



VMware ESXi NIC Optimization and Best Practices with EqualLogic SAN

A Dell Reference Architecture

Dell Storage Engineering
January 2014

Revisions

Date	Description
January 2014	Initial release

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.

© 2013 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.

PRODUCT WARRANTIES APPLICABLE TO THE DELL PRODUCTS DESCRIBED IN THIS DOCUMENT MAY BE FOUND AT: <http://www.dell.com/learn/us/en/19/terms-of-sale-commercial-and-public-sector> Performance of network reference architectures discussed in this document may vary with differing deployment conditions, network loads, and the like. Third party products may be included in reference architectures for the convenience of the reader. Inclusion of such third party products does not necessarily constitute Dell's recommendation of those products. Please consult your Dell representative for additional information.

Trademarks used in this text:

Dell™, the Dell logo, Dell Boomi™, Dell Precision™, OptiPlex™, Latitude™, PowerEdge™, PowerVault™, PowerConnect™, OpenManage™, EqualLogic™, Compellent™, KACE™, FlexAddress™, Force10™ and Vostro™ are trademarks of Dell Inc. Other Dell trademarks may be used in this document. Cisco Nexus®, Cisco MDS®, Cisco NX-OS®, and other Cisco Catalyst® are registered trademarks of Cisco System Inc. EMC VNX®, and EMC Unisphere® are registered trademarks of EMC Corporation. Intel®, Pentium®, Xeon®, Core® and Celeron® are registered trademarks of Intel Corporation in the U.S. and other countries. AMD® is a registered trademark and AMD Opteron™, AMD Phenom™ and AMD Sempron™ are trademarks of Advanced Micro Devices, Inc. Microsoft®, Windows®, Windows Server®, Internet Explorer®, MS-DOS®, Windows Vista® and Active Directory® are either trademarks or registered trademarks of Microsoft Corporation in the United States and/or other countries. Red Hat® and Red Hat® Enterprise Linux® are registered trademarks of Red Hat, Inc. in the United States and/or other countries. Novell® and SUSE® are registered trademarks of Novell Inc. in the United States and other countries. Oracle® is a registered trademark of Oracle Corporation and/or its affiliates. Citrix®, Xen®, XenServer® and XenMotion® are either registered trademarks or trademarks of Citrix Systems, Inc. in the United States and/or other countries. VMware®, Virtual SMP®, vMotion®, vCenter® and vSphere® are registered trademarks or trademarks of VMware, Inc. in the United States or other countries. IBM® is a registered trademark of International Business Machines Corporation. Broadcom® and NetXtreme® are registered trademarks of Broadcom Corporation. QLogic is a registered trademark of QLogic Corporation. Other trademarks and trade names may be used in this document to refer to either the entities claiming the marks and/or names or their products and are the property of their respective owners. Dell disclaims proprietary interest in the marks and names of others.



Table of contents

Revisions.....	2
Acknowledgements.....	5
Feedback	5
Executive summary	5
1 Introduction.....	6
1.1 Audience.....	6
2 Technical overview.....	7
3 Test configurations and methodology	8
3.1 Base SAN configuration	8
3.2 I/O execution and evaluation	10
3.3 Test case sequence	10
3.3.1 Test case sequence	11
4 Results and analysis.....	12
4.1 Baseline testing.....	12
4.1.1 Jumbo Frames.....	12
4.1.2 Flow control.....	13
4.1.3 Delayed ACK algorithm.....	14
4.2 Other available performance tuning options.....	17
4.2.1 Large Receive Offload	17
4.2.2 TCP Checksum Offloading.....	19
4.3 Broadcom BCM57810 iSOE mode performance results	22
5 Best practice recommendations.....	23
5.1 Broadcom BCM57810 Standard mode recommended configuration.....	23
5.2 Broadcom BCM57810 iSOE mode recommended configuration	24
5.3 Intel X520 recommended configuration	24
6 Conclusion.....	25
A Test configuration details.....	26
B Network adapter and TCP stack configuration details	27
B.1 Broadcom BCM57810 adapter options standard mode	27
B.2 Configuring Broadcom BCM57810 adapter parameters in standard mode	28
B.3 Broadcom BCM57810 adapter options iSOE mode	28



B.4	Configuring Broadcom BCM57810 adapter parameters in iSOE mode	29
B.5	Intel X520 adapter options	29
B.6	Configuring Intel X520 adapter properties.....	29
C	I/O parameters	30
D	Installing and configuring the EqualLogic MEM.....	31
	Additional resources	33



Acknowledgements

This best practice white paper was produced by the following members of the Dell Storage team:

Engineering: Pavel Viltres

Technical Marketing: Guy Westbrook

Editing: Camille Daily

Additional contributors: Mike Kosacek, Clay Cooper, Donald Williams, and David Sevene

Feedback

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



SISfeedback@Dell.com

Executive summary

This reference architecture explores the configuration options available for improving Dell EqualLogic PS Series SAN performance using the Broadcom BCM57810 or Intel X520 10 GbE network adapters and VMware ESXi 5.5 on a Dell PowerEdge 12th generation server. Recommended operating system and NIC configurations are given based on the results of SAN performance testing.



1 Introduction

Dell EqualLogic PS Series arrays provide a storage solution that delivers the benefits of consolidated networked storage in a self-managing iSCSI storage area network (SAN) that is affordable and easy to use, regardless of scale.

In every iSCSI SAN environment, numerous storage host configuration options can have an effect on overall SAN performance. These effects can vary based on the size and available bandwidth of the SAN, the host/storage port ratio, the amount of network congestion, the I/O workload profile and the overall utilization of system resources at the storage host. One setting might greatly improve SAN performance for a large block sequential workload yet have an insignificant or slightly negative effect on a small block random workload. Another setting might improve SAN performance at the expense of host CPU utilization.

This technical paper quantifies the effect on iSCSI bandwidth and throughput caused by managing several configuration options within the Broadcom and Intel 10 GbE adapter properties and the VMware ESXi 5.5 TCP stack while executing several common SAN workloads. From the results, recommended configurations for an EqualLogic PS Series SAN are given for each tested NIC type.

In order to focus on the pure SAN performance benefits of the tested configuration options, Data Center Bridging (DCB) and Broadcom Switch Independent Partitioning, also known as NIC Partitioning (NPAR) were excluded from the testing.

Note: The performance data in this paper is presented relative to baseline configurations and is not intended to express maximum performance or benchmark results. Actual workload, host to array port ratios, and other factors may also affect performance.

1.1 Audience

This technical white paper is for storage administrators, SAN system designers, storage consultants, or anyone who is tasked with connecting and configuring a host server to EqualLogic PS Series storage 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 TCP/IP and iSCSI standards as defined by the Internet Engineering Task Force (IETF).



2 Technical overview

iSCSI SAN traffic takes place over an Ethernet network and consists of communication between PS Series array member network interfaces and the iSCSI initiator of storage hosts. The Broadcom BCM57810 NetXtreme II and the Intel X520 10 GbE network adapters were used as the iSCSI initiators during this project.

The Broadcom BCM57810 network adapter features iSCSI Offload Engine (iSOE) technology which offloads processing of the iSCSI stack to the network adapter. When using iSOE mode, the network adapter becomes a host bus adapter (HBA) and a host-based, software iSCSI initiator is not utilized. This is as compared to standard (or non-offload) mode in which the network adapter functions as a traditional Ethernet NIC and requires a software iSCSI initiator to connect to iSCSI storage.

The Intel X520 does not offer iSCSI hardware or offload capability and therefore always requires a software iSCSI initiator to connect to an iSCSI storage array.

The following three initiator modes of operation were tested:

- Broadcom BCM57810 standard mode
- Broadcom BCM57810 iSCSI Offload Engine (iSOE) mode
- Intel X520

[Appendix B](#) provides a detailed list of the tested configuration options and default values for each NIC type as well as for the ESXi 5.5 TCP stack.



3 Test configurations and methodology

This section addresses the reasoning behind SAN design decisions and details the SAN configuration. Performance testing methodology, test case sequence, and results analysis are also explained.

3.1 Base SAN configuration

Every effort was made to simplify and optimize the test configurations so that the performance effects of each option could be isolated. The following configuration and design elements helped to achieve this goal.

- Eight RAID 10 volumes within a single storage pool, evenly distributed across array members
- An isolated SAN with no LAN traffic
- DCB was not used
- The NIC bandwidth was not partitioned (NPAR disabled)
- Load balancing (volume page movement) was disabled on the array members

Load balancing is recommended for production environments because it can improve SAN performance over time by optimizing volume data location based on I/O patterns. It was disabled for performance testing to maintain consistent test results. It is enabled by default.

See [Appendix A](#) for hardware and software infrastructure details.

A basic SAN design was chosen with a redundant SAN fabric and an equal number of host and storage ports. Having a 1:1 host/storage port ratio is ideal from a bandwidth perspective. This helped to ensure that optimal I/O rates were achieved during lab testing. Figure 1 shows only the active ports of the PS Series array members.



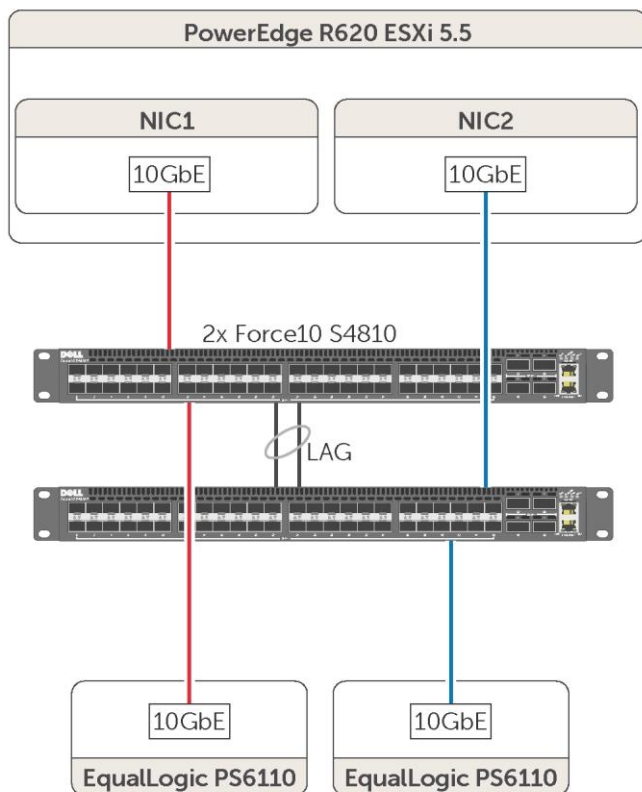


Figure 1 Physical diagram of the base SAN configuration

- Two switches
- Two array members each with a single port
- A single ESXi5.5 host with two 10 GbE NIC ports
- A single Windows 2008 R2 VM with two SCSI controllers connected to two SAN volumes
- A 1:1 storage/host port ratio

See [Appendix A](#) for test configuration details

3.2 I/O execution and evaluation

The host was restarted before each test to confirm configuration changes were in effect. After each boot completed, the even distribution of iSCSI connections across host and storage ports and of active array member ports across SAN switches was confirmed.

To generate the desired workloads for each test, Vdbench was run from the VM's guest OS.

The following Vdbench workloads were run:

- 8 KB transfer size, random I/O, 67% read (8KB random mix)
- 256 KB transfer size, sequential I/O, 100% read (256KB seq read)
- 256 KB transfer size, sequential I/O, 100% write (256KB seq write)
- For specific tests, 64 KB transfer size, random I/O, 67% read, and 128 KB transfer size random I/O, 67% read were also run.

For every test case, each Vdbench workload was run three times for twenty-minute durations. Results presented are an average of the three tests.

Vdbench IOPS results were used to evaluate the 8KB, 64KB, and 128KB random workload performance. Vdbench bandwidth results were used to evaluate 256K sequential workload performance. Host and array member retransmission rates and CPU utilization were also examined.

See [Appendix C](#) for a list of Vdbench parameters.

3.3 Test case sequence

The performance and effect on system resources of adapter and ESXi 5.5 TCP stack options were evaluated using the test cases listed below.

The baseline configuration was created after installing the EqualLogic Multipathing Extension Module for VMware (MEM) version 1.2 on an ESXi host. MEM installation gives the option to apply the EqualLogic Best Practice settings. By default, MEM implements the following changes:

- Jumbo frames enabled (9000 bytes)
- iSCSI initiator Delayed ACK disabled
- iSCSI initiator Advanced Setting LoginTimeout parameter set to 60 seconds (the default value is 5 seconds)

The baseline test configuration also enabled flow control.

Note: For both Broadcom and Intel network adapters Flow Control was enabled by default during testing. However, this default value may change depending on the NIC driver version and NIC type.



The following VMware Knowledge Base article number 1013413 describes how to configure flow control for an ESXi host:

http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=1013413

In this document, **Baseline** refers to the settings resulting after installing MEM 1.2 and applying the EqualLogic Best Practices setting.

Each subsequent test case consisted of a single option that was toggled from the baseline configuration to show its relative effect.

3.3.1 Test case sequence

The following tables show the test case sequence used to evaluate the effect of tested configuration options for the Broadcom BCM57810 Standard mode, BCM57810 iSOE mode, and Intel X520. **Bold text** indicates the changed value for each test scenario.

Table 1 Test case sequence

Test case	Frame size	Flow Control	Adapter settings	ESXi 5.5 iSCSI initiator settings	Comments
1	Jumbo	On	Default	MEM installation best practices	Baseline configuration
2	Standard	On	Default	MEM installation best practices	Standard frames performance effect
3	Jumbo	On	Default	MEM installation best practices, then enabling Delayed ACK	Delayed ACK Enabled performance effect
4	Jumbo	On	Default	MEM installation best practices, then, disabling Large Receive Offload (LRO)	LRO Disabled performance effect
6	Jumbo	On	Ethtool Rx-Checksumming Off	MEM installation best practices	Ethtool Rx-Checksumming Off performance effect

Bold text indicates the changed value for each test scenario.



4 Results and analysis

All test case performance results for each NIC mode and workload combination are presented in this section. For the sake of analysis, a 5% margin of error is assumed and only performance differences greater than this were acknowledged as significant.

4.1 Baseline testing

The following settings were used as a baseline configuration for further testing and evaluation.

4.1.1 Jumbo Frames

Jumbo frames enable Ethernet frames with payloads greater than 1500 bytes. In situations where large packets make up the majority of traffic, jumbo packets can reduce CPU utilization and improve wire efficiency.

Figure 2-Figure 4 show the dramatic effect that utilizing standard frames instead of jumbo frames can have on large block workloads. Significant throughput decreases were observed for both read and write of large block workloads with both the Broadcom and Intel network adapters in all supported operating modes. The small block workload was unaffected.

Dell recommends using jumbo frames (9000 byte packets) with NICs connected to EqualLogic iSCSI storage.

Broadcom 57810 - standard versus jumbo frames

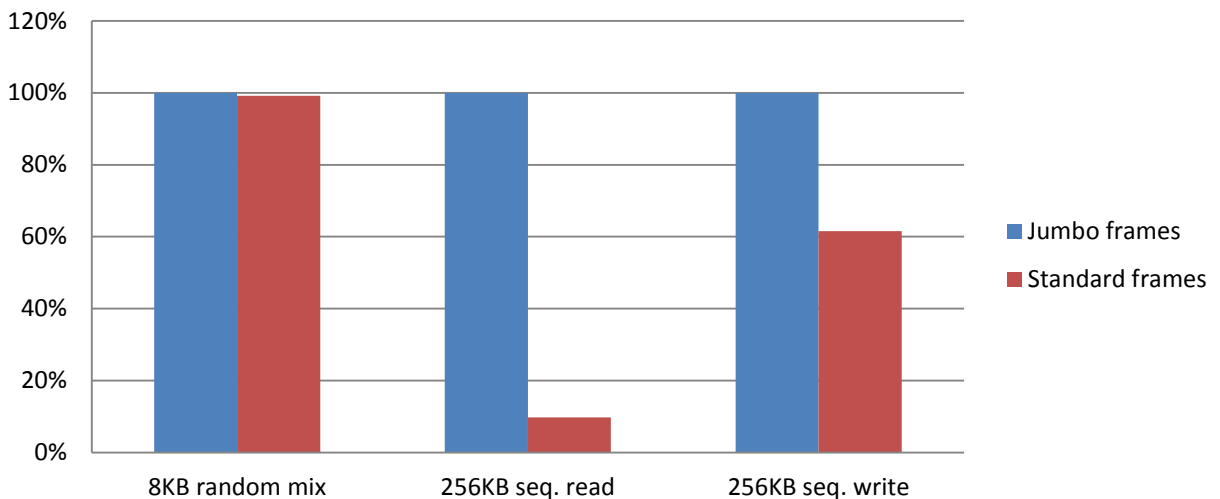


Figure 2 Performance of Jumbo versus Standard frames for Broadcom 57810 in Standard mode

Broadcom 57810 iSOE - standard versus jumbo frames

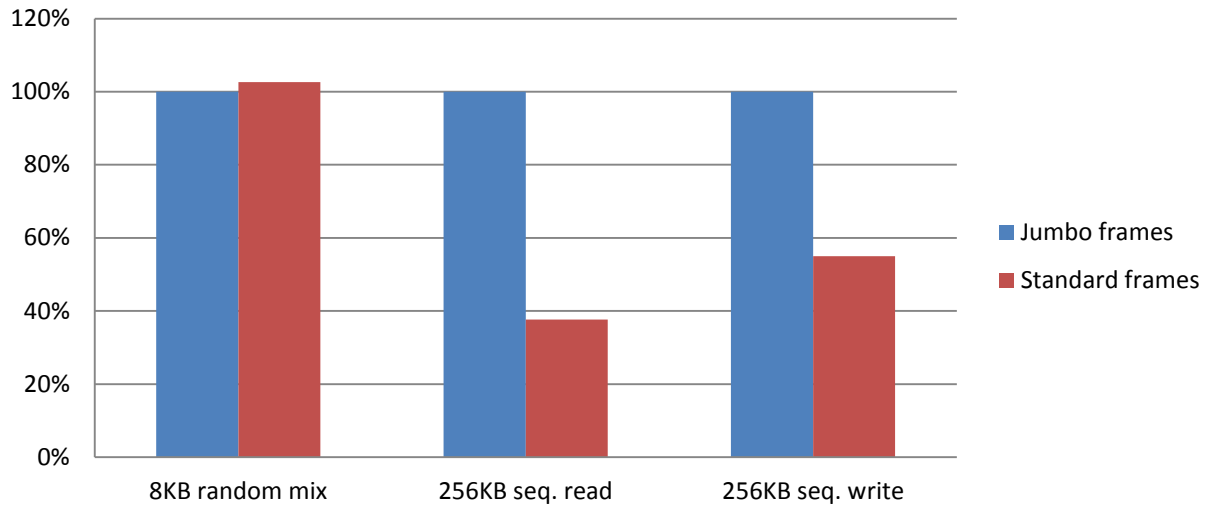


Figure 3 Performance of Jumbo versus Standard frames for Broadcom 57810 in iSOE mode

Intel X520 -- standard versus jumbo frames

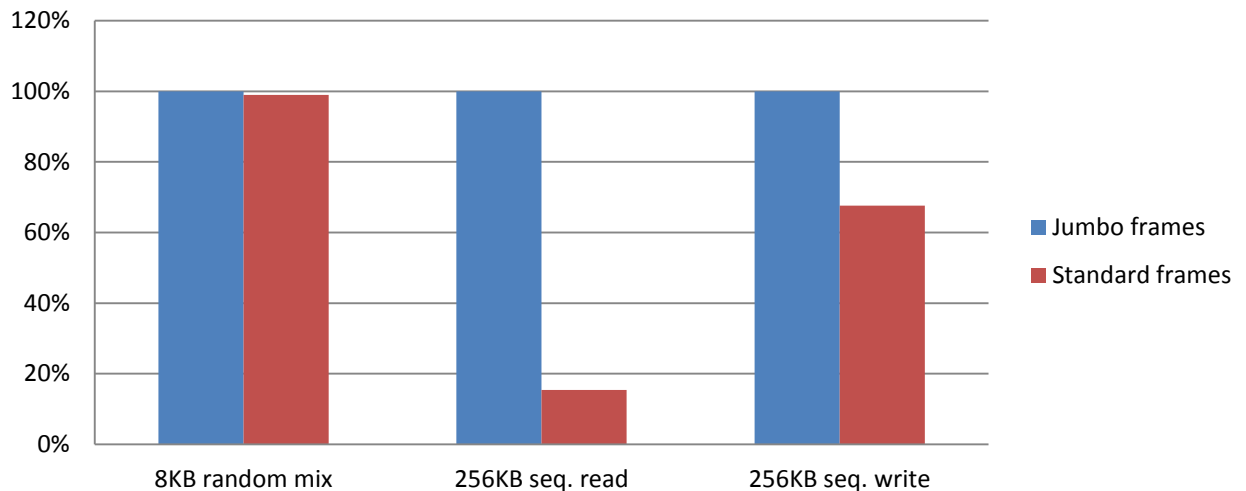


Figure 4 Performance of jumbo versus standard frames for Intel X520

4.1.2 Flow control

Flow control, defined in the IEEE 802.3x standard for Ethernet, is a link-level mechanism that enables the adapter to respond to or to generate flow control (PAUSE) frames, helping to regulate network traffic. It is recommended that flow control be enabled as it is well-known to be of benefit in a congested network environment. The flow control setting for both the Broadcom 57810 and Intel X520 is enabled by default. However, this default value could change depending on NIC driver version and the NIC type used.

For flow control configuration in an ESXi host instructions, refer to the VMware Knowledge Base article number 1013413 at http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=1013413

4.1.3 Delayed ACK algorithm

The delayed ACK algorithm is a technique to improve TCP performance by combining multiple ACK responses into a single response. When installing the EqualLogic Multipathing Extension Module (MEM) for VMware version 1.2 on an ESXi host an option is provided to apply the Dell EqualLogic best practice settings. EqualLogic baseline settings included this option, which disabled Delayed ACK on the iSCSI initiator.

The Delayed ACK setting can also be changed at the iSCSI initiator Advanced Settings. When using the Broadcom 57810 in iSOE mode, it is important to remember that each iSOE offload adapter is an iSCSI initiator and therefore the setting must be applied to each initiator.

For more information on the Delayed ACK concept and its effect on performance see the EqualLogic Tech Report “*Best Practices when implementing VMware vSphere in a Dell EqualLogic PS Series SAN Environment*” at

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

Figures 5-7 illustrate the effect of the Delayed ACK setting on the workloads tested. With both Broadcom and Intel network adapters, and in all NIC supported operating modes, enabling this parameter shows the greatest effect on the performance of large block sequential reads.

Dell recommends disabling Delayed ACK. This is the default setting when MEM 1.2 (or later) is configured to use the Dell Recommended Best Practices. (See [Appendix D](#))



Broadcom 57810 Standard mode - Delayed ACK

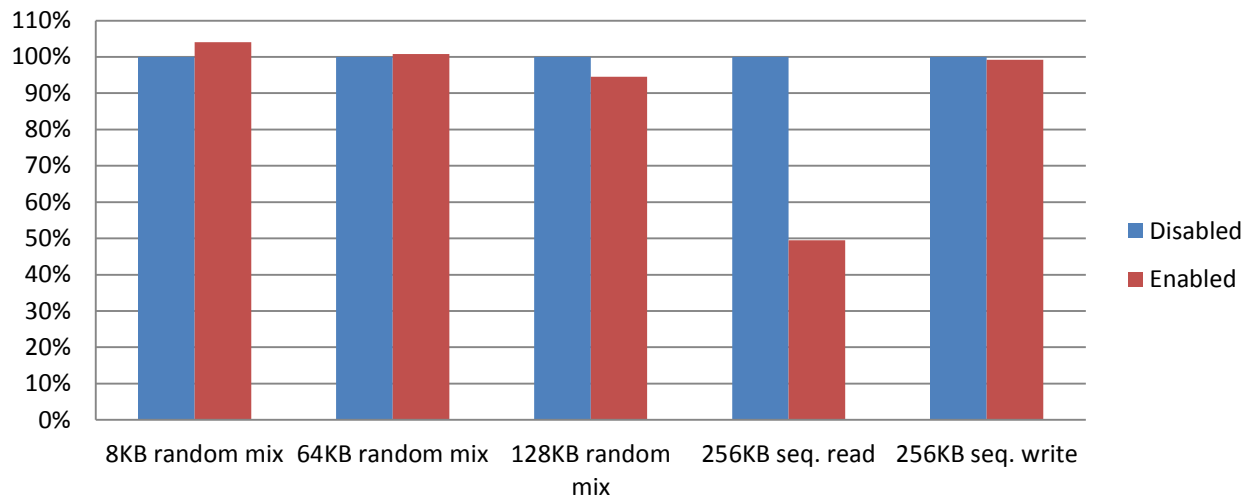


Figure 5 Performance of Delayed ACK configured with a Broadcom 57810 in Standard mode.

Broadcom 57810 iSOE mode - Delayed ACK

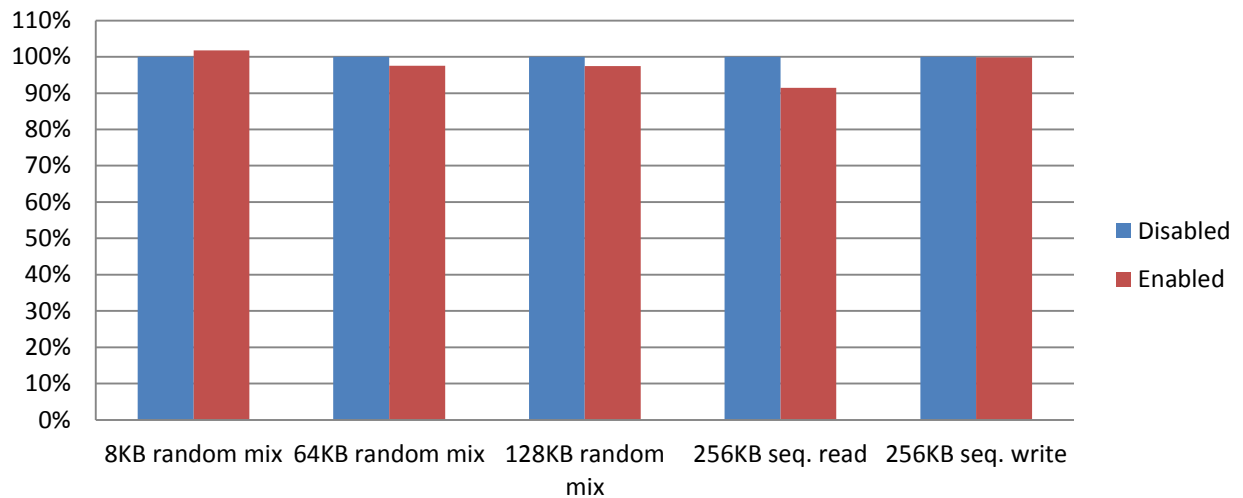


Figure 6 Performance of Delayed ACK configuration with a Broadcom 57810 in iSOE mode

Intel X520 - Delayed ACK

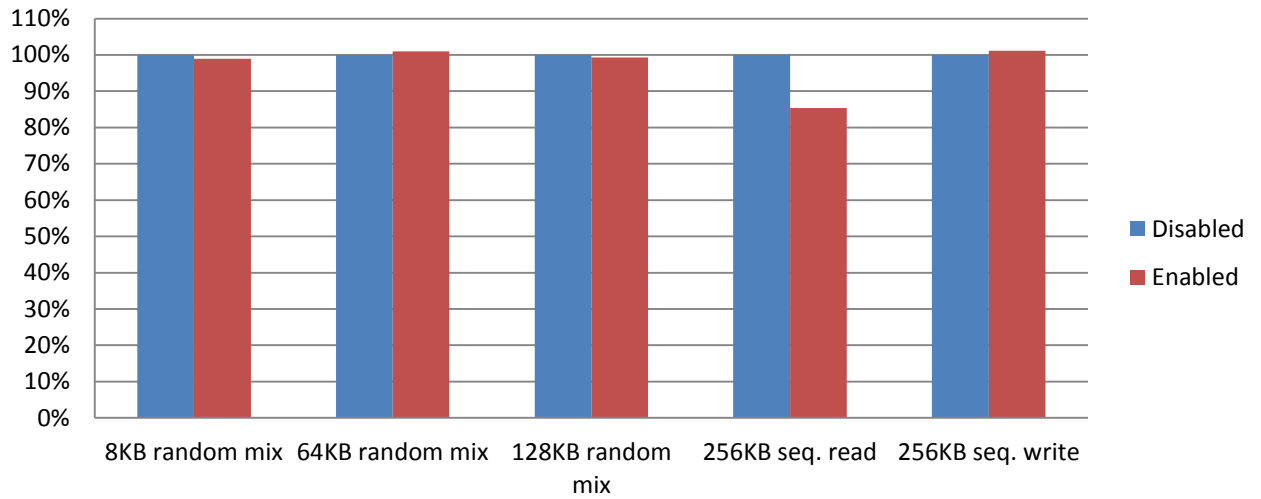


Figure 7 Performance of Delayed ACK configuration with an Intel X520

4.2 Other available performance tuning options

There are other options available for both the network adapter and the ESXi 5.5 TCP stack which can have an effect on SAN performance under specific circumstances. This section explains these additional configuration options and discusses the results of the performance testing. It is important to understand that the performance results described below may not translate to all EqualLogic PS Series SAN environments. The material presented below identifies setting and workload combinations that have a clearly positive or negative impact on iSCSI SAN performance. It is recommended that each potential configuration change be evaluated in the environment prior to implementation.

4.2.1 Large Receive Offload

Similar to Delayed ACK, Large Receive Offload (LRO) works by aggregating packets into a buffer before the received data is sent to the TCP stack. With iSCSI storage, this additional latency inserted into the process could potentially reduce performance for some workloads.

For more information about LRO in ESXi, see VMware Knowledge Base article number 2055140 at http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=2055140

Figures 8-10 illustrate the effect of disabling LRO on the ESXi 5.5 host. With both NICs that were tested, and in all NIC supported operating modes, the effects of this parameter are evident primarily for a large block sequential read I/O workload.

Dell recommends verifying that LRO is enabled (the default) for most environments.

Broadcom 57810 Standard mode - Large Receive Offload

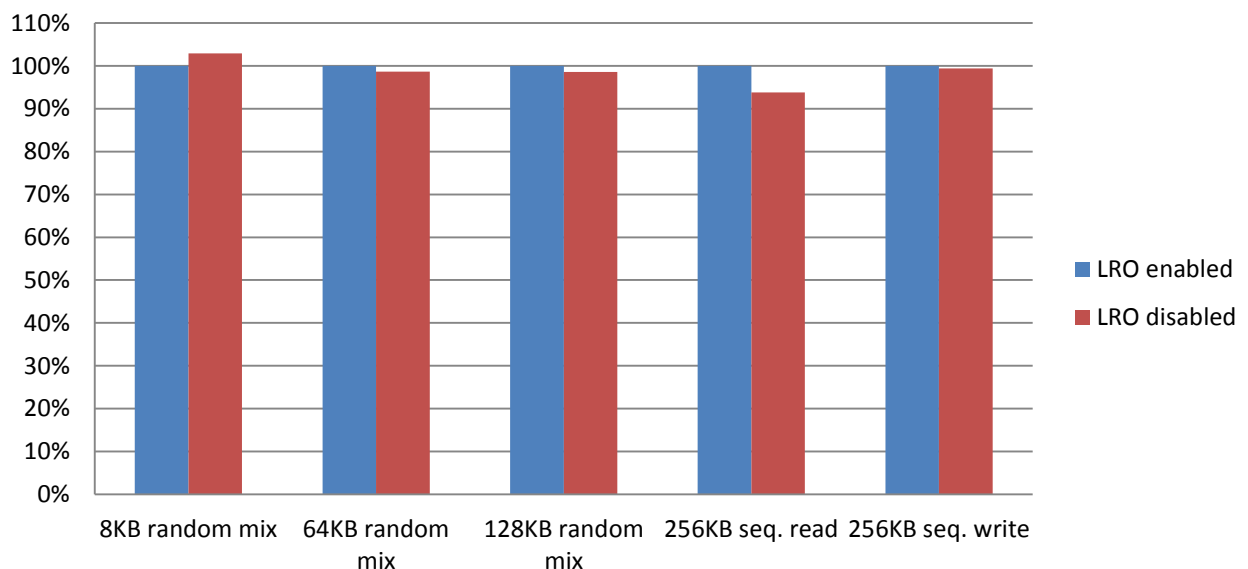


Figure 8 Performance of LRO configuration with a Broadcom 57810 in Standard mode

Broadcom 57810 iSOE mode - Large Receive Offload

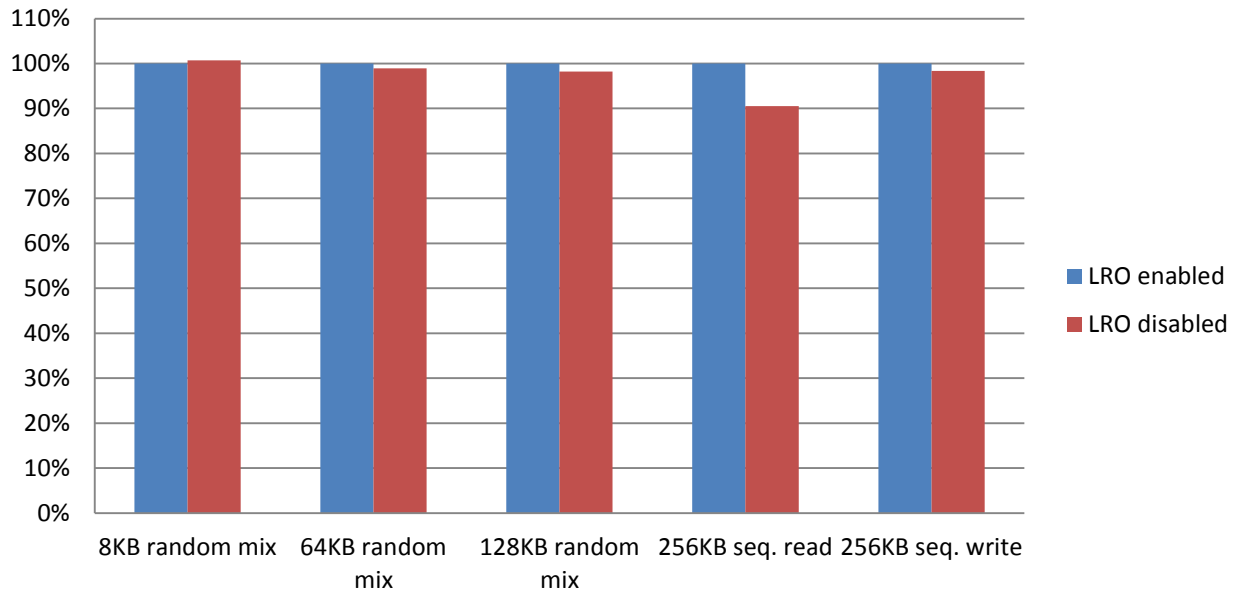


Figure 9 Performance of LRO configured when using a Broadcom 57810 in iSOE mode

Intel X520 - Large Receive Offload

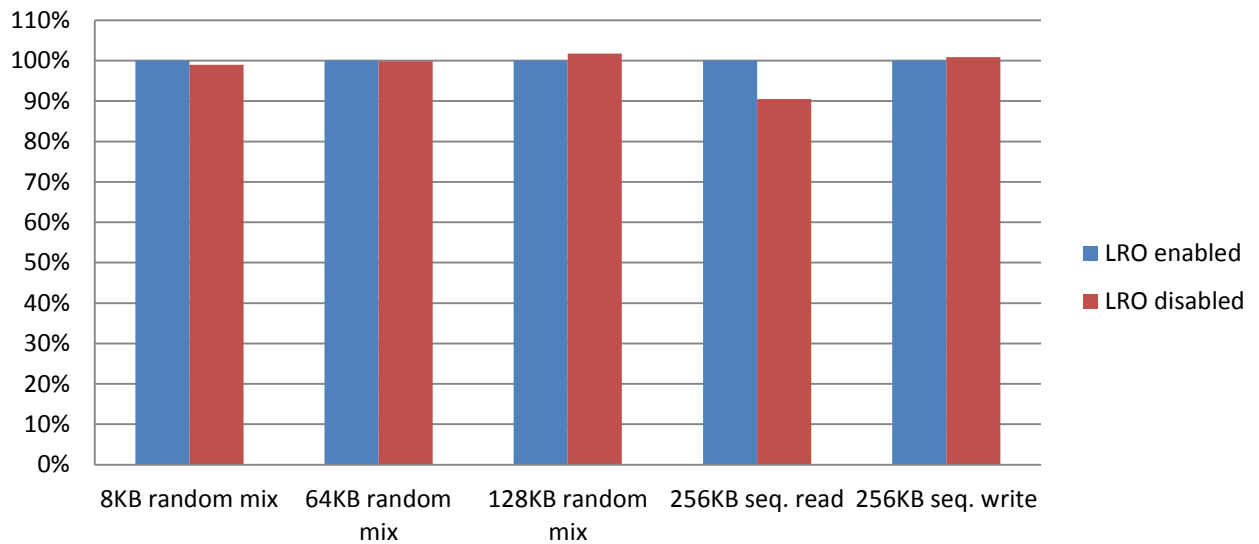


Figure 10 The performance effects of configuring LRO when using the Intel X520

4.2.2 TCP Checksum Offloading

Checksum Offload is a NIC feature that allows the TCP/IP/UDP checksums for send and receive traffic to be calculated by the adapter hardware rather than by the host CPU. In high-traffic situations, this can allow a system to handle a larger number of connections more efficiently than if the host CPU were forced to calculate the checksums. This property is inherently a hardware property and would not benefit from a software-only implementation. An adapter that supports Checksum Offload advertises this capability to the operating system so that the checksum does not need to be calculated in the protocol stack.

More information about TCO in ESXi can be found in the VMware Knowledge Base article number 2052904 at

http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=2052904

For all tested adapters, and in all NIC supported operating modes, the following default TCO settings were used.

Table 2 TCO settings

Offload parameters for vmnic:	Setting
rx-checksumming	On
tx-checksumming	On
tcp segmentation offload	On

The **tx-checksumming** and **tcp segmentation offload** settings cannot be changed because that function is not implemented in the driver versions used in testing.

Figures 14-16 show the results obtained after disabling **rx-checksumming**. They demonstrate that large block sequential read I/O is most significantly affected when changing this parameter.

The results of the tests illustrate that both Broadcom and Intel NIC drivers are tuned to deliver the best performance for most environments using the default settings.



Broadcom 57810 Standard mode - Rx Checksum Offload

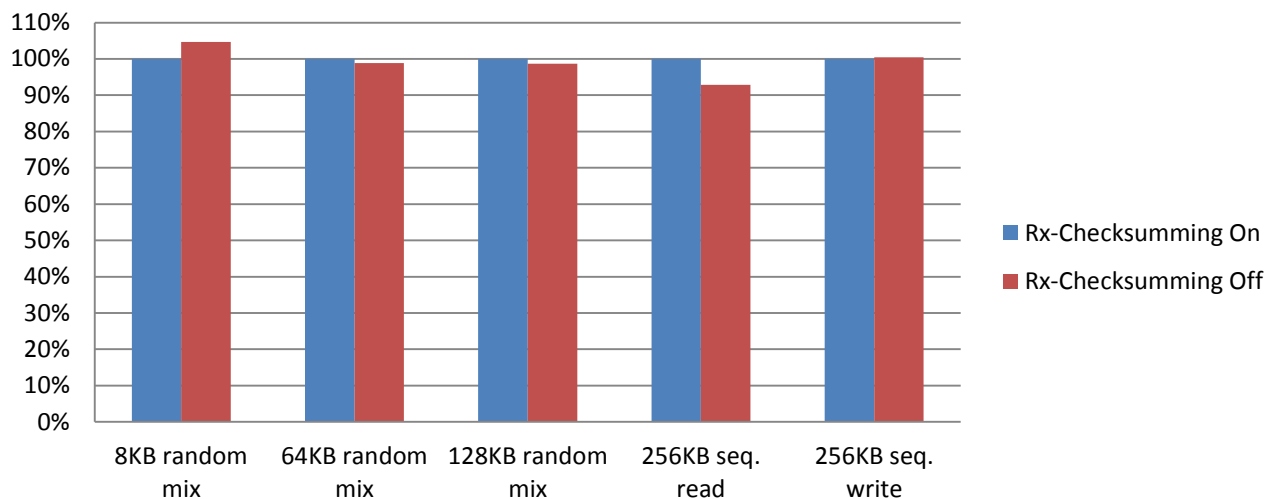


Figure 11 Configuring receive checksum offload for Broadcom 57810 in Standard mode

Broadcom 57810 iSOE mode - Rx Checksum Offload

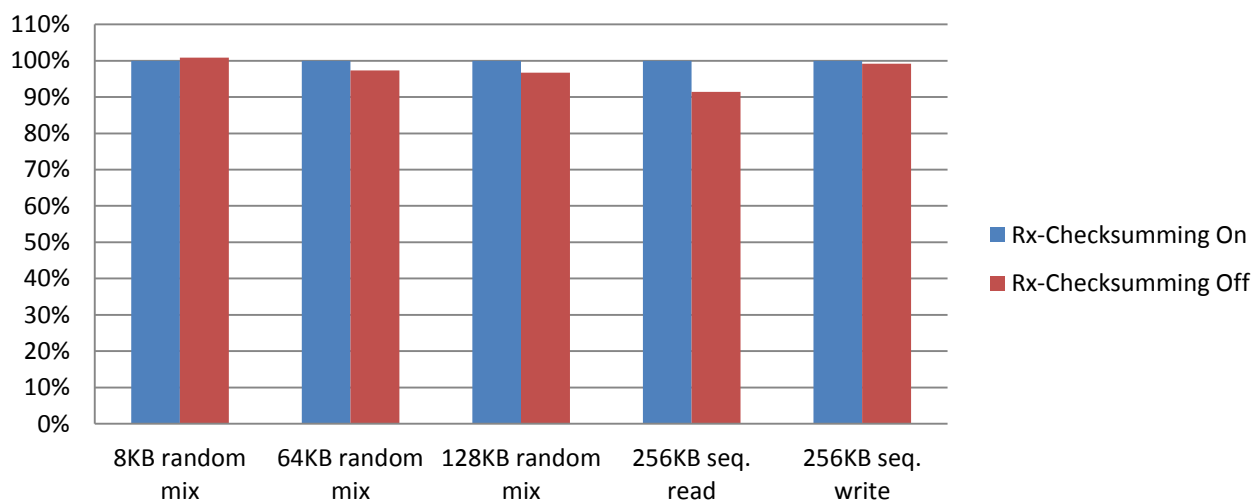


Figure 12 Configuring receive checksum offload for Broadcom 57810 in iSOE mode

Intel X520 - Rx Checksum Offload

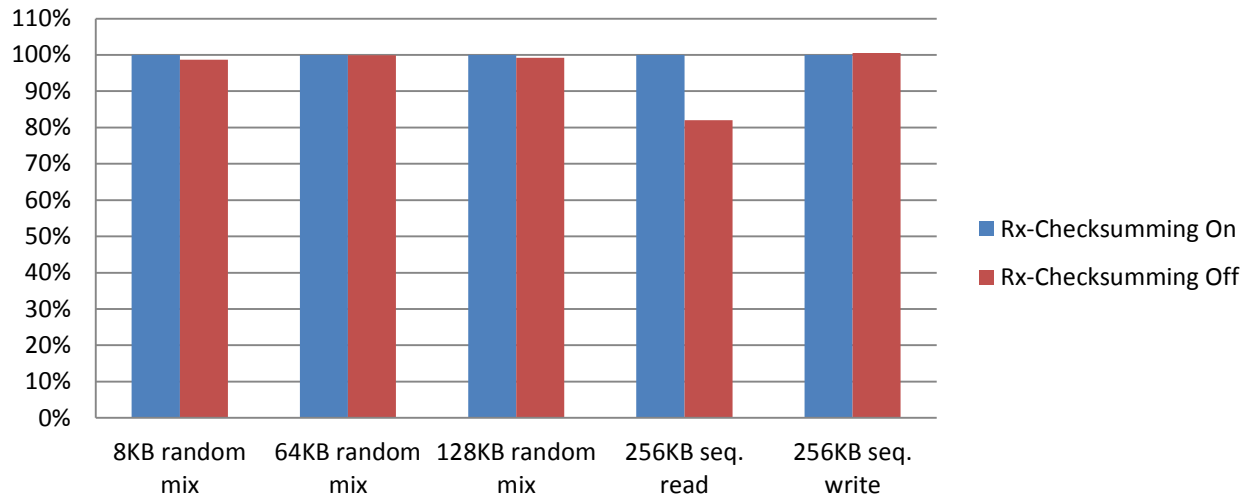


Figure 13 Configuring receive checksum offload for Intel X520

4.3 Broadcom BCM57810 iSOE mode performance results

This section compares the performance results using iSOE mode versus those using Standard mode with the five tested workloads and the Broadcom BCM57810 network adapter. Graph in Figure 14 illustrates the baseline configuration test results.

Broadcom 57810 - ISOE Vs. Standard mode

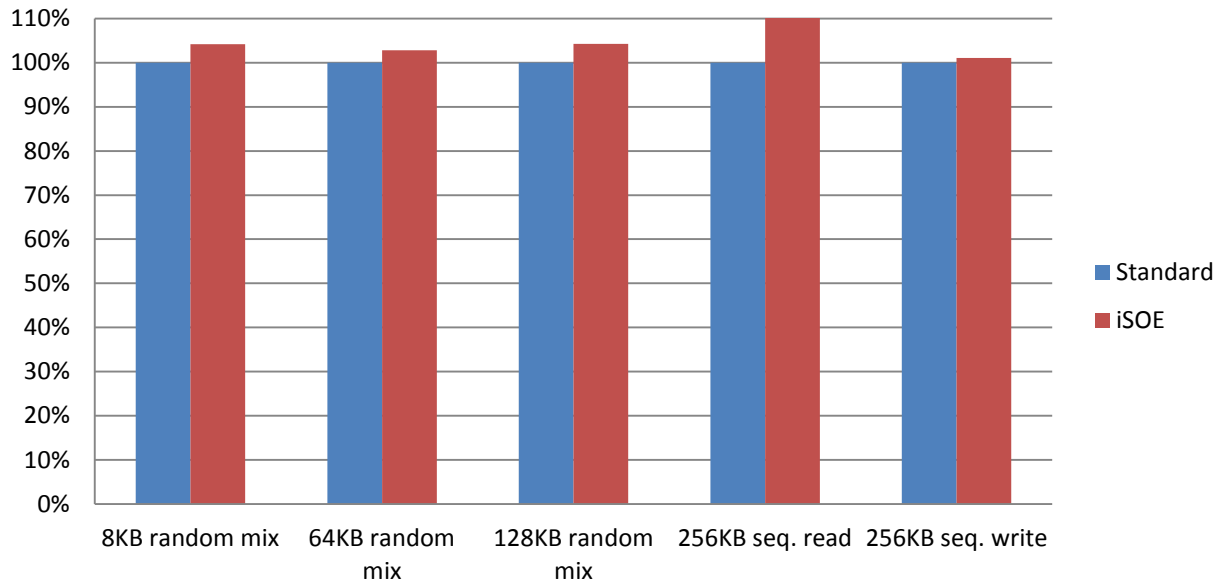


Figure 14 Performance effects of the Broadcom 57810 Standard mode versus iSOE mode

A small performance improvement was obtained when using the BCM57810 network adapter in iSOE mode. The large block sequential read workload demonstrated a 13% improvement over Standard mode performance.

5 Best practice recommendations

In this section, recommended configurations are given based on the performance results and analysis detailed in [Section 4](#). Non-default settings are recommended only when a significant change in performance for one or more workloads was detected, or when a setting is known to provide an advantage during network congestion or heavy CPU utilization. Only the non-default settings are mentioned.

For a complete list of tested options and default values as well as instructions on making configuration changes to the storage host, see [Appendix B](#).

5.1 Broadcom BCM57810 Standard mode recommended configuration

Based on the performance results and analysis for each workload, the following NIC and OS configuration changes are recommended for vSphere ESXi 5.5.

Table 3 Broadcom BCM57810 recommended configuration

Setting	Default value	Recommended value
Flow Control	Rx & Tx Enabled *	Rx & Tx Enabled
Jumbo packet **	1500	9000
iSCSI software initiator Delayed ACK **	Enabled	Disabled

Notes:

* Flow Control Default value may be different depending on NIC driver version and NIC type.

** Jumbo frames and Delayed ACK will be set to the recommended value during MEM installation.



5.2 Broadcom BCM57810 iSOE mode recommended configuration

Based on the performance results and analysis for each workload, the following NIC configuration changes are recommended for vSphere ESXi 5.5.

Table 4 Broadcom BCM57810 iSOE mode recommended configuration

Setting	Default value	Recommended value
Flow Control	Rx & Tx Enabled *	Rx & Tx Enabled
Jumbo packet **	1500	9000
iSCSI hardware initiator Delayed ACK **	Enabled	Disabled

Notes:

* Flow Control Default value may be different depending on NIC driver version and NIC type.

** Jumbo frames and Delayed ACK will be set to the recommended value during MEM installation. Since every HBA is an initiator, the Delayed ACK setting has to be configured for each one separately.

5.3 Intel X520 recommended configuration

Based on the performance results and analysis for each workload, the following NIC configuration changes are recommended for vSphere ESXi 5.5.

Table 5 Intel X520 recommended configuration

Setting	Default value	Recommended value
Flow Control	Rx & Tx Enabled *	Rx & Tx Enabled
Jumbo packet **	1500	9000
iSCSI software initiator Delayed ACK **	Enabled	Disabled

Notes:

* Flow Control Default value may be different depending on NIC driver version and NIC type.

** Jumbo frames and Delayed ACK will be set to the recommended value during MEM installation.



6 Conclusion

For an EqualLogic PS Series SAN, enable jumbo frames and flow control for both the Broadcom 57810 and the Intel X520 is recommended. Disabling iSCSI initiator Delayed ACK advance setting parameter also may improve performance for large blocks workloads.

The installation of EqualLogic Multipathing Extension Module for VMware (MEM) version 1.2 is highly recommended since it simplifies the SAN setup process and it gives the option to apply the EqualLogic best practice settings which enable jumbo packets and disable Delayed ACK.

With both Broadcom and Intel network adapters, and in all supported operation modes, flow control is required and is typically enabled by default. However, this default value might be different depending on NIC driver version and NIC type.

Instructions for how to configure flow control in an ESXi host are in the VMware Knowledge Base article number 1013413 at http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=1013413

The Broadcom BCM57810 iSOE mode is another compelling option since it exhibited performance benefits during every workload.

One thing to consider when using iSOE is the difference in administration procedures. When using the standard mode, an iSCSI software initiator can be bonded to several physical NICs. However once iSOE is enabled, each physical network adapter connected to the SAN (HBA) must be managed as an independent iSCSI initiator. The MEM setup script can help simplify this administration difference with its provided configuration scripts.

Leave all other NIC settings at their default unless instructed by Dell Technical Support.

Check www.support.dell.com for network adapter firmware updates, and www.vmware.com/downloads for NIC driver updates.



A Test configuration details

Hardware	Description
Server and software	<p>Dell PowerEdge R620 server:</p> <ul style="list-style-type: none"> • ESXi 5.5 Build# 1331820 • BIOS version: 2.0.19 • iDRAC firmware: 1.45.45 • (2) Intel Xeon E5-2620 • 32GB RAM • Dual Broadcom 57810 10GbE <ul style="list-style-type: none"> ◦ Driver bnx2x and bnx2i - 1.78.28.v55.2 ◦ Firmware 7.6.15 • Dual Intel x520 10GbE <ul style="list-style-type: none"> ◦ Driver ixgbe - 3.18.7iov ◦ Firmware 14.5.9 • Dell EqualLogic Multipathing Extension Module for VMware vSphere (MEM) version 1.2 • Server local storage datastore • Host is connected to two SAN Volumes using one of the two following iSCSI Initiator types: <ul style="list-style-type: none"> ◦ iSCSI Software adapter – For Broadcom 57810 standard mode and for Intel X520 tests ◦ Two Broadcom iSCSI Adapters – For Broadcom 57810 iSOE mode tests • Two 100 GB SAN volumes presented as VMFS5 datastores <p>Virtual Machine:</p> <ul style="list-style-type: none"> • Windows 2008 R2 Guest OS • 1vCPU • 4096 MB RAM • Three SCSI Controllers, each one controlling one of the three VMDK virtual hard disks: <ul style="list-style-type: none"> ◦ VM OS was installed in the host local storage datastore virtual disk ◦ One 100 GB VMDK virtual disk in each of the two SAN volume datastores used for testing
SAN switches	<p>(2) Dell Force10 s4810</p> <ul style="list-style-type: none"> • Firmware: 9.2.0.0
SAN array members	<p>(2) Dell EqualLogic PS6110XV</p> <ul style="list-style-type: none"> • (2) 10GbE controllers • Firmware: 7.0 R367801 • Two 100GB volumes



B Network adapter and TCP stack configuration details

This section provides details about the configuration options and default settings of the network adapter properties and the ESXi 5.5 TCP stack.

B.1 Broadcom BCM57810 adapter options standard mode

The following table lists the tested adapter options for the Broadcom BCM57810 NetXtreme II 10 GigE NIC in standard mode along with the default value.

Table 6 Broadcom BCM57810 standard mode adapter options

Setting	Default value	Recommended value
Flow control	Rx & Tx Enabled	Rx & Tx Enabled
Jumbo packet	1500	9000
Large Receive Offload	Enabled	Enabled
TCP/UDP Checksum Offload	Rx & Tx Enabled	Rx & Tx Enabled
Delayed ACK	Enabled	Disabled



B.2 Configuring Broadcom BCM57810 adapter parameters in standard mode

Broadcom BCM57810 network adapter and ESXi 5.5 TCP stack properties must be set in different phases. Jumbo frames and disabled Delayed ACK are established when installing and configuring EqualLogic MEM for VMware. (See [Appendix D](#))

NIC flow control default settings can be verified and changed from the ESXi CLI console, so physical or remote SSH access to the server is required.

B.3 Broadcom BCM57810 adapter options iSOE mode

The following table lists the tested adapter options for the Broadcom BCM57810 NetXtreme II 10 GigE NIC in iSOE mode along with the default value.

Table 7 Broadcom BCM57810 iSOE mode adapter options

Setting	Default value	Recommended value
Flow control	Rx & Tx Enabled	Rx & Tx Enabled
Jumbo packet	1500	9000
Large Receive Offload	Enabled	Enabled
TCP/UDP Checksum Offload	Rx & Tx Enabled	Rx & Tx Enabled
Delayed ACK	Enabled	Disabled



B.4 Configuring Broadcom BCM57810 adapter parameters in iSOE mode

Broadcom BCM57810 network adapter and ESXi 5.5 TCP stack properties must be set in different phases. Jumbo frames and disabled Delayed ACK are established when installing and configuring EqualLogic MEM for VMware.

NIC flow control default settings can be verified and changed from the ESXi CLI console. Physical or remote SSH access to the server is required.

[Appendix D](#) shows how to install and configure MEM in interactive mode.

B.5 Intel X520 adapter options

The following table lists the tested adapter options for the Intel X520 10 GigE NIC along with the default value.

Table 8 Intel X520 adapter options

Setting	Default value	Recommended value
Flow control	Rx & Tx Enabled	Rx & Tx Enabled
Jumbo packet	1500	9000
Large Receive Offload	Enabled	Enabled
TCP/UDP Checksum Offload	Rx & Tx Enabled	Rx & Tx Enabled
Delayed ACK	Enabled	Disabled

B.6 Configuring Intel X520 adapter properties

Intel X520 adapter and ESXi 5.5 TCP stack properties must be set in different phases. Jumbo frames and disabled Delayed ACK are established when installing and configuring EqualLogic MEM for VMware.

NIC flow control default settings can be verified and changed from the ESXi CLI console. Physical or remote SSH access to the server is required.

[Appendix D](#) shows how to install and configure MEM in interactive mode.



C I/O parameters

Vdbench SAN workloads were executed using the following parameters in the parameter file.

Common parameters:

```
hd=default  
hd=one, system=localhost
```

iSCSI volumes:

```
sd=A-a, lun=\\.\PhysicalDrive1, threads=1  
sd=A-b, lun=\\.\PhysicalDrive2, threads=1
```

8KB random 67% read workload:

```
wd=wd1, sd=A-*, seekpct=random, rdpct=70, xfersize=8k
```

256KB sequential read workload:

```
wd=wd1, sd=A-*, seekpct=seq, rdpct=100, rhpct=100, xfersize=262144
```

256KB sequential write workload:

```
wd=wd1, sd=A-*, seekpct=seq, rdpct=0, whpct=100, xfersize=262144
```

64KB random 67% read workload:

```
wd=wd1, sd=A-*, seekpct=random, rdpct=70, xfersize=64k
```

128KB random 67% read workload:

```
wd=wd1, sd=A-*, seekpct=random, rdpct=70, xfersize=128k
```

Runtime options:

```
rd=rd1, wd=wd*, elapsed=1200, interval=30, iorate=max
```



D Installing and configuring the EqualLogic MEM

This appendix provides instructions for installing and configuring the EqualLogic MEM for VMware in an ESXi5.1 or ESXi5.5 host.

1. Place the ESXi host in Maintenance Mode, right click on the ESX host from vSphere Client, and select **Enter Maintenance** mode.

Note: The appropriate VMware vSphere CLI (vCLI 5.1 or 5.5) environment must be installed on the management server.

2. From the vCLI command prompt, change to the directory where the MultiPathing Extension Module (MEM) was unzipped.
3. Type the following command:
setup.pl --install --server=<ESXi host IP address> --bundle=<bundle file name>

Where <bundle file name>=the zip file, such as dell-eql-mem-esx5-1.2.0.365964.zip

4. When prompted, login as root with the correct password. It may take several minutes.
5. When the setup is complete, **Clean install was successful** is displayed.
6. Type the following command in the vSphere CLI command prompt:
setup.pl --configure --server=<ESXi host IP address>
7. Enter username and password when prompted.
8. Provide the following answers to the configure wizard questions:
 - a. For the switch type, press **[Enter]** to select **vSwitch** (the default).
 - b. For the vSwitch name, press **[Enter]** to accept the default value.
 - c. To identify the NICs, enter the VMNIC name that will handle the iSCSI traffic, separated by space.
Example: vmnic4 vmnic5
9. Enter the IP addresses for every VMNIC.
10. Enter the netmask.
11. MTU size is 9000.
12. Select the default prefix by pressing **[Enter]**.
13. Select the SW iSCSI or HW iSCSI and press **[Enter]**.
14. If prompted, answer yes to enable the iSCSI initiator.
15. Enter the PS Group IP address.
16. Enter CHAP user and CHAP secret information and press **[Enter]**, or just press **[Enter]** to deactivate CHAP.
17. Answer **yes** to apply Dell EqualLogic Best Practice settings.
18. Answer **yes** to proceed with the installation.



For more information, the *ESXi 5.5 Host Configuration Guide* includes step-by-step instructions for installing and configuring these parameters at

<http://en.community.dell.com/dell-groups/dtcmedia/m/mediagallery/20094619/download.aspx>.

Specifically, Section 3 of the *ESXi 5.5 Host Configuration Guide* provides instructions for installing and configuring MEM in interactive mode.



Additional resources

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

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

Referenced or recommended Dell publications:

- EqualLogic Configuration Guide:
<http://en.community.dell.com/dell-groups/dtcmedia/m/mediagallery/19852516/download.aspx>
- Rapid EqualLogic Configuration Portal:
<http://en.community.dell.com/techcenter/storage/w/wiki/3615.rapid-equallogic-configuration-portal-by-sis.aspx>
- EqualLogic Compatibility Matrix (ECM):
<http://en.community.dell.com/techcenter/storage/w/wiki/2661.equallogic-compatibility-matrix.aspx>
- EqualLogic Switch Configuration Guides:
<http://en.community.dell.com/techcenter/storage/w/wiki/4250.switch-configuration-guides-by-sis.aspx>
- The latest EqualLogic firmware updates and documentation (site requires a login):
<http://support.equallogic.com>
- Dell Manuals:
<http://www.dell.com/support/Manuals/us/en/555/ProductSelector>

For EqualLogic best practices white papers, reference architectures, and sizing guidelines for enterprise applications and SANs, refer to Storage Infrastructure and Solutions Team Publications at <http://dell.to/sM4hJT>.

