



Best Practices for a DCB-enabled Dell M-Series Blade Solution with EqualLogic PS-M4110

A Dell EqualLogic Reference Architecture

Dell Storage Engineering
July 2013

This document has been archived and will no longer be maintained or updated. For more information go to the [Storage Solutions Technical Documents](#) page on Dell TechCenter or contact support.

© 2013 Dell Inc. All Rights Reserved. Dell, the Dell logo, PowerEdge, and other Dell names and marks are trademarks of Dell Inc. in the US and worldwide. Broadcom® is among the trademarks of Broadcom Corporation and/or its affiliates in the United States certain other countries and/or the EU. Microsoft®, Windows®, Windows Server®, and Hyper-V® are either trademarks or registered trademarks of Microsoft Corporation in the United States and/or other countries. All other trademarks mentioned herein are the property of their respective owners.



Table of contents

Acknowledgements	5
Feedback	5
Executive summary	5
1 Introduction.....	6
1.1 Audience.....	6
1.2 Terminology	6
2 Typical deployment scenarios.....	8
2.1 Scenario 1: Remote or branch office with iSCSI traffic.....	8
2.2 Scenario 2: Converged infrastructure environment.....	8
2.3 Scenario 3: Replication to remote EqualLogic storage	8
2.4 Scenario 4: Comparison of RAID 10 versus RAID 50.....	9
3 Test components and configuration	10
3.1 Dell Force10 MXL switches	12
3.2 PowerEdge M620 blade server	14
3.3 Dell EqualLogic PS-M4110 storage array	15
3.4 Other components	15
3.4.1 External Dell PowerEdge rack servers	15
3.4.2 Remote Dell EqualLogic PS series storage arrays	16
4 Analysis of deployment scenarios	17
4.1 Scenario 1: Remote or branch office with iSCSI traffic.....	17
4.1.1 Results and analysis	17
4.2 Scenario 2: Converged infrastructure environment.....	19
4.2.1 Results and analysis	19
4.3 Scenario 3: Replication to remote EqualLogic storage	20
4.3.1 Results and analysis	21
4.4 Scenario 4: Comparison of RAID 10 versus RAID 50.....	22
4.4.1 Results and analysis	22
5 Conclusions and best practice recommendations.....	24
A Configuration details	25
A.1 Dell Force10 MXL blade switch setup.....	25
A.2 Dell PowerEdge M620 blade server	28



A.3	Broadcom 57810 network adapter:	28
A.4	Dell EqualLogic PS-M4110 storage arrays	29
A.5	Dell PowerEdge M1000e chassis.....	29
	Additional resources	30



Acknowledgements

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

Engineering: Nirav Shah

Technical Marketing: Guy Westbrook

Editing: Margaret Boeneke

Additional contributors: Mike Kosacek

Feedback

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



SIFeedback@Dell.com

Executive summary

This white paper describes the configuration and several deployment scenarios for a solution consisting of a Dell™ PowerEdge™ M1000e Blade Enclosure, Dell PowerEdge M620 servers, and Dell EqualLogic™ PS-M4110 Blade Array storage. This solution provides a complete server, storage, and networking solution in a single M1000e blade chassis form factor. The paper also analyzes the scalability of the solution in relation to the number of work threads and the number of servers in an environment. It describes the effects of converged traffic and replication activities on the performance of the storage system. Finally, the paper compares the pros and cons of RAID 10 versus RAID 50 on the Dell EqualLogic PS-M4110 Blade Array storage.



1 Introduction

The Dell PowerEdge M1000e Blade Enclosure, PowerEdge M620 Blade Servers, and Dell EqualLogic PS-M4110 Blade Array storage work together to create a blade solution that fits deployments where server, storage, and networking needs are consolidated in a single M1000e blade chassis form factor. This solution provides a modular building block for growing an infrastructure – virtual or physical –while minimizing the data center impact in terms of power, cooling, and rack space. The modular blocks of storage, servers, and network switches can be used to create a variety of solutions from a self-contained datacenter unit for a small business to modular building-block units for enterprise class datacenters.

This paper contains an explanation of the highly integrated blade server, storage, and networking components and how they benefit the solution. Data Center Bridging (DCB) allows convergence of both LAN and SAN over a single network fabric, simplifying network deployments and saving on component costs. Configuration of DCB on the Dell Force10 MXL blade switches and Broadcom® 57810S 10 GbE Network adapters is explained in detail.

A comparison of performance and scalability with both Windows Server® 2012 and Microsoft® Hyper-V® 2012 environments is explored. Finally, the effects of RAID type and storage data replication are discussed.

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 white paper is intended for server and storage administrators, SAN system designers, storage consultants, or anyone who is involved in the design or implementation of data centers using hosted or hypervised operating systems. 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).

1.2 Terminology

The following terms are used throughout this document.

Broadcom Advance Controller Suite (BACS) – A windows management application for Broadcom network adapters.

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

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



Enhanced Transmission Selection (ETS) - (IEEE 802.1Qaz) Provides administrators with the ability to group multiple classes of services together and then define guaranteed minimum bandwidth allocation from the shared network connection.

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

Open Shortest Path First (OSPF) – A link-state routing protocol that calls for the sending of link-state advertisements (LSAs) to all other routers within the same hierarchical area.

Threads - Maximum number of concurrent outstanding I/O generated by the load generator tool (Vdbench).



2 Typical deployment scenarios

This paper examines four different configurations for solutions created using the Dell blade portfolio. The configurations described in this paper are designed to help deploy Dell blade servers, storage, and networking in the scenarios described in this section.

2.1 Scenario 1: Remote or branch office with iSCSI traffic

A remote or branch office configuration typically requires an easily manageable, compact, and scalable solution for server, storage, and networking needs. The environment may host applications such as databases, email, file shares, and directory services performing a variety of workloads on the storage arrays. The iSCSI storage traffic generated by these applications can vary from small block to large block and from random to sequential workloads. The most common workload generated by these application is small block random I/O. A Dell blade solution for a remote/branch office consisting of Dell PowerEdge M620 blade servers, Dell Force10 MXL blade switches and Dell EqualLogic PS-M4110 storage arrays can provide all server, storage, and networking needs in a single, compact M1000e chassis. A Dell blade solution also offers the flexibility to expand by adding additional blade servers, storage, or networking fabrics as business needs dictate.

This scenario demonstrates the scalability of the Dell blade solution as application servers are added and as workloads increase.

2.2 Scenario 2: Converged infrastructure environment

When DCB is enabled on the Force10 MXL blade switches, a remote or branch office environment is able to consolidate or converge multiple traffic types – in this case LAN and SAN - into a single physical network environment. Sharing a single switch fabric can lead to lower infrastructure costs and simpler administration.

At the same time, many storage applications have minimum bandwidth and latency requirements making the storage (iSCSI) performance mission critical. The ability to guarantee each application or traffic type a minimum amount of networking resources is an essential part of the ETS features within the DCB environment. ETS settings on the Force10 MXL blade switch can be configured to guarantee a minimum amount of bandwidth for the iSCSI storage traffic.

This scenario demonstrates the ability of Dell blade solution to manage LAN and SAN traffic on the same network fabric while maintaining required performance levels for iSCSI storage.

2.3 Scenario 3: Replication to remote EqualLogic storage

Auto-Replication is an enterprise disaster recovery feature included with EqualLogic arrays. It provides the ability to replicate data to peer EqualLogic storage arrays situated in remote locations without setting the volumes offline. Auto-Replication provides disaster recovery in cases where the original volume is destroyed or becomes unavailable. This feature provides point-in-time replicas between primary and secondary arrays. Volumes can be configured to replicate data on a schedule or manually, as needed.



After the initial synchronization of a volume replica, only new or changed data is replicated on subsequent point-in-time copies.

Details about the Dell EqualLogic auto-replication feature are in the white paper “Dell EqualLogic Auto-Replication Best Practices and Sizing Guide” available at:

<http://en.community.dell.com/techcenter/storage/w/wiki/2641.dell-equallogic-auto-replication-best-practices-and-sizing-guide-by-sis.aspx>.

This scenario analyzes the effects of an application workload on a volume while replication is in progress and also examines the effect of additional LAN traffic while replication is running.

2.4 Scenario 4: Comparison of RAID 10 versus RAID 50

EqualLogic Storage arrays with firmware version 6.0 support RAID 10, RAID 50, and RAID 6. There are advantages and disadvantages of each RAID type. This scenario compares the capacity and performance for RAID 10 and RAID 50 which are typically recommended with arrays containing drives of 2 TB or less.

A detailed analysis of how to choose a RAID policy can be found in the technical report “EqualLogic PS Series Storage Arrays: Choosing a Member RAID Policy” at

<http://www.dellstorage.com/WorkArea/DownloadAsset.aspx?id=1066>.

In this scenario, the capacity and performance of RAID 10 and RAID 50 were analyzed on an EqualLogic M4110 blade storage array. For the performance testing, both Microsoft Hyper-V 2012 and Windows server 2012 environments were compared. The performance data that is presented is based on iSCSI SAN storage traffic from running Vdbench on four PowerEdge M620 servers connected to volumes on the EqualLogic M4110 storage arrays.



3 Test components and configuration

The Dell PowerEdge M1000e is designed for more efficient use of space and power, as well as easier deployment and administration. The blade enclosure is modular and can be populated with blade server modules, blade storage modules, blade network switch or pass-through modules, or other supported blade modules as needed.

In the configuration used for testing, the Dell PowerEdge M1000e chassis was populated with eight Dell PowerEdge M620 blade servers, two Dell EqualLogic M4110 blade storage arrays, and two Dell Force10 MXL blade switches in Fabric A. The Dell PowerEdge M1000e blade chassis had redundant chassis management controllers and was connected to the management network for setup and management purposes.

For configuration details, refer to the *Dell PowerEdge M1000e Systems Configuration Guide* at: http://downloads.dell.com/Manuals/all-products/esuprt_ser_stor_net/esuprt_poweredge/poweredge-m1000e_Reference%20Guide3_en-us.pdf

A pair of Dell Force10 MXL switches in Fabric A were DCB enabled and were used for both iSCSI/SAN traffic and LAN traffic needs. Broadcom 57810 10 GbE mezzanine cards were used as 10 GbE network adapters on the Dell PowerEdge M620 server for iSCSI/SAN network and LAN network connectivity. There were also two external rack mounted Dell PowerEdge servers, which were used to simulate LAN traffic on the test configuration. This section shows the configuration of the components used in the tests.

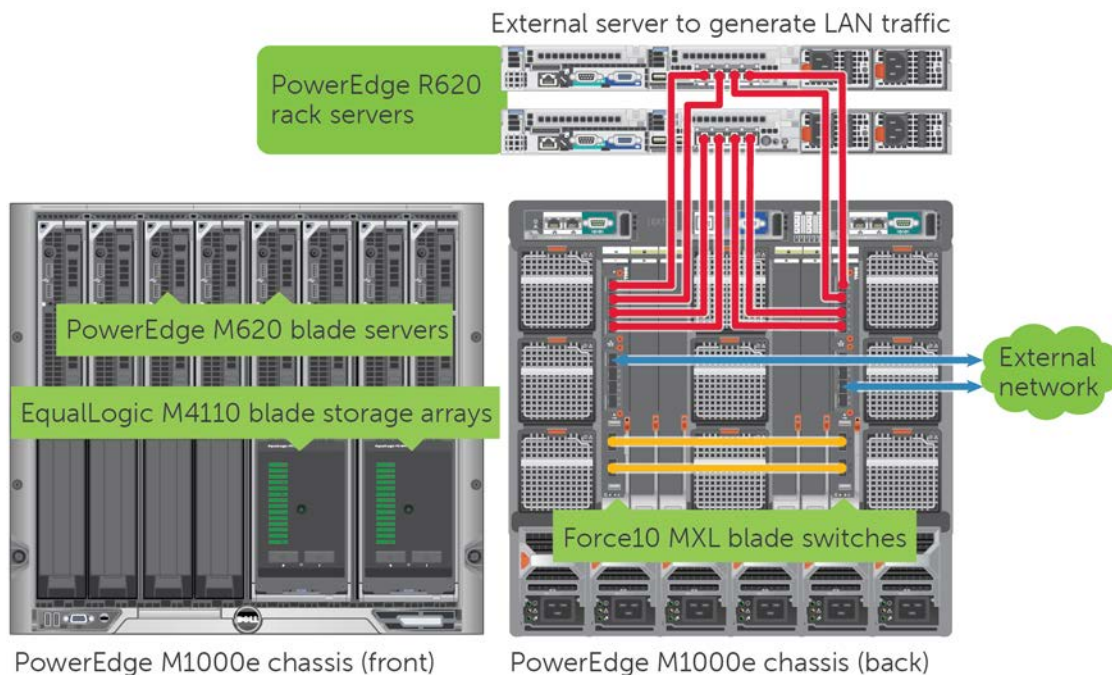


Figure 1 Dell blade solution with external PowerEdge rack mount servers

Figure 2 shows the remote EqualLogic PS storage arrays that were used as outbound replication partners for the EqualLogic PS-M4110 storage arrays in the Dell blade solution. These remote storage arrays were used for the data/volume replication test.

Remote EqualLogic PS Series storage array (used as outbound replication partner)

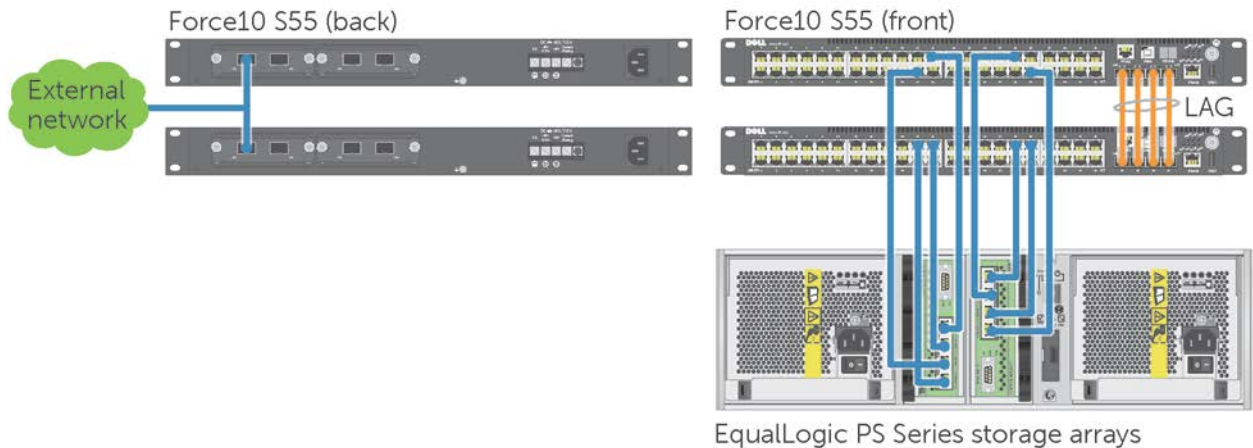


Figure 2 Remotely located EqualLogic PS Series storage

Figure 3 shows the logical connectivity diagram of the solution configuration. All of the connections between the Dell blade solution components are shown below, along with the connections to the external rack servers and the remote EqualLogic PS Series arrays.

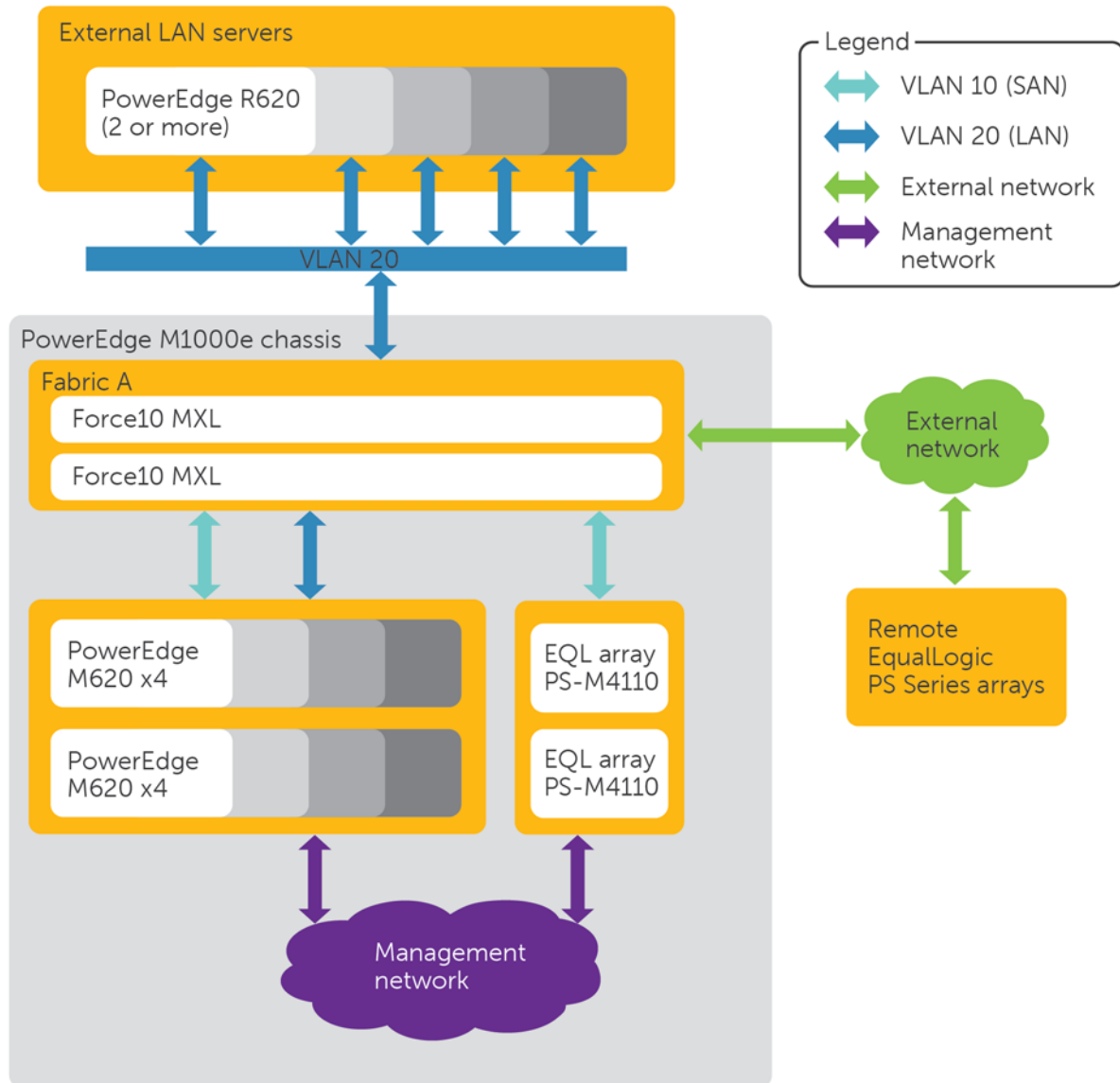


Figure 3 Logical connectivity diagram for the test configuration

3.1 Dell Force10 MXL switches

The Force10 MXL switch is a modular switch that is compatible with the PowerEdge M1000e blade chassis and provides 1/10/40 GbE ports addressing the diverse needs of most environments such as a large enterprise data center, a high performance computing environment, or branch offices. The MXL switch supports 32 internal 1/10 GbE ports, as well as two fixed 40 GbE QSFP+ ports and offers two bays for optional FlexIO modules. FlexIO modules can be added as needed to provide additional 10 GbE or 40 GbE ports.

For additional information on the Force10 MXL switch see the Manuals and Documents for the Force10 MXL Blade at: <http://www.dell.com/support/Manuals/us/en/04/Product/force10-mxl-blade>

The Force10 MXL blade switch is DCB capable, enabling it to carry multiple traffic classes, and it allows configuration of a lossless traffic class for iSCSI. The test environment had DCB enabled with the iSCSI traffic class configured in a lossless priority group to ensure priority of iSCSI/SAN traffic. LAN and other types of traffic were in a separate (lossy) priority group. For our tests, ETS was configured to provide a minimum of 50% bandwidth for iSCSI traffic and 50% for other or LAN traffic.

Two blade switches were used in Fabric A of the PowerEdge M1000e chassis for High Availability (HA). The two Force10 MXL switches were connected with a Link Aggregation Group (LAG) to enable communication and traffic between them.

Force10 MXL blade switch

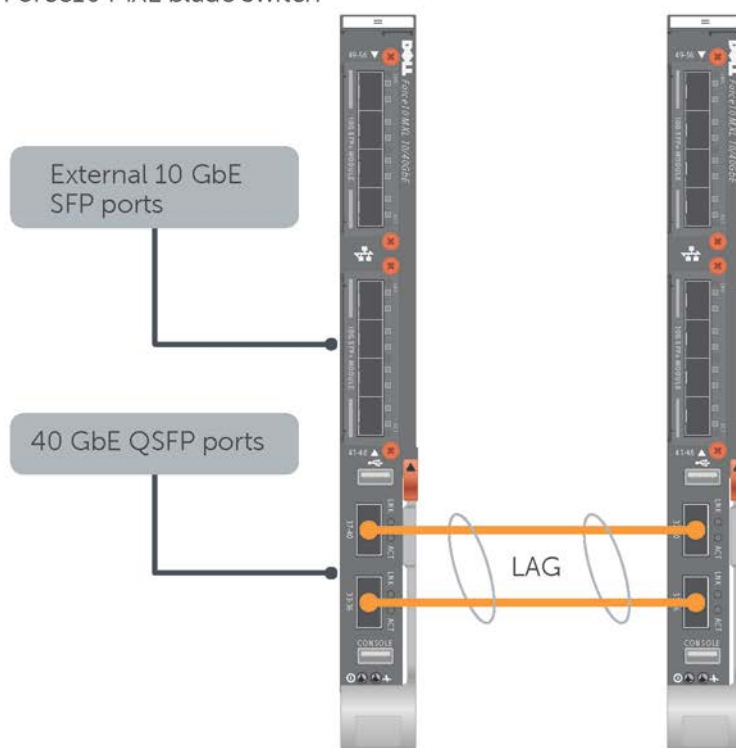


Figure 4 LAG connection between two Force10 MXL blade switches in a PowerEdge M1000e blade chassis fabric

The blade switch is internally connected with the blade servers through the backplane of the blade chassis, and it has optional I/O modules to provide an additional 16 external 10 GbE ports (ten gigabit 0/41 to ten gigabit 0/56) along with two fixed external 40 GbE QSFP ports (fortygigE 0/33 and fortygigE 0/37). The two external 40 GbE ports are used for a LAG between the two Force10 MXL switches. The external 10 GbE ports are used for connectivity with external PowerEdge R620 Rack servers (used to simulate LAN traffic) as well as for the connection to the switches designated for the external network (used for storage replication testing).

These configuration steps were used to configure the Force10 MXL switch:

- Enable switch ports and spanning tree on switch ports
- Configure LAG and QSFP ports for LAG.
- Enable DCB on the switch and create the DCB policy.
 - Create two VLANs, one used for iSCSI SAN traffic and the other used for LAN traffic.
 - Configure priority group and policies.
 - Configure ETS values.
 - Assign priority class to switch ports.
- Configure external connection for outbound replication.
- Save switch settings.

More configuration details are available in the appendix in Section A.4.

The iSCSI traffic class was assigned 50% bandwidth and LAN traffic was assigned 50% bandwidth through ETS settings.

3.2 PowerEdge M620 blade server

Each PowerEdge M620 blade server had a Broadcom 57810S Dual-Port 10 GBE KR blade Network daughter card connected to Fabric A, which was used for all iSCSI/SAN and LAN Ethernet traffic. Each Dell PowerEdge M620 server ran either Windows Server 2012 Operating System (OS) or Windows Server 2012 virtual machines on Microsoft Hyper-V Server 2012 depending on the test scenario.

Configuration for the Virtual Machine on the Microsoft Hyper-V Server 2012 environment is as follows:

- The Hypervisor used was Windows Hyper-V 2012.
- A single Windows Server 2012 virtual machine was created on each server running Microsoft Hyper-V server 2012.
- Microsoft Hyper-V Server 2012 used the Microsoft software iSCSI initiator to connect to volumes created on the Dell EqualLogic PS-M4110 blade storage. These volumes were passed to the Windows Server 2012 virtual machine as virtual disks.
- The NDIS function of the Broadcom 57810 on Hyper-V server 2012 was shared with the Virtual machine for LAN traffic connectivity.

The iSCSI offload engine (iSOE) was enabled on the Broadcom 57810S. The separate NDIS function was used for LAN connectivity. This resulted in one function for LAN traffic and one function for iSCSI SAN traffic. Even though iSOE mode was used for iSCSI, both Windows Server 2012 and Hyper-V Server use the Microsoft software iSCSI initiator to manage and establish sessions to volumes on the EqualLogic PS-M4110.



The NICs were configured as shown in Table 1. For detailed configuration of the Broadcom 57810S NIC, see appendix Section A.3.

For details on configuring the switch and network adapters for DCB, please refer to:
<http://en.community.dell.com/techcenter/storage/w/wiki/4355.configuring-dcb-with-equallogic-sans.aspx>

Table 1 Parameters on the NDIS and iSCSI partitions

Parameters	NDIS VDB client	iSCSI offload Adapter
IP address	10.10.20.XXX (LAN IP address)	10.10.10.XXX (iSCSI / SAN IP address)
VLAN ID	VLAN 20	VLAN 10
Jumbo frame	No change	MTU 9216
DCB settings	N/A	DCBx in "willing mode"

3.3 Dell EqualLogic PS-M4110 storage array

Two EqualLogic PS-M4110 modular blade storage arrays were installed into a PowerEdge M1000e chassis. The PS-M4110 blade storage arrays were configured on Fabric A. Both arrays were configured as a single group and placed into a single pool. All volumes created from the pool were automatically distributed across both array members.

The PS-M4110 storage arrays were configured using RAID 10, unless specified.

A general configuration guide for Dell EqualLogic PS series storage arrays can be found at:
<http://en.community.dell.com/dell-groups/dtcmedia/m/mediagallery/19852516/download.aspx>

3.4 Other components

This paper has specific tests for converged iSCSI traffic, LAN traffic, and storage replication to a remote site. These tests require additional components in the test environment and were configured as explained in the following sections:

3.4.1 External Dell PowerEdge rack servers

Two external PowerEdge rack servers were used to simulate LAN traffic generated on the Dell blade solution. Each of these external servers were populated with two 10 GbE dual port network adapters to generate LAN traffic. These ports were connected to the external 10 GbE ports on the Force10 MXL blade switch. The 10 GbE port on the external PowerEdge rack servers was configured to use VLAN 20 which was configured for LAN traffic in the test environment.



3.4.2 Remote Dell EqualLogic PS series storage arrays

The remote EqualLogic PS Series storage array was used as the outbound replication partner for data replication from the Dell blade solution. The remote setup consisted of EqualLogic PS Series 1 Gb storage arrays connected to the external network using Force10 S55 1 GbE switches as shown in Figure 2. There was a 10Gb connection between the Force10 S55 1 GbE switch for the remote storage and the Force10 MXL blade switches.

Note: Since LAN and WAN speeds between deployments can vary greatly, replication performance results may also differ from site to site.

This configuration with remote EqualLogic PS Series storage arrays was used for testing data replication only. A typical production environment may also have additional servers/hosts connected on the remote site for the purpose of data recovery from the replicated volumes in the event of a disaster at the primary site.



4 Analysis of deployment scenarios

4.1 Scenario 1: Remote or branch office with iSCSI traffic

For the scalability and performance test, either four or eight PowerEdge M620 servers were used, depending on the test. Similarly, two or four volumes were assigned to each server.

To demonstrate the performance in virtual OS environments, the test configuration used Windows 2012 Hyper-V as the Hypervisor. A single virtual machine running Windows 2012 as the guest operating system was created on each of the Hyper-V servers. The four volumes were presented to the Guest OS as virtual hard drives. iSOE mode was enabled on the Broadcom 57810 network adapter and the NDIS partition of the Broadcom 57810 was shared with the virtual machines for LAN traffic. The entire solution was DCB enabled, although only iSCSI traffic was generated for test scenario 1.

For these tests, a random small block (8K) workload with a distribution of 67% read and 33% write was generated with Vdbench.

Scalability of the Dell solution with EqualLogic PS-M4110 storage arrays was tested by increasing the number of work threads per server and the number of servers running a constant number of threads on the storage arrays. Two tests were run to measure scalability in terms of increase in customer workload and increase in number of servers in a customer deployment.

To test the scalability related to the number of threads running per server, the number of threads was increased from a single thread and increased the number of threads until the threshold of a 20 ms response time was reached. To test the scalability of servers running constant threads, the number of servers was increased from one until the threshold of a 20 millisecond response time was reached. The test for server scalability ran 10 threads per server.

Both of these scenarios were tested in a native Windows Server 2012 environment and a Windows 2012 Hyper-V environment.

4.1.1 Results and analysis

The following section describes the effects of an increased workload (represented by the increased number of threads or servers) for small block random I/Os.

4.1.1.1 Scaling work threads on Dell EqualLogic PS-M4110

Figure 5 shows the performance of the small block random workload when scaling the number of threads from one to 10 on four blade servers in both a Windows Server 2012 environment and a Hyper-V virtualized environment. Here, IOPS for the workload generated are compared to the number of threads running from four blade servers to four volumes per server simultaneously. The results show that there is a linear increase in the IOPS as the number of work threads by the blade servers is increased.



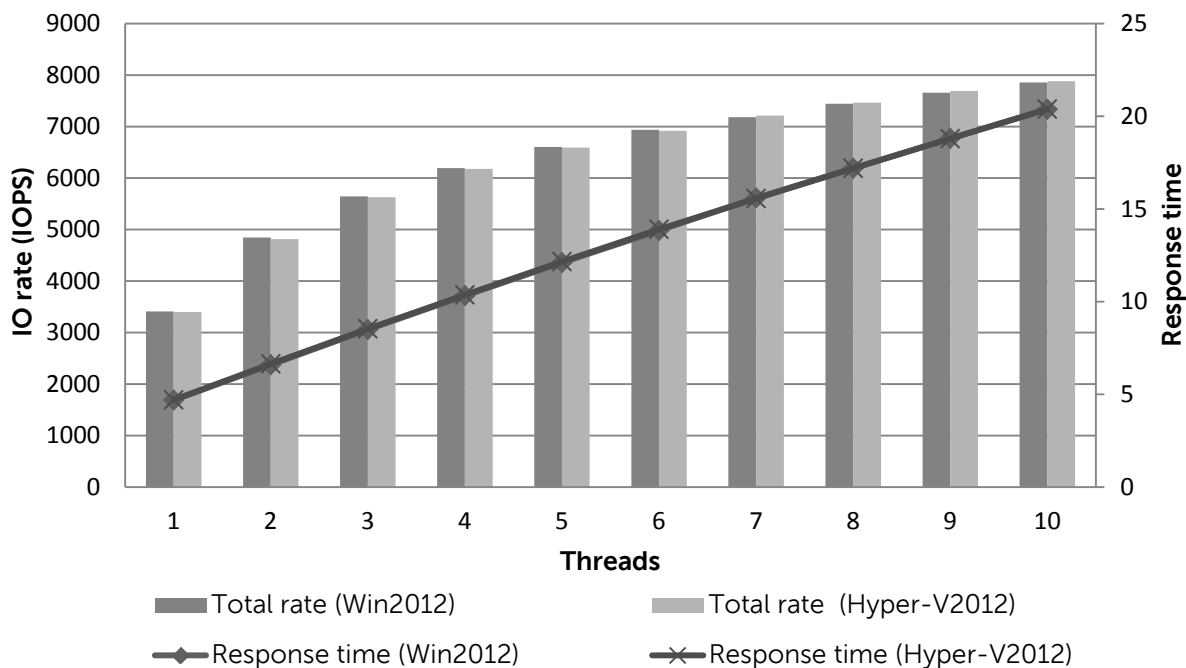


Figure 5 Performance Scaling (threads) for small block random workload in physical and virtual environments

4.1.1.2 Scaling the number of servers on the EqualLogic PS-M4110 storage solution

The M1000e PowerEdge blade chassis with M620 servers allows starting with a small deployment of one or more M620 blade servers and, depending on the growth and demand, expanding the M1000e blade chassis with up to 12 blade servers and two M4110 EqualLogic blade storage arrays. This ability to increase the number of servers depending on the configuration requirements makes the M1000e blade chassis a good fit for many types of deployments.

Figure 6 below demonstrates scalability for a random small block workload generated when the number of servers in the blade chassis is increased from one to eight. The Vdbench workload generation tool was run against two volumes on both the Windows Server 2012 and Hyper-V environments. The I/O performance of the Windows Server 2012 and Hyper-V virtual machine is very similar and scales linearly up to eight servers. The latency of the IOPS generated can also be seen in Figure 6.

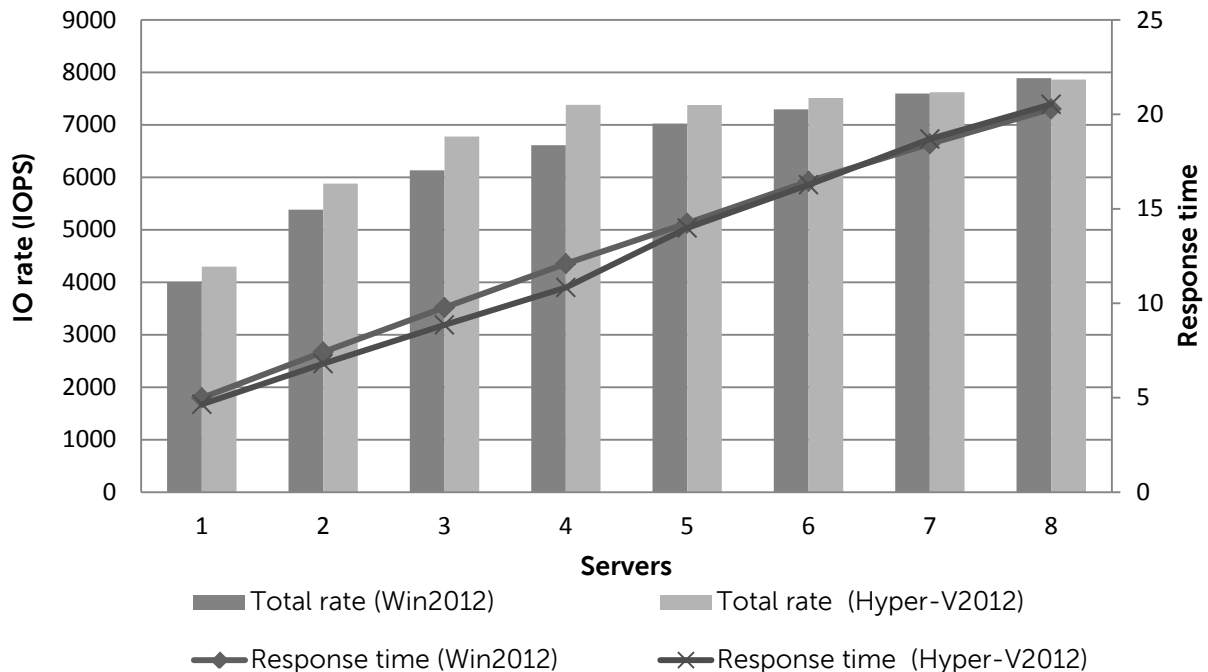


Figure 6 Performance Scaling (servers) for small block random workload in physical and virtual environments

4.2 Scenario 2: Converged infrastructure environment

The focus for this test scenario was to demonstrate the DCB capabilities of the Dell blade solution and to show how DCB with ETS can manage the bandwidth of multiple traffic classes while ensuring iSCSI traffic is lossless. Vdbench was used to generate iSCSI traffic from the blade server to the storage array and iPerf was used to simultaneously generate LAN traffic from the blade server to the external rack servers (to simulate clients reading data from the server). The test used large block sequential iSCSI traffic which is bandwidth intensive and similar to traffic generated by backups, audio or video streaming, and other multimedia applications.

The test used VLAN 10 for iSCSI traffic and VLAN 20 for LAN traffic. The ETS setting for the MXL switches was at 50% for iSCSI traffic and 50% for other types of traffic (LAN traffic).

Figure 3 provides a logical diagram of the configuration showing the external R620 servers and the blade chassis with all of its components.

4.2.1 Results and analysis

Figure 7 shows the throughput generated for iSCSI traffic alone, LAN traffic alone, and finally converged traffic. This shows how the DCB and ETS settings guaranteed the bandwidth for iSCSI and LAN traffic for the converged network configuration. The iSCSI traffic alone used 70% of the theoretical line rate of the 10Gb network while the LAN traffic was able to use above 85%.

Note: The calculations used to figure the theoretical line rate are as follows:
Line rate for two 10 Gb ports= $2 * (10\text{Gbit} / (8 * 1024)) = 2560 \text{ MBps}$

The results of the converged network test shows that the bandwidth is divided equally between iSCSI and LAN traffic types because of the 50%-50% ETS settings on the switch. This demonstrates that ETS was properly enforcing the configured bandwidth settings. No retransmits or dropped packets were seen during the test which shows that the DCB and ETS settings prevented any loss of packets for the lossless iSCSI traffic class.

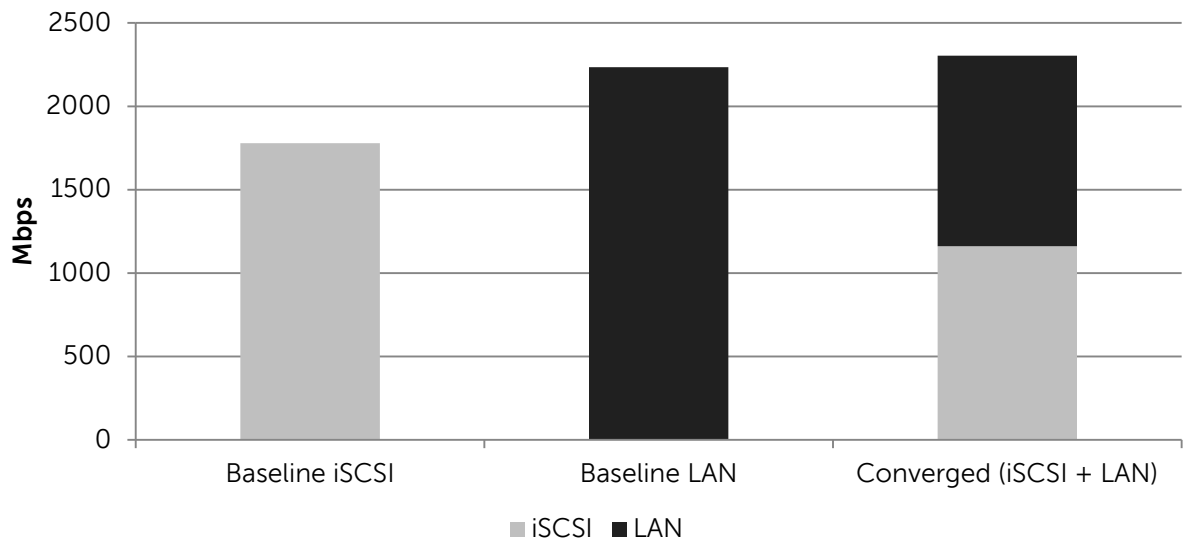


Figure 7 Sequential read performance with LAN and iSCSI traffic

4.3 Scenario 3: Replication to remote EqualLogic storage

The purpose of this test scenario was to see if replication time to an external site was affected by LAN and SAN traffic within the blade chassis.

Figure 3 provides a logical diagram showing the Dell blade solution connected with the external network and the remote EqualLogic PS Series arrays while Figure 2 provides additional details for the remote array connectivity. Two PowerEdge R620 rack mount servers were used to simulate LAN traffic on the blade solution. DCB was enabled with the same settings as used in Scenario 2.

To simulate data replication to a remote site for backup or disaster recovery, four volumes were configured for replication. The replication was manually started and the time required to complete the replication was measured for three scenarios:

- Four volumes replicated while there was no other network traffic on the Dell blade solution (used as the baseline).
- Four volumes replicated simultaneously with SAN traffic.
- Four volumes replicated simultaneously with SAN and LAN traffic.

Each of these tests was run five times and the average of the five runs was analyzed.

iSCSI traffic was generated by running Vdbench from four blade servers, each with a single volume with a mix of 75% small block random iSCSI workload and 25% large block sequential iSCSI workload.

LAN traffic was generated from the two external rack mounted PowerEdge servers connected to the PowerEdge M620 blade servers. Each external server simulated LAN traffic using the iPerf load generation tool.

4.3.1 Results and analysis

Figure 8 shows the proportional replication time for both tests compared to the baseline of time taken for four volume replication without any iSCSI or data traffic. The 4 volume replication baseline is shown at 100% in the below graph.

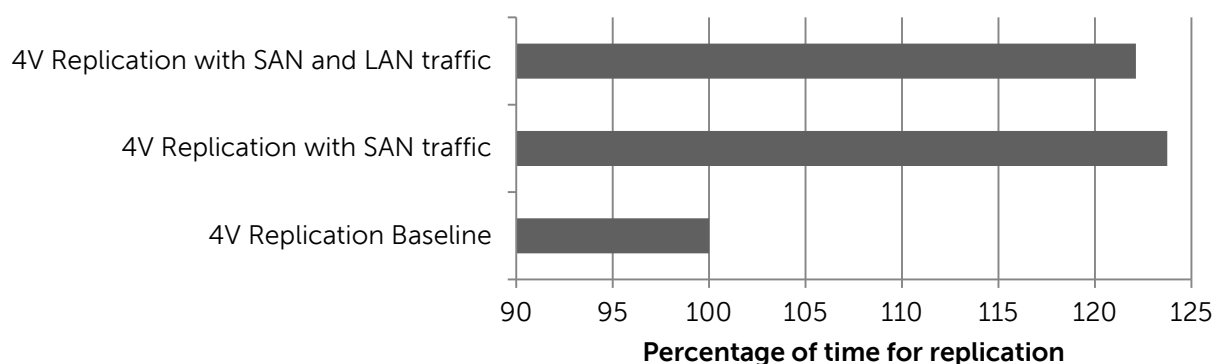


Figure 8 Difference in time taken for replication

When the storage only workload was run, replication time for the 4 volumes increased by about 22.1% on average.

When a converged traffic workload was run, replication time for the 4 volumes increased by about 23.7% on average, which is very similar to the scenario with storage traffic only.

Note: The test did not use a WAN emulator or distance simulator, so depending on the actual distance and link speeds, the time taken for replication may vary.

Note: The external network is not DCB enabled, so replication traffic exiting the blade chassis is no longer in a lossless queue.

During the replication activity, the storage I/O on the SAN was reduced by 25% with about the same percentage increase in response time. This was an expected behavior considering the system was running

near 80% of its max I/O capability and then replication activity was added. With a low workload, the impact would be minimal. Dell recommends scheduling replication during times where there is less LAN and SAN activity, such as at non-work hours to prevent an impact on performance.

4.4 Scenario 4: Comparison of RAID 10 versus RAID 50

This scenario compared the performance of RAID 10 and RAID 50 by useable storage capacity for each RAID type and by performance for each tested workload. The two M4110 EqualLogic storage arrays were configured first as RAID 10 and then as RAID 50. Four volumes were created per blade server. A total of four blade servers were running workloads. The number of Vdbench threads was simultaneously incremented from 1 to 10 threads on each of the 4 servers.

The same workload profiles that were used previously were used for this test as well:

- A. Small block random (8K block size, 67% read/33% write, random access)
- B. Large block sequential read (256K block size, 100% sequential read access)
- C. Large block sequential write (256 block size, 100% sequential write access)

These workloads were run in both Windows Server 2012 and Windows 2012 Hyper-V environments.

4.4.1 Results and analysis

Due to the method of disk redundancy used, the storage capacity available with RAID 50 is higher than with RAID 10 for the same number and size of disks. The useable capacity for RAID 10 was 40% lower compared to RAID 50, as is represented in Figure 9 (The storage capacity is calculated including hotspares).

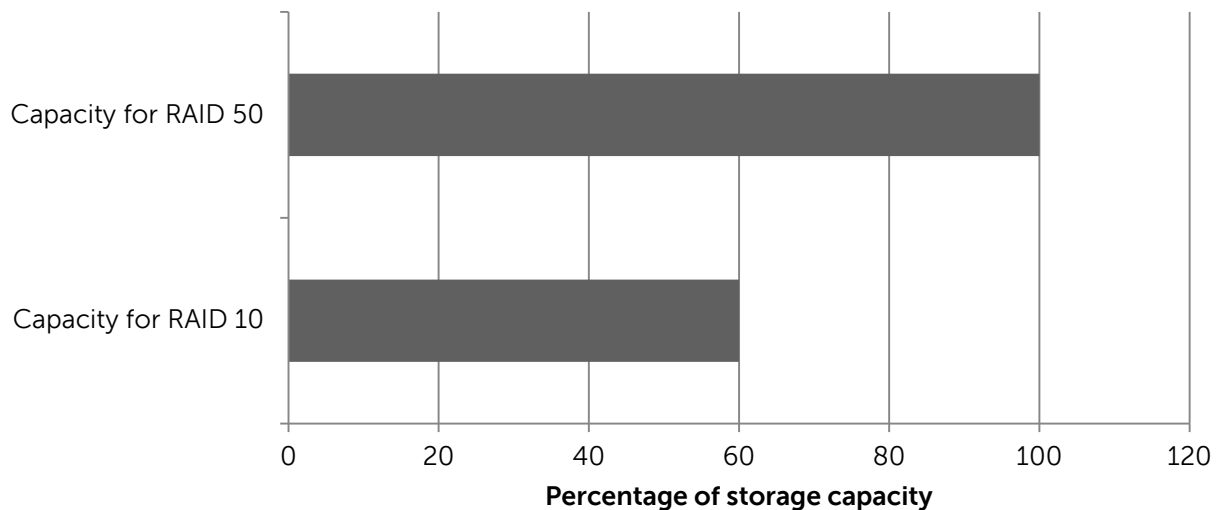


Figure 9 Comparison of storage capacity between RAID 10 and RAID 50



Volumes configured as RAID 10 showed a higher I/O performance and lower latency compared to RAID 50 for the same small block random workload, as illustrated in Figure 10. RAID 10 provided a 20% to 40% higher I/O rate and 17% to 63% lower response time versus RAID 50.

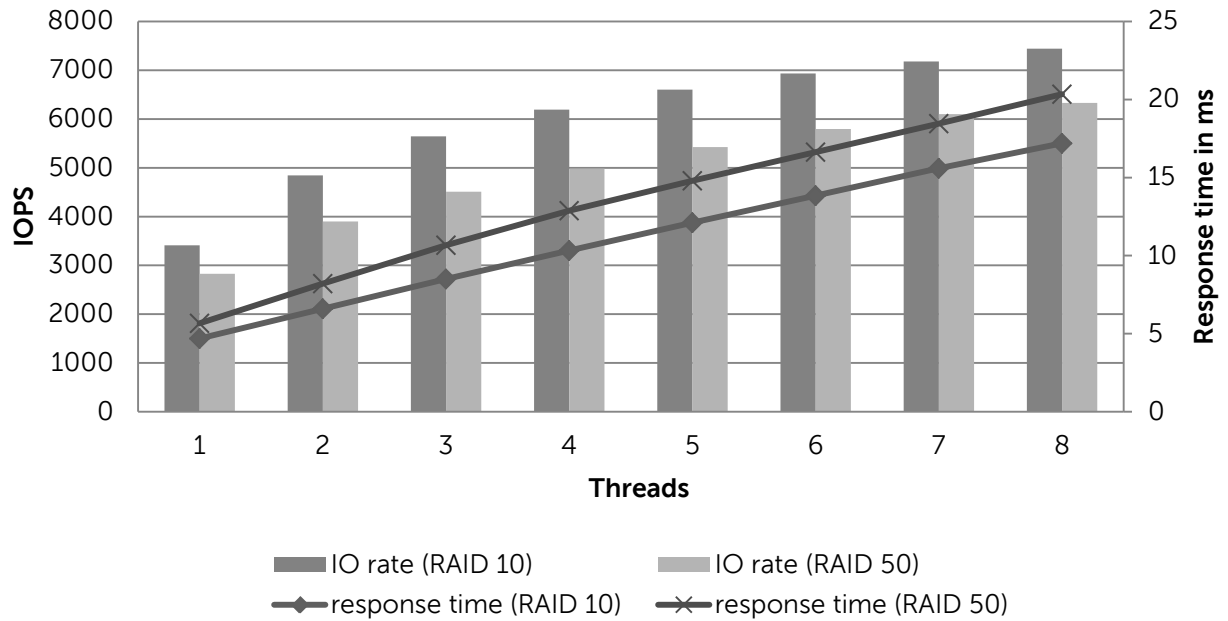


Figure 10 Performance comparison between RAID 10 and RAID 50 with small random workload

For the large block sequential read and write workloads, a performance difference of less than 10% between RAID 10 and RAID 50 was observed.

The tests conclude that RAID 10 is suited for environments expecting a higher percentage of small block random traffic, otherwise configure RAID 50 for maximum capacity on the storage arrays.

5 Conclusions and best practice recommendations

The analysis of work thread scalability for the blade solution on physical Windows Server 2012 and virtual Hyper-V 2012 environments indicates superior performance for applications relying on small block random workloads. Scalability in regards to increasing the blade servers to accommodate growth for an infrastructure is linear. The performance for small block random workload shows a linear increase in IOPS with an increase in the number of servers generating the workload requests. Overall, the scalability analysis demonstrates the capability of the blade solution to scale as needs increase.

Enabling DCB on the Force10 MXL network switch and Broadcom 57810 network adapters allows the Dell blade solution to converge LAN and SAN traffic onto a single network fabric while prioritizing mission critical iSCSI SAN traffic for application servers. A converged network infrastructure can lead to reduced infrastructure cost and simpler deployment and administration.

Replication activities may cause lower storage I/O performance and higher latency. This may in turn, affect application response times. Higher application server workloads within the blade chassis can affect available network bandwidth, and therefore replication time as well, resulting in a longer duration to complete the replication. Dell recommends scheduling replication during times where iSCSI SAN traffic is minimized, such as during non-work hours or hours of low application server activity.

Finally, use storage space according to workload requirement and configure RAID types on the storage accordingly. The tests conclude that RAID 10 is suited for environments expecting a higher percentage of small block random traffic, otherwise configure RAID 50 for maximum capacity on the storage arrays.



A Configuration details

A.1 Dell Force10 MXL blade switch setup

These steps are for configuring the Dell Force10 MXL blade switch with switch firmware version FTOS Version: 8.3.16.2. These instructions are for informational purposes only; for the most current configuration guide for the Dell Force10 MXL blade switch, refer to the switch configuration guide located here: <http://en.community.dell.com/dell-groups/dtcmmedia/m/mediagallery/20279157/download.aspx>.

Refer to Switch configuration guides (SCG) for current information on switch configurations for EqualLogic arrays. They are available at:
<http://en.community.dell.com/techcenter/storage/w/wiki/4250.switch-configuration-guides-by-sis.aspx>.

Table 2 Dell Force10 MXL blade switch configuration steps

Configuration Steps	Configuration Commands
1) Enable switch port	<pre>FTOS(conf)#interface range tengigabitethernet 0/1 - 32 , tengigabitethernet 0/41 - 56 FTOS(conf-if-range-te-0/1-32,te-0/41-56)#switchport FTOS(conf-if-range-te-0/1-32,te-0/41-56)#spanning-tree rstp edge-port FTOS(conf-if-range-te-0/1-32,te-0/41-56)#no shutdown FTOS(conf-if-range-te-0/1-32,te-0/41-56)#exit</pre>
2) Configure LAG	<pre><u>Configure Port Channel for LAG:</u> FTOS(conf)#interface Port-channel 1 FTOS(conf-if-po-1)#mtu 12000 FTOS(conf-if-po-1)#switchport FTOS(conf-if-po-1)#no shutdown FTOS(conf-if-po-1)#exit <u>Configure QSFP port for LAG:</u> FTOS(conf)#interface range fortyGigE 0/33 , fortyGigE 0/37 FTOS(conf-if-range-fo-0/33,fo-0/37)#mtu 12000 FTOS(conf-if-range-fo-0/33,fo-0/37)#no shutdown FTOS(conf-if-range-fo-0/33,fo-0/37)#flowcontrol rx on tx off FTOS(conf-if-range-fo-0/33,fo-0/37)#port-channel- protocol lacp FTOS(conf-if-range-fo-0/33,fo-0/37-lacp)#port-channel 1 mode active FTOS(conf-if-range-fo-0/33,fo-0/37-lacp)#exit FTOS(conf-if-range-fo-0/33,fo-0/37)#exit FTOS (conf)#exit</pre>



3) Configure DCB (DCB is enabled by default.)	<u>Enable DCB:</u> FTOS# configure FTOS(conf)# dcb enable FTOS(conf)# dcb stack-unit all pfc-buffering pfc-ports 56 pfc-queues 2 FTOS(conf)# exit FTOS# copy running-config startup-config FTOS# reload
	<u>Create Tagged VLAN:</u> FTOS(conf)# interface vlan 10 <i>(For iSCSI/SAN traffic)</i> FTOS (conf-if-vl-10)# tagged tengigabitethernet 0/1-32 FTOS (conf-if-vl-10)# tagged tengigabitethernet 0/41-56 FTOS (conf-if-vl-10)# tagged port-channel 1 FTOS (conf-if-vl-10)# exit FTOS(conf)# interface vlan 20 <i>(For LAN traffic)</i> FTOS (conf-if-vl-20)# tagged tengigabitethernet 0/1-32 FTOS (conf-if-vl-20)# tagged tengigabitethernet 0/41-56 FTOS (conf-if-vl-20)# tagged port-channel 1 FTOS (conf-if-vl-20)# exit
	<u>Configure priority group and policies:</u> FTOS(conf)# dcb-input pfc FTOS(conf-dcb-in)# pfc priority 4 FTOS(conf-dcb-in)# exit FTOS(conf)# priority-group iSCSI FTOS(conf-pg)# priority-list 4 FTOS(conf-pg)# set-pgid 1 FTOS(conf-pg)# exit FTOS(conf)# priority-group OTHER FTOS(conf-pg)# priority-list 0-3,5-7 FTOS(conf-pg)# set-pgid 2 FTOS(conf-pg)# exit FTOS(conf)# dcb-output ets FTOS(conf-dcb-out)# priority-group iSCSI qos-policy iSCSI FTOS(conf-dcb-out)# priority-group OTHER qos-policy OTHER FTOS(conf-dcb-out)# exit FTOS (conf)# service-class dynamic dot1p
	<u>Apply priorities to switch port:</u> FTOS(conf)# interface range tengigabitethernet 0/1 - 32 , tengigabitethernet 0/41 - 56 FTOS(conf-if-range-te-0/1-32,te-0/41-56)# dcb-policy input pfc FTOS(conf-if-range-te-0/1-32,te-0/41-56)# dcb-policy output ets FTOS(conf-if-range-te-0/1-32,te-0/41-56)# protocol lldp FTOS(conf-if-range-te-0/1-32,te-0/41-56-lldp)# exit FTOS(conf-if-range-te-0/1-32,te-0/41-56)# exit



	<pre> FTOS(conf)#interface range fortyGigE 0/33 , fortyGigE 0/37 FTOS(conf-if-range-fo-0/33,fo-0/37)#dcb-policy input pfc FTOS(conf-if-range-fo-0/33,fo-0/37)#dcb-policy output ets FTOS(conf-if-range-fo-0/33,fo-0/37)#exit FTOS(conf)#exit </pre>
	<p><u>Configure ETS values:</u></p> <pre> FTOS#configure FTOS(conf)#qos-policy-output iSCSI ets FTOS(conf-qos-policy-out)#bandwidth-percentage 50 FTOS(conf-qos-policy-out)#exit FTOS(conf)#qos-policy-output OTHER ets FTOS(conf-qos-policy-out)#bandwidth-percentage 50 FTOS(conf-qos-policy-out)#exit </pre>
4) Configure external OSPF network	<pre> FTOS#configure FTOS(conf)#interface tengigabitethernet 0/44 FTOS(conf-if-te-0/44)#ip address 192.168.1.1/30 FTOS(conf-if-te-0/44)#mtu 9216 FTOS(conf-if-te-0/44)#no shutdown FTOS(conf-if-te-0/44)#exit FTOS(conf)#interface Vlan 10 FTOS(conf-if-vl-10)#ip address 10.10.10.1/24 FTOS(conf-if-vl-10)#mtu 9216 FTOS(conf-if-vl-10)#no shutdown FTOS(conf-if-vl-10)#exit FTOS(conf)#interface Vlan 20 FTOS(conf-if-vl-20)# ip address 10.10.20.1/24 FTOS(conf-if-vl-20)#mtu 9216 FTOS(conf-if-vl-20)#no shutdown FTOS(conf-if-vl-20)#exit FTOS(conf)#router ospf 1 FTOS(conf-router_ospf-1)#network 10.10.10.0/24 area 0 FTOS(conf-router_ospf-1)#network 10.10.20.0/24 area 0 FTOS(conf-router_ospf-1)#network 192.168.1.0/30 area 0 FTOS(conf-router_ospf-1)#exit </pre>
5) Configure OSPF	<pre> FTOS#configure FTOS(conf)#router ospf 1 FTOS(conf-router_ospf-1)#network 10.10.10.0/24 area 0 FTOS(conf-router_ospf-1)#network 10.10.20.0/24 area 0 FTOS(conf-router_ospf-1)#network 192.168.1.4/30 area 0 FTOS(conf-router_ospf-1)#exit FTOS(conf)#exit FTOS # </pre>



6) Save Configuration	FTOS#copy running-config startup-config
-----------------------	---

A.2 Dell PowerEdge M620 blade server

Table 3 Dell PowerEdge M620 server components:

Dell PowerEdge M620 server components.	Version / Model number
BIOS version	Dell Inc. 1.4.9
OS version	Microsoft Windows Server 2012 Datacenter (6.2.9200 Build 9200)
HIT kit version	Version 4.5

A.3 Broadcom 57810 network adapter:

Table 4 Broadcom 57810 network adapter firmware and driver:

Broadcom 57810	Version / Model number
Firmware version	7.4.8
Driver family version	7.4.29.0
iSCSI Driver version	7.4.4.0
NDIS Driver version	7.4.23.0

The following steps were used to configure the Broadcom 57810S blade Converged Mezzanine Card on the blade servers:

1. Install Broadcom Advance Control Suite (BACS) on the blade server.
2. Go to Configuration for the Network adapter and under Resource Reservation click Configure and enable Ethernet/NDIS and iSCSI protocols.
3. This creates two functions for the 10 GbE port, NDIS VDB client (for LAN traffic) and iSCSI Adapter (for iSCSI/SAN traffic).
 - a. NDIS VDB client
 - i. Configure the IP address and other required parameters.
 - ii. Assign the VLAN ID used for LAN traffic (VLAN 20).
 - b. iSCSI Adapter
 - i. Configure IP address for the iSCSI Adapter through BACS.



- ii. Assign the VLAN ID used for iSCSI Traffic (VLAN 10).
- iii. Enable Jumbo frames on the iSCSI adapter by setting the value of MTU to 9216.
- iv. The Broadcom 57810S converged Network adapter is by default in willing mode for DCBx and will accept the configuration from the MXL switch.

A.4 Dell EqualLogic PS-M4110 storage arrays

Table 5 Dell EqualLogic M4110 storage array components

Dell EqualLogic M4110 storage array components	Version / Model number
Firmware version	Storage Array Firmware V6.0.1
RAID configuration	RAID 10 (unless specified)

A.5 Dell PowerEdge M1000e chassis

Table 6 Dell PowerEdge M1000e Chassis

Dell PowerEdge M1000e chassis components	Version / Model number
CMC firmware	4.20



Additional resources

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

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

Referenced or recommended Dell publications:

- Dell EqualLogic Configuration Guide:
<http://en.community.dell.com/dell-groups/dtcmedia/m/mediagallery/19852516/download.aspx>
- Dell Force10 MXL Switch Configuration Guide
<http://en.community.dell.com/techcenter/storage/w/wiki/4250.switch-configuration-guides-by-sis.aspx>

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

- <http://dell.to/sM4hJT>





This white paper is for informational purposes only. The content is provided as is, without express or implied warranties of any kind.