



Dell EqualLogic Configuration Guide

Dell Storage Engineering

February 2014
Version 14.4

This white paper is for informational purposes only, and may contain typographical errors and technical inaccuracies. The content is provided as is, without express or implied warranties of any kind.

© 2014 Dell Inc. All rights reserved. Reproduction of this material in any manner whatsoever without the express written permission of Dell Inc. is strictly forbidden. For more information, contact Dell.

Dell, the DELL logo, and the DELL badge, PowerConnect™, Dell™ EqualLogic™, PowerEdge™ and PowerVault™ are trademarks of Dell Inc. Broadcom™ is a registered trademark of Broadcom Corporation. Cisco® is a registered trademarks of Cisco Systems, Inc. and/or its affiliates in the U.S. and other countries. Intel™ is a registered trademark of Intel Corporation in the U.S. and other countries. Microsoft™, Windows™, Windows Server™, and Active Directory™ are either trademarks or registered trademarks of Microsoft Corporation in the United States and/or other countries.



Revision history

Revision	Date	Changes
14.4	February 2014	Section 9.7 Vertical port failover behavior in PS6210 controllers New information in 8.1.2 on how to ensure a secure network environment New information for Controller type 15 in Table 5.
14.3	October 2013	Added Purpose, Policies and Limitations, General Statement, and Policies for qualifying systems Eliminated redundant information that is available in other documents and provided links to that documentation Added High Availability chapter
14.2	June 2013	Small text updates.
14.1	March 2013	Added section 13 Data Center Bridging Added volume limits for Synchronous Replication in Table 3 Modified section 8.1.2 and section 8.2 Changed all references in tables in section 7 to read PS41x0 and PS61x0 Updated B – EqualLogic Upgrade Paths Updated the Related Publications
13.4	December 2012	Updates: New section numbering Updated Preparing for Firmware Upgrade and Controller Failover Added note box to reflect 10Gb support only on PS4110 and PS6110 ports Updated for PS-M4110XS and PS65x0 Hybrid Arrays Added note box to reflect no support for Direct Attach Storage (DAS) New Sections: Appendix D: Upgrade Paths for EqualLogic PS Series Arrays
13.3	September 2012	New sections: Section 7 Capacity Planning.

All significant changes in the newest version of this document are highlighted using blue text.



Table of contents

1	Purpose.....	7
1.1	Dell statement of support.....	7
1.2	General statement	7
1.3	Audience.....	7
2	Policies and limitations.....	8
2.1	Policies for qualifying systems.....	8
3	Related publications.....	9
4	PS Series storage arrays.....	11
4.1	Array models.....	11
4.2	PS Series supported configuration limits.....	13
4.3	Controller types in all models prior to PS4100/PS6100.....	15
4.4	Controller types in PS4100/PS6100 models.....	16
4.5	Array model PS-M4110.....	17
4.5.1	Controller type in PS-M4110 model.....	17
4.5.2	Configuration options.....	17
4.5.3	Networking considerations and guidelines.....	18
5	Controller firmware.....	20
5.1	About member firmware.....	20
5.2	Firmware upgrade considerations.....	20
6	RAID policies.....	22
7	Capacity planning.....	23
7.1	RAID 6 drive layouts and total reported usable storage.....	23
7.2	RAID 10 drive layouts and total reported usable storage.....	26
7.3	RAID 50 drive layouts and total reported usable storage.....	28
7.4	RAID 5 drive layouts and total reported usable storage.....	30
7.5	Array RAID configurations and associated RAID sets.....	31
8	EqualLogic SAN design.....	32
8.1	General requirements.....	32
8.1.1	Implementation of standards.....	32
8.1.2	General requirements and recommendations.....	33
8.1.3	Quality of service (QoS).....	34



8.2	Ethernet switches and infrastructure	35
8.2.1	Connecting SAN switches in a Layer 2 network.....	36
8.2.2	Sizing inter-switch connections.....	39
8.2.3	Comparing inter-switch connection types.....	39
9	Building a high-availability SAN.....	41
9.1	Multi-path I/O.....	41
9.1.1	EqualLogic MPIO Requirements	41
9.1.2	EqualLogic MPIO General Recommendations.....	42
9.2	Redundant SAN configuration.....	42
9.3	Controller redundancy in all models.....	43
9.4	Basic controller failover behavior in all models.....	44
9.4.1	Basic controller failover for models prior to PS4100/PS6100	44
9.4.2	Controller failover behavior for: PS41x0/PS61x0	46
9.4.3	Failover behavior in the PS-M4110	47
9.5	Vertical port failover behavior in PS4100/PS6100 controllers.....	48
9.6	Vertical port failover behavior in PS4110/PS6110 controllers.....	51
9.7	Vertical port failover behavior in PS6210 controllers.....	54
10	Mixed speed environments - Integrating 1GbE and 10GbE SANs	56
10.1	Mixed speed SAN best practices	56
11	Blade server chassis integration.....	58
11.1	Designing a SAN using blade chassis I/O modules with arrays	59
12	FS Series NAS Configuration.....	60
12.1	FS7500 connection paths.....	60
12.2	FS7600/FS7610 connection paths.....	63
13	Data Center Bridging (DCB).....	66
13.1	DCB Overview	67
13.2	DCB requirements for EqualLogic	67
13.3	VLANs for iSCSI.....	68
A	Network ports and protocols.....	70
A.1	Required ports and protocols.....	70
A.2	Optional ports and protocols	70
B	Upgrade paths for EqualLogic PS Series arrays.....	72



Abstract

This configuration guide provides technical guidance for designing and implementing Dell EqualLogic PS Series storage solutions in iSCSI SAN environments.

This guide includes:

- Details on the capacity and capabilities of the EqualLogic PS Series family and the FS Series family of NAS Appliances
- iSCSI SAN design guidelines for optimizing redundancy, availability, and performance of EqualLogic storage arrays

With the Dell™ EqualLogic™ PS Series of 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 be able to leverage the advanced features provided by an EqualLogic array, a robust, standards-compliant iSCSI SAN infrastructure must be created. While almost any industry standard, enterprise class Ethernet switch can be used to connect hosts to the EqualLogic-based virtual storage volumes, some switches, network interface controllers (NICs), iSCSI Host Bus Adapters (HBAs), and iSCSI initiators can directly impact the overall performance of the SAN storage solution. The *Dell EqualLogic Configuration Guide* is provided as an aid to help storage administrators determine how best to build an iSCSI infrastructure for use within an EqualLogic SAN solution. This document focuses on network configuration, host integration, and other topics that help to ensure a smooth deployment with optimum PS Series SAN and FS Series NAS appliance performance.

Unless otherwise stated, recommendations in this document are applicable to both file and block environments.

This document should be used strictly as a guide in planning an EqualLogic storage solution. It should not be considered a statement of support for any specific configuration. Actual viability of any configuration will depend on the capabilities of the individual components (switches, network interface controllers, initiators, etc.) that make up the SAN infrastructure.



1 Purpose

The goal of this guide is to provide a single reference for technical information, links to other product and technical information, and recommended Dell EqualLogic SAN design methodologies.

This document is for informational purposes only and is offered As Is.

This document is not intended to be used as:

- A document for statement of support for any specific configuration
- Approval for specific configurations or reference architectures

1.1 Dell statement of support

The scope of products available for Ethernet networking is too broad to allow Dell to certify every device for use within an EqualLogic infrastructure. The goal is not to restrict which infrastructure components you can integrate with the EqualLogic SAN, but to provide a clear statement of how Dell supports devices that have not been tested or sold by Dell.

1.2 General statement

In general, Dell allows any infrastructure component to be used within a Dell EqualLogic SAN solution, assuming it meets the minimum standards as described in the EqualLogic Configuration Guide required to be used within a Dell EqualLogic SAN infrastructure. Support resolves customer issues with the SAN solution until it has been determined that the infrastructure component is identified as causing the issue. If the identified component is not included in a Dell support agreement or Dell does not have a predefined support relationship (Cooperative Support Agreement), you may be required to directly contact that component provider for further support. Policies and procedures for support of EqualLogic products are explained in the Dell support statement which can be found at the link in Policies and limitations section below.

1.3 Audience

This configuration guide is intended for storage administrators, SAN/NAS system designers, storage consultants, or anyone who is considering purchasing or has purchased EqualLogic PS Series storage or FS Series Array appliances for use in a production SAN. 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 Ethernet standards as defined by the Institute of Electrical and Electronic Engineers (IEEE) as well as all TCP/IP and iSCSI standards as defined by the Internet Engineering Task Force (IETF).



2 Policies and limitations

- This document is being provided for informational purposes only and may change at any time.
- This version supersedes and replaces all previous versions.
- The information included in this document is intended as a guide for planning the configuration of systems for EqualLogic infrastructure and networking environments.
- It is not intended to be the sole resource for system configurations.
- For more information or questions not found in this document, please see relevant product documentation, contact your Sales Representative, or contact your customer Service and/or technical support representative.

2.1 Policies for qualifying systems

- Policies and procedures for support of EqualLogic products are set forth in the Dell support statement, which can be found at the following link:
<http://en.community.dell.com/techcenter/storage/w/wiki/4089.equallogic-support-policy-statement-updated-05172013.aspx>
- A list of validated hardware and software components are provided in the *EqualLogic Compatibility Matrix* (ECM), which can be found at the following link:
<http://en.community.dell.com/techcenter/storage/w/wiki/2661.equallogic-compatibility-matrix-07262013.aspx>



3 Related publications

The following locations provide additional background and technical details supporting configuration of EqualLogic SANs.

- EqualLogic Product Documentation (requires support login)
<https://eqlsupport.dell.com/support/download.aspx>
To access the Administration guides and other product documentation, follow this [link](#) and follow these steps:
 1. Select **PS Series Firmware**.
 2. Select the current firmware version.
 3. Select the Download Page of Dell EqualLogic PS Series Firmware.
 4. Select your guide under **Documentation**.

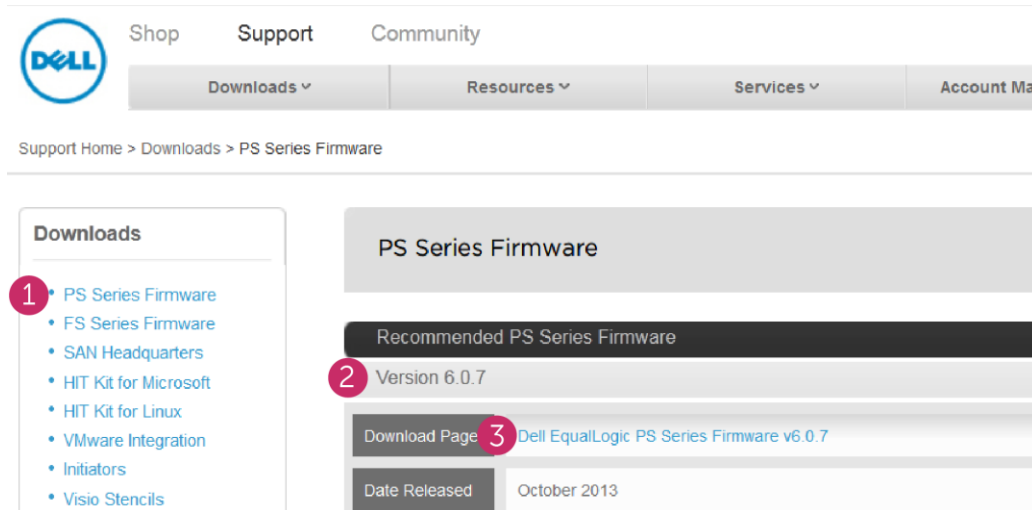


Figure 1 EqualLogic Support Download page

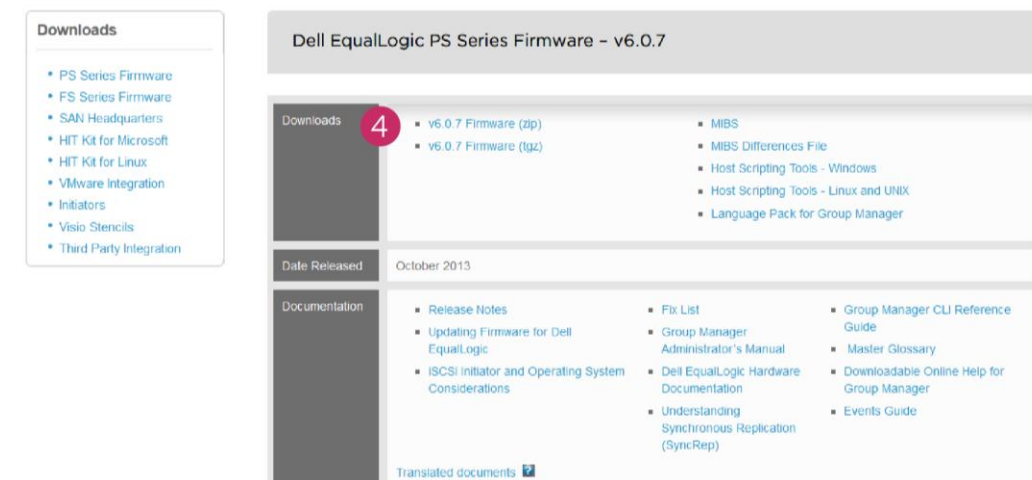


Figure 2 EqualLogic Support Documents



- *EqualLogic Compatibility Matrix*, including recommended switches and supported iSCSI initiators
<http://en.community.dell.com/dell-groups/dtcmedia/m/mediagallery/19856862/download.aspx>
- EqualLogic Technical Content
<http://en.community.dell.com/techcenter/storage/w/wiki/2660.equallogic-technical-content.aspx>
- Rapid EqualLogic Configuration Portal
<http://en.community.dell.com/techcenter/storage/w/wiki/3615.rapid-equallogic-configuration-portal-by-sis.aspx>
- Switch Configuration Guides
<http://en.community.dell.com/techcenter/storage/w/wiki/4250.switch-configuration-guides-by-sis.aspx>
- EqualLogic Technical Reports (requires support login)
https://eqlsupport.dell.com/support/tech_reports.aspx
- PS Series Hardware Documentation (requires support login)
<https://eqlsupport.dell.com/support/resources.aspx?id=2495>
- Dell EqualLogic Group Manager Online Help - This link provides browser-based access to the Group Manager firmware.
<http://psonlinehelp.equallogic.com/en/V6.0/groupmanager.htm>
- [VMware ESXi NIC Optimization and Best Practices with EqualLogic SAN](#)
- [Best Practices for Configuring DCB with Windows Server and EqualLogic Arrays](#)
- [Red Hat Enterprise Linux 6.3 NIC Optimization and Best Practices with EqualLogic SANs](#)



4 PS Series storage arrays

PS Series storage SANs provide a peer storage architecture comprised of one or more independent arrays. Each array contains its own controllers, cache, storage, and interface ports. Grouped together, they can create one or more single instance storage pools that are based on the IETF iSCSI standard. The capabilities of the PS Series solution are provided through a unique set of software and hardware capabilities that, when used cooperatively, can provide a full featured SAN solution. The following sections provide additional insight into specifications and functions of PS Series arrays.

4.1 Array models

The array models available prior to introduction of the PS4100/PS6100 family are shown in Table 1 below.

Table 1 Array Models prior to PS4100/PS6100

Array Model	Drive Type	Number of Drives
PS4000E	SATA	16
PS4000X, PS4000XV	SAS	16
PS5000E	SATA	16
PS5000X, PS5000XV	SAS	16
PS5500E	SATA	48
PS6000E	SATA	16
PS6000X, PS6000XV	SAS	16
PS6000S, PS6010S	SSD	16
PS6010E	SATA	16
PS6010X, PS6010XV	SAS	16
PS6000XVS, PS6010XVS	SAS / SSD	8 / 8
PS6500E	SATA	48
PS6510E	SATA	48
PS6510X	SAS	48
PS6500ES	SAS / SSD	41 SAS + 7 SSD
PS6510ES	SAS / SSD	41 SAS + 7 SSD

Starting with the introduction of the PS4100 and PS6100 family of arrays, configurations using 2.5" and 3.5" disks are available.



Table 2 PS4100/PS6100 array models

Array model	Drive type	Number of drives
PS4100E	3.5" SAS 7.2K RPM	12
PS4100X	2.5" SAS 10K RPM	24
PS4100XV	2.5" SAS 15K RPM	24
PS4100XV	3.5" SAS 15K RPM	12
PS6100E	3.5" SAS 7.2K RPM	24
PS6100X	2.5" SAS 10K RPM	24
PS6100XV	2.5" SAS 15K RPM	24
PS6100XV	3.5" SAS 15K RPM	24
PS6100S	SSD	12 or 24
PS6100XS	SSD + SAS 10K RPM	7 SSD + 17 SAS
PS4110E	3.5" SAS 7.2K RPM	12
PS4110X	2.5" SAS 10K RPM	24
PS4110XV	2.5" SAS 15K RPM	24
PS4110XV	3.5" SAS 15K RPM	12
PS6110E	3.5" SAS 7.2K RPM	24
PS6110X	2.5" SAS 10K RPM	24
PS6110XV	2.5" SAS 15K RPM	24
PS6110XV	3.5" SAS 15K RPM	24
PS6110S	SSD	12 or 24
PS6110XS	SSD + SAS 10K RPM	7 SSD + 17 SAS
PS-M4110E	7.2K RPM NL-SAS 2.5"	14
PS-M4110X	10K RPM SAS 2.5"	14
PS-M4110XV	15K RPM SAS 2.5"	14
PS-M4110XS	SSD + 10K RPM 2.5" SAS	5 SSD + 9 SAS



4.2 PS Series supported configuration limits

The Dell EqualLogic PS6xxx Series provides the full range of features and capabilities available with the EqualLogic PS Series storage products. The Dell EqualLogic PS4xxx Series provides a subset of features and capabilities, targeted at remote office and small to medium business storage deployments. The supported configuration limits for a PS Series group are provided in Table 3.

Table 3 Supported configuration limits

Configuration	PS4000/PS4100 and PS-M4110 groups only ^a	All other groups ^b
Volumes and replica sets per group	512 ^h	1024
Volume size ^c	15 TB	15 TB
Volumes enabled for replication (outbound) ^d	32	256
Snapshots and replicas per group	2048	10,000
Snapshots per volume	128	512
Replicas per volume	128	512
Volumes that have Synchronous Replication Enabled	4	32
Schedules (snapshot or replication) per volume or volume collection	16	16
Persistent Reservation registrants per volume	96	96
Replication partners per group	16	16
Replication partners per volume	1	1
Members per group	2	16 ^a
Members per pool	2	8
Pools per group	2	4
Volumes per collection	8	8
Collections per group (snapshot and replication)	100	100
Volume connections (each time an iSCSI initiator connects to a volume counts as a connection) ^{e,f}	512 per pool ^g 1024 per group with 2 pools	1024 per pool ^h 4096 per group with 4 pools
Access control records per volume and its snapshots	16	16




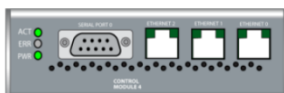
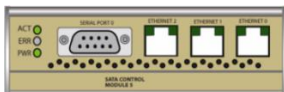
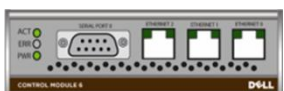




Configuration	PS4000/PS4100 and PS-M4110 groups only ^a	All other groups ^b
Simultaneous management sessions (any combination of GUI, telnet, or scripting sessions)	7	7
Thin provisioning ^l limits (minimum allocation)	10% of volume size	10% of volume size
Administrator accounts per group	100	100
<p>(a) A group can contain a maximum of two PS4000, PS4100, and/or PS-M4110 arrays.</p> <p>(b) Includes groups of mixed array types and all other group types except groups containing only PS40x0, PS41x0, and PS-M4110 arrays. When a group contains a mix of PS40x0, PS41x0, and PS-M4110 arrays and other array models, the higher limits prevail.</p> <p>(c) Practical maximum volume size is operating system-specific. A PS Series group can create and present volumes up to 15 TB.</p> <p>(d) Up to 16 simultaneous replication operations can take place.</p> <p>(e) To avoid service outages or possible loss of connectivity during failovers, Dell recommends increasing initiator timeouts for groups with more than 512 connections. See the iSCSI Initiator and Operating Systems Considerations document for more information.</p> <p>(f) Inbound replication connections count towards the total number of connections.</p> <p>(g) With firmware version 5.1, maximum number of volumes per group and maximum number of connections per pool increased from 256 to 512</p> <p>(h) Up to 1024 connections per pool are supported for all groups except groups containing only PS40x0, PS41x0, and PS-M4110 arrays. This provides increased scalability in virtualized environments. These connections may be distributed across a maximum of 512 volumes.</p> <p>(i) Volumes automatically created and managed by the NAS Service inside the NAS Reserve pool used by an FS Series appliance are fully allocated at creation and cannot be thin provisioned. Once those volumes are created (when the NAS Reserve is created they cannot be shrunk).</p>		



4.3 Controller types in all models prior to PS4100/PS6100

Array controllers can be identified and differentiated by the controller "type" designation. Each controller type will have a different colored label to help quickly identify the controller type. Table 4 lists each Dell EqualLogic controller along with some characteristics about each.

Table 4 Array controller types – All models prior to PS4100/PS6100





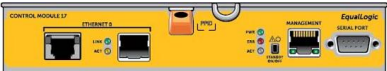
Controller type	Faceplate	Network interfaces	Storage type	Notes
Type 3 SAS Type 3 SATA		3 x 1GbaseT	SAS SATA	PS3000 – PS5000 1GB Cache Cannot mix Type 3 SAS with Type 3 SATA
Type 4		3 x 1GbaseT	SAS	PS3000 – PS5000 1GB Cache Cannot mix Type 3 SAS with Type 4 controller
Type 5		3 x 1GbaseT	SATA	PS3000 – PS5000 1GB Cache Cannot mix Type 3 SAS with Type 5 controller
Type 6		3 x 1GbaseT	SATA	PS5500 only 2GB Cache
Type 7		4 x 1GbaseT	SAS SATA SSD	PS6000 – PS6500 2GB Cache Cannot mix SAS and SATA drives in one array
Type 8		2 x 1GbaseT 1 x 10/100Mb mgt	SAS SATA	PS4000 only 2GB Cache Cannot mix SAS and SATA drives in same array
Type 9		2 x 1GbaseT 1 x 10/100Mb mgt	SAS SATA	2 nd generation PS4000 2GB Cache Cannot mix SAS and SATA drives in same array
Type 10		2 x 10GB SFP+ 1 x 10/100Mb mgt	SAS SATA SSD	10Gb Ethernet PS6010 – PS6510 2GB Cache



4.4 Controller types in PS4100/PS6100 models

The new controller types available in the PS4100 and PS6100 model arrays became available starting in August 2011. Table 5 lists each Dell EqualLogic controller along with some characteristics.

Table 5 PS4100/PS6100 controller types


Controller Type	Faceplate	Network Interfaces	Storage Type	Notes
Type 11		4 x 1GbaseT 1 x 10/100Mb mgt	SAS NL-SAS SSD	PS6100 Only 4GB Cache Supports vertical port failover Cache to Flash memory destaging
Type 12		2 x 1GbaseT 1 x 10/100Mb mgt	SAS NL-SAS	PS4100 Only 4GB Cache Supports vertical port failover Cache to Flash memory destaging
Type 14		1 x 10GbE SFP+ 1 x 1GbaseT 10/100Mb mgt	SAS NL-SAS SSD	PS6110 Only 4GB Cache Support for vertical port failover Cache to Flash memory destaging SFP+ OR 10GBaseT used at any given time
Type 15		2 x 10GbE SFP+ 2 x 1GbaseT 10/100Mb mgt	SAS NL-SAS SSD	16GB Memory per controller, 10GbaseT ports –can auto-negotiate to 1G. Cache to Flash (C2F) destaging. New battery backup for cache. Hot swappable, Active-hot stand by redundancy.
Type 17		1 x 10GbE SFP+ 1 x 1GbaseT 1 x 10/100Mb mgt	SAS NL-SAS	PS4110 Only 4GB Cache Supports vertical port failover Cache to Flash memory destaging SFP+ OR 10GBaseT used at any given time



4.5 Array model PS-M4110

4.5.1 Controller type in PS-M4110 model

The PS-M4110 controller is designed based on a modified version of the PS4100 Controller. Host and SAS cards are combined to form a single unit fixed I/O module, connecting to the M1000e chassis infrastructure.

Controller Type	Storage blade image	Network Interfaces	Storage type	Notes
Type 13	 A photograph of a storage blade server component, showing a front panel with multiple drive bays and a control panel.	2 x 10Gb Ports, (One Per Controller), Connected through the Backplane (No Label) Each of the two ports has an active link and a standby link to the fabric switches in the backplane of the M1000e chassis.	SAS NL-SAS	Dual, hot-pluggable 10GbE controllers 2GB of memory per controller, Cache to Flash Design 1 x dedicated management port – accessible via CMC 14x 2.5" 6Gb/s SAS disk drives

4.5.2 Configuration options

The PS-M4110 has four basic configuration options. It can be configured on Fabric A or Fabric B, and each Fabric Configuration can use a 10Gb KR switch or a 10Gb KR Pass-Thru Module (PTM). Figure 3 depicts a basic configuration using Dell Networking MXL switches, however any supported switches can be used in this configuration.



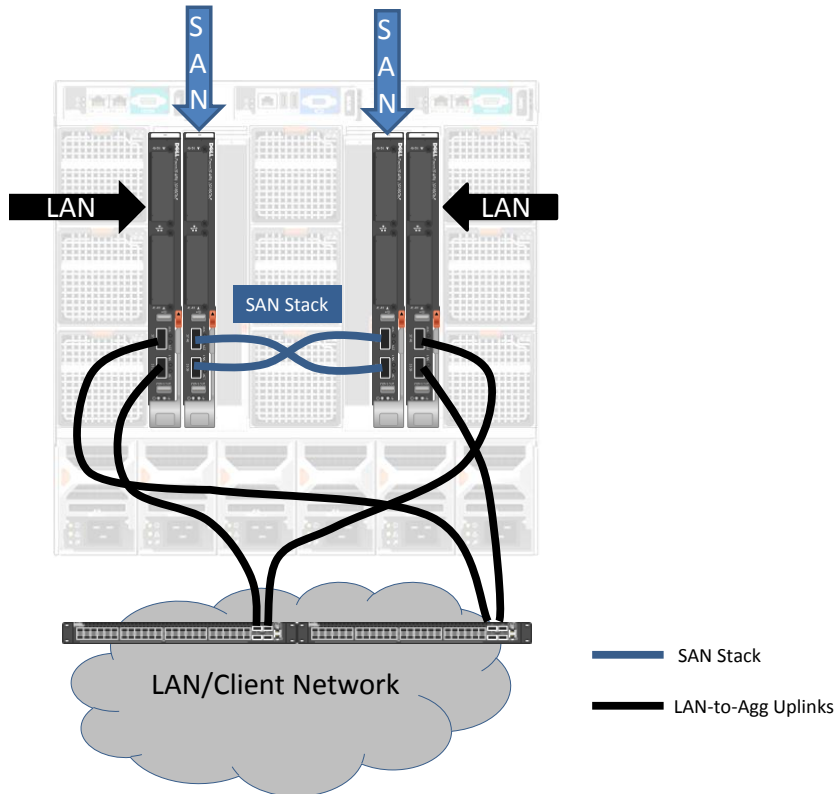


Figure 3 Basic PS-M4110 configuration for data center-in-a-box

4.5.3 Networking considerations and guidelines

Supported M-Series I/O modules

- 10G KR is the only supported I/O Module (IOM)
- Switches: The list of supported switches can be found in the ECM at the following link:
<http://en.community.dell.com/techcenter/storage/w/wiki/2661.equallogic-compatibility-matrix-07262013.aspx>

The following are basic networking recommendations for implementing the PS-M4110 Storage Blade.

- IOMs must be Interconnected
- External switches can be used to provide interconnection if rack mounted arrays are needed.
- Must use "Single Fabric" I/O module Placement; supported fabrics are "A" and "B" only
- Fabric A not supported on older M1000e systems (Mid-plane v1.0); Version should be physically marked on the chassis or you can check it using the CMC command. Upgrading is allowed but generally not recommended because of the required downtime and required professional services.
- It is recommended that PS-M4110 be placed into their own pool (not mixed with rack mount arrays)



For a more in-depth understanding of PS-M4110 Storage Blade, see the following white papers:

- *Dell EqualLogic PS-M4110 Blade Array Technical Article*
<http://en.community.dell.com/techcenter/storage/w/wiki/4134.dell-equallogic-ps-m4110-blade-array-technical-article.aspx>
- *Best Practices for DCB-Enabled Dell M-Series Blade Solution with EqualLogic PS-M4110*
http://en.community.dell.com/techcenter/extras/m/white_papers/20422608.aspx
- *Dell PowerEdge M1000e Blade and EqualLogic PS series SAN Design Best Practices*
<http://en.community.dell.com/dell-groups/dtcmedia/m/mediagallery/20198539/download.aspx>
- *PowerEdge M1000e Blade and EqualLogic PS Series SAN Design Best Practices Using Force10 Switches*
<http://en.community.dell.com/dell-groups/dtcmedia/m/mediagallery/20305343/download.aspx>



5 Controller firmware

5.1 About member firmware

Each control module in a group member must be running the same version of the PS Series firmware. Firmware is stored on a compact flash card or a microSD card on each control module.

Dell recommends the following:

- Always run the latest firmware to take advantage of new features and fixes.
- All group members must run the same firmware version. If you are adding a new array to a group, update the group to the latest firmware before adding the new member.
- Dell does not support groups running with mixed-firmware versions, except when you are in the process of upgrading the firmware of the entire group, and then only for as long as it takes to do the upgrade.
- You can upgrade member firmware to a higher version or downgrade member firmware to a lower version. In some firmware releases, firmware downgrades are not allowed. See the Dell EqualLogic PS Series Storage Arrays Release Notes for version-specific information about the firmware running on your group member(s).
- For instructions on how to update the firmware, see the document *Updating Firmware for Dell EqualLogic PS Series Storage Arrays*, available at the EqualLogic customer support site: <https://eqsupport.dell.com/support/download.aspx?id=1502>

Each EqualLogic PS Series array runs a core operating system in firmware that provides all of the PS Series features and functionality. The firmware version is defined using a version number and will be updated from time to time as new features are added or for general maintenance improvements.

The firmware version number takes the following form: "**X.Y.Z**":

- "**X**" is used to identify the "major" release number. This changes when there are very significant changes in the features and functionality.
- "**Y**" is the "minor" release number. This changes as new minor features or improvements in existing features are made for a given major release.
- "**Z**" denotes maintenance releases. This changes whenever a maintenance release level is released to resolve any identified technical or usability issues.

5.2 Firmware upgrade considerations

Before beginning a firmware upgrade process, review the following documentation. These documents are available from the Dell EqualLogic Support site at

<https://eqsupport.dell.com/support/download.aspx?id=6442455075&langtype=1033> (Support ID required for login access). These documents include:

- Release Notes and Fix List for the firmware version to which you are upgrading



- Release Notes for any FS Series appliances you are operating
- *Dell EqualLogic PS Series Storage Arrays iSCSI Initiator and Operating System Considerations*, available at: <http://en.community.dell.com/dell-groups/dtcmmedia/m/mediagallery/20371245/download.aspx>
- *Updating Firmware for Dell EqualLogic PS Series Storage Arrays*, available at: <https://eqlsupport.dell.com/support/download.aspx?id=1502>



6 RAID policies

Each array in an EqualLogic array group is configured with a single RAID policy. Arrays (or group members) within the same storage pool that have the same RAID policy cooperatively work to host volumes by distributing those volumes over multiple arrays. Two things that are defined by the RAID policy are:

- RAID level
- Hot-spare configuration

Each array implements a default RAID policy that includes a hot-spare.

The following report contains information about performance characteristics of the RAID policies under normal and failure conditions and help when choosing the RAID policy that meets the requirements for specific applications and environments. For a complete understanding of supported RAID policies for PS Series Group Members, see the report at the link below:

<http://en.community.dell.com/dell-groups/dtcmedia/m/mediagallery/19861480.aspx>



7 Capacity planning

7.1 RAID 6 drive layouts and total reported usable storage

RAID 6 (striped set with dual distributed parity) combines N disks in an arrangement where each stripe consists of N-2 disks capacity for data blocks and two disks capacity for parity blocks. Each parity block generates parity using a different view of the data blocks depending on the RAID 6 implementation. RAID 6 can tolerate up to two drive failures per RAID stripe set at the same time without data loss. RAID 6 is not recommended for workloads consisting mainly of random writes. Using a RAID 6 policy, Table 6 shows the drive layouts that are enforced based on the number of drives in each array/hot spare configuration, and the total usable storage available for each model.



Table 6 RAID 6 drive layouts and total storage available with hot spares (in GB)

Disk drives	Hot spare		No hot spare											
6	5 Data/Parity + 1 Hot-spare		6 Data/Parity											
7	6 Data/Parity + 1 Hot-spare		7 Data/Parity											
8	7 Data/Parity + 1 Hot-spare		8 Data/Parity											
12 ^(f)	11 Data/Parity + 1 Hot-spare		12 Data/Parity											
14	13 Data/ Parity + 1 Hot-spare		14 Data/Parity											
16	15 Data/ Parity + 1 Hot-spare		16 Data/Parity											
8 + 8 ^(c)	15 Data/Parity ^(d) + 1 Hot-spare		16 Data/Parity											
24 ^{(a)(f)}	23 Data/Parity + 1 Hot-spare		24 Data/Parity											
7 + 17 ^(e)	23 Data/Parity + 1 Hot-spare		24 Data/Parity											
48 ^(a)	47 Data/Parity + 1 Hot-spare		N/A ^(b)											
<p>(a) 24 and 48 drive arrays implement multiple RAID 6 sets within a single chassis. (b) 48 drive arrays cannot implement a no hot-spare RAID policy. (c) PS60x0 XVS with 8x 100GB SSD and 8x 450GB SAS. (d) One 450GB hot spare, 2x 450 GB parity and 2x 100GB SSD parity. In the event of a drive failure (SSD or SAS), the RAID set is reconstructed using the 450GB SAS hot spare. (e) One 600GB hot spare, 2x 600 GB parity and 2x 400 GB SSD parity. In the event of a drive failure (SSD or SAS), the RAID set is reconstructed using the 600 GB SAS hot spare. (f) PS4100/PS6100 models only</p>														
Total reported usable storage when using hot spares: All models prior to PS4100/PS6100														
Drive Qty / Size	50 ^(a)	100 ^(a)	74	146	250	300	400	450	500	600	750	1000	2000	3000
7 ^(b)	-	-	-	-	888	-	-	-	-	-	-	-	-	-
8 ^(c)	-	-	-	-	1111	-	-	-	-	-	-	-	-	-
14 ^(d)	-	-	660	-	2224	-	3566	-	4460	-	6695	-	-	-
16	577	1155	-	1696	2895	3475	4638	5220	5800	6964	8708	11613	23633	35717
8 + 8 ^(e)	-	-	-	-	-	-	-	2539	-	-	-	-	-	-
48	-	-	-	-	-	-	-	-	18062	21677	-	36136	69260	107151
Total reported usable storage when using hot spares: PS41x0/PS61x0														
Drive Qty / Size	146	200 ^(f)	300	400 ^(f)	500	600	900	1000	2000	3000				
6	-	-	-	-	1331	1597	-	2672	5457	8253				
12	1177	1658	2406	3328 ^(f)	4014	4812	7229	8038	16363	24719				
14	1434	-	2938	-	4906	5890, 4286 ^(h)	8841	9824	-	-				
24	2478	3512	5084	7034 ^(f)	8478	10178	15278	16947	34550	52202				
7 + 17	-	-	-	-	-	-	9318 ^(g)	-	-	-				
48	-	-	-	-	-	-	-	-	-	-				
<p>(a) Solid State Disk (b) Legacy PS50E (c) PS4000E and PS6000E only (d) Legacy PS70E, PS100E, PS100X, PS200E, PS300E and PS400E</p>					<p>(e) PS60x0 XVS with 8x100GB SSD and 8x450GB SAS (f) PS6100S (SSD) only (g) PS6100XS only (h) M4110XS only</p>									



Total reported usable storage when using hot spares: PS6210										
Drive Qty / Size	146	300	400	600	800	900	1200	2000	3000	4000
12	-	-	-	10182*	-	-	-	-	-	69857*
24	2478	5079 5084*	7035	10179	14076	15278 15280*	18614	34550	52204	69857

*SED



7.2 RAID 10 drive layouts and total reported usable storage

Using a RAID 10 policy, Table 7 shows the drive layouts that are enforced based on the number of drives in each array/hot spare configuration, and the total usable storage available for each model.

RAID 10 (mirrored sets in a striped set) combines two high performance RAID types: RAID 0 and RAID 1. A RAID 10 is created by first building a series of two disk RAID 1 mirrored sets, and then distributing data over those mirrors. RAID 10 implementations can tolerate one drive failure per mirror pair.

Table 7 RAID 10 drive layouts and total reported usable storage available with hot spares (in GB)

Disk drives	Hot spare		No hot spare											
6	4 Data (2 mirrors) + 2 Hot-spares		6 Data (3 mirrors)											
7	6 Data (3 mirrors) + 1 Hot-spare		N/A ^(a)											
8	6 Data (3 mirrors) + 2 Hot-spares		8 data (4 mirrors)											
12 ^(c)	10 Data (5 mirrors) + 2 Hot-spares		12 data (7 mirrors)											
14	12 Data (6 mirrors) + 2 Hot-spares		14 data (7 mirrors)											
16	14 Data (7 mirrors) + 2 Hot-spares		16 data (8 mirrors)											
24 ^(c)	22 Data (11 mirrors) + 2 Hot-spares		24 data (12 mirrors)											
48	46 Data (23 mirrors) + 2 Hot-spares		N/A ^(b)											
(a) Not supported. An even number of drives is required by RAID 10 (b) 48 drive arrays cannot implement a no hot-spare policy. (c) PS4100/PS6100 models only														
Total reported usable storage when using hot spares: All models prior to PS4100/PS6100														
Disk drive qty / Size	50 ^(a)	100 ^(a)	74	146	250	300	400	450	500	600	750	1000	2000	3000
7 ^(b)	-	-	-	-	666	-	-	-	-	-	-	-	-	-
8 ^(c)	-	-	-	-	666	-	-	-	-	-	-	-	-	-
14 ^(d)	-	-	395	-	1333	-	2135	-	2671	-	4013	-	-	-
16	309	621	-	913	1556	1867	2492	2805	3117	3744	4683	6247	12749	19230
48	-	-	-	-	-	-	-	-	10647	12779	-	21306	40840	63191
Total reported usable storage when using hot spares: PS41x0/PS61x0														
Disk drive qty / Size	146	200 ^(e)	300	400 ^(e)	500	600	900	1000	2000	3000				
6	-	-	-	-	887	1064	-	1781	3635	5498				
12	650	920	1331	1843	2222	2672	4014	4454	9103	13762				
14	781	-	1599	-	2670	3207	4816	5353	-	-				
24	1433	2027	2938	4065	4904	5888	8841	9820	19998	30221				
(a) Solid State Disk (b) Legacy PS50E (c) PS4000E and PS6000E only (d) Legacy PS70E, PS100E, PS100X, PS200E and PS300E (e) PS6100S (SSD) only														



Total reported usable storage when using hot spares: PS6210										
Drive Qty / Size	146	300	400	600	800	900	1200	2000	3000	4000
12	-	-	-	5890*	-	-	-	-	-	40443*
24	1434	2939 2938*	4065	5888	8144	8837 8841*	-	19999	30218	40443

*SED



7.3 RAID 50 drive layouts and total reported usable storage

Table 8 shows the drive layouts that are enforced when using a RAID 50 policy based on the number of drives in each array/hot spare configuration and the total usable storage available for each model.

RAID 50 (RAID 5 sets in a striped set) is created by first creating two or more RAID 5 sets and then striping data over those RAID5 sets. RAID 50 implementations can tolerate a single drive failure per RAID5 set.

Table 8 RAID 50 drive layouts and total reported usable storage available with hot spares (in GB)

Disk drives	Hot spare		No hot spare											
6	5 Data/Parity + 1 Hot-spare		N/A ^(a)											
7	6 Data/Parity + 1 Hot-spare		N/A ^(a)											
8	6 Data/Parity + 2 Hot-spares		8 data/parity											
12 ^(d)	10 Data/Parity + 2 Hot-spares		12 data/parity											
14	12 Data/Parity + 2 Hot-spares		14 data/parity											
16	14 Data/Parity + 2 Hot-spares		16 data/parity											
24 ^(d)	22 Data/Parity + 2 Hot-spares		24 data/parity											
48 ^(b)	46 Data/Parity + 2 Hot-spares		N/A ^(c)											
(a) RAID 50 requires an even number of disk drives. A 7 drive configuration without hot-spare would result in odd number of disk drives. (b) 48 drive arrays implement stripes across multiple RAID 5 sets within a single chassis. (c) 48 drive arrays cannot implement a no hot-spare policy. (d) PS4100/PS6100 models only														
Total reported usable storage when using hot spares: All models prior to PS4100/PS6100														
Disk Drive Qty / Size	50 ^(a)	100 ^(a)	74	146	250	300	400	450	500	600	750	1000	2000	3000
7 ^(b)	-	-	-	-	888	-	-	-	-	-	-	-	-	-
8 ^(c)	-	-	-	-	888	-	-	-	-	-	-	-	-	-
14 ^(d)	-	-	660	-	2224	-	3566	-	4460	-	6695	-	-	-
16	532	1066	-	1566	2671	3207	4280	4817	5353	6427	8037	10719	21819	32972
48	-	-	-	-	-	-	-	-	18062	21677	-	36136	69260	107151
Total reported usable storage when using hot spares: PS41x0/PS61x0														
Disk Drive Qty / Size	146	200 ^(e)	300	400 ^(e)	500	600	900	1000	2000	3000				
6	-	-	-	-	1781	2129	-	3563	7280	11008				
12	1044	-	2129	2949	3563	4280	6430	7137	14571	21975				
14	1304	-	2670	-	4459	5353	8036	8930	-	-				
24	2355	3328	4815	6666	8038	9646	14474	16087	32727	49455				
(a) Solid State Disk (b) Legacy PS50E (c) PS4000E and PS6000E only (d) Legacy PS70E, PS100E, PS100X, PS200E and PS300E (e) PS6100S (SSD) only														



Total reported usable storage when using hot spares: PS6210										
Drive Qty / Size	146	300	400	600	800	900	1200	2000	3000	4000
12	-	-	-	9646*	-	-	-	-	-	66180*
24	2355	4813 4816*	6666	9646	13335	14479 14475*	-	32727	49459	66180

*SED



7.4 RAID 5 drive layouts and total reported usable storage

RAID 5 (striped disks with distributed parity) will combine N disks in an arrangement where each stripe consists of N-1 disks that contain data blocks plus 1 disk that contains a parity block. For each stripe, the parity block will be placed on a different disk ensuring that the parity blocks are not located on a single disk in the RAID set. RAID 5 implementations can tolerate a single drive failure without data loss.

Table 9 shows the drive layouts that are enforced when using a RAID 5 policy based on the number of drives in each array/hot spare configuration, and the total usable storage available for each model.

Note: RAID 5 is not supported for PS6210.

Table 9 RAID 5 drive layouts and reported usable storage available with hot spares (in GB)

Disk drives		Hot spare		No hot spare											
6		5 Data/Parity + 1 Hot Spare		RAID Policy for RAID 5 without a hot spare is not supported											
7		6 Data/Parity + 1 Hot-spare													
8		7 Data/Parity + 1 Hot-spare													
12 ^(a)		11 Data/Parity + 1 Hot-spare													
14		13 Data/Parity + 1 Hot-spare													
16		15 Data/Parity + 1 Hot-spare													
24 ^{(a)(b)}		23 Data/Parity + 1 Hot-spare													
48 ^(b)		46 Data/Parity + 2 Hot-spares													
(a) PS4100/PS6100 models only (b) 24 and 48 drive arrays implement multiple RAID 5 sets within a chassis															
Total reported usable storage when using hot spares: All models prior to PS4100/PS6100															
Drive qty / Size	50 ^(a)	100 ^(a)	74	146	250	300	400	450	500	600	750	1000	2000	3000	
7 ^(b)	-	-	-	-	1111	-	-	-	-	-	-	-	-	-	
8 ^(c)	-	-	-	-	1333	-	-	-	-	-	-	-	-	-	
14 ^(d)	-	-	792	-	2671	-	4281	-	5354	-	-	-	-	-	
16	621	1244	-	1827	3118	3744	4995	5622	6247	7500	9378	12508	25456	38461	
48	-	-	-	-	-	-	-	-	19452	23345	-	38916	74580	115394	
Total reported usable storage when using hot spares: PS41x0/PS61x0															
Drive qty / Size	146	200 ^(e)	300	400 ^(e)	500	600	900	1000	2000	3000					
6	-	-	-	-	1781	2129	-	3563	7280	11008					
12	1300	1843	2672	3696	4454	5355	8038	8929	18186	27473					
14	1565	-	3206	-	5353	6426	9645	10719	-	-					
24	2744	3880	5620	7772	9379	11253	16855	18728	38184	57698					
(a) Solid State Disk (b) Legacy PS50E			(c) PS4000E and PS6000E only (d) Legacy PS70E, PS100E, PS100X, PS200E, and PS300E						(e) PS6100S (SSD) only						



7.5 Array RAID configurations and associated RAID sets

The tables show a logical drive layout when an array is initialized for the first time. The actual physical layout of drives can change and evolve due to maintenance and administrative actions. Spare drives can move as they are utilized to replace failed drives and newly added drives become the spares. It is not possible to determine which physical drives are associated with each RAID set. This information is dynamic and maintained by the EqualLogic firmware.

Table 10 shows the RAID set relationship for each RAID type in a 24-drive configuration.

Table 10 EqualLogic PS Series array RAID types and RAID set relationships

RAID policy	Spare disks	RAID set relationship	Best practice
RAID 6	1 Spare Disk	(10+2) (9+2)	Yes
RAID 10	2 Spare Disks	(6+6) (5+5)	Yes
RAID 50	2 Spare Disks	(5+1, 5+1) (4+1, 4+1)	For selected configurations
RAID 5	1 Spare Disk	(12+1) (9+1)	Not for business critical data

Table 11 shows the RAID set relationship for each RAID type in a 48-drive configuration.

Table 11 EqualLogic PS Series array RAID types and RAID set relationships

Raid policy	Spare disks	Raid set relationship	Best practice
RAID 6	1 Spare Disk	(12+2, 12+2, 12+2) (3+2)	Yes
RAID 10	2 Spare Disks	(7+7, 7+7, 7+7) (2+2)	Yes
RAID 50	2 Spare Disks	(6+1, 6+1, 6+1, 6+1, 6+1, 6+1) (3+1) ((5((3+1)	For selected configurations
RAID 5	2 Spare Disk	(12+1, 12+1, 12+1) (6+1)	Not for business critical data

Table 12 shows the RAID set relationship for each RAID type in a 16-drive configuration.

Table 12 EqualLogic PS Series array RAID types and RAID set relationships

Raid policy	Spare disks	Raid set relationship	Best practice
RAID 6	1 Spare Disk	(13+2)	Yes
RAID 10	2 Spare Disks	(7+7)	Yes
RAID 50	2 Spare Disks	(6+1, 6+1)	For selected configurations
RAID 5	1 Spare Disk	(14+1)	Not for business critical data



8 EqualLogic SAN design

An EqualLogic iSCSI SAN can be operated in any network that supports the industry standards and IP subnet design guidelines described in this section. Because of this flexibility, there are many network design and configuration choices that can affect SAN performance. The following sections provide details related to network design and configuration to support the use of an EqualLogic SAN.

Note: With the addition of FS Series NAS functionality, the EqualLogic product family now provides an iSCSI based unified file and block storage platform. An EqualLogic SAN can now be used to provide block level access (direct iSCSI access to PS Series arrays) or file system level access (via the FS Series appliance) using NFS or CIFS protocols and the Dell FluidFS scale-out file system.

Unless otherwise stated, SAN recommendations in this document are applicable to both file and block environments.

8.1 General requirements

8.1.1 Implementation of standards

EqualLogic SANs are based on industry standards. The following standards are required to support all host to target communications and member to member communications in an EqualLogic SAN:

- IETF Standards
 - IETF RFC1122 "Requirements for Internet Hosts – Communications Layers"
 - IETF RFC1123 "Requirements for Internet Hosts – Application and Support"
 - IETF RFC3270 "Internet Small Computer Systems Interface (iSCSI)"
- IEEE Standards
 - 802.1
 - 802.3

iSNS support

An Internet Storage Name Service¹ (iSNS) server can support discovery, management and configuration of group members by dynamically updating information about the iSCSI target names for group volumes. Once the IP address of an iSNS server is entered in an iSCSI initiator's configuration utility, the setting is persistent across initiator sessions. A PS Series group can be configured to register with up to three iSNS servers.

Note: Starting with Firmware V4.1.4, volume and snapshot identifiers are no longer automatically published to iSNS servers. This applies to new volumes and snapshots as well as volumes and snapshots that existed before the group was upgraded to V4.1.4.

¹ The Internet Storage Name Service (iSNS) specification: <http://tools.ietf.org/html/rfc4171>



8.1.2 General requirements and recommendations

For EqualLogic PS Series Arrays, the following general SAN design requirements apply:

- To ensure a secure network environment, Dell strongly recommends the following:
 - The network environment in which the group resides should be secure from network attacks such as packet sniffing or connection hijacking. This includes network hardware such as switches and routers. Firewalls and network isolation should be employed to protect resources. Network security can be achieved in a number of ways depending on the level of assurance needed and the types of network attacks to be expected. Security mechanisms include physical security, application level encryption (SSL/TLS and SSH), and/or IP level encryption (IPSec).
 - Network protocols and services used by the group should be secure and not subject to attacks. This includes RADIUS authentication, Network Time Protocol (NTP), syslog, and outbound SMTP traffic.
- For all members (arrays) in a given SAN Group all ports should be connected to the same subnet. This allows the arrays to communicate with each other as a group of peer members. The arrays must be in the same subnet as the EqualLogic Group IP address.

Note: Hosts can be in a different subnet as long as those hosts have layer 3 routing available to the subnet containing the arrays and the group's well known IP address.

- It is strongly recommended that a physically separated network be used for iSCSI traffic and that this network not be shared with other traffic types.

Note: If there is a requirement to share the same physical networking infrastructure with other non-iSCSI traffic then Data Center Bridging (DCB) is the recommended method for sharing networking resources.

- Rapid Spanning Tree Protocol must be enabled if the SAN infrastructure has more than two switches in a non-stacked configuration, and portfast must be enabled on all edge device ports (hosts, FS Series appliances and arrays).
- Port density requirements to support fully redundant configurations and maximum SAN throughput are as follows:
 - **PS4x00 family:** 2x 1GbE ports per controller = 4x 1GbE ports total
 - **PS4110 family:** 1x 10GbE port per controller = 2x 10GbE ports total
 - **PS5x00 family:** 3x 1GbE ports per controller = 6x 1GbE ports total
 - **PS6x00 family:** 4x 1GbE ports per controller = 8x 1GbE ports total
 - **PS6510 family:** 2x 10GbE ports per controller = 4x 10GbE ports total
 - **PS6110 family:** 1x 10GbE port per controller = 2x 10GbE ports total
 - **FS7500 NAS:** 12x1GbE ports per controller node (four client LAN and eight iSCSI SAN) + 1 100/1000Mb port per controller node for IPMI interconnection
 - **FS7600 NAS Appliance: 16x1GbE Ethernet ports per appliance** = 8 x 1Gb Ethernet ports per NAS controller for client connectivity and 8 x 1GbE Ethernet ports per NAS controller for SAN connectivity
 - **FS7610 NAS Appliance: 8x10GbE Ethernet ports per appliance** = 4 x 10GbE SFP+ Ethernet ports per NAS controller for client connectivity and 4 x 10GbE SFP+ Ethernet ports per NAS controller for SAN connectivity



- At least two iSCSI SAN ports per host (block level iSCSI access) are required for fully redundant SAN connectivity. Host ports can be 1GbE or 10GbE and the the host speed should match the array port speed.
- Quality of Service (QoS) based on what is traditionally designated as IEEE 802.1p is not currently supported for use with EqualLogic SANs. QoS and Class of Service designations must be disabled.
- All switches within the SAN must be interconnected such that there is always a path from any Ethernet port on one array to all other Ethernet ports on all other arrays in the group.
- All switches and host network controllers within the infrastructure must have (at a minimum, receive) flow control enabled for optimal performance.
- Any EqualLogic SAN group that is required to send or receive replication traffic to/from another SAN group must have an uninterrupted communications path (ie. "visibility") between each group.
- To prevent a switch failure from also disabling all paths between a host and its connected volumes, all ports from each controller need to be connected to at least two different physical switches. This also applies LAN side connections in FS Series NAS appliance configurations. (Refer to the connection diagrams in Section 12.1)
- For PS4100/PS6100 family arrays, split the vertical port pair connections between two switches to ensure 100% bandwidth capability is maintained in the event of a vertical port failover event.
- Do not block IPv6 traffic on the SAN internal switches when utilizing FS Appliances FS76x0 uses IPv6 for internal communication and setup only; not for user data. Client communication occurs over IPv4
- For SANs connecting to an FS Series appliance, all switches in the SAN must have jumbo frames enabled.
- Take advantage of your switch's VLAN capabilities. You may optionally create a VLAN dedicated to iSCSI traffic (even on dedicated switches). If necessary, create a second VLAN for management traffic. The actual VLAN configuration of your iSCSI SAN will be dictated by your SAN network design requirements and the features of the iSCSI SAN switches being used.
- Jumbo frames should be enabled for best performance. If you choose to use jumbo frames then all nodes in the SAN fabric must have jumbo frames enabled.
- For best performance and reliability, we recommend that all interconnection paths between non-stacking switches (LAGs) use a dynamic link aggregation protocol such as LACP

8.1.3 Quality of service (QoS)

QoS is described as either of the following:

- The ability to provide different priority levels to different applications, users, or data flows, or to guarantee a certain level of performance to a data flow.
- A network function implemented in some routers and switches that provides high priority for certain types of latency-sensitive traffic (for example, VoIP) and lower priority for other types of traffic (for example, web or http).

PS Series arrays are designed to provide I/O as fast as your network can support it. Therefore, using QoS with iSCSI traffic does not produce expected or desirable results on the SAN. Also, QoS rules can affect how well—or even whether—replication succeeds between PS Series groups. If you plan to use QoS, Dell



recommends that you use it only on WANs, where bandwidth is shared with other applications and the PS Series array uses it for time-insensitive replication traffic.

Dell recommends against using QoS on the SAN switches.

8.2 Ethernet switches and infrastructure

Any switch used in an EqualLogic SAN should meet the requirements listed in this section.

Note: For a complete list of tested and validated components see the EqualLogic Compatibility Matrix at the following link: [EqualLogic Compatibility Matrix](#)

An EqualLogic SAN consists of one or more hosts connected to one or more PS Series arrays through a switched Ethernet network.

Note: The minimum network configuration for a PS Series Array consists of a connection between Eth0 on each control module and a network switch. To increase performance and availability, configure multiple network interfaces on an array and connect them to multiple switches. EqualLogic does not support Direct Attached Storage (DAS) configurations.

To support a high performance Dell EqualLogic SAN, switches must meet the following general requirements:

- **Low latency:** Switches with relatively high latency may cause SAN throughput performance to degrade, and under high load conditions they could increase the risk of dropped connections.
- **Non-blocking backplane design:** SAN Switches should be able to provide the same amount of backplane bandwidth to support full duplex communication on ALL ports simultaneously.
- **Adequate buffer space per switch port:** In addition to supporting data transfers between the hosts and the SAN, Dell EqualLogic arrays also use the SAN to support inter-array communication and data load balancing. For this reason, the more buffer space per port that a switch can provide the better. Due to the multitude of buffer implementations used by switch vendors, Dell cannot provide definitive guidelines as to how much is enough. Port buffers should be designed such that data is not lost when traffic reaches extreme levels. Due to the clustered storage traffic patterns used by the EqualLogic SAN architecture, switches that support "cut-through" mode are not suitable for use in an EqualLogic SAN and may actually result in lower overall SAN performance.
- **Support for IEEE 802.3x flow control (passive and/or active) on ALL ports:** Switches and network interface controllers used in an EqualLogic SAN must be able to passively respond to any "pause" frames received.
- **Support for Jumbo Frames:** This is not a requirement, however most iSCSI SAN implementations should benefit from using jumbo frames and therefore it is recommend to enable Jumbo Frames. The actual impact on SAN throughput when using jumbo frames will depend on your workload's I/O characteristics.
- **Support for Rapid Spanning Tree protocol (IEEE 802.1w), or edgeport or Cisco "portfast" functionality if the SAN infrastructure will consist of more than two switches:** For SAN



infrastructures consisting of more than 2 non-stacking switches, R-STP must be enabled on all ports used for inter-switch trunks. All non-inter-switch trunk ports should be marked as “edge” ports or set to “portfast”.

- **Support for unicast storm control:** iSCSI in general, and Dell EqualLogic SANs in particular can send packets in a very “bursty” profile that many switches could misdiagnose as a virally induced packet storm. Since the SAN should be isolated from general Ethernet traffic, the possibility of actual viral packet storms occurring is non-existent. In an EqualLogic SAN, the switches must always pass Ethernet packets regardless of traffic patterns.
- **Support for Stacking:** A switch interconnection is required to link all switches in the SAN infrastructure together. Some Ethernet switches feature dedicated stacking interfaces. A good rule of thumb for a dedicated stacking link bandwidth would be a minimum 20 Gbps full-duplex.
- **Support for VLAN functionality.**

Note: It is recommended to use a physically separated network dedicated to iSCSI traffic that is not shared with other traffic. If sharing the same physical networking infrastructure is required, then use Data Center Bridging (DCB) for EqualLogic SAN.

- **Support for creating Link Aggregation Groups (LAG):** For non-stacking switches, the ability to bind multiple physical ports into a single logical link for use as an interconnection is required. The switch should support designating one or more ports for interconnection (via LAGs). The switch should support creation of LAGs consisting of at least eight 1Gbps ports or at least two 10Gbps ports. Variants of LAG (such as VLT, vPC, etc.) are also supported.

Note: For 1GbE SANs, using non-stacking switches to connect three or more EqualLogic arrays into a single group may negatively impact SAN I/O throughput performance.

8.2.1 Connecting SAN switches in a Layer 2 network

EqualLogic storage requires that the SAN be configured as a single Layer 2 network. Layer 2 refers to the data link layer in the OSI model (http://en.wikipedia.org/wiki/OSI_model). When more than one SAN switch is required, each switch connected to the array group members will be in the same subnet. These switches must be interconnected to provide a single switched Ethernet fabric. Figure 4 shows the two common methods for interconnecting switches, using either stacking switches or non-stacking switches.



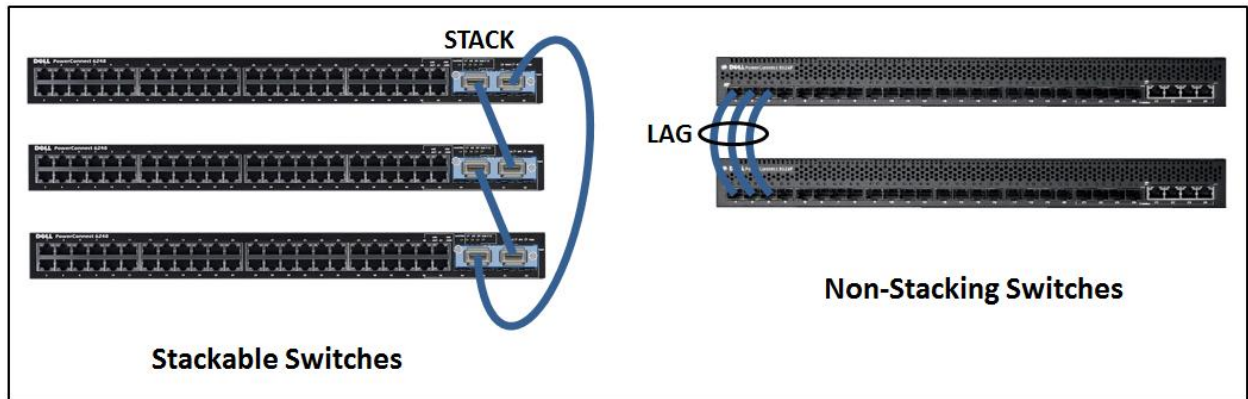


Figure 4 Switch Interconnects

8.2.1.1 Stacking switches

Stacking switches provides a simple method for creating a switch interconnection within a Layer 2 network infrastructure. Stacking is typically accomplished using a vendor proprietary, high-bandwidth, low-latency interconnect that allows two or more switches to be connected in such a way that each switch becomes part of a larger, virtual switch. Most stackable switches will provide a set of dedicated stacking ports or Installation of an optional stacking module may be required. Some switches also provide virtual stacking modes using standard Ethernet ports.

8.2.1.2 Non-stacking switches

Non-stacking switches do not have a dedicated switch interconnect. In this case to create an interconnect between the switches you must utilize one or more ports on each switch to create a LAG. This type of switch interconnection should use link aggregation functions (if provided by the switch) to aggregate multiple Ethernet ports into a single, logical high bandwidth interconnect. There are several options depending on the vendor of the switch. Table 13 describes the most common options.

Note: For best performance and reliability, we recommend that all interconnection paths between non-stacking switches (LAGs) use a dynamic link aggregation protocol such as LACP.

Table 13 Link aggregation types

Link aggregation type	Notes
Static	Static link aggregation defines a set of links that provide a point to point connection between two switches. These links may or may not provide failover redundancy or traffic load management.
LACP	Link Aggregation Control Protocol is based on IEEE 802.3ad or IEEE 802.1AX. LACP is a dynamic LAG technology that automatically adjusts to the appearance or disappearance of links within the defined LACP group.
PAgP	Port Aggregation Protocol (PAgP) is a Cisco Systems® proprietary networking protocol, which is used for the automated, logical aggregation of Ethernet switch ports, known as an etherchannel. This means it can only be used between Cisco switches and/or switches from licensed vendors.
Vendor Proprietary	Several switch vendors may provide additional link aggregation options that are completely proprietary or may be extensions to one of the two previously mentioned solutions. In most cases, this type of link aggregation solution is designed to reduce or eliminate the need – and the overhead – of the Spanning Tree Protocol that must be used in the two previous options. If available, these proprietary options should be considered. They may be very useful in allowing the network administrator to create a more efficient multi-switch layer 2 network infrastructure for a SAN. Be aware that these proprietary solutions must support the transmission of IEEE 802.1x flow control and jumbo frames (if used) to properly support an EqualLogic SAN. Examples of this type of interconnect include Dell Networking or Force10 Virtual Link Trunking (VLT) and Cisco Virtual Port Channel (VPC).

8.2.1.3 Using a LAG to Connect Stacked Switches

In some situations it may become necessary to expand the EqualLogic SAN by using more than one single switch stack. For example, you can link multiple stacks by creating a multi-link LACP based LAG between the switch stacks. A simplified stack plus LAG switch configuration is illustrated in Figure 5. You should consider the following recommendations when designing this type of SAN:

- If possible, use 10Gb or higher connections for all links between each stack.
- Distribute the links across multiple switches within each stack.
- Use LACP or another type of dynamic link aggregation protocol.
- Perform tests to determine the best hashing algorithm to use within the LAG (port channel).

Note: A multi-stack SAN infrastructure as described in this section may provide added reliability to the SAN environment. But, it may also introduce additional latency and the potential for lower throughput. The SAN designer will have to carefully consider the performance and reliability implications.



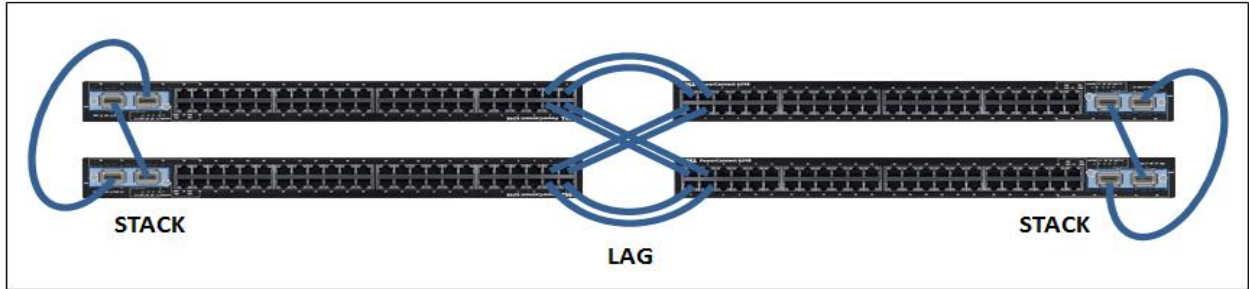


Figure 5 Using a LAG to interconnect switch stacks

8.2.2 Sizing inter-switch connections

Use the guidelines in Table 14 as a starting point for estimating inter-switch connection sizes.

Table 14 Switch Interconnect Design Guidelines

Connection speeds	Interconnection guidelines
1GbE switches attached to 1GbE array controllers	<p>1-5 arrays: 1Gb of inter-switch bandwidth per active array controller port (up to the aggregated maximum bandwidth of the inter switch links).</p> <p>6+ arrays: Use 1-5 array rule, then add 1Gb of additional bandwidth for each array added</p>
10GbE switches attached to 10GbE array controllers	<p>PS6010/PS6510 or PS4110/6110 (Random Small Block Workloads):</p> <p>1-5 arrays: 20 – 30Gb of inter-switch bandwidth between each switch</p> <p>6+ arrays: At least 6Gb of inter-switch bandwidth per array between each switch</p> <p>PS4110/PS6110 (Large Block Sequential Workloads):</p> <p>1-5 arrays: 10Gb of bandwidth per active array controller port (up to the aggregated maximum bandwidth of the IST).</p> <p>6+ arrays: Use 1-5 array rule, then add 10Gb of additional bandwidth for each array added</p>

8.2.3 Comparing inter-switch connection types

Table 15 provides details and recommendations for selecting interconnect options.

Table 15 Stacking versus LAG

Interconnect type	Primary purpose	Analysis
Stacking	Create a larger, logical switch within an isolated physical location.	<p><u>Advantages:</u></p> <ul style="list-style-type: none"> • Easier to manage multiple switches as single switch • Possibly higher bandwidth than using link aggregation and Ethernet • Not limited by Ethernet standards • Firmware upgrades will reset the entire stack <p><u>Concerns:</u></p> <ul style="list-style-type: none"> • Proprietary, cannot be used to interconnect switches from different vendors • Increases cost of switch • Stack bandwidth overload risk <p><u>Recommendation:</u></p> <ul style="list-style-type: none"> • Simplest way to scale a storage network in a single location on a single subnet • May provide lower latency and higher bandwidth than a LAG • Understand the stacking technology and limit stack size in accordance with total throughput requirements and connection counts.
LAG	Create a data path between switches in one location or subnet with those in another location or subnet	<p><u>Advantages:</u></p> <ul style="list-style-type: none"> • Leverages Ethernet standard extensions • Can be used to interconnect switches from different vendors • Can use Link Aggregation Protocols (LACP/EtherChannel) to pool multiple 1GbE or 10GbE links into a single logical link providing bandwidth and redundancy • Can upgrade switch firmware on each switch independently <p><u>Concerns:</u></p> <ul style="list-style-type: none"> • Some solutions limited to eight port link aggregation group • Spanning Tree Protocol must be used if more than two switches are used causing some links to be "blocked" and reducing bandwidth availability • LAG bandwidth overload risk <p><u>Recommendation:</u></p> <ul style="list-style-type: none"> • Use when stacking is not available • Use when connecting to aggregation/core switching infrastructure • Use when switches are from different vendors • Use when highest availability is required



9 Building a high-availability SAN

Designing a redundant SAN requires the availability of redundant NICs or HBAs on each server. A redundant NIC configuration on the server requires at least two NICs.

The information provided here does not address all of the possible variations in a SAN.

All information is presented using a set of basic reference designs that make the following assumptions:

- The SAN network is physically isolated from all other network traffic
- The examples are based on best practice design principles.
- Unless otherwise stated, all reference designs will provide end-to-end host to volume redundant paths
- A minimal number of switches will be illustrated to allow for the design concept to be understood. Actual implementations will vary depending on your network infrastructure requirements.
- If sharing physical switches with other non-SAN traffic, we assume all switches are VLAN capable.

Note: The PS4110 and PS6110 PS Series Arrays do not support dual speed connectivity. They are not designed to support the auto-negotiate feature. The connectivity must be 10Gb ONLY - This is the only supported configuration.

9.1 Multi-path I/O

There are generally two types of multi-path access methods for communicating from a host to an external device. For general networking communications, the preferred method of redundant connections is the teaming of multiple NICs into a single, virtual network connection entity. For storage, the preferred method of redundant connection is the use of Multi-Path IO (MPIO). Though some storage solution can and do support either method for iSCSI connectivity, EqualLogic requires the use of MPIO to enable multiple NIC/HBA connections to be utilized for access to an EqualLogic SAN.

The MPIO ToolKits, for available Operating Systems, can be found at the link below: (for example, the Host Integration ToolKit for Microsoft Windows and Linux or the Multipath Extension Module for VMware ESXi).

<https://eqsupport.dell.com/support/download.aspx>

Click on the Host Integration ToolKit for your Operating System.

Note: EqualLogic does not support NIC teaming on interfaces connected to the SAN.

9.1.1 EqualLogic MPIO Requirements

The following host port requirements must be met to use MPIO with EqualLogic SANs:

- At least two Ethernet ports are required on each host.
- The host operating system must have a supported MPIO driver or service available.
- The ports used for MPIO cannot be "teamed" to other ports.
- The ports must be the same speed



- The ports must be assigned IP addresses on the same subnet

9.1.2 EqualLogic MPIO General Recommendations

Follow this general set of guidelines for configuring MPIO on a host:

- Configure volume access controls to use standard iSCSI IQN names
- For a more secure configuration you can use the IQN name plus the CHAP security ID.
- On each array enable at least two ports for host connectivity.
- Install the Dell provided MPIO extension features, if available for the host operating system
- For Microsoft Windows, install the Device Specific Module (DSM) found in the Host Integration Toolkit for Windows.
- For VMware vSphere ESXi, install the EqualLogic Multipathing Extension Module.
- For Linux (supported versions), use the MPIO functionality provided by the Host Integration Toolkit for Linux.
- For other operating systems, use the native MPIO functionality.

9.2 Redundant SAN configuration

In a redundant iSCSI SAN, each component of the SAN infrastructure has a redundant connection or path. The following figures show example connection paths necessary to create a Redundant SAN.

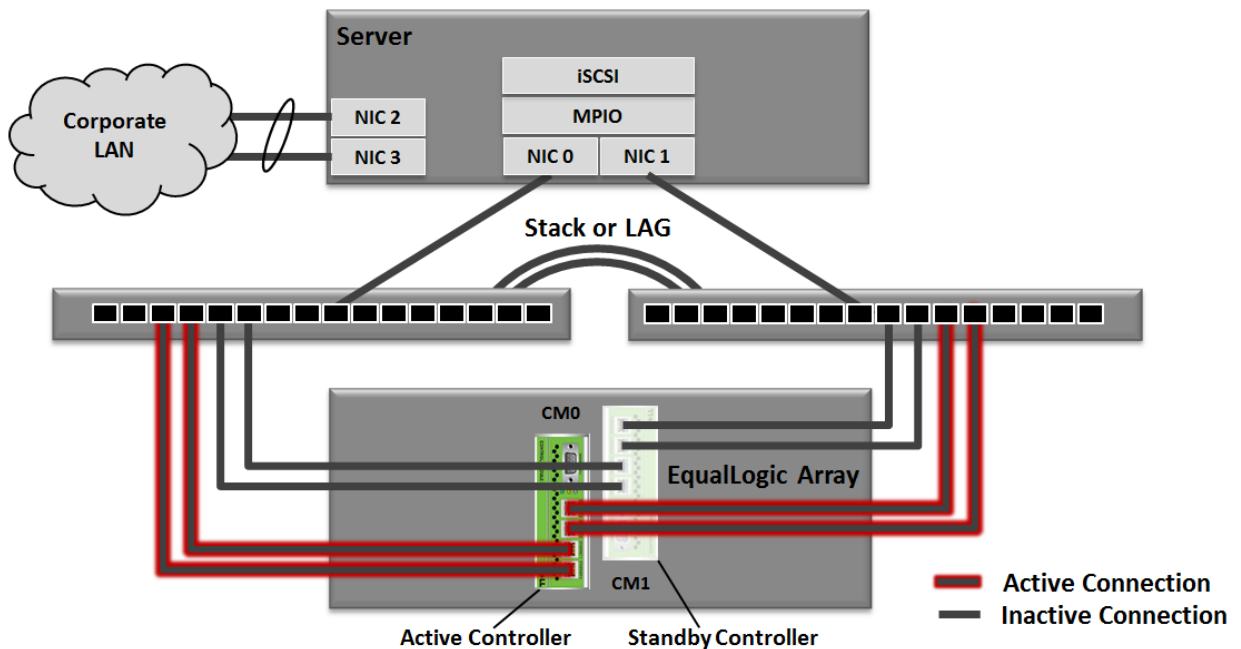


Figure 6 Redundant SAN Connection Paths: PS3000 to PS6500 Family Arrays

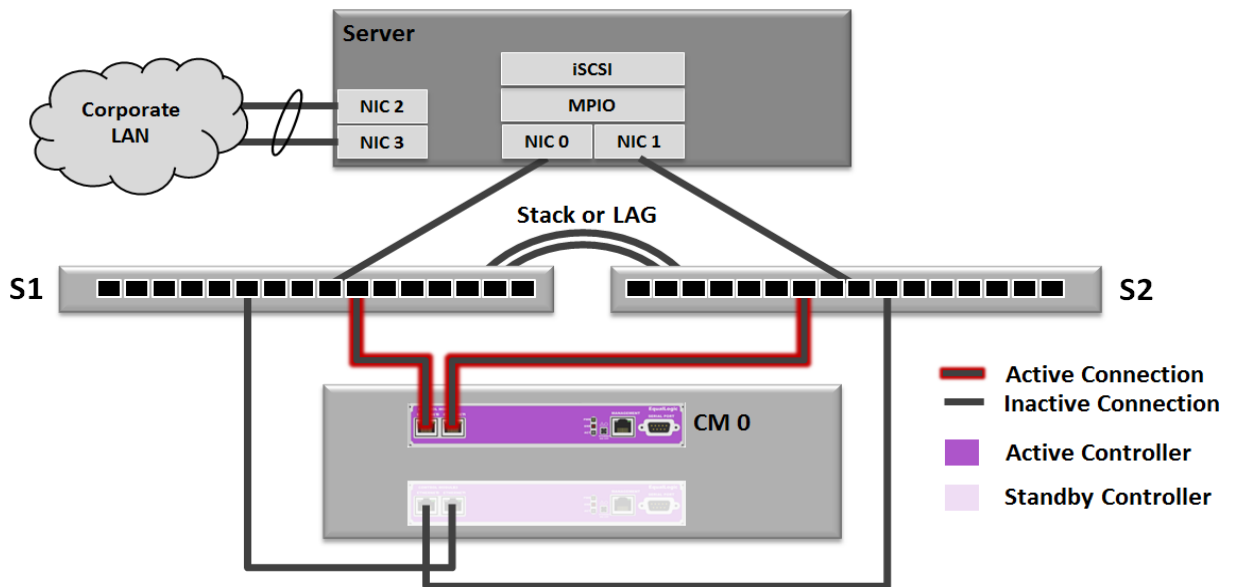


Figure 7 Redundant SAN Connection Paths: PS4100

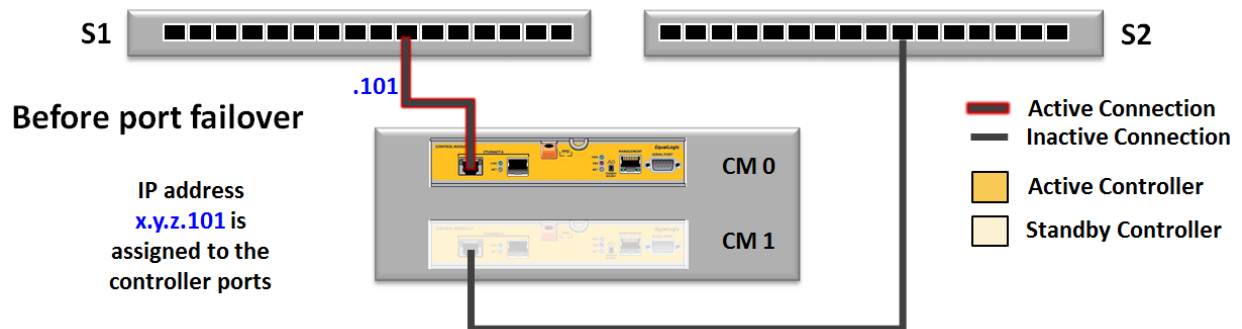


Figure 8 Redundant SAN Connection Paths: PS4110/PS6110

9.3 Controller redundancy in all models

Each array can be configured with either a single controller, or dual redundant controllers. The single controller configuration will provide the same level of I/O performance as a dual controller configuration. The dual controller configuration provides for redundancy. Dual controllers provide redundancy in the case of a controller failure. Host connections will fail over to the redundant controller.

The Active Controller is the controller which is processing all disk and network I/O operations for the array. A second controller in dual controller configurations will always be in a passive operating mode. In this mode, the secondary controller will exhibit the following characteristics:



- Each of its Ethernet ports are electrically inactive (active lights are off), unless a vertical port failover has occurred.
- The passive controller cache mirrors the cache of the active controller.

9.4 Basic controller failover behavior in all models

To support redundant controller fail over, each Ethernet port on the active controller that is connected to the SAN must have its corresponding port on the passive controller also connected to the same SAN network. In the event of a controller failure, the passive controller will immediately activate and continue to process all data requests to the array. The following changes occur during fail over:

- The backup controller automatically enables each of the corresponding Ethernet ports that were enabled on the failed primary controller
- The IP addresses that were assigned to each of the failed controller Ethernet ports are reassigned to the corresponding ports on the second controller.

9.4.1 Basic controller failover for models prior to PS4100/PS6100

A link failure on one or more of the Ethernet ports on the active controller does not constitute a controller failure. For this reason, it is important to ensure that ports from each controller are connected to at least two different switches. This will prevent a switch failure from also disabling all paths between a host and its connected volumes.

It is also critical that port connections for both controllers are configured so that the corresponding ports on each controller are connected to the SAN. If port 0 on the active controller is the only port connected to the SAN, then you must use port 0 on the passive controller for the other connection to the SAN. This is illustrated in the partially connected scenario shown in Figure 9.



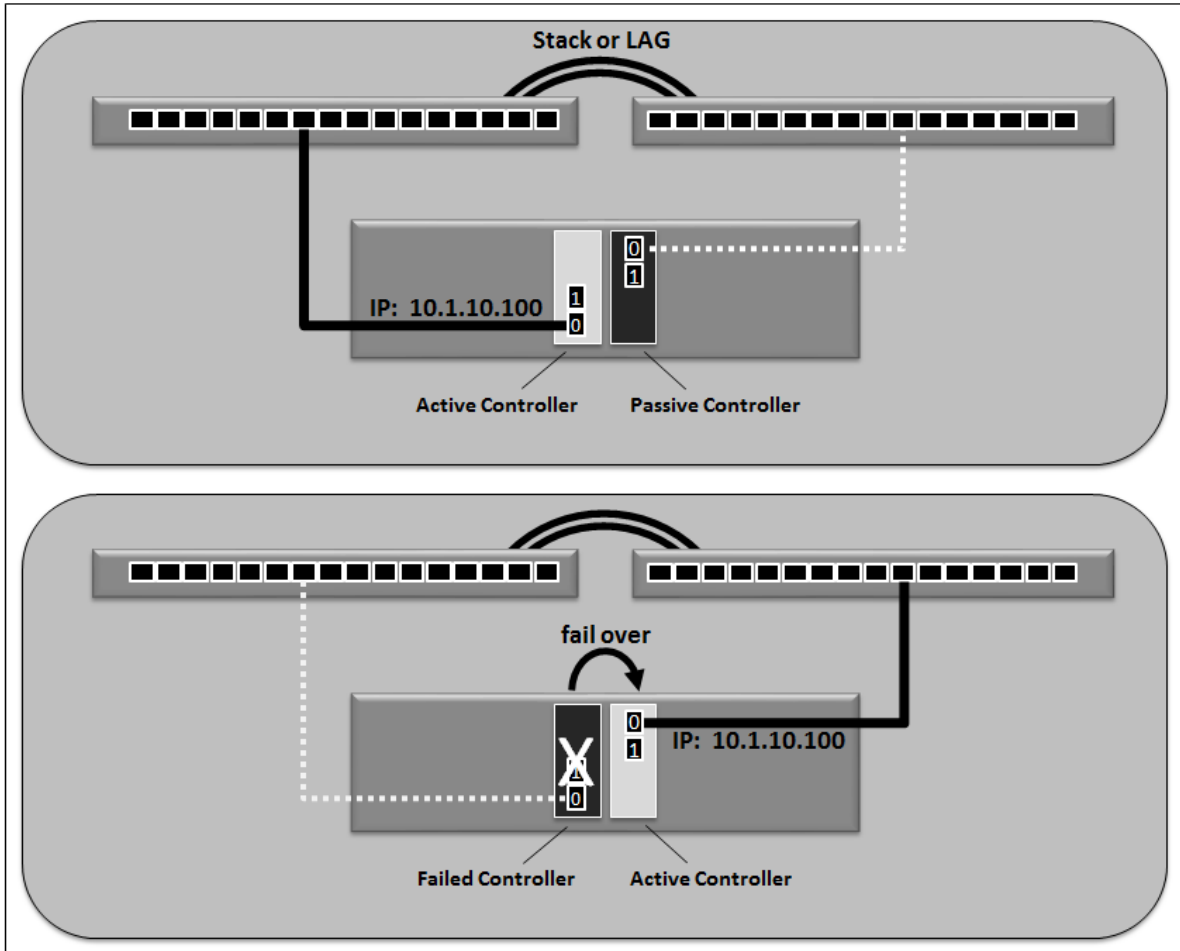


Figure 9 Partially connected controller failover

Note how IP addresses are reassigned to their corresponding ports during the failover processes shown in Figure 9 and Figure 10.

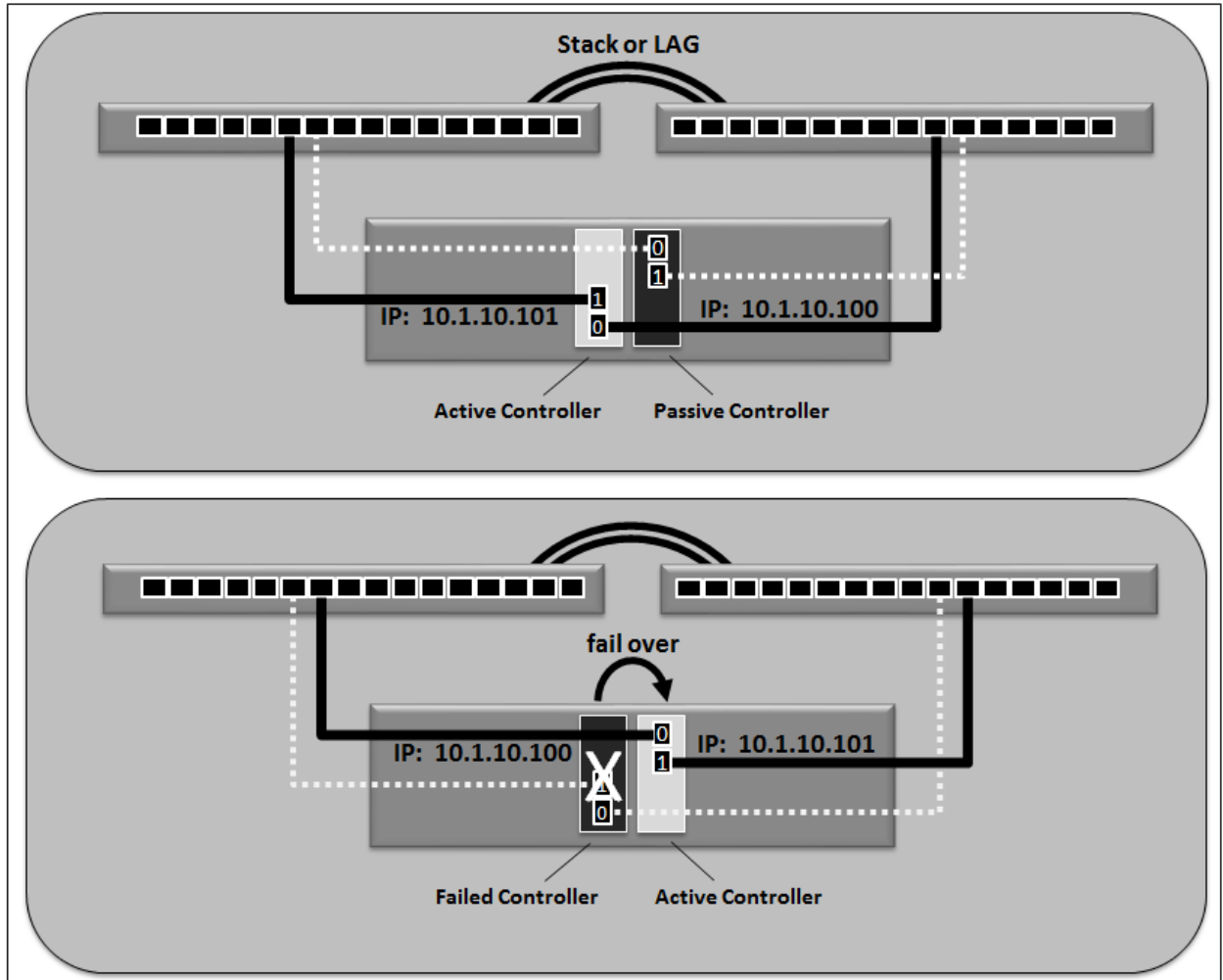


Figure 10 Fully connected controller failover

9.4.2 Controller failover behavior for: PS41x0/PS61x0

In the event of a controller failure the following changes occur:

- The passive controller immediately activates and continues to process all data requests to the array.
- Vertical port failover ensures that IP addresses assigned to each of the failed controller Ethernet ports apply to the corresponding ports on the second controller.

As stated in Section 9.4.1, controller models prior to the PS4100/PS6100 required having cabled connections present on corresponding controller ports for controller failover to work correctly. The vertical port failover feature in PS41x0/PS61x0 controllers eliminates this requirement, enabling partially cabled configurations to support controller failover too.

It is important to ensure that ports from each controller are connected to at least two different switches. This will prevent a switch failure from also disabling all paths between a host and its connected volumes.

We illustrate controller failover behavior for the PS4100 family controller in Figure 11. Controller failover behavior for the PS6100 (4 port) controller family is identical.

Note: To prevent a switch failure from also disabling all paths between a host and its connected volumes, you should make sure that ports from each controller are connected to at least two different switches. You should also split vertical port pair connections between two switches to ensure 100% bandwidth capability is maintained in the event of a vertical port failover event. Both of these guidelines are illustrated in Figure 11.

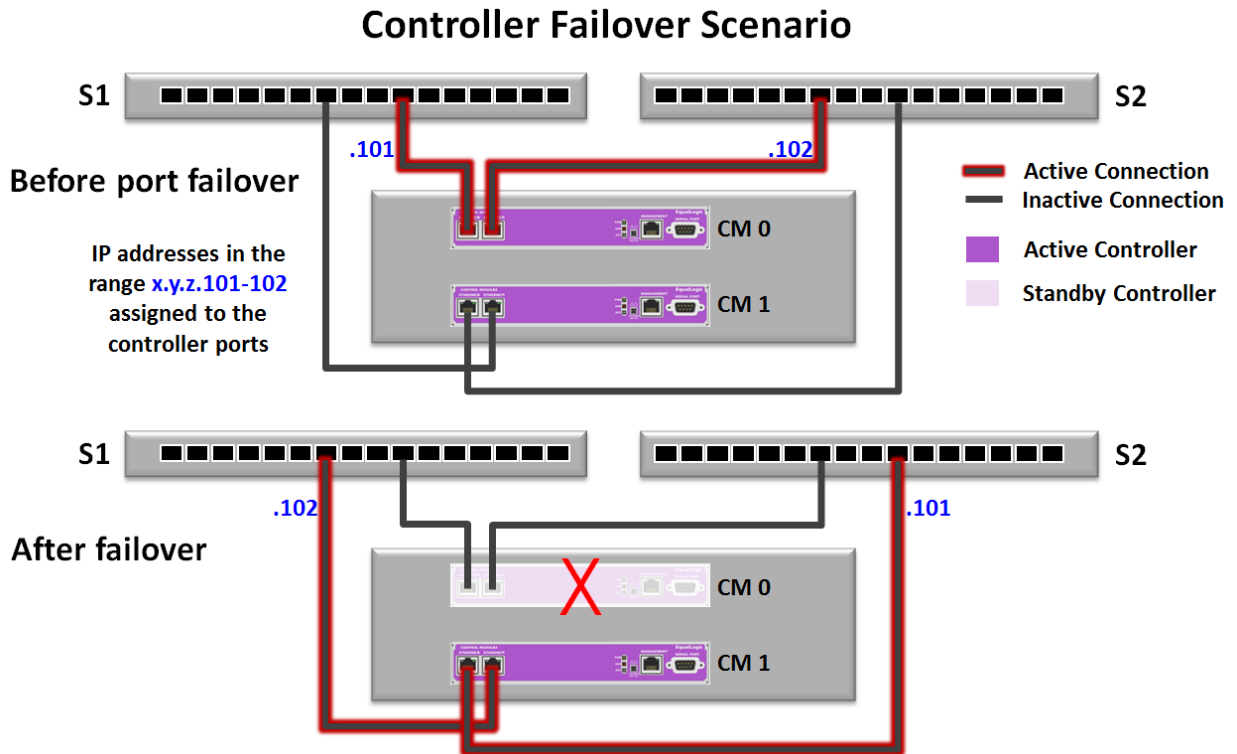


Figure 11 Controller failover process and optimal connection paths

9.4.3 Failover behavior in the PS-M4110

Each PS-M4110 array is configured with dual controllers, active and passive connections, and redundant fabrics to provide redundancy and prevent volume connections between hosts and SAN from being dropped in the event of a failure.

There are three failure scenarios that are protected by the array:

- In the event of a link failure, the active link of the active port goes down, and the standby link for the same port takes over.
- In the event of a switch failure, both active and passive ports automatically link to the redundant fabric switch.
- If the active controller fails, the passive controller takes over.

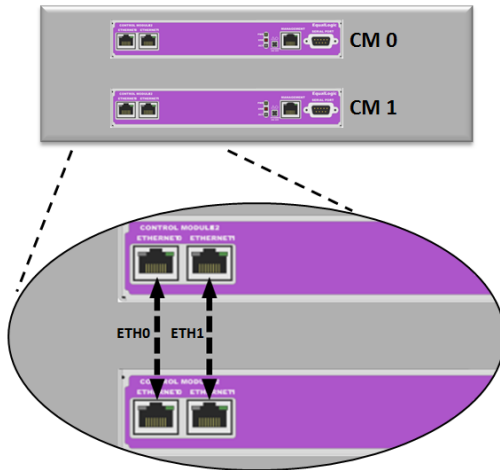


9.5 Vertical port failover behavior in PS4100/PS6100 controllers

In PS Series controllers prior to PS4100/6100 families, a link failure or a switch failure was not recognized as a failure mode by the controller. Thus a failure of a link or an entire switch would reduce bandwidth available from the array. Referring to Figure 12 or Figure 13, assume that CM0 is the active controller. In vertical port failover, if CM0 senses a link drop on the local ETH0 port connection path, it will automatically begin using the ETH0 port on the backup controller (CM1) without incurring a controller module failover.

Vertical port failover is bi-directional. If CM1 is the active controller then vertical port failover will occur from CM1 ports to CM0 ports if necessary.

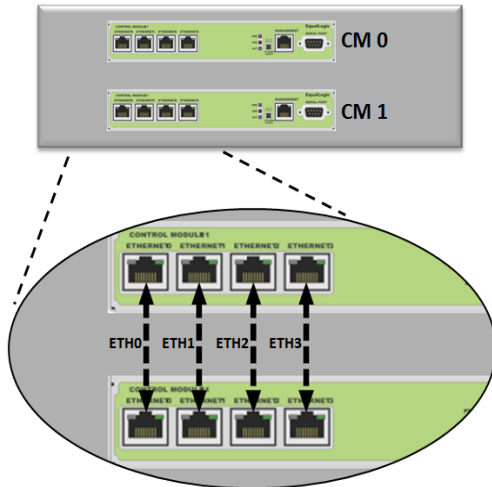




PS4100 Vertical Port Failover

- Both controller ports support vertical failover

Figure 12 PS4100 vertical port failover



PS6100 Vertical Port Failover

- All controller ports support vertical failover
- A maximum of two ports can failover at one time

Figure 13 PS6100 vertical port failover

With PS4100/PS6100 family controllers, vertical port failover can ensure continuous full bandwidth is available from the array even if you have a link or switch failure. This is accomplished by combining corresponding physical ports in each controller (vertical pairs) into a single logical port from the point of view of the active controller. In a fully redundant SAN configuration, you must configure the connections as shown in Section 9.2.

In a redundant switch SAN configuration, to optimize the system response in the event you have a vertical port failover you must split the vertical port pair connections between both SAN switches. The connection paths illustrated in Figure 14 and Figure 15 show how to alternate the port connection paths between the two controllers. Also note how IP addresses are assigned to vertical port pairs.



PS4100: Vertical Port Failover Scenario

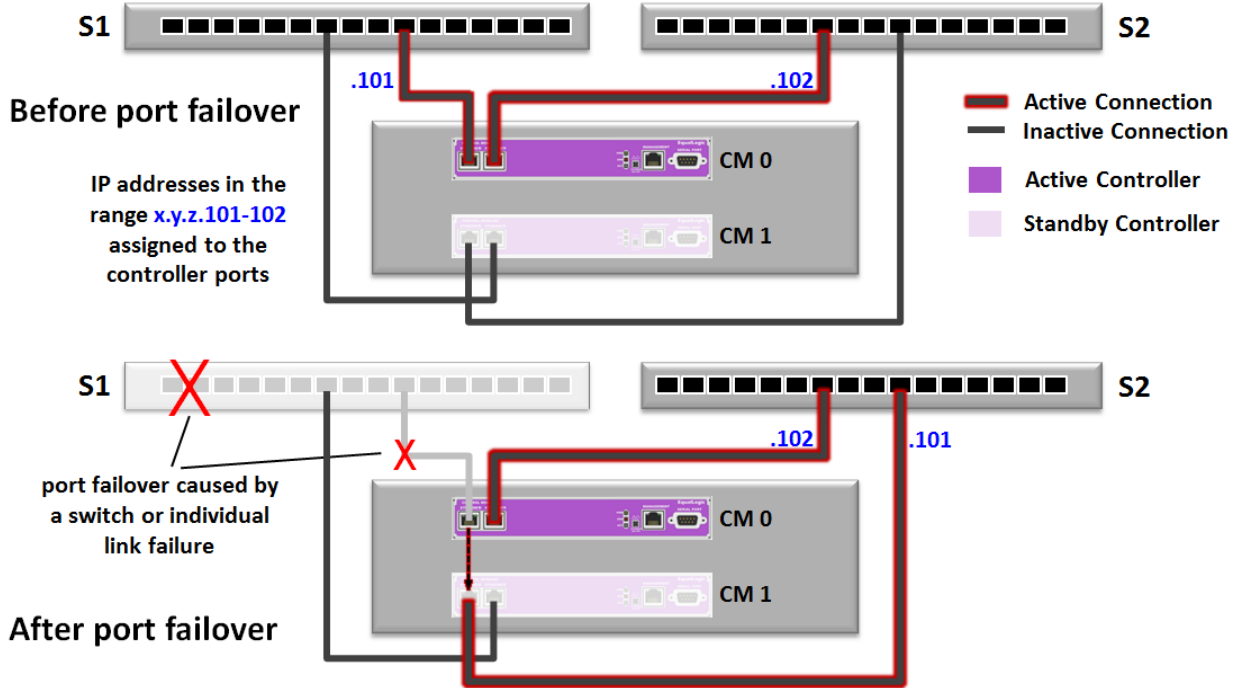


Figure 14 PS4100 vertical port failover and optimal connection paths

IMPORTANT: By alternating switch connection paths between ports in a vertical port pair, port failover allows the array to maintain 100% bandwidth capability in the event of a switch failure.

PS6100: Vertical Port Failover Scenario

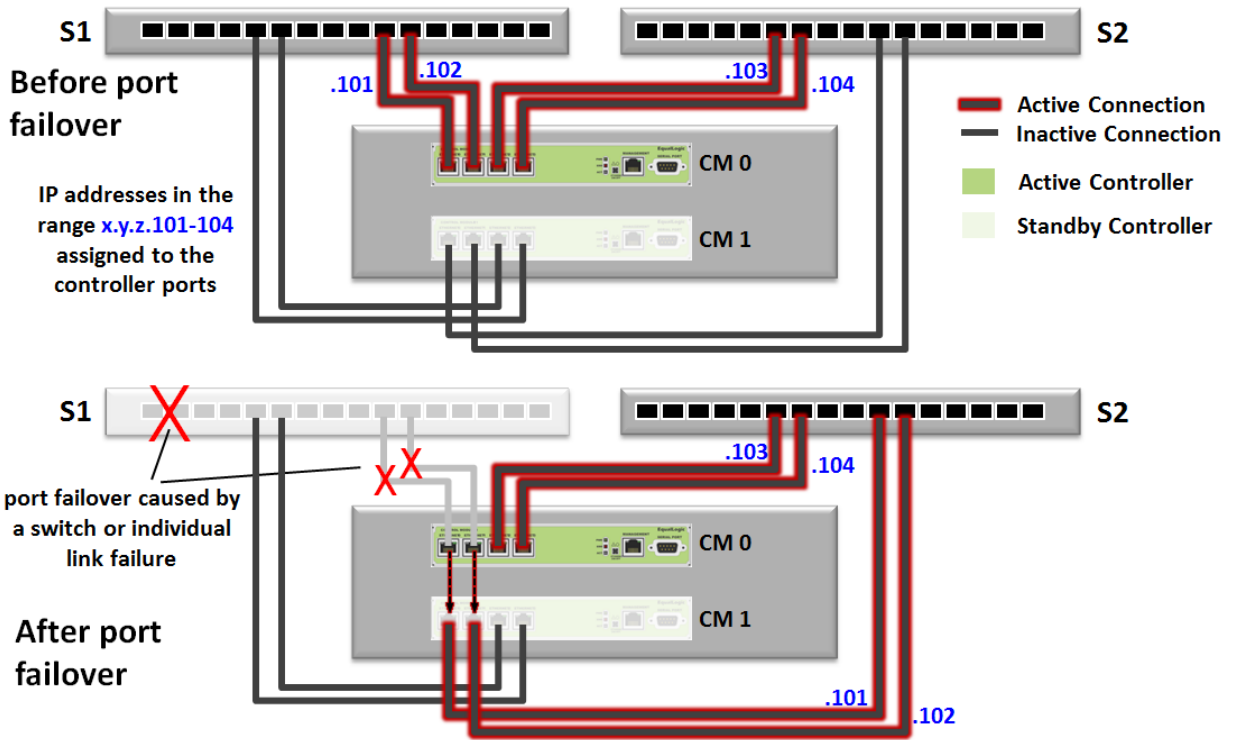


Figure 15 PS6100 vertical port failover process and optimal connection paths

9.6 Vertical port failover behavior in PS4110/PS6110 controllers

In PS Series controllers prior to PS4110/6110 families, a link failure or a switch failure was not recognized as a failure mode by the controller. This caused a failure of a link or an entire switch to reduce bandwidth available from the array. Referring to Figure 12 or Figure 13, assume that CM0 is the active controller. In vertical port failover, if CM0 senses a link drop on the local ETH0 port connection path, it will automatically begin using the ETH0 port on the backup controller (CM1). Vertical port failover is bi-directional. If CM1 is the active controller then vertical port failover will occur from CM1 ports to CM0 ports if necessary.



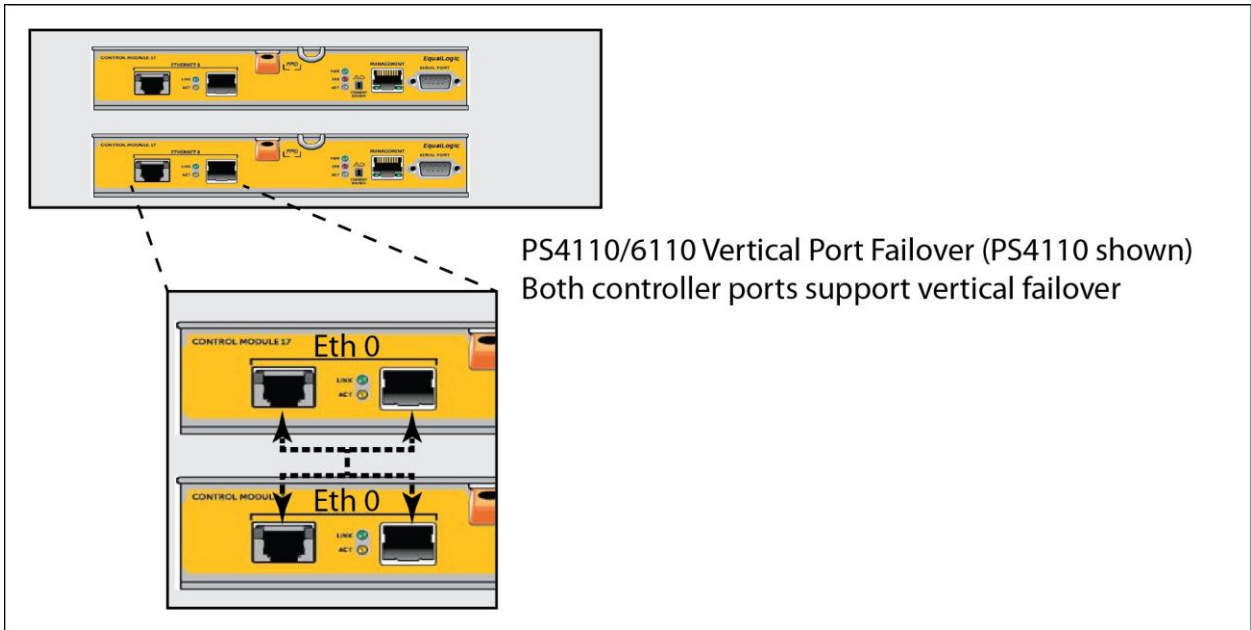


Figure 16 4110/6110 vertical port failover

With the PS4110/PS6110 family of controllers, vertical port failover can ensure continuous full bandwidth is available from the array even if you have a link or switch failure. This is accomplished by combining 10GbE “eth0” ports in each controller into a single logical port from the point of view of the active controller. In a fully redundant SAN configuration, you must configure the connection as shown in Figure 17.

In a redundant switch SAN configuration, to optimize the system response in the event you have a vertical port failover you must connect either the SFP+ cable or the 10Gbaset cable from each controller to a different switch in the SAN network. While it is a supported configuration, it is not recommended to connect both the SFP+ and 10Gbaset ports on each controller at the same time. In this scenario, the SFP+ port will always be the preferred active port and this preference cannot be changed by the user. This preference is not guaranteed in future array firmware releases.

The connection paths illustrated in Figure 17 show how to connect the port connection paths between the two controllers. Also note how IP addresses are assigned to the vertical port pair.



PS4110/PS6110: Vertical Port Failover Scenario

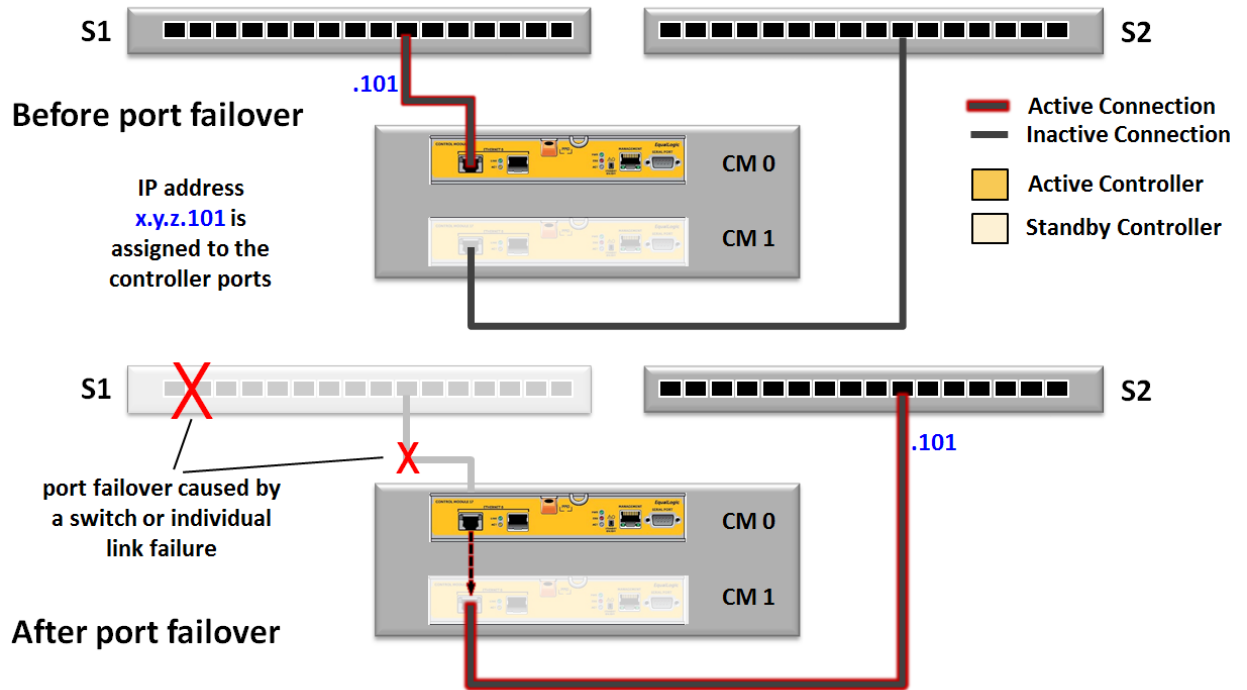


Figure 17 4110/6110 Vertical port failover scenario

9.7 Vertical port failover behavior in PS6210 controllers

In PS Series controllers prior to PS4110/6110/6210 families, a link failure or a switch failure was not recognized as a failure mode by the controller. This caused a failure of a link or an entire switch to reduce bandwidth available from the array. Referring to Figure 13, assume that CM0 is the active controller. In vertical port failover, if CM0 senses a link drop on the local ETH0 port connection path, it will automatically begin using the ETH0 port on the backup controller (CM1). Vertical port failover is bi-directional. If CM1 is the active controller then vertical port failover will occur from CM1 ports to CM0 ports if necessary.

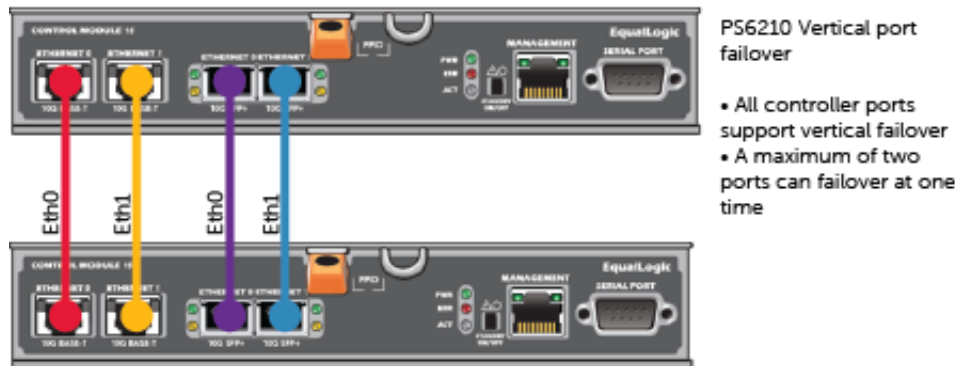


Figure 18 PS6210 vertical port failover

For more information about vertical port failover, see the PS6210 Hardware Owner's Manual at <https://eqsupport.dell.com/support/download.aspx> (requires support login). With the PS6210 family of controllers, vertical port failover can ensure continuous full bandwidth is available from the array even if you have a link or switch failure. This is accomplished by combining each 10GbE "eth0" and "eth1" ports in each controller into a single logical port from the point of view of the active controller. In a fully redundant SAN configuration, you must configure the connection as shown in Figure 18.

In a redundant switch SAN configuration, to optimize the system response in the event you have a vertical port failover you must connect either the SFP+ cable or the 10GbaseT cable from each controller to a different switch in the SAN network. While it is a supported configuration, it is not recommended to connect both the SFP+ and 10GbaseT ports on each controller at the same time. In this scenario, the SFP+ port will always be the preferred active port and this preference cannot be changed by the user. This preference is not guaranteed in future array firmware releases.

The connection paths illustrated in Figure 18 show how to connect the port connection paths between the two controllers. Also note how IP addresses are assigned to the vertical port pair.

PS6210: Vertical Port Failover Scenario

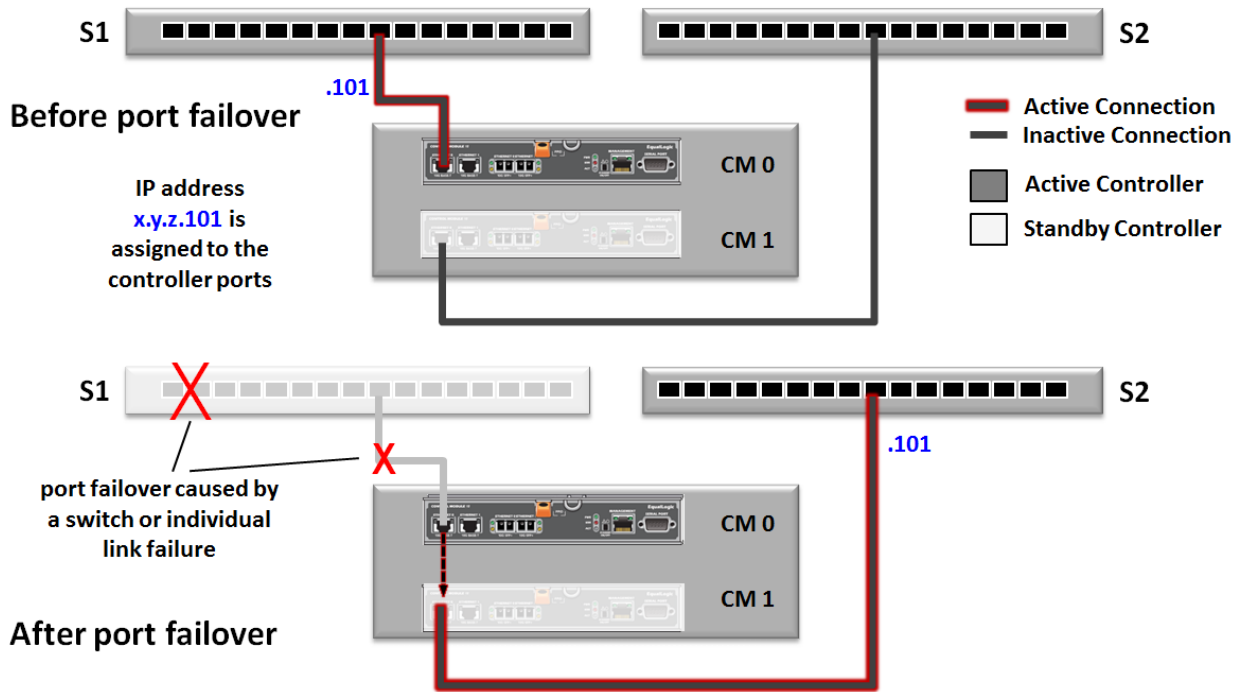


Figure 19 PS6210 Vertical port failover scenario

10 Mixed speed environments - Integrating 1GbE and 10GbE SANs

With the introduction of 10GbE, there will be situations that require 1Gb arrays and 10Gb arrays coexisting in the same SAN infrastructure. EqualLogic PS Series arrays support operation of 1Gb and 10Gb arrays within the same group. This section summarizes mixed speed SAN design guidelines that are presented in much more detail in the following publications:

- *Best Practices for Deploying Mixed 1Gb/10 Gb Ethernet SANs using Dell EqualLogic Storage Arrays and separate 1Gb and 10Gb switches:*
<http://en.community.dell.com/techcenter/storage/w/wiki/2640.deploying-mixed-1-gb-10-gb-ethernet-sans-using-dell-equallogic-storage-arrays-by-sis.aspx>
- *Best Practices for Deploying a Mixed 1 Gb/10 Gb Ethernet SAN using Dell EqualLogic Storage Array with Dell Networking 8100:*
<http://en.community.dell.com/techcenter/storage/w/wiki/4419.mixed-speed-devices-within-a-10-gb-equallogic-san-using-pc-8100-series-switches-by-sis.aspx>
- *Best Practices for Deploying a Mixed 1 Gb/10 Gb Ethernet SAN using Dell EqualLogic Storage Array with Dell Networking s4820T:*
http://en.community.dell.com/techcenter/extras/m/white_papers/20384458.aspx

The potential advantages in running a mixed speed (1GbE and 10GbE) EqualLogic SAN include:

- Not all of the application workloads on a SAN will require storage I/O performance that the 10Gb arrays provide. Thus, SAN administrators will have additional storage tiering flexibility based on array I/O performance.
- The PS Series Group Manager will allow the SAN administrator to still manage both types of arrays within the same SAN group.
- The ability to mix 1Gb and 10Gb arrays supports seamless operational coexistence during migration to a 10Gb SAN.

10.1 Mixed speed SAN best practices

The following list summarizes the important SAN design considerations for integrating 10Gb EqualLogic arrays into existing 1Gb EqualLogic SANs.

- When integrating 10Gb switches into your existing 1Gb switching environment, how you interconnect the mixed speed switches (split vs. straight uplink) does not have a significant impact on performance as long as the uplinks are sized appropriately to your workloads.
 - If your 1Gb switches are configured as a stack then you should use the split interconnect pattern as described in [Best Practices for Deploying a Mixed Speed 1Gb/10Gb Ethernet SAN using Dell EqualLogic Storage Arrays](#)
 - If your 1Gb switches are not stacked, then you must use the straight interconnect pattern



- When connecting 1Gb switches and 10Gb switches together you must always be aware of where Rapid Spanning Tree is going to block links to make sure that 10Gb traffic (i.e. EqualLogic inter-array data flow) never crosses the 1Gb switch.
- You must configure pools and volumes in a way that minimizes impact to IO performance.
 - Where possible, always connect 1Gb hosts only to 1Gb arrays and 10Gb hosts only to 10Gb arrays (except when performing migration tasks). Intermixing speeds may cause oversubscription of ports and lead to high latency or high retransmits.
 - When adding 10Gb arrays, place them in separate pools from your 1Gb arrays.



11 Blade server chassis integration

Integrating the PowerEdge M1000e Blade Server Solution (or any third party blade chassis implementation) requires additional SAN design considerations. 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 an intervening chassis midplane interface. Each fabric is associated with different interfaces on a given blade server as described in Table 16. Each blade server has a “LAN on Motherboard” capability that is mapped to the I/O modules located in the Fabric A I/O modules slots on the M1000e chassis and only supports 1Gb or 10Gb Ethernet networking depending on the blade server model. 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 I/O modules.

Figure 20 illustrates the layout of the three fabric I/O modules located on the back of the M1000e chassis.

Table 16 M1000e fabric mapping

	LOM/MLOM	Mezzanine B	Mezzanine C
Fabric	A	B	C

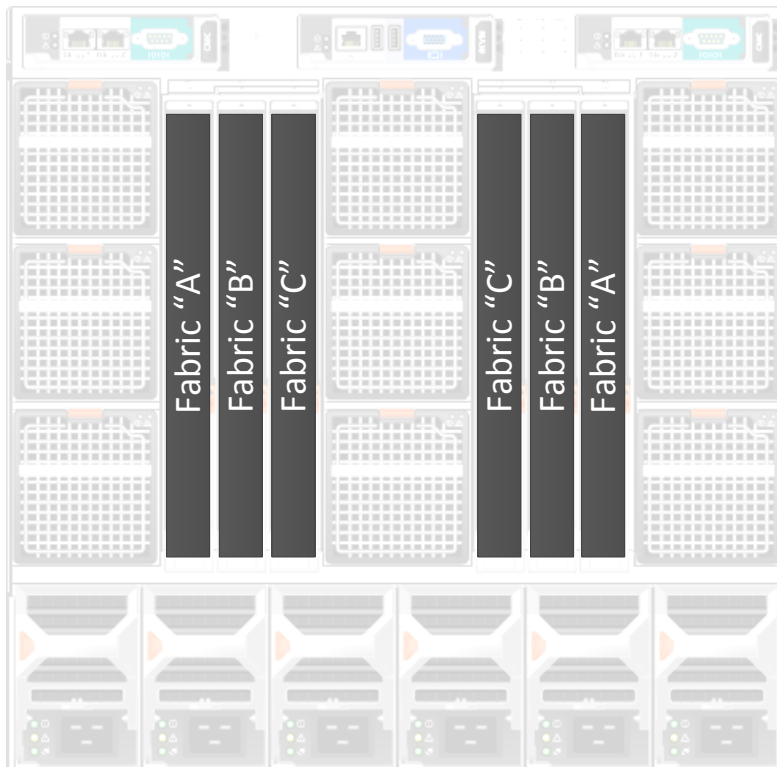


Figure 20 Blade I/O modules and M1000e Chassis



11.1 Designing a SAN using blade chassis I/O modules with arrays

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

1. **Blade IOM switch only (Direct-Attached)** – Network ports of both the hosts and storage are connected to the M1000e blade IOM switches. No Top of Rack (ToR) switches are required. The switch interconnect can be a stack or a LAG, and no uplink is required.
2. **ToR switch only (Pass-Through)** – Network ports of both the hosts and the storage are connected to external ToR switches. 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 VLT.
3. **Blade IOM switch with ToR switch (Two-Tier)** – 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 VLT and should connect the ToR switch to better facilitate inter-array member traffic. An uplink stack, LAG or VLT interconnect from the blade IOM switch tier to the ToR switch tier is also required.

For a detailed description of integrating blade chassis with EqualLogic arrays, see the following link:

<http://en.community.dell.com/techcenter/storage/w/wiki/4137.dell-powerededge-m1000e-blade-and-equallogic-ps-series-san-design-best-practices-by-sis.aspx>

For a detailed description of integrating blade chassis with EqualLogic arrays in a Data Center Bridging (DCB) environment, see the following link:

<http://en.community.dell.com/techcenter/storage/w/wiki/4451.poweredge-m1000e-blade-and-equallogic-ps-series-san-design-best-practices-using-force10-switches-by-sis.aspx>






12 FS Series NAS Configuration

In this section we provide detailed connection diagrams demonstrating how to setup fully connected iSCSI SAN and client LAN connection paths for the FS7500 and FS7600/FS7610 appliances.

Note: It is recommended to keep the client and SAN side networks physically separate and deploy at least two switches on both sides to provide redundancy in the event of a switch failure.

Table 17 lists the basic functional details for each FS Series product.

Table 17 FS Series Models

FS Series model	System components	I/O connections
FS7500 	2 x 1U NAS appliance; dual active-active controller configuration 1 x 1U Battery backup power supply unit (BPS)	Client LAN: 8x1GbE per appliance (4 per controller node) iSCSI SAN: 8x1GbE per system (4 per controller node) Controller Interconnect: 8x1GbE per system (4 per controller node)
FS7600 	2U NAS appliance with two active/active NAS controllers.	Client LAN: 8x1GbE per appliance (4 per controller node) iSCSI SAN: 8x1GbE per system (4 per controller node)
FS7610 	2U NAS appliance with two active/active NAS controllers.	Client LAN: 4x10GbE per appliance (2 per controller node) iSCSI SAN: 4x10GbE per system (2 per controller node)

12.1 FS7500 connection paths

The FS7500 appliance is comprised of two peer system controller nodes. In the required fully cabled configuration, each controller node requires thirteen separate Ethernet cable connections. Thus a single FS7500 appliance comprised of two controller nodes requires a total of 26 Ethernet connections (four connecting to the client LAN switches and nine connecting to the iSCSI SAN fabric switches.)



Figure 21 shows the client LAN connection paths.

Note: While it is possible to operate an FS7500 appliance in a partially cabled configuration, this configuration is not supported by Dell. You should use a fully cabled configuration in a production environment. You will also need to provision the required switch port count on the iSCSI SAN and client LAN sides of the system to support a fully connected configuration.

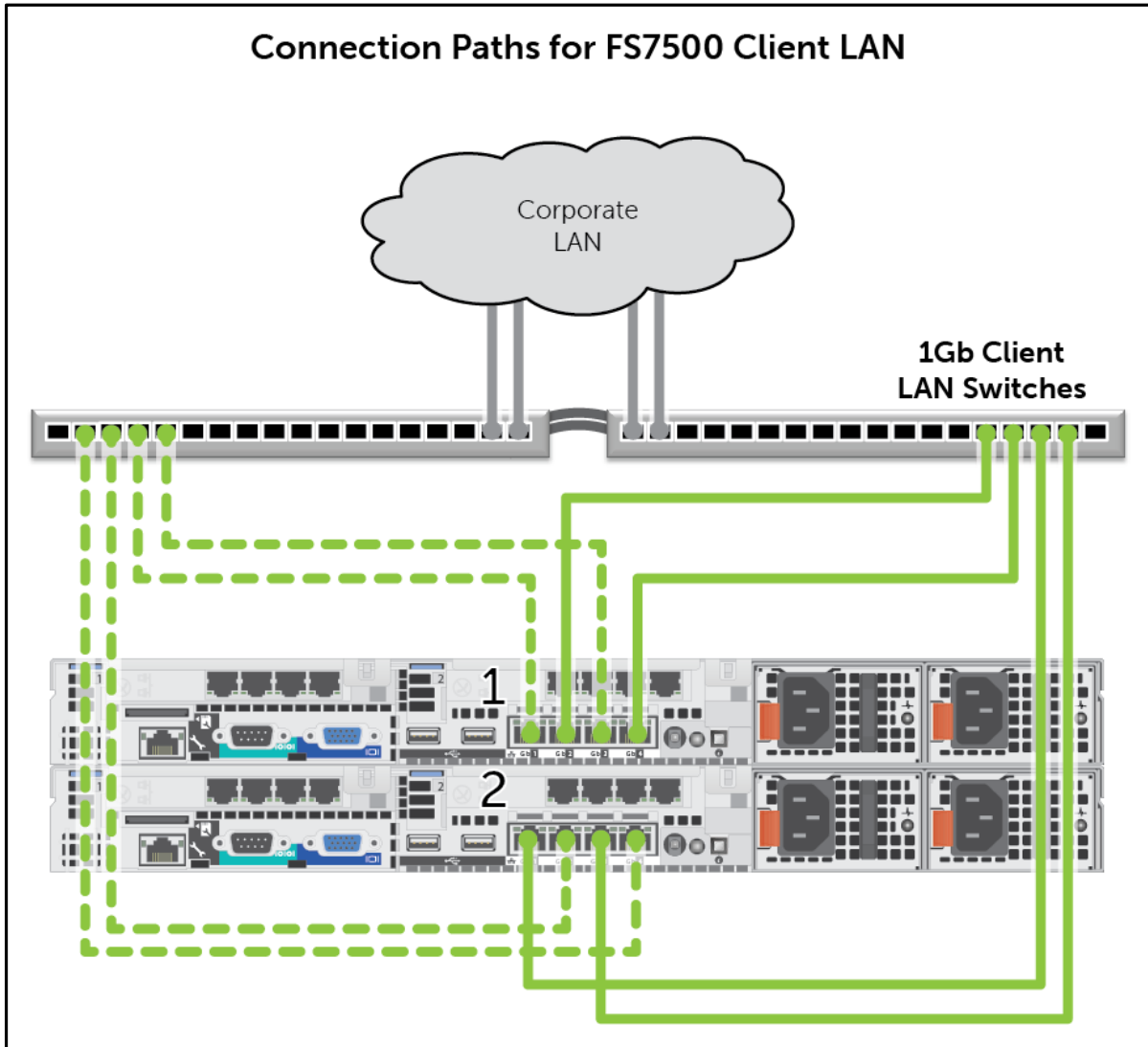


Figure 21 Connection Paths for FS7500 Client LAN

Figure 22 below shows the iSCSI SAN, IPMI, and node interconnect paths. Pay careful attention to how the controller ports alternate between redundant switch paths.

Note: With the exception of the IPMI connection paths, corresponding ports on each controller node must connect to the same SAN switch. This connection pattern is shown in Figure 22.



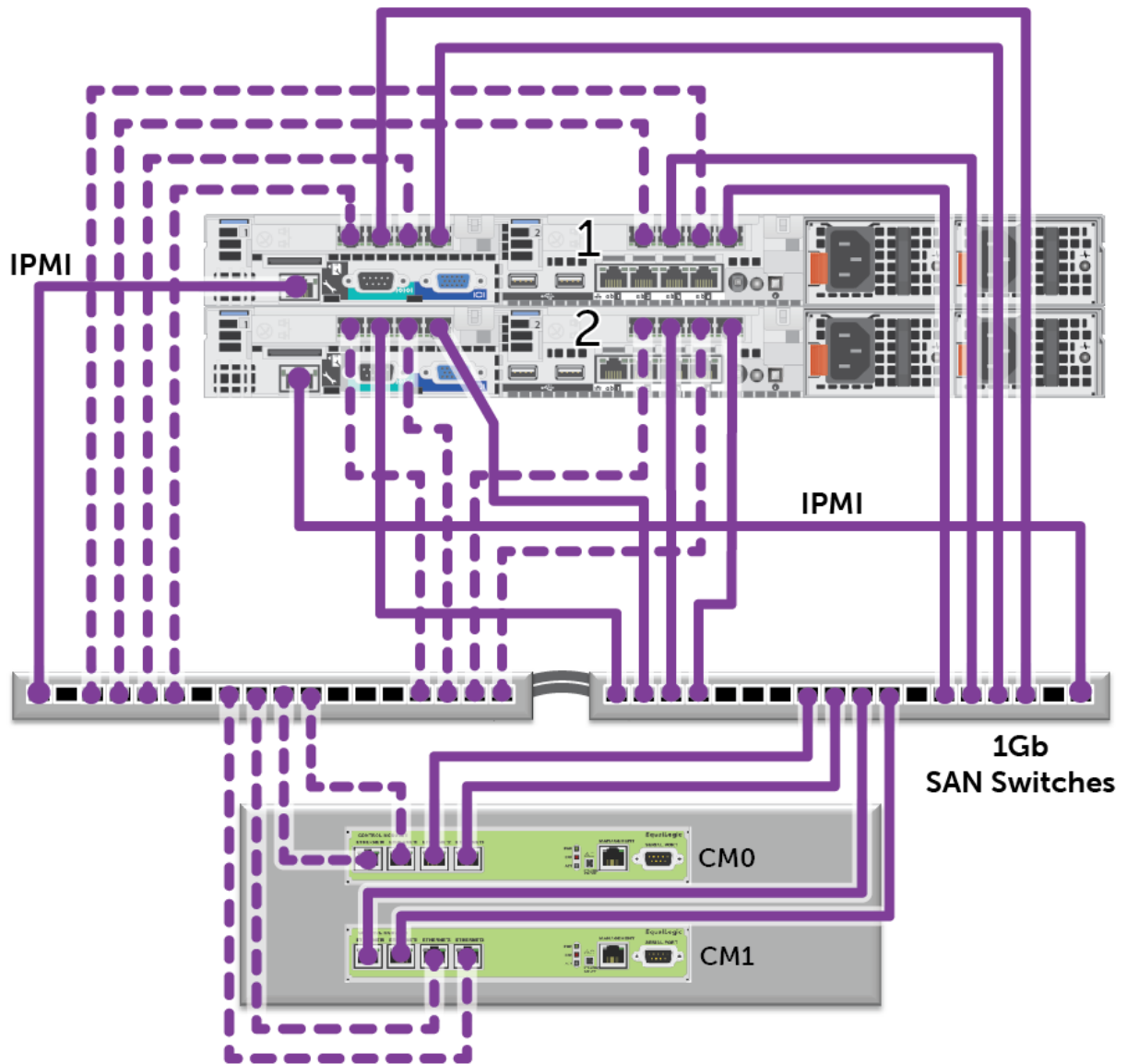


Figure 22 Connection Paths for FS7500 iSCSI SAN, IPMI and Controller Interconnect

The inter-switch connection sizing guidelines provided in Section 8.2.2 also apply to FS7500 SAN design. The FS7500 mirrors the write cache between controller nodes. To accomplish this, all write operations are transmitted across the controller interconnect. Thus, it is very important that you follow connection pattern shown in Figure 22 to ensure corresponding ports are connected to the same switch. This connection pattern prevents controller interconnect traffic from crossing the inter-switch connection.

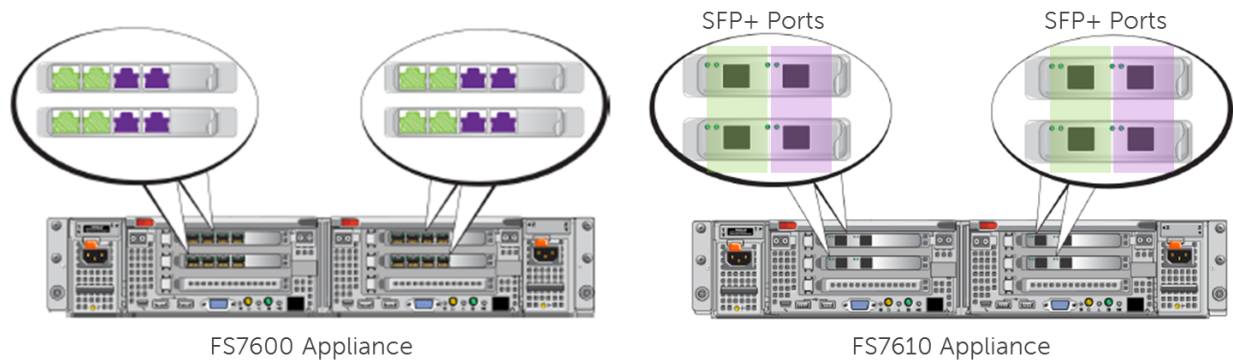


12.2 FS7600/FS7610 connection paths

The Dell EqualLogic NAS appliances require the following networks:

- Client network: Used for client access to the NFS exports and CIFS shares hosted by the NAS cluster.
- SAN/internal network: Used for internal communication between the controllers and communication between the controllers and the EqualLogic PS Series SAN. The SAN and Internal networks use the same set of switches.

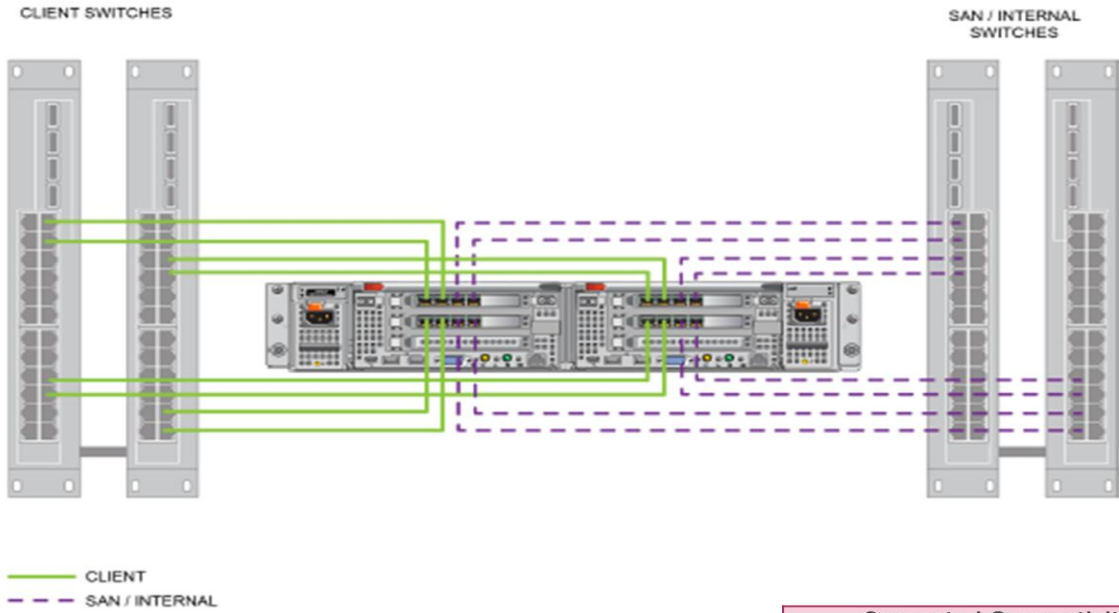
It is recommended to keep the client and SAN side networks physically separate and deploy two switches on both sides to protect against a switch failure.



Network	Network Usage	Ethernet Port Numbers
Client network	The Client network allows the clients (workstations, PCs, etc.) to access network shares.	Left port in all NICs
SAN / Internal networks	-The SAN network connections allow communication between the NAS cluster and the PS Series group (SAN). The SAN and internal networks use two different IP address ranges, but they exist on the same switch or VLAN. -The internal network connections allow communication between the two controllers.	Right port in all NICs

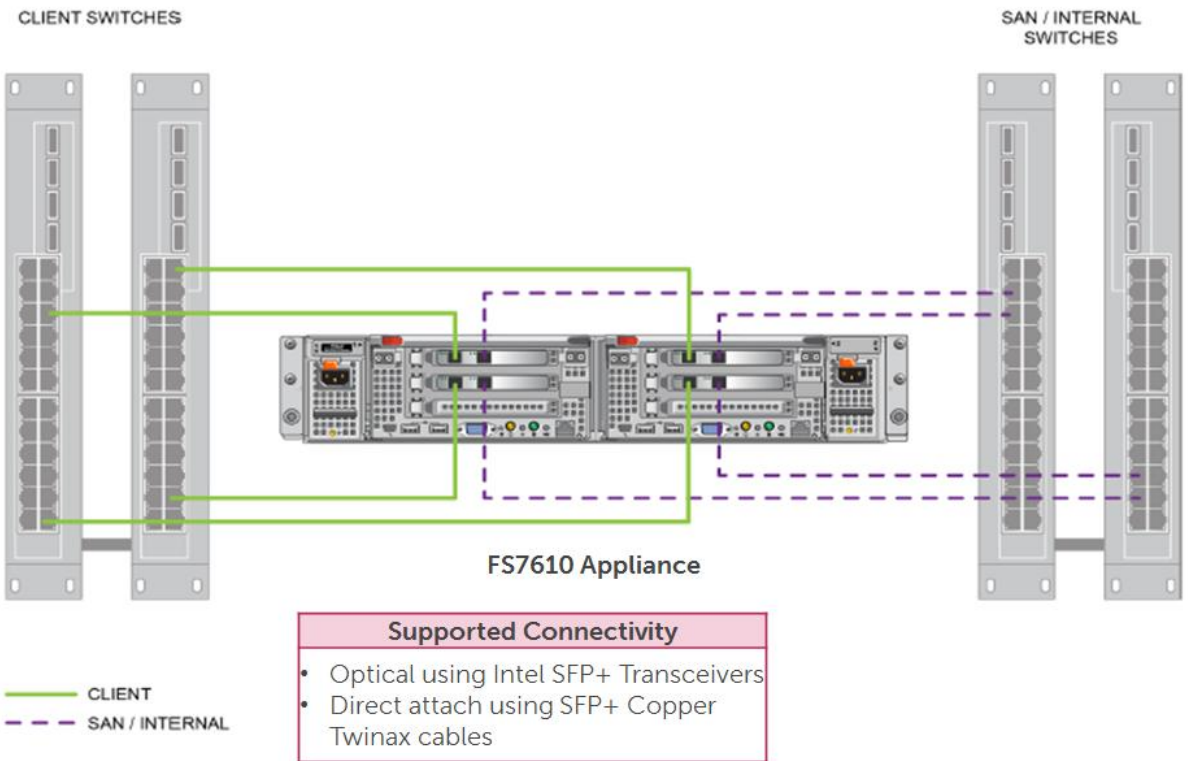
Figure 23 FS7600 and FS7610 networks

See figures below for network connections.



Supported Connectivity
<ul style="list-style-type: none"> • 1Gb Ethernet, RJ45 connector

Figure 24 FS7600 network



Supported Connectivity
<ul style="list-style-type: none"> • Optical using Intel SFP+ Transceivers • Direct attach using SFP+ Copper Twinax cables

Figure 25 FS7610 network



Installation/Expansion

- If installing FS7500/FS76x0 into an existing EqualLogic SAN, verify the existing LAN and SAN networks meet the minimum requirements.
- All NAS appliances in a NAS cluster must be of the same speed. Appliances with different connectivity cannot be mixed in a NAS cluster.
 - An FS7500 cluster can be expanded by using another FS7500 or an FS7600 but NOT an FS7610.
- Do not block IPv6 traffic on the SAN internal switches when utilizing FS Appliances. The FS76x0 uses IPv6 for internal communication and setup only; not for user data. Client communication occurs over IPv4.

Network

- The two iDRAC IPMI ports on the FS7500 are 100Mb. The SAN switches must be able to support 100Mb in addition to 1Gbps speed.



13 Data Center Bridging (DCB)

The enhancement to the Ethernet Specifications (IEEE 802.3 specifications) called Data Center Bridging (DCB) enables bandwidth allocation and lossless behavior for storage traffic when the same physical network infrastructure is shared between storage and other traffic.

The network is the fundamental resource that connects the assorted devices together to form the datacenter Ethernet infrastructure. These devices include the server hardware (along with the operating system and the applications that run on the host) and the storage systems that host application data. Sharing this Ethernet infrastructure with multiple traffic types (LAN and SAN) requires a fairness mechanism to provide bandwidth allocation and flow control for each type of traffic. Without such a mechanism, the Local Area Network (LAN) traffic and the Storage Area Network (SAN) traffic would have to be separated onto their own dedicated networks to ensure consistent performance and reliability.

When the SAN and LAN networks are shared, all traffic is equal, unless Quality of Service (QoS) and/or Class of Service (CoS) is used.

Note: Traditional or Non-DCB QoS (IEEE 802.1p) is not supported for EqualLogic implementations.

In a shared network environment, LAN and SAN traffic can impact each other and QoS may not solve this because of QoS implementation differences with vendors. These different implementations include:

- Number of Queues managed
- Relative priorities between queues
- Bandwidth reserved per queue
- Congestion management methods

Also traditional QoS lacks selective flow control for each traffic type. Regular IEEE 802.3 PAUSE will pause the entire link and not selectively pause LAN or SAN traffic. This flow control ability is important for reducing congestion on switches and to enable fair sharing of resources between traffic types.

Other methods of network sharing:

- VLANs offer port-level security and segregation, but do not provide guaranteed bandwidth or QoS.
- NIC partitioning (NPAR) manages traffic within the host. Once the network data exits the NIC to the switch, any QoS/bandwidth management enforced by NPAR is not honored by the switch.

Note: It is not recommended to share network infrastructure without DCB. DCB is the only recommended method of converging SAN and LAN in an EqualLogic Storage environment.



13.1 DCB Overview

DCB is a collection of Standards designed to improve networking and management in the Data Center that enables iSCSI SANs or FCoE SANs or both to converge with regular server LAN traffic on the same physical infrastructure. DCB aims to increase operational efficiency, constrain costs, and ease network management.

There are primarily three progressive versions of DCB:

- Cisco, Intel, Nuova (CIN) DCBX
- Converged Enhanced Ethernet (CEE) DCBX or baseline DCBX
- Institute of Electrical and Electronic Engineers (IEEE) DCB.

DCB technologies based on standards include:

- PFC – Priority based Flow Control (802.1Qbb)
- ETS – Enhanced Transmission Selection (802.1Qaz)
- CN – Congestion Notification (802.1Qau)
- DCBx – Data Center Bridging Capability eXchange
 - Support for iSCSI application protocol priority with DCBX (also known as iSCSI TLV)

Note: DCB uses 10Gb and higher Ethernet only.

I/O convergence using DCB for Dell EqualLogic iSCSI storage is the direction for “Converged iSCSI” in a lossless Ethernet environment.

13.2 DCB requirements for EqualLogic

It is required that all devices in the EqualLogic SAN support DCB for iSCSI when this functionality is enabled. If any device in the SAN does not support DCB, then DCB needs to be disabled at the switches for the entire SAN. Once all devices in the SAN are DCB compliant, then DCB can be re-enabled. Switches and server CNAs/NICs that are designated as DCB Supported in the EqualLogic Compatibility Matrix have been fully validated by Dell to ensure compatibility for EqualLogic SANs.

The minimum requirements for switches and server-based CNA/NICs to support an end-to-end DCB solution with EqualLogic are:

- Data Center Bridging Exchange (DCBx) -DCB protocol that performs discovery, configuration, and mismatch resolution using Link Layer Discovery Protocol (LLDP)
- Application Priority (iSCSI TLV) - Switches must support configuration of a priority value for iSCSI protocol and advertisement to peer ports. Server NICs/CNAs: Must support iSCSI protocol in application priority (learned from the switch) and must support tagging outgoing iSCSI frames with this priority.
- ETS - Requires a dedicated traffic class or priority group for iSCSI priority



- PFC - Requires enabling PFC (no drop or lossless behavior) for iSCSI priority
- Switches: Configure dedicated TC/PG for iSCSI priority with allocated bandwidth and PFC enabled for iSCSI priority
- Server NICs/CNAs: Adhere to TC/PG mapping for iSCSI priority and PFC for iSCSI priority (learned from the switch)

Designing a converged network deployment with components that have no DCB support or partial DCB support is not recommended for end-to-end converged I/O. Configurations with components that have no DCB support or partial DCB support may result in loss of expected functionality or may be functional without meeting the converged network objectives.

Note: It is not recommended to share network infrastructure with DCB and Non-DCB components.

For a complete list of components that support the DCB standards, see the: [EqualLogic Compatibility Matrix](#)

Note: It is important to verify that all components in the SAN are listed in the EqualLogic Compatibility Matrix as DCB Supported, or that the components support all EqualLogic requirements for DCB.

Links to DCB papers:

Data Center Bridging: Standards, Behavioral Requirements, and Configuration Guidelines with Dell EqualLogic iSCSI SANs: <http://en.community.dell.com/techcenter/storage/w/wiki/4396.data-center-bridging-standards-behavioral-requirements-and-configuration-guidelines-by-sis.aspx>

EqualLogic DCB Configuration Best Practices: <http://en.community.dell.com/techcenter/storage/w/wiki/4355.configuring-dcb-with-equallogic-sans.aspx>

Dell PowerEdge M1000e Blade Enclosure and EqualLogic PS Series SAN Design Best Practices Using Force10 Switches: <http://en.community.dell.com/techcenter/storage/w/wiki/4451.poweredge-m1000e-blade-and-equallogic-ps-series-san-design-best-practices-using-force10-switches-by-sis.aspx>

13.3 VLANs for iSCSI

A non-default VLAN is required for operating prioritized lossless iSCSI traffic in a DCB enabled Ethernet infrastructure. Switch ports that are based on the IEEE 802.1Q VLAN specification forward frames in the default or native VLAN without tags (untagged frames). Normally these ports also receive frames in the default or native VLAN as untagged. Such a port is typically termed a "tagged port" or "trunk port" and all non-default VLAN frames are forwarded with the tags intact. Since the DCB priority information (PCP value) is encoded in the VLAN tag, this information will be lost if an end-device sends iSCSI traffic in a default or native VLAN and the switch receives it in a tagged or trunk port. As a result, DCB prioritization and traffic classification will be lost.



All devices in the iSCSI data path must have the same VLAN ID configured on the respective ports participating in the iSCSI network to ensure proper functioning. These devices include the server iSCSI NIC/CNA ports, EqualLogic arrays, and all switches on the iSCSI SAN.

Note: The VLAN ID for iSCSI can be set in the EqualLogic Group Manager interface or the storage array CLI. When DCB is enabled on the switch, it is necessary to configure a non-default VLAN on the array, switch, and all host ports that are part of the EqualLogic SAN. VLAN IDs 0 and 1 are not recommended, because these may be the default or reserved VLAN for some switches, and as such, may forward frames untagged (e.g. no VLAN tagging). VLAN tagging is required to fully support DCB.

For more information on the DCB requirements and configuration guidelines, see the following white paper: [*Data Center Bridging: Standards, Behavioral Requirements, and Configuration Guidelines with Dell EqualLogic iSCSI SANs*](#)

For more information on the DCB requirements for EqualLogic, to ensure that DCB is properly enabled and/or disabled across all devices, and to assist with identifying and resolving basic DCB configuration issues in the EqualLogic SAN, see the following white paper: [*EqualLogic DCB Configuration Best Practices*](#)

For a comprehensive understanding of integrating an M1000e Blade Chassis in a Full End-to-End Data Center Bridging environment, see the following white paper: [*M1000e Blade Enclosure and EqualLogic Arrays SAN Design using Force10 Switches*](#)



A Network ports and protocols

PS Series groups use a number of TCP and UDP protocols for group management, I/O operations, and internal communication. If you have switches or routers set to block these protocols, you may need to unblock them to allow management or I/O operations to work correctly. The required and optional protocols are listed in the following sections.

A.1 Required ports and protocols

Table 18 lists the ports and protocols required for operating an EqualLogic iSCSI SAN.

Table 18 Required ports and protocols

Type	Port	Protocol	Access
iSCSI			
TCP	3260	iSCSI	To the group IP address and all individual member IP addresses
EqualLogic Internal			
UDP	161	SNMP	Management operations
TCP	9876	Internal	iSCSI intra-system control
TCP	25555	Internal	Group communication
TCP	20002	Internal	Event logging

A.2 Optional ports and protocols

Table 19 lists the optional ports and protocols used for management and alerts. They are not required for correct array operation, however blocking a port may reduce specific functionality.



Table 19 Optional ports and protocols

Type	Port	Protocol	Access
CLI Management			
TCP	23	Telnet	To group IP address
TCP	22	SSH	To group IP address
Web Based Management			
TCP	80	HTTP	To group IP address
TCP	3002	GUI communication	To group IP address
TCP	3003	GUI communication (encrypted)	To group IP address
SNMP			
UDP	161	SNMP	To and from group IP address
Syslog			
UDP	514	Syslog	From group IP address
EqualLogic Diagnostics			
TCP	21	FTP	Software update and diagnostic procedures; to all individual member IP addresses
TCP	25	SMTP	E-mail and diagnostic notifications; from all individual member IP addresses to the configured SMTP server
NTP			
UDP	123	NTP	To external Network Time Protocol (NTP) sources for time-of-day clock synchronization



B Upgrade paths for EqualLogic PS Series arrays

Table 20 EqualLogic upgrade paths

End of sales life arrays:	Latest available conversion model	1Gb to 10Gb conversion availability	Drive upgrades availability
PS-50 thru PS3000	None	None	None
PS4000	Yes, PS6000	Yes, PS6010	Yes
PS5000	Yes, PS6000	Yes, PS6010	Yes
PS6000	None	Yes, PS6010	Yes
PS5500	Yes, PS6500	Yes, PS6510	Yes
PS6010	None	N/A	Yes

Currently shipping arrays:	Latest available conversion model	1Gb to 10Gb conversion availability	Drive upgrades availability
PS6500	None	Yes, PS6510	Yes
PS6510	None	N/A	Yes
PS4100	N/A	N/A	Yes
PS4110	N/A	N/A	Yes
PS6100	None	N/A	Yes
PS6110	None	N/A	Yes
PS-M4110	None	N/A	Yes

