Dell PowerEdge M1000e Blade Enclosure and Dell PS Series SAN Design Best Practices Using Dell S-Series and M-Series Networking

Dell EMC Storage Engineering
May 2017

Dell EMC Best Practices

Revisions

Date	Description
February 2013	Initial release
May 2017	Minor updates

Acknowledgements

This white paper was produced by the PG Storage Infrastructure and Solutions team of Dell Inc.

The team that created this white paper:

Clay Cooper, Guy Westbrook, and Camille Daily

Updates provided by:

Steven Lemons

The information in this publication is provided "as is." Dell Inc. makes no representations or warranties of any kind with respect to the information in this publication, and specifically disclaims implied warranties of merchantability or fitness for a particular purpose.

Dell believes the information in this document is accurate as of its publication date. The information is subject to change without notice.

Table of Contents

1 Introduction

Dell™ PS Series arrays provide a storage solution that delivers the benefits of consolidated networked storage in a self-managing iSCSI storage area network (SAN) that is affordable and easy to use, regardless of scale. By eliminating complex tasks and enabling fast and flexible storage provisioning, these solutions dramatically reduce the costs of storage acquisition and ongoing operation.

A robust, standards-compliant iSCSI SAN infrastructure must be created to leverage the advanced features provided by a PS Series array. When using blade servers in a Dell PowerEdge™ M1000e blade enclosure (also known as a blade chassis) as a host, there are a number of network design options for storage administrators to consider when building the iSCSI SAN. For example, one option would be to connect the PS Series ports to the modular switches within the M1000e blade chassis while another option would be to connect blade server network ports (by their corresponding networking module switch ports) to Top of Rack (ToR) switches residing outside of the M1000e blade chassis. After testing and evaluating a variety of different SAN design options, this technical white paper quantifies the ease of administration, the performance, the high availability, and the scalability of each design. From the results, recommended SAN designs and practices are presented.

The SAN designs in this paper converge both SAN and LAN into a single network fabric using Data Center Bridging (DCB). The final SAN design is the pre-integrated Active System 800, the first Active System offering from the Dell Active Infrastructure family of converged infrastructure offerings. It consists of PowerEdge M620 blade servers in a PowerEdge M1000e blade enclosure with M-Series I/O Aggregator modules uplinked to Dell EMC Networking S4810P ToR switches and PS Series PS6110 array members. Active System 800 is deployed and managed by the Active System Manager application.

For more information on the PowerEdge M1000e blade enclosure solution, PS Series SAN architecture, Data Center Bridging and Dell Active Infrastructure see Section 2 titled, “Concept Overview”.

1.1 Audience

This technical white paper is intended for storage administrators, SAN/NAS system designers, storage consultants, or anyone who is tasked with integrating a Dell M1000e blade chassis solution with PS Series storage for use in a production storage area network. It is assumed that all readers have experience in designing and/or administering a shared storage solution. Also, there are some assumptions made in terms of familiarity with all current and possibly future Ethernet standards as defined by the Institute of Electrical and Electronic Engineers (IEEE) as well as TCP/IP and iSCSI standards as defined by the Internet Engineering Task Force (IETF).

1.2 Terminology

This section defines terms that are commonly used in this paper and the context in which they are used.

Blade I/O module (IOM) switch – A switch that resides in an M1000e Fabric slot.

Blade IOM switch only – A category of SAN design in which the network ports of both the hosts and the storage are connected to the M1000e blade IOM switches, which are isolated and dedicated to the SAN. No external ToR switches are required. The switch interconnect can be a stack or a LAG and no uplink is required. The Dell EMC Networking MXL supports Virtual Link Trunking (VLT): mVLT and L2/L3 over VLT.

Blade IOM switch with ToR switch – A category of SAN design in which host network ports are internally connected to the M1000e blade IOM switches and storage network ports are connected to ToR switches. A switch interconnect stack, LAG or VLTi between each ToR switch is required. An uplink stack, LAG, or VLT LAG from the blade IOM switch tier to the ToR switch tier is also required.

Converged network adapter (CNA) – Combines the function of a SAN host bus adapter (HBA) with a general-purpose network adapter (NIC).

Data Center Bridging (DCB) – An enhancement of the IEEE 802.1 bridge specifications for supporting multiple protocols and applications in the data center. It supports converged infrastructure implementations to carry applications and protocols on a single physical infrastructure.

Data Center Bridging Exchange (DCBX) – Protocol standard for the discovery and propagation of DCB configuration between DCB-enabled switches, storage array members, and CNA.

Host/storage port ratio – The ratio of the total number of host network interfaces connected to the SAN divided by the total number of active PS Series array member network interfaces connected to the SAN. A ratio of 1:1 is ideal for optimal SAN performance, but higher port ratios are acceptable in specific cases. The host/storage port ratio can negatively affect performance in a SAN when oversubscription occurs, that is when there are significantly more host ports or significantly more storage ports.

Link aggregation group (LAG) – Where multiple switch ports are configured to act as a single high-bandwidth connection to another switch. Unlike a stack, each individual switch must still be administered separately and function as such.

Multiple switch tier SAN design – A SAN design with both blade IOM switches and ToR switches. Host and storage ports are connected to different sets of switches and an uplink stack, LAG, or VLT LAG is required. Blade IOM switch with ToR switch designs are multiple switch tier SAN designs.

Single switch tier SAN design – A SAN design with only blade IOM switches or ToR switches but not both. Both host and storage ports are connected to the same type of switch and no uplink is required. Blade IOM switch only and ToR switch only designs are single switch tier SAN designs.

Stack – An administrative grouping of switches that enables the management and functioning of multiple switches as if they were one single switch. The switch stack connections also serve as high-bandwidth interconnects.

Switch interconnect – An inter-switch link that connects either the two blade IOM switches or the two ToR switches to each other. A switch interconnect unifies the SAN fabric and facilitates inter-array member communication. It can be a stack, a LAG, or a VLTi.

Switch tier – A pair or more of like switches connected by a switch interconnect which together create a redundant SAN Fabric. A switch tier might accommodate network connections from host ports, from storage ports, or from both. If all switches in a switch tier are reset simultaneously, for example the switch tier is stacked and the firmware is updated, then the SAN is temporarily offline.

ToR switch – A top of rack switch, external to the M1000e blade chassis.

ToR switch only – A category of SAN design in which the network ports of both the hosts and the storage are connected to external ToR switches. For this architecture, 10 GbE pass-through IOM are used in place of blade IOM switches in the M1000e blade chassis. The switch interconnect can be a stack, a LAG, or a VLTi.

Uplink – A link that connects the blade IOM switch tier to the ToR switch tier. An uplink can be a stack, a LAG, or a VLT LAG. Its bandwidth must accommodate the expected throughput between host ports and storage ports on the SAN.

Virtual link trunking (VLT) – A Dell EMC Networking feature which enables the LAG of an edge device to link to two separate upstream switches (referred to as VLT peer switches) in a loop-free topology without the need for a Spanning Tree protocol.

VLT domain – Refers to the VLT peer switches, the VLT interconnect, and any VLT LAG.

VLT LAG – A port channel which links VLT peer switches to edge devices.

VLTi – The VLT interconnect used to synchronize states between VLT peer switches.

2 Concept Overview

2.1 PowerEdge M1000e blade chassis solution

The following section describes the M1000e blade chassis networking Fabrics consisting of I/O modules, a midplane, and the individual blade server network adapters.

2.1.1 Multiple Fabrics

Each M1000e can support up to three separate networking Fabrics that interconnect ports on each blade server to a pair of blade I/O modules within each chassis Fabric through a passive chassis midplane. Each Fabric is associated with specific interfaces on a given blade server. Each blade server has a LAN on Motherboard (LOM) or a Network Daughter Card (NDC) that is mapped to the blade IOM located in the Fabric A slots of the M1000e chassis. In addition, each blade server has two mezzanine sockets for adding additional networking options such as 1 Gb or 10 Gb Ethernet, Infiniband, or Fibre Channel cards. These mezzanine cards are mapped to either the Fabric B or the Fabric C blade IOM.

Figure 1 illustrates the layout of the three Fabric blade IOMs located in the back of the M1000e chassis.

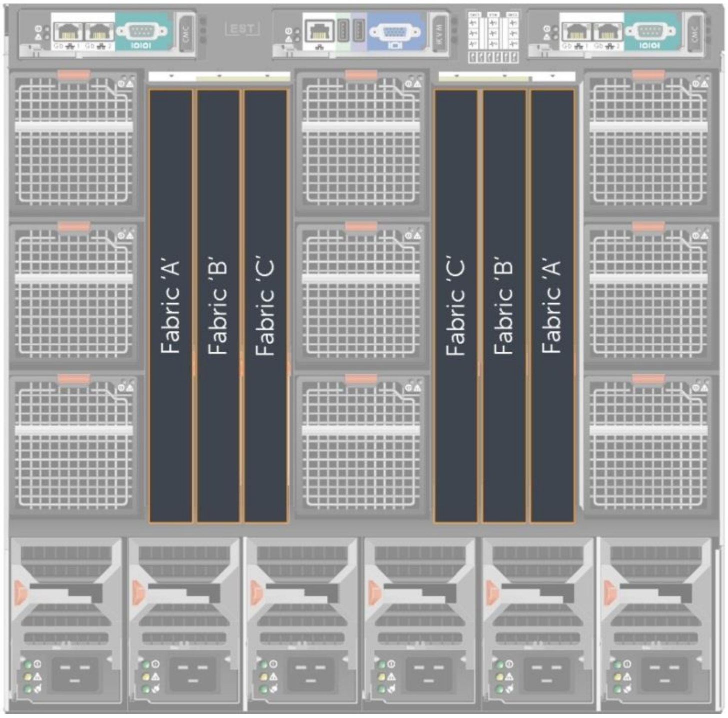

Figure 1 Blade I/O Modules and M1000e Chassis

2.1.2 M-Series Blade I/O modules

The following table lists the 1 GbE, 10 GbE & 40 GbE M-Series blade I/O module options and the number of ports available for PS Series SAN solution.

Table 1 1 GbE, 10 GbE & 40 GbE M-Series Blade I/O Module Port options for PS Series storage

	Internal blade server ports	1/10GbE (Base-T) External Ports	10 GbE External Ports	40 GbE External Ports
MXL	32 (10GbE)	4 (using module)	8 ports using QSFP+ breakout cables (up to 24 using modules)	2 integrated QSFP+ (up to 6 using modules)
I/O Aggregator	32 (10GbE)	4 (using module)	8 ports using QSFP+ breakout cables (up to 16 using modules)	2 integrated QSFP+ fixed in breakout mode (up to 6 using modules)
M8024-k	16 (10GbE)	2 (using module)	4 fixed SFP+ ports (1/10GB) (Add 4 more 10Gb ports using module)	N/A
M6348	32 (1GbE)	16 fixed (1GbE)	2 fixed SFP+ and 2 fixed CX4	N/A
M6220	16 (1GbE)	4 fixed (1GbE)	4 (using modules)	N/A
10 Gb Pass-through	16 (10GbE)	N/A	16 fixed SFP+ (supports 10GbE only)	N/A

2.2 PS Series SAN architecture

Figure 1 illustrates the basic SAN components involved when deploying an M1000e blade chassis with blade servers into a PS Series array SAN. When creating the SAN to connect blade server network ports to storage array member network ports, the SAN might consist of only blade IOM switches, only ToR switches, or both switch types together in two separate tiers. Note that the blade servers connect to the blade IOM switches internally with no cabling required. Blade servers can also be directly connected to ToR switches if the blade IOM switches are replaced with pass-through IOM.

It is a best practice to create a redundant SAN fabric using at least two switches, with PS Series array members and storage hosts each having connections to more than one switch. To unify the SAN fabric, the switches must be interconnected to create a single layer 2 SAN Fabric over which all PS Series array member network ports and host ports can communicate with each other. This switch interconnection can be a stack, a LAG, or a VLT interconnect (VLTi). SAN traffic will often have to cross the switch interconnect since iSCSI connections are distributed across all available host network ports by the PS Series Host Integration Toolkit. Assuming a worst case scenario of 100% of all SAN traffic crossing the switch interconnect in both directions (half going one way and half going the other) the interconnect bandwidth requirements are 50% of the aggregate bandwidth of all active PS Series array member ports. For example, four array members with one active 10 GbE port each would require 20 Gbps of interconnect bandwidth.

If there are blade IOM switches and ToR switches in the SAN, these two switch tiers will need to be connected by an uplink, which can be a stack (if there is stack compatibility between the blade IOM and ToR switch types), a LAG, or a VLT LAG (if there is a VLT interconnect between the ToR switches). Uplink bandwidth should be at least equal to the aggregate bandwidth of all active PS Series array member ports.

For a detailed description of each SAN design tested and evaluated to produce this paper, see Section 4 titled, “Tested SAN designs”.

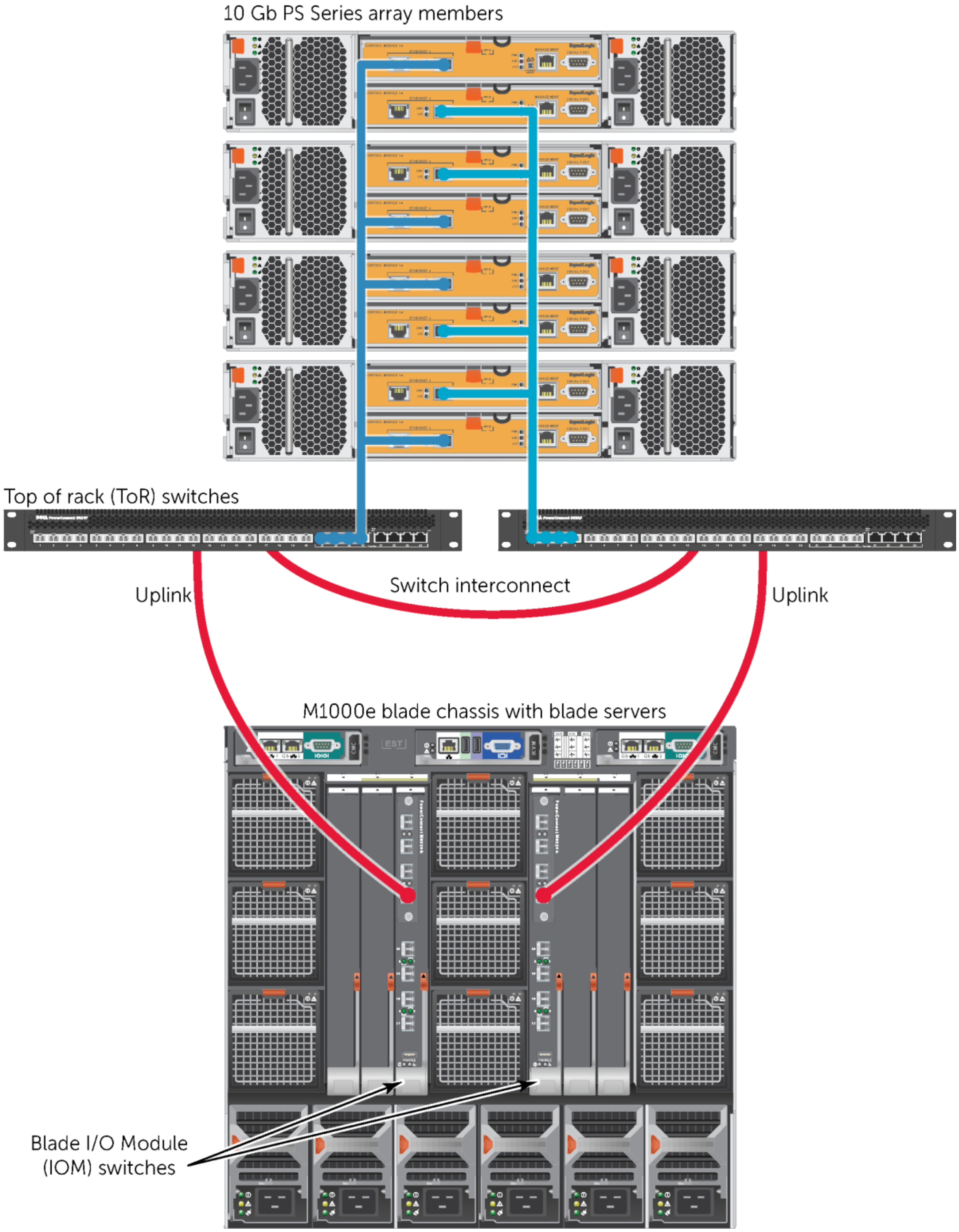
10 Gb PS Series array members
Top of rack (ToR) switches
Uplink
Switch interconnect
Uplink
M1000e blade chassis with blade servers
Blade I/O Module
(IOM) switches

Figure 2 An example PS Series SAN consisting of PS Series array members, an M1000e blade chassis with blade servers, and ToR and Blade IOM switches.

2.3 Data Center Bridging

DCB standards are enhancements to IEEE 802.1 bridge specifications to support multiple protocols and applications in the data center. They support converged infrastructure implementations to carry applications and protocols on a single physical infrastructure. For more information on using DCB with PS Series SANs, see the following resources:

- *Dell PS Series DCB Configuration Best Practice*:

  http://en.community.dell.com/dell-groups/dtcmedia/m/mediagallery/20305369

- Best Practices for Configuring DCB with Windows Server and EqualLogic Arrays:

  http://en.community.dell.com/techcenter/extras/m/white_papers/20438162

2.3.1 IEEE DCB technologies

- Data Center Bridging Exchange Protocol (DCBX), IEEE 802.1Qaz – Discovery and configuration of devices that support PFC, ETS
- Enhanced Transmission Selection (ETS), IEEE 802.1Qaz - bandwidth allocation per traffic class and sharing
- Priority Based Flow Control (PFC), IEEE 802.1Qbb – ability to pause individual traffic classes and enable lossless behavior
- Quantized Congestion Notification (QCN), IEEE 802.1Qau - end-to-end flow control in a L2 network to eliminate sustained congestion caused by long lived flows

2.3.2 Minimum requirements for iSCSI in converged I/O environments with DCB

Table 2 The minimum requirements for iSCSI in converged I/O environments with DCB

DCB Technology	Requirement	Standards Version	Purpose
DCBX	Required with support for iSCSI application priority	IEEE 802.1Qaz-2011	Discovery, configuration and mismatch resolution
ETS	Required with iSCSI mapped to dedicated traffic class or priority group	IEEE 802.1Qaz-2011	Minimum bandwidth allocation per traffic class during contention and additional bandwidth allocation during non-contention
PFC	Required with PFC turned on for iSCSI along with lossless queue support	IEEE 802.1Qbb-2011	Independent traffic priority pausing and enablement of lossless traffic classes

2.4 Dell Active Infrastructure

The Dell Active Infrastructure is a family of converged infrastructure offerings that combine servers, storage, networking, and infrastructure management into an integrated system that provides general purpose virtual resource pools for applications and private clouds. These systems blend intuitive infrastructure management, an open architecture, flexible delivery models, and a unified support model to allow IT to rapidly respond to dynamic business needs, maximize efficiency, and strengthen IT service quality.

Designed from the ground up as a converged infrastructure system, Active System integrates new unified infrastructure management, a new plug-and-play blade chassis I/O module – the PowerEdge M I/O Aggregator, modular servers and storage, and a converged LAN/SAN fabric. Key to Active System is Active System Manager, an intelligent and intuitive converged infrastructure manager that leverages templates to automate infrastructure on-boarding and re-configuration. This automation greatly simplifies and speeds up operations while also significantly reducing errors associated with manual configuration. The result is better infrastructure quality with fewer costly configuration errors.

2.4.1 Dell Active System Manager

Active System Manager is a workload-centric converged infrastructure manager that streamlines infrastructure configuration and on-going management. Active System Manager is a member of the Active Infrastructure Family and supports Active System compliant environments, which may be pre-integrated, based on a reference architecture or custom built using the Active System Matrix.

Below is a list of technical capabilities that customers can expect:

- Template-based provisioning and automated configuration to easily encapsulate infrastructure requirements and then predictably apply those requirements based on workload needs
- Management of the entire lifecycle of infrastructure, from discovery and on-boarding through provisioning, on-going management and decommissioning
- Workload failover, enabling rapid and easy migration of workload to desired infrastructure resources
- Wizard-driven interface, with feature-guided, step-by-step workflows
- Graphical logical network topology view and extended views of NIC partitions

2.4.2 Active System Manager

With its newest release, Active System Manager (ASM) continues to expand on ASM's intuitive user interface and its open and extensible architecture. In addition to Active System Manager for intuitive infrastructure management, and PowerEdge I/O Aggregator for simplified networking and a converged fabric (DCB over iSCSI), ASM supports PS Series storage showcasing fluid data and the latest PowerEdge blade servers for modular compute, all inside the most power efficient blade chassis on the market today, the M1000e. It provides an ideal foundation for private cloud and comes complete with turnkey integration services and unified single-number support. For more information on Active System Manager, see the *Dell Active System Manager Wiki*: http://www.dell.com/asmtechcenter

3 Summary of SAN designs and recommendations

This section provides the high level conclusions reached after the course of comprehensive lab testing and analysis of various PS Series array SAN designs which incorporate M1000e blade server hosts on a 10 GbE network. The following assumptions were made:

- Two SAN ports per host
- Two M-Series MXL, two M-Series I/O Aggregators, or two M-Series 10GbE pass-through I/O modules per blade chassis
- Tow Dell EMC Networking S4810P ToR switches (if the design includes ToR switches)

For complete results and recommendations see Section 5 titled, “Detailed SAN design analysis and recommendations”. For an illustration of each SAN design see Section 4 titled, “Tested SAN designs”.

Table 3 Summary of SAN designs and recommendations

	Switch tier topology	**Administration**	**Performance**	**High availability**	**Scalability**
MXL with LAG interconnect	Single	• No ToR switches required • Fewest cables • VLT not supported	• Equivalent performance and bandwidth	• Blade IOM switch failure reduces host ports by 50%	• Supports up to 16 array members using 2x 40 GbE expansion modules per blade switch • 4x 40 GbE ports available
S4810P with VLTi	Single	• No blade IOM switches required • VLT supported	• Equivalent performance and bandwidth	• ToR switch failure reduces host ports by 50%	• Supports up to 16 array members with no expansion modules required • 32x 10 GbE and 4x 40 GbE ports available
S4810P with VLTi / MXL with VLT LAG uplinks	Multiple	• Four switches to manage • Uplink configuration required	• Equivalent performance and bandwidth	• VLT LAG preserves host connectivity during ToR switch failure	• Supports up to 16 array members with no expansion modules required • 48x 10 GbE and 4x 40 GbE ports available

3.1 Administration

When reducing administrative overhead is the goal, a single switch tier design with a stacked interconnect is the simplest option. Because the storage is directly attached to the blade IOM switches, fewer cables are required than with the ToR switch only design, and the stacked interconnect allows the switches to be administered as a single switch.

If the availability of the SAN is critical, a LAG or VLTi interconnect will be preferred over stacking. If a switch interconnect is stacked, then a switch stack reload (required for tasks such as switch firmware updates) will temporarily make the SAN unavailable. In this case, SAN downtime for firmware updates would have to be scheduled.

DCB configuration should be configured at a single source switch at the core, aggregation, or ToR switch tier and allowed to flow down via DCBX to blade IOM switches, CNAs, and PS Series array members.

If ToR switches from a different vendor are used, the simplest choice is to implement the ToR only design by cabling M1000e pass-through IOM directly to the ToR switches. If multiple switch tiers are desired, plan for an uplink LAG using the high bandwidth ports of the blade IOM switches.

Taking advantage of the simplified behavior of the M-Series I/O Aggregator and features of Active System Manager can provide a solid foundation for a virtualized private cloud.

3.2 Performance

The throughput values were gathered during the performance testing of each SAN design with four hosts and four array members at three common workloads. Among all SAN designs, there were no significant performance differences during any of the three tested workloads.

3.3 High availability

ToR switch failures always collapse the fabric to a single switch as array member network ports failover to the remaining ToR switch. Host connectivity can be preserved during a ToR switch failure with redundant VLT LAG uplinks from the blade IOM switches. This is made possible by having a VLTi interconnect between the ToR switches, rather than a standard LAG. Stacked interconnects should be avoided because the SAN becomes unavailable during a switch stack reload.

Blade IOM switch failures always result in a loss of 50% of the host ports and in multiple-tier SAN designs a 50% loss in uplink bandwidth.

3.4 Scalability

All tested SAN designs can support up to 16 array members and provide adequate bandwidth within the SAN. While the Blade IOM only SAN design has no ports remaining for additional connectivity or uplinks when using 16 array members, the other three SAN designs have ample ports remaining for additional edge device connectivity and high bandwidth uplinks to a core switch. Considering the fact that the S-Series S4810 supports VLT and doesn’t require expansion modules, the simple ToR switch only SAN design is an excellent option. It creates a robust aggregation layer that accepts highly available VLT LAG from downstream edge devices and switches and also from the upstream Layer 3 core.

Note: The scalability data presented in this paper is based primarily on available port count. Actual workload, host to array port ratios, and other factors may affect performance.

4 Tested SAN designs

This section describes each tested M1000e blade chassis SAN design in detail including diagrams and a table for comparison of important values such as bandwidth, maximum number of supported array members, and the host to storage port ratio. The information below assumes a single M1000e chassis and 16 half-height blade servers with two network ports each.

There are three categories of SAN designs for M1000e blade chassis integration:

1. **Blade IOM switch only** – Network ports of both the hosts and storage are connected to the M1000e blade IOM switches. No ToR switches are required. The switch interconnect can be a stack or a LAG, and no uplink is required.
2. **ToR switch only** – Network ports of both the hosts and the storage are connected to external ToR switches. 10 GbE pass-through IOM switches are used in place of blade IOM switches in the M1000e blade chassis. The switch interconnect can be a stack, a LAG, or a VLTi.
3. **Blade IOM switch with ToR switch** – Host network ports are connected to the M1000e blade IOM switches and the storage network ports are connected to ToR switches. The switch interconnect can be a stack, a LAG, or a VLTi and should connect the ToR switch to better facilitate inter-array member traffic. An uplink stack, LAG or VLT LAG from the blade IOM switch tier to the ToR switch tier is also required.

4.1 Blade IOM switch only

This SAN design category includes configurations in which the PS Series array member ports are directly connected to the blade IOM switch ports within the blade chassis. For this SAN design, dual M-Series MXL switches in the M1000e chassis were used.

4.1.1 M-Series MXL with LAG interconnect

This SAN design provides 80 Gbps of interconnect bandwidth between the two MXL switches using two integrated 40 GbE QSFP ports on each switch to create a LAG. Since there is only a single tier of switches, there is no uplink to ToR switches. Using two 40 GbE QSFP expansion modules per MXL switch and 40 GbE to four 10 GbE breakout cables allows sixteen 10 GbE SFP+ external connections on each switch (32 total) which can accommodate the connection of sixteen 10 GbE PS series array members, each of which requires two ports for the active and passive controllers combined. However, when using 16 array members there will be no remaining 40 GbE QSFP ports available on the MXL switch for creating an uplink to a core switch, and the SAN itself would isolated. The host/storage port ratio with the maximum number of array members is 2:1.

The following diagram illustrates how four PS6110XV array members directly connect to the two MXL switches in Fabric C of the M1000e blade chassis and how the two MXL switches are connected by a LAG using two 40 GbE QSFP ports on each switch. This network design requires the use of two 40 GbE expansion modules in each of the MXL switches. Note that the port on the passive storage controller is connected to a different switch than the port on the active storage controller, ensuring that the port-based failover of the PS6110 array member will connect to a different switch upon port, cable or switch failure. The management network is shown for reference.

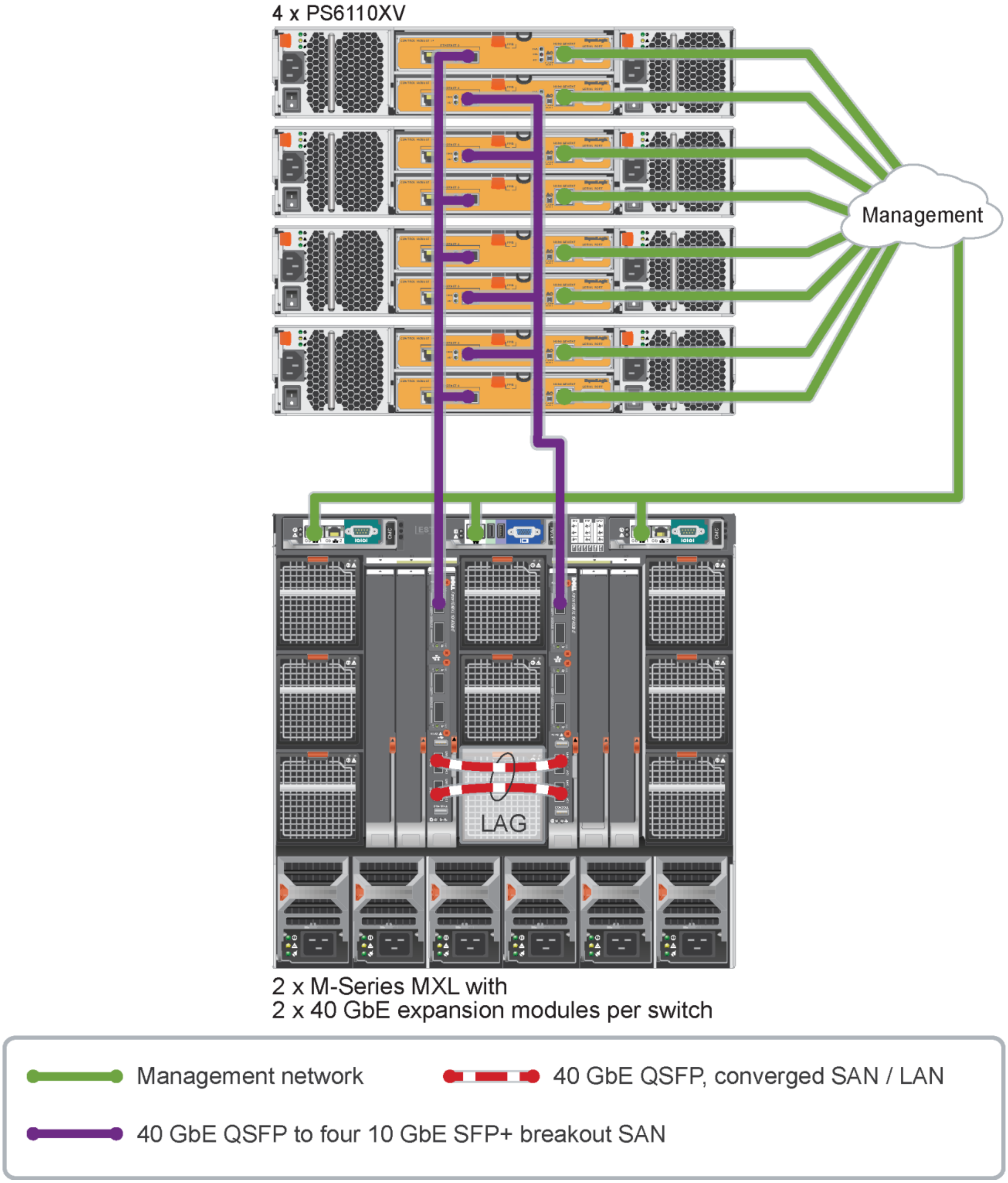
4 x PS6110XV
Management
LAG
2 x M-Series MXL with
2 x 40 GbE expansion modules per switch
Management network
40 GbE QSFP, converged SAN / LAN
40 GbE QSFP to four 10 GbE SFP+ breakout SAN

Figure 3 M-Series MXL with LAG interconnect

4.2 ToR switch only

These SAN designs include configurations where the storage ports and blade server host ports are directly connected to ToR switches. A 10 GbE pass-through IOM rather than a blade switch is used to connect the host ports to the ToR switches. For this SAN design, dual S-Series S4810P switches were used as the ToR switch.

Because a stacked interconnect makes the SAN unavailable during stack reloads and because the S4810P switches support VLT, a VLTi was used to connect the two ToR switches. When operating as VLT domain peers, the ToR switches appear as a single virtual switch from the point of view of any connected switch or server supporting Link Aggregation Control Protocol (LACP).

4.2.1 S-Series S4810P with VLTi

This SAN design provides 80 Gbps of interconnect bandwidth between the two S4810P switches using two 40 GbE QSFP ports on each switch to create a VLTi. Since the pass-through module is connecting to the ToR switches separately for each blade host there is no uplink required. Sixteen 10 GbE SFP+ ports on each ToR switch are required for the connections of the 16 hosts with two network ports each. The remaining 32 ports on each S4810P (64 ports total) can easily accommodate the connection of sixteen 10 GbE PS 6110 Series array members, each of which requires two ports for the active and passive controllers combined. Two 40 GbE QSFP ports per S4810P switch are available for uplinks to a core switch. The host/storage port ratio with the maximum number of array members is 2:1.

The following diagram illustrates how four PS6110XV array members directly connect to the two ToR S4810P switches and how the two switches are connected by a VLTi using two 40 GbE QSFP ports on each switch. It also shows the connection of four server blades each with two host ports to the S4810P switches using the 10 GbE pass-through IOM in Fabric C of the M1000e chassis. Note that the port on the passive storage controller is connected to a different switch than the port on the active storage controller. This ensures that the port-based failover of the PS6110 array member will connect to a different switch upon port, cable, or switch failure. The management network is shown for reference.

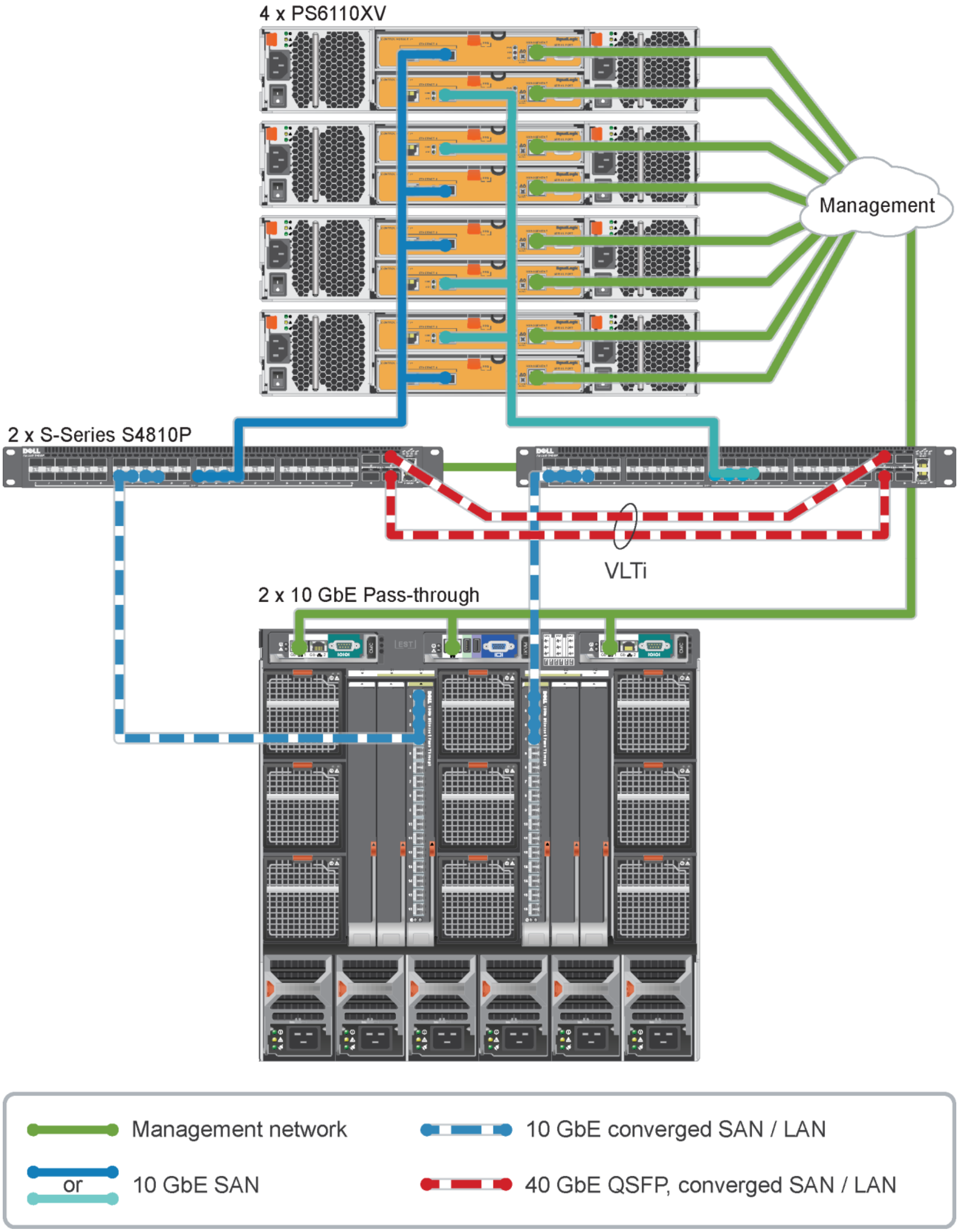

Figure 4 S-Series S4810P with VLTi

4.3 Blade IOM switch with ToR switch

These SAN designs include configurations in which the PS Series array member ports are connected to a tier of ToR switches while the server blade host ports are connected to a separate tier of blade IOM switches in the M1000e blade chassis.

With the multiple switch tier designs, it is a best practice to connect all array member ports to the ToR switches and not the blade IOM switches in the M1000e chassis. This allows the M1000e chassis to scale independently of the array members. The ToR switches rather than the blade IOM switches are interconnected by a stack, a LAG, or a VLTi to better facilitate inter-array member communication. The switch tiers themselves are connected by an uplink stack, LAG, or VLT LAG.

For the two SAN designs in this category, S-Series S4810P switches were used as the ToR switches while two different blade IOM switches were tested – the M-Series MXL and the M-Series I/O Aggregator.

Note that because the S4810P switch is not stack compatible with either the MXL or the I/O Aggregator, SAN designs with stacked uplinks were not possible.

Because a stacked interconnect makes the SAN unavailable during stack reloads and because the S4810P switches support VLT, a VLTi was used to connect the two ToR switches. When operating as VLT domain peers, the ToR switches appear as a single virtual switch from the point of view of any connected switch or server supporting LACP.

4.3.1 S-Series S4810P with VLTi / MXL with VLT LAG uplinks

This SAN design uses two of the four integrated 40 GbE QSFP ports on each S-Series S4810P to create a VLTi between each ToR switch. 40 GbE to four 10 GbE breakout cables were used to uplink the M-Series MXL switches to the ToR switches to save two integrated 40 GbE QSFP ports on each S-Series S4810P for VLT LAG uplinks to core switches. Because the ToR switches are operating as VLT domain peers, each MXL LAG can connect to both ToR switches in a redundant but loop-free topology.

This SAN design provides 160 Gbps of uplink bandwidth while providing 80 Gbps of interconnect bandwidth, more than enough to accommodate the maximum theoretical traffic of sixteen 10 GbE PS Series array members. Even with 16 array members, each of which requires two ports for the active and passive storage controllers combined, switch port count is not an issue. The host/storage port ratio with the maximum number of array members is 2:1.

The following diagram illustrates how four PS6110XV array members connect to the two ToR S4810P switches, how the ToR switches are interconnected with a VLTi, and how each MXL is connected to both ToR switches with a VLT LAG. Note that the port on the passive storage controller is connected to a different switch than the port on the active storage controller, ensuring that the port-based failover of the PS6110 array member will connect to a different switch upon port, cable, or switch failure. The management network is shown for reference.

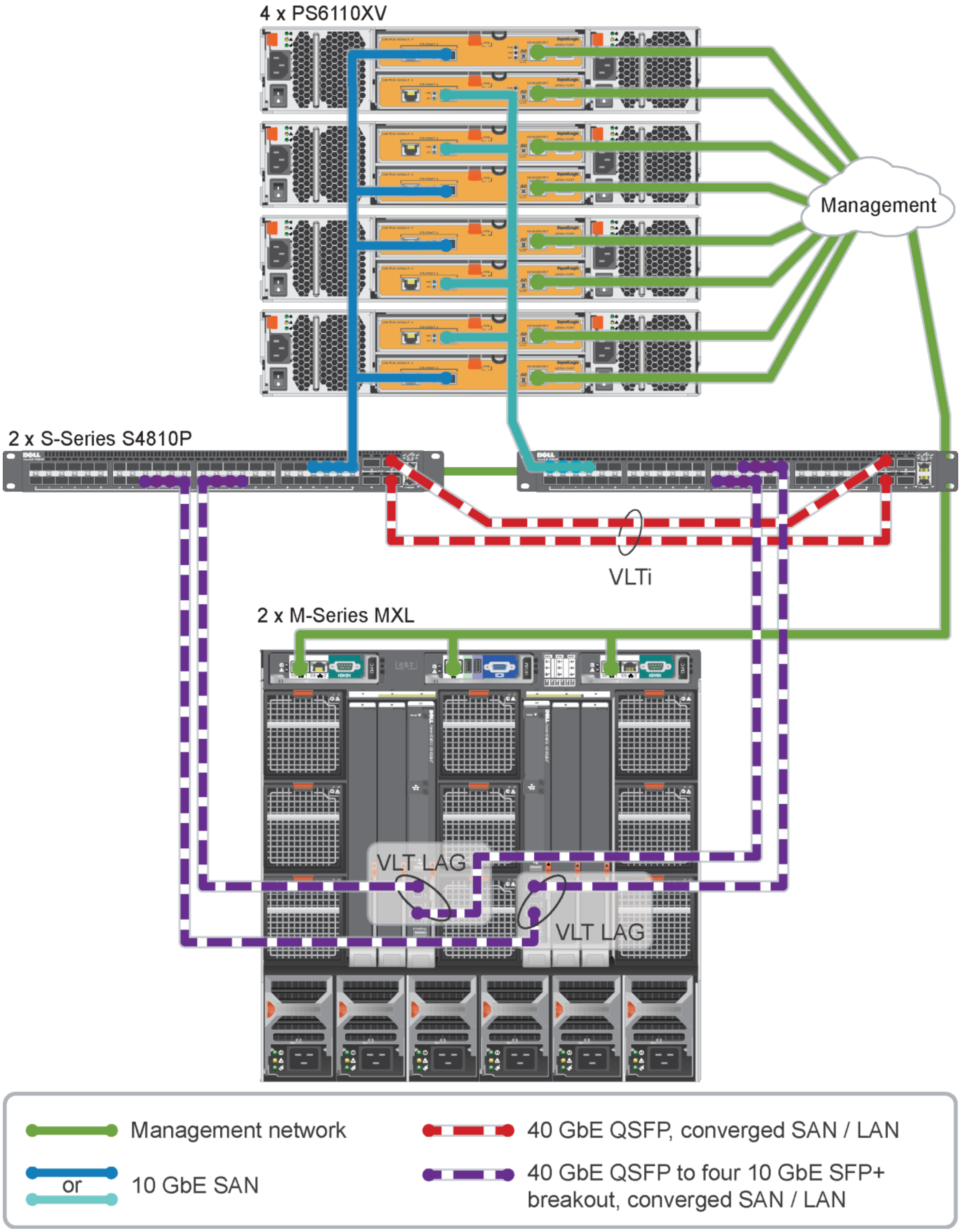

Figure 5 S-Series S4810P with VLTi / MXL with VLT LAG uplinks

4.3.2 S-Series S4810P with VLTi / M-Series I/O Aggregator with VLT LAG uplinks

This SAN design uses two of the four integrated 40 GbE QSFP ports on each S-Series S4810P to create a VLTi between each ToR switch. Eight 10 GbE SFP+ ports will need to be set aside on each S4810P switch to provide a 160 Gbps uplink from the two M-Series I/O Aggregators (IOA). This is because the IOA only supports using its integrated 40 GbE QSFP ports for interconnect stacking and because it only supports 40 GbE to four 10 GbE breakout cables for use with the other expansion module 40 GbE QSFP ports. Note that stacking the IOA is only supported when using Active System Manager for administrative setup. Because the ToR switches are operating as VLT domain peers, each IOA LAG can connect to both ToR switches in a redundant but loop-free topology. Two 40 GbE QSFP ports per S4810P switch are available for uplinks to a core switch.

This SAN design provides 160 Gbps of uplink bandwidth while providing 80 Gbps of interconnect bandwidth; more than enough to accommodate the maximum theoretical traffic of 16 10 GbE PS Series array members. Even with 16 array members, each of which requires two ports for the active and passive storage controllers combined, switch port count is not an issue. The host/storage port ratio with the maximum number of array members is 2:1.

The following diagram illustrates how four PS6110XV array members connect to the two ToR S4810P switches, how the ToR switches are interconnected with a VLTi, and how each IOA is connected to both ToR switches with a VLT LAG. Note that the port on the passive storage controller is connected to a different switch than the port on the active storage controller, ensuring that the port-based failover of the PS6110 array member will connect to a different switch upon port, cable, or switch failure. The management network is shown for reference.

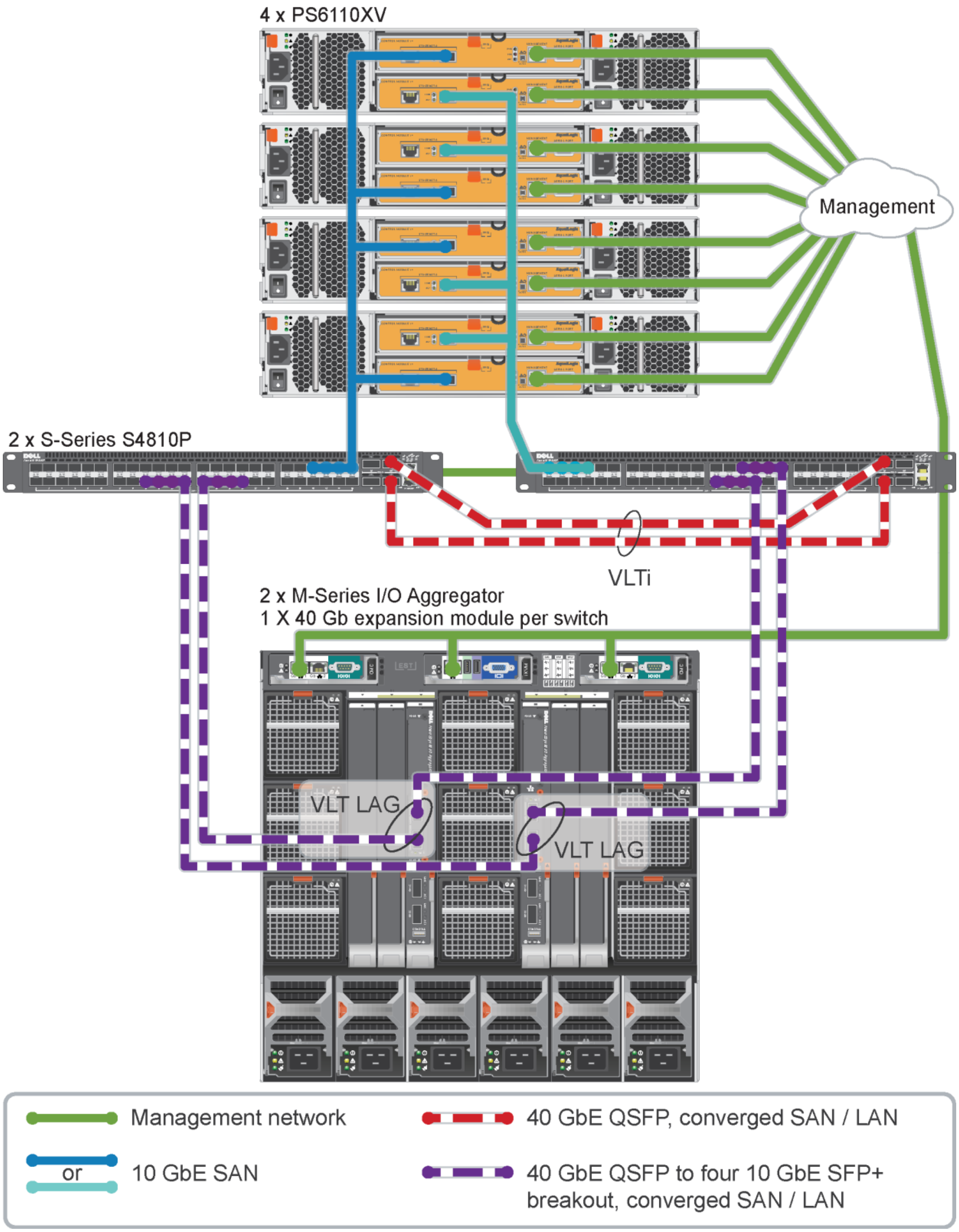
4 x PS6110XV
Management
2 x S-Series S4810P
VLTi
2 x M-Series I/O Aggregator
1 X 40 Gb expansion module per switch
VLT LAG
VLT LAG
Management network
or
10 GbE SAN
40 GbE QSFP, converged SAN / LAN
40 GbE QSFP to four 10 GbE SFP+
breakout, converged SAN / LAN

Figure 6 Active System 800 -- Force10 S4810P with VLTi / IOA with VLT LAG uplinks

4.4 Summary table of tested SAN designs

The following table assumes one fully populated M1000e blade chassis with 16 half-height blade servers each using two network ports (32 host ports total) and the maximum number of PS Series array members accommodated by the available ports of the array member switches -- either dual ToR S4810P switches or dual MXL switches in a single M1000e blade chassis I/O Fabric.

In single switch tier designs, increasing the number of total host ports per chassis decreases the number of ports available for array member port connection. Total host ports can be increased either by increasing the number of host ports per server blade or increasing the number of blade servers per chassis.

Table 4 A comparison of all tested SAN designs

	Host switch type	Array member switch type	Total uplink bandwidth	Total inter-connect bandwidth	Maximum number of hosts	Maximum number of arrays members	Port ratio with maximum hosts/array members
MXL with LAG interconnect	Blade	Blade	N/A	80 Gbps	16	8	4:1
S4810P with VLTi	ToR	ToR	N/A	80 Gbps	16	16	2:1
S4810P with VLTi / MXL with VLT LAG uplinks	Blade	ToR	160 Gbps	80 Gbps	16	16	2:1
S4810P with VLTi / IOA with VLT LAG uplinks	Blade	ToR	160 Gbps	80 Gbps	16	16	2:1

5 Detailed SAN design analysis and recommendations

The following section examines each M1000e blade chassis and PS Series SAN design from the perspectives of administration, performance, high availability, and scalability. In addition, SAN bandwidth, host to storage port ratios, and SAN performance and high availability test results are provided as a basis for SAN design recommendations.

5.1 Administration

In this section, SAN designs are evaluated by the ease of hardware acquisition and installation as well as initial setup and ongoing administration. Administrative tasks such as physical installation, switch configuration, and switch firmware updates play a role in determining the merits of a particular SAN design. The following paragraphs provide a list of common administration considerations and how each is affected by SAN design choice.

5.1.1 Uplinks and interconnects

One characteristic that all SAN designs share is the requirement for connections between switches. Even designs with a single tier of switches, like the blade IOM only designs, still have a switch interconnect. For multiple switch tier designs, the uplink between switch tiers needs sufficient bandwidth to prevent constraining the throughput of SAN traffic. High bandwidth ports or proprietary stacking ports are the best solution for an interconnect or an uplink and should be used whenever possible. The S-Series S4810P, the M-Series MXL, and the M-Series I/O Aggregator all have integrated 40 GbE QSFP ports which can be used to create high bandwidth interconnects. Keep in mind that the integrated ports on the IOA can only be used for a stacked interconnect. The high bandwidth integrated ports on the S4810P and MXL can also be used for creating uplinks. However, uplinks from the IOA can only be created using 40 GbE QSFP expansion modules and 40 GbE to 10 GbE breakout cables or using 10 GbE SFP+ expansion modules.

From an administrative perspective, a switch stack may be preferred because it allows the administration of multiple switches as if they were one physical unit. On the Dell EMC Networking switches, the initial stack is defined by configuring the correct cabling and completing a few simple steps. Then, all other tasks such as enabling flow control or updating firmware must be done only once for the entire stack.

One important thing to note is that the reloading of a switch stack will bring down all switch units in the stack simultaneously. This means that the SAN becomes unavailable during a stack reload if the switch interconnect is stacked. The resulting SAN downtime must be scheduled.

Another important note is that the S4810P is not stack compatible with either the MXL or the IOA, so for these devices creating stacked uplinks is not possible.

A second high bandwidth option for switch uplinks and interconnects is a link aggregation group (LAG). Multiple switch ports are configured to act as a single connection to increase throughput and provide redundancy, but each individual switch must still be administered separately. Creating a LAG between two Dell EMC Networking switches using LACP is a very straightforward process and administrative complexity is not a concern.

Lastly, there is a Dell EMC Networking feature called virtual link trunking (VLT) which enables switch interconnects and uplinks of a special type. For example, a VLT interconnect (VLTi) between two ToR S4810P switches enables an uplink LAG from a blade IOM switch to link to both ToR switches (referred to as VLT peer switches) in a loop-free topology. Spanning Tree protocol is still needed to prevent the initial loop that may occur prior to VLT being established. After VLT is established, RSTP may be used to prevent loops from forming with new links that are incorrectly connected and outside the VLT domain. When operating as VLT domain peers, the ToR switches appear as a single virtual switch from the point of view of any connected switch or server supporting LACP. This has many benefits including high availability, the use of all available uplink bandwidth, and fast convergence if either the link or the device fails. For more information on VLT, see the FTOS Configuration Guide for the S4810 System. The M-Series MXL and M-Series I/O Aggregator support Virtual Link Trunking (VLT): mVLT and L2/L3 over VLT.

5.1.2 M-Series MXL vs. M-Series I/O Aggregator

Though physically identical and very similar in functionality, the MXL and the IOA blade IOM switches run different firmware and have different features and behavior. The MXL switch requires complete configuration prior to deployment while the IOA has a set of default behavior meant to simplify deployment and to be easily integrated into an Active System converged infrastructure. IOA default behavior includes the following:

- All ports are active and configured to transmit tagged and untagged VLAN traffic
- External and internal port roles are set to receive DCB configuration from upstream switches and to propagate to connected blade servers
- All 40 GbE QSFP ports operate in 4x10GbE mode
- Stacking is disabled
- All uplink ports are configured in a single LAG (LAG 128)
- iSCSI optimization is enabled
- Jumbo frames are enabled

5.1.3 DCB configuration

It is a best practice to configure desired DCB settings at the core, aggregation, or ToR switches and allow DCBX to propagate the DCB configuration to edge devices such as initiators (NICs with software initiators/CNAs), PS Series array members, and downstream switches such as the blade IOM switches. Instructions for enabling and configuring DCB on an S-Series S4810P can be found in the following Dell TechCenter document:

http://en.community.dell.com/techcenter/storage/w/wiki/4250.switch-configuration-guides-for-ps-series-or-sc-series-sans

5.1.4 Active System Manager

Much of the complexity of deploying a blade chassis and PS Series SAN is alleviated by Active System Manager. ASM has the following features:

- Template-based provisioning and automated configuration to easily encapsulate infrastructure requirements and then predictably apply those requirements based on workload needs
- Management of the entire lifecycle of infrastructure, from discovery and on-boarding through provisioning, on-going management and decommissioning
- Workload failover, enabling rapid and easy migration of workload to desired infrastructure resources
- Wizard-driven interface, with feature-guided, step-by-step workflows
- Graphical logical network topology view and extended views of NIC partitions

Active System integration services provide a complete end-to-end deployment including switch configuration, DCB enablement, PS Series array member initialization, blade enclosure/server setup, and hypervisor installation.

5.1.5 Hardware requirements

The SAN design will determine the type and quantity of hardware and cabling required. Implementing a multiple tier switch SAN design will obviously require at least twice the number of switches as other more simple designs.

A blade IOM switch only SAN design requires the fewest cables, with only the array member ports and a single interconnect stack or LAG at the M1000e chassis to cable. The blade IOM switch with ToR switch SAN designs require the addition of two uplink stacks, LAGs or VLT LAGs. The ToR switch only designs (with pass-through IOM) need a stack, LAG or VLTi interconnect as well as a cable for each of the host ports; up to 32 cables for an M1000e chassis with 16 half-height blade servers with two host ports per server.

5.1.6 Using alternate switch vendors

While the choice of switches for use within an M1000e blade chassis is limited to the blade IOM product offering, ToR switches can be of any type or vendor. So for example if a SAN consisting of PS Series array members and an M1000e blade chassis were being deployed in a datacenter with an existing layer of non-Dell switches, there are blade IOM switch with ToR switch designs and ToR switch only designs which could accommodate such a scenario. For more information on PS Series SAN components see the *Dell Storage Compatibility Matrix* at http://en.community.dell.com/dell-groups/dtcmedia/m/mediagallery/20438558

5.1.7 Recommendations

In summary, when reducing administrative overhead is the goal, a single switch tier design with a stacked interconnect is the simplest option. Because the storage is directly attached to the blade IOM switches, fewer cables are required than with the ToR switch only design, and the stacked interconnect allows the switches to be administered as a single switch.

If the availability of the SAN is critical, a LAG or VLTi interconnect will be preferred over stacking. If a switch interconnect is stacked, then a switch stack reload (required for tasks such as switch firmware updates) will temporarily make the SAN unavailable. In this case, SAN downtime for firmware updates would have to be scheduled.

DCB configuration should be configured at a single source switch at the core, aggregation or ToR switch tier and allowed to flow down via DCBX to blade IOM switches, CNAs, and PS Series array members.

If ToR switches from a different vendor are used, the simplest choice is to implement the ToR only design by cabling M1000e pass-through IOM directly to the ToR switches. If multiple switch tiers are desired, plan for an uplink LAG using the high bandwidth ports of the blade IOM switches.

5.2 Performance

The second criterion by which SAN designs will be evaluated is their performance relative to each other. This section reports the performance results of each SAN design under three common I/O workloads.

Note: The results provided in this paper are intended for the purpose of comparing specific configurations used in our lab environment. The results do not portray the maximum capabilities of any system, software, and/or storage.

5.2.1 Test environment

In order to determine the relative performance of each SAN design we used the performance tool vdbench to capture throughput values at three distinct I/O workloads. Vdbench is "a disk and tape I/O workload generator for verifying data integrity and measuring performance of direct attached and network connected storage." (http://sourceforge.net/projects/vdbench/)

Since this project consists of network designs which converge SAN and LAN traffic using DCB, each vdbench test run was accompanied by a base level of LAN traffic generated by iperf, a tool used to measure maximum TCP and UDP bandwidth performance. (http://sourceforge.net/projects/iperf/)

Each performance test was conducted with the hardware and software listed below.

Note: All PS Series SAN best practices, such as enabling flow control and Jumbo frames, were implemented.
See Appendix A for more detail about the hardware and software infrastructure.

5.2.1.1 Hosts:

- Four PowerEdge M620 blade servers each with:
  - Windows Server 2008 R2 SP1
  - PS Series EqualLogic Host Integration Toolkit v4.0.0
  - Two Broadcom BCM57810 10Gb ports with iSCSI Offload Engine enabled and separate IP address configured for SAN and LAN

5.2.1.2 Storage:

- Four PS Series PS6110XV array members each with:
  - Firmware: 6.0.2
  - One active 10 GbE ports on the SAN
- Four iSCSI volumes dedicated to each host

Note: There were a total of eight host ports and four storage ports for a 2:1 ratio.

5.2.1.3 I/O

The following three vdbench workloads were defined:

- 8KB transfer size, random I/O, 67% read
- 256KB transfer size, sequential I/O, 100% read

- 256KB transfer size, sequential I/O, 100% write

Each vdbench workload was run for thirty minutes and the I/O rate was not capped (the vdbench "iorate" parameter was set to "max"). The throughput values used in the relative performance graphs are the sums of the values reported by each of the four hosts.

Each host ran one instance of iperf server and one instance of iperf client. Hosts ran iperf traffic in pairs such that one of the network ports on each host acted as an iperf server and the other network port acted as an iperf client, thus ensuring that LAN traffic was evenly distributed across the host network ports.

Table 5 An example distribution of iperf LAN traffic

Host 1 NIC1 (client) → Host 2 NIC1 (server)
Host 1 NIC2 (server) ← Host 2 NIC2 (client)

5.2.1.4 DCB

The following table lists the DCB configuration in place during performance testing.

Table 6 DCB configuration

Traffic Class	DCB Traffic Class	802.1p QoS	ETS Bandwidth Settings	PFC setting
iSCSI	1	4	60	Lossless
OTHER	2	0,1,2,3,5,6,7	40	Non-lossless

5.2.2 Bandwidth

The following table shows the uplink and interconnect bandwidth of each tested SAN design. Each of the single switch tier designs provide adequate interconnect bandwidth for the maximum number of array members that their port counts accommodate.

Table 7 A comparison of the bandwidth provided by all SAN designs

	Total uplink bandwidth	Total interconnect bandwidth
MXL with LAG interconnect	N/A	80 Gbps
S4810P with VLTi	N/A	80 Gbps
S4810P with VLTi / MXL with VLT LAG uplinks	160 Gbps	80 Gbps
S4810P with VLTi / IOA with VLT LAG uplinks	160 Gbps	80 Gbps

5.2.3 Results

The following three figures show the relative aggregate vdbench throughput of all four hosts within each SAN design at three different I/O workloads. Each throughput value is presented as a percentage of a baseline value. In each chart, the MXL with LAG interconnect design was chosen as the baseline value. All of the throughput values were achieved during a single thirty minute test run.

8 KB random I/O, 67% read workload

The following figure shows the aggregate vdbench throughput of all four hosts within each SAN design at an 8 KB random I/O, 67% read workload. All SAN designs yielded throughput results within 2% of the baseline value.

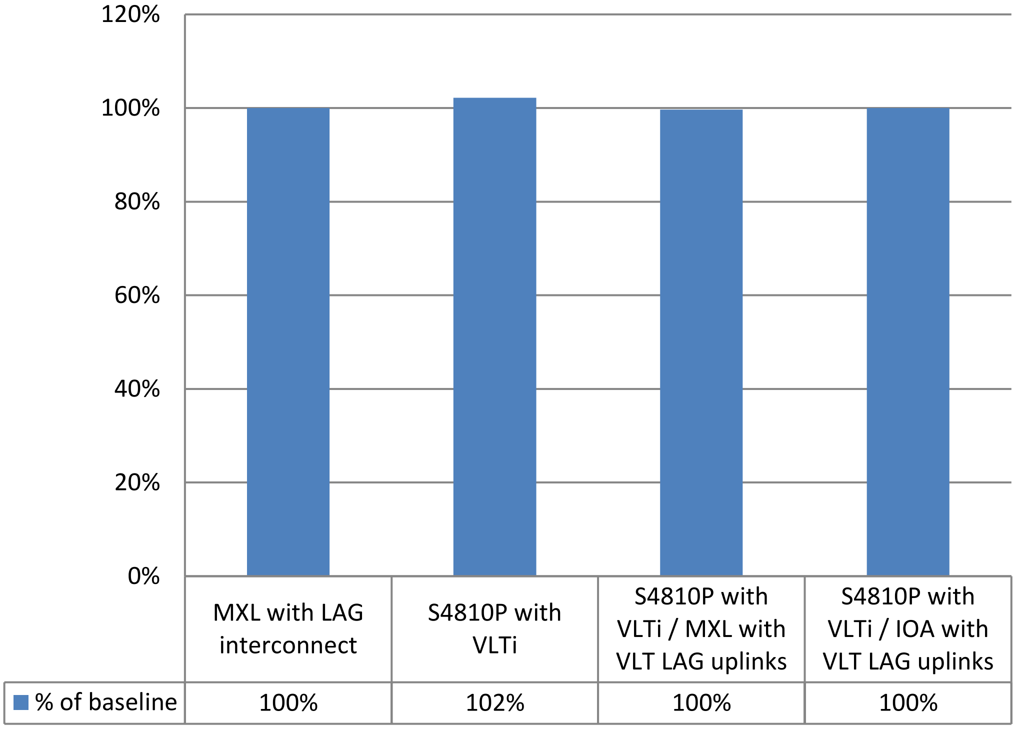

Figure 7 Aggregate vdbench throughput as a percentage of the baseline value in each SAN design during an 8 KB random I/O, 67% read workload

256 KB sequential I/O, read workload

The following figure shows the aggregate vdbench throughput of all four hosts within each SAN design at a 256 KB sequential I/O, read workload. All SAN designs yielded throughput results within 5% of the baseline value.

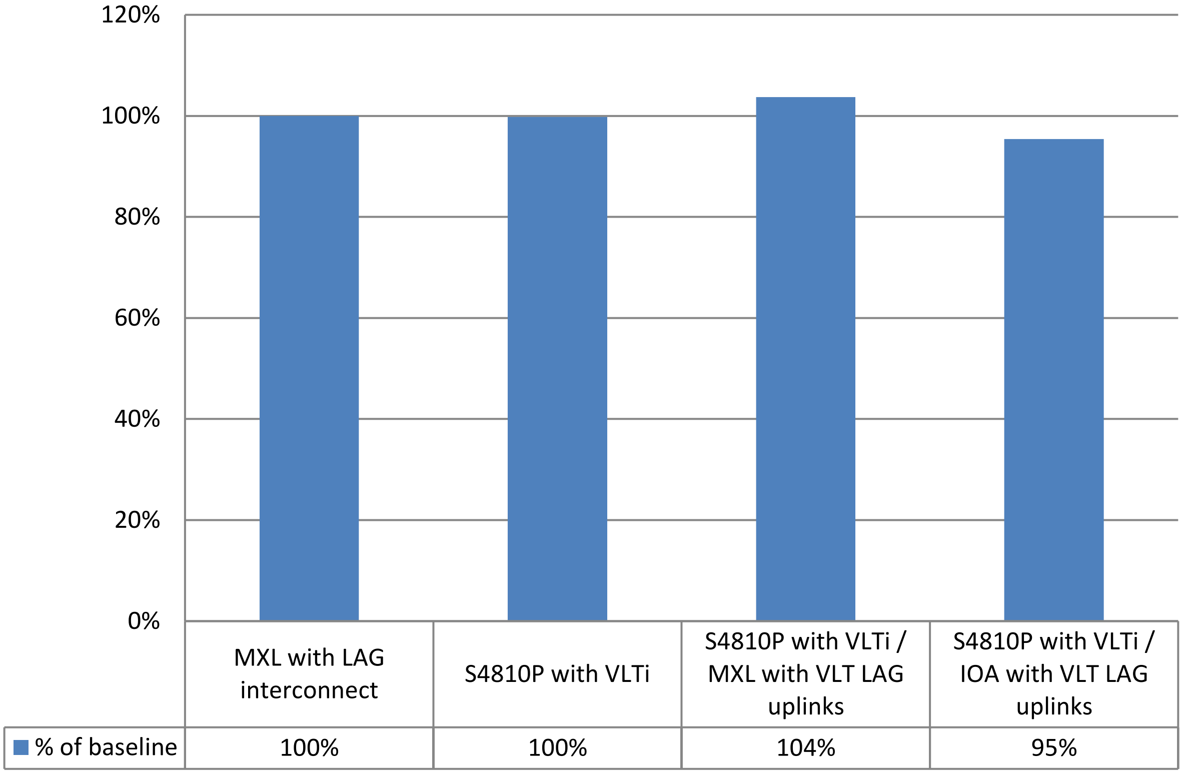

Figure 8 Aggregate Vdbench throughput as a percentage of the baseline value in each SAN design during a 256 KB sequential I/O, read workload

256 KB sequential I/O, write workload

The following figure shows the aggregate vdbench throughput of all four hosts within each SAN design at a 256 KB sequential I/O, write workload. All SAN designs yielded throughput results within 8% of the baseline value.

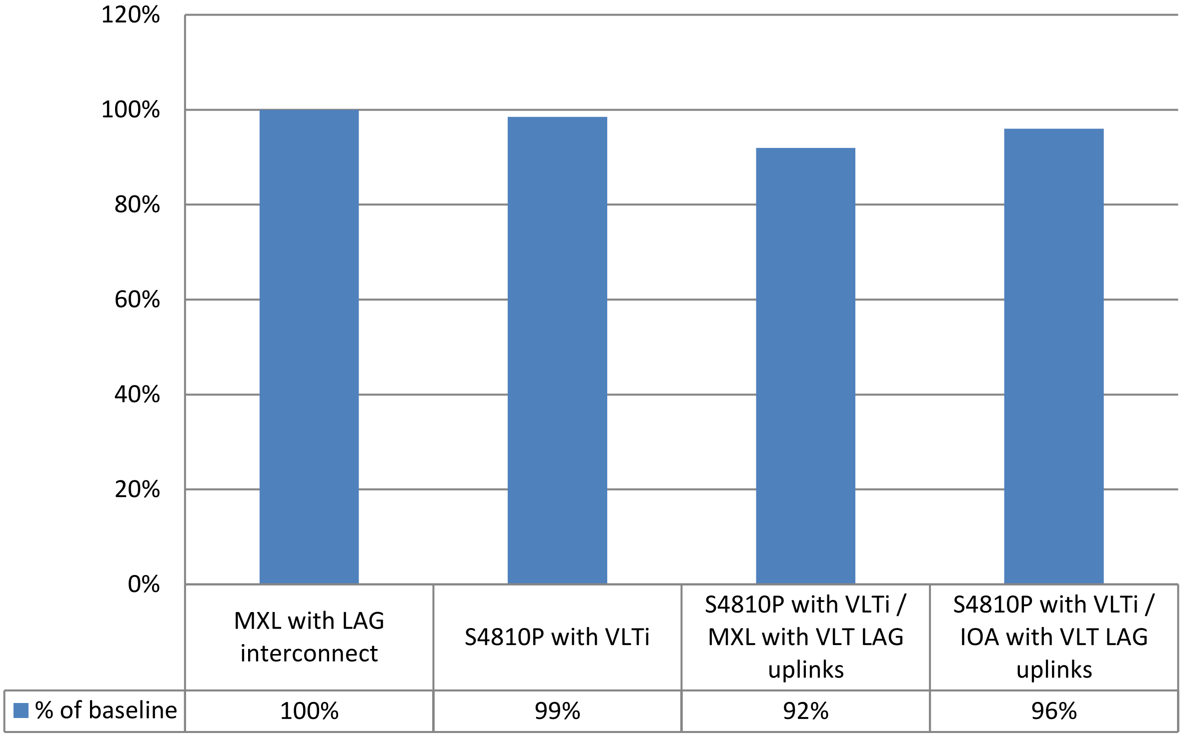

Figure 9 Aggregate Vdbench throughput as a percentage of the baseline value in each SAN design during a 256 KB sequential I/O, write workload

5.2.4 Recommendations

The throughput values were gathered during the performance testing of each SAN design with four hosts and four arrays members at three common workloads. Among all SAN designs, there were no significant performance differences during any of the three tested workloads.

5.3 High availability

The third criterion by which SAN designs will be evaluated is how each design tolerates a switch failure. This section quantifies how the loss of different switches within the SAN fabric affects the available bandwidth and the total number of connected host ports. The results below assume a single M1000e chassis and 16 half-height blade servers with two SAN ports each for a total of 32 host ports.

Note: Storage port disconnection is not addressed in the tables because the PS6110XV controller port failover ensures that no single switch failure will cause the disconnection of any array member ports.

Previous generations of PS Series arrays did not have individual port failover and a single port, cable or switch failure could reduce the number of connected array member ports.

To test SAN design high availability, an ungraceful switch power down was executed while the SAN was under load. The test environment was the same as the environment that was used during performance testing, and the workload was 256 KB sequential I/O write using vdbench. LAN traffic was generated with iperf.

In cases where host ports were disconnected, iSCSI connections were appropriately migrated to the remaining host ports. In these cases, even with the loss of 50% of the host ports there was at least a 1:1 host/storage port ratio.

5.3.1 ToR switch failure

The following table shows how each SAN design is affected by the loss of a ToR switch. Note that this failure is not applicable to the blade IOM switch only designs in which both host and storage ports are connected to blade IOM switches.

In the ToR switch only SAN design, a TOR switch failure reduces the number of connected host ports by 50% since the host port connect directly to the ToR switches using a pass-through module. In multiple-tier SAN designs, all host port connections can be maintained during a ToR switch failure when the ToR switch interconnect is a VLTi. This allows each blade switch uplink LAG to be distributed across both ToR switches, maintaining each blade switch’s connectivity in the event of a ToR switch failure. However, the loss of 50% of uplink bandwidth in the event of a ToR switch failure cannot be avoided in multiple switch tier SAN designs. Interconnect bandwidth becomes irrelevant after a ToR switch failure in all applicable SAN designs as all array member ports migrate to the remaining ToR switch to which all remaining host ports have either a direct or an uplinked connection.

Table 8 A comparison of the way each SAN designs tolerates a ToR switch failure

	Reduction in connected host ports	**Reduction in uplink bandwidth**	**Reduction in inter-connect bandwidth**
MXL with LAG interconnect	N/A	N/A	N/A
S4810P with VLTi	32-->16	N/A	80Gbps-->N/A**
S4810P with VLTi / MXL with VLT LAG uplinks	32-->32	160Gbps-->80Gbps	80Gbps-->N/A**
S4810P with VLTi / IOA with VLT LAG uplinks	32-->32	160Gbps-->80Gbps	80Gbps-->N/A**

**Interconnect bandwidth is no longer relevant because the ToR switch failure eliminates the interconnect.

5.3.2 Blade IOM switch failure

The following table shows how each SAN design is affected by the loss of a blade IOM switch. Note that this failure is not applicable to ToR switch only designs in which both host and storage ports are connected to the ToR switches.

In all applicable SAN designs, a blade IOM switch failure reduces the number of connected host ports by 50% and in the multiple switch tier SAN design the uplink bandwidth is also reduced by 50%. Interconnect bandwidth becomes irrelevant after a blade IOM switch failure in a single switch tier SAN design. This is because all array member ports migrate to the remaining blade IOM switch to which the remaining host ports a direct connection. However, the multiple switch tier SAN designs retain all interconnect bandwidth as the interconnect is between the ToR switches and not subject to a blade IOM switch failure.

Table 9 A comparison of the way each SAN designs tolerates a blade IOM switch failure

	Reduction in connected host ports	Reduction in uplink bandwidth	Reduction in inter-connect bandwidth
MXL with LAG interconnect	32-->16	N/A	80Gbps-->N/A**
S4810P with VLTi	N/A	N/A	N/A
S4810P with VLTi / MXL with VLT LAG uplinks	32-->16	160Gbps-->80Gbps	80Gbps-->80Gbps
S4810P with VLTi / IOA with VLT LAG uplinks	32-->16	160Gbps-->80Gbps	80Gbps-->80Gbps

**Interconnect bandwidth is no longer relevant because the switch failure eliminates the interconnect

5.3.3 Recommendations

ToR switch failures always collapse the fabric to a single switch as arrays member network ports failover to the remaining ToR switch. Host connectivity can be preserved during a ToR switch failure with redundant VLT LAG uplinks from the blade IOM switches made possible by having a VLTi interconnect between the ToR switches, rather than a standard LAG. Stacked interconnects should be avoided because the SAN becomes unavailable during a switch stack reload.

Blade IOM switch failures always result in a loss of 50% of the host ports and in multiple-tier SAN designs a 50% loss in uplink bandwidth.

5.4 Scalability

The final criterion by which SAN designs will be evaluated is scalability. Note that the scalability data presented in this section is based primarily on available port count. Actual workload, host to array port ratios, and other factors may affect performance.

The following tables show the number of host ports and array members, the host/storage port ratio, the number of blade IOM switch expansion modules required and most importantly the total number of SAN ports available for additional edge device connectivity or for a high bandwidth uplink from the converged SAN/LAN to a core switch. The number of available SAN ports includes both ToR switches (or blade IOM switches if no ToR switches are present) and assumes the following:

- Two M-Series MXL, two M-Series I/O Aggregator or two M-series pass-through I/O modules per blade chassis, and if applicable, two S-Series S4810P ToR switches
- 16 blade servers with two network ports each
- 16 array members with one active and one passive network port
- 80 Gbps of interconnect bandwidth
- 160 Gbps of uplink bandwidth between host and storage tier if applicable

Table 10 Scalability information for all SAN designs

	Host ports with 16 blade servers	**Maximum number of arrays members**	**Host / storage port ratio**	**Blade switch expansion modules required**	**Total available SAN port count**
MXL with LAG inter-connect	32	16	2:1	2x 40 GbE per switch	4 x 40 GbE
S4810P with VLTi	32	16	2:1	None	32 x 10 GbE ***and*** 4 x 40 GbE
S4810P with VLTi / MXL with VLT LAG uplinks	32	16	2:1	None	64 x 10 GbE ***or*** 48 x 10 GbE ***and*** 4 x 40 GbE**
S4810P with VLTi / IOA with VLT LAG uplinks	32	16	2:1	1x 40 GbE per switch	48 x 10 GbE ***and*** 4 x 40 GbE**

** Requires the use of 40 GbE QSFP to 10 GbE SFP+ breakout cables from the blade IOM switch to the ToR switch

A pair of M-Series MXL switches can accommodate up to 16 PS Series array members with no ToR switch using one 40 GbE QSFP expansion modules per switch. However, when using 16 array members there will be no remaining 40 GbE QSFP ports for creating an uplink to a core switch, creating an isolated SAN.

Two ToR S-Series S4810P switches and two 10GbE pass-through modules in the blade chassis can accommodate up to 16 array members with no expansion modules required. Furthermore, this SAN design has an additional thirty-two 10 GbE ports and four 40 GbE ports available for edge device connectivity and high bandwidth uplinks with no extra hardware needed.

Two ToR S-Series S4810P and two M-Series MXL switches can also accommodate up to 16 array members with no expansion modules required. This SAN design has the highest number of available ports -- an additional sixty-four 10 GbE ports or, if 40 GbE QSFP to 10 GbE SFP+ breakout cables are used to uplink the MXL to the S4810P, forty-eight 10 GbE ports and four 40 GbE ports.

Finally, two ToR S-Series S4810P and two M-Series I/O Aggregator switches can accommodate up to 16 array members, however one 40 GbE QSFP expansion modules per switch and 40 GbE QSFP to 10 GbE SFP+ breakout cables between the IOA and S4810P are required. With this SAN design, forty-eight 10 GbE ports and four 40 GbE ports are available.

5.4.1 Recommendations

In summary, all SAN designs can support up to 16 array members while providing not only adequate bandwidth within the SAN. While the Blade IOM only SAN design has no ports remaining for additional connectivity or uplinks when using 16 array members, the other three SAN designs have ample ports remaining for additional edge device connectivity and high bandwidth uplinks to a core switch. Considering the fact that the S-Series S4810 supports VLT and doesn’t require expansion modules, the simple ToR switch only SAN design is an excellent option. It creates a robust aggregation layer that accepts highly available VLT LAG from downstream edge devices and switches and also from the upstream Layer 3 core.

A Solution infrastructure detail

The following table is a detailed inventory of the hardware and software configuration in the test environment.

Table 11 A detailed inventory of the hardware and software configuration in the test environment

Solution configuration - Hardware components:		**Description**
Blade Enclosure	PowerEdge M1000e chassis: CMC firmware: 4.20	Storage host enclosure
10 GbE Blade Servers	(4) PowerEdge M620 server: Windows Server 2008 R2 SP1 BIOS version: 1.4.9 iDRAC firmware: 1.23.23 (2) Intel® Xeon® E5-2650 128GB RAM Dual Broadcom 57810S-k 10 GbE CNA Driver v7.4.23.0 Firmware v7.4.8 PS Series EqualLogic Host Integration Toolkit v4.0.0	Storage hosts for configs: M-Series MXL with LAG interconnect S-Series S4810P with VLTi S-Series S4810P with VLTi / MXL with VLT LAG uplinks S-Series S4810P with VLTi / M-Series I/O Aggregator with VLT LAG uplinks
10 GbE Blade I/O modules	(2) M-Series MXL FTOS v8.3.16.2 (2) 40 GbE QSFP expansion module (2) M-Series I/O Aggregator FTOS v8.3.17.0 (1) 40 GbE QSFP expansion module (2) Dell 10 Gb Ethernet Pass-through module	IO modules for configs: M-Series MXL with LAG interconnect S-Series S4810P with VLTi S-Series S4810P with VLTi / MXL with VLT LAG uplinks S-Series S4810P with VLTi / M-Series I/O Aggregator with VLT LAG uplinks
10 GbE TOR switches	(2) S-Series S4810P FTOS v8.3.12.0	Switches for configs: S-Series S4810P with VLTi S-Series S4810P with VLTi / MXL with VLT LAG uplinks S-Series S4810P with VLTi / M-Series I/O Aggregator with VLT LAG uplinks
10 GbE Storage	(4) PS Series PS6110XV: (24) 146GB 15K SAS disks – vHN63 (2) 10 GbE controllers Firmware: v6.0.2	Storage arrays for configs: M-Series MXL with LAG interconnect S-Series S4810P with VLTi S-Series S4810P with VLTi / MXL with VLT LAG uplinks S-Series S4810P with VLTi / M-Series I/O Aggregator with VLT LAG uplinks

B Technical support and resources

Dell.com/support is focused on meeting customer needs with proven services and support.

Dell TechCenter is an online technical community where IT professionals have access to numerous resources for Dell EMC software, hardware and services.

Dell.com/StorageResources on Dell TechCenter provide expertise that helps to ensure customer success on Dell EMC Storage platforms.

B.1 Related resources

See the following referenced or recommended Dell publications:

- *Dell Storage Compatibility Matrix*:

  http://en.community.dell.com/dell-groups/dtcmedia/m/mediagallery/20438558

- Switch Configuration Guides for PS Series or SC Series SANs:

  http://en.community.dell.com/techcenter/storage/w/wiki/4250.switch-configuration-guides-for-ps-series-or-sc-series-sans

- Dell Active System Manager Wiki:

  http://www.dell.com/asmtechcenter

- *Dell PS Series DCB Configuration Best Practice*:

  http://en.community.dell.com/dell-groups/dtcmedia/m/mediagallery/20305369

- *Best Practices for Configuring DCB with Windows Server and EqualLogic Arrays*:

  http://en.community.dell.com/techcenter/extras/m/white_papers/20438162