CAPACITY STUDY OF A VIRTUAL WEB-SERVER FARM

USING DELL™ POWEREDGE™ BLADE SERVERS, DELL EQUALLOGIC™ STORAGE AND VMWARE vSPHERE 4

October 2009

Dell Virtualization Solutions Engineering
www.dell.com/virtualization

Feedback: solutionfeedback@dell.com

Contents

Executive Summary

This paper provides insight into the behavior and performance of a web server workload running in a virtualized environment using Dell™ PowerEdge blade servers, Dell EqualLogic™ PS Series arrays, and VMware vSphere™. The goal of the evaluation is to measure the capacity and demonstrate the suitability of the virtual infrastructure for demanding real-world applications. Various aspects such as dynamic resource scheduling, dynamic power management, high availability, and serviceability are explored in detail. The architecture of the virtual infrastructure is based on the Dell Business Ready Configurations for virtualization. By keeping the virtual infrastructure resource utilization at typical datacenter utilization, we observed that the test environment could host around 128 web server virtual machines (each with 300 concurrent users) with enough additional capacity to support dynamic resource scheduling, dynamic power management, virtual machine availability and serviceability.

1 Introduction

As the use of virtual machines (VMs) grows inside datacenters, it becomes increasingly important for administrators to understand how a virtual environment will behave under normal workloads and as it scales across several physical servers. In such an environment, the impact of common virtualization operations like dynamic resource scheduling becomes important considerations in the design of the overall architecture. This paper provides insights for these topics by deploying a common web server workload in a virtual infrastructure and documenting the observed results. The data presented in this paper may be used for capacity planning of similar workloads in a virtualized environment, or for understanding the general behavior of a virtual infrastructure consisting of hundreds of VMs.

This paper examines the behavior and capacity of the virtual infrastructure under normal operating conditions and realistic resource utilization levels. Virtual machines are created to represent a real world web server workload and the web server VMs are scaled, while keeping a typical datacenter CPU utilization level (around 60%). The number of virtual machines and the number of users connected to the web server are measured to represent the capacity of the virtual web server farm. Another important aspect of virtual infrastructures is the impact on the power efficiency of the Data Center, to that end this paper also provides insight into savings through the implementation of dynamic power management. Finally, the robustness of the solution is examined by simulating failure scenarios where VMs are restarted on different hosts due to a failure, or migrated due to maintenance scenarios.

Note that this paper is not meant to provide exact guidelines for all virtual architectures, nor is it meant to specify exact per-server virtual machine capacities or performance. Several factors such as workload fluctuations, architecture customizations, or specific user requirements can change the overall performance and scalability of a virtualized environment. For a specific sizing of your datacenter requirements, contact your Dell sales or service representative.

The architecture of the virtual web-server farm is based on Dell Business Ready Configurations. Dell Business Ready Configurations for virtualization provide a good starting point for customers who are looking for simple and adaptive virtual infrastructure for their applications. Details of the test environment are covered in the next section.

2 Test Environment

This section describes the specification and configuration of the test environment.

2.1 Specification

This section describes the details of the hardware, hypervisor, virtual machines, and workload used in the test environment. Figure 1 shows the high level architecture of the test environment. The configuration consists of a Dell blade chassis (M1000e), eight Dell PowerEdge™ M610 blade servers and four Dell EqualLogic PS6000XV arrays. The four arrays form a storage group giving a total capacity of 10.35 TB. The M1000e blade chassis is equipped with six Cisco 3130 I/O modules. In order to provide enough uplink ports to connect all 4 arrays, two intermediate Cisco 3750 rack mount switches are used.

2.1.1 Hardware Specification

Table 1 below provides the specifications of the test environment.

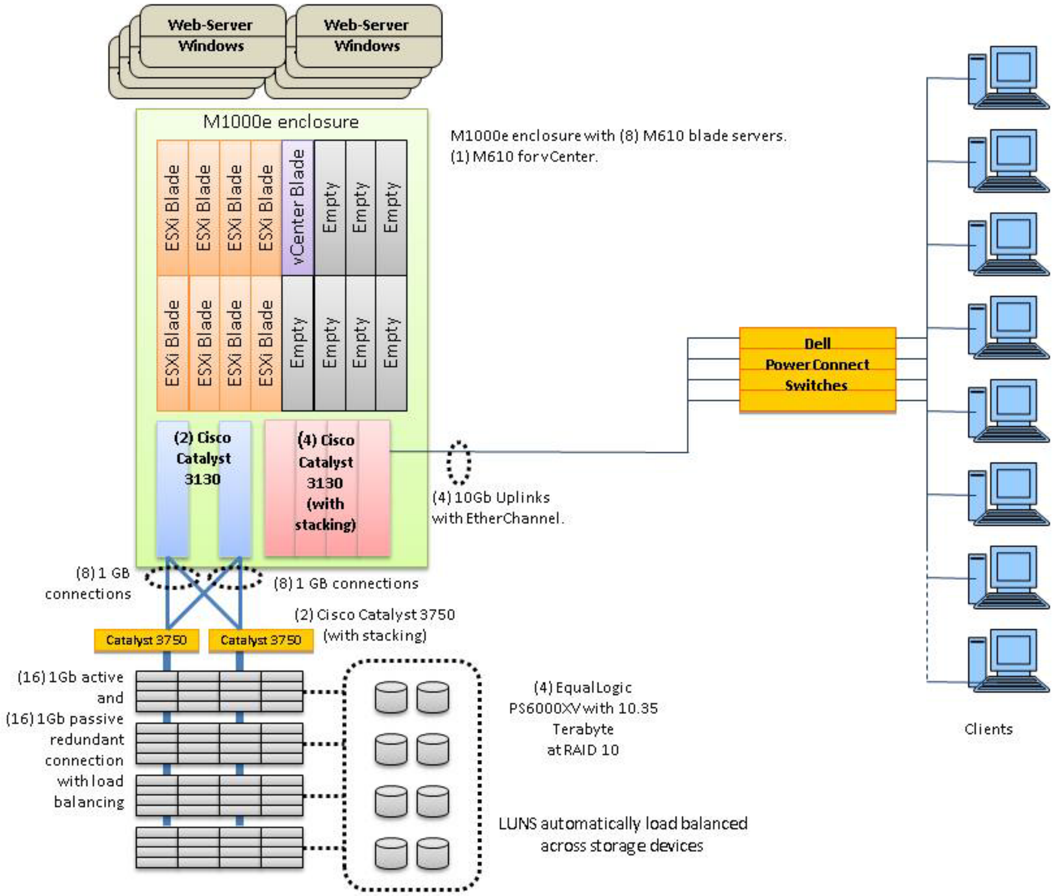

Figure 1: High level architecture of the test environment

Table 1: Hardware

Servers	
Server	Eight PowerEdge M610 blade servers
Processors per server	Two Intel® Xeon (Nehalem) E5520, 2.26GHz, 8M Cache Quad core processors
Memory per server	24 GB (6 x 4 GB, DDR3 DIMMS) 1066MHz DDR3
Network adapters per server	One Broadcom 5709 Dual Port LAN on Motherboard, Two Broadcom 5709 Dual Port I/O Mezzanine Cards
Hypervisor per server	ESXi 4.0.0 (Build 173507)
M1000e Enclosure Configuration	
I/O module for A1	Cisco M 3130X (two 10 Gb uplinks)
I/O module for B1	Cisco M 3130G with (4) 1000BAS-T SFP modules
I/O module for C1	Cisco M 3130X (two 10 Gb uplinks)
I/O module for A2	Cisco M 3130X (two 10 Gb uplinks)
I/O module for B2	Cisco M 3130G with (4) 1000BAS-T SFP modules
I/O module for C2	Cisco M 3130X (two 10 Gb uplinks)
Management	Chassis Management Controller (CMC)
KVM	Integrated Avocent keyboard, video, and mouse (iKVM) switch
Management Server	
Server	One PowerEdge M610 blade server

Processor	Two Intel® Xeon (Nehalem) E5520, 2.26GHz, 8M Cache Quad core processors
Memory	8 GB (12 x 4 GB, DDR3 DIMMS) 1066Mhz
Virtualization Management Software	VMware vCenter Version 4.0.0 (Build 162856)
Storage	
Storage Array	Four EqualLogic PS6000XV arrays
Drives	Each array has (16) 450 GB, 15K RPM SAS drives
RAID	RAID 10
Storage volumes	Eight storage volumes shared by all blades

2.1.2 Hypervisor Specification

Each blade has VMware ESXi 4.0 (build 173507) installed. VMware vCenter (Version 4.0.0 build 162856) is used to manage the blade servers. The eight blade servers are a part of a vSphere cluster. Among the available vSphere features Dynamic Resource Scheduling (DRS), Distributed Power Management (DPM) and VMware High Availability (HA) are enabled in the cluster.

2.1.3 Web-server Virtual Machine Specification and Workload Characteristics

The virtual machine is configured to simulate a typical ecommerce web server. Table 2 provides the configuration selected for this study.

Table 2: Virtual Machine Specification

Virtual Machine Specification	
vCPU	1 vCPU
RAM	1 GB
Disk	10 GB virtual disk for OS 10 GB virtual disk with PVSCSI driver for E-commerce workload data
Guest OS	Microsoft Windows Server 2003 32bit Enterprise Edition
Web Server	Microsoft IIS/PHP web server
NIC	1 virtual NIC with VMXNet3 driver
VMware tools	Yes

Table 3 shows the workload characteristics of the ecommerce workload of a single virtual machine. The workload characteristics are averaged over 128 VMs.

Table 3: Workload Characteristics

Workload characteristics of a single Web-server VM		
Number of simultaneous connections	300	
Average CPU utilization	4% of an 8 core M610 blade server	
Web Requests per second	70-90	
	Read	Write
IOPS	16.7	60.1
Average I/O packet size (KB)	15.4	6.7
Average IO (Bandwidth KB/s)	257	401.1

2.1.4 Client Specification

There are 128 virtual machines hosted on 16 PowerEdge M605 blade servers acting as clients. The following table describes the specification of the client environment.

Table 4: Client Specification

Client Host Configuration	
Number of Hosts	16 x PowerEdge M605
CPU	Two AMD Opteron 2350
RAM	32 GB

Disk	10 GB virtual disk for each client in a shared PS5500E

Client Virtual Machine Configuration	
Number of Clients	128 (16 per host)
vCPU	1 vCPU
RAM	1 GB
Disk	10 GB virtual disk for OS
Guest OS	Microsoft Windows Server 2003 32bit Enterprise Edition
NIC	1 virtual NIC
VMware tools	Yes

2.2 Configuration

This section describes the key configuration details of the environment.

Cisco Switch setup and configuration: As shown in Figure 1, the storage is connected to the servers via Cisco Catalyst switches. Each PS6000XV array has 4 active and 4 passive connections to the Cisco Catalyst 3750 switches. The Cisco Catalyst 3750 switches are stacked together and connected to four Cisco Catalyst 3130 I/O modules using 16 Gigabit links, as shown in figure 1. The internal Cisco Catalyst 3130 I/O modules are not stacked. For the virtual machine traffic, the test environment has 4 stacked Cisco Catalyst 3130 I/O modules. These switches are connected to clients via four 10 Gigabit Ethernet links combined into a single EtherChannel.

Virtual Network Configuration: The network configuration in the ESX server is shown below. Two virtual switches are used, as shown in figure 2. *vSwitch0* is created for virtual machine traffic, management and VMotion. The management and VMotion traffic are separated from virtual machine traffic using VLANs.

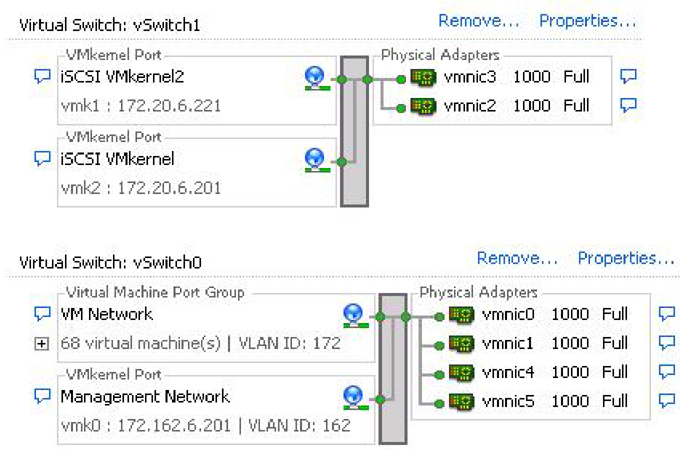

Figure 2: Network configuration of the test environment

A second virtual switch, *vSwitch1*, is created for iSCSI traffic. Two ESX kernel interfaces are created for iSCSI connection.

iSCSI Storage Configuration: The ESX servers are configured with built in software initiator. The virtual disks of all the VMs (both OS and data) are evenly distributed across 8 storage volumes. Each volumes is configured with round-robin load balancing.

VMware HA, DRS and DPM Configuration: The cluster has HA enabled, with host monitoring, admission control, and configured to tolerate one host failure. VMware DRS is enabled in "Fully Automated" mode, and is configured to apply priority 4 or higher recommendations. VMware DPM is in automatic mode and configured to apply priority 3 or higher recommendations. VMware DRS technology does automatic, policy based, load balancing at the VM level. The VMs are automatically distributed among hosts based on loads conditions and load balancing policy.

3 Methodology

The objective of this study is to measure the capacity of the virtual infrastructure farm in a typical configuration and resource utilization. The achieve this, we keep the average CPU utilization to around 60% and measure how many virtual machines of a real life workload is supported.

We create virtual machines to represent a real life workload; in this case we create web server virtual machines. Each web server VM supports 300 users or simultaneous connections. The web server VMs are scaled until the CPU utilization of the virtual infrastructure is around 60%. This is a typical level that data centers run. Measurements are made to answer the following:

- How does the workload VMs scale while keeping the CPU utilization to a typical level (~60%)?
- How does the quality of service experienced by users change?
- How does the VMware DPM improve the total energy consumption of the cluster?
- How long does it take to recover from host failures?
- How long does it take to put a loaded host in maintenance mode?

As the number of VMs are scaled the CPU utilization, active memory and QoS experienced by the user are measured. The cluster has a total of 64 cores; CPU utilization is the average utilization across all 64 cores. The active memory is the total memory used by all 8 servers. The QoS is measured as the percentage of users experiencing a good average response time, which is defined as 1.4 seconds or better.

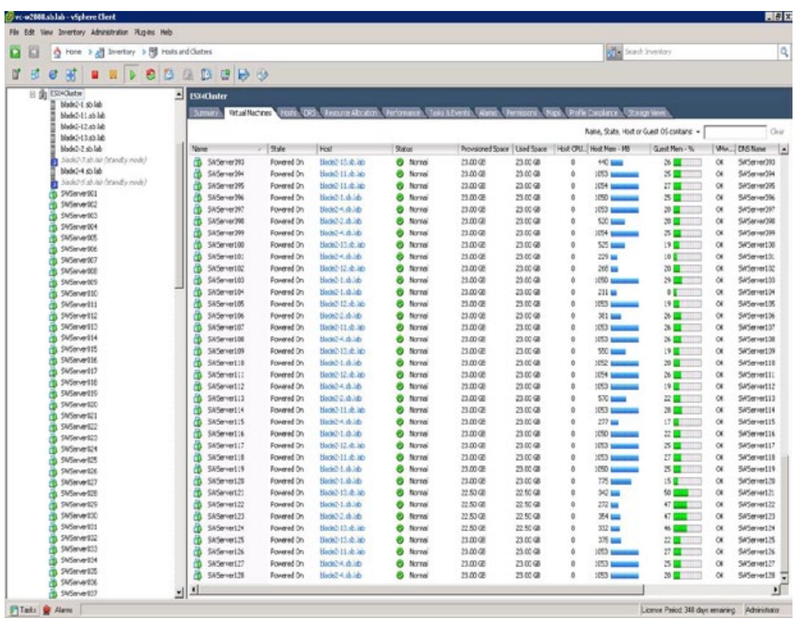

Figure 3: 128 Web-server Virtual Machines shown by vCenter

VMware DRS monitors resource and power utilization in the cluster regularly (every 5 min by default). DPM is used to optimize power utilization. During periods of low utilization, DPM can automatically consolidate VMs to fewer hosts. In the process, some hosts are automatically powered off to conserve energy. When the demand increases the hosts are automatically brought online (powered-on) and VMs distributed among them by DRS.

Initially, all 8 physical hosts in the cluster are in standby mode. DRS, DPM and HA services are enabled. Groups of sixteen web-server virtual machines are brought online at a time on the cluster. When the first VM is powered on, HA will automatically power on 2 hosts to meet the HA constraint. As the load scales, DPM automatically powers on additional hosts to meet the demand. In order to compare the energy savings of DPM, the experiment is repeated with DPM disabled. Figure 3 shows 128 VMs on eight hosts.

The iSCSI traffic in the PS6000XV arrays is measured using SANHQ. SANHQ is a performance monitoring application available at no additional cost with the EqualLogic PS arrays. (For more information see www.dell.com/equallogic). Figure 4, shows a sample screenshot of SAN HQ application monitoring the iSCSI storage. The VM network traffic is monitored with [StableNET express](#).

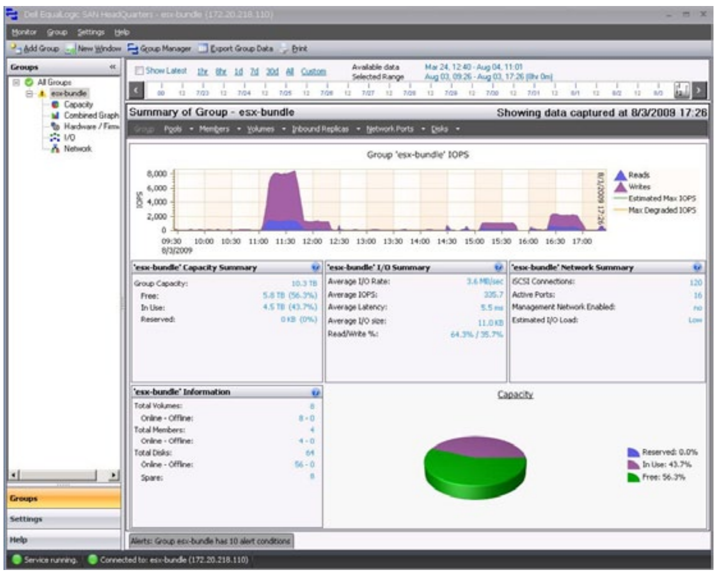

Figure 4: iSCSI storage overview shown in SANHQ

4 Observations

This section describes the observations made in the study.

4.1 Virtual infrastructure capacity for web server workload

As shown in Figure 5, when the number of web-server VMs increased from 16 to 128 and the numbers of users increased from 4800 to 38400, the total cluster CPU utilization and active memory increases proportional to the number of VMs. At 128 VMs the total CPU utilization is 68% leaving a 32% buffer to handle spikes in user request or for VMotion, VMware HA and other virtualization operations. Similarly the active memory is at 61.2% of the total capacity of 191GB.

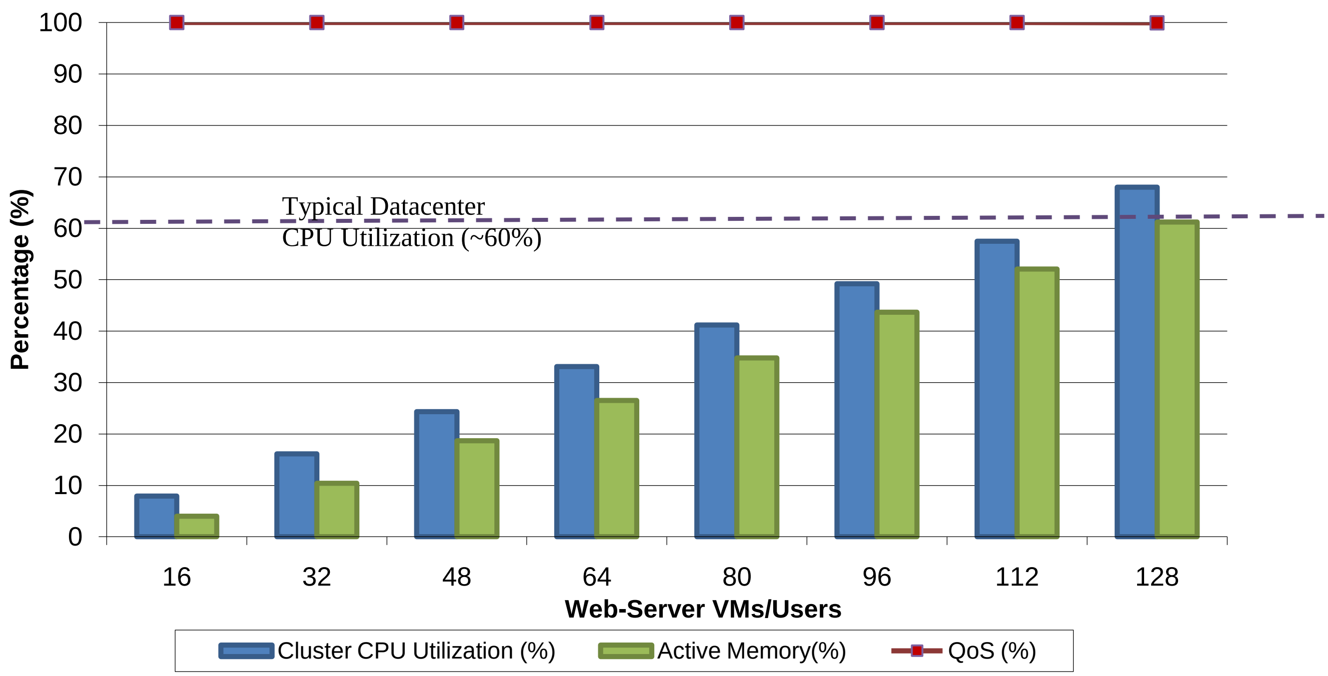

Figure 5: Capacity of the virtualized web server farm

QoS is measured as the percentage of users experiencing a good average response time, which is defined as 1.4 seconds or better. Note that the QoS is 100% and the average CPU utilization is 68% at 128 VMs. Hence large number of web server virtual machines (and number of users) may be supported by the cluster. The objective of this study is to measure the capacity of the virtualized web server infrastructure and typical data center resource utilization.

4.2 Energy Efficiency

In order to determine how much power is saved by using DPM, the tests are repeated with VMware DPM disabled. Initially, the 8 blade servers are powered on, but idle. The VMs are scaled from 16 to 128 as described previously, and power measurements are taken after every 16 VMs are brought online. As shown in figure 6, without VMware DPM the total power consumption of the cluster varies from 2888 to 3384 watts.

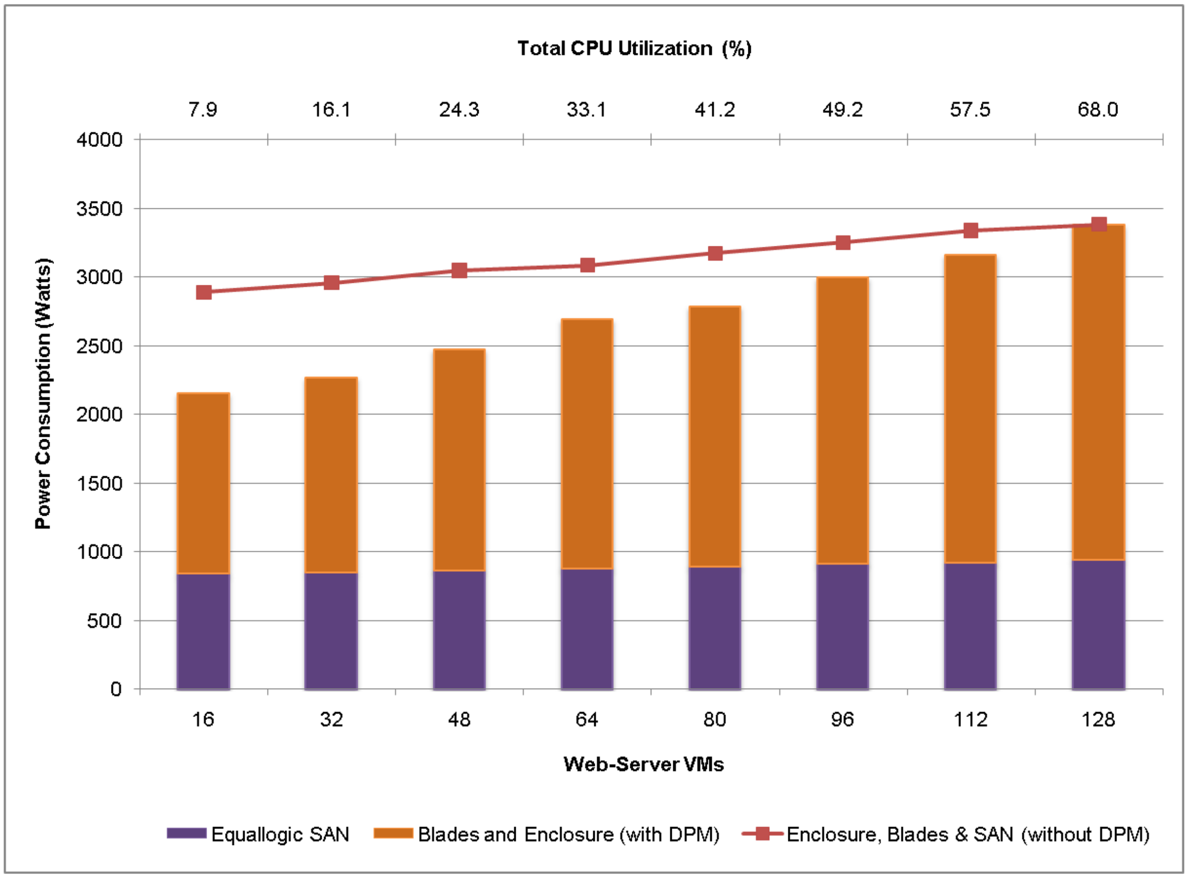

Figure 6: Reduced power consumption due to VMware DPM

The same experiment is repeated with VMware DPM enabled. First two blade servers are powered on along with 16 VMs. The other six hosts in the cluster are in standby mode. The total cluster power consumption is 2156 watts. As the demand is increased to 32 VMs (9600 users), the CPU utilization on the host server increases above 80%. This triggers the VMware DPM to power-on one more host. DRS redistributes the VMs with the newly added host, balancing the load among all the hosts. Similarly, as more VMs are brought online, DPM automatically powers-on hosts to satisfy the incremental demand, until all the hosts are turned on at 128 VMs (38400 users). The EqualLogic SAN power consumption is relatively constant; it varies from 840 to 940 watts as the number of VMs increase from 16 to 128.

To understand the power saving with DPM, let us assume a typical 24 hour cycle, where the CPU utilization varies from 10% during off-peak hours to 60% during working hours. Figure 7 shows the cluster CPU utilization and corresponding power saving for the entire 24 hour cycle. The percentages of power savings at different load conditions are shown. At low demands the total power savings is above 20%. The average power saving in our assumed 24 hour cycle is 13.1%.

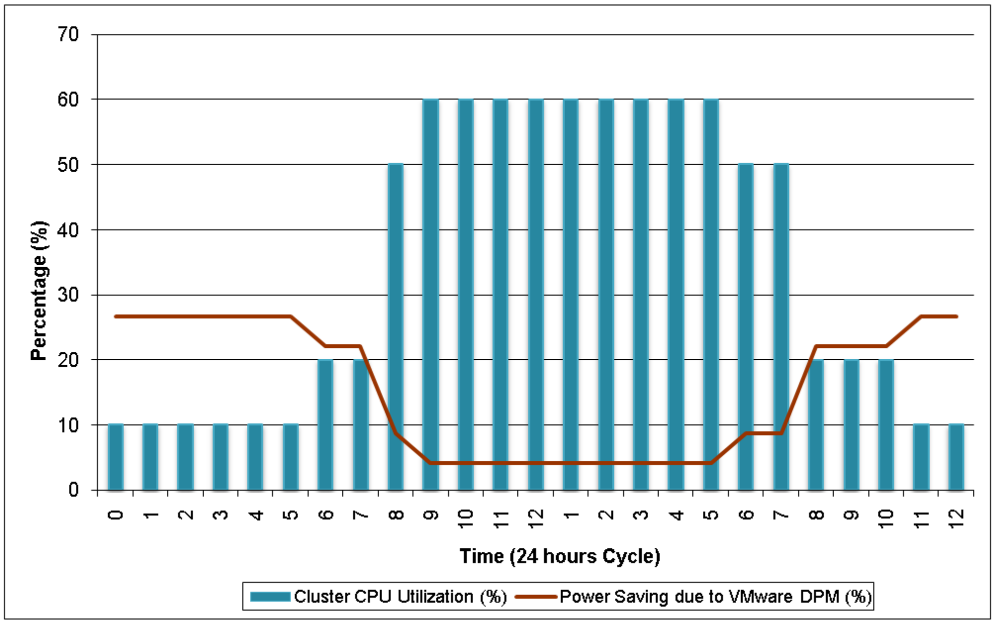

Figure 7: Power savings in a typical 24 hour cycle

4.3 High Availability and Serviceability

VMware HA provides protection for virtual machines against ESX host failures. In case of an ESX host failure, the virtual machines on the failed host are restarted on other available hosts in the cluster (figure 8).

The robustness of the virtual infrastructure is demonstrated by simulating single and dual server failure scenarios and measuring associated recovery time. The recovery time is the total time needed to restart the virtual machines on other hosts under loaded conditions (128 VMs supported 38400 users are running on the eight node cluster.) To simulate a server failure one of the 8 servers, with 16 virtual machines, is manually shutdown. It was observed that the 16 virtual machines are restarted in other hosts in **4 minutes**. Similarly, it was observed that a simultaneous two server failure recovered in **7 minutes**. There is was no significant loss in QoS after the failover.

Given the agility of a virtual infrastructure, it is easy to put a server in maintenance mode. Servers are put in maintenance mode in the eight node cluster under loaded conditions, which is at 128VMs and supporting 38400 users. The average time to place a server in maintenance mode is less than **7 minutes**. The servers are put in maintenance mode without significant loss in the QoS, because the cluster was designed to operate at around 60% CPU utilization at full load.

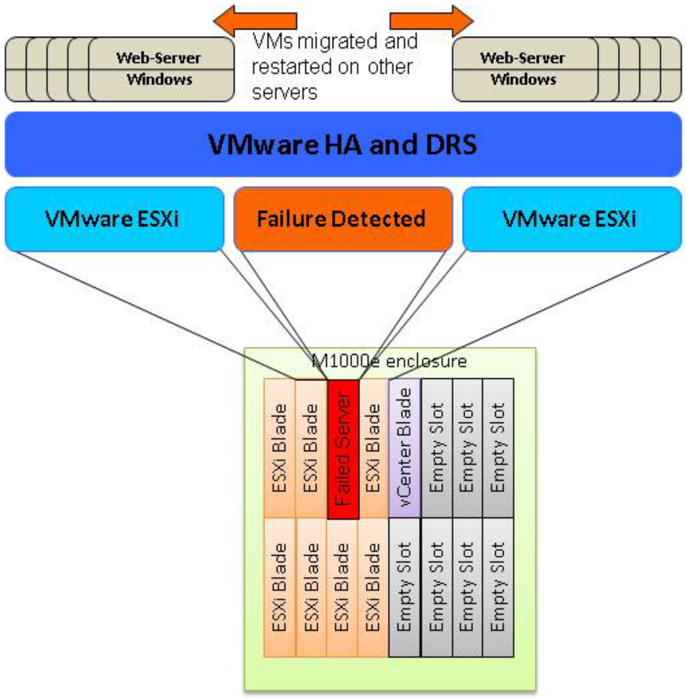

Figure 8: The ESX VMs are migrated from a failed blade server to other blades using VMware HA

4.4 Virtual Machine Network and Backend Storage Traffic

In the following section the observed network and backend storage traffic data is provided. The data can be used for capacity planning purposes.

Figure 9 shows the virtual machine network traffic. The data received (RX) shows the total network traffic due to the HTTP requests made by all the clients and data transmit (TX) is the total network traffic due to the content served by the web-servers. The large difference in the received and transmitted data is expected due to the nature of the e-commerce web-serving workload. Typically, the amount of content served is much larger compared to the size of the request. The average data received varies from 0.11 to 0.14 MB/sec per VM, and the average data transmitted varies from 3.15 to 3.80 MB/sec per VM. In a given VM, on average the ratio of transmit to received data is 28.9.

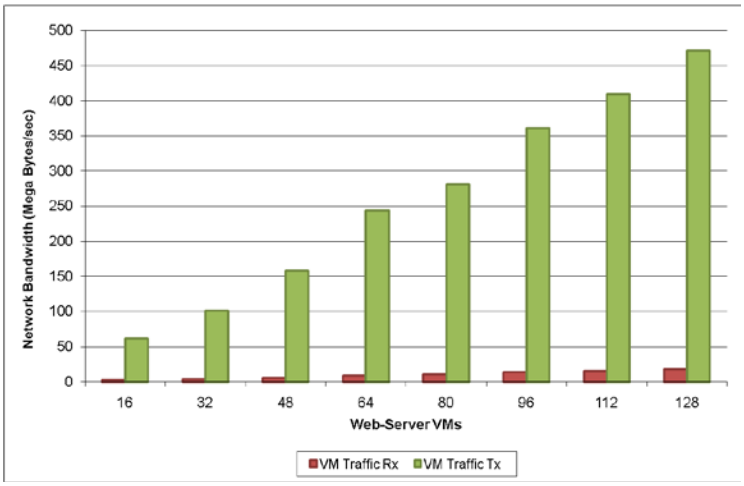

Figure 9: Virtual Machine Send and Receive network traffic

In Figure 10 on average the iSCSI SAN Tx is 25.7% and iSCSI SAN Rx traffic is 74.3%. The total traffic per VM varies from 0.51 to 0.53 MB/sec. Figure 11 shows the total IOPS registered; on average the Read IOPS are 16% and write IOPS are 84%. On average the total IOPS per VM is 69.4.

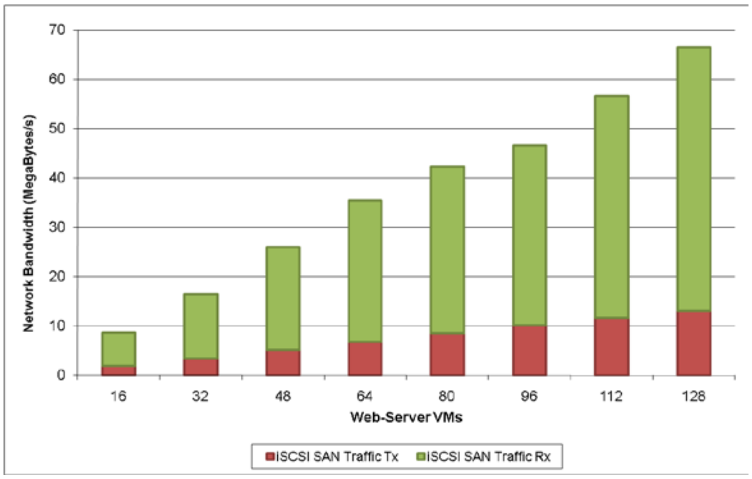

Figure 10: iSCSI SAN send and receive traffic

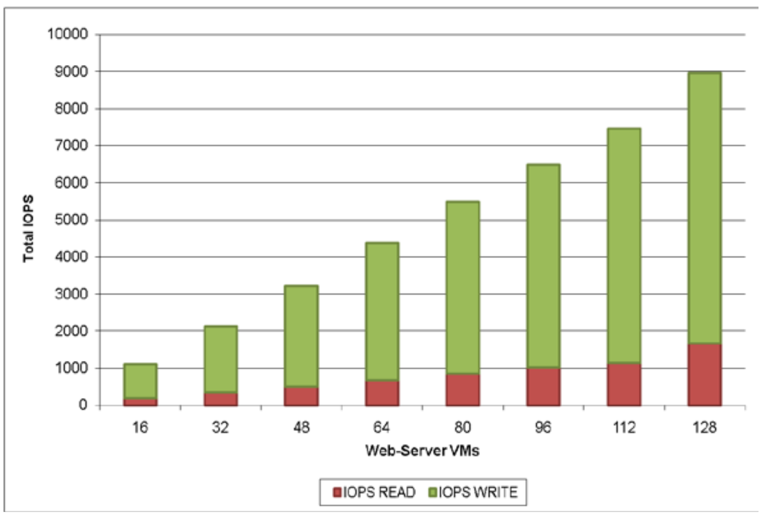

Figure 11: The total IOPS generated in iSCSI

5 Conclusion

The virtual infrastructure used in this study, consisting of Dell PowerEdge blade servers, Dell EqualLogic iSCSI storage arrays and VMware vSphere 4.0, can support 128 web-server VMs (each with 300 users) while the CPU and memory utilization are around the typical datacenter utilization levels. The network and storage subsystems have enough capacity to handle real-world workloads. The additional CPU and memory capacity can be used to handle spikes in usage or support virtualization operations like DRS, DPM, and HA enabling easy management and serviceability. Up to 30% power savings is observed due to VMware DRS and DPM during periods of low demands. VMware HA provides high-availability and ability to recover from failures quickly. Administrators and solution architects can use this information to size and scale their needs accordingly.

6 References

1. Business Ready Configurations for Virtualization – A Solution Guide for VMware vSphere 4 on Dell PowerEdge blade servers and Dell EqualLogic storage
2. Business Ready Configurations for Virtualization – Networking Best Practices for VMware vSphere 4 on Dell PowerEdge Blade servers
3. PS Series Array Network Performance Guidelines
4. vSphere Availability Guide: http://www.vmware.com/pdf/vsphere4/r40/vsp_40_availability.pdf