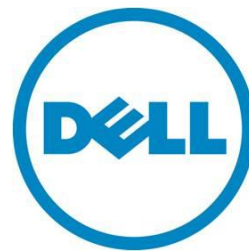# Microsoft® SQL Server® Database Consolidation using Dell™ EqualLogic™ PS-M4110 blade arrays, PowerEdge™ M620 blade servers and Force10 MXL Switches

This whitepaper describes the capabilities of PS-M4110 blade arrays in consolidating SQLServer databases in virtualized environments

Gopu Narasimha Reddy

Dell Database Solutions Engineering

January 2013

Dec 2012| Rev 1.0

Microsoft SQL Server Database Consolidation using Dell EqualLogic PS-M4110 blade arrays, PowerEdge M620 blade servers and Force10 MXL Switches

Contents

Executive summary

In order to reduce overall IT costs, many customers use virtualized platforms to run their workloads. This technical paper describes the pre-engineered virtualization platform designed for SQL Server database consolidation and rapid provisioning of SQL database instances. This reference configuration combines the core capabilities of Windows Server®, Hyper-V® and uses Dell PowerEdge blade servers, EqualLogic blade storage (PS M4110) and high performing Dell Force10 10Gbps blade switches at the backend.

Objectives

The main objectives of this technical paper are:

- Understand the Dell EqualLogic PS-M4110 blade storage capabilities to run the SQL workloads in Virtualized environments.

- Test it to see how many SQL VMs can be consolidated using PS-M4110 and PowerEdge M620 blades.

- Provide reference architecture with guidelines for consolidating SQL databases using Dell EqualLogic PS-M4110, PowerEdge M620 blades and Force10 MXL IOMs.

Introduction

Workload consolidation is a popular technique used to reduce the overall IT costs for customers. Many customers are migrating their workloads from physical to virtualized platforms to reduce the hardware and operating costs, and to increase the utilization of existing hardware.

As a popular database, many times customers have numerous SQL Server instances/databases, and sprawl is becoming problem. Dell, as an End-to-End solution provider, has introduced many pre-engineered and validated virtualized platforms for server consolidation. In this paper, we will discuss one such virtualized platform for SQL Server database consolidation where all the hardware components exist within a single blade M1000e blade chassis.

Overview of Dell Reference Architecture for SQL Server Database Consolidation

The virtualized reference architecture proposed for SQL Server database consolidation is built using the following hardware and software components.

Blade Chassis: We used the PowerEdge M1000e modular blade enclosure to host the PowerEdge M620 blade servers, EqualLogic blade storage and MXL SAN switches. M1000e can take advantage of its world-class design and provides efficiency, effortless scalability, and powerful yet easy to use management tools.

Servers: We used PowerEdge M620 blades in the proposed referral architecture. M620 is a 2-socket half-height blade server powered by Intel® Xenon® E5-2600 processors, designed for maximum performance for various workloads such as database, exchange and virtual environments. It supports

up to 786 GB Memory when all the memory channels are populated with 32G memory DIMMs. The following table lists the hardware details of blade servers used for the current reference architecture.

Table 1.   Hardware components: M620 Blade configuration

Components	Details
Server Model	2* PowerEdge M620
Processors per blade server	2* Intel Xeon E5-2643 @3.3 GHz,4 Cores
Memory	128 GB (16x 8GB, DDR3 @ 1600 Mhz)
Internal storage	2* 146G 15,000 RPM SAS drives configured in a RAID1
Raid Controller	1* H310 integrated Mini Raid Controller

Storage: - Dell EqualLogic recently introduced the PS-M4110 10G-blade storage array that is a scalable and easy-to-manage blade array, with redundant storage controllers designed to fit inside M1000e blade chassis. The introduction of this blade array enables a fully virtualized solution that integrates storage, servers and networking within one blade chassis.

The EqualLogic PS-M4110 blade array is available in four different configurations, which enables a true data-center-in-a-box solution to address a variety of environments including highly virtualized environments. The following table briefly describes the four PS-M4110 array configurations.

Table 2.   PS-M4110 Configuration Details

Model	Differentiation Details

PS-M4110E	Drives: 7.2K RPM NL-SAS 2.5 inch
	Capacity:
	  14 TB Per Array
	  28 TB Per Group (with 2 Arrays in one Group)
	  56 TB Per M1000e Chassis (with 4 Arrays in two Groups)
	Controllers: Single or Dual controller Option

PS-M4110X	Drives: 10K RPM SAS 2.5 inch
	Capacity:
	  12.6 TB Per Array
	  25.5 TB Per Group (with 2 Arrays in one Group)
	  50.4 TB Per M1000e Chassis (with 4 Arrays in two Groups)
	Controllers: Single or Dual controller Option

PS-M4110XV	Drives: 15K RPM SAS 2.5 inch
	Capacity:
	   4.2 TB Per Array
	   8.4 TB Per Group (with 2 Arrays in one Group)
	   16.8 TB Per M1000e Chassis (with 4 Arrays in two Groups)
	Controllers: Single or Dual controller Option

PS-M4110XS	Drives: 10K RPM SAS & SSD 2.5 inch
	Capacity: 6* 900G 10K Drives and 5* 400G SSD drives
	   7.4 TB Per Array
	   14.8 TB Per Group (with 2 Arrays in one Group)
	   29.6 TB Per M1000e Chassis (with 4 Arrays in two Groups)
	Controllers: Dual controller Option Only

Within a single M1000e blade chassis, we can install up to four PS-M4110 blade arrays grouped in to two groups each group having two blade arrays. In the current reference architecture, we used two PS-M4110XV arrays each fully populated with 14* 300G 15000 RPM SAS 2.5-inch SAS drives. In each array, two hard drives are used as hot spares. The two arrays are grouped into a single storage group. When both the arrays configured in a RAID 50 storage group, ~ 5.2TB of usable space is created. The following table describes the configuration details of storage used in the current reference architecture.

Table 3.   Hardware components: PS- M4110 Blade Array configuration

Components	Details
Storage Arrays	2* PS-M4110 10G arrays with redundant 10G iSCSI storage controllers
Disks per storage Array	14* 300G 15000 RPM SAS disks in each Array
Hot Spares	2* hot spares on each Array.
Usable RIAD storage Capacity	Around 5.2 TB with RAID5

The PS-M4110 receives power from, and makes all its network connections through, the M1000e modular blade enclosure. For increased availability, the 10GbE Ethernet ports on both PS-M4110 control modules are internally and automatically connected to each redundant M1000e IO module (IOM) of the configured fabric; one port is active and other port is passive. For example, if a PS-M4110 is configured for Fabric B, and if both the B1 and B2 IOMs are installed, the Ethernet ports from each

control module are connected to both B1 and B2 IOMs. This provides four potential Ethernet paths from the switches to the array. However, only one of those Ethernet paths is active at any given time. Figure 1 shows the basic connectivity of single PS-M4110 storage array when B1 and B2 IOMs are populated.

Figure 1.   PS-M4110 Storage Array Connectivity
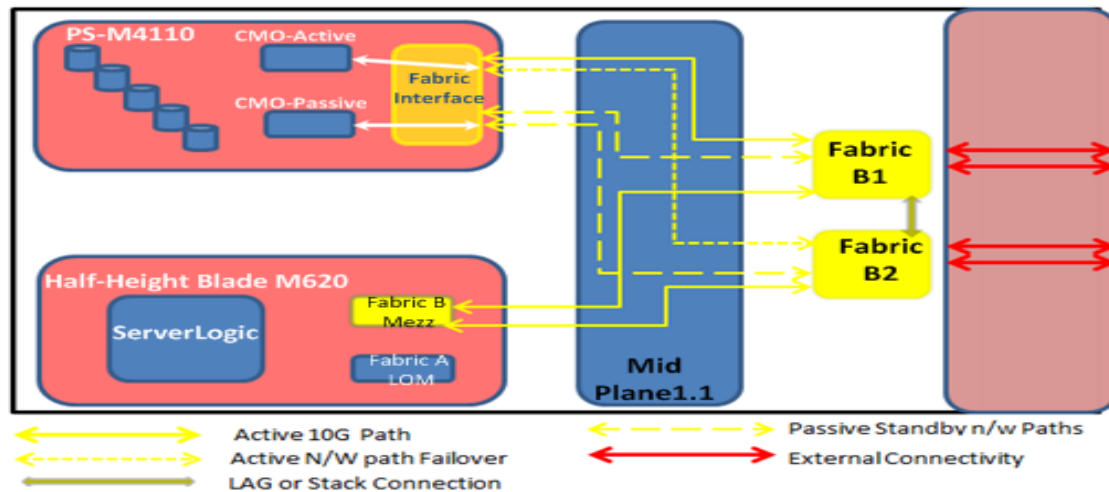
In a dual control module array, if the active control module fails, the secondary module automatically takes over and becomes active. The control module failover is transparent to applications.

Networking

The M1000e blade enclosure provides three fabrics for LAN & iSCSI traffic. Each fabric has two redundant I/O modules referred to as A1, A2, B1, B2, C1 and C2 respectively. The I/O modules can be populated with either 1G/10G iSCSI or 8G FC IO modules. In the current reference architecture, we used dedicated fabrics to separate the LAN traffic from the SAN traffic. Another alternative is to merge the LAN and SAN traffic using the Datacenter Bridging (DCB) feature through the same Mezzanine card.

Connecting to LAN

In current configuration, we used two Dell 10G Pass-Through switches in the A1 and A2 slots for LAN connectivity. This allows the blade hosts and the VMs to communicate with an external public network through a Top of Row network switch. We recommend that the LAN traffic is segregated into the following networks for better management:

- **Public and Management network**: used for Hyper-V host public connectivity.

- **Migration network**: used for live migration of VMs from one Hyper-V host to other.

- **Private network**: used for private network for the clustered Hyper-v Hosts.

- **VMs Public Network**: used for VMs public connectivity.

We used a Broadcom BCM57810 card in slot A of each blade server, which is a dual-port 10GbE card and supports NPAR, TOE, iSCSI offload, and more. As mentioned earlier, we used 10G Pass-

Through switches in A1 and A2 slots, and connected each respective blade ports from each IOM (A1 & A2) to a Top of Row switch.

By using the Broadcom Advanced Control Suit 4 (BACS4), we have teamed the two LAN ports on each Hyper-V host and created four virtual NICs for the networks as mentioned above and each VNIC is assigned an appropriate VLAN ID. On the Top-Of-Rack (TOR) switch, all of the above mentioned networks need to be segregated using specific Vlans with a corresponding IP address.

Connecting to SAN using Dell Force10 MXL Switches

The Dell Force10 MXL 10/40GbE switches are designed for M1000e blade servers to provide high-speed network connectivity. In conjunction with Mezzanine cards and network daughter cards (NDC), these switches can be configured to serve the complete LAN and SAN traffic (using Datacenter Bridging) of all the blade servers within the M1000e enclosure. However, in this configuration, we have used the Force10 MXL switches exclusively for the SAN traffic.

These switches are equipped with 2 fixed QSFP+ ports, two optional expansion slots for outgoing traffic and 32 internal (server-facing) ports. The two optional Flex I/O modules enhance the switch's flexibility and port scalability. For complete details on the Dell Force10 switches refer to the following link: dell.com/us/enterprise/p/force10-networking

Interconnecting the Dell Force10 MXL IOMs

When using a PS-M4110 inside an M1000e enclosure, the I/O modules are interconnected (Stacked or LAGged together). This interconnection is achieved by configuring LAG or stack on the appropriate switch ports. The LAG or Stack links must have sufficient bandwidth to handle the iSCSI traffic. For more details on how to LAG the MXL switches refer to the following link: en.community.dell.com/dell-groups/dtcmedia/m/mediagallery/20196885/download.aspx

In the current configuration, the two MXL switches are LAGged with each other as shown in Figure 2. We used the two-40GbE ports of the base-module on each switch and LAGged them together.

Figure 2.   Fabric Configuration
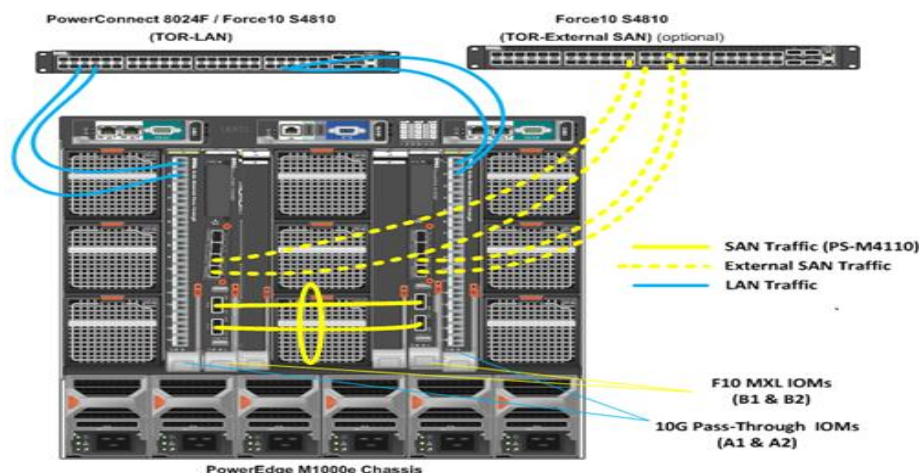
Licensing Costs on Virtualized Environments

Microsoft SQL Server 2012 offers two licensing options – one based on the number of cores (rather than the processors in a given machine) and the other based on number of Users or Devices.

For highly virtualized environments, it may be beneficial if all physical host cores are licensed with the Enterprise edition; this allows customers to deploy an unlimited number of virtual machines with reduced licensing costs. For more details on virtualized SQL Server 2012 licensing, refer to the following link: download.microsoft.com/download/7/3/C/73CAD4E0-D0B5-4BE5-AB49-D5B886A5AE00/SQL_Server_2012_Licensing_Reference_Guide.pdf

In this configuration, if we license all the eight physical host cores, we can run up to 8 virtual machines with a maximum supported number of virtual CPUs.

The Dell Difference

The proposed modular reference configuration has many advantages, and the following section highlights some of the tangible benefits the configuration offers.

Cost-effective Database Consolidation Platform

The proposed reference architecture is deployed on a single Dell blade chassis. This provides multiple advantages for the data center environment. Some of the green benefits of the configuration are:

- Reduced rack space/datacenter foot print
- Reduced power consumption

The proposed reference configuration also enables cost savings on the SQL Server software licenses. The high performing four core processors recommended in the configuration, along with the physical core-licensing scheme may be beneficial for deploying a highly dense consolidation environment with unlimited SQL Server virtual machines while incurring minimal software costs.

Easy and Flexible Installation

The database blades and the PS-M4110 storage arrays may be plugged into any of the available slots in the blade chassis. The whole database consolidation reference configuration may be brought up without any manual outside wiring, other than the customer environment integration.

Smooth, Scalable and Highly available infrastructure

The reference architecture allows the flexibility to have a scalable (Server and Storage) and highly available infrastructure by using clustered Hyper-v hosts. The M1000E blade chassis can accommodate up to sixteen M620 blades. If needed, additional external chassis can be added to add more servers to the fabric using the high bandwidth stacking options. Each M1000E enclosure can have up to four installed PS-M4110 arrays; each storage group can have a maximum of two PS-M4110 arrays. If needed, additional external PS-6110 arrays can be added to the group for enhanced scalability and performance; the peer storage architecture allows for seamless options and maximum scalability. For more details on the peer storage architecture, refer to the following link: equallogic.com/products/default.aspx?id=5823

Centralized Management

The M1000E Chassis Management Controller (CMC) enables centralized management of the whole infrastructure from a single console. It is a hardware and software solution that enables an IT

administrator to take inventory, perform configuration and monitoring tasks, remote power on/off blades and enable alerts for events on servers and components in the blade chassis. For more details on CMC, refer to the following link: en.community.dell.com/techcenter/systems-management/w/wiki/1987.dell-chassis-management-controller.aspx

Dell EqualLogic SAN tools at no additional Cost

Dell EqualLogic provides the Auto-Snapshot manager (ASM) that helps customers to manage SQL server databases at no additional cost.

Simplified SQL server Management using Dell EqualLogic Auto-Snapshot Manager

The Dell EqualLogic Auto-Snapshot Manager for Microsoft (ASM/Microsoft Edition) is an integration tool that helps with the administration of SQL Server environments at no additional cost to Dell EqualLogic Customers. The latest release introduces several new features including HIT Groups, Remote Installation and Upgrades, and additional PowerShell scripting tools that further enhance and enrich the installation, maintenance, and management of Microsoft SQL server environments with Dell EqualLogic iSCSI SANs. For more details on ASM/ME, refer to the following link: i.dell.com/sites/content/business/solutions/whitepapers/en/Documents/sql-server-data-protection.pdf

Setting up Highly available Modular Infrastructure with Hyper-v Cluster

The following sections describe how the Hyper-v hosts are configured for virtualized environments.

Clustered Hyper-V hosts

We used a PERC H310 integrated Mini Controller card to manage the locally attached M620 blade server disks. The two disks are configured for RAID 1, and presented to the host where the Windows 2008 R2 Operating System is installed. Hyper-V feature is installed on each host and are clustered using the Microsoft Windows 2008 R2 clustering feature to achieve high availability in case of any node failure.

Cluster Shared Volumes (CSV) Requirements

CSV volumes are used to store the Virtual Machines (VMs) .vhd files. CSV provides shared access to the disk and a storage path for I/O fault tolerance (dynamic I/O redirection). If the storage path on one node becomes unavailable, the I/O for that node is rerouted through a server message block (SMB) to another node. This feature can use any cluster communications network and further increases the need for high-speed networks. We created one single large volume of 1TB and presented it to all the nodes in the cluster. This volume is configured as CSV in the Failover clustering. All the VM related files (VHDs, Configuration files etc.) that belong to the VMs hosted on the cluster are stored on the single CSV.

Live Migration networking setup

When planning a Live Migration network configuration, it is recommended to either have a dedicated network card or use VLANs to isolate the live migration traffic from other LAN traffic. We

used a separate network for live migration in order to separate this traffic from other LAN traffic such as public, private etc.

Test Methodology

In order to show case the PS-M4110 database consolidation capabilities in virtualized environments, we considered the small-to-medium customers and conducted several tests and analyzed the results. We configured guest machines on the two PowerEdge M620 Hyper-V nodes, and installed SQL instances on each guest VM. The following table describes the hardware and software components used for the proposed reference architecture.

Table 4.   Highly available reference architecture details

Hardware Components	Details
Server	2 * PowerEdge M620
Processors	2* Intel Xeon Sandy Bridge CPUs E5-2643 @3.3 GHz,4 Cores per Server
Total Cores per Server	8
Total Logical Processors (HT Enabled)	16 per Server
Total Installed Memory	128GB @ 1600 Mhz per Server
Network Adapters	Fabric A : BroadCom BCM57810 Fabric B : BroadCom BCM57810
MultiPathing Software	Dell EqualLogic Hit Kit 4.0
MultiPathing Policy	Least Queue Depth(Default)
Blade Storage Array	2 * PS-M4110
Disks Hot Spares	28* 300G, 6Gbps 15k SAS drives 2 per PS-M4110 Array
Network IO modules	Fabric A1 & A2: Two Dell 10GbE Pass-Through IOMs Fabric B1 & B2: Two Force10 MXL IOMs
Blade Enclosure	M1000e with Mid plane version 1.1
Operating System	Windows 2008 R2 SP1 Enterprise Edition
Hypervisor	Microsoft Windows 2008 R2 hypervisor
Workload and Bench Mark	OLTP and Dell Quest TPC-E Bench Mark

VM Configuration

In the test configuration, the two PS-M4110 storage arrays were used to host the SQL databases present on all the guest VMs. Each guest VM is presented with 100 GB volume to store database data files and transaction log files. The following table describes the VM configuration for small-to-medium workloads.

Table 5.   Medium Guest Virtual Machine Configuration

Component	Medium VM
Processor	2 VCPU
Memory	4 GB
Storage	OS: 100 GB VHD on PS-M4110 Group Database: 100GB Volume on PS-M4110 Group.
Operating System	Windows 2008 R2 SP1 Enterprise Edition
Database size	25 GB
Database Software	SQL Server 2012 Standard Edition
Workload	TPC-E Benchmark
Users	20

We selected the TPC-E benchmark to stress the database because it has a more OLTP database-centric workload with enhanced schema complexity. The TPC-E transactions are more CPU & I/O intensive and as random as other TPC benchmarks. A workload of 20 users was selected as it stressed an individual VM configuration to its maximum efficiency. The same workload was used on all the VMs created.

Results and Analysis

Figure 3 shows the cumulated Database Transactions/Sec and Database Batch requests/sec achieved with VMs one through six.

Figure 3.   Database Transactions/Sec and Batch Request/Sec
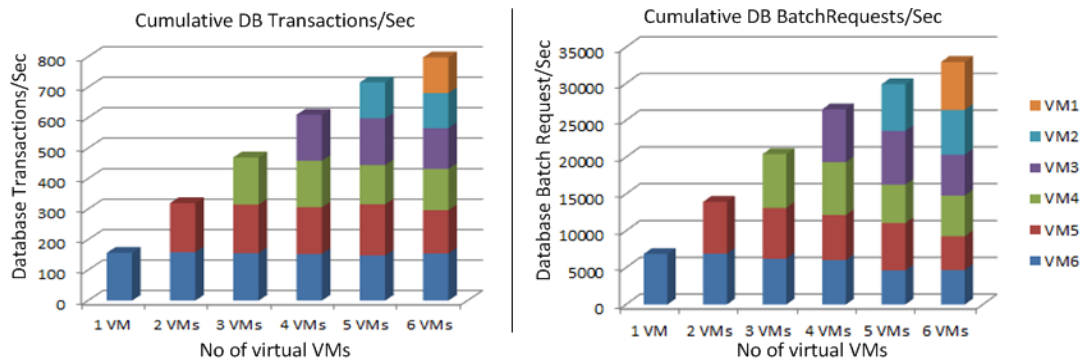
We observed that there is an almost linear scalability for Database Transactions/Sec and Batch Requests/Sec when adding VM 1 to 5. However, on adding the 6th VM, we noticed an overall reduction in the average Transactions/Sec and Batch Requests/Sec for the VMs. This led to the conclusion that we reached the maximum limit of VMs supported by the EQL group at the backend.

Using the SAN Head Quarters Monitoring tool, we observed that the EqualLogic storage group is delivering around 4000 Read IOPs with a latency of 9 ms and 1500 Write IOPS with 2ms latency when all six VMs run simultaneously.

Figure 4 shows how the Average Disk ms/Transfer behaved when adding multiple VMs.

Figure 4.   Disk Response time behavior
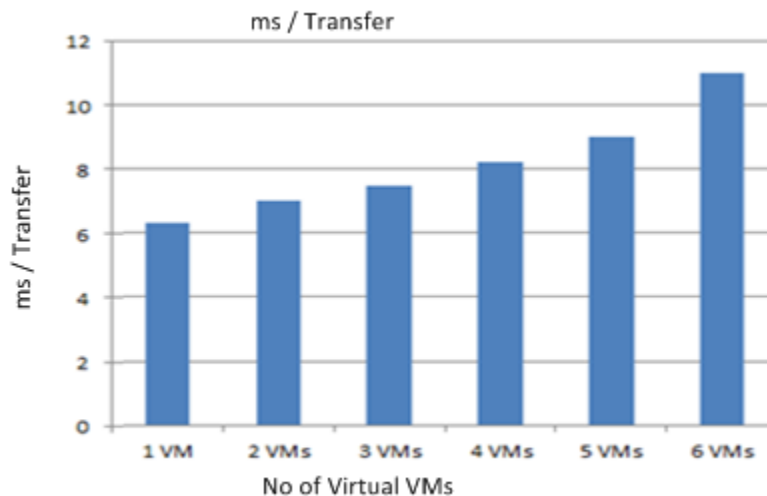
We observed that when one virtual machine is running, each disk transfer takes around 6 ms, and this number kept increasing as we added more virtual machines. When six virtual machines run simultaneously, each disk transfer takes more than 10ms with constant or reduced overall database transactions/sec.

 NOTE: In all of the above test scenarios, the same storage group formed with two PS-M4110 arrays is used to store all VM's .Vhd files and database data and log files.

Managing Virtualized Environments: Backups, Recovery & Monitoring

Dell continues to identify areas to expand its offerings such as monitoring, backup and recovery, and many more. There are various solutions to help recover the infrastructure in the desired granularity, and the following sections describe some of the Dell offerings for managing virtualized platforms.

Dell AppAssure for Backup and Recovery of Hosts, Guests and Applications

Dell AppAssure provides innovative capabilities for complete server, data, and application protection for complex physical, virtual and private cloud infrastructure. AppAssure 5 sets a new standard for unified data protection by combining backup, replication, and recovery in a single solution.

Dell AppAssure 5 offers the following technologies

- Live Recovery: an instant recovery technology that offers near-continuous access to data volumes on virtual or physical servers.

- Recovery Assure: allows automated recovery testing and verification of the backups. It provides 100% recoverability of applications and backups in virtual and physical environments.

- Universal Recovery: gives more flexibility in restorations. We can restore the backups from physical to virtual, virtual to physical, virtual to virtual, and physical to physical.

- Global Deduplication: Provides true global deduplication that dramatically reduces storage capacity requirements.

    For more details on this product, refer to the following link: appassure.com/server-backup-replication-and-recovery/windows-server-backup/

Dell Quest LiteSpeed for SQL Database Backups & Recovery

Dell Quest LiteSpeed product for SQL Server databases offers unique capabilities in database backup, compression, speed, control, and management. LiteSpeed compression technology reduces the storage requirements for database backups by ~85%. By using both deduplication and compression it can significantly reduce backup times and storage costs with smaller, more intelligently managed backups. It offers the following benefits to the customers:

- Fast and advanced data compression

- Object level recovery

- Transaction log reader

- Centralized management and reporting

- Robust encryption

For more details on this product, refer to the following link: quest.com/litespeed-for-sql-server/

Dell vFoglight for monitoring virtualized platforms

Dell Quest vFoglight is a best-in-class solution for performance monitoring and capacity management for virtual environments built with either Hyper-v or VMware®. With vFoglight, we can monitor the entire virtualized infrastructure through detailed graphical displays. With its automatic alerting system

and with expert advice and remediation, a customer can resolve issues 75% faster than with alternate solutions. vFoglight is the only multi-hypervisor virtualization management solution that combines infrastructure performance monitoring, capacity management, cost allocation, and service management in a single, extensible, and customizable solution. Some of the features and benefits that vFoglight can offers are:

- Asset and configuration tracking: the ability to track the movement of VMs and their configuration changes.

- Capacity trending, forecasting and alerting: resource consumption limits based on the historical growth rates to help plan for future needs.

- Chargeback and showback: the ability to allocate infrastructure costs helps us to see how groups and workloads are consuming resources.

- Detailed architectural representation: high-level, graphical displays highlight problem areas to help you quickly diagnose and fix problems in the virtual infrastructure.

For more details on this product, refer to the following link: quest.com/vfoglight/

EqualLogic SAN HeadQuarter for monitoring EqualLogic storage Groups

SAN HeadQuarter is a key Dell EqualLogic product that monitors, manages, and reports every aspect of multiple PS Series groups in a single system. EqualLogic SAN HQ provides consolidated performance and robust event monitoring across multiple EqualLogic groups so that you have a well-tuned EqualLogic SAN to meet and surpass your business requirements. In addition you can:

- Improve performance by identifying performance bottlenecks
- Effectively allocate group resources by identifying under-utilized assets within the SAN
- Improve availability by proactively identifying problems within your EqualLogic environment.
- Assist in future planning by identifying requirements for storage growth

Quick recovery and advanced data protection with Dell EqualLogic Auto-Snapshot Manager for Microsoft (ASM/ME)

The EqualLogic PS Series is tightly integrated with Microsoft Volume Shadow Copy Services (VSS), enabling comprehensive backup and recovery for SQL Server deployments. This integration is called the Auto Snapshot Manager for Microsoft (ASM/ME) and is available to all EqualLogic customers at no additional cost.

Utilizing Dell EqualLogic's Auto-Snapshot Manager for Microsoft (ASM/ME), administrators can make full point-in-time copies of one or many SQL Server databases simultaneously — accomplished online, usually in a few seconds, without any downtime or database performance impact. Administrators can create and manage SQL Server snapshots, clones or replicas for full or differential copies that can be instantly restored. This enables administrators to setup realistic testing and development environments quickly.

Database protection tasks can be scheduled to proceed automatically, limiting data loss, and scheduled to take place as frequently as needed while minimizing required disk space. Should an outage or disaster occur, the SQL Server environment can return to operation quickly.

Summary

Dell EqualLogic PS-M4110 blade storage array is an innovative product from Dell that fits inside the M1000e blade chassis. The PS-M4110 blade array enables a fully-virtualized solution that integrates storage, servers, and networking within one blade chassis. In this technical paper, a reference architecture was outlined and evaluated for Microsoft SQL Server database consolidation in virtualized environments. Within a single M1000e blade chassis, two PS-M4110 blade arrays were used and configured to host the VMs and the SQLServer databases. For medium OLTP workloads, this reference architecture was able to support up to five VMs (each VM stressed up to its maximum extent) with linear scalability in terms of transactions/second with response time not more than 10ms. The addition of more VMs caused a non-linear scalability, with slightly increased response time.

The EqualLogic PS-4110 blade array may be right choice for small and medium customers who want to consolidate SQL Server databases in a virtualized environment. This reference architecture can be scaled up to four PS-M4110 arrays (grouped into two groups) within a single M1000e array. For further storage needs, there is an option to add additional external EqualLogic storage arrays. The major advantages that the EqualLogic blade arrays bring into the virtualized modular implementations are:

- Data center in a box: an infrastructure that completely fits within a single M1000e blade chassis and therefore provides:
    - Reduced rack space/datacenter foot print.
    - Reduced power consumption.
    - No cabling required between storage & Hosts.
- Easy to setup and configure the storage in less time.
- Provides a scalable (Server and Storage) and highly available infrastructure using clustered Hyper-v hosts.
- EqualLogic SAN tools such as EqualLogic ASM/ME, SAN HeadQuarters are freely available for the customers at no additional cost.
- Cost-effective solution with lower software licensing costs.
- 40Gbps networking options with Force10 MXL.
- Centralized management using CMC.