Dell EqualLogic Best Practices Series

SAN Design Best Practices for the Dell PowerEdge M1000e Blade Enclosure and EqualLogic PS Series Storage (1GbE)

A Dell Technical Whitepaper

Storage Infrastructure and Solutions Engineering
Dell Product Group
August 2012

Table of Contents

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 **Margaret Boeneke**

Feedback

We encourage readers of this publication to provide feedback on the quality and usefulness of this information by sending an email to SISfeedback@Dell.com.

SISfeedback@Dell.com

1 Introduction

With the Dell™ EqualLogic™ PS Series storage arrays, Dell provides 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 operations.

To leverage the advanced features provided by an EqualLogic array, a robust, standards-compliant iSCSI storage area network (SAN) infrastructure must be created. When using blade servers in a Dell PowerEdge™ M1000e blade enclosure (also known as a blade chassis) as hosts, there are a number of network design options for storage administrators to consider when building the iSCSI SAN. For example, the PS Series array member network ports can be connected to the switches within the M1000e blade chassis or the blade server network ports can be connected to top of rack (TOR) switches residing outside of the 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.

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 EqualLogic 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.

TOR switch – A "top of rack" (TOR) switch, external to the M1000e blade chassis.

Blade IOM switch – A blade I/O module (IOM) switch, residing in an M1000e fabric slot.

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.

LAG – A link aggregation group (LAG) in which 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 functions as such.

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

ISL – An inter-switch link that connects either the two blade IOM switches or the two TOR switches to each other. An ISL can be a stack or a LAG.

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 ISL can be a stack or a LAG, and no uplink is required.

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, 1GbE pass-through IOM are used in place of blade IOM switches in the M1000e blade chassis. The ISL can be a stack or a LAG.

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. An ISL stack or LAG between each blade IOM switch and/or between each TOR switch is required. An uplink stack or LAG from the blade IOM switch tier to the TOR switch tier is also required.

Switch tier – A pair or more of like switches connected by an ISL 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 if the switch tier is stacked and the firmware is updated, then the SAN is temporarily offline.

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.

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 or LAG is required. Blade IOM switch with TOR switch designs are multiple switch tier SAN designs.

Host/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 to 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.

1.2.1 Terminology illustration

Figure 1 illustrates the basic SAN components involved when deploying an M1000e blade chassis with blade servers into an EqualLogic 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 connected to TOR switches if the blade IOM switches are replaced with pass through IOM.

If only TOR switches or only blade IOM switches are used, this paper will refer to the SAN as a single switch tier design. When both switch types are used then the SAN will be referred to as a multiple switch tier design. In multiple switch tier SAN designs the multiple switch tiers will need to be connected by an uplink, which can be either a stack or a LAG. For both one and multiple switch tier

SAN designs a like switch pair will need to be interconnected by an inter-switch link or ISL. Like the uplink, the ISL can be a stack or a LAG. The ISL is necessary to create a single layer 2 SAN fabric over which all PS Series array member network ports can communicate with each other.

For a much more detailed description of each SAN design that was tested and evaluated see Section 4 – Tested SAN designs.

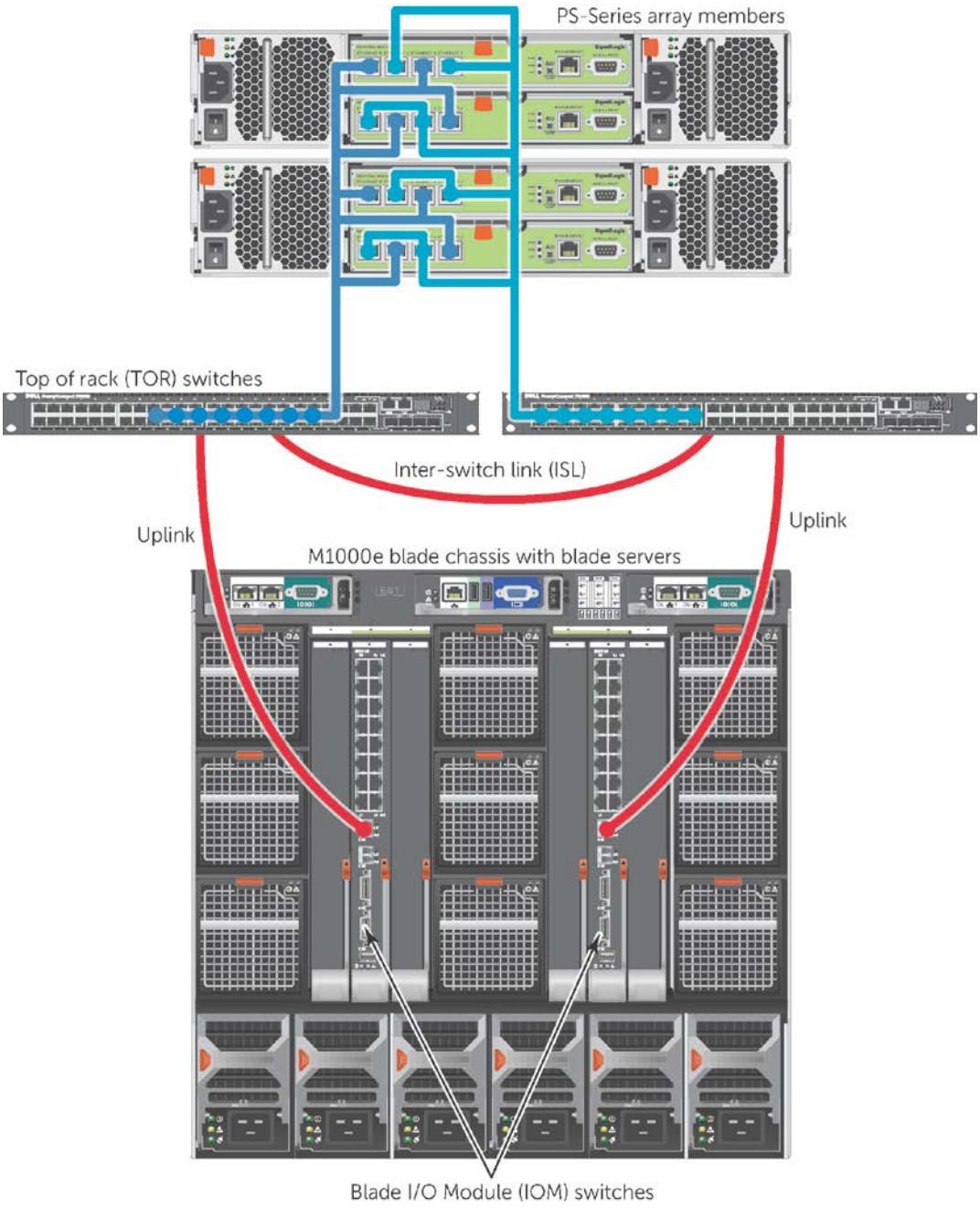

Figure 1 An example EqualLogic SAN consisting of PS Series array members, an M1000e blade chassis with blade servers, and TOR and Blade IOM switches.

2 Overview of M1000e blade chassis solution

The following section describes the M1000e blade chassis networking fabrics consisting of IO modules, a midplane, and the individual blade server network adapters.

2.1 Multiple fabrics

Each M1000e can support up to three separate networking "fabrics" that interconnect ports on each blade server to a pair of blade IO modules within each chassis fabric through a passive chassis midplane. Each fabric is associated with specific interfaces on a given blade server as described in Table 2. Each blade server has a LAN on Motherboard (LOM) capability that is mapped to the blade IOM located in the Fabric A slots in the M1000e chassis. In addition, each blade server has two mezzanine sockets for adding additional networking options such as 1Gb or 10Gb Ethernet, Infiniband, or Fibre Channel cards. These mezzanine cards are mapped to either the Fabric B or the Fabric C blade IOM.

Figure 2 illustrates the layout of the three fabric blade IOM located in the back of the M1000e chassis.

Table 1 M1000e Fabric Mapping

	LOM	Mezzanine B	Mezzanine C
Fabric	A	B	C

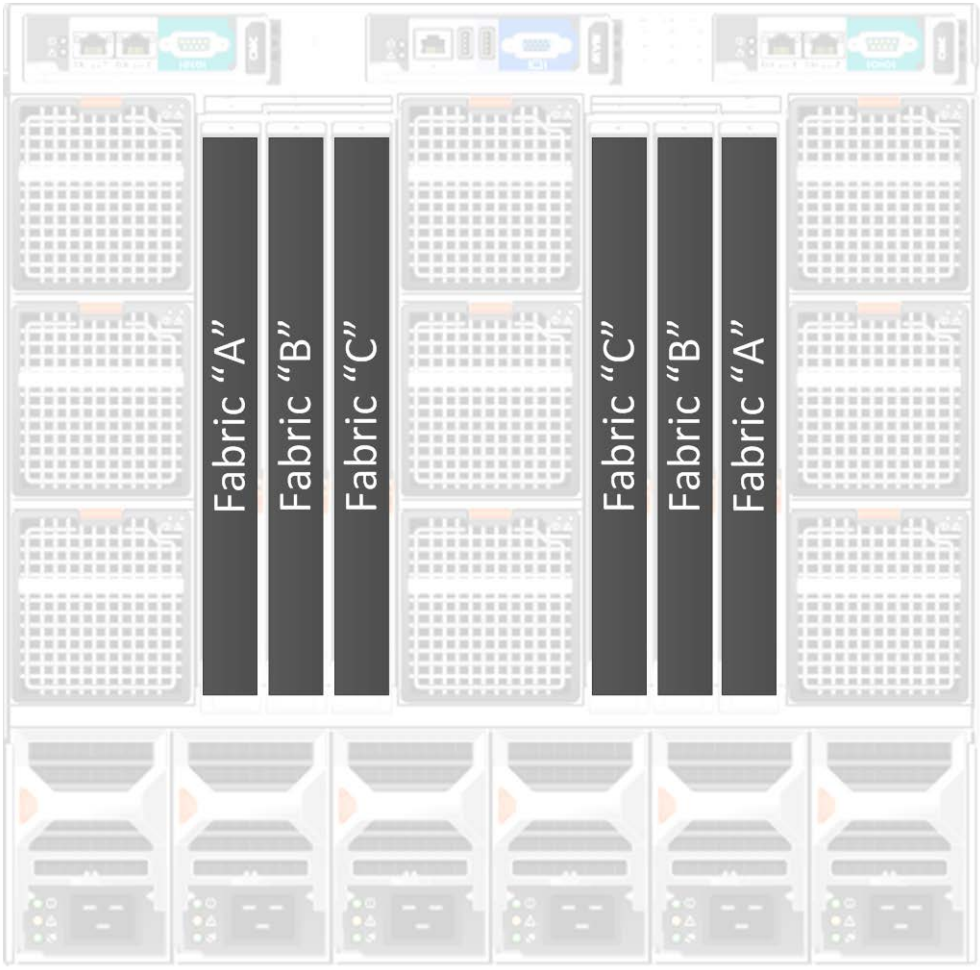

Figure 2 Blade IO Modules and M1000e Chassis

2.2 Blade IO modules

The following table lists the 1GbE blade IO module options (available at the time of this publication) and the number of ports available for EqualLogic SAN solutions.

Table 2 1GbE Blade IO Module options for EqualLogic

	1GbE external ports	10GbE uplink ports	Stacking ports
PowerConnect M6220	4	Up to 4	Up to 2
PowerConnect M6348	16	2	2
Cisco Catalyst Blade Switch 3032	8	N/A	N/A
Cisco Catalyst Blade Switch 3130G	8	N/A	2
Cisco Catalyst Blade Switch 3130X	Up to 8	Up to 2	2
1GbE Pass-through	16	N/A	N/A

3 Summary of SAN designs and recommendations

The following section provides the high level conclusions reached after the course of comprehensive lab testing and analysis of various EqualLogic PS Series array SAN designs which incorporate M1000e blade server hosts on a 1GbE network.

For complete results and recommendations see Section 5 - Detailed SAN design analysis and recommendations. For an illustration of each SAN design see Section 4 – Tested SAN designs.

Green cells indicate the recommended SAN design within each design category based on all factors considered during testing, while orange cells indicate designs that for various reasons, such as SAN availability or uplink bandwidth, might not be preferred.

Table 3 Summary of SAN designs and recommendations

	Switch tier topology	Ease of administration	Performance	High availability	Scalability
Blade IOM with ISL stack	Single	• A single switch stack to manage • No TOR switches required • Fewest cables • During ISL stack reload SAN is unavailable	• Suitable for smaller scale SAN	• Blade IOM switch failure reduces host ports by 50%	• One chassis with two blade IOM switches accommodates four array members • Additional chassis must be connected with ISL
Blade IOM with ISL LAG	Single	• Two switches to manage • No TOR switches required • Fewest cables	• Suitable for smaller scale SAN	• Blade IOM switch failure reduces host ports by 50%	• One chassis with two blade IOM switches accommodates four array members • Additional chassis must be connected with ISL
TOR with ISL LAG	Single	• Two switches to manage • No blade IOM switches required • Most cables	• Suitable for smaller scale SAN	• TOR switch failure reduces host ports by 50%	• Two TOR switches accommodate eight array members • Additional TOR switches must be connected with ISL

Blade IOM and TOR with 4-way stack	Multiple	• A single switch stack to manage • Uplinks required between tiers • During stack reload, SAN is unavailable	• The least amount of uplink bandwidth (32Gbps)	• Blade IOM switch failure reduces host ports by 50% • Blade IOM or TOR switch failure reduces uplink bandwidth by 50%	• Two TOR switches accommodate twelve array members • Recommended switch stack size limits design to two chassis
Blade IOM and TOR with 3-way LAG	Multiple	• Four switches to manage • Uplinks required between tiers	• Mid-range uplink bandwidth (40Gbps)	• Blade IOM or TOR switch failure reduces host ports by 50% • Blade IOM or TOR switch failure reduces uplink bandwidth by 50%	• Two TOR switches accommodate twelve array members • Beyond two blade chassis and a blade IOM ISL is required or TOR switches must be added to accommodate blade IOM switch uplinks
Blade IOM and TOR with uplink stacks	Multiple	• Two switch stacks to manage • Uplinks required between tiers • Separate stacks allow for SAN availability • TOR switches must be stack-compatible with blade switches	• The most uplink bandwidth (64Gbps)	• Blade IOM switch failure reduces host ports by 50% • Blade IOM or TOR switch failure reduces uplink bandwidth by 50%	• Two TOR switches accommodate twelve array members • Recommended switch stack size accommodates up to five blade chassis

Blade IOM and TOR with ISL stacks	Multiple	• Two switch stacks to manage • Uplinks required between tiers • During stack reloads, SAN is unavailable	• Mid-range uplink bandwidth (40Gbps)	• Blade IOM switch failure reduces host ports by 50% • Blade IOM or TOR switch failure reduces uplink bandwidth by 50%	• Two TOR switches accommodate twelve array members • Beyond two blade chassis and a blade IOM ISL is required or TOR switches must be added to accommodate blade IOM switch uplinks

3.1 Ease of administration

When reducing administrative overhead is the goal, a single switch tier design with an ISL stack 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. For multiple switch tier SAN designs, the 4-way stack is the easiest to setup and maintain.

If the availability of the SAN is critical, then an ISL LAG configuration may be preferred over stacking. If a switch tier ISL is stacked, then the SAN is temporarily unavailable during a switch stack reload. A multiple switch tier design that avoids this is the uplink stack with ISL LAG design, which provides some administrative benefit while ensuring the SAN is always available.

If TOR switches from a different vendor are used, then 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 as the blade IOM switches will not be stack-compatible with the TOR switches from a different vendor.

3.2 Performance

The throughput values were gathered during the performance testing of each SAN design with four hosts and two arrays members at two common workloads. For a smaller scale SAN, there were no significant performance differences among the tested SAN designs.

For larger scale SANs, it is more likely that a two tier SAN design will be chosen because this accommodates the largest number of member arrays (12). In this case, only the uplink stacks SAN design provides adequate bandwidth for both the uplink (64Gbps) and the ISL (40Gbps)

3.3 High availability

In both the external and internal switch failure scenarios, the blade IOM switch with TOR switch and an uplink stack SAN design retained as many or more host port connections while retaining the highest amount of uplink bandwidth of any other applicable SAN design.

3.4 Scalability

For blade IOM switch only SAN designs, scaling the number of array members is only possible with the addition of M1000e blade chassis and scaling beyond three blade chassis is not recommended due to increased hop-count and latency over the ISL connection.

TOR switch only SAN designs are somewhat more scalable in the sense that they allow up to eight arrays with two TOR switches and one chassis, but just as with the blade IOM switch only designs, the SAN traffic increasingly relies on ISL connections as the number of switches grows.

For larger scale PS Series SAN deployments, the blade IOM switch with TOR switch designs can accommodate a far higher number of array members without the need to add blade chassis or TOR switches. Among these designs, the only the uplink stack design provides adequate uplink and ISL bandwidth while easily accommodating the largest number of chassis by incorporating additional blade IOM switches into the two uplink stacks.

Note that 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.

3.5 Conclusion

Of all SAN designs, the blade IOM switch with TOR and uplink stack design had the best combination of ease of administration, performance, high availability, and most of all scalability. The only caveat to the uplink stack design is that blade IOM and TOR switches must be from the same product family to be stack-compatible.

Two tier stack SAN designs are significantly superior to single switch tier designs in scalability, the former having TOR switch ports dedicated entirely to storage array members.

For smaller scale SAN deployments, single switch tier designs can be more cost effective, with fewer switches to manage.

4 Tested SAN designs

The following 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. All 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 the storage are connected to the M1000e blade IOM switches. No TOR switches are required. The ISL 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. 1GbE pass-through IOM are used in place of blade IOM switches in the M1000e blade chassis. The ISL can be a stack or a LAG.
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. An ISL stack or LAG between each blade IOM switch and/or between each TOR switch is required. An uplink stack or 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 EqualLogic PS Series array member ports are directly connected to the blade IOM switch ports within the blade chassis. In these scenarios, dual PowerConnect M6348 switches in the M1000e chassis were used. Two SAN designs of this type were tested:

1. M6348 switches connected with an ISL stack
2. M6348 switches connected with an ISL LAG

4.1.1 M6348 switch with ISL stack

This SAN design provides 32Gbps of ISL bandwidth between the two M6348 switches using the two integrated 16Gb stack ports on each switch for dual stack connections. Since there is only a single tier of switches, there is no uplink to external switches. The 16 external ports on each M6348 (32 ports total) can accommodate the connection of four 1GbE PS series array members, each of which require eight ports for the active and passive controllers combined. The host/storage port ratio with the maximum number of array members is 2:1.

The following diagram illustrates how the two PS6100XV array members directly connect to the two M6348 switches in Fabric B of the M1000e blade chassis and how the two M6348 switches are stacked using the integrated stacking ports. Each array member controller connects to both M6348 switches for SAN redundancy. Note that the corresponding port on the passive controller is connected to a different switch than the port on the active controller, ensuring that the port-based failover of the PS6100 array member will connect to a different switch upon port, cable or switch failure. Management and host LAN networks are shown for reference.

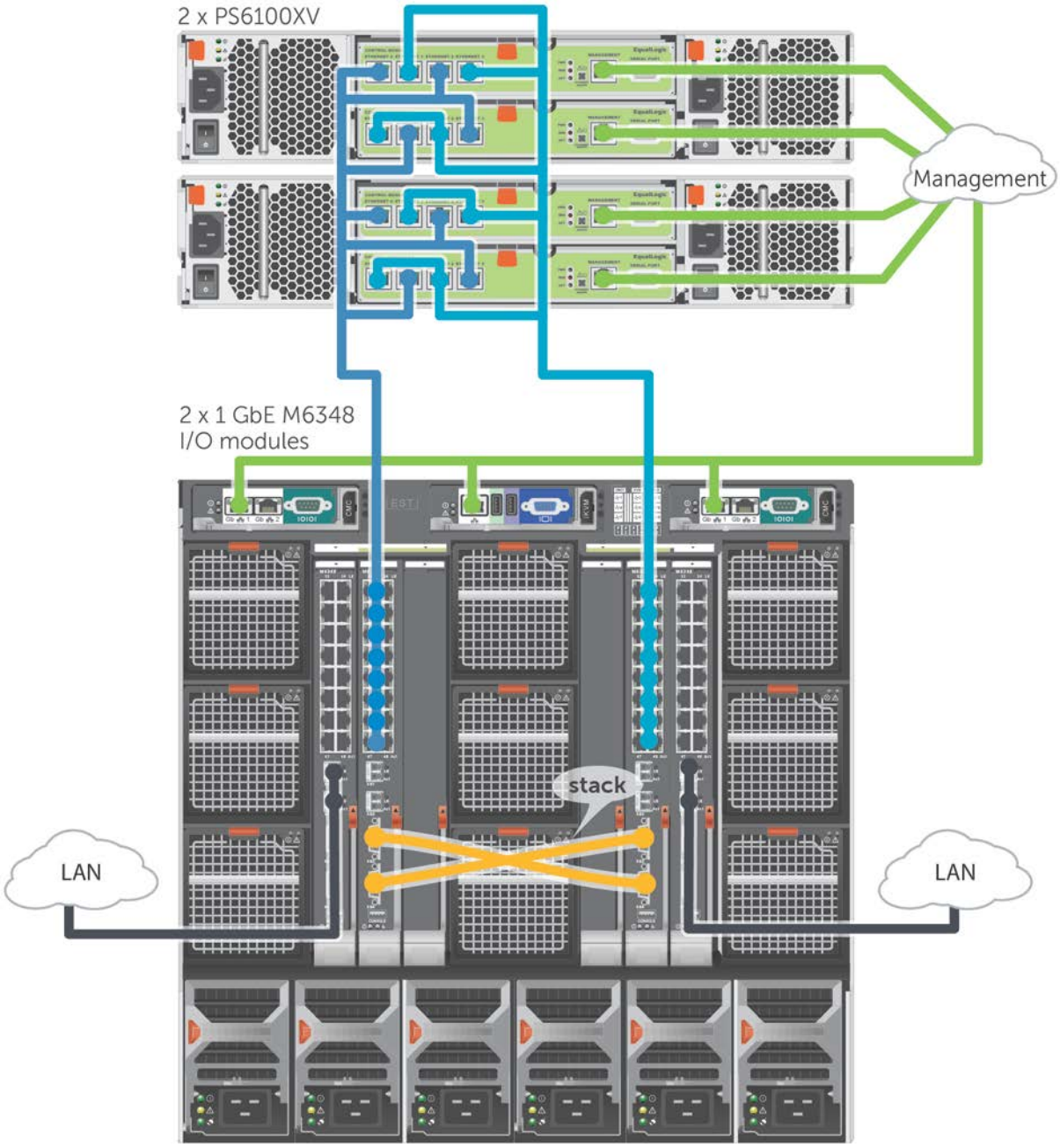

Figure 3 Blade IOM switch only with ISL stack

4.1.2 M6348 switch with ISL LAG

This SAN design provides 20Gbps of ISL bandwidth between the two M6348 switches using the two integrated 10GbE SFP+ ports to create a LAG. Since there is only a single tier of switches, there is no uplink to external switches. The 16 external ports on each M6348 (32 ports total) can accommodate the connection of four 1GbE PS series array members, each of which require eight ports for the active and passive controllers combined. The host/storage port ratio with the maximum number of array members is 2:1.

The following diagram illustrates how the two PS6100XV array members directly connect to the two M6348 switches in Fabric B of the M1000e blade chassis and how the two M6348 switches are connected by a LAG using the integrated 10GbE SFP+ ports. Note how each array member controller connects to both M6348 switches for SAN redundancy. Also note that the corresponding port on the passive controller is connected to a different switch than the port on the active controller, ensuring that the port-based failover of the PS6100 array member will connect to a different switch upon port, cable or switch failure. Management and host LAN networks are shown for reference.

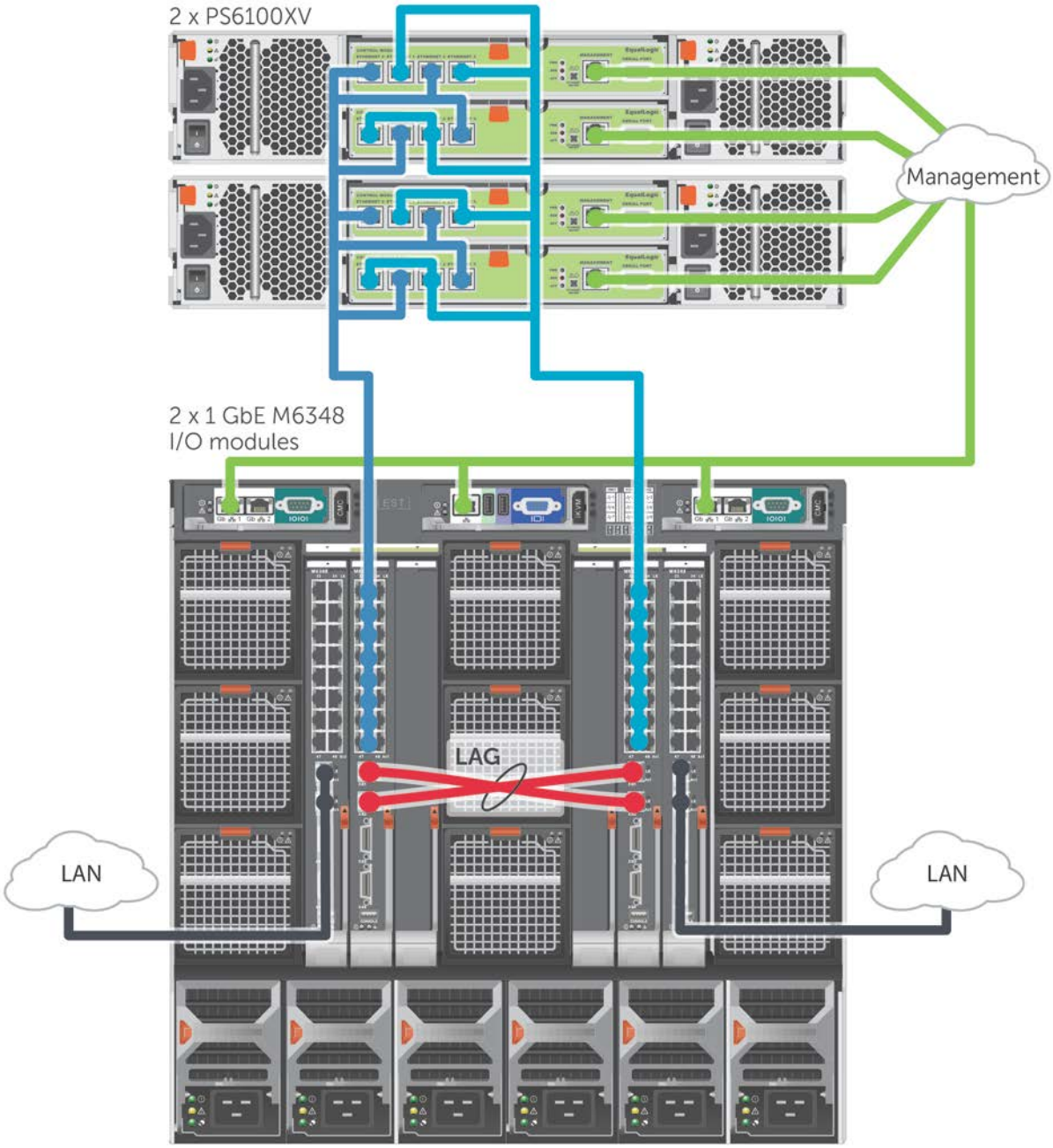

Figure 4 Blade IOM switch only with ISL LAG

4.2 TOR switch only

These SAN designs include configurations where the blade server host ports are directly connected to TOR switches using 1GbE pass-through IOM in the M1000e blade chassis. The storage ports are also connected to the TOR switches, in this case a pair of PowerConnect 7048. One SAN design of this type was tested:

1. PC7048 switches connected with an ISL LAG

Note that because an ISL stack is not a recommended configuration due to a lack of SAN availability during stack reloads, only one single switch tier design with ISL stack was tested – Blade IOM switch only – and the TOR switch only with ISL stack design was excluded.

4.2.1 PC7048 switch with ISL LAG

This SAN design provides 20Gbps of ISL bandwidth between the two PC7048 switches using two 10GbE SFP+ expansion module ports to create a LAG. Since there is only a single tier of switches, there is no uplink from the blade IOM switches. The remaining 32 ports on each PC7048 (64 ports total) can accommodate the connection of eight 1GbE PS series array members, each of which require eight ports for the active and passive controllers combined. The host/storage port ratio with the maximum number of array members is 1:1.

The following diagram illustrates how two PS6100XV array members directly connect to the two TOR PC7048 switches and how the two switches are connected by an ISL LAG using the expansion module 10GbE SFP+ ports. It also shows the connection of four server blades each with two host ports to the PC7048 switches using the 1GbE pass-through IOM in Fabric B of the M1000e chassis. Note how each array member controller connects to both PC7048 switches for SAN redundancy. Also note that the corresponding port on the passive controller is connected to a different switch than the port on the active controller, ensuring that the port-based failover of the PS6100 array member will connect to a different switch upon port, cable or switch failure. Management and host LAN networks are shown for reference.

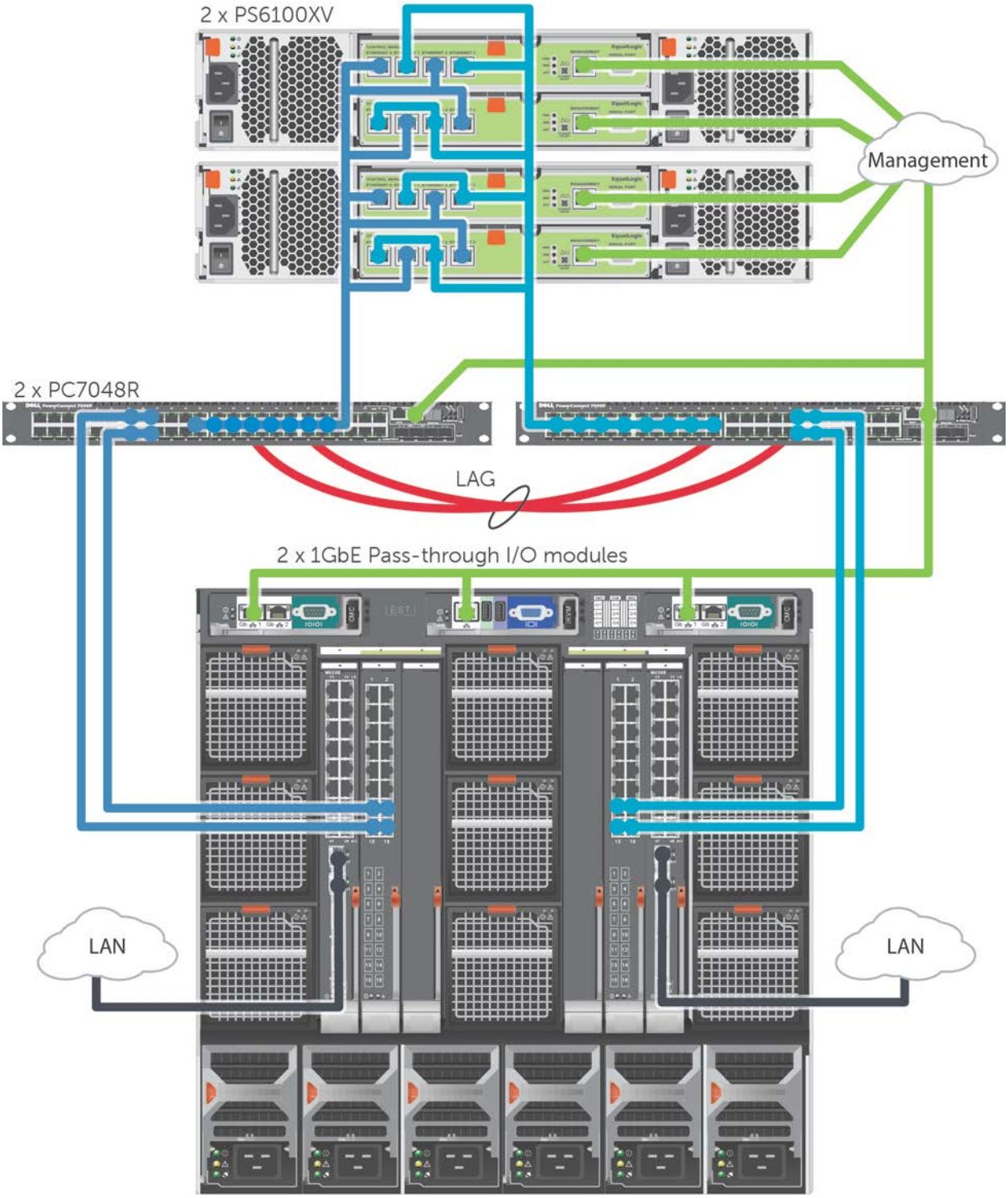

Figure 5 TOR switch only with ISL LAG

4.3 Blade IOM switch with TOR switch

These SAN designs include configurations in which the EqualLogic 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 independly of the array members. The switches within each switch tier are connected to each other by an ISL stack or LAG. It is also a best practice to have the ISL span the TOR switches connecting the array members to better facilitate inter-array member communication. The switch tiers themselves are connected by an uplink stack or LAG.

In this case the TOR switches in the storage tier are PowerConnect 7048 and the blade IOM switches in the host tier are PowerConnect M6348 residing in the M1000e chassis. Note that because the PC7048 switch and the M6348 switch are in the same switch product family, stacking the uplink between switch tiers is possible. Four network designs of this type were tested:

1. All four M6348 and PC7048 switches interconnected with an administrative stack – a *four-way stack*
2. All four M6348 and PC7048 switches interconnected with two uplink LAGs and one ISL LAG between the TOR switches – a *three-way LAG*
3. Each M6348/PC7048 pair connected with uplink stacks, and a single ISL LAG connecting both pairs in a mesh pattern – *uplink stacks*
4. Each switch tier pair connected in ISL stacks, with a single uplink LAG connecting both tiers in a mesh pattern – *ISL stacks*

4.3.1 M6348/PC7048 switches with four-way stack

This SAN design uses the 16Gb stacking ports provided by the PC7048 expansion module and the stacking ports integrated into the M6348 to create a "completed ring" stack of all four switches. It provides 32Gbps of uplink bandwidth between the storage tier of PC7048 switches and the host tier of M6348 switches, while providing 32Gbps of ISL bandwidth. As with all of the multiple switch tier designs, all 48 ports on each PC7048 (96 ports total) are available to the storage and can accommodate the connection of 12 1GbE PS series array members, each of which require eight ports for the active and passive controllers combined. The host/storage port ratio with the maximum number of array members is 2:1.

The following diagram illustrates how two PS6100XV array members connect to the two TOR PC7048 switches and how all four switches are interconnected in a "completed ring" stack using the available stack ports of each switch. This network design limits the uplink bandwidth to a total of 32Gbps (16Gbps for each switch) due to the M6348 having only two stacking ports. Each array member controller connects to both PC7048 switches for SAN redundancy. Note that the corresponding port on the passive controller is connected to a different switch than the port on the active controller, ensuring that the port-based failover of the PS6100 array member will connect to a different switch upon port, cable or switch failure. Management and host LAN networks are shown for reference.

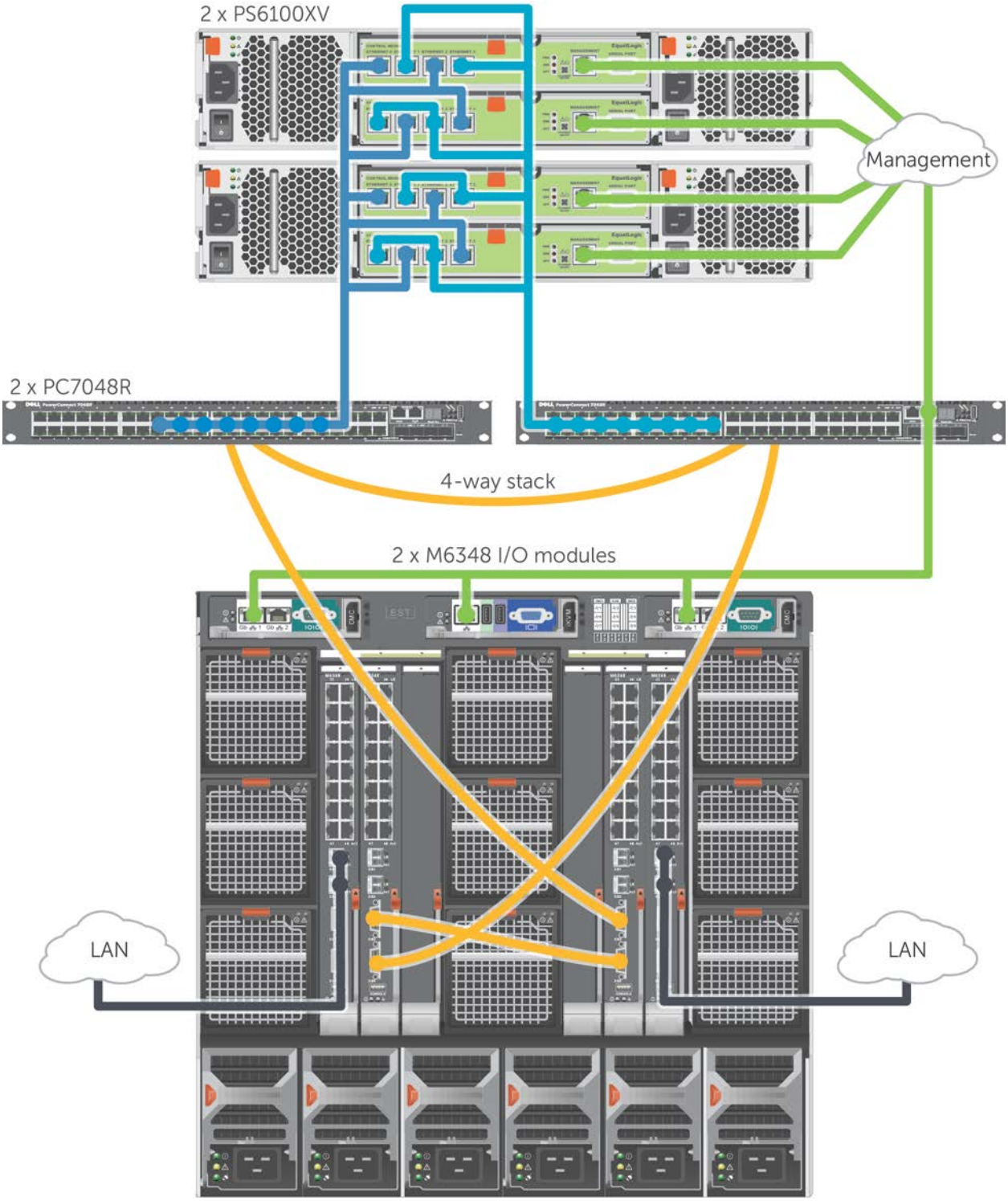

Figure 6 Blade IOM switch with TOR switch and a four-way stack

4.3.2 M6348/PC7048 switches with three-way LAG

This SAN design uses the 10GbE SFP+ ports provided by the PC7048 expansion module and the 10GbE SFP+ ports integrated into the M6348 to setup two separate uplink LAGs and one ISL LAG between the TOR PC7048 switches. It provides 40Gbps of uplink bandwidth between the storage tier of PC7048 switches and the host tier of M6348 switches, while providing 20Gbps of ISL bandwidth. Although a four-way LAG is possible by creating an additional ISL between the M6348 switches, it requires the use of one 10GbE SFP+ port on each of the M6348 leaving only two total 10GbE SFP+ ports with which to create the uplink. Since this would only allow for 20Gbps of uplink bandwidth the four-way LAG design was not considered.

As with all of the multiple switch tier designs, all 48 ports on each PC7048 (96 ports total) are available to the storage and can accommodate the connection of 12 1GbE PS series array members, each of which require eight ports for the active and passive controllers combined. The host/storage port ratio with the maximum number of array members is 2:1.

The following diagram illustrates how two PS6100XV array members connect to the two TOR PC7048 switches and how all four switches are interconnected with two uplink LAGs and one ISL LAG between the TOR switches using the available SFP+ ports of each switch. This network design requires the use of two 10GbE uplink expansion modules in each of the PC7048 switches. Each array member controller connects to both PC7048 switches for SAN redundancy. Note that the corresponding port on the passive controller is connected to a different switch than the port on the active controller, ensuring that the port-based failover of the PS6100 array member will connect to a different switch upon port, cable or switch failure. Management and host LAN networks are shown for reference.

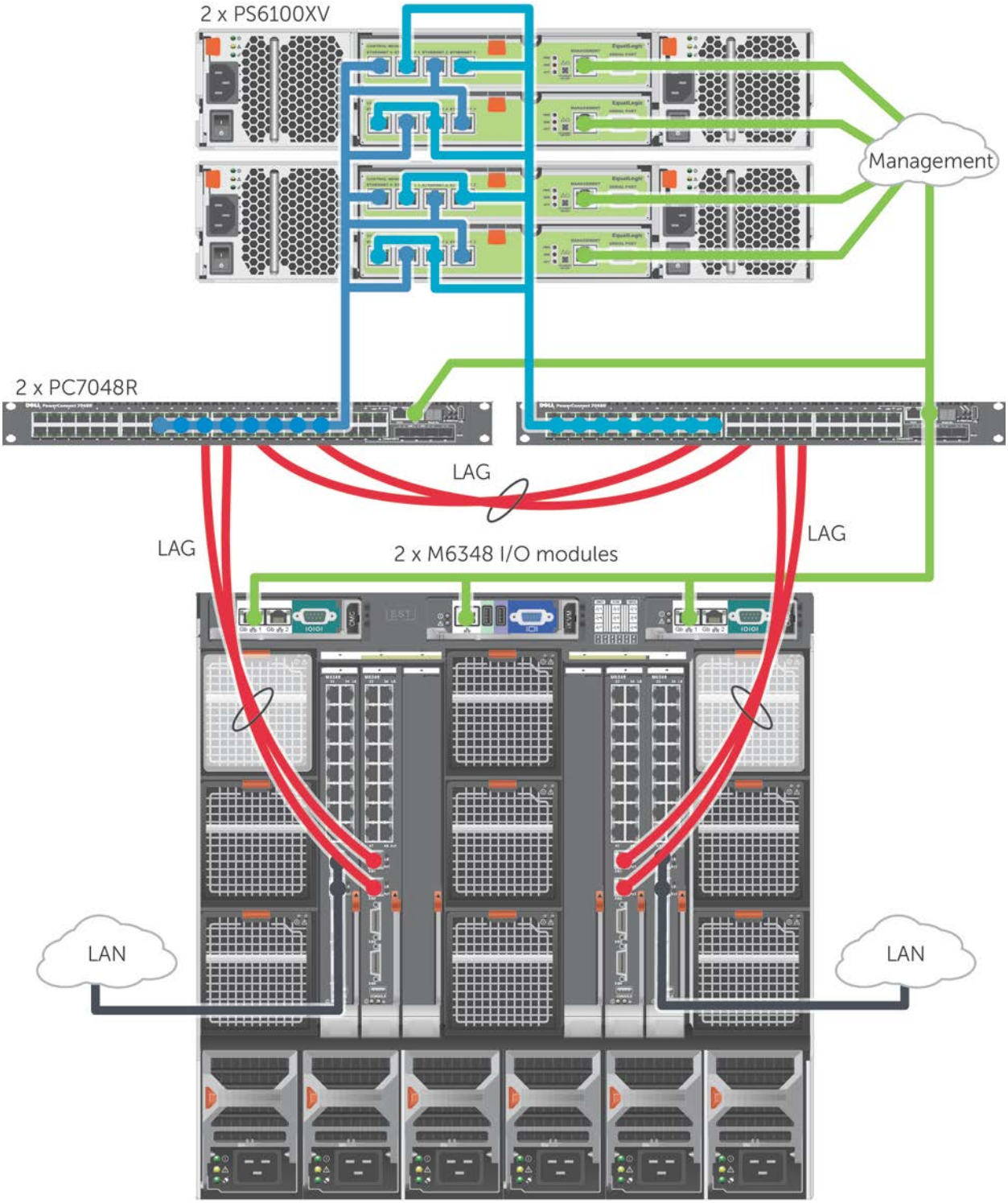

Figure 7 Blade IOM switch with TOR switch and a three-way LAG

4.3.3 M6348/PC7048 switches with uplinks stacked

This SAN design uses the 16Gb stacking ports provided by the PC7048 expansion module and the stacking ports integrated into the M6348 to configure two uplink stacks between each pair of M6348 and PC7048 switches. Additionally, one ISL LAG connecting each uplink stack in a mesh pattern is created using the 10GbE SFP+ ports provided by the PC7048 expansion module and the 10GbE SFP+ ports integrated into the M6348. It provides 64Gbps of uplink bandwidth between the storage tier of PC7048 switches and the host tier of M6348 switches, while providing 40Gbps of ISL bandwidth. As with all of the multiple switch tier designs, all 48 ports on each PC7048 (96 ports total) are available to the storage and can accommodate the connection of 12 1GbE PS series array members, each of which require eight ports for the active and passive controllers combined. The host/storage port ratio with the maximum number of array members is 2:1.

The following diagram illustrates how two PS6100XV array members connect to the two TOR PC7048 switches and how all four switches are interconnected by two uplink stacks and one ISL LAG using the available stacking ports and SFP+ ports of each switch. This network design requires the use of one stacking and one 10GbE uplink expansion module in each of the PC7048 switches. Each array member controller connects to both PC7048 switches for SAN fabric redundancy. Note that the corresponding port on the passive controller is connected to a different switch than the port on the active controller, ensuring that the port-based failover of the PS6100 array member will connect to a different switch upon port, cable or switch failure. Management and host LAN networks are shown for reference.

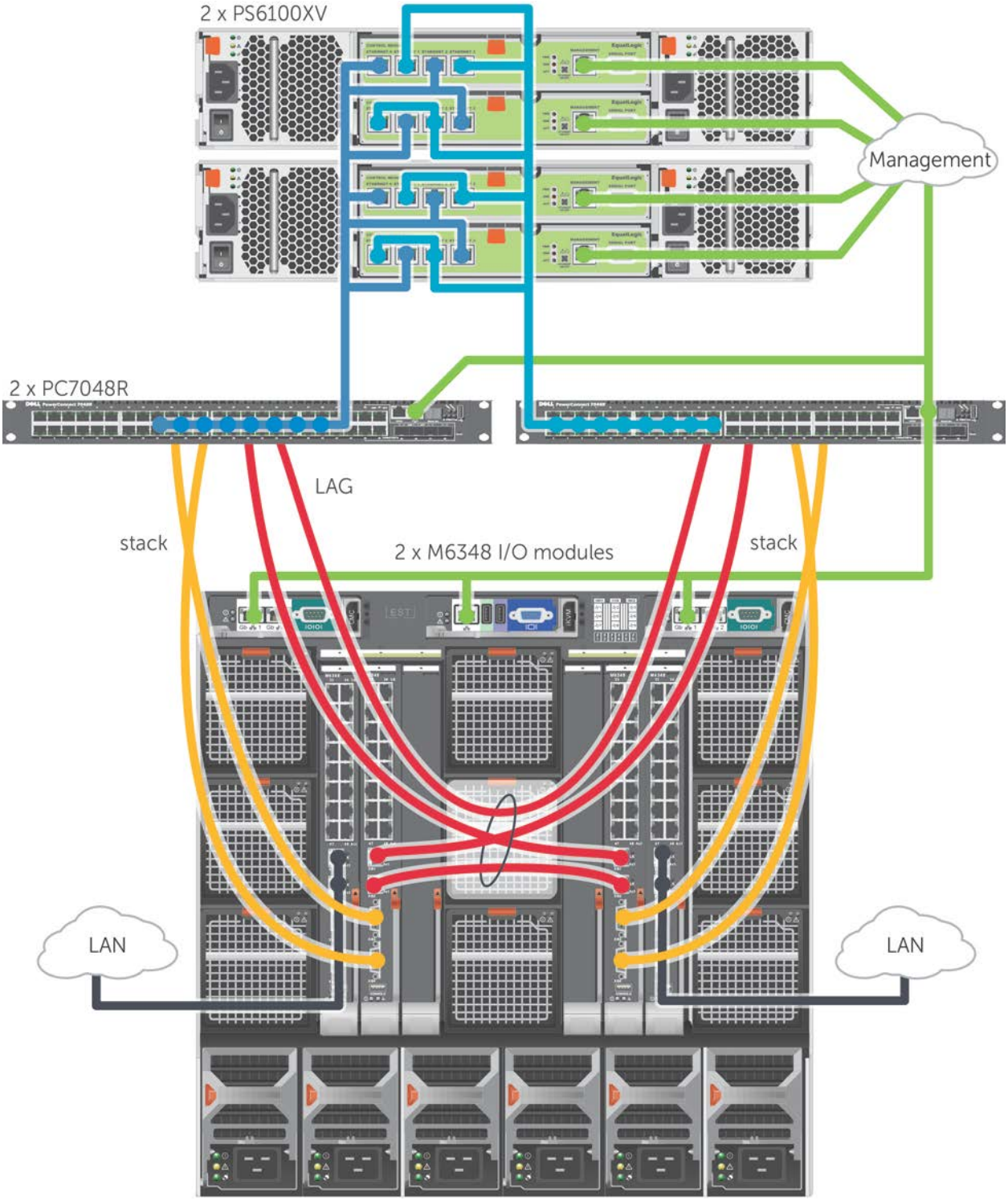

Figure 8 Blade IOM switch with TOR switch and uplink stacks

4.3.4 M6348/PC7048 switches with ISL stacks

This SAN design uses the 16Gb stacking ports provided by the PC7048 expansion module and the stacking ports integrated into the M6348 to setup two ISL stacks between each pair of like switches. Additionally, an uplink LAG connecting each switch tier in a mesh pattern is created using the 10GbE SFP+ ports provided by the PC7048 expansion module and the 10GbE SFP+ ports integrated into the M6348. It provides 40Gbps of uplink bandwidth between the storage tier of PC7048 switches and the host tier of M6348 switches, while providing 64Gbps of ISL bandwidth. As with all of the multiple switch tier designs, all 48 ports on each PC7048 (96 ports total) are available to the storage and can accommodate the connection of twelve 1GbE PS series array members, each of which require eight ports for the active and passive controllers combined. The host/storage port ratio with the maximum number of array members is 2:1.

The following diagram illustrates how two PS6100XV array members connect to the two TOR PC7048 switches and how all four switches are interconnected by multiple switch tier stacks and one uplink LAG using the available stacking ports and SFP+ ports of each switch. This network design requires the use of one stacking and one 10GbE uplink expansion module in each of the PC7048 switches. Each array member controller connects to both PC7048 switches for SAN fabric redundancy. Note that the corresponding port on the passive controller is connected to a different switch than the port on the active controller, ensuring that the port-based failover of the PS6100 array member will connect to a different switch upon port, cable or switch failure. Management and host LAN networks are shown for reference.

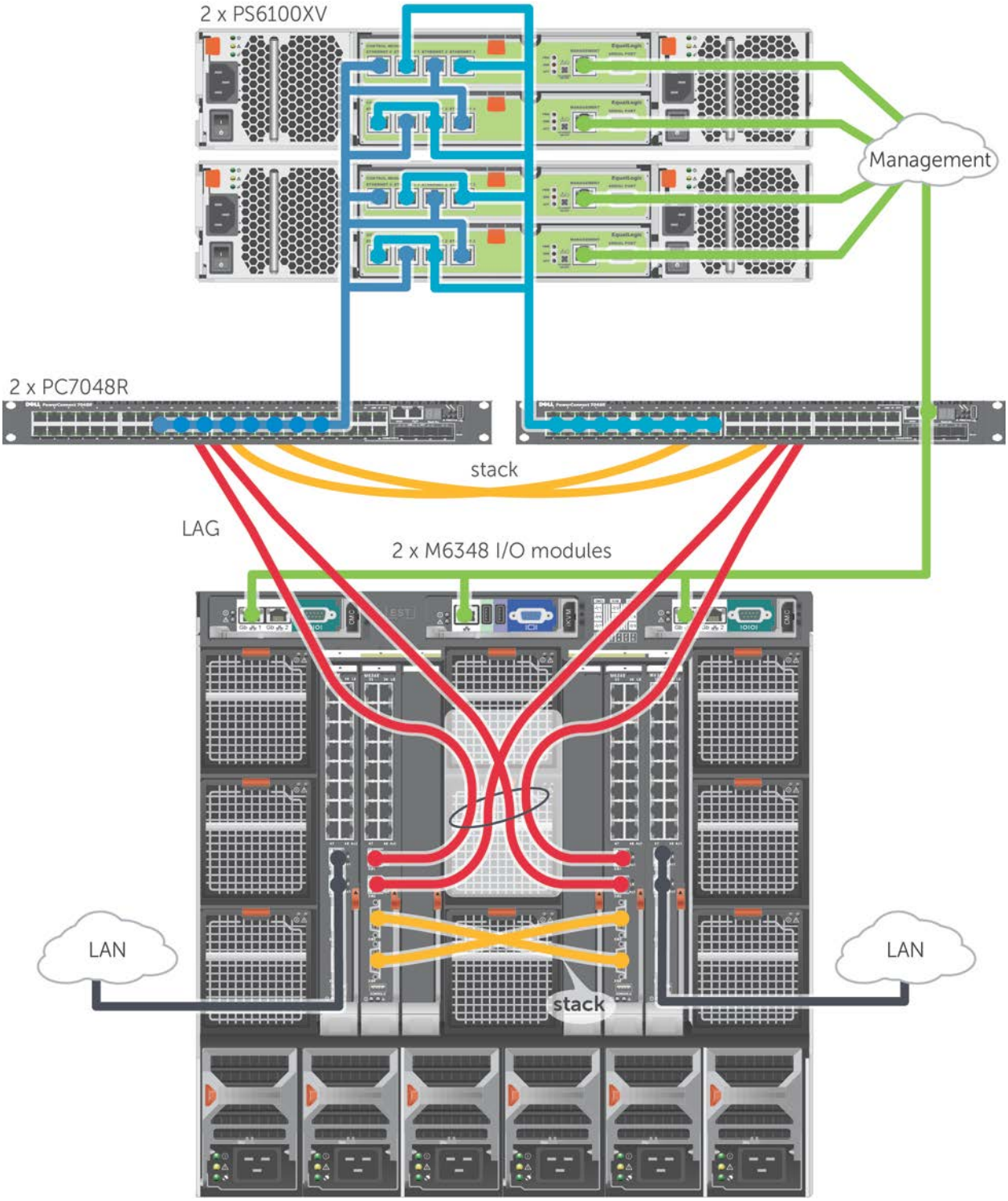

Figure 9 Blade IOM switch with TOR switch and ISL stacks

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 PC7048 switches or dual M6348 switches in a single M1000e blade chassis IO 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.

Green cells indicate the recommended SAN design within each design category based on all factors considered during testing, while orange cells indicate designs that might not be preferred.

Table 4 A comparison of all tested SAN designs

	Host switch type	Array member switch type	Total uplink bandwidth	Total ISL bandwidth	Maximum number of hosts	Maximum number of arrays members	Port ratio with maximum hosts/array members
Blade IOM with ISL stack	Blade	Blade	N/A	32Gbps	16	4	2:1
Blade IOM with ISL LAG	Blade	Blade	N/A	20Gbps	16	4	2:1
TOR with ISL LAG	TOR	TOR	N/A	20Gbps	16	8	1:1
Blade IOM and TOR with 4-way stack	Blade	TOR	32Gbps	32Gbps	16	12	.67:1
Blade IOM and TOR with 3-way LAG	Blade	TOR	40Gbps	20Gbps	16	12	.67:1
Blade IOM and TOR with uplink stacks	Blade	TOR	64Gbps	40Gbps	16	12	.67:1
Blade IOM and TOR with ISL stacks	Blade	TOR	40Gbps	64Gbps	16	12	.67:1

5 Detailed SAN design analysis and recommendations

The following section examines each M1000e blade chassis and EqualLogic 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. Section 5.1 provides a list of common administration considerations and how each is affected by SAN design choice.

5.1.1 Stack vs. LAG

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, will still have an ISL between switches. For multiple switch tier designs, the uplink between switch tiers needs sufficient bandwidth to prevent constraining the throughput of SAN traffic. 10GbE SFP+ ports or proprietary stacking ports are the best solution for an ISL or an uplink and should be used whenever possible. The PC7048 switch can provide up to four 10GbE SFP+ ports or four 16Gb stacking ports using two expansion modules in the back of the switch. The M6348 blade switch has two 10GbE SFP+ ports and two 16Gb stacking ports built into the switch chassis.

PowerConnect switch stacking ports can be used to create an administrative stack with other PowerConnect switches in the same product family. The PC7048 and the M6348 switches are stack compatible. From an administrative perspective, a switch stack is preferred because it allows the administration of multiple switches as if they were one physical unit. First, on the PowerConnect 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 thing to note with this configuration is that a switch stack reset will bring down all switch units simultaneously and if switches within a tier are stacked together, then the SAN becomes unavailable. The resulting SAN downtime must be scheduled.

The alternative inter-switch connection to the administrative switch stack 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 PowerConnect switches is very straightforward and administrative complexity is not a concern.

In a multiple switch tier SAN design, stack and LAG may be combined to achieve the benefits of both options. See Section 4 for diagrams of multiple switch tier SAN designs which use both a stack and a LAG.

5.1.2 Hardware requirements

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

The blade IOM switch only SAN design requires the fewest cables, with only the array member ports and a single ISL stack or LAG at the M1000e chassis to cable. The blade IOM switch with TOR switch SAN designs require the addition of two or more stacks or LAGs, and finally the TOR switch only designs (with pass-through IOM), while needing only one ISL stack/LAG, requires 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.

5.1.3 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 as long as uplinks are not stacked. So for example if a SAN consisting of EqualLogic PS Series array members and an M1000e blade chassis were being deployed in a datacenter with an existing layer of non-PowerConnect switches, there are blade IOM switch with TOR switch designs and TOR switch only designs which could accommodate such a scenario.

Note that multiple switch tier SAN designs that do require uplink stacks – the four-way stack and the uplink stack designs – would only be possible with stack-compatible Dell TOR switches.

Also note that while setting up a LAG between two PowerConnect is very straightforward, configuring a LAG between a PowerConnect and a switch of a different vendor might be more complex. While there is an industry-standard protocol for link aggregation, LACP, many switch vendors have their own unique implementations and additional diligence might be required to ensure a properly functioning LAG. Thus even two tier SAN designs without an uplink stack – the 3-way LAG and the switch tier stack designs – require a bit more administrative planning when TOR switches of alternate switch vendors are used.

5.1.4 Recommendations

In summary, when reducing administrative overhead is the goal, a single switch tier design with an ISL stack 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. For multiple switch tier SAN designs, the 4-way stack is the easiest to setup and maintain.

If the availability of the SAN is critical, then an ISL LAG configuration may be preferred over stacking. If a switch tier ISL is stacked, then a switch stack reload (required for tasks such as switch firmware updates) will temporarily reset the entire switch tier making the SAN unavailable during that time. In this case, SAN downtime for firmware updates would have to be scheduled. A multiple switch tier design that avoids this is the uplink stack with ISL LAG design, which provides some administrative benefit while ensuring the SAN is always available.

If TOR switches from a different vendor are used, then 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 as the blade IOM switches will not be stack-compatible with the TOR switches from a different vendor.

5.2 Performance

The second criterion by which SAN designs will be evaluated is their performance relative to each other. Section 5.2 reports the performance results of each SAN design under two common IO workloads.

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."*

Vdbench is available at: http://sourceforge.net/projects/vdbench/

Each performance test was conducted with the following hardware and software. Note that all EqualLogic 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. See Appendix B for a list of Vdbench parameters.

Hosts:

- Four PowerEdge M610 blade servers each with:
  - Windows Server 2008 R2
  - Dell EqualLogic Host Integration Toolkit v4.0.0
  - Two 1GbE ports on the SAN

Storage:

- Two EqualLogic PS6100XV array members each with:
  - Firmware: 5.2.2 R229536
  - Four active 1GbE ports on the SAN

- Four iSCSI volumes dedicated to each host

Note that there were a total of eight host ports and eight storage ports for the ideal 1:1 ratio.

The following two Vdbench workloads were defined:

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

Each Vdbench workload was run for one hour 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.

5.2.2 Bandwidth

All SAN designs provide different amounts of ISL bandwidth between the two switches within each switch tier. While single switch tier designs have host and storage ports connected to the same

switches, multiple switch tier SAN design require an uplink stack or LAG between switch tiers. Each multiple switch tier design provides a different amount of uplink bandwidth between the host and storage switch tiers.

Uplink bandwidth should be at least equal to the aggregate bandwidth of all active PS Series array member ports. For example, twelve array members with four active ports each would require 48Gbps of uplink bandwidth. Choosing a SAN design that maximizes uplink bandwidth is of particular importance for larger scale SAN.

ISL bandwidth is also important. Since it is a best practice to create a redundant SAN fabric with at least two switches in each switch tier, SAN traffic will often have to cross the ISL. Assuming a worst case scenario of 100% of all SAN traffic crossing the ISL in both directions (half going one way and half going the other) the ISL bandwidth requirements are 50% of the uplink bandwidth. The ISL bandwidth of each SAN design should be considered accordingly.

The following table shows the uplink and ISL bandwidth of each SAN design. Each of the single switch tier designs (highlighted in green) provide adequate uplink and ISL bandwidth for the maximum number of array members that their port counts accommodate. However, only one of the multiple switch tier designs provide adequate uplink and ISL bandwidth. Only the uplink stacks design (also highlighted in green) provides more than the 48Gbps of uplink bandwidth and 24Gbps of ISL bandwidth required by twelve array members. The other three multiple switch tier designs (highlighted in orange) do not provide adequate uplink bandwidth for up to twelve member arrays.

Green cells indicate the recommended SAN design within each design category based on all factors considered during testing, while orange cells indicate designs that might not be preferred.

Table 5 A comparison of the bandwidth provided by all SAN designs

	Switch tier topology	Total uplink bandwidth	Total ISL bandwidth	Maximum number of host ports	Maximum number of array member active ports	Port ratio with maximum hosts/array members
Blade IOM only with ISL stack	Single	N/A	32Gbps	32	16	2:1
Blade IOM only with ISL LAG	Single	N/A	20Gbps	32	16	2:1
TOR only with ISL LAG	Single	N/A	20Gbps	32	32	1:1
Blade IOM and TOR with four-way stack	Multiple	32Gbps	32Gbps	32	48	.67:1
Blade IOM and TOR with three-way LAG	Multiple	40Gbps	20Gbps	32	48	.67:1
Blade IOM and TOR with uplink stacks	Multiple	64Gbps	40Gbps	32	48	.67:1
Blade IOM and TOR with ISL stacks	Multiple	40Gbps	64Gbps	32	48	.67:1

5.2.3 Results

The following two figures show the relative aggregate Vdbench throughput of all four hosts within each SAN design at two different I/O workloads. Each throughput value is presented as a percentage of a baseline value. In each chart, the blade IOM switch only with ISL LAG design was chosen as the baseline value. All throughput values were achieved during a single one hour test run and are not an average of multiple test runs.

5.2.3.1 8KB random I/O, 67% read workload

The following figure shows the aggregate Vdbench throughput of all four hosts within each SAN design at an 8KB random I/O, 67% read workload. All SAN designs yielded throughput results within 10% of the baseline value.

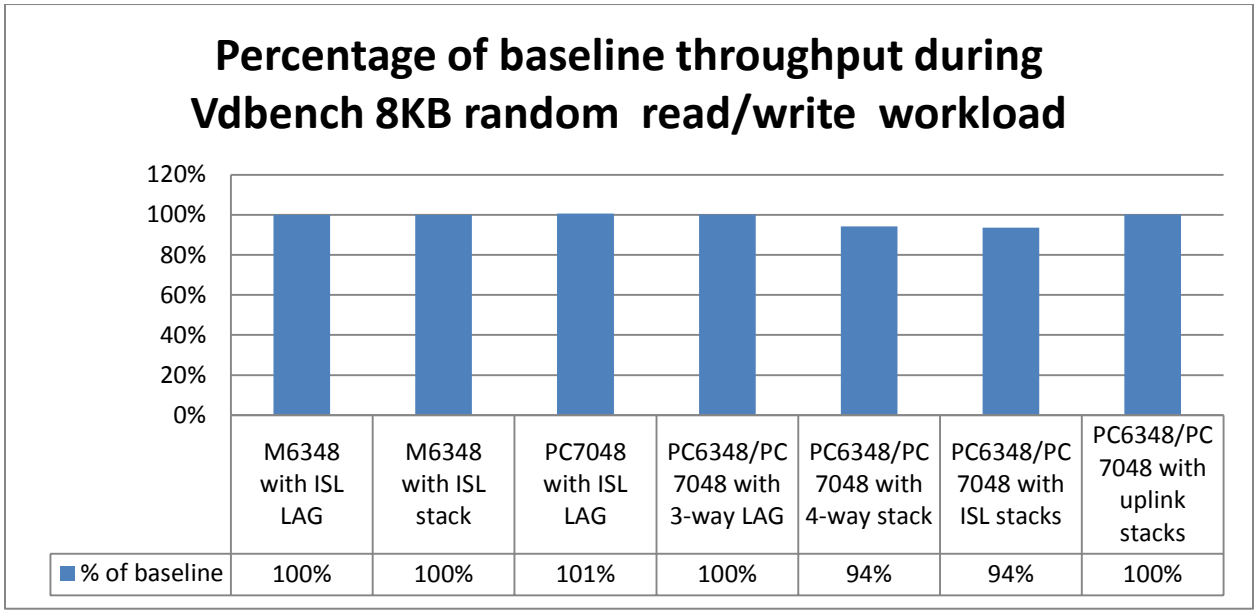

Figure 10 Aggregate Vdbench throughput as a percentage of the baseline value in each SAN design during an 8KB random I/O, 67% read workload

5.2.3.2 256KB sequential I/O, read workload

The following figure shows the aggregate Vdbench throughput of all four hosts within each SAN design at a 256KB sequential I/O, read workload. All SAN designs yielded throughput results within 10% of the baseline value.

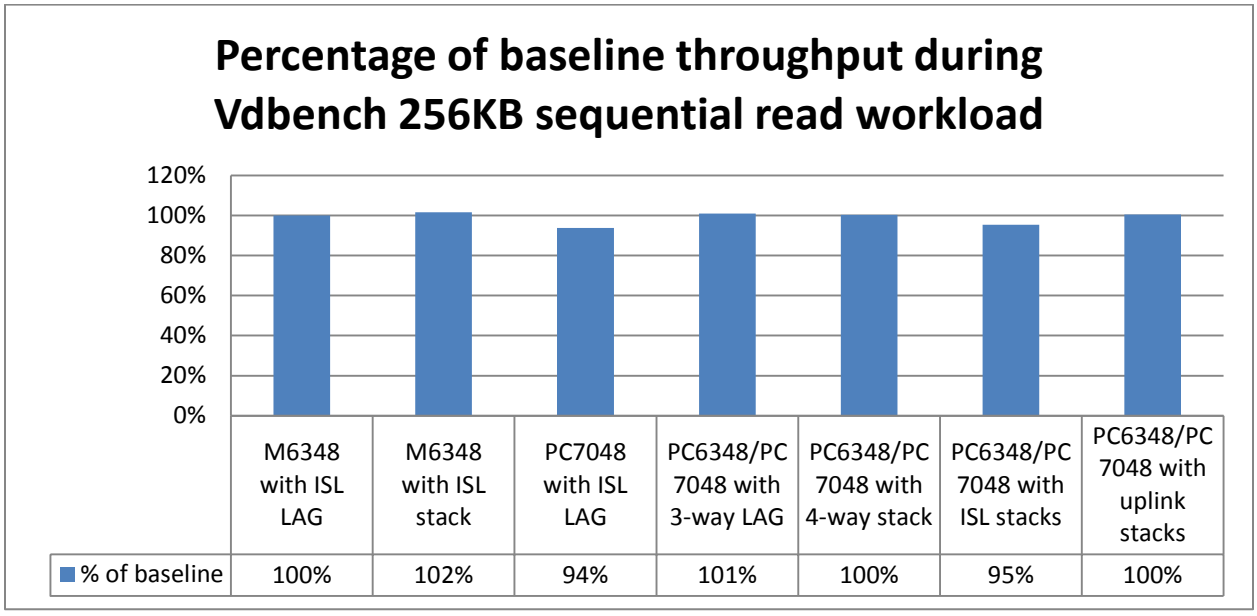

Figure 11 Aggregate Vdbench throughput as a percentage of the baseline value in each SAN design during a 256KB sequential I/O, read workload

5.2.4 Recommendations

The throughput values were gathered during the performance testing of each SAN design with four hosts and two arrays members at two common workloads. For a smaller scale SAN, there were no significant performance differences among the tested SAN designs.

For larger scale SANs, it is more likely that a two tier SAN design will be chosen because this accommodates the largest number of member arrays (12). In this case, only the uplink stacks SAN design provides adequate bandwidth for both the uplink (64Gbps) and the ISL (40Gbps)

5.3 High availability

The third criterion by which SAN designs will be evaluated is how each design tolerates a switch failure. Section 5.3 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 that storage port disconnection is not addressed in the tables because the PS6100XV 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 256KB sequential I/O write using Vdbench.

In all cases, Vdbench I/O continued without error and no iSCSI volume disruptions were observed. In cases where host ports were disconnected, iSCSI connections were appropriately migrated to the

remaining host ports. Also in these cases, the loss of 50% of the host ports reduced the Vdbench throughput by the same amount, as would be expected.

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 all applicable SAN designs, a TOR switch failure reduces the uplink bandwidth by 50%, however the 4-way stack, uplink stack, and ISL stack designs do not suffer a loss of 50% of the host ports. It is definitely noteworthy that out of all SAN designs only the uplink stack design (highlighted in green) keeps all host ports connected through a TOR switch failure and at 32Gbps still retains sufficient uplink bandwidth to accommodate the 32 host ports.

All applicable SAN designs retain enough ISL bandwidth to accommodate the expected ISL traffic of the remaining host ports. As discussed in Section 6.2.2 this is normally about 50% of the expected uplink throughput.

Green cells indicate the recommended SAN design within each design category based on all factors considered during testing, while orange cells indicate designs that might not be preferred.

Table 6 A comparison of the way each SAN designs tolerates a TOR switch failure

	Reduction in connected host ports	Reduction in uplink bandwidth	Reduction in ISL bandwidth
Blade IOM only with ISL stack	N/A	N/A	N/A
Blade IOM only with ISL LAG	N/A	N/A	N/A
TOR only with ISL LAG	32 → 16	N/A	20Gbps → N/A**
Blade IOM and TOR with four-way stack	32 → 32	32 → 16Gbps	32 → 16Gbps
Blade IOM and TOR with three-way LAG	32 → 16	40 → 20Gbps	20Gbps → N/A**
Blade IOM and TOR with uplink stacks	32 → 32	64 → 32Gbps	40 → 20Gbps
Blade IOM and TOR with ISL stacks	32 → 32	40 → 20Gbps	64 → 32Gbps

**ISL bandwidth is no longer relevant because the switch failure eliminates the ISL. Note that this happens in conjunction with the loss of the 50% of the host ports connected to the remaining TOR switch.

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 host ports (and hence expected throughput) by 50%. All multiple switch tier designs suffer a 50% reduction in uplink bandwidth, but since the host port number and expected throughput are reduced by the same

percentage the loss of bandwidth is not a factor. Of all multiple switch tier designs, the uplink stack (highlighted in green) retains by far the most uplink bandwidth.

All applicable SAN designs retain enough ISL bandwidth to accommodate the expected ISL traffic of the remaining host ports. As discussed in Section 5.2.2 this is normally about 50% of the expected throughput between the host and storage ports.

Green cells indicate the recommended SAN design within each design category based on all factors considered during testing, while orange cells indicate designs that might not be preferred.

Table 7 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 ISL bandwidth
Blade IOM only with ISL stack	32 → 16	N/A	32Gbps → N/A**
Blade IOM only with ISL LAG	32 → 16	N/A	20Gbps → N/A**
TOR only with ISL LAG	N/A	N/A	N/A
Blade IOM and TOR with four-way stack	32 → 16	32 → 16Gbps	32 → 16Gbps
Blade IOM and TOR with three-way LAG	32 → 16	40 → 20Gbps	20 → 20Gbps
Blade IOM and TOR with uplink stacks	32 → 16	64 → 32Gbps	40 → 20Gbps
Blade IOM and TOR with ISL stacks	32 → 16	40 → 20Gbps	64 → 32Gbps

**ISL bandwidth is no longer relevant because the switch failure eliminates the ISL. Note that this happens in conjunction with the loss of the 50% of the host ports connected to the failing blade IOM switch.

5.3.3 Recommendations

In both the TOR and blade IOM switch failure scenarios, the blade IOM switch and TOR switch with uplink stacks SAN design retained as many or more host port connections while retaining the highest amount of uplink bandwidth of any other applicable SAN design.

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. Section 5.4.1 will list the maximum number of array members and the resulting host/storage port ratios for each SAN design. Section 5.4.2 will discuss how each SAN design would accommodate additional M1000e blade chassis or PS Series array members.

5.4.1 Host / array member port ratios for single chassis

The following table shows the maximum number of array members supported by each SAN design assuming a single M1000e chassis and 16 half-height blade servers with two M6348 switches or two pass-through IO modules, two SAN ports per host and, if applicable, two 48-port TOR switches. Note that the blade IOM switch only SAN designs allow the fewest array members per blade chassis and hence have the highest host/storage port ratio, 2:1. TOR switch only designs support twice as many

array members as the blade IOM switch only designs, yielding an ideal 1:1 host/storage port ratio. The TOR switches in the multiple switch tier designs are entirely dedicated to storage and can accommodate up to 12 array members. With that many array members, it begins to make sense to add chassis to increase the host/storage port ratio.

Table 8 A comparison of all SAN designs with a single M1000e chassis

	Switch tier topology	Host switch type	Array member switch type	Maximum number of array members	Port ratio with maximum hosts/array members
Blade IOM only with ISL stack	Single	Blade	Blade	4	2:1
Blade IOM only with ISL LAG	Single	Blade	Blade	4	2:1
TOR only with ISL LAG	Single	TOR	TOR	8	1:1
Blade IOM and TOR with four-way stack	Multiple	Blade	TOR	12	.67:1
Blade IOM and TOR with three-way LAG	Multiple	Blade	TOR	12	.67:1
Blade IOM and TOR with uplink stacks	Multiple	Blade	TOR	12	.67:1
Blade IOM and TOR with ISL stacks	Multiple	Blade	TOR	12	.67:1

5.4.2 Adding blade chassis or array members

M1000e blade chassis, switches, and PS Series array members will scale at different rates because of the differing numbers of ports available to storage ports and the location of the switches in each SAN design.

With blade IOM switch only designs, in order to scale the number of array members a new blade chassis must be added to provide additional switches for the storage to connect to. As with a single chassis, each additional blade chassis supports up to four new array members on a single IOM fabric. Once a second blade chassis is added, it makes sense to distribute the four active ports of each array member across the four blade IOM switches to minimize ISL traffic. However, once there are three or more blade chassis, there is a much greater possibility of increased hop-counts and latency as ISL traffic across the blade IOM switches increases. Blade IOM switch only SAN designs are not recommended for more than three M1000e blade chassis. It should also be mentioned that each storage pool should consist of array members that are connected to the same M1000e chassis so that inter-array member traffic does not span more than one ISL.

With the TOR switch only SAN design, scaling the number of TOR switches allows for the addition of blade chassis or array members or both. Each additional 48-port TOR switch allows for one and a half chassis worth of host ports (at 32 host ports per chassis) or six array members or some combination, such as one chassis and two array members. As with the blade IOM switch only SAN design, the TOR

switch only design will experience increased hop-counts and latency as the number of switches increases. Also, each storage pool should consist of array members that are all connected to the same pair of adjacent switches so that inter-array member traffic does not span more than one ISL.

For both single tier switch SAN designs, an ISL stack or LAG will need to extend between the switches (whether blade IOM or TOR) as their number scales.

In the case of the multiple switch tier SAN designs, chassis number can scale independently of the number of TOR switches, while scaling the number of array members requires additional TOR switches. However, given the low host/storage port ratios of the multiple switch tier designs due to their support of up to 12 array members, it is much more likely that blade chassis number will be scaled without ever needing to scale beyond two TOR switches. As the blade IOM switch tier scales beyond the two switches in a single blade chassis, it is necessary to discuss how previous multiple switch tier SAN designs would scale.

The completed ring of an all-way stack only allows for two uplink connections for a total of 32Gbps of bandwidth. Furthermore, the recommended maximum stack size of six switches for M6348 and stack-compatible switches effectively limits this SAN design to two blade chassis.

Both an all-way LAG and an ISL stacks SAN design would suffer from a similar fundamental weakness -- after the ISL stack or LAG between the two TOR switches was in place there would be only four 10GbE SFP+ ports with which to make uplink connections, for a maximum of 40Gbps of uplink bandwidth. The only way to increase the uplink bandwidth when using an uplink LAG would be to increase the number of TOR switches. Furthermore, with only two TOR switches, the four blade IOM switches in two blade chassis could each only have a single non-redundant uplink cable to the TOR switch tier. Beyond two blade chassis and four blade IOM switches there would be more blade IOM switches than uplink ports available on the two TOR switches, and traffic would increasingly have to traverse a blade IOM ISL.

The uplink stacks SAN design can actually be scaled rather easily and would provide 64Gbps of uplink bandwidth, well in excess of even the 48 active ports of 12 array members. This can be accomplished by adding one blade IOM switch from each new M1000e blade chassis to each of the existing uplink stacks, up to six switches in each stack. Thus the uplink stacks SAN design could be scaled to five blade chassis, 80 blade hosts, 160 host ports and 12 array members with 48 active storage ports for a 3.33:1 host/storage port ratio. While this host/port ratio is not ideal from a performance perspective, it is clear that the uplink stacks design is the most scalable multiple switch tier design with respect to chassis/blade number.

The following diagram shows how an uplink stacks SAN design could accommodate an additional chassis. The completed ring stack of each uplink could be extended to include five chassis total. The ISL LAG could be extended to include additional M6348 10GbE SFP+ ports as ISL bandwidth requirements increased.

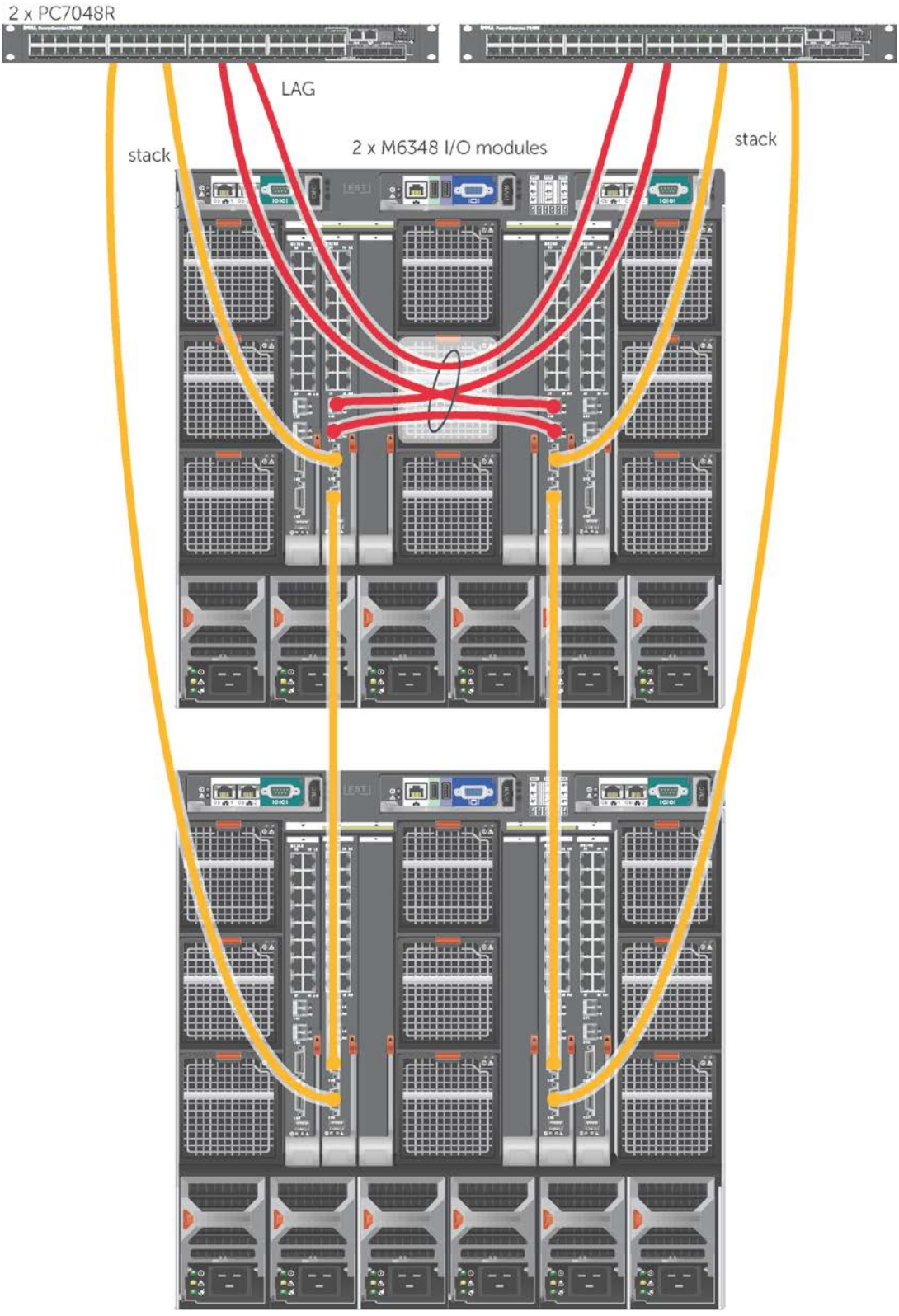

Figure 12 Incorporating an additional M1000e chassis into the blade IOM switch with TOR switch and uplink stacks SAN design

5.4.3 Recommendations

For blade IOM switch only SAN designs, scaling the number of array members is only possible with the addition of M1000e blade chassis and scaling beyond three blade chassis is not recommended due to increased hop-count and latency over the ISL connection.

TOR switch only SAN designs are somewhat more scalable in the sense that they allow up to eight arrays with two TOR switches and one chassis, but just as with the blade IOM switch only designs, the SAN traffic increasingly relies on ISL connections as the number of switches grows.

For larger scale PS Series SAN deployments, the blade IOM switch with TOR switch designs can accommodate a far higher number of array members without the need to add blade chassis or TOR switches. Among these designs, the only the uplink stack design provides adequate uplink and ISL bandwidth while easily accommodating the largest number of chassis by incorporating additional blade IOM switches into the two uplink stacks.

For additional information on multi-blade chassis SAN designs, including the number of supported array members and the maximum recommended stack size for different blade chassis IOM switches, see the Blade Server Chassis Integration section of the Dell EqualLogic Configuration Guide (ECG).

The ECG is available at: http://en.community.dell.com/techcenter/storage/w/wiki/2639.equallogic-configuration-guide.aspx

Appendix A Solution infrastructure detail

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

Table 9 A detailed inventory of the hardware and software configuration in the test environment

Solution configuration - Hardware components:		Description
Blade Enclosure	Dell PowerEdge M1000e chassis: CMC firmware: 4.00	Storage host enclosure
1GbE Blade Servers	(4) Dell PowerEdge M610 server: Windows Server 2008 R2 SP1 BIOS version: 6.1.0 iDRAC firmware: 3.35 (2) Intel® Xeon® X5650 24GB RAM Fabric B – Quad 5709s M 1GbE Driver v14.2.2 Firmware v6.2.16 Only the first (2) ports will be used Dell EqualLogic Host Integration Toolkit v4.0.0	Storage hosts for configs: Blade IOM switch only with ISL stack Blade IOM switch only with ISL LAG TOR switch only with ISL LAG Blade IOM switch and TOR switch with 4-way stack Blade IOM switch and TOR switch with 3-way LAG Blade IOM switch and TOR switch with uplink stacks Blade IOM switch and TOR switch with ISL stacks
1GbE Blade IO modules	(2) Dell PowerConnect M6348 Firmware v4.2.1.3 (2) Dell Ethernet Pass-through module	IO modules for configs: Blade IOM switch only with ISL stack Blade IOM switch only with ISL LAG TOR switch only with ISL LAG Blade IOM switch and TOR switch with 4-way stack Blade IOM switch and TOR switch with 3-way LAG Blade IOM switch and TOR switch

		with uplink stacks Blade IOM switch and TOR switch with ISL stacks
1GbE External switches	(2) Dell PowerConnect 7048 Firmware v4.2.1.3 Stacking module 10Gb uplink module (SFP+)	Ethernet switches for configs: TOR switch only with ISL LAG Blade IOM switch and TOR switch with 4-way stack Blade IOM switch and TOR switch with 3-way LAG Blade IOM switch and TOR switch with uplink stacks Blade IOM switch and TOR switch with ISL stacks
1GbE Storage	(2) Dell EqualLogic PS6100XV: (16) 146GB 15K SAS disks – vHN62 (2) Quad port 1GbE controllers Firmware: 5.2.2 R229536	Storage arrays for configs: Blade IOM switch only with ISL stack Blade IOM switch only with ISL LAG TOR switch only with ISL LAG Blade IOM switch and TOR switch with 4-way stack Blade IOM switch and TOR switch with 3-way LAG Blade IOM switch and TOR switch with uplink stacks Blade IOM switch and TOR switch with ISL stacks

Appendix B Vdbench parameters

Vdbench workloads were executed using the following parameters in the parameter file, where "N" is the number of iSCSI volumes under load.

Common parameters:

```
hd=default

hd=one,system=localhost
```

iSCSI volumes:

```
sd=sd3,host=*,lun=\\.\PhysicalDrive3,size=1m,threads=5

sd=sd4,host=*,lun=\\.\PhysicalDrive4,size=1m,threads=5

sd=sd5,host=*,lun=\\.\PhysicalDrive5,size=1m,threads=5

sd=sd6,host=*,lun=\\.\PhysicalDrive6,size=1m,threads=5
```

8KB 67% read, random I/O workload:

```
wd=wd1,sd=(sd3-sd6),xfersize=8k,rdpct=67
```

256KB read, sequential I/O workload:

```
wd=wd1,sd=(sd3-sd6),xfersize=262144,rdpct=100,seekpct=sequential
```

256KB write, sequential I/O workload:

```
wd=wd1,sd=(sd3-sd6),xfersize=262144,rdpct=0,seekpct=sequential
```

Runtime options:

```
rd=rd1,wd=wd1,iorate=max,elapsed=3600,interval=30
```

Additional resources

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

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

Referenced or recommended Dell publications:

- Dell EqualLogic Configuration Guide: http://en.community.dell.com/techcenter/storage/w/wiki/2639.equallogic-configuration-guide.aspx

For EqualLogic best practices white papers, reference architectures, and sizing guidelines for enterprise applications and SANs, refer to Storage Infrastructure and Solutions Team Publications at:

- http://dell.to/sM4hJT