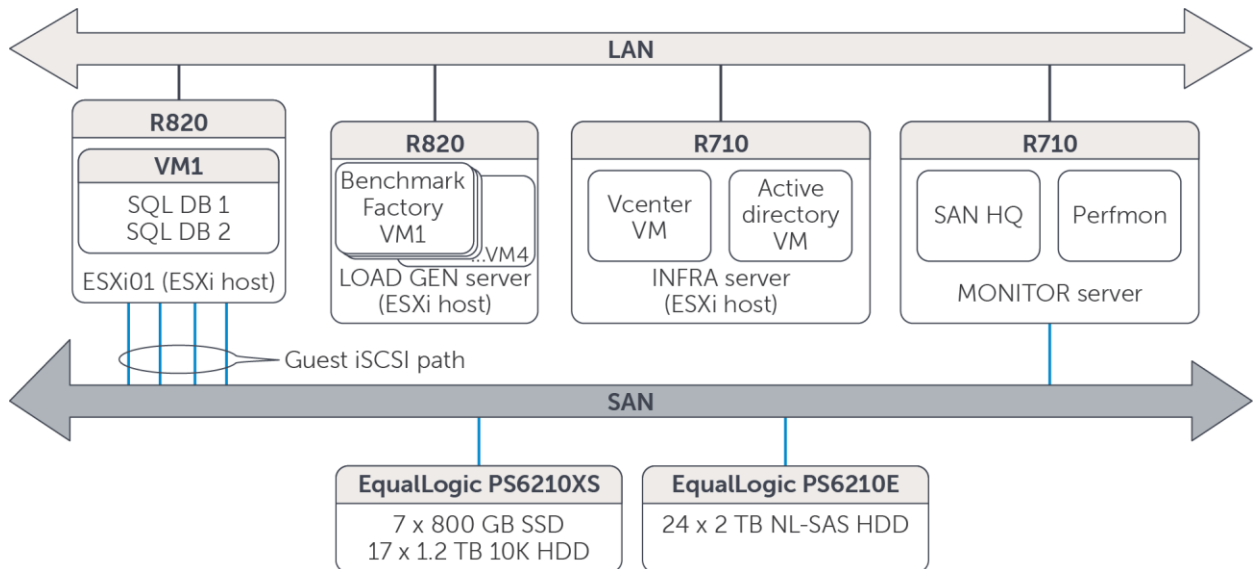


Figure 1 High-level overview of test configuration

Key design details of the test system configuration shown in Figure 1 include:

- Two Dell PowerEdge R820 server were used to host the virtualized Microsoft SQL Server 2012 and the virtualized QBMF clients
 - Microsoft SQL Server 2012 Enterprise Edition was installed on a Microsoft Windows Server 2008 R2 SP1 Enterprise edition virtual machine.
 - Another virtual machine with Microsoft Windows Server 2008 R2 SP1 Enterprise Edition was used for running a synthetic workload using IOMeter.
 - Four virtual machines were created on the second PowerEdge R820 to host the QBMF clients.
 - > These virtual machines were setup with guest iSCSI initiators running the EqualLogic Host Integration Toolkit – Microsoft Edition (HIT/ME) to provide access to the iSCSI SAN.
- Two Dell PowerEdge R710 servers were used to host the infrastructure and monitoring services which included VMware vCenter and EqualLogic SAN HQ, among others.
- iSCSI SAN Networking was provided over two Dell Networking S4810 switches which were stacked to provide high availability, redundancy, and multi-pathing.
- LAN services for communication between the various virtual machines was provided by two Dell Networking S4810 switches, which were stacked.

The logical test configuration diagram shown below was used for validating the SQL Server 2012 OLTP on PS210XS solution.



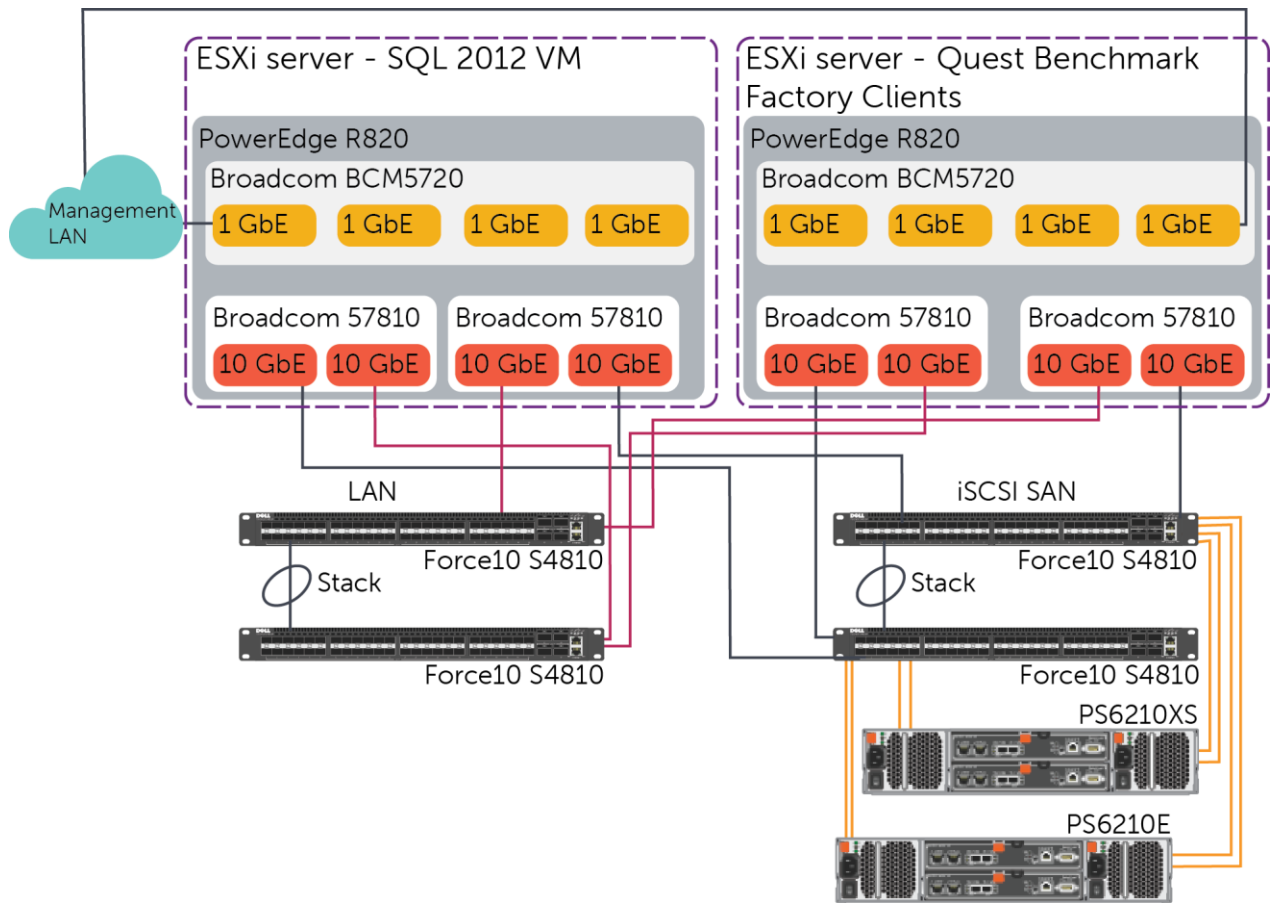


Figure 2 SQL Server 2012 on PS6210XS test configuration, including LAN and iSCSI SAN connectivity

3.2 Storage configuration

The EqualLogic PS6210XS hybrid array hosted the volumes used to simulate OLTP I/O using IOMeter and the database volumes for deploying a SQL Server 2012 OLTP type database.

- EqualLogic PS6210XS hybrid array:
 - 7x 800 GB SSDs and 17x 1.2 TB 10K SAS HDDs with one hard drive serving as a hot spare
 - Dual 2 x 10 GbE controller running firmware version 7.0.2
 - RAID configuration: RAID 6 – Accelerated (Only supported RAID policy on EqualLogic hybrid arrays)
- EqualLogic PS6210E array:
 - 24 x 2 TB 7.2K NL-SAS HDDs with two hard drives serving as hot spare
 - Dual 2 x 10 GbE controller running firmware version 7.0.2
 - RAID configuration: RAID 6 with one hot spare.

The EqualLogic PS6210XS hybrid arrays offer excellent performance and capacity due to the unique combinatory of SSD and 10K SAS drives within the same chassis. The PS6210XS storage array performance

for OLTP environments also depends on various factors such as the total capacity utilized on the array, frequency of data access and also the locality of the data reference.

The PS6210XS arrays offer a total usable capacity of about 18.18 TB distributed across the seven 800 GB SSD drives and 17 1.2 TB 10K SAS drives. The useable capacity offered by the SSDs is about 3.6 TB or about 20% of the total usable capacity which is also typically equal to or more than the measured OLTP usage patterns for the most active portion of the total capacity available. The array attempts to position the frequently accessed data on the SSDs to provide the best performance. If the entire database cannot be accommodated on the SSDs, the data is spread across both SSDs and the 10K SAS drive. Although the data is spread across the SSDs and SAS drives, the PS6210XS arrays automatically monitor the frequently accessed data continuously and move the frequently accessed data to SSDs to achieve better performance.

If all the allocated data is accommodated in the space offered by the SSDs, the storage array performance will be better. To ensure that the dataset used to validate the solution was similar to a realistic environment, more than 30% of the total array capacity was utilized. This ensured that the data was spread across both SSD and SAS drives.

The PS6210E arrays with 2 TB hard drives offer a total usable capacity of about 32.9 TB in RAID 6 configuration. This array was primarily used to host the volumes that held the backups for the SQL databases.

- Eight volumes each of 500 GB capacity were used to simulate OLTP I/O using IOMeter
- Two SQL Server 2012 databases were also deployed, each using seven 150 GB capacity volumes
 - Six of these volumes were used to store SQL Server data files and one stored the log file.
- The total utilized capacity for these 22 volumes was 5.9 TB (32.8%) and the free space on the member was 12.22 TB (67.2%).
- One 1 TB volume was created on the PS6210E in a separate EqualLogic Pool to host the SQL database backups.

Table 1 Volume layout on EqualLogic arrays

Volume name	Array	Capacity	Purpose
IOM-1 – 8	PS6210XS	8 x 500 GB	IOMeter Simulation
SQL-DB1-Data 1-6	PS6210XS	6 x 150 GB	Database volume for DB 1
SQL-DB1-Log	PS6210XS	1 x 150 GB	Database log volume for DB 1
SQL-DB2-Data 1-6	PS6210XS	6 x 150 GB	Database volume for DB 2
SQL-DB2-Log	PS6210XS	1 x 150 GB	Database log volume for DB 2
SQL-Backup	PS6210E	1 x 1 TB	Volume to store database backup



4 SQL Server 2012 OLTP I/O performance study

Typical OLTP type database transactions for financial organizations like banks and brokerage firms are represented by I/O with 8K block size and a mix of 90% reads and 10% writes. A series of I/O profiling studies were run using IOMeter to get a baseline performance of the PS6210XS storage array while simulating OLTP like database I/O operations. Section 4.1 below provides more details about the test configuration and results using the IOMeter tool.

Once the baseline performance was determined, a SQL Server 2012 database was installed and populated with an OLTP like database schema using QBMF. Real-world database transactions were run on this OLTP like database. Section 4.2 below has more information about the test configuration and results.

Additional tests showed that the PS6210XS array was able to provide performance benefits compared to the PS6110XS by adding another database to the SQL Server without adding hardware to the configuration.

4.1 I/O profiling study with IOMeter

The IOMeter tool simulated SQL Server OLTP database transactions (I/O mix of 90% Read and 10% Write) without deploying any SQL Server databases.

4.1.1 Test criteria

The test criteria used for the study was:

- The storage array disk access latencies (read and write) remained below 20ms
- TCP retransmits on the storage network remained below 0.5%

4.1.2 IOMeter test configuration

The test used 8K I/O block size and specific queue depths (number of outstanding I/Os) to determine the maximum IOPS the storage array would sustain within the 20ms latency with read and write latencies measured separately.

The configuration parameters for the tests are shown in Table 2 below.



Table 2 Test parameters: I/O workload studies

Configuration parameters	
EqualLogic SAN	1 x PS6210XS (7x 800 GB SSDs and 17x 1.2 TB 10K SAS HDDs)
RAID type	RAID 6 Accelerated
I/O mix	8K Block size; 90% Read / 10% Write; Queue depth: 20
I/O tool	IOMeter (8 workers mapped to 8 volumes)
Storage configuration	8 x 500 GB volumes 14 x 150 GB volumes (additional allocations to ensure that data is spread across both SSDs and HDDs)

The results collected from the tests are illustrated in Section 4.1.3.

4.1.3 IOMeter test results

The total array capacity utilization was around 6 TB so that the data was not confined to the SSDs, but was spread across both the SSDs and the 10K HDDs. A large amount of simultaneous I/O requests were simulated on the storage array by setting the queue depth to 20 for the eight workers which in turn were mapped to eight individual volumes on the storage array.

Additional IOMeter testing, with smaller storage volumes and the above workload, showed that PS6210XS array was able to provide nearly 51,000 IOPS.

The queue depth was kept constant throughout the test to simulate a heavy load on the EqualLogic array.

The test result showed 8,100 IOPS at the beginning and then gradually increased to around 26,169 IOPS over a duration of about 4.5 hours.



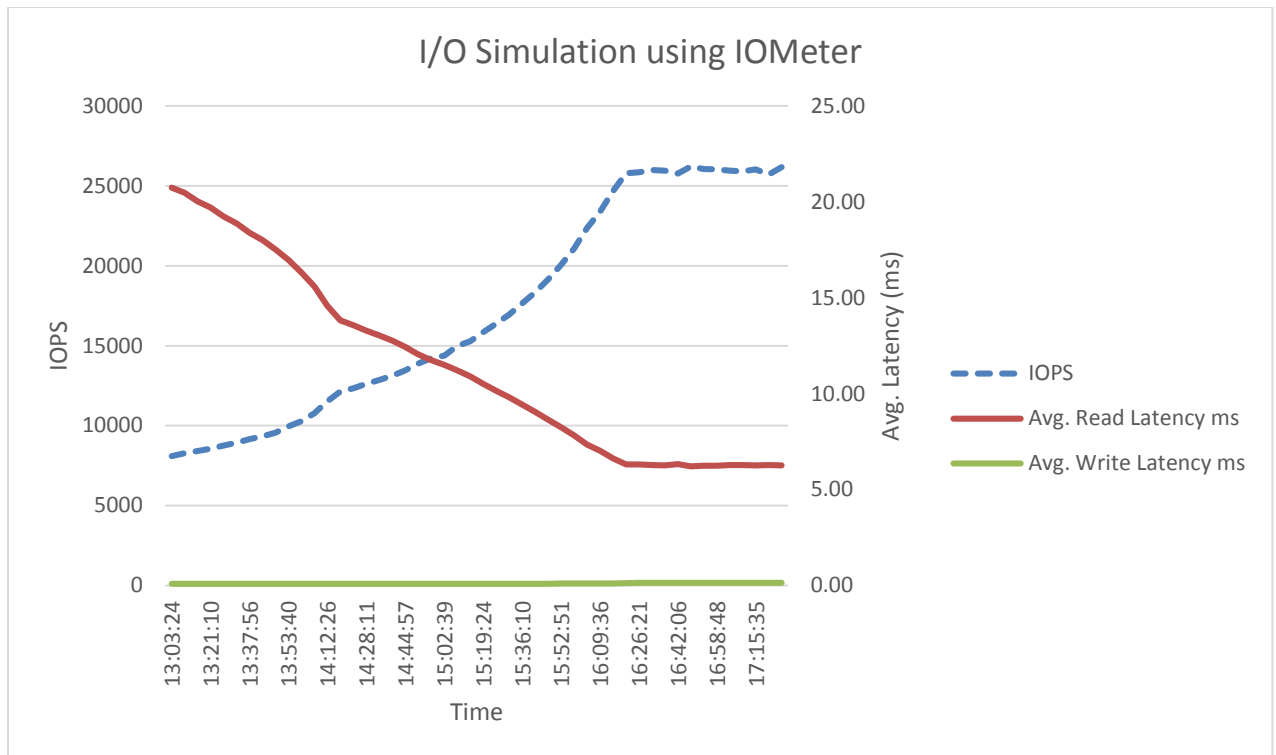


Figure 3 IOMeter OLTP I/O simulation on PS6210XS

As seen in the chart above, the IOPS at the end of the test was nearly three times the number at the beginning of the test because of the automatic tiering technology built in to the PS6210XS hybrid arrays.

The maximum IOPS for the same I/O workload on the previous generation PS6110XS hybrid array was around 16,000 IOPS, and for a non-hybrid PS6110XV array was about 10,000 IOPS compared to the 26,000 IOPS on the PS6210XS.

4.2 SQL Server 2012 OLTP Workload

This section contains details about the database performance characterization tests that were executed with QBMF. An OLTP like database was created and populated using QBMF to simulate a real-world SQL Server OLTP database I/O transactions.

4.2.1 Test criteria

The test criteria used for this study was:

- The storage array disk access latencies (read and write) remained below 20ms to ensure that database application user experience was not degraded.
- Database server CPU utilization remained below an 80% average to ensure the SQL Server database had sufficient resources for database processing.
- TCP retransmits on the storage network remained below 0.5%



4.2.2 OLTP test configuration

The 300 GB OLTP like database was created using QBMF which corresponds to a scaling factor of 36.

Storage arrays usually perform well with read I/O. Also, the memory on the server helps by caching most of the read I/O, which in turn helps improve the overall read performance. A more realistic test for any storage array is when more write I/O transactions are introduced.

The write transactions for the OLTP like workload were increased by modifying the weights associated with the database schema. Figure 4 shows the write intensive transaction mix used to run this test compared to the default transaction mix of QBMF.

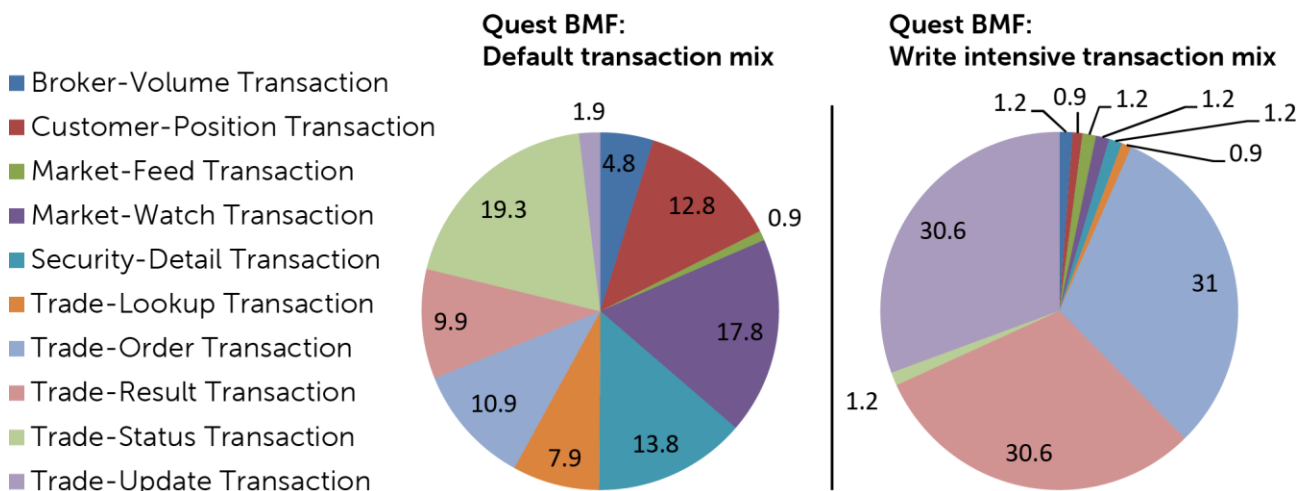


Figure 4 Modified transaction mix to push more write I/O transaction – QBMF

The buffer cache size plays an important role in SQL Server database performance and the amount of I/O that is processed by the backend storage array. The SQL Server buffer cache size was set to 12 GB in this test to force more read I/O operations to the backend storage array.

For more details on buffer cache recommendations and its effects on database performance, refer to the document: OLTP I/O Profile Study with Microsoft SQL 2012 Using EqualLogic PS Series Storage, here: <http://en.community.dell.com/dell-groups/dtcmedia/m/mediagallery/20308518/download.aspx>

In order to demonstrate the improved performance and capacity of the PS6210XS hybrid arrays, the tests were first performed with a single database and then repeated by creating an additional database of the identical size. Adding a second database also meant that another QBMF client was added to run the identical workload on the second database simultaneously. These results are presented in Section 4.2.4.

The configuration parameters for the tests are shown in Table 3 below.

Table 3 Test parameters: OLTP workload studies

Configuration parameters	
EqualLogic SAN	1 x PS6210XS (7x 800 GB SSDs and 17x 1.2 TB 10K SAS HDDs)
RAID type	RAID 6 Accelerated
I/O mix	8K Block size; 90% Read / 10% Write; Queue depth: 20
I/O tool	Quest Benchmark Factory; Scaling factor: 36
Storage configuration	7 x 150 GB volumes (SQL Server 2012 database volumes, primary) 7 x 150 GB volumes (SQL Server 2012 database volumes, secondary) 8 x 500 GB IOMeter volumes (additional volumes to ensure that data is spread across both SSDs and HDDs)

4.2.3 OLTP test results with one database

The database size was around 500 GB but eight additional volumes were created and populated with additional data so that the total array capacity utilization was around 6 TB. This ensured that the data was spread across both SSDs and 10K HDDs. The OLTP workload database transactions were simulated using multiple agents from different machines to simulate real-world customer environment.

To ensure the results were comparable to the IOMeter tests, the following configuration settings were used in the SQL Server database environment to exercise an I/O mix of 90% read / 10% write on the backend storage array.

- OLTP workload transactions were simulated using the modified transaction mix as illustrated in Figure 5 to determine the performance from a PS6210XS array when exposed to write intensive I/O workload, which is a worst-case scenario.
- The number of simultaneous users accessing the database was increased from five to 30 in increments of five.
- Each iteration was run for a duration of 30 minutes.
- The buffer cache size for the SQL Server database was set to 12 GB to push more I/O to the backend storage.



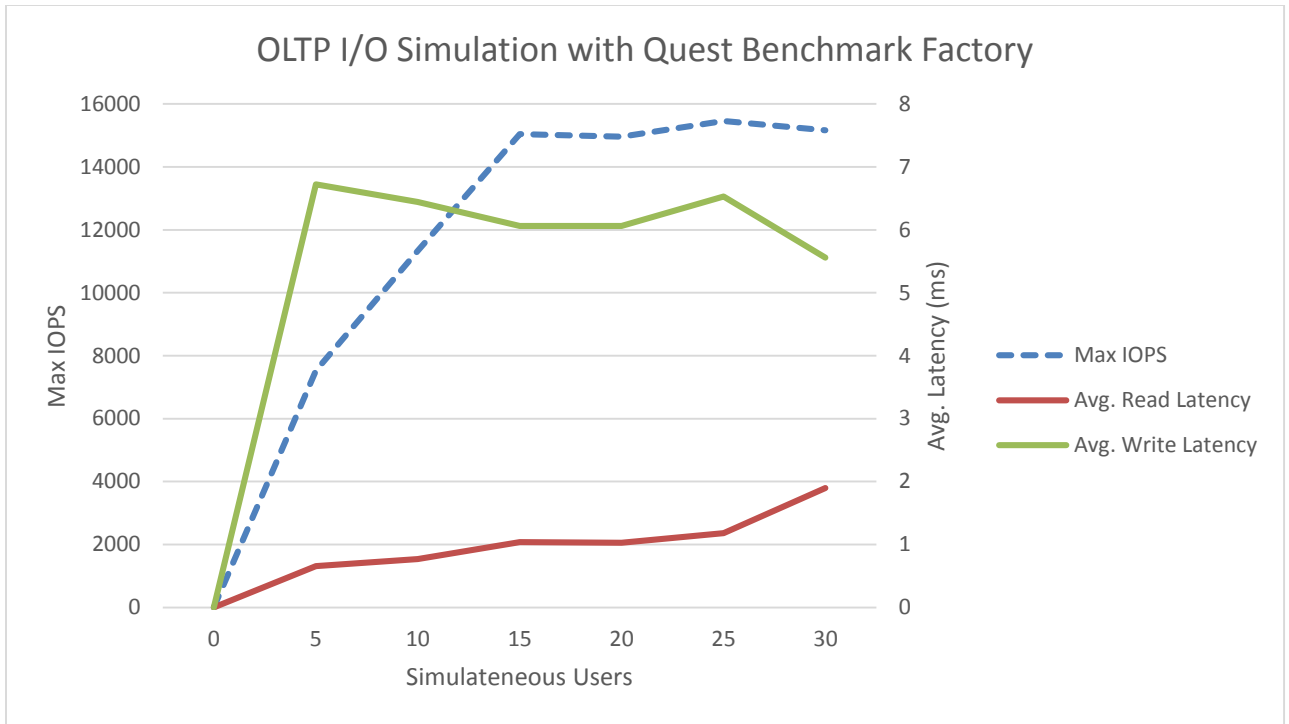


Figure 5 SQL Server 2012 OLTP workload on PS6210XS with one database

The test results showed that the maximum performance reached was nearly 12,000 IOPS which corresponds to a workload of 25 simultaneous users.

The performance of the PS6210XS was nearly 30% better than the performance of the OLTP workload running on PS6110XS hybrid array as seen in a previous paper available here:

http://en.community.dell.com/techcenter/extras/m/white_papers/20439011/download.aspx

4.2.4 OLTP test results with two databases

The test described in Section 4.2.3 was repeated with two active SQL Server 2012 databases. The SQL Server database environment had the following configuration settings.

- OLTP transactions were simulated using the modified transaction mix as illustrated in Figure 4 to determine the performance from a PS6210XS when exposed to write intensive I/O workload from two databases.
- The setup for this test was identical to the test in Section 4.2.3 with the exception of having two databases instead of one.

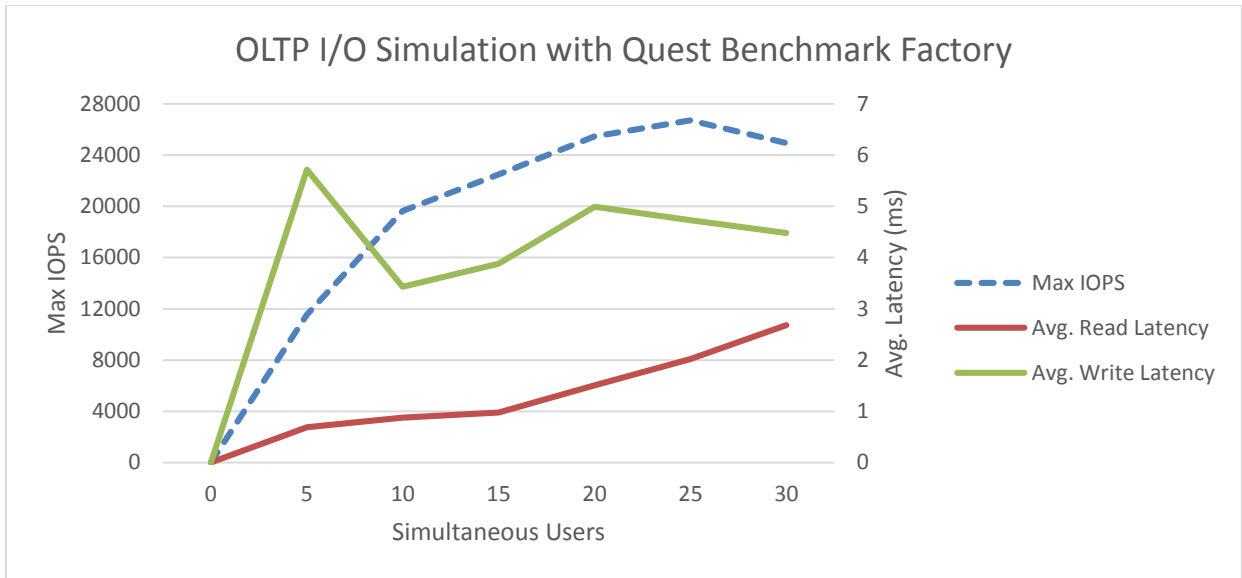


Figure 6 SQL Server 2012 OLTP workload on PS6210XS with two databases

The test showed nearly 26,700 IOPS that corresponded to a workload of 25 users for each of the databases.

From the results in Section 4.2.3 and 4.2.4 it can be clearly seen that the PS6210XS arrays can handle nearly twice the amount of workload, which provides up to 26,700 IOPS while maintaining less than 10 ms latencies for reads and writes.



5 Best practice recommendations

5.1 Storage

Use the following best practices when configuring Dell EqualLogic storage arrays for an OLTP solution.

- When possible, use high performance hybrid arrays or even all-SSD based arrays to host the SQL Server database volumes. This will allow for consistent and sustained performance numbers. For the OLTP test configuration in this paper, PS6210XS hybrid arrays were used to host the database volumes to obtain the best performance from the overall system
- Store database backups and server virtual machines on separate volumes that may be hosted on a high capacity array like the PS6210E or the PS6510E.

5.2 Network infrastructure

The following are the network infrastructure design best practices:

- OLTP workloads tend to be random in nature with smaller block sizes. It is highly recommended to isolate LAN traffic from iSCSI SAN traffic either by using separate infrastructures or by using DCB based 10 Gb Ethernet networks.
- Design redundant SAN component architectures. This includes the NICS on the servers and the switches for the storage network (including server blade chassis switches and external switches).
- Make sure that the server NIC ports and storage NIC ports are connected so that any single component failure in the SAN will not disable access to any storage array volumes.
- Enable flow control on both the server NICs and switch ports connecting the server and storage ports.
- Enable jumbo frames on the server ports and switch ports.
- On iSCSI SAN switches, spanning tree should be disabled on switch ports connected to end devices like server and storage ports. Portfast or its equivalent should be enabled in the switch configuration for these ports.

General recommendations for EqualLogic PS Series array network configuration and performance are provided in the document titled *EqualLogic Configuration Guide* at <http://en.community.dell.com/dell-groups/dtcmedia/m/mediagallery/19852516/download.aspx>

5.3 VMware vSphere ESXi Server/VM

The lab test environment for this paper was comprised of a VMware ESXi server to host the virtualized SQL Server database as well as the QBMF workload simulation, VMware vCenter and other infrastructure virtual machines. The following best practices are applicable for running VMware ESXi based virtual machines in conjunction with EqualLogic storage and/or SQL Server environments.

For these tests, iSCSI SAN storage access was configured for Windows based virtual machines to use a direct access path and the guest OS (Windows) iSCSI initiator.



- Use VMXNET3 guest network adapters for the virtual machines to provide the best performance.
- Enable jumbo frames for the virtual switches assigned to the iSCSI SAN traffic.
- Enable TCP Segmentation Offload and Large Receive Offload in the guest VM NICS for iSCSI SAN traffic.
- When using guest iSCSI initiators, install the Dell EqualLogic Host Integration Tools for Windows within the guest OS. This installs the EqualLogic DSM for the Windows Server MPIO framework that provides multi-path optimizations tailored to the EqualLogic PS arrays.
- When using the VMware ESXi software iSCSI initiators instead of the guest iSCSI initiators, install the EqualLogic Multipathing Extension Module (MEM) for vSphere to provide iSCSI connection management and enhanced load balancing.

5.4 SQL Server 2012 best practices

The following best practices were implemented for the SQL Server 2012 deployment used in this paper.

5.4.1 Database volumes

- Use basic disk type for all EqualLogic volumes
- Use default disk alignment provided by Windows Server 2008 or higher
- Use NTFS file system with 64KB allocation unit for SQL Server database and log partitions

5.4.2 Buffer cache size

SQL Server buffer cache highly optimizes the disk access. The more database pages held in the buffer cache, the more efficient SQL Server will be in responding to the queries which in turn improves application response times. The change in buffer cache sizes changes the logical I/O access pattern, which changes the physical I/O access pattern on the storage array. The impact of the buffer cache size on the read/write percentage, application response times, IOPS, and I/O latency needs to be considered when planning for SQL Server storage sizing. Different customer environments will have different sizing requirements. The appropriate buffer cache size for a specific database size can only be determined by monitoring an existing setup.

Details on the impact of buffer cache size on SQL I/O pattern can be found in the white paper *OLTP I/O Profile Study with Microsoft SQL 2012 Using EqualLogic PS Series Storage* at <http://en.community.dell.com/dell-groups/dtcmedia/m/mediagallery/20308518/download.aspx>.

5.4.3 Table partitioning

Table partitioning requires careful analysis of the existing data access patterns and current operational procedures. A successful execution of table partition requires these steps:

- Plan the partition
 - Select a table or tables that can benefit from the increased manageability and availability of partitioning
 - Select the column or column combination that will be the base of the partition



- Select the type of partitioning to be used
- Specify the partition boundaries in a partition function
- Create a partition storage plan for file groups using a partition scheme

Table partitioning is a recommended best practice to improve query processing so that the SQL Server can efficiently fetch relevant data (see the following note). SQL Query efficiency and I/O throughput improvements depend on the partitioning scheme chosen, which is heavily dependent on the data and workload characteristics. Typically, partitioning will offer benefits when implemented on frequently accessed tables. However, partitioning may not be beneficial and can yield performance degradation if the partitioning scheme is not favoring the query behavior. So the query behavior has to be taken into account when designing the table partitions. In addition to that, the table index partition alignment also plays a major role in query performance improvement. If the indexes are not properly aligned with the table partitions, this may create a CPU overhead, which results in increased query response time. So the table indexes also need to be aligned with the partitions to yield the maximum benefits. It would be useful to monitor and understand the existing database setup and the query processing using a SQL Profiler before planning on implementing the table partitions.

Note: For additional reference, see the MSDN SQL Server Development Center, “Partitioned Table and Index Strategies Using SQL Server 2008”: <http://msdn.microsoft.com/en-us/library/dd578580.aspx>

5.4.4 Files and file groups

A database file is a physical allocation of space and can be primary (.mdf), secondary (.ndf), or log (.ldf). Database objects can be grouped in filegroups for allocation, performance, and administration purposes. There are two types of filegroups: user-defined filegroups and the primary filegroup. Either of them can be the default filegroup. The primary file is assigned to the primary filegroup. Secondary files can be assigned to user filegroups or the primary filegroup. Log files are never a part of a file group. Log space is managed separately from data space. Microsoft’s recommendations are:

- If the primary filegroup is set as default, the size or the auto grow setting needs to be carefully planned to avoid running out of space.
- Microsoft recommends, with larger database deployments that are easily administrated and for performance reasons, to define a user-defined filegroup as the default. In addition, create all secondary database files in user-defined filegroups so that user objects do not compete with system objects for space in the primary filegroup.

5.4.5 Data file growth

If sufficient space is not initially assigned to a database, the database could grow continuously and performance would be affected. Performance is improved if the initial file size and the percent growth are set to a reasonable size to avoid the frequent activation of the auto grow feature. Microsoft’s recommendations are:



- Leave the auto grow feature on at database creation time to avoid running out of space and also to let SQL Server automatically increase allocated resources when necessary without DBA intervention.
- Set the original size of the database to a reasonable size to avoid the premature activation of the auto grow feature.
- Set the auto grow increment to a reasonable size to avoid the frequent activation of the auto grow feature.

5.4.6 Transaction log file growth

The transaction log is a serial record of all modifications and their executions that occurred in the database. SQL Server uses the transaction log for each database to recover transactions. The log file size depends on the recovery models and the frequency of the log backups. If the transaction logs are frequently backed up, theoretically a larger transaction log might not be needed. However, SQL Server can clear only the inactive portion of the log, so if there are active transactions, they might not allow the log to clear immediately. This means that SQL Server will require a larger transaction log. The most preferred recovery model is FULL to minimize downtime and data loss. Microsoft's recommendations for the transaction log are:

- Place the log and the data files into separate volumes to avoid both random and sequential I/O going to the same volume.
- Set the original size of the transaction log to a reasonable size to avoid constant activation of the auto grow feature, which creates new virtual files and stops logging activity as space is added.
- Set the auto grow percent to a reasonable but small enough size to avoid frequent activation of the auto grow feature and to prevent stopping the log activity for too long a duration
- Use manual shrinking rather than automatic shrinking.

5.4.7 Tempdb file growth

The tempdb database is a global resource that holds all the temporary tables and stored procedures for all users connected to the system. The tempdb database is recreated every time SQL Server starts so the system starts with a clean copy of the database. Determining the appropriate size for tempdb in a production environment depends on many factors like the existing workload and the SQL Server features that are used. Microsoft recommends the following:

- Set the recovery model of tempdb to SIMPLE. This model automatically reclaims log space to keep space requirements small.
- Pre-allocate space for all tempdb files by setting the file size to a value large enough to accommodate the typical workload in the environment. This prevents tempdb from expanding too frequently, which can affect performance. The tempdb database should be set to auto grow to increase disk space for unplanned exceptions.
- Set the file growth increment to a reasonable size to avoid the tempdb database files from growing by too small a value. Microsoft recommends the following general guidelines for setting the FILEGROWTH increment for tempdb files.



Table 4 Tempdb file growth increment

Tempdb file size	File growth increment
0 to 100 MB	10 MB
100 to 200 MB	20 MB
200 MB or more	10%



6 Conclusion

This white paper showcases the benefits of the Dell EqualLogic PS6210XS hybrid arrays. The test results clearly illustrate the benefits of automatic tiering of data offered by the PS6210XS arrays which significantly improved the performance of SQL Server based OLTP database applications

Optimal operation of an OLTP application can be achieved when the applicable best practices laid out in this paper are adhered to. It must be ensured that the entire ecosystem including server, storage and networking resources are sized and configured appropriately to meet the workload performance requirements.

The internal tiering mechanism of the PS6210XS monitors the I/O activity continuously and identifies the frequently accessed data. This data is automatically moved to the SSD tier which results in significant performance improvements in terms of increased IOPS and reduced read/write latencies. The SSD tier of the PS6210XS hybrid array is able to use a portion of the SSD space as write cache which further helps in reducing the latencies seen for write requests.

The benefits of automatic tiering depend entirely on the amount of data being accessed frequently and its locality of reference.

The key findings from this white paper are:

- Typical database applications tend to have specific regions of data which are accessed more frequently than other data objects.
- The test results also show that the PS6210XS hybrid arrays provide significant improvement in performance compared to the previous generation PS61xx family of arrays. The maximum IOPS for an OLTP type of I/O workload on a PS6210XS showed a performance increase of nearly 125% compared to the previous generation PS6110XS array.



A Solution components

This section contains an overview of the configurations used throughout the testing described in this document.

Table 5 Test configuration: Hardware components

Solution configuration - Hardware components:		Description
SQL Server 2012 and VM for running IOMeter	BIOS Version: 1.2.6 4x 8 Core Intel Xeon E5-4620 Processors 2.19 GHz 128 GB RAM, 4x 146 GB 15K SAS internal disk drives 2 x Broadcom 57810 10 GbE dual-port NIC - firmware version 7.2.14	1 PowerEdge R820 server running VMware ESXi v5, hosting a SQL Server database virtual machine. This server also hosted a VM for running IOMeter workload.
LoadGen Server (Quest Benchmark Factory – OLTP Workload Database execution)	BIOS Version: 1.2.6 4 x 8 Core Intel Xeon E5-4620 Processors 2.19 GHz 128 GB RAM, 4 x 146 GB 15K SAS internal disk drives 2 x Broadcom 57810 10 GbE dual-port NIC - firmware version 7.2.14	1 PowerEdge R820 server running VMware ESXi v5, hosting 4 virtual machines. 1 virtual machine hosted the console for running Quest Benchmark Factory and other 3 for running agents.
Infra server and Monitoring server	BIOS Version: 6.1.0 Quad Core Intel Xeon X5570 Processor 2.26 GHz 48 GB RAM, 2x 146 GB 15K SAS internal disk drives Broadcom 5709c 1 GbE quad-port NIC (LAN on motherboard) – firmware version 6.4.5 Broadcom 57810 10 GbE dual-port NIC – firmware version 7.2.14	2 Dell PowerEdge R710 Server running VMware ESXi v5, hosting VMs for hosting SAN-HQ and vCenter.
Network	2x Dell Force10 S4810 switches Firmware: 9.1.0.0	2 switches stacked together for iSCSI SAN connectivity.
Storage	1x Dell EqualLogic PS6210XS: 7x 800 GB SSDs and 17 x 1.2 TB 10K SAS Firmware: 7.0.2 RAID 6 Accelerated 1x Dell EqualLogic PS6210E: 24x 2 TB NL-SAS HDDs Firmware: 7.0.2 RAID 6 with 1 Hot Spare	Backend storage for SQL server 2012 database volumes. Additional storage for SQL Server Database backups.



Table 6 Test configuration: Software components

Test configuration	Software components
Operating systems	<p>Host: VMware vSphere ESXi Server v5</p> <p>Guest: Microsoft Windows Server 2008 R2 Enterprise Edition (virtual machine):</p> <ul style="list-style-type: none"> • MPIO enabled using EqualLogic DSM for Windows when using guest iSCSI initiator • EqualLogic Host Integration Toolkit(HIT) v4.6.0 installed
Applications	Microsoft SQL Server 2012 Enterprise Edition
Monitoring Tools	<p>Dell EqualLogic SAN Headquarters version 3.0</p> <p>Microsoft Windows Perfmon</p>
Simulation Tools	Quest Benchmark Factory for databases version 6.9.3



B Additional resources

Support.dell.com is focused on meeting your needs with proven services and support.

DellTechCenter.com is an IT Community where you can connect with Dell Customers and Dell employees for the purpose of sharing knowledge, best practices, and information about Dell products and installations.

Referenced or recommended Dell publications:

- Dell EqualLogic Configuration Guide:
<http://en.community.dell.com/dell-groups/dtcmedia/m/mediagallery/19852516/download.aspx>
- OLTP performance comparison between the new generation Dell EqualLogic PS6210XS hybrid array and the previous generation Dell EqualLogic PS6110XS hybrid array.
http://en.community.dell.com/techcenter/extras/m/white_papers/20438106/download.aspx

Referenced or recommended Microsoft publications:

- SQL Server 2012 Best Practices with PS6110XS
<http://en.community.dell.com/dell-groups/dtcmedia/m/mediagallery/20321740/download>
- OLTP I/O Profile study with Microsoft SQL 2012 and EqualLogic storage
<http://en.community.dell.com/dell-groups/dtcmedia/m/mediagallery/20308518/download.aspx>
- PS Series Groups Deploying Microsoft SQL Server in an iSCSI SAN
http://www.equallogic.com/uploadedfiles/Resources/Tech_Reports/tr1013-sql-server.pdf
- EqualLogic PS Series Architecture: Hybrid Array Load Balancer
<http://en.community.dell.com/dell-groups/dtcmedia/m/mediagallery/20349614/download.aspx>

