Best Practices for Configuring the IBM System Storage™ DS3300 and an IP SAN

Best Practices IP SAN Guide
June 2009

Table of Contents

Introduction

Servers have traditionally used the Small Computer Systems Interface (SCSI) to provide block level access to various peripheral devices such as scanners, tape drives, disk drive drives and storage subsystems. As the need for increased speed, reliability and distance arose for block-level access, the Fibre Channel industry used the SCSI-3 protocol as the I/O layer to accommodate these requirements. Fibre Channel Storage Area Networks (FC SANs) have since become the defacto standard for many enterprise and corporate data center environments. Unfortunately, the cost and complexity of FC SANs, have made this a non-viable solution for many smaller, less complex organizations with smaller budgets.

In IP based networks, SCSI has not been generally available until the release of the iSCSI protocol. With iSCSI, existing TCP/IP (Transmission Control Protocol/Internet Protocol) networks can be utilize to transmit block oriented data across your Local Area Network (LAN) or even further using your Wide Area Network (WAN). The SCSI design uses an initiator/target design where a device such as a host will initiate data transfers to various target devices such as the IBM System Storage™ DS3300.

The iSCSI protocol transports inputs and outputs (I/O) from a host to a storage device over an IP network using SCSI commands generated by the file system of a server. The iSCSI protocol packages the SCSI commands so that they can move through an IP network and once at their destination the commands are then executed. For example a host wishing to write a file to a storage device would use normal SCSI read and write commands to a centralized storage subsystem but the iSCSI protocol would package the commands and transport them transparently to the host, the host file system and the host applications.

While similar on the surface to the more common Networked Attached Storage that uses the Common Internet File System (CIFS) or the Network File System (NFS), iSCSI provides raw storage to a server and it's applications as opposed to a file oriented storage share.

Disclaimer

The steps in this document were performed on a test setup. There was no live data involved. This document implies no warranties. These instructions are not guaranteed to work in every situation. For the latest information on the DS3300, refer to the DS3300 Installation Guide found at http://www.ibm.com/servers/storage/support.

Planning your IP SAN

Before beginning the implementation of an IP SAN, review this section to understand the various concepts and components used in such a network.

The iSCSI Storage Subsystem

The iSCSI protocol is based upon the network-standard client-server model. Rather than "clients" and "servers", however, iSCSI features "initiators" and "targets". An initiator acts as a client device, and a target acts as a server. A connection must be established between the initiator and the target before data can be sent.

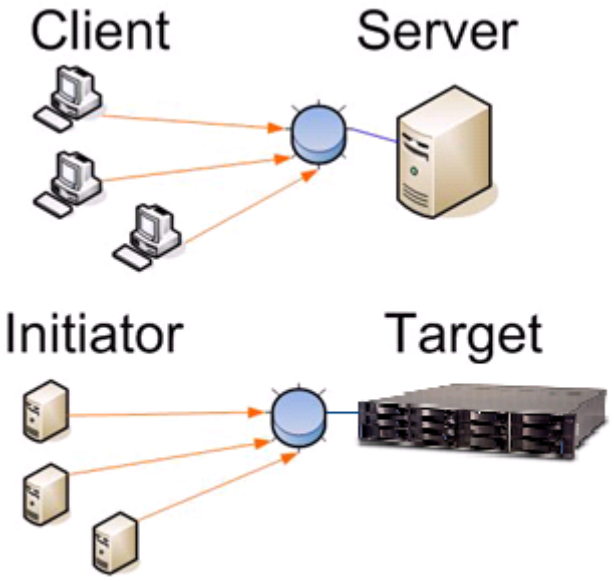

Figure 1: iSCSI Model

Targets listen for connection requests via TCP port 3260.This can be adjusted if required, but port 3260 is the default and there is little reason to alter it. Once a connection is established, other Ethernet ports may be created to handle subsequent communications as each target portal is configured independently for a TCP/IP listening port. An iSCSI connection is called a *session*. An active session is required for data transfer.

iSCSI names

While iSCSI devices use IP addresses like any other device on a network, these devices are identified primarily by an iSCSI "name". As it is difficult to track and remember IP addresses, particularly if they are new IPv6 addresses, it is much more convenient to address devices by a name.

This also allows devices to be accessed even when they are sharing IP addresses in a cluster, or when they have a changing IP when using DHCP. The same idea is used on networks today for standard devices such as servers. DNS (Domain Name Server) applications, for example, translate server names into actual IP addresses. And websites have their own URL (Universal Resource Locater) address, independent of actual IP address. It's much easier to remember and type "www.ibm.com", for example, than to know and type the particular IP address and port of the server you need to access.

iSCSI name breakdown

There are two primary types of iSCSI names, the "iqn" (iSCSI Qualified Name) name, and the "eui" (IEEE EUI-64 identifier) name. The "iqn" format was developed strictly for iSCSI. It is broken down into 4 main components: type, date, naming authority, and a unique string.

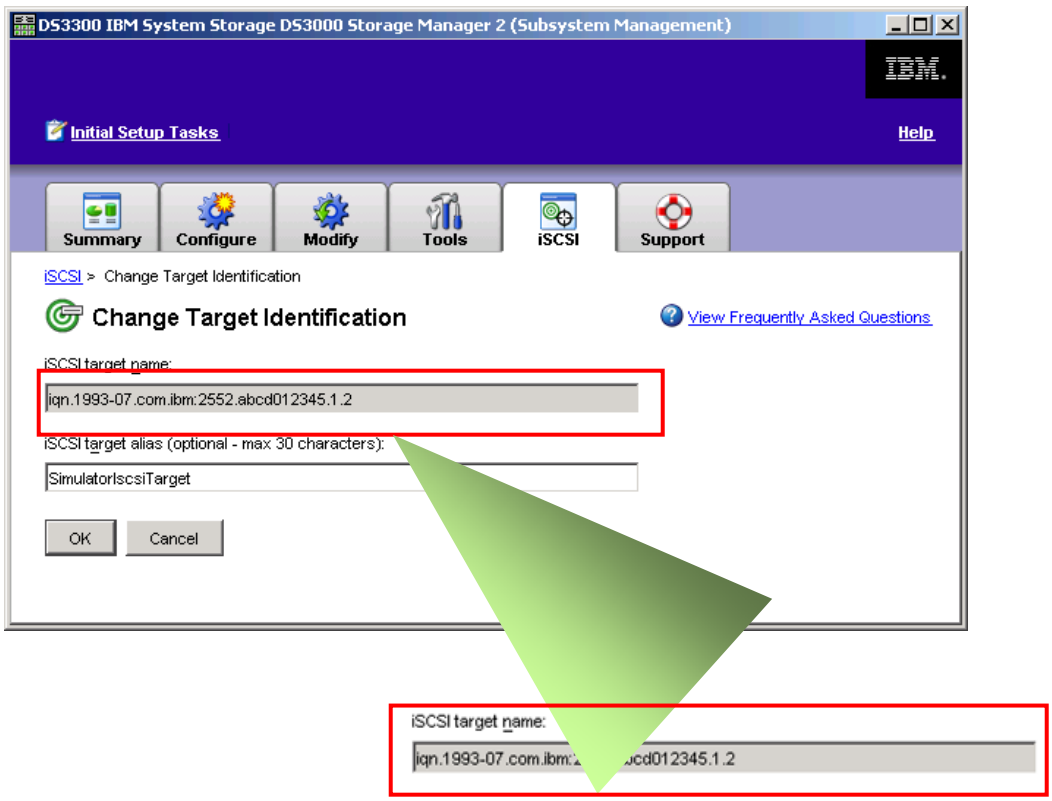

Figure 2: iSCSI name example

In Figure 2, you can see the four sections. The first section, the type, is the "iqn" string indicating that this is an iSCSI qualified name. A period delineates the next field, which is a date code. This is a date that corresponds to the next field, which is a reversed domain name. The date should be the point in time at which the domain name was registered. The combination of the domain and date indicate that the name is based on a registered authority. A colon separates the first three fields (the qualification) from the unique string that identifies the device. In this example, the IBM controller has generated a name based upon the controller WWN. Many initiators and targets allow the user to alter the iSCSI target name (Note: the DS3300 does not support the changing of the iSCSI target name) . Be careful when doing so, as improper names can cause conflicts. The alias option is provided so that more user-friendly names can be used alongside the official iSCSI name.

Discovery

While separating the iSCSI name from the actual IP address of a device adds flexibility to the scheme, it also requires an additional step when connecting to an iSCSI device. This additional step is the discovery session, when an initiator must identify a target and link a particular IP address and port to a specific iSCSI name. The simplest discovery session is performed when the initiator is given a specific IP address and port (typically 3260). A connection request is sent to that IP and port, and the iSCSI device

(if the address and port are legitimate) can reply with iSCSI name information to establish a more permanent session. However, it is not always desirable to make the IP addresses of all iSCSI devices known to all users due to maintaining security or bandwidth across the network. You may also wish to have a subset of iSCSI devices accessible by only a certain group of initiators. To manage this type of configuration, the iSNS (Internet Storage Name Service) is available.

iSNS

Much like a DNS (Domain Name Server), the iSNS acts as a go-between for clients and servers. The iSNS should run on a server that has a well-known IP address. Both initiators and targets can register with the iSNS, and the service will manage the discovery sessions between the two types. The iSNS IP address is either stored with the initiators and targets or obtained from a DHCP server. This way, they can register their current IP with the service at any time. Rather than trying to discover a specific target, an initiator can simply request information from the iSNS and find all the targets known to the service. Additionally, an iSNS can separate devices into domains.

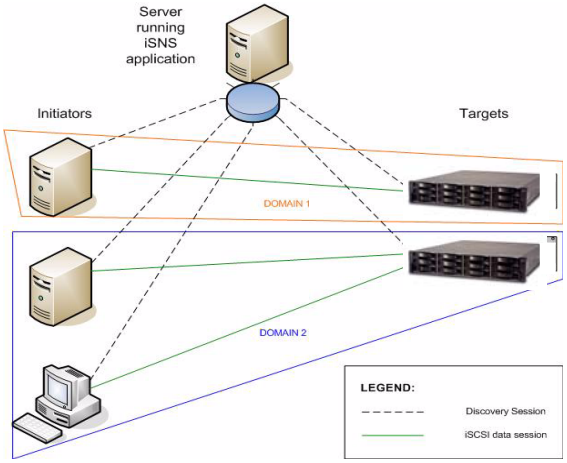

Figure 3: iSNS application

The iSNS can be set to present only certain targets to certain initiators. This is useful in environments that need to use the iSNS service, but don't want all iSCSI devices to be visible to all servers.

Portals

Another important concept is the idea of iSCSI "portals". A given iSCSI device may have only one name, but many IP addresses or physical Ethernet connections. This is critical for high availability systems. Why create a redundant, powerful, iSCSI device that is subject to failure due to a single bad RJ-45 connection? An iSCSI portal is the actual connection to the network for a given iSCSI device. Every iSCSI device requires at least one portal, but it may be advantageous for a device to have multiple portals. Multiple access points to the network will increase reliability, and add flexibility by making it possible to give a device access to multiple subnets without relying on a single connection. While portals typically have a unique IP address, it is possible to group multiple portals with a virtual IP address, giving physical redundancy without logical redundancy. It is also possible to develop IO drivers to take advantage of multiple iSCSI portals to a given iSCSI device, improving performance by spreading data transfer over more physical connections.

Security

Since iSCSI devices are on the network, they need security protocols to protect the information stored on them. While initiators and targets can register with the iSNS and can provide some layer of security, more security measure are required to prevent malicious users from attacking the device via random IP queries and other methods.

• CHAP
The Challenge Handshake Authentication Protocol (CHAP) is the standard security measure for iSCSI devices. When enabled, any device with CHAP protection will "challenge" any application that tries to establish a connection. If the querying application (typically another iSCSI device) cannot respond with a particular string, the connection is refused. This essentially acts as password protection. The "password" for an iSCSI device using CHAP is called the *CHAP secret*. Initiators and targets may both have unique secrets. Any iSCSI device may have its own unique CHAP secret. Note that if CHAP protection is enabled, the secret must be known to any device trying to access the protected iSCSI node, including the iSNS if used.

Other Security

The iSCSI protocol works with all other types of internet security, such as VPN (Virtual Private Network), and ACL (Access Control List). These security measures have no discernable effect on iSCSI operation, as they apply to lower levels of the standard OSI network model.

Implementation Concepts

Since iSCSI is designed to work with existing network hardware, it is possible to operate a system without adding anything besides the iSCSI storage device itself. Software exists for a variety of platforms that will emulate an initiator using the standard network connection on a server, whether it's part of the motherboard or a NIC (Network Interface Card) in an expansion slot. However, many vendors are also offering NICs that manage iSCSI connections without the overhead of the additional software. In practice, there are several ways to approach building an iSCSI solution.

Hardware

• iSCSI HBAs
When using a QLogic iSCSI HBA or any other brand HBA, you should check the manufacturer's website for information and firmware updates required to support your iSCSI SAN. Additionally, you should check the operating system website for iSCSI SAN support information. Operating system sites such as Microsoft, RedHat, SUSE and VMware will post iSCSI SAN information and patches. When searching these websites suggested keywords might be: iSCSI SAN, boot-from-SAN, QLogic, VMware, or Windows 2008. These are only suggestions and not the full set of search words or vendors.

• TOE NICs
Another option is a card that handles some of the network traffic, but still relies on another program to perform the actual iSCSI protocol work. These TOE (TCP/IP Offload Engine) cards can improve performance by relieving the server's CPU of extra cycles that would otherwise be spent creating network packets. However, they are not as efficient as true iSCSI HBA cards.

• Onboard TOE ports
A standard on some servers includes an onboard Ethernet ports that provide TOE support. This is an advanced network card offers features such as Receive-Side Scaling (RSS) and TCP/IP Segmentation Offload (TSO). RSS balances the network load across multiple CPUs and TSO breaks down data into smaller segments that pass through the network connection and reduces CPU overhead.

	NIC	TOE	iSCSI Offload HBA
Definition	Network Interface Card – provides Ethernet connectivity	TCP/IP Offload Engine – a specialized NIC that provides additional functionality	iSCSI HBA - an HBA that provides Ethernet connectivity and additional functionality
What "work" do they offload	Physical and data link communication	TCP/IP, physical and data link communication	iSCSI read/write processing, TCP/IP processing, physical and data link communication

	NIC	TOE	iSCSI Offload HBA
Work that must be done by the CPU	iSCSI protocal management and TCP/IP protocol management	iSCSI protocal management	None
Leverage SW based-initiator?	Yes	Yes	No – HW based
Location	Installed within the server, typically on the motherboard	Installed in server, usually as an expansion card	Installed in server, usually as an expansion card
iSCSI performance	Adequate	Good	Best
Additional notes	Easiest to use	Can help with other networking applications besides iSCSI	Requires more setup time and configurations

Figure 4: Comparison between different adapter technologies

• Switches

Carefully consider the switch options for the DS3300. The switches in Layer 2 and the switches in Layer 3 each offer different advantages.

The switches in Layer 2 effectively provide the same functionality as a bridge. They are similar to multiport bridges in that switches in Layer 2 learn and the switches forward frames on each port. The major difference between a switch in Layer 2 and a bridge is the involvement of hardware that ensures the multiple switching paths inside the switch can be active simultaneously.

The switches in Layer 3 operate as routers with the fast-forwarding performed by hardware. IP forwarding typically involves a route lookup, which includes the following actions:

- Decrementing the time to live (TTL) count
- Recalculating the checksum
- Forwarding the frame with the appropriate MAC header to the correct output port

Routers use routing protocols such as the following:

- Open Shortest Path First (OSPF)
- Routing Information Protocol (RIP)

Using routing protocols, routers communicate with other Layer 3 switches or routers and build their routing tables. The routing tables determine the route for an incoming packet.

The best solution is to use Multipath I/O (MPIO) which allows you to create multiple connections from each server's iSCSI initiator to your storage subsystem (specify the Ethernet connection and path the data should take.) MPIO requires that each Ethernet

card connected to your DS3300 in your server has its own IP address so you can set paths in different subnets.

Software

Nearly every major OS vendor now provides iSCSI software free of charge. The standard product is a software initiator. This is a program that runs in the background (as a UNIX daemon or Windows service) and captures network frames from a given network interface, looking for iSCSI commands. This initiator program will also present iSCSI sessions to the OS as target storage devices, in the same manner as a HBA would. And, a good software initiator provides an inexpensive and simple way to start using iSCSI.

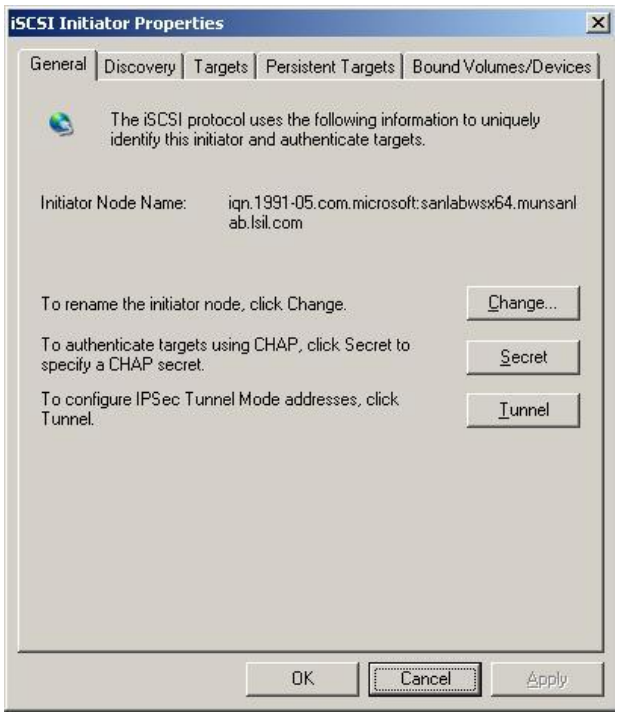

Figure 5: Microsoft iSCSI Initiator Main Screen

Additionally, most OS vendors provide an iSNS program. Like the software initiator, these also are usually free of charge. And like the software initiator, this program will consume resources on the server it is running. However, it does not have nearly the same performance impact as the software initiator, since the only function of the iSNS server is to point initiators to targets. Once a session between an initiator and target has been established, the iSNS is out of the picture, even if it was used during the discovery session to link the iSCSI devices.

IP SAN Planning

The DS3300 is available in two configurations. This first is a simplex configuration, which contains one storage controller. The second offering is a duplex configuration with dual active controllers. Dual active is defined as one controller owning a specific logical drive, while the other controller will "own" another logical drive. Allocating the logical drives among both controllers will ensure that the load is equally balanced across both storage controllers and maximizes potential performance.

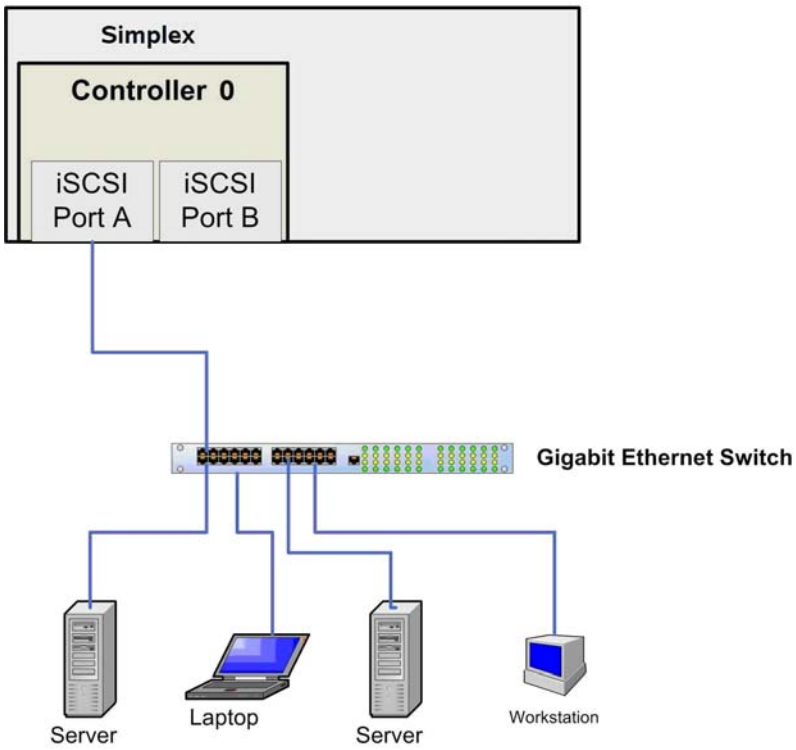

Figure 6: Non-redundant IP SAN implementation

As shown in Figure 6, this is the non redundant and most simple implementation of an IP SAN. The benefits are low infrastructure costs, but if any component within the IP SAN fails, access to the data will be lost.

Figure 7 shows a fully redundant iSCSI SAN using the DS3300 with duplex controllers. Such a configuration requires a larger investment, but if any component fails, the data will be still accessible. This Setup also allows you to use the high availability (HA) features of the DS3300 such as online firmware upgrade and load balancing between the two NICs in the Server to a logical drive owned by Controller 0, for example. This is also the preferred option for using the DS3300 in an IP SAN.

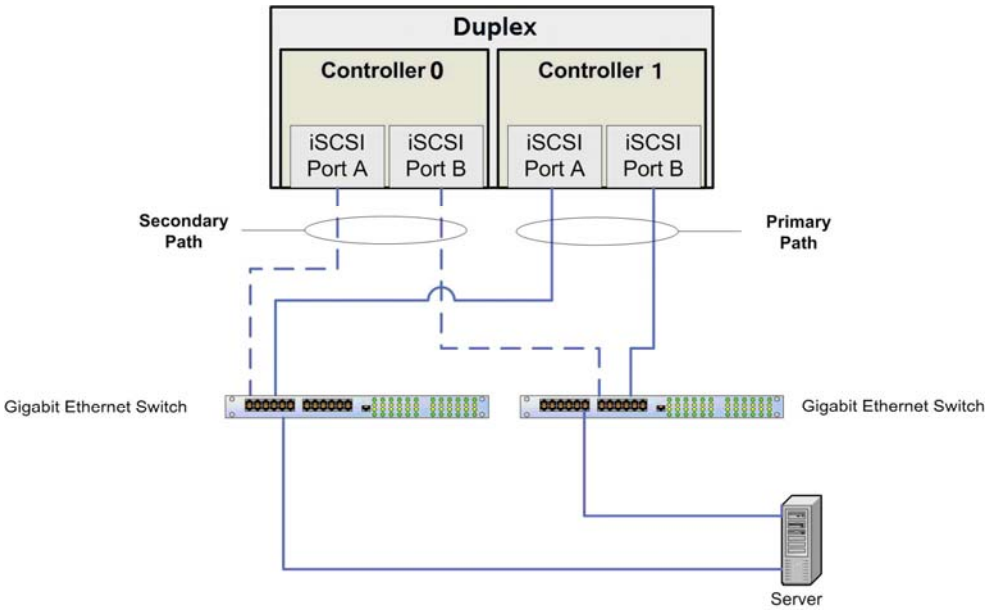

Figure 7: Highly available IP SAN configuration

Installing the DS3300 in an IP SAN

The IBM System Storage™ DS3300 Installation

Before the installation of the DS3300 storage subsystem in an IP SAN environment, make sure all of the components you are intending to use are qualified with the DS3300. To verify your components, refer to the following document:

http://www-03.ibm.com/systems/storage/disk drive/ds3000/pdf/interop.pdf

Ensure the storage subsystem is cabled correctly. Figure 8 displays the cabling of simplex and duplex controllers. The simplex configuration includes one path to the system and drive expansion enclosures. If any component in the simplex configuration goes offline, access to data will be lost. To avoid any potential loss of access to data, utilize the redundant component and high availability of the DS3300. Set up the duplex DS3300 and the EXP3000 expansion enclosures in a redundant configuration to maximize your reliability and availability.

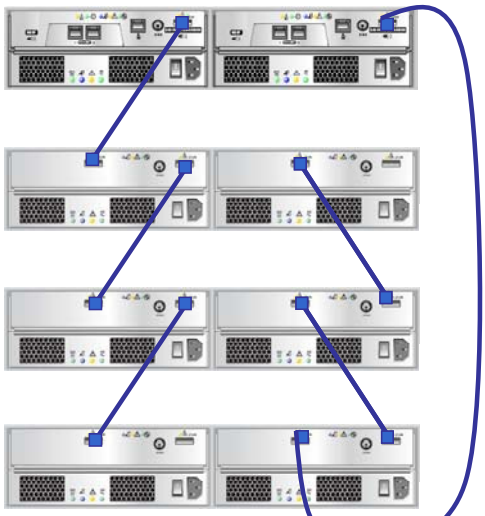

Figure 8: Duplex Configuration

Power Up/ Power Down

A very important hint: Make sure the DS3300 with attached EXP3000 expansion enclosures is powered up and shut down in the correct sequence:

Power down:

- Turn off the DS3300.
- Turn off the EXP3000 enclosure(s).

Power on:

- Turn on the EXP3000 expansion enclosure(s). Wait for the enclosure status LED to light blue.
- Turn on the DS3300 and wait for the status LED to indicate that the unit is ready:
  - If the status LEDs are lit as a solid amber color, the DS3300 is still coming online.
  - If the status LEDs are blinking amber, there is an error that can be viewed using the DS3000 Storage Manager.
  - If the status LEDs light a solid blue, the DS3300 is ready.

The DS3300 storage subsystem provides the value of storage consolidation without the cost and complexity of Fibre Channel network. The DS3300 can consolidate up to 32 fully redundant hosts, expand to support up to 48 TB of data and provides task-based installation, intuitive management, and advanced data protection. The DS3300 has redundant storage subsystem controllers, fans, and power supplies.

Interface Options

You can use either of the following two methods to configure the DS3300 storage subsystem:

DS3000 Storage Manager software – The DS3000 Storage Manager is an intuitive, task-based management interface significantly reducing the complexity of installation, configuration, management and diagnostic tasks. DS3000 ensures a friendly user interface from set-up to administration. With DS3000 software, a system can concurrently support multiple RAID levels, various array sizes, and one or more logical drives per array

Server Cabling

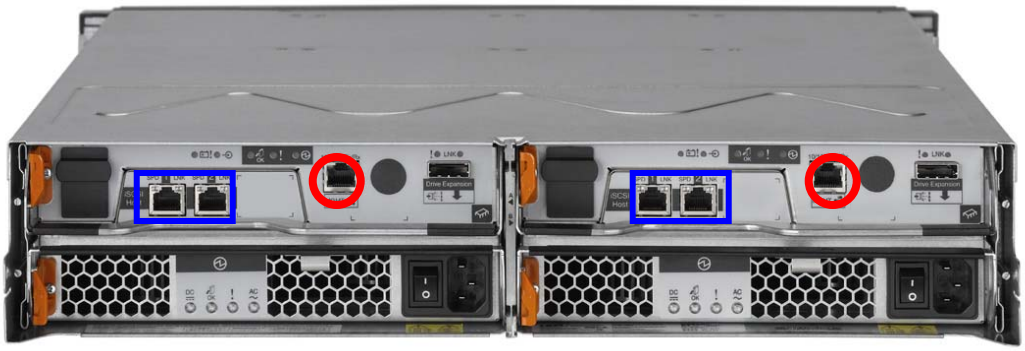

*Figure 9: B*ack of the DS3300

The four iSCSI ports of the DS3300 are marked blue here. Cable the System according to you selected configuration. The red Ethernet ports are for Management.

The Management IP Addresses are set by DHCP. If there is no DHCP server the default IP address for the Management ports of the controllers are:

- Controller 0: 192.168.128.101

- Controller 1: 192.168.128.102
- Port 2463 (registered with IANA; http://www.iana.org/assignments/port-numbers)
- IBM DS3000 Storage Manager uses TCP for connections
- IBM DS3000 Storage Manager uses UDP broadcast on local subnet for automatic discovery

The Microsoft iSCSI initiator has to be installed first, along with the Microsoft StorPort patch KB932755. The iSCSI Software initiator can be downloaded from Microsoft. There are different builds of the iSCSI initiator for the various platforms (x86, x64 and IA64).

http://www.microsoft.com/downloads/details.aspx?FamilyID=12cb3c1a-15d6-4585-b385-befd1319f825&DisplayLang=en

The Storport Patch can be found at:

http://support.microsoft.com/kb/932755/

Install the iSCSI Initiator on each server that will be attached to the DS3300 subsystem. If you are using the fully redundant configuration, ensure that the Microsoft MPIO Multipath Support for iSCSI checkbox is marked. A reboot of the server required once this installation has been completed.

At the Microsoft iSCSI Initiator Installation screen (Figure 10), check all Installation Option boxes except for Microsoft MPIO. IBM installs MPIO with a DS3300 Storage Manager and if it is checked at this step, it will not work appropriately.

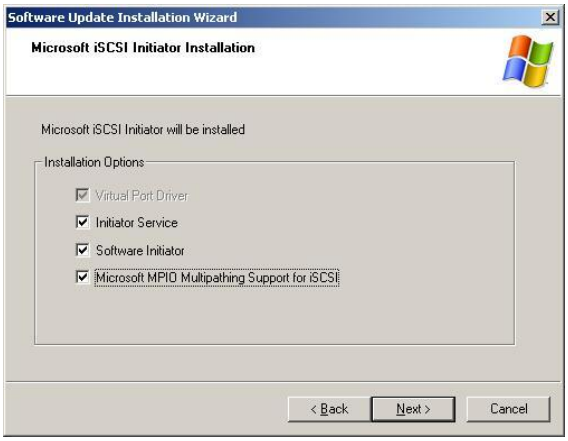

Figure 10: Microsoft iSCSI Initiator Installation Screen

Now the DS3000 Storage Manager can be installed on a Windows host.

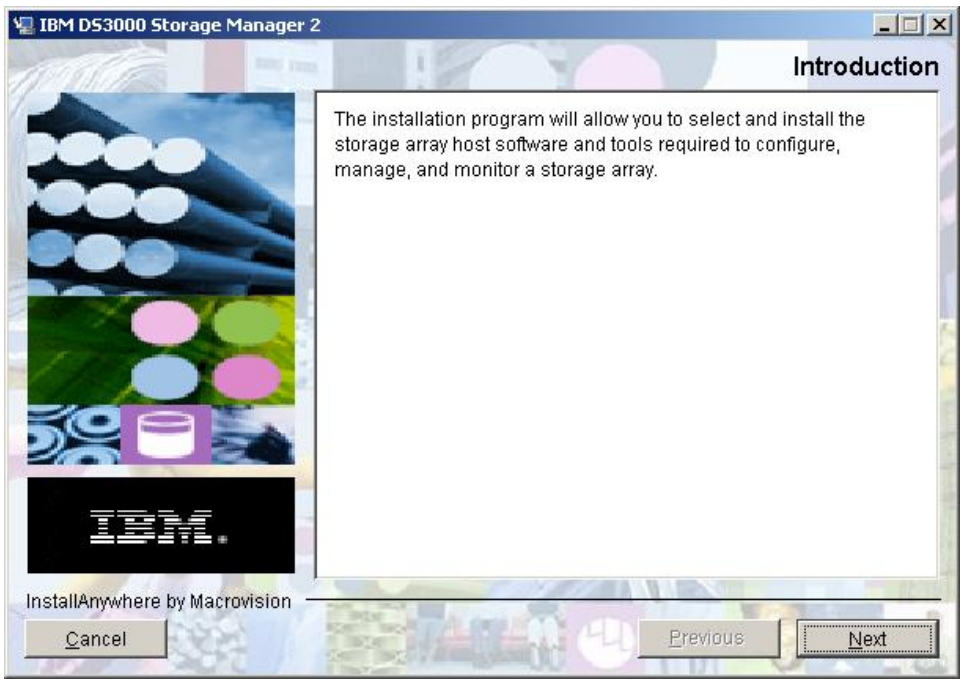

Figure 11: Install Screen of IBM DS3000 Storage Manager

To setup the iSCSI host Port IP Addresses, click on the iSCSI Tab in the DS3000 Storage Manager. You can set all four IP Addresses by selecting one after another in the "iSCSI host port:" scroll down button and then click finish. Note that all four host ports need different IP addresses. The factory default addresses for the iSCSI Host ports are:

- Controller 0 iSCSI Port 0: 192.168.130.101
- Controller 0 iSCSI Port 1: 192.168.131.101
- Controller 1 iSCSI Port 0: 192.168.130.102
- Controller 1 iSCSI Port 1: 192.168.131.102

In our example we used:

- 192.168.0.190 Controller 0 Port 1
- 192.168.0.191 Controller 0 Port 2
- 192.168.0.192 Controller 1 Port 1
- 192.168.0.193 Controller 1 port 2

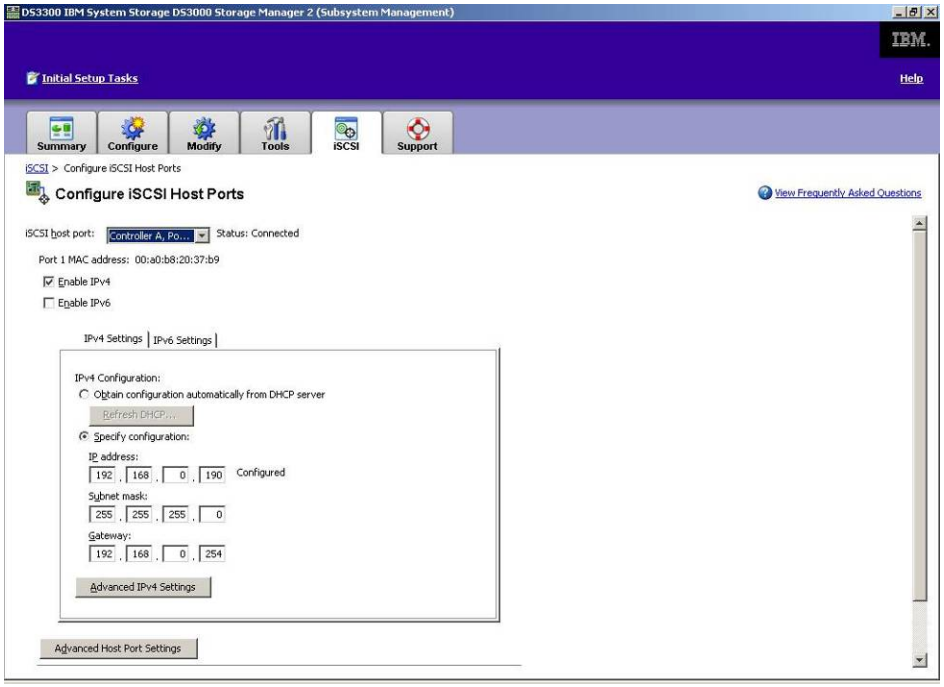

Figure 12: DS3000 Storage Manager iSCSI Host Port Configuration Screen

The “Advanced” button will let you choose the following options:

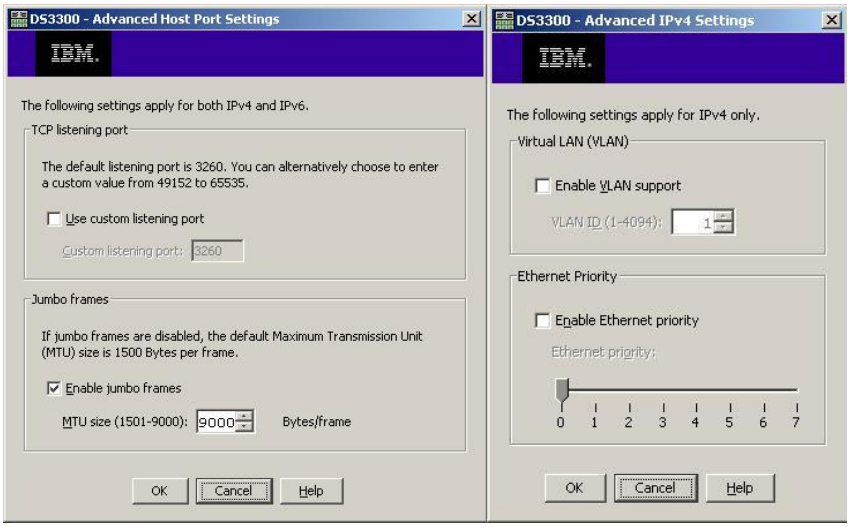

Figure 13: The DS3000 Storage Manager iSCSI Host Port Advanced Option Screen

VLAN

If a separate network is not feasible, as a minimum, configure a separate VLAN for the iSCSI traffic. Configure a VLAN within the switch and on the storage subsystem using the DS3000 Storage Manager software.

Ethernet Priority

Table 1 shows the general guidelines a network manager uses to determine network priorities.

Table 1: General Guidelines for Network Priorities

Priority	Description
0	Lowest priority
ran1-4	Ranges from loss-eligible traffic to controlled-load applications, such as streaming multimedia and business-critical traffic.
5-6	Delay-sensitive applications such as interactive video and voice
7	Highest priority reserved for network-critical traffic. Do not use with the DS3300.

Ethernet priority should be used in an isolated LAN environment where there are multiple hosts and systems. If you have one host and one system it shouldn’t be

necessary. It can also be used in a public LAN (with or without a VLAN). However, other LAN traffic *will* be impacted. If you use Ethernet priority, you need to make sure that the switches are capable of this option.

Jumbo Frames

Most enterprise gigabit Ethernet equipment provides some support for jumbo frames. Enabling jumbo frames has the following effects:

- Accelerates iSCSI performance by about five percent
- Reduces server CPU utilization by two percent to three percent with standard or smarter NICs

Because TCP off-load engine (TOE) cards or HBAs already perform off-loading, the CPU savings from jumbo frames is negligible when the frames are used with a TOE or HBA. However, jumbo frames should still accelerate performance.

When using jumbo frames, ensure that all of the devices on your iSCSI network—including switches, initiators and targets—are configured to use the same maximum frame size. Jumbo-frame sizes are not standardized and sizes can vary from 1501 bytes to 9000 bytes.

If your servers or DS3300 is set to a maximum frame size that is larger than your switches are set to, your DS3300 might appear to be working perfectly. However, if you start performing large data transfers that exceed the switch's maximum frame, disk drive I/O errors may occur.

Again, this is an area where tuning is required. It is recommended that if jumbo frames are utilized that you make sure to setup and configured in ALL devices.

After completing these steps the DS3300 should be visible to the attached servers.

The Host Install on Windows 2003 Server

We have configured the IP Addresses of the two NICs in the Server as the following.

- 192.168.0.198 Network Interface Card 1
- 192.168.0.199 Network Interface Card 2

The Microsoft iSCSI Initiator will be configured now

NOTE: The Microsoft iSCSI Software Initiator doesn't support "aliases" for the iSCSI initiator node name.

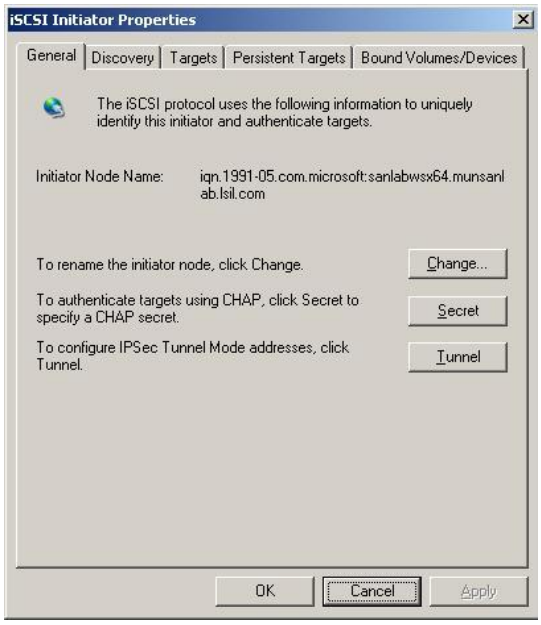

Figure 14: Microsoft iSCSI Initiator General Properties

CHAP is an optional iSCSI authentication method where the storage subsystem (target) authenticates iSCSI initiators on the host server. Two types of CHAP are supported:

Target CHAP – The storage subsystem authenticates all requests for access issued by the iSCSI initiator(s) on the host server using a CHAP secret.

Mutual CHAP – Both the storage subsystem and the iSCSI initiator authenticate each other. Mutual CHAP will impact performance.

NOTE: CHAP is an optional feature and is not required by iSCSI. However, without CHAP authentication, any host connected to the same IP network as the

storage subsystem can read from and write to the storage subsystem. It is recommended to set up CHAP on each with a different user and password on each to prevent access by other hosts.

If you are using CHAP for authentication, enter the Target secret here. Note that the secret needs to match the secret you entered in the DS3000 Storage Manager:

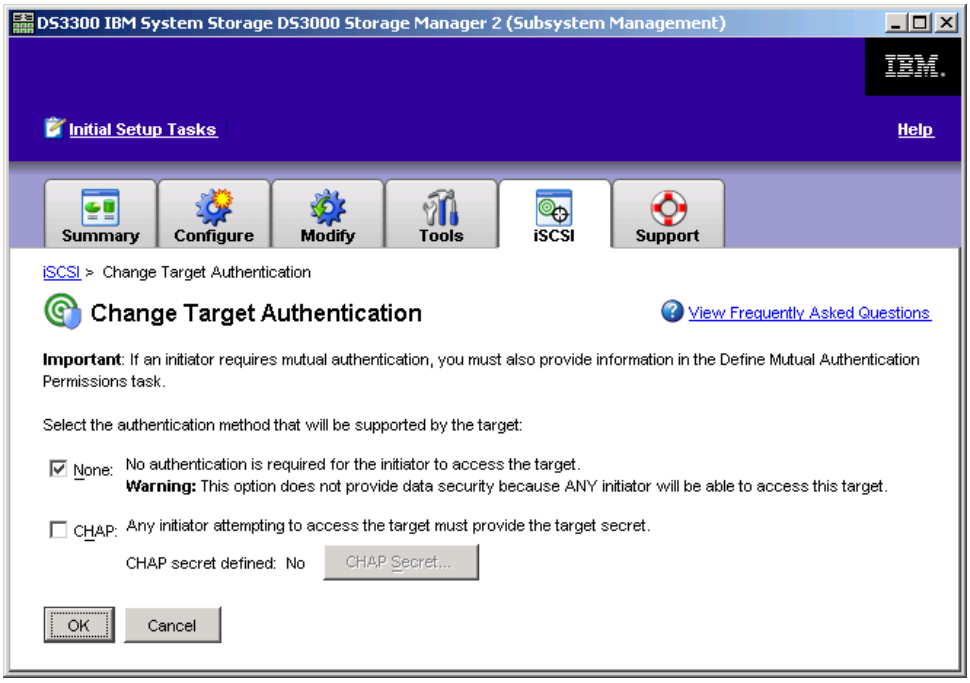

Figure 15: DS3000 Storage Manager iSCSI Target Authentication Screen

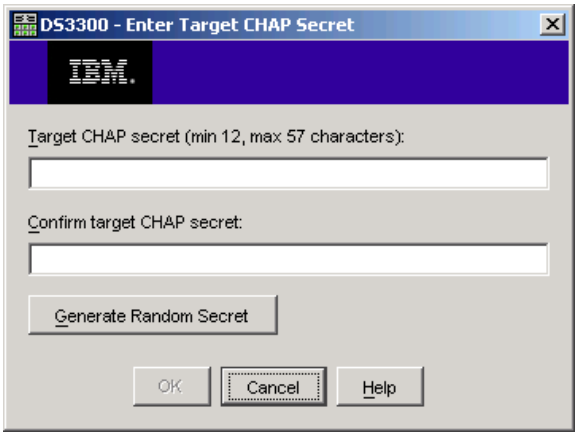

Figure 16: DS3000 Storage Manager CHAP Secret Input Screen

You can also generate a Random Secret by clicking the "Generate Random Secret" button. As with all passwords a longer and more complex password provides better security than a shorter and simpler one.

The following steps are important to make sure failover and load balancing will work correctly:

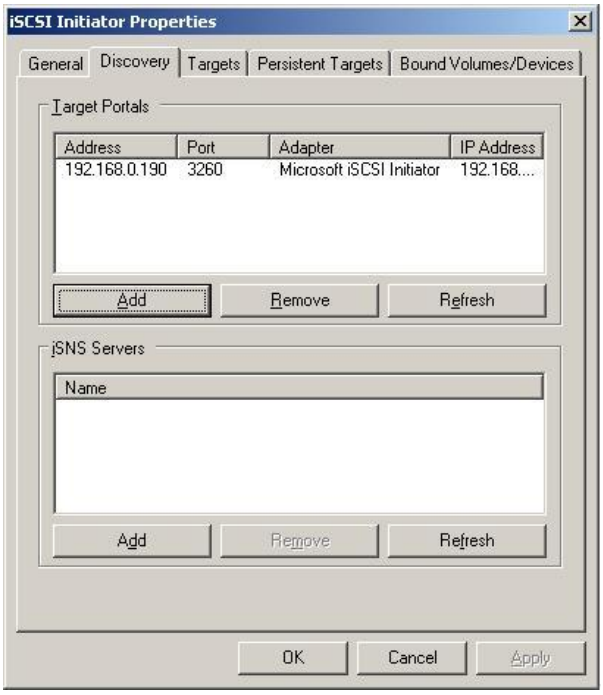

Figure 17: Add the first port of Controller 0 to the target portals.

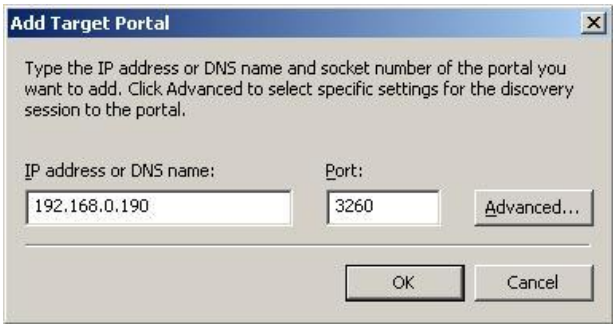

Figure 18: Select the advanced tab

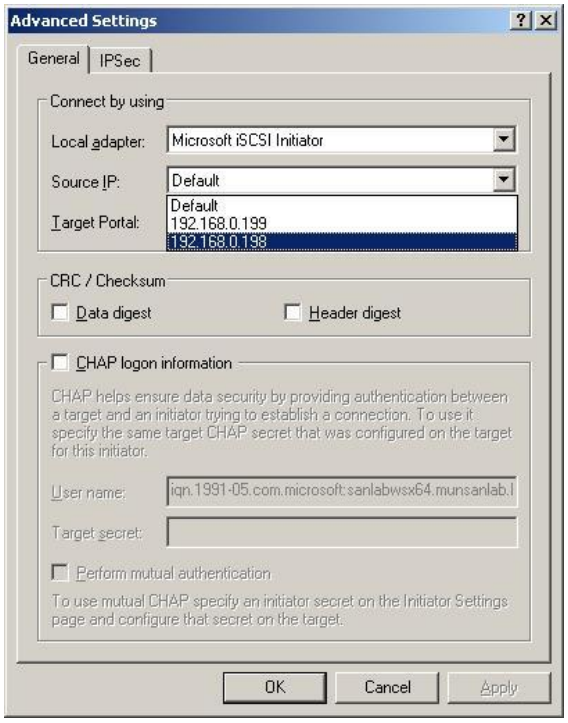

Figure 19: Select the Source IP and Target Portal

Select the local adapter as Microsoft iSCSI initiator. As source IP select the IP Address of NIC 1, as Target Portal the IP Address of Controller 0 Port 0. Target portals for both controller 0 and 1 will respond to the discovery session and should look like this:

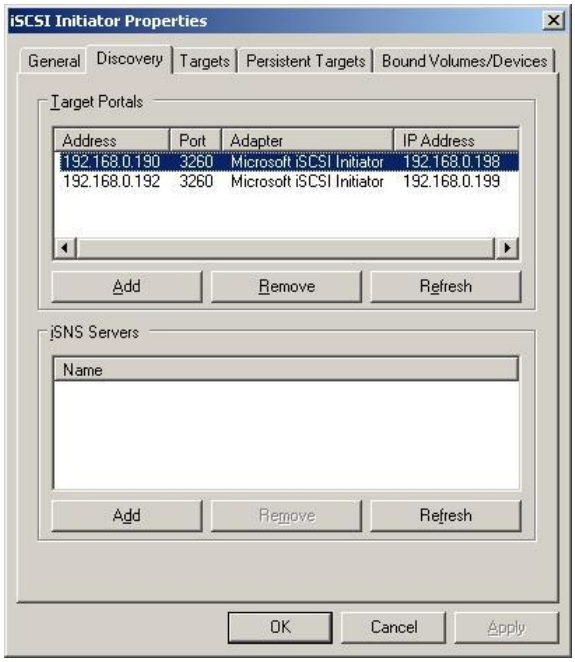

Figure 20: Microsoft iSCSI Initiator Discovery Properties

The next step is to log in to our selected target portals. This will enable us to actually use the DS3300 subsystem. Note that these steps are also very important to make sure failover and load balancing will work correctly.

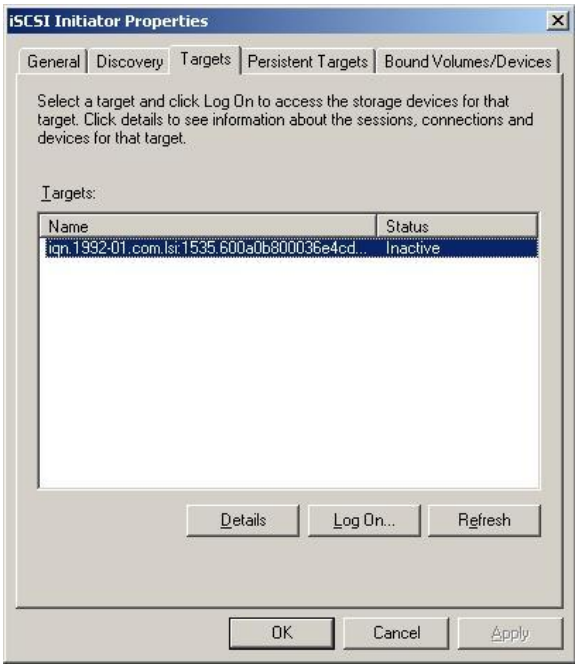

Figure 21: Click on Log on and enter the following

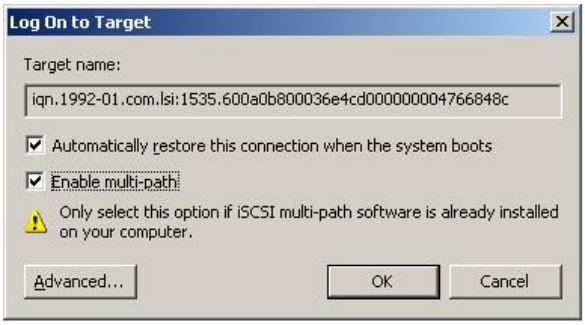

Figure 22: Confirm the checkbox "Enable multi-path" is checked. Click the advanced button

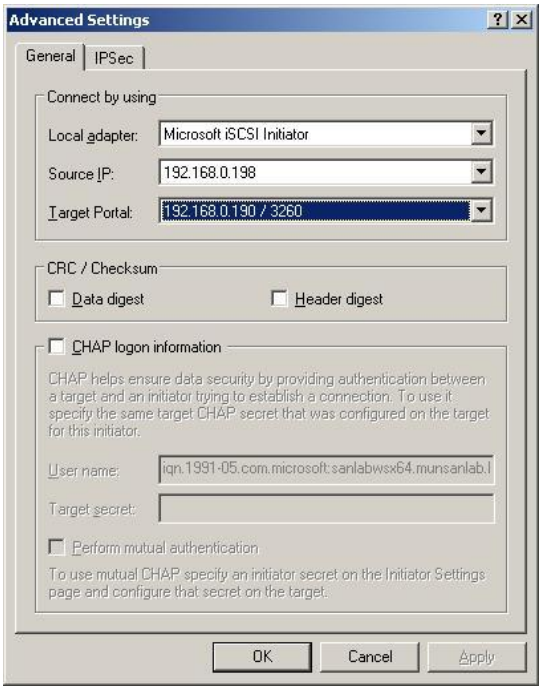

Figure 23: Select the Microsoft iSCSI Initiator as Local Adapter and use NIC1 as Source IP and Controller 0 Port 0 as Target Portal

NOTE: You need to logon twice: the second one will be the Source IP of NIC2 and as the Target Portal the IP Address of Controller 1 Port 0.

If you are finished a click on details you should see something like Figure 24:

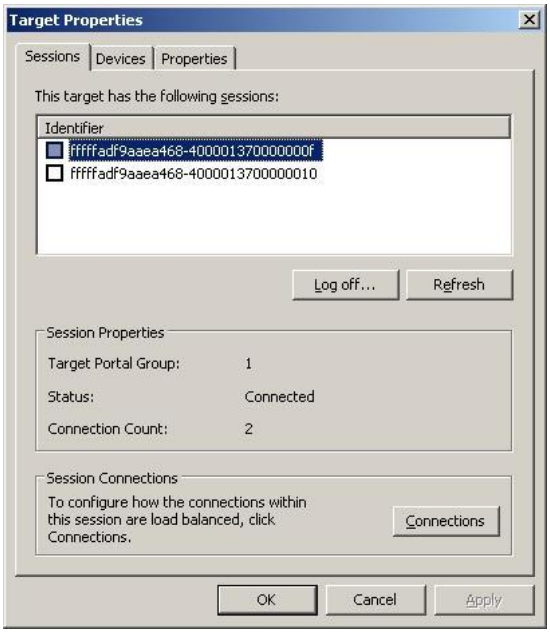

Figure 24: Target Properties Window

Select connections for each logon and add paths by clicking on the Add button on the bottom. Select the second NIC (192.168,0,199) as the Source and the Controller 0 Port 1 (192.168.0.191) as the Target. This ensures that both NICs have access to Controller 0. Repeat this for the other session to make sure that both NICs also see both Ports of Controller 1. It is recommended to keep the Load Balance Policy at Round Robin.

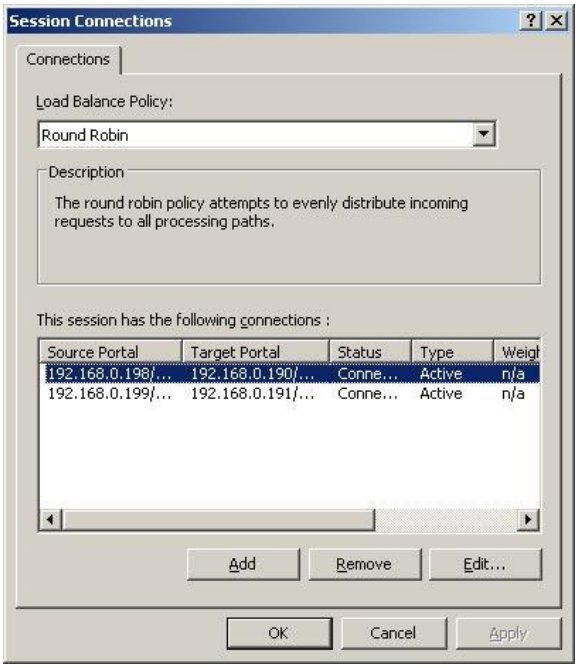

Figure 25: Session Connections Window

NOTE: The first logon using NIC1 to Controller 0 Port 0 should have as second path NIC2 to Controller 0 Port1. The second logon using NIC2 to Controller 1 Port 0 should have as second Path NIC1 to Controller 1 Port 1. This will enable load balancing between both NICs and the two ports of each Controller

Note that each logon uses two connections:

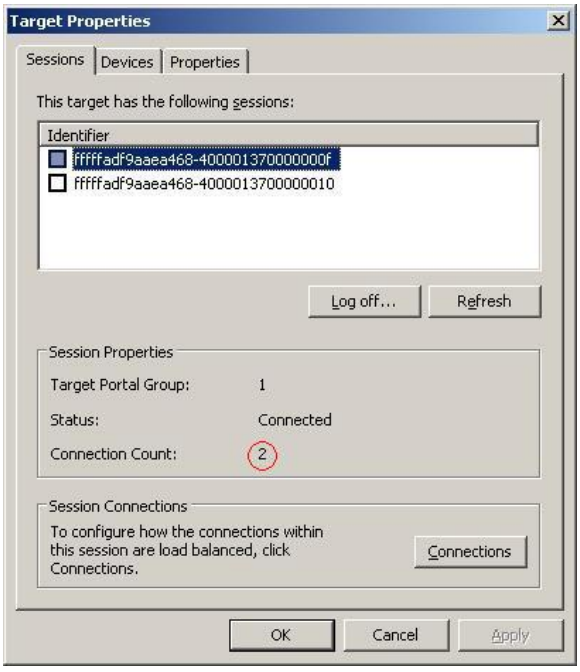

Figure 26: Target Properties Window with two Connections per session

After finishing this setup, click ok. You should see the DS3300 universal XPort in the Windows Device Manager:

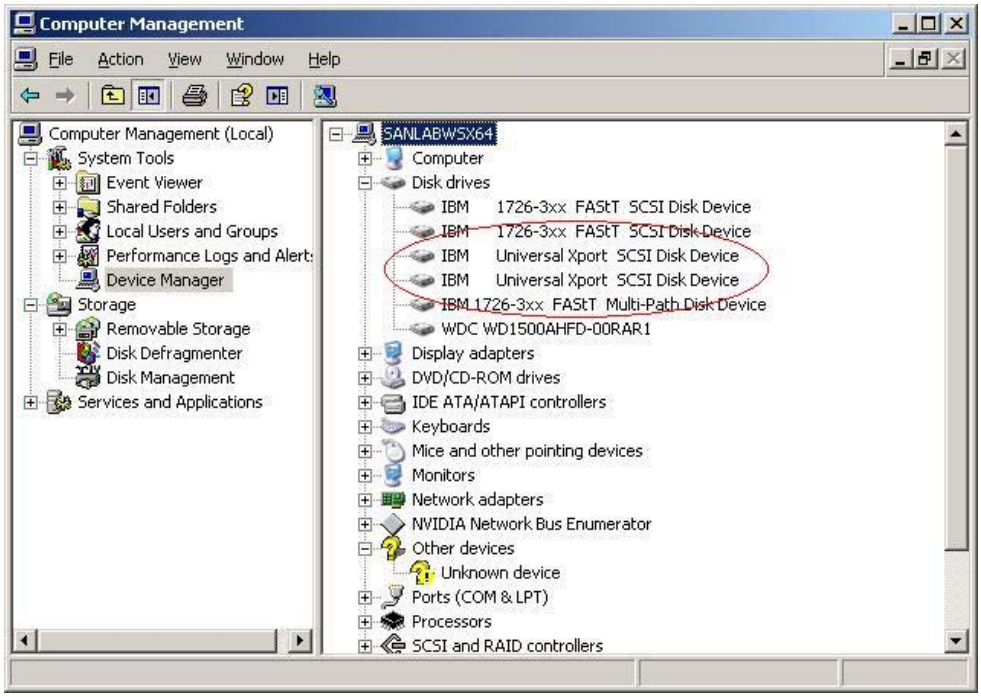

Figure 27: Windows Computer Management showing the two Universal XPort Devices

This indicates that both controllers are now visible by Windows. The DS3000 Storage Manager comes with a full context agent, which will submit all necessary information like iSCSI Initiator Node Name and OS version to theDS3300 so that the configuration of the logical drive masking is much easier. To enable the agent, we need to restart the Agent Service in the Computer Management.

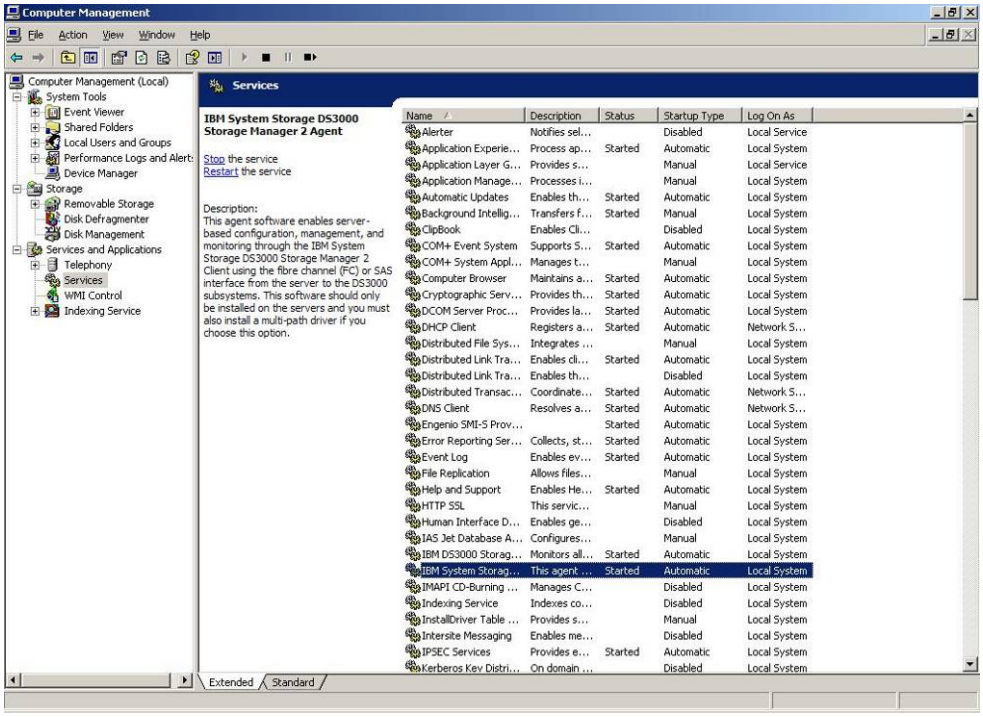

Figure 28: Windows Computer Management Services Screen

Choosing the Optimal DS3300 Configuration

After these initial steps, it is now time for the configuration of the DS3300 subsystem. Several items are important to consider when developing your configuration.

Disk Drive Types and Speeds

A single DS3000 storage subsystem can support both SAS and SATA drives. Each of these drive types offer different price and performance points and which one is best will depend upon your specific requirements. SATA drives offer high capacities and lower performance but at a better cost per megabyte. SAS drives offer higher level I/O performance and reliability but a higher cost per megabyte.

To ensure we are picking the right drive for our environment, we first need to look at the different performance requirements. There are two primary performance metrics – IOPS and MB/s:

IOPS Measures Random, Small-block I/O

Important for transaction-based applications such as OLTP, databases, Exchange, web servers, file servers, achieving high IOPS requires a streamlined controller design to handle large amounts of small I/Os.

Drive factors: Dependent on the # of drives – more drives, more IOPS. Drive-limited configurations can result in bottlenecks that create similar performances in systems with different maximum capabilities.

SAS vs. SATA – With disk drive drives some of the key performance enablers are seek time, latency, rotational velocity, command queuing – in all cases SAS drives have the advantage over SATA drives.

MB/s Measures Sequential, Large-block I/O

Important for throughput-intensive applications such as video, seismic processing, high performance computing (HPC), throughput rates are heavily dependent on the internal controller bandwidth and the number of host/drive ports.

Drive factors: Maximum rates can typically be reached with a smaller number of drives as compared to maximum I/O rates.

SAS vs. SATA – Again some of the key performance enablers for disk drive drives are data transfer rate, maximum I/O size, command queuing – in these areas there is only a slight SAS advantage.

Performance-oriented applications – **IOPS and MB/s** – will be best served by SAS drives although SATA may be sufficient for throughput applications

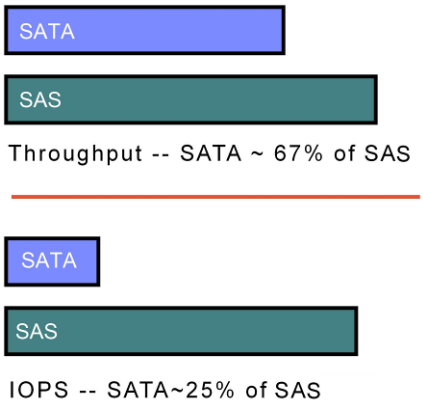

Figure 29: Performance Comparisons between SAS and SATA Drives

Based on these performance results, Exchange, Oracle, OLTP, web applications and most file-services, SAS drives will provide the optimal outcome. And as Exchange, Oracle and OLTP have the highest IOPs requirements of these applications, they are better served by using 15K SAS drives

Archiving, backup to disk drive and some file services may find SATA drives adequate as the requirements with these applications typically require sequential large block I/O rather than small random I/O. With SATA drives, high capacity can be achieved at a more desirable price/performance point.

RAID Levels

Choosing the right RAID Level is very critical when it comes to data protection and performance. In this Guide, the implications of RAID5 vs. RAID 1 (or RAID 10) will be discussed.

Small I/O's are generally used with RAID 5, RAID 6, RAID 1 or RAID 1/0

Large I/O's are generally used with RAID 3 or RAID 5, or RAID 6

What is a small/large I/O size?

- Small I/O < 32K
- Large I/O > 256K
- Grey area >= 32K and <= 256K

RAID 1, 1/0, 5 and 6 benefit from concurrent I/O's

Number of disk drives to use with RAID

High bandwidth applications

RAID 5 in 8+1 drive configurations

RAID 6 in 8+2 drive configurations

High I/O rate Applications

RAID 5 in 4+1 or 8+1 configurations

RAID 6 in 4+2 or 8+2 configurations

RAID 1/0 striped in multiple of 2 drives

More spindles will provide higher I/O rates

Capacity requirements will limit the number of spindles

When creating striped logical drives, select the appropriate stripe unit size. Always select an even stripe width I/O when possible.

Even stripe width I/O calculation:

Application I/O = (# of drives * segment size).

e.g. The application I/O size = 1MB. Now the stripe size of the logical drive should match the application I/O size. Therefore if we use a RAID5 4+1 the number of data drives is 4. Therefore, the segment size should be 256K to match the application I/O size.

Logical Drive Management with the DS3300

In this section we will create a logical drive with DS3000 Storage Manager and map it to our previously configured server. We will also take a look at how we can use the DS3300 dynamic features to expand a logical drive online and make the additional space visible with Microsoft Windows 2003.

Configuring the Host Access

Before we start creating a logical drive, we want to make sure that our Windows 2003 server is already configured within the DS3300. This is needed for partitioning which allows you to restrict which hosts will have access to a logical drive.

You should use the automatic Host Access Configuration. As the full context agent already submitted all the information, the screen should look like the following. Select your Host and use the "add" button to move the Host to the "selected Hosts" section. Confirm your selection with OK. Now the Host is configured and ready to use for the DS3300 partitioning. If you select the "view details" you will shown the iSCSI Initiator Node Name of the Server, as well as the OS.

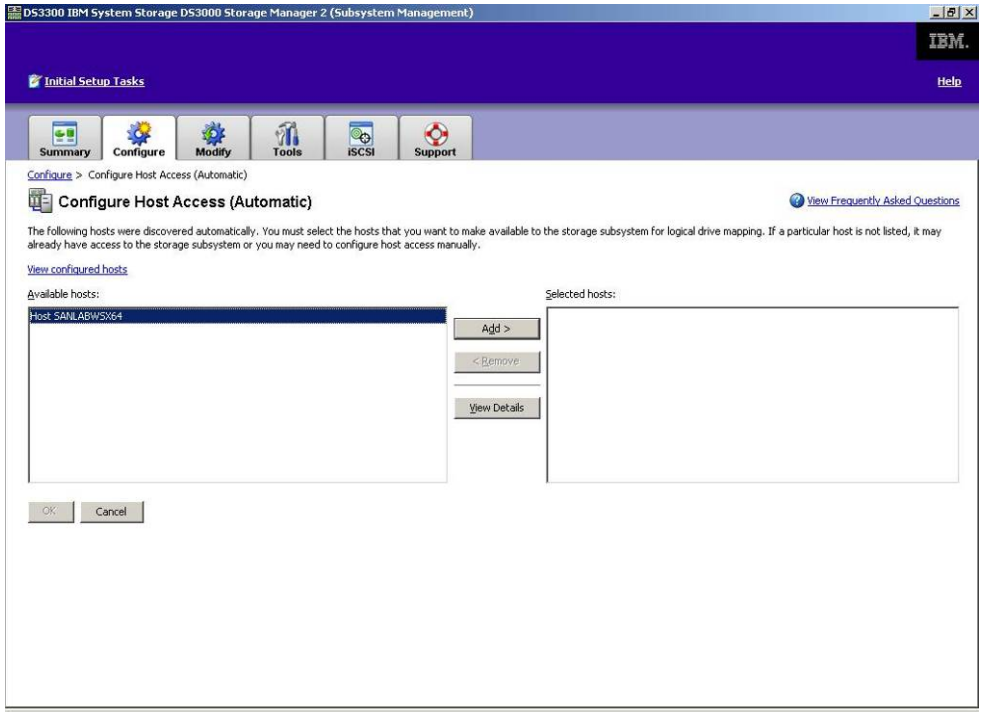

Figure 30: DS3000 Storage Manager Host Access Configuration Screen

Creating a Logical Drive

We will use the DS3000 Storage Manager to create our logical drive.

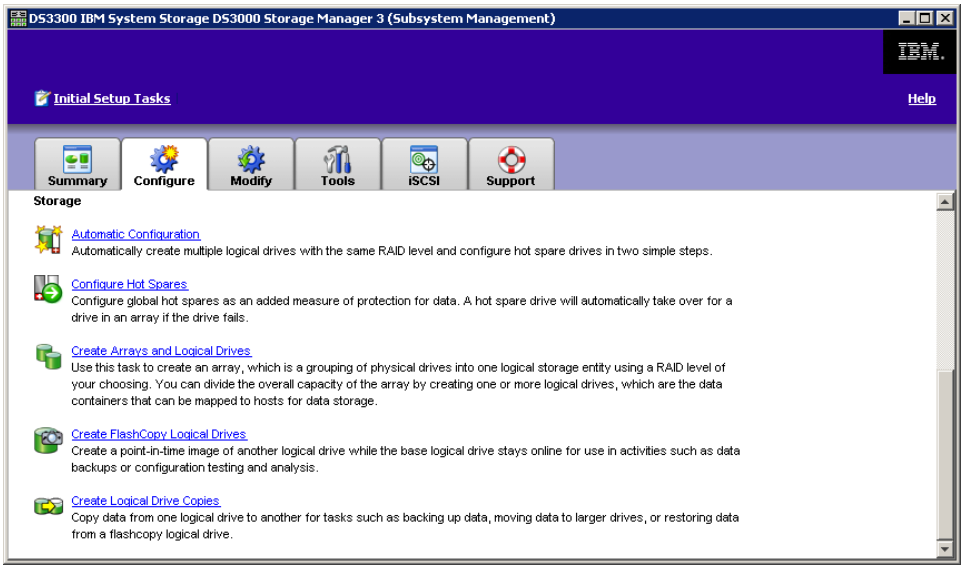

Figure 31: DS3000 Storage Manager Create Arrays and Logical Drives Configuration Screen Part 1

In this section we will choose the appropriate RAID level and the numbers of drives for our logical drive. In our case we will stick with the recommendation of a 4+1 RAID 5.

The next section is where you specify the size of the logical drive in the just created 4+1 RAID 5. You can create up to 32 logical drives in a single RAID 5 array. However in high performance applications it is recommended not to create more than three logical drives. This is due to the possible result in disk drive heads thrashing and a decrease in logical drive performance.

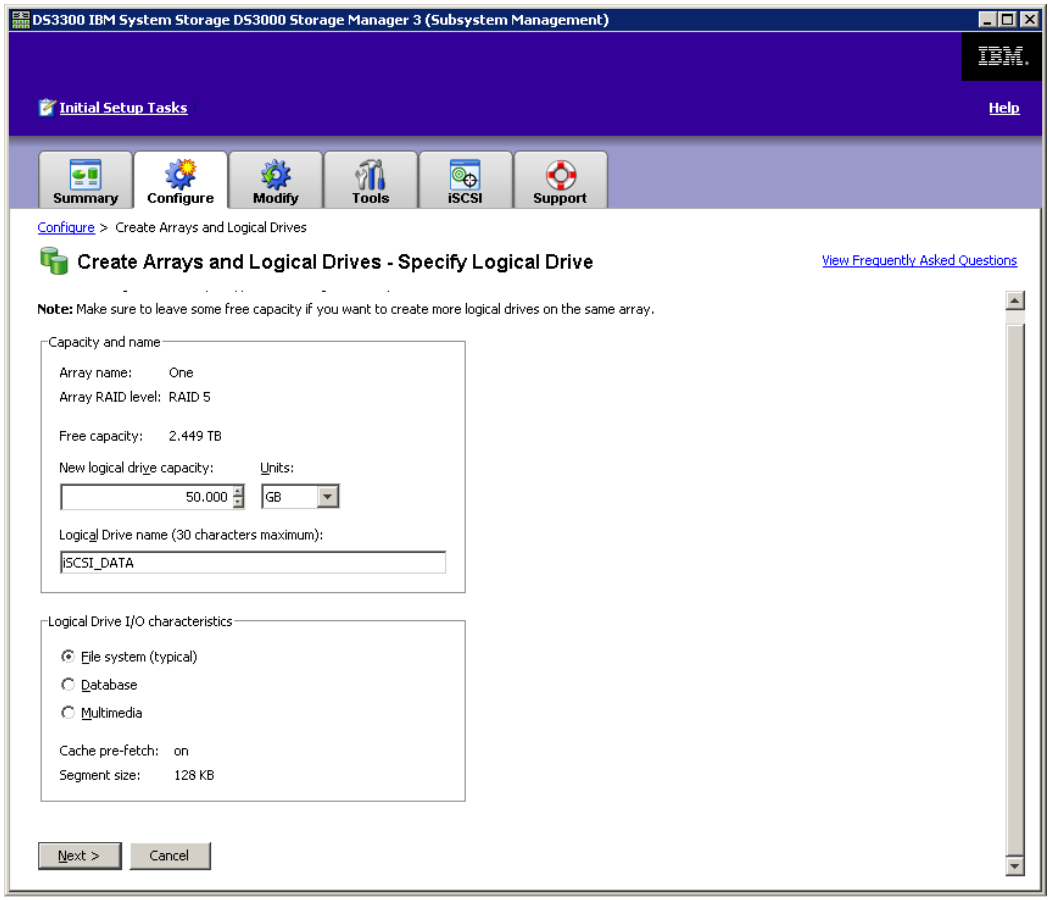

Figure 32: DS3000 Storage Manager Logical Drive Configuration Screen Part2

Additionally you can specify some of the logical drive IO characteristics. This will modify some parameters regarding the segment size and cache settings for this logical drive. If you should later wish to change any of these parameters, they can be modified using the DS3000 Storage Manager command line interface. Dependent on your application, you should choose File System (general purpose), Database (such as Exchange, SQL, Oracle) or Multimedia (Audi/Video, large sequential IO).

The next Window will ask us which Server should be accessing the logical drive. Select the Server already configured and assigns the LUN 0.

NOTE: Start with the LUN 0 and use consecutive LUNs for each server. There are some OS's that won't see any LUNs which are non consecutive (Linux for example.)

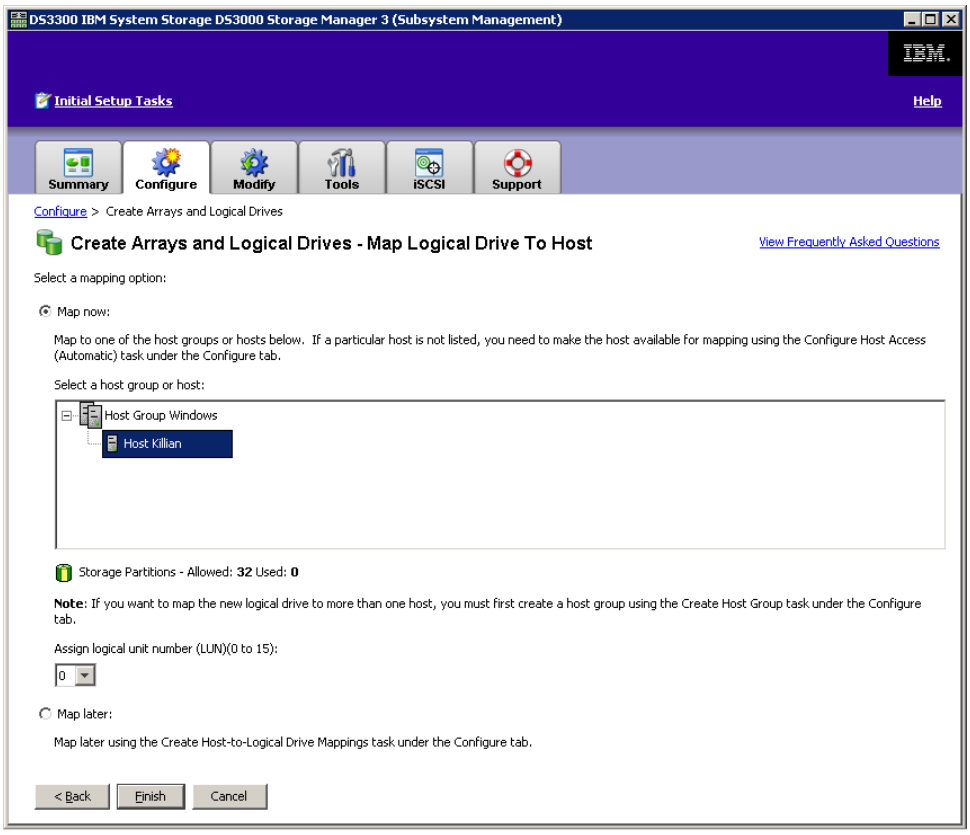

Figure 33: DS3000 Storage Manager Logical Drive Configuration Screen Part 3

The logical drive is now available to our Windows 2003 Server. If you open up the Windows disk drive Management, you will now see the new 50GB DS3300 logical drive.

Right click on a drive and start the Windows Initialization of the drive. Do not upgrade this drive to a dynamic drive. We recommend only using basic drives in an iSCSI environment. We will later show how we can expand a basic drive using the Microsoft utility disk drivepart.

After initializing the logical drive, we will create a partition and format the logical drive with a file system.

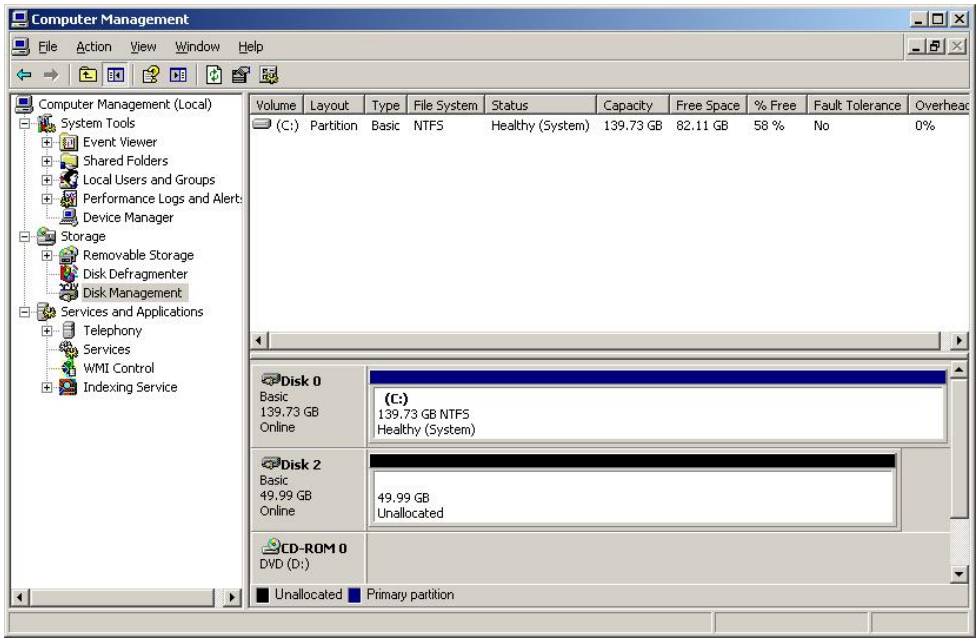

Figure 34: Windows Computer Management Screen Disk drive Management Window

Now right click on the unallocated space to create the partition. Follow the Wizard until you come to the section where Windows asks you to format it

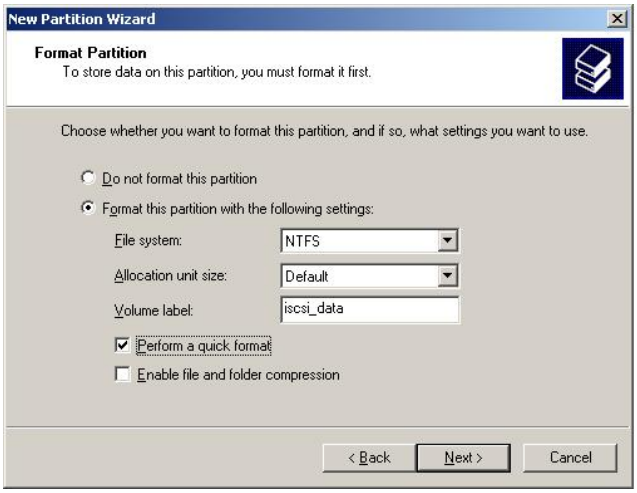

Figure 35: Windows New Partition Wizard – Format Partition Window

Now the logical drive is ready to use and we can copying data to it.

Helpful hints for using the DS3300 on Windows

This section will give you some hints on how to verify that load balancing and failover is working correctly on the Windows 2003 Server.

Load-Balancing and Failover

The simplest way to check for load balancing in the fully redundant configuration is to copy data onto the logical drive and open the Windows Task manager and see that the I/O is equally balanced between the two NICs:

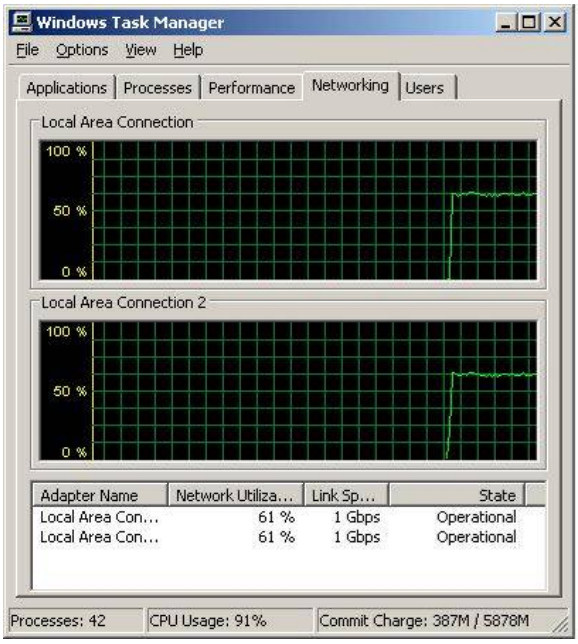

Figure 36: Windows Task Manager – Networking TAB

NOTE: As Microsoft MPIO will do the load balancing, dsmutil will only show one path per controller.

For failover there is a utility with the DS3000 Storage Manager MPIO called dsmutil and it is located in the following directory (on x64 Servers):

```
C:\Program Files (x86)\DSMDrivers\ds4dsm>dsmutil
```

This utility has several options:

```
-e error_level

-g target_id

-s ["failback" | "preferred" | "avt" | "busscan"]
```

We will use the "-g" option. As our DS3300 is currently the only storage subsystem attached, it should have target_id 0. The output of dsmutil –g0 should be the following:

```
Hostname    = sanlabwsx64
Domainname  = munsanlab.lsil.com
Time        = GMT Tue Feb 12 09:53:20 2008
MPP Information:
----------------
      ModuleName: DS3300                                SingleController: N
 VirtualTargetID: 0x000                                    ScanTriggered: N
     ObjectCount: 0x000                                       AVTEnabled: N
             WWN: 600a0b800036e4cd000000004766848c             RestoreCfg: N
    ModuleHandle: 0xFFFFFADF9AB1F920                        Page2CSubPage: Y
 FirmwareVersion: 6.70.4.0
   ScanTaskState: 0x00000000
Controller 'A' Status:
-----------------------
ControllerHandle: 0xFFFFFADF9A93CC10                   ControllerPresent: Y
    UTMLunExists: Y (007)                                         Failed: N
   NumberOfPaths: 1                                     FailoverInProg: N
                                                           ServiceMode: N

    Path #1
    ---------
 DirectoryVertex: 0xFFFFFADF9875C908                           Present: Y
       PathState: OPTIMAL
 P08P00I00
Controller 'B' Status:
-----------------------
ControllerHandle: 0xFFFFFADF9AAE5120                   ControllerPresent: Y
    UTMLunExists: Y (007)                                         Failed: N
   NumberOfPaths: 1                                     FailoverInProg: N
                                                           ServiceMode: N

    Path #1
    ---------
 DirectoryVertex: 0xFFFFFADF9875CA70                           Present: Y
       PathState: OPTIMAL
 P08P00I01
Lun Information
---------------
    Lun #0 - WWN: 600a0b800036e4cd00000f9a47affbca
    ----------------
       LunObject: 0x0                                  CurrentOwningPath: A
  RemoveEligible: N                                       BootOwningPath: A
   NotConfigured: N                                      PreferredPath: A
        DevState: OPTIMAL                          NeedsReservationCheck: N
                                                              TASBitSet: Y
                                                               NotReady: N
                                                                   Busy: N
                                                              Quiescent: N
    Controller 'A' Path
    --------------------
   NumLunObjects: 1                                    RoundRobinIndex: 0
         Path #1: LunPathDevice: 0xFFFFFADF9C1E2290
                      IoCount: 0
                     DevState: OPTIMAL
                  RemoveState: 0x0  StartState: 0x0  PowerState: 0x0
    Controller 'B' Path
```

```
    --------------------
   NumLunObjects: 1                                    RoundRobinIndex: 1
         Path #1: LunPathDevice: 0xFFFFFADF98793410
                   IoCount: 0
                  DevState: OPTIMAL
                 RemoveState: 0x0   StartState: 0x0   PowerState: 0x0
```

This output is very complex, but it will provide some helpful information about our path status. We will briefly discuss the most important sections:

```
Controller 'A' Status:
----------------------
ControllerHandle: 0xFFFFFADF9A93CC10                ControllerPresent: Y
    UTMLunExists: Y (007)                                     Failed: N
   NumberOfPaths: 1                                    FailoverInProg: N
                                                          ServiceMode: N

    Path #1
    ---------
 DirectoryVertex: 0xFFFFFADF9875C908                          Present: Y
       PathState: OPTIMAL
 P08P00I00
```

This indicates that we have one path to Controller 0 that is present and optimal. The same applies to our Controller 1.

As we go into the LUN level we should see the same results:

```
Lun #0 - WWN: 600a0b800036e4cd00000f9a47affbca
    ----------------
       LunObject: 0x0                                CurrentOwningPath: A
  RemoveEligible: N                                     BootOwningPath: A
   NotConfigured: N                                      PreferredPath: A
        DevState: OPTIMAL                          NeedsReservationCheck: N
                                                              TASBitSet: Y
                                                               NotReady: N
                                                                   Busy: N
                                                              Quiescent: N
    Controller 'A' Path
    --------------------
   NumLunObjects: 1                                    RoundRobinIndex: 0
         Path #1: LunPathDevice: 0xFFFFFADF9C1E2290
                   IoCount: 0
                  DevState: OPTIMAL
                 RemoveState: 0x0   StartState: 0x0   PowerState: 0x0
    Controller 'B' Path
    --------------------
   NumLunObjects: 1                                    RoundRobinIndex: 1
         Path #1: LunPathDevice: 0xFFFFFADF98793410
                   IoCount: 0
                  DevState: OPTIMAL
                 RemoveState: 0x0   StartState: 0x0   PowerState: 0x0
```

LUN 0 has path to both Controllers and both are OPTIMAL. This information confirms that the logical drive is optimally configured for failover.

Expanding a Logical Drive with Windows Basic Disk drives

We will now look how daily data growth can be handled with the online features of the IBM System Storage DS3300 and features of Microsoft Windows. Microsoft's description on how the dispart utility works can be found at: http://support.microsoft.com/kb/325590.

The DS3300 includes standard the capability of expanding a logical drive with host I/O still active. This can be done with the DS3300 CLI utility. Here is an example of the syntax to expand the logical drive we created earlier:

```
C:\Program Files (x86)\IBM_DS3000\client>smcli -n DS3300 -c "set logicaldrive
["Iscsi_Data"] addcapacity=10 GB;"

Performing syntax check...

Syntax check complete.

Executing script...

Script execution complete.

SMcli completed successfully.

C:\Program Files (x86)\IBM_DS3000\client>
```

We have chosen to expand the capacity by 10GB. To add a different amount of free capacity from the array, change the section in the add capacity field to the desired value.

NOTE: If the free capacity in your array is insufficient, you can use the DS3000 Storage Manger to add additional drives to the group and increase the available free capacity. This not only increases the available capacity but can increase the performance as well as the controller will restripe the data across the additional drives. However this process may take a while and it cannot be stopped once started.

After the operation is complete, use the Windows utility disk drivepart to extend this basic disk drive. Before using this command, rescan the disk drives in the Windows Disk drive Management Console to ensure that the host now sees the added capacity.

```
C:\>disk drivepart

Microsoft Disk drivePart version 5.2.3790.3959

Copyright (C) 1999-2001 Microsoft Corporation.

On computer: SANLABWSX64

DISK DRIVEPART>
```

A logical drive needs to be selected before we can start. However this time we need to select the associated logical drive on the logical drive, as we want to extend the partition and the file system. To find out which logical drive we need to select, use the list logical drive command:

```
DISK DRIVEPART> list volume
```

```
  Volume ###  Ltr  Label        Fs     Type        Size     Status     Info
  ----------  ---  -----------  -----  ----------  -------  ---------  --------
  Volume 0     E   iscsi_Data   NTFS   Partition     50 GB  Healthy
  Volume 1     D                CDFS   DVD-ROM      422 MB  Healthy
  Volume 2     C                NTFS   Partition    140 GB  Healthy    System
```

We see that Volume 0 is our 50GB iSCSI logical drive. We will now select this disk drive:

```
DISK DRIVEPART> select volume 0

Volume 0 is the selected volume.
```

To extend this logical drive, simply type “extend”:

```
DISK DRIVEPART> extend

Disk drivePart successfully extended the volume.
```

To check if the logical drive has grown, use the list logical drive command again:

```
DISK DRIVEPART> list volume

  Volume ###  Ltr  Label        Fs     Type        Size     Status     Info
  ----------  ---  -----------  -----  ----------  -------  ---------  --------
* Volume 0     E   iscsi_Data   NTFS   Partition     60 GB  Healthy
  Volume 1     D                CDFS   DVD-ROM      422 MB  Healthy
  Volume 2     C                NTFS   Partition    140 GB  Healthy    System

DISK DRIVEPART>
```

Also check with the Windows Disk drive Management if the disk drive has grown:

The logical has successfully expanded.

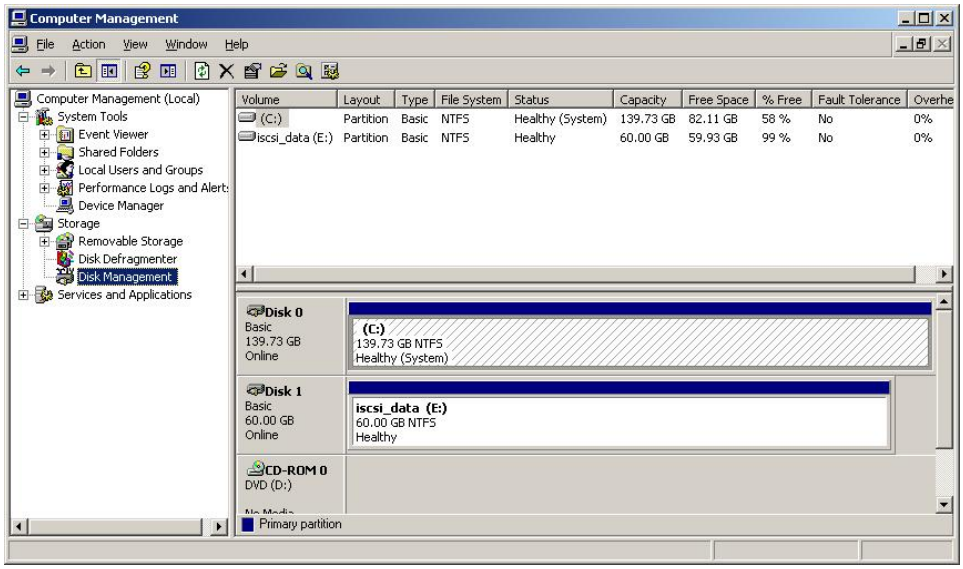

Figure 37: Windows Disk drive Management Window

Installing the DS3300 in an VMware ESX 3.5 Environment

The following section describes the necessary steps to setup the IBM DS3300 iSCSI storage subsystem with VMware ESX Server Version 3.5.

Before beginning, refer to the VMware ESX SAN Support matrix, which can be found at:

http://www.vmware.com/resources/compatibility/pdf/vi_san_guide.pdf

When planning to boot from a SAN it is important to research the components and operating systems involved. When using a QLogic iSCSI HBA or any other brand HBA you should check the manufacturer's website for information and firmware updates required to boot from a SAN. Additionally, you should check the operating system website for boot-from-SAN information. Operating system sites such as Microsoft and VMware will post boot-from-SAN information and patches. When searching these websites suggested keywords might be: boot-from-SAN, QLogic, VMware, or Windows 2008. These are only suggestions and not the full set of search words or vendors.

There are different possibilities to use the DS3300 storage array in a VMware environment:

- Use the DS3300 logical drive with the ESX Server and a vmfs File System
- Use the DS3300 logical drive with the ESX Server and a raw device mapping
- Use the DS3300 logical drive with the Virtual Machine (Install the Software iSCSI Initiator on the Virtual Machine and bypass the ESX)

Every different use has its advantages and disadvantages. The easiest way to use the DS3300 is to use the logical drive with the ESX Server and a vmfs file system. This is required for the boot partition of a virtual machine. It makes special functions of VMware very easy, such as VMotion. But the content can only be read from an ESX Server. Using a “raw-device mapping” allows you to allocate space through VMware to the Virtual Machine without adding a vmfs to the Logical Drive. This can be useful, if the data on the logical drive needs to be read by stand-alone Servers. However all the multipathing functionality and maintenance stays with ESX. Bypassing the ESX and mapping the logical drive directly into the virtual machine by installing the Software iSCSI Initiator directly into the virtual machine can leverage the built in features of the Software iSCSI Initiator (e.g. multipathing) but required a storage partition for every virtual machine. As the DS3300 is only capable of a maximum of 32 partitions, this can be a restriction as well as the more difficult implementation of VMware features such as VMotion. Moreover the logical drive can be easily moved between virtual machines and physical servers.

Before the planning, you should carefully think about those options and decide which is best for your environment. Applications such as databases or mail systems are best with raw-device mappings or direct connects into the virtual machines as the underlying configuration can be better optimized for those applications. Regular file services can benefit from the vmfs features, as well as optimized resource allocation.

The best configuration from an iSCSI network standpoint and the only under VMotion supported configuration is the fully redundant configuration:

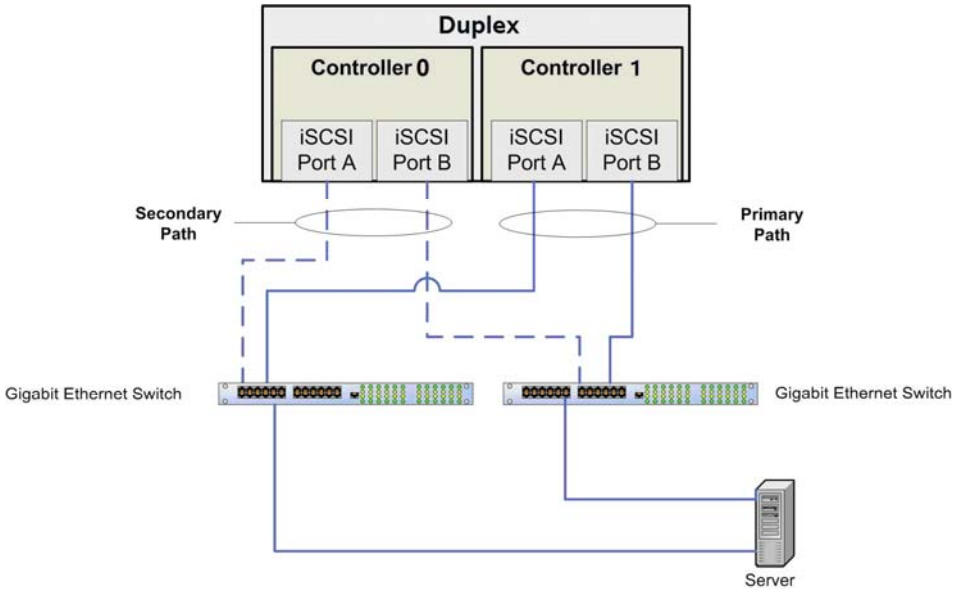

Figure 38: VMware fully redundant iSCSI SAN configuration

Setup the VMware iSCSI Software Initiator

Logon to VMware Infrastructure Client:

Figure 39 : VMware Infrastructure Client Log-on

This brings you to the VMware start screen :

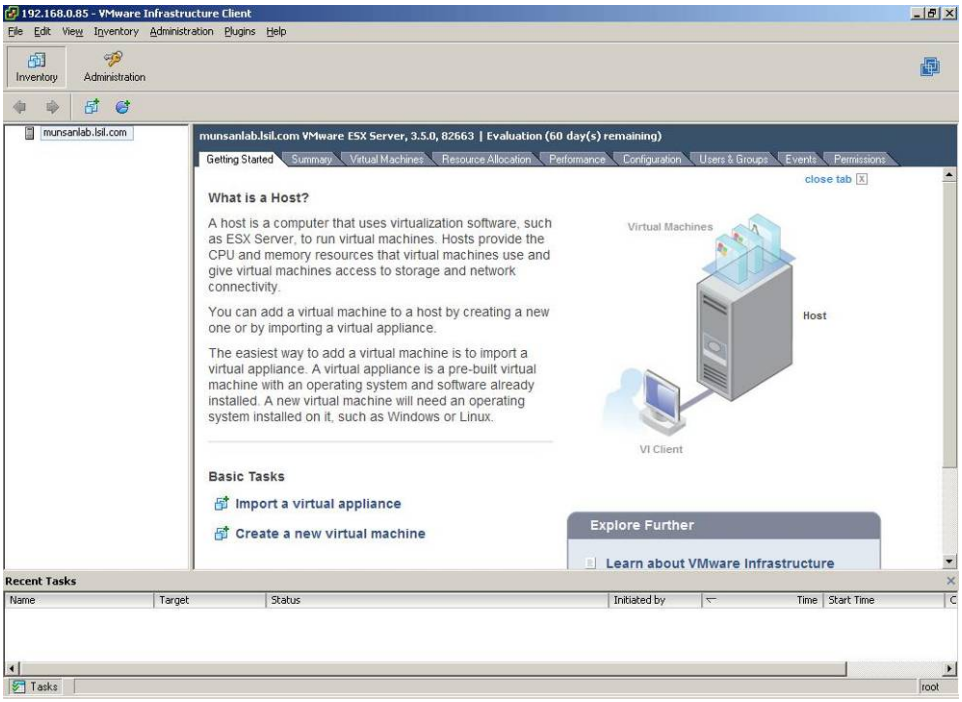

Figure 40: VMware ESX 3.5 Start Screen

Select Networking in the Configuration Tab. We need to make sure, that the VMkernel is installed. If not, click on the *Add Networking* in the upper right corner:

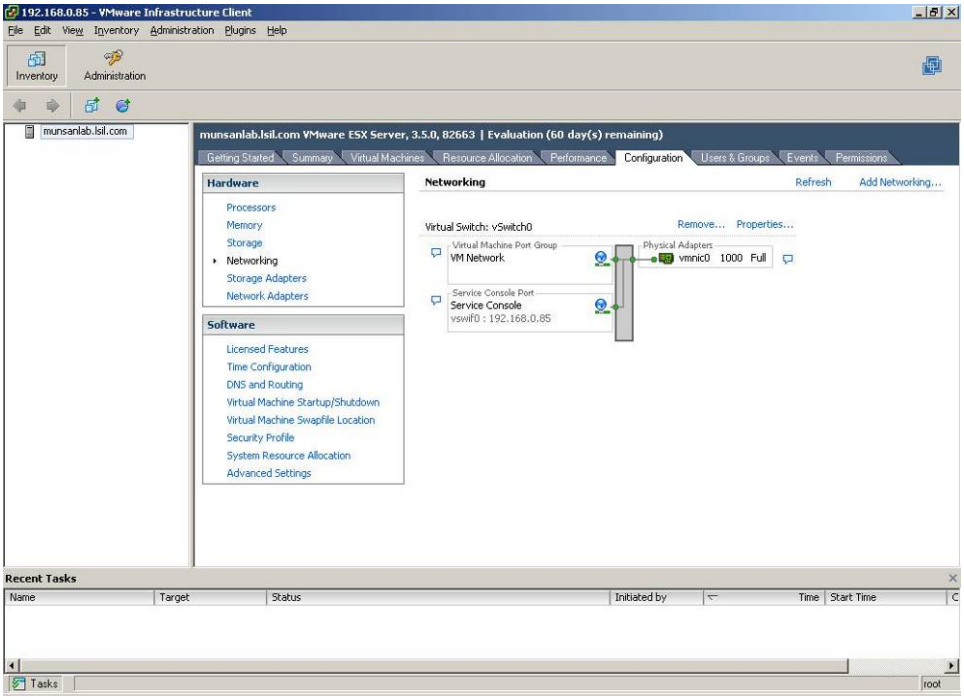

Figure 41: VMware ESX Networking Configuration Screen

The Add Network Wizard provides VMkernal guidance setup. Select VMkernel:

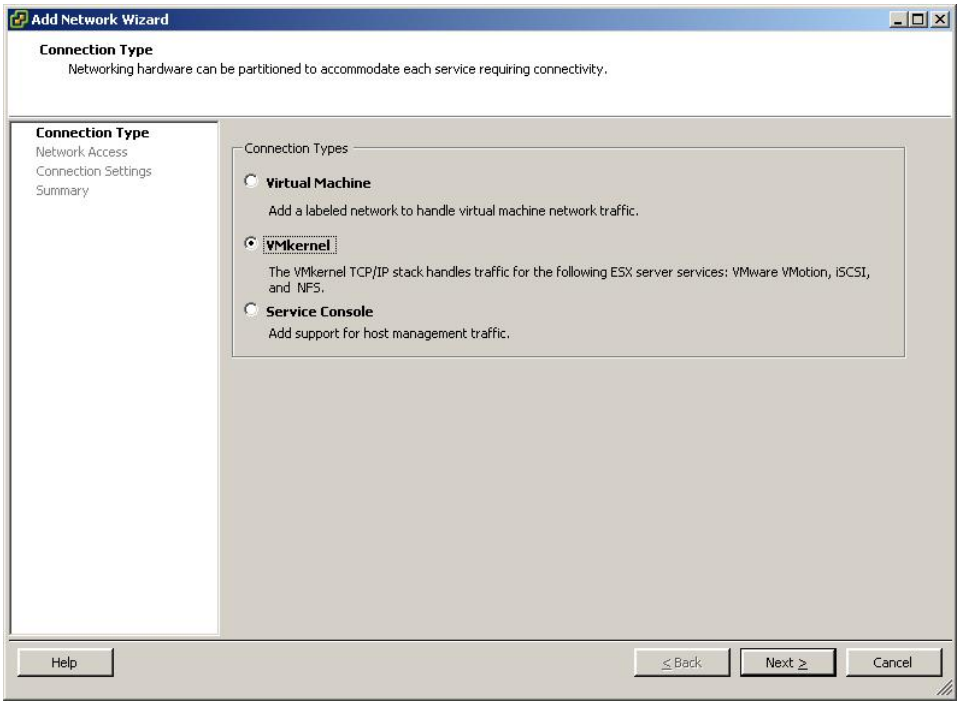

Figure 42: Add Network Wizard

Then select the virtual switch you want the VMkernel to connect through:

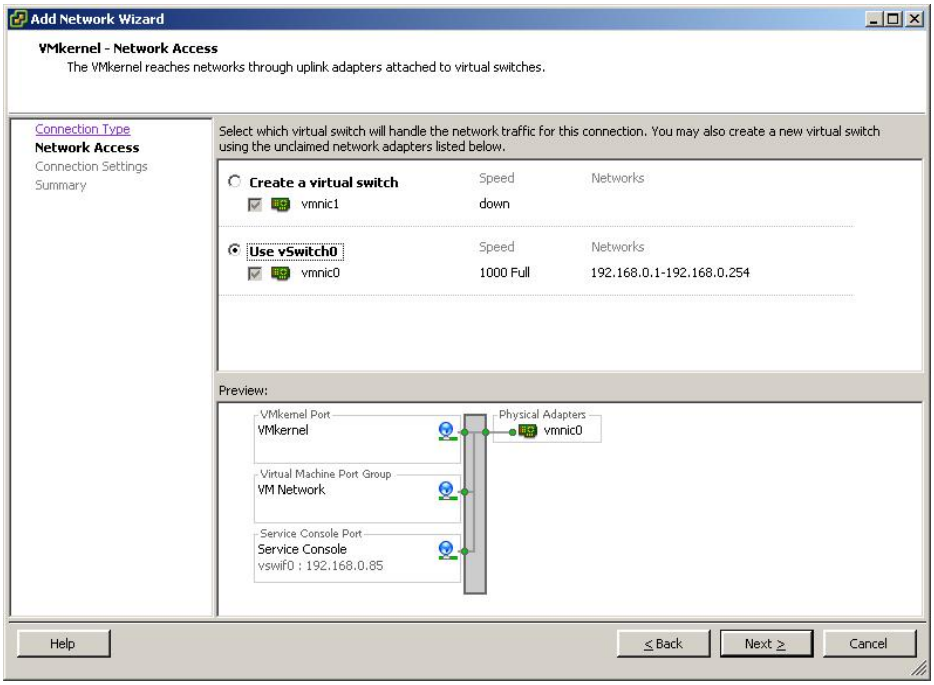

Figure 43: VMkernel Network Access Screen

Now enter the IP-Address of the VMkernel:

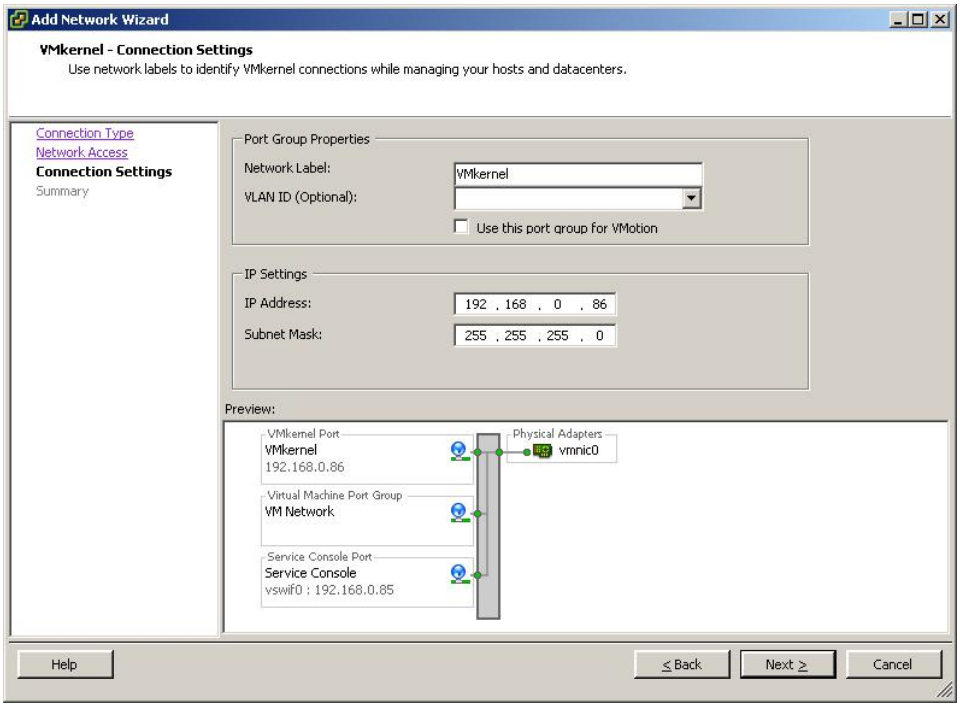

Figure 44: VMkernel – Connection Settings

The last screen of provides an overview of your settings. Click finish.

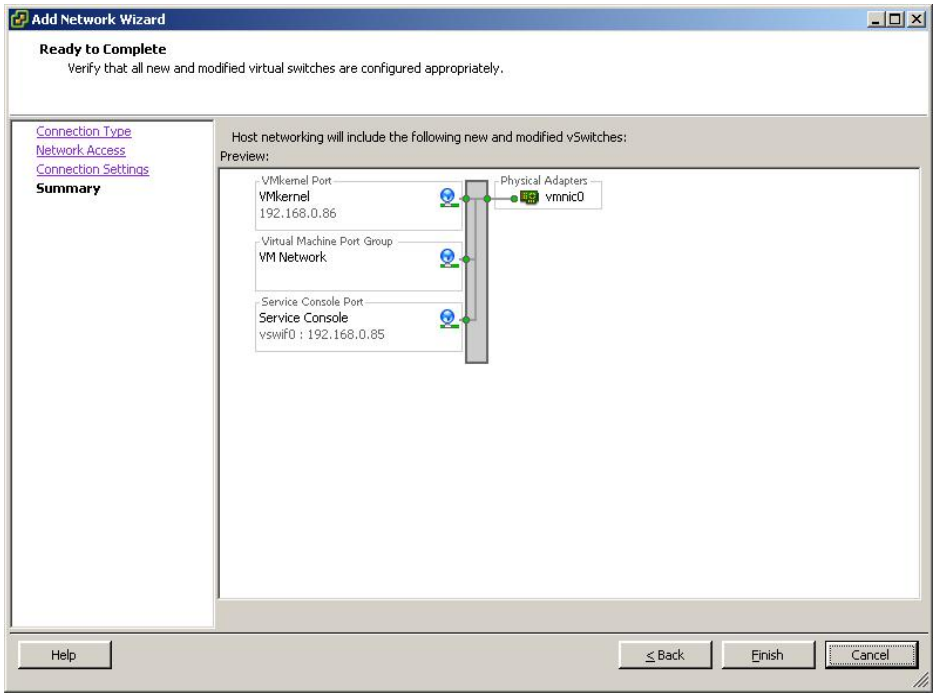

Figure 45: VMkernel Summary Screen

Before we can use the iSCSI initiator, we have to ensure that the VMware firewall settings allow iSCSI traffic. To enable this, select the *security profile* link in the configuration tab:

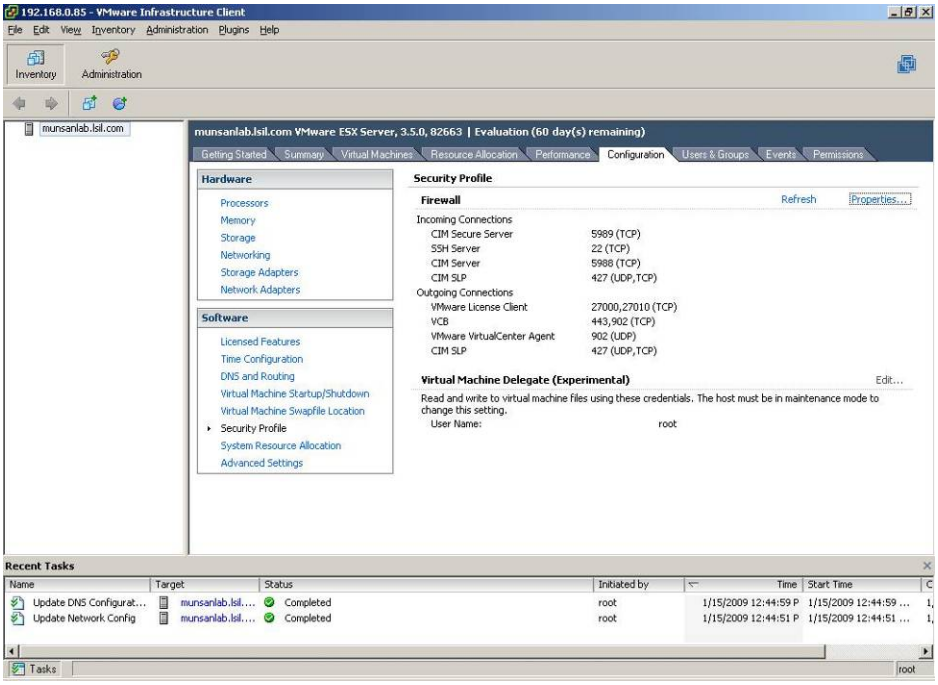

Figure 46: Security Profile Screen

Now click on *properties* in the upper right corner. This will lead us to the configuration screen. Make sure, that the Software iSCSI client box is checked:

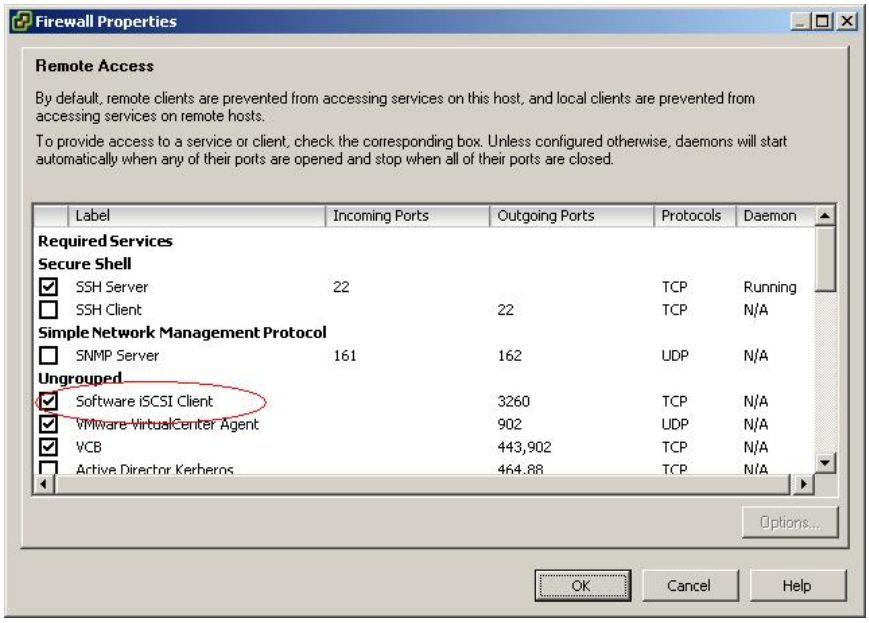

Figure 47: Firewall Properties

After enabling the Software iSCSI client, we should see the iSCSI HBA in the Storage Adapters list under the configuration tab:

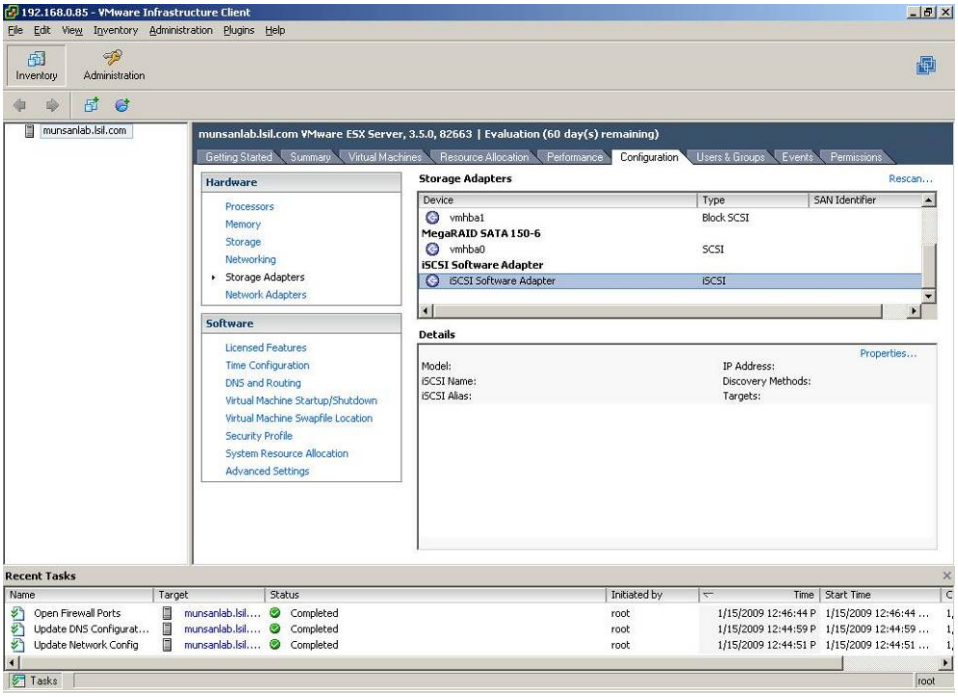

Figure 48: Storage Adapters Screen

Next, enable the iSCSI Software Adapter and enter the target discovery as well as the CHAP credentials if necessary. In order to accomplish this, we will click on the *properties* link in the iSCSI Software Adapter details:

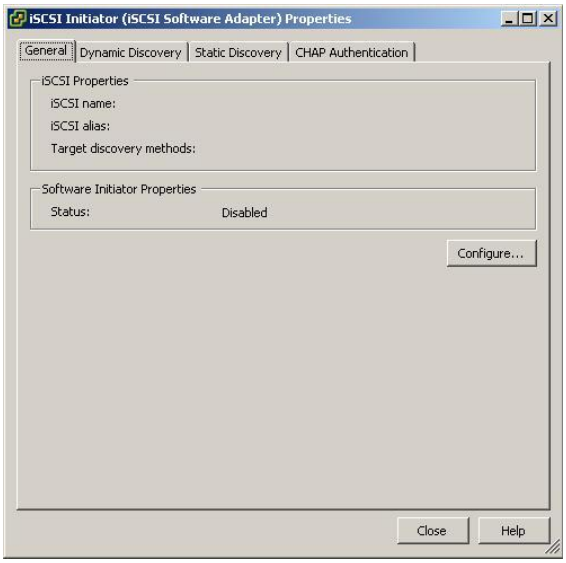

Figure 49: iSCSI Initiator Properties

Click on configure to enable the Adapter and selct "enable" at the General Properties window.

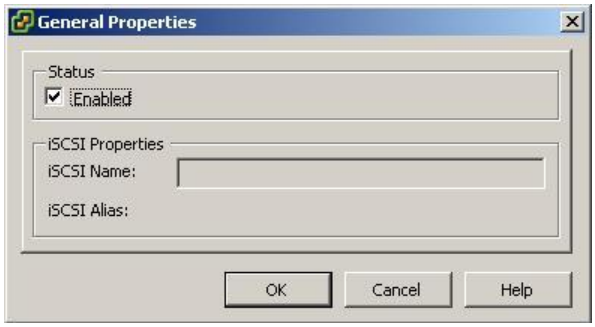

Figure 50: General iSCSI Properties

We will now add the previously assigned Controller A1 IP-Address into the Dynamic Discovery tab:

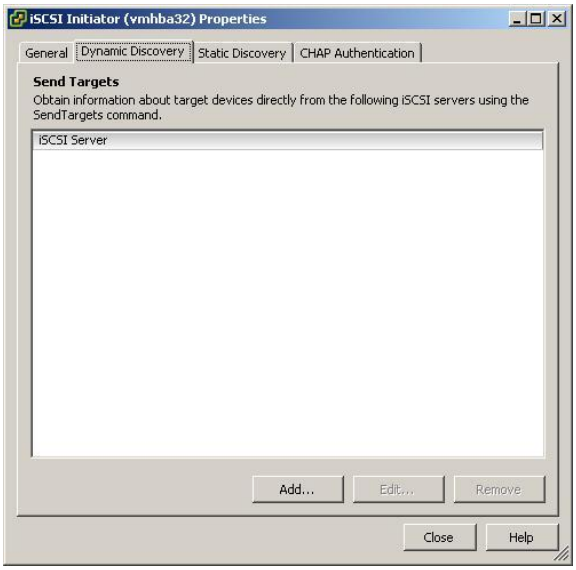

Figure 51: Dynamic Discovery

We only need to enter one IP-Address of our Controller. VMware will then automatically discover the other three:

Figure 52: Add Send Targets Server

If CHAP-secrets are being used in your environment, we can enter them in the CHAP Authentication tab. Here is a view of the screen:

Figure 53: CHAP Authentication

The basic iSCSI VMWare configuration is done. We will continue now with the setup of the DS3300 in order to work with VMware.

Configuring the DS3300 for VMWare

As VMware does not have a host-content based Agent, we have to manually create the host in the DS3300 Storage Manager. Click on *Configure Host Access (Manual)* in the Configure Tab of the DS3300 Storage Manager:

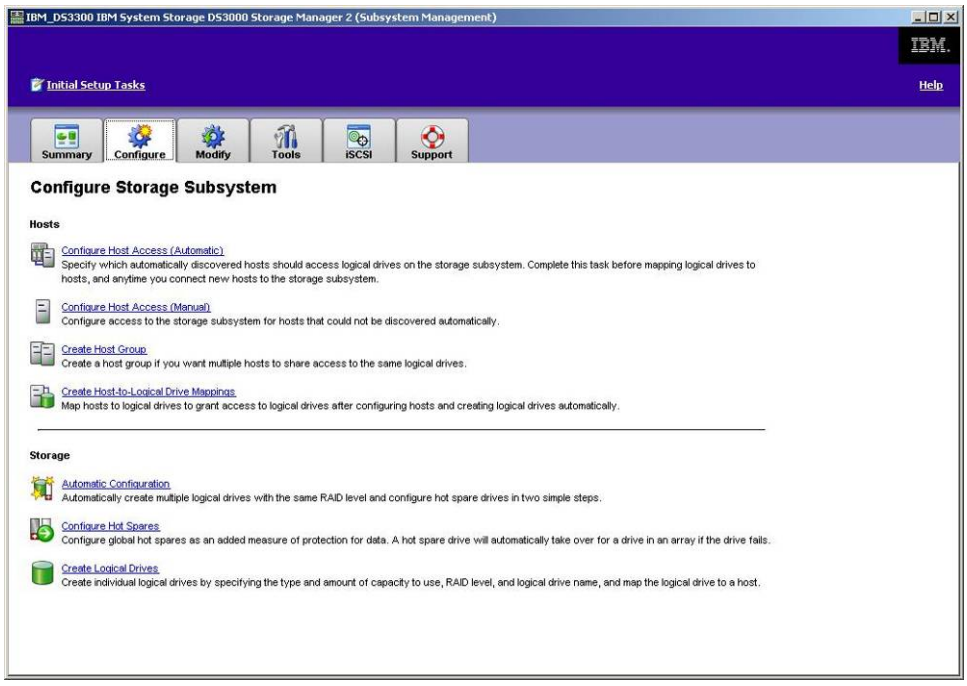

Figure 54: IBM DS3000 Storage Manager Configure screen

Enter the Host name (typically the Name you have given to the Server) and select the Operating System LNXCLVMWARE for the ESX 3.5 Server:

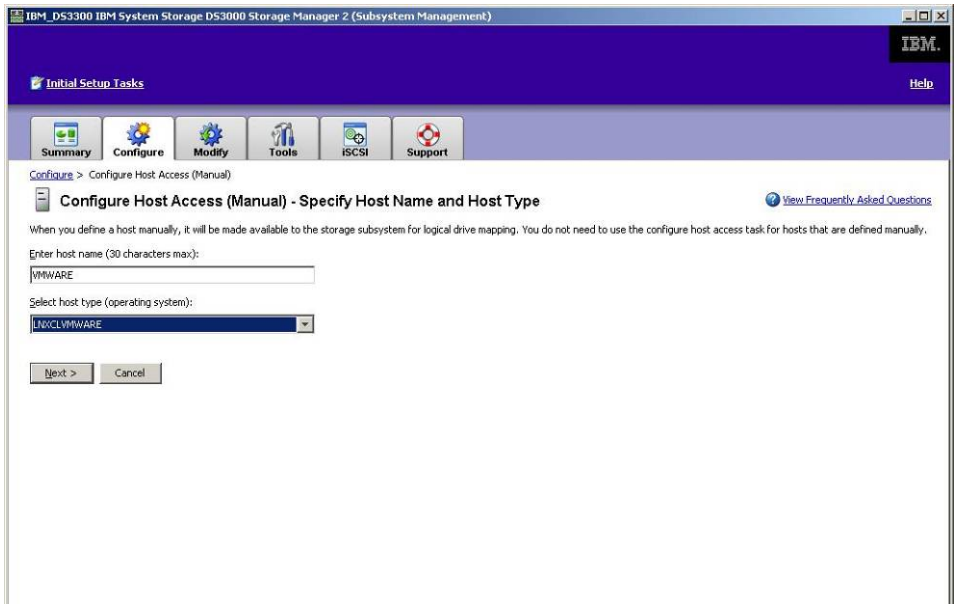

Figure 55: Host Specification Window

The iSCSI Initiator name of the VMware server should appear in the left selection box. Click on add to select this Initiator name. It now should be in the right hand box.

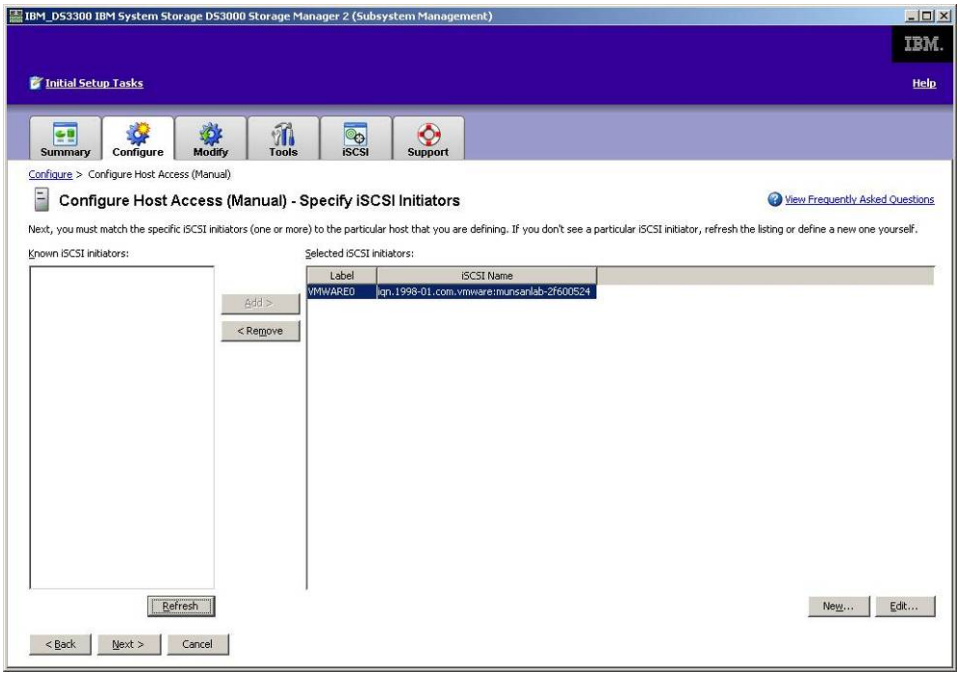

Figure 56: Specify iSCSI Initiator

The next step is to tell the DS3300 Software if this is a stand-alone VMware Server or a VMotion system. Select either Yes (for VMotion) or No (for stand-alone).

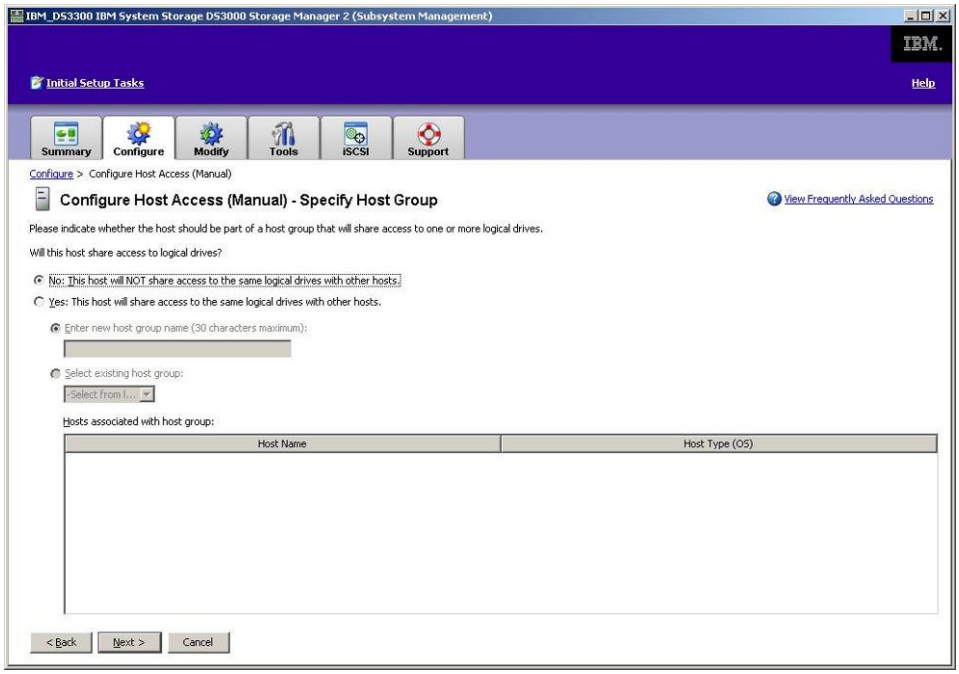

Figure 57: Specify Host Group

The confirmation screen will give us a quick overview on the selected options. Once we hit finish, the Host will be created and is ready to use. Before we can go back to VMware we need to create a logical drive on the DS3300 for use in VMware:

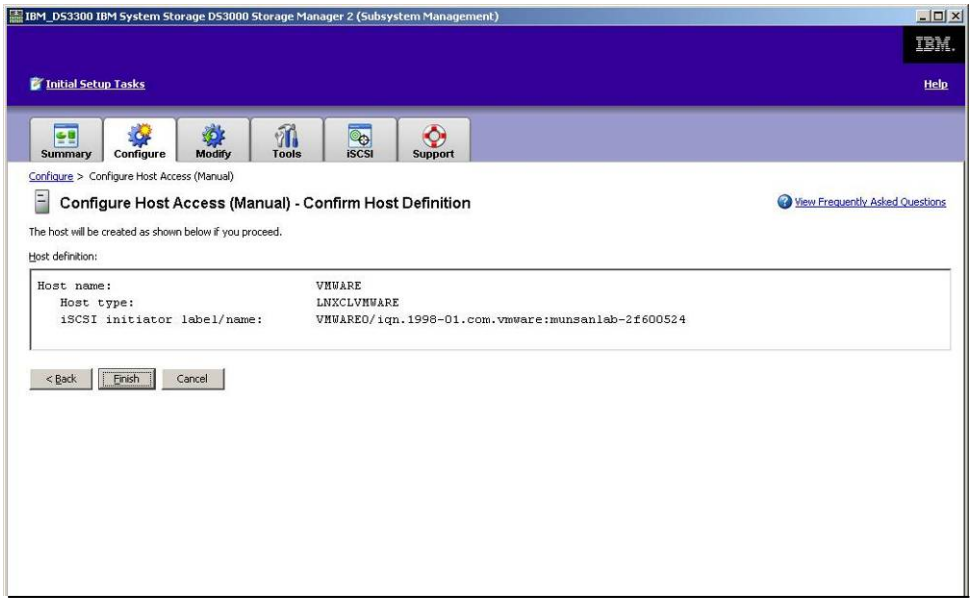

Figure 58: Confirmation Window

Again in the configure tab of the DS3300 we will click on the link *Create Logical Drive*. This will start a Wizard to create a logical drive. If you need information on creating a logical drive, refer to section logical drive Management with the DS3300. Map the logical drive to the VMware Server:

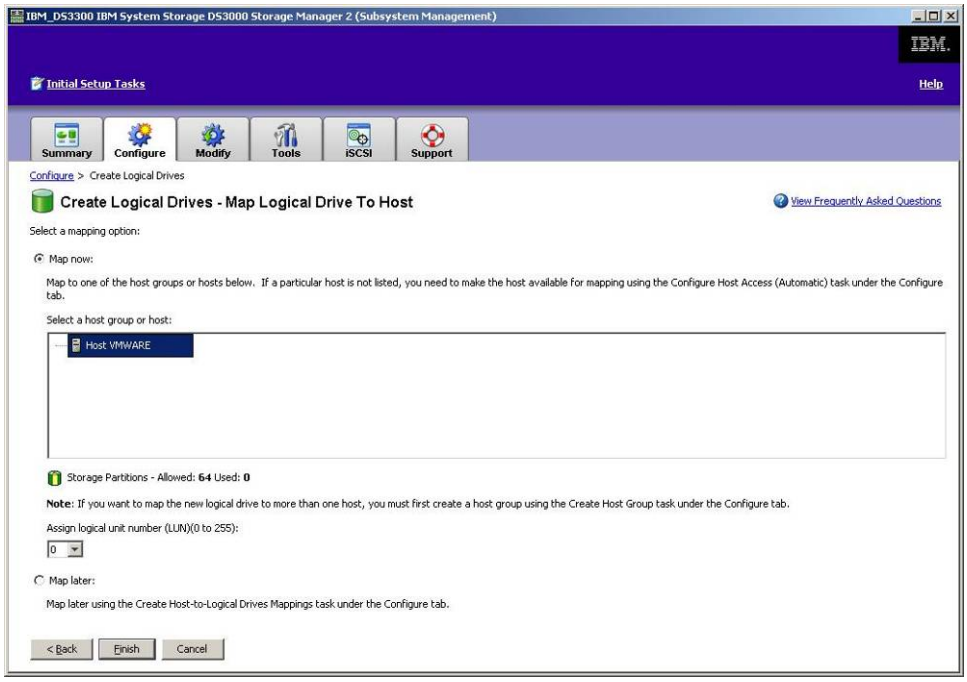

Figure 59: Map Logical Drive to Host Window

After clicking finish, the logical drive is visible to our VMware ESX Server.

Creating a Datastore on the DS3300 Logical Drive

We now need to rescan the iSCSI Software Adapter by clicking the *Rescan* link in the Configuration Tab under Storage Adapters:

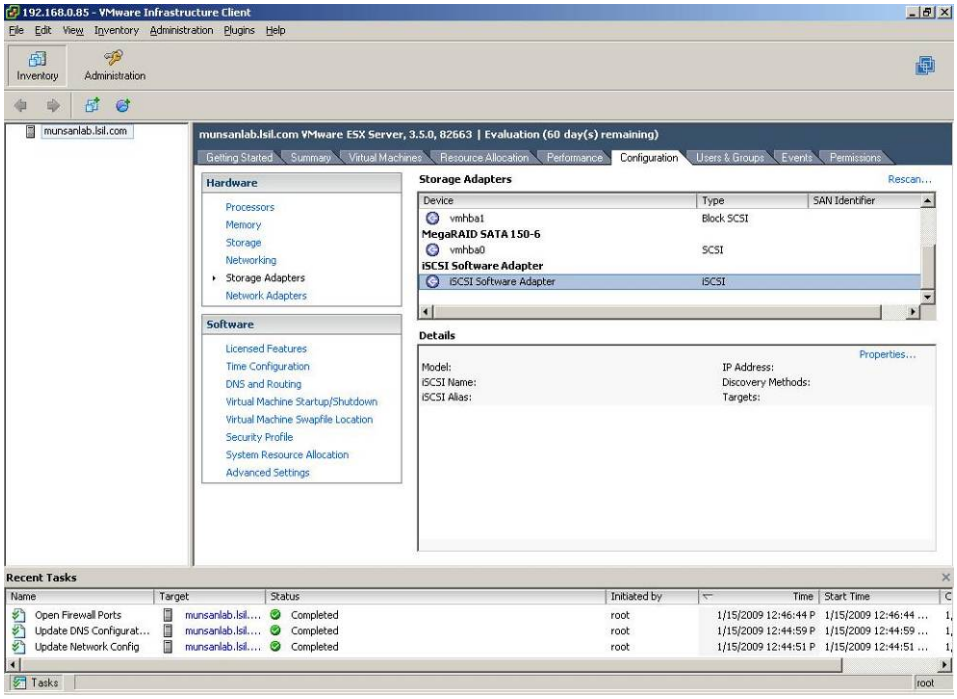

Figure 60: VMware Storage Adapters Configuration Screen

We want to select all devices and file systems for rescan:

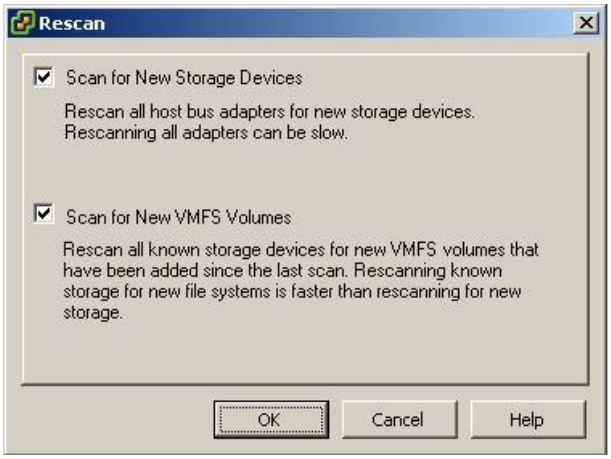

Figure 61: VMware Rescan Window

And, our newly created logical drive is visible under VMware:

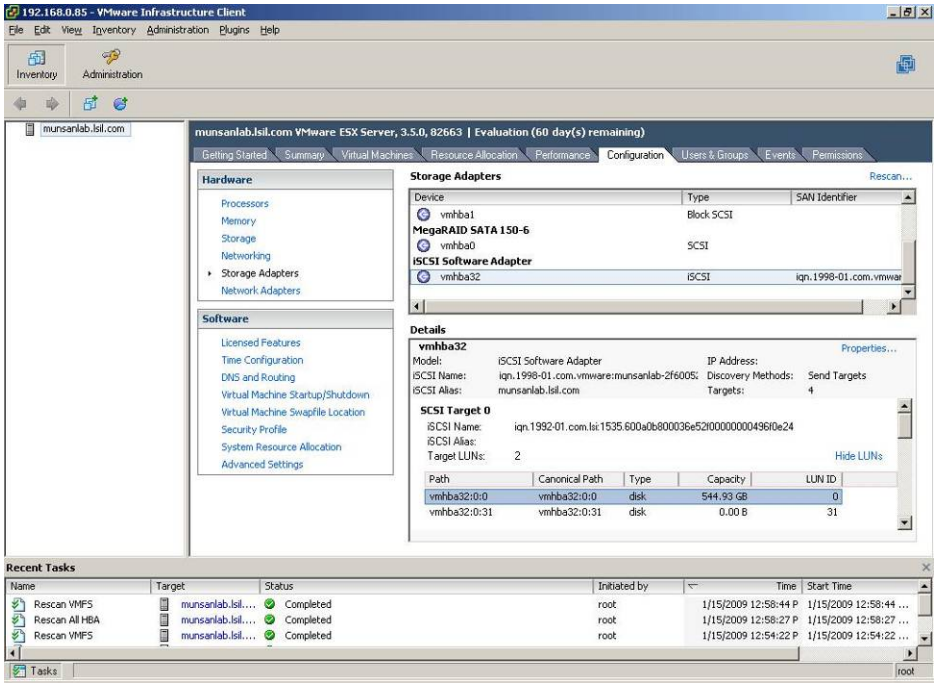

Figure 62: Storage Adapter Configuration Window

We now switch the view to the storage link in the hardware section of this tab and click on the Add Storage Link to create a vmfs on our logical drive:

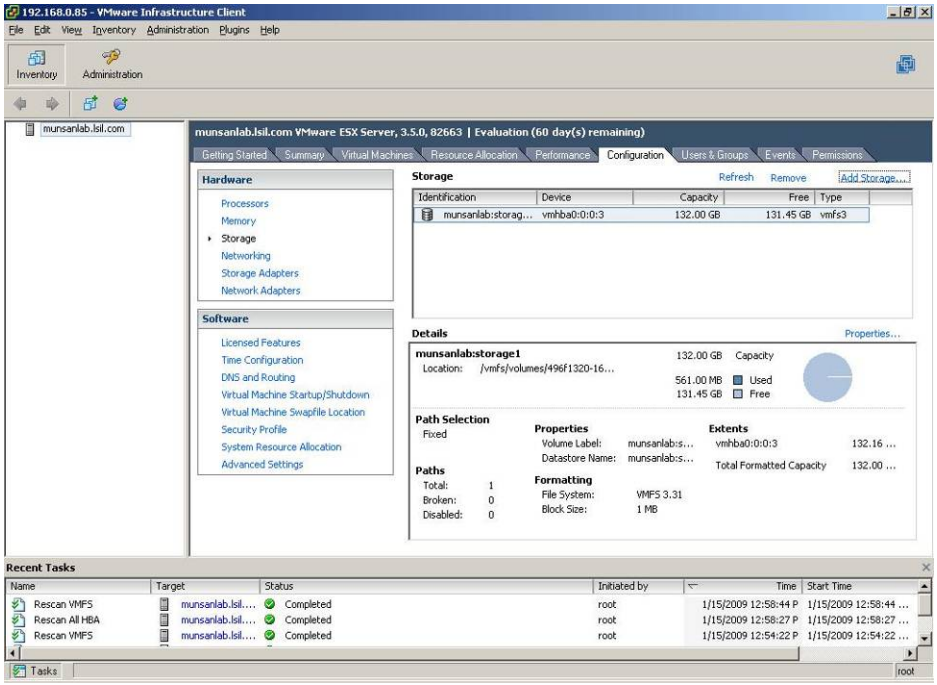

Figure 63: Storage Configuration Tab

This will start the Add Storage Wizard which will guide us through the creation of the vmfs file system. It is best to use the default settings:

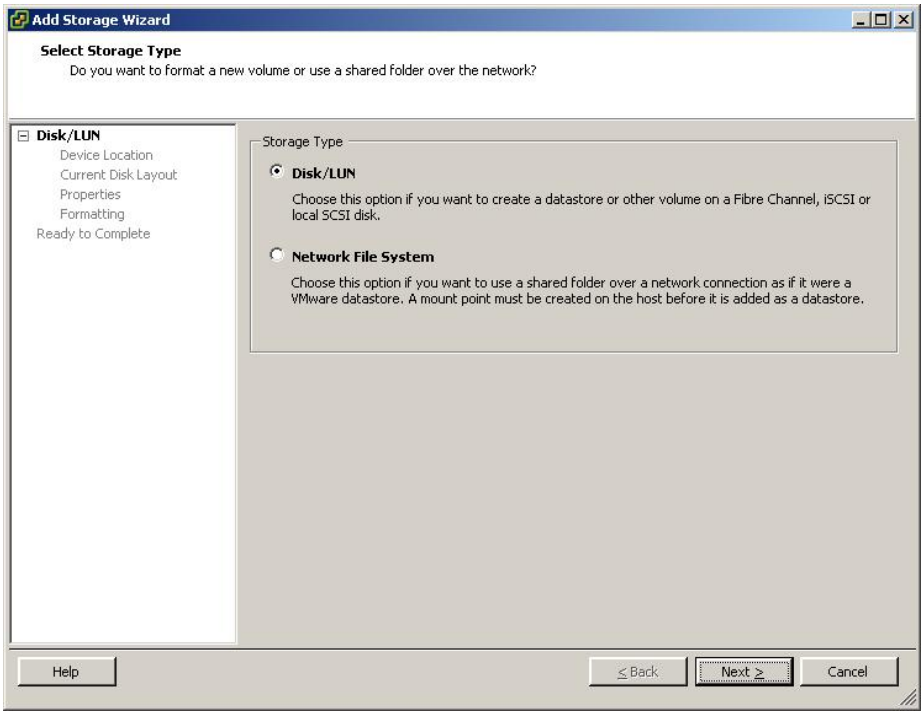

Figure 64: Add Storage Wizard

Select the recently created iSCSI logical drive:

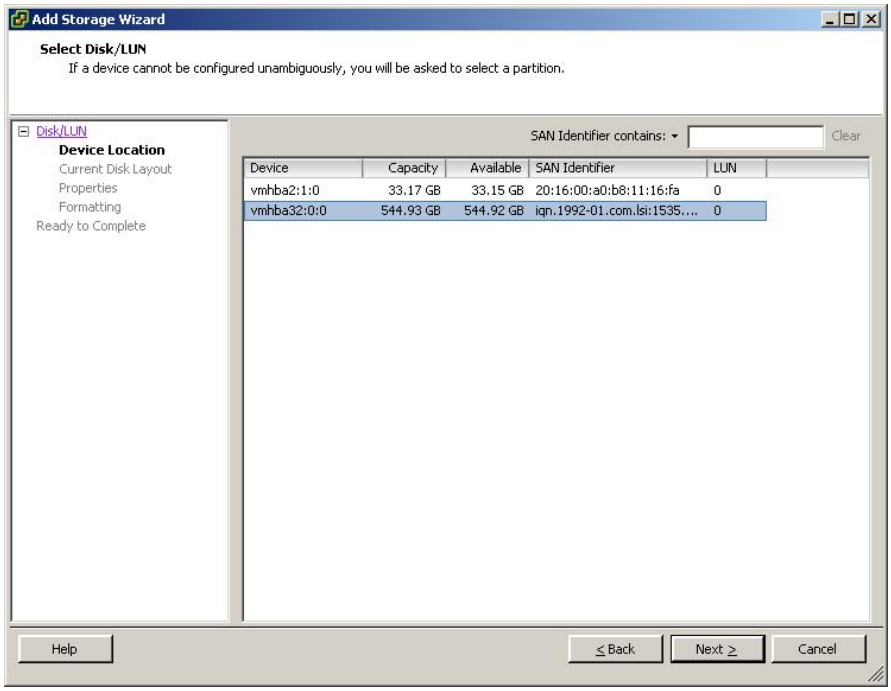

Figure 65: Select Disk / LUN Screen

We will now see the content of this drive. As it has never been used, this drive is blank:

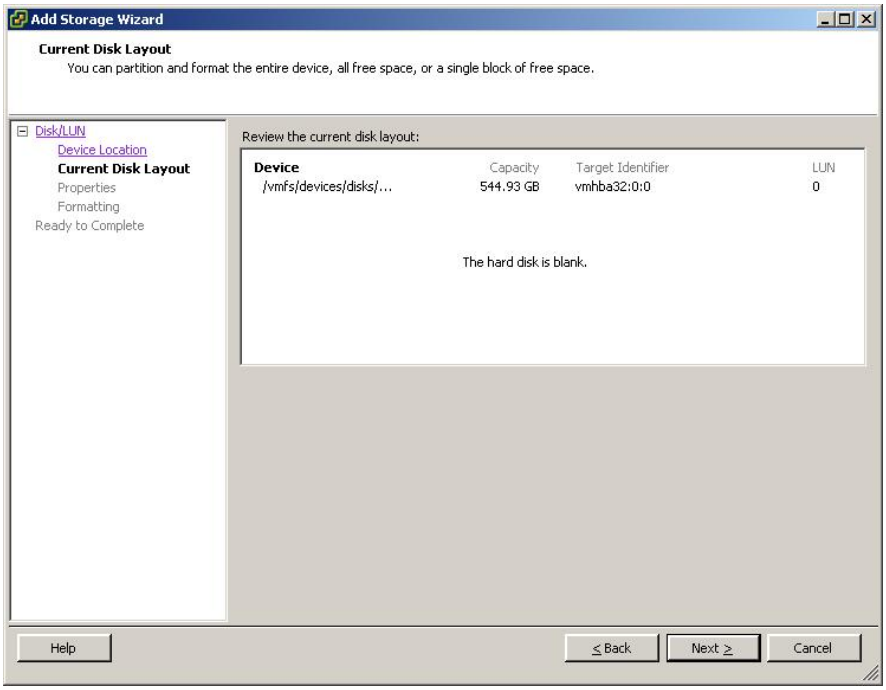

Figure 66: Current disk drive layout screen

Enter a name for the datastore:

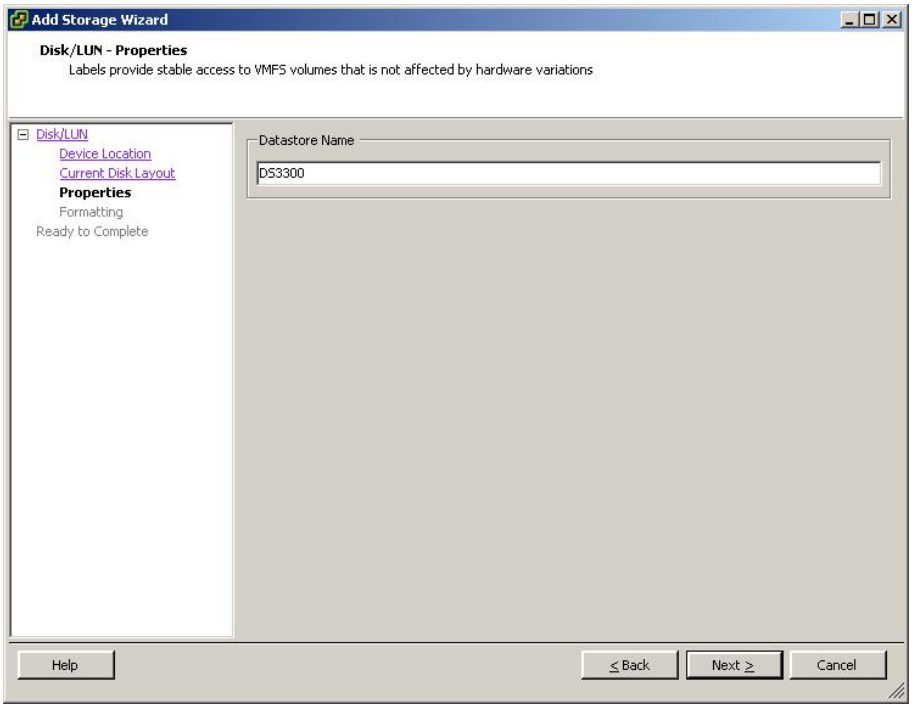

Figure 67: Disk drive/LUN Properties Window

We now specify the formation properties of this Logical Drive. Use the the settings to maximize your capacity, by selecting a blocksize equal or higher than the size of your disk:

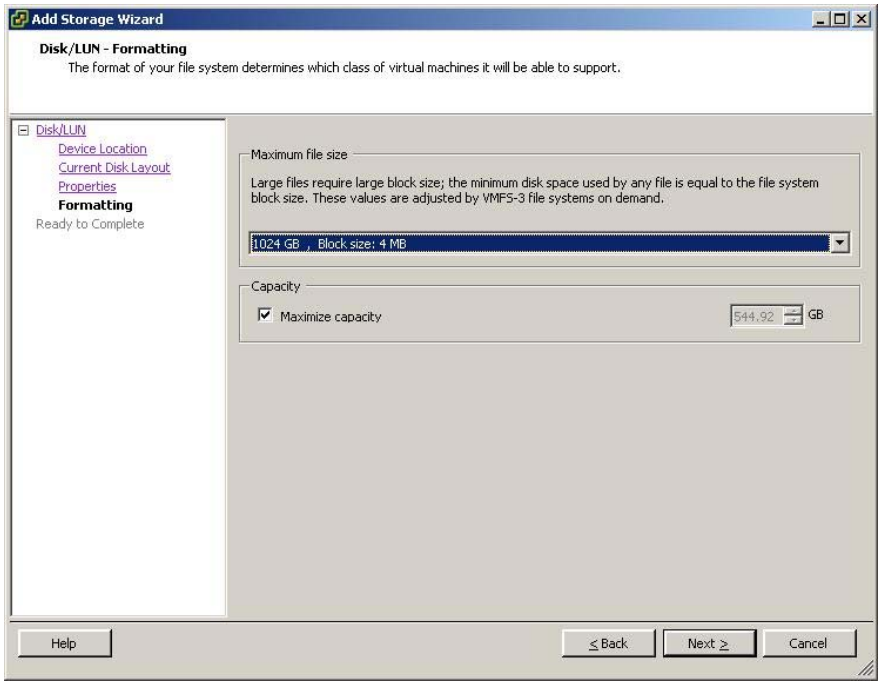

Figure 68: Disk drive/LUN Formatting Window

We will now get a summary and after clicking on finished, we are finished adding our logical drive to our VMware ESX Server:

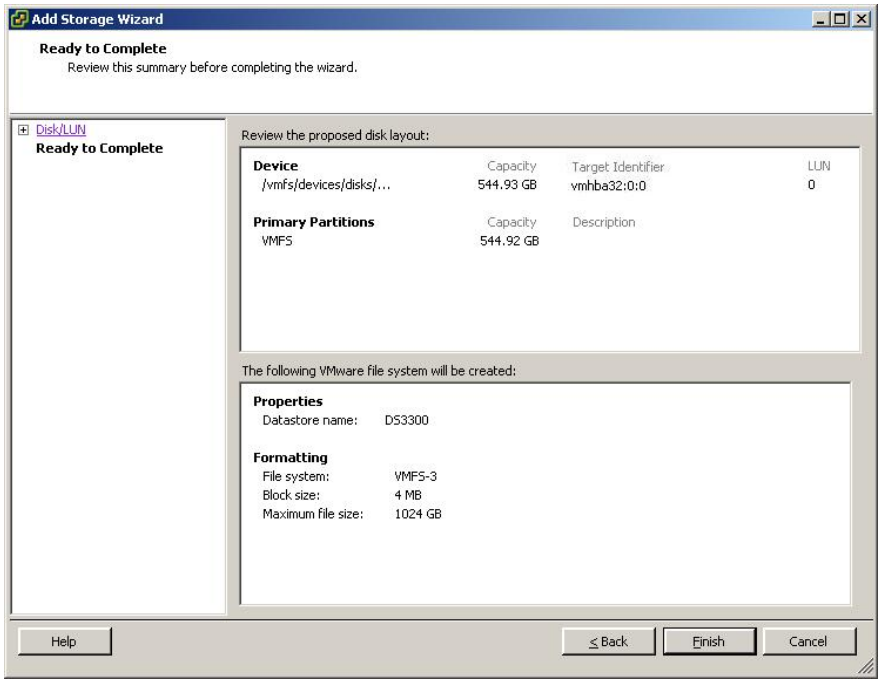

Figure 69: Add Storage Wizard Summary Screen

We will now see the logical drive in our Storage Summary screen and it is ready to use:

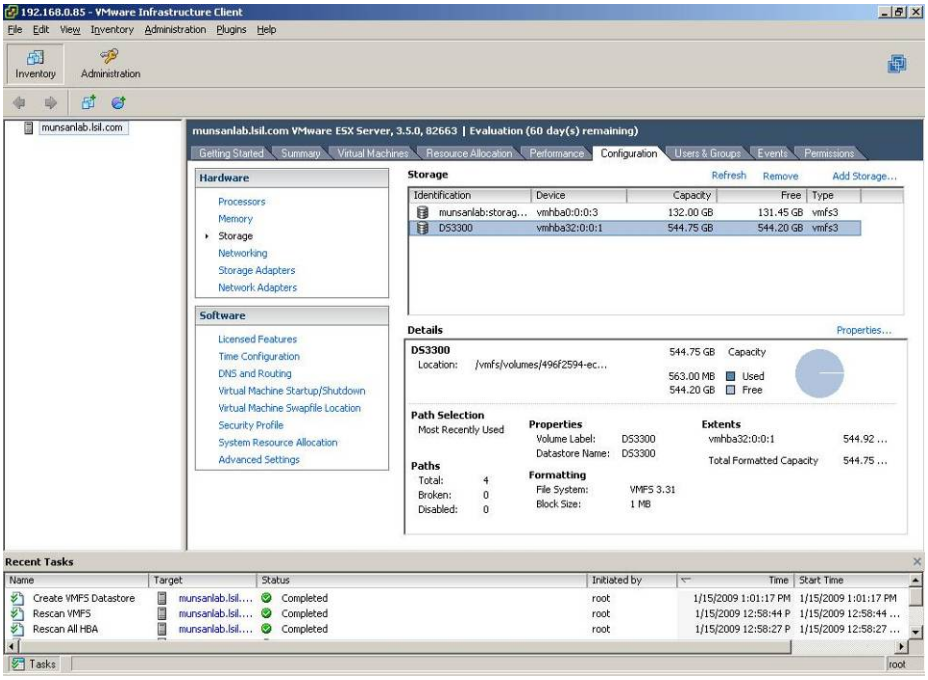

Figure 70: Storage Configuration Tab

Before we start creating a Virtual Machine on this Datastore, we want to make sure the path-failover configuration is correct by clicking on properties in the Details pane:

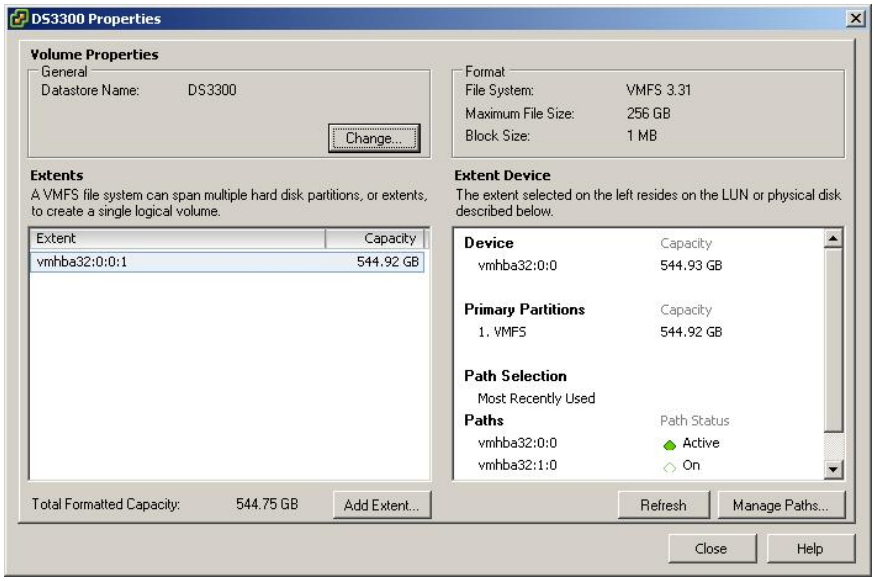

Figure 71: Datastore Properties Window

By clicking on the Manage Paths button, we should see all four paths to our controllers with one marked active and the other three marked on.

NOTE: It is recommended for best performance to use more than one Logical Drive per ESX Server, as ESX 3.5 does not support load-balancing on the iSCSI Software Initiator. Make sure that the paths to all Logical Drives are balanced between all available iSCSI paths, e.g. Logical Drive 1: Active – on – on – on, Logical Drive 2: on - Active – on – on, Logical Drive 3: on – on – Active – on, Logical Drive 4: on – on – on - Active

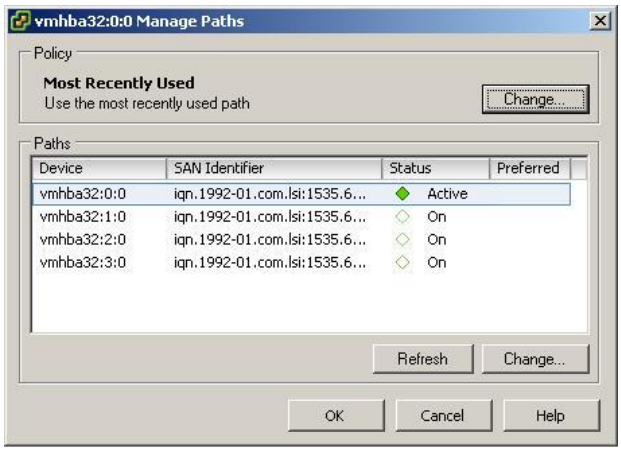

Figure 72: VMware Datastore Path Management

In order to make fail-over work correctly, especially in VMotion environments, make sure the Path-Policy is set to MRU – Most Recently Used. You can change this option if not set correctly:

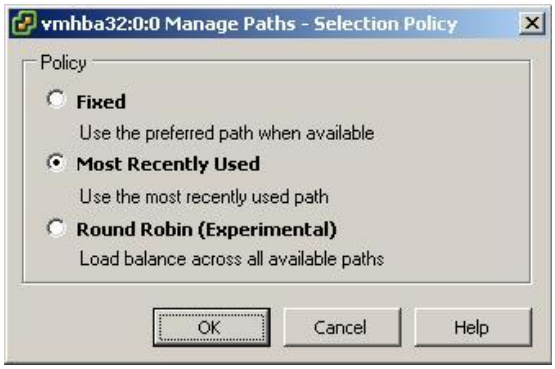

Figure 73: Path Selection Policy Window

The setup of our Logical Drive is now complete. We now can start adding a Virtual Machine.

Creating a Virtual Machine on the DS3300 Datastore

We now want to create a Virtual Machine with the Guest Operating System Windows Server 2003. To do this we click on the *Getting Started Tab* and select the link create *New Virtual Machine*:

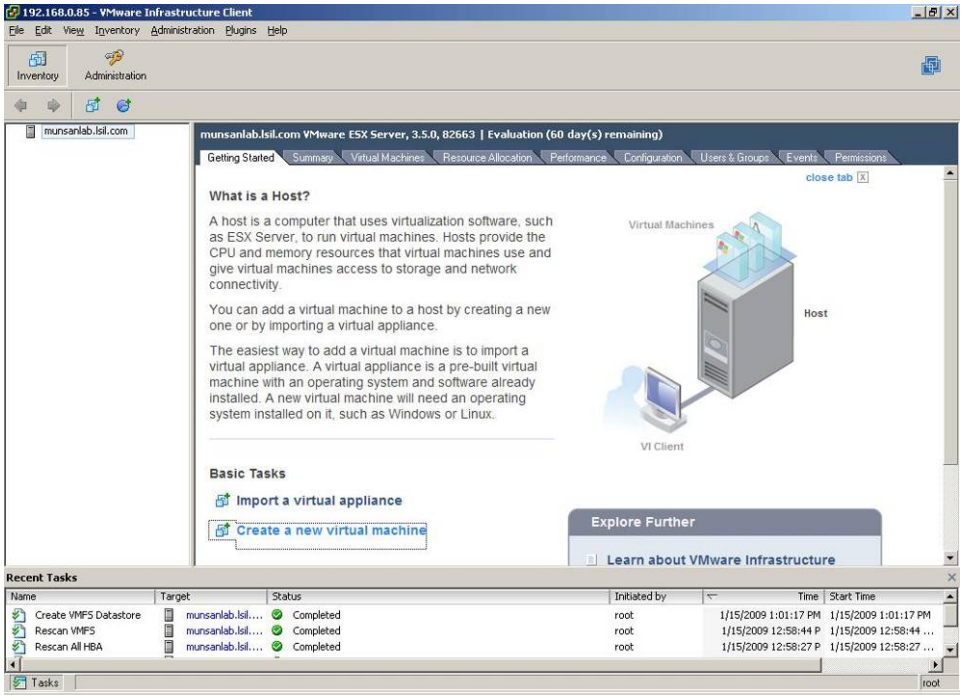

Figure 74: VMware Getting Started Screen

This will start the new Virtual Machine Wizard:

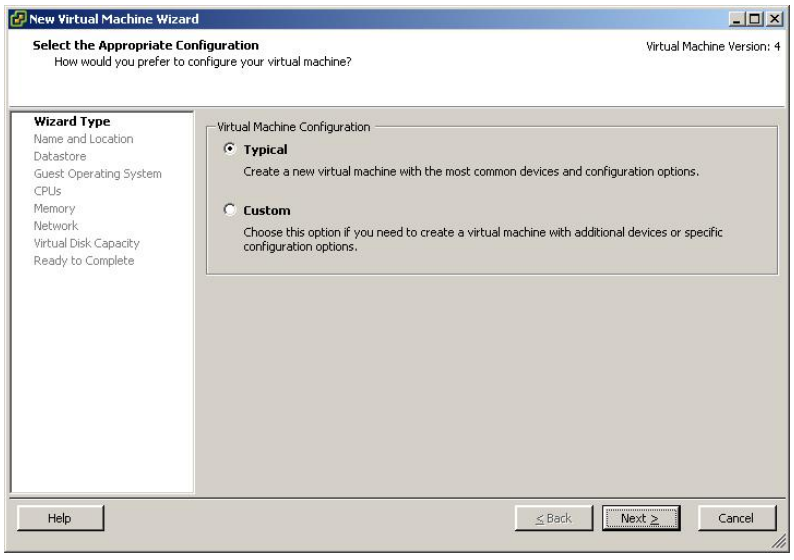

Figure 75: New Virtual Machine Wizard

Select the Name for the New Virtual Machine:

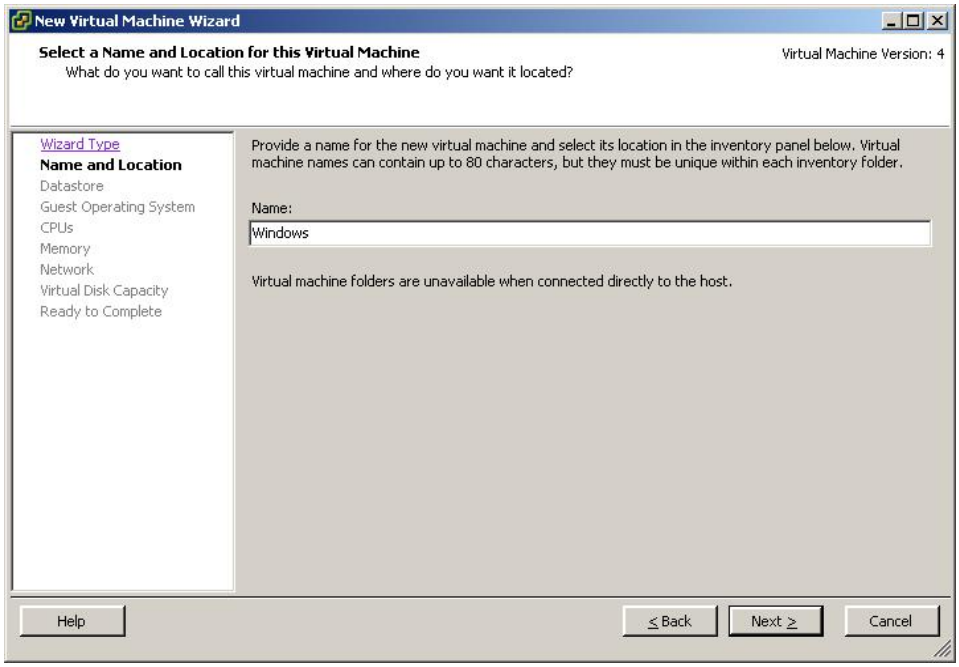

Figure 76: Name and Location of the New Virtual Machine

Now choose the newly created Datastore for our Virtual Machine:

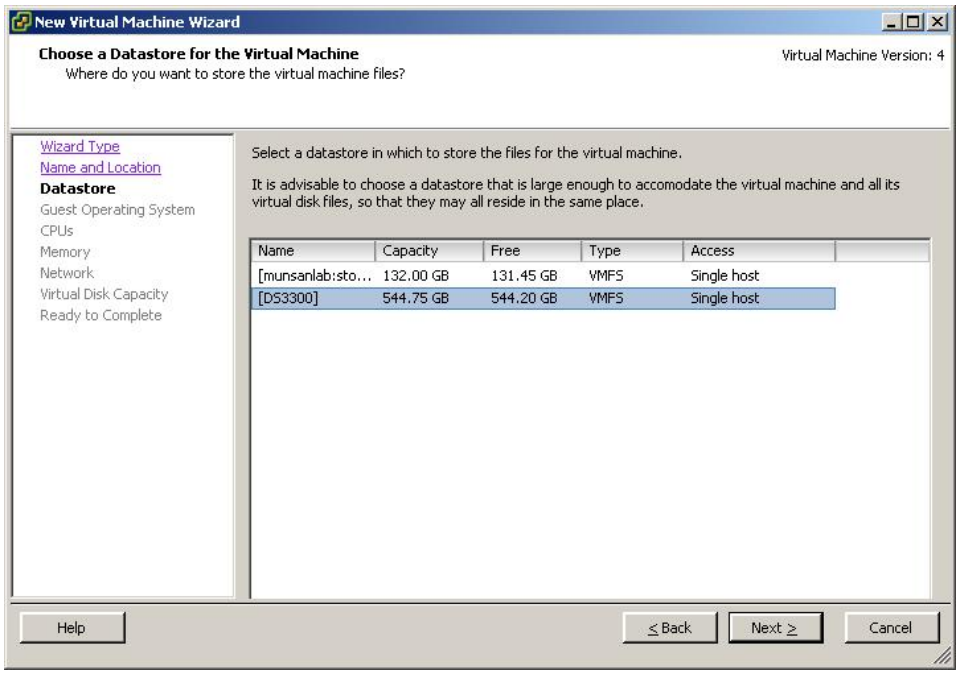

Figure 77: Datastore Selection Window

Select the Guest Operating System:

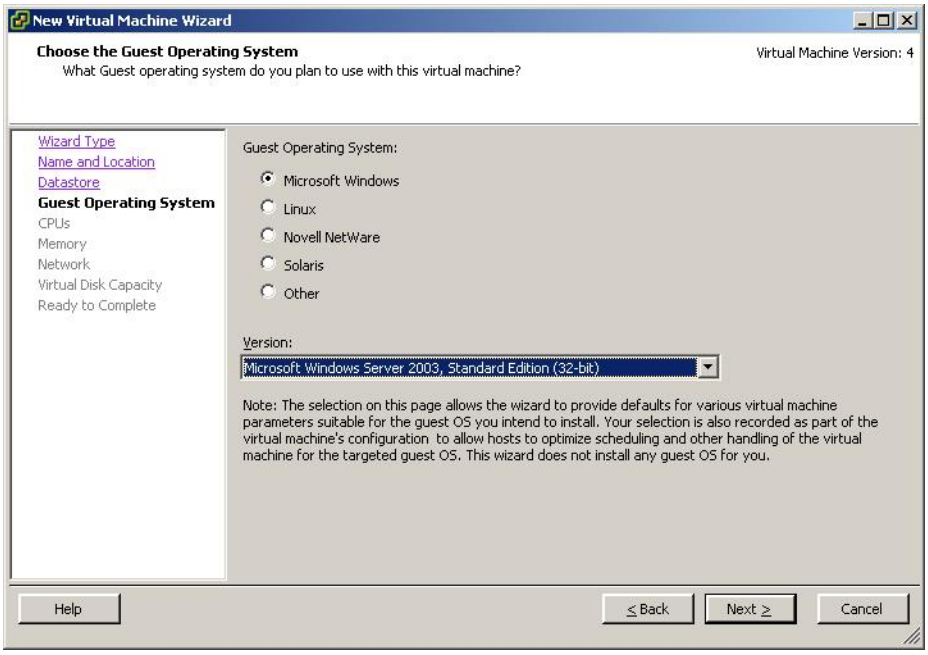

Figure 78: Guest Operation System Selection

Select the number of virtual processors as well as the guest OS memory and the networking you would like added:

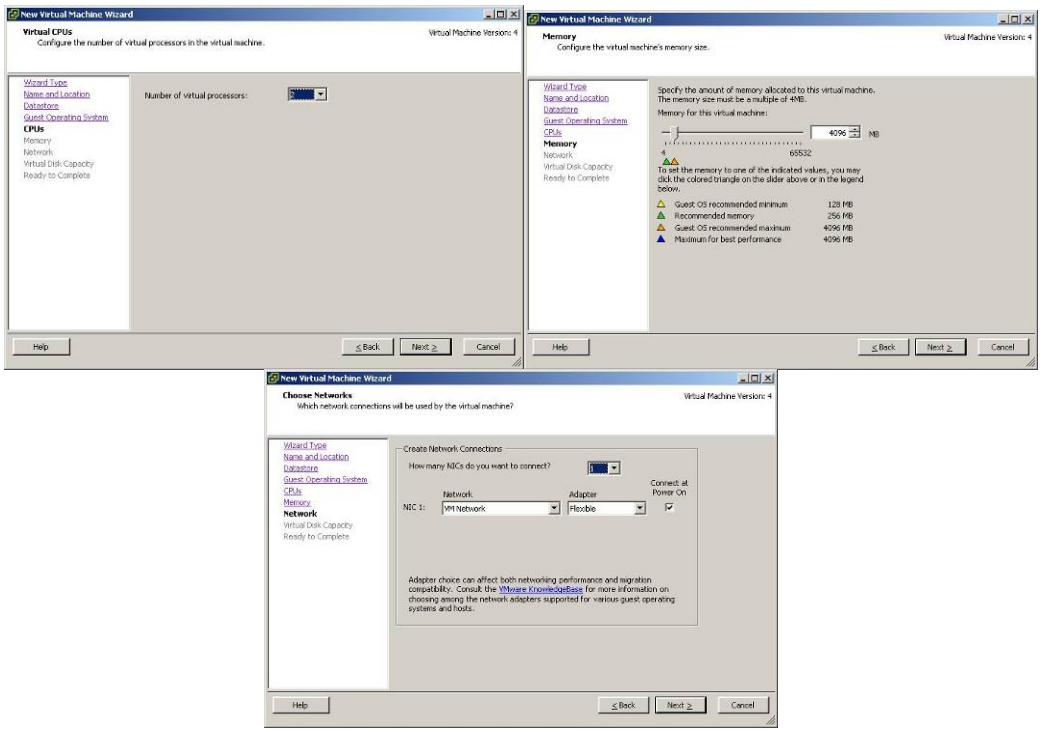

Figure 79: CPU, Memory and Network Selection

Now, select the amount of disk drive space for your guest OS. It is recommended to use this space only for the guest OS and put the applications running on those virtual machines on a separate virtual disk drive, raw-device mapping or direct mapping to the virtual machine. You can add space / virtual disk drives after you created the virtual machine.

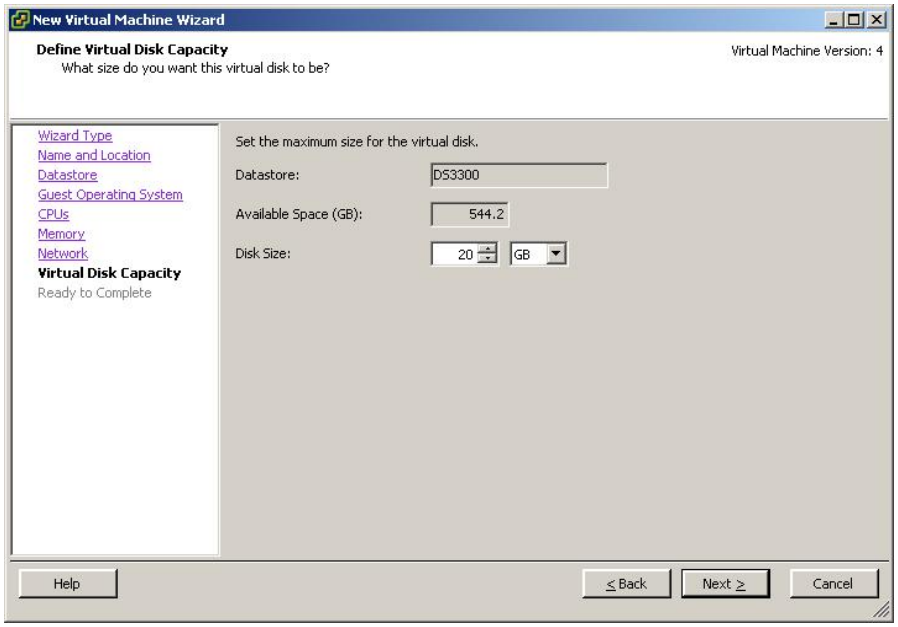

Figure 80: Disk drive Space Selection Screen

After the confirmation and Summary screen we have created the virtual machine and it is ready to install.

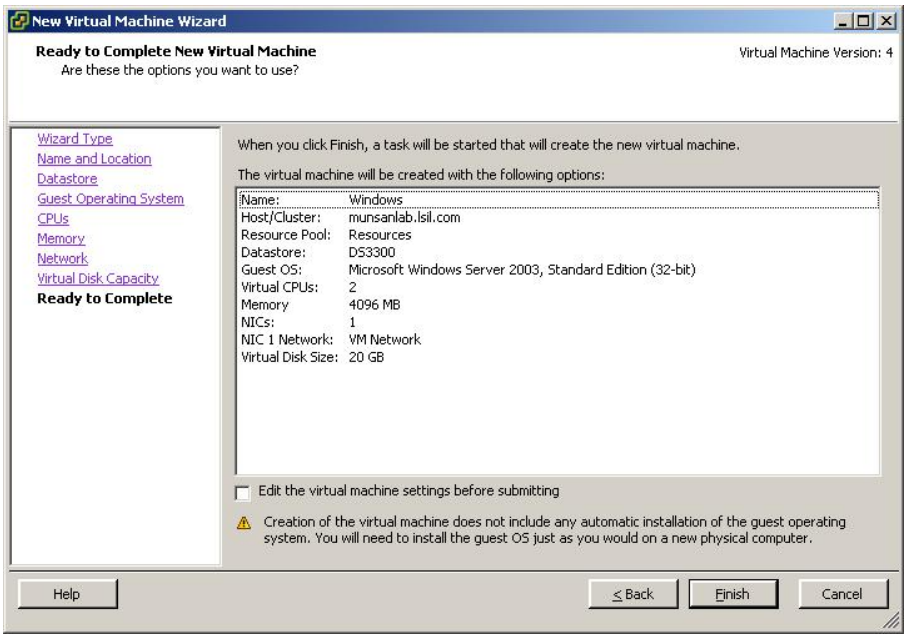

Figure 81: Virtual Machine Wizard Summary Screen

You can power up and install the virtual machine by right-clicking on the virtual machine and selecting “Power Up”. Assign the CD-ROM or ISO Image prior to powering up the virtual machine.

NOTE: It is recommended to convert your guest OS CR-ROM into an ISO Image and upload this onto the VMware Datastore. Installations are much faster from an ISO Image

You can add the CD-ROM or ISO-Image by clicking on the Add button.

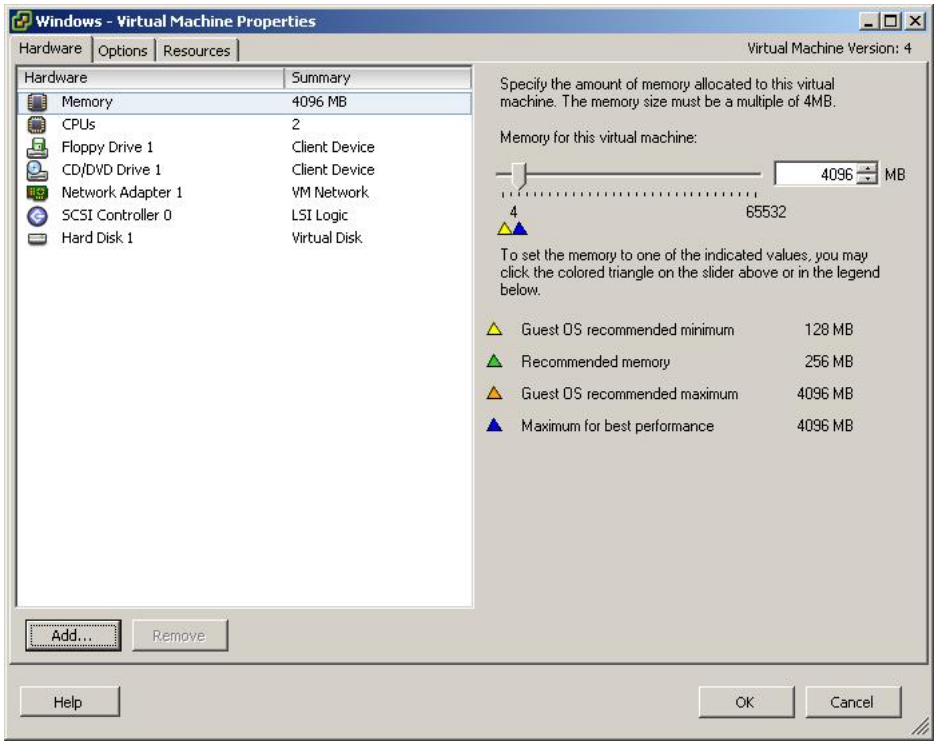

Figure 82: Virtual Machine Properties

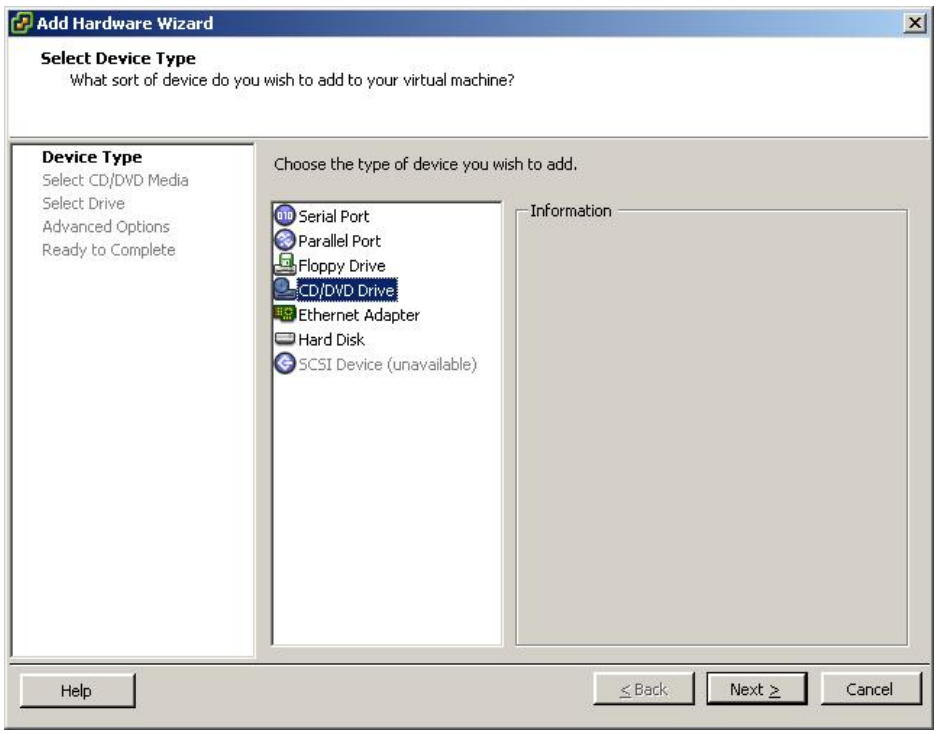

Figure 83: Add CD/DVD Drive Screen

If you are using an ISO-Image, click on Browse and select the formerly uploaded ISO-Image of your guest OS. You can then start installing your virtual machine.

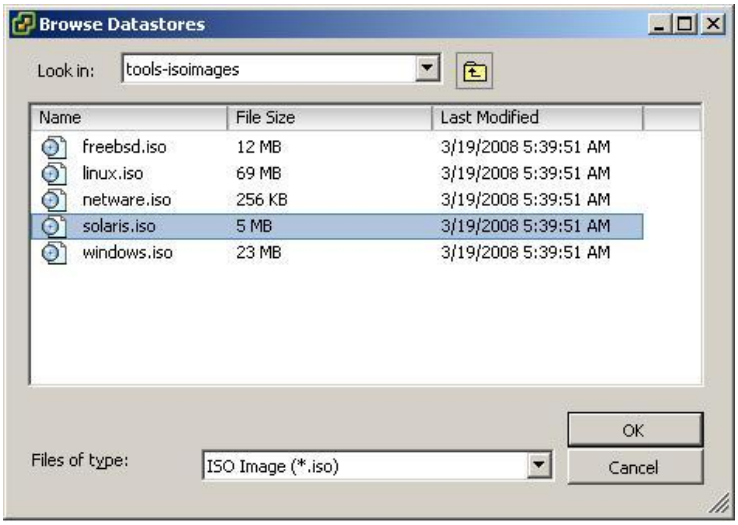

Figure 84: Browse Datastore

Installing the DS3300 in an Linux Environment

The following section describes the necessary steps to setup the DS3300 storage subsystem with Redhat Linux 4 and SuSE SLES 10.

How to install the iSCSI Software Initiator

Make sure your hardware is working properly and you already have IP addresses assigned to your server. You need to install the iSCSI Software Initiator onto the Linux System. Your OS CD may be required to install those packages. To do so follow these instructions:

Redhat:

Click on Applications – System Settings – Add Remove Applications

Select Network Services and click on Details:

Figure 85: Package Management

Open the Extra Package selection and click on iscsi-initiator-utils. Hit the close button when done to close this selection and then click on update.

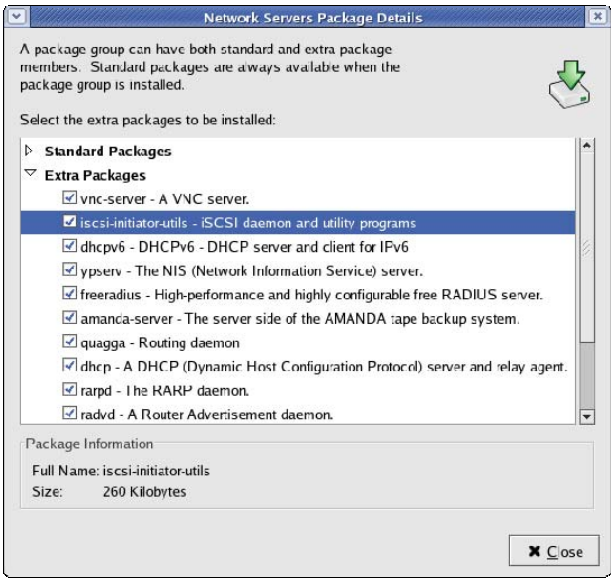

Figure 86: Package Management

SLES10: Click on Software – Software Management. In the Search, search for "iscsi":

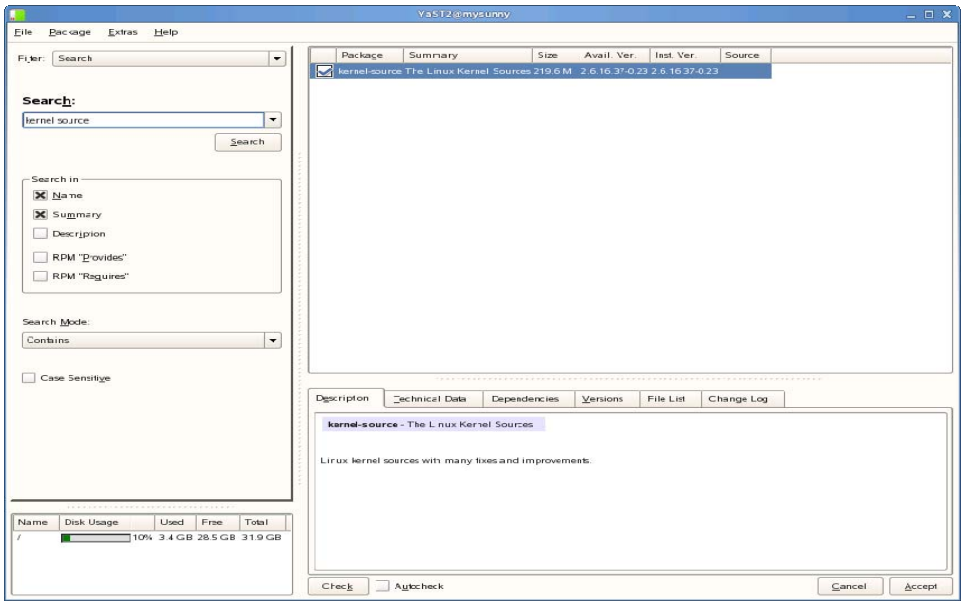

Figure 87: Software Management

Check the package box and the hit accept to install the software.

How to install the IBM DS3000 Storage Manager

Mount the cdrom with the IBM Storage Manager. Execute the Install Program in the Install Subdirectory. It is called, depending on the version, SMIA-LINUX(64)-02X.XX.XX.XX.bin. Follow the instructions on the screen and install all packages.

NOTE: The mpp-RDAC failover driver needs to be downloaded seperately. We need to manual install the mpp-RDAC package.

First, configure the IP Addresses for the DS3300 iSCSI ports. We will stick to the same addresses in our previous section. Refer to page 17 on how to do this.

We will now add the Target IP-Addresses configured for the DS3300 and start the iSCSI daemon.

RedHat:

Edit the /etc/iscsi.conf

```
DiscoveryAddress=192.168.0.190
DiscoveryAddress=192.168.0.192
HeaderDigest=never
DataDigest=never
LoginTimeout=15
IdleTimeout=15
PingTimeout=5
ConnFailTimeout=144
AbortTimeout=10
ResetTimeout=30
Continuous=no
InitialR2T=no
ImmediateData=yes
MaxRecvDataSegmentLength=65536
FirstBurstLength=262144
MaxBurstLength=16776192
```

Now we can start the iSCSI daemon by:

```
#/etc/init.d/iscsi start
```

SLES10:

Start the iSCSI daemon by

```
#/etc/init.d/open-iscsi start
```

Then discover the DS3300 with the following command:

```
#iscsiadm –m discovery st –p 192.168.0.190

[2de368] 172.168.10.14:3260 via sendtargets
[52d0e8] 172.168.10.15:3260 via sendtargets
```

The hex number in brackets are the available paths. We will now log onto every path available:

```
#iscsiadm –m node –r 2de368 –l
#iscsiadm –m node –r 52d0e8 –l
```

We are now ready to configure the Host in the DS3300 Storage Manager. As Linux will come with a Host context agent we need to restart the Agent, so it can give the Information to the DS3300 about our iSCSI Node Name and Host OS:

```
#/etc/init.d/SMAgent stop
#/etc/init.d/SMAgent start
```

We can now automatically add the host with the IBM DS3300 Storage Manager and create a logical drive to map it to our Linux Host. To do so, follow the steps from page 33.

Installing the mpp-RDAC failover driver

After the Logical Drive is mapped to our Server we need to install the mpp-RDAC multipath driver.

Locate the RDAC package and untar it

example:

```
#tar -zxvf rdac-LINUX-<version>-source.tar.gz
```

It creates a directory 'linuxrdac-<version> with all the files in it

```
#cd linuxrdac-<version>
#make install
```
-- Compiles the linuxrdac driver & creates a MPP's image (INITRD) by name mpp-<kernel-version>.img under directory '/boot'

If the host/system has running kernel-version: 2.6.9-34.ELsmp, then mpp's image name would be 'mpp-2.6.9-34.ELsmp.img

```
#ls -al /boot/mpp*

-rw-r--r--  1 root root 2741550 Aug  4 22:03 /boot/mpp-2.6.9-34.ELsmp.img
```

Redhat:

Modify the boot loader '/etc/grub.conf' and save it

Original bootloader 'kernel-smp' entry looks like,

```
title Red Hat Enterprise Linux AS (2.6.9-34.ELsmp)
    root (hd0,0)
    kernel /vmlinuz-2.6.9-34.ELsmp ro root=/dev/VolGroup00/LogVol00 rhgb quiet
    initrd /initrd-2.6.9-34.ELsmp.img
```

Copy all the above 4 lines, paste another entry and modify it

Modify the above entry for RDAC as shown below,

```
title Red Hat Enterprise Linux AS (2.6.9-34.ELsmp) w/ MPP

        root (hd0,0)

        kernel /vmlinuz-2.6.9-34.ELsmp ro root=/dev/VolGroup00/LogVol00 rhgb quiet

        initrd /mpp-2.6.9-34.ELsmp.img
```

Default entry 'default=n' should be modified accordingly so that host can be booted up with mpp's image

Reboot the host

SuSE SLES10:

Modify the boot loader '/boot/grub/menu.lst' and save it

Original bootloader 'kernel-smp' entry looks like:

```
title Linux

        root (hd0,0)

        kernel (hd0,0)/vmlinuz root=/dev/sda3 vga=0x314 selinux=0 splash=silent
        resume=/dev/sda2 elevator=cfq showopts

        initrd (hd0,0)/initrd
```

Copy all the above 4 lines, paste another entry and modify it

Modify the above entry for RDAC as shown below,

```
title LinuxMPP

        root (hd0,0)

        kernel (hd0,0)/vmlinuz root=/dev/sda3 vga=0x314 selinux=0 splash=silent
        resume=/dev/sda2 elevator=cfq showopts

        initrd (hd0,0)/mpp-2.6.5-7.244-smp.img
```

Default entry 'default=n' should be modified accordingly so that host can be booted up with mpp's image

Reboot the host

Access the Logical drives with Linux

To make sure mpp-RDAC is installed an working properly you should issue the following command:

```
#/opt/mpp/lsvdev

        Array          LUN            sd device
        ---------------------------------------

        DS3300         0              /dev/sdb
```

You can also check the redundant paths by using the command mppUtil –g0.

The device is now ready to use. If you want to create a standard partition type:

```
#fdisk drive /dev/sdb
```

NOTE: If you need optimal performance, look carefully at the next section, as it will show how to create aligned partition with Linux. It is generally recommended to align your partitions.

Now follow the steps to create your partition and create the file System of your choice (ext2, ext3, ReiserFS)

Create a mount point and mount your Logical drive:

```
#mkdir /DS3300
#mount /dev/sdb1 /DS3300
```

Your Logical Drive is now ready to use.

Persistent Device Naming and Partition Alignment

Persistent device naming identifies logical drives by labels instead of the device name only. As device names can change when other logical drives are removed and server is rebooted, this is highly recommended for mppRDAC multipathing drivers.

There is the utility "udev" for Kernel 2.6. More information is available at:
http://kbase.redhat.com/faq/FAQ_85_8082.shtm

http://www.redhat.com/docs/manuals/enterprise/RHEL-5-manual/Virtualization-en-US/ch-virt-lun-persistence.html

http://www.kernel.org/pub/linux/utils/kernel/hotplug/udev.html

Partition alignment can be useful when you need to maximize performance. As described in Appendix A for Exchange, this should also be done under Linux.

It looks like Linux places partition start on track #1 (2nd track) but in fact Linux places partition start on sector 32 or sector 63. The start sector depends on track size (sectors per track) and the track size depends on logical drive size.This may result in mis-alignment with the segment size of the logical drive. The partition start sector should be adjusted to prevent mis-alignment. To achieve this, the ppartition should start at least at the first sector of a segment, or even better, the partition starts at the first sector of a stripe set. This is ideal for "Full Stripe Writes", good for storage block IO and a good choice for Oracle ASM.

Example of a default partition with mis-alignment:

```
# fdisk drive /dev/sdc

Command (m for help): p
Disk drive /dev/sdc: 513.5 GB, 513556348928 bytes
255 heads, 63 sectors/track, 62436 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes

Device Boot      Start         End      Blocks   Id  System
/dev/sdc1               1       62436  501517138+  83  Linux
```

```
Command (m for help): u
Changing display/entry units to sectors


Command (m for help): p

Disk drive /dev/sdc: 513.5 GB, 513556348928 bytes
255 heads, 63 sectors/track, 62436 cylinders, total 1003039744 sectors
Units = sectors of 1 * 512 = 512 bytes

Device Boot    Start          End                   Blocks         Id      System
/dev/sdc1             63             1003034339     501517138+     83      Linux
```

It looks like the partition starts at cylinder 1, but actually we are switching to sectors and we will see that it starts at sector 63. This can degrade performance as writes and reads will probably span two drives instead of one.

To create an aligned partition:

```
# fdisk drive /dev/sdb

Command (m for help): u
Changing display/entry units to sectors
Command (m for help): p
Disk drive /dev/sdc: 143.4 GB, 143457779712 bytes
255 heads, 63 sectors/track, 17441 cylinders, total 280190976 sectors
Units = sectors of 1 * 512 = 512 bytes

Device Boot      Start         End      Blocks   Id  System

Command (m for help): n
Command action
e   extended
p   primary partition (1-4) p
Partition number (1-4): 1
First sector (63-280190975, default 63): 8192
Last sector or +size or +sizeM or +sizeK (32768-280190975, default 280190975):

Using default value 280190975

Command (m for help): u
Changing display/entry units to sectors
Command (m for help): p
Disk drive /dev/sdc: 143.4 GB, 143457779712 bytes
255 heads, 63 sectors/track, 17441 cylinders, total 280190976 sectors
Units = sectors of 1 * 512 = 512 bytes


Device Boot    Start          End                   Blocks         Id      System
/dev/sdc1      8192           280190975             140091392      83      Linux

Command (m for help): w
The partition table has been altered!
Calling ioctl() to re-read partition table.
Syncing disk drives.
```

Now you can create your file system and it will be configured for best performance.

Installing the DS3300 in a Windows 2008 Environment

The following section was taken from the following IBM Redbook: IBM System Storage DS3000: Introduction and Implementation Guide - http://www.redbooks.ibm.com/redpieces/abstracts/sg247065.html?Open

Note: If you are configuring the system for clustering, see the chapter on Windows Clusters.

This chapter describes how to install the Storage Manager software for Windows Server in a standard (non-cluster) configuration.

Before you install the software, read the following information:

For correct installation, make sure that you have completed all preparation tasks

Always check for a readme file on any installation media. The readme file contains important information that was not available when this document was prepared.

This version of the Storage Manager software does not support 64-bit versions of Windows XP. All administrator functions that use the Storage Manager software must be performed from a 32-bit management station.

Configure the event monitor on only one management station to prevent the receipt of duplicate event messages. Note that duplicate alerts are also sent if the Enterprise Management window and the SMmonitor are running simultaneously.

Do not restart the system during the installation process. You will restart the system after you install all Storage Manager software components.

Installation Process

You can install the Storage Manager software components either on host servers only or on both host servers and management stations.

Management station: A management station is a system that you use to manage the storage subsystem. You can attach it to the storage subsystem in either of the following ways:

Through a TCP/IP Ethernet connection to the controllers in the storage subsystem

Through a TCP/IP network connection to a host server with host-agent software installed on it, which in turn is directly attached to the storage subsystem through the I/O path

You *must* install the Storage Manager SMclient software on a management station.

Host server: A host server (or host) is a server that is connected directly to the storage subsystem through an I/O path.

Important: Make sure that you install the host bus adapter and device driver before you install the Storage Manager software.

Installing the Storport Miniport Host Bus Adapter Device Driver

An IBM Fibre Channel, iSCSI, or SAS host bus adapter provides the interface between a host server and a DS3000 storage subsystem. IBM Fibre Channel, iSCSI, and SAS host bus adapters are high-performance, direct memory access, bus-master host adapters that are designed for high-end systems. The Fibre Channel host bus adapters support all Fibre Channel peripheral devices that support private-loop, direct-attach, and fabric-loop attachment. The IBM host bus adapter device driver enables the operating system to communicate with the host bus adapter.

The Storage Manager software provides the multi-path support for a Fibre Channel, iSCSI, or SAS host bus adapter device driver based on the Microsoft Storport miniport device-driver model. The Storport miniport device-driver model was introduced in the Microsoft Windows Server 2003 release as a replacement for the SCSIport miniport device driver model. It is the only supported device driver model for Windows Server 2003 x64 editions, which support the AMD64 and EM64T servers.

To support Storport miniport device drivers, Service Pack 2 and the latest Storport miniport hotfix must be installed in the Windows Server 2003 operating system. You can download the latest Storport hotfix from http://www.support.microsoft.com/kb/932755/en-us/. See the Storage Manager readme file for Microsoft Windows operating systems for other requirements, such as controller firmware versions or other Microsoft updates, and for information about the latest versions of the hotfix.

Storage Manager software for Windows Server 2003 and 2008

The Storage Manager software package contains the following components for Microsoft Windows Server 2003 and Windows Server 2008:

Storage Manager Agent (SMagent)

Storage Manager Utility (SMutil)

Storage Manager Client (SMclient)

Storage Manager redundant disk array controller (RDAC)

Note: The Storage Manager RDAC installs the multipath I/O (MPIO) Device Specific Module (DSM) multipath driver for Windows. The Storage Manager RDAC is different from the Linux multipath proxy (MPP) multipath driver.

Install the Storage Manager Client (SMclient) on the management station. Installing the SMclient on a host management station is optional.

Install the following components only on the host:

Storage Manager RDAC (MPIO)

Storage Manager Agent (SMagent)

Storage Manager Utility (SMutil)

Installing the Storage Manager Host Software Packages

This section describes how to install different components of Storage Manager host software on management stations and host servers that run Windows operating systems.

Before you install the software, make sure that the management station or host server has at least 220 MB of available disk space for the installation software package, the temporary files during installation, and the final files after the installation.

Important: When you install SMclient on a stand-alone host and manage storage subsystems through the I/O path and not through the network, you must install the TCP/IP software on the host and assign an IP address to the host.

All four components of the Storage Manager host software, SMclient, Storage Manager RDAC (MPIO), SMagent, and SMutil, are packaged in a single host software package installer. You can install up to four packages by using this host software package installation program.

Note: The Microsoft Windows XP and Windows Vista operating systems support the Storage Manager Client package only. Do *not* install any other Storage Manager software packages on Windows XP or Windows Vista. MPIO is *not* supported on Windows XP or Windows Vista.

To install the Storage Manager host software packages on either a host server or management station, complete the following steps.

Important: These installation instructions are not for host servers with SAN-boot or Remote-boot disk. SAN-boot and Remote-boot disk are supported only by DS3300 and DS3400 storage subsystems.

1. Before you install this software, close all other programs.
2. Insert the *IBM System Storage DS3000 Support* CD into the CD drive

Note: If you downloaded the Storage Manager host software package for Windows from the DS3000 support Web site, you must extract the files from the downloaded .zip file into a specific directory From the Windows desktop, double-click the **My Computer** icon. The My Computer window opens.

3. Double-click the CD drive that contains the *Support* CD. The CD window opens.
4. Select the applicable directory on the CD for your operating system architecture.

Directory	Operating System
Windows_x86_32bit	Windows Server 2003 x86 32-bit edition (IA32) Server, Enterprise Edition and DataCenter Windows Server 2008 x86 32-bit edition (IA32) Server, Enterprise Edition and DataCenter
Windows_x64_64bit	Windows Server 2003 x64 64-bit edition (AMD64 and EM64T) Server, Enterprise Edition and DataCenter Windows Server 2008 x64 64-bit edition (AMD64 and EM64T) Server, Enterprise Edition and DataCenter

5. To start the host software package installer, double-click the executable (.exe) file, such as SMIA-WSxxx.exe. The InstallAnywhere Installation wizard window opens while the software is loading. Follow the instructions in each window of the wizard.

6. When the Select Installation Type window opens, you can select one of the following options

   Typical (Full Installation): This default selection installs all of the packages on the system. This is a safe choice if you do not know which installation type to select. This option installs all four host software components: SMclient, MPIO, SMagent, and SMutil. This is the default selection for a host running a supported server operating system such as Microsoft Windows Server, SUSE Linux Enterprise Server, or Red Hat Enterprise Linux.

   Management Station: This selection installs the software that is required to configure, manage, and monitor a storage subsystem. This option is for the workstation or management computer. This option installs the SMclient program only. This is the default selection for a host running a supported client operating system such as Microsoft Windows XP or Windows Vista.

   Host: This selection installs the storage subsystem server software. Use this type of installation for the host (server) that is connected to the storage subsystem (Windows Server 2003 and Windows Server 2008 only). This option installs all host software components except the SMclient program.

   Custom: This selection enables you to customize the features that are to be installed.

Important: You must install the MPIO package before you can install and use the Storage Manager Agent.

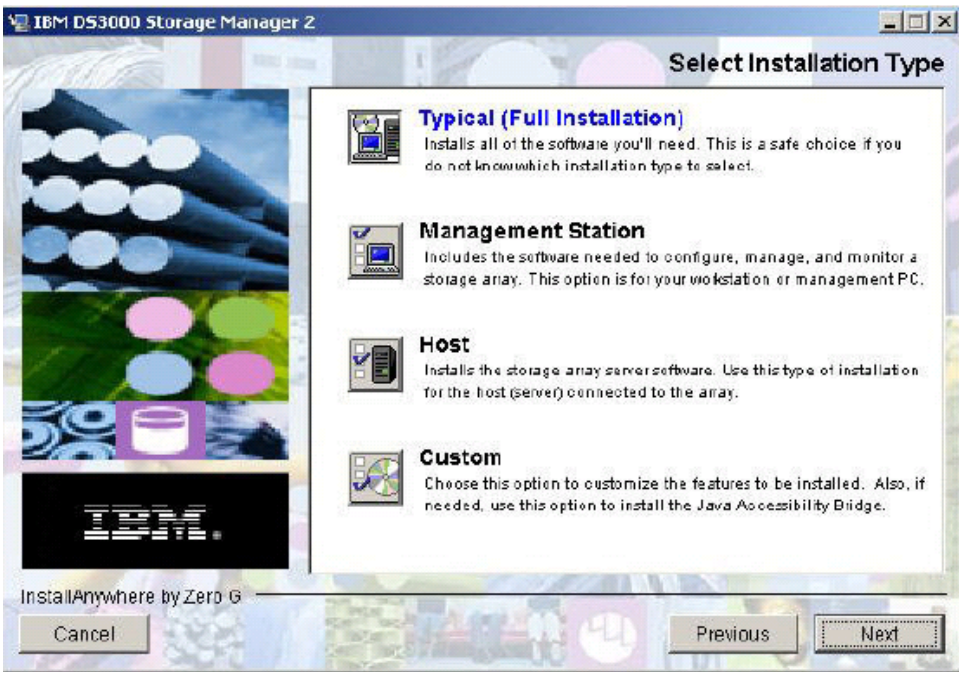

Figure 88:Select Installation Type window

7. Click **Next**. If there are previously installed Storage Manager software packages, the Overwrite Warning window opens.

8. If the Overwrite Warning window opens, click **OK**. The Automatically Start Monitor window opens. This is the event monitor service that will monitor the specified DS3000 storage subsystems and forward any critical alerts from those subsystems when the SMclient program is not running.

9. On the Automatically Start Monitor window, select the applicable option for your system

Notes:

To enable automatic ESM firmware synchronization, you must enable the Event Monitor. Select **Automatically Start Monitor**. For more information, see "Automatic ESM firmware synchronization"

If you start the Event Monitor Service on multiple systems, you might receive duplicate error messages from the same storage array. To avoid receiving duplicate error messages, start the Event Monitor on only one system that will run continuously.

10. Click **Next**. The Pre-Installation Summary window opens.

11. Click **Install**. The Installing IBM DS3000 Storage Manager window opens while the software is being installed. The Installation/Remove status window might also open throughout the installation process. After the software is installed, the Install Complete window opens.

Important: If you cancel an installation before the installation is completed (while the progress bar is still visible), the installation might not clean up the canceled installation correctly, and the host-software installation wizard creates an installation log. Also, the Add/Remove Program window may show that the program is already installed. However, when you try to uninstall it, an uninstallation error is displayed, and you are asked whether you want to remove the entry from the Add/Remove program list. Click **Yes** to remove the entry. Then, you must delete the .xml file from one of the following directories:

> **For Windows Server 2003 32-bit edition or Windows Server 2008 32-bit edition:** Windows_boot_drive_letter:\Program Files\Zero G Registry
>
> **For Windows Server 2003 64-bit edition:**
> Windows_boot_drive_letter:\Program Files (x86)\Zero G Registry

The Windows_boot_drive_letter:\Program Files\Zero G Registry directory might be hidden from the normal Windows Explorer view. If this is the case, change the settings of the Windows Explorer to **Show hidden files and folders**. If you cancel the installation *before* the progress bar is visible, the installation wizard cleans up the canceled process correctly. You do not have to do anything before you start the wizard again.

12. Make sure that the **Yes, restart my system** option is selected. Click **Done**.

Several files and program packages are installed to the directory that you specified previously in the procedure. The default directories are as follows:

> **For Windows Server 2003 32-bit edition or Windows Server 2008 32-bit edition:** Windows_boot_drive_letter:\Program Files\IBM_DS3000
>
> **For Windows Server 2003 64-bit edition:**
> Windows_boot_drive_letter:\Program Files (x86)\IBM_DS3000

The installation is completed, and Windows is restarted.

Important: If you repeatedly cancel an installation or uninstallation before the process is completed and you try to install the software again, the installation process might not work, and the software might not be installed after the installation process is completed. The Install Complete window indicates where the software is installed, but the software is not there. If this problem occurs, delete the .xml file from the applicable default directory

To make sure that the software packages were installed, go to "Verifying the installation"

Verifying the Installation

To make sure that the host-software package installer ran successfully, complete the following steps:

1. Select **Start** → **Settings** → **Control Panel**. The Control Panel window opens. Select **Add/Remove programs**. The Add/Remove Programs Properties window opens.
2. Find the IBM DS3000 Storage Manager Host Software entry. The entry has the applicable version number, for example, 02.17.x5.00.
3. If the installation was successful (no failure reported), go to "Completing the Storage Manager software, installation, and configuration". Otherwise, repeat the steps in "Installing the Storage Manager host software packages" on page 77. If the failure remains, contact your technical-support representative.

Creating or deleting logical drives

Use the following procedures to add or delete logical drives in a standard (noncluster) or cluster configuration.

Standard (noncluster) configuration for Windows Server 2003 and Windows Server 2008

Use the following procedures to create or delete logical drives in a standard configuration.

Create Logical Drives

When you create logical drives with the Storage Manager software, you must add the new logical drives to the Windows Server 2003 or Windows Server 2008 operating system. Each logical drive is recognized by Windows Server as a single disk drive.

After you create logical drives, to add them to the Windows Server 2003 operating system, complete the following steps:

1. Right-click **My computer**; then, click **Manage** → **Device Manager** → **Disk Drives**.
2. Click **Actions** → **Scan for new or removed hardware**.

After you add the logical drives, run the SMdevices utility that is provided with the Storage Manager software. The SMdevices utility identifies logical drives by their associated operating-system device names. For more information, see "Using the SMdevices utility" and the online help.

Note: For Windows Server 2008, before disk initialization, make sure that the disk device is Online. In the Computer Management window, click **Disk Management**. The disk devices are displayed in the bottom area of the window. Right-click the disk device that you want to put online and select **Online**.

Deleting logical drives

Attention: You must use the Disk Administrator to delete the cluster physical disk resources, delete operating-system partitions, and unassign drive letters before you

delete logical drives or reset the configuration in the Storage Manager software. This action avoids damage to the registry information.

Before you delete logical drives with the Storage Manager, complete the following steps:

1. Use the Disk Administrator to delete any partitions and to unassign drive letters that are associated with the logical drives.
2. If possible, restart the system to remove the configuration information

Booting from a SAN with the DS3300

One of the most common solutions for all SAN products is a remote boot (or Boot from SAN) solution. This is done to get a lower cost of solution for disaster recovery. It also allows for lower TCO by removing the need for internal drives which have a low disk drive utilization. You are able to move to shared disk drives which can support multiple server installations allowing for greater disk drive utilization.

When planning to boot from a SAN it is important to research the components and operating systems involved. When using a QLogic iSCSI HBA or any other brand HBA you should check the manufacturer's website for information and firmware updates required to boot from a SAN. Additionally, you should check the operating system website for boot-from-SAN information. Operating system sites such as Microsoft and VMware will post boot-from-SAN information and patches. When searching these websites suggested keywords might be: boot-from-SAN, QLogic, VMware, or Windows 2008. These are only suggestions and not the full set of search words or vendors.

To enable boot from SAN capabilities with the DS3300 you would do the following:

1. At the configure tab, select Create Logical Drives

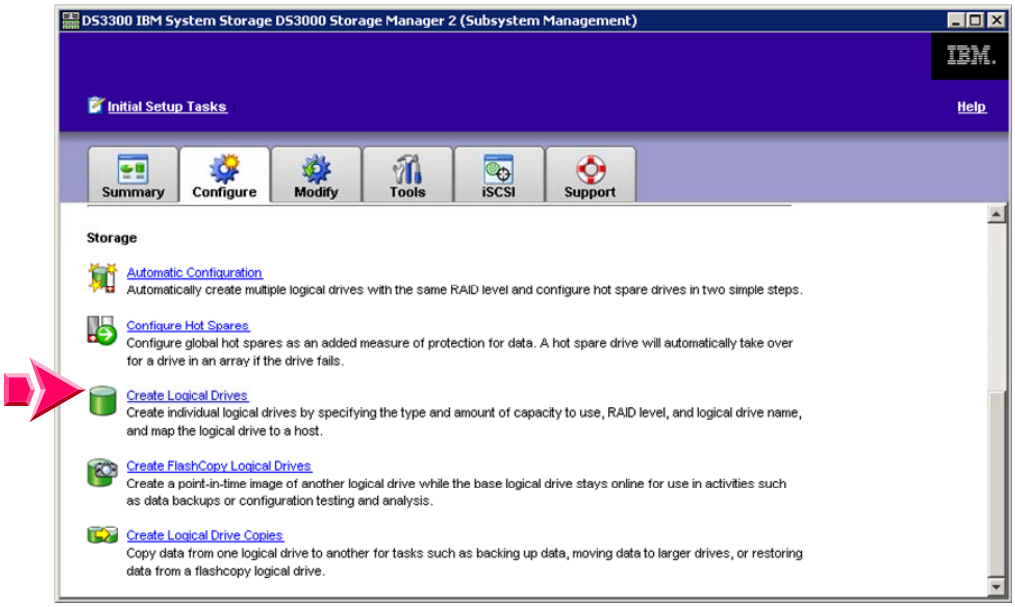

Figure 89: IBM DS3000 Storage Subsystem Window

3. Install a QLogic iSCSI Single Port PCIe HBA for IBM System x or a QLogic iSCSI Dual Port PCIe HBA for IBM System x in the host
4. Boot the host up and press Ctrl-Q when the banner appears stating "Press <Ctrl-Q> for Fast!Util"
5. Once you are in the Fast!Util configuration screen, select the HBA port to boot from

6. Select Configuration Settings
7. Select Host Adapter Settings
8. Select Initiator IP Settings
9. Enable either the IPv4 or IPv6 options depending on if you are booting from an IPv4 or IPv6 target
10. Determine if you are going to obtain a DHCP lease. If you do skip to step 14. Otherwise continue to step 11
11. Select IPv4 address and type in the static IP address that you wish to use
12. Select the IPv4 subnet mask that you wish to use
13. Select the IPv4 gateway that you wish to use
14. Select ESC
15. Select ESC again
16. Select iSCSI Boot Settings
17. Select Adapter Boot Mode and set it to Manual
18. Select Primary Boot Device Settings
19. Select IPv4 or IPv6 for the boot device
20. Select Target IP and enter the IP address of the target portal to boot from
21. Select ESC
22. Select ESC again
23. Select Save Changes
24. Scroll down to reinitialize adapter and select Enter
25. Return to the DS3000 Storage Manager window

26. Return to the Configure tab and select Configure Host Access (Manual)

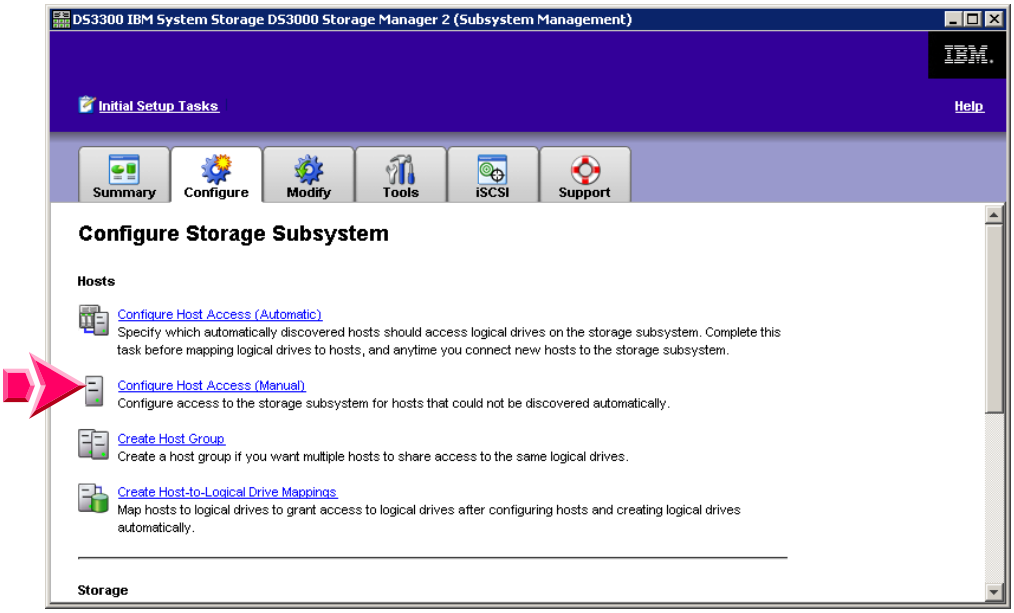

Figure 90: IBM DS3000 Storage Subsystem Window

27. Enter a unique name for the host, select the OS type that will be installed, click Next

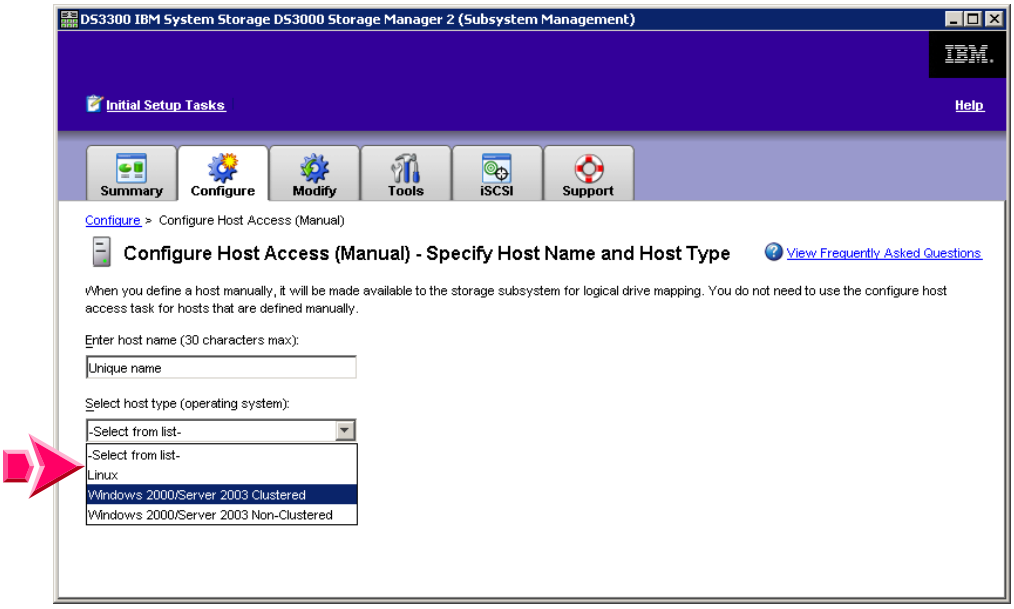

Figure 91: IBM DS3000 Configure Host Access Window

28. Select the adapter iqn listed under “Known iSCSI Initiators”

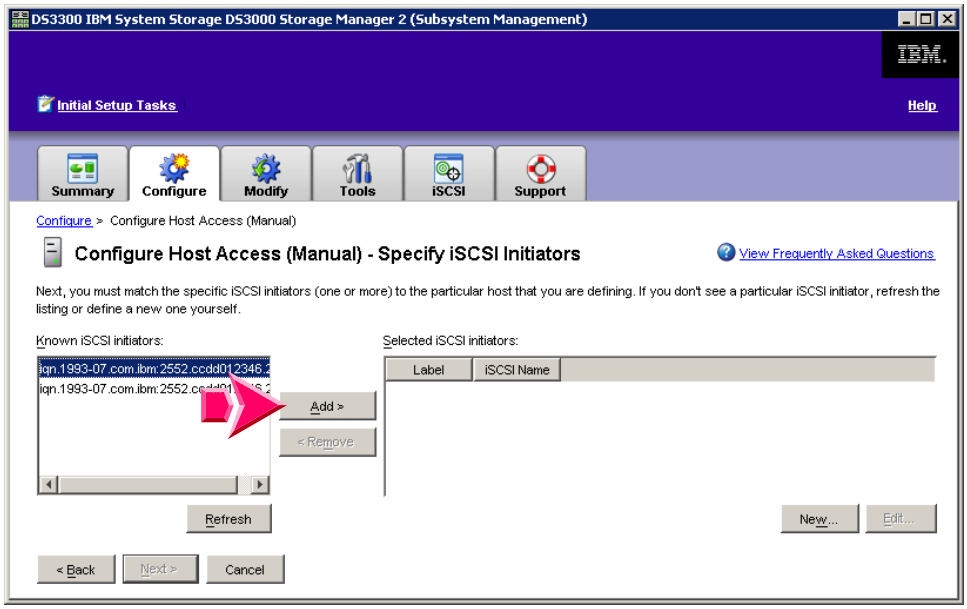

Figure 92: IBM DS3000 Configure Host Access Window

29. Select Add to move it to the selected iSCSI initiator column.

30. Select Next, Next again, and then Finish.

31. Return to the Configure Tab and select Create Host-to-Logical Drive Mappings

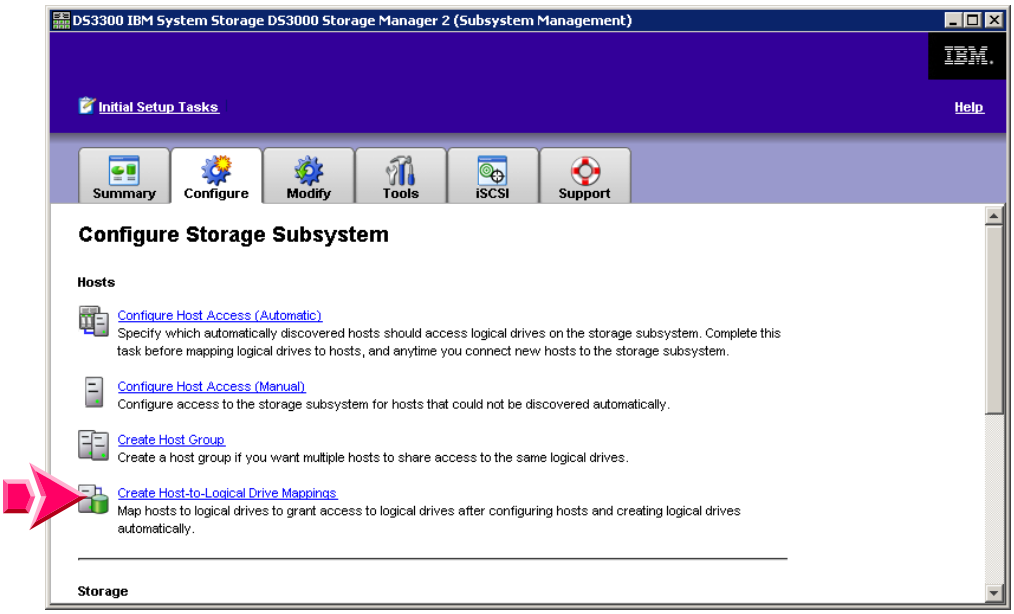

Figure 93: IBM DS3000 Storage Subsystem Window

32. Select the host that was created in steps 25 through 30

33. Select Next

34. Select the logical drive created in steps 1 and 2
35. Select Finish
36. Return to the Fast!Util on the host
37. Select Configuration Settings
38. Select iSCSI Boot Settings
39. Select Primary and then Enter
40. The device list will then appear. Select the controller and logical drive to boot from
41. Select Enter to save the changes
42. Press ESC and select reboot the host
43. Begin the OS installation

Note: If installing Windows 2003, press F6 to add the device driver support disk driveette for the QLogic iSCSI Single Port PCIe HBA for IBM System x or the QLogic iSCSI Dual Port PCIe HBA for IBM System x. Windows 2008 includes an embedded driver that can be used during the OS install.

Appendix A: Exchange Best Practice

For Exchange it is best to use a RAID 1 for the Log DB and a RAID 10 for the Storage Groups. The number of drives per Storage group is dependent on the number of users. It Is recommended to use one storage group per RAID system and one log file per RAID system. Use more storage groups (Exchange 2003/2007) to achieve more performance. Create a storage group for each new database until the maximum number of storage groups has been created. Spread the load of mailboxes across as many stores and storage groups as possible. This will lead to more granularities and improves performance, manageability and restore times.

Aligning Exchange 2003 I/O with Storage Track Boundaries

Windows 2003 has an internal structure called the master boot record (MBR) that limits the maximum number of hidden sectors to 63. Specifically, the partition table inside of the MBR maintains Starting Sector and Ending Sector values, which are only six bits in length. Therefore, their maximum value is 63 due to the limited number of bits and the fact that sector enumeration begins at 1, not 0. This characteristic of the MBR causes the default starting sector for disk drives that report more than 63 sectors per track to be the 64th sector. As a result, when programs transfer data to or from these disk drives, misalignment can occur at the track level, with allocations beginning at a sector other than the starting sector. This misalignment can defeat system optimizations of I/O operations designed to avoid crossing track boundaries. Additional information regarding the MBR and its internal structures can be found in the Disk drive Concepts and Troubleshooting article at

http://technet.microsoft.com/en-us/library/aa998219.aspx

In the case of a physical disk drive that maintains 64 sectors per track, Windows 2003 always creates the partition starting at the 64th sector, thus misaligning it with the underlying physical disk drive. To be certain of disk drive alignment, you can use disk drivepart, which is a utility provided by Microsoft in the Windows 2000 Resource Kit. disk drivepart is a command line utility that has the ability to explicitly set the starting offset in the master boot record (MBR). By setting the starting offset, one can ensure track alignment and improve disk drive performance. Microsoft Exchange 2000 server writes data in multiples of 4 KB I/O (4 KB for the databases and up to 32 KB for streaming files), so it is important that the starting offset be a multiple of 4KB. Failure to do so may result in a single I/O spanning two tracks, causing performance degradation.

Figure 94 shows a partition created without using disk drivepart, resulting in a partition that is not aligned with the underlying physical disk drive.

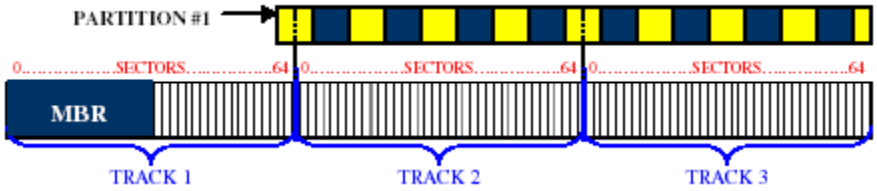

Figure 94: View of a Misaligned Partition that was not created with disk drivepart

Breaking out the partition into 4 KB blocks, as the colors do above, clearly shows that one out of every eight blocks would be straddling a track boundary. This will greatly reduce the performance of the physical disk drive, and thus the application.

This figure shows the case where disk drivepart was used to create the partition along a track boundary of the underlying physical disk drive.

Figure 95: View of a properly Aligned Partition Created with disk drivepart

In this case, there are no 4 KB blocks straddling a track boundary.

Open a command prompt and run the utility disk drivepart:

```
C:\>disk drivepart

Microsoft Disk drivePart version 5.2.3790.3959
Copyright (C) 1999-2001 Microsoft Corporation.
On computer: SANLABWSX64

DISK DRIVEPART>
```

Now to find out which disk drive we should use, use the “list disk drive” command

```
DISK DRIVEPART> list disk drive


  Disk drive ###  Status      Size     Free     Dyn  Gpt
  --------  ----------  -------  -------  ---  ---
  Disk drive 0    Online        50 GB    50 GB
  Disk drive 2    Online       140 GB  8033 KB
```

As you can see, disk drive 0 has 50GB of free space. This disk drive has not yet been partitioned. If you open the disk drive Management console, disk drive 0 is the same as the created DS3300 disk drive. You can also confirm this using the “detail disk drive” command.

```
DISK DRIVEPART> select disk drive 0
Disk drive 0 is now the selected disk drive.
DISK DRIVEPART> detail disk drive
IBM 1726-3xx  FAStT  Multi-Path Disk drive Device
Disk drive ID: 9BC170BE
Type   : iSCSI
Bus    : 0
Target : 0
LUN ID : 0
There are no volumes.
```

To create a partition with alignment, it is necessary to use the create partition command with the align option. This option specifies the offset in KByte.

To achieve the best performance with Microsoft Exchange we recommend the setting of the partition alignment to 64Kbyte.

Now this is how the command looks like:

```
DISK DRIVEPART> create partition primary align=64


Disk drivePart succeeded in creating the specified partition.
```

The disk drivepart is now complete.

In the Windows Disk drive Management console we see now a disk drive ready for formatting.

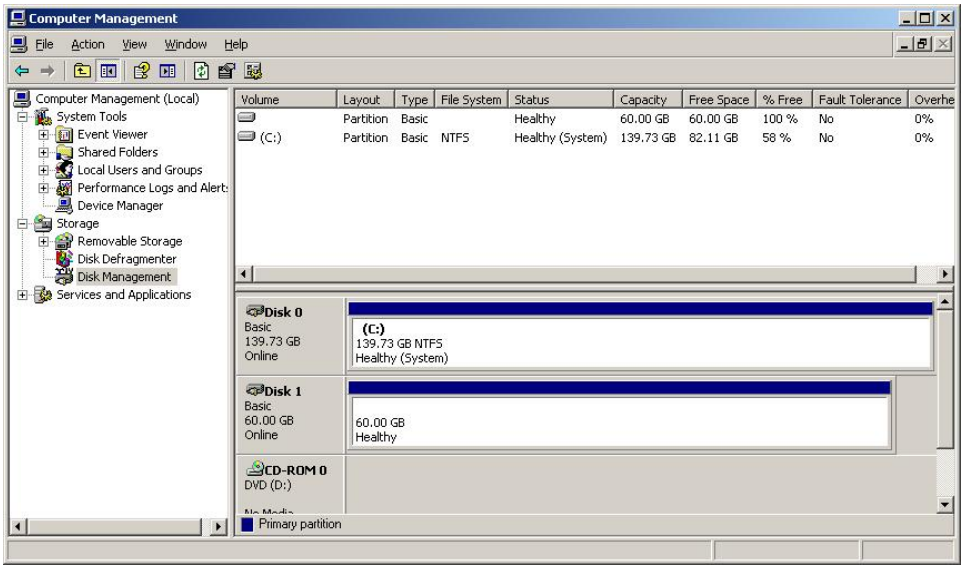

Figure 96: Windows Disk drive Management Window

You need to assign a drive letter to the partition before formatting. :

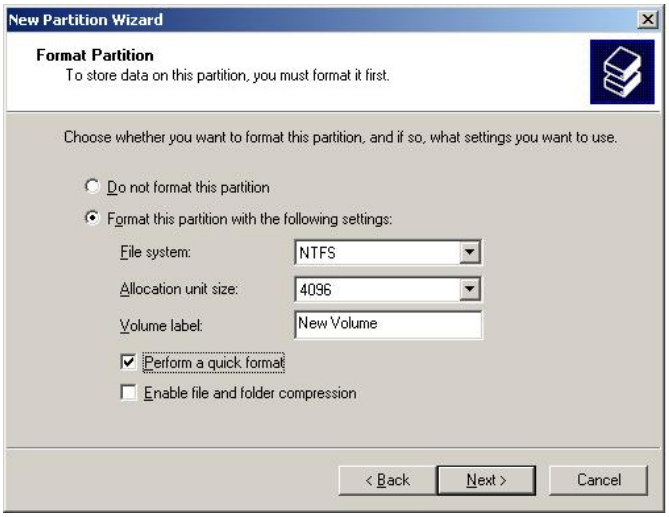

Figure 97: Microsoft New Partition Wizard – Format Partition Window

The Microsoft utility Jetstress can be used to measure performance for an Exchange 2003 system. You can use System Monitor, Event Viewer, and Exchange Server Database Utilities together with Jetstress to verify that your disk drive subsystem meets or exceeds the performance criteria you establish. You should never run Jetstress on a live production server. Don't use Jetstress to measure disk drive performance on a test server whose hardware is significantly different from production equipment's hardware.

Appendix B: References

General DS3300 Support Documents URL, http://www-947.ibm.com/systems/support/supportsite.wss/selectproduct?taskind=7&brandind=5000028&familyind=5356498&typeind=0&modelind=0&osind=0&psid=sr&continue.x=1

IBM Redbooks URL, http://www.redbooks.ibm.com/abstracts/sg247065.html?Open

IBM, 2008, *Installation and Support Guide for Microsoft Windows Server, Linux, Novell Netware and VMware ESX Server*, ftp://ftp.software.ibm.com/systems/support/system_x_pdf/46m1364.pdf

Microsoft, 2007. *How to Align Exchange I/O with Storage Track Boundaries*, http://technet.microsoft.com/en-us/library/aa998219.aspx

Microsoft, 2007. *Planning Storage Configurations for Exchange 2007*, http://technet.microsoft.com/en-us/library/bb124518.aspx

Microsoft, 2008. *Microsoft Storage Technologies – iSCSI*, http://www.microsoft.com/windowsserver2003/technologies/storage/iscsi/default.mspx (FAQs and more information on Microsoft iSCSI)

Microsoft Download Center, 2007. *Microsoft iSNS Server 3.0*, http://www.microsoft.com/downloads/details.aspx?familyid=0dbc4af5-9410-4080-a545-f90b45650e20&displaylang=en

Microsoft Exchange Team Blog, 2007. *Exchange 2007 Mailbox Server Role Storage Requirements Calculator*, http://msexchangeteam.com/archive/2007/01/15/432207.aspx

Cisco, undated. *Flooding Control*, http://www.cisco.com/univercd/cc/td/doc/product/lan/c2900xl/29_35xu/olhelp/floodctl.htm

Appendix C: DS3300 Setup Worksheet

IBM DS3300 SETUP WORKSHEET

Customer Location:	
Storage Subsystem Name:	
Machine Type:	1726
Serial Number:	
Date Installed:	

Storage Subsystem Name	Management Method	Controller: Ethernet & IP addresses and host name		Host IP address and host name
		Controller A	Controller B	
DS3300 Unit Finance EXAM	Direct PLE	Hardware Ethernet address = 00a0b8020420 IP address = 192.168.128.101 Host = Denver_a	Hardware Ethernet address = 00a0b80000d8 IP address = 192.168.128.102 Host = Denver_b	EXAM
Engineering EXAM	Host-Agent PLE	EXAM	PLE	IP address = 192.168.2.22 Host = Atlanta
DS3300 Unit		Hardware Ethernet address = IP address = Host =	Hardware Ethernet address = IP address = Host =	
				IP Address = Host =
				IP Address = Host =
				IP Address = Host =
				IP Address = Host =
				IP Address = Host =

Appendix D: Terminology

Bandwidth – The amount of data the storage subsystem can process over time. Bandwidth is measured in megabytes per second (MB/s).

Initiator – The system component that originates an input/output (I/O) command over an I/O bus or network. I/O adapters, network interface cards, and intelligent controller device I/O bus control application specific integrated circuits (ASICs) are typical initiators. A peripheral device is a target. (*The Dictionary of Storage Networking Terminology*)

Latency – The interval of time between submitting a request and receiving a response.

MPIO (Multi Path I/O) – In computer storage, an arrangement whereby there is more than one logical path between the central processing unit (CPU) in a computer system and its storage devices. This path is routed through the buses and bridge devices that connect the CPU and its storage devices. If one controller, port, or switch fails, the operating system can route I/O through the remaining controller so that work can continue.

NIC (Network Interface Card) – An adapter that connects an intelligent device to a network. Usually called a network interface card or Ethernet NIC network.

Node – An addressable entity connected to an input/output (I/O) bus or network. Used primarily to refer to computers, storage devices, and storage subsystems. The component of a node that connects to the bus or network is a port. (*The Dictionary of Storage Networking Terminology*).

Portal – A service that links initiators and targets when their IP addresses are not known.

Response time – The interval of time between submitting a request and receiving a response.

Session – A group of iSCSI connections. As each connection is established, it can negotiate its own unique parameters.

Target – The storage destination. In this document, the target is the IBM System Storage™ DS3300.

TCP/IP (Transmission Control Protocol/ Internet Protocol – Shorthand for a suite of protocols that includes Transmission Control Protocol (TCP), Internet Protocol (IP), User Datagram Protocol (UDP), and Internet Control Message Protocol (ICMP). This is the

basic set of communication protocols used on the Internet. (*The Dictionary of Storage Networking Terminology*)

Throughput – The number of individual I/Os the storage subsystem can process over time. Throughput is measured in I/Os per second (IOPS).

TOE (TCP/IP Offload Engine) – A NIC that offloads processing of TCP/IP segments from the host CPU. For systems with high CPU usage, a TOE might provide better performance because it provides more CPU cycles that can be used for other processes.

Legal Disclaimers and Trademarks

This publication is provided "AS IS." IBM product information is subject to change without notice.

No part of this document may be reproduced or transmitted in any form without written permission from IBM Corporation. Product data has been reviewed for accuracy as of the date of initial publication. Product data is subject to change without notice.

This document could include technical inaccuracies or typographical errors. IBM may make changes, improvements or alterations to the products, programs and services described in this document, including termination of such products, programs and services, at any time and without notice. Any statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only. The information contained in this document is current as of the initial date of publication only and is subject to change without notice. IBM shall have no responsibility to update such information.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products in connection with this publication and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products.

IBM makes no representations or warranties, expressed or implied, regarding non-IBM products and services, including those designated as ServerProven.

IBM is not responsible for the performance or interoperability of any non-IBM products discussed herein. Performance data for IBM and non-IBM products and services contained in this document was derived under specific operating and environmental conditions. The actual results obtained by any party implementing such products or services will depend on a large number of factors specific to such party's operating environment and may vary significantly. IBM makes no representation that these results can be expected or obtained in any implementation of any such products or services.

MB, GB and TB equal 1,000,000, 1,000,000,000 and 1,000,000,000,000 bytes, respectively, where referring to storage capacity. Actual storage capacity will vary based upon many factors and may be less than stated. Some numbers given for storage capacities give capacity in native mode followed by capacity using data compression technology.

THE INFORMATION IN THIS DOCUMENT IS PROVIDED "AS-IS" WITHOUT ANY WARRANTY, EITHER EXPRESSED OR IMPLIED. IBM EXPRESSLY DISCLAIMS ANY

WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NONINFRINGEMENT. IBM products are warranted according to the terms and conditions of the agreements (e.g., IBM Customer Agreement, Statement of Limited Warranty, International Program License Agreement, etc.) under which they are provided.

References in this document to IBM products, programs or services does not imply that IBM intends to make such products, programs or services available in all countries in which IBM operates or does business. Consult your local IBM business contact for information on the product or services available in your area. Any reference to an IBM program or product in this document is not intended to state or imply that only that program may be used. Any functionally equivalent program or product that does not infringe IBM's intellectual property rights may be used instead. It is the user's responsibility to evaluate and verify the operation of any non-IBM product, program or service.

IBM's customer is responsible for ensuring its own compliance with legal requirements. It is the customer's sole responsibility to obtain advice of competent legal counsel as to the identification and interpretation of any relevant laws and regulatory requirements that may affect the customer's business and any actions the customer may need to take to comply with such laws. IBM does not provide legal advice or represent or warrant that its services or products will ensure that the customer is in compliance with any law.

The provision of the information contained herein is not intended to, and does not, grant any right or license under any IBM patents or copyrights. Inquiries regarding patent or copyright licenses should be made, in writing, to:

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785
U.S.A.