Dell EqualLogic Best Practices Series

Best Practices for Deploying a Mixed 1Gb/10Gb Ethernet SAN using Dell EqualLogic Storage Arrays

A Dell Technical Whitepaper

Jerry Daugherty

Storage Infrastructure and Solutions
Dell Product Group
March 2011

Table of Contents

1 Introduction

The goal of this paper is to present network design recommendations and best practices for integration of 1GbE and 10GbE iSCSI SANs running Dell™ EqualLogic™ storage arrays.

In 2010 Dell introduced the EqualLogic PS6010/PS6510 series iSCSI storage arrays. These are the first EqualLogic PS Series storage arrays to be equipped with 10 Gigabit Ethernet (10GbE) controllers. By combining EqualLogic 10GbE storage arrays with industry standard Ethernet switches and the latest 10Gb Ethernet NIC options available with Dell™ PowerEdge™ servers, Dell customers now have the ability to build end-to-end 10GbE high-performance iSCSI Storage Area Network (SAN) infrastructures.

Dell recognizes that in many situations you will not be able to build stand-alone 10GbE SANs from scratch. Instead, you may need to take an incremental approach toward integration of 10GbE components into existing 1GbE SAN environments. There are many reasons that may lead to this approach. For example:

- You may be adding new EqualLogic PS6010/PS6510 series arrays equipped with 10GbE controllers into existing 1GbE EqualLogic SANs.
- You may be adding new servers that are equipped with 10Gb NICs, to be used in converged networking projects.
- Your application I/O throughput requirements are reaching the limits of a 1GbE environment.

1.1 Audience

This paper is intended for Solution Architects, Storage Network Engineers, System Administrators, and IT Managers. We assume the reader has working knowledge of network switch deployment and operation, network administration and EqualLogic-based iSCSI SAN operation.

1.2 The rest of this paper

The rest of this paper contains the following sections:

2 Integration strategies

Before we begin describing integration strategies in detail we first want to define the fully redundant SAN designs that we will use as building blocks for integration and testing. Figure 1 shows a 1GbE redundant SAN design using a pair of stacked Dell™ PowerConnect™ 6248 switches connecting to one or more EqualLogic PS6000 series 1GbE storage arrays. Figure 2 shows a similarly designed 10GbE SAN using a pair of PowerConnect 8024F switches connecting to one or more EqualLogic PS6010 series 10GbE storage arrays.

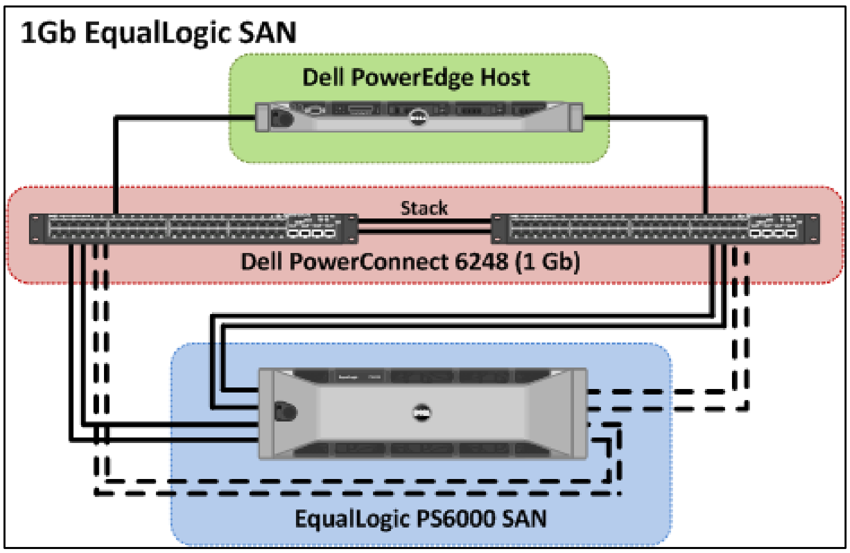


Figure 1 Baseline design: 1GbE SAN

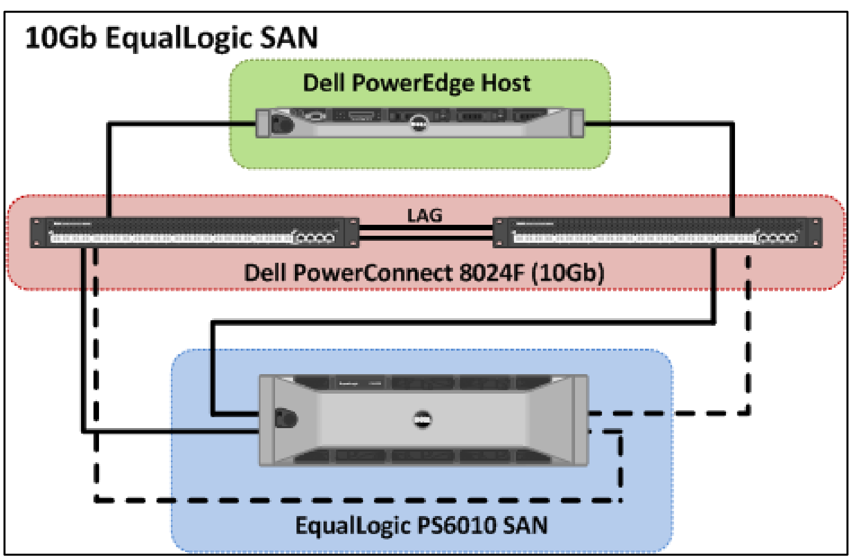


Figure 2 Baseline design: 10GbE SAN

The key differences between the 1GbE and 10GbE SAN designs shown in Figure 1 and Figure 2 are:

- The PowerConnect 8024F switches do not provide a stacking option. Instead of a creating a stack, the PC8024F switches are connected using a link aggregation group (LAG) comprised of one or more 10GbE links.
- The EqualLogic PS6000 series is equipped with two 4-port 1GbE controllers. The EqualLogic PS6010 series is equipped with two 2-port 10GbE controllers.

2.1 Design considerations for a mixed speed EqualLogic SAN

To design mixed speed EqualLogic SANs, there are two key integration questions that you must focus on:

- *How do you properly design a mixed speed switched Ethernet network fabric?*
  As you probably know already, proper design and configuration of the Ethernet switching fabric for 1GbE SANs is critical to achieving and maintaining performance and reliability requirements. The same is true for 10GbE SANs. When you begin to integrate 10GbE components into an existing SAN consisting of 1GbE arrays you must first carefully consider how you design the integration points for the 10GbE switches. This design decision can be critical to the performance and reliability of the integrated SAN.

- *How should you configure storage pools (array membership) and volume layout within those pools when combining 1Gb and 10Gb arrays together in the same SAN?*
  EqualLogic's Peer Storage functionality allows you to combine multiple arrays (or "members") into a single logical storage pool. Physical storage of volumes is automatically distributed across the pool members. This behavior means that you will also have to carefully consider pool membership and volume layout when mixing 1GbE and 10GbE storage arrays together in the same SAN.

Note: For additional information, see the following documents:

> - *Integrating EqualLogic PS6x10 Arrays with Existing SANs*, available at: http://www.equallogic.com/resourcecenter/assetview.aspx?id=9447
> - *Dell EqualLogic Configuration Guide*, available at: http://www.delltechcenter.com/page/EqualLogic+Configuration+Guide

There are many ways to design an iSCSI SAN that integrates 1GbE and 10GbE networks together. A significant factor affecting the design of a mixed speed iSCSI SAN is the feature set and port densities provided by the Ethernet switches you will be using in your SAN. To create a redundant, mixed speed iSCSI SAN, at a minimum we recommend that you start with dual 1GbE and dual 10GbE switches.

Figure 3 shows a design pattern for combining 1Gb and 10Gb EqualLogic SANs. The interconnect between the mixed speed components in the SAN is accomplished using 10GbE LAG groups as uplink connections between the 1Gb switches and the 10Gb switches.

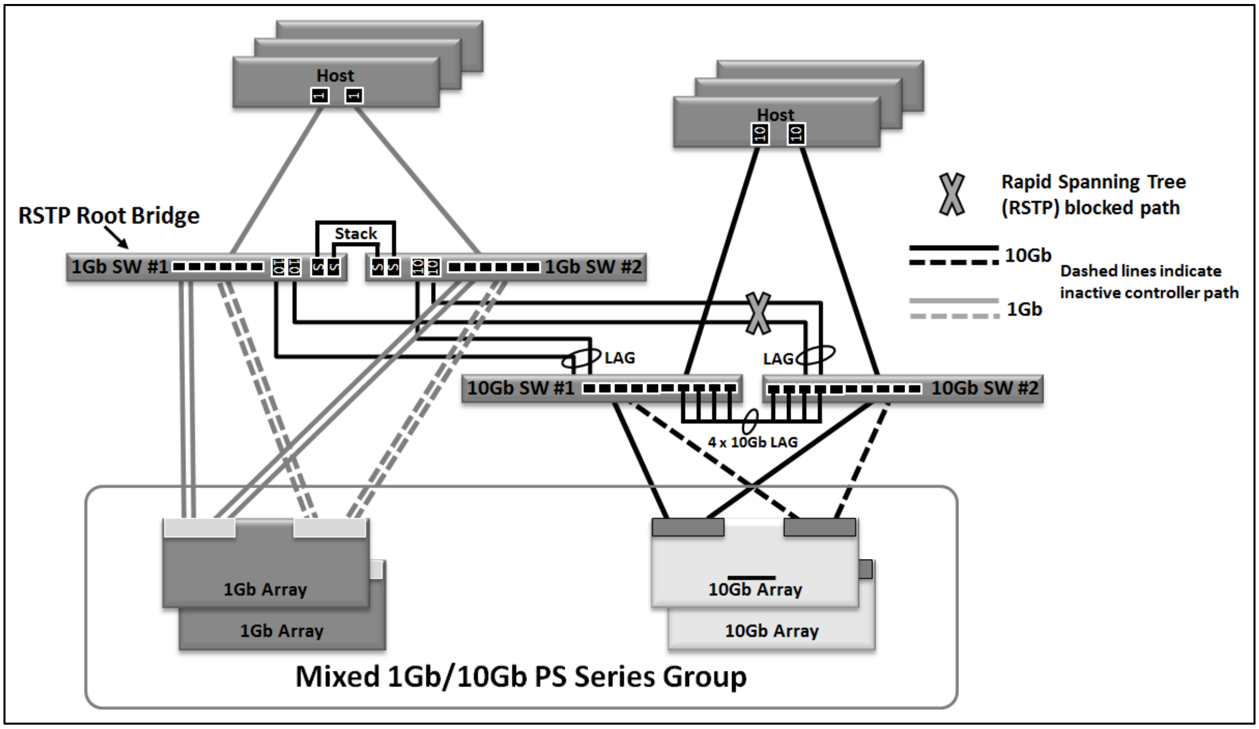


Figure 3 Mixed speed EqualLogic SAN

Referring to Figure 3, there are some important design considerations to be aware of:

- Each of the 1Gb switches is configured with one dual-port 10GbE uplink module and one stacking module. The 10GbE uplink modules are used for creating 20Gb LAG uplinks to the 10Gb switches.
- We assume the 10Gb switches are not stackable. Instead, a LAG connection is created between them.
- The 20Gb LAG uplinks between the 1Gb and 10Gb switches are cross-connected so that each 10Gb switch physically connects to both switches in the 1Gb stack.
- The 10Gb switches are connected together using multiple 10Gb links. This provides us a path for devices connected directly to the 10Gb switches to communicate without having to go through the slower 1Gb switches.

The LAG between the 10Gb switches creates a loop in the network, Rapid Spanning Tree Protocol (RSTP) will compensate for this by blocking paths as necessary. You should be aware of which switch is the root bridge in the spanning tree. For the mixed speed SAN design shown in Figure 3 the root bridge is **1Gb SW#1**. Based on this information you can assign costs to links to ensure that the desired link configuration is achieved. For the network design in Figure 3 we want to prevent RSTP blocking of the inter-switch trunk between the 10Gb switches, thus causing some 10Gb traffic to traverse the slower 1Gb switches. We manually assigned a high link cost to one of the 20Gb uplink LAGs to insure this doesn't happen. Note the location of the RSTP path in Figure 3.

You may be wondering "why not just attach my 10Gb arrays directly to the 10Gb uplink ports on the PowerConnect 6248 switches?" While this looks like a good idea for small installations, for performance and reliability considerations we recommend using this connection strategy for lab or test environments only. One of the primary reasons for this is that the buffering available on most 1 Gb switches was never designed to have utilization devices connecting to these ports.

Scalability is a second reason for not directly attaching 10Gb arrays to those ports. In the case of the two-member stack shown in Figure 3, you can only attach one 10Gb PS6010 array into the environment. This is because two active 10Gb ports and 2 passive 10Gb ports will exhaust the available uplink ports. What happens when you will need to add more 10Gb arrays? In this scenario, doing so may require shutdown of mission critical storage infrastructure components to facilitate necessary equipment upgrades and SAN design changes.

3 Switch integration testing

In this section we present test results when comparing two different connection patterns for the 1Gb to 10Gb uplinks between the different speed switches: Split (or cross) connect vs. Straight interconnect. Figure 4 below illustrates the difference between the split and straight connection scenarios.

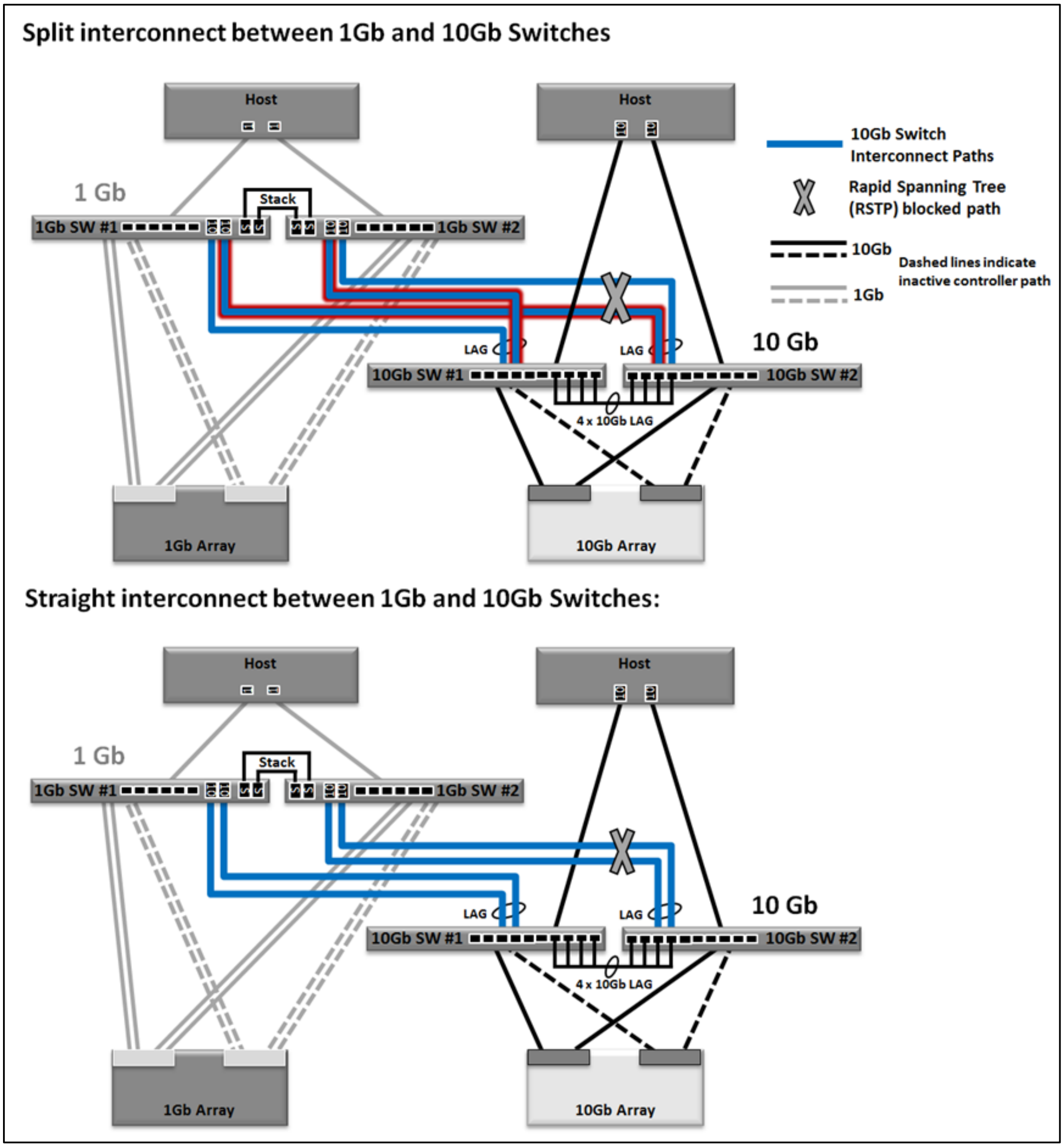


Figure 4 Split vs. straight interconnect between 1Gb and 10Gb switches

> **Note:** If you are using switches that do not support a stacking mode then you must use the straight interconnect pattern as shown in Figure 4.

3.1 Comparing split-uplink vs. straight uplink performance using 1Gb host initiator

In the previous section we described considerations and two options for connecting 10 Gb switches into an existing 1Gb SAN, In this section we present results from a series of IO performance tests designed to determine the best option.

In the first series of tests we tested the following **1Gb initiator IO paths**:

A. One PowerEdge R710 with two 1Gb SAN ports connecting to one **EqualLogic PS6000 1Gb array** via stacked PowerConnect 6248 switches.
B. One PowerEdge R710 with two 1Gb SAN ports connecting to one **EqualLogic PS6010 10Gb array** via stacked PowerConnect 6248 switches and PowerConnect 8024F switches using a **split-uplink configuration**.
C. One PowerEdge R710 with two 1Gb SAN ports connecting to one **EqualLogic PS6010 10Gb array** via stacked PowerConnect 6248 switches and PowerConnect 8024F switches using a **straight-uplink configuration**.

Part A provides an IO performance baseline for the pure 1Gb SAN configuration. Part B and Part C allow us to compare IO performance between the split-uplink and straight-uplink scenarios.

The EqualLogic Host Integration Toolkit was installed on each server and Multi-path IO was enabled with default settings used (mode = least queue depth, maximum two connections per slice, maximum six connections per volume). Each of the target arrays contained four 100 GB volumes. For more system configuration details see Appendix A .

During each test we gathered throughput (MB/s), IOs per second (IOPs), and latency (milliseconds) data at the host initiator. We also monitored TCP retransmission rates to ensure that it never went above 0.5% during the test runs. We used VDBench (http://vdbench.org) for I/O generation. VDBench was selected because of its ability to run multiple initiators simultaneously while rolling up the results into a single performance metric, as well as the granularity and detail of the metrics it provides.

We tested three IO workload types: sequential read, sequential write and a random read/write mix. Each test run ran a single workload type for 15 minutes. We completed three separate test runs for each workload type, and then computed an average of the results of all three runs. Details on the workload types are shown in Table 1.

Workload Type	Block Size	Read/Write Ratio
Sequential Read	256 K	100%/0%
Sequential Write	64 K	0%/100%
Random Read/Write	8 K	67%/33%

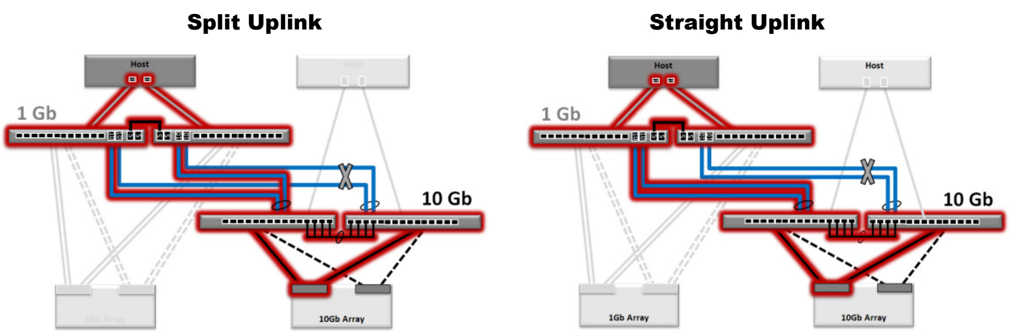
Table 1 – Workloads

Figure 5 illustrates the IO flow paths and shows the results for 1Gb initiator test sequence. TCP retransmit rates remained well below our 0.5% during these test runs. (See Appendix A for full test result data.)

The performance comparison data in Figure 5 is for two 1Gb host ports accessing four volumes on a single 10Gb array, with data flow traversing the uplink path between the 1Gb and 10Gb switches. **Note the percent differences in IO performance calculated for each workload are negligible.** Again, throughput for both sequential workloads was almost at line rate for the two 1Gb SAN ports on the initiator. In those test runs the traffic flowing across the 20Gb uplinks to the 10Gb switches was well below the available bandwidth.

Performance comparison: Split vs. straight uplinks between 1Gb and 10Gb switches

- **1Gb initiator (host) connecting to 10Gb array**



I/O Type		% Difference from Split to Straight Uplink
Seq. Write	IOPS	**+1.12**
	MB/s	**+1.12**
	Latency (ms)	**-1.14**
Seq. Read	IOPS	**-0.01**
	MB/s	**-0.01**
	Latency (ms)	**+0.01**
Random R/W	IOPS	**+0.44**
	MB/s	**+0.45**
	Latency (ms)	**-0.45**

Figure 5 – Results: 1Gb initiator uplink performance comparison

3.2 Comparing split-uplink vs. straight uplink performance using 10Gb host initiator

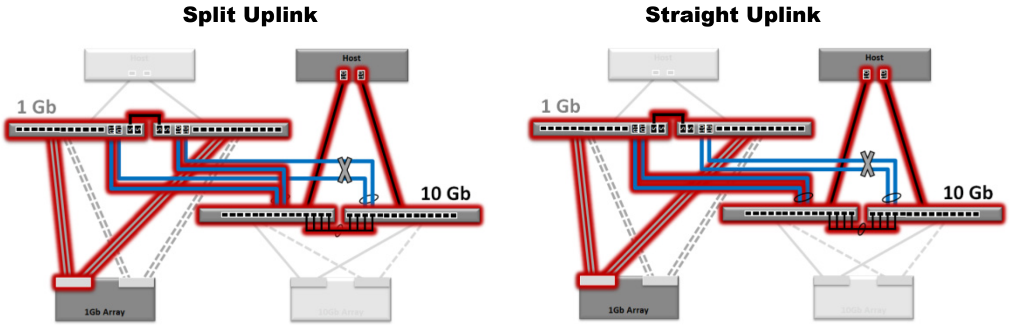
In the previous section we presented results when using 1Gb host initiators. In this section we present results for using a 10Gb host initiator instead, for the same series of tests. We tested the following **10Gb initiator IO paths**:

D. One PowerEdge R710 with two 10Gb SAN ports connecting to one **EqualLogic PS6010 10Gb array** via PowerConnect 8024F switches.
E. One PowerEdge R710 with two 10Gb SAN ports connecting to one **EqualLogic PS6000 1Gb array** via PowerConnect 8024F switches and stacked PowerConnect 6248 switches using a **split-uplink configuration**.
F. One PowerEdge R710 with two 10Gb SAN ports connecting to one **EqualLogic PS6000 1Gb array** via PowerConnect 8024F switches and stacked PowerConnect 6248 switches using a **straight-uplink configuration**.

Part D provides an IO performance baseline for the pure 10Gb SAN configuration. Part E and Part F allow us to compare IO performance between the split-uplink and straight-uplink scenarios. The conduct of this test series was the same as for the 1Gb host initiator series in the previous section.

Figure 6 illustrates the IO flow paths and shows the results for 10Gb initiator test sequence. (See Appendix B for full test result data.) The performance comparison data corresponds to two 10Gb host ports accessing four volumes on a single 1Gb array, with data flow traversing the uplink path between the 1Gb and 10Gb switches. **Note the percent differences in IO performance calculated for each workload are negligible.** TCP retransmits peaked at 0.165%, staying well below the 0.5% limit. The measureable rate of retransmits is consistent with having a high bandwidth link saturating a lower bandwidth link.

Performance comparison : Split vs. straight uplinks between 1Gb and 10Gb switches



I/O Type		% Difference from Split to Straight Uplink
Seq. Write	IOPS	**-0.80**
	MB/s	**-0.81**
	Latency (ms)	**+0.78**
Seq. Read	IOPS	**+1.23**
	MB/s	**+1.24**
	Latency (ms)	**-1.30**
Random R/W	IOPS	**-1.16**
	MB/s	**-1.15**
	Latency (ms)	**+1.15**

Figure 6 – Results: 10Gb initiator uplink performance comparison

3.3 Analysis and conclusion: split vs. straight uplink connection paths

Our test results show that there is almost no difference in performance between the split and the straight uplink connection patterns. The largest IO performance differences we measured were still less than 2%.

Use the split uplink pattern when using stackable switches

In our test system we used stackable 1Gb switches (two Dell PowerConnect 6248 switches). ***When using stackable 1Gb switches we recommend using the split uplink pattern*** (see Figure 4). In the event of a switch failure in the 1Gb stack, split uplinks will minimize the time necessary for the network to converge to a stable configuration. If you are using straight uplinks and a switch fails in the 1Gb stack, then you will have to wait first for the stack to reset, and then for Rapid Spanning Tree convergence to complete. If you are using split uplinks you will only have to wait for the stack to reset. Rapid Spanning Tree convergence does not occur in the split uplink scenario because the split uplink is recognized as a single link entity by spanning tree.

You must use the straight uplink pattern with using non-stackable switches
If you are using switches that do not support a stacking mode then you must use the straight interconnect pattern.

> **Note:** For development of this paper we did not conduct testing using non-stacking switches. We are providing these additional design guidelines in case you are using non-stacking switches.

For this scenario, note the following design differences as shown in Figure 7:

- A LAG is used to create the connection between **1Gb SW#1** and **1Gb SW#2**.
- A high rapid spanning tree link cost is assigned to the 1Gb switch LAG (note the location of the RSTP blocked path). Doing this prevents 10Gb inter-switch traffic from having to pass through the 1Gb switch LAG.

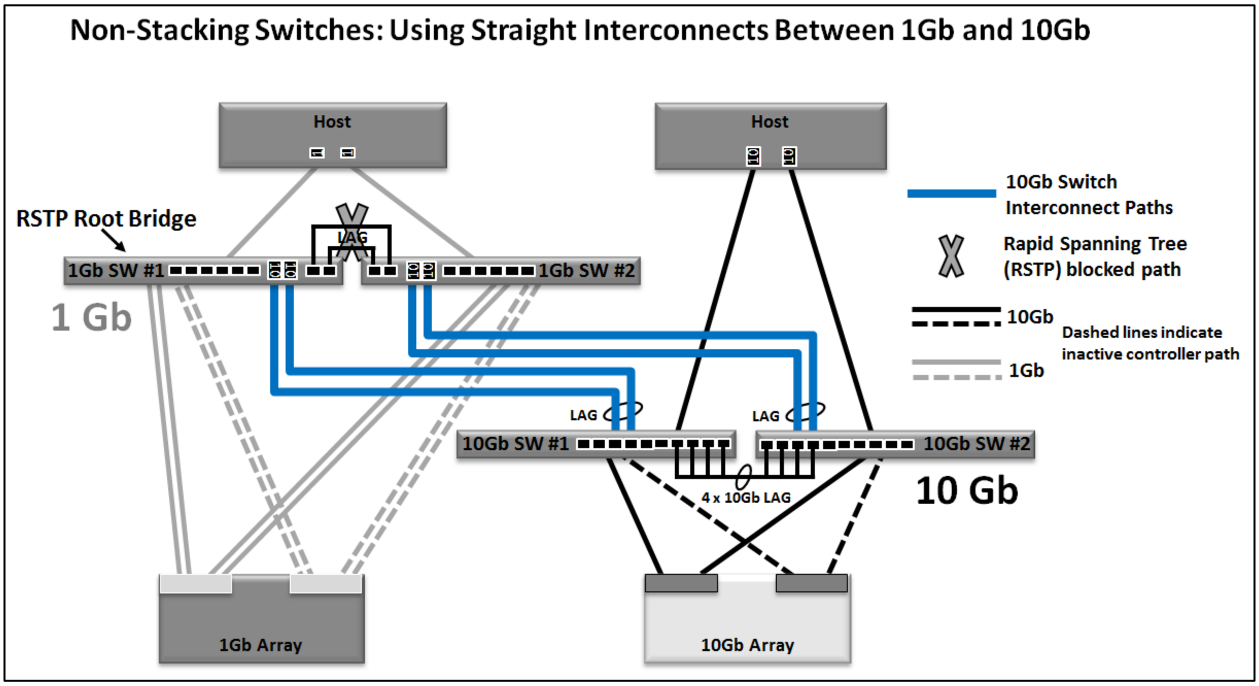


Figure 7 – Non-stacking switches: using straight interconnects between 1Gb and 10Gb

> **Note:** For both straight and split uplink connection patterns it is important to keep in mind as you grow your architecture that you make sure the inter-switch links are appropriately sized for storage and inter-array traffic[1].

For all of the test cases described in the following sections we use the Split Uplink pattern and stacked 1Gb switches.

4 Pool architecture and volume placement testing

When integrating 10Gb arrays into existing 1Gb EqualLogic SANs you must also reconsider how you locate volumes within the available pools. In this section we focus on answering two key questions:

- Will you get better performance having 1Gb initiators access volumes on 10Gb arrays?
- Do you place 10Gb arrays in the same pools as existing 1Gb arrays?

The EqualLogic PS series arrays are very flexible in regards to their configuration—up to and including the placement of arrays into storage pools. The pool is a very important construct within an EqualLogic SAN, as they control the load balancing of volumes across multiple arrays, redundancy among those arrays, and inter-array communication.

It is possible to place arrays with different speed controllers in the same pool. But is it advisable to do so? This question has become important, as customers now have the ability to add 10Gb arrays into existing SANs.

To provide answers to these questions we ran the three IO workloads defined earlier against different configurations of volumes and pools. As before, we ran each workload three times for 15 minutes, then calculated average IO performance data. The goal was to measure performance differences between same speed pools and the mixed speed pools.

4.1 Comparing pool design and volume placement when using a 1Gb host initiator

We tested the following **1Gb initiator IO paths**:

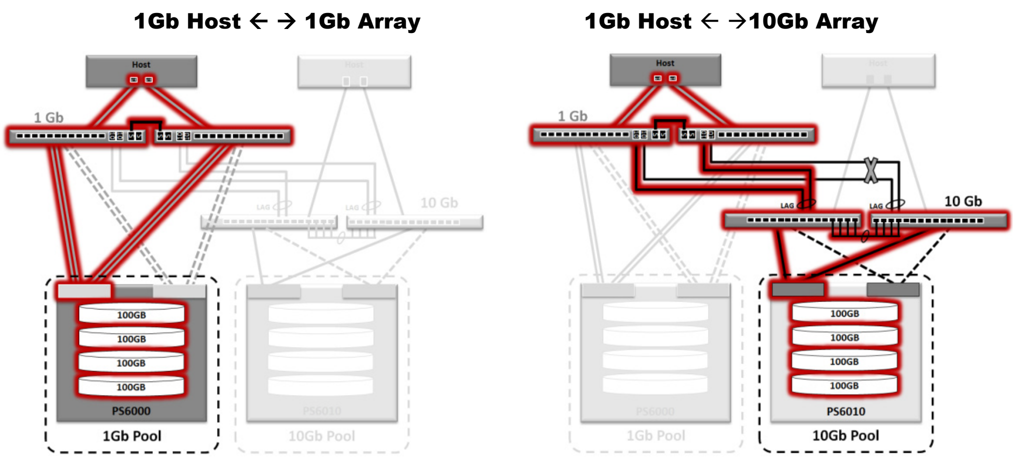
A. One PowerEdge R710 with two 1Gb SAN ports connecting to four volumes in one **EqualLogic PS6000 1Gb array** via stacked PowerConnect 6248 switches.
B. One PowerEdge R710 with two 1Gb SAN ports connecting to four volumes in one **EqualLogic PS6010 10Gb array** via stacked PowerConnect 6248 switches and PowerConnect 8024F switches
C. One PowerEdge R710 with two 1Gb SAN ports connecting simultaneously to **separate speed pools:** two volumes in one **EqualLogic PS6000 1Gb array** and two volumes in one **EqualLogic PS6010 10Gb array**.
D. One PowerEdge R710 with two 1Gb SAN ports connecting simultaneously to four volumes stored in a **mixed speed pool** consisting of one **EqualLogic PS6000 1Gb array** and one **EqualLogic PS6010 10Gb array**

Parts A and B provide IO performance data for comparing the difference when connecting 1Gb initiators to 1Gb vs. 10Gb arrays.

Parts C and D provide IO performance data for comparing IO performance when a 1Gb initiator simultaneously connects to volumes in separate speed pools vs. connecting to the same set of volumes in a single mixed speed pool

Figure 8 illustrates the IO flow paths and shows the results when comparing 1Gb initiator accessing 1Gb targets vs. 10Gb targets.

Performance comparison: **1Gb host** connecting to 1Gb array vs. 10Gb array



I/O Type		% Difference from 1Gb targets to 10Gb targets
Seq. Write	IOPS	**-9.28**
	MB/s	**-9.28**
	Latency (ms)	**+8.49**
Seq. Read	IOPS	**+4.90**
	MB/s	**+4.89**
	Latency (ms)	**-5.13**
Random R/W	IOPS	**+1.10**
	MB/s	**+1.11**
	Latency (ms)	**-1.12**

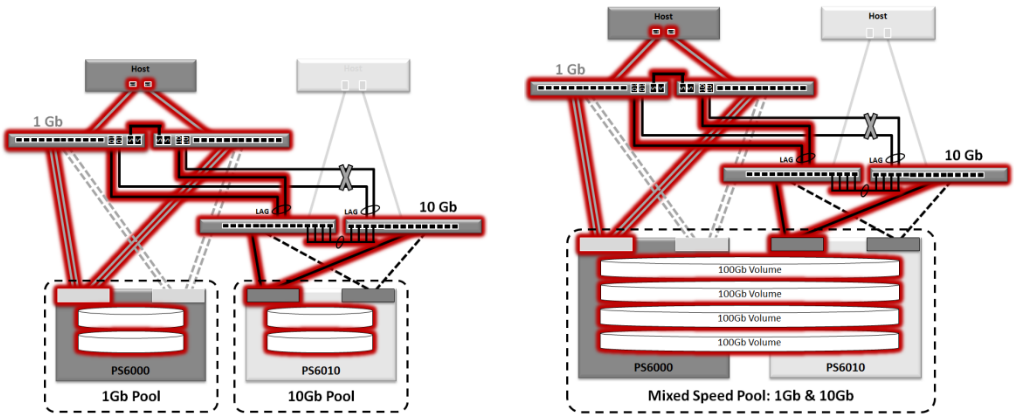
Figure 8 – Results: 1Gb initiator connecting to 1Gb targets vs. 10Gb targets

In Figure 8 we measured some differences in sequential IO performance. After moving the target volumes over to the 10Gb array there was a small decrease in sequential write performance and a small increase in sequential read performance. There was negligible performance difference for random IO.

Figure 9 illustrates the IO paths and shows the results when comparing 1Gb initiator performance between connecting to volumes in separate speed pools vs. connecting to the same set of volumes in a single mixed speed pool.

Performance comparison: **1Gb host** connecting to mixed speed targets in separate pools vs. connecting to mixed speed targets in a single pool

1Gb & 10Gb targets in separate speed pools | **10Gb targets in a mixed speed pool**



I/O Type		% Difference from 1Gb targets to 10Gb targets
Seq. Write	IOPS	-0.06
	MB/s	-0.06
	Latency (ms)	-0.06
Seq. Read	IOPS	-2.07
	MB/s	-2.07
	Latency (ms)	+0.31
Random R/W	IOPS	-2.17
	MB/s	-2.18
	Latency (ms)	+1.78

Figure 9 – Results: 1Gb initiator connecting to separate speed vs. mixed speed pools

As you can see from the results in Figure 9, we measured very little difference in IO performance at the 1Gb initiator when connecting to separate speed pools vs. mixed speed pools.

From this series of 1Gb initiator test results we can conclude that the placement of the volumes within a single pool vs. separating them out by speed does not have a significant impact on performance. In all cases, throughput was limited by the 1Gb bandwidth available to the host. Latency and IOPs provided similar results.

4.2 Comparing pool design and volume placement when using a 10Gb host initiator

We repeated the same test sequence as in the previous section using a **10Gb initiator** on the host:

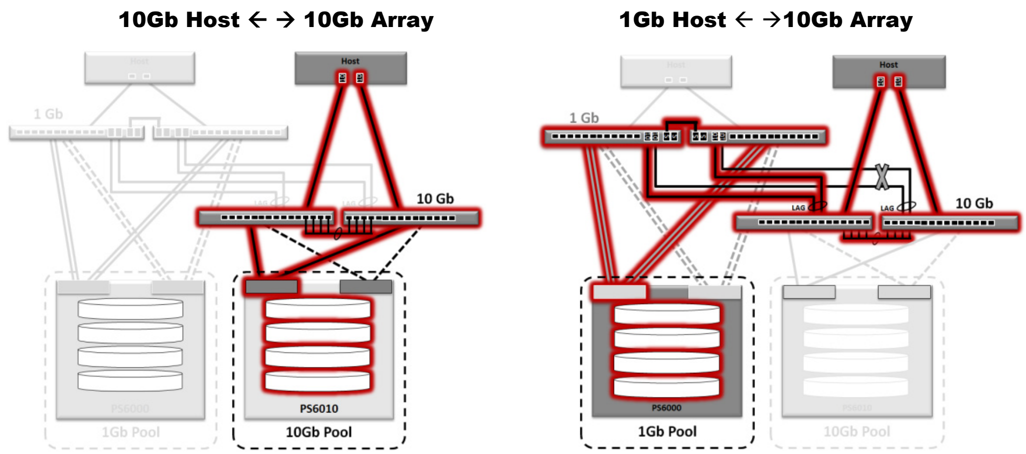
E. One PowerEdge R710 with two 10Gb SAN ports connecting to four volumes in one **EqualLogic PS6010 10Gb array** via PowerConnect 8024F switches.
F. One PowerEdge R710 with two 10Gb SAN ports connecting to four volumes in one **EqualLogic PS6000 1Gb array** via PowerConnect 8024F switches and stacked PowerConnect 6248 switches.
G. One PowerEdge R710 with two 10Gb SAN ports connecting simultaneously to **separate speed pools:** two volumes in one **EqualLogic PS6000 1Gb array** and two volumes in one **EqualLogic PS6010 10Gb array**.
H. One PowerEdge R710 with two 10Gb SAN ports connecting simultaneously to four volumes stored in a **mixed speed pool** consisting of one **EqualLogic PS6000 1Gb array** and one **EqualLogic PS6010 10Gb array**

Parts E and F provide IO performance data for comparing the difference when connecting 10Gb initiators to 1Gb vs. 10Gb arrays.

Parts G and H provide IO performance data for comparing IO performance when a 10Gb initiator simultaneously connects to volumes in separate speed pools vs. connecting to the same set of volumes in a single mixed speed pool.

Figure 10 illustrates the IO flow paths and shows the results when comparing 10Gb initiator access to 1Gb targets vs. 10Gb targets. As expected, we measured significant differences in sequential IO performance when moving the target volumes from the 10Gb array to the 1Gb array.

Performance comparison: **10Gb host** connecting to 10Gb array vs. 1Gb array

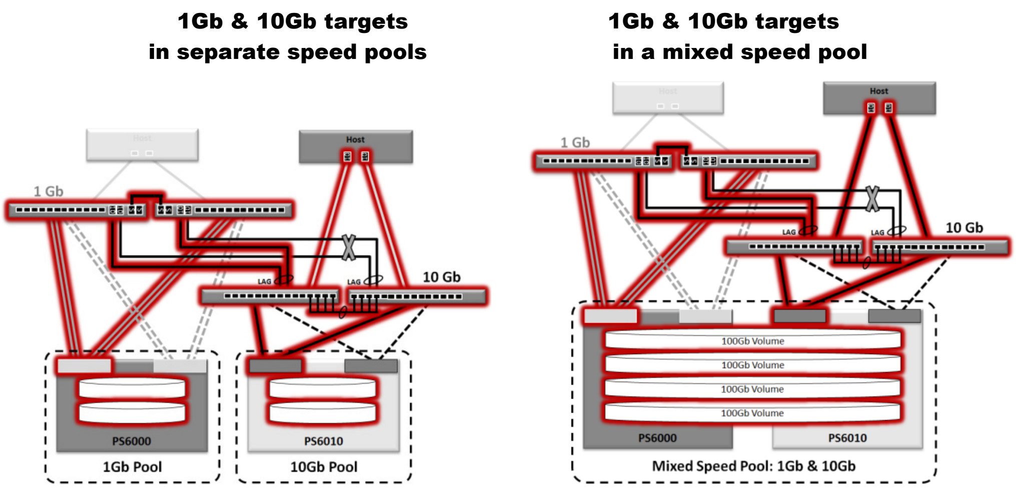


I/O Type		% Difference from 1Gb targets to 10Gb targets
Seq. Write	IOPS	-17.57
	MB/s	-17.57
	Latency (ms)	+14.95
Seq. Read	IOPS	**-130.28**
	MB/s	**-130.28**
	Latency (ms)	**+56.60**
Random R/W	IOPS	+4.38
	MB/s	+4.38
	Latency (ms)	-4.59

Figure 10 – Results: 10Gb initiator connecting to 10Gb targets vs. 1Gb targets

Figure 11 illustrates the IO paths and shows the results when comparing 10Gb initiator performance between simultaneous connection to volumes in separate speed pools vs. connecting to the same set of volumes in a single mixed speed pool.

Performance comparison: **10Gb host** connecting to mixed speed targets in separate pools vs. connecting to mixed speed targets in a single pool



I/O Type		% Difference from 1Gb targets to 10Gb targets
Seq. Write	IOPS	**-45.99**
	MB/s	**-45.99**
	Latency (ms)	**+31.20**
Seq. Read	IOPS	**-37.48**
	MB/s	**-37.48**
	Latency (ms)	**+23.84**
Random R/W	IOPS	+0.27
	MB/s	-1.42
	Latency (ms)	+1.40

Figure 11 – Results: 10Gb initiator connecting to separate speed vs. mixed speed pools

The results in Figure 11 show a significant decrease in performance when using mixed speed pools for sequential IO workloads vs. separate speed pools.

4.3 Analysis of pool design and architecture test results

For 1Gb initiators, it is apparent that the placement of volumes in either separate, single speed pools or a single mixed speed pool is almost irrelevant from a performance standpoint. One thing to be considered though is the migration of those servers to 10Gb as the SAN expands in the future. If this is likely, then the recommendations for 10Gb initiators should be followed. The performance at the 1Gb level would not be affected, but it would obviate the need to either recreate volumes or juggle volumes around on arrays in the event of a 10Gb initiator upgrade.

The second question of whether 1Gb initiators will benefit from accessing 10Gb volumes is also answered positively. No matter the decision on placement of volumes, performance increases can be had by accessing new volumes on the 10Gb arrays.

For 10Gb initiators, the placement of those volumes in separate speed pools is much more critical. As seen in the test results, the accessing of data located only on 1Gb arrays from a 10Gb initiator performs much less. This result does not hold as true when accessing volumes of both speeds. Remember that these tests were single initiator tests only. Any performance penalties from accessing the lower speed arrays would likely be worsened when multiple 10Gb initiators are in play. This is especially true of latency and TCP retransmits.

Based on this analysis, we will be using the separate single-speed pool architecture for the scalability testing.

5 Best practice recommendations

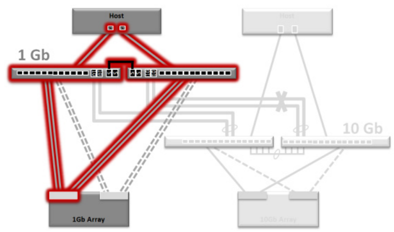
The results presented in this paper show that you must pay attention to some important SAN design considerations when integrating of 10Gb EqualLogic arrays into and existing 1Gb EqualLogic SAN.

- When integrating 10Gb switches into your existing 1Gb environment, how you interconnect the mixed speed switches (split vs. straight uplink) does not have a significant impact on performance as long as the uplinks are sized appropriately to your workloads.
- When connecting 1Gb switches and 10Gb switches together you must always be aware of where Rapid Spanning Tree is going to block links to make sure that 10Gb traffic (i.e. EqualLogic inter-array data flow) never crosses the 1Gb switch.
- You must configure pools and volumes in a way that minimizes impact to IO performance.
    - If you have predominately 1Gb initiators, start upgrading your arrays to 10Gb for comparable or better performance across almost all situations.
    - If you have predominately 10Gb initiators, you should only access data and volumes residing on 10Gb arrays (from those initiators). You may see high latency and retransmit rates when 10Gb initiators connect to 1Gb targets.
    - When adding 10Gb arrays, place them in separate pools from your 1Gb arrays.

Appendix A Uplink comparison: full test results

The baseline 1Gb IO performance data in Figure 12 corresponds to two 1Gb host ports accessing four 100GB volumes on a single 1Gb array. During the baseline test the throughput generated by both sequential workloads was almost equal to line rate for the two 1Gb SAN ports on the host initiator. Note that during the baseline test all traffic was isolated to the PC6248 stack, as all the 1Gb ports were directly connected – no traffic crossed the uplink ports.

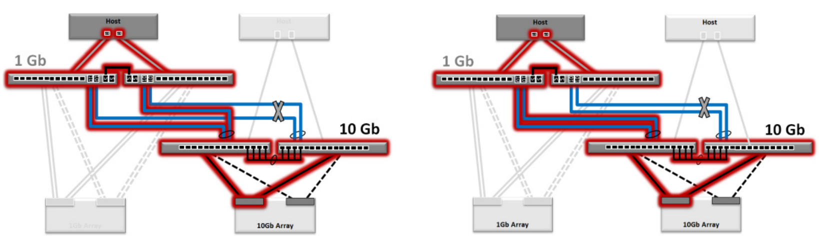
Baseline SAN I/O Performance for **1Gb Initiator**:



I/O Type	Seq. Write	Seq. Read	Random R/W
IOPs	3531.71	860.22	3652.06
MB/s	220.73	215.06	28.53
Latency (ms)	13.586	74.376	8.757

Performance comparison: Split vs. straight uplinks between 1Gb and 10Gb switches

Split Uplink **Straight Uplink**

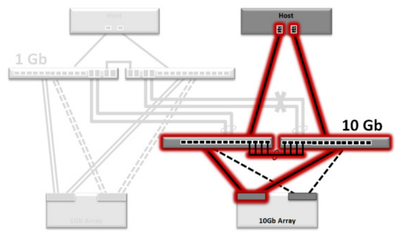


I/O Type		Split Uplink	Straight Uplink	% Difference
Seq. Write	IOPS	3231.85	3268.43	+1.13
	MB/s	201.99	204.28	+1.13
	Latency (ms)	14.847	14.679	-1.13
Seq. Read	IOPS	904.50	904.43	-0.01
	MB/s	226.13	226.11	-0.01
	Latency (ms)	70.744	70.754	+0.01
Random R/W	IOPS	3692.86	3709.37	+0.45
	MB/s	28.85	28.98	+0.45
	Latency (ms)	8.66	8.62	-0.45

Figure 12 – Full test results: 1Gb initiator uplink performance comparison

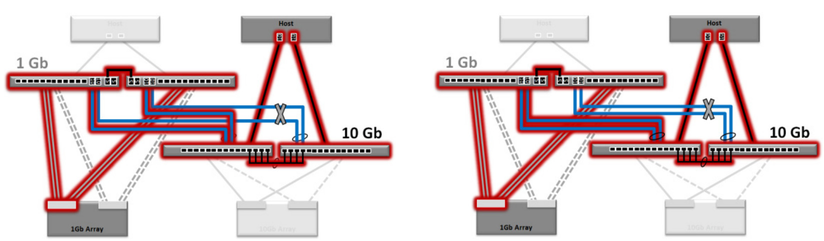
The baseline 10Gb IO performance data in Figure 13 corresponds to two 10Gb host ports accessing four 100GB volumes on a single 10Gb array. Note that during the baseline test all traffic was isolated to the PC8024F switches, as all of the 10Gb ports were directly connected – no traffic crossed the uplink ports to the PC6248 switches. Throughput for these tests was stable, but not near line rate for 10Gb links.

Baseline 10Gb SAN I/O Performance for **10Gb Initiator**



I/O Type	Seq. Write	Seq. Read	Random R/W
IOPs	4151.26	3117.93	3590.33
MB/s	259.45	779.48	28.05
Latency (ms)	11.556	20.520	8.907

Performance comparison : Split vs. straight uplinks between 1Gb and 10Gb switches



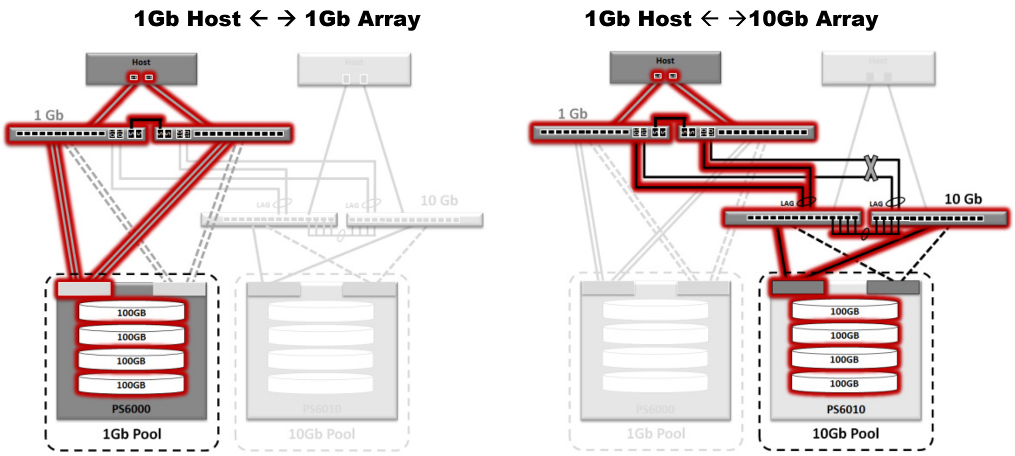
I/O Type		Split Uplink	Straight Uplink	% Difference
Seq. Write	IOPS	3530.80	3502.62	-0.80
	MB/s	220.68	218.91	-0.80
	Latency (ms)	13.588	13.695	+0.79
Seq. Read	IOPS	1353.99	1370.92	+1.25
	MB/s	338.50	342.73	+1.25
	Latency (ms)	47.285	46.679	-1.28
Random R/W	IOPS	3754.83	3711.61	-1.15
	MB/s	29.33	29.00	-1.15
	Latency (ms)	8.517	8.615	+1.16

Figure 13 – Full test results: 10Gb initiator uplink performance comparison

Appendix B Pool/volume placement: full test results

Performance comparison:

1Gb host connecting to 1Gb array vs. 10Gb array



I/O Type		Target: 1Gb Array	Target: 10Gb Array	% Difference
Seq. Write	IOPS	3531.71	3231.85	-8.49
	MB/s	220.73	201.99	-8.49
	Latency (ms)	13.586	14.847	+9.28
Seq. Read	IOPS	860.22	904.50	+5.15
	MB/s	215.06	226.13	+5.15
	Latency (ms)	74.376	70.744	-4.88
Random R/W	IOPS	3652.06	3692.86	+1.12
	MB/s	28.53	28.85	+1.12
	Latency (ms)	8.757	8.66	-1.11

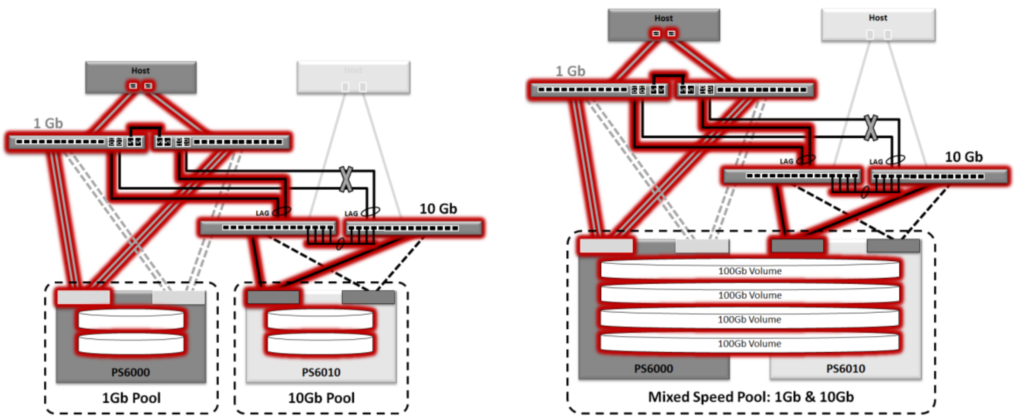
Figure 14 – Full test results: 1Gb initiator connecting to 1Gb targets vs. 10Gb targets

Performance comparison:

1Gb host connecting to mixed speed targets in separate pools vs.
connecting to mixed speed targets in a single pool

1Gb & 10Gb targets in separate speed pools

10Gb targets in a mixed speed pool



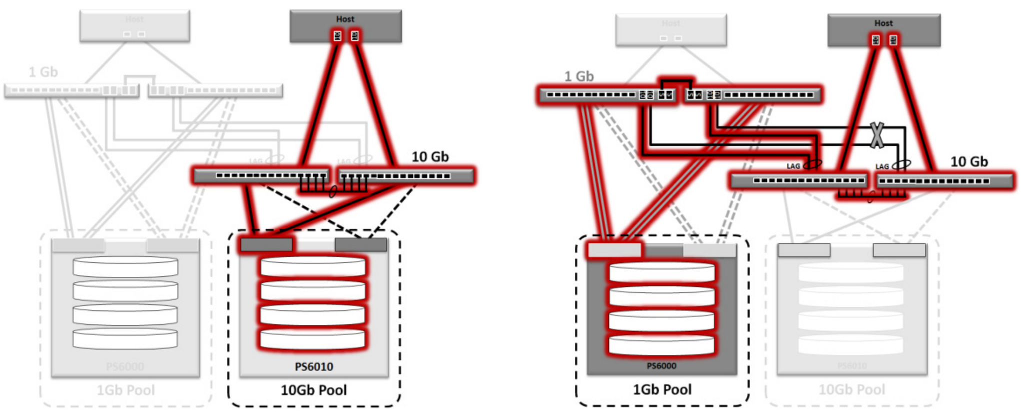
I/O Type		Separate Speed Pools	Mixed Speed Pool	% Difference
Seq. Write	IOPS	3687.70	3685.55	-0.06
	MB/s	230.49	230.35	-0.06
	Latency (ms)	13.01	13.01	-0.06
Seq. Read	IOPS	923.25	904.57	-2.02
	MB/s	230.81	226.14	-2.03
	Latency (ms)	70.57	70.78	+0.31
Random R/W	IOPS	5344.48	5231.00	-2.12
	MB/s	41.76	40.87	-2.13
	Latency (ms)	6.00	6.11	+1.82

Figure 15 – Full test results: 1Gb initiator connecting to separate speed vs. mixed speed pools

Performance comparison:

10Gb host connecting to 10Gb array vs. 1Gb array

10Gb Host ← → 10Gb Array **1Gb Host ← →10Gb Array**



I/O Type		Target: 10Gb Array	Target: 1Gb Array	% Difference
Seq. Write	IOPS	4151.26	3530.80	-14.95
	MB/s	259.45	220.68	-14.95
	Latency (ms)	11.55	13.58	+17.58
Seq. Read	IOPS	3117.93	1353.99	**-56.57**
	MB/s	779.48	338.50	**-56.57**
	Latency (ms)	20.52	47.285	**+130.44**
Random R/W	IOPS	3590.33	3754.83	+4.58
	MB/s	28.05	29.33	+4.58
	Latency (ms)	8.90	8.51	-4.39

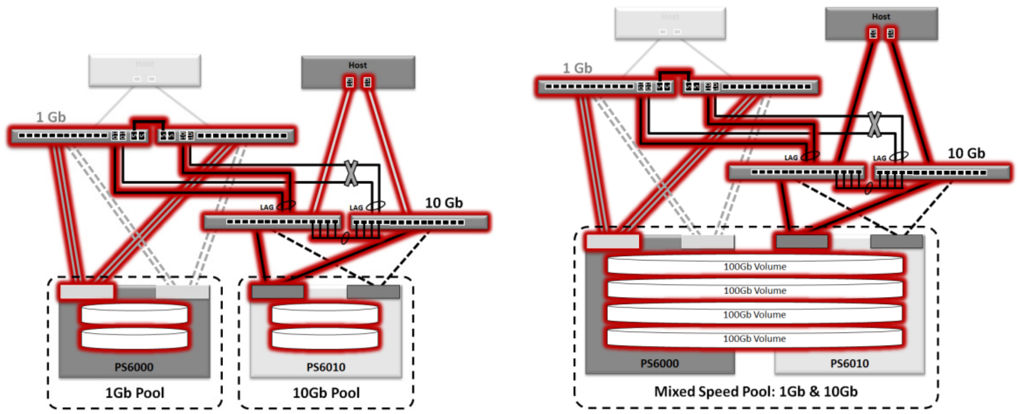
Figure 16 – Full test results: 10Gb initiator connecting to 10Gb targets vs. 1Gb targets

Performance comparison:

10Gb host connecting to mixed speed targets in separate pools vs. connecting to mixed speed targets in a single pool

1Gb & 10Gb targets in separate speed pools

1Gb & 10Gb targets in a mixed speed pool



I/O Type		Separate Speed Pools	Mixed Speed Pool	% Difference
Seq. Write	IOPS	7670.07	5253.69	**-31.50**
	MB/s	479.38	328.35	**-31.50**
	Latency (ms)	6.28	9.13	**+45.35**
Seq. Read	IOPS	3427.19	2492.79	**-27.26**
	MB/s	856.79	623.20	**-27.26**
	Latency (ms)	19.75	25.945	**+31.31**
Random R/W	IOPS	5394.02	54.08	+0.27
	MB/s	42.14	41.55	-1.40
	Latency (ms)	5.92	6.01	+1.42

Figure 17 – Full test results: 10Gb initiator connecting to separate speed vs. mixed speed pools

Appendix C Performance testing configuration

Table 2 lists the major hardware components for the test system configuration used throughout this document:

Test Configuration	Hardware Components
1 Gigabit iSCSI Initiators	• 10 x Dell PowerEdge R710 Servers: ○ 2 x Quad Core Intel® Xeon® X5570 Processors ○ BIOS 2.1.15 ○ 32 GB RAM, 8 M Cache, 2.93 GHz ○ 2 x 146GB 10K SAS internal disk drives ○ Broadcom 5709 1GbE quad-port NIC (LAN on motherboard) ▪ Driver 5.2.22.0
10 Gigabit iSCSI Initiators	• 10 x Dell PowerEdge R710 Servers: ○ 2 x Quad Core Intel® Xeon® X5570 Processors ○ 32 GB RAM, 8 M Cache, 2.93 GHz ○ 2 x 146 GB 10K SAS internal disk drives ○ Broadcom NetXtreme II 57711 10GbE NIC, Dual-Port ▪ Driver 5.0.38.0
Network	• 2 x Dell PowerConnect 6248 1Gb Ethernet Switch ○ Firmware 3.2.0.7 • 2 x Dell PowerConnect 8024F 10Gb Ethernet Switch ○ Firmware 3.1.4.5
Storage	• 5 x Dell EqualLogic PS6000E: ○ Firmware 4.3.7 ○ 16 x 450GB SATA disks with dual 4 port 1 GbE controllers ○ 14 SATA disks configured as RAID 10, two hot spares • 4 x Dell EqualLogic PS6010E: ○ Firmware 4.3.7 ○ 16 x 450GB SATA disks with dual 2 port 10 GbE controllers ○ 14 SATA disks configured as RAID 10, two hot spares

Table 2 – Test Configuration Hardware Components

Table 3 lists the major software components for the test system configuration used throughout this document:

Test Configuration	Software Components
Initiators	• Windows 2008 R2 Server ○ MS iSCSI Initiator ○ EqualLogic Host Integration Toolkit version 3.4.2 ○ VDBench version 5.0.2 IO Generation software
Monitoring Tools	• EqualLogic SAN Headquarters version 2.0 • EQL_Monitor: Dell developed PERL-based monitoring script • PerfMon

Table 3 – Test Configuration Software Components

Related publications

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

- *Integrating EqualLogic PS6x10 Arrays with Existing SANs*

  http://www.equallogic.com/resourcecenter/assetview.aspx?id=9447

- *EqualLogic Configuration Guide*

  http://www.delltechcenter.com/page/EqualLogic+Configuration+Guide

- *Dell EqualLogic PS Series Network Performance Guidelines*

  http://www.equallogic.com/resourcecenter/assetview.aspx?id=5229