Protecting Microsoft SQL Server databases with EqualLogic Synchronous Replication

ABSTRACT

EqualLogic SyncRep allows users to protect their Microsoft SQL Server based mission-critical application data within a campus environment by simultaneously writing every data change to two different data volumes that may be located in two different building. This technical report provides a proof point for protecting SQL Server 2008 based OLTP application data in a campus using the SyncRep feature of the EqualLogic PS Series arrays from Dell.

October 2012

Preface

EqualLogic PS Series arrays are easy to use and they optimize resources and simplify operations by automating many performance enhancement and network load balancing tasks. Additionally, PS Series arrays offer all-inclusive array management software, host software, and firmware updates at no additional license fees.

Audience

The information in this guide is intended for system architects and administrators who wish to protect their Microsoft SQL Server 2008 based mission-critical application data using EqualLogic SyncRep.

Related Documentation

For detailed information about PS Series arrays, groups, volumes, array software, and host software, log in to the [Documentation page](#) at the customer support site.

Dell Online Services

You can learn about Dell products and services from the following web based resources:

1. Visit http://www.dell.com or the URL specified in any Dell product information.

2. Use the locale menu or click on the link that specifies your country or region.

Dell EqualLogic Storage Solutions

To learn more about Dell EqualLogic products and new releases being planned, visit the Dell EqualLogic TechCenter site: http://delltechcenter.com/page/EqualLogic. Here you can also find articles, demos, online discussions, technical documentation on infrastructure and application best practices, and more details about the benefits of our product family.

For an updated Dell EqualLogic compatibility list please visit the following URL: https://support.equallogic.com/compatibility

TABLE OF CONTENTS

REVISION INFORMATION

The following table describes the release history of this Technical Report.

Report	Date	Document Revision
1.0	October 2012	Initial Release

The following table shows the software and firmware used for the preparation of this Technical Report.

Vendor	Model	Software Revision
Dell^{TM}	Dell EqualLogic PS6100X	6.0.0
Dell^{TM}	Dell EqualLogic SAN Headquarters	2.2.11.6301
Dell^{TM}	Force 10 S-Series S60 Switches	OS Version: 1.0 Application Software Version: 8.3.3.7
Dell^{TM}	HIT Kit	V4.0
Microsoft®	SQL Server 2008, 64-bit	SP1
Microsoft®	Windows Enterprise Server 2008 R2, 64-bit	SP1
Microsoft®	iSCSI Initiator	v6.1.7601.17514
Microsoft®	Perfmon	6.1.7601
VMware®	VMware vSphere	5.0 Update 1
Quest Software®	Benchmark Factory for Databases	6.7

EXECUTIVE SUMMARY

Online Transaction Processing (OLTP) applications are by far some of the most mission critical applications in an enterprise. Protecting an organization's OLTP application data and making it highly available during planned system maintenance like software or hardware upgrades as well as in unplanned disasters like network or power outage and building fire or flood is of paramount importance. This technical report validates and documents an efficient disaster recovery (DR) and high availability (HA) solution for a high transaction volume Microsoft® SQL Server® based OLTP environment in an extended campus-wide datacenter. The solution uses the EqualLogic SyncRep feature available with PS series firmware v6.0 or above. SyncRep protects data by simultaneously writing volume data to two different storage pools in the same PS Series Group.

The primary objectives of this Technical Report are to:

- Document a tested and a validated SQL Server based application data protection solution using the EqualLogic SyncRep feature.
- Provide acceptable application and storage system performance test results with and without the SyncRep feature enabled in a SQL Server based OLTP environment under load in order to permit a relative comparison of the effect of enabling SyncRep on OLTP performance.
- Demonstrate the effectiveness of SyncRep during planned switchovers and unplanned failovers (as simulated in testing) in a SQL Server based environment.

INTRODUCTION

Microsoft SQL Server is the database of choice for mission critical OLTP applications in many small, mid-size and large enterprises. Downtime for such applications often means severe revenue loss due to interruption in business processes. Thus, it is extremely important that OLTP systems remain online all the time without any disruption and data loss. It is therefore becoming a common practice to design a high availability and disaster recovery solution for a mission critical SQL Server application environment. Storage based replication can provide an efficient and easy to use HA and DR solution to protect such mission critical OLTP application environments.

In addition to snapshots, clones and asynchronous replication, Dell EqualLogic also provides a synchronous replication solution to efficiently protect EqualLogic storage systems data without significantly impacting the application performance.

This technical report describes a tested and validated solution for designing a campus high availability and disaster recovery solution suitable for Microsoft SQL Server OLTP application deployments on EqualLogic PS Series using storage-based synchronous replication technology.

WHAT IS EQUALLOGIC SYNCHRONOUS REPLICATION (SYNCREP)?

Synchronous Replication (SyncRep) is a feature included in EqualLogic PS Series firmware v6.0 or above at no additional cost. SyncRep allows simultaneous volume data writing to two different storage pools in the same PS Series Group resulting in two hardware-independent copies of volume data. With SyncRep enabled, the PS series storage will only acknowledge write completion to the application when the write to

both the active and alternate volumes is hardened in the protected cache of the PS series members. If one pool is not available due to a power failure or other resource outage, volume data from the other pool can still be obtained. Listed below are some of the salient features/characteristics of SyncRep:

- SyncRep is the process of simultaneously copying all writes to a volume in a storage area network to a second volume so that an additional copy of data is available at all times.
- SyncRep writes data to the active and alternate sites synchronously before acknowledgement is sent back to the host (initiator) so that the data remains current at both locations.
- SyncRep is highly desirable for applications that demand a zero recovery point objective (RPO) – no data loss.
- In addition to SyncRep a high availability and disaster recovery server and application failover automated process is highly desirable for applications that demand a short recovery time objective (RTO).
- SyncRep provides Quick Recovery of data if primary volumes become unavailable.
- SyncRep is a firmware feature and is operating system, as well as application, agnostic.

For more information on SyncRep, refer to [TR1085](TR1085).

Requirements and considerations for using SyncRep

Requirements and considerations for using SyncRep in an EqualLogic SAN are listed below. Refer to the EqualLogic PS Series Storage Release Notes Version 6.0 for more information:

- Requires PS Series Array firmware v6.0.0 or higher.
- Requires an EqualLogic PS series group, containing two pools, each containing at least one member.
- Requires fast and reliable connectivity between pools.
- The same volume cannot be configured for both Auto-Replication and Synchronous Replication.
- Volume rethinning (unmap) is not recommended on volumes where SyncRep is used. *Please see the PS Series firmware v 6.0 release notes for more information.*
- Up to 32 SyncRep enabled volumes are supported in a PS6xx0 Series group (PS4xx0 only groups support up to 4 SyncRep volumes).
- Requires a minimum of 5% snapshot reserve.
- SyncRep cannot be enabled for volumes bound to single member arrays.
- Merge of members which have SyncRep enabled volumes is not allowed

SYNCREP TERMINOLOGY AND STATES

SyncRep feature is enabled on a per-volume basis. For SyncRep-enabled volumes, the volume data exists simultaneously in two different pools. Snapshots will exist in whichever pool was the active pool at the time the snapshot was created.

SyncActive – The pool to which iSCSI initiators connect when reading and writing volume data.

SyncAlternate – When volume data is written to the SyncActive pool, the group simultaneously writes the same data to this pool. The roles of the SyncActive and SyncAlternate pools can be switched. Host initiators do not make connections to this pool during normal operation.

Switchover – When the SyncActive and SyncAlternate pools are switched the former SyncActive pool then becomes the SyncAlternate pool and vice-versa. No iSCSI configuration changes are required. During the switch, host connections are temporarily logged out using standard iSCSI features. The iSCSI initiators on the hosts can reconnect automatically and transparently when the switch has completed. After the switch completes, the volume is automatically set online and synchronous replication resumes.

In sync - Both the SyncActive and the SyncAlternate pools contain the same volume data. Switching roles is allowed.

Paused - An administrator may pause and later resume SyncRep. While SyncRep is paused, the volume is still online, and initiators can connect to and write to the SyncActive copy of the volume. For example, this could happen in a maintenance window during which the SyncAlternate pool is taken offline. If data is written to the volume while synchronous replication is paused, it is written only to the SyncActive pool, and the two pools are out of sync, however, the group tracks all volume writes while synchronous replication is paused and when the administrator resumes synchronous replication the tracked changes are written to the SyncAlternate copy of the volume.

Out of sync – The SyncActive pool and SyncAlternate pool may not contain the same volume data; the SyncActive pool contains the most recent volume data. A SyncActive volume can become out of sync in following situations:

(1) if synchronous replication is paused manually,
(2) if the SyncAlternate pool becomes unavailable or has no free space,
(3) when the snapshot reserve in the SyncAlternate pool is full (but only when the snapshot space recovery policy sets volumes offline if the snapshot reserve is depleted), or
(4) if there is insufficient bandwidth between the SyncActive and SyncAlternate pools to keep up with the write demand on the volume.
(5) if there is a disaster or a fault in the SyncActive pool

Failover - Failover is a pool relationship change that is required because of a disaster or fault occurring in the SyncActive pool. In such situations, the SyncRep status automatically becomes out of sync and Failover to the SyncAlternate pool is used to bring the SyncAlternate copy of data online and can only be invoked when the SyncRep status is *out of sync*. This operation option is usually performed on the SyncAlternate pool when the SyncActive pool is unavailable.

TEST CONFIGURATION AND TOOLS

In order to test a typical synchronous replication scenario in a campus environment a set of volumes containing a SQL Server 2008 database was configured with SyncRep from one EqualLogic storage pool (Pool 1) to another pool (Pool 2) within the same group (group A) that was separated by a simulated distance of one kilometer. We used Quest Benchmark Factory for Databases[1] to simulate a TPC-C style workload on the test system using a SQL Server 2008 database.

Figure 1 illustrates the overall test configuration used for the benchmark testing. Multiple virtual machines (VMs) running Benchmark Factory for Databases software concurrently executed the TPC-C benchmark at increasing user loads. All transactions performed database access through one Microsoft SQL Server. Each Benchmark Factory client accessed its own database located on the PS series group consisting of two pools with either one or three EqualLogic PS6100X arrays in each pool, as described below.
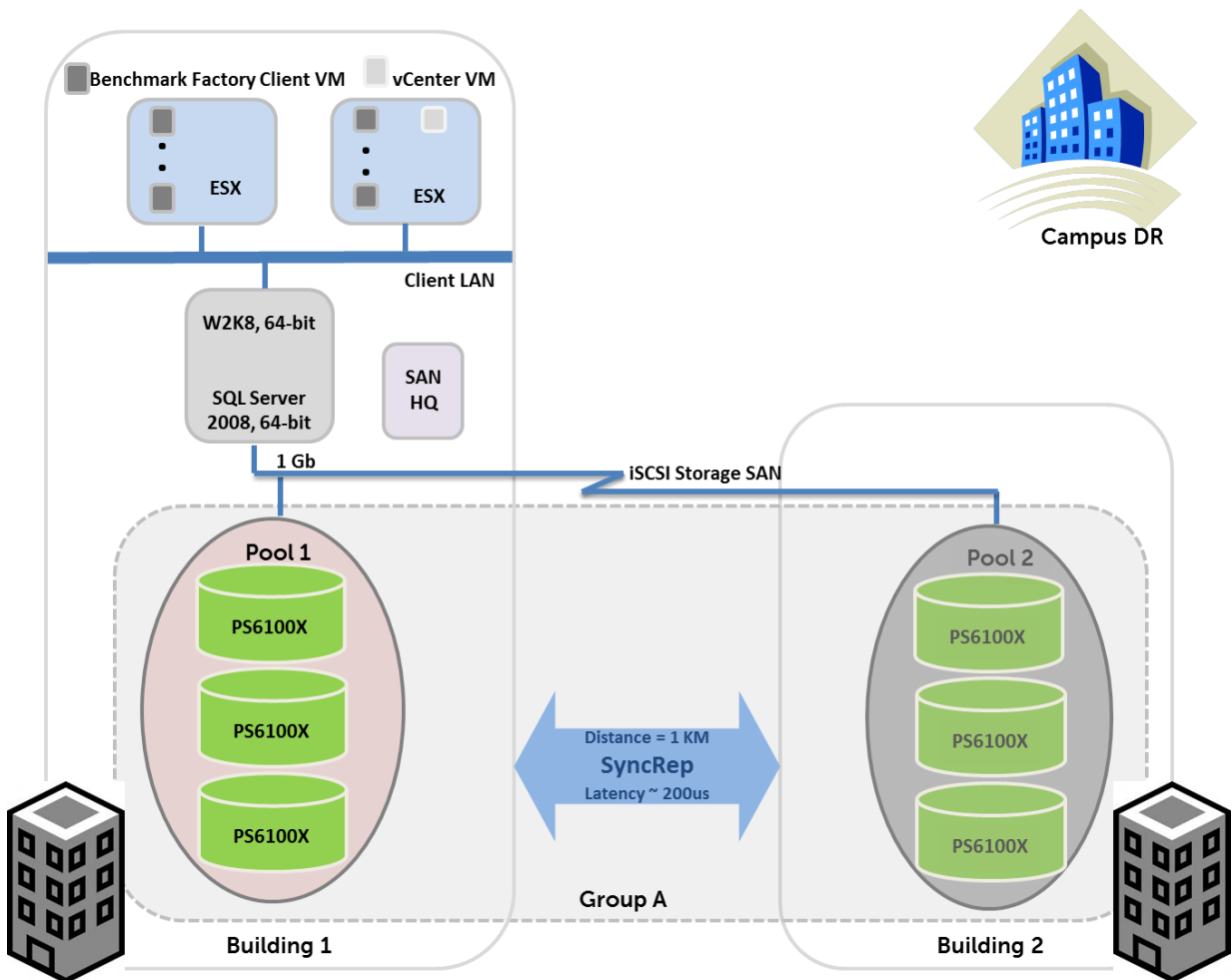

Figure 1: Illustration of a synchronously replicated campus site

---

[1] Quest Benchmark Factory for Databases is a load generator that can simulate OLTP workloads running on a database at various user load settings.

Test goals:

    A. Simulate a TPC-C style database transactional workload while gathering application transactions per second (TPS) and end-to-end transaction response time data under increasing user load on a single PS Series array
        a. Without the SyncRep feature enabled
        b. With the SyncRep feature enabled

    B. Measure any performance impact of synchronous replication on the production database supporting the TPC-C transaction load for a specific number of concurrent virtual users
        a. On a group with a single PS Series array in a storage pool
        b. On a group with three PS Series arrays in a storage pool

    C. Measure and compare the system performance when performing a planned switchover and an unplanned failover on the production database supporting TPC-C transaction load for a specific number of concurrent virtual users.

Please note that the goal of this white paper is to understand the impact of SyncRep on EqualLogic array performance in a SQL Server based OLTP environment rather than to establish the highest possible TPS that could be obtained in a SyncRep configuration. Consequently, in all the test results reported in this whitepaper the SQL Server was configured to use a maximum of 4 GB of cache, thereby forcing high I/O rates to the storage arrays even under low TPS load. The best way to interpret these results is not to focus on the absolute TPS results achieved, but to focus on the percent difference between TPS results achieved with and without SyncRep enabled.

BENCHMARK FACTORY TPC-C

TPC-C is an OLTP benchmark that simulates an order-entry environment where a number of concurrent end-users execute transactions against a database. The benchmark is comprised of transactions which include entering and delivering orders, recording payments, checking the status of orders, and monitoring the level of stock at the warehouses. The performance metric of this benchmark can be expressed in terms of transactions per second (TPS); defined as the number of orders that can be fully processed per second.

The testing employed Quest Software's Benchmark Factory® for Databases for the TPC-C benchmark load generation. One Benchmark Factory client was run against one database under test for a total of four clients running simultaneously against the four databases hosted by the EqualLogic group. A single ESXi 5.0 host (Dell™ PowerEdge™ R610 server with dual-socket hex-core Intel® Xeon® 5675, 3.07 GHz processors, and 72 GB memory) with four client VMs. Each Benchmark Factory client utilized a Windows® Server 2008 R2 64-bit VM.

A Benchmark Factory TPC-C schema was created and populated with approximately 200 GB of data in tables plus indexes, for each of the four databases. This ensured approximately 94% of the 520 GB allocated per database was in use and that almost 73% (2.87 TB) of a single PS6100X array's capacity was in actual use thus simulating a typical end-user database and array utilization scenario.

SQL SERVER

Microsoft SQL Server 2008 SP1, 64-bit, provided the database services for the TPC-C clients. The SQL Server was hosted on a Dell PowerEdge R610 server with dual-socket hex-core Intel Xeon X5675, 3.07 GHz processors and 72 GB memory. As mentioned in the callout text box on Page 5, the SQL Server was configured to use a maximum of 4GB of cache.

The SQL Server utilized the EqualLogic MPIO DSM for Windows from version 4.0 of the EqualLogic Host Integration Toolkit for Microsoft and was connected to the SAN via four 1 GbE NICs connected to a Dell™ Force 10™ S60 switch.

EQUALLOGIC STORAGE ARRAYS

The EqualLogic PS6100X arrays under test were connected to the SQL Server via Dell Force 10 S-Series S60 switches. This network was dedicated for the storage traffic. Each EqualLogic PS6100X array had twenty-four, 10,000 rpm, 300 GB SAS drives and was configured in RAID 10 for maximum possible OLTP database performance.

EQUALLOGIC SAN HEADQUARTERS

The I/O characteristics of the EqualLogic PS6100X arrays were monitored by the EqualLogic SAN Headquarters (SANHQ) software. SANHQ is a storage monitoring solution designed for and included with the Dell EqualLogic PS Series iSCSI SAN arrays as part of its all-inclusive software package. It delivers consolidated and centralized storage performance, capacity, and network monitoring. SANHQ can dramatically reduce the time required to identify performance bottlenecks, correlate them to events in the virtual and/or physical server infrastructure, and develop remediation strategies.

OVERALL TEST CONFIGURATION

Figure 2 shows the system test configuration used for the benchmark testing. Multiple VMs using Benchmark Factory for Databases software concurrently executed the TPC-C benchmark at increasing user loads. All transactions accessed the database through one Microsoft SQL Server. Each Benchmark Factory client accessed its own database located on the EqualLogic group.

The test configuration consisted of a single group "Group A" with two pools, "Pool 1" and "Pool 2". There were two array configurations that were tested:

1. For scenario A, we used one PS6100X array assigned to Pool 1 that was the SyncActive pool and another single PS6100X array assigned to Pool 2 that was the SyncAlternate pool.
2. For scenarios B and C, we used three PS6100X arrays assigned to Pool 1 that was the SyncActive pool and three more PS6100X arrays assigned to Pool 2 that was the SyncAlternate Pool.

In both cases, the arrays were connected to a pair of Dell Force 10 S60 switches. Pool 1 and Pool 2 were connected by a 1 kilometer single mode Fibre connection with approximately 200 microsecond latency to create a campus distance between the members in the SyncActive and SyncAlternate pools.
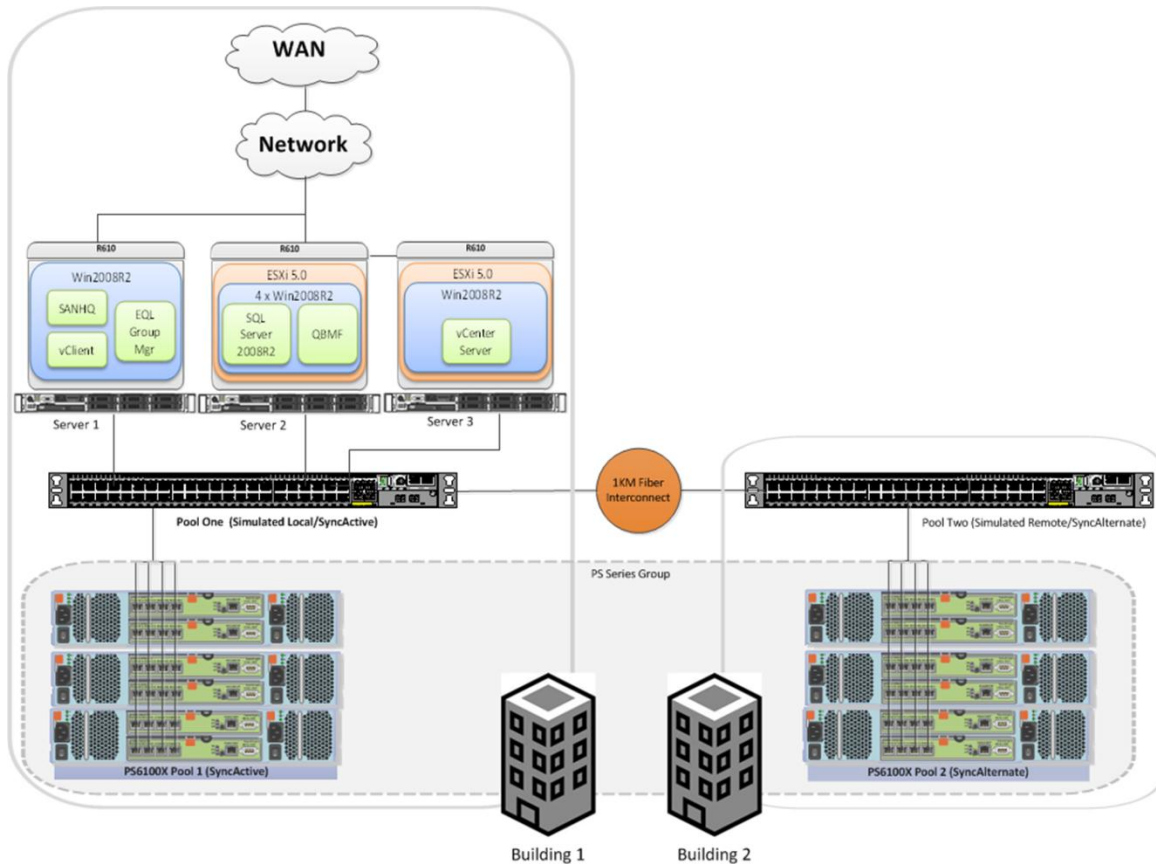
Figure 2: EqualLogic test system configuration for a synchronously replicated campus site

Note: This environment was setup just for the test lab purposes. Please refer to the [Reference Architectures](#) *documents for a complete understanding of how to set up a highly-available configuration.*

Description of test system configuration:

- SQL Server 2008 Enterprise Edition was installed in a Windows Enterprise Server 2008 R2 virtual machine hosted on a Dell PowerEdge R610 server running VMware vSphere® ESXi5.0 with default settings.
- The virtual machine that hosted the SQL Server was configured to use a maximum of 4 GB of cache for SQL Server application, thereby forcing more I/O to the storage arrays as noted in the callout text box on Page 5.
- Since the SyncRep implementation for the EqualLogic PS series uses two pools in the same PS Series group all members were in the same VLAN/subnet (a "flat network" is required).
- The SyncActive and SyncAlternate pools were configured with Force 10 S60 switches and were not shared with Client LAN traffic. Buffers and Flow control was configured according to published reference architectures.
  For more information please see Dell TechCenter: [Reference Architectures](#)
- A 1 kilometer single mode Fibre connection was installed in between the two Force 10 S60 switches.

To measure the performance of consolidated OLTP applications, sixteen 85 GB volumes and eight 20 GB volumes were created. Four databases were then hosted on

7

this group. Each database was assigned eight volumes. Of the eight volumes, six were data volumes, one was a log volume, and one was a tempdb volume. All eight volumes for a single database were part of a single PS series volume collection.[2] Volume collection is useful when an operation needs to be performed simultaneously on multiple volumes. This feature helps maintain consistency in SQL Server environments because SQL database files are typically spread across multiple volumes. The TPC-C benchmark test was then run on these databases.

Since there were four databases, there were a total of four volume collections each with eight volumes that were part of a single database. Four SQL Server databases were created, populated, and split into multi part MDF's across these volumes in the following manner:

Database	Volume Name	Size (GB)
QBMFDB01	QBMF01-Vol01-Data01	85
	QBMF01-Vol01-Data02	85
	QBMF01-Vol01-Data03	85
	QBMF01-Vol01-Data04	85
	QBMF01-Vol01-Data05	85
	QBMF01-Vol01-Data06	85
	QBMF01-Vol01-TLog	20
	QBMF01-Vol01-TmpDB	20

Table 1: Database layout for SQL Server stored on PS6100X arrays

Similar volumes were created on QBMFDB02, QBMFDB03, and QBMFDB04 to bring to the total number of volumes to 32.

---

[2] Synchronous replication (SyncRep) volume collections are a unique type of volume collection that can only contain synchronous replication-enabled volumes. Like other volume collections, you can clone or take snapshots of synchronous replication collections. However, synchronous replication collections have the additional capability of supporting synchronous replication operations. In synchronous replication volume collections, all volumes in the collection behave as a single object. The volumes will simultaneously reside in the same SyncActive and SyncAlternate pools. When synchronous replication switchovers are performed on the collection, they occur for all volumes in the collection. Synchronous replication switchovers cannot be performed on individual volumes in the collection. Synchronous replication collections are only considered to be in sync when all volumes are in sync, and go out of sync if a single volume within the collection ceases to be in sync. This is particularly important in a SQL Server database environment so that data consistency is maintained on all volumes that are part of a single database.

The following three scenarios were tested using this configuration:
  A. For scenario A, a TPC-C style database transactional workload was generated while gathering transaction per second (TPS) and end-to-end transaction response time data under increasing user load on a single PS Series array. Four TPC-C benchmark clients simultaneously ran against each of these four databases. The workload was increased until the log latency exceeded 5 milliseconds. Next, the benchmark was repeated for databases hosted on the same configuration with SyncRep enabled.
  B. For scenario B, tests were run to measure the performance impact of synchronous replication on the production database supporting the TPC-C transaction load for a specific number of concurrent virtual users on a group with single PS Series array in a storage pool replicating to another pool with one member. This test was then repeated to measure the performance impact on a group with three PS Series arrays in a storage pool replicating to another storage pool of three members within the same group.
  C. Finally, for scenario C, tests were also performed to measure the performance impact on the production database volumes supporting the TPC-C transaction load during switchover and failover operations. These tests were conducted for a specific number of concurrent virtual users.

The analysis compared the total number of concurrent users supported, transactions completed and IOPS delivered at the same transaction response time to determine the performance impact when the data is synchronously replicated between two EqualLogic pools. Sufficient server and network resources were available during all the tests.

TEST RESULTS

During all the test runs the default settings of the TPC-C benchmark, including transaction weightings and think times, were used.

A. Application Transactions/Second and Application Response time

The TPC-C benchmark was run against four databases hosted on a PS series group consisting of a single EqualLogic PS6100X array (Scenario A). Aggregate number of users was stepped from 100 to 3,000 in increments of 100. Each step was run for 30 minutes. Figure 3 illustrates the measured relationship of aggregate number of users and the corresponding transactions per second (TPS) and average transaction response time.

For the tests with synchronous replication enabled, the group consisted of two PS6100X arrays. However, the benchmark test was run on a single PS6100X array and all writes were replicated using the SyncRep feature to the second PS6100X array,

which was in a different storage pool but within the same PS Series group. (Scenario A).
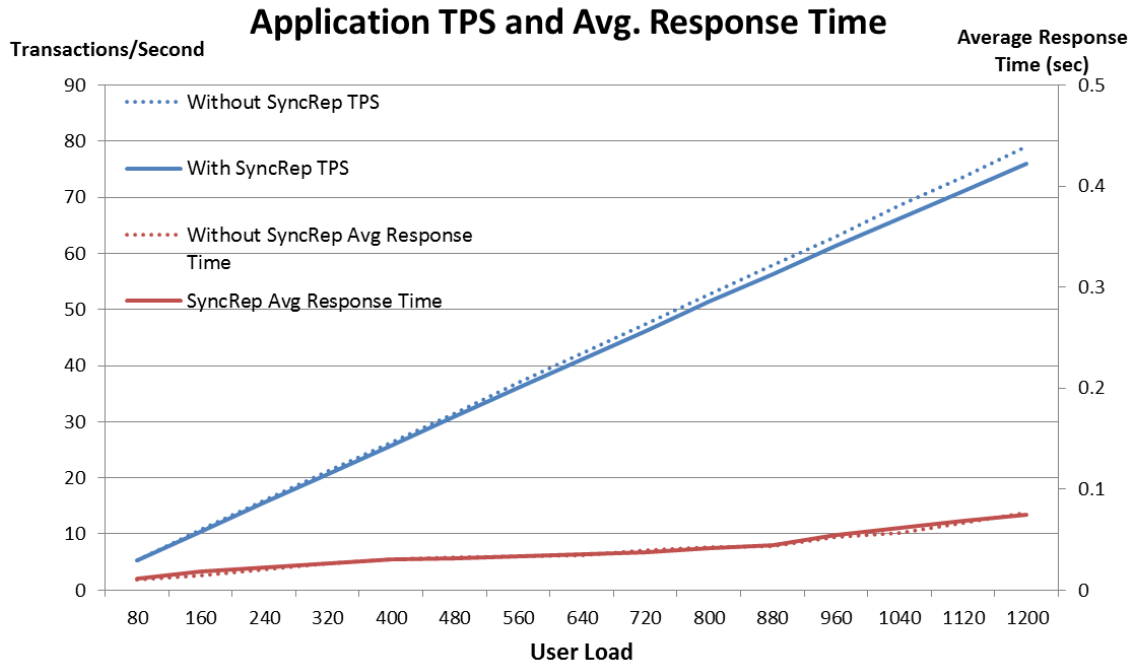

Figure 3: Application TPS and Average transaction response time for a single member pool

In Figure 3, the dotted lines represent results from test runs without SyncRep functionality enabled. The solids lines represent results from SyncRep test runs. As seen in the chart, there was negligible impact on the application's average transactions per second or the average transaction response time when the same database volumes were being replicated to Pool 2, which was separated from Pool 1 by a 1 kilometer fiber optic link.

Table 2 provides a snapshot of the test results that are shown in figures 3, 4, and 5 for a specific user load:

	Performance Results					
	TPS	Transaction Response Time (sec)	Storage Read I/O Latency (ms)	Storage Write I/O Latency (ms)	Storage Read IOPs	Storage Write IOPs
Without SyncRep	79	0.077	~10	~8	267	493
With SyncRep	76	0.075	~10	~16	304	974

Table 2: Summary of TPS, storage IOPs and storage latency for a specific virtual user workload

As mentioned in the text box on Page 5, the best way to interpret the results in Figure 3 and Table 2 is not to focus on the absolute TPS results achieved, but to focus on the percent difference between TPS results achieved with and without SyncRep.

Figure 4 and Figure 5 shows the SAN HQ charts for the above tests with the highest user workload that is summarized in Table 2 above. Figure 4 is without SyncRep enabled and Figure 5 with SyncRep enabled. Note that the storage write IOPs in the

SyncRep test run doubled from that of the test run without SyncRep enabled. This is due to the fact that with SyncRep enabled, the writes that are written to the volume in SyncActive pool are simultaneously written to the volume in the SyncAlternate pool. Also, note that the write latency doubles from test 1 to test 2. However, there is negligible impact on the application TPS and the application average response time as seen in Figure 3 above.

Note that the read/write percentage differs by small margin from one test to another due to the pseudo-random nature of the benchmark that was run.
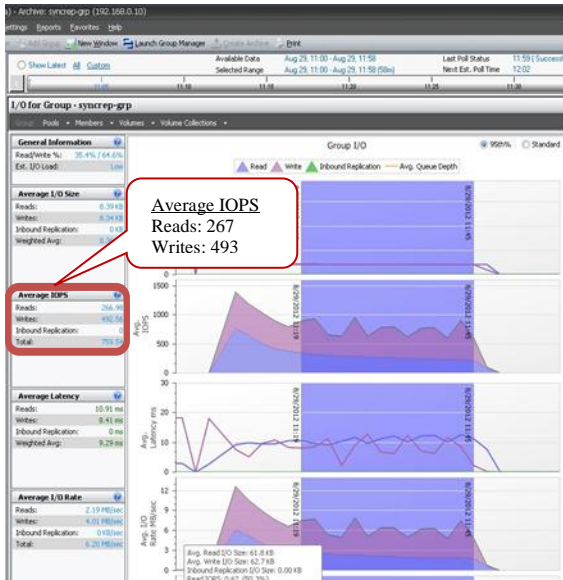

Figure 4: SANHQ metrics for 1 PS6100X at 1200 user load <u>without SyncRep</u> enabled
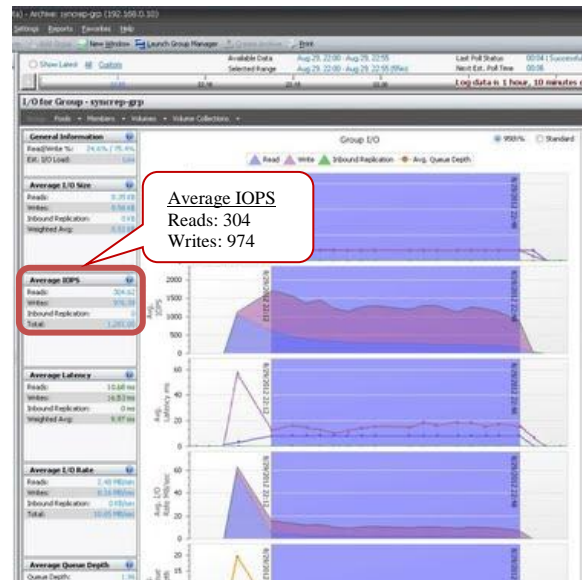
Figure 5: SANHQ metrics for 1 PS6100X at 1200 user load <u>with SyncRep</u> enabled

B. SAN Scaling

Tests were performed to measure performance impact of synchronous replication on the production database supporting the TPC-C transaction load for a specific number of concurrent virtual users. The first test was run on a group with a single PS Series array in a storage pool replicating using SyncRep to another pool with one member (Scenario A). This test was then repeated to measure the performance impact on a group with three PS Series arrays in a storage pool replicating using SyncRep to another storage pool of three members within the same group (Scenario B).

In these tests we measured the scalability of I/O performance as the number of EqualLogic PS6100X arrays (or members) increased within the group while SyncRep was enabled.

The test demonstrates the same phenomenon as seen in the test using single arrays: When SyncRep is enabled on a volume, the system will simultaneously write volume data across two different storage pools in the same PS Series Group, resulting in twice the write IOPs on the system for the same workload ensuring two hardware

independent copies of the same data. Since the write IOPs double, this will result in increasing the write latency on the storage system.

The configuration parameters for these tests are shown in Table 3:

Configuration Parameters	
EqualLogic SAN	6xPS6100X (2.5", 24 10K SAS drives,300 GB)
Array configuration#1	2xPS6100X (two EqualLogic storage Pools with one member in each Pool) A 1200 virtual user workload was run against a single member that was part of the SyncActive Pool. The second pool with one member was the SyncAlternate pool and was used just for replication.
Array configuration#2	6xPS6100X (two EqualLogic storage Pools with three members in each Pool) A 4500 virtual user workload was run against three members that were part of the SyncActive Pool. The second pool, also with three members, was the SyncAlternate pool and was used just for replication.
Volume Configuration	32 volumes (8 volumes per database, 4 databases in total running OLTP workloads)
RAID Policy	RAID 10

Table 3: Configuration parameters for SAN Scaling tests

The results collected from the SAN scaling studies are reported in Figure 6 and Figure 7, below. Figure 6 shows relative storage read and write IOPs performance for 1 PS6100X array and 3 PS6100X array configurations. Figure 7 shows the corresponding transactions per second for the same configurations.

For these tests, a typical OLTP data workload running two types of virtual users (1x users on the configuration with 1x PS6100X and 3x users on the configuration with 3x PS6100X) were executed to get the read and write IOPs with configurations consisting of one and three PS6100X arrays respectively. The IOPs numbers were obtained while staying within the generally accepted disk latency limit of 20ms for data volumes (for both read and write latencies, measured separately) and 5ms for log volumes.
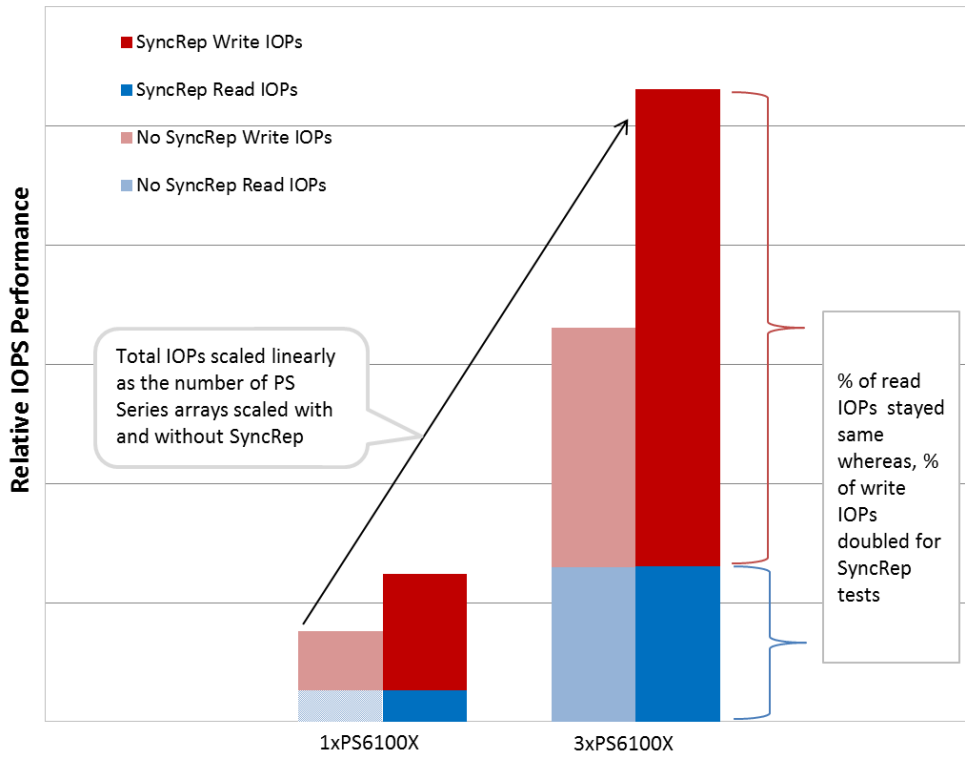
Total IOPs scale as the Members scale

- SyncRep Write IOPs
- SyncRep Read IOPs
- No SyncRep Write IOPs
- No SyncRep Read IOPs

Relative IOPS Performance

Total IOPs scaled linearly as the number of PS Series arrays scaled with and without SyncRep

% of read IOPs stayed same whereas, % of write IOPs doubled for SyncRep tests

1xPS6100X   3xPS6100X

Figure 6: Read IOPs stay constant while write IOPs double with SyncRep on single array as well on a three array configuration.
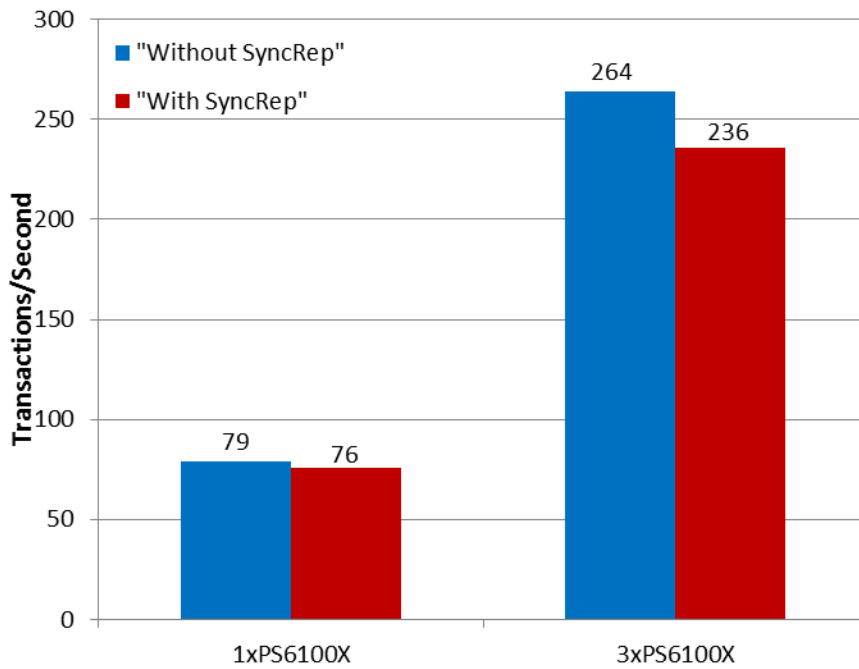
Figure 7: Application TPS with and without SyncRep feature enabled

Once again, as mentioned in the text box on Page 5, the best way to interpret the results in Figures 6 and 7 is not to focus on the absolute TPS results achieved, but to focus on the percent difference between TPS results achieved with and without SyncRep.

Key observations from these results can be summarized as follows:
1. IOPS scaled linearly with the addition of more arrays. This is as expected because the EqualLogic Peer Storage architecture scales the available storage resources linearly to provide increased IOPs with increasing resources.
2. For the SyncRep test, the storage write IOPs double as viewed from a whole group perspective when compared to the write IOPs of the test without SyncRep for both the single array and the three arrays tests. This is due to the fact that all writes made to the active volume in the SyncActive pool are replicated to the alternate volume in the SyncAlternate pool.
3. The application transactions per second (TPS) drops slightly for the SyncRep tests when compared to the equivalent test without SyncRep enabled. There was negligible impact on application response time as seen in Figure 3 above.

C. Changing Pool Relationships

Tests were also performed to measure and compare the system performance during both switchover and failover operations on the production database supporting the TPC-C transaction load for a specific number of concurrent virtual users.

Note that there are two operations that can change the iSCSI initiator access from the SyncActive to SyncAlternate pool – switchover and failover.

Switchover – This operation is typically used for *planned maintenance events*. The volumes in the SyncActive and SyncAlternate pools must be in an in-sync status in order to initiate a switchover.

Failover – In order to avoid data loss, this operation should be initiated only when *a fault has occurred in the pool with the SyncActive volumes*. In such situation, the SyncRep status automatically becomes out of sync and Failover to SyncAlternate pool is used to bring the SyncAlternate copy of the data online. This can only be invoked when the SyncRep status is out of sync. This operation is usually performed on the SyncAlternate pool when the SyncActive pool is unavailable.

Note that once the volume is out-of-sync, the option to switchover the volume is replaced with an option to failover the volume. Failovers can only be performed if the volume is offline. In the case of a failover, once the primary pool has recovered from the failure and the volumes are back in sync, the pool relationship can be changed to switch back to the original SyncActive pool.

Both Switchover and Failover operations will force the host iSCSI initiator to reconnect to the volume through an async logout due to the use of the same IP (the group IP) and volume IQN string to connect to then volume in the SyncAlternate pool.

Switchover

A pool switch for a SyncRep volume swaps the roles of the SyncActive and SyncAlternate pools. As part of the operation, the SyncActive pool becomes the SyncAlternate pool and the SyncAlternate pool becomes the SyncActive pool. When a pool switch takes place, iSCSI initiators are temporarily logged out of the volume. They can log back in again after the switch is complete. Depending on the initiator configuration, the initiators normally will automatically attempt to reconnect. After the switchover completes, the volume is automatically set online and synchronous replication resumes.

Typically, the administrator will choose to switch the synchronous replication pool configuration if a failure is imminent in the SyncActive pool, or if there is maintenance that needs to be performed on the array hardware in the SyncActive pool. The roles of the pools may be switched at any time, even if a failure has occurred, provided that the volume is in sync.

In our tests we performed a switchover on all 32 volumes spread across 3x PS6100X members while we were running the TPC-C benchmark on the four SQL Server databases.

The following steps were taken to switchover the four volume collections that contained the 32 volumes.

1. To initiate a switchover for a volume collection, select the collection in the left pane and click on "Switch collection to SyncAlternate" in the activities pane.
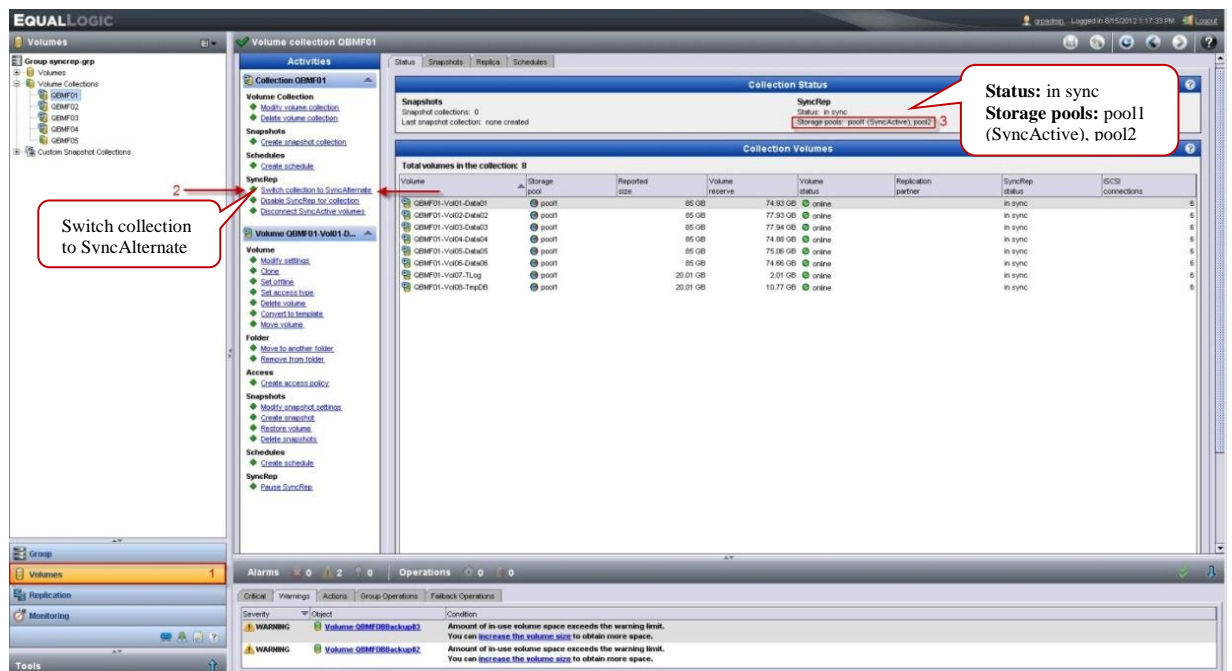

Figure 8: PS Series Group Manager GUI showing SyncRep user option for Switchover

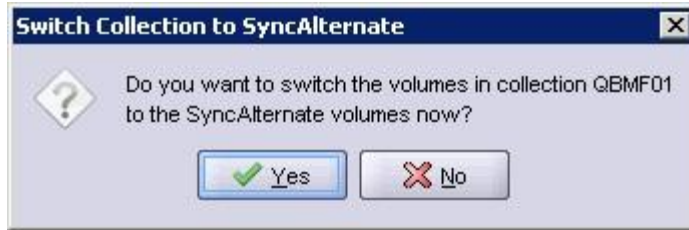2. Once you initiate the switchover, you will be prompted to confirm that activity, click "yes" to continue.

Figure 9: Switchover warning message

The switch was performed with four Quest Benchmark Factory instances running a TPC-C OLTP workload against four SQL databases and completed almost instantaneously with no noticeable impact on the operations. There were no errors encountered during the switchover and the transactions that were taking place during the switchover completed successfully.

The following is a screenshot of Quest Benchmark factory during the switchover
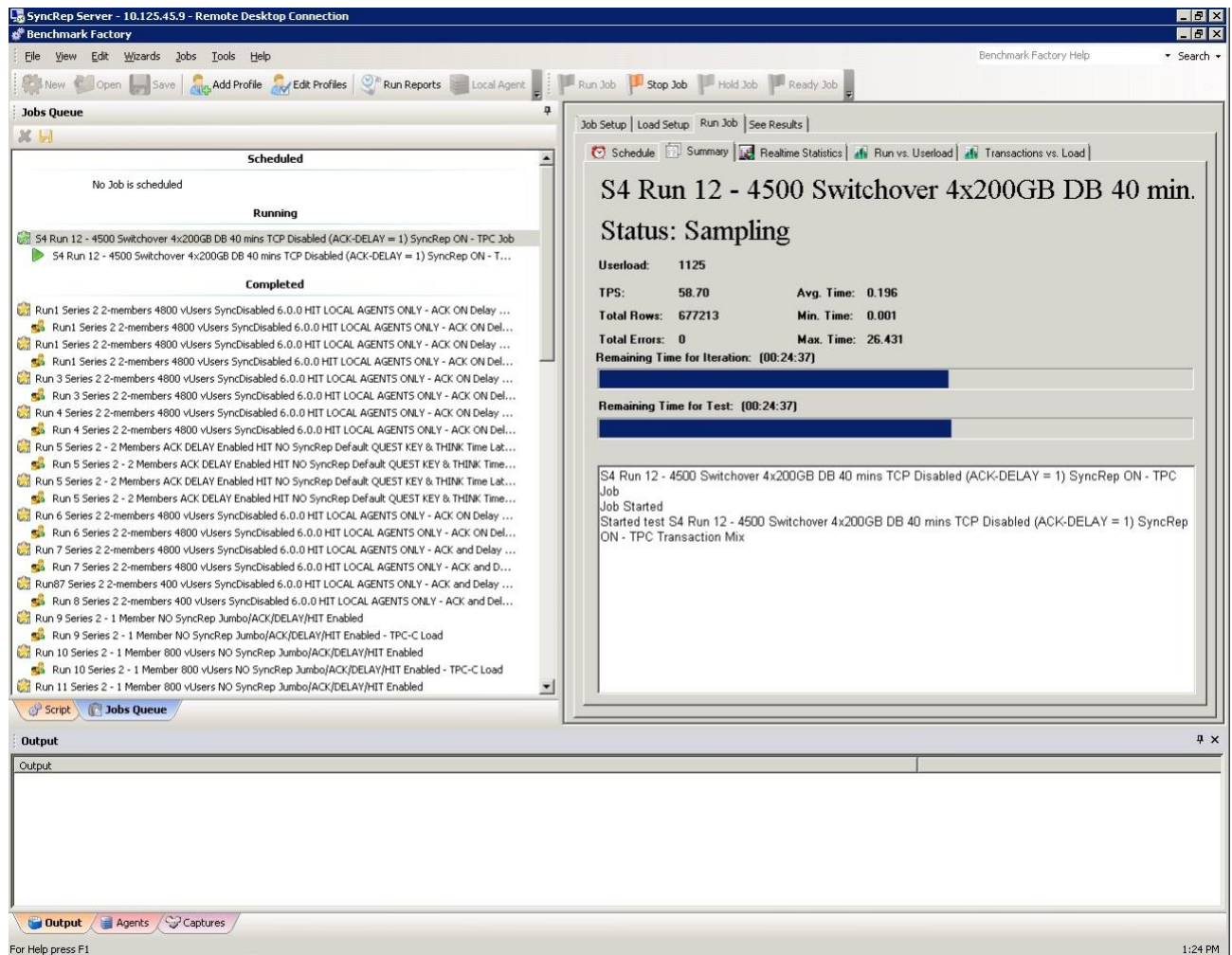

Figure 10: Quest Benchmark factory during the switchover of a volume collection

Once the switchover is complete, you can examine the properties of the volumes and verify that Pool 1, which was initially SyncActive, is now SyncAlternate Pool.
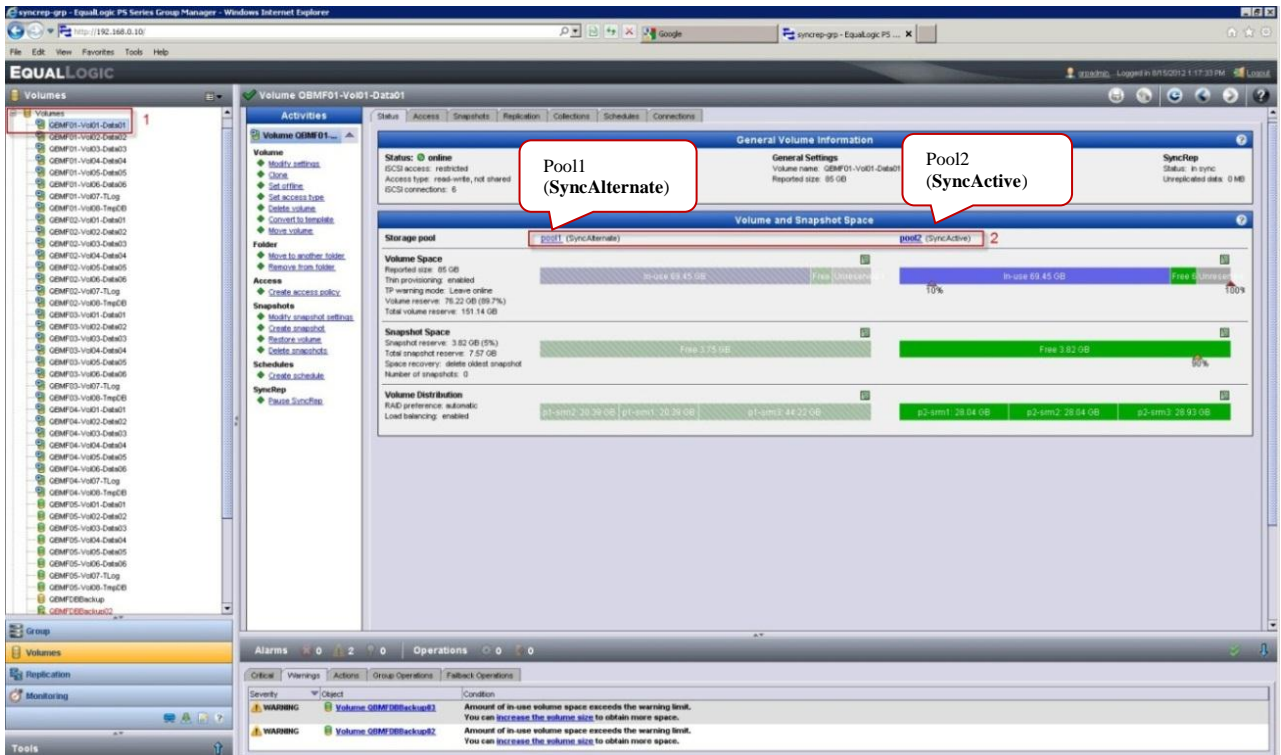
Figure 11: PS Series Group Manager GUI after switchover completes

Table 4 summarizes the test results that were captured during the switchover tests:

	End-User Experience Attributes			TPC-C Benchmark Performance Results		
	Transaction Response Time	Storage Read or Write I/O Latency	Concurrent Users	Log latency	TPC-C TPS	TPC-C IOPS
SyncRep	< 0.30 s	15 ms	4500	≤ 5ms	236	5,226
SyncRep during Switchover	< 0.21 s	12 ms	4500	≤ 5ms	237	5,108

Table 4: Summary of Switchover test results

Figure 12 and Figure 13 show the SAN HQ charts for the same tests. Identical number of virtual users were running concurrently on Pool 1 with three PS6100X arrays that were replicating all 32 volumes to Pool 2 with three additional PS6100X arrays during normal SyncRep and during Switchover operation. Figure 12 shows the performance of the arrays under normal operations whereas; Figure 13 shows the performance of the arrays during a switchover operation. Note that there is negligible impact on TPC-C IOPs and storage latency performance during the switchover operation.
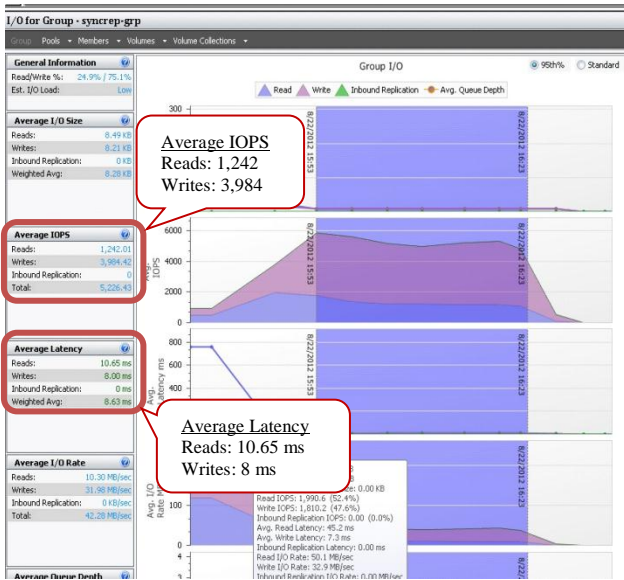
Figure 12: SANHQ metrics for 3 PS6100X at 4500 user load with SyncRep enabled
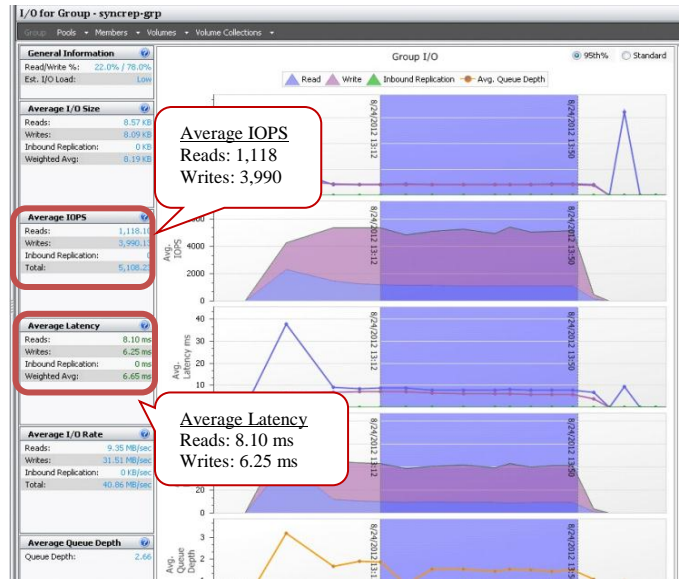
Figure 13: SANHQ metrics for 3 PS6100X at 4500 user load with SyncRep enabled during switchover operation

Failover

Failover is a pool relationship change that is required because of a fault occurring in the SyncActive pool. The "Failover to SyncAlternate" option is used to bring the SyncAlternate copy of data online,

The SyncRep Failover is a manual process. Please refer to TR1085 for detailed step-by-step instructions for performing a manual failover.

Note: If you failover to the SyncAlternate while the volume is out of sync, any changes that were written to the SyncActive volume since the volumes went out of sync will be lost. The state of the SyncActive volume at the time of failover will be saved in a snapshot in the SyncActive pool but that snapshot may be deleted due to policy and cannot be guaranteed to be there. Hence, failovers should only be performed under extraordinary circumstances (i.e., when old data is better than not having access to the latest data) or in cases where you are sure that the failover will destroy not any necessary data.

Once the volumes have been failed over, set them back online, which will restore initiator access. The new SyncActive volume (formerly SyncAlternate) accepts, writes and tracks changes made while the original SyncActive volume is offline. Depending on the host configuration and the application requirements, it may be necessary to restart hosts or applications after a failover.

Figure 14 shows the screenshots of the iSCSI Initiator and MSSQL Server Management Studio shows that the volumes and the database were back online and I/O to the volumes resumed.
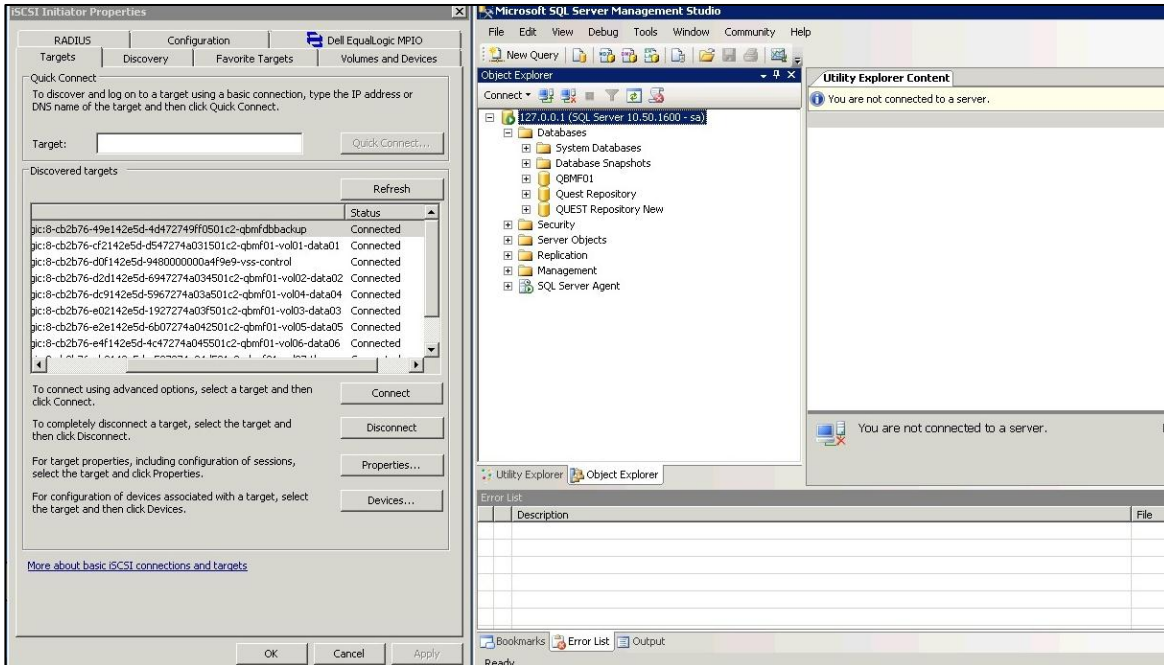
18

Figure 14: iSCSI Initiator and MSSQL Server Management Studio after failover

CONCLUSION

Many OLTP applications in real production environments also exhibit characteristics similar to the TPC-C benchmark workload tested in the simulated environment at Dell Labs. Protecting such SQL Server environments without compromising application performance is very desirable however, it can be difficult to achieve. Nevertheless, in this whitepaper we have shown that a SQL Server based OLTP application environment on EqualLogic PS Series PS6100X arrays with SyncRep can achieve an outstanding level of availability and protection of transaction data with no loss and without impacting the application performance.

Technical Support and Customer Service

Dell support service is available to answer your questions about PS Series SAN arrays.

Contacting Dell

1. If you have an Express Service Code, have it ready.
   The code helps the Dell automated support telephone system direct your call more efficiently.

2. If you are a customer in the United States or Canada in need of technical support,
   call 1-800-945-3355. If not, go to Step 3.

3. Visit support.equallogic.com.

4. Log in, or click "Create Account" to request a new support account.

5. At the top right, click "Contact Us," and call the phone number or select the link for the type of support you need.

Warranty Information

The MODEL array warranty is included in the shipping box. For information about registering a warranty, visit http://support.dell.com/EqualLogic.