




Dell EqualLogic PS Series Storage Arrays
iSCSI Initiator and Operating System Considerations
Optimizing Your SAN Environment for High Availability



Notes, cautions, and warnings

-  **NOTE:** A NOTE indicates important information that helps you make better use of your product.
-  **CAUTION:** A CAUTION indicates either potential damage to hardware or loss of data and tells you how to avoid the problem.
-  **WARNING:** A WARNING indicates a potential for property damage, personal injury, or death.

© 2018 Dell Inc. or its subsidiaries. All rights reserved. Dell, EMC, and other trademarks are trademarks of Dell Inc. or its subsidiaries. Other trademarks may be trademarks of their respective owners.

2018 - 03

Rev. 110-6299-EN-R1

Contents

About This Manual.....	5
Audience.....	5
Related Documentation.....	5
Related Sites.....	5
Online Services.....	5
Dell EqualLogic Storage Solutions.....	6
Technical Support and Customer Service.....	6
Contacting Dell.....	6
Proactive Maintenance Support.....	6
1 About High-Availability iSCSI Data Transfer.....	7
High-Availability Procedures	7
Update Initiators and Drivers.....	7
2 Configure Your Hosts for High Availability.....	8
Host Integration Tools.....	8
Operating System Recommendations.....	8
Test Modifications in a Non-Production Environment.....	8
Increase Transmit and Receive Resource Settings.....	8
Operating System-Specific Configuration Information.....	8
3 Configure Your Network for High Availability.....	9
General Network Requirements.....	9
Design Network for Redundancy.....	9
Ports.....	9
Switches.....	9
FS Series Switch Requirements.....	11
Dell Networking N-Series Switching MPIO With M-LAG Advisory.....	11
Hardware Initiators.....	11
A Configure Your Microsoft Windows Operating System.....	12
Procedure Summary.....	12
1. Install the Recommended Microsoft Hotfixes.....	12
2. Install the Host Integration Tools for Microsoft.....	13
3. Set the Registry Entries.....	13
Use the Registry Editor Carefully.....	13
4. Verify Microsoft Windows Guidelines.....	16
Application-Specific Events.....	16
Microsoft Windows Applications.....	16
Microsoft Exchange Server.....	16
Set Delay Between Login Attempts.....	16
Enable ICMP Echo Requests for iSCSI Initiators in Windows.....	17



B Configure Your VMware Environment.....	18
Set iSCSI Login Timeout Value on vSphere 5.1 and 5.0.....	18
Storage Heartbeat on vSphere 5.0, 4.1, and 4.0.....	18
C Configure Your Linux Operating System.....	20
Configure SCSI Device Settings Using the udev Facility.....	20
Increase SCSI Device Timeout Value to Avoid Disk I/O Timeouts.....	20
Increase SCSI Device Read-Ahead Buffer Size.....	20
Change SCSI Device I/O Scheduler Algorithm.....	20
Configure iSCSI Settings Using the /etc/iscsi/iscsid.conf File.....	20
Change FastAbort Setting in iSCSI Configuration File.....	20
Change node.startup Setting to Avoid Login Storms.....	20
Configure Linux Networking.....	21
Disable Generic Receive Offload.....	21
Enable Flow Control.....	21
Enable Jumbo Frames.....	21
Avoid ARP Flux.....	21
Loosen IPv4 Reverse Path Filtering.....	21
Summary of Recommended Changes.....	22
D Configure Your Solaris 10 Operating System.....	23
Zero-Length Volumes.....	23
E Configure Your IBM AIX Operating System.....	24
iSCSI Initiators Lose Volume Connectivity.....	24



About This Manual

Dell EqualLogic PS Series arrays optimize resources by automating performance and network load balancing. Additionally, PS Series arrays offer all-inclusive management firmware, host software, and free firmware updates. Dell EqualLogic FS Series appliances, combined with PS Series arrays, offer a high-performance, high-availability, scalable NAS solution.

This document describes configuration information, limitations, and other considerations required to architect a Dell EqualLogic SAN for high availability and preparation for non-disruptive array firmware updates.

Audience

The information in this guide is intended for experienced administrators who are responsible for managing a PS Series group. To use the information in this manual, you must have an understanding of the following topics:

- PS Series group administration
- Operating environment administration for:
 - Windows (or Windows Server) systems
 - VMware
 - Linux systems
- Basic network and SAN administration

Related Documentation

For more information about PS Series arrays, FS Series appliances, groups, volumes, array firmware, and host software, log in to the [Documentation page](#) on the customer support site.

Related Sites

- The tech center site:
 - Provides an extensive compilation of application and networking best practices and reference architectures for Dell EqualLogic. See: <http://en.community.dell.com/techcenter/storage/>
 - Contains the *PS Series Configuration Guide*, available from: <http://en.community.dell.com/techcenter/storage/w/wiki/equallogic-configuration-guide>
- The support site contains the latest Dell EqualLogic firmware compatibility information. See: eqlsupport.dell.com/support/download.aspx?id=6442454231&langtype=1033

Online Services

To learn about Dell products and services:

1. Visit dell.com or the URL specified in any Dell product information.
2. Use the locale menu, or click the link that specifies your country or region.



Dell EqualLogic Storage Solutions

To learn more about Dell EqualLogic products and new releases being planned, visit the Dell EqualLogic TechCenter site: [Dell EqualLogic TechCenter](#). You can also see articles, demos, online discussions, and more details about the benefits of our product family.

Technical Support and Customer Service

Dell support service is available to answer your questions about PS Series SAN arrays and FS Series appliances.

Contacting Dell

Dell provides several online and telephone-based support and service options. Availability varies by country and product, and some services might not be available in your area. To contact Dell for sales, technical support, or customer service issues, go to dell.com/support.

Proactive Maintenance Support

If you have a Dell proactive service agreement, your Dell Technical Support representative can help you analyze your SAN environment and ensure it is correctly set up for high availability.

Proactive Maintenance provides maintenance events on supported Dell systems. These events can assist with analyzing performance issues and can help reduce the likelihood of future incidents due to incompatible hardware, software, BIOS, and firmware versions. Proactive Maintenance services can easily be scheduled through Dell Technical Support and can be delivered at any time (24 hours a day, 7 days a week).

For more information about Proactive Maintenance, visit dell.com/services.



About High-Availability iSCSI Data Transfer

When you optimize your SAN environment for high availability (HA), you increase the fault tolerance of the SAN. Fault tolerance provides protection from component failures and minimizes service disruptions.


Many causes of service disruptions can occur in a data center. Some of the more common disruptions are:

- Hardware failure in servers, NICs, cables, switches, storage controllers, and disks
- Software failure such as bugs or application crashes
- Updates to software, driver, or hardware firmware

In most cases, if your data center is designed correctly for redundancy, the SAN can tolerate service disruptions. In some cases, however, a redundant design is not sufficient to allow applications and operating systems to continue to operate during disruptions.

For example, problems with software or updates to software and firmware can cause disruptions that redundancy cannot manage. In these cases, other best practices are typically required to tune applications and operating systems to continue to operate without being affected.

This document describes the actions that you can take to ensure a highly available iSCSI data transfer that minimizes the possibility of service issues with Dell EqualLogic SANs.

 **NOTE: If you have a Dell proactive service agreement, your Dell Technical Support representative can help you analyze your SAN environment and ensure that it is set up correctly for high availability.**

For more information about Proactive Maintenance, visit dell.com/services.

High-Availability Procedures

You must perform certain procedures to optimize your Dell EqualLogic SAN environment for high availability (HA) and non-disruptive operation.

Read the following sections and then perform the recommended actions:

- [Update Initiators and Drivers](#) is recommended before you begin the configuration process.
- [Configure Your Hosts for High Availability](#) contains general operating system recommendations and guidelines to configure settings on your servers to ensure high availability.
- [Configure Your Network for High Availability](#) outlines the general best practices that are required to configure a redundant iSCSI network.

Update Initiators and Drivers

Dell recommends that you keep initiators and drivers up to date. A list of supported initiators can be found in the *PS Series Configuration Guide*. Apply the recommended modifications to your operating system and iSCSI initiators as described in this guide. See:

<http://en.community.dell.com/techcenter/storage/w/wiki/equallogic-configuration-guide>



Configure Your Hosts for High Availability

This section provides information about Host Integration Tools and operating system recommendations.

Host Integration Tools

Dell has automated a portion of the configuration process in the Dell EqualLogic Host Integration Tools kits. These kits provide integrated software such as MPIO modules, enhanced data protection, automated provisioning, and alerts for Windows, VMware, and Linux hosts. If installed, this software implements changes (such as configuration options on the hosts) that incorporate Dell's recommended best practices.

You can download the Dell EqualLogic Host Integration Tools kits from the eglsupport.dell.com website, and install them on supported operating systems.

Operating System Recommendations

This section contains detailed information about settings and configuration options that you should apply to hosts running in your Dell EqualLogic iSCSI SAN to ensure high availability (HA) at the host level.

Test Modifications in a Non-Production Environment

If you are not running any of the operating systems or iSCSI initiators discussed in this document, or if you are unsure of your applications' HA tolerance, Dell recommends that you first test the suggested modifications in a non-production environment.

Increase Transmit and Receive Resource Settings

A PS Series group might deliver more network I/O than iSCSI initiators can handle, resulting in dropped packets and retransmissions. Some iSCSI initiators, however, let you increase the transmit and receive resource settings, which can help minimize this problem.

For example, on a Windows system, you might be able to increase the initiator's network interface settings for receive and transmit descriptors. On a UNIX or Linux system, increase the transmit and receive resources for your network ports.

Operating System-Specific Configuration Information

For configuration information specific to your operating system, see the following sections:

- [Configure Your Microsoft Windows Operating System](#)
- [Configure Your VMware Environment](#)
- [Configure Your Linux Operating System](#)
- [Configure Your Solaris 10 Operating System](#)
- [Configure Your IBM AIX Operating System](#)



Configure Your Network for High Availability

A Dell 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, many network design and configuration choices can affect SAN performance.

For additional network design and configuration information to support the use of a Dell EqualLogic SAN, see the *PS Series Configuration Guide*. This guide can be downloaded from:

<http://en.community.dell.com/techcenter/storage/w/wiki/equallogic-configuration-guide>

General Network Requirements

This section outlines the general best practices for configuring a redundant iSCSI network.

Design Network for Redundancy

For Dell EqualLogic PS Series arrays, follow these general SAN design requirements.

Ports

- Dell recommends that, whenever possible, the arrays should be in the same subnet as the group's IP address. For all members (arrays) in a given SAN group, all array Ethernet ports should be connected to the same subnet. This connection allows the arrays to communicate with each other as a group of peer members.
- At least two iSCSI SAN ports per host must be connected to two different switches in a SAN fabric for fully redundant SAN connectivity.

Switches

- All switches within the SAN must be interconnected such that a path is always present from any Ethernet port on one array to all other Ethernet ports on all other arrays in the group. The preferred method of interconnect is stacking the switches.
- Enabling jumbo frames on all switches in the SAN is a recommended best practice as long as the rest of the infrastructure (such as initiators and NICs) supports jumbo frames as well. This requirement applies to SANs connected to either a PS Series array or an FS Series appliance.
- Rapid Spanning Tree Protocol must be enabled if the SAN infrastructure has more than two switches in a nonstacked configuration, and portfast must be enabled on all edge device ports (hosts, and FS Series appliances and PS Series arrays).
- To prevent a switch failure from also disabling all paths between a host and its connected volumes, all ports from each controller must be connected to at least two different switches. This requirement also applies to FS Series NAS appliance configurations.

For PS4100/PS6100 family arrays, split the vertical port pair connections between two switches to ensure that 100-percent bandwidth capability is maintained in the event of a vertical port failover. Split the connections as shown in [Figure 1. Redundant SAN PS4100](#) and [Figure 2. Redundant SAN PS6100](#).



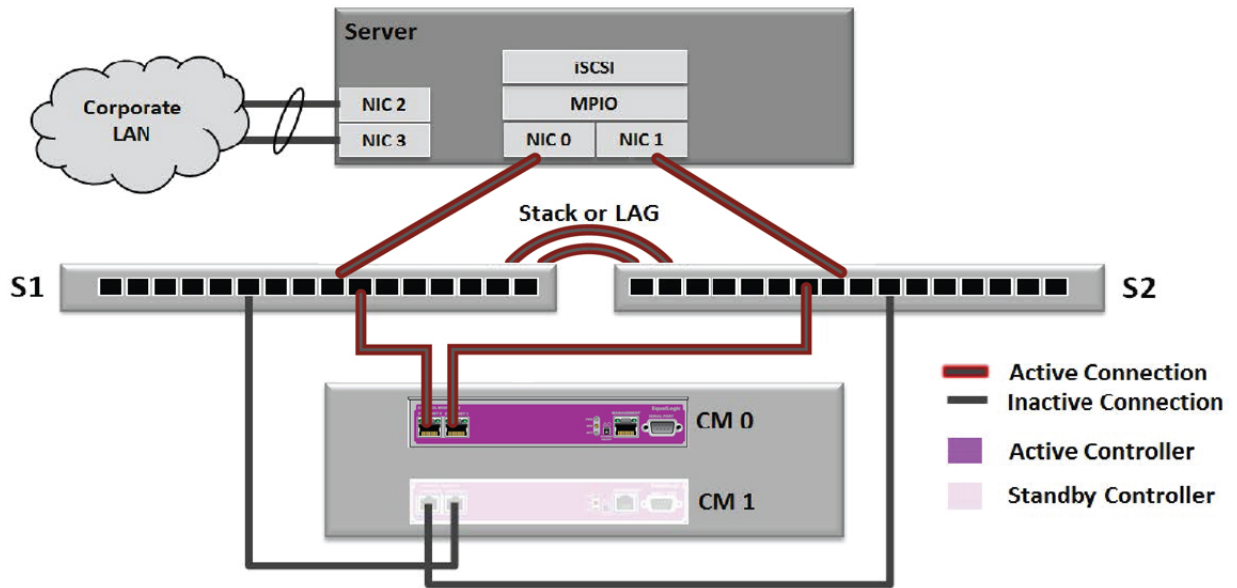


Figure 1. Redundant SAN PS4100

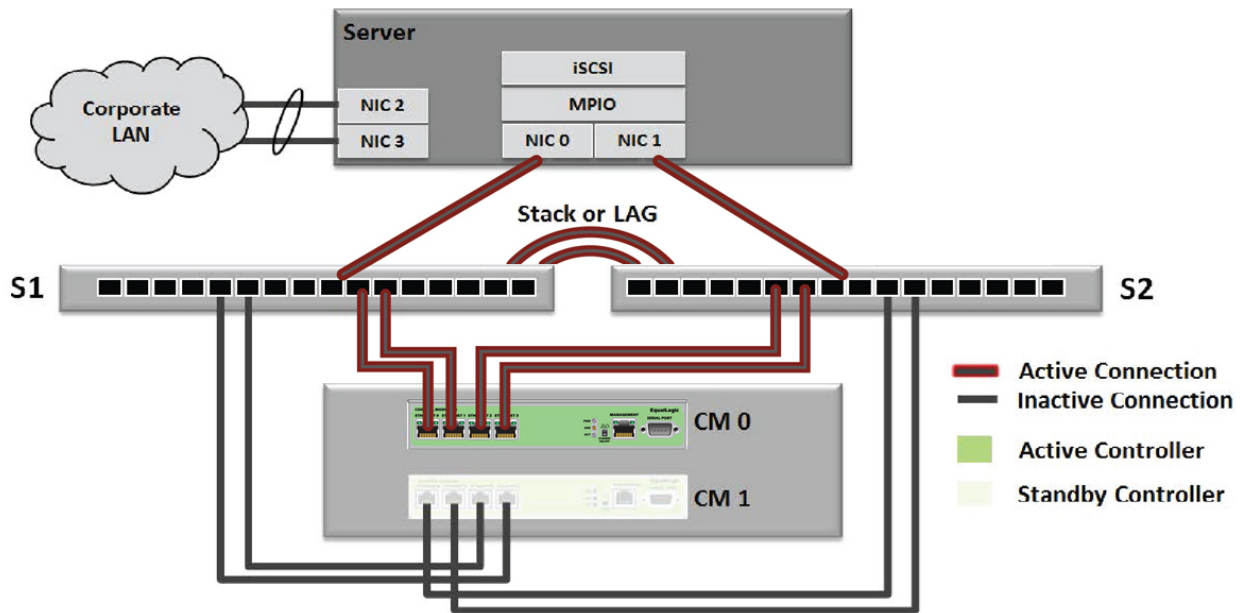


Figure 2. Redundant SAN PS6100

- Both link aggregation group (LAG) or stacking interconnects can be used as the bridging mechanism between redundant switches in the EqualLogic SAN, but each method has a different set of considerations that might make one approach more preferable than the other. For more information about LAG and stacking considerations, see the *PS Series Configuration Guide* at: <http://en.community.dell.com/techcenter/storage/w/wiki/equallogic-configuration-guide>
- Any EqualLogic SAN group that is required to send or receive replication traffic to or from another SAN group must have an uninterrupted communications path (such as “visibility”) between each group.

See Dell’s switch configuration guides for specific switch vendor and model settings and best practices. Examples and detailed instructions for configuring SAN network switches and hosts can be obtained through the Rapid EqualLogic Configuration Portal or by reviewing sample SAN reference architectures on the Dell TechCenter website:



FS Series Switch Requirements

- Keep the client and SAN side networks physically separate, and deploy two switches on each side to protect against a switch failure.
- Do not block IPv6 traffic on your SAN/internal switches if you have FS series appliances attached.

Dell Networking N-Series Switching MPIO With M-LAG Advisory

For EqualLogic SAN designs, both switches must be in the same network broadcast domain and thus must have a common data plane. Every array port must be able to talk to every other array port in the SAN group. M-LAG does not provide this functionality and is thus not suitable for use as a design component. Continue to use either switch stacking or standard link aggregation group (LAG) constructs to interconnect switches.

Hardware Initiators

Most hardware initiators are referred to as converged network adapters (CNAs). CNAs support a variety of enhancements such as iSCSI Offload and DCB, making them more than just a hardware initiator. Of the major CNAs supported by EqualLogic, the most common ones are QLogic, Broadcom, Emulex, and Intel.



Configure Your Microsoft Windows Operating System


This section describes how to configure settings on your Microsoft Windows servers to ensure high availability.

Procedure Summary

To ensure high availability, configure the settings on your Microsoft Windows servers as follows. If you have installed the Host Integration Tools for Microsoft, skip step 3.

1. Install the recommended Microsoft hotfixes described in [1. Install the Recommended Microsoft Hotfixes](#).
2. Install the Host Integration Tools for Microsoft and proceed to step 4. This kit can be downloaded from eglsupport.dell.com.
3. (Optional) Set the registry entries if the Host Integration Tools for Microsoft kit is not installed. See [3. Set the Registry Entries](#).
4. Check the Microsoft guidelines. See [4. Verify Microsoft Windows Guidelines](#).

1. Install the Recommended Microsoft Hotfixes

 **NOTE:** Not all hotfixes are necessary for all versions of Windows. Dell recommends that you evaluate each fix to determine whether it applies to your environment, then install the latest Microsoft hotfixes that apply.

[Table 1. Recommended Microsoft Hotfixes](#) lists the hotfixes recommended by Microsoft.

Table 1. Recommended Microsoft Hotfixes

Hotfix/Link	Description of Issue
Windows 2012	
social.technet.microsoft.com/wiki/contents/articles/15576-hyper-v-update-list-for-windows-server-2012.aspx	List of Hyper-V hotfixes for Windows Server 2012
support.microsoft.com/kb/2784261	List of hotfixes and updates for Windows Server 2012-based failover clusters
support.microsoft.com/kb/2869606	LUN for a storage device is lost when multiple paths are lost concurrently in Windows
support.microsoft.com/kb/2871085	I/O failures occur when you create a snapshot for a large volume
support.microsoft.com/kb/2957560	Windows freezes or loses cluster disk reservation during a stress test in iSCSI initiator
support.microsoft.com/kb/2965750	Performance issues occur over a high-latency network if the Windows scale option is removed from retransmitted SYN packets
Windows 2016	

Hotfix/Link	Description of Issue
http://support.microsoft.com/kb/4056890	Provides a listing of the Windows 2016 hotfixes and recommended Microsoft updates.

To obtain update rollups from Microsoft, go to update.microsoft.com/.

Refer to support.microsoft.com/kb/119591 for information about downloading driver software updates and other support files.

2. Install the Host Integration Tools for Microsoft


Dell has automated a portion of this process in the EqualLogic Host Integration Tools for Microsoft (HIT/Microsoft) kit. You can download this kit from the eqlsupport.dell.com/support website, and install it on supported operating systems.

Dell EqualLogic Host Integration Tools for Microsoft provide integrated software such as MPIO modules, enhanced data protection, automated provisioning, and alerting. If installed, this software implements changes that follow Dell's recommended best practices for several storage-related configuration settings in the registry.

3. Set the Registry Entries

 **NOTE:** If you have not installed the Host Integration Tools for Microsoft, modify or create the registry entry if it does not exist. Host Integration Tools for Microsoft automatically set these values.

Use the Registry Editor Carefully

 **CAUTION:** To configure the settings recommended in this section, use the Registry Editor. Using the Registry Editor incorrectly might cause serious problems that require reinstallation of the operating system. Dell and Microsoft cannot guarantee that you will be able to solve problems that result from using the Registry Editor incorrectly.

[Table 2. Registry Parameter Values](#) summarizes the keys to change and their recommended values.

For the parameters in the table that refer to an <Instance Number>, at least one subkey is named 0000, or 0001, or 0002, and so on. To identify which <Instance Number> subkey to modify, expand each subkey until you find one that has a subkey called Parameters.

 **NOTE:** For all the registry values in this table, the type is DWORD.

Table 2. Registry Parameter Values

Parameter	Recommended Value (Decimal)
All Windows Servers	
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\Class\{4d36e97b-e325-11ce-bfc1-08002be10318}\<Instance Number>\Parameters\DelayBetweenReconnect	1
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\Class\{4d36e97b-e325-11ce-bfc1-08002be10318}\<Instance Number>\Parameters\EnableNOPOut	1
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\Class\{4d36e97b-e325-11ce-bfc1-08002be10318}\<Instance Number>\Parameters\ErrorRecoveryLevel\	2
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\Class\{4d36e97b-e325-11ce-bfc1-08002be10318}\<Instance Number>\Parameters\FirstBurstLength	65536



Parameter	Recommended Value (Decimal)
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\Class\{4d36e97b-e325-11ce-bfc1-08002be10318}\<Instance Number>\Parameters\ImmediateData	1
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\Class\{4d36e97b-e325-11ce-bfc1-08002be10318}\<Instance Number>\Parameters\InitialR2T	0
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\Class\{4d36e97b-e325-11ce-bfc1-08002be10318}\<Instance Number>\Parameters\IPSecConfigTimeout\	60
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\Class\{4D36E97B-E325-11CE-BFC1-08002BE10318}\<Instance Number>\Parameters\Link DownTime	35
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\Class\{4d36e97b-e325-11ce-bfc1-08002be10318}\<Instance Number>\Parameters\MaxBurstLength	262144
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\Class\{4d36e97b-e325-11ce-bfc1-08002be10318}\<Instance Number>\Parameters\MaxConnectionRetries	4294967295
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\Class\{4d36e97b-e325-11ce-bfc1-08002be10318}\<Instance Number>\Parameters\MaxPendingRequests	255
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\Class\{4d36e97b-e325-11ce-bfc1-08002be10318}\<Instance Number>\Parameters\MaxRecvDataSegmentLength\	65536
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\Class\{4d36e97b-e325-11ce-bfc1-08002be10318}\<Instance Number>\Parameters\MaxRequestHoldTime	60 (if using MPIO) 300 (non-MPIO)
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\Class\{4d36e97b-e325-11ce-bfc1-08002be10318}\<Instance Number>\Parameters\MaxTransferLength\	262144
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\Class\{4d36e97b-e325-11ce-bfc1-08002be10318}\<Instance Number>\Parameters\NetworkReadyRetryCount\	10
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\Class\{4d36e97b-e325-11ce-bfc1-08002be10318}\<Instance Number>\Parameters\PortalRetryCount\	5
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\Class\{4d36e97b-e325-11ce-bfc1-08002be10318}\<Instance Number>\Parameters\SrbTimeoutDelta\	15
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\Class\{4d36e97b-e325-11ce-bfc1-08002be10318}\<Instance Number>\Parameters\TCPConnectTime\	15
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\Class\{4d36e97b-e325-11ce-bfc1-08002be10318}\<Instance Number>\Parameters\TCPDisconnectTime\	15



Parameter	Recommended Value (Decimal)
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Control\Class\{4d36e97b-e325-11ce-bfc1-08002be10318}\<Instance Number>\Parameters\WMIRequestTimeout\	30
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\Disk\TimeOutValue	60
Only Windows Servers with MPIO	
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\iScsiPrt\Parameters\IoTimeoutValue\	60 (if using MPIO) 300 (non-MPIO)
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\mpio\Parameters\AdditionalCriticalWorkerThreads	0
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\mpio\Parameters\AdditionalDelayedWorkerThreads	0
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\mpio\Parameters\DiskPathCheckDisabled	0
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\mpio\Parameters\DiskPathCheckEnabled	1
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\mpio\Parameters\DiskPathCheckInterval\	15
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\mpio\Parameters\FlushHealthInterval\	1440
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\mpio\Parameters\GatherHealthStats\	1
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\mpio\Parameters\LinkDownTime\	35
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\mpio\Parameters\MaxRequestHoldTime\	300
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\mpio\Parameters\PathRecoveryInterval\	60
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\mpio\Parameters\PathVerificationPeriod\	20
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\mpio\Parameters\PathVerifyEnabled\	0
HKEY_LOCAL_MACHINE\System\Current ControlSet\Services\mpio\Parameters\PDORemovePeriod	240
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\mpio\Parameters\RetryCount\	3
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\mpio\Parameters\RetryInterval\	1
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\mpio\Parameters\TcpMaxDataRetransmissions\	4
HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\mpio\Parameters\UseCustomPathRecoveryInterval\	1



Parameter	Recommended Value (Decimal)
HKEY_LOCAL_MACHINE\System\Current ControlSet\Services\mpio\Parameters\UseCustomPathRecoveryInterval	1

4. Verify Microsoft Windows Guidelines

Application-Specific Events

Applications that run in an iSCSI environment are sometimes directly affected by logging various events. The following section discusses specific applications that might be affected by iSCSI SAN events and settings.

 **NOTE:** You might notice some iSCSI SAN messages in the event log. These events have no impact on operations.

Microsoft Windows Applications

While most of the applications running in Windows rely on the OS and the software initiator to manage and track storage access, a few applications inform or log occurring I/O events or take independent actions based on the availability of the storage subsystem in use.

You can expect some events to be tracked in your Windows systems while access to the back-end SAN has been temporarily delayed due to a failover event.

Microsoft Exchange Server

The Exchange Server database engine (ESE) reports every I/O that requires more than 60 seconds to the event log, and as of Exchange Server version 2010 takes additional actions as shown in [Table 3. I/Os Requiring More Than 60 Seconds](#).

Table 3. I/Os Requiring More Than 60 Seconds

Exchange Version	I/O Type	I/O Times	Behavior
Exchange Server 2010 RTM	Completed	>60 seconds	Write to event log; ESE performs clean-page overwrite on pages affected by slow I/O
Exchange Server 2010 SP1	In flight	>60 seconds	Write to event log
		>4 minutes	Terminate wininit.exe process and bugcheck the server
	Completed	>30 seconds	Write to event log; ESE performs clean-page overwrite on pages affected by slow I/O

For more information about this behavior, refer to Microsoft Technet at:

technet.microsoft.com/en-us/library/ff625233.aspx

Set Delay Between Login Attempts

The `DelayBetweenReconnect` parameter sets the delay between each attempt to log in again. The default value is 5 seconds, but Microsoft recommends setting it to 1 second.



Enable ICMP Echo Requests for iSCSI Initiators in Windows

By default, Windows Firewall does not allow an incoming ICMP Echo message (also known as an ICMP Echo Request message). Dell recommends changing Firewall settings to allow ICMP echo for ICMPv4, or for ICMPv6 if using an IPv6 initiator. See the Windows Firewall documentation for instructions on creating inbound custom rules to allow ICMPv4 and ICMPv6 Echo Request packets.



Configure Your VMware Environment

This section describes how to configure settings on your VMware products, including vSphere and ESX hosts, to ensure high availability. For more detailed information, see Tech Report TR1091, *Best Practices for Implementing VMware vSphere in a Dell EqualLogic PS Series SAN Environment*. This tech report is available from:

http://en.community.dell.com/techcenter/extras/m/white_papers/20434601/download.aspx

Set iSCSI Login Timeout Value on vSphere 5.1 and 5.0

Under vSphere 5.x, VMware allocates 5 seconds for an iSCSI session to log in to an iSCSI target. In a normally healthy environment, this period of time is sufficient.

However, this 5-second timeout is not long enough to process the large number of iSCSI login requests that occur after a network failure, switch reboot, or controller failover. Dell recommends that the iSCSI **Login Timeout** value be increased to 60 seconds.

You can set the iSCSI **Login Timeout** value using either of the following methods.

- From the vSphere Client GUI:
 - a. Click **Configuration** → **Storage Adapters**.
 - b. Select the storage adapter that you want to modify, then click **Properties**.
 - c. On the iSCSI Initiator **Properties** page, click **Advanced**.
 - d. On the **Advanced Settings** page, enter **60** in the **Login Timeout** field and click **OK**.
- From the ESXi CLI:
 - a. Enter the following command:

```
esxcli iscsi adapter param set --adapter=<vmhbaX> --key=LoginTimeout --value=60
```
 - b. Reboot the server for the changes to take effect.

NOTE:

- For dependent HBAs, such as the Broadcom NetXtreme II 57711E, the iSCSI **Login Timeout** value must be updated for each HBA configured for iSCSI.
- For vSphere 5.0 VMware, patch ESXi500-201112001 is required for the iSCSI **Login Timeout** value to be editable.

Storage Heartbeat on vSphere 5.0, 4.1, and 4.0

In the VMware virtual networking model, certain types of vmkernel network traffic are sent out on a default VMkernel port for each subnet. The iSCSI multipathing network configuration requires that the iSCSI VMkernel ports use a single physical NIC as an uplink. If the physical NIC that is being used as the uplink for the default VMkernel port goes down, network traffic that is using the default VMkernel port will fail. This traffic includes vMotion traffic, SSH access, and ICMP ping replies.

Though iSCSI traffic is not directly affected by this condition, a side effect of the suppressed ping replies is that the EqualLogic PS Series group will not be able to accurately determine connectivity during the login process, and therefore a suboptimal placement of iSCSI sessions can occur. In some scenarios, depending on array, server, and network load, logins might not be completed in a timely manner. To prevent this issue from occurring, create a highly available VMkernel port on the iSCSI subnet serving as the default VMkernel port for such outgoing traffic.

 **NOTE:** This recommendation for using Storage Heartbeat applies to vSphere 4.1 and 5.0 only.

[Table 4. Storage Heartbeat Summary](#) lists a summary of heartbeats.

Table 4. Storage Heartbeat Summary

ESX/ESXi Version	iSCSI Login Timeout Value	Storage Heartbeat	Multipathing Policy
ESX/ESXi 4.0	Not applicable ¹	Yes	RR
ESX/ESXi 4.1	Not applicable ¹	Yes	MEM ⁴ , RR
ESX/ESXi 4.1 update 3	60	Yes	MEM ⁴ , RR
ESXi 5.0	60 ²	Yes	MEM ⁴ , RR
ESXi 5.1	60	No ³	MEM ⁴ , RR
ESXi 5.5	60	No ³	MEM ⁴ , RR

1. Under ESXi 4.0 and 4.1, the iSCSI **Login Timeout** value is set to 15 seconds and cannot be changed.
2. VMware patch ESXi500-201112001 is required to enable editing of the iSCSI **Login Timeout** value.
3. Due to improvements in later ESXi releases, it is no longer necessary to implement the Storage Heartbeat VMkernel port.
4. The Dell EqualLogic Multipathing Extension Module for vSphere (MEM) depends on the vSphere Storage APIs for Multipathing, which are available only under Enterprise or Enterprise Plus licensing.



Configure Your Linux Operating System

This section describes how to configure recommended settings on your Linux hosts to ensure high availability. It also shows how the Host Integration Tools for Linux (HIT/Linux) configure these recommended settings when installed on your Linux hosts.

Configure SCSI Device Settings Using the udev Facility

The Linux udev utility can be used to dynamically modify device settings every time a new device is added or changed. The following examples show ways that the HIT/Linux kit modifies the device settings with udev rules.

Increase SCSI Device Timeout Value to Avoid Disk I/O Timeouts

The default disk timeout should be increased to 60 seconds. The specific udev rule implemented by HIT/Linux is installed in the file named `/lib/udev/rules.d/99-eqlsd.rules`.

```
ACTION!="remove", SUBSYSTEM=="block", ENV{ID_VENDOR}=="EQLOGIC", RUN+="/bin/sh -
c `echo 60 > /sys/%p/device/timeout`"
```

Increase SCSI Device Read-Ahead Buffer Size

The default SCSI device read-ahead buffer setting of 128KB is increased to 1024KB. The specific udev rule implemented by HIT/Linux is installed in the file named `/lib/udev/rules.d/38-equallogic.rules`.

```
ATTR{queue/read_ahead_kb}="1024"
```

Change SCSI Device I/O Scheduler Algorithm

The default SCSI device I/O scheduler algorithm of CFQ (Completely Fair Queuing) is set to noop. The specific udev rule implemented by HIT/Linux is installed in the file named `/lib/udev/rules.d/38-equallogic.rules`.

```
ATTR{queue/scheduler}="noop"
```

Configure iSCSI Settings Using the `/etc/iscsi/iscsid.conf` File

The following changes should be made in the Linux iSCSI configuration file `/etc/iscsi/iscsid.conf`. Any settings in this file are propagated to iSCSI targets on discovery or rediscovery.

Change FastAbort Setting in iSCSI Configuration File

The default iSCSI initiator setting for FastAbort is set to Yes. EqualLogic targets prefer that this iSCSI configuration parameter be set to No. The HIT/Linux software automatically implements this change.

```
node.session.iscsi.FastAbort = No
```

Change `node.startup` Setting to Avoid Login Storms

The default iSCSI initiator setting for `node.startup` is set to automatic, which means that all previously discovered iSCSI targets will be logged in to at system boot. This behavior can cause a surge of iSCSI login activity during system reboots,

which can place a resource burden on both the EqualLogic arrays and the Linux host. Dell recommends that you change the setting to manual. The HIT/Linux eqltune utility recommends this change, but does not automatically implement it.

```
node.startup = manual
```

Configure Linux Networking

Dell recommends that you configure multipathing for Ethernet interfaces on a Linux host to connect to an EqualLogic group.

This configuration of multiple NICs per subnet requires that several Linux networking parameters be modified from their default values for reliable implementation. The HIT/Linux eqltune utility, when run in fix mode, will automatically configure the Linux operating system to operate correctly with this configuration.

 **NOTE: Apply these commands, changes, and settings to each Ethernet interface (for example: eth1, eth2).**

Disable Generic Receive Offload

The Generic Receive Offload (GRO) feature, implemented by some Linux network drivers, causes poor read performance results. This feature can be disabled (turned off) using the ethtool utility.

```
# ethtool ethX -K gro off
```

Enable Flow Control

Flow control is often negotiated properly, but a network that supports full flow control (RX and TX pause names on) is recommended. If your network infrastructure is set up to support flow control, you could benefit from forcing flow control to be on. The HIT/Linux eqltune utility recommends this change, but does not automatically implement it.

```
# ethtool --pause ethX autoneg off rx on tx on
```

To make this change persistent across reboots, add the command to the `/etc/rc.local` directory.

Enable Jumbo Frames

If supported by your networking infrastructure, enabling jumbo frames (that is, an MTU of 9000) enhances iSCSI throughput under heavy loads. The HIT/Linux eqltune utility recommends this change, but does not automatically implement it.

In `/etc/sysconfig/networking/devices/ifcfg-ethX`, add this line:

```
MTU=9000
```

Avoid ARP Flux

The default value of 0 (zero) allows NICs to reply to ARP requests on any other NIC, which can result in false positives to externally originating ping tests, as well as unexpected ARP resets. This condition is known as ARP Flux. To avoid ARP Flux in a configuration with multiple NICs per subnet, apply the following parameters in `/etc/sysctl.conf`:

```
net.ipv4.conf.ethX.arp_ignore = 1  
net.ipv4.conf.ethX.arp_announce = 2
```

Loosen IPv4 Reverse Path Filtering

To allow multiple iSCSI connections in a configuration with multiple NICs per subnet, reverse path filtering must be loosened. To loosen the filtering, use the following setting:

```
net.ipv4.conf.ethX.rp_filter=2
```

 **NOTE: This setting is required only on RHEL version 6 and version 7 and on newer ubuntu and debian kernels.**



Summary of Recommended Changes

[Table 5. Summary of Recommended Changes](#) summarizes the changes to Linux settings that are recommended by Dell.

Table 5. Summary of Recommended Changes

Area	Setting	Default Value	Recommended Value (EQL)
SCSI block devices	Device timeout	30 seconds	60 seconds
Per SCSI device	read-ahead buffer	128KB	1024KB
Per SCSI device	I/O scheduler	CFQ	noop
iSCSI settings	node.session.iscsi.FastAbort	Yes	No
iSCSI settings	node.startup	automatic	manual
Ethernet devices	Generic Receive Offload	on	off
Ethernet devices ¹	flow control ¹	auto-negotiate	on
Ethernet devices ¹	jumbo frames ¹ (MTU size)	1500	9000
Sysctl tunable - ARP	net.ipv4.conf.ethX.arp_ignore	0	1
Sysctl tunable - ARP	net.ipv4.conf.ethX.arp_announce	0	2
Sysctl tunable - ARP ²	net.ipv4.conf.ethX.arp_filter	0	2

¹ Verify that the network infrastructure supports these two settings (flow control and jumbo frames) before implementing them. To enable pause auto-negotiation, use the `ethtool -p ethX autoneg off` command.

² This setting is required only on RHEL version 6 and version 7 and on newer ubuntu and debian kernels.

[Table 6. Optional Changes](#) lists other changes to Linux settings that Dell suggests should be applied to `/etc/sysctl.conf`.

Table 6. Optional Changes

Area	Setting	Recommended Value (EQL)
Sysctl tunable	net.core.rmem_max	16777216
Sysctl tunable	net.core.wmem_max	16777216
Sysctl tunable	net.ipv4.tcp_rmem	8192 87380 16777216
Sysctl tunable	net.ipv4.tcp_wmem	4096 65536 16777216
Sysctl tunable	net.core.wmem_default	262144
Sysctl tunable	net.core.rmem_default	262144
iSCSI settings	node.sessions.cmd_max	1024
iSCSI settings	node.session.queue_depth	128
iSCSI settings	iscsi.MaxRecvDataSegmentLength	131072



Configure Your Solaris 10 Operating System

You can use the default initiator configuration for Oracle Solaris 10 operating systems. No changes are required. However, you might experience the following issue.

Zero-Length Volumes

By default, the Solaris 10 initiator automatically connects to all volumes that it discovers. If it connects to a zero-length volume (such as the vss-control volume), it panics. To avoid this issue:

- Make sure the vss-control volume on the group has access control records that are configured to prevent the Solaris initiator from discovering it.
- Make sure the volume does not allow unrestricted access.

It is also possible that, after the EqualLogic Remote Setup Wizard is run, the vss-control ACL uses CHAP. Using the CHAP discovery filter would also prevent Solaris from discovering it. You can use the EqualLogic CLI command **grpparams discovery-use-chap disable** to disable filtering based on CHAP.



Configure Your IBM AIX Operating System

IBM AIX systems might experience the following issue.

iSCSI Initiators Lose Volume Connectivity

In some cases, after a load-balancing operation completes, AIX iSCSI initiators lose volume connectivity. This issue can be corrected by applying a patch to your AIX system. For more information, contact your IBM customer support provider and reference incident number IZ90885, or visit:

ibm.com/support/docview.wss?uid=isg1IZ90885&crawler=1



NOTE: As a short-term workaround, before applying the patch, disable the PS Series group's load-balancing feature.