Dell™ EqualLogic™ Storage Systems

Integrating EqualLogic PS6x10 Arrays with Existing SANs

A Dell Technical White Paper

Table of Contents

1 Introduction

With the introduction of 10Gb Ethernet-based EqualLogic™ PS 6x10 series of iSCSI arrays, administrators need to determine what is the best way to take advantage of this high performing storage solution. Should they build out separate storage area network (SAN) groups for their 10 gigabit arrays, or can they integrate their new arrays into their existing 1 gigabit EqualLogic SAN group.

Leveraging Dell's award winning EqualLogic PS Series SAN array technology, the 10Gb Ethernet-based PS6010 and PS6510 series of arrays provide a scalable, high-performance, tiered storage solution to meet almost any application requirement. This paper describes strategies and best practices for integrating and managing the 10 gigabit Ethernet (10GbE) PS Series arrays into existing EqualLogic SANs consisting of one gigabit Ethernet (1GbE) PS Series arrays. While the recommendations described here are not the only possible architecture options, they are options that can provide the flexibility to grow the SAN as needed while continuing to leverage existing server and storage resources as administrators migrate to newer 10GbE solutions.

As a basis for the ensuing discussion, we will first review the 10GbEstandard and how 10GbE is different from future Enhanced Ethernet standards. Next, we will describe a basic 1GbE networked storage environment that consists of Dell™ PowerEdge™ Servers with 1GbE host adapters connected to an existing EqualLogic SAN group consisting of 1GbE switches and PS Series arrays. We will then discuss the requirements, recommendations and procedures for introducing 10GbE infrastructure components into the existing SAN switching infrastructure, integration of 10GbE PS Series into the existing EqualLogic SAN Group and finally adding 10GbE hosts to the SAN.

2 Ten Gigabit Ethernet Technology Overview

2.1 What is 10Gb Ethernet?

Ethernet has been around since the mid 1970's. Over that time, it has transformed itself in many areas including speed – moving from 1Mbps, 10Mbps, 100Mbps and then to 1Gbps – as well as in terms of features and functionality by adding extensions such as moving from a broadcast, bus architecture to a switched architecture, implementing vLANs, jumbo frames and full-duplex communications among many other innovations. Most of these features and improvements have been managed by the IEEE standards body and its 802 Working Group.

So, how is 10GbE different from the current implementation of the 1Gb Ethernet standard…other than speed? Not as much as you would probably think. At its core, 10GbE is still based on IEEE 802.3, the main standard for Ethernet. In its current implementation, the 802.3ae and later the 802.3an standards define how 10GbE must be implemented. In general, 10GbE added a new physical connection interface called SFP+ as well as introducing a 10Gbase-T standard along with more stringent cabling standards. Last, 10GbE now supports full-duplex communications only as opposed to half-duplex or shared, collision detection implementations.

The objectives for 10GbE had to meet the following requirements:

- Preserve the 802.3 frame format
- Preserve the minimum and maximum frame size limitations
- Support forwarding between various speeds of previous Ethernet implementations…including 10Mbps Ethernet

- Support Full Duplex only (no CSMA/CD, or half-duplex)
- Support ISO 11801 media (Category 6A and Category 7 copper)
- Provide multiple Physical Layer implementations such as LR, SR fiber and copper

From a storage networking perspective, we are interested in specifics around real-world implementations of 10GbE components. It is important to understand that while the 802 standard states that various physical layer implementations are supported; actual vendor implementations may be based on a subset of the available options. In particular, most initial implementations of switches, iSCSI targets, and host network controllers, are based on the SFP+ form factor. Why would this be the case? Primarily, this is because of power requirements. SFP+ solutions currently have a lower power load per port than 10Gbase-T solutions, though this is changing as improvements in manufacturing are moving 10Gbase-T to lower power levels. SFP+ also provides customers with flexibility in choosing the type of cable used to interconnect devices since it is a modular, pluggable standard. The SFP+ based Direct Access Copper solution (DAC or 10GSFP+Cu) that is available is a fairly short-haul solution restricted to cable lengths of less than 15 meters, though most vendors are certifying these cables at much shorter distances…typically in the 3-7 meter range. Also, this standard copper SFP+ cable includes the SFP+ interface as part of the cable as opposed to the optical cable options. Fiber optic SFP+ solutions will require the purchase of separate interface modules that match the cable standard being used. The various SFP+ cable options are listed in Table 1: SFP+ Connector and Cable Types. The discussions and recommendations in this paper do not differentiate between any of the SFP+ physical implementation and should apply to any of the various SFP+ cabling solutions.

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

Table 1: SFP+ Connector and Cable Types

2.2 How is this different from Data Center Bridging?

The availability of 10GbE has prompted the development of several new standards proposals for the data center, as organizations begin to deploy unified networking—that is, a single network fabric for all networking traffic (LAN, SAN, Infiniband, etc). These new standards go by several names: “Converged Enhanced Ethernet,” “Data Center Ethernet,” and the industry-standard term “Data Center Bridging,” or DCB.

While Ethernet is generally a very reliable networking technology, packets can, and do get lost during transmission for various reasons such as network congestion and high loads on servers, switches and other network connected devices. For many applications, dropped packets can affect performance.

This is especially true of a storage area network. Most networks today rely on TCP/IP to provide a guaranteed delivery service on top of Ethernet.

DCB is a set of proposed extensions to the Ethernet standard that promote two primary capabilities to the Ethernet standard: the ability to share the available network with multiple traffic types and to provide a "lossless" Ethernet standard without the need for TCP/IP. By adding these extensions to the Ethernet standard, all types of traffic (be it iSCSI, NFS, TCP/IP, Infiniband or Fibre Channel over Ethernet) can gain the benefits of a these new capabilities. DCB is still (at the time of this writing) in the draft stage of development, but 10GbE can be, and currently is being deployed in data centers everywhere.

3 Integrating 10Gb PS Series Arrays

Now that we have a basic understanding of 10GbE, let us look at how to integrate the 10GbE based EqualLogic products such as the PS6010 and PS6510 arrays into an existing EqualLogic SAN group that consists of arrays using 1GbE controller technology. In this section, we will define a typical 1GbE EqualLogic SAN configuration then look at how to add 10GbE switches, 10GbE EqualLogic arrays and 10GbE hosts into the mixed technology SAN. In the next section, we will discuss what 10GbE arrays allow you to do and how to take advantage of the new technology once it is integrated into the SAN.

3.1 The Starting Point

Before talking about the various 10GbE components to be integrated into a 1GbE SAN and the deployment and management strategies, we have to have a starting point…a typical 1GbE PS Series SAN. Figure 1: 1GbE EqualLogic SAN Infrastructure illustrates a typical, fully redundant EqualLogic SAN reference infrastructure. While Figure 1 looks very complex, each connection has a role to play in providing a redundant yet scalable networked storage solution.

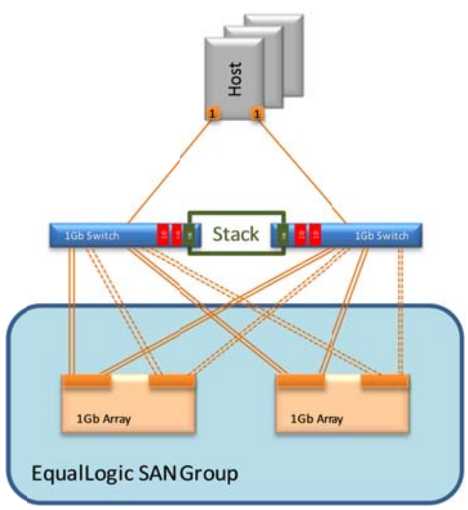

Figure 1: 1GbE EqualLogic SAN Infrastructure

Switches from different vendors will provide different capabilities and offer differing ways to expand a network infrastructure using theses switches. Basically, there are two ways to connect switches together within a layer-2 network…proprietary stacking and standards based inter-switch linking using standard Ethernet ports. Many switches provide both capabilities and they are typically used for different purposes. Table 2 compares the two technologies and describes the primary role each plays within a network infrastructure.

Interconnect	Primary Purpose	Pros/Cons/Bottom Line
Stacking	Creating a larger, logical switch within an isolated physical location	Pros: ➢ Easier to manage multiple switches as single switch ➢ Higher bandwidth than using link aggregation and Ethernet ➢ Not limited by Ethernet standards Cons: ➢ Proprietary, cannot be used to interconnect switches from different vendors ➢ Increases cost of switch Bottom Line: ➢ Best way to scale a storage network in a single location on a single subnet providing lowest latency and best bandwidth relative to Inter-switch Linking
Inter-switch Linking	Creating a data path between switches in one location or subnet with those in another location or subnet	Pros: ➢ Leverages Ethernet standard extensions ➢ Can be used to interconnect switches from different vendors ➢ Can use Link Aggregation Protocols (LACP/EtherChannel) to pool multiple 1GbE or 10GbE links into a single logical link providing bandwidth and redundancy Cons: ➢ Most solutions limited to 8 port link aggregation group ➢ Spanning Tree Protocol must be used if more than two switches are used causing some links to be "blocked" reducing bandwidth availability Bottom Line: ➢ Use when stacking is not available ➢ Use when connecting to aggregation/core switching infrastructure ➢ Use when switches are from different vendors

Table 2: Stacking vs. Inter-Switch Linking

A SAN consists of three major components: hosts, switches, and storage targets. For our 1GbE SAN starting point, we will make some general assumptions about the configuration of each of these components, but keep in mind that your configuration may vary. Table 3 describes each SAN component used and its configuration prior to integrating the 10GbE components.

Component	Configuration
Switches	The SAN consists of two 1GbE stackable switches that have 24 or 48 1GbE ports and at least two 10GbE "uplink" ports. Switches such as the PowerConnect 6224, Cisco® Catalyst® 3750, and others offer this type of configuration at varying costs. Each switch is configured as follow: ➢ Stacking option to allow switches to be configured into a single, logical switch while providing redundancy and simplified switch management. ➢ Jumbo Frames have been enabled on all ports ➢ Spanning Tree Protocol is currently disabled since these switches are not currently connected to any other switches. ➢ Flow Control has been enabled on all Ethernet ports on the switch to allow the switch to manage packet flow between hosts and arrays.
PS Arrays	The SAN consist of two EqualLogic PS6000 arrays fully connected using the following best practices for a redundant, scalable SAN ➢ Connect two ports from each array controller to each switch. This will result in 2 ports from one controller (solid orange lines in Figure 1) from each array going to the left-hand switch and 2 ports from the second controller (dashed orange lines in Figure 1) from each array going to the right-hand switch.
Hosts	Each host is configured as follows: ➢ Two 1GbE ports dedicated to SAN connectivity. One port from each host connects to each switch providing a fully redundant path between the host and the arrays within the SAN. ➢ Microsoft® Windows® operating system with the Microsoft® iSCSI initiator software installed ➢ Dell's EqualLogic Host Integration Toolkit is installed and the PS Series MPIO Device Specific Module (DSM) has been configured to balance traffic on all host Ethernet ports on the SAN's IP subnet.

Table 3: Initial SAN Component Configuration Details

3.2 Strategies for Integrating 10Gb Switches

When looking at ways to implement 10GbE based storage technology, there are two basic choices that are available – replace or integrate. The next two sections will look at each of these options a little differently since this document is really focused on the second option – to integrate 1GbE and 10GbE technologies.

3.2.1 Replacing Existing 1GbE SAN Components

Replacing the current 1GbE SAN components can present differing challenges depending on the component being replaced, but the biggest challenge could be in replacing the actual storage arrays. Depending on the storage solution, this could be as simple as changing the array controllers within each array or as complex as having to migrate all of the data from the older arrays to newer 10GbE arrays. Regardless, any "rip and replace" process will require extensive planning and in most cases will require some downtime as individual components are replaced. EqualLogic PS Series arrays have a distinct advantage here.

A core feature of virtual storage – and EqualLogic SANs – is the ability to move volumes from one storage pool (set of arrays) to another storage pool. In fact, EqualLogic makes it even easier by providing a "Delete Array" command that automatically moves any volumes hosted by the array in question to other arrays within the SAN group as long as there is adequate free storage available. This feature, along with the ability to have 1Gb arrays in the same SAN with 10Gb arrays, means that EqualLogic provides a simpler process for migrating to a 10Gb solution than a complete replacement of the entire storage solution that is required with most vendors. Once the data has been migrated, the 1Gb arrays can be removed from the SAN group and repurposed.

While this is a viable option, many customers will want to continue to take advantage of their 1Gb arrays, whether for lower priority data storage, test/dev environments or archival disk-to-disk backup solutions. The primary way to do this is to take advantage of the same capability just mentioned – supporting Groups with both 1Gb and 10Gb arrays – to seamlessly integrate 10Gb arrays into the existing Group and immediately taking advantage of this higher performance storage in parallel with the existing storage.

3.2.2 Integrate 10Gb w/ existing 1Gb

In many customer environments, there will be a need to have 1Gb arrays and 10Gb arrays coexist in the same SAN infrastructure and same EqualLogic SAN group. There are many advantages to this solution. In many cases, only a small number of applications will need the additional performance that a 10Gb array will provide. By integrating 10Gb arrays into the existing SAN group with 1Gb arrays, the administrator will have the flexibility of advanced storage tiering capabilities providing the ability to migrate volumes supporting applications in need of additional performance from existing 1Gb arrays to new, higher-performance 10Gb arrays and allowing the administrator to continue to manage both types of storage arrays within the same SAN group. The next several sections will provide more insight into just how to integrate 10Gb arrays into your existing EqualLogic SAN group.

3.3 Preparing the Network Infrastructure

When considering integrating 10Gb networking components into an existing 1Gb SAN infrastructure, planning is one of the most important steps for success. Several considerations must be made when planning this integration including:

- How many 10Gb arrays will need to be integrated?

  The number of arrays will help determine the number of ports need from the candidate switches. The number of ports required per array will vary by model, but in most cases, at least two 10Gb ports from each array controller for a total of four 10Gb ports per array (2 controllers per array) will need to be connected to the SAN infrastructure to ensure that all arrays and hosts have a redundant path through the SAN infrastructure.

- How many 10GbE hosts will be connecting?

  Again, this will help determine the number of ports that will be needed in the final solution. Each host will require two 10GbE ports.

- Do the existing 1Gb switches have any 10Gb uplink ports and if so, are they SFP+ compatible?

  Each switch vendor has many different models of their 1Gb switches. Each model family will have different features depending on the target market. One feature in more robust, higher performance switch families is the integration of several 10GbE ports that can be used as uplink ports to other switches or to support 10Gb edge devices (like a host or array). The number of 10GbE ports available and the socket/cable types supported will vary from vendor to vendor. In many cases, these switches will provide between two and four 10Gb ports. Older switch models may not support SFP+ as a connector type, while newer models may offer SFP+

as an option. Since most newer 10GbE switches and EqualLogic arrays use the SFP+ standard for connections, the 1GbE switches currently in use will need to have this option available as well.

If the existing 1Gb switches do not have integrated 10Gb ports, integration of 10Gb will not be possible without replacing the existing 1Gb switches. Options in this scenario include purchasing 10Gb switches that have options to integrate multiple 1Gb ports or that have dual speed sensing ports and optional 1000Base-T SFP+ modules, or looking at 10Gb switches that have a stacking technology that is compatible with existing 1Gb switches. In most cases, the new 10Gb switches will have to be from the same vendor as the existing 1Gb switches.

3.3.1 10Gb Switches

When considering candidate 10Gb switches, it is helpful to understand the options and features that are typically available in the current offerings. As 10Gb standards progress and mature, additional features will be added over time by all of the switch vendors, but initial implementations will have some restrictions. If you compare the past transition from 100 megabit Ethernet to 1GbE, you will see a gradual expansion of new features and technologies that were introduced over time. Initial 1GbE switches, when first introduced, had very limited features and capabilities. Almost all of the switches were non-stacking, layer 2 switches. Many switches at the time did not offer 1GbaseT connections, but used proprietary connections.

The transition to 10GbE switches appears to be following a similar course. 10Gb ports were initially introduced as uplink ports on 1Gb switches and used a variety of non-standard physical connection types such as XFP, XENPAK, and X2. As the 10GbE standard has matured, an industry migration to a more standard SFP+ solution has taken place and over the next several years, the standard will more than likely migrate to a standard 10Gbase-T connection type for copper connections similar to current 1Gbase-T switches today.

Likewise, the set of features on current 10Gb switches is somewhat limited. There are few vendors offering switches in the “edge” switch segment that offer more advanced features such as dedicated stacking functionality, layer 3 routing and larger port counts. To get these features with current offerings will require purchasing more advanced “aggregate” or “core” switches. This will present some challenges when designing larger SANs that require more ports as some of these ports will need to be used to inter-connect switches to each other in a redundant fashion. With these limitations, we need to also look at how these limitations affect the overall SAN design with respect to EqualLogic. The next several sections will discuss the use of Link Aggregation Groups to connect switches together, how Spanning Tree will affect these inter-switch links (ISLs) and how EqualLogic PS Series design also affects ISL and other SAN design requirements.

3.3.1.1 Link Aggregation and Inter-Switch Linking

Since most 10Gb switches currently do not provide any high bandwidth stacking functionality, these switches will need to be interconnected using the existing 10Gb ports on each switch in conjunction with industry standard link aggregation functionality to combine multiple ports into a single logical connection. This provides a single logical pipe that provides greater than 10 gigabits of bandwidth between switches. To create these link aggregation groups (LAG), switch vendors typically provide support for the link aggregation control protocol (LACP) as defined by IEEE standard 802.3ad (since upgraded to 802.3AX).

3.3.1.2 Spanning Tree

Spanning Tree Protocol (STP) is a protocol that ensures that an Ethernet network infrastructure has only one path to travel between two points in the Layer 2 network. It basically removes loops from the network which can result in packets being sent repeatedly to switch nodes in the network. As the number of switches is increased in a network to allow for the addition of new arrays or hosts, these new switches must be interconnected with the existing switches. To ensure that there is a redundant path between a host and each array, multiple links will be needed between these switches…inevitably creating physical loops within the SAN network infrastructure.

STP will block some of these physical links to ensure that there are not *logical* loops within the network. Figure 2 provides an example of how STP affects the logical network with three switches that provides redundant physical paths between the host and array. The result of the effects of STP is that there will be several ports on each switch dedicated to inter-switch connections that will not be used when running in a normal operational mode. The more switches in the network, the more inter-switch links that potentially will be blocked.

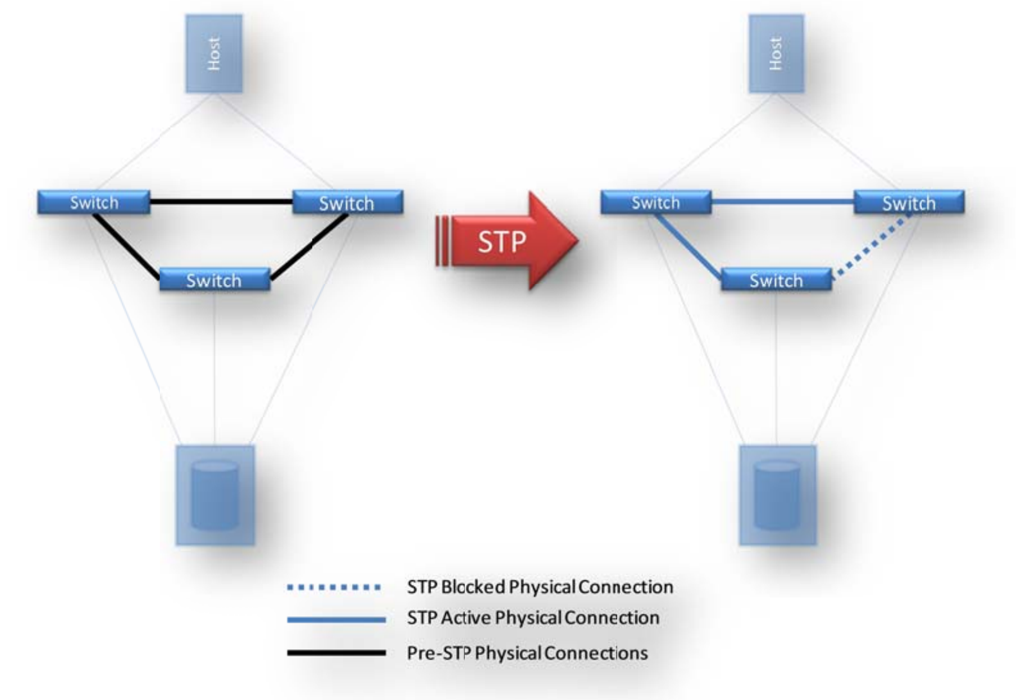

Figure 2: Effects of Spanning Tree on a Network

3.3.1.3 EqualLogic PS Series Network Requirements

Connecting an EqualLogic PS array to the SAN infrastructure will require connecting each controller to at least two different switches. Each controller on a 1GbE array will have between two Ethernet ports (PS4000) or more Ethernet ports (PS5000 & PS6000). Each controller on a 10GbE array will have two Ethernet ports that can be connected to the SAN infrastructure.

EqualLogic has the following requirements within the SAN:

- Within a given EqualLogic SAN group, all arrays must be on the same IP subnet

- For a routed layer 3 SAN infrastructure, hosts may be on different subnets from the SAN arrays
- All switches within the SAN must be interconnected such that there is always a path from any Ethernet port on one array to any other Ethernet port on all other arrays.
- When the SAN infrastructure consists of more than two switches, Rapid Spanning Tree Protocol must be enabled
- All switches and host network controllers within the infrastructure must have flow control enabled for optimal performance
- Jumbo frames may be enabled. If enabled, all switches and host network controllers must have jumbo frames enabled.
- Port requirements for fully redundant, maximum throughput EqualLogic arrays connection are as follows:
  - PS4000 Family – 2x 1GbE ports per controller = 4x 1GbE total ports
  - PS5x00 Family – 3x 1GbE ports per controller = 6x 1GbE total ports
  - PS6x00 Family – 4x 1GbE ports per controller = 8x 1GbE total ports
  - PS6x10 Family – 2x 10GbE ports per controller = 4x 10GbE total ports
- Port requirement for fully redundant iSCSI host connection is 2x ports per host. Ports can be either 1GbE or 10GbE

4 Mixed Speed SAN Infrastructure

4.1 Terminology

The following sections will illustrate and describe several SAN architecture configurations that might be considered when integrating 10Gb PS Series arrays into an existing SAN infrastructure consisting of 1Gb PS Series arrays. The actual viability of each design will depend on the features and capabilities of the candidate switch(s) and may not be suitable for all switch options.

Each illustration will use the following color scheme to ensure that each illustration is consistent:

Symbol/Color	Definition
	Indicates 1Gb Ethernet components such as cables, NICs, HBAs, or array controllers
	1Gb PS Series Array
	Indicates 10Gb Ethernet components such as cables, NICs, HBAs, or array controllers
	10Gb PS Series Array
	Indicates proprietary stacking connections as defined by switch vendor. Typically these stacking solutions provide high-bandwidth switch interconnects outside of the Ethernet standard protocols.
	Hosts connected to SAN
	Link Aggregation Group used as inter-switch link

Table 4: Diagram Symbol/Color Scheme

4.2 Recommended Connection Strategies

4.2.1 Attaching a single 10GbE Array to Existing 1GbE Modular Switches

Using the reference 1Gb PS Series SAN configuration defined earlier, one possible method of integrating a 10Gb PS array into the existing SAN infrastructure might be to utilize any available 10Gb "uplink" ports that might exist on the 1Gb switches. Several vendors integrate 2 or more 10GbE ports (usually as an optional expansion module) for use as inter-switch links between switches. If the switch vendor provides an SFP+ option for these ports, then they could be used to connect to a 10GbE PS array as illustrated in Figure 3. Depending on the number of 10GbE ports available and the internal design of the switch, it might be possible that one 10GbE array could be connected to the 1GbE switch infrastructure using these 10GbE ports. This configuration is only recommended for SAN groups with a need for only one 10GbE arrays as the 1GbE switches have not been designed for extensive 10GbE traffic.

Using standard Equallogic best practices on connecting an array to the SAN infrastructure in a fully redundant fashion will require that each port from the active controller be connected to two different switches and that each port from the standby controller be connected to two different switches (can be the same two switches as used to connect the active controller). This is illustrated in Figure 3 where the red solid lines represent the active network connections and the red dashed lines represent the connections to the standby controller.

> **Note:** This solution should be used for evaluation or very limited 10GbE storage access. Actual performance is very dependent on the internal design of the switch in terms of how the 10GbE ports are managed. Many 24-port solutions utilize a single switch fabric management computer chip (ASIC) to manage all of the ports. This typically means that this single ASIC is managing all of the "front-end" 1GBE ports, any stack interface, and any optional 10GbE ports. As these high-performance interfaces are used, they will typically cause a drop in overall performance of other interfaces.
>
> Remember that these switches were primarily designed as 1GbE switches, and not meant for extensive 10GbE edge device connection. Test this option thoroughly before any production deployments.

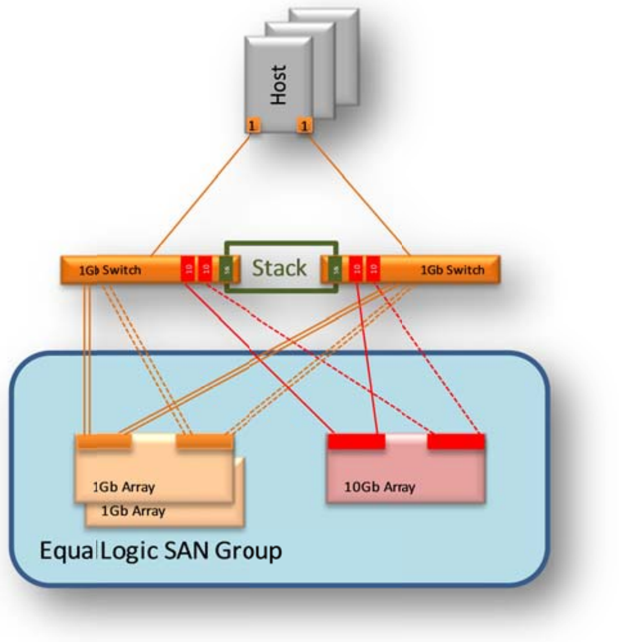

Figure 3: Single 10Gb Array Using 10Gb Uplink Ports on 1Gb Switches

4.2.2 Adding 10GbE Blades to Existing 1GbE Chassis Switches

When the existing iSCSI SAN infrastructure consists of scalable chassis switches, it may be possible to add new switch modules that provide 10GbE ports. Many 1Gb switch chassis vendors provide these 10GbE modules to support higher bandwidth interfaces for interconnecting switches in many data center core-edge architectures, but these same ports may be used for 10GbE edge connections such as 10GbE arrays or hosts. There are two important factors to consider with this solution.

First, does the switch support all the standard Ethernet and iSCSI protocols on these 10GbE ports? These ports must meet the requirements of a switch for use with EqualLogic arrays (see *PS Series Array Network Performance Guidelines* on equallogic.com). This includes support for jumbo frames, flow control, etc.

Second, and just as important, is considering oversubscription. Many chassis switches can have limitations on the size of the connection between the blade module and the chassis backplane. If there is a module that has ten 10GbE ports (100Gb of bandwidth within the module), but only has a module to chassis interface that can support 50Gb of bandwidth, then there is a 2:1 oversubscription between blade modules within the switch chassis. While this may or may not be a performance inhibitor, it is important to understand the amount of oversubscription within the chassis. The higher this oversubscription is the lower the overall potential performance of the SAN compared to a switch with lower oversubscription.

4.2.3 Adding Dedicated 10GbE Switches to the 1GbE Infrastructure

If more than one 10GbE arrays or if hosts with 10GbE NICs need to be connected to the SAN, then it is recommended that dedicated 10GbE switches be added to the current 1GbE SAN infrastructure. The ability to do this with existing 1GbE switches will depend on these switches having 10Gb uplink ports available to provide inter-switch connections between the 1GbE and 10GbE switches. Figure 4 illustrates how the 10GbE switches would leverage any existing 10GbE uplink ports on the 1GbE switches to allow for integration of the 10GbE SAN infrastructure.

As shown in Figure 4, with two 1GbE stacked switches and two 10GbE non-stacking switches, the following inter-switch links need to be created:

- LAG1(2 10GbE links):
  - 10Gb Switch #1 → 1Gb Switch #1
  - 10Gb Switch #1 → 1Gb Switch #2
- LAG2 (2 10GbE links):
  - 10Gb Switch #2 → 1Gb Switch #1
  - 10Gb Switch #2 → 1Gb Switch #2
- LAG3 (2 10GbE links)
  - 10Gb Switch #1 → 10Gb Switch #2
  - 10Gb Switch #1 → 10Gb Switch #2

As discussed in Section 3.3.1.2, Spanning Tree will require that one or more of the inter-switch connecting LAGs be blocked resulting in that LAG not being actively used for data transmission. Looking at Figure 4, the LAG descriptions above, it would be desirable to have either LAG1 or LAG2 be the "blocked" link. In fact, in Figure 4, LAG2 is illustrated in gray to represent the link being blocked by STP. It is preferable for one of these two links to be the blocked link to ensure that all of the 10Gb array ports can directly communicate with each other without traversing one of the 1GbE switches. Remember, that while the 1GbE switches have 10GbE ports, they are not designed as 10GbE switches and will not be able to handle high volumes of 10GbE traffic. It is better to isolate the majority of the 10GbE traffic within the 10GbE designed switches. The only traffic that should traverse the 1GbE-to-10GbE links is traffic between 1GbE hosts and 10GbE arrays (and vice versa) as well as any data movement from 1GbE arrays to 10GbE arrays.

While STP will block one of the links logically, the link does provide physical redundancy in the switch interconnects should one of the switches or links fail. If provided by the switch, use a "link cost" or "link weight" setting to assign each of the inter-switch LAGs a cost to help control which inter-switch LAG will get blocked. By assigning LAG3 a "cheaper" cost, it should remain active when the SAN is operating normally.

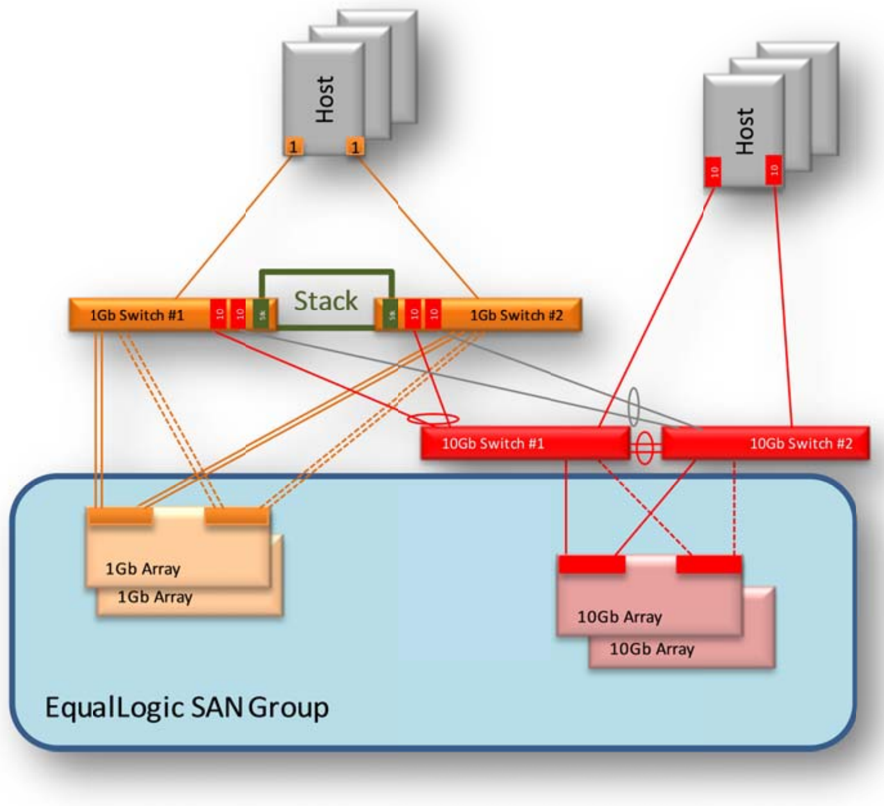

Figure 4: Connecting 10GbE Switches to Existing 1GbE SAN Infrastructure (2x 10GbE switches)

As the number of switches increases, the number of inter-switch links increases as does the number of blocked inter-switch links. As the need for larger numbers of 10GbE ports becomes greater, you should consider more advanced switch offerings that provide larger port counts such as a dedicated modular chassis switch or look at putting hosts and arrays on sets of switches or configuring the switch infrastructure as a Layer 3 routed SAN instead of the simpler Layer 2 SAN configuration.

Figure 5 illustrates a sample infrastructure where a 3rd 10GbE switch is needed. In this configuration, the 10Gb switches are configured into a "ring" using inter-switch links. By definition, a ring is a loop and loops are managed by spanning tree protocol. In fact, this configuration has two different loops, the original look between the 10GbE switches and the 1GbE switches, and the new loop created by the 10GbE ring. STP will block one of the links on the first loop and also block one of the links in the 10GbE ring. Any one of the three 10GbE switch → 10GbE switch links can be blocked without much difference in the performance.

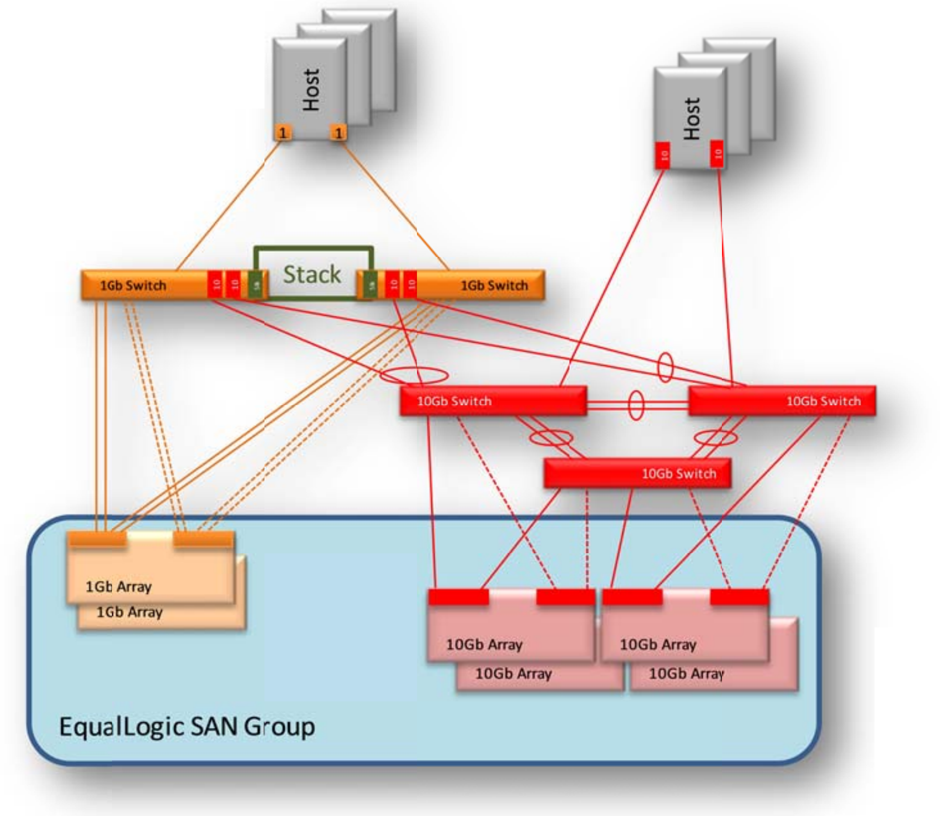

Figure 5: Connecting 10GbE Switches to Existing 1GbE SAN Infrastructure (3x 10GbE switches)

In a three switch configuration, as Figure 5 shows, distributing the array port-to-switch connections help distribute the load across the three 10GbE switches. If more than 3 switches is required, it might be advantageous to focus array connections onto a subset of the 10GbE switches and isolate the host-to-switch connections to a different subset of 10GbE switches. This is illustrated in Figure 6 below.

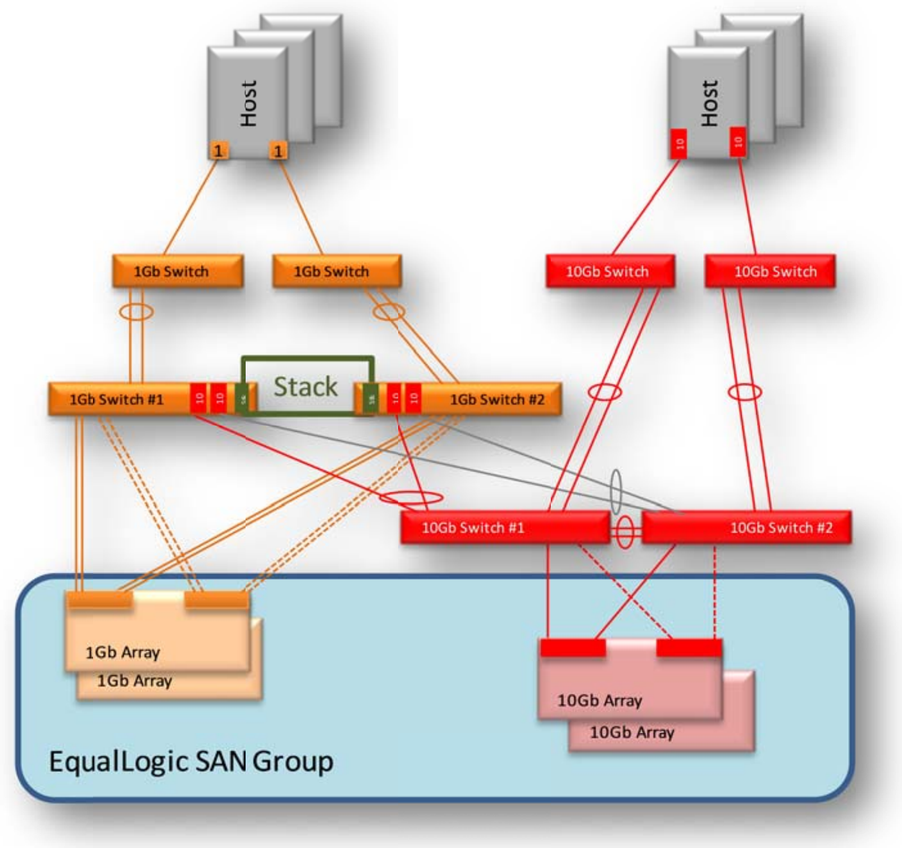

Figure 6: Example of a Multi-Switch SAN

4.3 What About Blades?

A blade server solution such as the PowerEdge M1000e actually works similarly to Figure 6 in the discussion above. The "host" switches are actually the PowerConnect M8024 10 Gigabit Blade IO Modules and external PowerConnect 8024F switches would be used in the "array" switches. Figure 7 illustrates a SAN configuration that consists of M1000e Blade Chassis with PowerConnect M8024 blade IO modules along with external PowerConnect 8024F 10GbE switches for 10GbE array connectivity. In Figure 7, M1000e blade chassis that contain 1GbE IO modules would connect to existing 1GbE switches and newer M1000e blade chassis that use the M8024 10GbE IO modules would connect to the external 8024F switches as illustrated. As in the previous configuration, this allows the EqualLogic SAN group to provide storage services to both 1GbE and 10GbE blade server hosts.

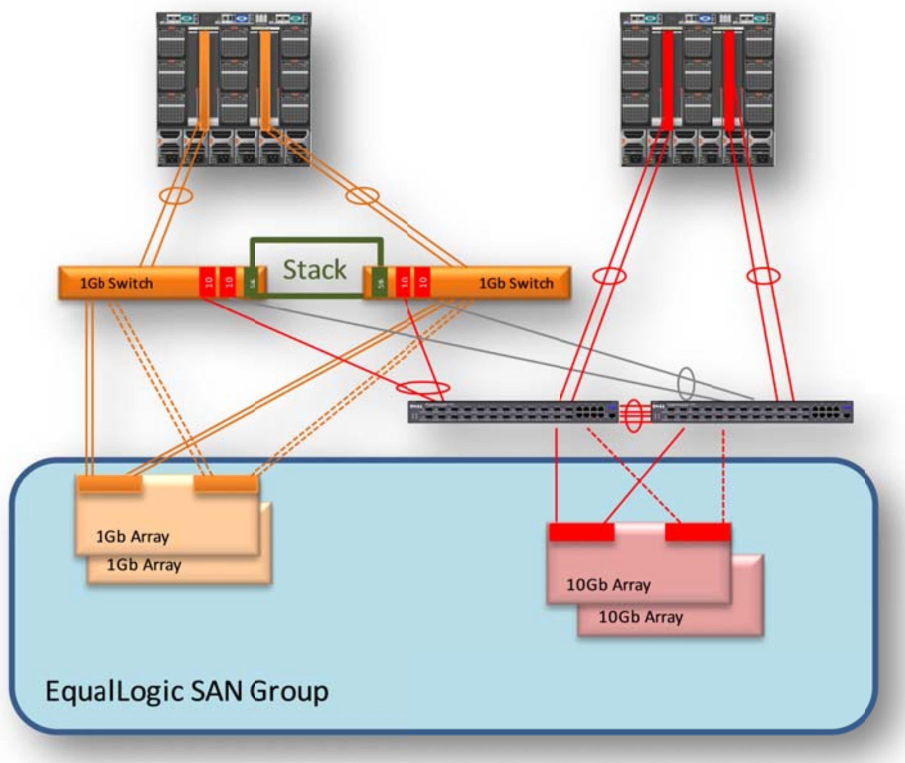

Figure 7: SAN Configuration Using M1000e Blade Chassis

5 Conclusion

One of the great value propositions of the EqualLogic PS Series "scale out" deployment model is the ability to easily and quickly integrate the latest storage technologies offered by the product into an existing SAN group composed of older, even multi-generation arrays. With the introduction of 10GbE arrays, Dell has again raised the bar on performance and scalability that administrators can take advantage of by implementing some straightforward network modifications.

Introducing 10Gb Ethernet into existing SAN infrastructures may not be an obvious choice or one that requires administrators to consider deploying separate network infrastructures, thus creating more complex deployments, and even more complex storage management strategies, but that is not necessarily the case. By understanding some basic concepts, 10GbE switches can be integrated together with 1GbE switches to create a multiple speed SAN allowing the use of both mainstream 1GbE technology and newer, higher performance 10GbE technology.

Important factors in designing a strategy for integrating 10GbE components into your existing 1GbE SAN should be considered. Most importantly, being aware of oversubscription of the various links that connect the switches together and understanding the affects of spanning tree. Depending on the switches being, the features they provide, and the internal design of the switches will dictate what will actually be possible and how you would interconnect the switches. Not all of the recommendations discussed here will work with all switches, but understanding the concepts here should provide you with some guidance and enable your SAN design to be a success.