Demartek iSCSI Deployment Guide 2011

Overview

Because of an increased interest in iSCSI (Internet SCSI), Demartek has produced this ***Demartek iSCSI Deployment Guide 2011***, one in a series of technology deployment guides. This guide can be found on our website in our iSCSI Zone or by searching the Internet for "Demartek iSCSI Deployment Guide" using any well-known Internet search engine.

Audience

This guide is designed for managers and technical professionals within IT departments who are exploring the possible benefits of iSCSI technology or who are looking for actual deployment examples of iSCSI storage solutions.

Objectives of this Guide

This guide is designed to provide basic information about iSCSI and practical guidance for planning and deploying iSCSI technology and products. Although this guide contains some iSCSI marketplace information and discusses specific vendor products, it is primarily a technical document. It focuses on iSCSI primarily, but not exclusively, in a Microsoft Windows environment and is an updated version of the 150-page *Microsoft iSCSI Storage Deployment White Paper* that we published for Microsoft in June 2007, which is available on our website.

Because iSCSI includes server, networking, and storage technologies, this guide provides information and guidance in each area. A basic understanding of each of these areas is needed to successfully deploy iSCSI technology.

This guide is intended to be used as a reference and is divided into sections including iSCSI marketplace data, iSCSI technology areas, and specific vendor products in the area of network adapters and storage targets. There are screen shots and information from actual deployments of these products. Most of the work was performed in the Demartek lab in Arvada, Colorado, USA.

About Demartek

Demartek has its own lab and the vast majority of the research work we do involves running hardware and software solutions in our lab with our staff. The Demartek lab is equipped with servers, network infrastructure, and storage, and supports 1Gb Ethernet, 10Gb Ethernet, iSCSI, Fibre Channel, FCoE, SSDs, and a variety of other technologies.

Demartek produces highlight videos of public evaluations and deployment guides. Links to these videos are available on our web site and are posted on the Demartek channel on YouTube.

Table of Contents

Basic Storage Architectures

Storage for computers can be architected in several ways, each satisfying specific needs. Here we provide an overview of these architectures and show where iSCSI technology fits.

Direct Attached Storage (DAS)

Direct Attached Storage (DAS) is probably the most well-known form of computer storage. In a DAS implementation, the host computer has a private connection to the storage and almost always has exclusive ownership of the storage. The host computer accesses the storage in a "block" fashion, which means that it directly addresses blocks on the storage device. This implementation is relatively simple and usually inexpensive. Potential disadvantages are that the distance between the host computer and the storage are frequently short, such as inside a computer chassis or within a rack or adjacent rack. Some DAS implementations require that the host computer be taken offline when adding or removing storage devices, such as a boot drive directly connected to a motherboard storage interface. SATA is a common DAS interface.

Network Attached Storage (NAS)

Network Attached Storage (NAS) devices, also known as file servers, share their storage resources with clients on the network in the form of "file shares" or "mount points." The clients use network file access protocols such as CIFS/SMB or NFS to request files from the file server. The file server then uses block protocols to access its internal storage to satisfy the requests. Because NAS operates on a network, the storage can be very far away from the clients. Many NAS solutions provide advanced features such as snapshot technologies, global namespace, SSD caching and more.

Storage Area Network (SAN)

SAN architecture provides a way to use block access methods over a network such as Ethernet or Fibre Channel to provide storage for host computers. The storage in a SAN is not owned by one server but is accessible by all of the servers on the network. This SAN storage can be carved into logical storage pools or volumes that can be assigned to particular host servers. These logical volumes are independent of the geometries or components of the storage hardware. The storage appears to host servers and applications in the same way that DAS storage appears, but because SAN storage uses a network, storage can be a long distance away from the host servers.

SAN architectures use block Small Computer System Interface (SCSI) protocol for sending and receiving storage data over their respective networks. Fibre Channel (FC) SANs implement the SCSI protocol within the FC frames. Internet SCSI (iSCSI) SANs implement the same SCSI protocol within TCP/IP packets. Fibre Channel over Ethernet (FCoE) is a newer interface that encapsulates the Fibre Channel protocol within Ethernet packets using a relatively new technology called Data Center Bridging (DCB), which is a set of enhancements to traditional Ethernet and is currently implemented with some 10GbE infrastructure. Because each of these technologies allow applications to access storage using the same SCSI command protocol, it is possible to use all of these technologies in the same enterprise, or to move from one to the other. Generally speaking, applications running on a host server cannot tell the difference between Fibre Channel SAN storage, FCoE SAN storage, and iSCSI SAN storage. In fact, applications generally cannot tell the difference between DAS storage and SAN storage.

There has been much debate over Fibre Channel vs. iSCSI. Some people focus on the lower entry price points available for iSCSI SANs, while others focus on high reliability and availability of Fibre Channel SANs. The conventional wisdom in this debate no longer holds true in many cases. While it is true that at the low end iSCSI tends to be less expensive than Fibre Channel, as performance, reliability and high availability features and designs are included in iSCSI solutions, the iSCSI price advantage diminishes. Over the past few months we tested various solutions for ease of use and found that some implementations of Fibre Channel are just as easy, or in fact easier to use, than some iSCSI implementations. We have also found several iSCSI solutions that are designed for performance, high reliability, and high availability. Much work has been done to increase performance, reliability, and ease-of-use with both of these SAN technologies. The real issues are satisfying the needs of the business or application, working with existing infrastructure and expertise, maintaining service-level agreements, and staying within budgets. Fibre Channel and iSCSI technology can meet these requirements, and there is room for both in current and future IT environments.

For the larger enterprises that have implemented SAN technology, most have implemented Fibre Channel technology. These enterprises typically demand proven technology, have the need for high bandwidth storage solutions, have the budgets to pay for more expensive hardware to meet their performance and reliability needs, and typically have full-time staff dedicated to storage management. Some of these enterprises continue to invest in Fibre Channel storage solutions and plan to do so for the foreseeable future. However, some of these enterprises are also investing in iSCSI storage solutions, especially with 10GbE technology, for their virtualized server environments.

Smaller enterprises are often attracted to iSCSI technology because of its lower entry price point, and grow their iSCSI SAN as their needs change.

There is more to choosing a storage system than selecting the host interface. Regardless of the type of interface, several other factors need to be considered, including the number and type of disk drives, amount and type of SSD technology, management software, advanced features, support from the vendor, and several other factors. Advanced features of modern storage systems may include various forms of replication, thin provisioning, compression, data de-duplication, caching, automated storage tiering, and more.

Unified Storage

Unified storage combines NAS and SAN technologies into a single, integrated solution. These unified storage solutions provide both block and file access to the shared storage environment. These often provide simplified management by combining the management of all storage, regardless of the transport or “plumbing” into a single management console.

iSCSI Marketplace

Demartek focuses primarily on real world, hands-on experience with various technologies, but we decided that providing some iSCSI marketplace data would help provide additional perspective. The following marketplace data is provided by IDC and is reprinted with their permission.

iSCSI

The desire for lower-cost, less complicated networking infrastructures continues to compel solution providers and users to consider iSCSI. Adoption of iSCSI scales from the SMB market to large datacenters. In addition, server virtualization becomes one of the major driving forces for iSCSI adoption.

iSCSI adoption should increase the TAM in the SMB market segment and provide additional opportunity in large corporations. However, increases in TAM may be offset by lower cost per gigabyte associated with the systems. The impact of 10GbE on iSCSI demand is still uncertain, but may lead to growing competition with FCoE in converged IT infrastructure environments in the late years of the forecast.

-Worldwide Enterprise Storage Systems 2010-2014 Forecast Update: December 2010 (IDC#226223 December 2010)

10Gbps Ethernet

The evolution of 10Gbps technology will expedite the adoption of iSCSI SAN in larger organizations and play an increasingly important role in NAS and clustered storage environments.

Long term, the availability of 10Gbps Ethernet will spur interest in FCoE and iSCSI SAN, with FCoE being the preferred alternative for many large enterprises with FC SANs. IDC expects a modest shift of some revenue from FC to the iSCSI SAN segment in out years, with FCoE becoming more important in storage in 2011.

-Worldwide Enterprise Storage Systems 2010-2014 Forecast Update: December 2010 (IDC#226223 December 2010)

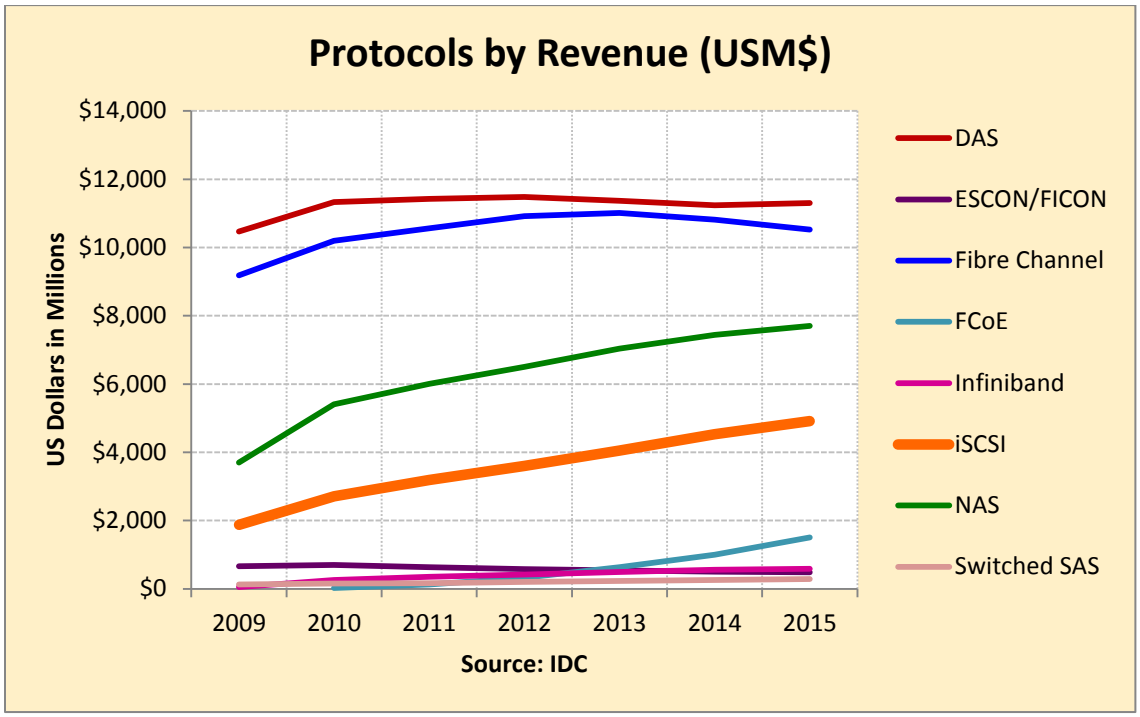
While the overall market revenue saw a substantial decline, the iSCSI SAN segment posted a healthy factory revenue growth rate, 24.4% YoY in 2009. While this rate of growth was slower than in previous years for this segment, it was still a good indicator of end user interest in the Ethernet-based alternative to, generally, more expensive FC SAN.

- Worldwide Disk Storage Systems 2009 Vendor Share: Year in Review (IDC#225847 December 2010)

IDC forecasts that of the established segments, iSCSI SAN will grow to be a $4.9 billion dollar market, at a 9.04% compound annual growth rate (CAGR) for the 2011-2015 period. iSCSI SAN is also forecasted to grow at a higher rate than the established than Fibre Channel (0.65% CAGR) and NAS (7.34% CAGR).

-IDC Forecast Pivot

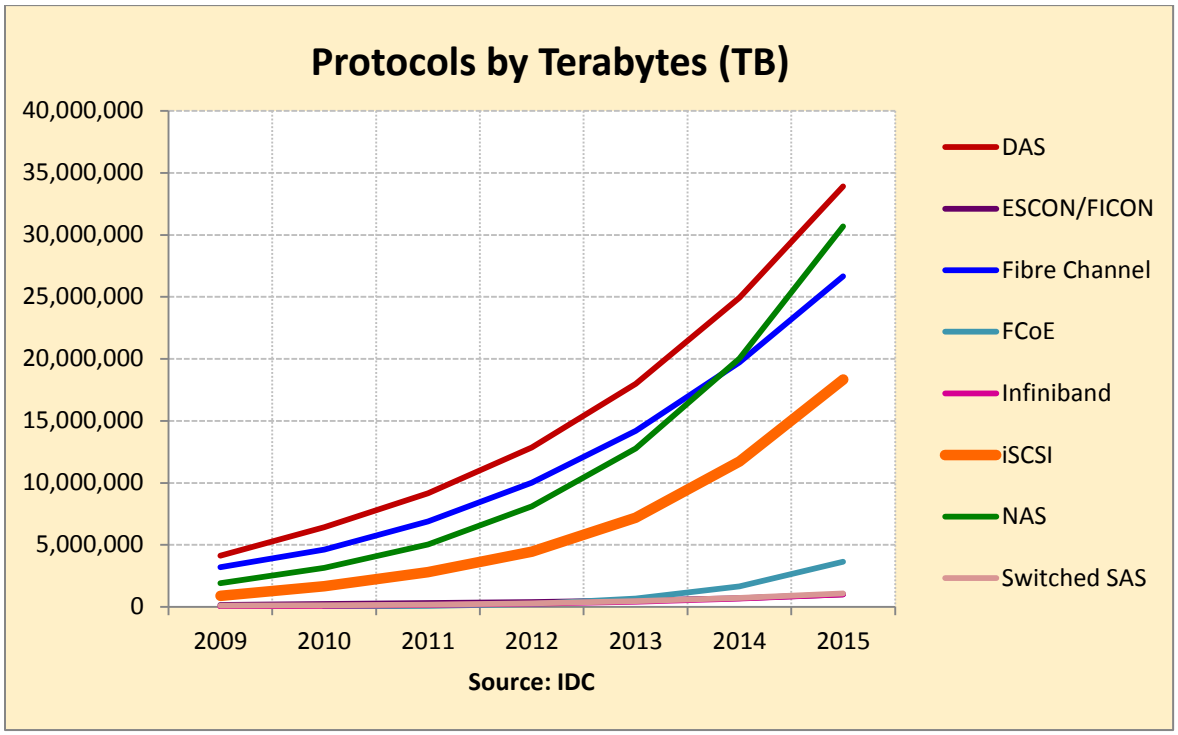
Protocols by Annual Sum of Customer Revenue (USM$)



	Annual							
Protocol	2009	2010	2011	2012	2013	2014	2015	2010-15 CAGR
DAS	$10,473	$11,333	$11,425	$11,483	$11,373	$11,237	$11,299	-0.1
ESCON/FICON	$664	$700	$631	$575	$545	$504	$481	-7.2
Fibre Channel	$9,181	$10,193	$10,558	$10,923	$11,008	$10,816	$10,529	0.7
Fibre Channel over Ethernet (FCOE)		$23	$129	$322	$635	$1,000	$1,507	130.2
Infiniband	$53	$260	$352	$417	$491	$561	$590	17.8
iSCSI	$1,874	$2,709	$3,185	$3,600	$4,052	$4,526	$4,910	12.6
NAS	$3,702	$5,406	$6,006	$6,501	$7,034	$7,439	$7,704	7.3
Switched SAS (SAS SAN)	$127	$153	$168	$200	$230	$261	$289	13.6

Source: IDC, reprinted with permission

Protocols by Annual Sum of Terabytes (TB)



	Annual							
Protocol	2009	2010	2011	2012	2013	2014	2015	2010-15 CAGR
DAS	4,122,487	6,430,702	9,151,908	12,845,016	17,991,661	24,923,806	33,927,686	39.5
ESCON/FICON	162,947	234,462	308,371	386,632	517,702	700,648	995,693	33.5
Fibre Channel	3,207,733	4,612,718	6,882,335	10,013,628	14,184,303	19,707,653	26,659,284	42.0
Fibre Channel over Ethernet (FCOE)		17,521	71,971	227,173	682,544	1,648,667	3,629,966	190.6
Infiniband	27,312	79,891	153,748	255,605	421,900	676,586	1,005,964	66.0
iSCSI	900,432	1,657,018	2,801,610	4,438,713	7,186,983	11,725,440	18,337,428	61.7
NAS	1,906,285	3,141,448	5,044,362	8,109,973	12,784,235	20,021,300	30,701,486	57.8
Switched SAS (SAS SAN)	93,253	129,788	174,808	284,321	451,896	717,717	1,087,626	53.0

Source: IDC, reprinted with permission

iSCSI Technology and Definitions

Internet SCSI (iSCSI) is an industry standard developed to enable transmission of SCSI block storage commands and data over an existing IP network by using the TCP/IP protocol. The encapsulated SCSI commands and data can be transmitted over a local area network (LAN) or a wide area network (WAN). As with traditional SCSI, an iSCSI storage solution requires at least one "initiator" residing on the application server and at least one "target" residing on the storage.

Initiators

Initiators are host servers that need access to storage resources. This is in keeping with the standard SCSI protocol where initiators issue commands to storage devices and wait for replies. iSCSI initiators are included with or are available for every major desktop and server operating system. iSCSI initiators can be implemented in software such as in the operating system or in hardware such as in a network adapter. iSCSI initiators can access multiple targets simultaneously.

Targets

Targets are devices that provide storage resources for iSCSI initiators. These follow the SCSI protocol for targets by responding to requests from initiators. Targets can be implemented as dedicated hardware devices such as disk arrays or as combination hardware and software in a server or appliance solution. Targets can respond to multiple iSCSI initiators simultaneously.

iSCSI Qualified Name (IQN)

Every iSCSI initiator and target has a unique iSCSI qualified name (IQN). The IQN is the standard way of identifying iSCSI initiators and targets. The official standard for IQN naming is listed in RFC 3720. Some iSCSI initiator and target configuration applications do not require rigid adherence to standard IQN naming rules. Examples of standard IQ naming are listed below.

- iqn.1991-05.com.microsoft:dmrtk-srvr-b.lab.demartek.com
- iqn.1991-05.com.microsoft:dmrtk-srvr-g
- iqn.1992-08.com.netapp:sn.118050892

Some iSCSI initiator target configuration applications can use an IP address instead of an IQN. However, best practice is to use the IQN so that if the IP address of the initiator or target changes, the iSCSI session logon sequences do not have to be changed.

iSCSI Session

Sessions are established between iSCSI initiators and iSCSI targets when the iSCSI initiator performs a "logon" or "connect" with the target. Once established, the IQN and IP address information of the initiator and target for each session are maintained.

Multi-Path I/O (MPIO)

Many targets support multipath I/O (MPIO) in an iSCSI session that uses multiple IP addresses on the initiator, target, or both. MPIO can be used for failover or load-balancing. Microsoft provides a general multipath driver, known as MPIO, which works for iSCSI, Fibre Channel and SAS interfaces. MPIO is a feature that can be installed from the "Add Features" wizard in the Server Manager for Windows Server 2008 R2. Some targets require the installation of special drivers or device specific modules (DSM) on the initiator host to support MPIO.

Multiple Connections

Some targets support multiple connections using the same initiator and target IP address within the same session, which may improve performance. Generally, MPIO and multiple connections are not supported at the same time for the same target.

Challenge Handshake Authentication Protocol (CHAP)

CHAP is a protocol that iSCSI initiators and targets can use to authenticate with each other using a shared secret (similar to a password). CHAP can be either one-way or mutual. With one-way CHAP, one side uses the shared secret to compute a one-way hash that is transmitted to the other side during the login process. The other side computes a hash of its shared secret and if the hash matches, it accepts the authentication from the sender. With mutual CHAP, both sides perform this operation to authenticate each other. This provides additional authentication security above and beyond what may be present using TCP/IP.

Although not required for iSCSI operation, the best practice is to use one-way CHAP at a minimum and mutual CHAP for higher levels of security.

IPsec

IPsec authenticates and encrypts the packets at the IP packet layer sent between the iSCSI initiator and iSCSI target. Beginning with Microsoft Windows Vista and Windows Server 2008, IPsec configuration works with the Windows firewall and group policy.

Networking Technologies

Because iSCSI operates over TCP/IP and Ethernet, it is important to understand several of these networking concepts. Some of the technologies mentioned below apply to Ethernet networking in general and specifically for iSCSI.

Line Speed

In theory, iSCSI can use any speed of Ethernet, however the best practice is to use gigabit Ethernet or faster. We have tested iSCSI using Wi-Fi in the Demartek lab, and it does work. However we do not recommend using iSCSI over Wi-Fi in enterprise environments for security, performance, and latency reasons.

10-Gb Ethernet (10GbE) prices are slowly dropping, making it a viable choice, especially in virtualized server environments. The big server vendors are beginning to offer 10GbE as an option on the newer server motherboards, which we believe will increase the adoption of 10GbE. Some even plan to offer a choice of copper or fiber-optic connectors.

Cabling

Because iSCSI runs over Ethernet, the iSCSI cabling requirements are the same as they are for Ethernet. For gigabit Ethernet, Cat5E or Cat6 cabling can be used. For 10GbE, Cat6a and Cat7 cabling can be used and Cat6 cabling can be used up to 55 meters. Fiber-optic cabling can also be used in gigabit and 10GbE environments.

Additional copper and fiber-optic cabling and connector details for iSCSI and other storage interfaces are available on the Demartek Storage Interface Comparison page on the Demartek website. This reference page can also be found by searching the Internet for "Demartek Storage Interface Comparison" using any well-known Internet search engine.

Network Adapters and Features

The best practice for iSCSI is to use a good server-class network adapter in order to achieve reliable performance. In the past, we tested iSCSI storage workloads using cheap desktop-class network adapters that did not support many of the features listed below, and found that the performance was poor. In some cases under heavy load, the host server froze and the only way to recover was to power off the server and re-start it. The same iSCSI storage workloads ran well when using server-class network adapters.

Good server class network adapters implement many of the following features and provide various degrees of hardware assistance. It is recommended that these types of network adapters be used for iSCSI traffic. We tested several 1Gb and 10Gb network adapters from several vendors that work well with iSCSI solutions, each having a separate section in this guide:

- Broadcom
- Chelsio
- Emulex
- Intel
- QLogic

Jumbo Frames

Jumbo frames (jumbo packets) can be used to allow more data to be transferred with each Ethernet packet, reducing server overhead for such things as protocol, checksums, etc. The standard maximum transmission unit (MTU) payload size in an Ethernet packet is 1500 bytes. The most common size for jumbo frames is 9000 bytes. Depending on the manufacturer, some network devices support jumbo frame sizes in the range of 4000 bytes to 14,000 bytes. Jumbo frames are typically supported with 1GbE and 10GbE switches, network adapters, and iSCSI targets, but are not generally supported with fast Ethernet (100Mb/s) devices.

When jumbo frames are used, the same packet size setting must be set in all the adapters, switches, and iSCSI targets in the data path in order for them to be successfully transmitted and received.

Receive-Side Scaling (RSS)

Early TCP/IP implementations did not handle multicore systems properly but routed all incoming packets to one core of the system. Receive-Side Scaling, sometimes known as "Scalable I/O," balances incoming packet handling across the CPU cores and requires some intelligence in the network adapter to be supported. RSS support is a recommended best practice for iSCSI.

Also see the "Scalable Networking Pack" notes below.

Partial Offload Functions

Most good server-class network adapters support a number of partial offload functions that include hardware assistance for several stateless functions needed for TCP/IP processing. These include:

- TCP and UDP checksum offload
- Large Send Offload (LSO)
- Large Receive Offload

These partial offload functions may support IPv4, IPv6, or both.

TCP Chimney Offload

TCP Chimney offload, also known as TCP/IP Offload Engine (TOE), transfers TCP/IP protocol processing from the host CPU to a network adapter that supports full TCP/IP offload processing, sometimes known as a TOE card. Both TCP/IPv4 connections and TCP/IPv6 connections can be offloaded if the network adapter supports this feature.

TCP Chimney offload may not be able to coexist with other applications or services that rely on lower-level networking services. Services that may not be able to take advantage of TCP Chimney offload include IPsec, network load balancing, and others.

In Windows Server 2008 R2 the default behaviors for adapters that support TCP/IP offload are:

- On 10 GbE network adapters, TCP connections are offloaded by default
- On 1 GbE network adapters, TCP connections are not offloaded by default

To offload TCP connections on a 1 GbE network adapter, you must explicitly enable TCP offloading, which can be performed by using the "netsh" command.

Also see the "Scalable Networking Pack" notes below.

Virtual Machine Queue

Virtual machine queue (VMQ) is a feature found in some network adapters that works with servers running Windows Server 2008 R2 with Hyper-V. VMQ offloads some of the packet handling and sorting required in a Hyper-V virtual machine environment, enabling data packets to be delivered directly to the guest virtual machine with minimal handling by the hypervisor, using Direct Memory Access (DMA) techniques.

For guest virtual machines running Windows 7 or Windows Server 2008 R2, this feature is automatically supported. For guest virtual machines running Windows Vista or Windows Server 2008, the integration services setup disk must be installed. Guest virtual machines running earlier versions of Windows do not support VMQ.

iSCSI Offload

Some server-class network adapters support iSCSI offload, which is a combination of the stateless offload functions, TCP Chimney offload, and the iSCSI stack offload. These network adapters, in effect, become host bus adapters (HBAs), and appear to Windows as storage controllers.

Because these adapters offload a great deal of the network processing, they can reduce overall host CPU utilization while providing excellent performance.

There were relatively few of these for iSCSI offload adapters for 1Gb Ethernet, but with the growth of 10Gb Ethernet, server virtualization, and multicore servers, several vendors are now providing 10Gb iSCSI offload adapters. We tested some of these adapters for this report.

NetDMA

NetDMA offloads the network subsystem memory copy operation to a dedicated DMA engine. NetDMA must be enabled in the BIOS and cannot be used with TCP Chimney offload.

Also see the "Scalable Networking Pack" notes below.

Scalable Networking Pack

In 2006, Microsoft released the "Scalable Networking Pack" for Windows Server 2003 and Windows XP 64-bit. This free download included Receive-Side Scaling, TCP Chimney offload and NetDMA support. These features became a standard part of the operating system beginning with Windows Server 2008.

iSCSI Initiators

The primary interface from a host server into the iSCSI infrastructure is the iSCSI initiator. The iSCSI initiators can be implemented in software or hardware. Software initiators can be provided in the operating system or through third-party applications. Hardware initiators are usually provided with specialized network adapters. The hardware initiators usually provide their own TCP/IP stack and iSCSI stack.

Microsoft iSCSI Initiator

Beginning with Windows Vista and Windows Server 2008, the iSCSI initiator was included as part of the Windows operating system. For older versions of Windows, including Windows 2000, Windows XP, and Windows Server 2003, the Windows iSCSI initiator was available as a free download.

The iSCSI initiator on Windows operating systems is available with a command-line interface and a graphical user interface (GUI). Other operating systems such as Linux and UNIX variants include a command-line interface to the iSCSI initiator. Graphical user interfaces for the iSCSI initiator in these other environments are also available.

The iSCSI initiator interface is used to discover iSCSI targets, establish sessions with iSCSI targets and make other configuration adjustments for the iSCSI session. Demartek has produced a short video highlighting the Microsoft iSCSI initiator graphical user interface from the perspective of a 10-year-old girl. This video is available on the Demartek channel on YouTube.

With Windows 7 and Windows Server 2008 R2, the Microsoft iSCSI initiator has several tabs, which are:

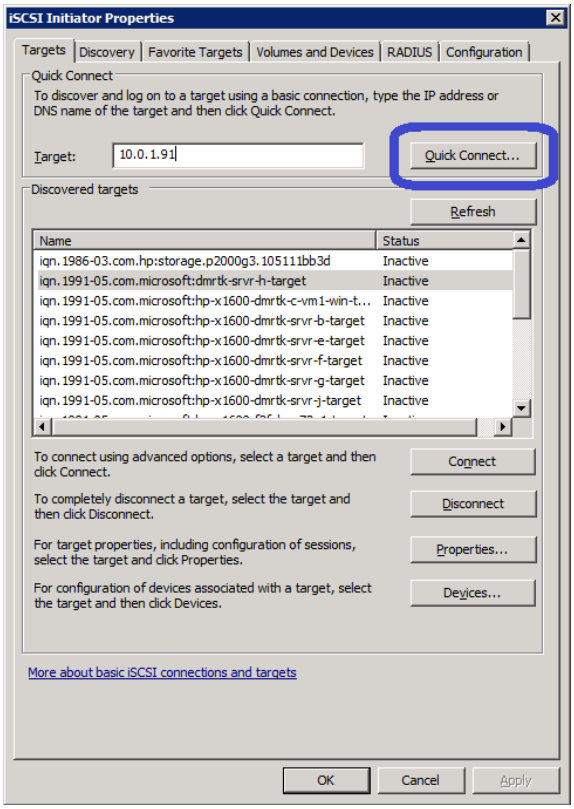
- **Targets** – has the Quick Connect option and provides for regular connection and disconnection options. The Targets tab also allows the properties for the connection to be queried and provides details about the iSCSI devices.
- **Discovery** – provides the list of iSCSI targets for discovery purposes and allows for registration with an iSNS server.
- **Favorite Targets** – provides the list of favorite targets that will automatically be connected when the computer restarts.
- **Volumes and Devices** – provides the list of volumes and devices associated with SCSI targets and gives the option to have the volume automatically mounted when the computer restarts.
- **RADIUS** – provides a list of RADIUS servers that should be used to perform authentication of iSCSI connections.
- **Configuration** – provides the full IQN for this initiator and is the location for the CHAP and IPsec settings.

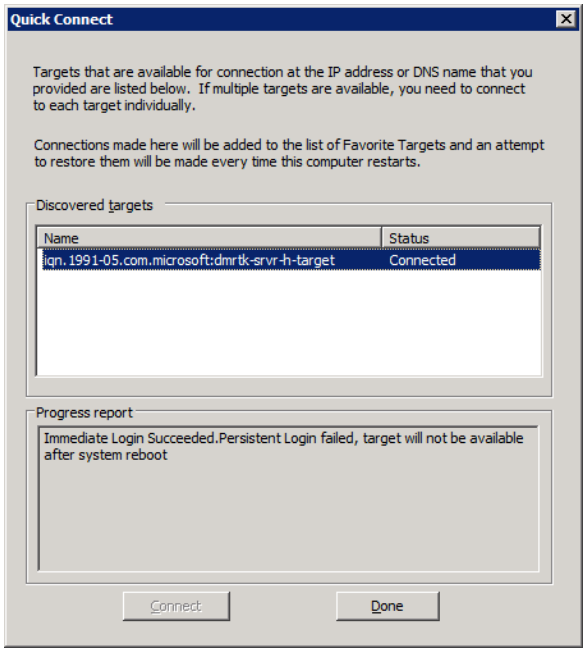
The Microsoft iSCSI initiator supports IPv4 and IPv6.

Quick Connect

The Microsoft iSCSI software initiator provides an option known as "Quick Connect" that performs the iSCSI logon to an iSCSI target with a minimum number of keystrokes and mouse clicks. This is intended for connections to storage devices that do not require advanced settings.

The user enters an IP address on the first screen, clicks "Quick Connect," and the second screen shows the results of the immediate login. The LUNs or disk volumes are immediately made available to the host server and are ready to be formatted, etc.





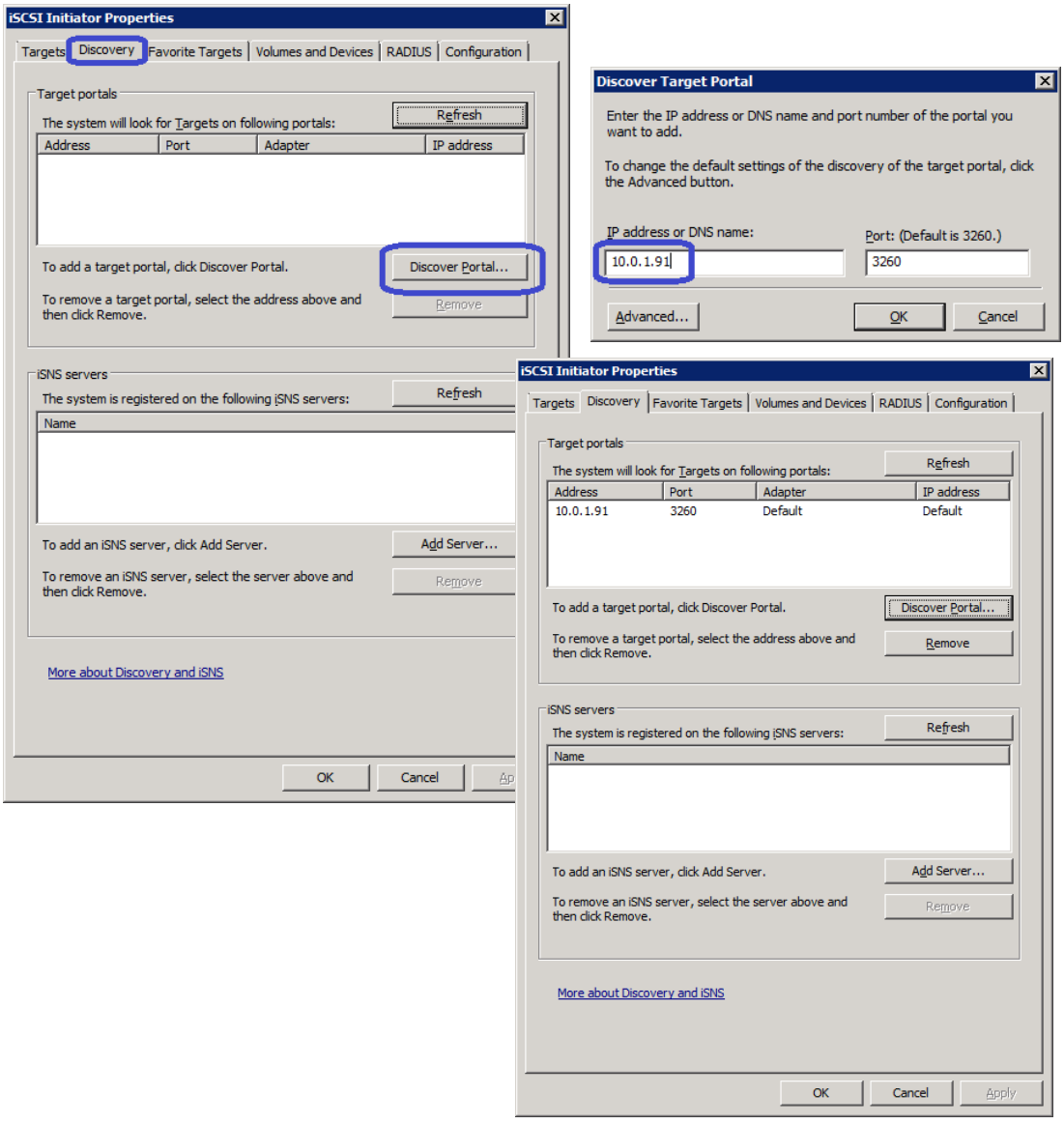
Standard iSCSI Login Process

The standard login process uses several of the tabs shown on the Microsoft iSCSI initiator interface. These can be used to provide additional options for login and subsequent connections.

A typical logon process uses the discovery and targets tabs to establish a connection with an iSCSI target.

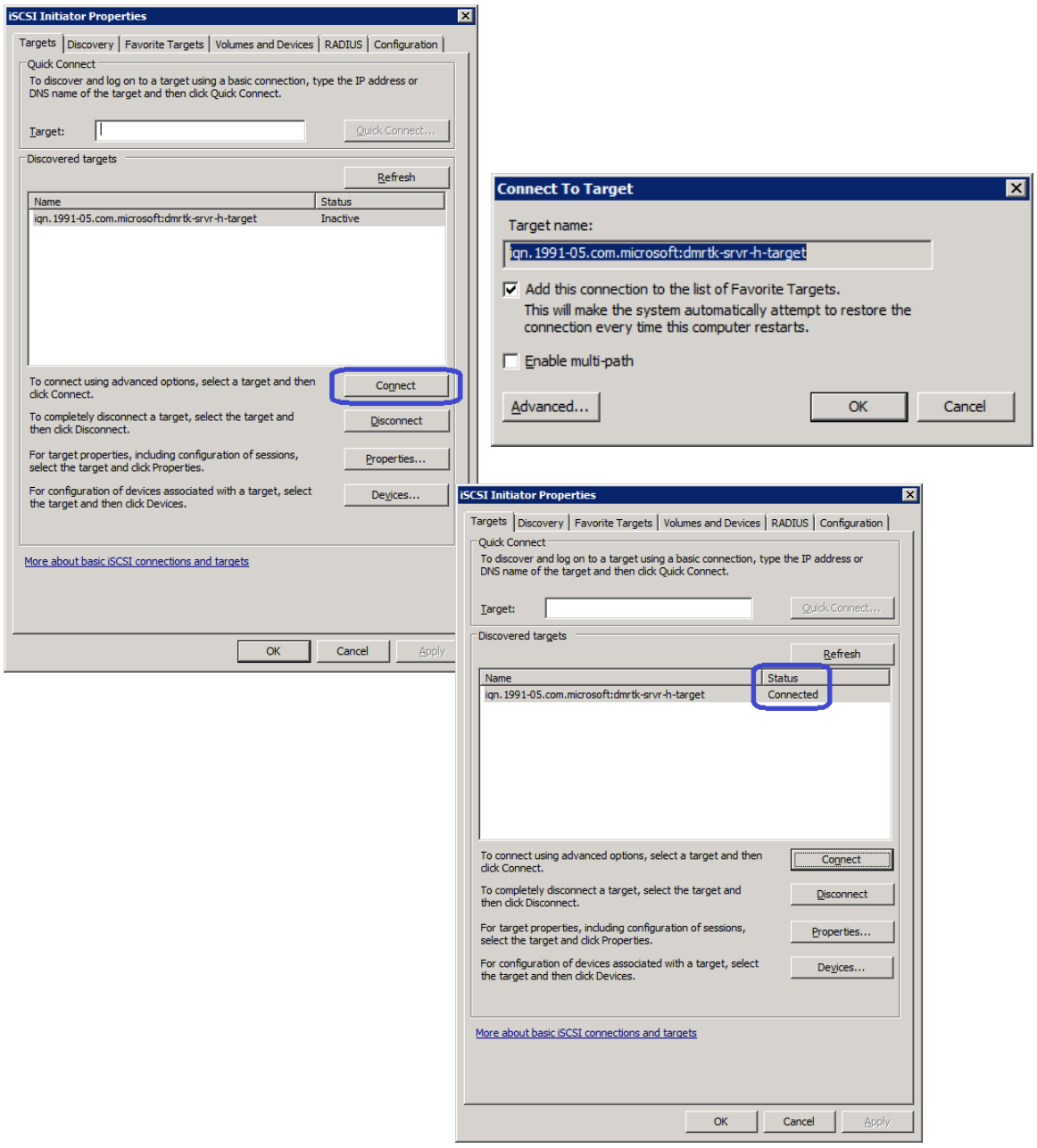
1. Enter the IP address of the iSCSI target in the "Discover Portal" section of the "Discovery" tab
2. Connect to the iSCSI target on the "Targets" tab

Step 1 – Discovery



Step 2 – Connect

These steps show an iSCSI login using a single path between the initiator and target. Multiple paths may also be specified by checking the "Enable multi-path" box and entering specific initiator and target IP addresses to be used for each path. The "Advanced" button allows for setting advanced options such as CHAP, IPsec, RADIUS authentication settings, etc.

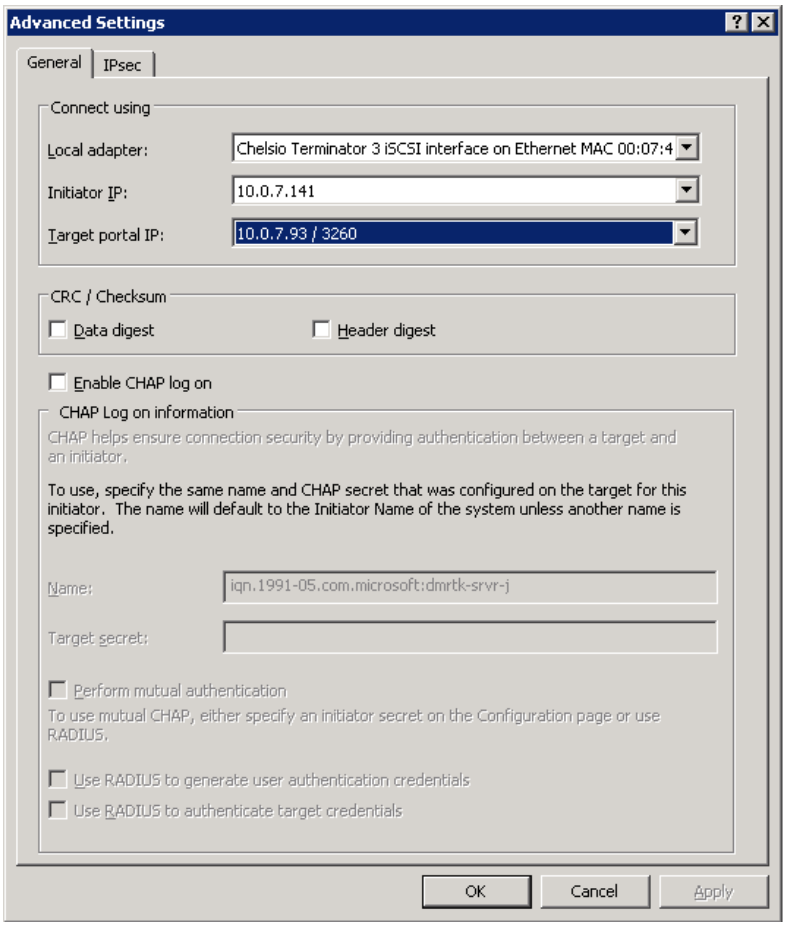


Hardware Initiators

Hardware initiators may provide their own application interface directly to the adapter hardware for performing login and other iSCSI management functions, or may use the Microsoft iSCSI software initiator interface as a front end to their hardware.

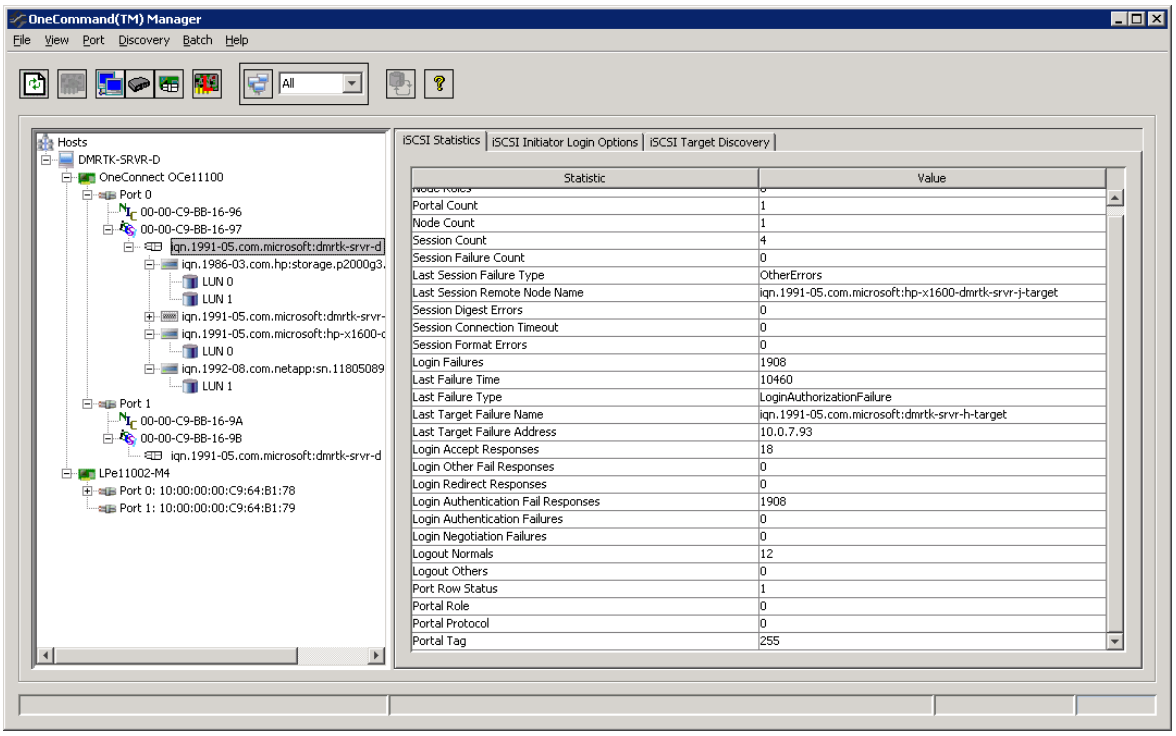
Hardware Initiators that use the Microsoft Initiator Interface

In this example, notice that the Chelsio adapter is listed as the local adapter in the Microsoft iSCSI software initiator interface. Here, the Microsoft iSCSI software initiator is merely acting as a front-end interface for the Chelsio management software and adapter. Full details regarding the Chelsio adapter are located in the Chelsio adapter section below.



Hardware Initiators that use their own Native Interface

In this example, the Emulex iSCSI offload adapter is managed by its own software, known as OneCommand Manager. This software manages the iSCSI initiator including all the login and target discovery options. The software provides additional features and is explained in the Emulex adapter section below.



iSCSI Targets

iSCSI targets are available from many server and storage vendors, and span the full range from very low to very high-end storage systems. Some of these storage solutions are based on the Microsoft iSCSI target software and run on a Microsoft Windows server platform. Some solutions run on an appliance with its own operating system and storage behind it. Some solutions are in the form of traditional disk arrays. Some of these disk arrays are purpose-built for iSCSI only, and others have iSCSI as one of several choices of host interfaces.

Hardware Targets

Several complete iSCSI storage solutions are included in the storage target sections below. Some of these solutions are targeted toward small and medium business, while others are targeted at enterprise data centers. There is a separate section in this report for solutions from:

- Hewlett-Packard (HP)
- NetApp
- Xiotech

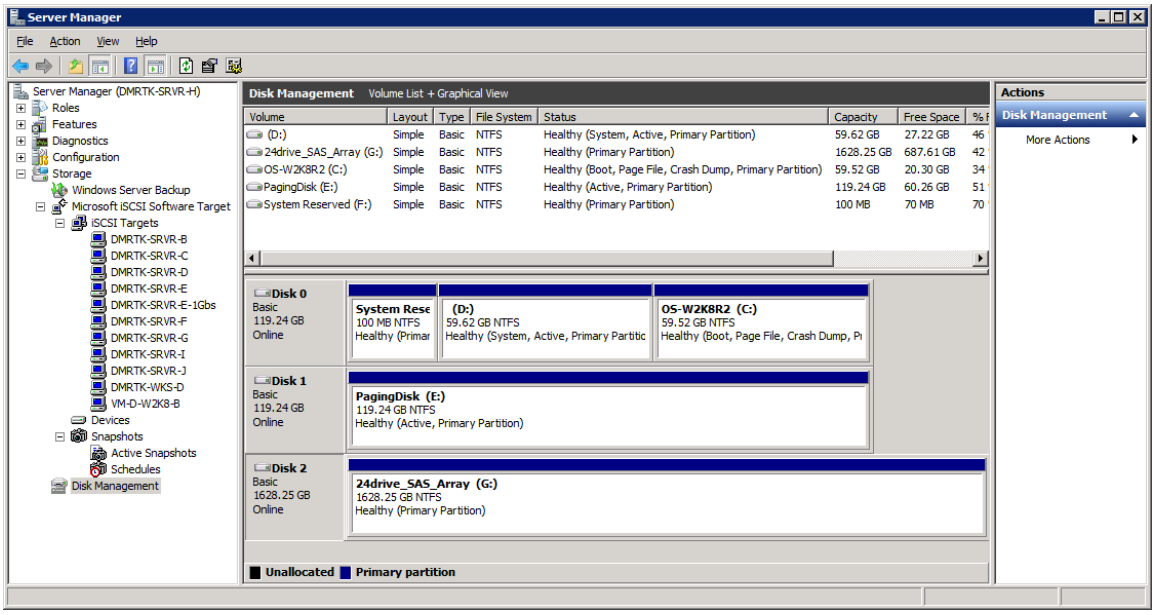
Software Targets

Ordinary servers can be made into iSCSI targets by implementing iSCSI target software and dedicating some of that server's storage to the iSCSI target software. Software targets provide a great deal of flexibility in the ways they can be implemented. For example, they can be clustered and scaled to any size that is supported by the operating system. In addition, the software targets can use either low-end or high-end storage systems behind the software targets, such as direct attached storage or other SAN storage. Advanced features such as replication, compression, de-duplication, encryption, or others can be combined with software targets.

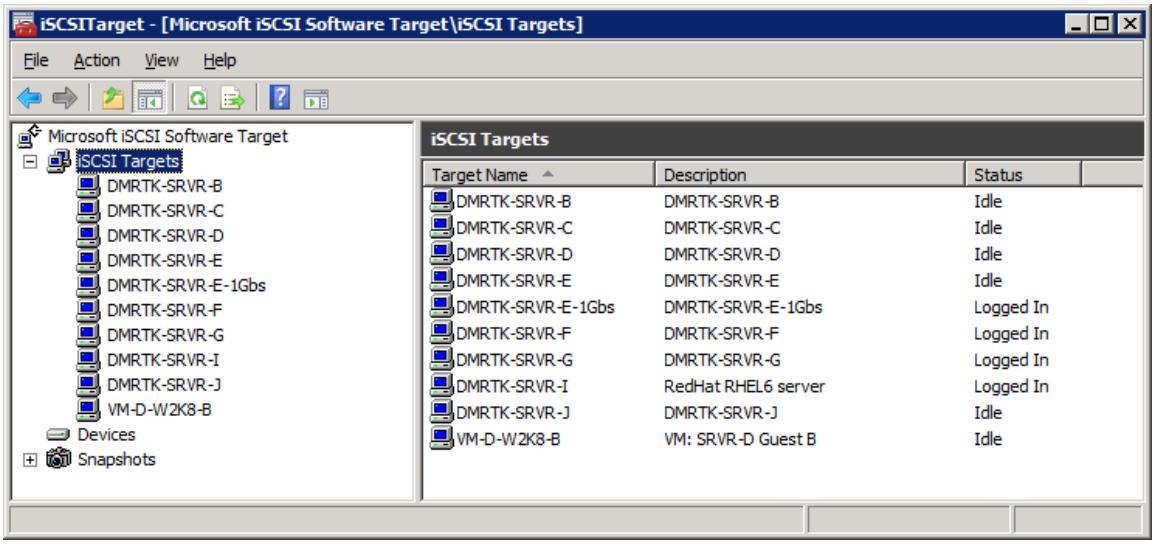
There are several iSCSI target software solutions available today. For example, Microsoft released its iSCSI Software Target version 3.3 in April 2011.

Microsoft iSCSI Software Target

We installed the Microsoft iSCSI target software version 3.3 onto one of our existing Windows Server 2008 R2 platforms. The Microsoft iSCSI software target management plug-in fits into Server Manager. We configured the software target to provide storage for several iSCSI clients, as shown below.

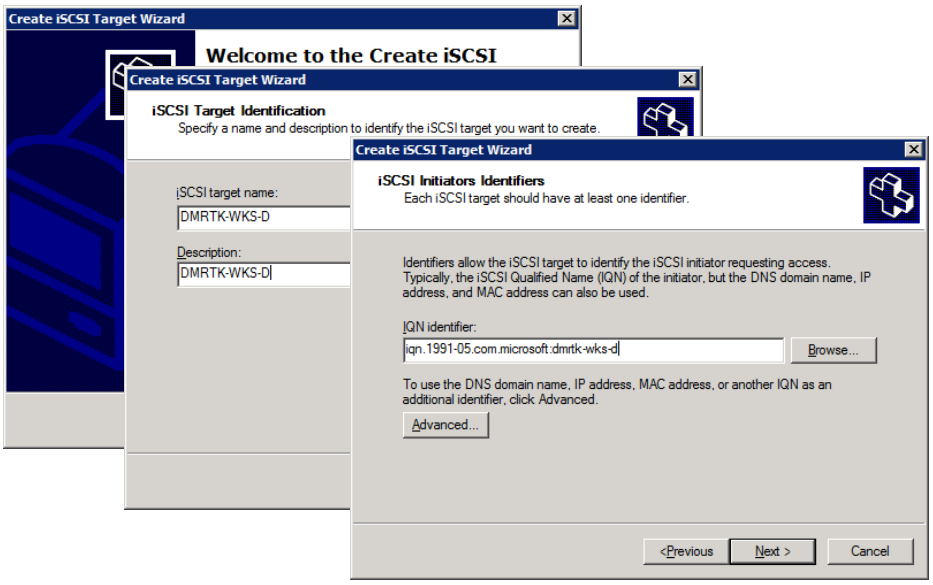


There is also a separate Microsoft iSCSI Software Target application found under Administrative Tools that can be used to manage the iSCSI software target environment.

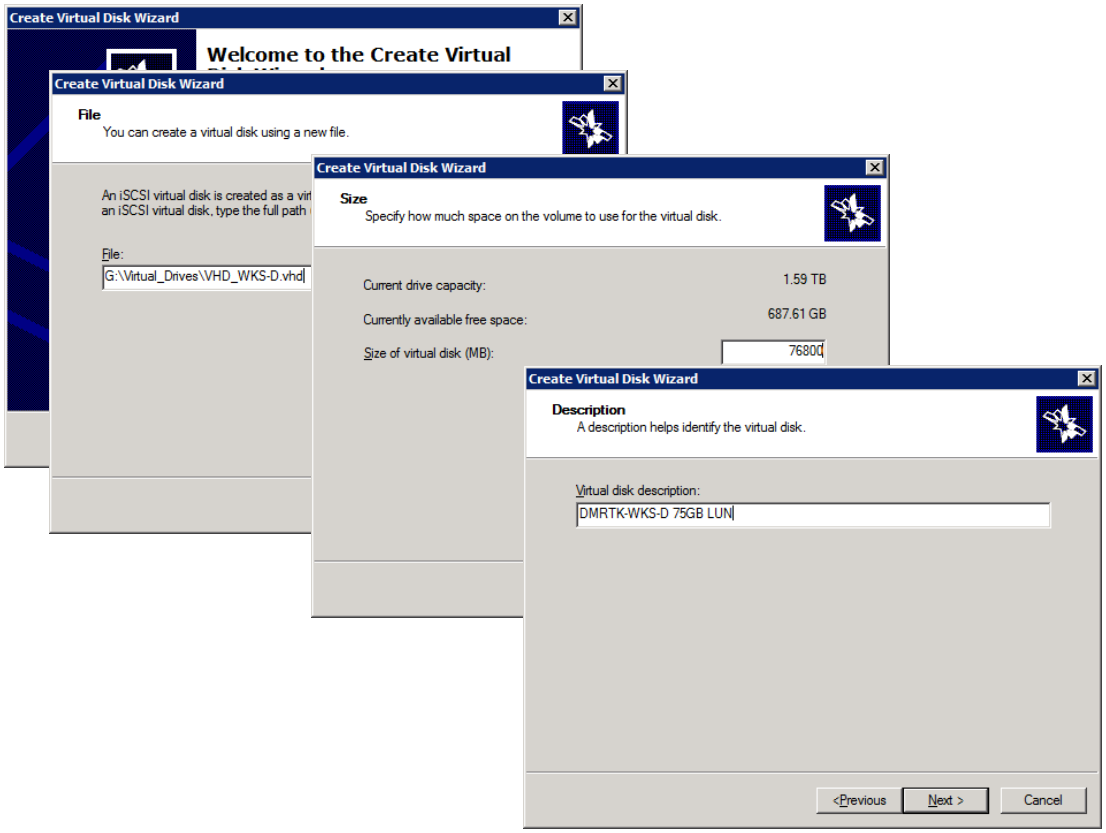


A wizard is used to create individual storage targets that can be assigned to individual iSCSI initiators. A second wizard is used to create the virtual disk for that iSCSI initiator.

Step 1 – Create the individual storage target



Step 2 – Create the Virtual Disk that becomes the storage for the individual initiator



Internet Storage Name Service (iSNS)

The Internet Storage Name Service (iSNS) protocol is used for interaction between iSNS servers and iSNS clients to facilitate automated discovery, management, and configuration of iSCSI devices on TCP/IP networks. It can also be used to discover Fibre Channel devices using iFCP gateways. The Microsoft iSNS Server only supports the discovery of iSCSI devices.

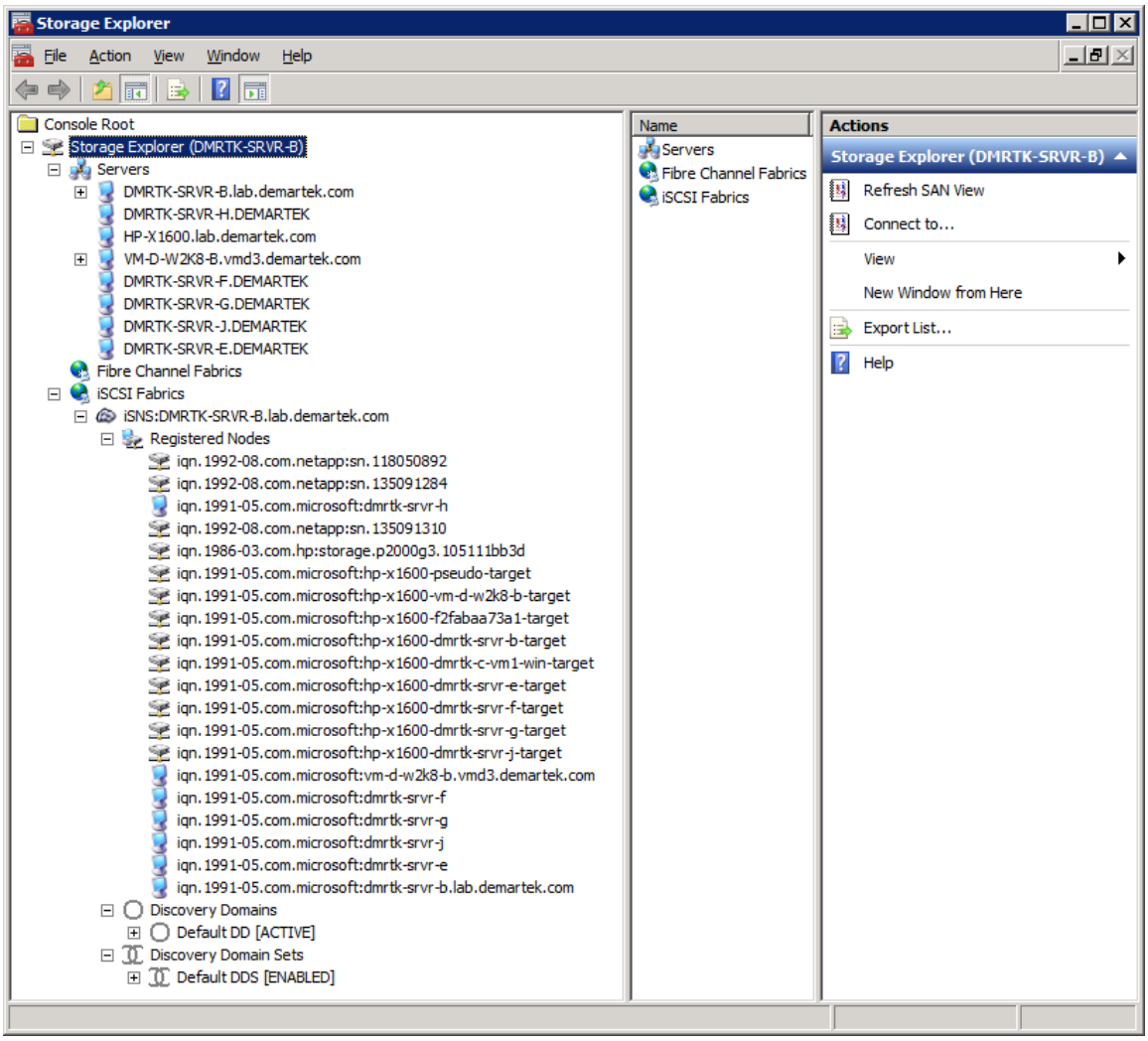
Deploying iSNS is not required for basic iSCSI functionality, however, its value increases as the number of initiators and targets in the iSCSI SAN fabric increases. As new iSCSI initiators are added to the fabric, they can be provided with the iSNS server IP address which allows them to automatically discover all the available targets without having to manually discover each target separately.

The iSNS is a feature that can be installed from the "Add Features" wizard in the Server Manager for Windows Server 2008 R2. Simply check the "Internet Storage Name Server" check-box and proceed with the installation. For best results, the iSNS server should be a member of a domain, although other servers using the iSNS service do not have to be members of a domain.

Windows Storage Explorer

Windows Storage Explorer is a management application found under Administrative Tools on Windows Servers that can display the iSCSI initiators and targets that have been discovered in the iSCSI SAN fabric via the iSNS server. The following screenshot was taken from the Storage Explorer running on the iSNS server used in the preparation for this report.

The registered nodes include iSCSI initiators running on physical and virtual servers and iSCSI targets. Some iSCSI targets are listed only once and other iSCSI targets have separate entries for each logical target that has been configured.



Windows Storage Explorer is available for server versions of the Windows operating system. It is also available for Windows 7 Professional, Enterprise, or Ultimate editions as part of the Remote Server Administration Tools (RSAT) free download. The RSAT download is available at http://technet.microsoft.com/en-us/library/ee449470(WS.10).aspx.

Virtualization and iSCSI

iSCSI technology is supported with popular server virtualization solutions such as VMware vSphere, Microsoft Hyper-V, and others. It is especially useful when the ability to move virtual machine guests from one physical server to a different physical server is highly desired. iSCSI can be used to store the VM image itself and can be used to provide separate volumes or LUNs to that guest operating system for use by various applications.

From a guest virtual machine viewpoint, iSCSI volumes or LUNs are straightforward and simple to access and to use. Most of the major applications that might be running in a virtual machine support iSCSI storage. It is possible to connect a VM guest directly to an iSCSI target without requiring any intervention from the hypervisor using either Microsoft Hyper-V or VMware ESX/ESXi. iSCSI volumes connected in this way can be added or removed without requiring a reboot.

Microsoft Windows Server 2008 R2 with Hyper-V

We used three different methods to configure iSCSI LUNs for guest virtual machines in a Windows Server 2008 R2 with Hyper-V environment. These are:

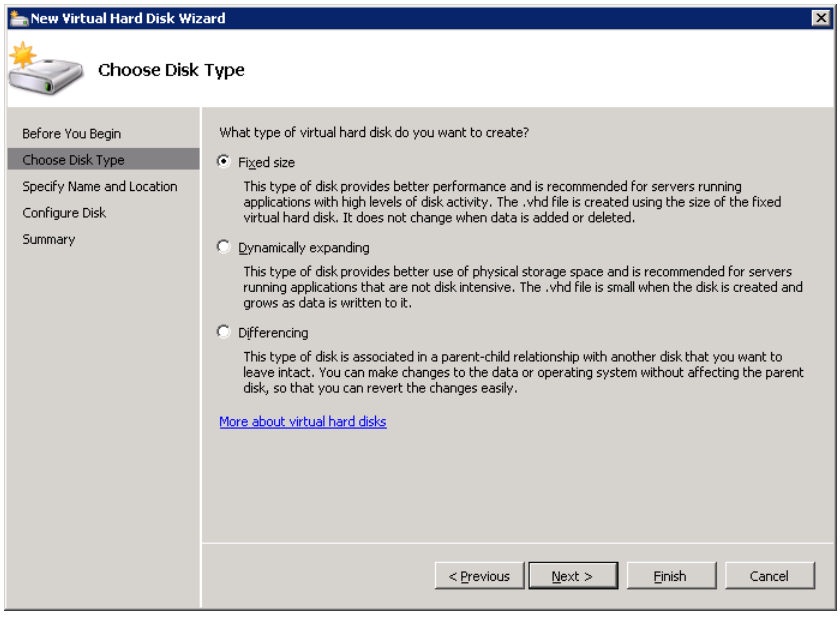
1. Virtual Hard Disk (VHD) method
2. Pass-through method
3. Direct from guest method

The first two methods require that the iSCSI volume be managed by, and be visible to, the hypervisor. The iSCSI volume created by the third method is not visible to the hypervisor.

Virtual Hard Disk (VHD) Method

In the VHD method, an iSCSI volume is allocated to the parent (Hyper-V) partition in the usual manner with the Microsoft iSCSI initiator or any other hardware or software iSCSI initiator. This iSCSI volume is brought online, initialized and formatted, and folders may be created, if desired.

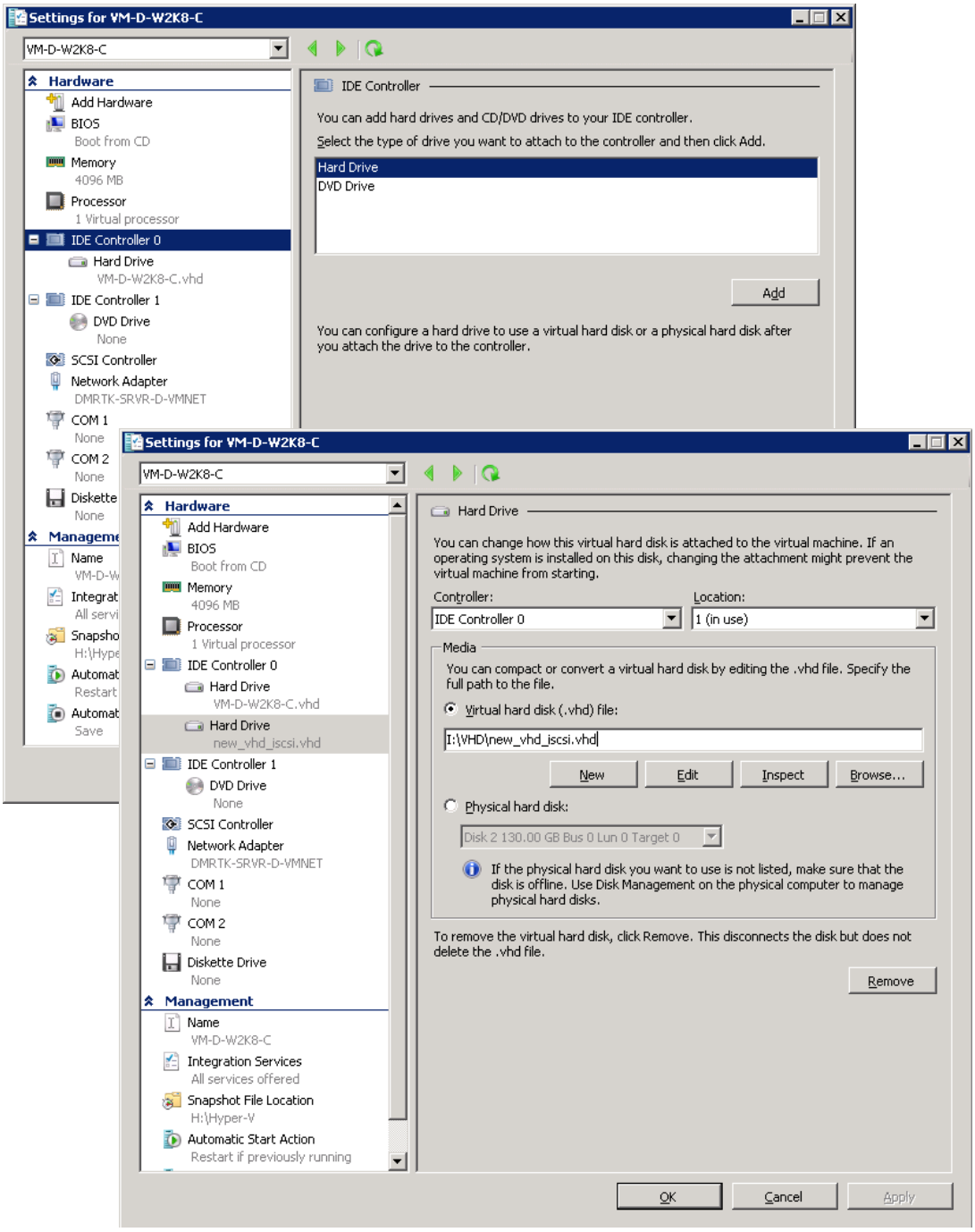
In the Hyper-V management role, we created a new hard disk using the "Actions" pane. There are three choices for the type of VHD: Fixed size, Dynamically expanding and Differencing. We chose fixed size.



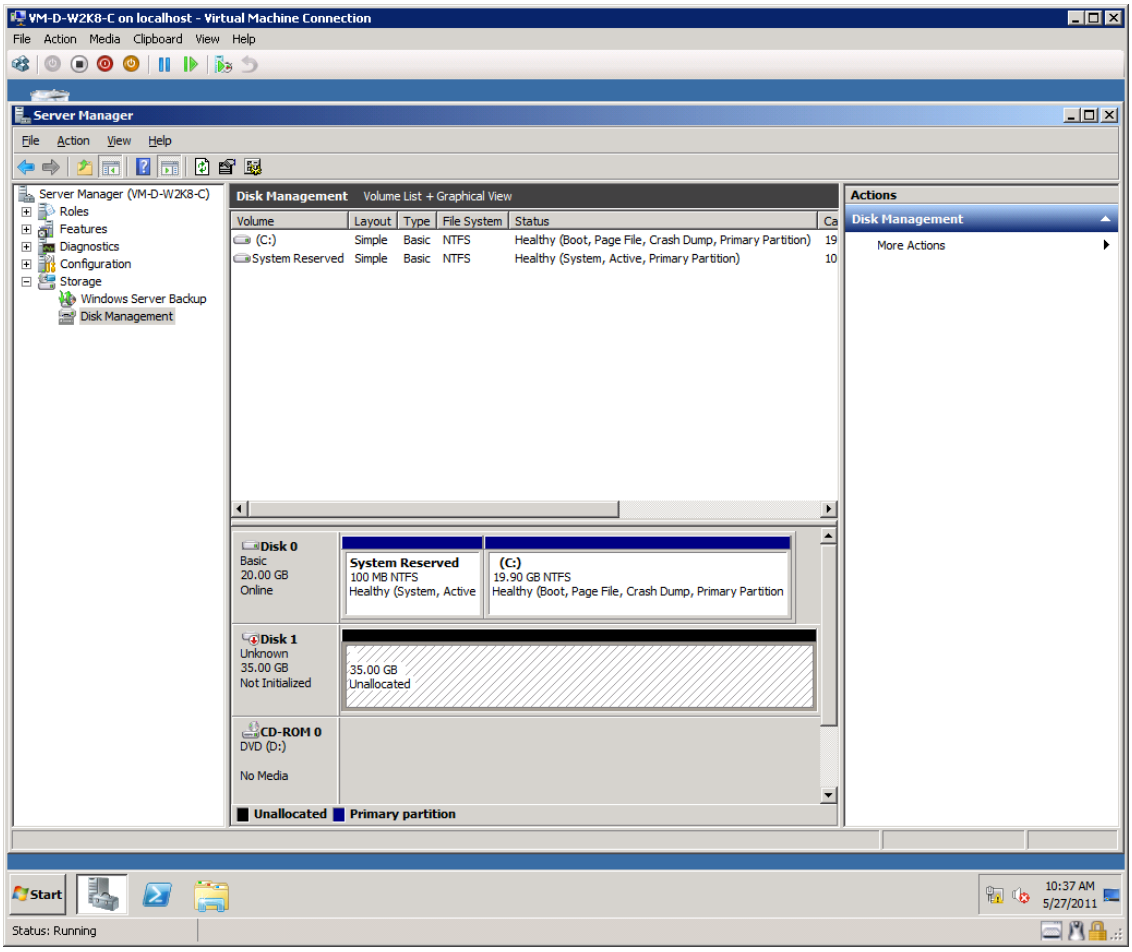
We placed the new VHD onto the iSCSI volume we just created, and allocated 35GB to this VHD.

In the Hyper-V management role, we select a virtual machine that is not running, and open "Settings." We choose to add a

hard drive under the IDE Controller, and specify the location of the VHD we just created on the new iSCSI volume.



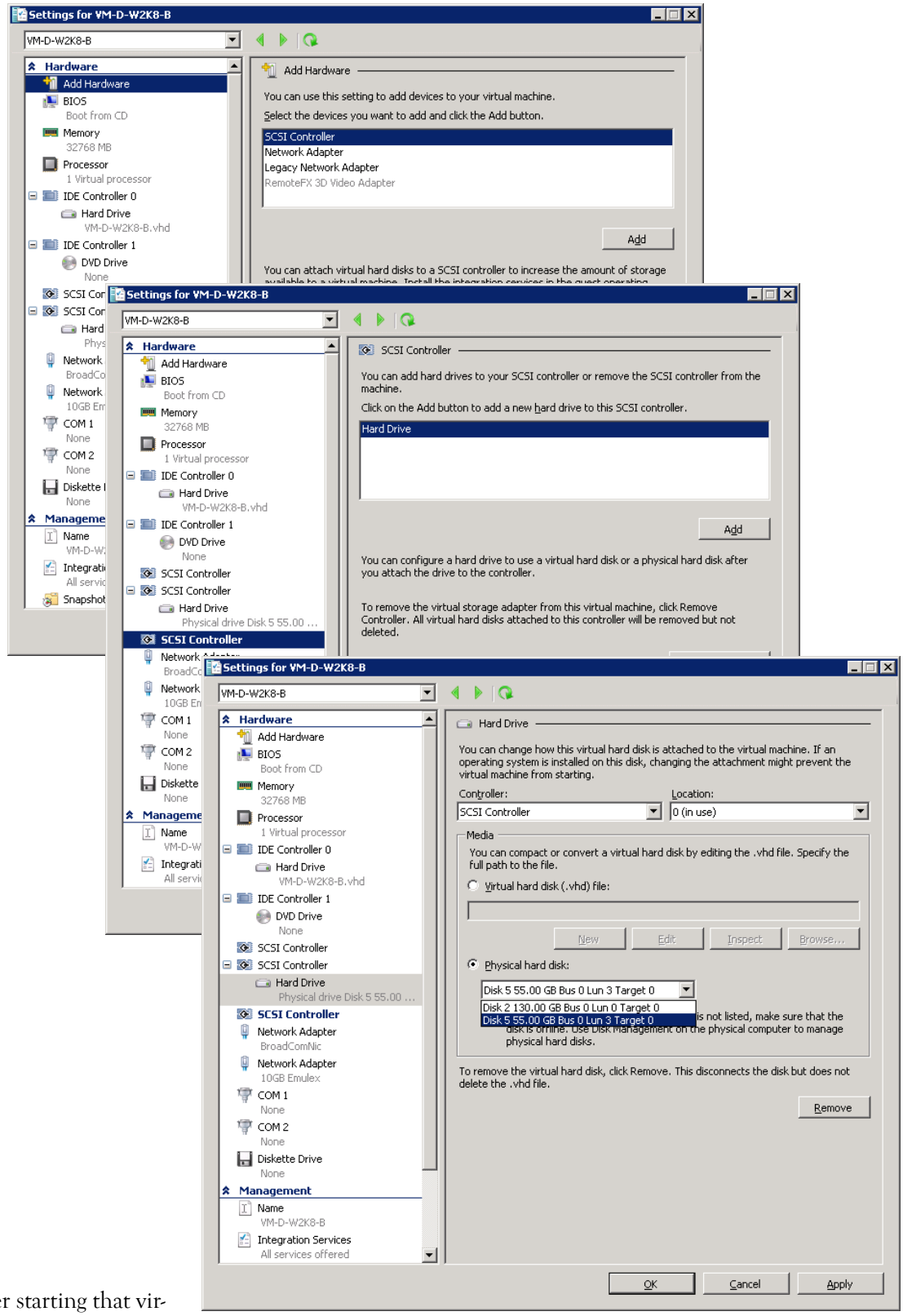
After starting that virtual machine, the new 35GB volume appears in Disk Manager for that virtual machine.



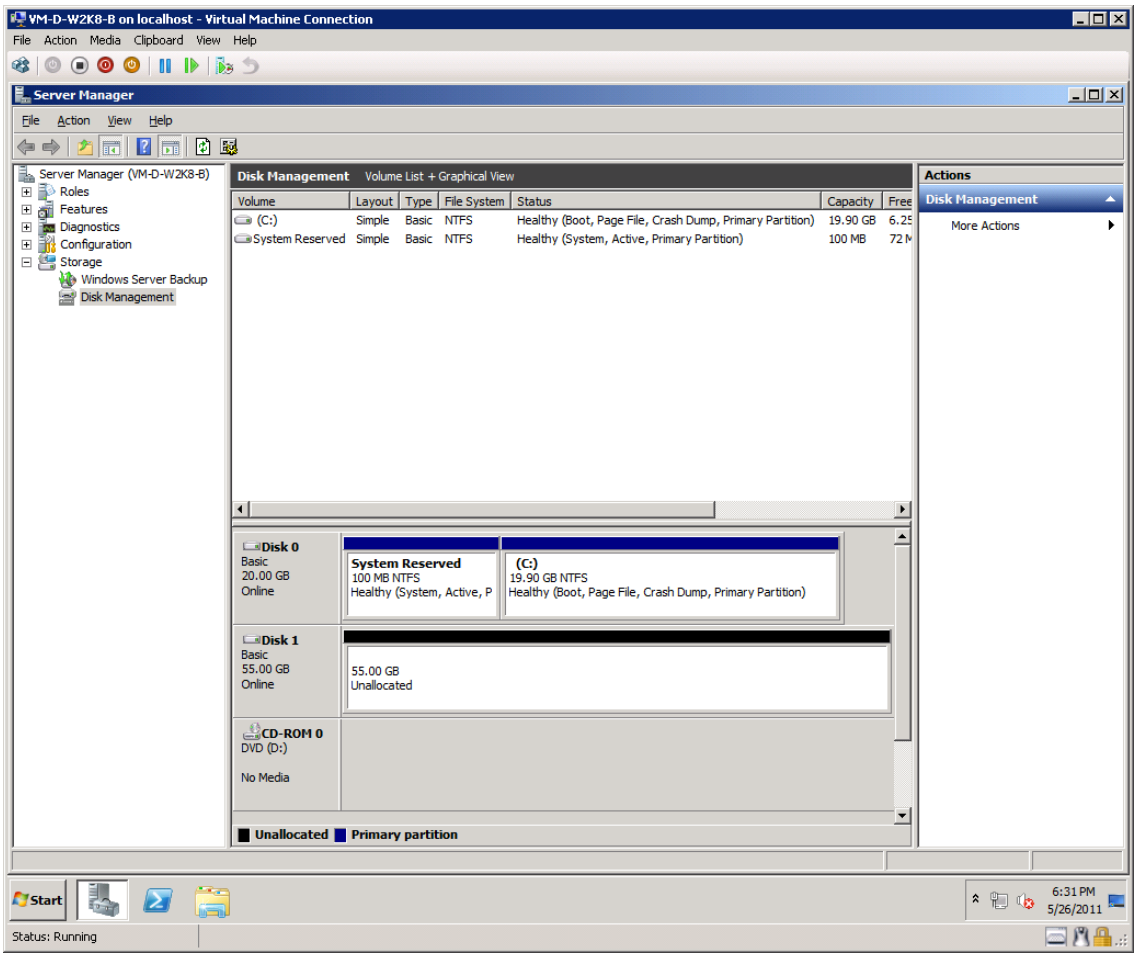
Pass-through Method

The pass-through method for allocating iSCSI LUNs to guest virtual machines requires that the iSCSI LUN be allocated on the parent partition in the usual manner, which is similar to the VHD method. The difference is that the iSCSI LUN is assigned to the guest by using the "settings" function for the guest VM in the Hyper-V management role, while the guest is not running.

The iSCSI LUN is given to the guest by the "Add Hardware" function and selecting the "SCSI Controller". In the SCSI Controller section, select "Hard Drive", then "Add". Choose "Physical hard disk." We created a 55GB LUN for this guest.

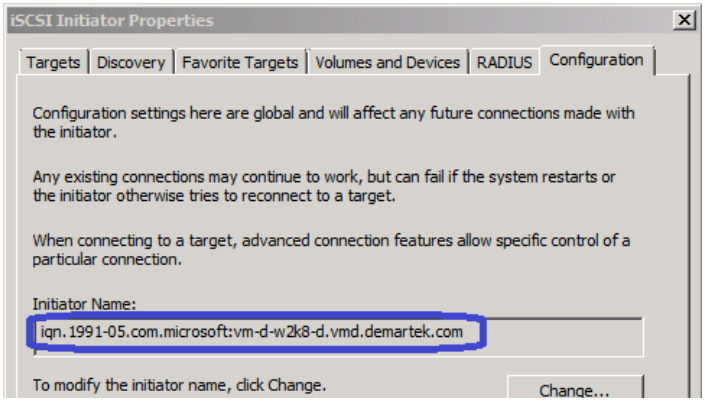


After starting that virtual machine, the new 55GB volume appears in Disk Manager for that virtual machine.

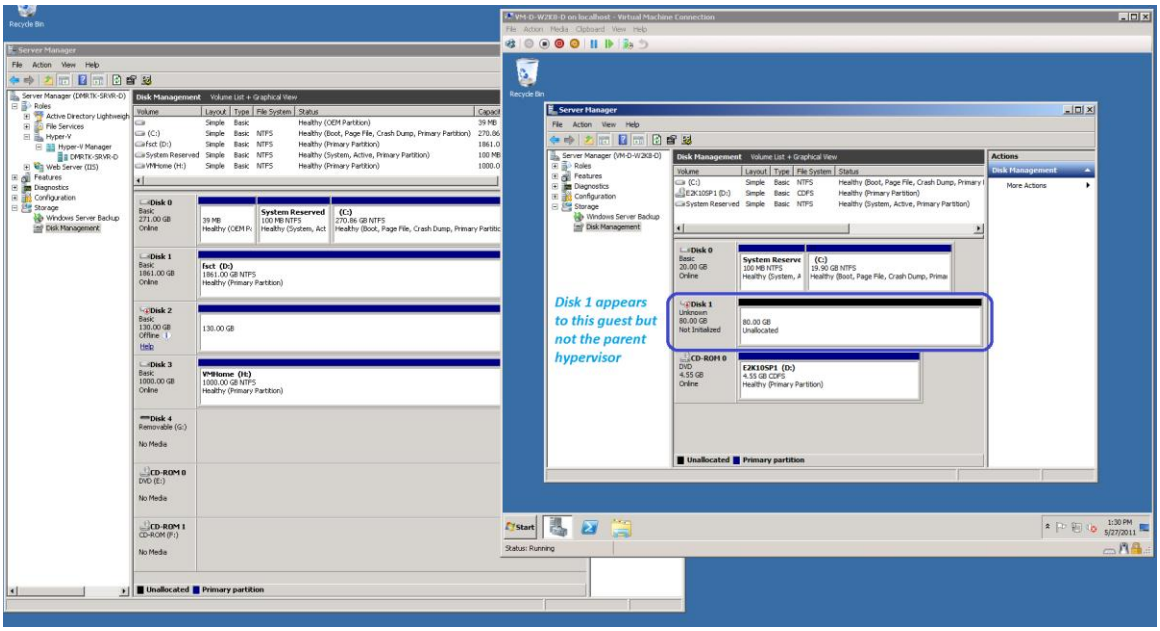


Direct from Guest Method

The direct from guest method is performed entirely within the guest operating system. It uses the normal iSCSI initiator steps of discovering the target portal and connecting to the target in the standard manner. As long as the guest can access the target using one of its network adapters, this process works the same as it does in physical servers.



Notice that the guest VM name is included as part of the IQN.
The 80GB iSCSI LUN appears as one of the disks for the guest, but does not appear to the hypervisor in the either the Disk Manager or the iSCSI initiator for the hypervisor.


Disk 1 appears to this guest but not the parent hypervisor

VMware vSphere 4.1

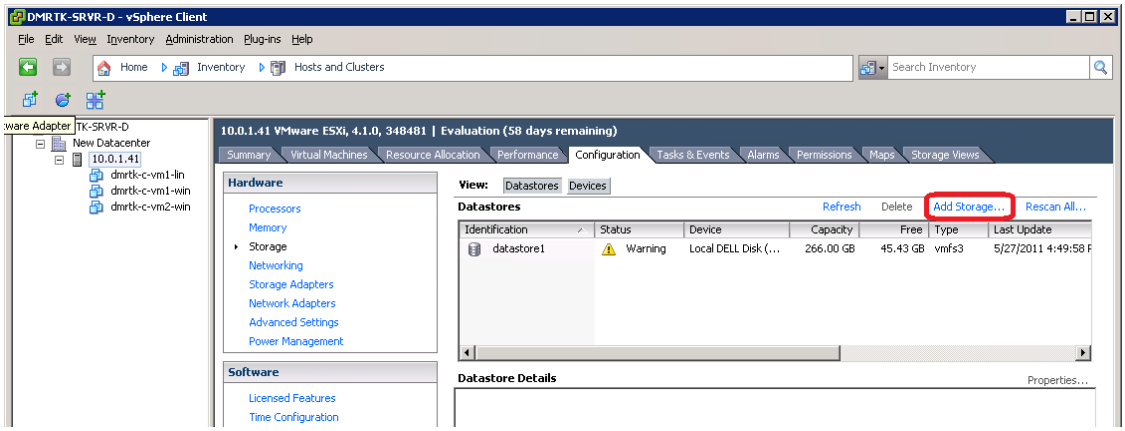
We used three different methods to configure iSCSI LUNs for guest virtual machines in a VMware vSphere 4.1 environment. These are:

1. Virtual Hard Disk (VHD) method
2. Pass-through method
3. Direct from guest method

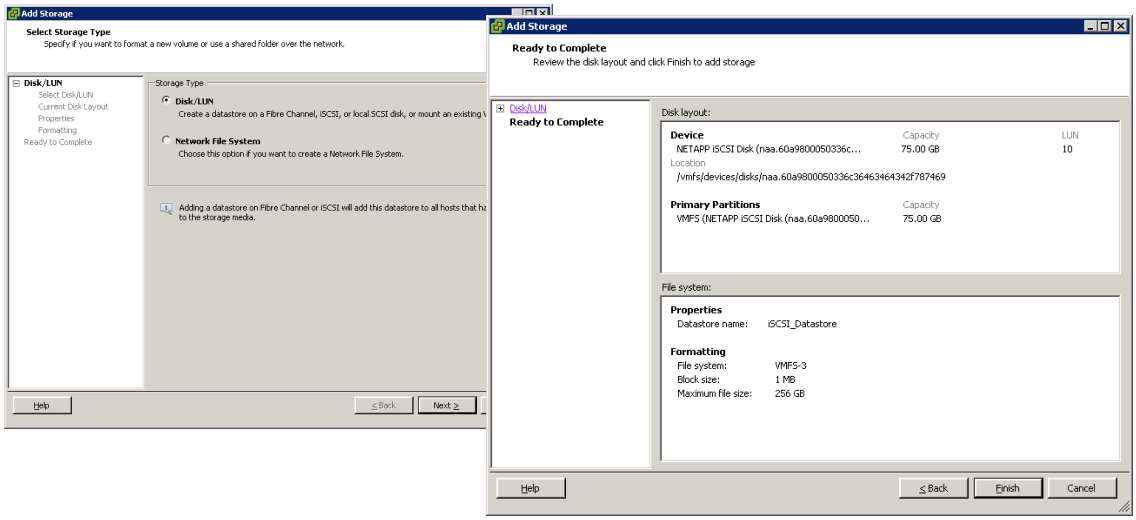
Virtual Hard Disk (VHD) method

Using the virtual hard disk method, an iSCSI LUN is created by the hypervisor (ESX or ESXi) via the vSphere console in the usual manner. The network adapter is already installed into the hypervisor system following the installation instructions provided by the network adapter vendor and is available to and configured for the VM kernel.

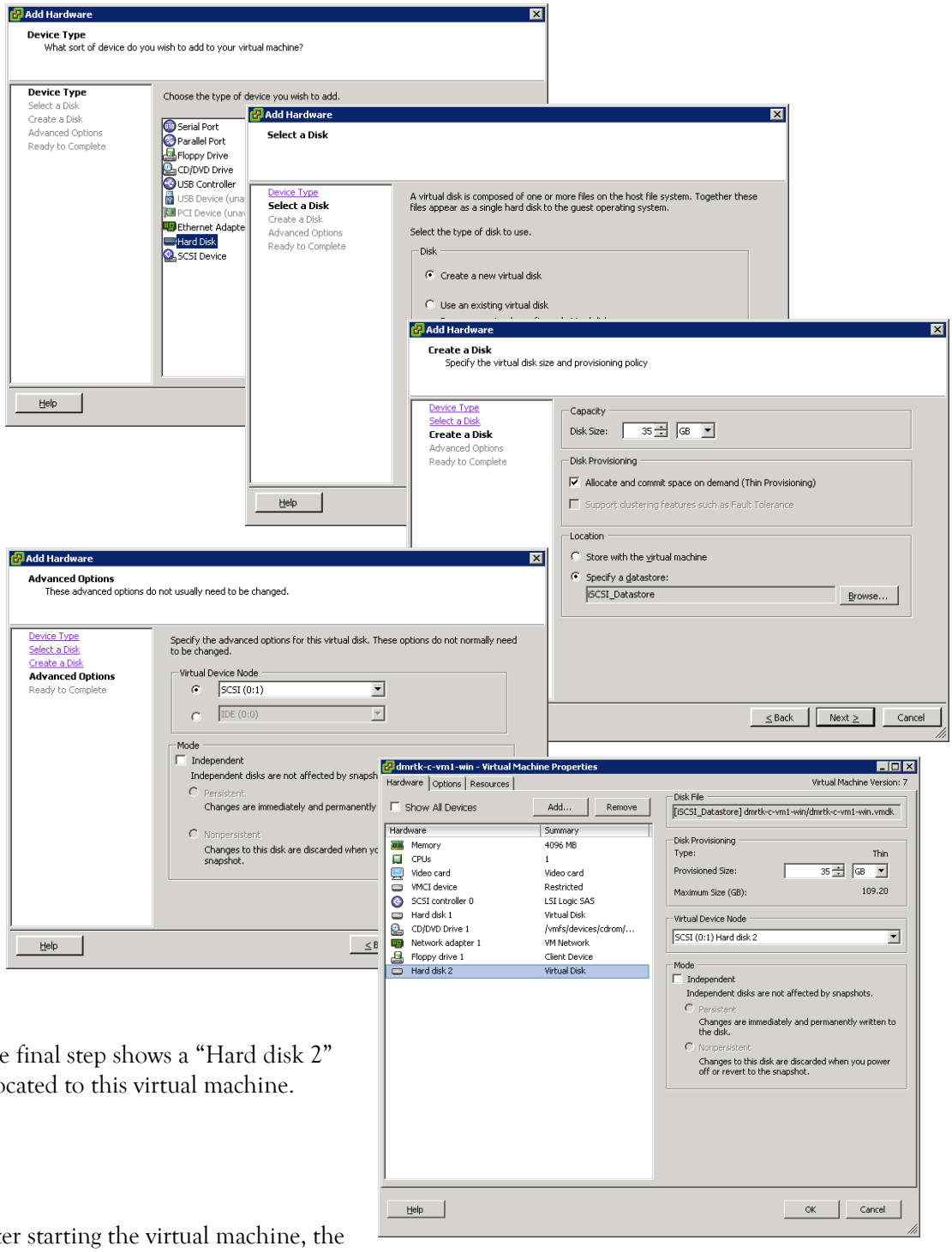
From the vSphere console Configuration tab, select "Storage", and then click "Add Storage".



The wizard steps through the process of selecting the Disk/LUN, providing the size of the LUN, name of the datastore, and the block size for VMFS. Here we select a 75GB LUN that has been created on one of the iSCSI targets. The first and last steps are shown below.

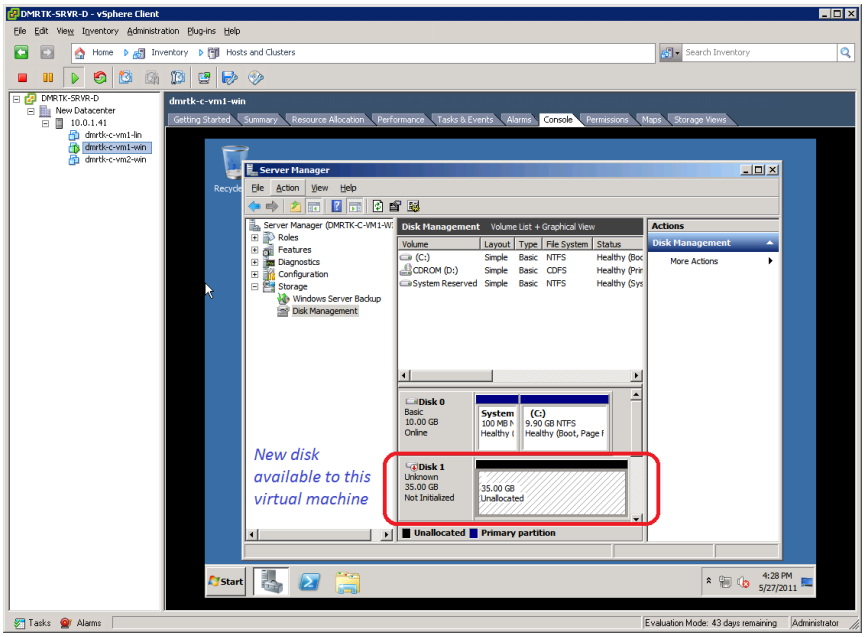


At this point, the new datastore (iSCSI LUN) is allocated to the virtual machine by selecting the virtual machine properties and adding hard disk hardware. A new virtual disk is created and linked to the datastore that we created previously. We selected the "on-demand" (thin provisioning) option.



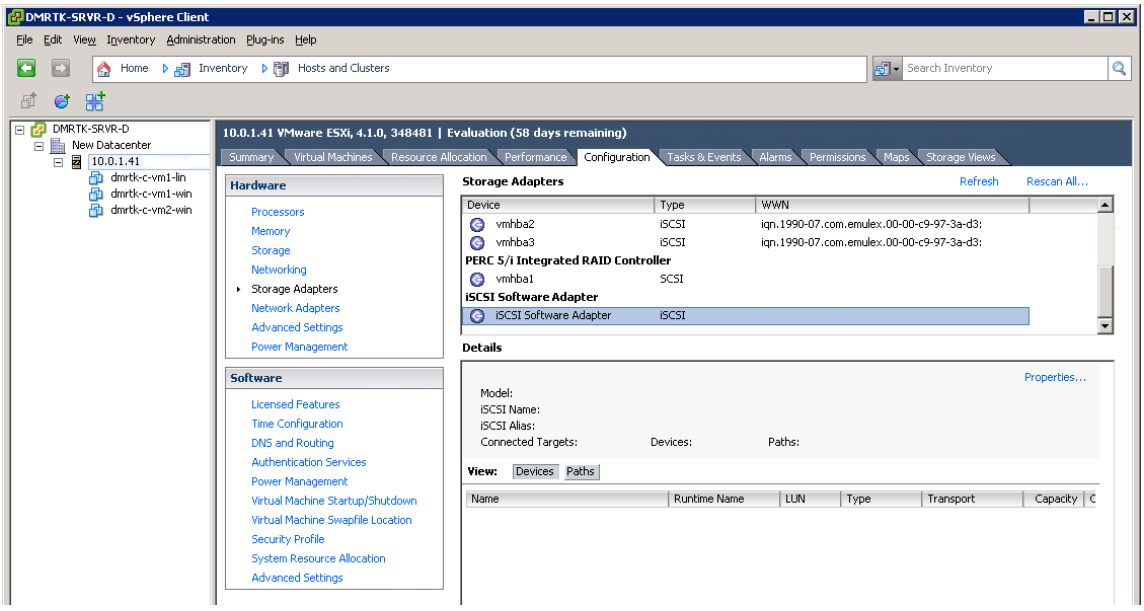
The final step shows a "Hard disk 2" allocated to this virtual machine.

After starting the virtual machine, the new 35GB LUN is now available to it.

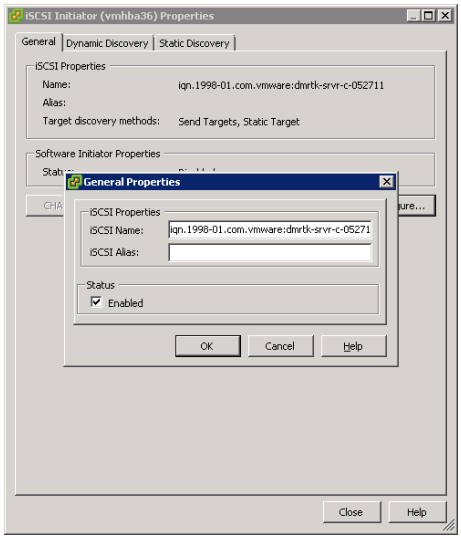


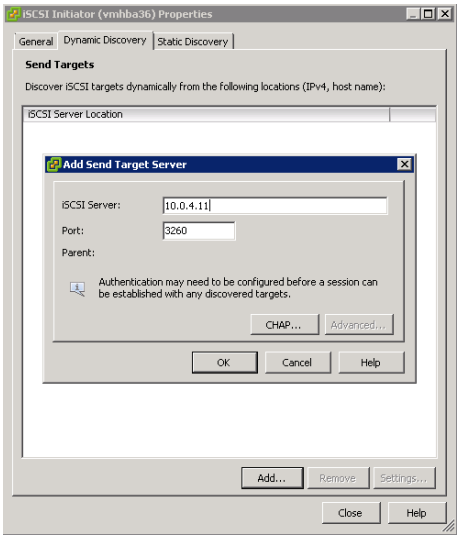
Pass-through Method

The pass-through method requires several steps to be completed. The network adapter is already installed into the hypervisor system following the installation instructions provided by the network adapter vendor and is available to and configured for the VM kernel. In this example, our network adapter is identified as an iSCSI adapter.

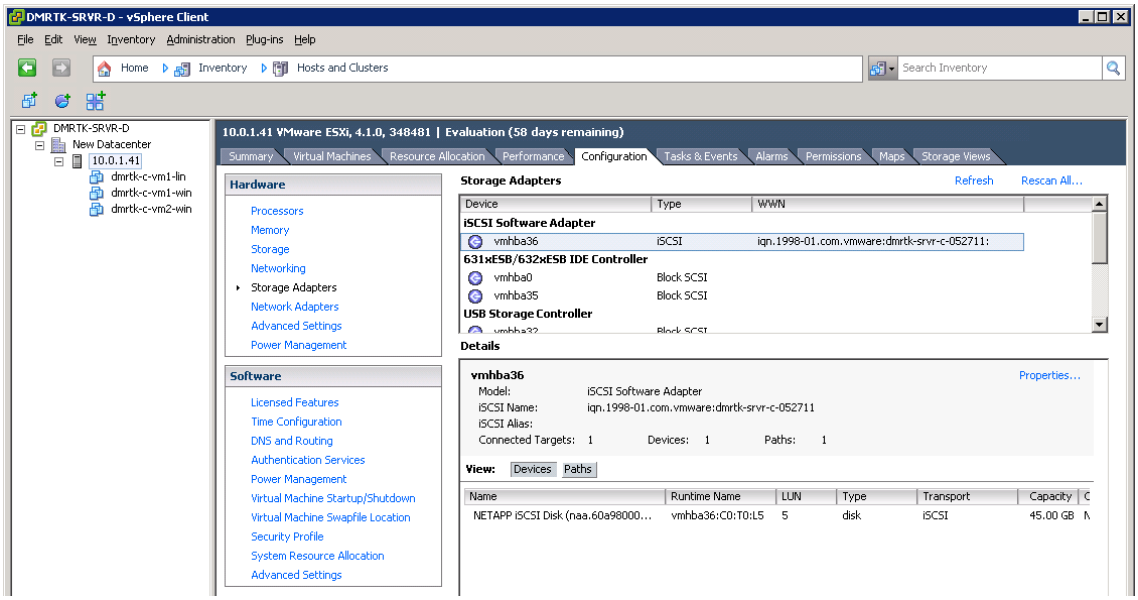


The iSCSI name must be configured and enabled, and provided to the target. To complete this step, select "Properties", then Configure, then Enable. After the IQN has been provided to the target, it can be discovered using the "Dynamic Discovery" tab.

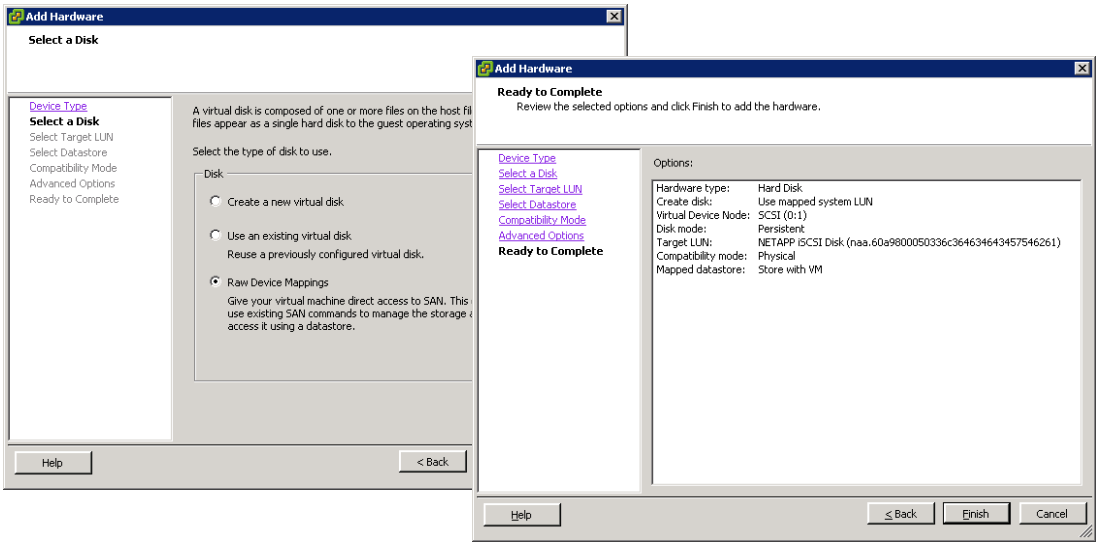




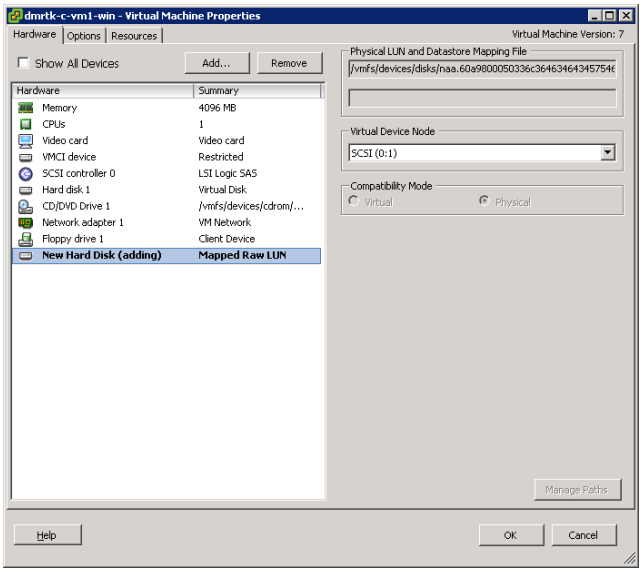
This process makes the iSCSI LUN visible and available to the hypervisor.



Now the iSCSI LUN is ready to allocate to the virtual machine. This process is similar to the virtual hard disk method above, except that this iSCSI LUN will be allocated using "raw device mappings" and is mapped with physical device mappings that are stored with the virtual machine. Selected screenshots from this process are shown below.

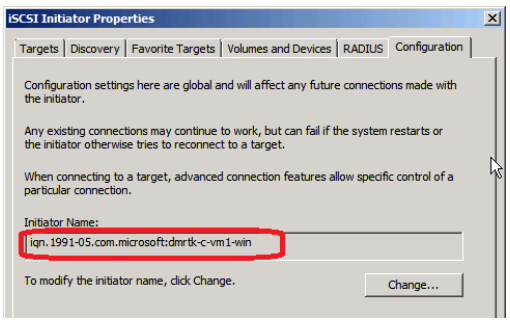


The iSCSI LUN is now available to the virtual machine and is described as a "Mapped Raw LUN".



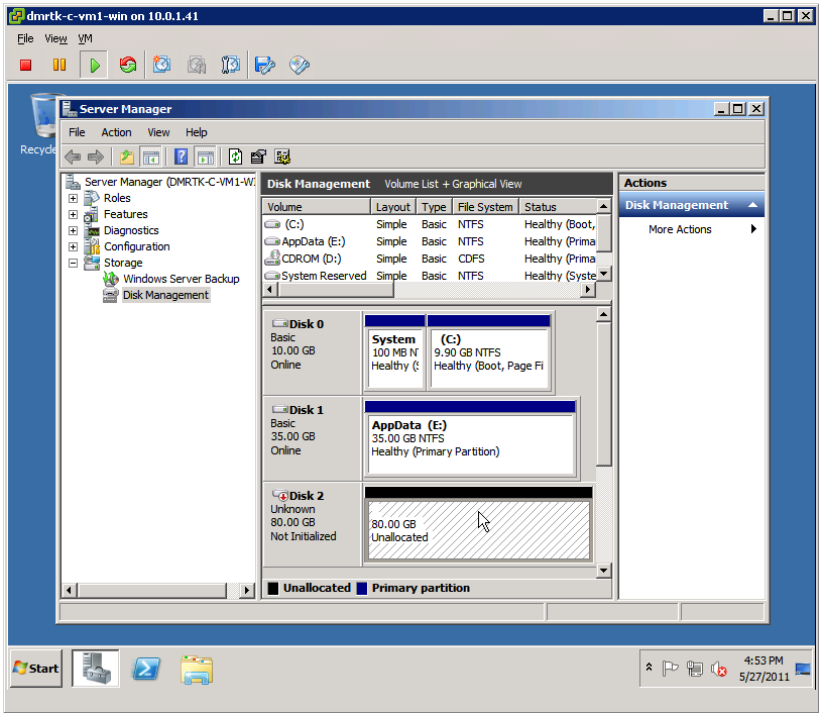
Direct from Guest Method

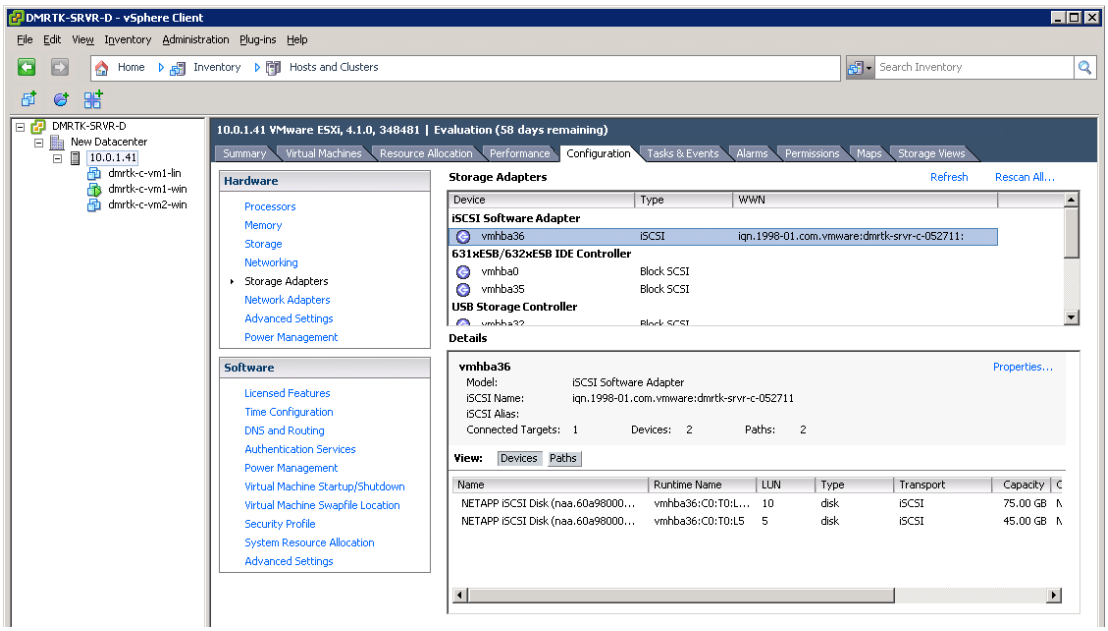
The direct from guest method is performed entirely within the guest operating system. It uses the normal iSCSI initiator steps of discovering the target portal and connecting to the target in the standard manner. As long as the guest can access the target using one of its network adapters, this process works the same as it does in physical servers.



Notice that the guest VM name is included as part of the IQN.

The new 80GB iSCSI LUN is visible to the virtual machine but not the hypervisor.


dmrtk-c-vm1-win on 10.0.1.41
File View VM
Server Manager
File Action View Help
Server Manager (DMRTK-C-VM1-W
Roles
Features
Diagnostics
Configuration
Storage
Windows Server Backup
Disk Management
Disk Management Volume List + Graphical View
Volume Layout Type File System Status
(C:) Simple Basic NTFS Healthy (Boot,
AppData (E:) Simple Basic NTFS Healthy (Prima
CDROM (D:) Simple Basic CDFS Healthy (Prima
System Reserved Simple Basic NTFS Healthy (Syste
Disk 0
Basic
10.00 GB
Online
System
100 MB N
Healthy (
(C:)
9.90 GB NTFS
Healthy (Boot, Page Fi
Disk 1
Basic
35.00 GB
Online
AppData (E:)
35.00 GB NTFS
Healthy (Primary Partition)
Disk 2
Unknown
80.00 GB
Not Initialized
80.00 GB
Unallocated
Unallocated Primary partition
Actions
Disk Management
More Actions
Start
4:53 PM
5/27/2011


DMRTK-SRVR-D - vSphere Client
File Edit View Inventory Administration Plug-ins Help
Home Inventory Hosts and Clusters
Search Inventory
DMRTK-SRVR-D
New Datacenter
10.0.1.41
dmrtk-c-vm1-lin
dmrtk-c-vm1-win
dmrtk-c-vm2-win
10.0.1.41 VMware ESXi, 4.1.0, 348481 | Evaluation (58 days remaining)
Summary Virtual Machines Resource Allocation Performance Configuration Tasks & Events Alarms Permissions Maps Storage Views
Hardware
Processors
Memory
Storage
Networking
Storage Adapters
Network Adapters
Advanced Settings
Power Management
Software
Licensed Features
Time Configuration
DNS and Routing
Authentication Services
Power Management
Virtual Machine Startup/Shutdown
Virtual Machine Swapfile Location
Security Profile
System Resource Allocation
Advanced Settings
Storage Adapters
Refresh
Rescan All...
Device Type WWN
iSCSI Software Adapter
vmhba36 iSCSI iqn.1998-01.com.vmware:dmrtk-srvr-c-052711:
631xESB/632xESB IDE Controller
vmhba0 Block SCSI
vmhba35 Block SCSI
USB Storage Controller
Details
vmhba36
Properties...
Model: iSCSI Software Adapter
iSCSI Name: iqn.1998-01.com.vmware:dmrtk-srvr-c-052711
iSCSI Alias:
Connected Targets: 1 Devices: 2 Paths: 2
View: Devices Paths
Name Runtime Name LUN Type Transport Capacity
NETAPP iSCSI Disk (naa.60a98000... vmhba36:C0:T0:L... 10 disk iSCSI 75.00 GB
NETAPP iSCSI Disk (naa.60a98000... vmhba36:C0:T0:L5 5 disk iSCSI 45.00 GB

Measuring iSCSI Performance

Performance measurement for iSCSI has been the subject of much discussion. Because iSCSI operates both as a network protocol and a storage protocol, iSCSI performance can be affected by many factors. From the networking perspective, iSCSI is a multilayer protocol, including the iSCSI, TCP and IP layers, and can run over single or multiple 1Gb or 10Gb links. It is subject to all the same factors that affect network performance. From the storage perspective, performance is affected by all of the usual storage performance factors, including the number, speed and type of the disk drives, as well as various RAID settings, the backend architecture of the storage unit or system, and the processing power of the storage controller, among other things. Obviously the host server platform, including CPU and memory, plays a role. The application workload and its use of storage also has a significant effect on performance.

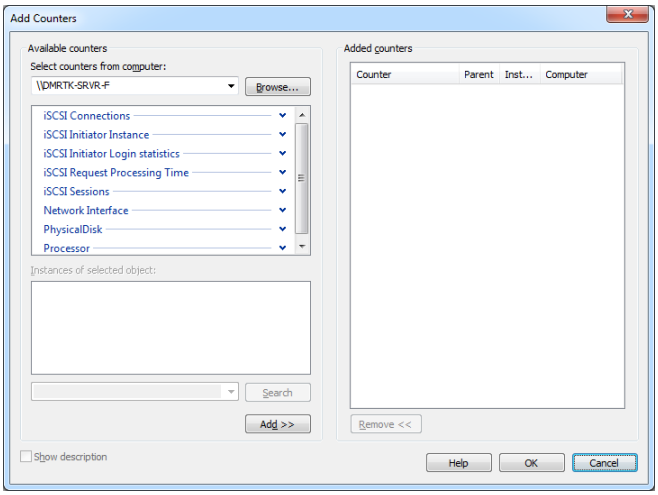
With the increasing use of 10Gb Ethernet for iSCSI storage, concerns about performance for iSCSI storage systems are decreasing. In our lab tests, it has not been difficult to completely consume all the bandwidth of a 1Gb iSCSI connection between one host and one iSCSI storage target. When this happens, it suggests that the network is the bottleneck. This also explains one of the best practices for iSCSI regarding separating iSCSI storage traffic from regular LAN traffic, at least for 1Gb networks. However we have not yet been able to consume all the bandwidth of a 10Gb iSCSI connection between one host and one iSCSI storage target using real-world applications and real-world storage targets. In one test, we were able to achieve a sustained rate of approximately 35% of the bandwidth of a 10Gb iSCSI connection between one host running a single application and a three rack-unit (3U) iSCSI storage target that used a combination of solid-state storage and hard disk drive technology. This reinforces our opinion that 10Gb networks and solid-state storage technologies work well together. In another test, we were able to achieve a sustained rate of more than 50% of the bandwidth of a 10Gb iSCSI connection. These examples show that 10GbE connections used for iSCSI allow for plenty of room for growth for application performance.

Some vendors specialize in iSCSI storage and offer products that have 1Gb and 10Gb iSCSI host interfaces. Several vendors provide storage systems that have multiple choices for the host interfaces including 10Gb iSCSI and 8Gb Fibre Channel. Some vendors add 10Gb FCoE as a choice for the host interface. With these types of choices for the host interfaces for the same storage system, the interface is much less of an issue with respect to performance of a storage system.

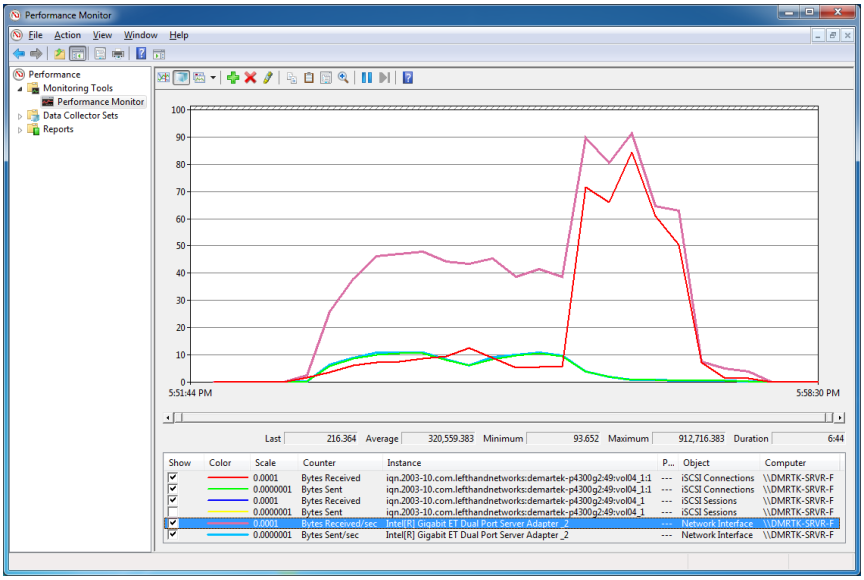
Slower-speed WAN environments can be used for iSCSI storage where latency and bandwidth are not a major concern. For example, replication to a disaster recovery site can be performed via iSCSI. This type of replication can be performed by a host server or directly by some storage systems.

Windows Performance Monitor (Perfmon)

When an iSCSI initiator is present on a Windows operating system, several performance counters are automatically loaded into the native Windows performance monitor application, known as Perfmon. At a minimum, when we measure iSCSI performance on a Windows system, we include all the iSCSI counters, the network interface counters, the physical disk counters, and the processor counters in a Perfmon trace, as shown in the screenshot below.



Below is a sample of a Perfmon trace from a Windows Server 2008 R2 host system using a single one gigabit connection to an iSCSI storage system. The trace of the iSCSI connection and the network interface trace show some similarities in performance. The network interface shows a slightly higher number of bytes received per second than the iSCSI connection because the iSCSI connection statistics do not include some over the TCP/IP overhead.



Exchange Jetstress Performance Testing

Microsoft provides Exchange Jetstress as a tool to stress test storage systems to determine if a particular configuration of a storage system is capable of adequately supporting an Exchange Server environment. Jetstress provides a Pass/Fail rating based on combination of the achieved IOPS, database read latencies, log write latencies, and some other factors. Jetstress can be used by end users as an informal stress test tool and by vendors for a more formal submission to Microsoft for the Exchange Solution Reviewed Program (ESRP). The latest vendor ESRP submissions are available at http://technet.microsoft.com/en-us/exchange/ff182054.aspx.

Microsoft updates Exchange Server periodically with new versions. The current and previous two versions of Exchange Server are Exchange Server 2010, Exchange Server 2007, and Exchange Server 2003. Customers move to these new versions at various rates. Some have moved to the 2010 version, some have moved to the 2007 version and many are still running the 2003 version. Each of these versions of Exchange Server have different I/O profiles, due to different architectures (32-bit vs. 64-bit), different database page sizes and other changes made over time. In order to help administrators get a better understanding of the I/O profiles of these versions of Exchange Server, we have compiled the Exchange Server 2003 vs 2007 vs 2010 I/O Comparison Summary on our website. For example, the achieved database IOPS results for Jetstress 2010 use a blocksize of 32KB.

We used Jetstress to stress test some of the iSCSI storage solutions in this guide. These test results are intended to provide a basic understanding of the potential performance of the storage solution using an application workload. The tests were deliberately configured differently for each storage solution and adapter so that these results would not be construed to be a direct competitive performance benchmark study.

Network Adapter Performance Tests

For the performance measurement of each of the 10Gb network adapters, we used the following configuration for the host initiator server and the storage target. We used Perfmon to capture the aggregate disk bytes/second statistics as viewed from the initiator host.

NOTE: The performance tests for the adapters were not an attempt to obtain maximum performance of the adapter, but a representative sample of real application performance using one host initiator server connected to one or two real-world storage targets. These storage targets use either SSDs or spinning hard drives. We believe that higher performance is possible with any of the adapters that we tested.

The adapter performance test results show the performance of 10GbE adapters, either with or without full iSCSI offload. These test results

Host initiator server specifications:

- Qty. 2: Intel Xeon X5570 processors, 2.93 GHz, 8 total cores, 16 total logical processors
- 24GB RAM
- Windows Server 2008 R2 Enterprise Edition

Storage target specifications:

- Qty. 2: Intel Xeon E5540 processors, 2.53 GHz, 8 total cores, 16 total logical processors
- 48GB RAM

- Windows Server 2008 R2 Enterprise Edition
- Microsoft iSCSI Target software version 3.3
- Internal SATA SSD boot drive
- LSI MegaRAID 9260-8i RAID controller connected to 8x SmartModular 200GB SAS SSD configured as 1.45TB RAID0 stripe, 64KB format allocation
- Motherboard LSI 2008 SAS controller connected to 4x Intel 32GB SLC SSD configured as 115GB RAID0 stripe, 64KB format allocation

iSCSI Best Practices

Here are a few best practices for deploying iSCSI. This is not necessarily an exhaustive or complete list, neither are these absolute requirements in order to use iSCSI.

Networking

- Deploy iSCSI on gigabit or faster networks.
- Using server-class network adapters for iSCSI traffic
  - Enable RSS, stateless offload and other server-class network adapter features
- Use one-way or mutual CHAP to increase the security between iSCSI initiators and targets.
- Segregate iSCSI storage traffic from regular LAN traffic.
  - This may not be required on 10Gb networks for performance reasons but still might be a good idea for security reasons.
- Use non-blocking switches for iSCSI storage traffic.
- Disable unicast storm control on iSCSI ports. Most switches have unicast storm control disabled by default.
- Unbind or disable File and Print Sharing protocols on the adapters dedicated to iSCSI traffic

Applications

- Microsoft Exchange Server and SQL Server support the use of iSCSI storage in cluster and non-cluster configurations.
  - Follow the usual storage guidelines for separating database disk volumes from log disk volumes.

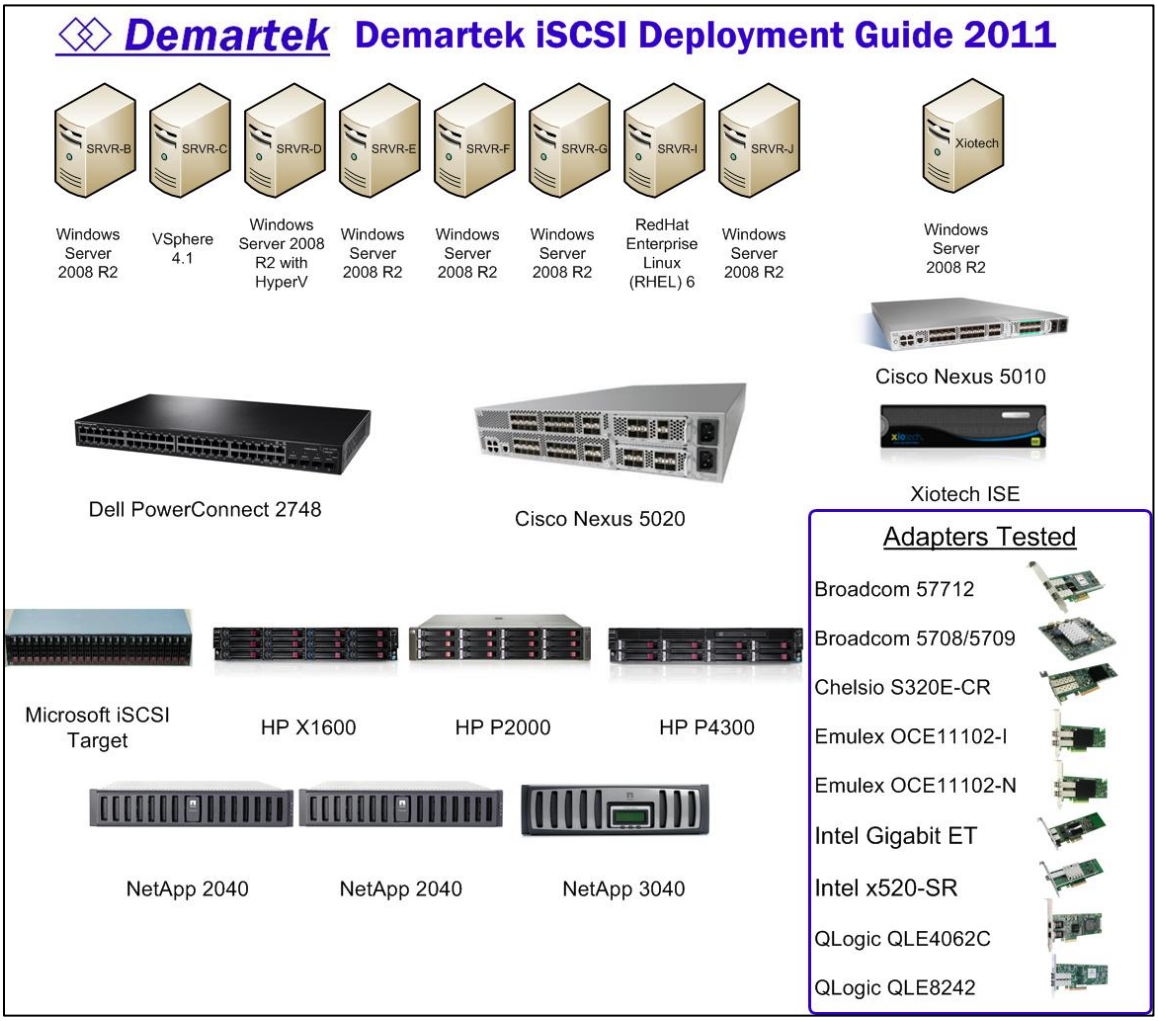
Multi-Path I/O

- Use MPIO to provide load balancing and failover capabilities
- For high-availability enterprise environments, configure redundant adapters and switches for iSCSI traffic, similar to high availability Fibre Channel storage environments.

Test Environment for this Guide

To run the tests required to produce this guide, we used nine physical servers in the Demartek lab in Arvada, Colorado, USA. Eight of the servers were used as initiators and one server was made into a target by installing the Microsoft iSCSI target software. One desktop workstation was also used for some of the tests. Six other target storage devices were provided by other vendors for this test. All eight initiator servers and seven target storage devices were run in our lab simultaneously, with some initiators and targets connected via 1Gb interfaces, some with 10Gb interfaces and some with both. Configurations of the servers, adapters and storage targets were changed to perform various tests, so the diagram below does not show all the network interconnections.

We ran other iSCSI tests in a nearby vendor lab using their storage device and host servers in a 10Gb environment.



Vendor Products Tested for this Guide

Specific network adapter and storage products were tested to produce this deployment guide. A vendor-specific section in this document provides additional information for these products. These were:

Network Adapters

- Broadcom BCM5709C and BCM57712
- Chelsio S320E-CR
- Emulex OCe11102-I and OCe11102-N
- Intel Gigabit ET and X520-SR
- QLogic QLE4062 and QLE8242

Storage Targets

- Hewlett-Packard (HP) X1600, P2000 and P4000
- NetApp FAS2040 and FAS3040
- Xiotech ISE and ISE-Hybrid

Initiator Specifications

The initiator servers were configured with various combinations of 1Gb and 10Gb network adapters and were connected to different storage targets at different times. Some of the storage targets had 1Gb host interfaces, some had 10Gb host interfaces, and some had both types of host interface.

Most of the initiator servers were physical servers running Windows Server 2008 R2. One server was running VMware vSphere 4.1 with guest virtual machines. One server was running Microsoft Hyper-V with guest virtual machines. One server was running Red Hat enterprise Linux (RHEL) 6. The desktop workstation was running Windows 7 Ultimate 64-bit.

DMRTK-SRVR-B

- **Operating System:** Windows Server 2008 R2
- **Processor:** Qty. 1 – Intel Pentium D 950, dual-core, 3.4GHz – Total logical processors: 2
- **RAM:** 8GB
- **Storage:** Internal 7200 RPM SATA disk drives

DMRTK-SRVR-C

- **Operating System:** VMware vSphere 4.1
- **Processor:** Qty. 2 – Intel Xeon E5345, quad-core, 2.33GHz – Total logical processors: 8
- **RAM:** 48GB
- **Storage:** Internal SAS disk array, 6x 15K RPM SAS disk drives

DMRTK-SRVR-D

- **Operating System:** Windows Server 2008 R2 with Hyper-V
- **Processor:** Qty. 2 – Intel Xeon E5345, quad-core, 2.33GHz – Total logical processors: 8
- **RAM:** 48GB
- **Storage:** Internal SAS disk array, 4x 15K RPM SAS disk drives, 4x 7200 RPM SATA disk drives

DMRTK-SRVR-E

- **Operating System:** Windows Server 2008 R2
- **Processor:** Qty. 2 – Intel Xeon E5345, quad-core, 2.33GHz – Total logical processors: 8
- **RAM:** 16GB
- **Storage:** Internal SAS disk array, 8x 15K RPM SAS disk drives

DMRTK-SRVR-F

- **Operating System:** Windows Server 2008 R2
- **Processor:** Qty. 2 – Intel Xeon E5345, quad-core, 2.33GHz – Total logical processors: 8
- **RAM:** 16GB
- **Storage:** Internal SAS disk array, 8x 15K RPM SAS disk drives

DMRTK-SRVR-G

- **Operating System:** Windows Server 2008 R2
- **Processor:** Qty. 2 – Intel Xeon E5345, quad-core, 2.33GHz – Total logical processors: 8
- **RAM:** 32GB
- **Storage:** Internal SAS disk array, 5x 15K RPM SAS disk drives

DMRTK-SRVR-I

- **Operating System:** RedHat Enterprise Linux (RHEL) 6
- **Processor:** Qty. 2 – Intel Xeon E5320, quad-core, 1.86GHz – Total logical processors: 8
- **RAM:** 8GB
- **Storage:** Internal 7200 RPM SATA disk drives

DMRTK-SRVR-J

- **Operating System:** Windows Server 2008 R2
- **Processor:** Qty. 2 – Intel Xeon X5570, quad-core, 2.93GHz – Total logical processors: 16
- **RAM:** 16GB
- **Storage:** Internal SAS disk array, 2x 10K RPM SAS disk drives, 3x 150GB SSDs

DMRTK-WKS-D

- **Operating System:** Windows 7 Ultimate 64-bit
- **Processor:** Qty. 1 – Intel Core i7 860, quad-core, 2.8GHz – Total logical processors: 8
- **RAM:** 8GB
- **Storage:** Internal SATA SSD boot drive, 7200 RPM SATA disk drives

COSNAS10E-N2 (Xiotech Lab)

- **Operating System:** Windows Server 2008 R2
- **Processor:** Qty. 2 – Intel Xeon E5620, quad-core, 2.40GHz – Total logical processors: 16
- **RAM:** 48GB
- **Storage:** Internal SAS disk array, 1x 10K RPM SAS disk drive

Target Specifications

Two of the storage targets were based on Windows Storage Server 2008 R2. One storage target was a server running the Microsoft iSCSI Target Software 3.3. The other five targets were vendor-specific proprietary designs.

DMRTK-SRVR-H

- **Operating System:** Windows Server 2008 R2 with Microsoft iSCSI Target Software 3.3

- **Processor:** Qty. 2 – Intel Xeon E5540, quad-core, 2.53GHz – Total logical processors: 16
- **RAM:** 48GB
- **Storage:** Internal SATA SSD boot drive, 6Gb SAS RAID controller with 24x 15K RPM 6Gb SAS disk drives in external disk enclosure. For some tests, the spinning disk drives were replaced with 8x SmartModular SAS SSDs and 4x Intel X25-E SATA SSDs.
- **Host Interfaces:** 2x 1Gb and 2x 10Gb

HP X1600

- **Operating System:** Windows Storage Server 2008 R2 (with iSCSI target software)
- **Processor:** Qty. 1 – Intel Xeon E5520, quad-core, 2.27GHz – Total logical processors: 8
- **RAM:** 6GB
- **Storage:** Internal RAID controller with 2x 10K RPM boot drives and 12x 7200 RPM 1TB SATA disk drives
- **Host Interfaces:** 2x 1Gb and 2x 10Gb

HP P2000

- **Operating System:** Proprietary
- **Storage:** 24x 7200 RPM 6Gb SAS 500GB disk drives
- **Host Interfaces:** 2x 10Gb

HP P4000

- **Operating System:** Proprietary, clustered unit with 2 nodes
- **Storage:** 16x 15K RPM 6G SAS, dual-port, 450GB disk drives
- **Host Interfaces:** 4x 1Gb

NetApp FAS2040-1

- **Operating System:** ONTAP 8.0.1
- **Storage:** Disk shelf with 12x 7200 RPM 500GB disk drives
- **Host Interfaces:** 4x 1Gb

NetApp FAS2040-2

- **Operating System:** ONTAP 8.0.1
- **Storage:** Disk shelf with 12x 7200 RPM 500GB disk drives
- **Host Interfaces:** 4x 1Gb

NetApp FAS3040

- **Operating System:** ONTAP 8.0.1
- **Storage:** Disk shelf with 14x 10K RPM 300GB disk drives
- **Host Interfaces:** 4x 1Gb and 2x 10Gb

Xiotech ISE (Xiotech Lab)

- **Operating System:** Windows Storage Server 2008 R2 (with iSCSI target software)
- **Processor:** Qty. 2 – Intel Xeon E5620, quad-core, 2.40GHz – Total logical processors: 16
- **RAM:** 48GB
- **Storage:** Qty. 2 – Xiotech ISE DataPac
  - DataPac #1 – 20x Seagate 10K RPM, 600GB disk drives
  - DataPac #2 – 10x Seagate 10K RPM, 600GB disk drives and 10x SSD: 200GB SAS Smart Modular XceedIOPS
- **Host Interfaces:** 10Gb

Network Infrastructure

Two switches in the Demartek lab were used for the iSCSI testing. Various Cat5e and Cat6 cables were used for the 1Gb connections. The 10Gb cabling consisted of copper SFP+ cables and fiber-optic OM3 cables.

Switches

- Dell PowerConnect 2748 – 48x 1Gb ports
- Cisco Nexus 5020 – 40x 10Gb ports

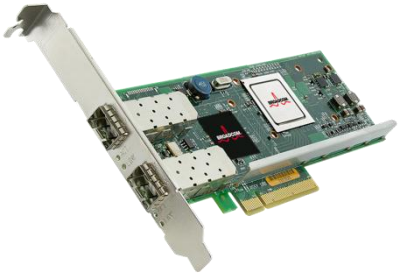
Adapters: Broadcom

Broadcom is a major producer of semiconductors for wired and wireless communications, and is one of the largest fabless semiconductor suppliers in the world. Their Infrastructure and Networking Group is responsible for many of the motherboard network interface controllers (NICs) found in servers from several major vendors. Broadcom supplies 1Gb and 10Gb network controllers and adapters.

We tested two Broadcom Ethernet server adapters for this report:

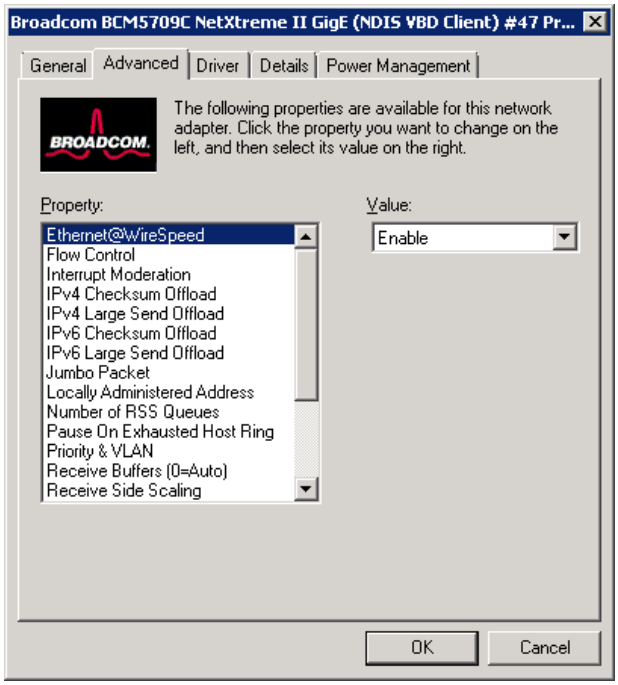
Broadcom BCM5709C NIC/TOE (1Gb)

BCM57712 NetXtreme II 10GigE (10Gb)



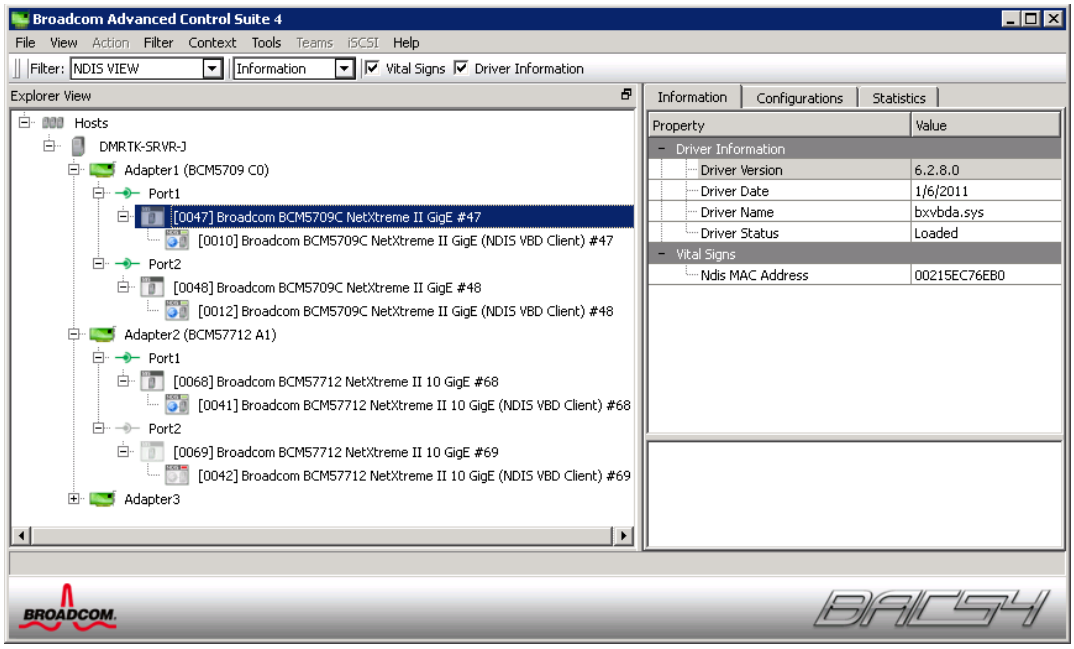
The Broadcom BCM57712 includes iSCSI offload functions and supports Fibre Channel over Ethernet (FCoE).

The basic functions of the Broadcom adapters can be managed from the properties functions of the Windows Device Manager. The Broadcom adapters we tested provided the advanced functions expected in server-class adapters, such as Large Send offload, RSS, TCP checksum offload, Teaming, etc.



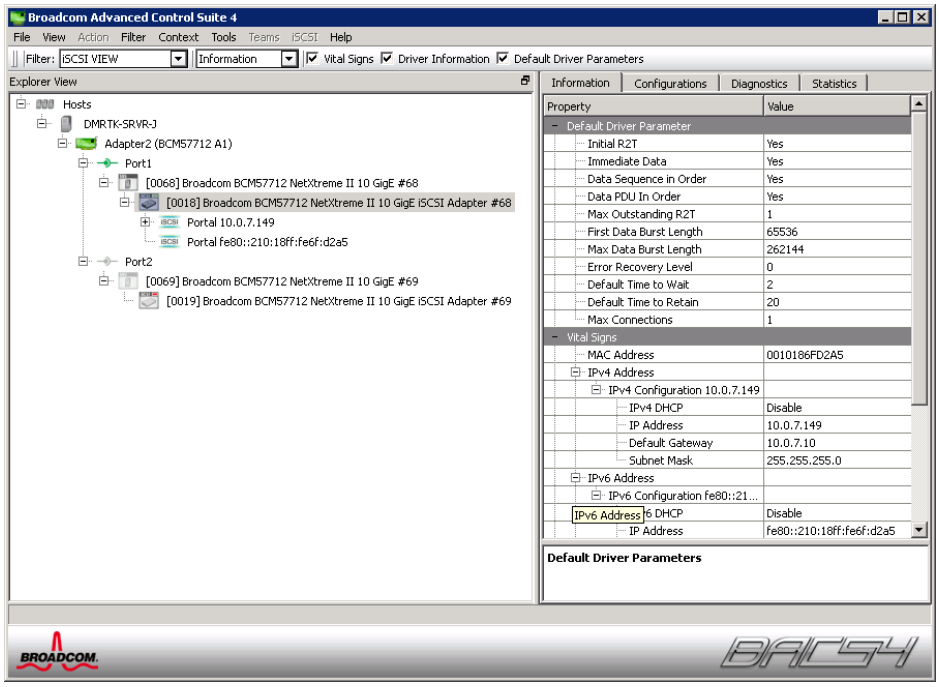
Broadcom Advanced Control Suite

Broadcom supplies the Broadcom Advanced Control Suite (BACS) to configure and manage the Broadcom adapters. BACS can perform diagnostics and other advanced functions on all the Broadcom adapters installed in a server. It can also report on non-Broadcom adapters that might be installed.

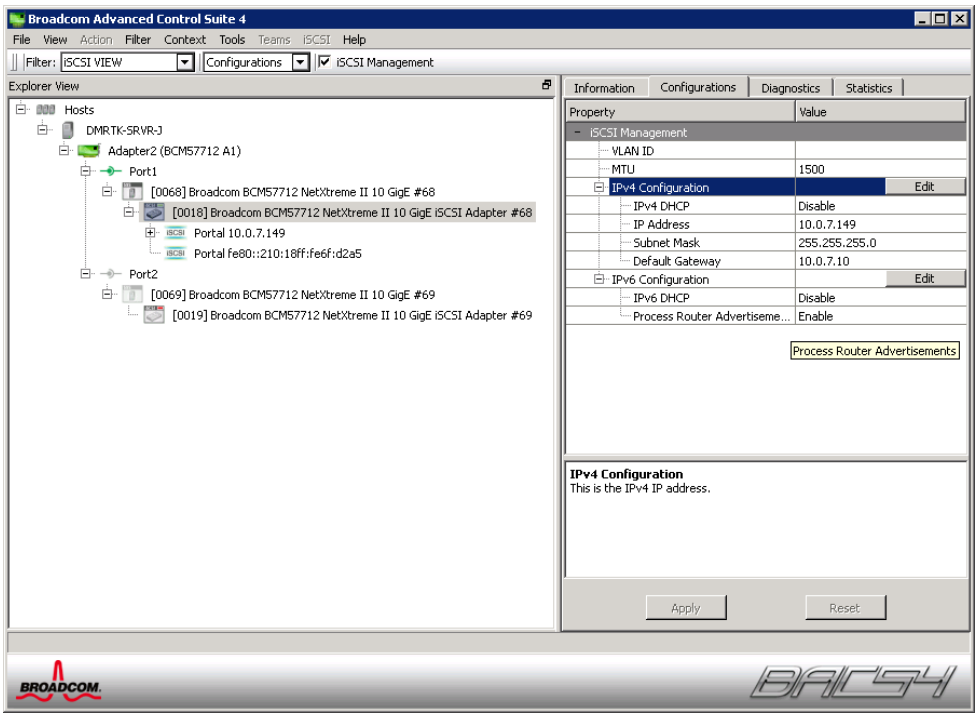


iSCSI Offload

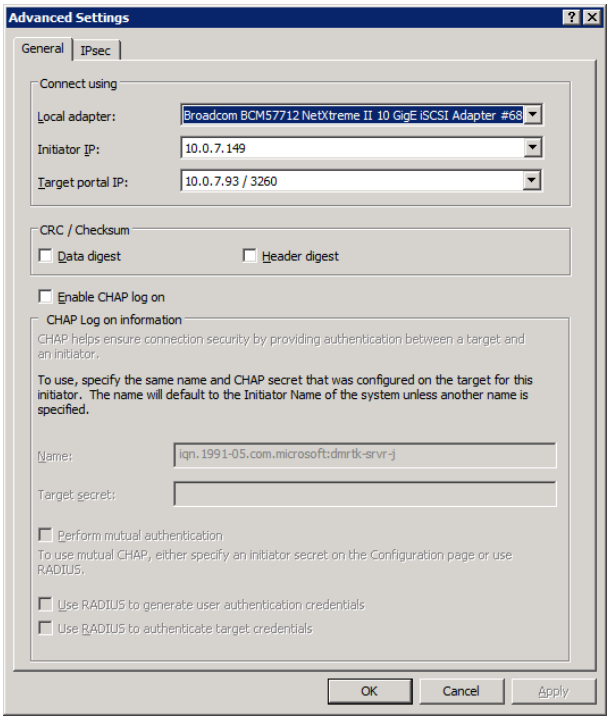
The iSCSI options are available in the "iSCSI View" filter view for the adapters that perform full iSCSI offload functions, such as the BCM57712.



The iSCSI controllers can be configured by clicking on the iSCSI Client, then the Configurations tab.



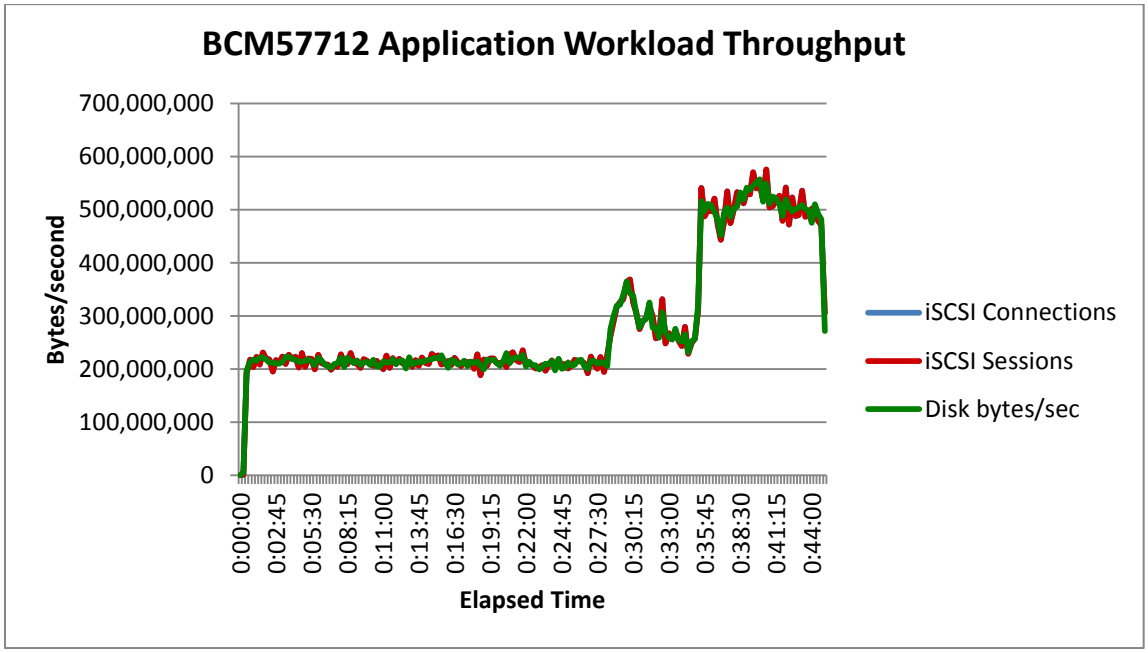
After configuring the Broadcom iSCSI offload card, then we selected it in the Microsoft iSCSI initiator using the advanced settings for the connect function.



Performance

The performance of network adapters in an iSCSI environment is dependent on several factors, including various software and hardware components in the host server, the network, and the storage target. Varying components or settings in any of these areas can affect performance, either positively or negatively.

Using a combination of Microsoft Jetstress 2010 and Windows ROBOCOPY workloads, we were able to generate the following throughput using the Broadcom BCM57712 10Gb adapter with the iSCSI offload features enabled. These workloads were run from the same physical server to two different storage targets. One target had SSDs and the other target had spinning disk drives. The throughput measurements varied as the application workloads varied.



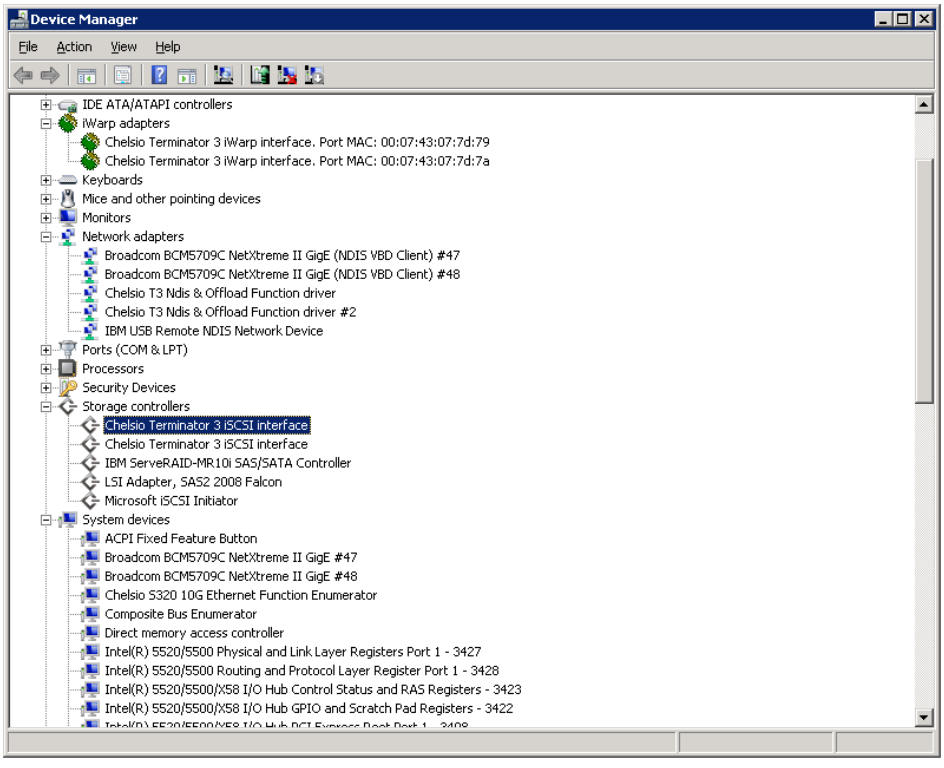
NOTE: This is not the maximum performance available for this adapter, but rather a representative sample that took advantage of the iSCSI offload features using real applications connected to real-world storage targets. Broadcom's published performance data can be found at the following link http://www.broadcom.com/collateral/pb/iSCSI_HBA-PB102-R.pdf. Broadcom also provides information on power consumption and CPU utilization at the following link: http://www.broadcom.com/docs/articles/Networks_Go_GrEEN.pdf.

Additional details about Exchange Jetstress and the initiator and target specifications used for this test are available in the Measuring iSCSI Performance section of this document.

Adapters: Chelsio

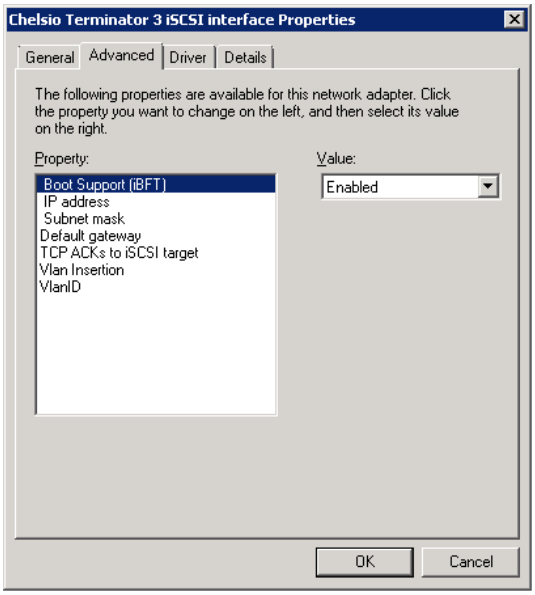
Chelsio Communications is a privately-held company that focuses on converged LAN, SAN and cluster traffic over 10Gb Ethernet. Chelsio adapters support fully offloaded TCP, iSCSI and iWARP protocols. Chelsio recently added Fibre Channel over Ethernet (FCoE) protocol support to their latest generation adapters. We tested their S320E-CR 10GbE adapter for this report.

The installation process for the Chelsio adapters is straightforward and requires little interaction. The Chelsio adapters are listed under network adapters, iWARP adapters, and storage controllers in the Windows Device Manager because of their combined NIC, iWARP and full iSCSI offload capabilities. They can be managed using the standard Windows device properties functions.



Chelsio also provided the T422-CR low-profile adapter that provides 2x10GbE and 2x1GbE ports. The iSCSI drivers for this adapter will be available soon.

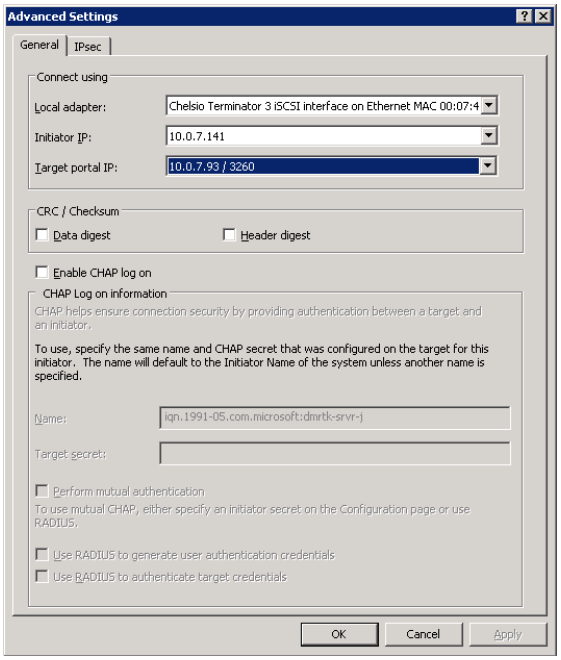
Their iSCSI features can be managed from the standard properties functions for this type of device.



Chelsio Hardware Initiator

After the Chelsio iSCSI initiator driver package is installed, then the Chelsio iSCSI offload hardware initiator and the Microsoft iSCSI software initiator appear as choices in the standard Windows iSCSI initiator management software. To take advantage of the Chelsio hardware features, we have selected the Chelsio iSCSI hardware initiator.

The other iSCSI login and configuration features are managed using the iSCSI initiator interface in the normal manner.



Performance

The performance of network adapters in an iSCSI environment is dependent on several factors, including various software and hardware components in the host server, the network, and the storage target. Varying components or settings in any of these areas can affect performance, either positively or negatively.

Using Microsoft Jetstress 2010 as a workload generator, we were able to generate the following throughput using the Chelsio S320E-CR 10Gb adapter with the iSCSI offload features enabled.

The Jetstress test parameters for this test were:

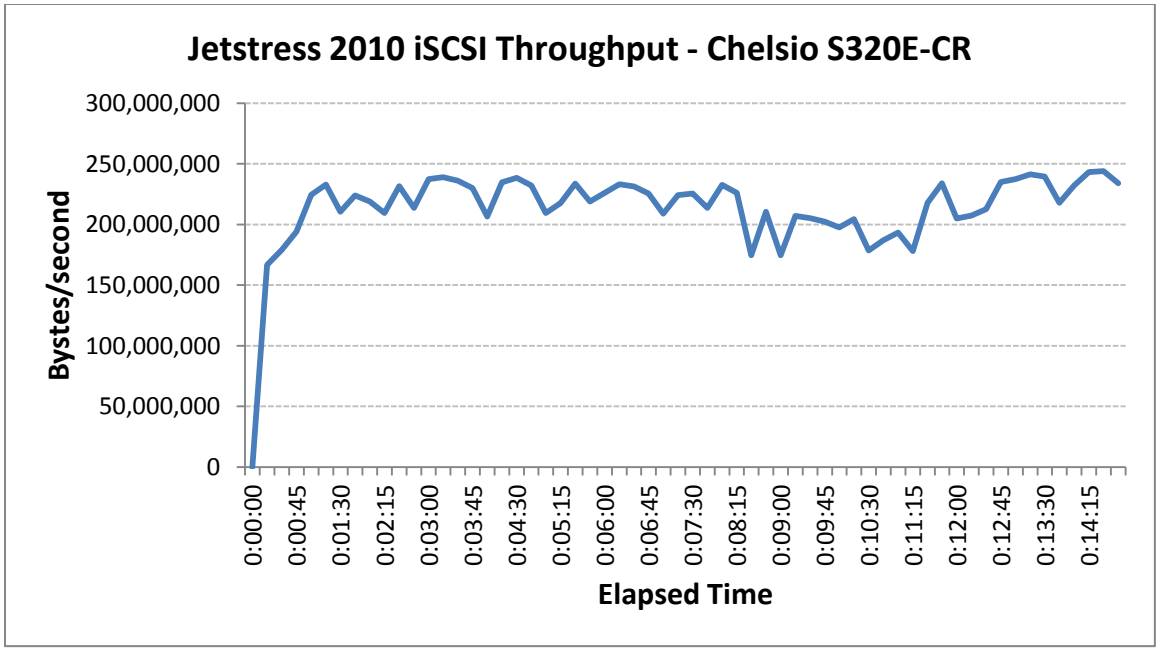
- Mailboxes=300
- Mailbox size=1000MB
- Storage Groups=3
- IOPS=0.36
- Threads=8

CPU utilization during this test, as reported by Jetstress was:

- Average: 9.063%
- Minimum: 6.645%
- Maximum: 11.286%

Achieved IOPS (32KB) were: 6200.476

We were able to sustain an average of greater than 200MB/sec for this test with the Chelsio adapter. This is not the maximum performance available for this adapter, but a representative sample that took advantage of the iSCSI offload features.

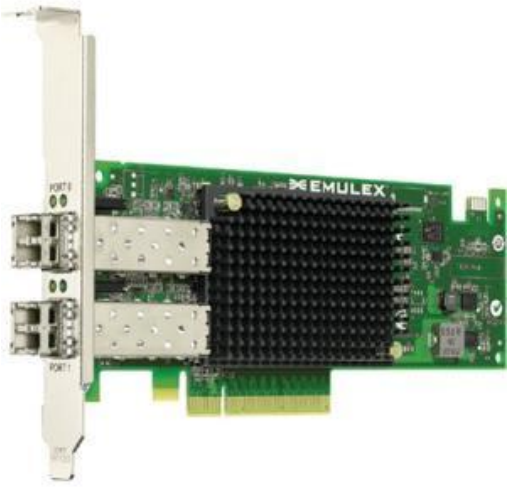


Additional details about Exchange Jetstress and the initiator and target specifications used for this test are available in the Measuring iSCSI Performance section of this document.

Adapters: Emulex

Emulex designs and builds a full line of storage adapters, network interface cards, controller chips, embedded bridges, and connectivity management software. Emulex is probably best known for their Fibre Channel host bus adapters (HBAs). In the last couple of years, Emulex has branched out into the Ethernet adapter market by providing 10Gb Ethernet network interface cards, full 10Gb iSCSI offload adapters, and 10Gb Fibre Channel over Ethernet (FCoE) converged network adapters (CNAs).

We tested their latest generation of 10Gb OneConnect™ Ethernet network interface cards and 10Gb iSCSI offload adapters for this report. These are the OCe11102-N and OCe11102-I adapters, which are 10Gb dual port adapters that include stateless TCP/IP offload and TCP chimney offload functions. The iSCSI model ("-I") also performs a full iSCSI offload. The NIC ("-N") model can be upgraded to the iSCSI model or the FCoE ("-F") CNA.



Both models support PCI express 2.0, and have either copper (10GBase-CR) or fiber-optic connectors (10GBase-SR). They also support iSCSI over DCB for optimized iSCSI performance.

Emulex has a history of providing comprehensive operating system support for all of its adapters and its 10Gb Ethernet adapters are no different. They provide support for Windows Server (including Hyper-V), VMware ESX, Red Hat Enterprise Linux Server, Novell SUSE Linux Enterprise Server, and Oracle Solaris.

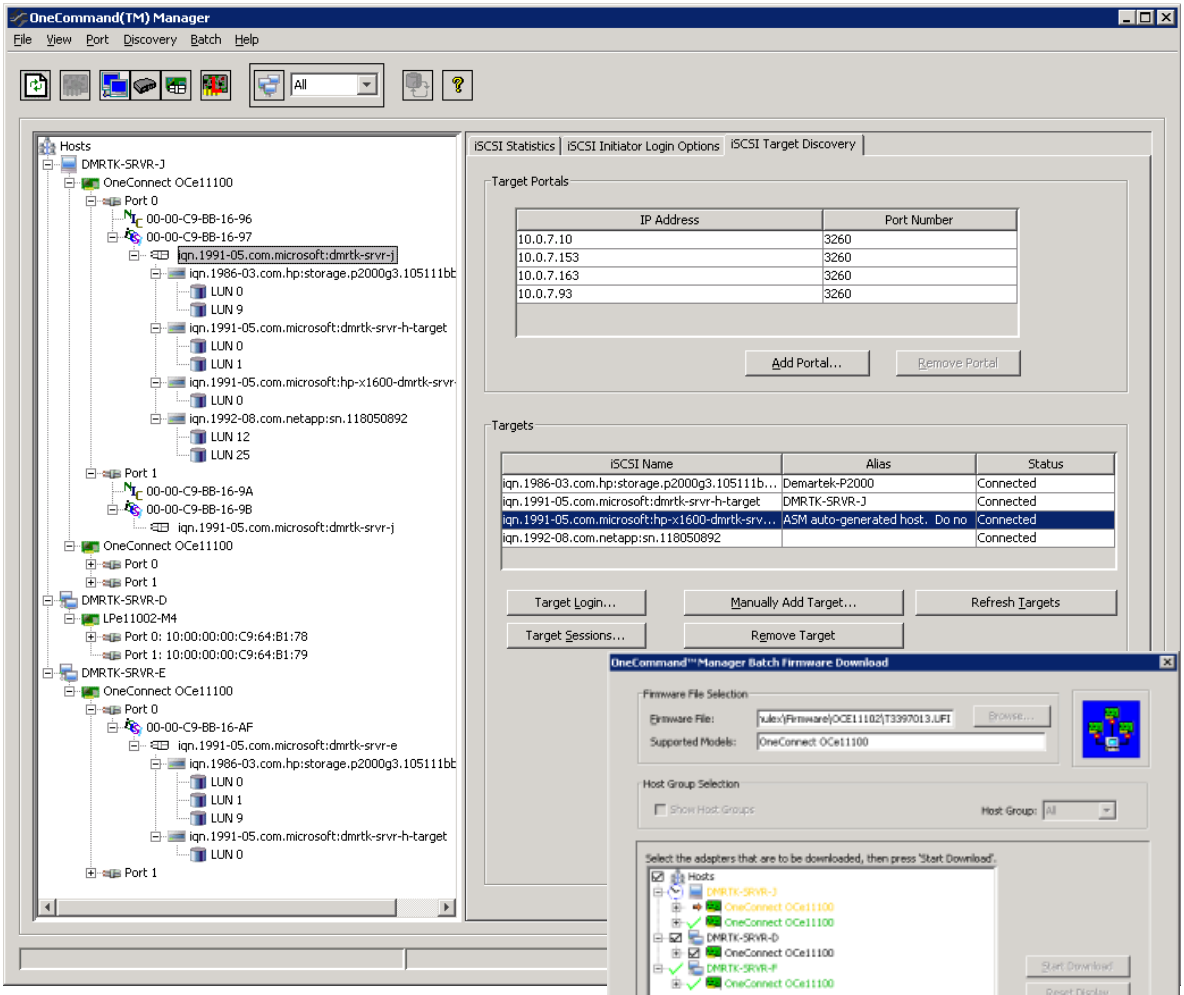
Emulex provides the OneCommand™ Manager management application for centralized management of its Fibre Channel and10Gb Ethernet adapters. OneCommand Manager offers a host of automation capabilities and diagnostic tools to improve efficiency and streamline administrative tasks enabling IT administrators to manage multiple Emulex adapters in multiple servers at the same time. OneCommand Manager can be installed in one of three management modes: local host management only, local host management plus remote access from other hosts, and full management of adapters on the local host and other hosts that allow it.

For VMware environments, Emulex also offers OneCommand Manager for VMware Plug-in that enables comprehensive control of Emulex Fibre Channel HBAs and 10Gb Ethernet adapters directly from the VMware vCenter management console.

Installation of the Emulex software is simple and straightforward. The Emulex installation utility identifies the Emulex adapters and automatically installs the proper operating system drivers for each adapter installed in that host server.

Topology View

The Emulex OneCommand Manager topology view provides a list of all the Emulex adapters installed within all visible hosts, including Ethernet adapters and Fibre Channel adapters. The iSCSI offload adapter management functions include the normal steps needed by an iSCSI initiator to configure the IP address of the initiator, set the CHAP shared secret, and perform target discovery and login. OneCommand Manager can also fully participate with an iSNS server.



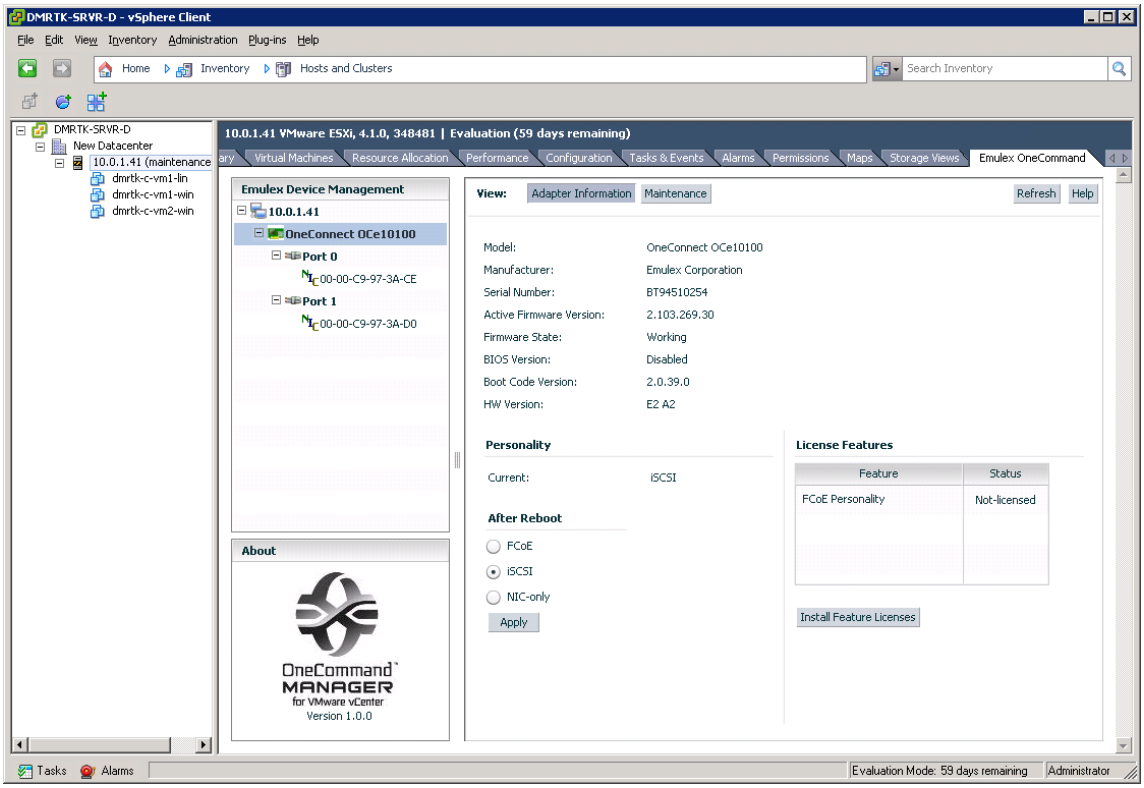
Firmware Updates

OneCommand Manager allows for both firmware and driver parameter updates. Firmware updates can occur individually for each adapter or in a "batch mode" fashion, which automatically updates all similar adapter models in any or all hosts visible to OneCommand Manager. The firmware is selected and OneCommand Manager provides a list of discovered adapters within the fabric which can use the selected firmware.

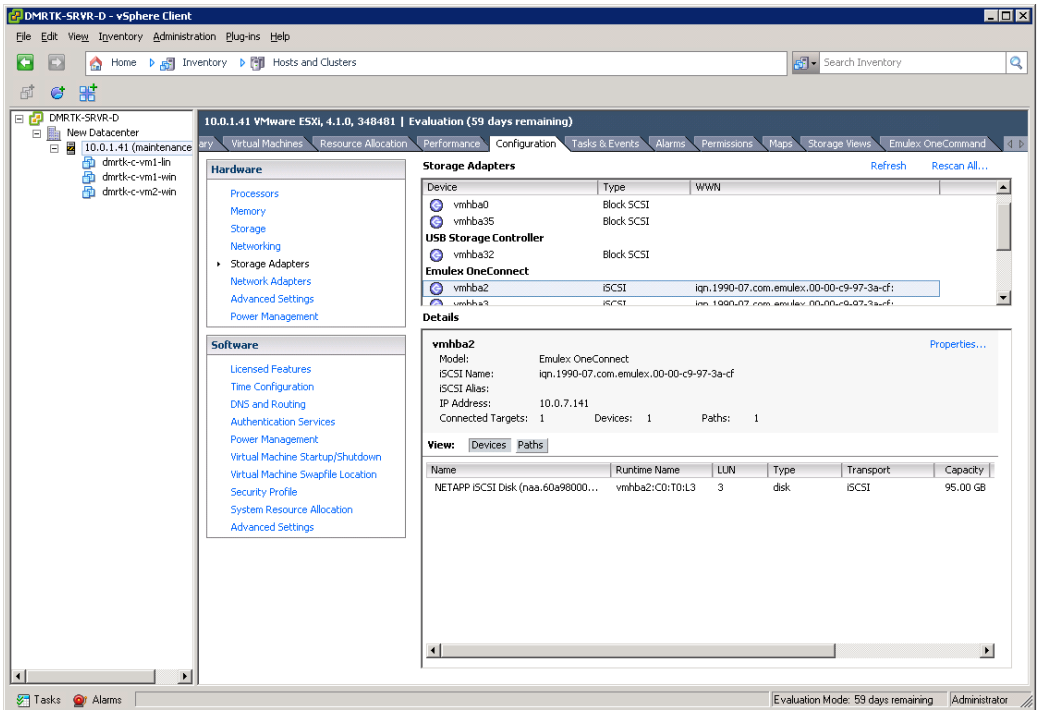
In the example above, we show updating firmware in four OCe11100 adapters in three different servers using a "batch mode" process that was initiated from one of those servers.

VMware vSphere Support

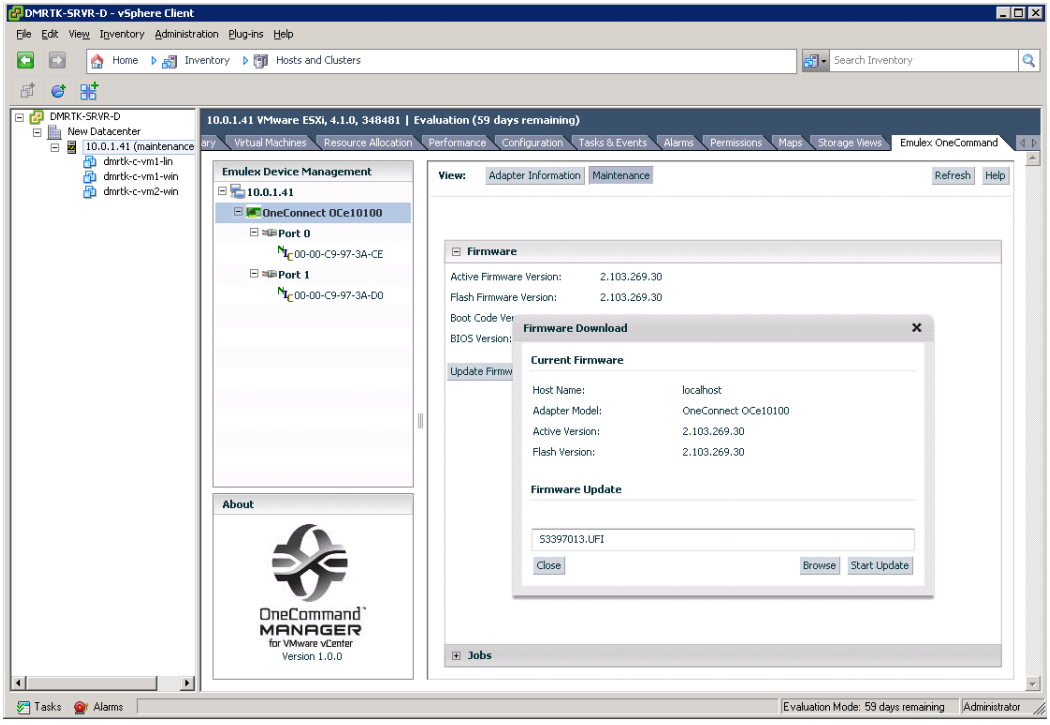
Emulex also provides the OneCommand Manager for VMware Plug-in that allows an Emulex adapter in a vSphere server to be managed from another server. In this example, we installed the vSphere client in a separate Windows server (dmrtk-srvr-d) and installed the VMware Plug-in into the vSphere 4.1 server (dmrtk-srvr-c).



The Emulex iSCSI offload adapter appears in the configuration tab of the vSphere client as a storage adapter, along with other storage adapters.



The Emulex VMware Plug-in also allows firmware to be updated in the Emulex adapter installed in the vSphere server.



Performance

The performance of network adapters in an iSCSI environment is dependent on several factors, including various software and hardware components in the host server, the network, and the storage target. Varying components or settings in any of these areas can affect performance, either positively or negatively.

Using Microsoft Jetstress 2010 as a workload generator, we were able to generate the following throughput using the Emulex OCe11102-N and OCe11102-I (iSCSI offload) 10Gb adapters.

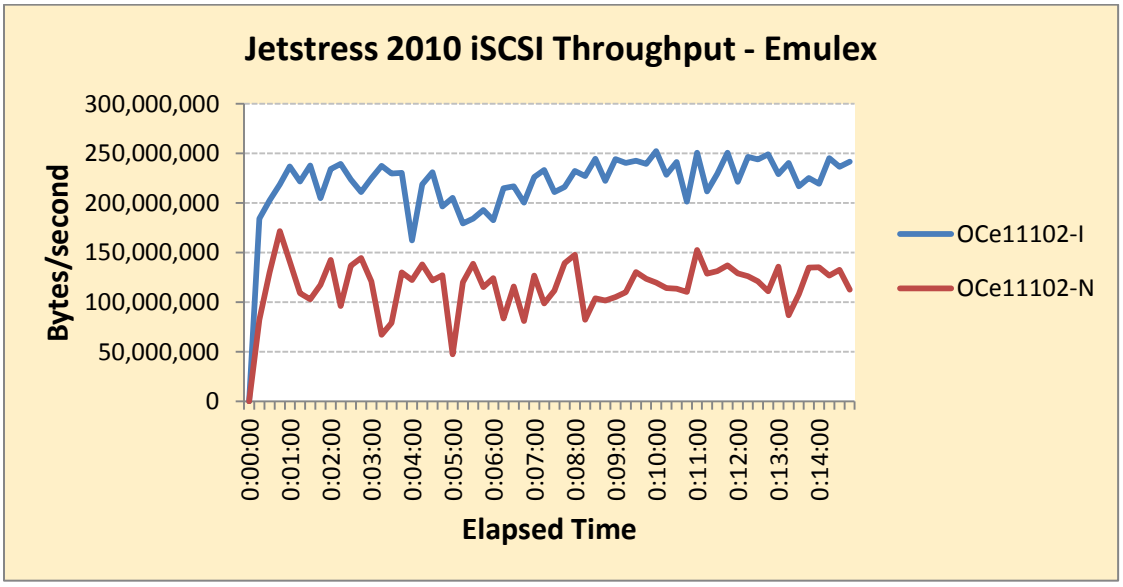
The Jetstress test parameters for this test were:

- Mailboxes=400
- Mailbox size=1000MB
- Storage Groups=4
- IOPS=0.30
- Threads=8

CPU Utilization	OCe11102-N	OCe11102-I
Average	6.525%	9.241%
Minimum	3.861%	7.048%
Maximum	10.332%	11.007%

IOPS (32KB)	OCe11102-N	OCe11102-I
Achieved IOPS (32KB)	3333.734	6239.027
IOPS/Avg. %CPU Util.	510.92	675.16

We were able to sustain an average of more than twice the throughput with the Emulex iSCSI offload adapter when compared to the Emulex 10Gb NIC and the iSCSI software initiator for Windows. This is not the maximum performance available for these adapters, but is a representative sample that shows the performance improvement using iSCSI offload.



Additional details about Exchange Jetstress and the initiator and target specifications used for this test are available in the Measuring iSCSI Performance section of this document.

Adapters: Intel

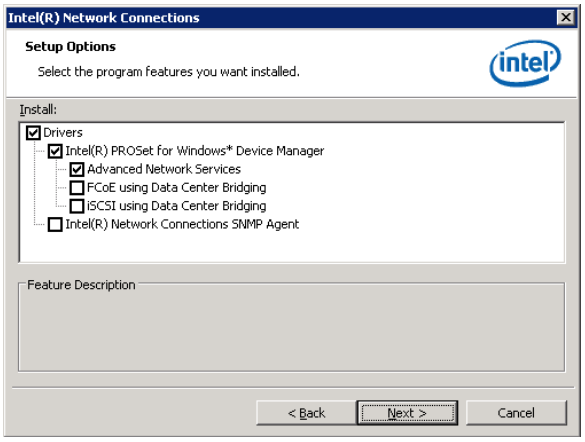
Intel® is widely known for its microprocessors, but has also had a robust Ethernet adapter and controller product line for many years. The 1Gb and 10Gb Intel Ethernet adapters are available in desktop and server versions. The server adapters support the native operating system iSCSI initiators.

We tested two Intel Ethernet server adapters for this report, and these support PCIe 2.0:

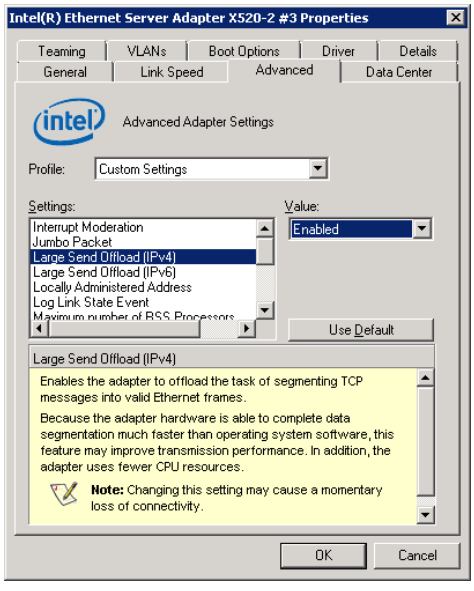
Intel Gigabit ET Server Adapter (1Gb)

Intel Ethernet X520 Server Adapter (10Gb)

The installation process for the Intel adapters is simple, straightforward and follows typical procedures. Intel includes a "Connections" CD that has drivers for all their Ethernet adapters. This driver installation software detects the adapter or adapters that are installed and provides a list of appropriate options. In addition, these drivers can be downloaded from the Intel website. No firmware is needed for the Intel adapters.



The Intel Ethernet adapters support a wide variety of operating systems and virtualization environments. The Intel Ethernet X520 Server Adapter (10GbE) also supports Fibre Channel over Ethernet (FCoE) and Data Center Bridging (DCB).



Management

Intel network adapters are managed using the standard operating system management tools and no additional management software is required. In Device Manager in Microsoft Windows, there are several tabs that provide access to various adapter settings. For most installations, the default options work well.

The Intel server adapters (including both adapters we tested) support all the advanced features such as Large Send offload, RSS, TCP checksum offload, Teaming,

etc., for both IPv4 and IPv6

The Intel adapters work with the Microsoft iSCSI initiator and are considered a software initiator. All the usual iSCSI target discovery and login features are supported and functioned as we expected.

Performance

The performance of network adapters in an iSCSI environment is dependent on several factors, including various software and hardware components in the host server, the network, and the storage target. Varying components or settings in any of these areas can affect performance, either positively or negatively.

Using Microsoft Jetstress 2010 as a workload generator, we were able to generate the following throughput using the Intel X520 10Gb adapter.

The Jetstress test parameters for this test were:

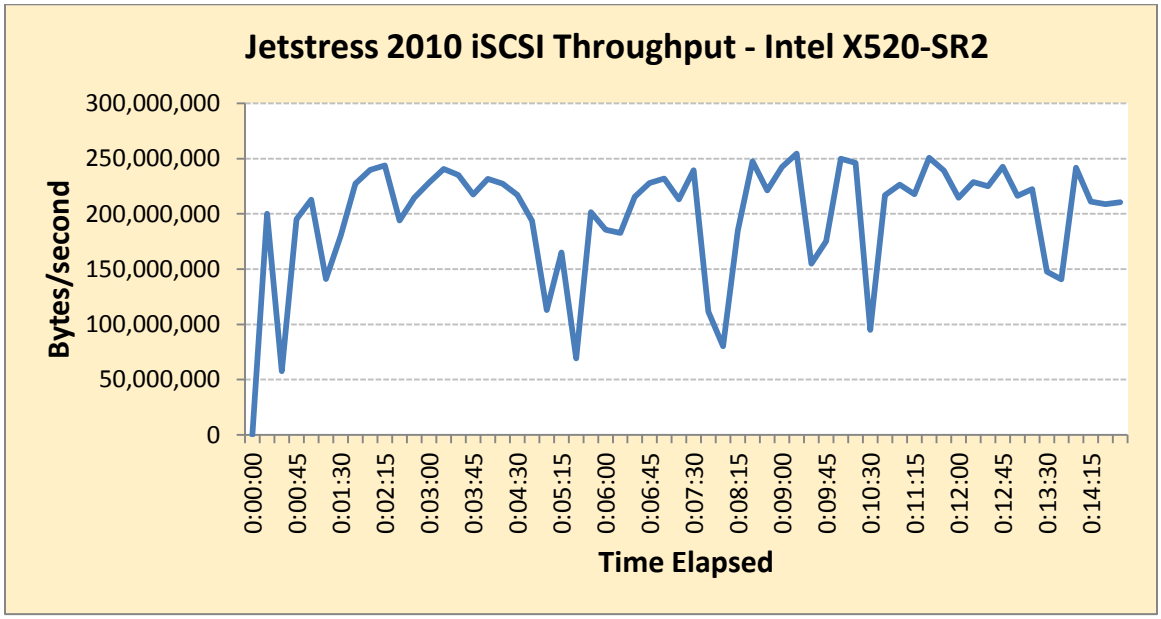
- Mailboxes=300
- Mailbox size=1000MB
- Storage Groups=3
- IOPS=0.24
- Threads=8

CPU utilization during this test, as reported by Jetstress was:

- Average: 7.663%
- Minimum: 3.118%
- Maximum: 9.962%

Achieved IOPS (32KB) were: 5754.108

We were able to sustain an average of approximately 200MB/sec for this test with the Intel X520 adapter. This is not the maximum performance available for this adapter, but a representative sample.



Additional details about Exchange Jetstress and the initiator and target specifications used for this test are available in the Measuring iSCSI Performance section of this document.

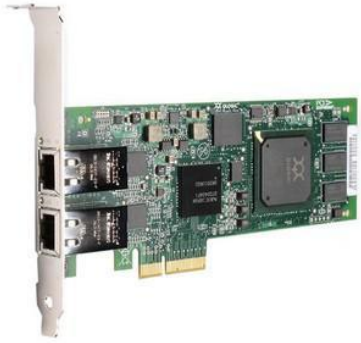
Additional performance data for the Intel Ethernet Server Adapter X520 is available in our evaluation report for this adapter on our website. For this evaluation, we ran several performance tests with storage targets with many spinning hard disk drives.

Adapters: QLogic

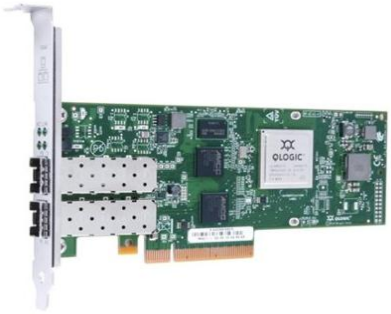
QLogic is a major provider of data, server, and storage networking infrastructure solutions. They are probably best known for their Fibre Channel HBAs and switches, but they also provide 10GbE adapters, Converged Network Adapters (CNAs), Infiniband products and ASICs. QLogic partners with many of the major server and storage vendors, providing them with adapters, switches, routers and ASICs.

We tested two QLogic Ethernet server adapters for this report

QLogic QLE4062-C (1Gb)



QLogic QLE8242 (10Gb)



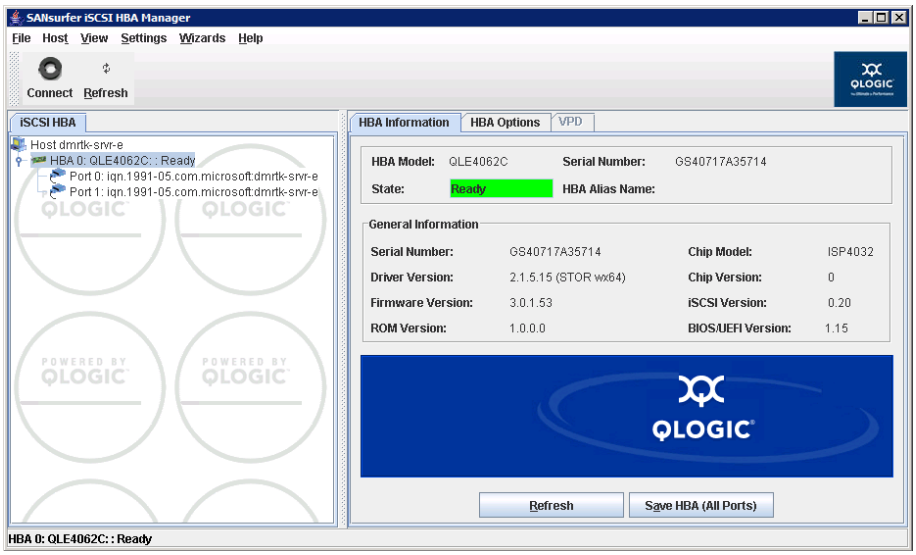
The QLogic QLE8242 includes iSCSI offload functions and supports Fibre Channel over Ethernet (FCoE).

QLogic provides two separate management software applications to manage these devices:

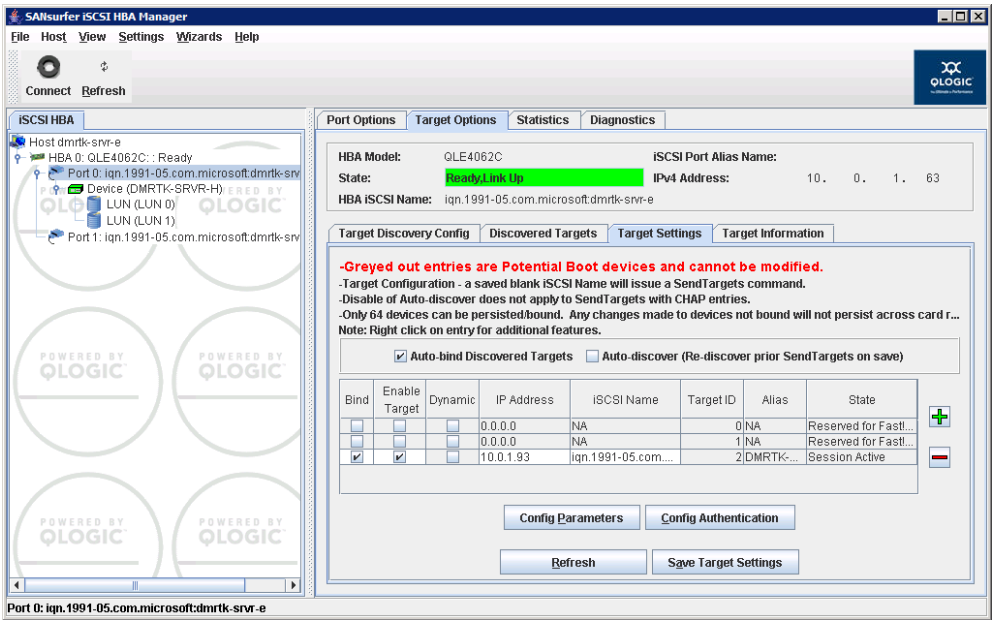
- SANsurfer
- QConvergeConsole (QCC)

SANsurfer

The QLogic QLE4062-C is an iSCSI HBA that has been on the market for a few years. It uses SANsurfer for its configuration and management.



The IP addresses for the ports can be set on the Network tab within the Port Options tab. After the IP addresses have been added, the iSCSI management features are accessed from the Target Options tab. The target options include target discovery, target login and other target-related options.



After the target has been discovered and the target login is complete, the LUNs (volumes) from that target become available to the operation system. These are accessed in the normal manner.

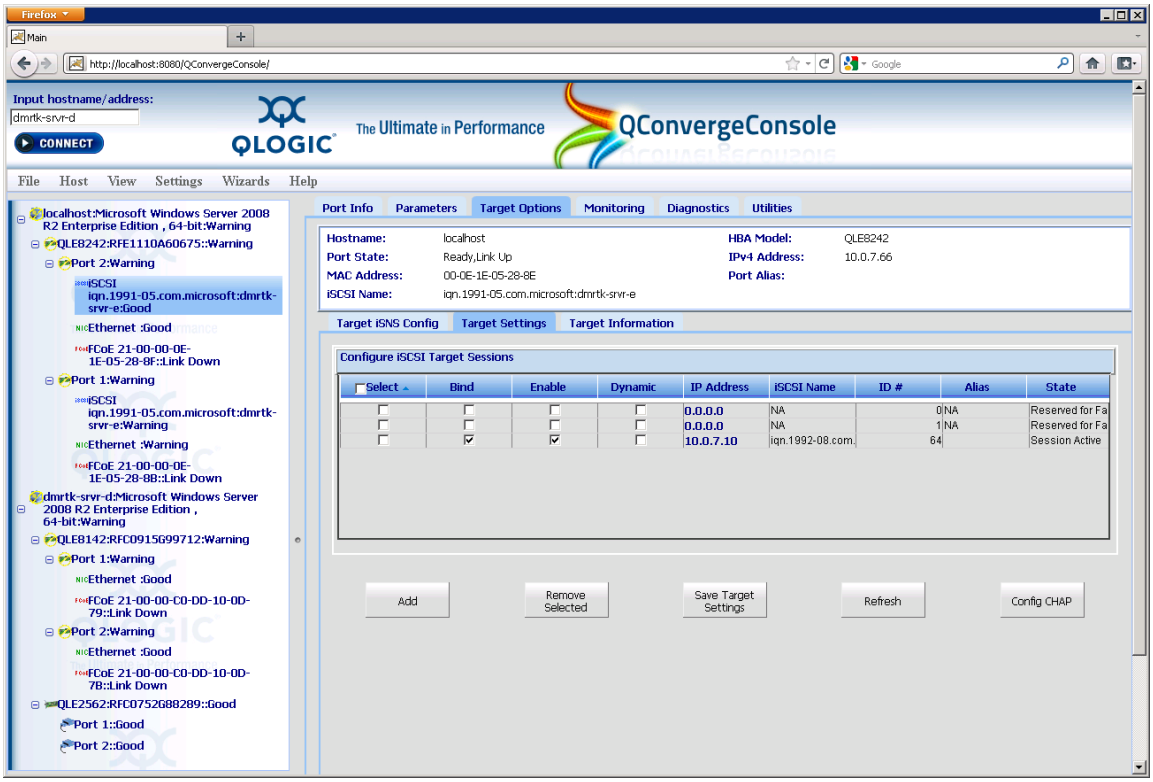
QConvergeConsole (QCC)

The QConvergeConsole (QCC) is the replacement for SANsurfer and is the management application for the newer QLogic Fibre Channel HBAs and 10Gb NICs and CNAs. These include:

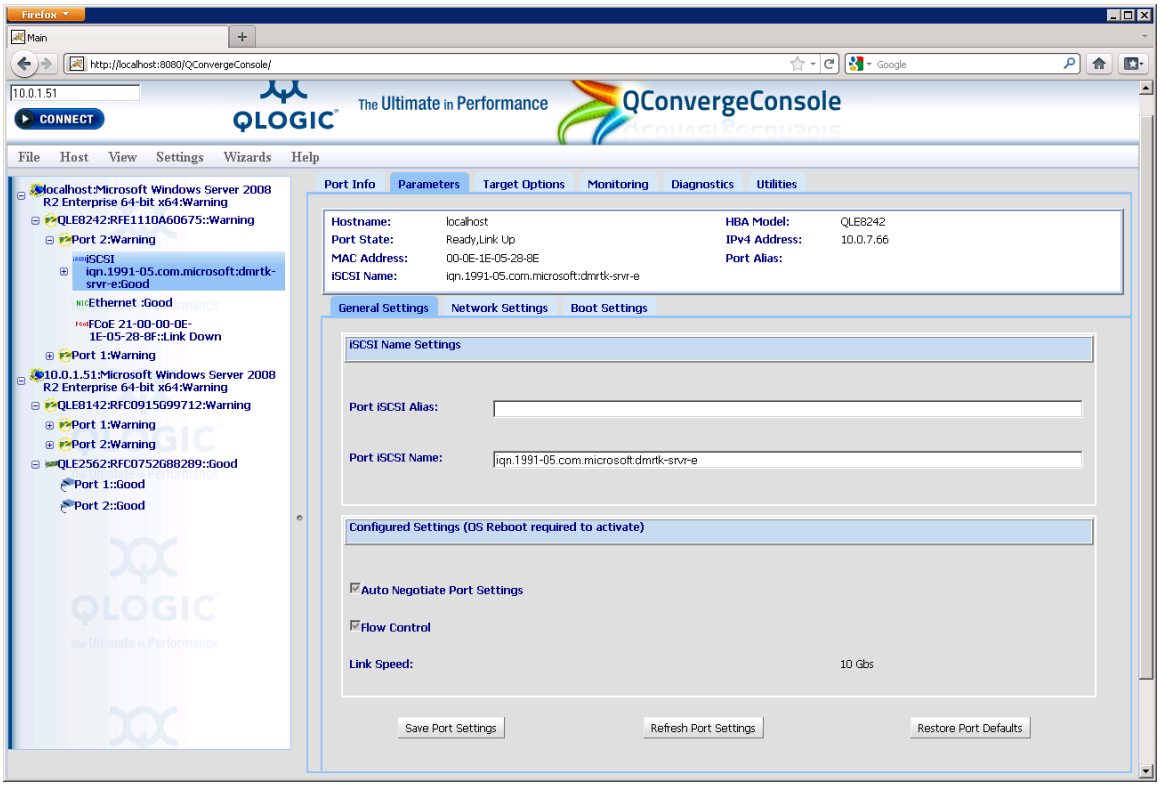
- QLE24xx and QLE25xx Fibre Channel HBAs (4Gb and 8Gb)
- QLE32xx series Intelligent Ethernet Adapters (10Gb)
- QLE81xx and QLE82xx Converged Network Adapters (CNAs)

The installation of QCC also requires the installation of the Apache Tomcat application server, which is included with the QCC package and installed if required. A web browser is the interface into QCC. QLogic also provides a command-line interface for QCC.

QCC will be familiar to anybody who has used SANsurfer in the past. The screens within QCC are organized similarly to SANsurfer, but the response is much faster than SANsurfer. QCC can manage QLogic adapters that are installed in other host servers, as shown on the left side of the screenshot below.



The iSCSI management features are somewhat similar to those previously described for SANsurfer.



As with SANsurfer, after configuring the IP address under the Network Settings tab then the target discovery and target login options, the LUNs from the target are available to the operating system in the usual manner.

Performance

The performance of network adapters in an iSCSI environment is dependent on several factors, including various software and hardware components in the host server, the network, and the storage target. Varying components or settings in any of these areas can affect performance, either positively or negatively.

Using Microsoft Jetstress 2010 as a workload generator, we were able to generate the following throughput using the QLogic QLE8242 10Gb adapter.

The Jetstress test parameters for this test were:

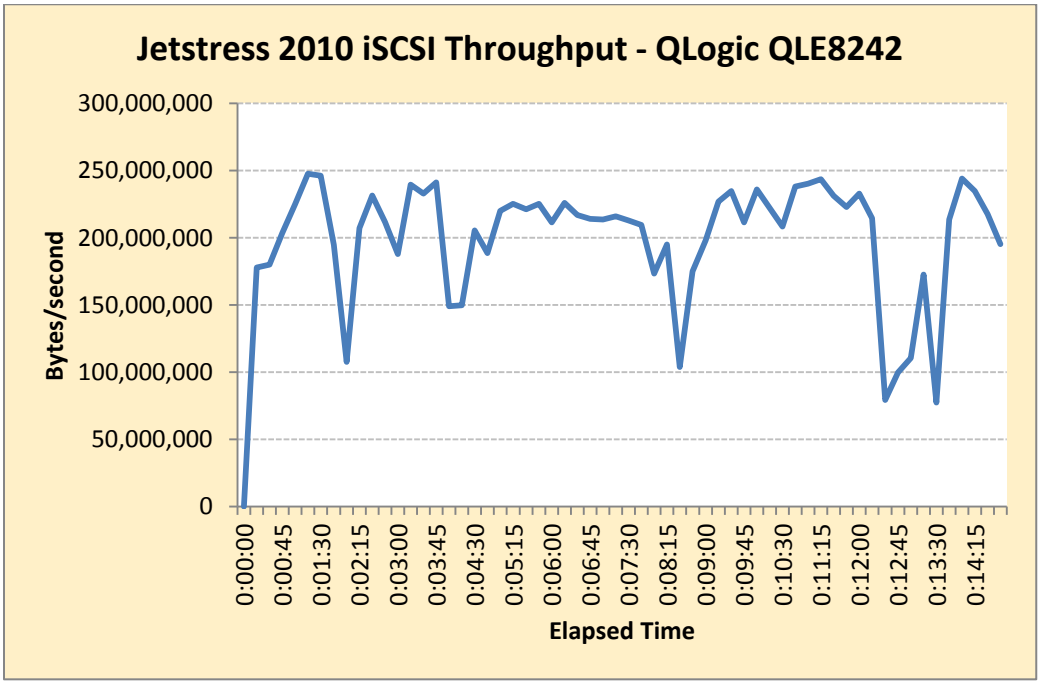
- Mailboxes=300
- Mailbox size=1000MB
- Storage Groups=3
- IOPS=0.30
- Threads=8

CPU utilization during this test, as reported by Jetstress, was:

- Average: 7.851%
- Minimum: 3.582%
- Maximum: 10.760%

Achieved IOPS (32KB) were: 5775.26

We were able to sustain an average of approximately 200MB/sec for this test with the QLogic adapter. This is not the maximum performance available for this adapter, but a representative sample.



Additional details about Exchange Jetstress and the initiator and target specifications used for this test are available in the Measuring iSCSI Performance section of this document.

Storage Targets: Hewlett-Packard (HP)

HP has a variety of iSCSI storage solutions designed to address different types and sizes of enterprises. We tested three different HP iSCSI storage solutions for this deployment guide.

- HP X1600 G2 Network Storage System
- HP P2000 G3 10GbE iSCSI MSA System
- HP P4000 G2 SAN Solutions

Some of these storage solutions, such as the HP P4000 series, are exclusively iSCSI, with options for 1Gb or 10Gb Ethernet interfaces. The other two product families are available with choices for host interfaces including 1Gb iSCSI, 10Gb Ethernet and 8Gb Fibre Channel.

Performance

Microsoft provides Exchange Jetstress as a tool to stress test storage systems to determine if a particular configuration of a storage system is capable of adequately supporting an Exchange Server environment. Microsoft provides the Exchange Solution Reviewed Program (ESRP), and HP has submitted several performance tests using the storage systems listed here. These performance results are available at http://technet.microsoft.com/en-us/exchange/ff182054.aspx.

HP X1600 G2 Network Storage System

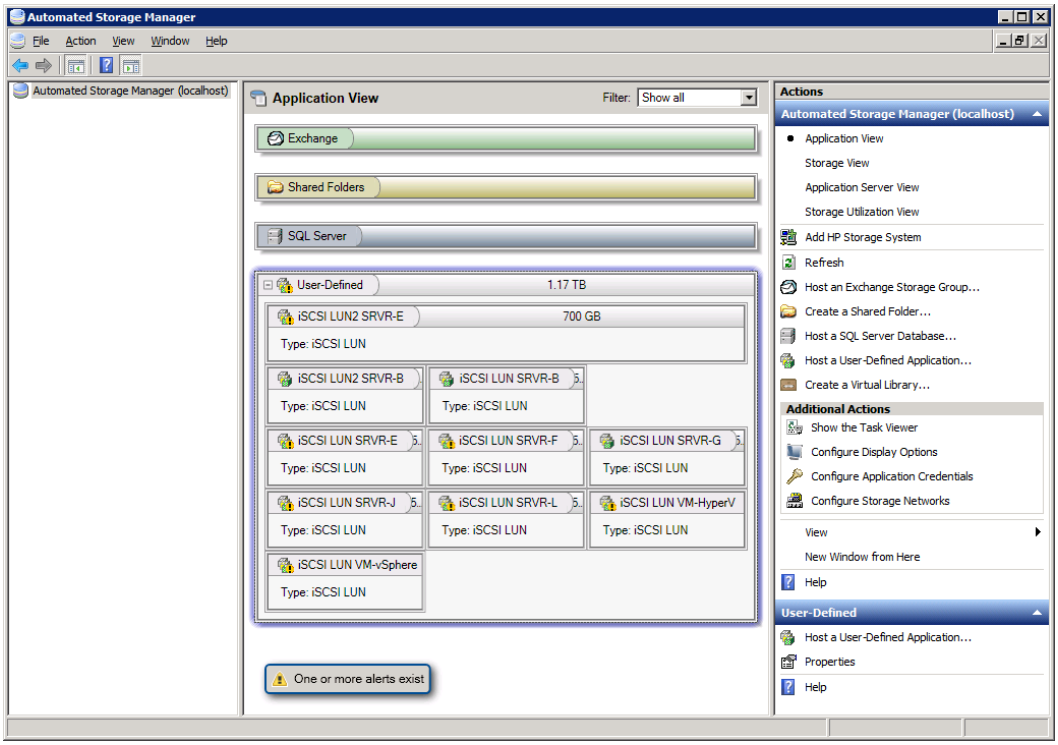
The HP X1600 G2 is a member of the HP X1000 G2 Network Storage System product family that operates as a file server or an iSCSI target, or both. The HP X1600 G2 is a NAS solution running Windows Storage Server 2008 R2 with the Microsoft iSCSI Software Target and the HP Automated Storage Manager (ASM) software. It is built on an HP ProLiant DL 180 G6 server chassis and has options for up to 12 large form factor (3.5-inch) or up to 25 small form factor (2.5-inch) disk drives in the front slots.

X1000 G2 NAS is currently available in a number of different preconfigured internal capacity configurations ranging from 4 TB to 24 TB. The X1000 G2 configuration can add additional NAS capacity by using external disk enclosures (such as HP's D2000). Additional information on the full X1000 G2 product family can be found at www.hp.com/go/x1000.

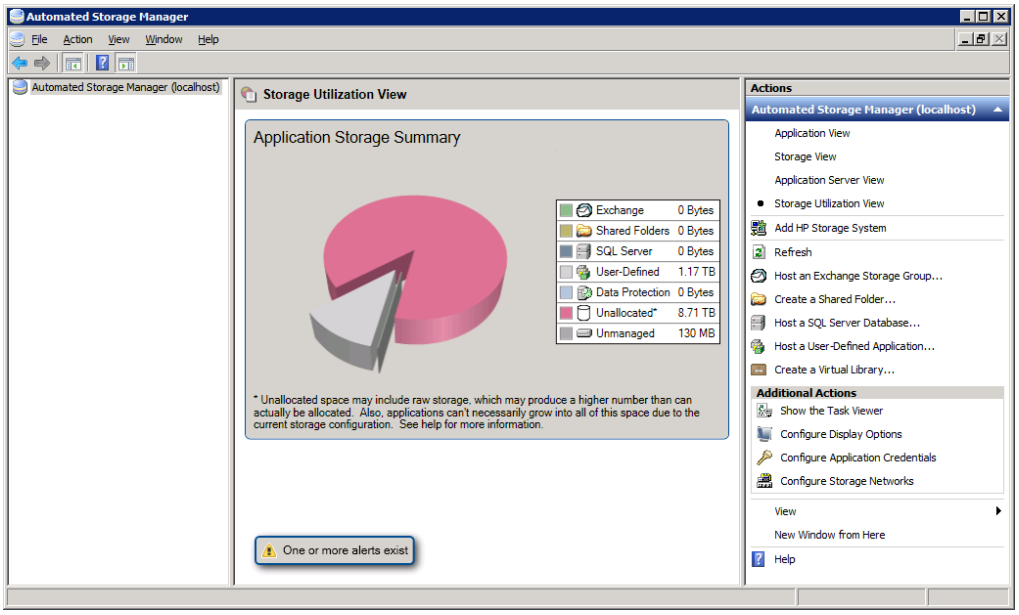
Administrators have two options for managing storage on the X1600. Because it is a Windows-based platform, standard Windows management tools can be used to configure its storage including file shares and iSCSI target devices. HP has also provided the Automated Storage Manager (ASM) application that can be used to manage and provision storage resources. ASM provides wizards that assist with the process of allocating and configuring storage for host application data and shared folders. These wizards can help with storage for Exchange Server, SQL Server, user-defined applications, shared folders, replication, snapshots, and more. Storage that has been allocated can be increased using the wizards as data growth occurs.

We configured several iSCSI LUNs that were assigned to several servers in the lab. ASM streamlines the process of creating or growing iSCSI LUNs using a wizard-based process. The wizard steps the administrator through one set of parameters that are used to create the iSCSI targets and the

LUNs associated with them. This process requires only a few mouse clicks to complete. This process can also be accomplished using the standard Microsoft iSCSI target management functions, but requires more steps to complete.



ASM also provides storage management functions, such as the storage utilization view, that helps administrators monitor their storage capacity utilization.



HP P2000 G3 iSCSI MSA Array

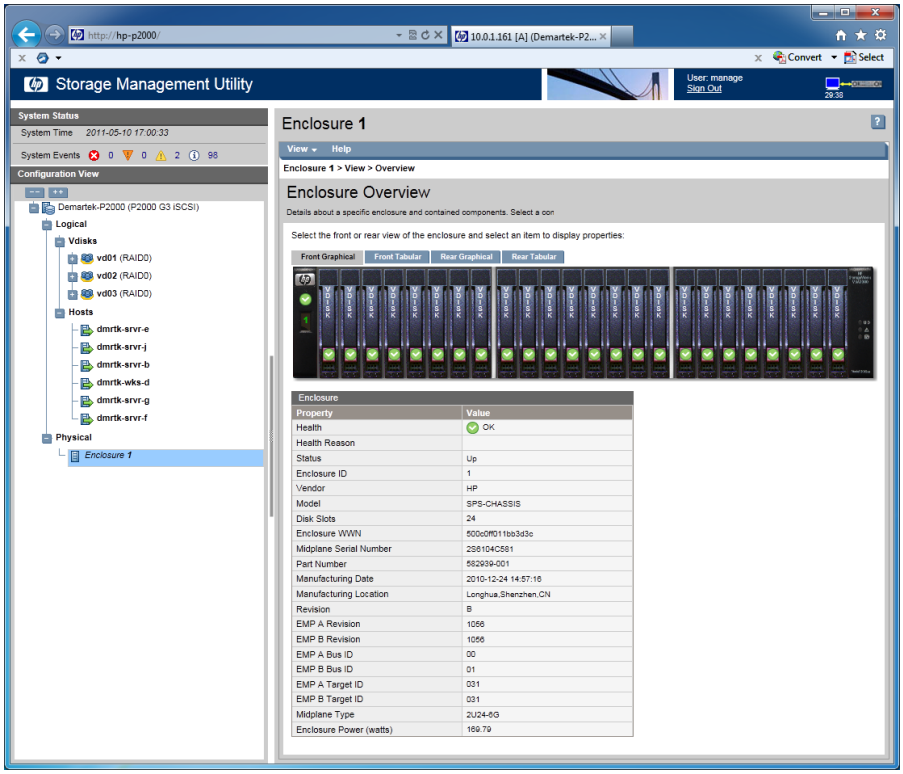
The HP P2000 is a dedicated storage system designed to meet the needs of small and medium business by providing performance with flexibility. Options are available in several areas:

- Host interfaces: It has an assortment of host interfaces available including 8Gb Fibre Channel, a Combo 8Gb FC/1GbE iSCSI, 6Gb SAS, and 1GbE or 10GbE iSCSI.
- Disk drives: Can accommodate up to 149 2.5-inch disk drives or 96 3.5-inch disk drives that can have SAS or SATA interfaces.
- Controllers: Available in single controller and dual controller models.

Snapshots and volume copy functions are included as standard features with the HP P2000 series arrays.

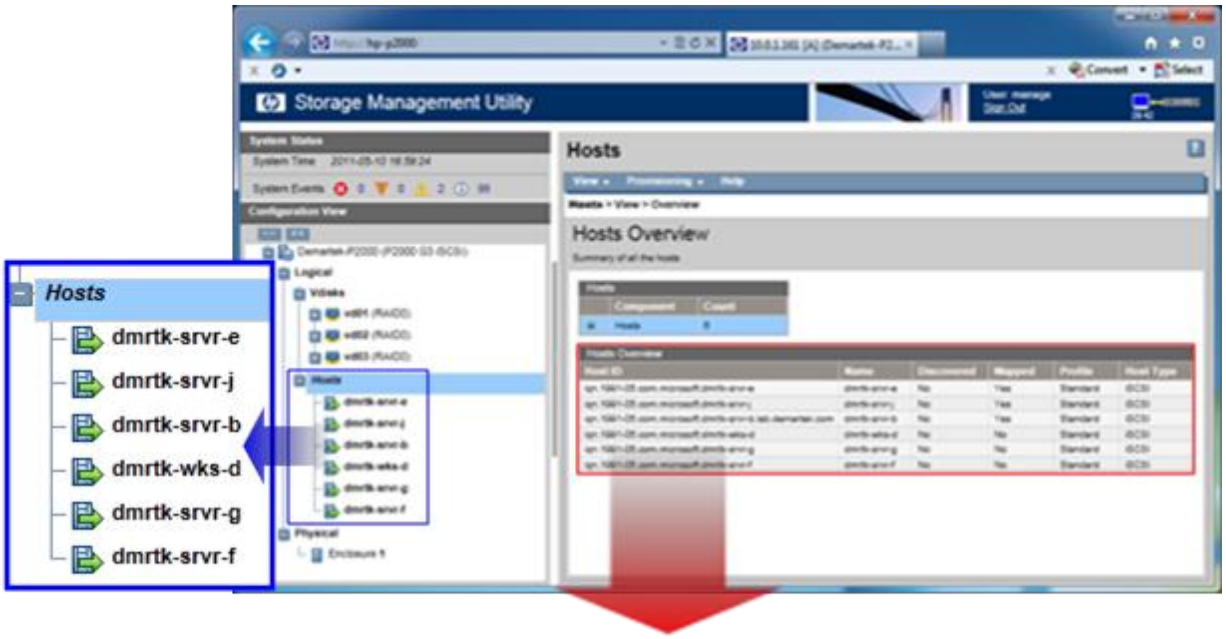
P2000 G3 MSA is currently available with capacities from as few as 2 drives up to a full capacity of 192 TB with today's latest drives. Check out the different connectivity and capacity options at: www.hp.com/go/p2000.

The HP P2000 management console uses an embedded web browser based storage management utility that includes photographs of the front and back of the hardware to assist in selecting the right components such as disk drives for storage groups, host interfaces, etc.



Storage is provisioned into "Vdisks" that are logical volumes that will ultimately be presented to the host initiators. A handy feature of the HP P2000 is the ability to create several volumes of the same size with a single provisioning step. For our tests we created three disk groups with eight drives in each disk group. Then in a single step we were able to provision 20 equally sized volumes on one disk group. We repeated this same provisioning step for each of the other two disk groups and were able to create 60 volumes in less than two minutes.

A very straightforward process is used to configure hosts for the HP P2000. We configured several hosts to have access to this storage.



Hosts Overview

Host ID	Name	Discovered	Mapped	Profile	Host Type
iqn.1991-05.com.microsoft:dmrtk-srvr-e	dmrtk-srvr-e	No	Yes	Standard	iSCSI
iqn.1991-05.com.microsoft:dmrtk-srvr-j	dmrtk-srvr-j	No	Yes	Standard	iSCSI
iqn.1991-05.com.microsoft:dmrtk-srvr-b.lab.demartek.com	dmrtk-srvr-b	No	Yes	Standard	iSCSI
iqn.1991-05.com.microsoft:dmrtk-wks-d	dmrtk-wks-d	No	No	Standard	iSCSI
iqn.1991-05.com.microsoft:dmrtk-srvr-g	dmrtk-srvr-g	No	No	Standard	iSCSI
iqn.1991-05.com.microsoft:dmrtk-srvr-f	dmrtk-srvr-f	No	No	Standard	iSCSI

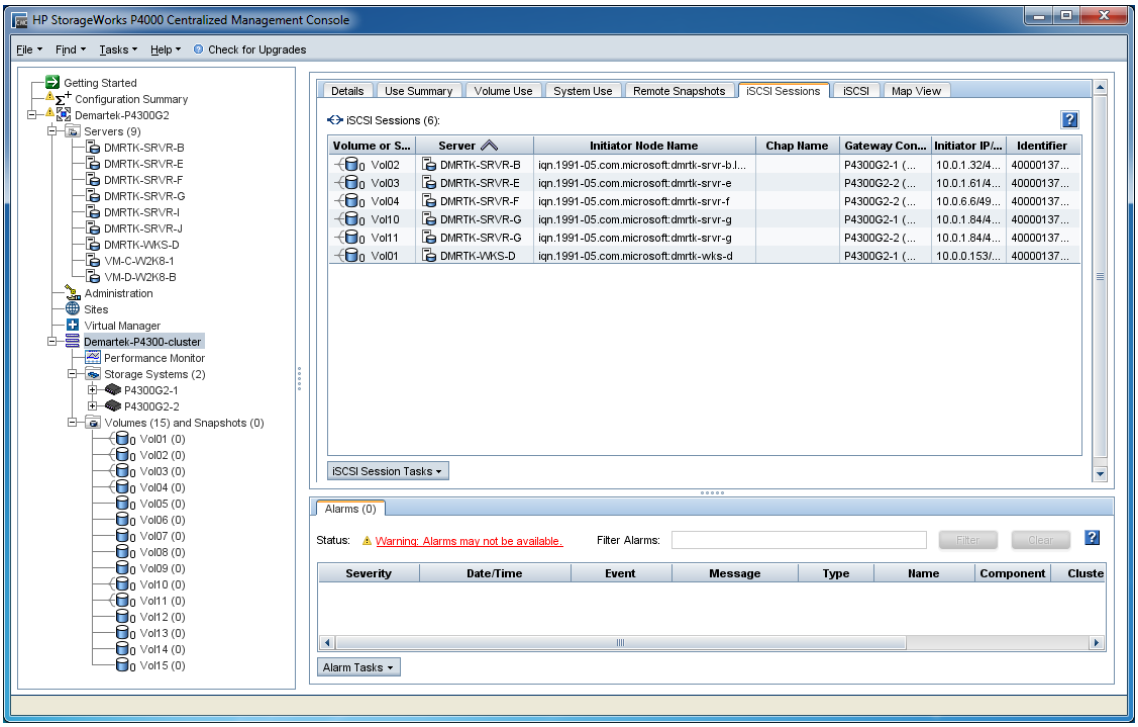
HP P4000 SAN Solutions

The HP P4000 SAN product family is a scale-out iSCSI storage solution. The HP P4000 storage solution uses storage nodes, which each contain CPU, disks, controllers, and network interfaces, and clusters them together, whereby, consolidating all of their critical resources into pools of storage. All of the available capacity and performance is aggregated and is available to every volume in the cluster. The HP P4000 series leverages network RAID technology which synchronously replicates data across a cluster of storage nodes. This volume level setting protects the data from any single point of failure, up to, and including a storage node. Volumes can be thin provisioned for space efficiency, and have thin provisioned, reservation-less, snapshots created for point-in-time copies of data. The HP P4000 is capable of replicating snapshots to other P4000 systems at primary or remote locations for disaster recovery, and these snapshots are also thinly provisioned.

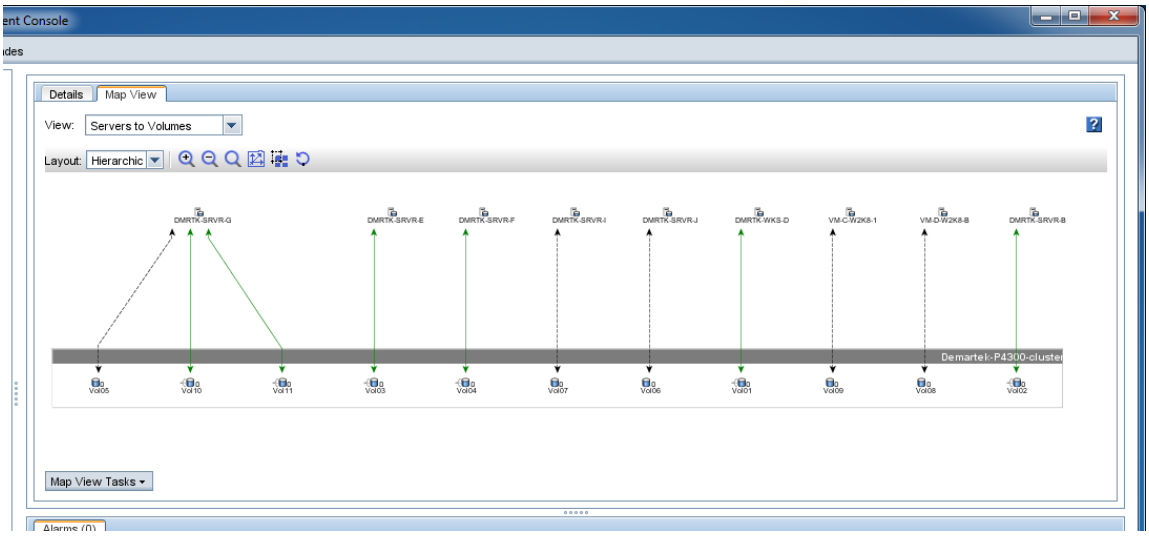
P4000 G2 SAN family is a scale-out technology that allows you to cumulatively add to a P4000 iSCSI SAN's overall capacity and performance and can be added non-disruptively. P4000 iSCSI SAN configurations start at 7.2 TB and can grow to 768 TB. To learn more about The P4000 G2 product family go to www.hp.com/go/p4000.

HP P4000 product family was formerly known as the LeftHand Networks storage system. HP announced its acquisition of LeftHand Networks in October 2008.

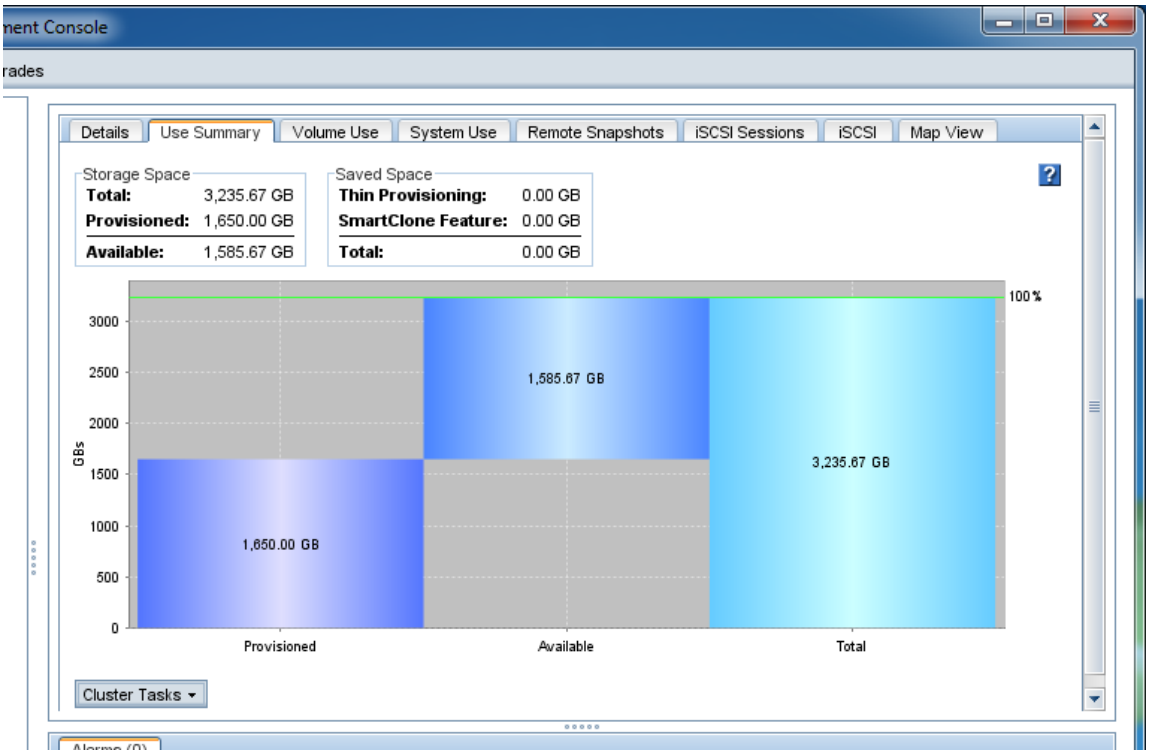
The HP P4000 series is managed with the Centralized Management Console (CMC) that can be run from a desktop or server platform. From the CMC, storage nodes, clusters, and volumes can be managed regardless of location. The CMC handles all the usual iSCSI host assignments and mappings, CHAP settings, etc.



The CMC has some interesting graphical representations of the host-to-volume mappings. These graphical representations can be rotated so that servers are above the storage, as depicted below, or can be rotated 90° in either direction or 180°.



Another useful graphical representation is the storage usage summary, shown below.


ment Console
rades
Details
Use Summary
Volume Use
System Use
Remote Snapshots
iSCSI Sessions
iSCSI
Map View
Storage Space
Total: 3,235.67 GB
Provisioned: 1,650.00 GB
Available: 1,585.67 GB
Saved Space
Thin Provisioning: 0.00 GB
SmartClone Feature: 0.00 GB
Total: 0.00 GB
100 %
3000
2500
2000
1500
1000
500
0
GBs
1,585.67 GB
3,235.67 GB
1,650.00 GB
Provisioned
Available
Total
Cluster Tasks

Storage Targets: NetApp

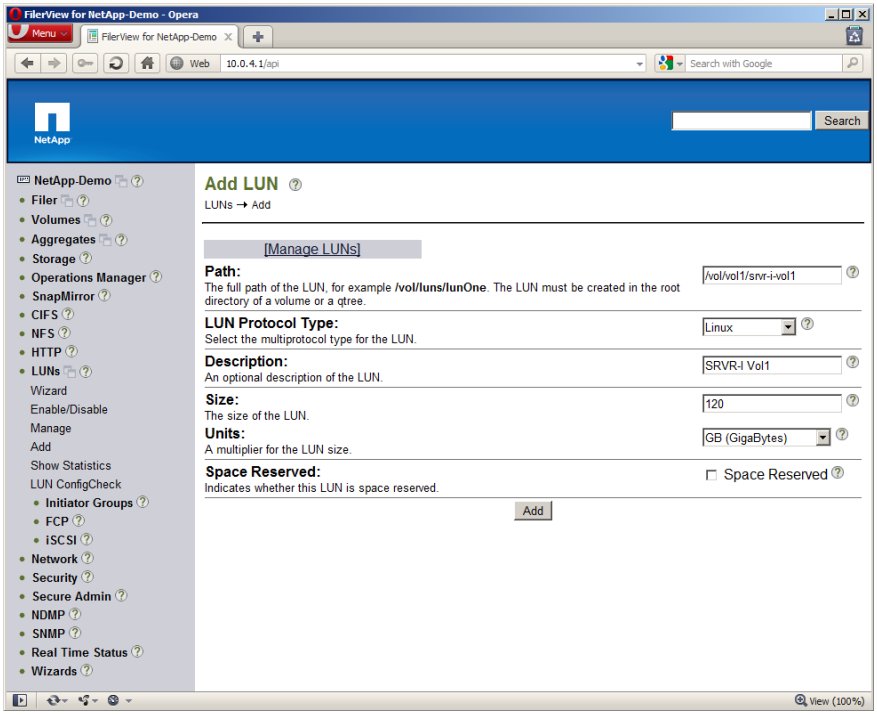
NetApp builds a complete family of unified storage systems that support file protocols such as NFS and CIFS, and block protocols including iSCSI, Fibre Channel, and FCoE. This product family includes the FAS2000 series, FAS3x00 series and the FAS6x00 series. These models scale from the low-end FAS2000 series that are designed for branch offices or departments all the way up to the FAS6x00 series designed for large enterprises with almost 3PB of raw storage capacity in one system. This full product line gives IT administrators and managers the ability to choose the performance and capacity that is needed.

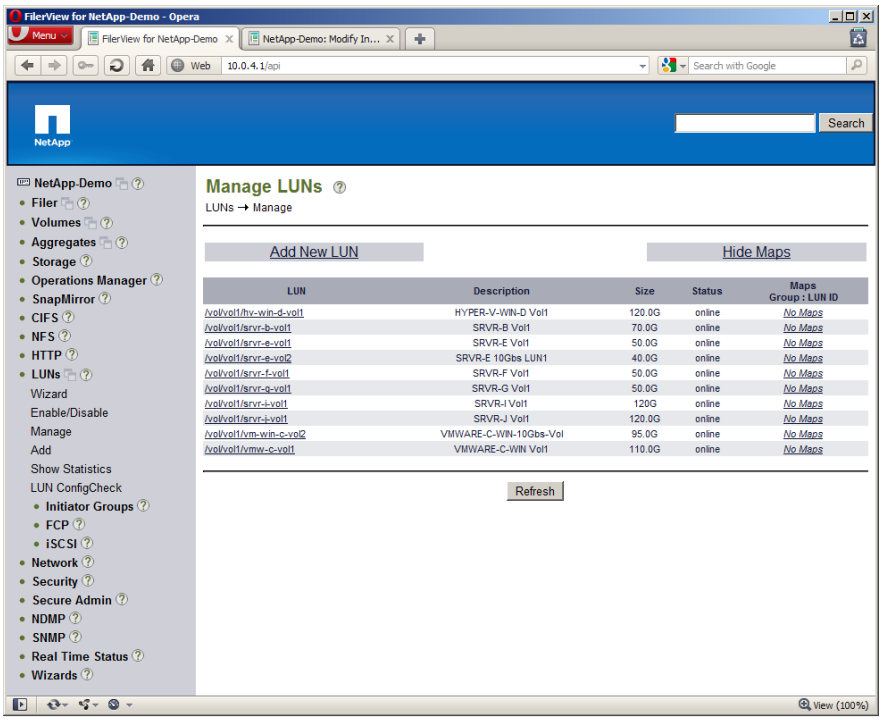
NetApp uses the Data ONTAP operating environment across the entire FAS product line. Data ONTAP provides a consistent management experience that not only supports multiple protocols, but includes storage efficiency features such as thin provisioning, compression, data de-duplication, and various forms of snapshot technologies. The newer models of the FAS product line also support Flash Cache, NetApp's solid-state caching technology that caches recently read user data and NetApp metadata in the storage controller, working with all storage protocols and NetApp software.

Data ONTAP can be used with a graphical user interface accessible through any web browser or through a command-line interface. The graphical user interface is intuitive and includes online help. Data ONTAP supports a rich set of iSCSI technologies, including multipath I/O, a wide variety of iSCSI software and hardware initiators, 1Gb and 10Gb interfaces, iSNS servers, multiple connections per session, and CHAP.

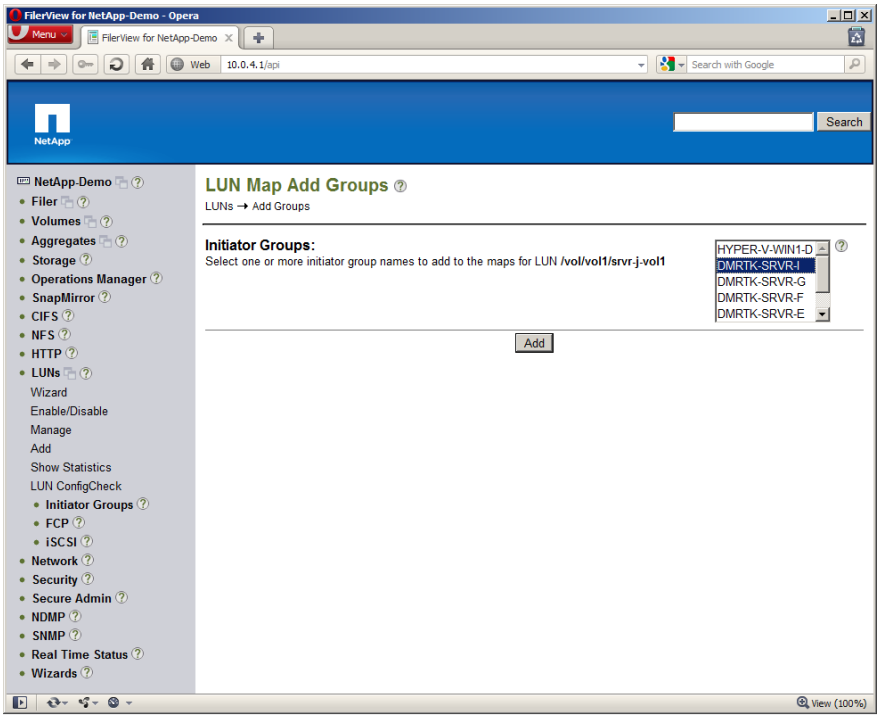
We tested two of the smaller models of the FAS product line in our lab with Data ONTAP 8.0.1, the FAS2040 and FAS3040. We tested and configured iSCSI LUNs, CHAP, adding iSCSI initiators configured the network interfaces and more. We used 1Gb connections on the FAS2040 and both 1Gb and 10Gb connections on the FAS3040. The Data ONTAP default RAID type is RAID-DP, also known as RAID6. RAID-DP provides protection in the event that two disk drives fail in the same raid group.

LUNs, for either iSCSI, Fibre Channel, or FCoE, are created within volumes that live on aggregates. LUNs are created within a volume and include the full volume path and can be space reserved.

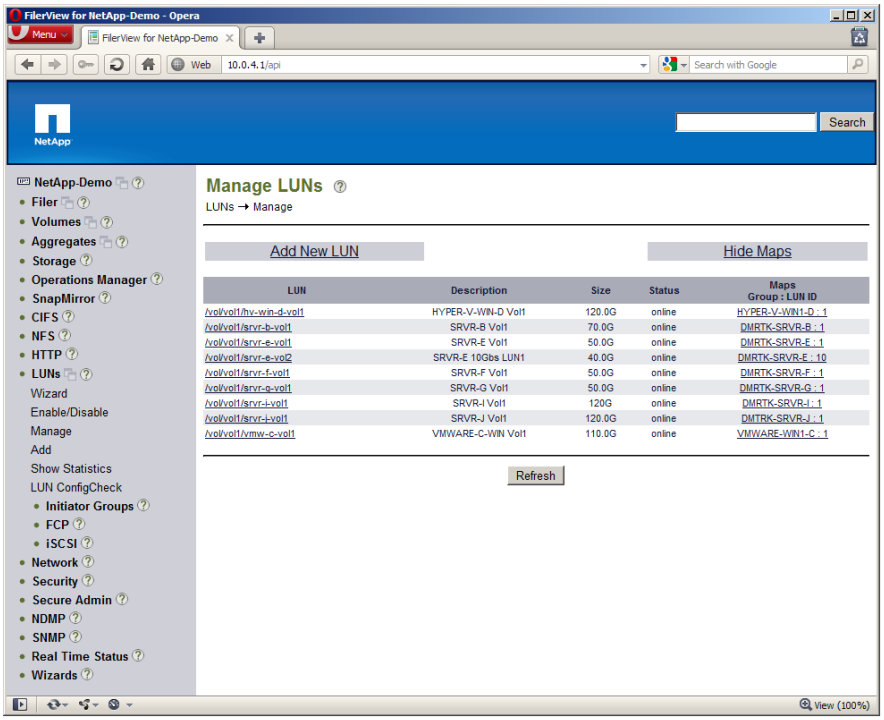




After creating the LUNs, they can be mapped to initiators or groups of initiators.



After all the LUNs are mapped, a full list of the LUNs can be displayed, showing which LUNs are mapped to which initiators and the LUN number for that initiator.



Performance Results

We ran an Exchange Jetstress test on a host server with a QLogic QLE4062-C 1Gb iSCSI HBA connected to a NetApp FAS3040. It should be noted that the FAS3040 is a smaller, older model and that the current models of FAS3x00 storage systems are capable of significantly higher performance. This particular model of FAS3040 contains one shelf (14 disk drives) of 10K RPM disk drives. Newer models of the FAS3x00 and FAS6x00 systems also support NetApp's Flash Cache, which significantly improves performance.

FAS3040

Exchange mailbox profile: mailboxes=500, mailbox size=600MB, IOPS=0.18 Threads=2 Storage Groups=3, total database size=300GB, LUN size=1TB, RAID-DP

Results

Achieved Transactional I/O per Second (32KB):	268.905 (almost triple the target)
Target Transactional I/O per Second:	90.0
Average Database read latency (all storage groups):	14.242 msec
Average Log write latency (all storage groups):	3.377 msec

Storage Targets: Xiotech

The Xiotech Intelligent Storage Element (ISE) storage solution combines a Microsoft Windows Storage Server 2008 R2 front-end management server with Xiotech's ISE blades to provide a dense, high-performance storage solution in a small footprint.

Xiotech's intellectual property is focused on their ISE technology that includes sealed drive canisters containing 10 or 20 disk drives (DataPac), battery modules, power and cooling units, and a pair of active-active managed reliability controllers inside each three rack-unit (3U) storage blade. These storage blades can contain up to 19 TB of raw capacity in a 3U form factor.

The Xiotech DataPac includes a mechanical design that dramatically reduces heat and vibration in order to extend the life of the hard disk drives and allows the drives to operate at maximum performance. In addition, Xiotech has developed advanced firmware that incorporates the ANSI T10-DIF (data integrity field) self-healing technology.

Xiotech recently announced their Hybrid ISE that contains 10 hard disk drives and 10 SSDs per DataPac that provides very high performance in a 3U form factor.

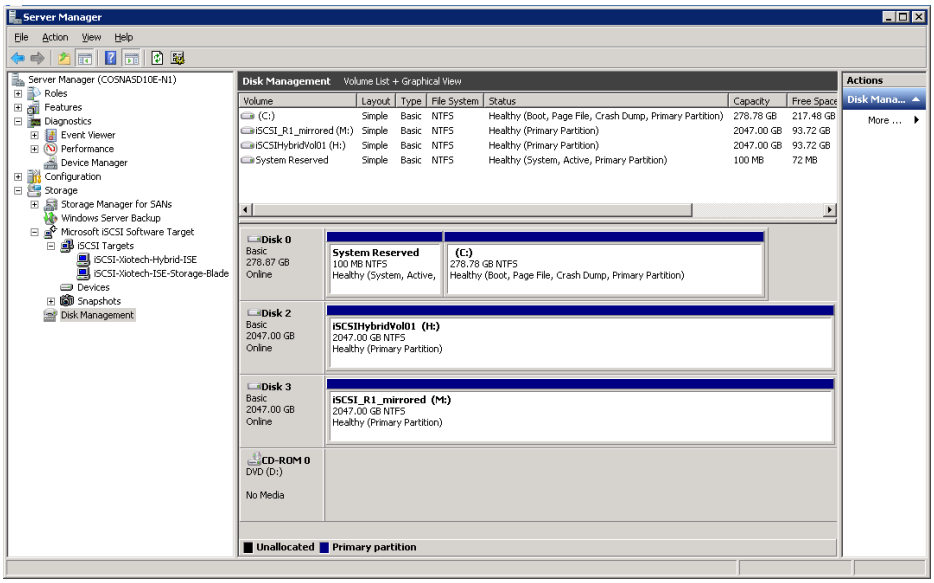
We had the opportunity to test their first-generation ISE storage system and to test an early version of the Hybrid ISE storage, both with 10Gb iSCSI host interfaces. Xiotech also offers 8Gb Fibre Channel host interfaces. Additional information on particular models and configurations is available at http://www.xiotech.com/products-services/ise-storage-blade/.

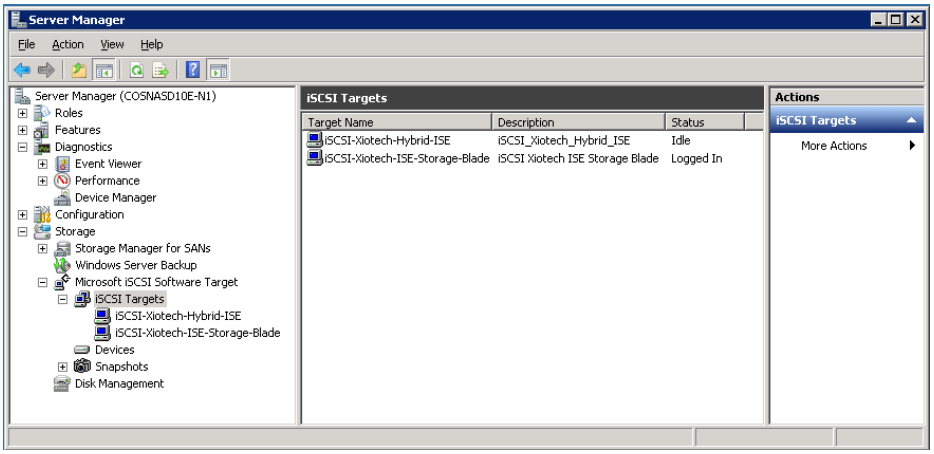
Configuration and deployment of the Xiotech ISE system will be very familiar to Windows server administrators, because it uses Windows Storage Server 2008 R2 for its primary management functions. This familiarity extends to the configuration of the iSCSI targets, due to its use of the Microsoft iSCSI software target that is included with Windows Storage Server 2008 R2.

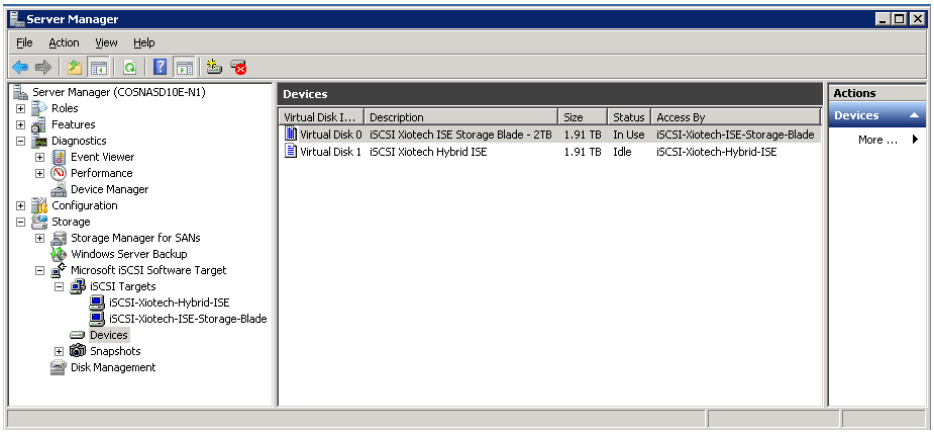
We configured two RAID1 LUNs of 2TB each, one on a first generation ISE DataPac and one on the new Hybrid ISE DataPac. We used the standard Microsoft iSCSI target and Disk Manager management applications to configure the storage on the storage management server. After configuring the storage target, we used the Microsoft iSCSI initiator on the host server to connect to the storage, one LUN at a time, so that we could run Exchange Jetstress to test the performance of each type of DataPac.

Target Configuration Summary

The following screenshots show the configuration of the Xiotech ISE target with two separate DataPacs. The Disk Management and Microsoft iSCSI Software Target configurations are shown.

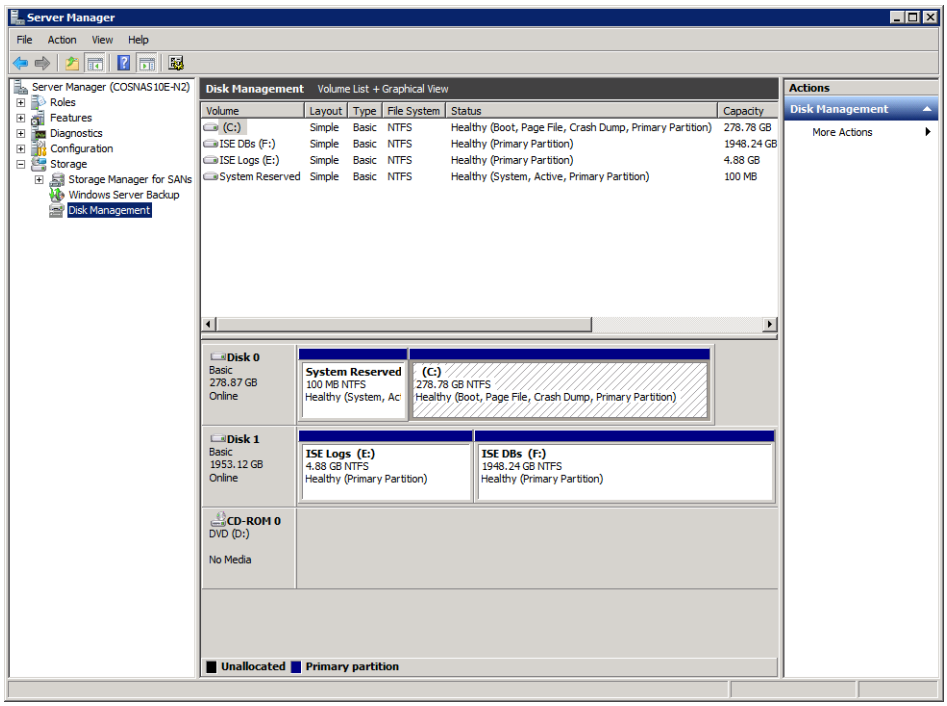




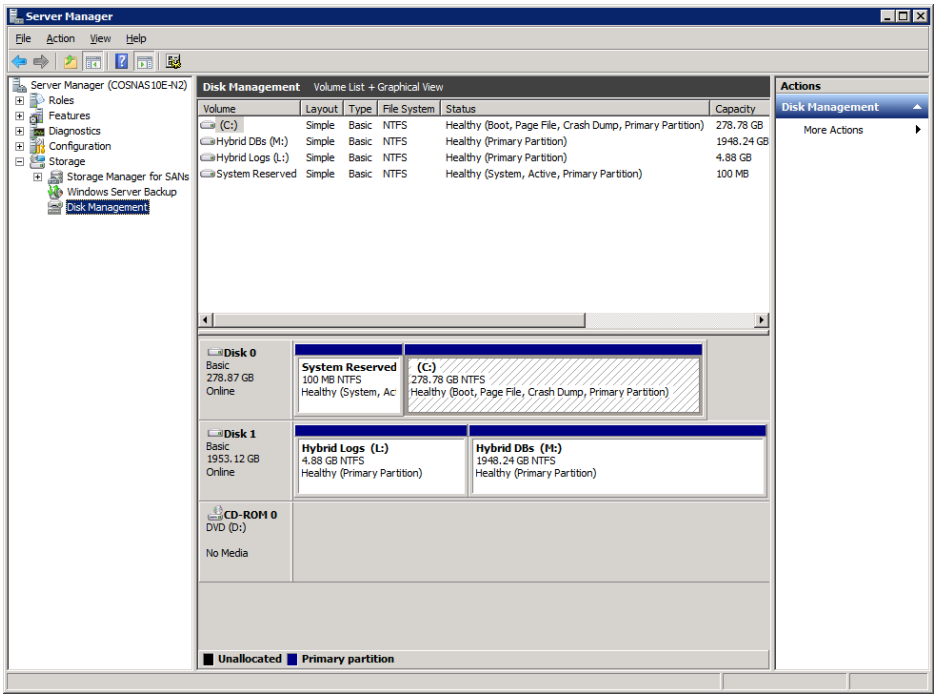


Initiator Configuration Summary

The following screenshots show the configuration from the initiator host server. The first configuration used drives E: and F: on the first generation ISE.



The second configuration used drives L: and M: on the Hybrid ISE.



Performance Results

We ran two different Exchange mailbox profiles using Microsoft Exchange Jetstress 2010, one for the first-generation ISE DataPac, and a second set of tests for the Hybrid ISE DataPac. As a standard feature, the Xiotech Hybrid ISE DataPac performs analysis of the I/Os going to the hard drives and uses solid-state drives to cache I/O's over time which accelerates the overall performance.

The historical best practice for running Exchange Jetstress, and for the full Exchange Server, is to configure the databases and the logs on separate disk groups. However, Xiotech stripes all logical volumes across all the drives in each DataPac, and we only used one DataPac at a time, so the databases and logs were on the same set of disk drives for these tests. Additional information about Jetstress is available in the Exchange Jetstress Performance Testing section of this report.

Test 1 – First Generation ISE DataPac (20 hard disk drives)

Exchange mailbox profile: mailboxes=1500, mailbox size=1000MB, IOPS=0.12 Threads=1 Storage Groups=19, total database size=1.5TB, LUN size=2TB, RAID10

Results

Achieved Transactional I/O per Second (32KB):	362.501 (more than double the target)
Target Transactional I/O per Second:	180.0
Average Database read latency (all storage groups):	11.005 msec
Average Log write latency (all storage groups):	4.426 msec

Test 2 – Hybrid ISE DataPac (10 hard disk drives plus 10 SSDs)

Exchange mailbox profile: mailboxes=200, mailbox size=1000MB, IOPS=0.30, Threads=32, Storage Groups=2, total database size=200GB, LUN size=2TB, RAID10

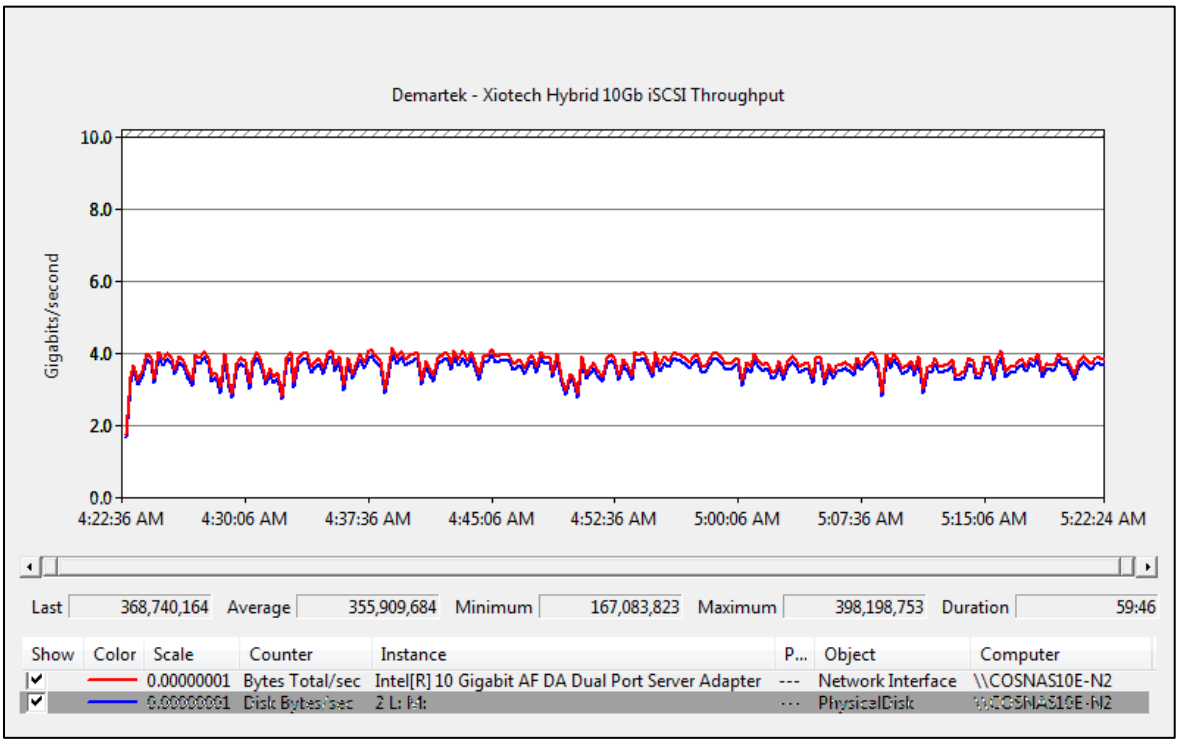
A series of 15 runs with the Exchange mailbox profile were run to observe the effects of the SSD caching.

Results

Achieved Transactional I/O per Second (32KB):	8934.418 to 10285.565 (15 runs)
Target Transactional I/O per Second:	60.0
Average Database read latency (all storage groups):	6.899 to 6.1845 msec (15 runs)
Average Log write latency (all storage groups):	5.1875 to 4.8245 msec (15 runs)

iSCSI Throughput

For the Hybrid ISE Jetstress tests, the best performance averaged 3.7 Gb/second sustained throughput on the 10Gb Ethernet link and 3.5 Gb/second on the iSCSI disk. The graph from Perfmon shows the 10GbE link and the disk throughput rates.


Demartek - Xiotech Hybrid 10Gb iSCSI Throughput
Gigabits/second
10.0
8.0
6.0
4.0
2.0
0.0
4:22:36 AM
4:30:06 AM
4:37:36 AM
4:45:06 AM
4:52:36 AM
5:00:06 AM
5:07:36 AM
5:15:06 AM
5:22:24 AM
Last 368,740,164 Average 355,909,684 Minimum 167,083,823 Maximum 398,198,753 Duration 59:46
Show Color Scale Counter Instance P... Object Computer
0.00000001 Bytes Total/sec Intel[R] 10 Gigabit AF DA Dual Port Server Adapter --- Network Interface \\COSNAS10E-N2
0.00000001 Disk Bytes/sec 2 L: H: --- PhysicalDisk \\COSNAS10E-N2

Legal and Trademarks