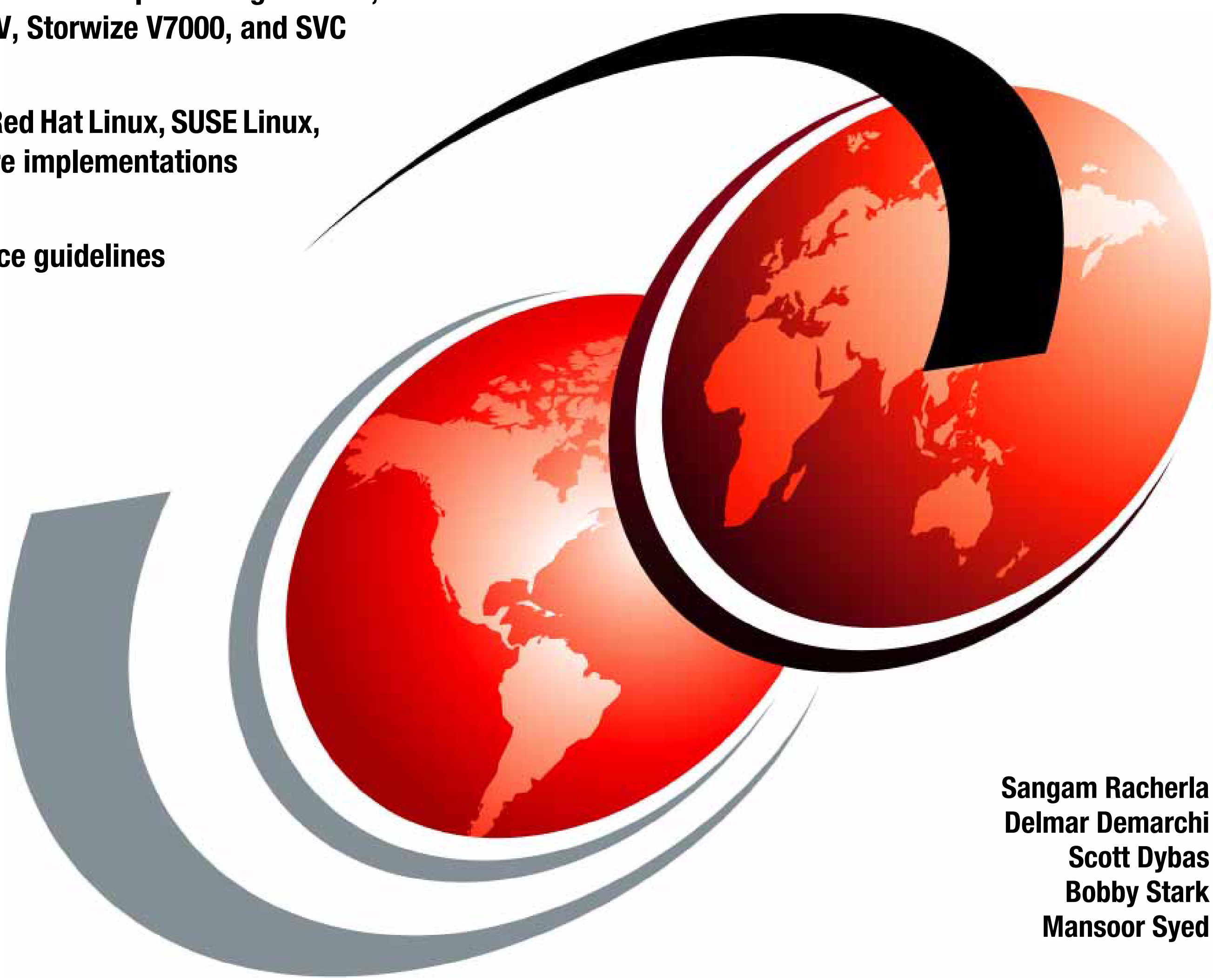
IBM

SAN Boot Implementation and Best Practices Guide for IBM System Storage

Implementation examples using DS5000, DS8000, XIV, Storwize V7000, and SVC

Windows, Red Hat Linux, SUSE Linux, and VMware implementations

Best Practice guidelines



Sangam Racherla
Delmar Demarchi
Scott Dybas
Bobby Stark
Mansoor Syed

Redbooks

ibm.com/redbooks

International Technical Support Organization

SAN Boot Implementation and Best Practices Guide for IBM System Storage

September 2012

SG24-7958-00

Note: Before using this information and the product it supports, read the information in “Notices” on page ix.

First Edition (September 2012)

- IBM System Storage DS5000 with Storage Manager 10.70.x (Firmware 7.70.x)
- IBM System Storage DS8000 with licensed machine code (LMC) level 6.6.xxx.xx (bundle version 86.0.xxx.xx)
- IBM XIV Storage System Hardware Version 2.5 and IBM XIV Storage System Software Version 10.2.2
- IBM System Storage SAN Volume Controller Version 6.1.0
- IBM Storwize V7000 Version 6.1.0

© Copyright International Business Machines Corporation 2012. All rights reserved.
Note to U.S. Government Users Restricted Rights -- Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

Contents

© Copyright IBM Corp. 2012. All rights reserved.

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

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 and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

© Copyright IBM Corp. 2012. All rights reserved.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US registered or common law trademarks owned by IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at http://www.ibm.com/legal/copytrade.shtml

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

AIX 5L™
AIX®
BladeCenter®
DB2®
DS4000®
DS8000®
Easy Tier®
FlashCopy®
IBM®
Language Environment®
Netfinity®
POWER6+™
POWER6®
PowerHA®
POWER®
Redbooks®
Redbooks (logo) ®
Storwize®
System p®
System Storage DS®
System Storage®
System x®
System z®
Tivoli®
XIV®
xSeries®
z/OS®

The following terms are trademarks of other companies:

Intel Xeon, Intel, Intel logo, Intel Inside logo, and Intel Centrino logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

LTO, the LTO Logo and the Ultrium logo are trademarks of HP, IBM Corp. and Quantum in the U.S. and other countries.

Microsoft, Windows NT, Windows, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

NetApp, and the NetApp logo are trademarks or registered trademarks of NetApp, Inc. in the U.S. and other countries.

Java, and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Red Hat, the Shadowman logo, Red Hat Enterprise Linux, RHEL, Red Hat Network, and RHN are trademarks of Red Hat, Inc., registered in the United States and other countries.

Other company, product, or service names may be trademarks or service marks of others.

Preface

Booting servers from a storage area network (SAN) is being used increasingly in complex data center environments today, due to its significant benefits over the traditional method of booting from local disks. SAN Boot enables organizations to maximize consolidation of their IT resources, minimize their equipment costs, and realize the considerable management benefits of centralizing the boot process.

In SAN Boot, you can deploy diskless servers in an environment where the boot disk is located on (often RAID-capable) storage connected to the SAN. The server (initiator) communicates with the storage device (target) through the SAN using the Fibre Channel host bus adapter (HBA).

The system downtime is greatly minimized in case a critical component such as a processor, memory, or host bus adapter fails and needs to be replaced. The system administrator needs to swap only the hardware and reconfigure the *HBA's BIOS*, *switch zoning*, and *host-port definitions* on the storage server. The system image still exists on the logical drive, therefore the server is fully operational after the hardware swap and configuration change is completed.

This IBM® Redbooks® publication can help you with the SAN Boot implementation. We present various SAN Boot scenarios using IBM System Storage® products that include DS5000, IBM DS8000®, IBM XIV®, and SAN Volume Controller (SVC). The operating systems that are covered include Windows 2008, Red Hat Linux, SUSE Linux, and VMware.

The topics covered in this book are limited to the SAN Boot implementations. For any additional information, see the materials listed in "Related publications" on page 451.

The team who wrote this book

This book was produced by a team of specialists from around the world working at the International Technical Support Organization, San Jose Center.

Sangam Racherla is an IT Specialist and Project Leader working at the International Technical Support Organization (ITSO), San Jose Center. He holds a degree in Electronics and Communication Engineering and has eleven years of experience in the IT field. He has been with the ITSO for the past eight years and has extensive experience installing and supporting the ITSO lab equipment for various Redbooks publication projects. His areas of expertise include Microsoft Windows, Linux, IBM AIX®, IBM System x®, and IBM System p® servers, and various SAN and storage products.

Delmar Demarchi is an IT Storage Specialist from IBM LabService Brazil team with expertise in the areas of System p, AIX, IBM PowerHA®, DS3/4/5000, DS8000, SVC, V7000, XIV, and SAN products. Delmar was involved in various projects that typically include IBM storage, HA, and DRP solutions implementation using SAN switches and SAN routers, as well as XIV, IBM DS4000, DS5000, DS8000, and Copy Services. Delmar has also executed several storage and server migrations and implementations, and holds a MBA Diploma in Business and Information Technology from the Getulio Vargas Foundation from Brazil.

© Copyright IBM Corp. 2012. All rights reserved.

Scott Dybas is a Senior Software Engineer in IBM Tucson. He has 21 years of experience in performance roles including IBM Information Archive, IBM Tivoli® Storage Productivity Center (TPC), IBM Enterprise and LTO tape drives, as well as support roles that included IBM DS8000, XIV, SVC, and SDD offerings. His areas of expertise include performance analysis, storage area networks, and open systems storage solutions.

Bobby Stark is a Field Technical Support Specialist (FTSS) for IBM Federal Intelligence in the Washington, DC area, and has over 10 years experience in SAN installations. He holds a US Patent (http://www.pat2pdf.org/patents/pat7930529.pdf) in Boot from SAN Failover. Before joining IBM Federal Intelligence, he worked with IBM xSeries® Lab Service for three years as a Consultant performing xSeries and SAN installation, while also fixing customer SAN problems. He was on the development team for IBM BladeCenter® for three years, testing the BladeCenter with the IBM SAN in Kirkland, WA. He was also the Team Lead for the Microsoft Cluster Certification test team, at the Center for Microsoft Technologies in Kirkland, WA for four years.

Mansoor Syed is a Storage Support Specialist for the IBM Entry and Midrange Storage Systems, in Research Triangle Park, North Carolina. He provides hardware and software product support for IBM System Storage midrange and high performance computing systems as well as IBM virtualized disk system solutions. Mansoor joined IBM as a Client Advocate in 2004 and since has contributed in various technical support teams for products ranging from BladeCenter, xSeries, and System x servers to High Availability and High Performance Cluster Solutions. His responsibilities also involve enhancing RAS packaging of several IBM products, training of support engineers, and management of hardware infrastructure for Support teams. Presently, Mansoor provides technical leadership in the Product Field Engineering team for IBM System Storage DS5000 and DS3000 series.

Thanks to the following people for their contributions to this project:

Alex Osuna
Karen Orlando
Jon Tate
Bertrand Dufrasne
Mary Lovelace
Ann Lund

International Technical Support Organization, San Jose Center

Mark S. Fleming
Robert P. Sidur
Mario David Ganem
LIbor Miklas
Matus Butora
Hrvoje Stanilovic
Antonio Dell'Apa
Alexander Watson

IBM

Corne Lottering

Saudi Business Machines

Now you can become a published author, too!

Here's an opportunity to spotlight your skills, grow your career, and become a published author—all at the same time! Join an ITSO residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

Find out more about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

We want our books to be as helpful as possible. Send us your comments about this book or other IBM Redbooks publications in one of the following ways:

- Use the online **Contact us** review Redbooks form found at:

  ibm.com/redbooks

- Send your comments in an email to:

  redbooks@us.ibm.com

- Mail your comments to:

  IBM Corporation, International Technical Support Organization
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400

Stay connected to IBM Redbooks

- Find us on Facebook:

  http://www.facebook.com/IBMRedbooks

- Follow us on Twitter:

  http://twitter.com/ibmredbooks

- Look for us on LinkedIn:

  http://www.linkedin.com/groups?home=&gid=2130806

- Explore new Redbooks publications, residencies, and workshops with the IBM Redbooks weekly newsletter:

  https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm

- Stay current on recent Redbooks publications with RSS Feeds:

  http://www.redbooks.ibm.com/rss.html

1

Introduction to SAN Boot

Servers today use a variety of data storage solutions, such as DAS (direct attach), NAS (network attach) and SAN (storage area network) for the purpose of file and data storage. Historically, these servers used their own internal storage devices from which to boot the operating systems.

As organizations grew to large scale data centers that can have up to thousands of systems, it became necessary to increase efficiency. One method was to take the local disk out of the servers and move it to shared storage devices. This method helped them increase efficiency, by creating centralized management of the operating systems as well as decreased power and cooling.

© Copyright IBM Corp. 2012. All rights reserved.

1.1 What SAN Boot is

SAN Boot is a technique that allows servers to utilize an operating system (OS) image installed on external SAN-based storage to boot up, rather than booting off their own local internal disk or direct attached storage.

1.1.1 SAN Boot concepts

Now that Internet Small Computer System Interface (iSCSI) has become more popular, we need to use the terms *remote boot* or *network boot* rather than *SAN Boot* because iSCSI is typically not a SAN. However, because the iSCSI and Fibre Channel worlds are merging together, we are using the term *SAN Boot* in this chapter to explain booting using both the Fibre Channel and iSCSI techniques.

1.1.2 Benefits of SAN Boot

Here are some benefits to the organization, by utilizing SAN Boot:

- Interchangeable servers:

  By allowing boot images to be stored on the SAN, servers are no longer physically bound to their startup configurations. Therefore, if a server happens to fail, it becomes very easy to replace it with another generic server and resume operations with the exact same boot image from the SAN (only minor reconfiguration is required on the storage subsystem). This quick interchange can help reduce downtime and increase host application availability.

- Provisioning for peak usage:

  Because the boot image is available on the SAN, it becomes easy to deploy additional servers to cope temporarily with high workloads.

- Centralized administration:

  SAN Boot enables simpler management of the startup configurations of servers. Rather than needing to manage boot images at the distributed level at each individual server, SAN Boot empowers administrators to manage and maintain the images at a central location in the SAN. This feature enhances storage personnel productivity and helps to streamline administration.

- Utilizing high-availability features of SAN storage:

  SANs and SAN-based storage are typically designed with high availability in mind. SANs can utilize redundant features in the storage network fabric and RAID controllers to ensure that users do not incur any downtime. Most boot images that are located on local disk or direct attached storage do not share the same protection. Using San Boot allows boot images to take advantage of the inherent availability built in to most SANs, which helps to increase availability and reliability of the boot image and reduce downtime.

- Efficient disaster recovery process:

  Assuming that data (boot image and application data) is mirrored over the SAN between a primary site and a recovery site, servers can take over at the secondary site in case a disaster destroys servers at the primary site.

- Reduced overall cost of servers:

  Placing server boot images on external SAN storage eliminates the need for local disk in the server, which helps lower costs and allows SAN Boot users to purchase servers at a reduced cost but still maintain the same functionality. In addition, SAN Boot minimizes the IT costs through consolidation, what is realized by electricity, floor space cost savings, and by the benefits of centralized management.

1.2 SAN overview

A SAN is categorized as a high-speed network, attaching servers and storage devices. For this reason, it is sometimes referred to as “the network behind the servers.” A SAN allows “any-to-any” connection across the network, using I/O elements such as routers, gateways, hubs, switches and directors.

For businesses, data access is critical and requires performance, availability, and flexibility. In other words, there is a need for a data access network that is fast, redundant (multipath), easy to manage, and always available. That network is a storage area network (SAN).

A SAN is a high-speed network that enables the establishment of switched, routed, or direct connections between storage devices and hosts (servers) within the specific distance supported by the designed environment. At the basic level, the SAN is a Fibre Channel (FC) network; however, new technology now enables this network to be routed or tunneled over many other networks as well.

The SAN can be viewed as an extension of the storage bus concept, which enables storage devices to be interconnected using concepts similar to that of local area networks (LANs) and wide area networks (WANs). A SAN can be shared between servers or dedicated to one server, or both. It can be local or extended over geographical distances.

For more information about designing SANs, see *Designing an IBM Storage Area Network*, SG24-5758.

1.3 What WWNNs and WWPNs are

In this section, we explain the concepts of World Wide Node Names (WWNNs) and World Wide Port Names (WWPNs).

1.3.1 Concepts

A WWNN is a World Wide Node Name; used to uniquely identify a device in a storage area network (SAN). Each IBM storage device has its own unique WWNN. For example, in a DS8000, each Storage Facility Image (SFI) has a unique WWNN. For SAN Volume Controller (SVC), each SVC Node has a unique WWNN.

A WWPN is a World Wide Port Name; a unique identifier for each Fibre Channel port presented to a storage area network (SAN). Each port on an IBM storage device has a unique and persistent WWPN.

IBM System Storage devices use persistent WWPN. It means that if a Host Bus Adapter (HBA) in an IBM System Storage device gets replaced, the new HBA will present the same WWPN as the old HBA. IBM storage uses a methodology whereby each WWPN is a child of the WWNN. It means that if you know the WWPN of a port, you can easily match it to the WWNN of the storage device that owns that port.

A WWPN is always 16 hexadecimal characters long. It is actually 8 bytes. Three of these bytes are used for the vendor ID. The position of vendor ID within the WWPN varies based on the format ID of the WWPN.

To determine more information, we actually use the first character of the WWPN to see which format it is (the vendor ID position is bold):

- 1 = IEEE 803.2 Standard Format (example 10:00:**00:00:c9**:2f:65:d6). It is an Emulex HBA WWPN because Emulex owns the 0000c9 company ID.
- 2 = IEEE 803.2 Extended Format (example: 21:00:**00:e0:8b**:90:90:04). It is a QLogic HBA WWPN because QLogic owns the 00e08b company ID.
- 5 = IEEE Registered Name Format (example 5**0:05:07:6**3:00:c7:01:99). It is an IBM WWPN because IBM owns the 005076 company ID, as shown in Example 1-1.

A typical Registered Name Format WWPN is shown in Example 1-1.

Example 1-1 Registered Name Format WWPN

```
50:05:07:63:00:c7:01:99.
```

We know that this WWPN uses the Registered Name Format because it starts with a ‘5’.

This Registered Name Format WWPN can be broken down into three sections (Table 1-1).

Table 1-1 Sections of a Registered Name Format WWPN

Section 1 Format ID	Section 2 –Vendor ID	Section 3 -Vendor Unique
5	0:05:07:6	3:00:c7:01:99

Here is a list of the Vendor IDs registered by the company:

- Vendor ID 001738 is registered to IBM (formerly registered to XIV).
- Vendor ID 005076 is registered to IBM.
- Vendor ID 00A0B8 is registered to Symbios Logic, now owned by LSI Logic, who manufacture the DS3000, DS4000® and DS5000.
- Vendor ID 0080E5 is registered to LSI Logic.

There are literally thousands of company IDs registered with the IEEE. To view the complete list, visit the following website.

http://standards.ieee.org/regauth/oui/oui.txt

1.3.2 IBM WWNN / WWPN range by type

In Table 1-2, we provide the WWNN and WWPN ranges by the IBM storage types.

Table 1-2 IBM WWNN / WWPN range by type

Machine type	WWNN	WWPN
DS3000/ DS4000/DS5000	20:0z:00:A0:B8:xx:xx:xy 20:0z:00:80:E5:xx:xx:xx	20:yy:00:A0:B8:xx:xx:xy 20:yz:00:80:E5:xx:xx:xx
2076 (IBM Storwize® V7000)	50:05:07:68:02:0x:xxxx	50:05:07:68:02:yx:xx:xx
2107 (DS8000 –all models)	50:05:07:63:0z:FF:Cx:xx	50:05:07:63:0z:yy:xx:xx
2145 (SVC)	50:05:07:68:01:0x:xxxx	50:05:07:68:01:yx:xx:xx
2810 (XIV)	50:01:73:8x:xx:xx:00:00	50:01:73:8x:xx:xx:yy:yy

In Table 1-2, x, y, and z have the following meanings:

- x: Determined by manufacturing and unique to each machine. On XIV it is the 5 digit serial number in hex.
- y: Determined by physical position within the machine.
- z: Determined by manufacturing.

1.4 What OS multipathing is

The term *multipathing* is a fault-tolerance technique used to spread the disks over multiple input/output (I/O) adapters. If you do not use multipathing, you will run into I/O path reliability problems trying to communicating to the SAN. Multipathing takes advantage of multiple paths between a host system and the storage subsystems (LUN) or set of LUNs, as shown in Figure 1-1. When an adapter fails, the system automatically reroutes I/O operations to another available path. This support also allows the balancing of the I/O load in multiple paths, preventing I/O bottlenecks.

Figure 1-1 demonstrates why multipath support is necessary. There are a number of single points of failure between the server bus and the first switch port. Beyond the switch port, there can be redundancy in both the fabric and the storage subsystem. Without multipathing, the internal bus cannot reroute I/O from a failed adapter.

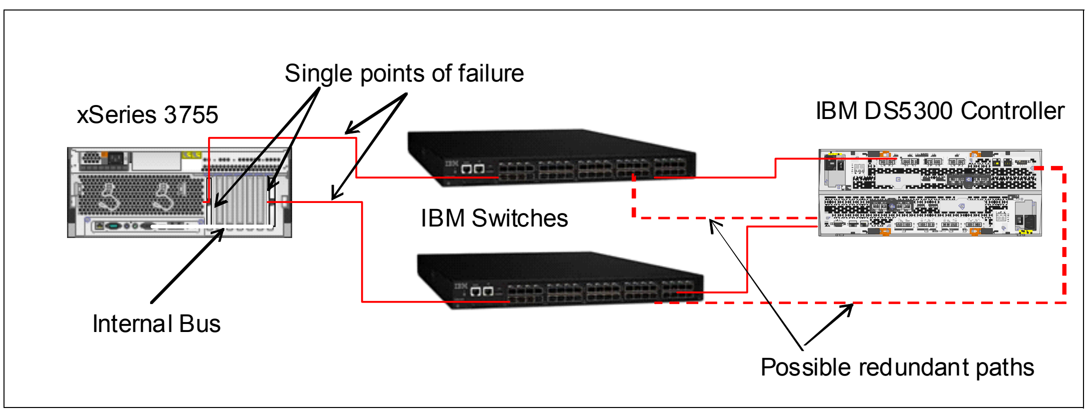


Figure 1-1 Possible failure points without multipathing

1.5 What switch zoning is

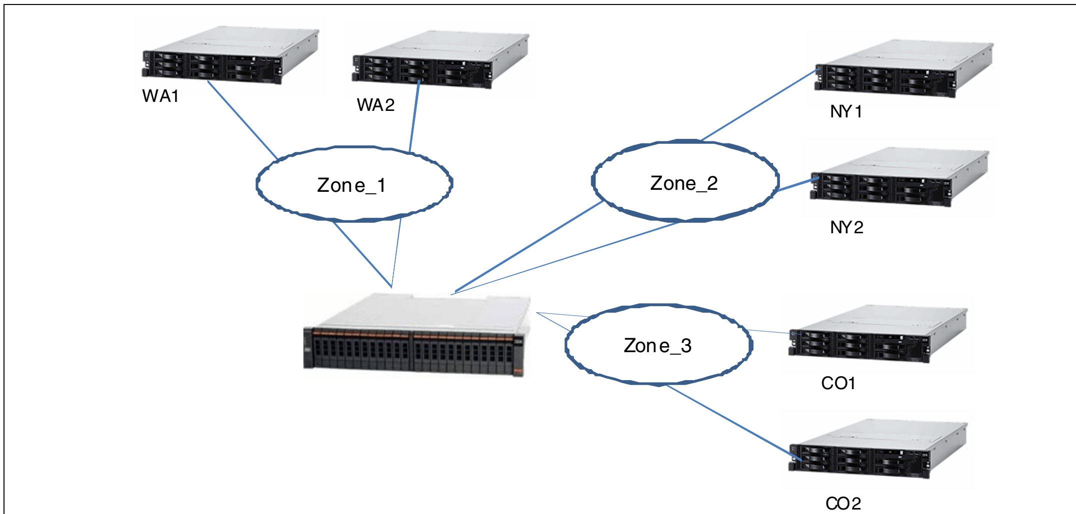
Zoning is mandatory when connecting FC hosts to a storage system through a Fibre Channel switch. By default, most FC switches come with (1) zone preconfigured. Zoning is configured on the SAN switch, and its purpose is to be boundary isolating and to restrict FC traffic to only those HBAs within a given zone.

A zone can be either a hard zone or a soft zone. Hard zones group the HBAs depending on the physical ports to which they are connected on the SAN switches. Soft zones group the HBAs depending on the World Wide Port Names (WWPNs) of the HBA. Each method has its merits, and you need to determine which is right for your environment.

Typically, you can use zones to do the following tasks:

- Provide security: Use zones to provide controlled access to fabric segments and to establish barriers between operating environments. For example, isolate systems with various uses or protect systems in a heterogeneous environment.
- Customize environments: Use zones to create logical subsets of the fabric to accommodate closed user groups or to create functional areas within the fabric. For example, include selected devices within a zone for the exclusive use of zone members, or create separate test or maintenance areas within the fabric.
- Optimize IT resources: Use zones to consolidate equipment logically for IT efficiency, or to facilitate time-sensitive functions. For example, create a temporary zone to back up non-member devices.

The diagram in Figure 1-2 shows a brief overview of a basic zoning configuration.



Zone Name	Aliases
Zone_1	WA1, WA2
Zone_2	NY1, NY2
Zone_3	CO1, CO2

Alias	WWPN
WA1	50:05:76:ab:cd:22:03:65
WA2	50:05:76:ab:cd:22:03:66
NY1	50:05:76:ab:cd:22:03:67
NY2	50:05:76:ab:cd:22:03:68
CO1	50:05:76:ab:cd:22:03:69
CO2	50:05:76:ab:cd:22:03:70

Figure 1-2 Basic zoning configuration

Important reminders:

Correct zoning helps avoid many problems and makes it easier to trace the cause of errors. Here are some examples of why correct zoning is important:

- Ensure that all zoning information is fully documented and that documentation is kept up to date. This information must be kept in a safe location and can be used for reference and planning purposes. If done correctly, the document can be used to assist in diagnosing zoning problems as well.
- When configuring World Wide Name (WWN) based zoning, it is important to always use the World Wide Port Name (WWPN), not the World Wide Node Name (WWNN). With many systems, the WWNN is based on the Port WWN of the first adapter detected by the HBA driver. If the adapter that the WWNN is based on happens to fail, and you based your zoning on the WWNN, then your zoning configuration becomes invalid. Subsequently, the host with the failing adapter then completely loses access to the storage attached to this switch.
- Keep in mind that you must update the zoning information if you ever need to replace a Fibre Channel adapter in one of your servers. Most storage systems such as the DS4000, Enterprise Storage Subsystem, and IBM Tape Libraries have a WWN tied to the Vital Product Data of the system unit, so individual parts can usually be replaced with no effect on zoning.

For more details on configuring zoning with your particular switch, see:

- *Implementing an IBM/Brocade SAN with 8 Gbps Directors and Switches*, SG24-6116
- *IBM SAN Survival Guide*, SG24-6143

1.6 What an HBA is

A Host Bus Adapter (HBA) is a Fibre Channel or iSCSI adapter card that can be added to the server in the PCIe or PCIx slots. It allows the server to connect to a Fibre Channel or iSCSI SAN or NAS storage subsystem.

1.6.1 Fibre Channel HBA

Fibre Channel HBAs are basically SCSI controllers that have the functionality to wrap the SCSI commands and data frames into Fibre Channel packets and transport them by Fibre Channel cables. Most HBAs can function attached to an arbitrated loop, or a Fibre Channel fabric, or in point to point mode directly attached to an array. There are several companies that manufactures HBAs, but in this book we only are using only two different companies in our testing; Emulex and QLogic.

For more information about what HBAs are supported, visit the IBM support website:

http://www-03.ibm.com/systems/xbc/cog

1.6.2 Fibre Channel architecture

The Fibre Channel architecture provides various communication protocols on the storage unit. The units that are interconnected are referred to as nodes. Each node has one or more ports.

A storage unit is a node in a Fibre Channel network. A host is also a node in a Fibre Channel network. Each port attaches to a serial-transmission medium that provides duplex communication with the node at the other end of the medium.

Storage unit architecture supports three basic interconnection topologies or network structures:

- Arbitrated loop
- Point-to-point
- Switched-fabric

> **Tip:** Only some IBM storage subsystems supports arbitrated loop, point-to-point, and switched-fabric topologies. Check the IBM Support website to find the correct configuration that can be used for your specific storage environment.
>
> See the *Installation and Planning Guide* for your storage subsystem for more information about the list of longwave and shortwave adapter cables and their distances.

1.6.3 iSCSI HBA

The iSCSI HBA is better known as the iSCSI initiator and can be either a physical iSCSI HBA inside a host server, or you can define a software iSCSI initiator by using an iSCSI stack on the Ethernet network adapter.

Software initiators

A configuration that uses software initiators includes the following components:

- Microsoft iSCSI Software Initiator or equivalent: The Microsoft Server 2008 already has the iSCSI software initiator built in.
- One or two Ethernet cards: There is no iSCSI card required on the server to implement a connection to an iSCSI SAN environment.

For more information about iSCSI on Microsoft, visit the Microsoft support pages:

http://technet.microsoft.com/en-us/library/ee338474(WS.10).aspx

A similar concept applies for UNIX platforms:

- The majority of UNIX systems (AIX, Linux, HP-UX, and so on) already have built-in SW packages to support iSCSI. Usually a certain level of OS is required to best utilize all features. Consult OS product support for details.
- Minimum one 10 Gbps Ethernet card for 10 Gbps iSCSI emulation and 10 Gbps capable LAN switches. In storage environments with required high data throughput, we do not recommend to utilize only 1 Gbps network.

Here are the software requirements for the most common UNIX platforms on IBM systems:

- AIX:

  http://www.ibm.com/developerworks/aix/library/au-iscsi.html

- Linux Red Hat:

  http://docs.redhat.com/docs/en-US/Red_Hat_Enterprise_Linux/6/html/6.1_Technical_Notes/index.html#iscsi-initiator-utils

- Linux SUSE:

  http://www.suse.com/documentation/sles10/book_sle_reference/?page=/documentation/sles10/book_sle_reference/data/sec_inst_system_iscsi_initiator.html

Hardware initiators

A configuration that uses hardware initiators includes the following components:

- One or two iSCSI cards for each server, which access the storage array and associated drivers.

  For BladeCenters, you use the QLogic iSCSI Expansion Card for IBM BladeCenter. The iSCSI Expansion Card option is a hardware initiator that provides iSCSI communication from the blade server to an iSCSI storage device. It delivers full TCP/IP Offload Engine (TOE) functionality to reduce CPU processing. For more information, see the Web link:

  http://www.ibm.com/common/ssi/rep_ca/4/897/ENUS105-194/ENUS105-194.PDF

  System x servers currently support these iSCSI TOE adapters:

  – QLogic QLE8142-SR-CK Dual port 10 GBps iSCSI HBA for System x
  – Emulex OCe10102-IX-D Dual port 10 Gbps iSCSI HBA for System x
  – QLogic QLE4060C 1 Gbps iSCSI Single Port PCIe HBA for IBM System x
  – QLogic QLE4062C 1 Gbps iSCSI Dual Port PCIe HBA for IBM System x
  – IBM iSCSI 1 Gbps Server TX Adapter (30R5201)
  – IBM iSCSI 1 Gbps Server SX Adapter (30R5501)
  – IBM iSCSI 1 Gbps Server Adapter (73P3601)

  For more information, see this website:

  http://www.ibm.com/support/entry/portal/docdisplay?brand=5000008&lndocid=MIGR-57073

- One or two Ethernet switches, preferably using two Ethernet switches for redundancy.

TCP Offload Engine benefits

We know that the processing of TCP packets from an Ethernet connection consumes many processor resources, and iSCSI protocol only adds another layer of processing. With the number of packets and their corresponding interrupts required for iSCSI, the software iSCSI packet processing can burden the host system with 50–65% CPU usage. Depending upon signaling options used, high CPU usage might even render certain host applications unusable.

Therefore, it is important that efficient iSCSI systems depend on a hardware TCP Offload Engine (TOE) to handle the transportation protocols of iSCSI. A TOE network interface card (NIC) is a dedicated interface card specifically designed for interfacing a server to the IP-SAN including iSCSI offloading and additionally TCP/IP encapsulation from the server processors. A hardware TOE implements the entire standard TCP and iSCSI protocol stacks on the hardware layer. This approach completely offloads the iSCSI protocol from the primary processors, leveraging storage communications efficiently and enabling applications to run faster and more reliable. By using the TCP Offload Engine, a single system can run multiple initiators for improved throughput.

Choosing between hardware and software initiators

Using hardware initiators offers the following key benefits:

- Their performance is noticeably faster.
- They do not interfere with data networking traffic if the network topology is designed to segregate storage traffic such as by using a separate set of switches to connect servers to iSCSI storage.
- The traffic that passes through them will not load the server's CPU to the same extent that might be the case if the storage traffic passed through the standard IP stack.
- It is possible to implement *iSCSI boot from SAN* with hardware initiators.

Using software initiators offers the following key benefits:

- The cost of the iSCSI hardware is avoided.
- It is possible to use one set of switches for both data and storage networking, avoiding the cost of additional switches but possibly impacting performance.
- It is possible to access other network storage devices such as NAS, NFS, or other file servers using the same network interfaces as are used for iSCSI.

Attention: The iSCSI software initiator support is limited to Ethernet adapters. Software initiators with physical Converged Network Adapters (CNA) are not supported.

1.6.4 Comparison: iSCSI versus Fibre Channel

The iSCSI protocol is a transport layer for SCSI over TCP/IP. Until recently, the standard IP protocol infrastructure (100 MBps Ethernet) was not able to provide the enough bandwidth and less latency to accommodate storage traffic. Dedicated infrastructure with respective communication protocols such as Fibre Channel Protocol (SCSI over Fibre Channel), was developed to achieve high volume data interchange over storage area networks. With the recent advances in Ethernet technology, it is now practical (from a performance and cost perspective) to access storage devices through an IP network. 10 Gigabit Ethernet is now widely available in many datacenters and the results are competitive to 4 and 8 Gbps Fibre Channel.

Similar to FC protocol, Fibre Channel over Ethernet (FCoE) and iSCSI allows storage to be accessed over a storage area network, allowing shared access to the devices. Opposed to a dedicated TCP and FC networks, the investment into Converged Network utilizing Converged Network Adapters (CNA) and convergence-capable LAN switches or SAN directors, clients significantly reduce the management cost by operating single network, thus saving on power, cooling, and floor space in the expensive datacenters. The key advantage of iSCSI over FCP is that iSCSI can utilize standard, off-the-shelf Ethernet network components. In addition, the network, that incorporates iSCSI SAN only, exploits a single kind of network infrastructure only (1 Gbps or 10 Gbps Ethernet) for both data and storage traffic, whereas use of FCP requires a separate type of infrastructure (Fibre Channel) and administration for the storage. Furthermore, FCoE and iSCSI based SANs can expand over arbitrary distances, and are not subject to distance restrictions that currently limit FCP. This concept helps clients consolidate their strategic datacenters with remote branch offices or departments into the single, centrally managed infrastructure.

Because an iSCSI and FCoE are designed to run on an IP network, it takes the advantage of existing features and tools that were already in place for IP networks. Today’s Ethernet network standards guarantee delivery of data and congestion control. Lossless Ethernet is one of the key requirements for implementation of storage networking on 10 Gbps IP-based networks. IPSec can be utilized to provide security for an iSCSI SAN, whereas a new security mechanism might need to be developed for the Fibre Channel. Service Location Protocol (SLP) can be used by iSCSI to discover iSCSI entities in the network. Thus, in addition to iSCSI running on standard, cheaper, off-the-shelf hardware, iSCSI also benefits from using existing, standard IP-based tools and services.

Figure 1-3 shows an example of typical datacenter networking utilizing FC and Ethernet components separately, opposed to the storage networking solutions that benefit from iSCSI or FCoE technology. As mentioned in the beginning of the chapter, we focus on FC and iSCSI configurations as FCoE is not natively supported by DS5000 family at the time of writing. There are no FCoE HICs available yet.

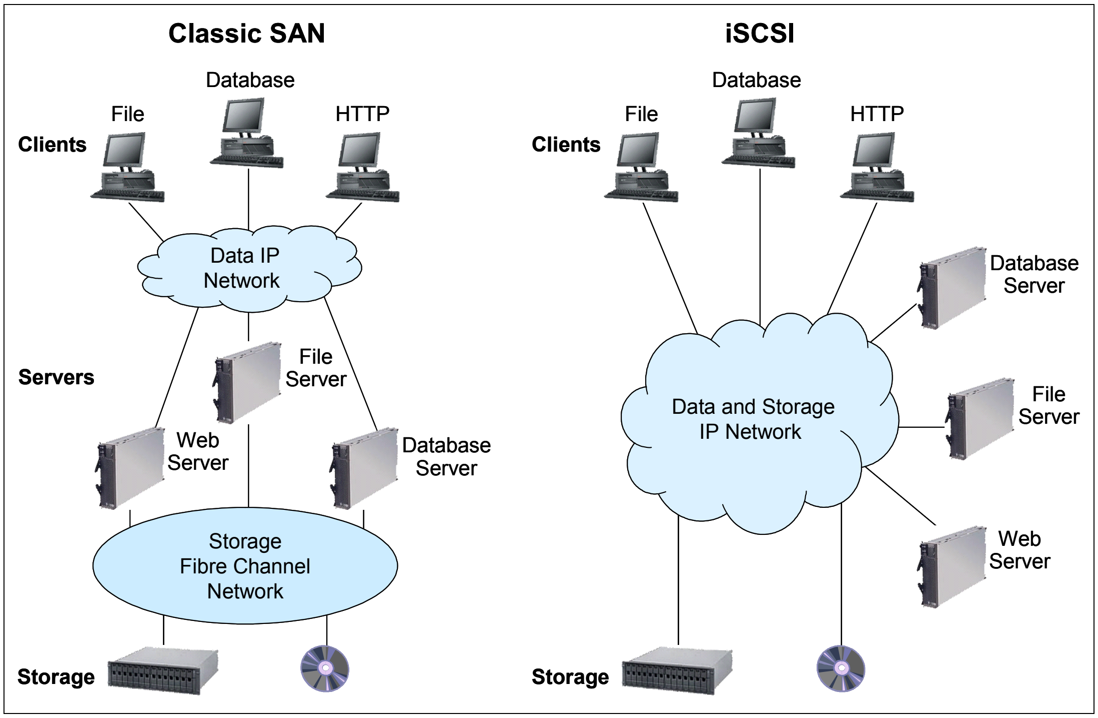


Figure 1-3 Typical FC SAN and an iSCSI SAN topology

Tip: Where the Boot from iSCSI or Fibre Channel SAN methods can serve and satisfy the requirements of several client environments, there are certain engineering and scientific applications that require the use of the local disk for better performance. Such applications might not be feasible for a Boot from iSCSI or Fibre Channel SAN solution; for example:

- Nastran (linear finite element analysis codes for engineering)
- Dytran (non-linear finite element analysis)
- Marc (another non-linear code)
- Fluent (fluid dynamics)
- Gaussian (computation chemistry)
- Amber (computational chemistry)
- GCG (computational chemistry)

Therefore, you need to consider the performance requirements for each application installed on your systems and analyze if those are feasible for SAN Boot.

1.7 SAN Boot implementation

Setting up and executing SAN Boot requires performing certain steps on both the server and the storage sides of the SAN. We first provide a brief overview of how a user can implement SAN Boot and how SAN Boot works.

SAN Boot relies on configuring servers with a virtual boot device, which enables the server to access the actual boot information stored on a specific LUN in the storage subsystem. The virtual boot device is configured on the HBA in the server and thus assumes that the HBA BIOS supports booting from SAN attached storage.

1.7.1 Basic concepts

Multiple LUNs can be configured with various boot images if there is a need for separate operating system images. Each LUN that contains a boot image for a specific operating system is referred to as a boot partition. Boot partitions are created, configured, and mapped just like normal storage partitions on the DS5000 storage subsystem.

Figure 1-4 shows a 4+P RAID 5 array that is carved up into boot partition LUNs. Each of these LUNs corresponds to a separate server. The LUNs can be connected to various homogeneous hosts, or they can be connected to heterogeneous hosts.

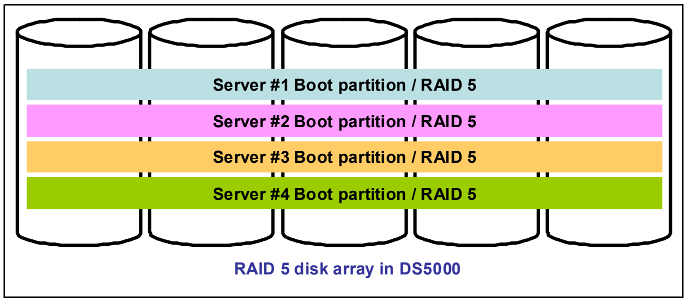


Figure 1-4 Boot partition LUNs on a RAID 5 array

Figure 1-5 shows five heterogeneous servers that use SAN Boot and store their boot images on the same DS5000 storage subsystem. However, this diagram shows the concept similar to Fibre Channel or iSCSI implementation.

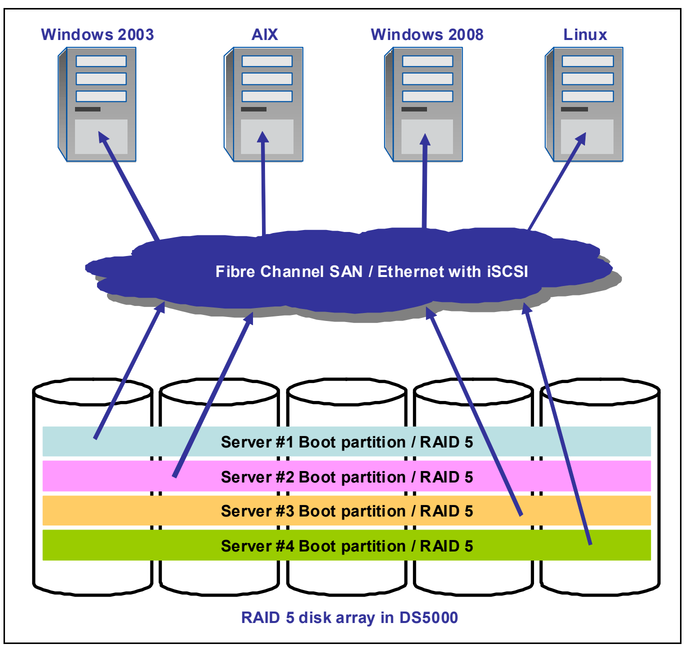


Figure 1-5 Logical diagram of SAN Boot partitions

In order to implement SAN Boot, you must first ensure that all your hosts are properly zoned to the storage subsystems in the SAN fabric. Zoning emulates a direct connection from the storage to each server. Storage partitioning, which needs to be done on the DS5000, only ensures that authorized hosts can access their dedicated storage units through the SAN.

The Host Bus Adapters (HBAs) and SAN-based storage subsystem both require setup. The specific syntax and steps differ from vendor to vendor for HBAs, so we take more of a general approach when describing the setup with the DS5000 storage subsystem.

The general steps are as follows:

1. Establish the physical connectivity between the HBA, switch, and DS5000 storage subsystem.
2. Provision LUNs on the DS5000 storage subsystem to handle host images. Create one LUN per boot image.
3. Select the proper port (HBA WWPN) and LUN from which the host must launch its boot.
4. Ensure that storage partitioning properly maps the appropriate LUN with the appropriate host server. It is also required that no other server can view that LUN.
5. Configure the HBAs of the hosts that are to utilize SAN Boot to point toward the external storage unit. The HBA BIOS tells the hosts that its boot image is now located on the SAN.
6. Install the boot image onto the appropriate LUN.

Here we list the requirements and guidelines for SAN Boot:

- SAN configuration, zoning of boot devices, and multipath configurations
- Active path to boot LUN
- Only one path to the boot LUN being enabled, prior to installing and enabling a multipath driver
- HBA BIOS selectable boot, or boot BIOS, must be enabled

These general steps can serve as a guide for how to configure SAN Boot. It is important that you analyze additional prerequisites and specific directions according to the particular host, HBA, fabric, and storage subsystem you might have.

Important:

- When installing the operating system, you can only have one path to the storage device (LUN). Because servers are generally equipped with two HBAs, and because most DS5000 storage subsystems have two RAID controllers, you must isolate (disconnect) the second HBA or do an appropriate zoning.
- For Windows and Linux based operating systems, the boot LUN must be assigned as LUN 0 when doing storage partitioning.
- The boot LUN must be accessible only by the host that is utilizing it to boot, which can be achieved through the storage partitioning feature of the DS5000 storage subsystem.
- For iSCSI configuration, follow the vendor instructions for hardware initiators, make sure your firmware version is at required level; for software initiators study the online operating system documentation for mandatory prerequisites.

1.7.2 Basic SAN Boot procedure

The basic steps for setting up the system to boot from SAN are as follows:

1. Establish the physical connectivity between the HBA, SAN switch, and storage subsystem.
2. Provision LUNs on the storage subsystem to accommodate host images. Create one LUN per OS boot image.
3. Select the proper HBA port (HBA WWPN) and LUN from which the host must launch its boot image.
4. Ensure that the appropriate LUN is mapped to the appropriate host server through storage partitioning. It is also required that no other host server can view this specific LUN.

5. Boot into the host server system BIOS or UEFi and disable the on board RAID controller.
6. Configure the HBAs on the hosts to point toward the external storage unit. The HBA BIOS will inform the host that its boot image is now located on the SAN.
7. Install OS onto the appropriate LUN.

Here we list the requirements and guidelines for SAN Boot:

- SAN configuration, zoning of boot devices, and multipath configurations
- Active path to boot LUN
- Only one path to the boot LUN enabled, prior to installing and enabling a multipath driver
- HBA BIOS selectable boot, or boot BIOS, must be enabled

These general steps can serve as a guide for how to configure SAN Boot. We cover QLogic, Emulex, and iSCSI HBAs in this book. We are covering the configuration setting on the storage subsystems as well.

> **Important:**
>
> When installing the operating system, you must only have one path to the storage device LUN. Because servers are generally equipped with two HBAs, and because most storage subsystems have two RAID controllers, you must isolate (disconnect) the second HBA or do the appropriate zoning.
>
> The boot LUN must be accessible only by the host that is utilizing it to boot, which can be achieved through the use of storage partitioning feature or equivalent on the storage subsystem.

1.8 Storage efficiency and data protection and retention

In this section, we introduce some new IBM concepts and features and explain how these features can be merged with the SAN Boot Technology.

We cover the following topics:

- Storage efficiency for SAN Boot
- Data protection and retention for SAN Boot

1.8.1 Storage efficiency for SAN Boot

In our current environment, customers are being charged with reducing I/T spending as much as possible. Storage infrastructure is one area where traditionally, customers simply bought more storage to deal with demand and never implemented different features or products to optimize their storage infrastructure. Today, according to CIOs, that trend cannot be sustained anymore. Customers must optimize their storage infrastructure to reduce costs or avoid/defer/minimize disk purchases.

Every industry needs an optimized storage infrastructure; even in today’s poor economic environment, analysts are reporting storage growth across the board. However, with this storage growth, all industries are looking at ways to do it as efficiently as possible. Here, a highly optimized storage infrastructure can be of help. Using Storage Virtualization, Storage Consolidation and Storage Resource Management (SRM) allows customers in each industry to attend the company's requests and manage their existing storage in a much better way.

The IBM Storage Products Portfolio focuses on providing solutions to help customers address the ever increasing cost pressures they face. This storage family has multiple levels of products to match against the customer's needs.

In this section, we present you with some IBM storage products and features, and briefly explain how they can be used for SAN Boot implementation.

Automated Tiering feature

Automated Tiering (IBM System Storage IBM Easy Tier and IBM Storage Tier Advisor) can monitor data throughput and automate the placement of data across the appropriate drive tiers to optimize current workload requirements. Sharing the same box with different application characteristics and overload, including SAN Boot LUNs, this feature can help to move heavy IOs to different disk pools and prevent downtimes and low performance.

This feature allows for the efficient use of solid state drives to increase performance by up to 200% on critical applications.

Using this automated feature, you can reduce costs by migrating less critical data to less expensive media and can optimize the labor involved with storage management.

Flexible delivery options

Consolidated, cloud-based flexible delivery models for storage can help you give your end users control of their storage needs, while improving accountability. It can give departments the ability to provision storage to meet changing demands and developers the ability to set up systems to build and test applications while also maintaining system reliability.

This flexibility can help the IT storage manager optimize their storage environment, mixing the technology to better match the user’s request. SAN Boot can benefit from it, especially for critical servers, speeding up their management, security, and capacity.

Mixing traditional spinning drives and ultra fast solid-state drives (SSDs)

This feature helps support tiered storage environments and consolidation of multiple storage systems onto single storage system. You can mix the SAN Boot LUNs with the Intensive I/O application in the same storage subsystem.

Physical capacity upgradeable without system disruption

This feature enables capacity upgrades to address the dynamic needs of your business. Depending on your environment, this feature can help you avoid business disruption in your SAN Boot environment.

IBM Full Disk Encryption drives

IBM FDE drives protect sensitive information from internal or external threats. This feature is intended to be used for servers with critical data and can prevent intrusive copy of data at the storage level.

Thin provisioning

Thin provisioning makes it possible to optimize the utilization of available storage.

Overview

Thin provisioning optimizes the utilization of available storage by allocating storage space as needed while improving storage utilization rates significantly. Relating to SAN Boot, thin provisioning can prevent unnecessary cost and have free space available for OS in case of critical and unexpected demands.

Concepts

Thin provisioning provides the following benefits:

- It facilitates reduced storage costs, Thin provisioning allows logical volumes to exceed physical capacity, while the latter need only to be large enough to contain the data actually written. This flexibility can offer dramatic savings in storage, power, and space costs.
- It helps provide simplified provisioning and space reclamation, which can ease the management of capacity allocation for application workload. Using IBM storage subsystem management tools, administrators can easily zero out space no longer in use; the system then releases that space into the general pool, allowing its reuse.
- Instant space reclamation is also enabled for supported applications that offer thin reclamation.

Virtualized storage

With virtualization, storage capacity can be moved across the storage infrastructure without interruption. You logically create disk storage so capacity can be efficiently allocated across applications and users. Virtualization can bring the following benefits to the datacenter:

- You can increase utilization of existing storage by up to 30%.
- You can move and manage data so hardware can be serviced, all without disrupting users.

The IBM SVC and Storwize V7000 can provide a single point-of-control for storage resources.

Dynamic data migration

Dynamic data migration provides the following benefits:

- You can migrate data among devices without taking applications offline.
- You can migrate and scale storage capacity without disrupting applications.

Pools the storage capacity of multiple storage systems on a SAN

Virtualization helps you manage storage as a resource to meet business requirements and not just as a set of independent storage subsystems:

- It helps administrators better deploy storage as required beyond traditional “SAN islands.”
- It can help increase utilization of storage assets.
- It insulates applications from physical changes to the storage infrastructure.

Visibility

With improved visibility into the data and data utilization in your storage environment, you can make more intelligent business decisions around efficiency, so that you respond faster to changing business demands while reducing cost.

It can be achieved through an easy-to-use graphical management interface, a single interface for a storage configuration that provides management and service tasks regardless of storage vendor. The XIV GUI is a perfect example, designed to make the storage interfaces even easier, more productive to use, and able to be accessed from anywhere on the network with a web browser.

The IBM Tivoli Storage Productivity Center (TPC) is a leader in helping organizations manage complex heterogeneous storage environments. Along with IBM solutions for storage consolidation such as SVC, it provides the broadest support for heterogeneous storage, integrates with virtualization and backup tools, and gives customers one of the most comprehensive views of their storage environment from a single pane of glass. Competitive solutions such as HP Storage Essentials, Symantec Command Central, and a host of smaller SRM vendors also do a good job of heterogeneous management, However, SVC with TPC is a much better value proposition for managing heterogeneous storage.

The TPC family offers comprehensive, enterprise-wide storage resource management solutions, that can be managed centrally through a single user interface, on a single platform, making storage management for our customers “Simple, Centralized, and Automated” to give the best ROI.

New Intelligent Performance Optimization Analytics from IBM research enables improved disk optimization of enterprise environments increasing overall availability of storage network and applications.

IBM Systems Director provides integrated management for physical and virtual server infrastructures, including monitoring for higher availability and operational efficiency.

IBM Tivoli Storage IBM FlashCopy® Manager application-aware snapshots performs near-instant application-aware snapshot backups, with minimal performance impact for IBM DB2®, Oracle, SAP, Microsoft SQL Server, and Microsoft Exchange.

SAN Boot implementations can take advantage of a centralized tool to manage its tasks. TPC can manage various storage subsystems, capture the utilization, and draw a new design to balance the performance. It can also manage complex features such as remote copies for SAN Boot LUNs.

Storage infrastructure optimization

Some IBM storage systems can be extended into a grid of similar systems adding capabilities for clustering, failover, site-switching automation, and replication supporting up to four sites. The XIV grid, as an example, is accomplished by combining XIV systems with IBM SAN Volume Controller for XIV software.

IBM FlashCopy and IBM FlashCopy SE

The IBM FlashCopy and FlashCopy SE feature helps to ensure data recovery in the event of source data loss with point-in-time data copies for backup or replication. It enables data to be copied in the background while making both source and copied data available to users almost immediately. Using the Clone and point-in-time copies, you can duplicate any internal SAN Boot LUN and build new servers instantly. This feature can help to reduce the downtime for tasks such as backups, updates, and upgrades in servers using SAN Boot LUNs.

IBM Metro Mirror, Global Mirror, Metro/Global Mirror, and Global Copy

IBM Metro Mirror, Global Mirror, Metro/Global Mirror, and Global Copy allow for data to be replicated to remote sites. This capability provides flexible options to support multiple two-site and three-site business continuity scenarios.

It provides users access to data during planned and unplanned system and site outages. This feature, working together with the SAN Boot process, can speed up the time to recover systems, specifically in Disaster/Recovery events.

This combination of storage and features can improve the storage utilization and help IT storage managers to better use their technologies, virtualized, scalable, and with security.

1.8.2 Data protection and retention for SAN Boot

In addition to increasing storage efficiency, we need to focus on data protection as well:

- Requirements to ensure that data is protected from disaster
- Demand that data be continuously accessible
- The need to adhere to business and legal requirements

In this section, we look at how SAN Boot implementations using IBM disk storage systems and IBM storage software solutions can help storage administrators and CIOs win the data war.

Backup and recovery

You back up your data so it can be recovered or restored in case of data loss or corruption. Your business cannot afford not to have key data available. But ongoing data growth places more and more demands on backup processes.

Solutions

IBM offers a wide range of effective, efficient, and secure backup and recovery solutions to ensure that key business information is always available.

IBM Tivoli Storage Manager

IBM Tivoli Storage Manager enables you to protect your organization's data from failures and other errors. It does so by storing backup, archive, space management, and bare-metal restore data, as well as compliance and disaster-recovery data, in a hierarchy of offline storage.

You can take advantage of a SAN Boot host solution using both IBM disk storage systems and IBM Tivoli Storage Manager. Thus, host backup and recovery processing benefits from high performance disk storage and advanced functionality such as FlashCopy to securely and efficiently meet ever demanding backup and recovery requirements.

Continuous operations

According to recent research, a one-hour IT system outage can cost a business $1.1M. Your IT systems must be able to prevent and avoid outages and recover from them quickly,

Solutions

IBM offers a range of reliable, high-performance solutions to ensure that a business can remain operational through component failures or significant disasters.

IBM Tivoli Storage FlashCopy Manager

Use of IBM Tivoli Storage FlashCopy Manager enables organizations to perform and manage frequent, near-instant, non-disruptive, application-aware backups and restores, making use of advanced FlashCopy snapshot technologies in IBM storage systems.

Just as IBM Tivoli Storage FlashCopy Manager takes advantage of IBM storage systems functionality, so can SAN Boot host solutions utilizing IBM storage systems. With the ability to take instantaneous FlashCopies of a host operating system, costly issues, which might be encountered with newly applied operating system or application code, can be quickly restored.

By centralizing the location of the host operating system storage with use of SAN Boot, System Administrators are also able to rapidly recover from host hardware errors. By simply re-assigning SAN Boot volumes to standby hardware, recovery times and outage costs are greatly reduced.

Control and automation

By automating management and enforcement of storage policies, you can keep storage infrastructure in its desired state efficiently. It can help you improve service quality, improve data protection, recovery, reliability, and integrity while reducing ongoing staff intervention.

Solutions

Automated lifecycle management tools can help you meet information retention mandates, automate provisioning, backup, and archive to optimize performance, utilization, lower costs, and build agility.

IBM Tivoli Storage Productivity Center

IBM Tivoli Storage Productivity Center Standard Edition is a single integrated solution designed to help you improve your storage TCO and ROI by combining the assets, capacity, performance, and operational management, traditionally offered by separate Storage Resource Management, SAN Management, and Device Management applications, into a single platform.

Implementing a SAN Boot host solution in conjunction with IBM Tivoli Storage Productivity Center can further improve and facilitate Storage Resource Management. By co-locating host operating storage on IBM storage systems, numerous benefits can be realized:

- Capacity can be monitored at the storage system level, not at the individual host level.
- Automation can be implemented to create capacity on demand storage solutions.
- Overall storage utilization can be planned for at the storage subsystem level.

Retention and compliance

Today, significant amounts of data must be saved to meet business or legal requirements.

Solutions

As the amount of information stored grows and grows, effective data archiving solutions can help you improve productivity, control both infrastructure and legal costs, and manage compliance and operational risks.

IBM Information Archive

IBM Information Archive is a simple and flexible archiving solution to help organizations of all sizes address their complete information retention needs: business, legal, or regulatory.

With a SAN Boot host solution, retention is enhanced by the functionality of the IBM storage system that the data resides on. Use of centralized storage subsystem functionality, instead of utilizing a mixture of functionalities found with individual host level storage, reduces administrative overhead.

In regard to compliance, SAN Boot host data can benefit from FlashCopy snapshots of data and the retention of that data in a locked and immutable state, pending movement to an archive environment such as IBM Information Archive.

2

SAN Boot implementation with IBM System Storage DS5000

In this chapter, we describe the host connectivity for the IBM System Storage DS5000 series storage subsystem and how to implement a host SAN Boot from DS5000 products. We address key aspects of host connectivity and review concepts and requirements for both Fibre Channel (FC) and Internet Small Computer System Interface (iSCSI) protocols.

The term *host* refers to a server running a supported operating system. This chapter covers tasks that pertain to SAN Boot implementation. The following operating systems are described:

- Windows Server 2008 R2
- Linux (RHEL, SUSE Linux Enterprise Server)
- AIX 6.1
- VMware ESX 4.1

© Copyright IBM Corp. 2012. All rights reserved.

2.1 DS5000 overview

The DS5000 series storage subsystems use Redundant Array of Independent Disks (RAID) technology, which is used to offer various levels of performance and protection for the user data from disk drive failures. DS5000 storage subsystem offer Fibre Channel (FC) interfaces to connect the host systems and external disk drive enclosures. With the DS5000, the host side connections can be up to 8 Gbps. Additionally, with the DS5000 series, there is also an iSCSI interface available for host side connections. It is covered in Chapter 6, “iSCSI SAN Boot implementation with IBM System Storage DS5000” on page 395.

For more details on features and family members of the DS5000 series, see these books:

- *IBM Midrange System Storage Hardware Guide*, SG24-7676
- *IBM Midrange System Storage Implementation and Best Practices Guide*, SG24-6363.

2.2 SAN configuration planning

Careful planning is essential to configuration of the SAN Boot process. This section provides a guideline to help you with the planning process.

Choosing the right equipment and software, and also knowing what the right settings are for a particular installation, can be challenging. Every installation needs to answer these questions and accommodate specific requirements.

Having a well thought-out design and plan prior to the implementation will help you get the most out of your investment.

During the planning process, you need to answer numerous questions about your environment:

- What tool do I use to manage and configure the DS5000 series hardware?
- How do I provision my storage for SAN Boot?
- How do I configure my host for SAN Boot?
- How do I map the LUN to the host?
- How do I configure my operating system
- What, if any, best practices, do I use for DS5000 series storage while doing SAN Boot?
- What, if any, troubleshooting setup do I need to know for doing SAN Boot?

2.3 IBM System Storage DS Storage Manager

The IBM System Storage DS® Storage Manager (also referred to as DS Storage Manager, or Storage Manager) software is the primary tool for managing, configuring, monitoring, updating firmware, and support data collection for the DS3000, DS4000, and DS5000 series storage subsystems, and repair procedures. This tool provides two interfaces, a user friendly graphical user interface (GUI), and a command line interpreter (`smcli`) interface for use with scripting to make repetitive work easy.

Here are some of the tasks that can be performed:

- Configuration of RAID arrays and logical drives
- Assigning logical drives to a host
- Expanding the size of the arrays and logical drives
- Converting from one RAID level to another

The tool can be used for troubleshooting and management tasks such as, checking the status of the storage subsystem components, updating the firmware of the RAID controllers, replacement procedures for failed components, including rebuilding drives for use, and managing the storage subsystem. Finally, it offers implementation and management capabilities for advanced premium feature functions such as FlashCopy, Volume Copy, and Enhanced Remote Mirroring. The DS Storage Manager software package also includes the required host software components for the specific host environments that are planned for.

The DS Storage Manager software level is closely tied to the level of the firmware code that is being run on the subsystem. Newer DS Storage Manager levels are designed to be backward compatible with current firmware levels for previous generations of products as well as earlier versions of firmware for the current product line. Newer firmware levels might require a newer version of the DS Storage Manager to be installed.

> **Tip:** Always consult the System Storage Interoperation Center for the latest supported host types and operating systems:
>
> http://www-03.ibm.com/systems/support/storage/config/ssic/displayesssearchwithoutjs.wss?start_over=yes

2.3.1 Storage Manager software

The DS Storage Manager software is now packaged as follows:

- *Host-based* software:
  - Storage Manager 10.70.x Client (SMclient):

    The SMclient component provides the GUI and the "smcli" interfaces for managing storage subsystems through the Ethernet network or from the host computer.
  - Storage Manager 10.70.x Runtime (SMruntime):

    The SMruntime is a Java runtime environment that is required for the SMclient to function. It is not available on every platform as a separate package, but in those cases, it has been bundled into the SMclient package.
  - Storage Manager 10.70.x Agent (SMagent):

    The SMagent package is an optional component that allows in-band management of the DS4000 and DS5000 storage subsystems.
  - Storage Manager 10.70.*x* Utilities (SMutil):

    The Storage Manager Utilities package contains command line tools for making logical drives available to the operating system for specific host environments.
  - Multipath drivers:

    The storage manager offers a choice of multipath drivers, RDAC, or MPIO. This choice might be limited depending on host operating systems. Consult the Storage Manager readme file for the specific release being used.

    RDAC or MPIO are Fibre Channel I/O path failover drivers that are installed on host computers. These are only required if the host computer has a host bus adapter (HBA) installed.

- *Controller-based* software:
  - Storage subsystem controller firmware and NVSRAM are always installed as a pair and provide the “brains” of the DS5000 storage subsystem.
  - Storage subsystem Environmental Service Modules (ESM) firmware controls the interface between the controller and the drives.

Storage subsystem drive firmware is the software that tells the specific drive types how to perform and behave on the back-end FC loops.

2.3.2 Installation of the Storage Manager

We first install the IBM System Storage DS Storage Manager software for managing the DS5000 storage subsystem.

> **Testing:** For the testing, we use Storage Manager SMIA-WSX64-10.70.35.25.exe downloaded from IBM Fix Central at `http://www-933.ibm.com/support/fixcentral/`. We are running the Storage Manager on a Microsoft Windows 2007 machine.

We also cover the changes needed in the IBM System Storage DS Storage Manager to run on a workstation. Here we only show the panels in which you need to make changes, and the “splash” panel. The complete installation procedure can be found in *IBM Midrange System Storage Hardware Guide*, SG24-7676.

Here we outline the procedure:

1. Figure 2-1 shows the first panel after the installer is launched.

Figure 2-1 IBM System Storage DS Storage Manager splash panel

2. In Figure 2-2, select **Custom**.

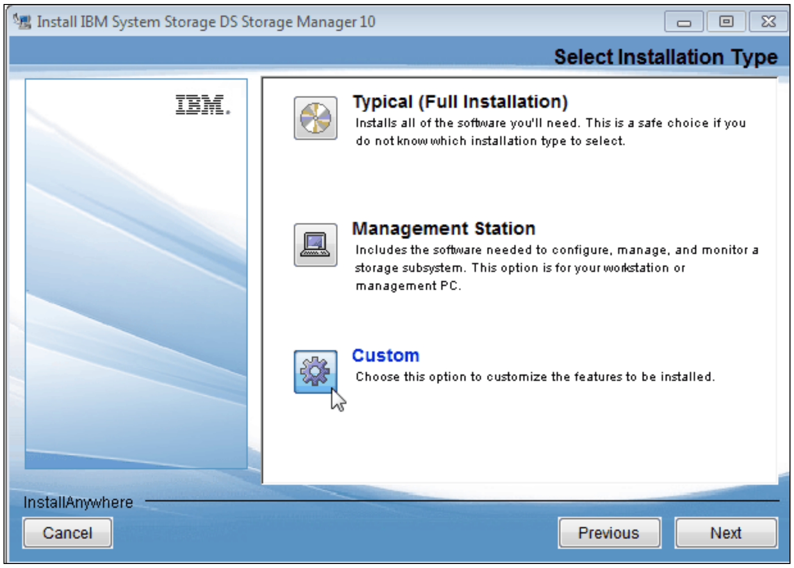


Figure 2-2 Selecting Custom

3. Uncheck **Support Monitor** and **Java Access Bridge** and click **Next** as shown in Figure 2-3.

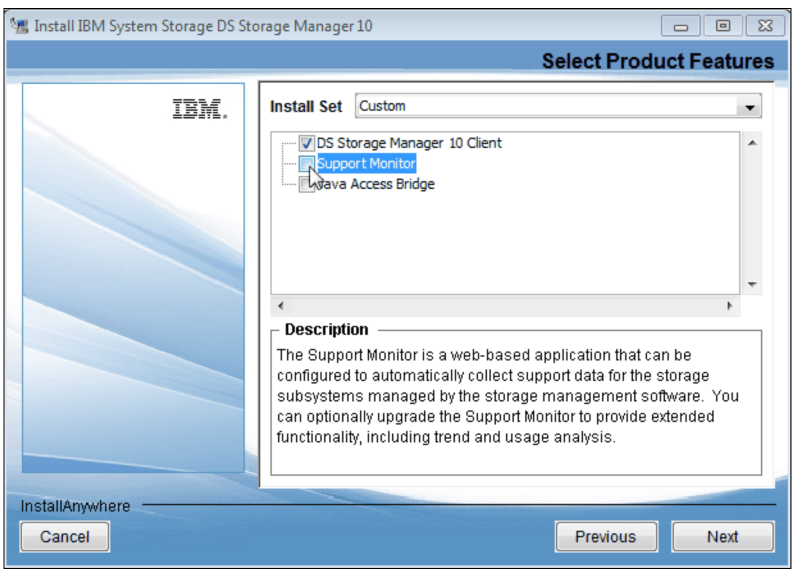


Figure 2-3 Selecting the Product Features

4. Select **Do not Automatically Start the Monitor** as seen in Figure 2-4. Click **Next**.

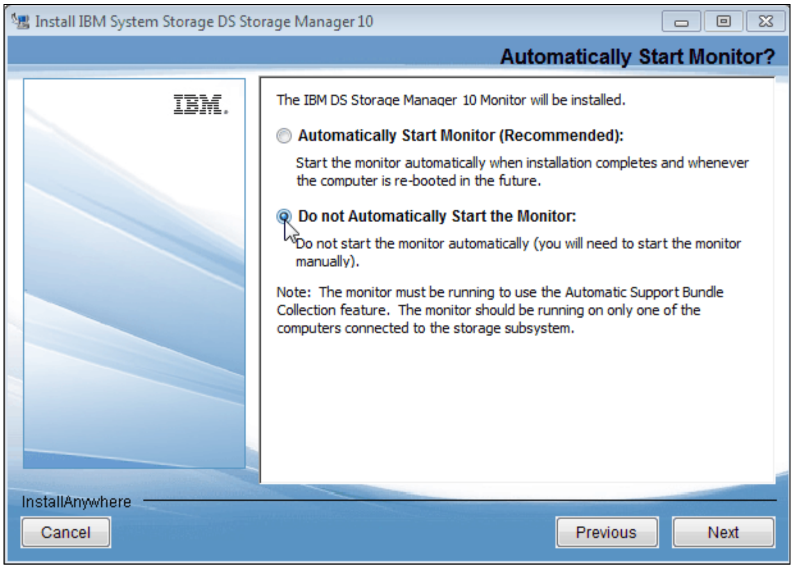


Figure 2-4 Change to Do Not Automatically Start the Monitor

5. After the installation, you should see the **Installation Complete** as shown in Figure 2-5.

> **Tip:** If you do not have Admin rights to your system to which you are installing IBM System Storage DS Storage Manager 10, it will fail to install.

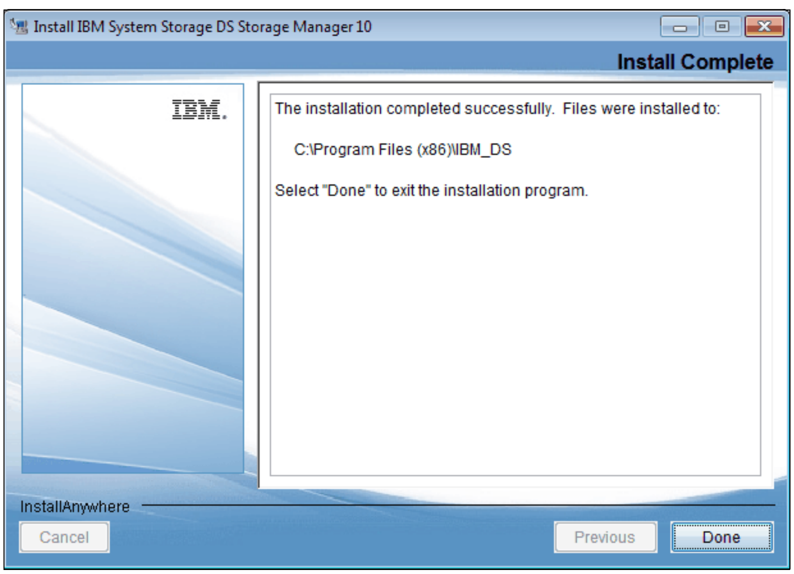


Figure 2-5 Installation complete

6. After finishing the installation, launch the IBM System Storage DS Storage Manager 10 from the appropriate location, depending on your operating system (OS).
7. When you have launched the IBM System Storage DS Storage Manager 10, you see the panel in Figure 2-6. You need to select **Manual** if you are going to define the IP address. Else, you can do an **Automatic** scan and let the storage scan the subnet and scan the fibre path.

> **Tip:** For Manual addition, you must know the DNS name or IP Address. For Automatic addition, it must be either connected to the Fibre Channel network or on the same subnet, or it will not detect a subsystem.

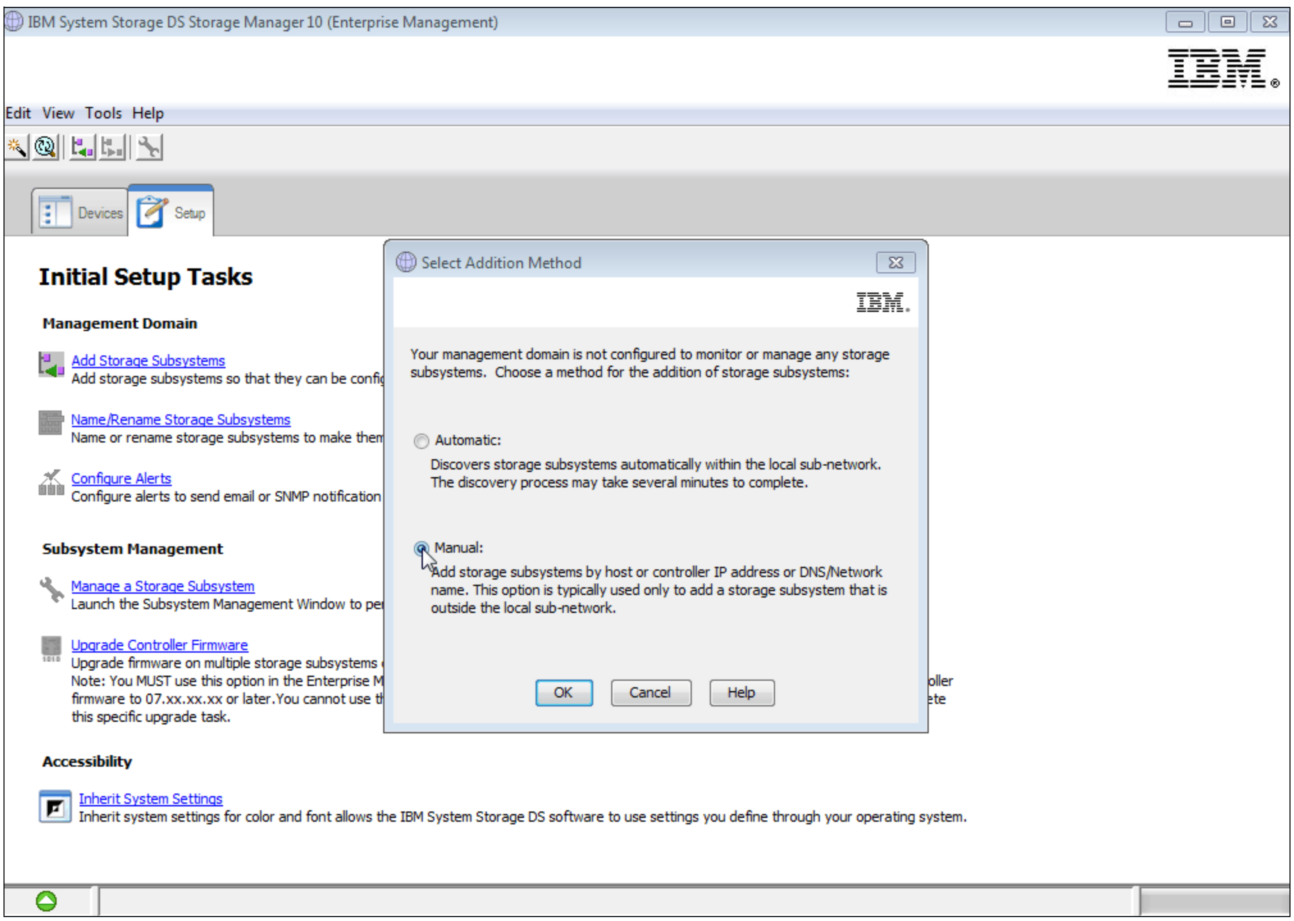


Figure 2-6 Selection of Automatic or Manual discovery

8. As shown in Figure 2-7, enter the IP address for your two DS5000 controllers for out-of-band management.

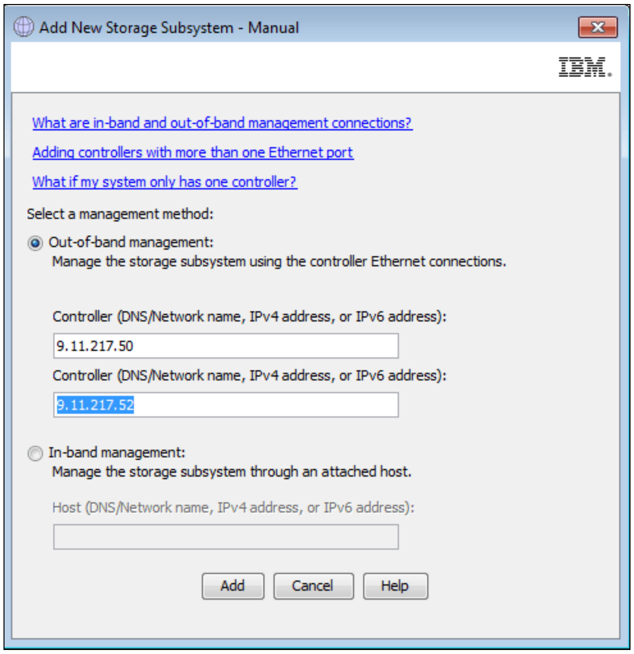


Figure 2-7 IBM System Storage DS Storage Manager 10 Out-of-band Management

9. After the system discovers the storage, you can see that it completes successfully, and asks you if want to add another one. At this panel (Figure 2-8), we select **No**, as we do not want to add more storage subsystems. Now that your storage manager is set up, we can cover setting up the storage subsystem in 2.8, “Storage provisioning” on page 50.

Tip: If you see a failure here, check your fibre cables or zoning if you are doing in-band management. If you are doing out-of-band management, check your IP address, VLAN, or ports for the possibility of a blocked port.

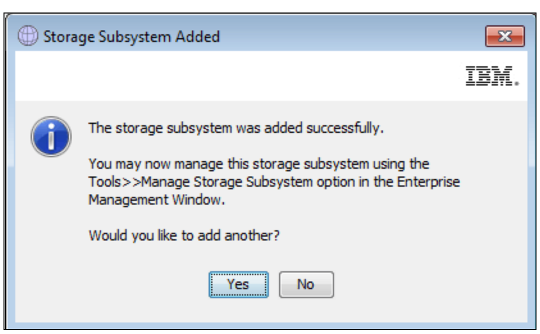


Figure 2-8 Storage was added successfully

2.4 Configuring the IBM System x based BIOS system

To configure the host to storage connection, we must make changes in the host system BIOS.

Configuration of the IBM System x BIOS is done as follows:

1. Power on the host system and interrupt the boot sequence to enter the system BIOS by pressing the F1 key.
2. Select **Devices and I/O ports** from the main menu, as shown in Figure 2-9.

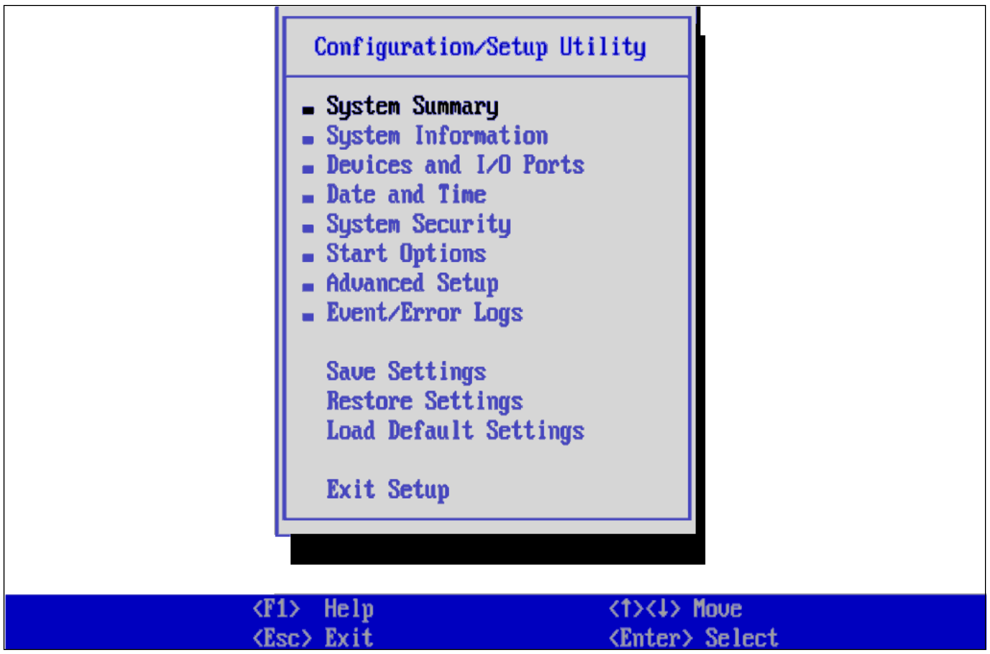


Figure 2-9 Disabling IDE Devices

3. Select **Planar SAS** from the next menu and disable the port, as shown in Figure 2-10.

> **Important:** This action will disable the onboard Serial Attached SCSI (SAS) to speed the boot process and make sure that the system does not try to boot off the onboard disks. Even if you remove the local disks, you still need to disable the Planar SAS.

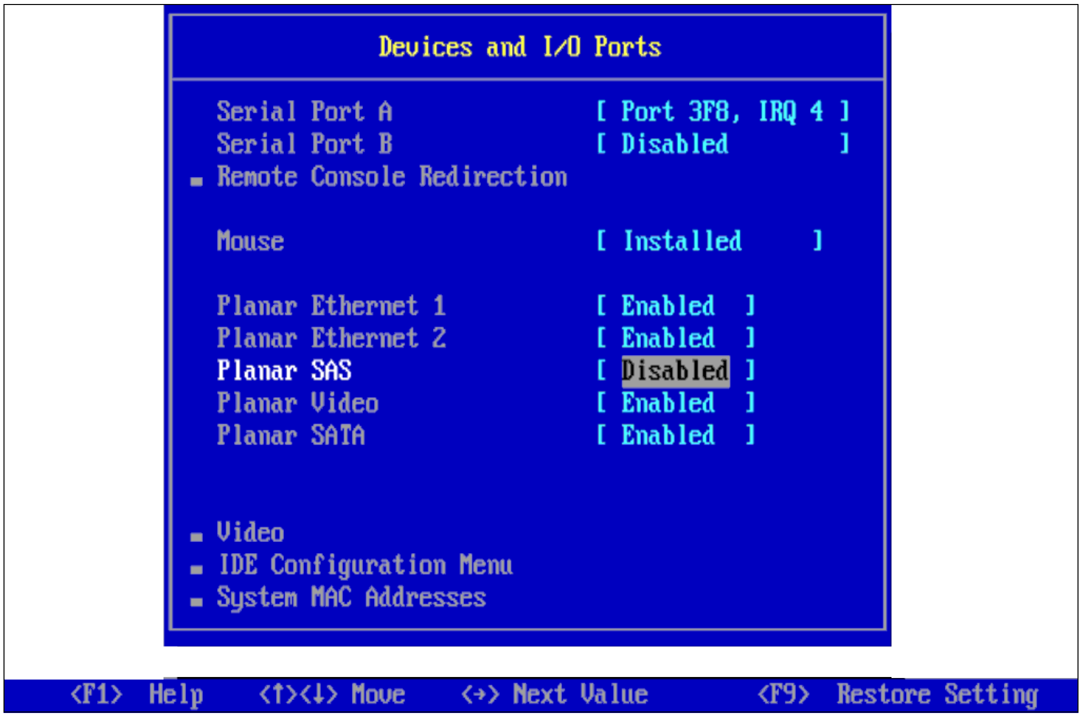


Figure 2-10 Disabling the Planar SAS

4. Press Esc to exit out of this menu and save the changes.
5. Restart the system. After the system is restarted, you are ready to configure the HBA adapter. The HBA configuration is described in the following sections:
   - 2.6, "QLogic HBA configuration"
   - 2.7, "Emulex HBA configuration"

2.5 Configuring IBM System x based UEFI system

Follow this procedure to make changes in the system Unified Extensible Firmware Interface (UEFI).

Configuration of the IBM System x UEFI is done as follows:

1. Power on the system and interrupt the boot sequence by entering the system UEFI by pressing the F1 key. Figure 2-11 shows what you will see after you enter the UEFI.

> **Tip:** Booting the system with a UEFI can take several minutes. Also, the appearance and options on how to set it up might be different, depending on your system.

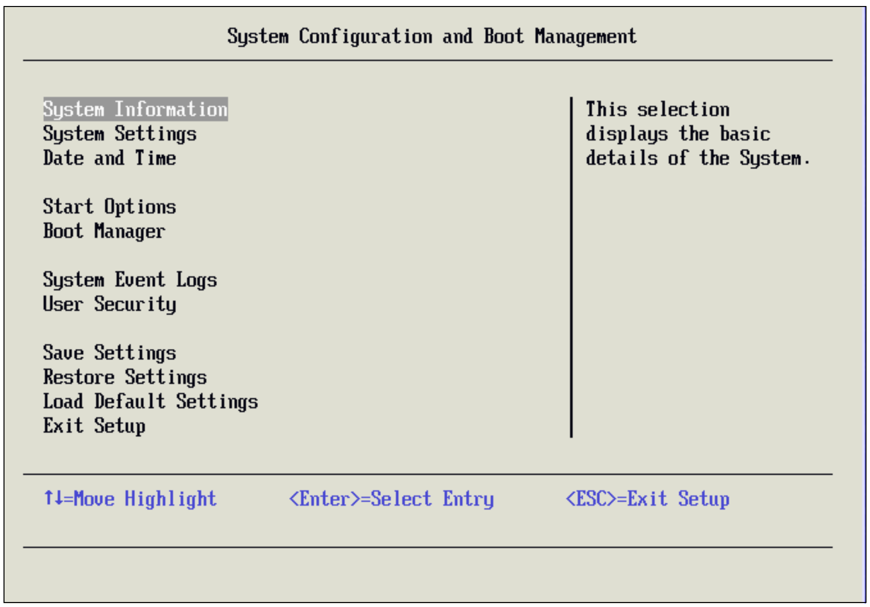


Figure 2-11 UEFI main panel

2. Select **System Settings** from the main menu. Then select **Devices and I/O Ports**, as shown in Figure 2-12.

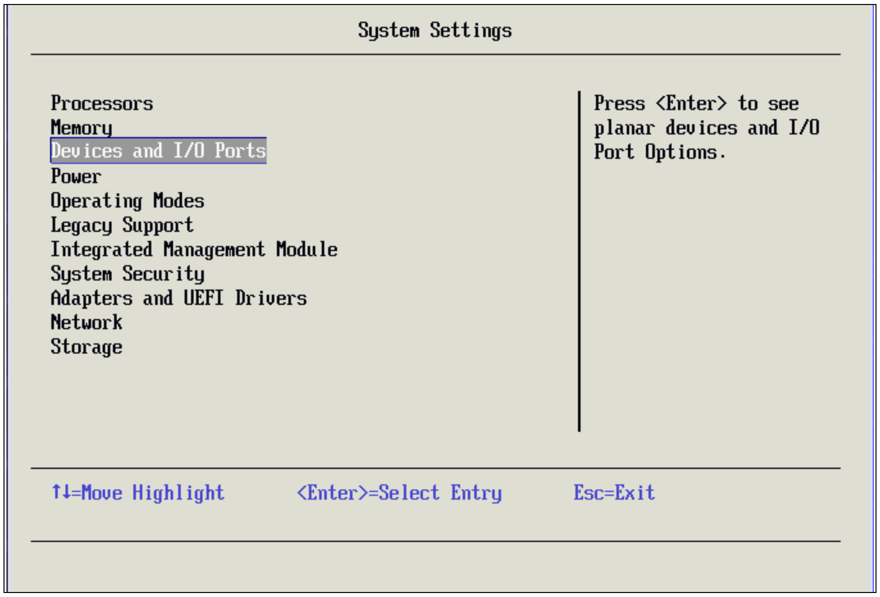


Figure 2-12 Selection of Devices and I/O Ports

3. Select **Planar SAS** and press Enter to enter the edit mode. Then select **Disable** to disable the onboard SAS controller as shown in Figure 2-13. It speeds up the boot process and keeps the system from booting to the onboard in case there are local drives in the system.

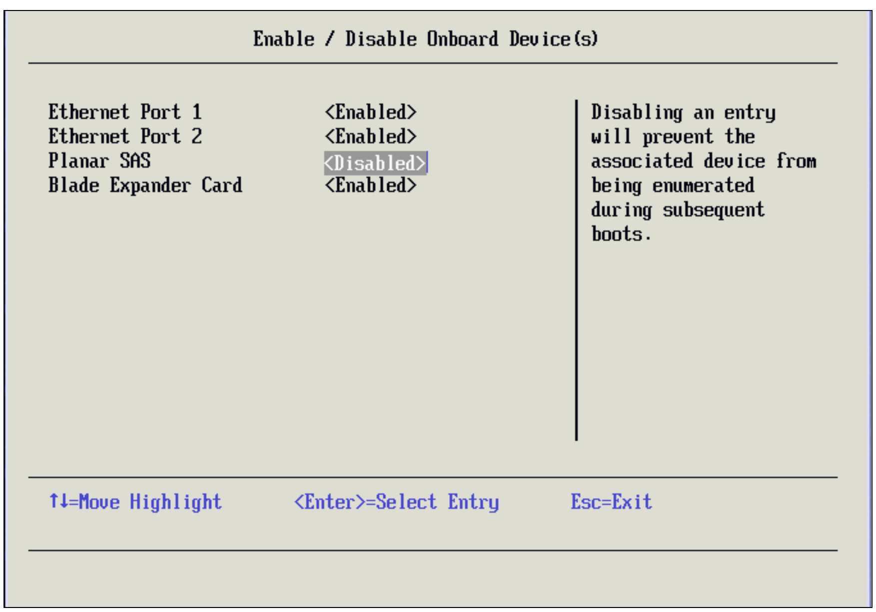


Figure 2-13 Disable onboard SAS controller

4. Press Esc to exit out of this menu and save the changes.
5. Power on or restart the system. After the system is restarted, you are ready to configure the HBA adapter. The HBAs is described in the following sections:
   – 2.6, "QLogic HBA configuration"
   – 2.7, "Emulex HBA configuration"

2.6 QLogic HBA configuration

In this section, we configure the QLogic HBA to communicate with the DS5000 storage subsystem. Ultimately this will allow for a broadcast over the SAN fabric, so that the storage subsystem can recognize the World Wide Port Number (WWPN) of the HBA.

> **Tip:** You need to do the zoning configuration on your SAN switch before moving on. We do not cover the SAN zoning steps in this chapter. Check with your switch manufacturer for the proper way to do zoning.
>
> See *Implementing an IBM b-type SAN with 8 Gbps Directors and Switches*, SG24-6116, for information regarding the IBM b-type SAN products.

1. During the boot process, press Ctrl+Q to enter the HBA BIOS configuration utility, as shown in Figure 2-14.

```
Press <CTRL-Q> or <ALT-Q> for Fast!UTIL
Firmware Version 4.03.01

Checking Adapter 0 Loop ID 254
No Fibre devices found!
ROM BIOS NOT INSTALLED

QLogic Corporation
QLE2562     PCI Fibre Channel ROM BIOS Version 2.02
Copyright (C) QLogic Corporation 1993-2008. All rights reserved.
www.qlogic.com

Press <CTRL-Q> or <ALT-Q> for Fast!UTIL

BIOS for Adapter 0 is disabled
Firmware Version 4.03.01

Checking Adapter 1 Loop ID 254
No Fibre devices found!
ROM BIOS NOT INSTALLED

<Alt-Q> Detected, Initialization in progress, Please wait...
```

Figure 2-14 invoke FC HBA BIOS

2. Select the first Fibre Channel adapter, as shown in Figure 2-15.

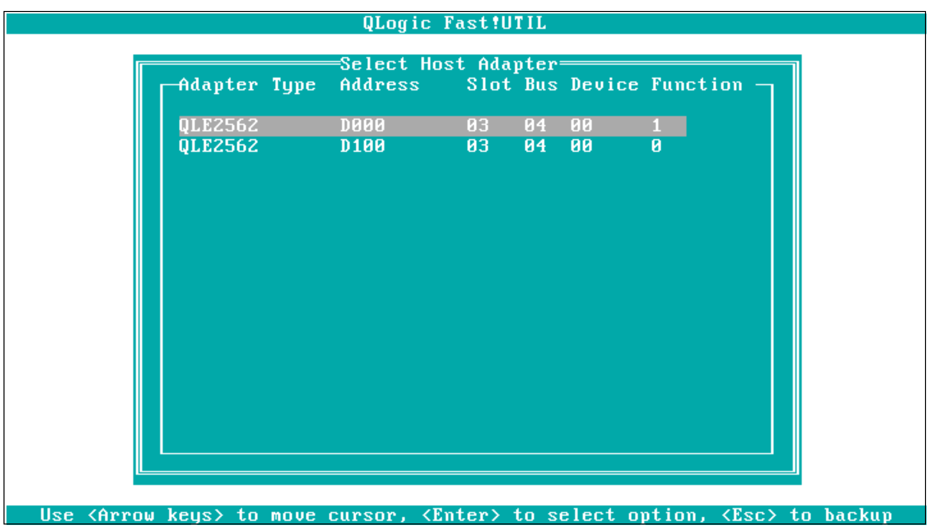


Figure 2-15 Select FC adapter

3. Select **Configuration Settings** as shown in Figure 2-16.

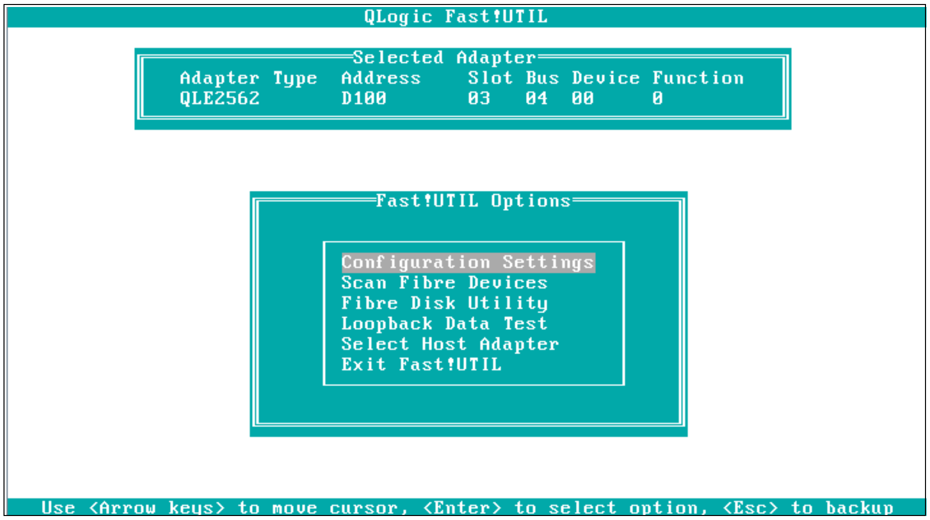


Figure 2-16 Selection of Configuration Settings

4. Select **Adapter Settings** to see the settings of the adapter that you have selected, as shown in Figure 2-17.

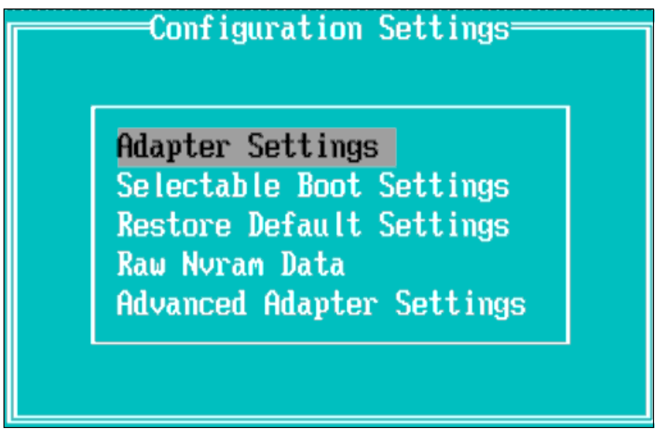


Figure 2-17 Selection of Adapter Settings

5. Change the Host Adapter BIOS to enabled (the default value is disabled), as shown in Figure 2-18.

> **Tip:** While on this panel, be sure to record the World Wide Port Name (WWPN) for this HBA (21:01:00:1b:32:3a:fe:e5). It will be needed while configuring storage partitioning on the DS5000 storage subsystem in 2.8.4, “Creating a host on DS5000 series” on page 58, and also for fabric zone configuration if it has not been done already. We also record the WWPN for the second HBA. It will be used in 2.8.6, “Adding the second path to the Host partition” on page 65.

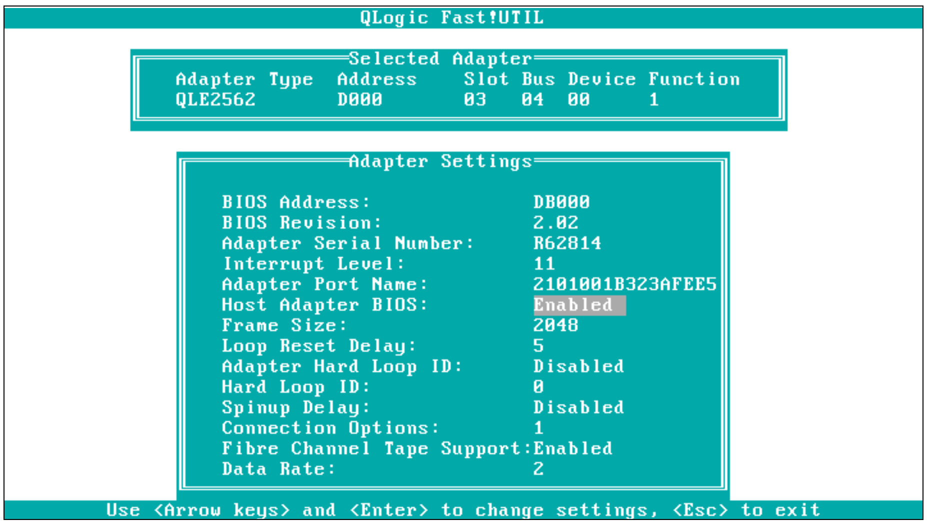


Figure 2-18 Enable FC BIOS

6. Press Esc two times until you see a red panel saying *Configuration settings modified*, as shown in Figure 2-19. Select **Save Changes** to save the changes made to the HBA settings.

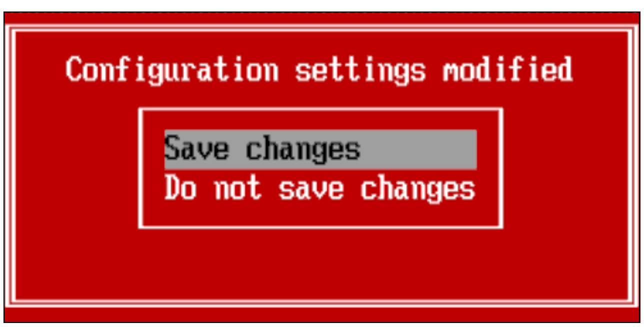


Figure 2-19 Save Changes

7. After the changes have been saved, press **Scan Fibre Devices**, as shown in Figure 2-20. It will scan the Fibre Channel path for the DS5000 storage subsystem. It will present the HBA to the storage subsystem so that it will show up in the storage configuration.

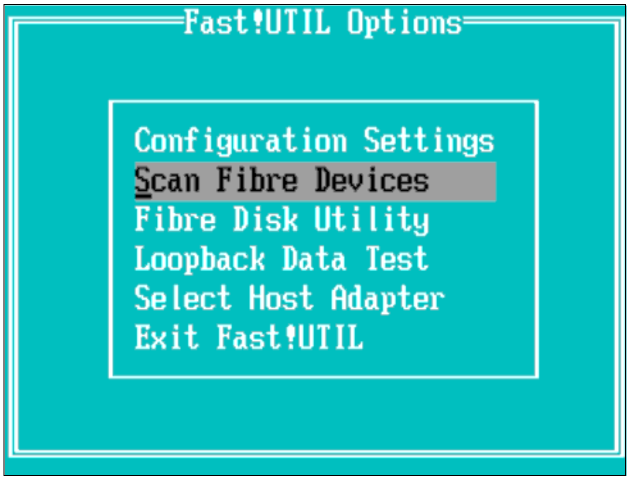


Figure 2-20 Scan for Fibre Devices

8. After you scan for fibre devices, you might not see any devices, as shown in Figure 2-21. It is possible that no storage LUNs have been configured on the storage subsystem for this specific host.

Tip: If you do see storage LUNs, the storage might already have been configured for this host server.

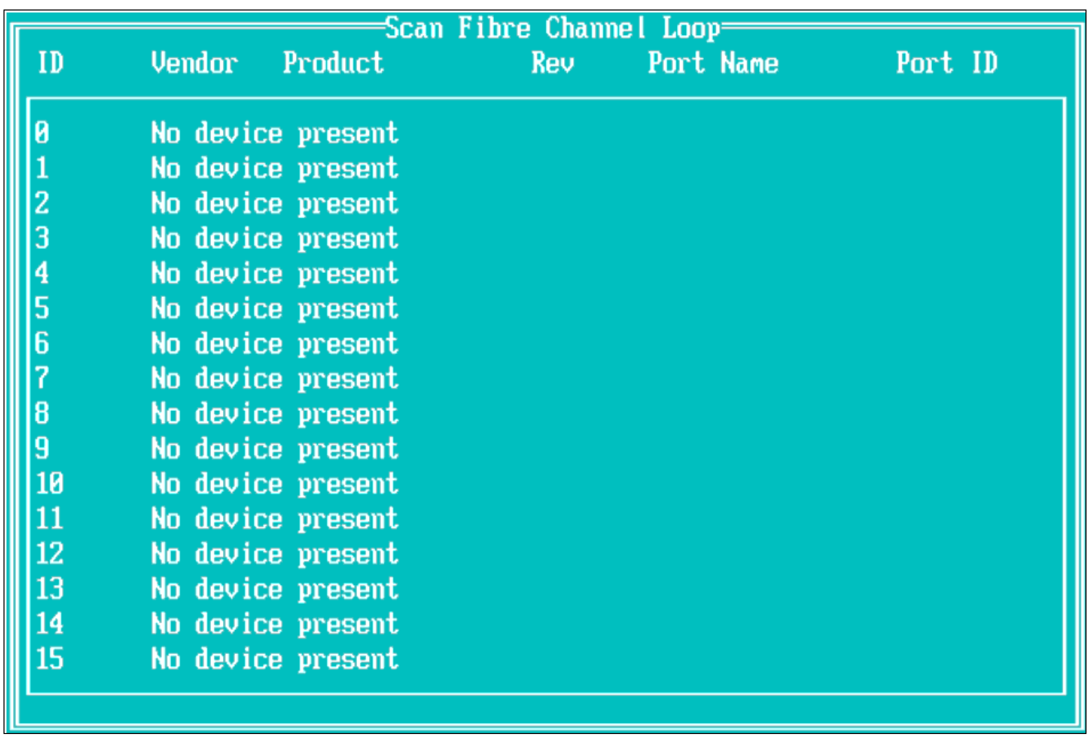


Figure 2-21 No device present yet

Leave the host system logged in at this panel, because we return later to complete the setup. At this point, you need to configure your storage if you have not done so already, which is covered in 2.8, "Storage provisioning" on page 50.

2.6.1 Configuring the QLogic HBA to use the Boot LUN

After you have your SAN configured as described in 2.8, “Storage provisioning” on page 50, you can proceed with completing the QLogic HBA setup listed in the following steps.

We are assuming that you left the host system at the panel Scan Fibre Device, as shown Figure 2-21 on page 36. If you are not at this panel, follow the instructions in 2.6, “QLogic HBA configuration” on page 33 to get to the Scan for Fibre Devices panel.

1. To rescan for the DS5000 storage subsystem, press Esc one time to go out of this panel. Select **Scan Fibre Devices** again, and you can see the storage subsystem as illustrated in Figure 2-22.

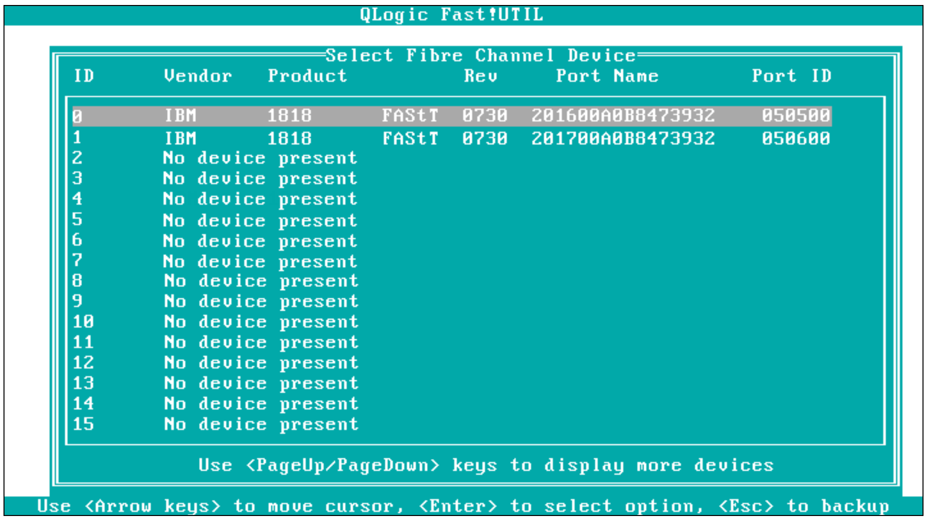


Figure 2-22 Scanning fibre Devices and Found DS5300

2. Now that the HBA can see the storage subsystem, we need to finish the configuration. Press Esc one time to get to the panel shown in Figure 2-23. Then select **Configuration Settings** and press Enter.

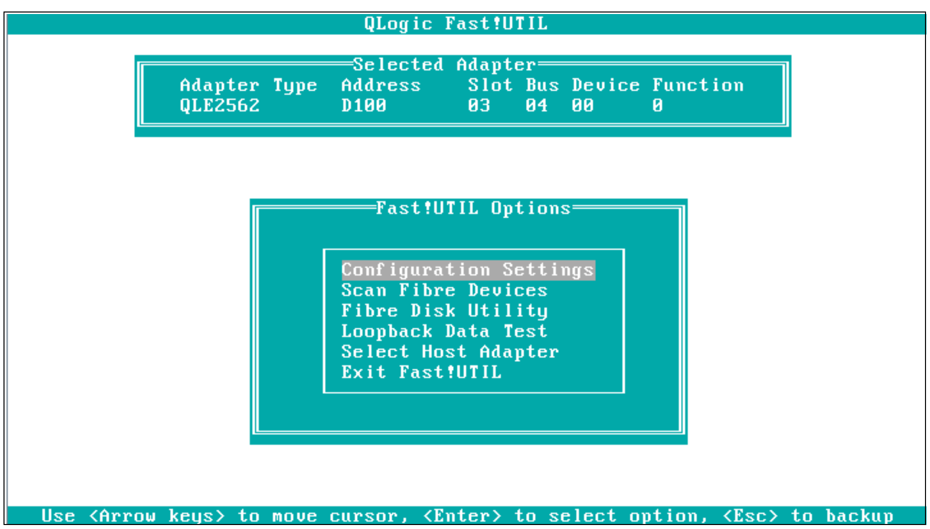


Figure 2-23 HBA Main panel selecting Configuration Settings

3. Select **Selectable Boot Settings** as shown in Figure 2-24.

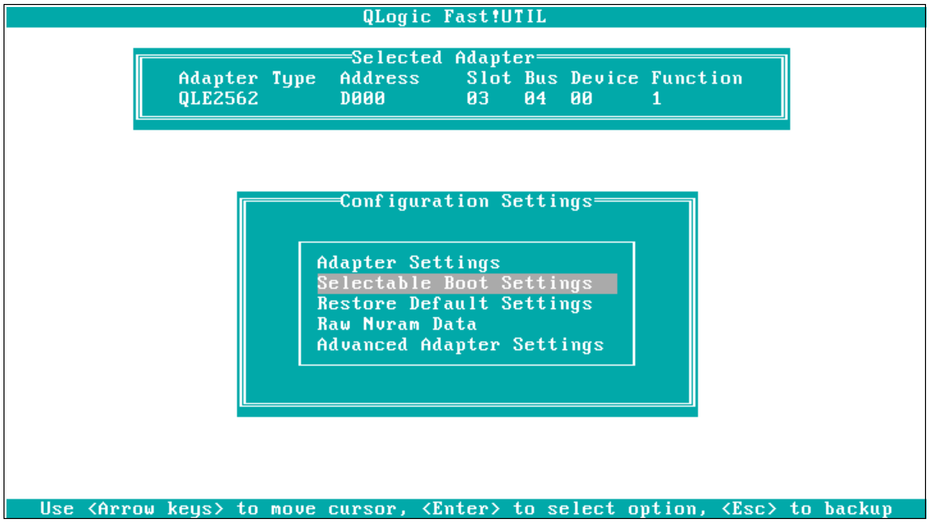


Figure 2-24 Selectable Boot Settings

4. Select **Selectable Boot Settings** and press Enter, as shown in Figure 2-25. Then **Enable** the **Selectable Boot**, because it is disabled by default. Arrow down to the **Primary Boot Port Name** and press Enter.

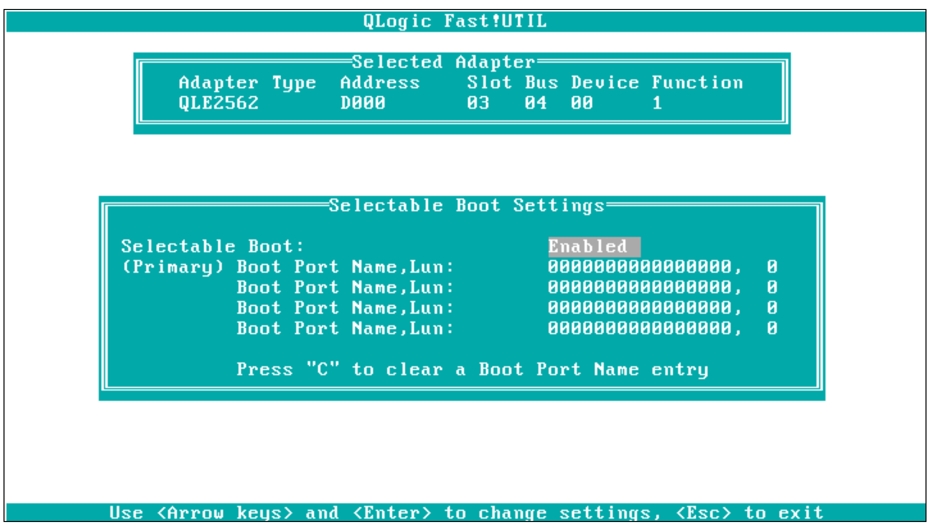


Figure 2-25 Enabling Selectable Boot

5. You can then see the product type 1818, which is the DS5300 machine type. Select it and press Enter, as shown in Figure 2-26.

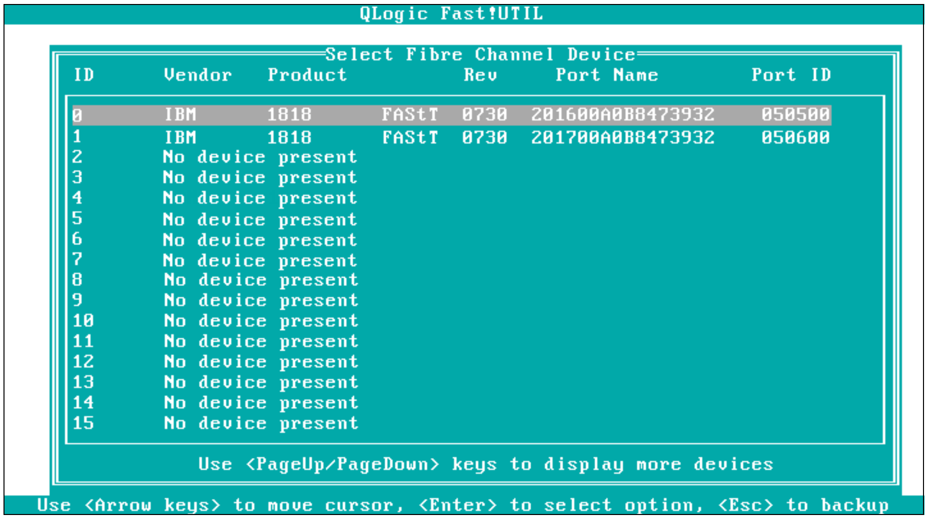


Figure 2-26 Selecting the Storage

6. Because we created the LUN as LUN 0 on the DS5000 storage subsystem, you can see the LUN number 0 here as shown in Figure 2-27. The LUN number displayed here is dependent on your LUN configuration on the storage. Select LUN 0 and press Enter.

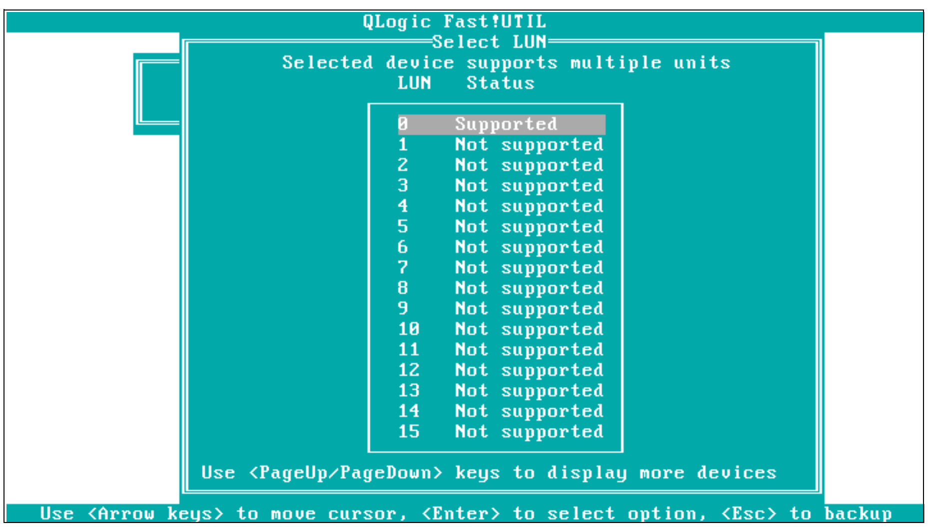


Figure 2-27 Selecting the LUN 0

7. Verify that all the information is correct after it goes back to Selectable Boot Settings; it should look like Figure 2-28. Notice that LUN number on the right side of the figure now matches what we did when we configured our host mappings in 2.8.5, “Creating the LUN to host mapping on DS5000 series” on page 63.

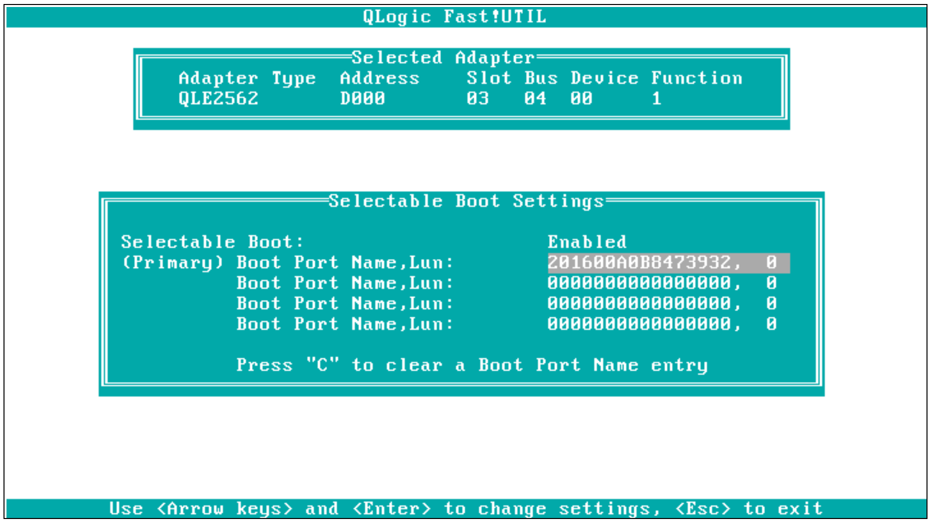


Figure 2-28 Verifying the mapping to the HBA

8. After verifying all the details, keep pressing Esc until it asks you to save your changes. Save the changes and reboot the system to let the changes take effect, as shown in Figure 2-29.

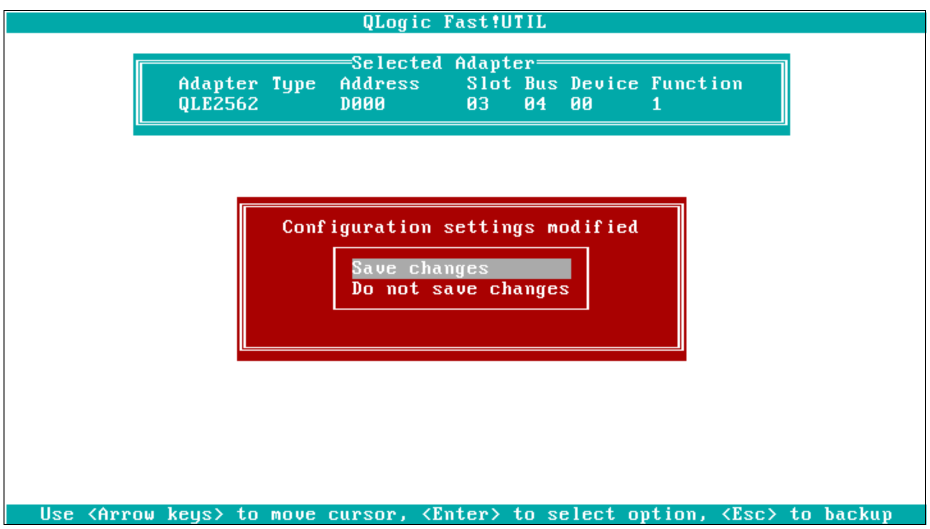
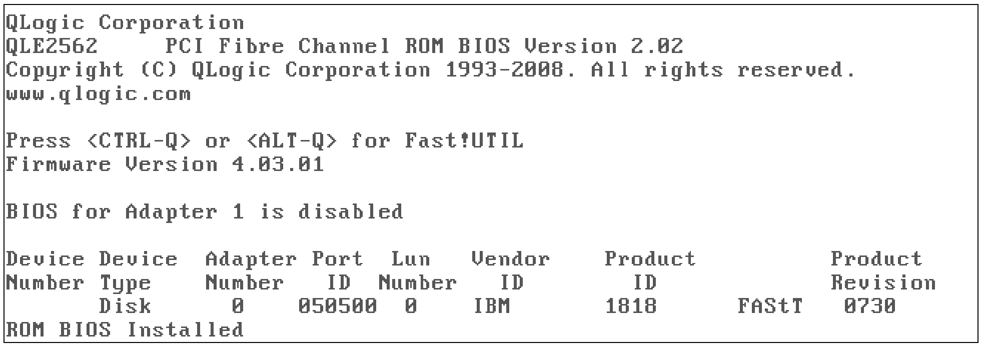


Figure 2-29 Saving QLogic HBA settings

9. When the system reboots, during the post startup, you can see a Disk 0 with a product type of 1818, as shown in Figure 2-30. When the host is able to see the disk on the storage device, you are ready to install your operating system. Depending on the operating system (OS) that you are going to install, you need to look at the setup instructions in this chapter for your specific OS.



```
QLogic Corporation
QLE2562     PCI Fibre Channel ROM BIOS Version 2.02
Copyright (C) QLogic Corporation 1993-2008. All rights reserved.
www.qlogic.com

Press <CTRL-Q> or <ALT-Q> for Fast!UTIL
Firmware Version 4.03.01

BIOS for Adapter 1 is disabled

Device Device   Adapter Port   Lun    Vendor    Product          Product
Number Type     Number   ID   Number    ID        ID             Revision
       Disk       0    050500  0     IBM       1818      FAStT    0730
ROM BIOS Installed
```

Figure 2-30   DS5300 Boot disk found

> **Important:** Unless you are installing VMWare ESX 4i update 1, do not install the second path until the OS is installed and the multi path driver is set up. If the second path is enabled prior to the multi path setup, the OS will see the drive twice.

2.6.2 Adding a second path to QLogic HBA

Before you add the second path, make sure that your OS is installed and multipathing is installed, or you can encounter some problems:

1. To install the second path on the QLogic HBA, use the same procedure described in 2.6.1, “Configuring the QLogic HBA to use the Boot LUN” on page 37, but this time, select the second path as shown in Figure 2-31.

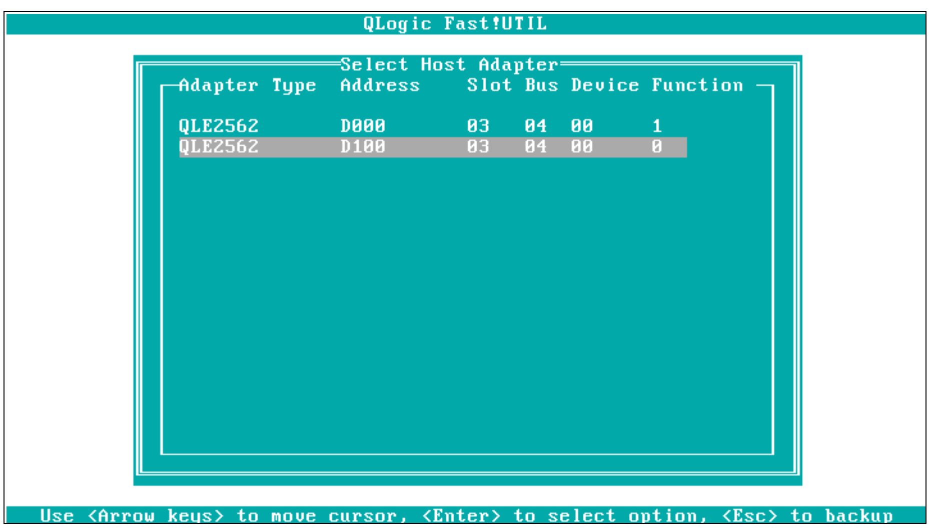


Figure 2-31   Selection of the second port

2. We need to add the second storage controller path as well. You can see that we have added both WWPNs for controller A and controller B, as shown in Figure 2-32.

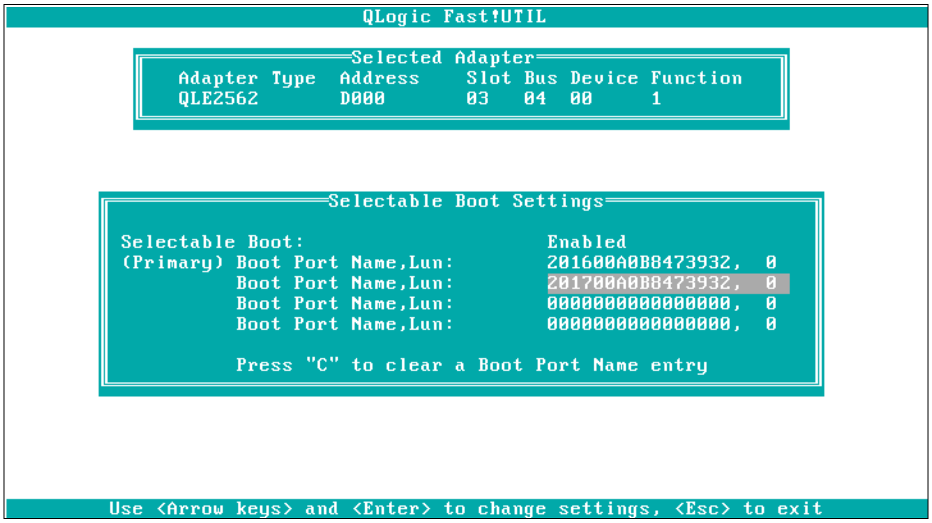


Figure 2-32 Adding the second controller to the QLogic

3. We can now reboot the host server and check the operating system to verify that the changes are applied.

2.7 Emulex HBA configuration

In this section, we configure the Emulex HBA to communicate with the DS5000 storage subsystem. Ultimately, this will allow for a broadcast over the SAN fabric, so that the storage subsystem can recognize the World Wide Port Number (WWPN) of the HBA.

> **Tip:** You must do the zoning configuration on your SAN switch before moving on. We do not cover the SAN zoning steps in this chapter. You need to check with your switch manufacturer about the proper way to do zoning.
>
> See *Implementing an IBM b-type SAN with 8 Gbps Directors and Switches*, SG24-6116, for information regarding the IBM b-type SAN products.

1. During the boot process, press **Alt+ E** or **Ctrl+E** to enter the BIOS configuration utility, as shown in Figure 2-33.

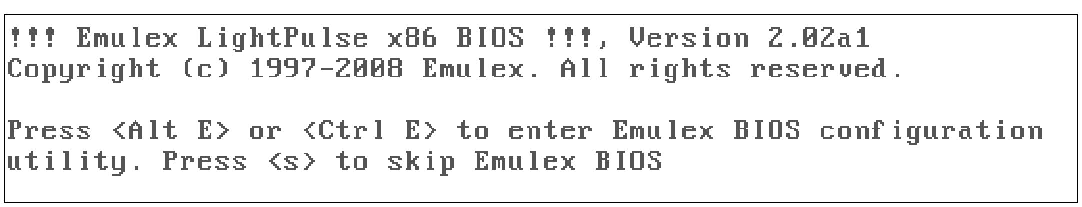


Figure 2-33 Emulex entering the Bios during post

2. Select the first Fibre Channel adapter that is listed by entering 1, as shown in Figure 2-34.

```
        Emulex LightPulse BIOS Utility, UB2.02a1
        Copyright (c) 1997-2008 Emulex. All rights reserved.

        Emulex Adapters in the System:

 1.  42D0494:          PCI Bus, Device, Function (04,00,01)
 2.  42D0494:          PCI Bus, Device, Function (04,00,00)




  Enter a Selection: _

Enter <x> to Exit
```

Figure 2-34 Selection of the Emulex adapter

3. To enable the adapter as a boot device, we need to **Enable** the Adapters Bios. When you are at this panel, enter **2** and press Enter, as shown in Figure 2-35.

```
    Adapter 01:           PCI Bus, Device, Function (04,00,01)

    42D0494:       Mem Base: FE980000  Firmware Version: US1.00A12
    Port Name: 10000000 C98E438B   Node Name: 20000000 C98E438B
    Topology: Auto Topology: Loop First (Default)
    The BIOS for this adapter is Enabled

1.  Configure Boot Devices
2.  Configure This Adapter's Parameters




  Enter a Selection: 2

Enter <x> to Exit       <d> to Default Values      <Esc> to Previous Menu
```

Figure 2-35 Selection of the option, Configure This Adapter's Parameters

4. Now enter **1** and press Enter as shown in Figure 2-36.

```
    Adapter 01:           PCI Bus, Device, Function (04,00,01)

    42D0494:       Mem Base: FE980000  Firmware Version: US1.00A12
    Port Name: 10000000 C98E438B   Node Name: 20000000 C98E438B
    Topology: Auto Topology: Loop First (Default)
    The BIOS for this adapter is Enabled

1.  Enable or Disable BIOS
2.  Change Default ALPA of this Adapter
3.  Change PLOGI Retry Timer (+Advanced Option+)
4.  Topology Selection (+Advanced Option+)
5.  Enable or Disable Spinup Delay (+Advanced Option+)
6.  Auto Scan Setting (+Advanced Option+)
7.  Enable or Disable EDD 3.0 (+Advanced Option+)
8.  Enable or Disable Start Unit Command (+Advanced Option+)
9.  Enable or Disable Environment Variable (+Advanced Option+)
10. Enable or Disable Auto Boot Sector (+Advanced Option+)
11. Link Speed Selection (+Advanced Option+)

  Enter a Selection: 1_

Enter <x> to Exit          <Esc> to Previous Menu
```

Figure 2-36 Section of Enable or Disable BIOs

5. Enter **1** and press Enter to **Enable** the HBA BIOS, as shown in Figure 2-37.

```
    Adapter 01:           PCI Bus, Device, Function (04,00,01)


    The BIOS is Disabled!!


    Enable Press 1, Disable Press 2:_










Enter <x> to Exit          <Esc> to Previous Menu
```

Figure 2-37 Enabling the Emulex BIOs

6. Press Esc twice, to get back to the main panel. Enter **1** to **Configure Boot Device** and press Enter, as shown in Figure 2-38.

> **Tip:** While on this panel, be sure to record the World Wide Port Name (WWPN) for this HBA, in our example it is 10:00:00:00:c9:8e:43:8b. It will be needed while configuring storage partitioning on the DS5000 storage subsystem in 2.8.4, “Creating a host on DS5000 series” on page 58, and also for fabric zone configuration, if it has not been done already. We record the WWPN for the second HBA later, as covered in 2.8.6, “Adding the second path to the Host partition” on page 65.

```
    Adapter 01:           PCI Bus, Device, Function (04,00,01)

    42D0494:       Mem Base: FE980000  Firmware Version: US1.00A12
    Port Name: 10000000 C98E438B   Node Name: 20000000 C98E438B
    Topology: Auto Topology: Loop First (Default)
    The BIOS for this adapter is Enabled

1.  Configure Boot Devices
2.  Configure This Adapter's Parameters

  Enter a Selection: 1

Enter <x> to Exit        <d> to Default Values       <Esc> to Previous Menu
```

Figure 2-38 Configure Boot Device

7. Enter **1** and press Enter, as that is the **Primary Boot**, as shown in Figure 2-39.

```
    Adapter 01:  S_ID: 05003A    PCI Bus, Device, Function (04,00,01)

    List of Saved Boot Devices:

1. Unused   DID:000000 WWPN:00000000 00000000 LUN:00  Primary Boot
2. Unused   DID:000000 WWPN:00000000 00000000 LUN:00
3. Unused   DID:000000 WWPN:00000000 00000000 LUN:00
4. Unused   DID:000000 WWPN:00000000 00000000 LUN:00
5. Unused   DID:000000 WWPN:00000000 00000000 LUN:00
6. Unused   DID:000000 WWPN:00000000 00000000 LUN:00
7. Unused   DID:000000 WWPN:00000000 00000000 LUN:00
8. Unused   DID:000000 WWPN:00000000 00000000 LUN:00

    Select a Boot Entry: 1

Enter <x> to Exit          <Esc> to Previous Menu
```

Figure 2-39 Selection of the primary boot

8. Next we see the storage subsystem, as shown Figure 2-40. One of them listed is a DS5300 with a machine type of 1818. We also see the LUN 1F, as this WWPN has not been set up on the storage subsystem.

```
    Adapter 01:  S_ID: 05003A    PCI Bus, Device, Function (04,00,01)

00. Clear selected boot entry!!
01. DID:050500 WWPN:201600A0 B8473932 LUN:1F    IBM     Universal Xport 0730
02. DID:050600 WWPN:201700A0 B8473932 LUN:1F    IBM     Universal Xport 0730














  Select The Two Digit Number of The Desired Boot Device:

Enter <x> to Exit        <Esc> to Previous Menu       <PageDn> to Next Page
```

Figure 2-40 Scanning for the SAN storage

Leave the host system logged in at this panel, as we return later to complete the setup. At this point you need to configure your storage on the storage subsystem, if you have not done so already. It is covered in 2.8, “Storage provisioning” on page 50.

2.7.1 Configuring the Emulex HBA to use the Boot LUN

After you have your storage subsystem configured in 2.8, “Storage provisioning” on page 50, then you can finish configuring the Emulex HBA, listed in the following steps.

We are assuming you left the Emulex HBA configuration at the panel, as just shown in Figure 2-40. If you are not at that panel, follow the instructions in 2.7, “Emulex HBA configuration” on page 42 to get to that panel.

1. You see an updated list of your storage, as shown Figure 2-41. The one here is a DS5300 with a machine type of 1818. We also see the LUN 0, but is listed as a two-digit number. Here, enter **01** and press Enter to select that storage. You are seeing both controller A and controller B.

Tip: You do not add 02 as of yet. We address that action in 2.7.2, “Adding a second path to the Emulex HBA” on page 50.

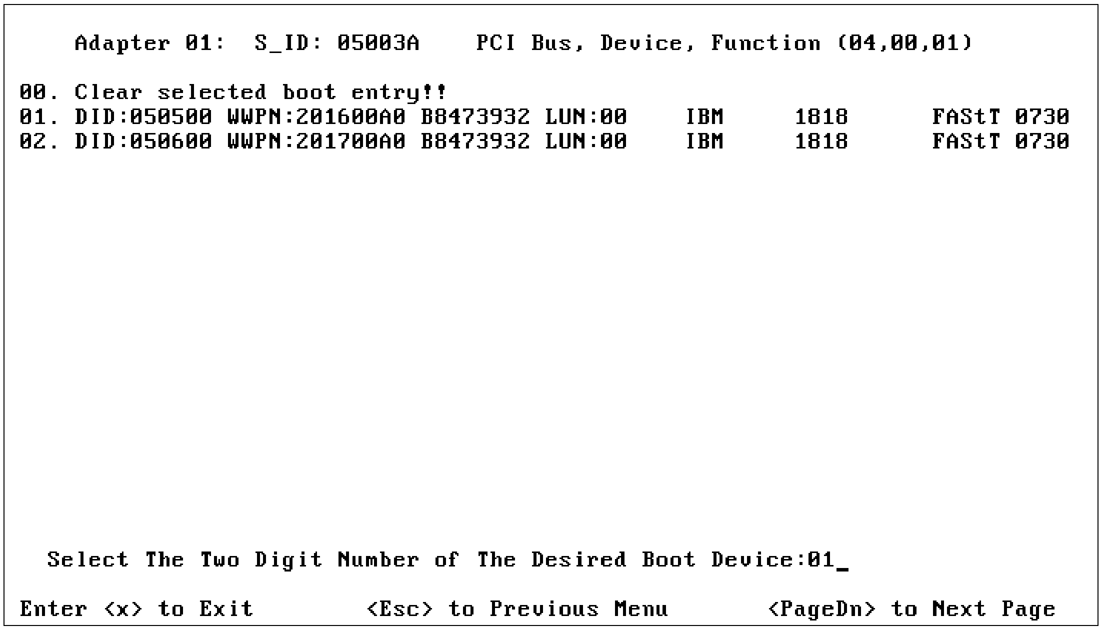


Figure 2-41 Emulex Scanning for Fibre Channel devices

2. You are now asked for the LUN number that you assigned to the LUN when you created it in 2.8.5, “Creating the LUN to host mapping on DS5000 series” on page 63. To add LUN 0, enter it as **00** and press Enter, as shown in Figure 2-42.

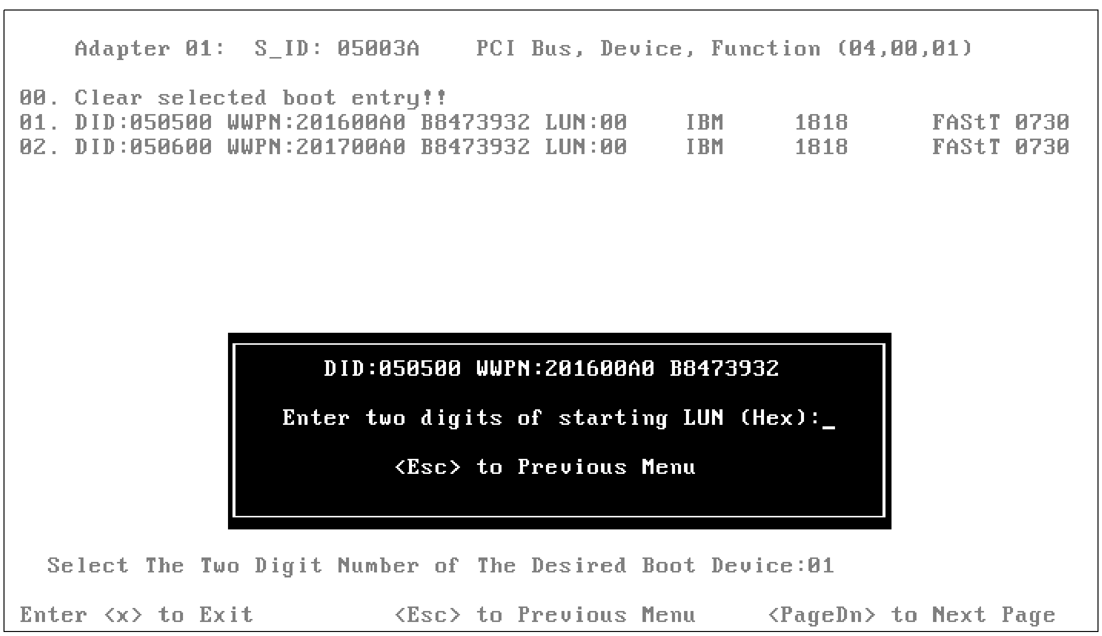


Figure 2-42 Emulex entering the LUN number

3. On the next panel, enter **01** and press Enter, as shown in Figure 2-43.

```
    Adapter 01:  S_ID: 05003A    PCI Bus, Device, Function (04,00,01)

    DID:050600 WWPN:201700A0 B8473932

01.      LUN:00              IBM     1818      FAStT 0730








  Enter a Selection: 01
  B#W: Boot number via WWPN. B#D: Boot number via DID
Enter <x> to Exit          <Esc> to Previous Menu
```

Figure 2-43 Emulex selection of the boot device

4. You are now asked if you want to use the Fibre Channel destination ID (DID), or the World Wide Port Number (WWPN). We select **1** and press Enter, because we are using only WWPNs, as shown in Figure 2-44.

```
    Adapter 01:  S_ID: 05003A    PCI Bus, Device, Function (04,00,01)

    DID:050500 WWPN:201600A0 B8473932

01.      LUN:00              IBM     1818      FAStT 0730


                 DID:050500 WWPN:201600A0 B8473932 LUN:00

                       1. Boot this device via WWPN
                       2. Boot this device via DID

                         <Esc> to Previous Menu
                           Enter a Selection:  1

  Enter a Selection: 01
  B#W: Boot number via WWPN. B#D: Boot number via DID
Enter <x> to Exit          <Esc> to Previous Menu
```

Figure 2-44 Emulex selection of DID or WWPN

You can now see the added LUN as shown in Figure 2-45.

```
    Adapter 01:  S_ID: 05003A    PCI Bus, Device, Function (04,00,01)


    List of Saved Boot Devices:

1. Used      DID:000000 WWPN:201600A0 B8473932 LUN:00  Primary Boot
2. Unused    DID:000000 WWPN:00000000 00000000 LUN:00
3. Unused    DID:000000 WWPN:00000000 00000000 LUN:00
4. Unused    DID:000000 WWPN:00000000 00000000 LUN:00
5. Unused    DID:000000 WWPN:00000000 00000000 LUN:00
6. Unused    DID:000000 WWPN:00000000 00000000 LUN:00
7. Unused    DID:000000 WWPN:00000000 00000000 LUN:00
8. Unused    DID:000000 WWPN:00000000 00000000 LUN:00






    Select a Boot Entry: _


Enter <x> to Exit          <Esc> to Previous Menu
```

Figure 2-45 Emulex with LUN shown

2.7.2 Adding a second path to the Emulex HBA

Before you add the second path, make sure that your operating system is installed and that multipathing is installed; otherwise, you can encounter some problems.

To install the second path on the Emulex HBA, you follow the steps in 2.7.1, “Configuring the Emulex HBA to use the Boot LUN” on page 46, but this time you select the second path, as well as setting up both paths on the second adapter (Figure 2-46). As you can see, there are two paths here; one of them goes to controller A and the other goes to controller B.

```
    Adapter 01:  S_ID: 05003A    PCI Bus, Device, Function (04,00,01)


    List of Saved Boot Devices:

1. Used      DID:000000 WWPN:201600A0 B8473932 LUN:00  Primary Boot
2. Used      DID:000000 WWPN:201700A0 B8473932 LUN:00
3. Unused    DID:000000 WWPN:00000000 00000000 LUN:00
4. Unused    DID:000000 WWPN:00000000 00000000 LUN:00
5. Unused    DID:000000 WWPN:00000000 00000000 LUN:00
6. Unused    DID:000000 WWPN:00000000 00000000 LUN:00
7. Unused    DID:000000 WWPN:00000000 00000000 LUN:00
8. Unused    DID:000000 WWPN:00000000 00000000 LUN:00


    Select a Boot Entry:

Enter <x> to Exit          <Esc> to Previous Menu
```

Figure 2-46 Emulex with second path added

2.8 Storage provisioning

Here we describe steps involved in DS5000 storage provisioning procedures. For detailed instructions of all the topics described here, see the *IBM Midrange System Storage Hardware Guide*, SG24-7676 and the *IBM Midrange System Storage Implementation and Best Practices Guide*, SG24-6363.

2.8.1 Configuring the DS5000 storage subsystem

The DS5000 storage subsystem configuration must be performed from the Remote Management workstation using the IBM System Storage DS Storage Manager 10 Client utility. Complete the following steps to configure a LUN for SAN Boot on the DS5000 storage subsystem.

> **Tip:** If you have not installed IBM System Storage DS Storage Manager 10 Client on your remote system yet, see 2.3.2, "Installation of the Storage Manager" on page 24.

Complete the following steps to configure the DS5000 Subsystem for SAN Boot:

1. To get started, launch the IBM System Storage DS Storage Manager 10 Client. After it is started, you see a panel that looks like the one in Figure 2-47. You need to *double-click* your storage subsystem.

   > **Attention:** You can have many storage subsystems list here, so make sure that you pick the correct one.

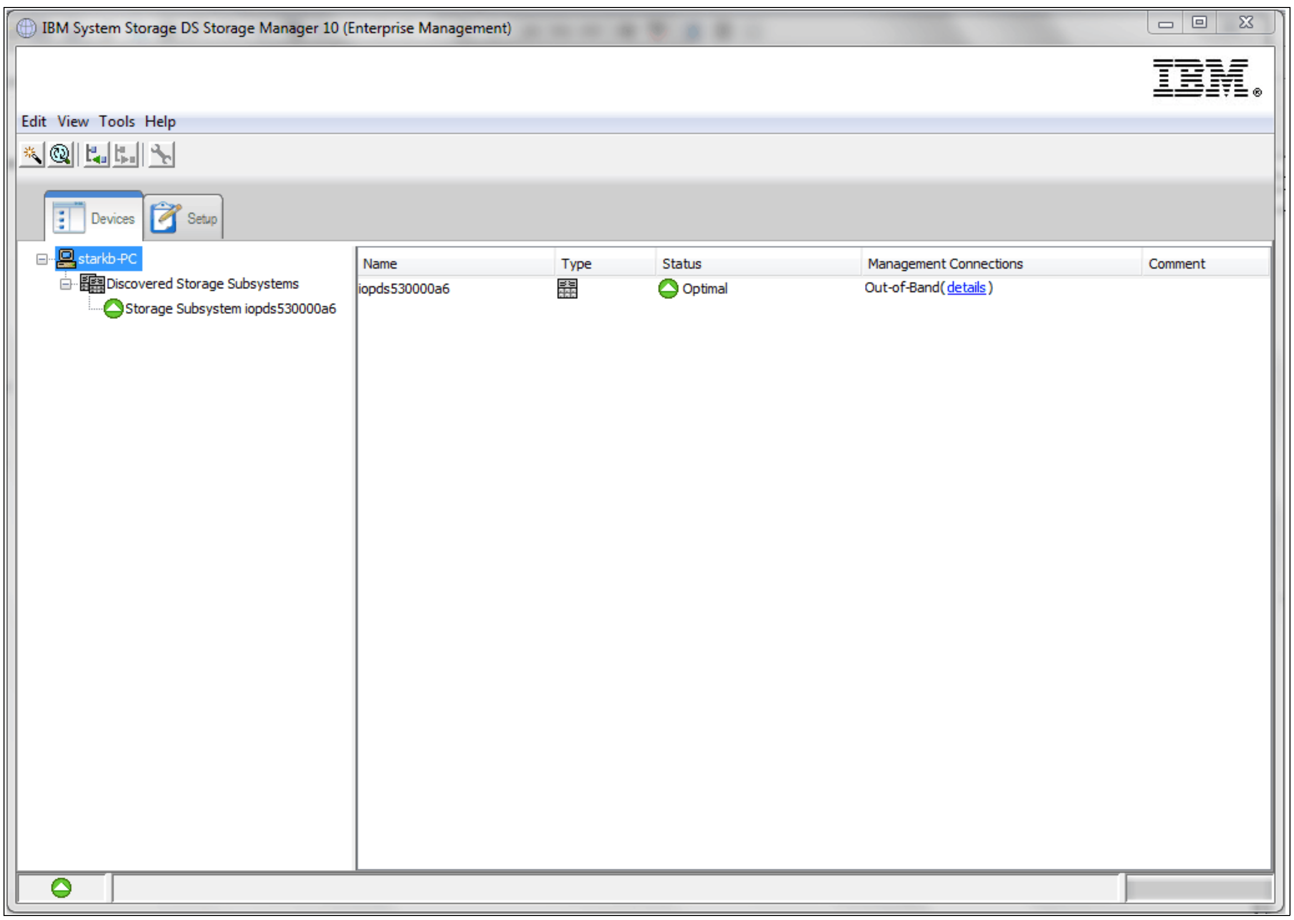


Figure 2-47 DS Storage Manager (Enterprise Management)

2. When you click your storage subsystem, you will see the **Summary** tab, if it is the first time you installed and opened the IBM System Storage DS Storage Manager 10 Client. If you have opened Storage Manager before, and went to a different tab, it will open at that tab and not the **Summary** tab.

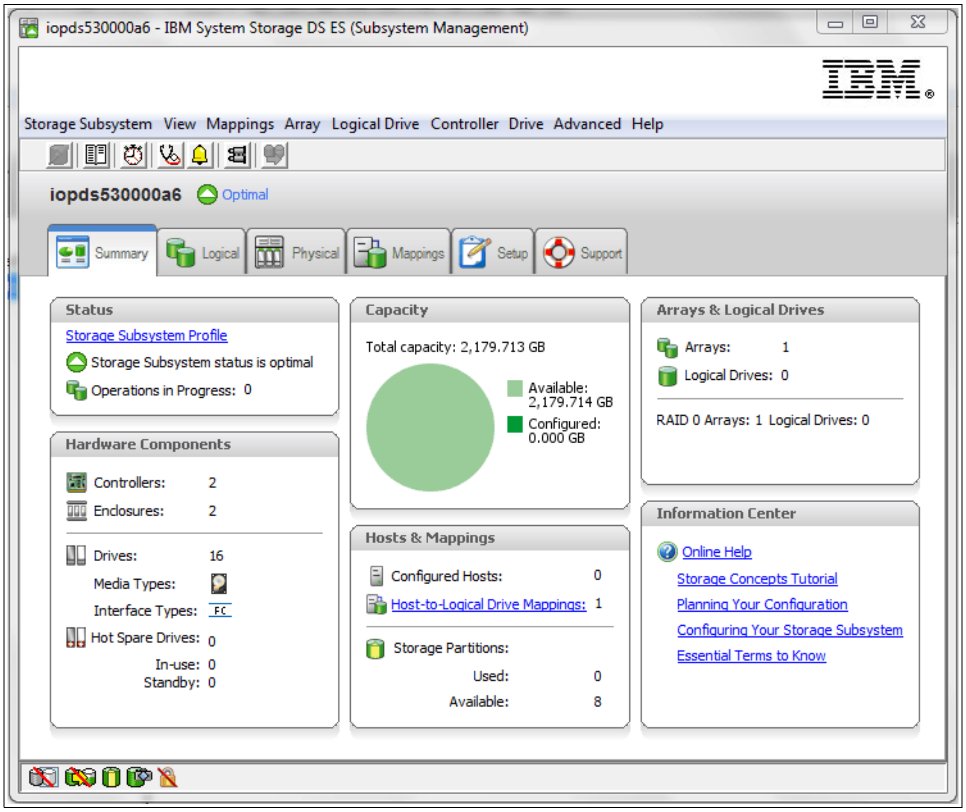


Figure 2-48 Summary tab

Tip: The Summary tab displays plenty of useful information, but here we are mostly going to look at the *Storage Partitions*, used and available. The basic configuration for the DS5000 series comes with eight partitions; additional partitions will be a License upgrade and must be purchased. The DS5000 series supports up to 512 partitions. For each Boot LUN, one (1) partition is needed.

2.8.2 Creating the RAID array on the DS5000 series

Follow this procedure to create a RAID array on the DS5000 storage subsystem:

1. Select the **Logical** tab. You need to create the array before you can create a logical drive. To create the array, *right-click* the **Total Unconfigured Capacity,** and click **Create Array**, as shown in Figure 2-49.

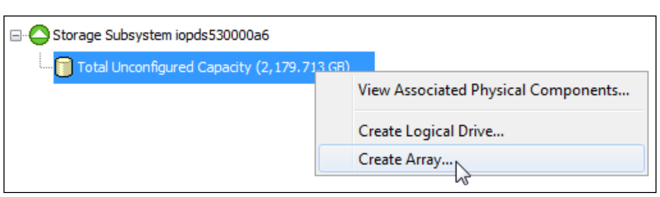


Figure 2-49 Creating the array

2. After you click **Create Array**, in the next panel, click **Next,** as it is an information panel only. After you click **Next**, you are asked to choose a name for your array. Array names cannot have spaces in the name and can only be used one time in the SAN.
3. You are also asked to either select the disks to be used or let the system select them for you. If you select **Automatic**, it will select the disks in consecutive order; it is fine if you only have one expansion tray. If you have more than one tray, you might want the disks to be spread across expansion trays. This method is called enclosure protection, as shown in Figure 2-50. For this installation, we select **Manual** to explore the various options of the manual configuration.

> **Tip:** Storage Enclosure protection is a best practice. In the event that you have an enclosure failure, you will not lose your data, as it is protected across the other enclosures.

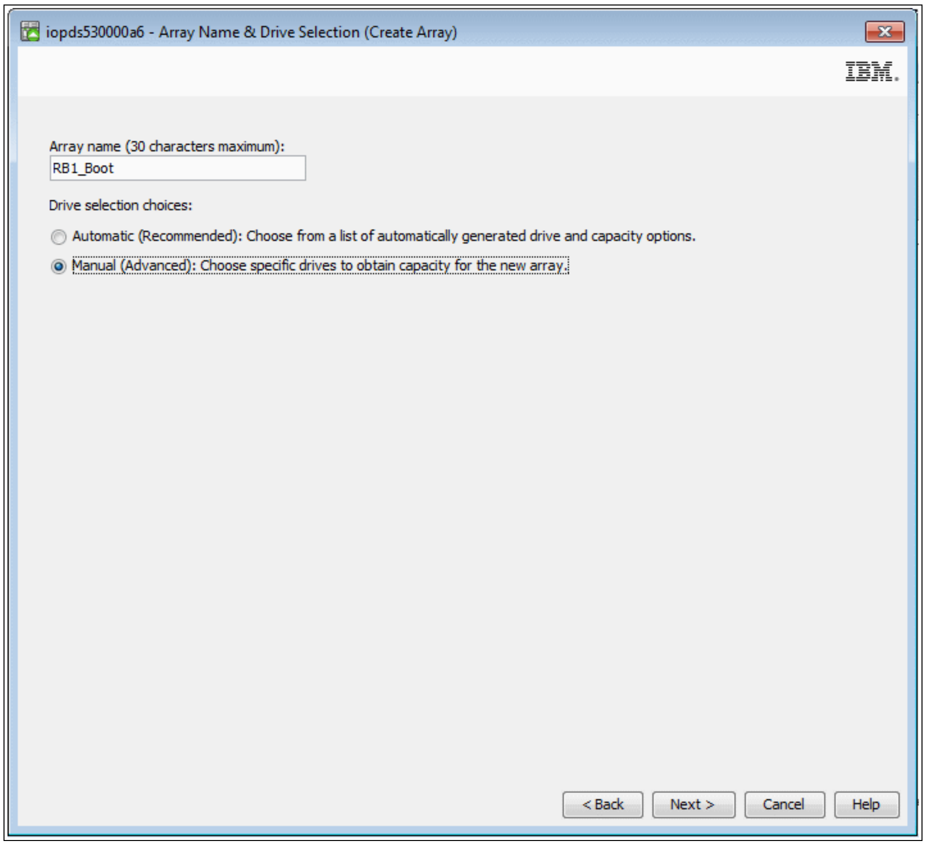


Figure 2-50 Creation of the name for the array

4. Next we select the RAID level, pick the drive type, and the number of drives that will be used for the SAN Boot LUN. For this installation and best practice for SAN Boot, we are going to use RAID 0, and we select the first two drives, as we only have one enclosure in our setup. After the drives are added, click the **Calculate Capacity** button and then click **Finish** as shown in Figure 2-51.

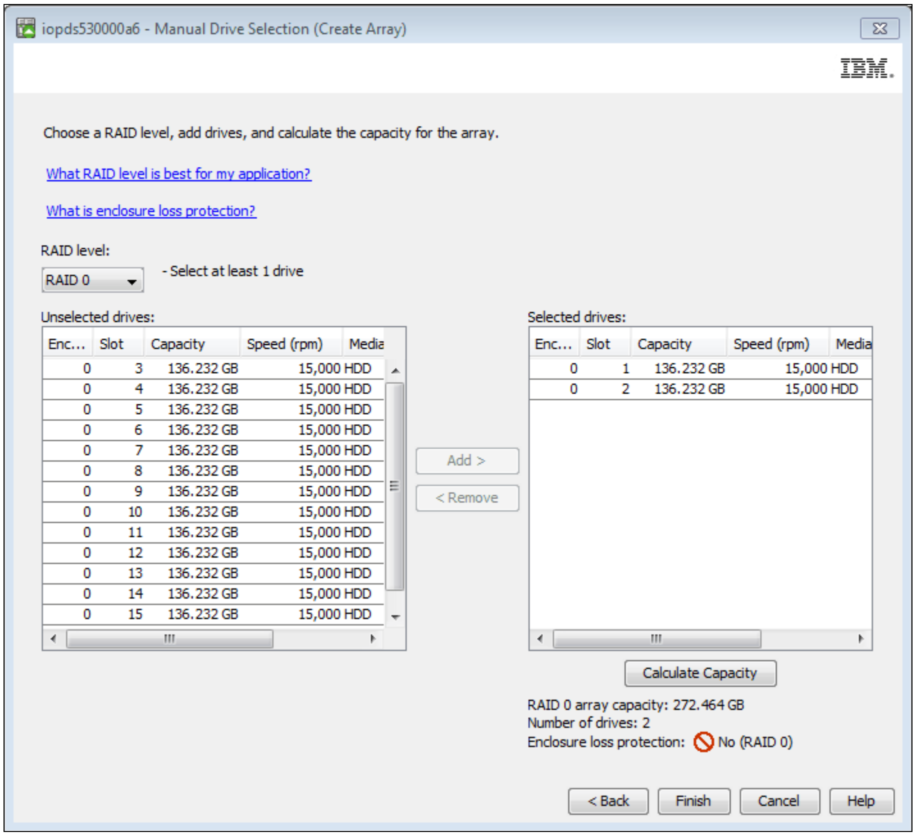


Figure 2-51 Selection of the RAID and drives

5. After you click **Finish**, you will see the panel shown in Figure 2-52. If you want to create more RAID Arrays, then you can select **Yes** and it will take you back to step 4 in this section. For our configuration, we select **No**.

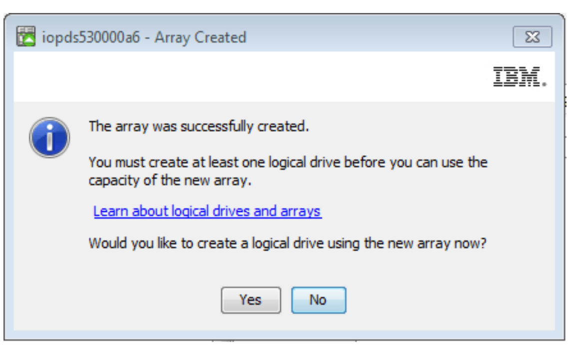


Figure 2-52 Raid creation finish panel

2.8.3 Creating a logical drive on the DS5000 series

Follow these steps to create a logical drive on the DS5000 storage subsystem:

1. To create the logical drive, in the **Logical** view, right-click the free space on the RB1_Boot array that was created, and click **Create Logical Drive**, as shown in Figure 2-53.

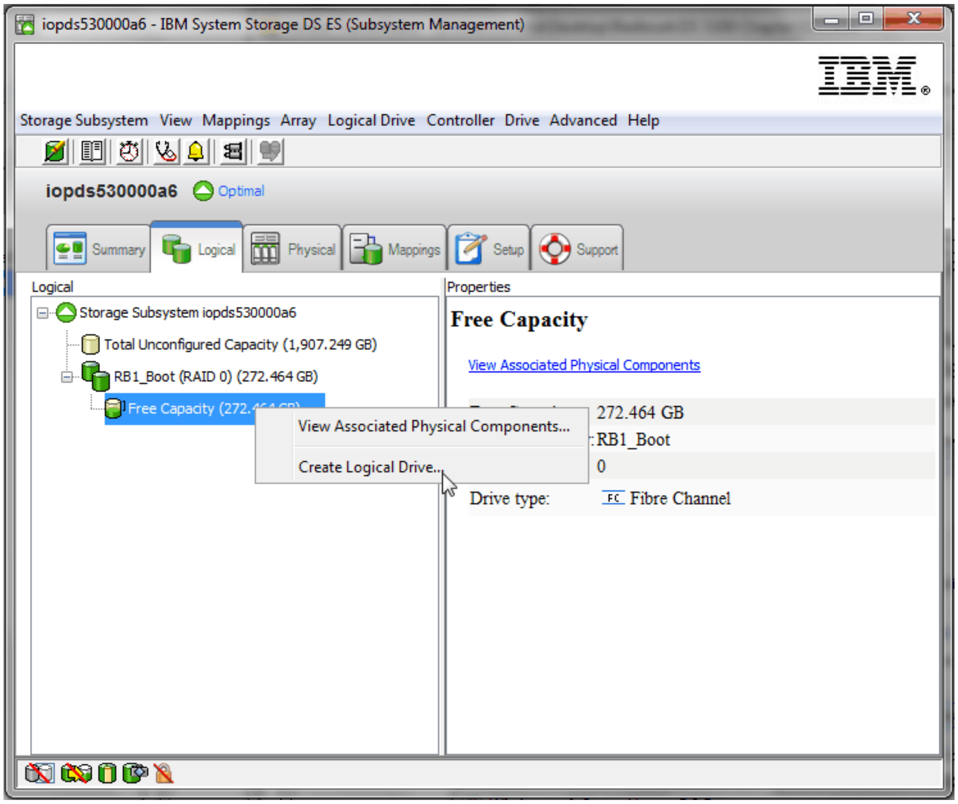


Figure 2-53 Selection create logical drive

2. After clicking **Create Logical Drive**, you will see an informational panel, click **Next** here. In the following panel as shown in Figure 2-54, you need to select the size of the drive you want to present as your SAN Boot LUN. Here you need to give the logical drive a name.

   The same policy as in the RAID creation applies here. The names cannot have any spaces and cannot use the same name anywhere in the SAN. You are given the choice of selecting **Use recommended setting** or **Customize settings** under the Advanced logical drive parameters. These parameters can also be changed later. We are going to select **Customize settings** here to show all the options that are available.

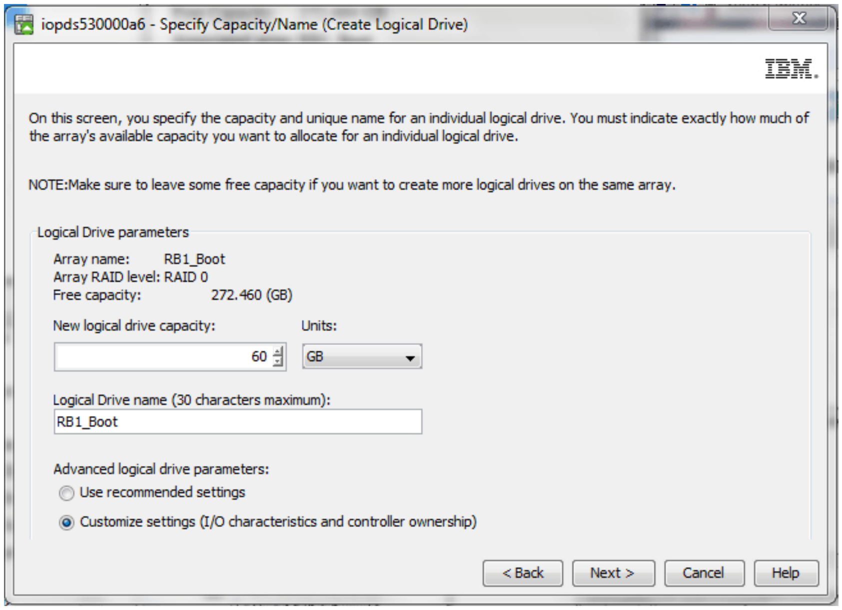


Figure 2-54 Creating the logical drive

3. In the Customize Advanced logical drive window, you will select the type of drive you require. For SAN Boot it will always be a ***File System*** type, even if will be using the boot LUN for running a Data base. It will be **File System** by default so you will just click **Next**, as shown in Figure 2-55.

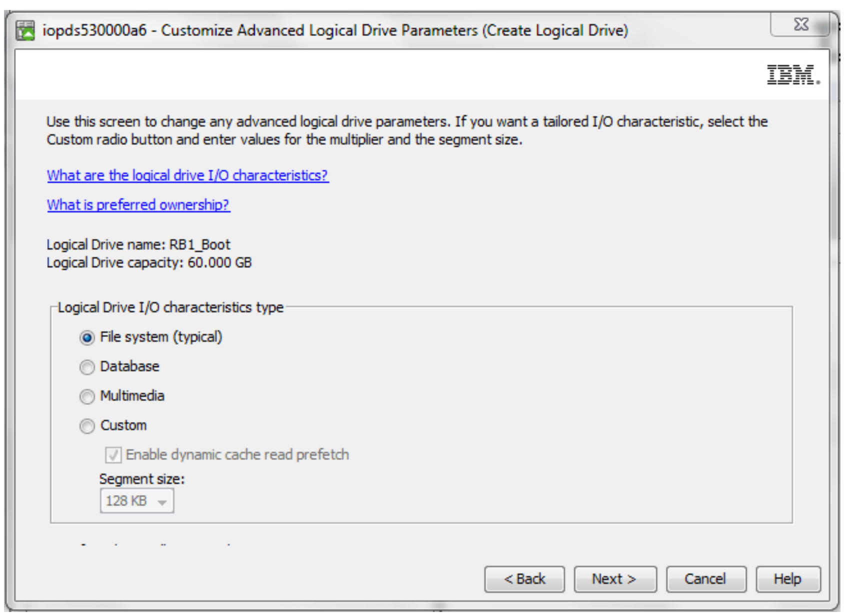


Figure 2-55 Selection of the file type

4. On the Specify Logical Drive-to-LUN Mapping window, we pick **Map** later using the mappings view, because the hosts are not yet created on the storage subsystem, as shown in Figure 2-56.

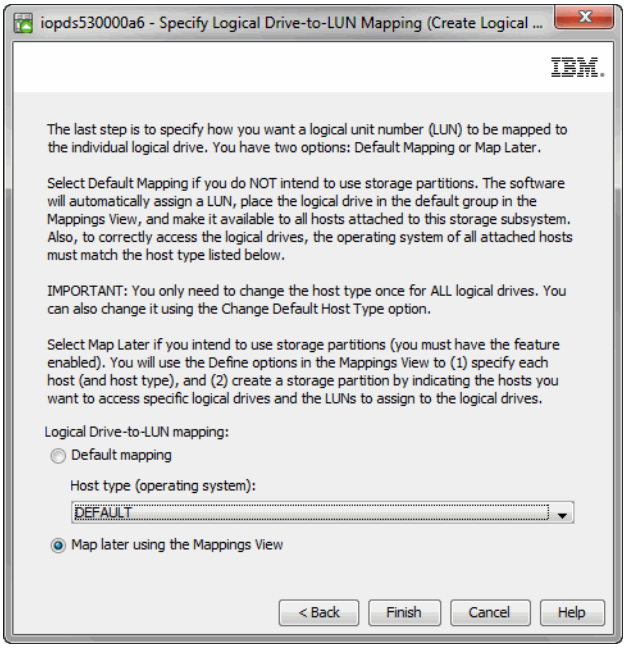


Figure 2-56 Default mapping

5. After you click **Next** on the Specify Logical Drive-to-LUN Mapping window, it will create the logical drive. When it is created, it asks you if you want to create another logical drive, but here we click **No**, as shown in Figure 2-57.

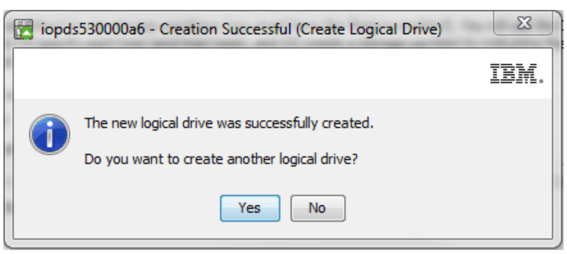


Figure 2-57 finished creating the logical drive

Tip: You need only to do Step 6 on page 58 if you are doing a direct connection to the SAN, with no switch, or you have your switch set up similar to the diagram in Figure 2-58.

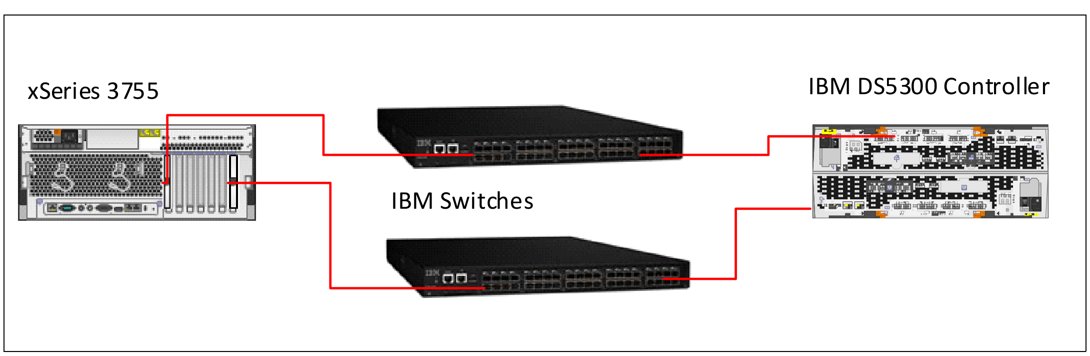


Figure 2-58 Switch with no redundancy

6. After the logical drive is created, you need to make sure that the logical drive is on the path that you will booting from. When creating the logical drives, the DS5000 series system staggers the ownership to controller A first and then B second. However, when we are configuring SAN Boot, we need to have all the logical drives that we are booting from, on controller A in our setup, as shown in Figure 2-59.

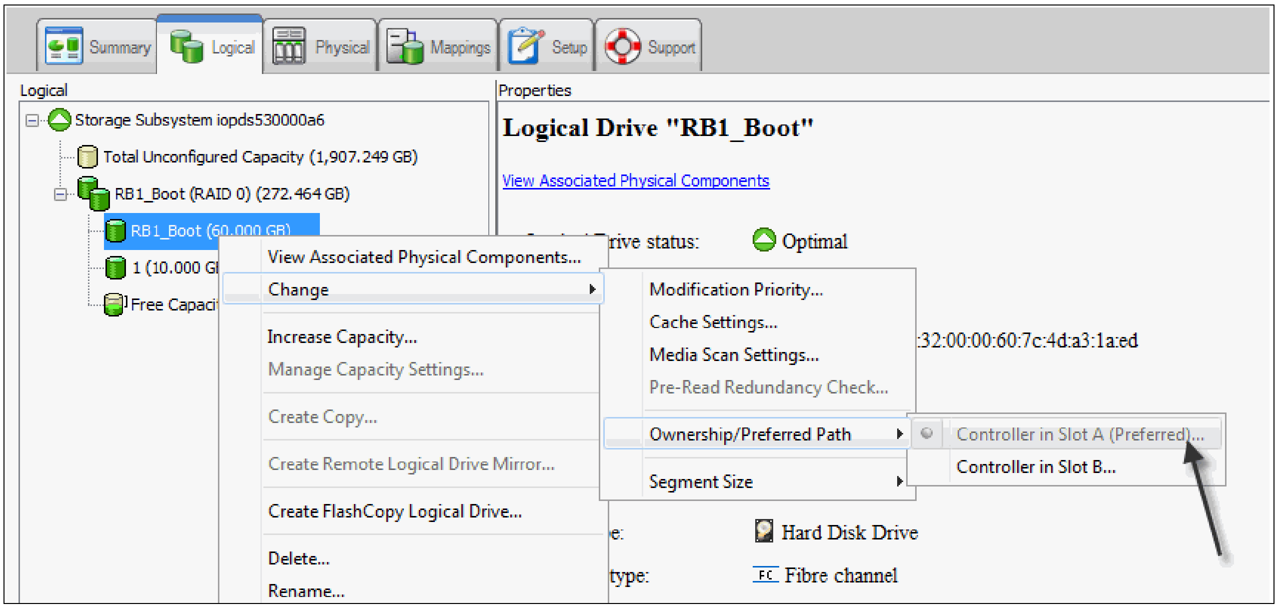


Figure 2-59 Changing the path to Controller A

2.8.4 Creating a host on DS5000 series

Follow this procedure to create a host on the DS5000 storage subsystem:

1. After the logical drive is created, click the **Mappings** tab. Here you will see the logical drive that you just created. You will see it under **Undefined Mappings**, as shown in Figure 2-60.

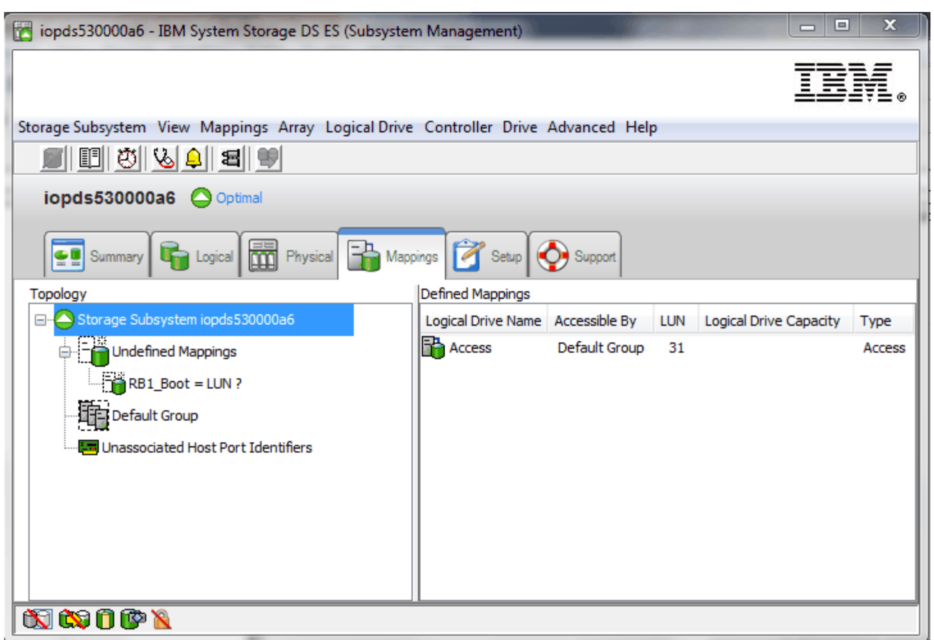


Figure 2-60 Undefined Mappings view

2. To create the host, right-click your storage subsystem, then click **Define** → **Host**, as shown in Figure 2-61.

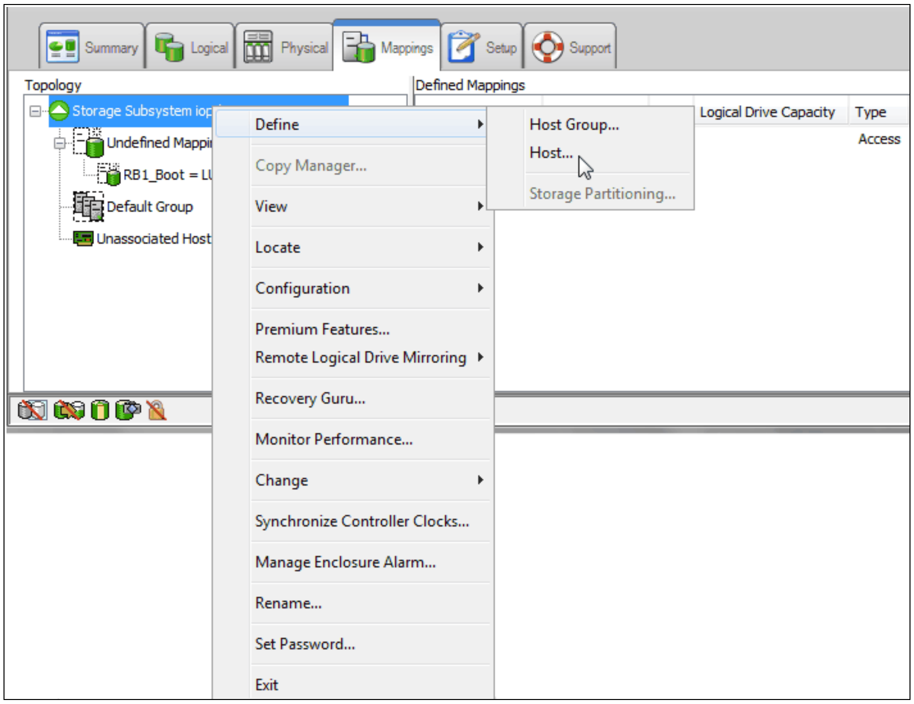


Figure 2-61 Defining the Host

3. When you specify the host name, the same policy applies: no spaces, and names that were not used before. You will see a question asking if you plan to use storage partitioning. We say **Yes** to this panel, as shown in Figure 2-62.

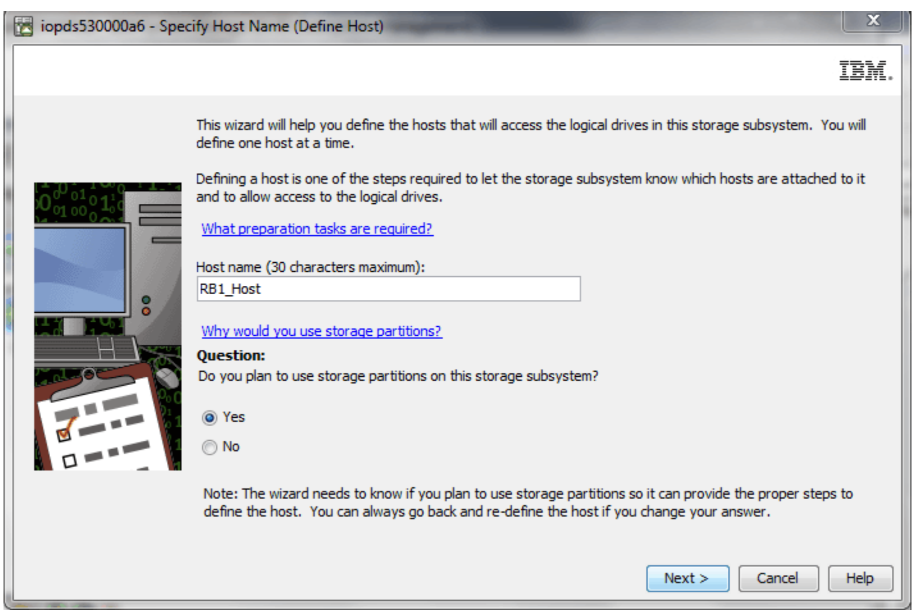


Figure 2-62 Creating the host name

4. We must now define the host port or ports. For SAN Boot, we only configure one port now and add the other later. It is where we need the WWPN recorded earlier in 2.6, “QLogic HBA configuration” on page 33 or 2.7, “Emulex HBA configuration” on page 42. You also need to create an Alias for the adapter. We used RB1_Host_Port1, because it is port1 on a 2 channel adapter. Click **Add**, as shown in Figure 2-63.

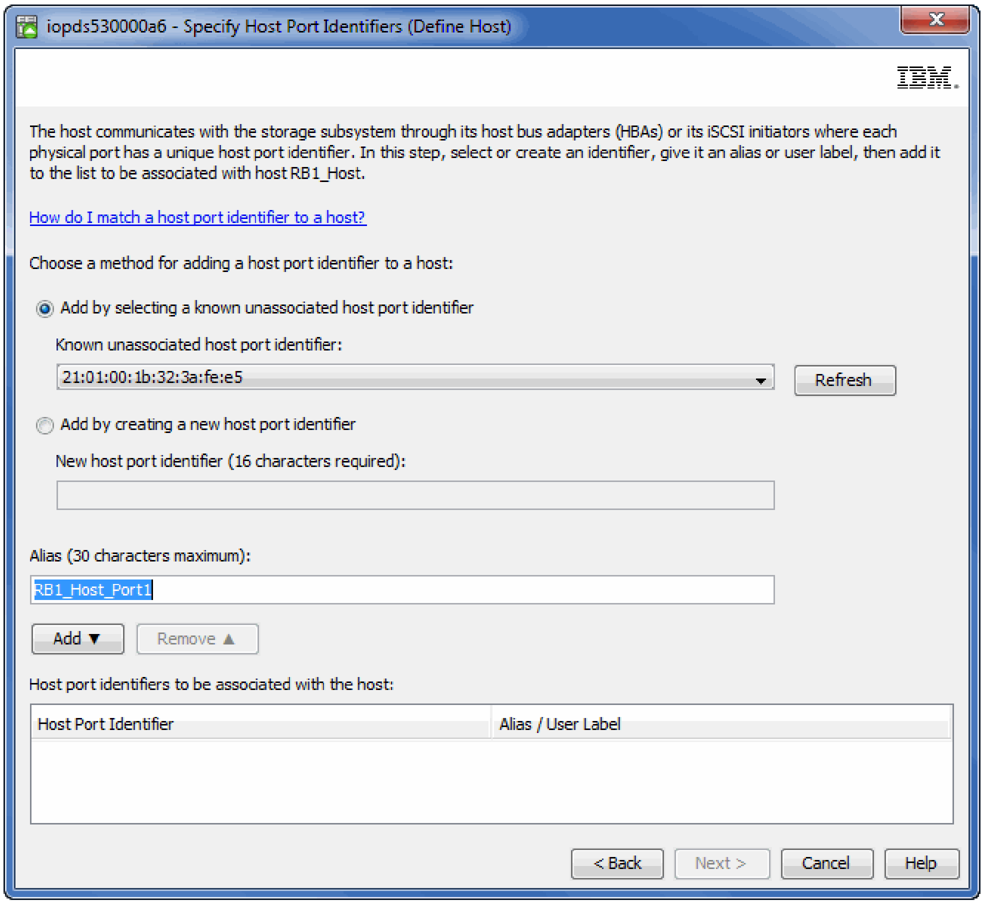


Figure 2-63 Creating the Host adapter

5. When you click **Add,** the host port identifier is added to the alias, as shown in Figure 2-64.

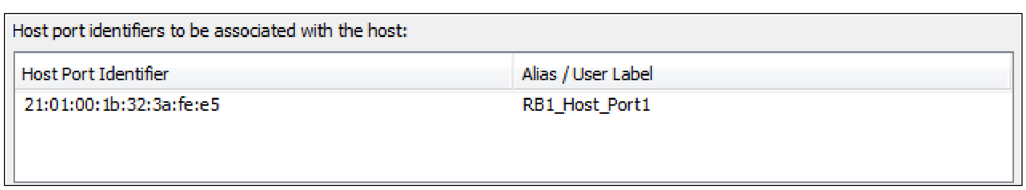


Figure 2-64 Added the host port

6. We now need to define what type of host this will be. It is used by the storage subsystem to understand how to present the storage to the host operating systems. It is important to define the correct host type. There are several other common operating systems available from the drop-down box. In our example, we are going to install Windows 2008 R2. This choice will be applied only for this host as shown in Figure 2-65.

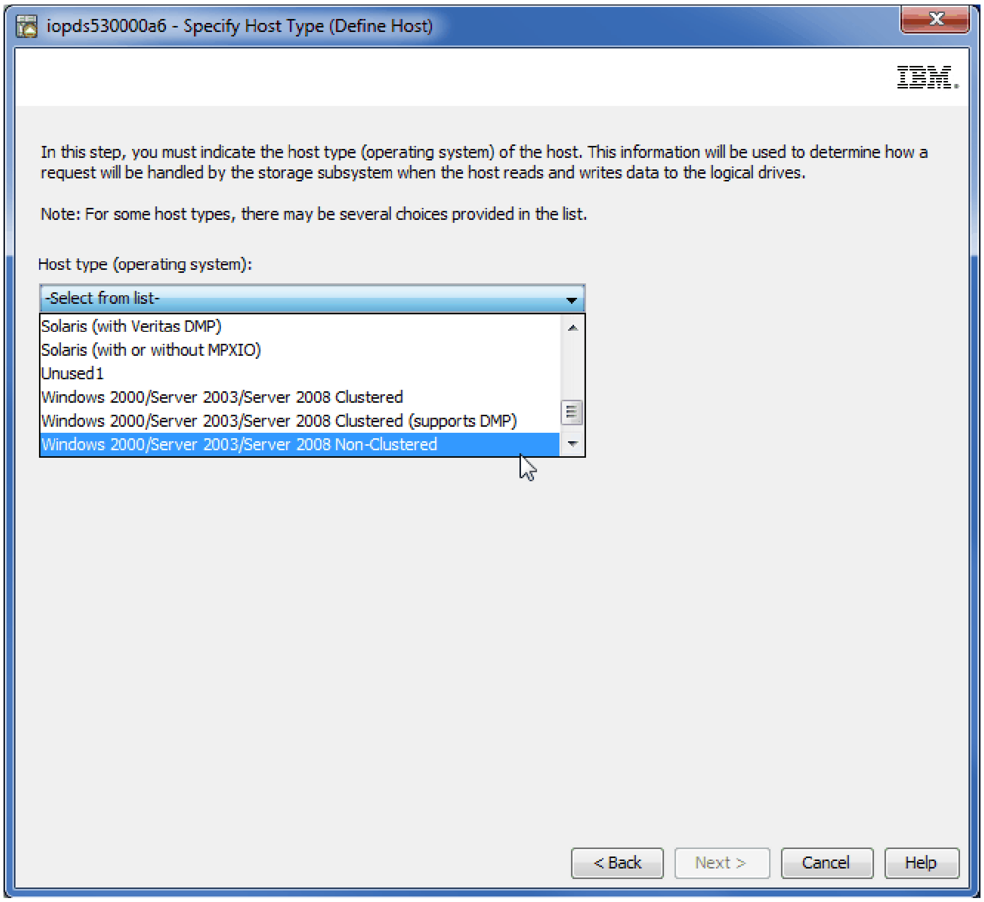


Figure 2-65 Selection of the host type

7. You are asked if this host will be part of a cluster, and if the host is going to share access. We select **No** because at this time, our installation will not share access, as shown in Figure 2-66.

> **Tip:** If our installation was going to be a Windows cluster or Linux cluster doing SAN Boot, we would select **Yes**.

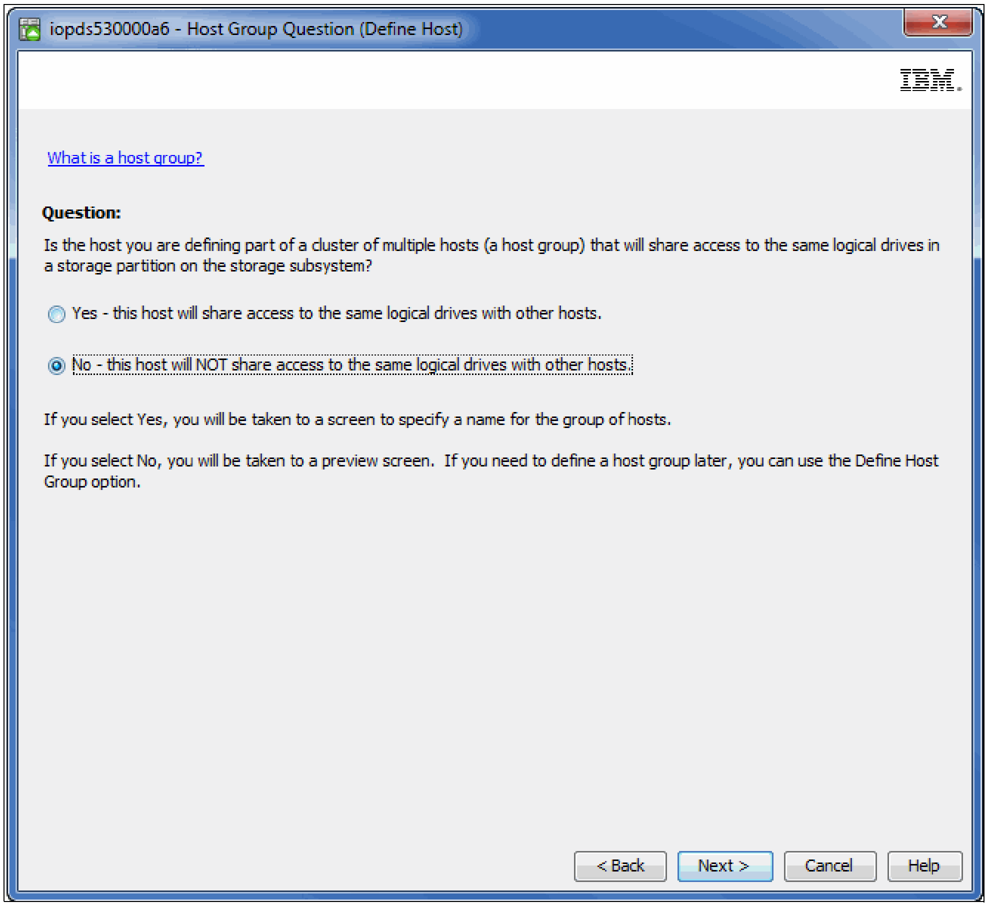


Figure 2-66 Shared access to the same logical drive

8. In the next panel, we review all the configuration changes, before we create the host. After reviewing, click **Finish**, as shown in Figure 2-67.

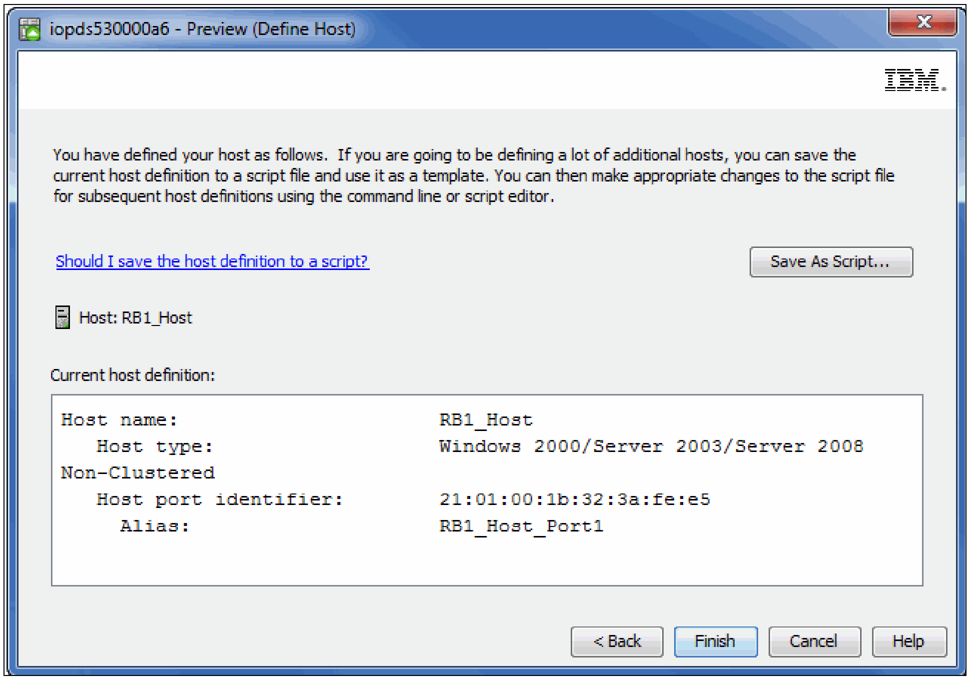


Figure 2-67 Verifying your host configuration

2.8.5 Creating the LUN to host mapping on DS5000 series

Follow this procedure to create the LUN to Host mapping on the DS5000 storage subsystem:

1. After the host is created, we can now assign the LUN to the host. To do it, you need to right-click the LUN that we created, and click **Define Additional Mapping**, as shown in Figure 2-68. In our example, it is RB1_Boot = LUN ?. The reason it appears in this way is because it does not have a LUN assigned to it.

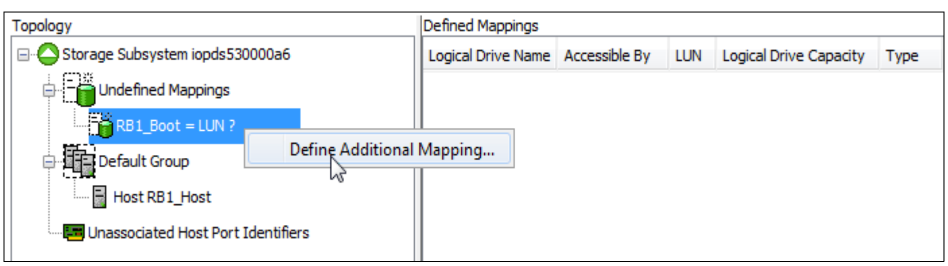


Figure 2-68 Define Additional Mapping

2. In the *Define Additional Mapping* window, select the host that we created previously RB1_Host. Check the LUN number. By default, it will be LUN 0, unless you have another LUN 0 assigned to this host previously. It is best practice to assign the boot LUN as LUN 0. Next, select the drive from the list in the **Logical Drive** list, as there could be more than just RB1_Boot listed. Then click **Add**, as shown in Figure 2-69.

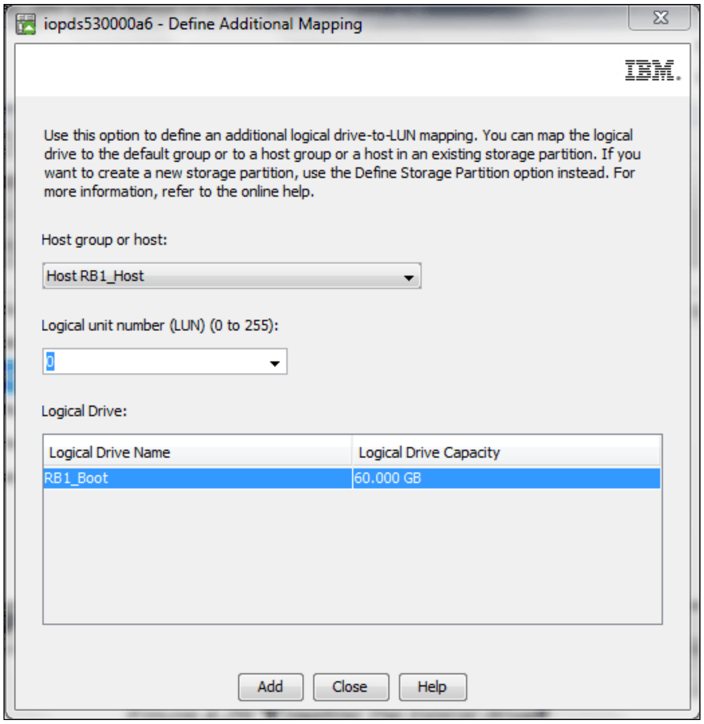


Figure 2-69 Assigning the LUN to the host

3. When the LUN is added, the resulting panel will be similar to the one shown in Figure 2-70.

> **Attention:** Notice that there are two logical drives, a LUN 0 and 31 mapped to the host. LUN 31 is an access LUN, and during the operating system (OS) installation, you see it as a 20 MB drive. The access LUN is read-only and cannot be used for data storage. If you choose the access LUN during the OS installation, the installation will fail.

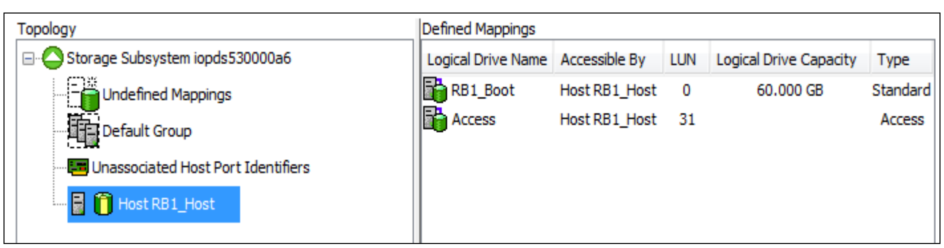


Figure 2-70 The created host with the assigned LUN

For detailed configuration steps and more information about the access LUN, see the *IBM Midrange System Storage Hardware Guide*, SG24-7676.

At this point, the storage configuration process for the primary path is complete. Ensure that the secondary path from the host system to the DS5000 controller B is disconnected respectively (not included in the fabric zone configuration). We cover it in 2.8.6, "Adding the second path to the Host partition" on page 65.

2.8.6 Adding the second path to the Host partition

Warning: Do not start this process until you have installed the operating system on the host and configured multipathing.

In this section, we show how to add the second HBA path to the already configured Host partition.

1. We start in the Mappings tab of the DS Manager. To add the second path to the host RB1_Host, you right-click the host and click **Manage Host Port Identifiers**, as shown in Figure 2-71.

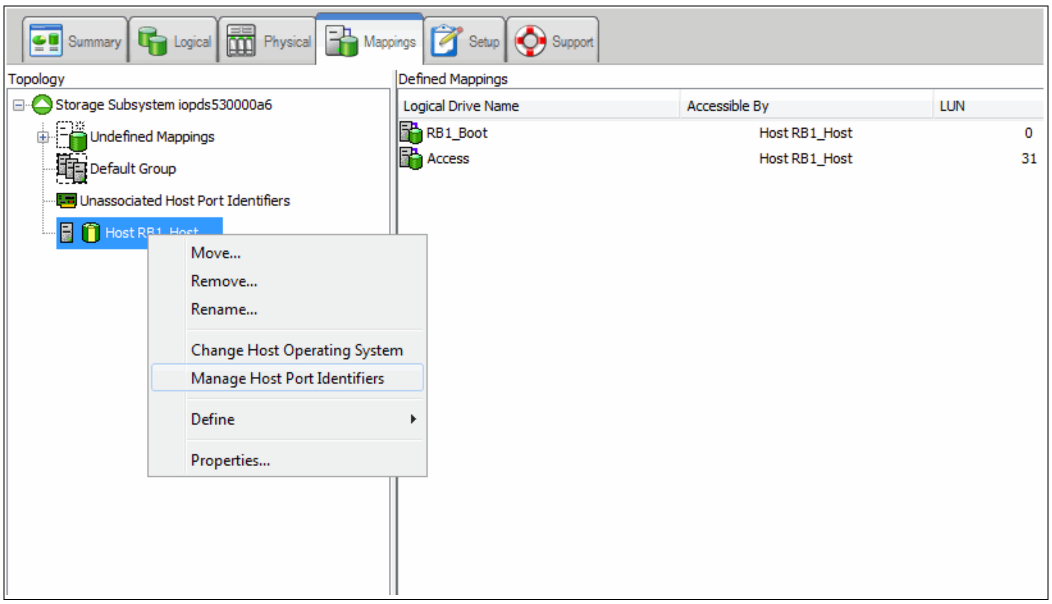


Figure 2-71 Adding WWPN to host

2. You can see a window, Manage Host Port Identifiers, as shown in Figure 2-72. Make sure that the host selected in the drop-down box matches the host to which you want to add the WWPN. Ours shows RB1_Host, which is the one to which we want to add the port. After verifying it, click **Add**.

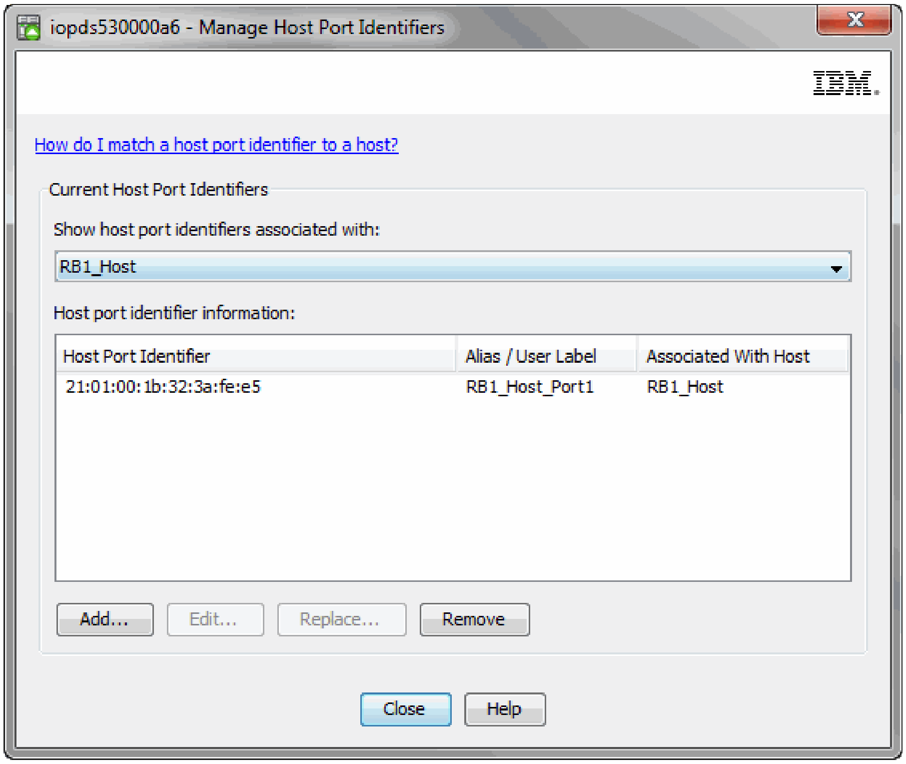


Figure 2-72 Add the WWPN

3. Select the correct WWPN from the drop-down in the *Know unassociated host port identifier*, as shown in Figure 2-73. In this case we select 21:00:00:1b:32:1a:fe:e5 as the one we are going to add.

   For easy identification, we create an alias for this WWPN. Keeping with our naming convention, we type the name as RB_Host_Port0.

> **Tip:** If you do not see the WWPN in the drop-down list, it can be caused by a number of problems, most commonly as follows:
>
> - WWPN already assigned
> - Zoning not done properly
> - Cable disconnected

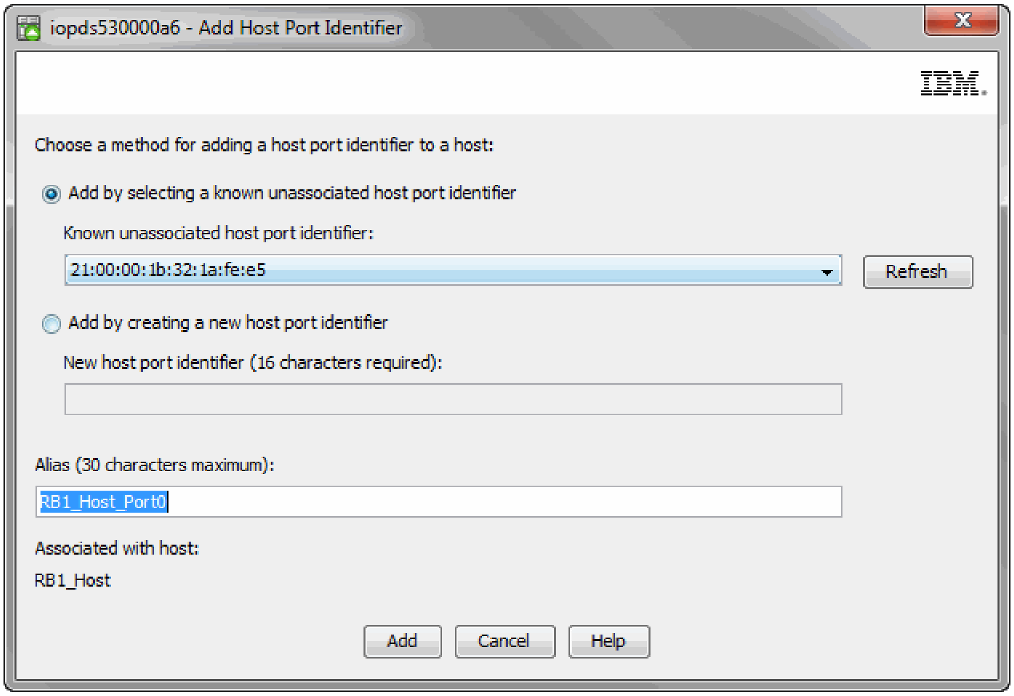


Figure 2-73 Creating the Alias to the WWPN

4. When the port is added, the resulting panel will be similar to the one shown in Figure 2-74. Verify the changes and click **Close**.

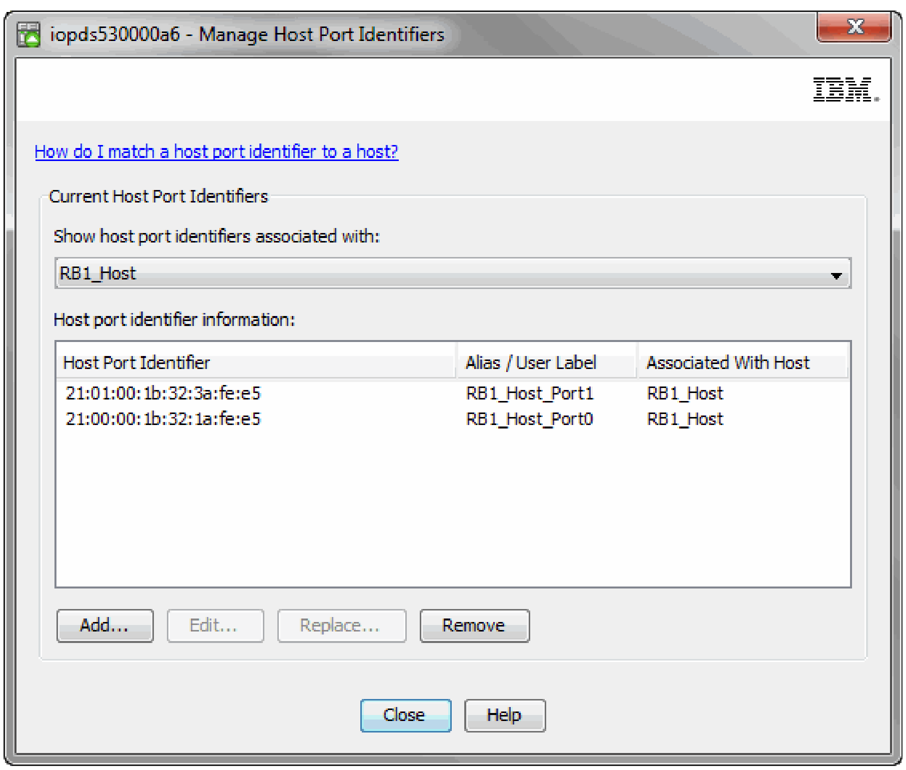


Figure 2-74 Added the WWPN and verified

5. Now that you have the second path added to the Storage Manager, you can go back to the HBA section to finish the configuration as documented in 2.6.2, “Adding a second path to QLogic HBA” and 2.7.2, “Adding a second path to the Emulex HBA”

2.8.7 Disabling AVT for Linux Host type

When using the DS5000 storage subsystem, Auto Logical Drive Transfer (ADT/AVT) mode is automatically enabled in the Linux storage partitioning host type. This mode causes contention when the RDAC driver is installed. If using the “Linux” host type, it must be disabled using the script that is bundled in this Linux RDAC web package or in the \Scripts directory of this DS Storage Manager version 10 support for Linux CD. The name of the script file is DisableAVT_Linux.scr. If opened in notepad, it contains the following coding, as shown in Example 2-1.

Example 2-1 DisableAVT_Linux.scr

```
// IMPORTANT: You Must REPLACE Z In The Script Commands In This File With The
// Appropriate Host Number/Index Of The Host Type That You Want To Disable AVT.
//
// Name: Disable AVT Script
// Date: 02-26-2003
// Revision:
// - 08-20-2004, the value z has been change to 5 so that the user does not need
//                       to make manual modifications to Disable avt for Linux
//                       het host region.
// - 01-30-2006, fix typo in comment field
//
//
// Comments:
// This script is to set the HostNVSRAMByte 0x24 to 0 for disabling AVT.
// Because in fw 5.x or higher, the AVT is Disable by host type; one
// must supply to correct number of the host type in order for the
// script to work correctly. Please refer to the table below for
// the cross reference between host type and host number.
//
// Host number/indexHost Type                        Note
//     0Windows Non-Clustered (SP5 or higher)
//     1Windows Clustered (SP5 or higher)
//     2Windows 2000 Non-Clustered
//     3Windows 2000 Clustered
//     4NetWare-IBMSANEnable by default
//     5Linux       Enable by default
//     6AIX
//     7HP-UX       Enable by default
//     8Solaris (Sparc)
//     9PTX           Enable by default
//    10Irix
//    11Netware FailoverEnable by default
//    12IBM TS SAN VCEEnable by default
//    13LNXCL
//
// Replace the value z in the following script commands with the appropriate host
// index/number.
//
// For Controller A
```

```
show " ";
show " Displayed old setting and then Disable AVT in controller A";
show controller [a] HostNVSRAMBYTE [5,0x24];
set controller [a] HostNVSRAMBYTE [5,0x24]=0x00;
// Now for Controller B
show " ";
show " Displayed old setting and then Disable AVT in controller B";
show controller [b] HostNVSRAMBYTE [5,0x24];
set controller [b] HostNVSRAMBYTE [5,0x24]=0x00;
// Verify the settings
show " ";
show " Displayed new setting for controllers A and B";
show controller [a] HostNVSRAMBYTE [5,0x24];
show controller [b] HostNVSRAMBYTE [5,0x24];
//
// For fw 5.3x.xx.xx and later ...,  you can reset the
// controller using these two script commands. Otherwise, you
// must manually reset the controller.
//
show " ";
show "reset controllers A and B";
reset Controller [a];
reset Controller [b];
```

To disable the AVT, follow these steps:

1. To verify if you are running the Linux Host type or not, in Storage Manager, right-click the host that was previously create in 2.8.4, “Creating a host on DS5000 series” on page 58 and click **Properties**. You will see the Host type as shown Figure 2-75.

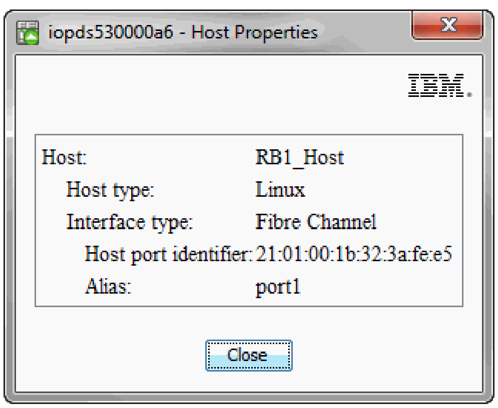


Figure 2-75 Verification of the Host type Linux

2. To execute a script on the storage, you go to the main storage device panel. When you are there, right-click the DS5000 storage subsystem that you will be using. You need to click **Execute Script**, as shown in Figure 2-76.

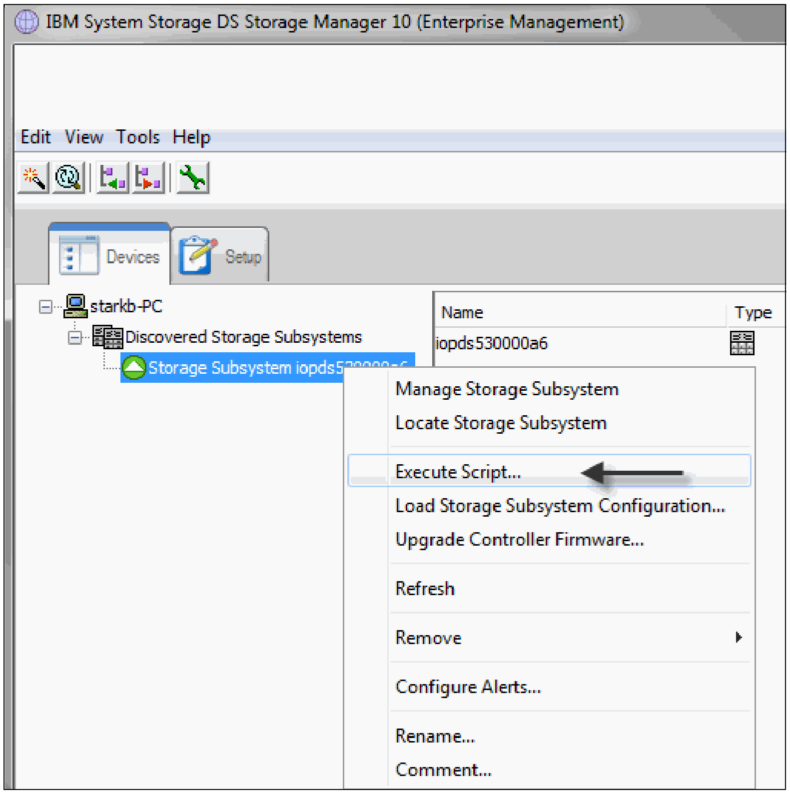


Figure 2-76 Main Storage manager panel and showing Execute Script

3. You will see a new window appear called Script Editor. To load the script, you need to click **File** → **Load Script**, as shown in Figure 2-77.

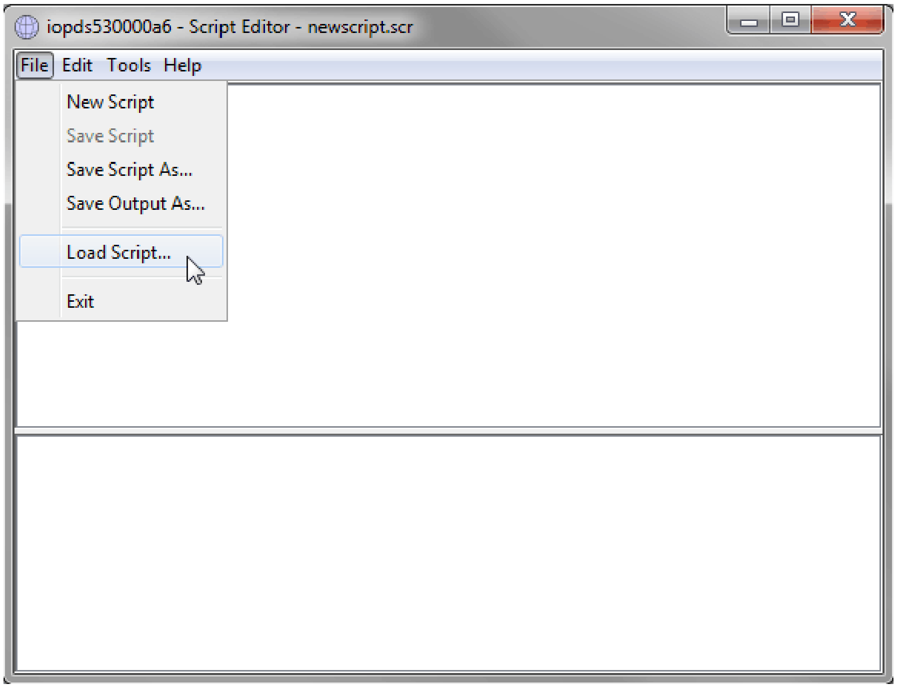


Figure 2-77 Selection of Load Script

4. After the script is loaded, to execute the script, click **Tools** → **Execute Only**, as shown in Figure 2-78.

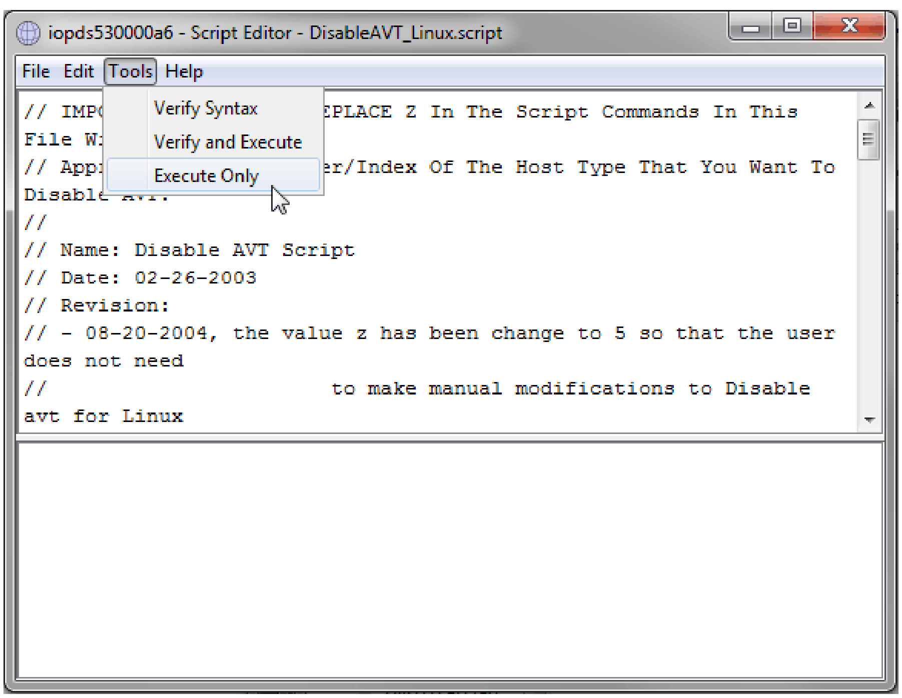


Figure 2-78 Selection of script execution

5. When the script completes, the results are displayed as shown in Figure 2-79.

```
 Displayed new setting for controllers A and B
Controller "a" Host Type Index 5 NVSRAM offset 0x24 = 0x0.
Controller "b" Host Type Index 5 NVSRAM offset 0x24 = 0x0.

reset controllers A and B
Script execution complete.
```

Figure 2-79 Successfully installed the script

See the *IBM System Storage DS Storage Manager version 10 Installation and Support Guide* for the appropriate operating system. or the Enterprise Window Online help for more information about how to execute a script.

http://www.ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5075652

2.9 FC SAN Boot implementation for Windows 2008 server

In this section, we cover the correct way to install Microsoft Windows 2008 R2 SP1 operating systems using SAN Boot while using a single path during the installation. We then continue explaining how to verify that your second path is up and working.

Microsoft support for booting from a SAN: See the following article for information about Microsoft support for booting from a SAN:

http://support.microsoft.com/kb/305547

The following steps illustrate the procedure for booting an IBM System x host system from an IBM System Storage DS5000 series storage subsystem, when installing Windows 2008 R2 SP1 server.

The iSCSI SAN Boot implementation using DS5000 is covered in Chapter 6, “iSCSI SAN Boot implementation with IBM System Storage DS5000” on page 395.

2.9.1 Limitations of booting from Windows 2008

When the server is configured with storage access and path redundancy, observe the following limitations:

- If there is a path failure and the host is generating I/O, the boot drive moves to the other path. However, while this transition is occurring, the system appears to halt for up to 30 seconds.
- If the boot device (LUN 0) is not on the same path as the bootable HBA port, you receive an `INACCESSIBLE_BOOT_DEVICE` error message.
- By booting from the DS5000 storage subsystem, most of the online diagnostic strategies are effectively canceled, and path problem determination must be done from the diagnostics panel, which can be accessed by pressing Ctrl+Q.

Important: The IDE disk devices must not be re-enabled.

2.9.2 Installation of the HBA driver

Follow these steps:

1. During the OS installation, if Windows 2008 R2 SP1 Standard does not recognize the HBA, it will not present the drive. Click **Load Driver** from the Windows installation panel, as shown in Figure 2-80, to load the appropriate HBA drivers. You need to check if your adapter is in the Windows 2008 R2 SP1 default list, or go to the HBA support website for the correct drivers.

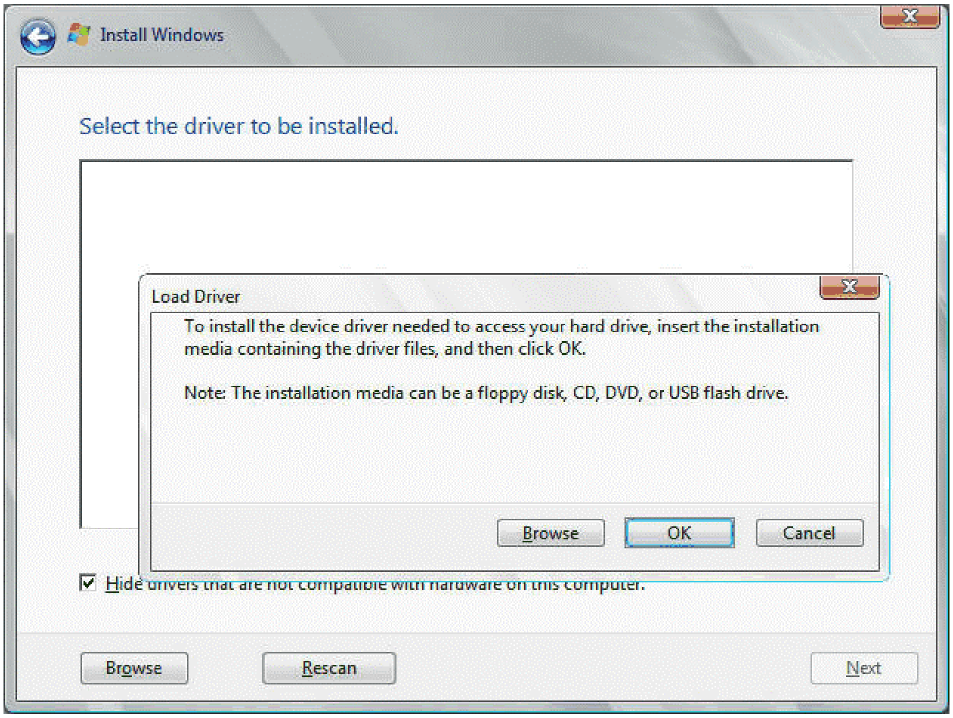


Figure 2-80 Missing HBA Driver

2. If the driver is found, it will be loaded as shown in Figure 2-81.

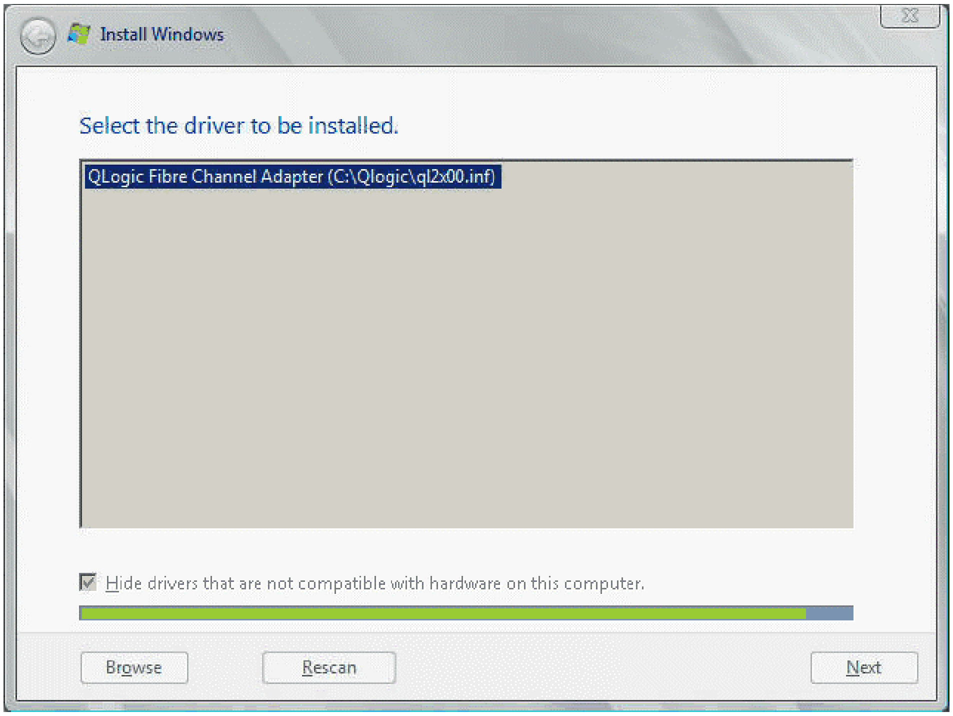


Figure 2-81 Loading HBA driver Windows 2008 R2 SP1

3. After the driver is loaded, you can see that the host found the logical drive we had previously created in 2.8.3, “Creating a logical drive on the DS5000 series” on page 55 with the same size. Notice that there are two other drives that are both 20.0 MB, which are shown as offline.

   These two drives are access LUNs on the DS5000 storage subsystem, and cannot be edited in any way. You need to select Disk 2 and click **New** to create a partition on the drive, as shown in Figure 2-82.

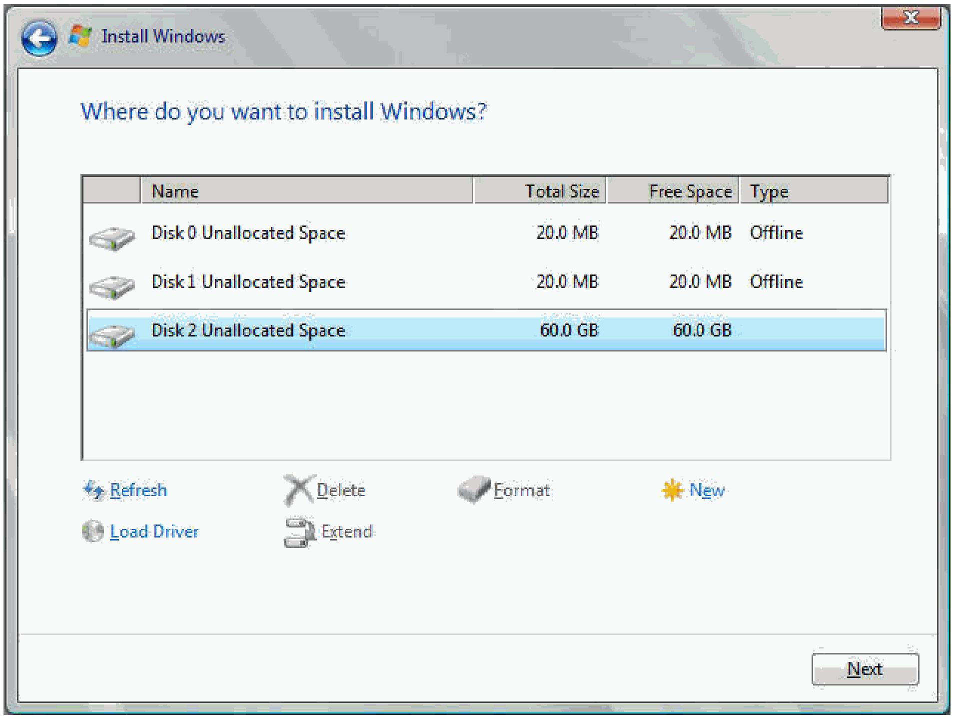


Figure 2-82 Discovering the disk

4. The new partition is created as shown in Figure 2-83. After creating this partition, follow the default OS installation procedure and complete the installation.

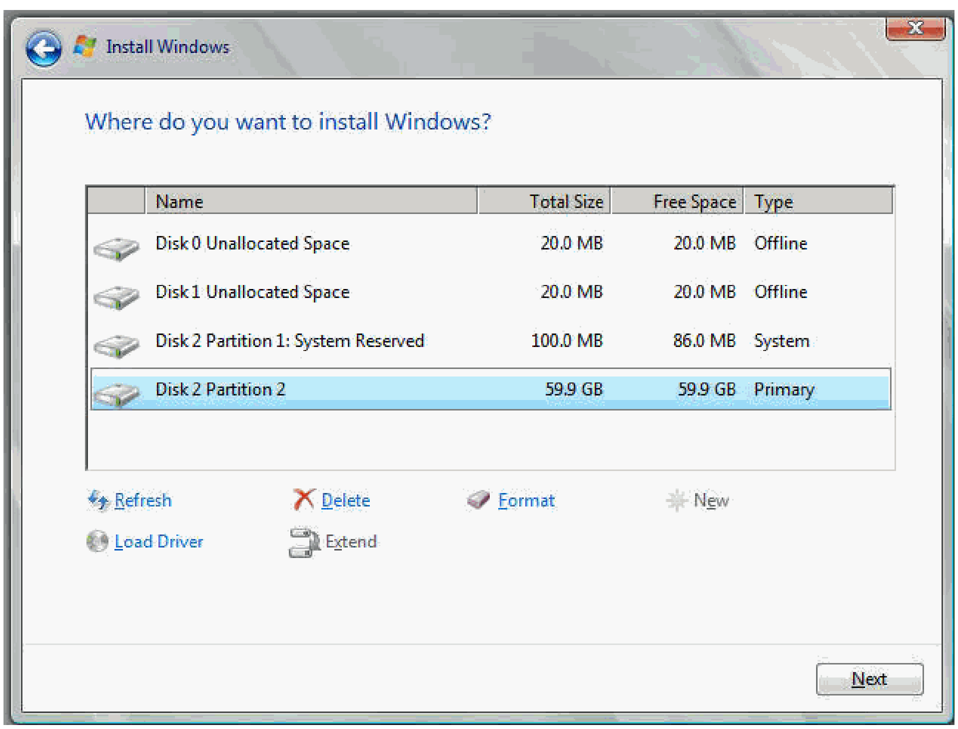


Figure 2-83 Creating the new partition

Tip: The rest of the installation procedure for the Windows OS is same as the stand-alone installation, and the panels for the rest of the installation are not shown here.

2.9.3 Updating the HBA driver and installing the multipath driver

In this section, we describe the correct way to install the multipath driver, as well as updating the HBA driver:

1. Open the *Device Manager* in Windows from the **Start Menu** → **Control Panel** → **System**. You see a window similar to Figure 2-84. You must see an *IBM 1818 FAStT Multi-Path Disk Device*, which is your boot LUN in this case. You will see two of them, due to the fact that one of them is seen from controller A and the other is from controller B. In our case, because the LUN was assigned to controller A, only this one is read/write. The other one is read-only. When we install the multipath driver, it will mask the two paths into one.

   The *Storage controllers* section lists the Fibre Channel and other controllers. If the driver is not installed correctly, you see an attention mark on the icon.

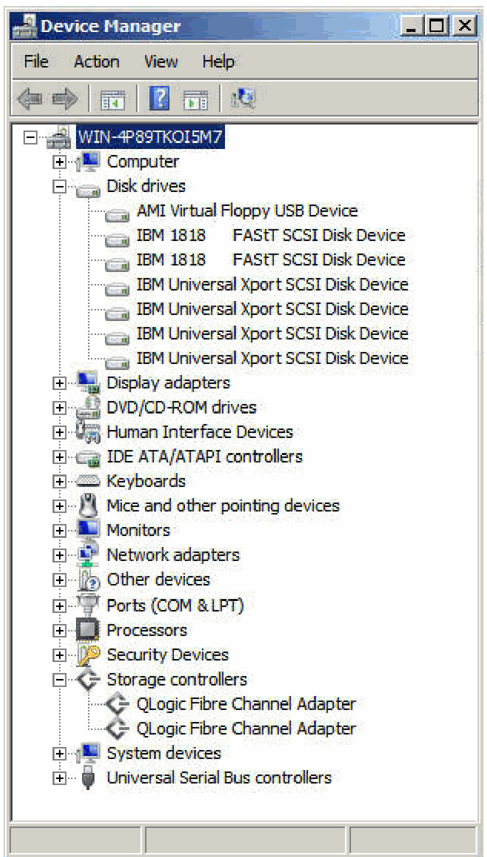


Figure 2-84 Device Manager with no multipathing

2. To install the multipath driver, you need to install IBM System Storage DS Storage Manager 10 on the host system. See 2.3.2, “Installation of the Storage Manager” on page 24 on how to install Storage Manager for the complete installation procedure. Here we list only the panels that are appropriate for the multipathing installation.

   During the Storage Manager installation, notice that a button called **Host** is shown in Figure 2-85. This button is only displayed if an HBA is detected on the system on which you are installing. You need to click **Host**, as shown in Figure 2-85.

Figure 2-85 IBM System Storage DS Storage Manager 10 - Select Installation Type

3. By selecting the Host option, the installation will include the multipath driver. When the installation is complete, reboot the host as shown in Figure 2-86. It is required to reboot the host to load the drivers needed for the installation.

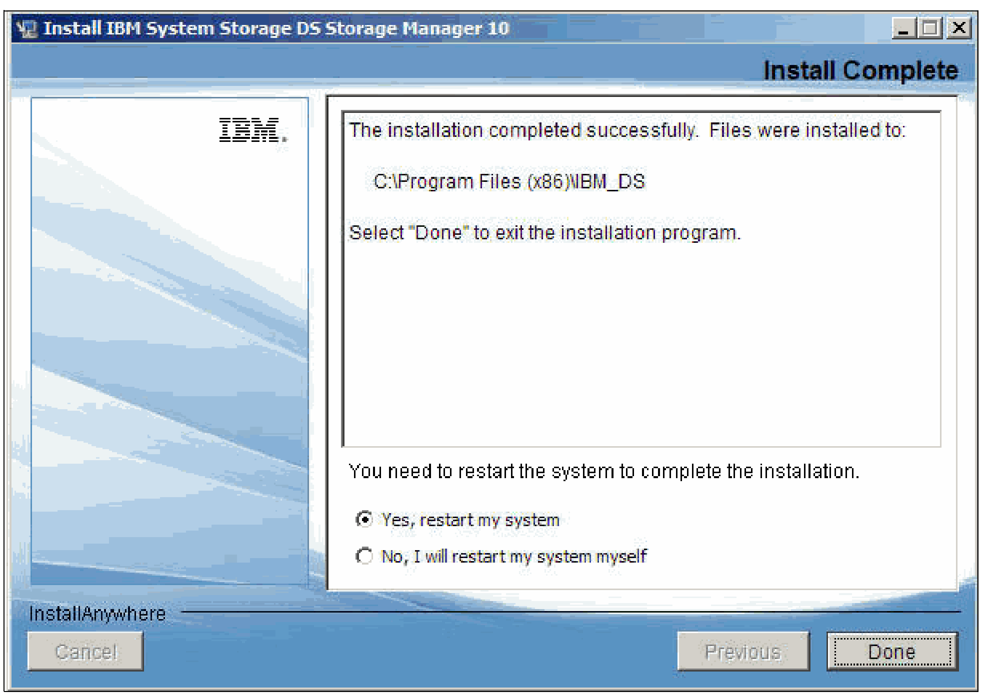


Figure 2-86 IBM System Storage DS Storage Manager 10 - Install Complete

Tip: At this point your paths and OS are up and running, and you can verify the paths using the procedure documented in “Windows 2008 R2 SP1 Multipath Verification”.

2.9.4 Windows 2008 R2 SP1 Multipath Verification

To verify the installation of the multipath driver:

1. After you have rebooted the Windows system, go to the Device Manager to check the devices. As you can see in Figure 2-87, it has added three new devices. Under the Storage controllers, it added the Microsoft Multi-Path Bus Driver, and under System devices, it added the **DS5000 Multi-Path** and **Disk Manager**. If you do not install IBM System Storage DS Storage Manager 10 on the host and just use the onboard multi-path, it will not function correctly.

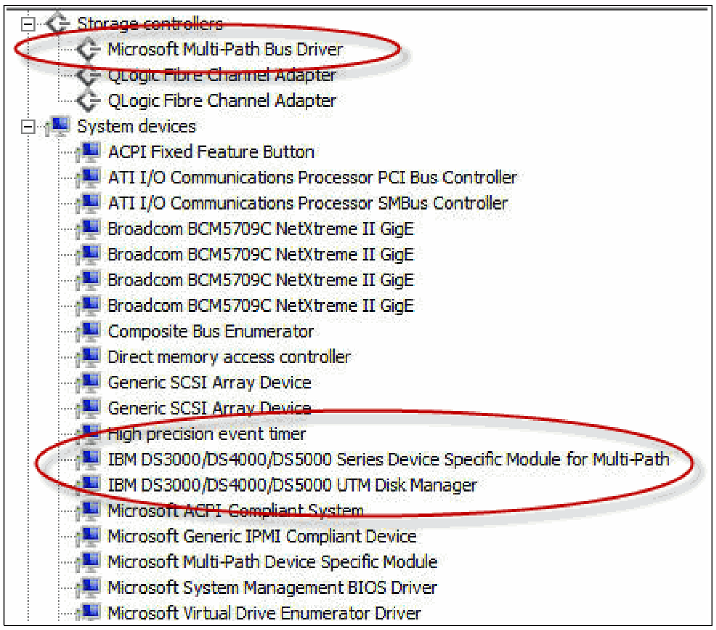


Figure 2-87 New multi-path shown in device manager

2. In Figure 2-88, you can see four IBM Universal Xport SCSI Disk Devices and one **IBM 1818 FAStT Multi-Path Disk Device.** The Xport disk is the two controllers shown as disks. The Multi-Path Disk is the LUN 0, and is our boot LUN. If we had created other disks from the storage, you will see more Multi-Path Disks with all having the same name here. It means that if you had one LUN 0 for the boot LUN, and four other disks LUN 1 to 4, then you should see five **IBM 1818 FAStT Multi-Path Disk Devices**.

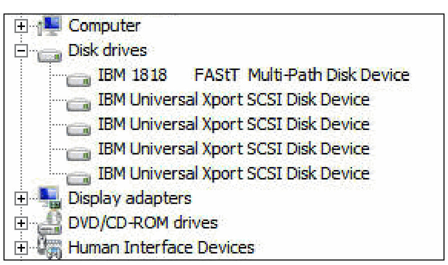


Figure 2-88 Showing the new disk devices

3. To get more information about the LUN, or to check or change multipathing, you can right-click the device and go to **Properties** as shown in Figure 2-89.

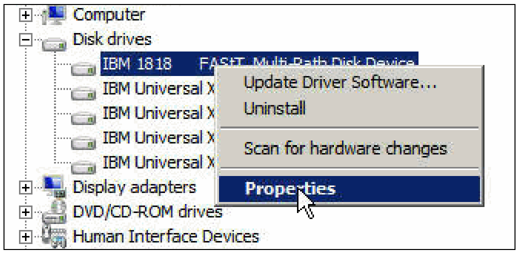


Figure 2-89 Properties of the Multi-Path Disk Device

4. On the **General** tab, you can see that it is presented as LUN 0 and is connected by ports 6,6,7,7 on Bus 0. It tells us that it is seeing the LUN on both paths, as shown in Figure 2-90. The 6,6 is both paths to controller A, and the 7,7 is both paths to controller B.

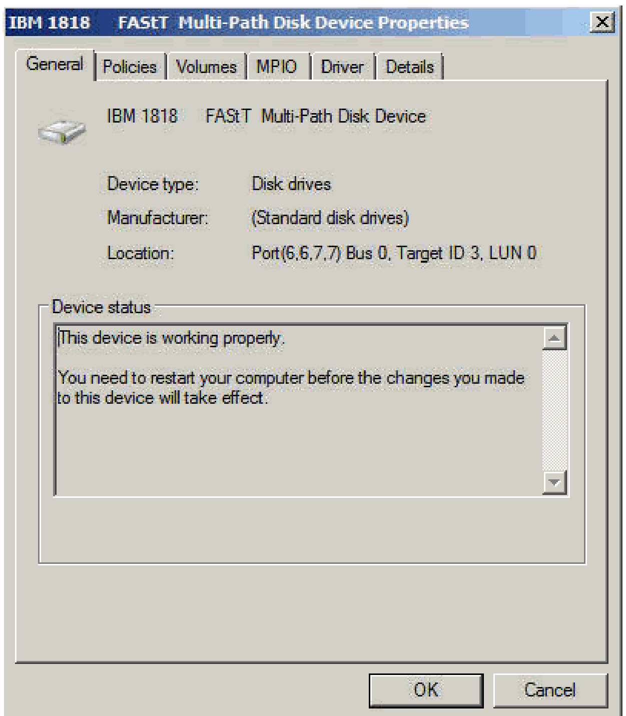


Figure 2-90 General Tab of the Disk Device

5. We can also check the volume information, by clicking the **Volumes** tab. To populate the information, you need to click **Populate** at the bottom of the window, as shown in Figure 2-91.

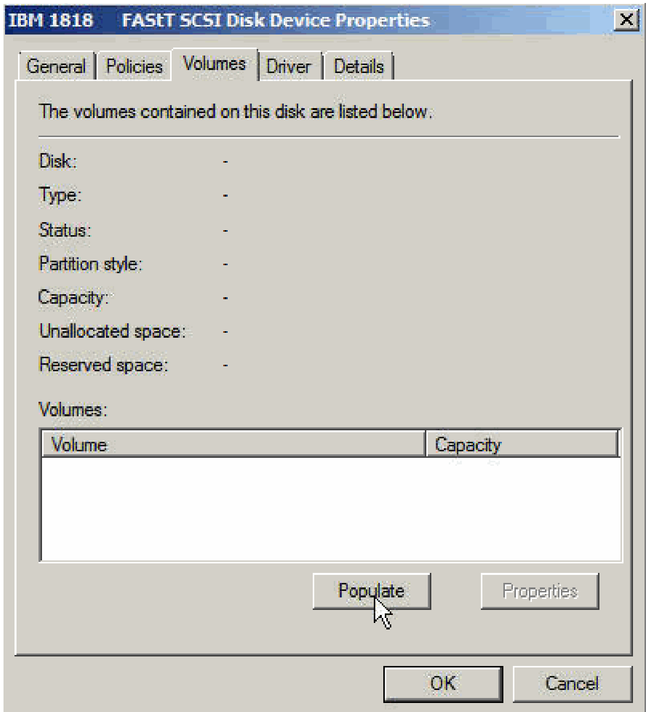


Figure 2-91 Volumes Tab

6. Here we see the LUN 0, which is our boot LUN, as shown in Figure 2-92.

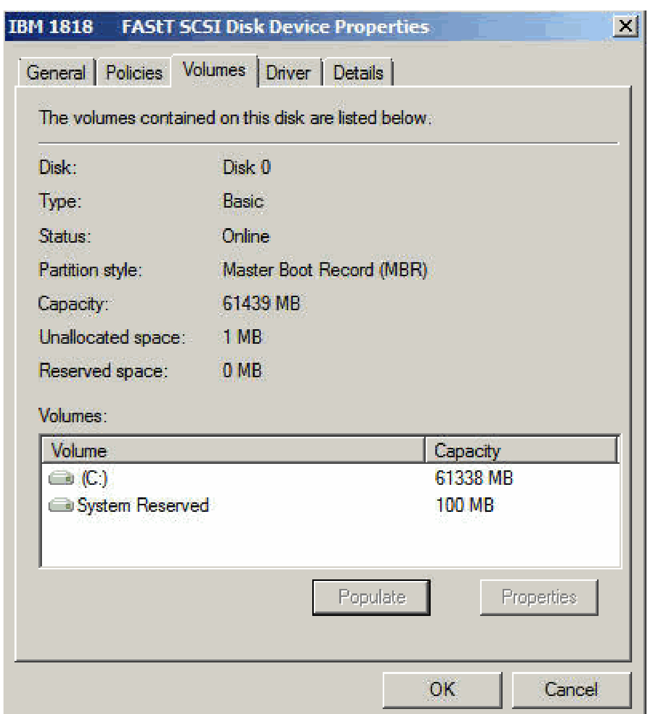


Figure 2-92 Volumes tab populated

7. On the MPIO tab, we have Least Queue Depth, which is the default setting, and we are not making a change to anything on this tab. Also on this tab are the paths, and we notice that there are two with a Path Id of 77060004. It is the path to controller B and is in standby status. Two with Path Id of 77060003 are in Active/Optimized state and are to controller A, again telling us that the multipathing is working, as shown in Figure 2-93.

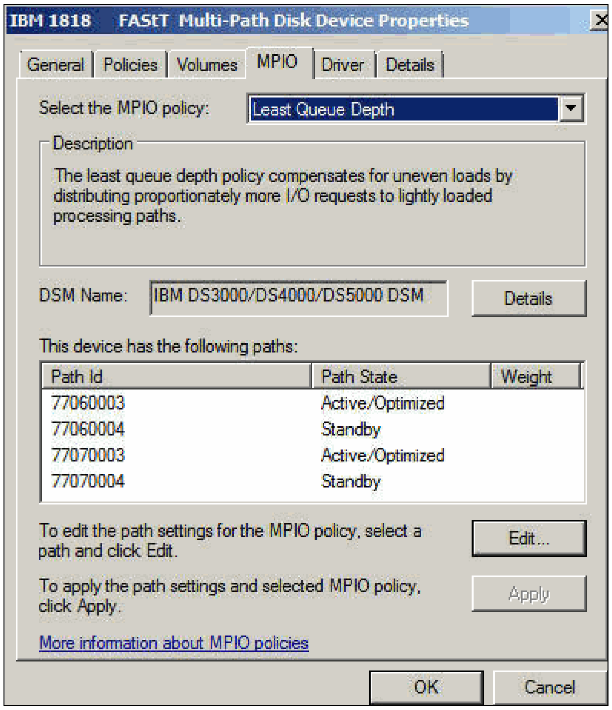


Figure 2-93 Verification of the MPIO tab

The installation and verification of the multipath on Windows 2008 R2 SP1 is now complete.

2.10 FC SAN Boot implementation for SUSE Linux Enterprise Server 11 SP1

The following steps illustrate the procedure for booting an IBM System x system from an IBM System Storage DS5000 series storage subsystem, with SUSE Linux Enterprise Server 11 SP1 installed. To install SUSE Linux Enterprise Server 11 SP1, you do not need to install it using a single path only, as SUSE Linux Enterprise Server 11 takes care of multipathing during the installation process. It means that when you follow the HBA instructions, you can configure both paths at the same time.

2.10.1 Installation steps

We are going to install SUSE Linux Enterprise Server 11 on a 60-GB SAN Boot LUN 0 that we created using the procedure described in 2.8, “Storage provisioning” on page 50. To configure the HBAs, follow the steps outlined in 2.6, “QLogic HBA configuration” on page 33 or 2.7, “Emulex HBA configuration” on page 42. We are using the QLogic HBA for this demonstration.

1. The first part of the installation is done with default settings, and we have not included the panels. When you get to the **Installation Settings** section, you will need to make changes. On the **Expert** tab, click the **Partitioning** hyperlink in the middle of the window, as shown in Figure 2-94.

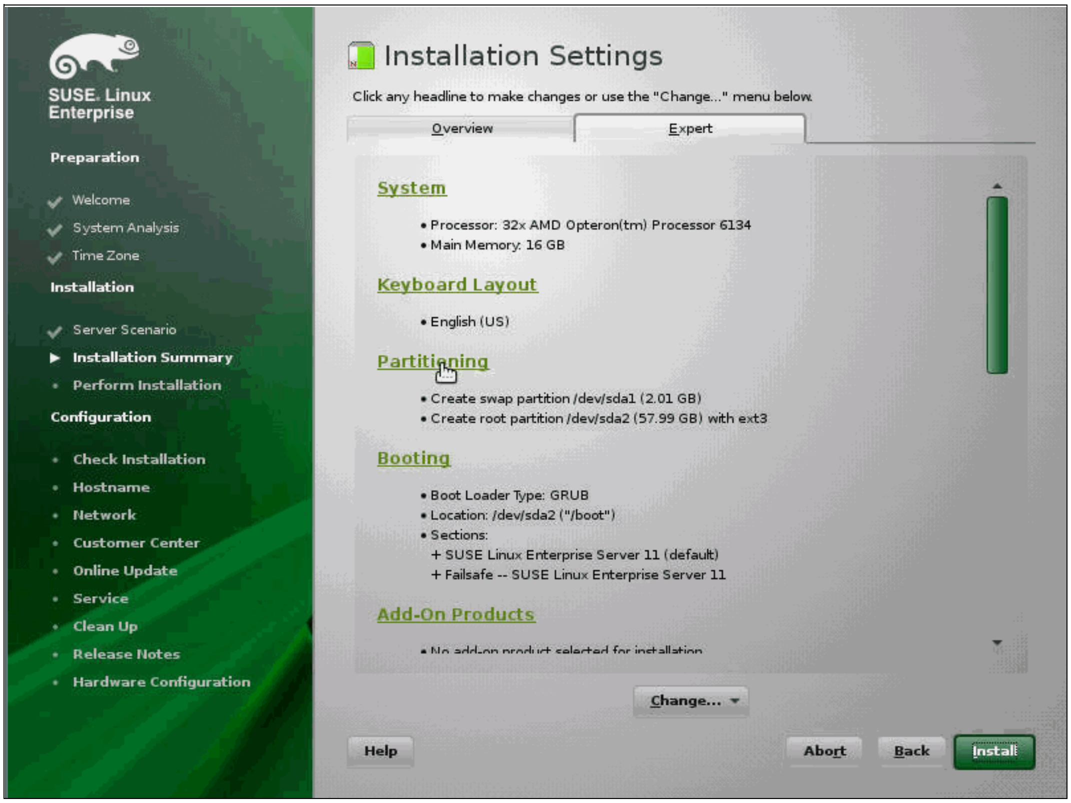


Figure 2-94 SUSE Linux Enterprise Server Expert tab

2. On the next panel, we select **Custom Partitioning** and click **Next**, as shown in Figure 2-95.

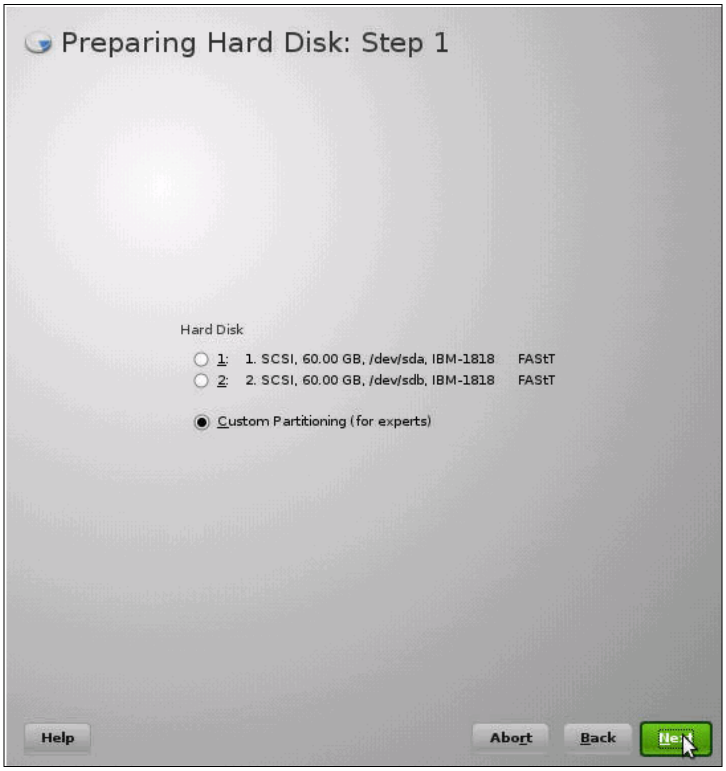


Figure 2-95 Selection of Customer Partitioning

3. In the Expert Partitioner panel, we select the **Hard Disks** Icon as shown in Figure 2-96.

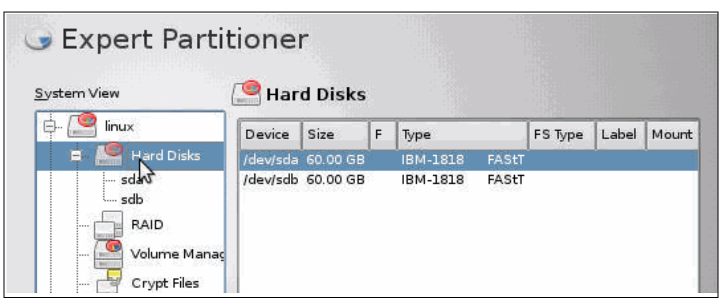


Figure 2-96 Selection of Expert Partitioner

4. After clicking **Hard Disks**, we click the **Configure** button at the bottom right corner of the panel, and select **Configure Multipath** as shown in Figure 2-97.

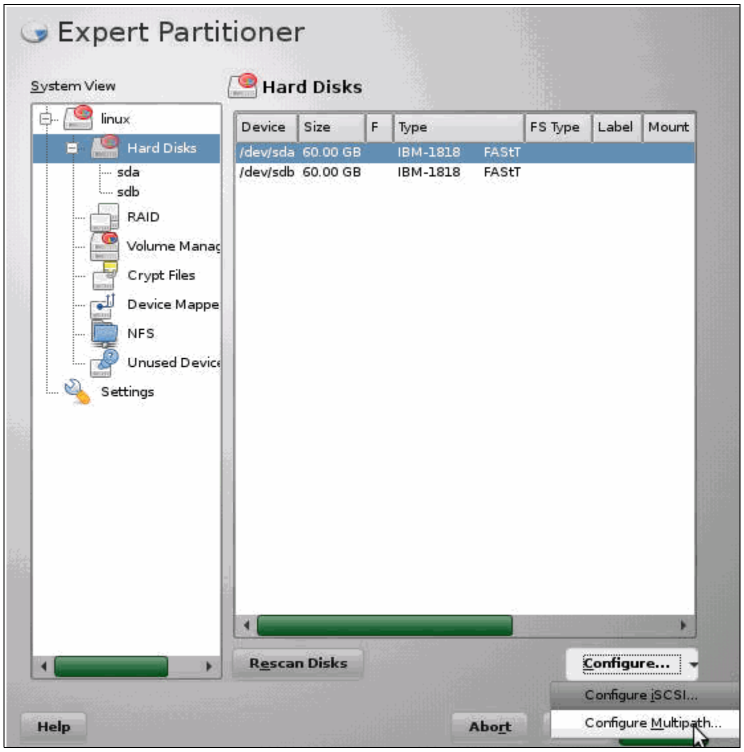


Figure 2-97 Selection of Multipath

5. On the next panel, we confirm the activation of the multipath by clicking **Yes**, as shown in Figure 2-98.

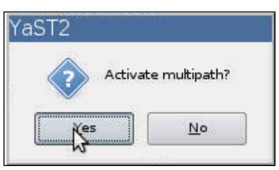


Figure 2-98 Activate multipath

6. You will see the drive that was created as shown in Figure 2-99. Now we need to add a new partition to this new drive before we continue with the OS installation.

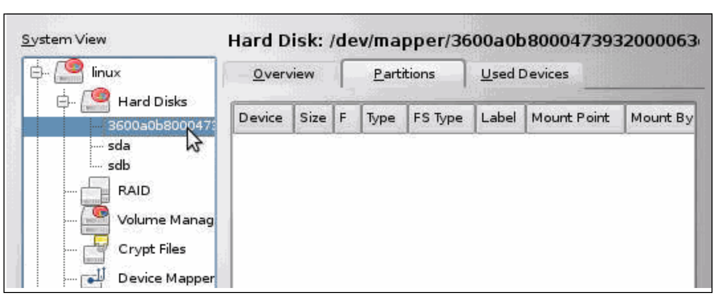


Figure 2-99 Selection of the new Mapper drive

7. Click **Add** at the bottom of the panel, as shown in Figure 2-100. It will add a new partition to the new disk, such as swap and root. The disk can be partitioned for your requirements. In this case, we created a 2.01-GB swap, and the rest of the space is allocated for the root.

Figure 2-100 Clicking Add to create the Partitions

8. You can see the two new partitions that we created, as shown in Figure 2-101. Verify the drive size and click **Accept** to accept the changes.

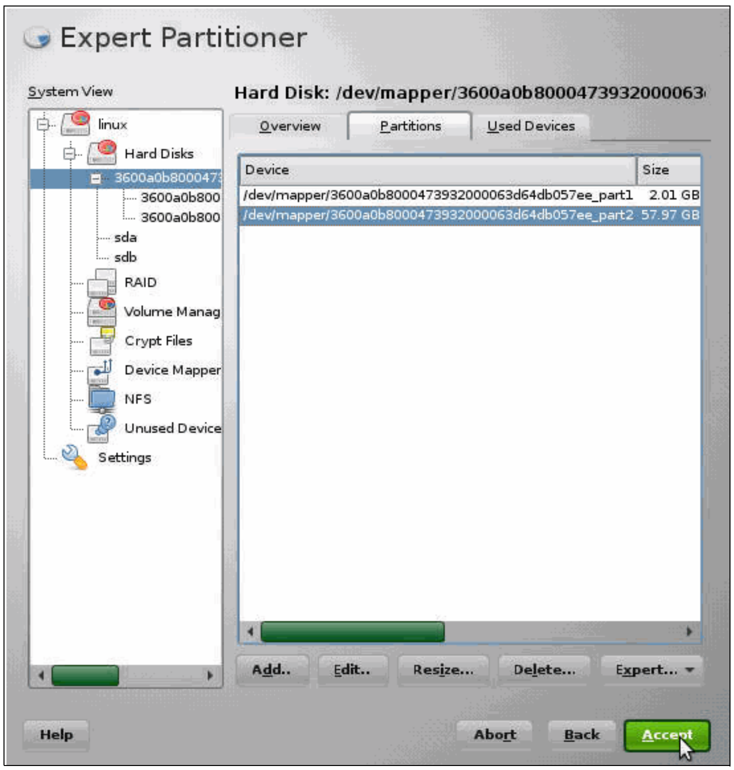


Figure 2-101 New Partitions shown

9. We now confirm the new location of the bootloader. Click **Yes** to confirm the new location, as shown in Figure 2-102.

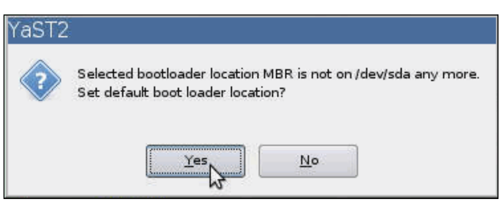


Figure 2-102 Confirming the new Bootloader location

10. We now need to update the SUSE Linux Enterprise Server Grub file with the Multipathing information. Click **Booting** to update the Grub file as shown in Figure 2-103. Here we see that the location is still showing /dev/sda (MBR).

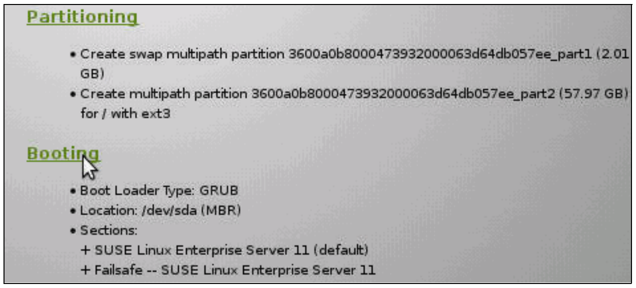


Figure 2-103 Selecting Booting to change the Grub

11. Click the **Boot Loader Installation** tab, as shown in Figure 2-104, select the SUSE Linux Enterprise Server 11, and click **Edit**.

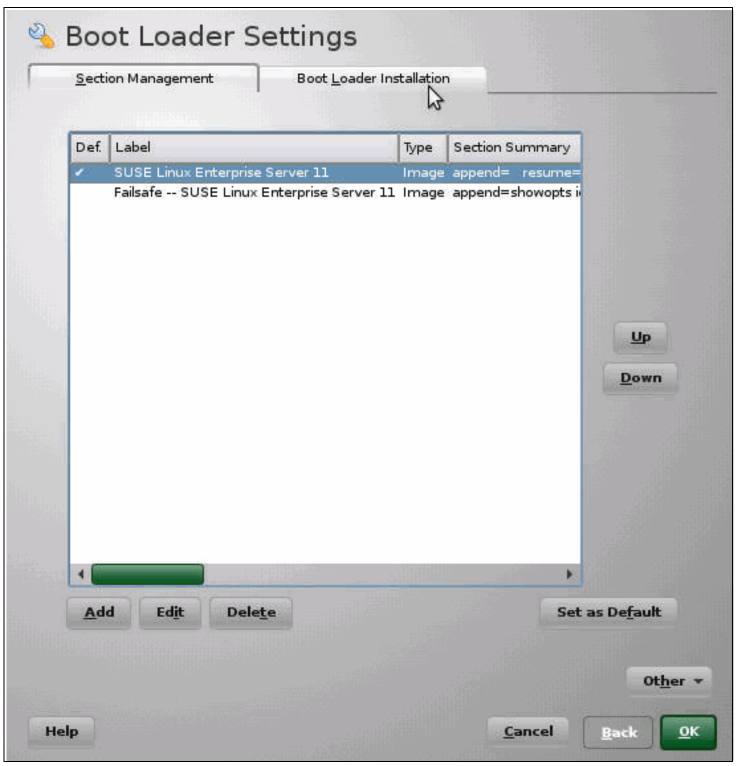


Figure 2-104 Selecting the Boot Loader Installation tab

12. We now click the **Boot Loader Installation Details**, as shown in Figure 2-105.

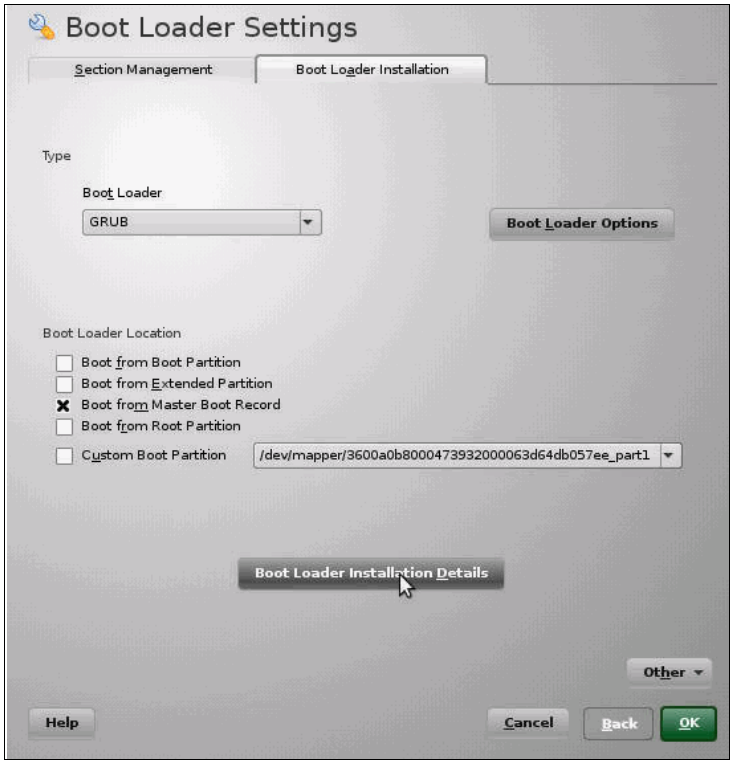


Figure 2-105 Selecting of the Boot Loader Installation details

13. When we are in the Boot Loader Device Map panel, we are not going to make any changes here. We need to go through each panel to build the Grub at the end by clicking **OK**, as shown in Figure 2-106.

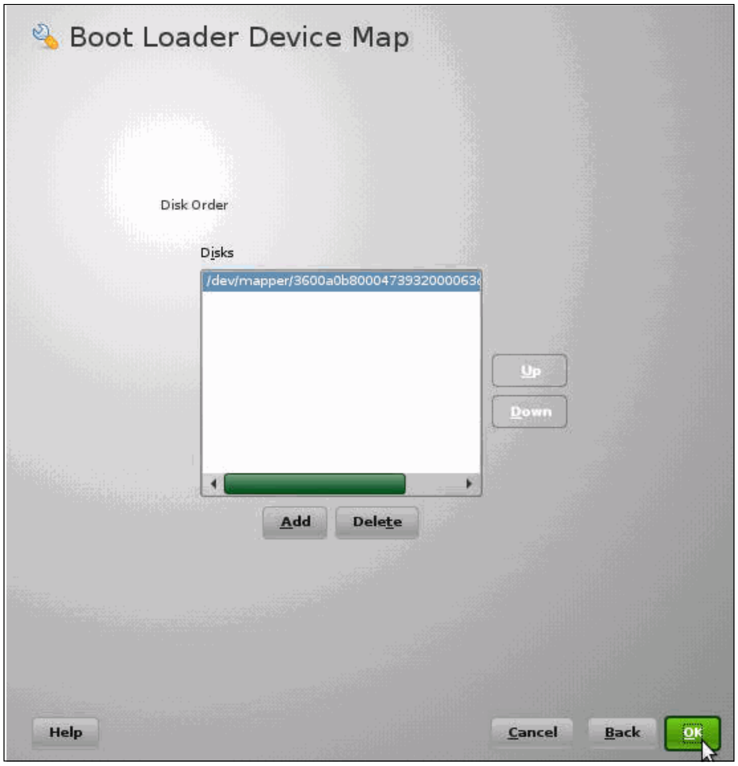


Figure 2-106 Boot loader Installation details

14. Now we see that the location was updated as shown in Figure 2-107. After this setting, the rest of the installation panels are left as default. The rest of the SUSE installation panels are not shown here.

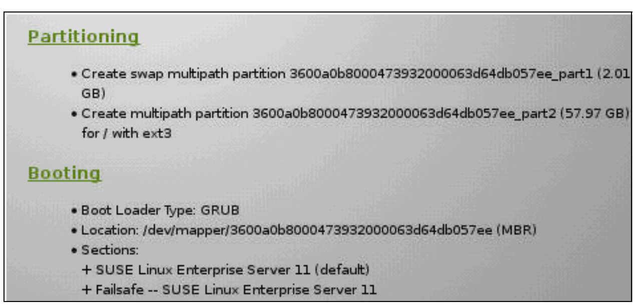


Figure 2-107 Verifying the new Location to the grub

2.10.2 SUSE Linux Enterprise Server 11 SP1 multipath verification of installation

To verify if the multipath is working, we need to open a terminal and run the **multipath -l** command**.** The results should look like the listing in Figure 2-108. We can also run the **df** command to return the disk size.

```
linux-gxhz:~/Desktop # multipath -l
3600a0b800047393200006644db7fe2e dm-0 IBM,1818      FAStT
[size=70G][features=1 queue_if_no_path][hwhandler=1 rdac][rw]
\_ round-robin 0 [prio=-2][active]
 \_ 6:0:3:0 sda 8:0   [active][undef]
 \_ 7:0:3:0 sdc 8:32  [active][undef]
\_ round-robin 0 [prio=-2][enabled]
 \_ 6:0:4:0 sdb 8:16  [active][undef]
 \_ 7:0:4:0 sdd 8:48  [active][undef]
linux-gxhz:~/Desktop #
```

Figure 2-108 Verification of multipath on SUSE Linux Enterprise Server

2.11 FC SAN Boot implementation for Red Hat Enterprise Linux 5.5 server

The following steps illustrate the procedure for booting an IBM System x host system from an IBM System Storage DS5000 series storage subsystem, with Red Hat Enterprise Linux (RHEL) 5.5 installed.

To install RHEL 5.5, we need to follow the single path installation method, which is the same as described for 2.9, “FC SAN Boot implementation for Windows 2008 server” on page 72. RHEL does not install multipath during the boot process. While it is possible to use the Linux multipathing, for our example, we use the LSI RDAC driver. Linux multipathing is not covered in this publication.

We also cover how to verify that your second path is active and working.

Tip: To install RHEL 5.5, you need to follow the single path installation method, which is the same as for Windows. RHEL does not install multipath during the boot process. We are not using the Linux multipathing; rather, we are using the LSI RDAC driver. You can use the Linux multipathing instead of the RDAC, but we are not covering it in this publication.

In the DS5000 storage subsystem, Auto Logical Drive Transfer (ADT/AVT) mode is automatically enabled in the Linux storage partitioning host type. This mode causes contention when an RDAC driver is installed. If using the “Linux” host type, it must be disabled using the script that is bundled in this Linux RDAC web package or in the \Scripts directory of this DS Storage Manager version 10 support for Linux CD. You need to follow the instructions in 2.8.7, “Disabling AVT for Linux Host type” on page 68 to disable AVT.

We are going to install RHEL 5.5 server on a 40 GB SAN Boot LUN 0 that we created using the procedure described in 2.8, “Storage provisioning” on page 50. To configure the HBAs, follow the steps outlined in 2.6, “QLogic HBA configuration” on page 33 and 2.7, “Emulex HBA configuration” on page 42. We are using the QLogic HBA for this demonstration.

Tip: We only show the panels of the RHEL installation that are different from a normal installation on the local hard drive.

Follow these steps:

1. We start with the RHEL 5.5 installation using the Graphical User Interface (GUI).
2. During the process, the installation will automatically detect the SAN Boot disk and we are presented with a pop-up for initializing the drive. The key point here is to check that it is drive sda, as that will be our bootable drive, as shown in Figure 2-109. Click **Yes** to initialize the drive. If you have other disks mounted, you might see other disks such as sdb or more.

Figure 2-109 Warning RHEL will initialize the sda

3. In the panel that asks you to change the sda partition, we leave the default selection and click **Next**. If you need to customize the partition, you can do so by clicking the **Advanced storage configuration** option as shown in Figure 2-110.

Figure 2-110 Not changing the partitions

4. We now need to confirm the removal of all Linux partitions (and ALL DATA on them). Here we need to make sure that it is the correct drive, and after it is verified, we click **Yes**, as shown in Figure 2-111.

Figure 2-111 RHEL confirm partition removal

5. On the next additional tasks panel, we select **Software Development.** You also need to click **Customize Now**, as shown in Figure 2-112. You only need to do this to one of the systems, as you need the development tools to compile the RDAC.

Tip: You can skip this step under the following circumstances:

- If you will not be using RDAC
- If you already have RDAC compiled of the kernel that you will be running

Figure 2-112 Customizing the Software Tools

6. The only other tool we need to add is the Legacy Software Development (Figure 2-113).

Figure 2-113 Adding the Legacy Software Development

7. After the Software Tools selection, the rest of the panels are configured according to your installation requirements. For this demonstration, we choose all default options.

2.11.1 Installation of the RDAC in RHEL 5.5

After the installation of RHEL, we need to install RDAC, which is the LSI multipath driver. Linux multipathing can also be used and is covered in 2.11.2, “Linux MPP driver” on page 95.

1. When the installation is complete and the machine is up and running, download the latest RDAC source from LSI Linux RDAC Driver Packages site at `http://www.lsi.com/rdac/`.

   For this example, version 09.03.0C05.0439 of the RDAC driver was used. as is shown in Example 2-2.

   Example 2-2 wget the rdac from LSI

   ```
   wget http://www.lsi.com/rdac/rdac-LINUX-09.03.0C05.0439-source.tar.gz
   ```

2. Extract the source by running the following command as shown in Example 2-3.

   Example 2-3 untar the RDAC Driver

   ```
   tar -zxf rdac-LINUX-09.03.0C05.0439-source.tar.gz
   ```

3. Navigate to the RDAC source directory as in Example 2-4.

   Example 2-4 CD to the untared location

   ```
   cd linuxrdac-09.03.0C05.0439
   ```

4. To install the RDAC source, run the following commands as shown in Example 2-5. See the Readme.txt file for details.

   Example 2-5 Compile the driver with the correct kernel

   ```
   make clean
   make
   make install
   ```

5. After running **`make install`**, we see a message similar to that in Example 2-6.

   Example 2-6 Message when RDAC is finished compiling

   ```
   You must now edit your boot loader configuration file, /boot/grub/menu.lst, to
   add a new boot menu, which uses mpp-2.6.18-194.el5.img as the initrd image.
   Now Reboot the system for MPP to take effect.
   The new boot menu entry should look something like this (note that it may
   vary with different system configuration):

   ...
         title Red Hat Linux (2.6.18-194.el5) with MPP support
         root (hd0,5)
         kernel /vmlinuz-2.6.18-194.el5 ro root=LABEL=RH9
         initrd /mpp-2.6.18-194.el5.img
   ...
   MPP driver package has been successfully installed on your system.
   ```

> **Tip:** Note the name of the new mpp initrd image that is created in your /boot directory. In this example, the name is mpp-2.6.18-194.el5.img.

6. Open the **/boot/grub/menu.lst** file with your preferred text editor.
7. Create a new boot entry by copying the first boot entry. Then edit the initrd line to point to the new initrd file so that the mpp initrd is booted. The new initrd line should look similar to Example 2-7.

   Make sure that the default is set to boot to the entry you just created (default=0 means the system is default to boot to the first boot entry). You need to change the (hd0,0), and the root to be the same as what is in your menu.lst.

   Example 2-7 Changes to the menu.lst

```
default=0
timeout=5
splashimge=(hd0,0)/boot/grub/splash.xpm.gz
hiddenmenu
title Red Hat Enterprise Linux Server with mpp support (2.6.18-194.el5)
   root (hd0,0)
   kernel /boot/vmlinuz-2.6.18-194.el5 ro root=/dev/VolGroup00/LogVol00 rhgb
quiet
   initrd/boot/mpp-2.6.18-194.el5.img
title Red Hat Enterprise Linux Server (2.6.18-194.el5)
   root (hd0,0)
   kernel /boot/vmlinuz-2.6.18-194.el5 ro root=/dev/VolGroup00/LogVol00 rhgb
quiet
   initrd/boot/initrd-2.6.18-194.el5.img
```

8. Reboot the system.
9. When the system is back up, verify that the mpp driver is installed correctly by running the following command from the command line (Example 2-8).

   Example 2-8 Verify the mpp driver

```
ls -lR /proc/mpp/
```

10. You should see that all the physical LUNs have been discovered and that the virtual LUNs have been created. For this example, because the environment has three physical LUNs, running **ls -lR /proc/mpp/** produced three virtual LUNs: virtualLun0.

    ls -1R /proc/mpp also shows us that controller A is connected to one physical LUN: LUN0 (see Example 2-114).

```
/proc/mpp/iopds530000a6:
total 0
dr-xr-xr-x 4 root root 0 Apr 26 23:08 controllerA
dr-xr-xr-x 4 root root 0 Apr 26 23:08 controllerB
-rw-r--r-- 1 root root 0 Apr 26 23:08 virtualLun0

/proc/mpp/iopds530000a6/controllerA:
total 0
dr-xr-xr-x 2 root root 0 Apr 26 23:08 qla2xxx_h0c0t4
dr-xr-xr-x 2 root root 0 Apr 26 23:08 qla2xxx_h1c0t4

/proc/mpp/iopds530000a6/controllerA/qla2xxx_h0c0t4:
total 0
-rw-r--r-- 1 root root 0 Apr 26 23:08 LUN0
-rw-r--r-- 1 root root 0 Apr 26 23:08 UTM_LUN31

/proc/mpp/iopds530000a6/controllerA/qla2xxx_h1c0t4:
total 0
-rw-r--r-- 1 root root 0 Apr 26 23:08 LUN0
-rw-r--r-- 1 root root 0 Apr 26 23:08 UTM_LUN31

/proc/mpp/iopds530000a6/controllerB:
total 0
dr-xr-xr-x 2 root root 0 Apr 26 23:08 qla2xxx_h0c0t5
dr-xr-xr-x 2 root root 0 Apr 26 23:08 qla2xxx_h1c0t5

/proc/mpp/iopds530000a6/controllerB/qla2xxx_h0c0t5:
total 0
-rw-r--r-- 1 root root 0 Apr 26 23:08 LUN0
-rw-r--r-- 1 root root 0 Apr 26 23:08 UTM_LUN31
```

Figure 2-114 Output of ls -lR /proc/mpp/

2.11.2 Linux MPP driver

This section describes how to install the MPP (RDAC) driver for a Linux configuration.

> **Important:** Before you install MPP, make sure that the partitions and LUNs are configured and assigned and that the correct HBA driver is installed.

Complete the following steps to install the MPP driver:

1. Download the MPP driver package from the IBM DS4000/DS5000 System Storage Disk Support website.
2. Create a directory on the host and download the MPP driver package to that directory.
3. Uncompress the file by typing the following command:

Example 2-9 Uncompress Linux MPP Driver

```
# tar -zxvf rdac-LINUX-package_version-source.tar.gz
```

   Here, **package_version** is the SUSE Linux Enterprise Server or RHEL package version number.

   Result: A directory called **linuxrdac-version#** or **linuxrdac** is created.

4. Open the README that is included in the **linuxrdac-version#** directory.
5. In the README, find the instructions for building and installing the driver and complete all of the steps.

Tip: Make sure that you reboot the server before you proceed to the next step.

6. Type the following command to list the installed modules:

Example 2-10 List the installed modules

```
# lsmod
```

7. Verify that module entries are included in the lsmod list, as follows:

   Module entries for SUSE Linux Enterprise Server or RHEL:

   – mppVhba
   – mppUpper
   – lpfc (or qla2xxx for BladeCenter configurations)
   – lpfcdfc (if ioctl module is installed)

Tip: If you do not see the mpp_Vhba module, the likely cause is that the server was rebooted before the LUNs were assigned, so the mpp_Vhba module was not installed. If it is the case, assign the LUNs now, reboot the server, and repeat this step.

8. Type the following command to verify the driver version:

Example 2-11 Verify the driver version

```
mppUtil -V
Result: The Linux multipath driver version displays.
```

9. Type the following command to verify that devices are configured with the RDAC driver.

Example 2-12 Verify that devices are configured

```
ls -1R /proc/mpp
```

Results: Output similar to the following example in Figure 2-115.

```
# ls -1R /proc/mpp
/proc/mpp:
total 0
dr-xr-xr-x    4 root     root            0 Oct 24 02:56 DS4100-sys1
crwxrwxrwx    1 root     root     254,   0 Oct 24 02:56 mppVBusNode

/proc/mpp/ DS4100-sys1:
total 0
dr-xr-xr-x    3 root     root            0 Oct 24 02:56 controllerA
dr-xr-xr-x    3 root     root            0 Oct 24 02:56 controllerB
-rw-r--r--    1 root     root            0 Oct 24 02:56 virtualLun0
-rw-r--r--    1 root     root            0 Oct 24 02:56 virtualLun1
-rw-r--r--    1 root     root            0 Oct 24 02:56 virtualLun2
-rw-r--r--    1 root     root            0 Oct 24 02:56 virtualLun3
-rw-r--r--    1 root     root            0 Oct 24 02:56 virtualLun4
-rw-r--r--    1 root     root            0 Oct 24 02:56 virtualLun5

/proc/mpp/ DS4100-sys1/controllerA:
total 0
dr-xr-xr-x    2 root     root            0 Oct 24 02:56 lpfc_h6c0t2

/proc/mpp/ DS4100-sys1/controllerA/lpfc_h6c0t2:
total 0
-rw-r--r--    1 root     root            0 Oct 24 02:56 LUN0
-rw-r--r--    1 root     root            0 Oct 24 02:56 LUN1
-rw-r--r--    1 root     root            0 Oct 24 02:56 LUN2
-rw-r--r--    1 root     root            0 Oct 24 02:56 LUN3
-rw-r--r--    1 root     root            0 Oct 24 02:56 LUN4
-rw-r--r--    1 root     root            0 Oct 24 02:56 LUN5

/proc/mpp/ DS4100-sys1/controllerB:
total 0
dr-xr-xr-x    2 root     root            0 Oct 24 02:56 lpfc_h5c0t0

/proc/mpp/ DS4100-sys1/controllerB/lpfc_h5c0t0:
total 0
-rw-r--r--    1 root     root            0 Oct 24 02:56 LUN0
-rw-r--r--    1 root     root            0 Oct 24 02:56 LUN1
-rw-r--r--    1 root     root            0 Oct 24 02:56 LUN2
-rw-r--r--    1 root     root            0 Oct 24 02:56 LUN3
-rw-r--r--    1 root     root            0 Oct 24 02:56 LUN4
-rw-r--r--    1 root     root            0 Oct 24 02:56 LUN5
```

Figure 2-115 Verify that devices

Tip: After you install the RDAC driver, these commands and man pages are available:

- mppUtil
- mppBusRescan
- mppUpdate
- RDAC

At this point, if you see the correct paths, the configuration is complete.

2.12 FC SAN Boot implementation for VMware ESXi 4

The following steps illustrate the procedure for booting an IBM System x system from an IBM System Storage DS5000 series storage subsystem, with VMware ESXi 4 installed.

Tip: To install ESXi 4 update 1, we do not need to install using single path, as ESXi takes care of multipathing during the installation process. It means that when you follow the HBA instructions, you configure both paths at the same time.

We are going to install ESXi 4 on a 60 GB SAN Boot LUN 0 that we created using the procedure described in 2.8, "Storage provisioning" on page 50. Both QLogic and Emulex HBAs are supported and for this installation, but we are using a QLogic HBA. The HBAs can be configured using procedure outlined in 2.6, "QLogic HBA configuration" on page 33 and 2.7, "Emulex HBA configuration" on page 42.

For more information about SAN Boot on ESX, see *VMware Implementation with IBM System Storage DS4000/DS5000*, REDP-4609.

2.12.1 Installation steps

After the HBA configuration, proceed with the ESXi 4 update1 installation as normal and no changes will be needed. We go through the setup to show the installation.

1. At the Boot Menu of ESXi 4, you need to click **ESXi Installer**, because we want to install it to a drive, as shown in Figure 2-116.

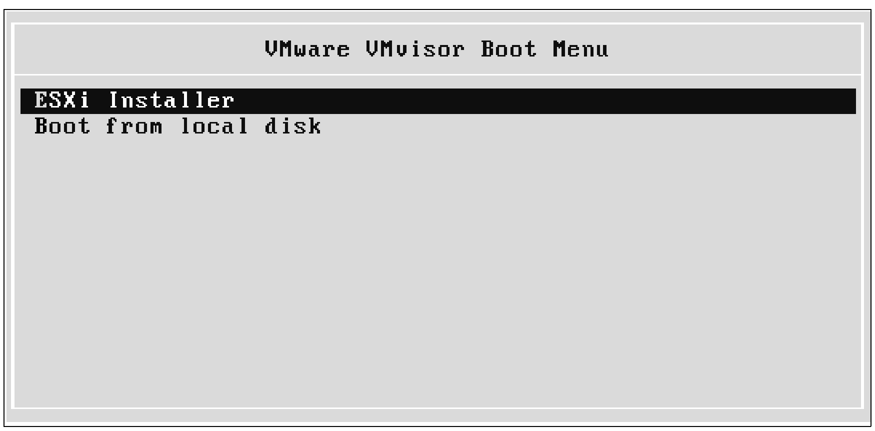


Figure 2-116 Boot Menu of ESXi

2. After the initialization of the ESXi is complete, press Enter to **Install**, as shown in Figure 2-117.

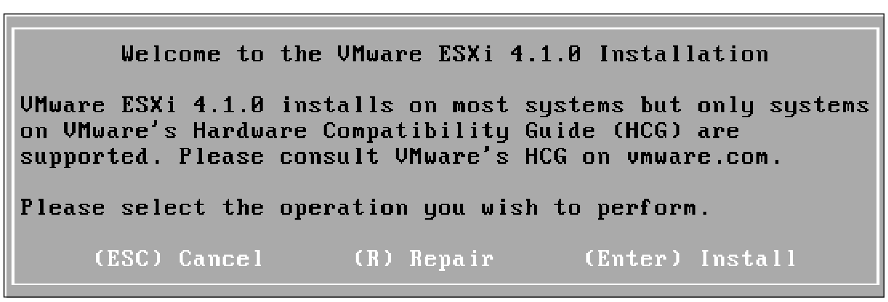


Figure 2-117 ESXi asking you to Install

3. Read and accept the license agreement, as shown in Figure 2-118.

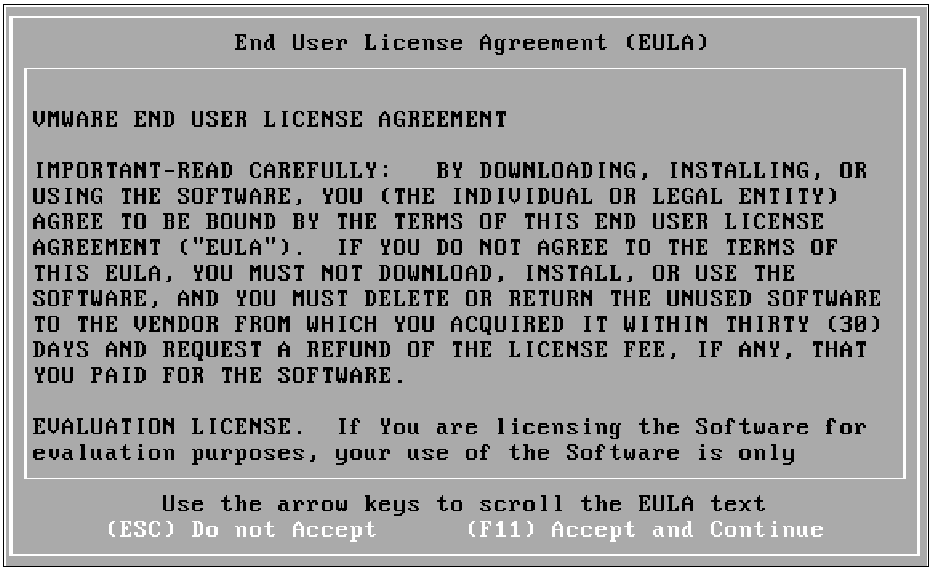


Figure 2-118 Accept the License Agreement

4. As we already disabled the local drives, we see no Local Drives here. We only see the Remote Storage, as shown in Figure 2-119 and we only see one instance of the drive to pick from. We select the 60 GB drive that we created for this installation and press Enter.

Attention: Do not select the Universal Xport drive, as that is the controller.

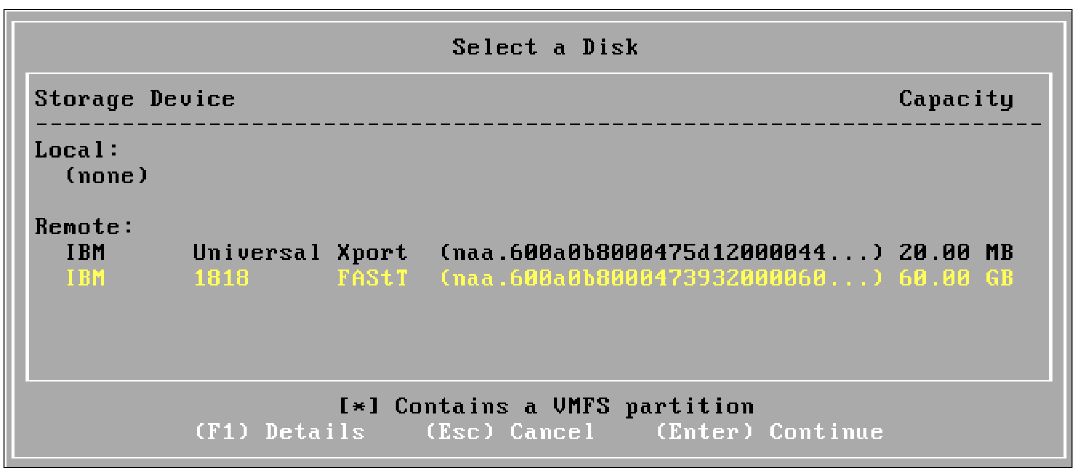


Figure 2-119 Selecting a Remote Disk

5. Confirm the ESXi installation by pressing F11. Be warned that the storage selected for this installation will be formatted as shown in Figure 2-120.

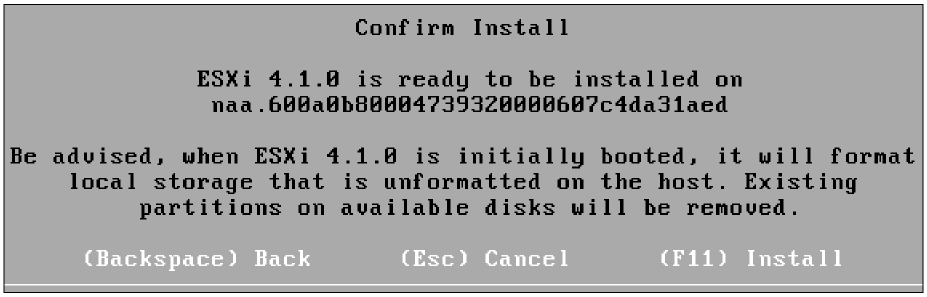


Figure 2-120 ESXi Confirm Install

6. After the installation in completed, we see a panel as shown in Figure 2-121. Press Enter to reboot the system, and the system will now boot from the SAN.

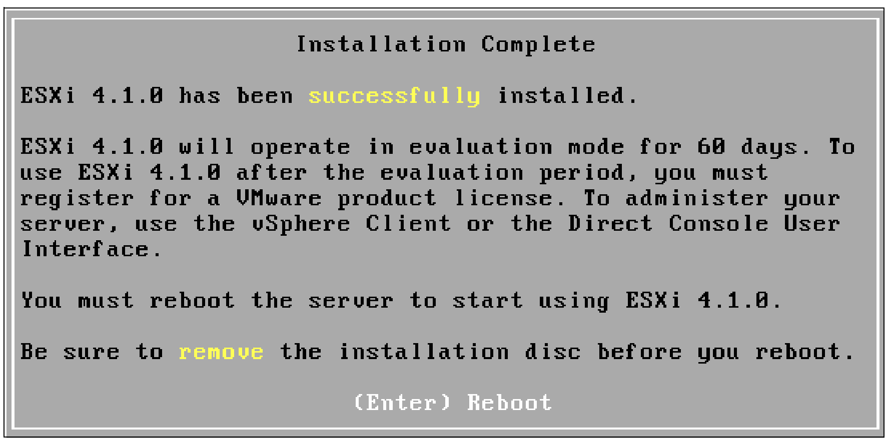


Figure 2-121 Successfully installed ESXi 4

2.12.2 Verifying the ESXi pathing

Follow this procedure to verify the ESXi pathing information:

1. After the installation of ESXi 4 update 1, we need to install Virtual Center to verify that the multipathing is working. In Virtual Center, click the **Summary** tab, and on the right side of the panel you will see the datastore, as shown in Figure 2-122.

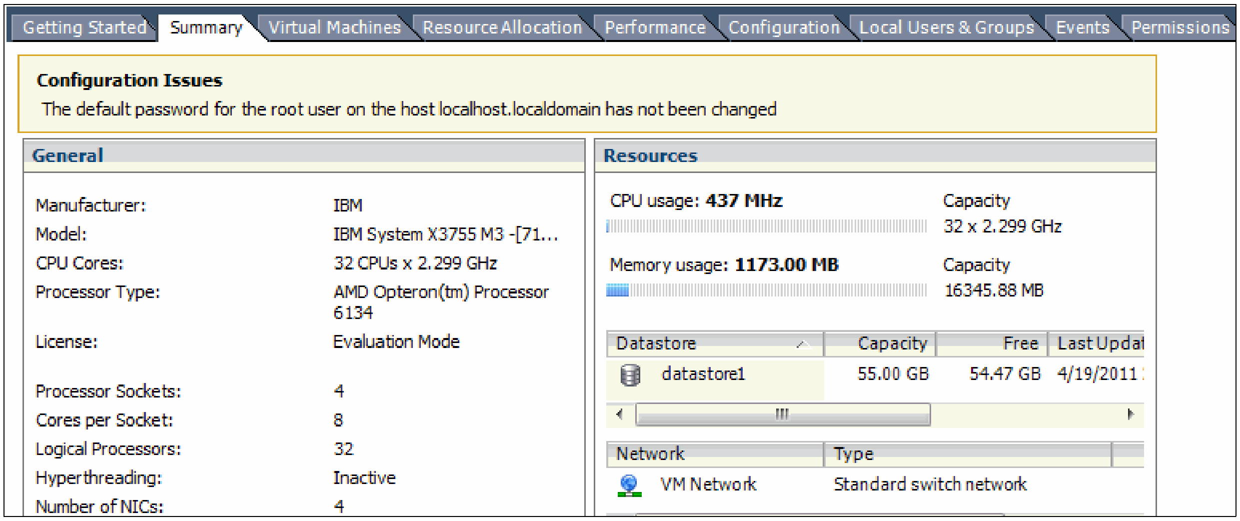


Figure 2-122 Summary tab in Virtual Center

2. To check the pathing, right-click the **datastore1** and then click **Properties**, as shown in Figure 2-123.

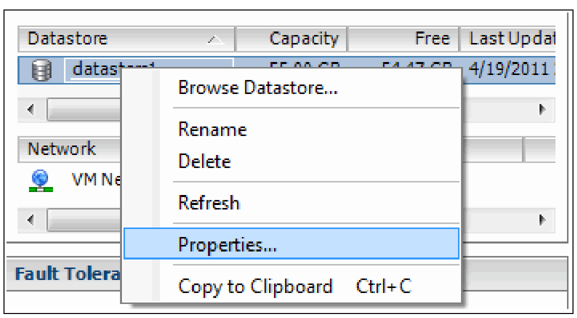


Figure 2-123 Properties of the datastore

3. Verify drive size and click the **Manage paths** button to check the multipath (Figure 2-124).

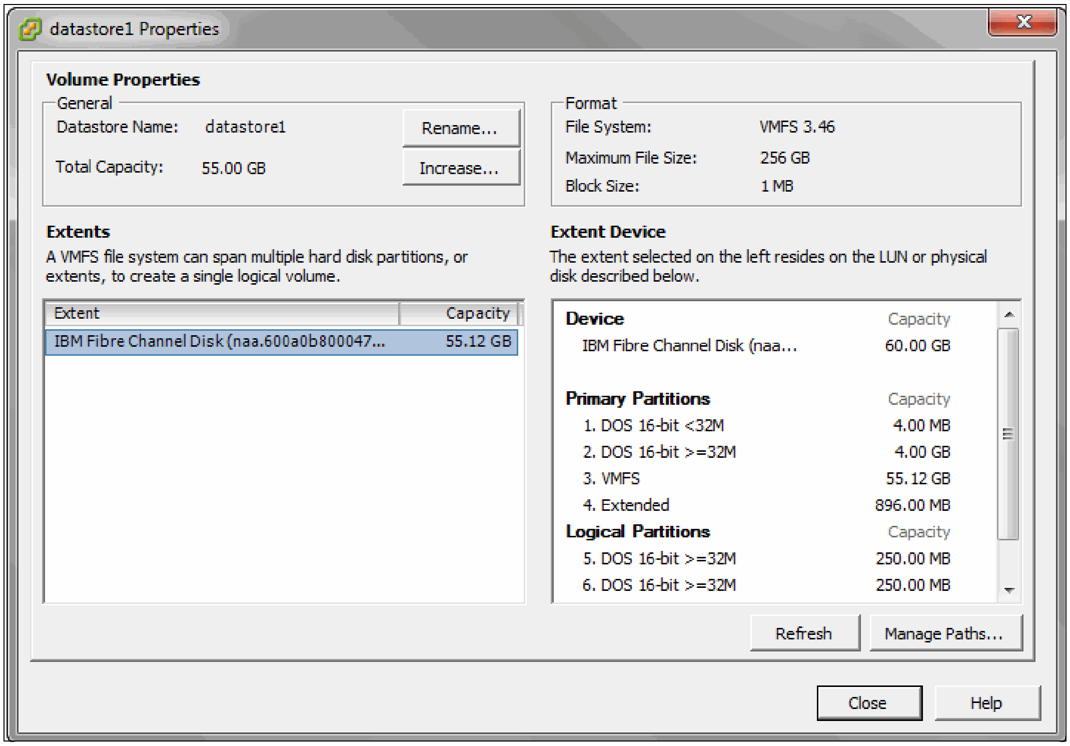


Figure 2-124 Manage paths and properties view

4. As shown in Figure 2-125, we see the two paths, therefore the multipathing is working.

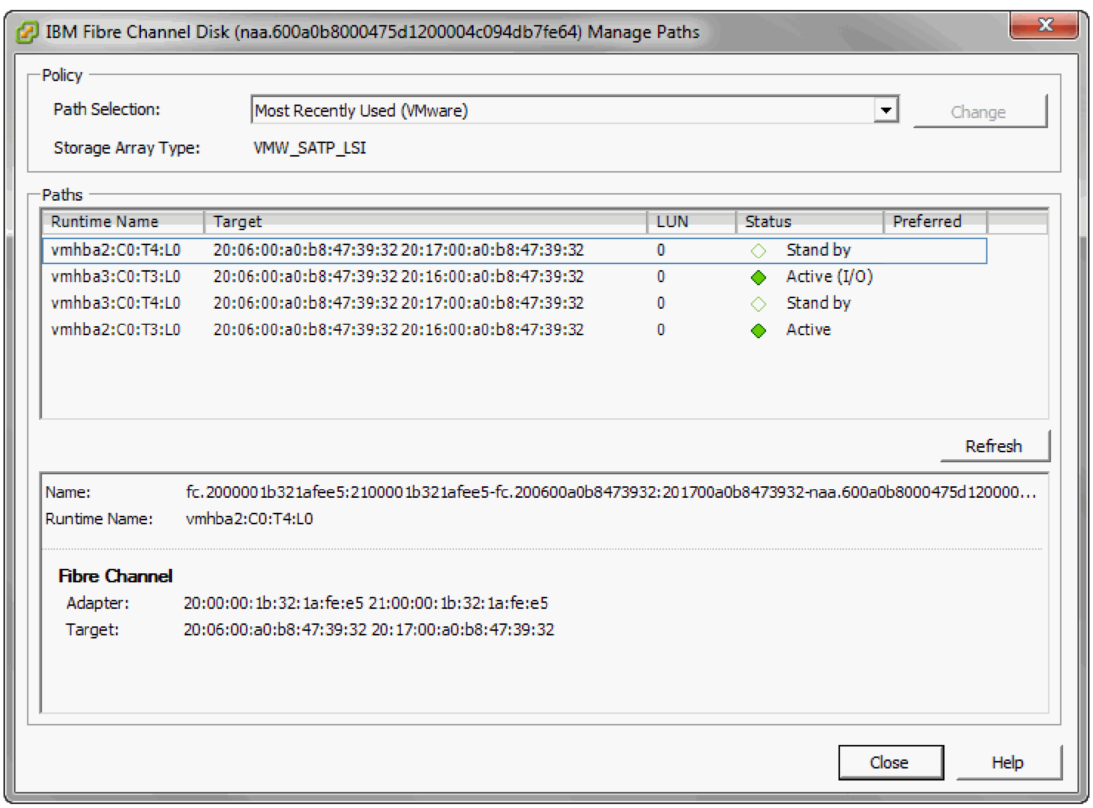


Figure 2-125 Multipathing view

2.13 Installation of AIX 6.1 for SAN Boot

This section describes the steps available to install or migrate the IBM AIX operating system using SAN Boot on the DS5000 series storage subsystems.

The subsections cover the following topics:

- 2.13.1, "AIX requirements for installation"
- 2.13.2, "Installing AIX 6.1 using a pre-existing installation"
- 2.13.3, "New Installation process for SAN Boot"
- 2.13.4, "New AIX installation with NIM"
- 2.13.5, "Identifying the logical drive to hdisk"
- 2.13.6, "AIX post installation tasks with DS5000 best practices"

2.13.1 AIX requirements for installation

This section covers the AIX requirements for installation.

Implementation checklist

Before proceeding with the AIX installation, check the following requirements:

- Plan your OS installation while keeping in mind what OS version, hardware, storage subsystems, SAN architecture, and multipathing will be used.
- Check the status of your Hardware Management Console (HMC) and hardware.
- Verify the micro code for your server, and firmware for your HBA, and make sure that it has been updated.
- Collect all installation media to be used before starting the process.
- Check the license agreements.
- Verify that the LUN assignments and host connectivity were properly done.
- Verify the SAN zoning for the host.
- Check for any special requirement and conduct your installation following the recommendations.

This section describes the various possibilities to install AIX in the SAN Boot environment.

> **Tip:** At the time of writing this book, only FC configuration is supported on AIX for SAN Boot and the newest Power Servers on the DS5000 Subsystem.

Implementation possibilities

Implementations of SAN Boot with AIX include the following capabilities:

- To implement SAN Boot on a system with an already installed AIX operating system, you have these possibilities:
  - Use the alt_disk_install system utility.
  - Mirror an existing SAN Installation to several other LUNs using Logical Volume Copy.
- To implement SAN Boot on a new system, you have these possibilities:
  - Start the AIX installation from a bootable AIX CD install package.
  - Use the Network Installation Manager (NIM).

The methods known as *alt_disk_install or mirroring* are simpler to implement than using the Network Installation Manager (NIM).

You can find more information about the differences about each process by accessing the Information Center, searching for Installation topics using the following link:

http://publib.boulder.ibm.com/infocenter/aix/v6r1/index.jsp?topic=/com.ibm.aix.baseadmn/doc/

2.13.2 Installing AIX 6.1 using a pre-existing installation

Follow this procedure for installing AIX 6.1 using a pre-existing installation.

Creating a boot disk with alt_disk_install

The following procedure is based on the **alt_disk_install** command, cloning the operating system located on an internal SCSI disk to the storage subsystem SAN disk.

> **Tip:** In order to use the **alt_disk_install** command, the following file sets must be installed:
>
> - bos.alt_disk_install.boot_images
> - bos.alt_disk_install.rte (for rootvg cloning)

It is necessary to have an AIX system up and running and installed on an existing SCSI disk. Otherwise, you must first install a basic AIX operating system on the internal SCSI disk and then follow this procedure:

1. Create a logical drive on the storage subsystem, big enough to contain all rootvg data, and make the appropriate zoning to see the storage subsystem from the AIX system.
2. Assuming that the source internal SCSI disk is hdisk0 and the target disk on the storage subsystem is hdisk3, you can use one of the following commands to clone hdisk0 to hdisk3:

   /usr/sbin/alt_disk_install –C –B -P all hdisk3
   /usr/sbin/alt_disk_copy -O -B -d hdisk3 /AIX 6.1/
   smitty alt_clone /appropriate input required/

   > **Attention:** The target disk must have the same or greater capacity than the source disk.

3. Figure 2-126 shows the result of executing **alt_disk_install** on our test IBM AIX 5L™ V6.1 system.

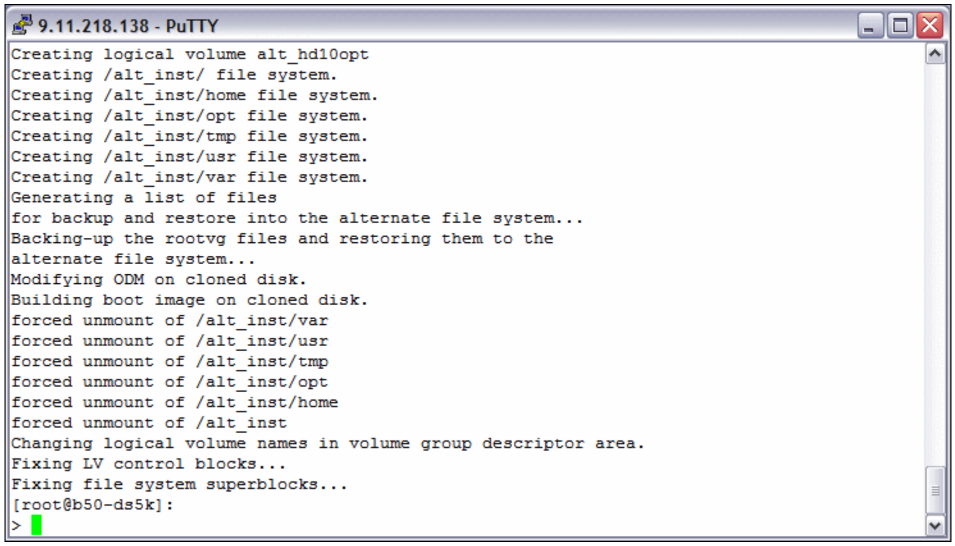


Figure 2-126 Executing alt_disk_install

> **Consideration:** If the rootvg is mirrored, remove the second copy before running the **alt_disk_install** command, or alternatively, use the other *mirroring* method, which is described in 2.13.3, “New Installation process for SAN Boot” on page 107.

4. The **alt_disk_install** command can run for several minutes. It creates a new volume group on the storage subsystem disk called altinst_rootvg, and it creates and copies all logical volumes from rootvg. You can verify the result with the **lspv** command, activating the volume group with **varyonvg altinst_rootvg** and **lsvg -l altinst_rootvg.** See Figure 2-127.

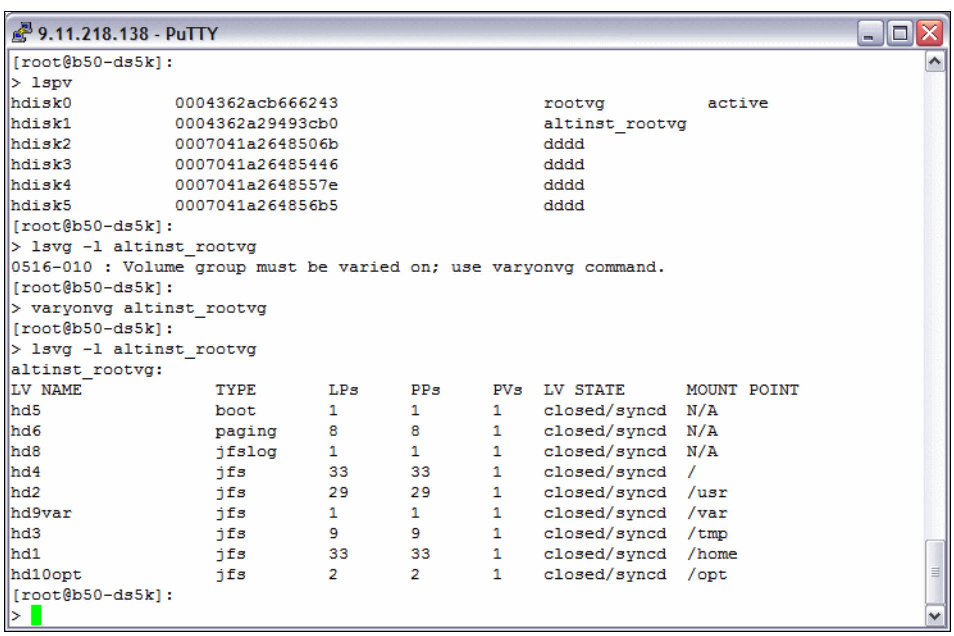


Figure 2-127 Checking the results of alt_disk_install

5. Clean up the alternate disk volume group and make hdisk0 a system disk using either of the following commands:

```
alt_disk_install -X
alt_rootvg_op -X / AIX 6.1/
```

6. Ensure the proper boot logical setting for the target disk:

```
/usr/lpp/bosinst/blvset -d /dev/hdisk3 -g level
```

7. If you find that the boot logical volume settings on the cloned disk are blank, update the cloned boot logical volume manually with the following command:

```
echo y | /usr/lpp/bosinst/blvset -d /dev/hdisk3 -plevel
```

8. Change the AIX boot list to force AIX to start on hdisk3 at the next reboot with a fallback or secondary boot disk of hdisk0 if there is any problem with hdisk3:

```
bootlist -m normal hdisk3 hdisk0
```

> **Attention:** This procedure requires a reboot, so run `shutdown -Fr`.

The system comes back up, booting from the storage subsystem rootvg disk.

9. Check which disk was used to boot by using the following command:

```
bootinfo -b
```

For more information about the use of the `alt_disk_install` command, see the *IBM eServer Certification Study Guide - AIX 5L Installation and System Recovery*, SG24-6183.

Alternatively, the cloned disk (hdisk3 in the previous example) can be remapped on the storage subsystem and used to boot up another AIX machine (or LPAR), a method that is significantly quicker than installing from CDROM on a second machine. An added advantage is that any customized configuration on the source AIX system such as user IDs and tuning parameters are duplicated on the new AIX system or LPAR. Be careful using this method on the new machine, because any network adapters will carry duplicated IP addresses causing network problems. This problem can be easily overcome by ensuring that the new machine does not have any physical network connectivity when it is initially booted using the cloned disk.

Mirroring ROOTVG for SAN Boot

The mirroring process is a basic concept and can be used to migrate the current internal disk for a SAN Boot disk. You can find the basic commands in the Redbooks publication, *IBM eServer Certification Study Guide - AIX 5L Installation and System Recovery*, SG24-6183.

These basic steps are covered in the aforementioned documentation:

- Configure the AIX server to recognize the new SAN Boot disk.
- Proceed with HBA configuration for AIX.
- Add the SAN Boot disk to the current ROOTVG.
- Using LVM Mirror, create and synchronize the system operation logical volumes.
- Remove the LV copy from the internal disk using LVM commands.
- Remove the internal disk from the ROOTVG.
- Configure the bootlist and bosboot to reflect the new configuration.

2.13.3 New Installation process for SAN Boot

To install AIX on storage subsystem disks, make the following preparations:

1. Update the system/service processor microcode to the latest level. For the downloadable versions, see the following website:

   http://techsupport.services.ibm.com/server/mdownload/

2. Be sure to update the FC adapter (HBA) microcode to the latest supported level.
3. Make sure that you have an appropriate SAN configuration. Check that the host is properly connected to the SAN, the zoning configuration is updated, and at least one LUN is mapped to the host.

> **Tip:** If the system cannot see the SAN fabric at login, you can configure the HBAs at the server prompt, `Open Firmware`.

> **Tip:** By default, AIX will install using native drivers and the MPIO package. Using it, the server can be configured with multiple paths before starting to install.

Installation procedure

Complete the Base Operating System (BOS) installation using the following procedure. In this procedure, we perform a new and complete base operating system installation on a logical partition using the partition's media device. This procedure assumes that there is an HMC attached to the managed system.

Prerequisites before starting the installation process

> **Tip:** The information in this how-to scenario was tested using specific versions of AIX 6.1. The results you obtain might vary significantly depending on your version and level of AIX.

Here are some prerequisites to consider before starting the installation process:

- For the installation method that you choose, ensure that you follow the sequence of steps as shown. Within each procedure, you must use AIX to complete some installation steps, while other steps are completed using the HMC interface.
- Before you begin this procedure, you should have already used the HMC to create a partition and partition profile for the client. Assign the SCSI bus controller attached to the media device, a network adapter, and enough disk space for the AIX operating system to the partition. Set the boot mode for this partition to SMS mode. After you have successfully created the partition and partition profile, leave the partition in the Ready state. For instructions about how to create a logical partition and partition profile, see the article, "Creating logical partitions and partition profiles," in *IBM PowerVM Virtualization Introduction and Configuration*, SG24-7940.

Activating and installing the partition

Perform the following steps on the HMC interface to activate and install the partition:

1. Activate the partition as follows:
   a. Insert the AIX Volume 1 installation media into the media device of the managed system.
   b. Check if DVD is already assigned to the current partition.
   c. Right-click the partition to open the menu.

d. Select **Activate**. The Activate Partition menu opens with a selection of partition profiles. Be sure the correct profile is highlighted.
e. Select **Open a terminal window or console session** at the bottom of the menu to open a virtual terminal (vterm) window.
f. Select **Advanced** to open the Advanced options menu.
g. For the Boot mode, select **SMS**.
h. Select **OK** to close the Advanced options menu.
i. Select **OK**. A vterm window opens for the partition.

2. In the SMS menu on the vterm, do the following steps:
   a. Press the 5 key and press Enter to select **5. Select Boot Options** as shown in Figure 2-128.

```
 Version EL340_075
 SMS 1.7 (c) Copyright IBM Corp. 2000,2008 All rights reserved.
-------------------------------------------------------------------------------
 Main Menu
 1.   Select Language
 2.   Setup Remote IPL (Initial Program Load)
 3.   Change SCSI Settings
 4.   Select Console
 5.   Select Boot Options

-------------------------------------------------------------------------------
 Navigation Keys:
X = eXit System Management Services

-------------------------------------------------------------------------------
 Type menu item number and press Enter or select Navigation key:5
```

Figure 2-128 SMS Basic Menu

   b. Press the 1 key and press Enter to select **1. Select Install/Boot Device**.
   c. Press the 3 key and press Enter to select **3. CD/DVD**.
   d. Select the media type that corresponds to the media device and press Enter.
   e. Select the device number that corresponds to the media device and press Enter. The media device is now the first device in the Current Boot Sequence list.
   f. Press the 2 key and press Enter to select **2. Normal Mode Boot**.
   g. Press the 1 key and press Enter to confirm your boot option.

   Afterwards, the system will boot your machine using the CD/DVD and load the basic Kernel to install the machine.

3. Boot from the AIX Volume 1, as follows:
   a. Select console and press Enter.
   b. Select language for BOS Installation menus, and press Enter to open the Welcome to Base Operating System Installation and Maintenance menu.
   c. Type 2 to select **Change/Show Installation Settings and Install** (as shown in Figure 2-129) in the Choice field and press Enter.

```
                         Welcome to Base Operating System
                          Installation and Maintenance

Type the number of your choice and press Enter.  Choice is indicated by >>>.

>>> 1 Start Install Now with Default Settings

    2 Change/Show Installation Settings and Install

    3 Start Maintenance Mode for System Recovery

    4 Configure Network Disks (iSCSI)


    88  Help ?
    99  Previous Menu

>>> Choice [1]: 2
```

Figure 2-129 Base Installation menu

4. Verify or Change BOS Installation Settings, as follows:
   a. Type 2 in the Choice field to select the **Change/Show Installation Settings option**.
   b. Type 1 for System Setting to select **New and Complete Overwrite** in the Choice field and press Enter.
   c. When the Change Disk(s) panel opens, you can change the destination disk for the installation. If the default shown is correct, type 0 in the Choice field and press Enter. To change the destination disk, do the following steps:
      i. Type the number for each disk you choose in the Choice field and press Enter. Do not press Enter a final time until you have finished selecting all disks. If you must deselect a disk, type its number a second time and press Enter.
      ii. Typing 77 two times, you can check if the Storage WWPN and LUN number associated with your machine is correct, see Example 2-13.

Example 2-13 Checking Storage WWPN and LUN ID

```
Change Disk(s) Where You Want to Install

Type one or more numbers for the disk(s) to be used for installation and press
Enter. To cancel a choice, type the corresponding number and Press Enter.
At least one bootable disk must be selected. The current choice is indicated
by >>>.
Name      Device Adapter Connection Location
                 or Physical Location Code
     1  hdisk1    U789C.001.DQDU764-P2-D5
>>>  2  hdisk17   ...-W50050768014052D1-LA000000000000
     3  hdisk6    ...-W5001738000D00142-L1000000000000
     4  hdisk0    U789C.001.DQDU764-P2-D8
     5  hdisk2    U789C.001.DQDU764-P2-D4
    06  MORE CHOICES...

>>>  0   Continue with choices indicated above
    55  More Disk Options
    66  Devices not known to Base Operating System Installation
    77  Display More Disk Information
    88  Help ?
    99  Previous Menu

>>> Choice [0]:
```

iii. When you have selected the disks, type 0 in the Choice field and press Enter. The Installation and Settings panel opens with the selected disks listed under System Settings.

> **Important:** Be sure that you have made the correct selection for root volume group, because the existing data in the destination root volume group will be destroyed during Base Operating System (BOS) installation.

d. If needed, change the primary language environment. Use the following steps to change the primary language used by this installation to select the language and cultural convention you want to use.

> **Tip:** Changes to the primary language environment do not take effect until after the Base Operating System Installation has completed and your system is rebooted.

i. Type 2 in the Choice field on the Installation and Settings panel to select the Primary IBM Language Environment® Settings option.

ii. Select the appropriate set of cultural convention, language, and keyboard options. Most of the options are a predefined combination, however, you can define your own combination of options.

e. After you have made all of your selections, verify that the selections are correct. Press Enter to confirm your selections and to begin the BOS Installation. The system automatically reboots after installation is complete.

5. Select Open terminal window to open a virtual terminal (vterm) window.

a. Type the model of your terminal as the terminal type.

b. In the License Agreement menu, select Accept License Agreements.

c. Select **Yes** to ACCEPT Installed License Agreements.
d. Press F10 (or Esc+0) to exit the License Agreement menu.
e. In the Installation Assistant main menu, select **Set Date and Time** (Figure 2-130).

```
                              Installation Assistant

Move cursor to desired item and press Enter.

  Set Date and Time
  Set root Password
  Configure Network Communications
  Install Software Applications
  Using SMIT (information only)


F1=Help             F2=Refresh          F3=Cancel           F8=Image
F9=Shell            F10=Exit            Enter=Do
```

Figure 2-130 Post installation menu - Installation Assistant

f. Set the correct date, time, and time zone. Press the F3 (or Esc+3) key to return to the Installation Assistant main menu.
g. Select **Set Root Password** and set a root password for the partition.
h. Select **Configure Network Communications**. Select **TCP/IP Startup**. Select from the Available Network Interfaces and press Enter. Enter the appropriate network information in the Minimum Configuration and Startup menu and press Enter. Use the F3 (or Esc+3) key to return to the Installation Assistant main menu.
i. Exit the Installation Assistant by typing F10 (or Esc+0).

The vterm window will display a login prompt at this time.

Switch the partition to Normal Mode

When possible, switch the partition to Normal Mode from the HMC, as follows:

1. Right-click the partition profile to open the menu. Be sure that the correct partition profile is highlighted.
2. Select **Properties** and then select the Settings tab.
3. For the Boot Mode, select **Normal** and select **OK** to close the Properties menu.
4. Right-click the partition and select **Restart Partition**.
5. Select **Immediate** for the Restart Options. Confirm that you want to restart the partition.

2.13.4 New AIX installation with NIM

Network Installation Manager (NIM) is a client server infrastructure and service that allows remote installation of the operating system, manages software updates, and can be configured to install and update third-party applications. Although both the NIM server and client file sets are part of the operating system, a separate NIM server must be configured that will keep the configuration data and the installable product file sets.

NIM preparations

We assume that the following preparations have been completed:

- The NIM environment is deployed and all of the necessary configuration on the NIM master is already done.
- The NIM server is properly configured as the NIM master, and the basic NIM resources have been defined.
- The Fibre Channel Adapters are already installed in the machine onto which AIX is to be installed.
- The Fibre Channel Adapters are connected to a SAN and, on the storage subsystem, have the designated logical drive defined and mapped to the host or NIM client.
- The target machine (NIM client) currently has no operating system installed and is configured to boot from the NIM server.

For more information about how to configure a NIM server, see *NIM from A to Z in AIX 5L*, SG24-7296.

NIM installation procedure

Prior the installation, you can modify the bosinst.data file (for a more automated install), where the installation control is stored. Insert your appropriate values at the following stanza:

```
SAN_DISKID
```

This stanza specifies the World Wide Port Name and a Logical Unit ID for Fibre Channel attached disks. See Example 2-14. The (World Wide Port Name) and (Logical Unit ID) are each in the format returned by the **lsattr** command, that is, “0x” followed by 1-16 hexadecimal digits. The ww_name and lun_id are separated by two slashes (//):

```
SAN_DISKID = <worldwide_portname//lun_id>
```

Example 2-14 LUN ID example

```
SAN_DISKID = 0x0123456789FEDCBA//0x2000000000000

You can specify PVID (the example uses an internal disk):
target_disk_data:
PVID = 000c224a004a07fa
SAN_DISKID =
CONNECTION = scsi0//10,0
LOCATION = 10-60-00-10,0
SIZE_MB = 34715
HDISKNAME = hdisk0
```

To create a BOS installation client using the NIM, follow these steps:

1. Enter the command:

   `# smit nim_bosinst`

2. Select the **lpp_source** resource for the BOS installation.
3. Select the **SPOT** resource for the BOS installation.
4. Select the **BOSINST_DATA** to use during installation option, and select a bosinst_data resource that is capable of performing a unprompted BOS installation.
5. Select the **RESOLV_CONF** or define the host to use for network configuration option, and select a resolv_conf resource.
6. Select the **Accept New License Agreements** option, and select **Yes**. Accept the default values for the remaining menu options.

7. Press Enter to confirm and begin the NIM client installation.
8. To check the status of the NIM client installation, for example, where client name is “va09,” enter the command:

   # lsnim -l va09

9. After checking that the NIM Server and Client are ready, proceed using SMS menu to boot and configure the LPAR to access NIM Master Server.
10. Access the partition console and select **1** to access the SMS Menu; see Figure 2-131.

```
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM

1 = SMS Menu 5 = Default Boot List
8 = Open Firmware Prompt 6 = Stored Boot List
```

Figure 2-131 Boot menu

11.; Select **2** for IPL Configuration Menu, as shown Figure 2-132.

```
PowerPC Firmware
Version SF220_001
SMS 1.5 (c) Copyright IBM Corp. 2000, 2003  All rights reserved.
-----------------------------------------------------------------------------
Main Menu

1. Select Language
2. Setup Remote IPL (Initial Program Load)
3. Change SCSI Settings
4. Select Console
5. Select Boot Options



-----------------------------------------------------------------------------
Navigation Keys:

X = eXit System Management Services
-----------------------------------------------------------------------------
Type the number of the menu item and press Enter or select Navigation Key:
```

Figure 2-132 SMS first menu

12. Be sure that the partition already was cabled to have access to the same LAN where the NIM Master is.

13. Select the NIC card cabled already to access the NIM Master. We select **1** in our case, as shown in Figure 2-133.

```
PowerPC Firmware
Version SF220_001
SMS 1.5 (c) Copyright IBM Corp. 2000, 2003  All rights reserved.
----------------------------------------------------------------------------
NIC Adapters
Device Slot Hardware Address
1. 10/100 Mbps Ethernet PCI Adapt Integ:U0.1-P1/E2 0002554f5c46
2. IBM 10/100/1000 Base-TX PCI-X 4:U0.1-P2-I4/E1 00145eb7f39d
----------------------------------------------------------------------------
Navigation keys:
M = return to Main Menu
ESC key = return to previous screen X = eXit System Management Services
----------------------------------------------------------------------------
Type the number of the menu item and press Enter or select Navigation Key: 1
```

Figure 2-133 NIC card list for remote IPL

14. Select **1** to configure IPs and Network Masks, as shown Figure 2-134.

```
PowerPC Firmware
Version SF220_001
SMS 1.5 (c) Copyright IBM Corp. 2000, 2003  All rights reserved.
----------------------------------------------------------------------------
Network Parameters
IBM 10/100/1000 Base-TX PCI-X Adapter: U0.1-P2-I4/E1
1. IP Parameters
2. Adapter Configuration
3. Ping Test
----------------------------------------------------------------------------
Navigation keys:
M = return to Main Menu
ESC key = return to previous screen X = eXit System Management Services
----------------------------------------------------------------------------
Type the number of the menu item and press Enter or select Navigation Key: 1
```

Figure 2-134 Setting the IP Parameters panel

15. Configure the IP defined for this partition in the /etc/hosts of NIM Master, IP of the NIM Master server, the gateway and mask, as shown in Figure 2-135.

```
PowerPC Firmware
Version SF220_001
SMS 1.5 (c) Copyright IBM Corp. 2000, 2003  All rights reserved.
----------------------------------------------------------------------------
IP Parameters
IBM 10/100/1000 Base-TX PCI-X Adapter: U0.1-P2-I4/E1
1. Client IP Address [9.3.58.217]
2. Server IP Address [9.3.58.194]
3. Gateway IP Address [9.3.58.194]
4. Subnet Mask [255.255.255.000]
----------------------------------------------------------------------------
Navigation keys:
M = return to Main Menu
ESC key = return to previous screen X = eXit System Management Services
----------------------------------------------------------------------------
Type the number of the menu item and press Enter or select Navigation Key:
```

Figure 2-135 Menu to configure the network to access NIM

16. After the correct IP parameters have been entered, return to the main menu to select which boot device you want to use.
17. In the main menu, select **5** (Select boot option)
18. Select 6 (Network) to define your boot by Network
19. Select the right NIC card to boot. We use **1** as described in the steps before.
20. Select **2** for Normal Mode Boot.
21. Select **Yes** to proceed with the normal installation process as per described in the using the NIM Master.

2.13.5 Identifying the logical drive to hdisk

Because a SAN allows access to a large number of devices, identifying the logical drive to hdisk for installation of AIX onto it can be difficult. Use the following method to facilitate the discovery of the lun_id to hdisk correlation:

1. For a new IBM POWER® machine that does not have AIX installed, map only the logical drive that you want to install AIX on. After AIX has been installed, then map all other logical drives required.
2. If more than one disk is assigned to an installed AIX host, you need to make sure you are using the right hdisk and its associated logical drive mapped on the storage subsystem:
   – Identify the hdisk by comparing the unique ID of the hdisk with the logical drive ID when viewing its properties on the storage subsystem. On AIX, use the **lsattr** command to identify the unique ID, for example, on hdisk9, which correlates to “logical drive ID” when you view its properties on the storage subsystem:

     ```
     lsattr -El hdisk9 | grep unique_id | awk '{print (substr($2,6,32))}'
     600A0B80004777D800006FC14AA6F708
     ```

   – The simplest method of checking on AIX system already running where native AIX MPIO is being used is to run the following command:

     mpio_get_config -Av

The output of this comman, as shown in Example 2-15, clearly details the hdisk - logical drive relationship.

Example 2-15 mpio_get_config -Av output

```
Frame id 0:
  Storage Subsystem worldwide name: 608e50017b5bc00004a955e3b
    Controller count: 2
    Partition count: 1
    Partition 0:
    Storage Subsystem Name = 'ITSO_5020'
        hdisk      LUN #   Ownership          User Label
        hdisk2         2   A (preferred)      Secured1
        hdisk3         3   A (preferred)      AIX_Boot
        hdisk4         4   A (preferred)      test3
        hdisk5        40   B (preferred)      p630_Test_40_64KB
        hdisk6        41   A (preferred)      p630_Test_41_8KB
        hdisk7        42   B (preferred)      p630_Test_42_512KB
        hdisk8        43   A (preferred)      p630_Test_43_128KB
```

– For AIX systems using SDDPCM with AIX MPIO, then issuing the equivalent command to the **mpio_get_config** will give the output in the same format as before:

sddpcm_get_config -Av

2.13.6 AIX post installation tasks with DS5000 best practices

After installing AIX, some tasks must be done to preserve the environment, especially when using SAN Boot.

HBA additional configuration

List the Fibre Channel devices and its attributes as shown in Figure 2-136.

```
# lsdev -Cc adapter|grep fcs
fcs0    Available 03-00 4Gb FC PCI Express Adapter (df1000fe)
# lsattr -El fscsi0
attach       switch       How this adapter is CONNECTED         False
dyntrk       no           Dynamic Tracking of FC Devices        True
fc_err_recov delayed_fail FC Fabric Event Error RECOVERY Policy True
scsi_id      0xc0700      Adapter SCSI ID                       False
sw_fc_class  3            FC Class for Fabric                   True
#
```

Figure 2-136 List of devices and default attributes for HBA0/FCS0

As a best practice, the recommendation for HBA is to use the following settings:

- DYNTRK=YES (AIX supports dynamic tracking of Fibre Channel (FC) devices.)

  Previous releases of AIX required for the user to unconfigure FC storage device and adapter device instances before making changes on the system area network (SAN) that might result in an N_Port ID (SCSI ID) change of any remote storage ports.

  If dynamic tracking of FC devices is enabled, the FC adapter driver detects when the Fibre Channel N_Port ID of a device changes. The FC adapter driver then reroutes traffic destined for that device to the new address while the devices are still online.

Events that can cause an N_Port ID to change include moving a cable between a switch and storage device from one switch port to another, connecting two separate switches using an inter-switch link (ISL), and possibly rebooting a switch.

Dynamic tracking of FC devices is controlled by a new fscsi device attribute, dyntrk. The default setting for this attribute is **dyntrk=no**. To enable dynamic tracking of FC devices, set this attribute to **dyntrk=yes**, as shown in Example 2-16.

Example 2-16 changing dyntrk

```
# chdev -l fscsi0 -a dyntrk=yes -a fc_err_recov=fast_fail -P

fscsi0 changed
```

- **FC_ERR_RECOV=FAST_FAIL**

  If the driver receives an RSCN from the switch, it could indicate a link loss between a remote storage port and switch. After an initial 15 second delay, the FC drivers query to see if the device is on the fabric. If not, I/Os are flushed back by the adapter. Future retries or new I/Os fail immediately if the device is still not on the fabric. If the FC drivers detect that the device is on the fabric but the SCSI ID has changed, the FC device drivers do not recover, and the I/Os fail with PERM errors. The default setting for this attribute is delayed_fail. To change the recovery model of FC devices, set this attribute to fc_err_recov=fast_fail, as shown in Example 2-17.

Example 2-17 changing recovery mode

```
# chdev -l fscsi0 -a dyntrk=yes -a fc_err_recov=fast_fail -P
fscsi0 changed
```

Installing the drivers and multipathing subsystem

By default, the AIX will be installed using native MPIO drivers and they must be updated with the right driver to access the DS5000 series subsystems.

AIX multipath drivers

AIX host system requires the MPIO failover driver for Fibre Channel path redundancy. The failover driver monitors I/O paths. If a component failure occurs in one of the Fibre Channel paths, the failover driver reroutes all I/O to another path.

> **Tip:** AIX supports both Redundant Disk Array Controller (RDAC) and Multiple Path I/O (MPIO).

Steps for downloading the AIX file sets

While the RDAC driver is being deprecated with AIX, it is still the default driver on AIX 5.2 and 5.3 for DS4000 devices. With AIX 6.1, the default driver is MPIO for DS5000 series devices. All future DS5000 models will only be supported with MPIO.

The AIX file set for MPIO is **devices.common.IBM.mpio.rte**.

Follow these steps to download the latest file set:

1. Go to the IBM Fix Central website: http://www.ibm.com/support/fixcentral/.
2. In the Product family drop-down menu, select **System p**. The Product drop-down menu displays.
3. In the Product drop-down menu, select **AIX.** A Version drop-down menu displays.

4. Select your AIX version. The Fix type drop-down menu displays.
5. In the Fix type drop-down menu, select **Fix packs** and click **Continue**. The Fix packs page displays.
6. From the Select a Technology Level drop-down menu, select the fix pack for your operating system version and click **Go**. A Fix packs table displays.
7. From the Fix packs table, click the link for the latest available version and click **Continue**. The Package download page displays.
8. Under Obtain package, click the radio button next to the download method you prefer and click **Continue**. The Download all file sets using applet window will appear. Click the **Download now** button to begin downloading your file set.

Installing the AIX multipath driver

After you install the client software and configure your storage subsystems, use these instructions to install the appropriate DS Storage Manager multipath device driver.

You must install the multipath driver on all AIX hosts that are connected to your storage subsystem. This section describes how to check the current multipath driver version level, update the multipath device driver, and verify that the multipath driver update is complete.

> **Tip:** MPIO is installed and configured as part of the base operating system installation. No further configuration is required. However, you can add, remove, reconfigure, enable, and disable devices (or device paths) using SMIT, Web-based System Manager, or the command line interface. The following commands help manage MPIO paths.

Steps for installing the multipath driver: Complete the following steps to update the multipath driver version (devices.common.IBM.mpio.rte) on an AIX system. Repeat these steps for all AIX systems that are connected to the storage subsystem.

Before you begin: Verify whether your AIX systems have the most recent multipath file sets by checking the Storage Manager README file for AIX. You need to perform this installation only on AIX systems that do not have the most recent multipath file sets.

1. Download the most recent file sets by completing the procedure described in “Steps for downloading the AIX file sets” on page 117.
2. Verify that the correct version of the software was successfully installed by typing the following command:

   Example 2-18 Verify that the correct version

   ```
   # lslpp -ah devices.common.IBM.mpio.rte
   ```

   The verification process returns a table that describes the software installation, including the installation package file set name, version number, action, and action status. If the verification process returns an error, contact your IBM technical support representative. If it does not return an error, then you are finished installing the updated multipath driver on this AIX system.
3. For each AIX host that is to be attached to a DS5000 series storage subsystem, repeat the foregoing steps in order to install and verify the multipath driver.

The implementation of the AIX SAN Boot procedures for IBM System Storage DS5000 is now complete.

3

SAN Boot implementation with IBM System Storage DS8000

In this chapter, we describe SAN Boot implementation using IBM System Storage DS8000 and the best practices to improve the security, performance, and reliability.

Careful planning is essential to configuration of the SAN Boot process. This chapter provides a guideline to help you with the planning process. Choosing the right equipment and software, and also knowing what the right settings are for a particular installation, can be challenging. Every installation needs to answer these questions and accommodate specific requirements. Having a well thought-out design and plan prior to the implementation will help you get the most out of your investment.

During the planning process, there are numerous questions that you need to answer about your environment before you start this implementation:

- What tools better match the requirement to help me to manage and configure the DS8000 series hardware?
- How should I configure my storage for SAN Boot?
- How do I configure my host for SAN Boot?
- How do I map the LUN to the host?
- How do I configure my operating system to use SAN Boot?
- What are the best practices for using DS8000 series while doing SAN Boot?
- What are the troubleshooting procedures I need to know for SAN Boot?

In this chapter we provide answers to these questions.

© Copyright IBM Corp. 2012. All rights reserved.

3.1 DS8000 overview

This section introduces the features, functions, and benefits of the IBM System Storage DS8000 storage subsystem. We used the latest generation of the DS8000 product line to give you a quick overview of all features and functions. Some of these features might be not available for some DS8000 older families. See your product specific documentation for more information regarding the older models of the DS8000.

This section only provides a basic overview of the DS8000 functions and features. We provide an overview of the following topics:

- The DS8000 Storage family
- Overall architecture and components
- DS8000 Storage characteristics helpful for SAN Boot

For more information, refer to *IBM System Storage DS8800: Architecture and Implementation*, SG24-8886.

3.1.1 The DS8000 Storage family

The IBM System Storage DS family is designed as a high performance, high capacity, and resilient series of disk storage systems. It offers high availability, multiplatform support, and simplified management tools to help provide a cost-effective path to an on demand world.

The IBM System Storage DS8000 series encompasses the flagship disk enterprise storage products in the IBM System Storage portfolio. The DS8800, which is the IBM fourth generation high-end disk system, represents the latest in this series, introducing the new small form factor 2.5-inch SAS disk drive technology, IBM POWER6+™ processors, as well as the new 8 Gbps disk adapter (DA) and host adapter (HA) cards.

The IBM System Storage DS8800, shown in Figure 3-1, is designed to support the most demanding business applications with its exceptional all-around performance and data throughput. Combined with the world-class business resiliency and encryption features of the DS8800, it provides a unique combination of high availability, performance, and security. Its tremendous scalability, broad server support, and virtualization capabilities can help simplify the storage environment by consolidating multiple storage systems onto a single DS8800.

Introducing new high density storage enclosures, the DS8800 model offers a considerable reduction in footprint and energy consumption, thus making it the most space and energy-efficient model in the DS8000 series.

Figure 3-1 IBM System Storage DS8800, the IBM fourth generation high-end disk system

The IBM System Storage DS8800 adds Models 951 (base frame) and 95E (expansion unit) to the 242x machine type family, delivering cutting edge technology, improved space and energy efficiency, and increased performance. Compared with its predecessors, the IBM System Storage DS8100, DS8300 and DS8700, the DS8800 is designed to provide new capabilities for the combination of price and efficiency that is right for all application needs. The IBM System Storage DS8800 is a high performance, high capacity series of disk storage systems. It offers balanced performance and storage capacity that scales linearly up to hundreds of terabytes.

Here are some highlights of the IBM System Storage DS8800:

- Robust, flexible, enterprise class, and cost-effective disk storage
- Exceptionally high system availability for continuous operations
- Cutting edge technology with small form factor (2.5-inch) SAS-2 drives, 6 Gbps SAS-2 high density storage enclosures
- 8 Gbps Fibre Channel host and device adapters providing improved space and energy efficiency, and increased performance

- IBM POWER6® processor technology
- Capacities currently from 2.3 TB (16 x 146 GB 15k rpm SAS drives) to 633 TB (1056 x 600 GB 10k rpm SAS drives)
- Point-in-time copy function with FlashCopy, FlashCopy SE
- Remote Mirror and Copy functions with Metro Mirror, Global Copy, Global Mirror, Metro/Global Mirror, IBM z/OS® Global Mirror, and z/OS Metro/Global Mirror with Incremental Resync capability
- Support for a wide variety and intermix of operating systems, including IBM i and System z®
- Designed to increase storage efficiency and utilization, ideal for green data centers

3.1.2 Overall architecture and components

From an architectural point of view, the DS8800 offers continuity with respect to the fundamental architecture of the predecessor DS8100, DS8300, and DS8700 models. It ensures that the DS8800 can use a stable and well-proven operating environment, offering the optimum in availability. The hardware is optimized to provide higher performance, connectivity, and reliability.

Figure 3-2 and Figure 3-3 show the front and rear view of a DS8800 base frame (model 951) with two expansion frames (model 95E), which is the current maximum DS8800 system configuration.

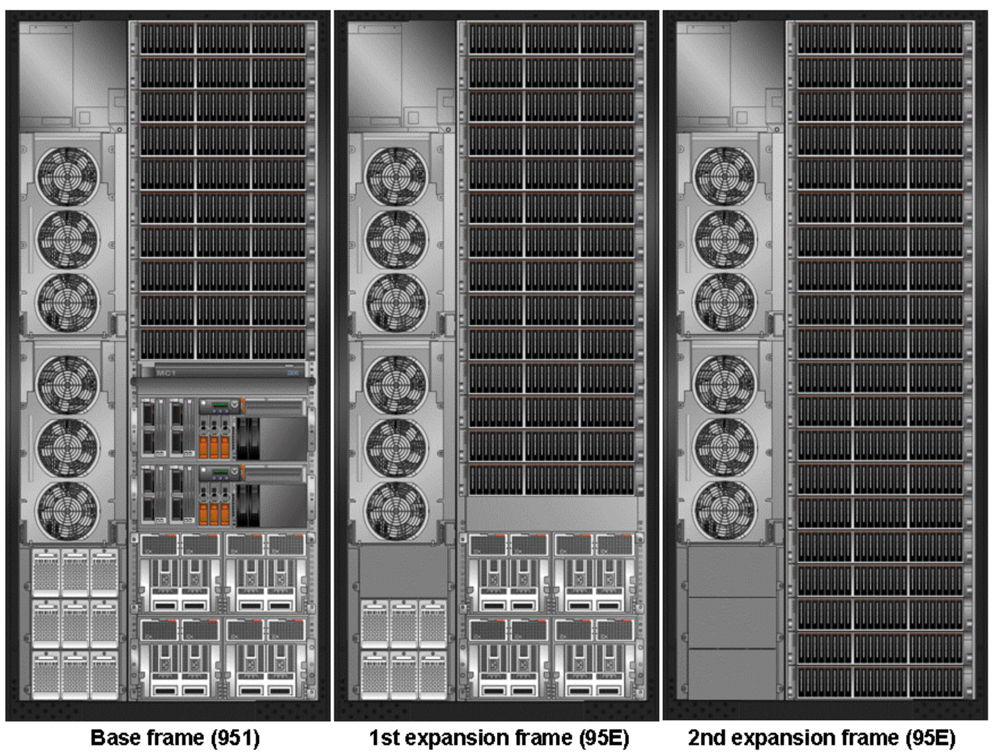


Figure 3-2 DS8800 base frame with two expansion frames (front view, 2-way, no PLD option)

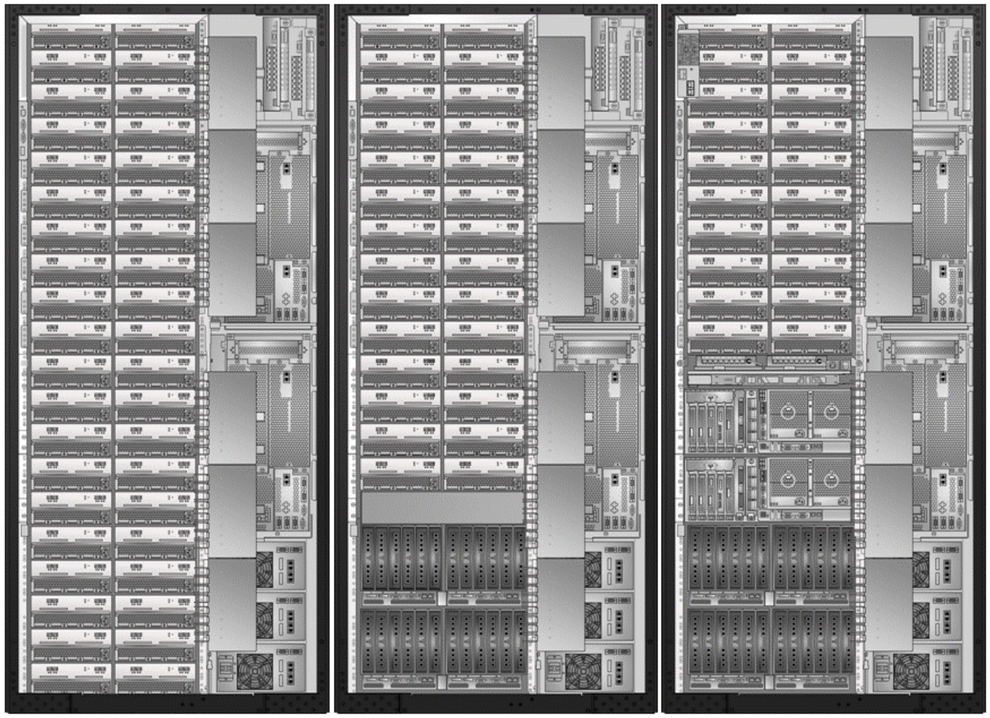


Figure 3-3 DS8800 base frame with two expansion frames (rear view, 2-way, no PLD option)

3.1.3 DS8000 Storage characteristics helpful for SAN Boot

The DS8000 series uses virtualization techniques to separate the logical view of hosts onto LUNs from the underlying physical layer, thus providing high configuration flexibility. Use these features to separate different configuration on same subsystem, for example, Open Systems and Mainframe.

Dynamic LUN/volume creation, deletion, and expansion

The DS8000 gives a high degree of flexibility in managing storage, allowing LUNs to be created and deleted non-disruptively. Also, when a LUN is deleted, the freed capacity can be used with other free space to form a LUN of a different size. A LUN can also be dynamically increased in size. Plan for the environment to have better all resources available on DS8000 Subsystems to gain performance and security.

Simplified LUN masking

The implementation of volume group-based LUN masking (as opposed to adapter-based masking, as on the ESS) simplifies storage management by grouping all or some WWPNs of a host into a Host Attachment. Associating the Host Attachment to a Volume Group allows all adapters within it access to all of the storage in the Volume Group.

Logical definitions: Maximum values

Here is a list of the current DS8000 maximum values for the major logical definitions:

- Up to 65280 logical devices
- Up to 2 TB LUNs
- Up to 1280 paths per FC port
- Up to 8000 process logins (509 per SCSI-FCP port)

3.2 Management and support tools

We have four different ways to access the DS8000. The first way is used by IBM support preferably and is called the Hardware Management Console. The other possibilities are by Storage Manager GUI, by System Storage Productivity Center (SSPC)/Tivoli Productivity Center (TPC). or by Command Line Interface (DSCLI).

3.2.1 Storage Hardware Management Console for the DS8800

The Hardware Management Console (HMC) is the focal point for maintenance activities. The management console is a dedicated workstation (mobile computer) that is physically located (installed) inside the DS8800 and can proactively monitor the state of your system, notifying you and IBM when service is required. It can also be connected to your network to enable centralized management of your system using the IBM System Storage DS Command-Line Interface or storage management software utilizing the IBM System Storage DS Open API. The HMC supports the IPv4 and IPv6 standards.

An external management console is available as an optional feature and can be used as a redundant management console for environments with high availability requirements.

3.2.2 IBM System Storage DS Storage Manager GUI

The IBM System Storage DS Storage Manager, Figure 3-4, is an interface that is used to perform logical configurations and Copy Services management functions.

The Tivoli Storage Productivity Center is required to remotely access the DS Storage Manager GUI. The DS Storage Manager can be accessed through the Tivoli Storage Productivity Center Element Manager from any network-connected workstation with a supported browser. It can also be accessed directly from the management console by using the browser on the hardware management console.

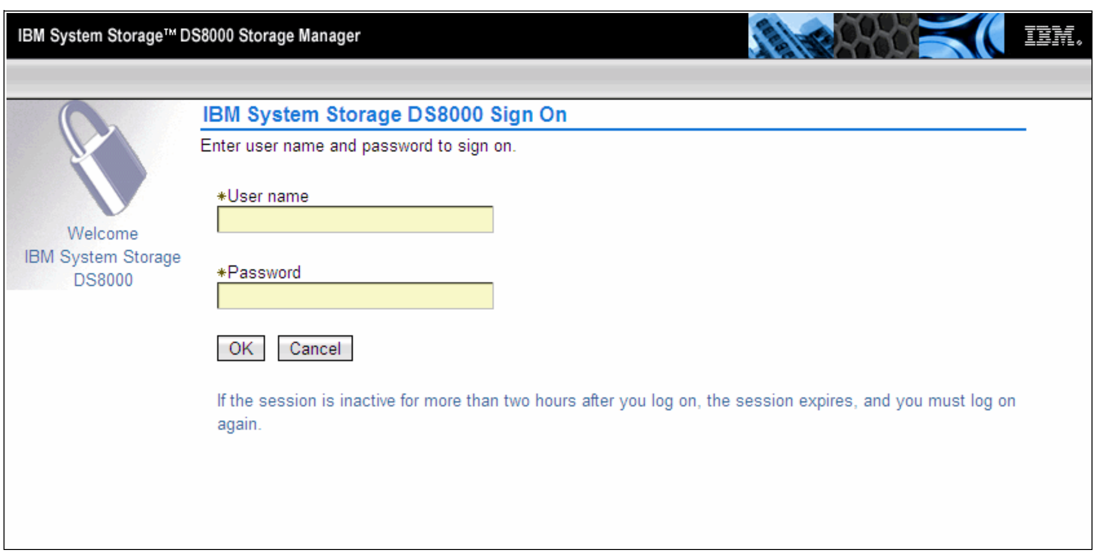


Figure 3-4 DS Storage Manager GUI Interface

3.2.3 DS command line interface for the DS8800

The command line interface (CLI) provides a full-function command set that allows you to check your Storage Unit configuration and perform specific application functions when necessary. For detailed information about DS CLI use and setup, see *IBM System Storage DS: Command-Line Interface User's Guide*, GC53-1127.

The following list highlights a few of the functions that you can perform with the DS CLI:

- Create user IDs that can be used with the GUI and the DS CLI.
- Manage user ID passwords.
- Install activation keys for licensed features.
- Manage storage complexes and units.
- Configure and manage Storage Facility Images.
- Create and delete RAID arrays, ranks, and Extent Pools.
- Create and delete logical volumes.
- Manage host access to volumes.
- Check the current Copy Services configuration that is used by the Storage Unit.
- Create, modify, or delete Copy Services configuration settings.
- Integrate LDAP policy usage and configuration.
- Implement encryption functionality.

Tip: The DSCLI version must correspond to the LMC level installed on your system. You can have more versions of DSCLI installed on your system, each in its own directory.

Figure 3-5 shows a sample panel of the DS CLI.

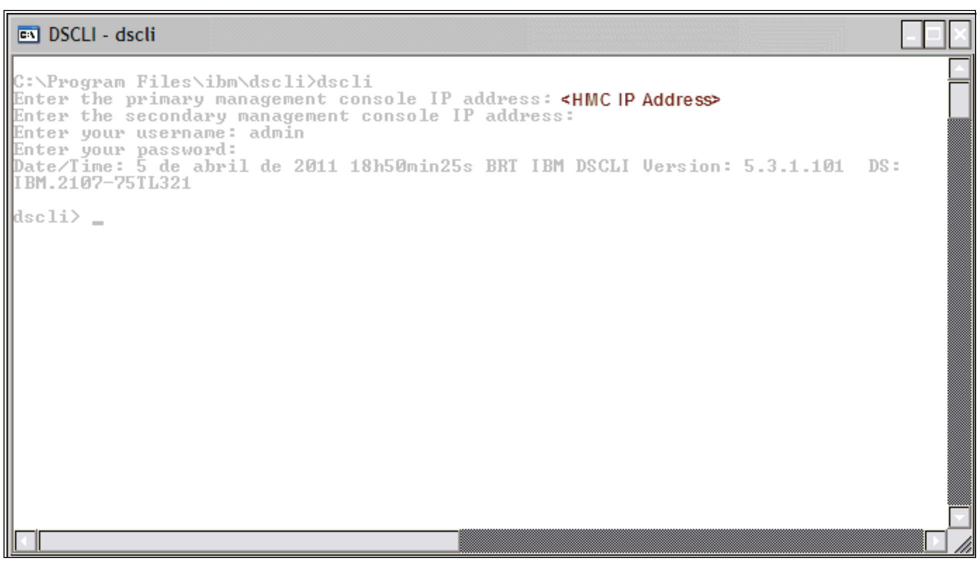


Figure 3-5 DSCLI interface

3.2.4 IBM System Storage Productivity Center management console

As the main focal point for configuration and management, the DS8800 leverages the IBM System Storage Productivity Center (SSPC), an advanced management console that can provide a view of both IBM and non-IBM storage environments. The SSPC can enable a greater degree of simplification for organizations confronted with the growing number of element managers in their environment. The SSPC is an external System x server with pre-installed software, including IBM Tivoli Storage Productivity Center Basic Edition.

Utilizing IBM Tivoli Storage Productivity Center (TPC) Basic Edition software, SSPC extends the capabilities available through the IBM DS Storage Manager. SSPC offers the unique capability to manage a variety of storage devices connected across the storage area network (SAN). The rich, user-friendly graphical user interface provides a comprehensive view of the storage topology, from which the administrator can explore the health of the environment at an aggregate or in-depth view.

Moreover, the *TPC Basic Edition*, which is pre-installed on the SSPC, can be optionally upgraded to *TPC Standard Edition*, which includes enhanced functionality. Such functions include monitoring and reporting capabilities that can be used to enable more in-depth performance reporting, asset and capacity reporting, and automation for the DS8000, These features can also be used to manage other resources, such as other storage devices, server file systems, tape drives, tape libraries, and SAN environments.

Figure 3-6 shows the login panel of the SSPC console. By accessing the SSPC console, you can launch the Element Manager Console, a Web based interface to manage the DS8000 storage subsystem.

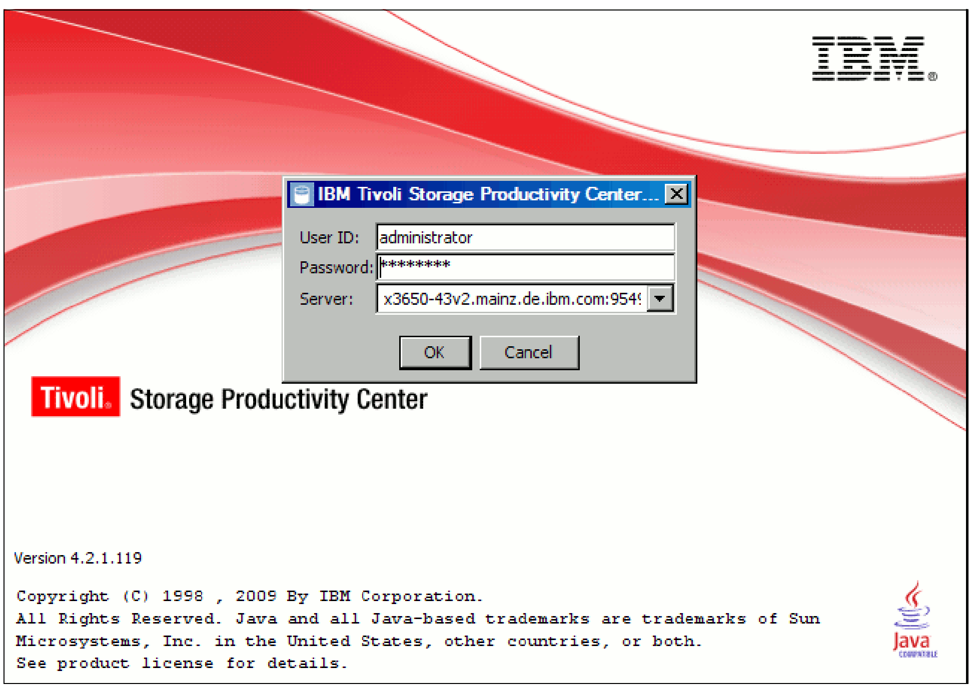


Figure 3-6 SSPC/TPC login interface example

Select the DS8000 Subsystem and launch the interface as shown in Figure 3-7.

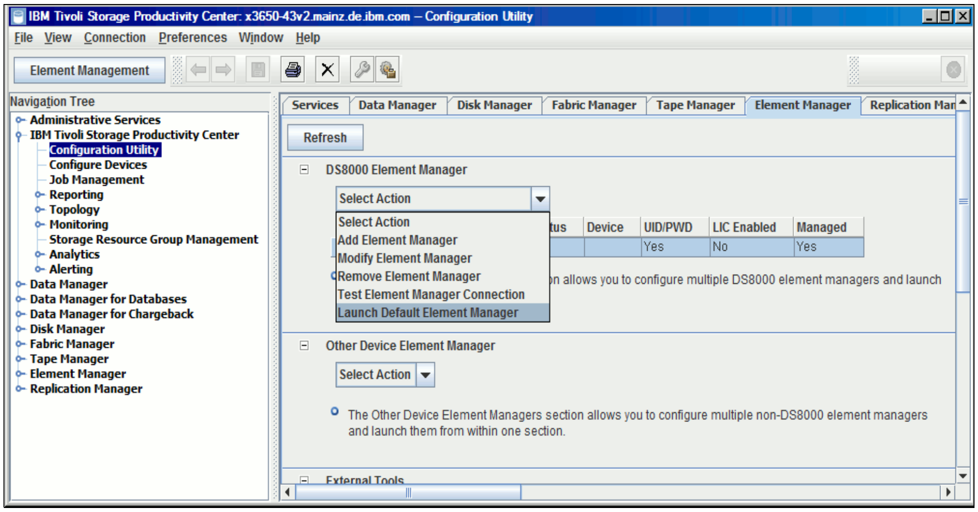


Figure 3-7 Launch the Element Manager

3.3 SAN configuration for the DS8000 Subsystem

As a best practice and to improve the security, all SAN Administrators need to think about how to implement and guarantee the operation for their environment.

For optimal performance and security, use the following guidelines if you are attaching multiple paths from a single host system to I/O ports on a host adapter of a storage image, as shown in Figure 3-8:

- Use attached I/O ports on different host adapters.
- Use multiple physical adapters.
- Do not use all the ports on each host adapter.

Disregarding these path considerations can affect the performance and availability of a storage image.

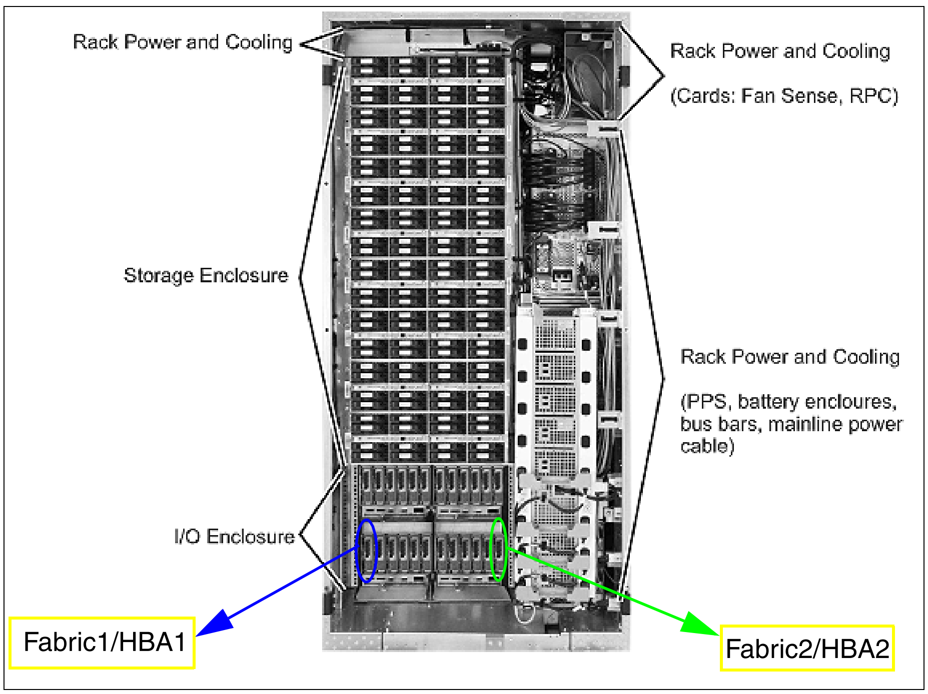


Figure 3-8 Example to connect a different HBA to different port on different host card and cage

> **Tip:** In summary, to prevent any disruption in the service, plan to use at least two ports on the storage subsystem, in two different cards, in two different cages. Split your hosts and SAN Boots to access different paths, balancing the throughput for all available paths at DS8000.

3.3.1 Fibre Channel architecture

Fibre Channel architecture provides various communication protocols on the storage unit. The units that are interconnected are referred to as nodes. Each node has one or more ports.

A DS8000 storage subsystem is a node in a Fibre Channel network. Each port on a storage unit Fibre Channel host adapter is a Fibre Channel port. A host is also a node in a Fibre Channel network. Each port attaches to a serial transmission medium that provides duplex communication with the node at the other end of the medium.

DS8000 storage subsystem architecture supports three basic interconnection topologies or network structures for open systems:

- Arbitrated loop
- Point-to-point
- Switched-fabric

> **Tip:** DS8000 supports the arbitrated loop, point-to-point, and switched-fabric topologies. To change the topology, configure the ports.

For a list of longwave and shortwave adapter cables and their distances, see the *IBM System Storage DS8700 and DS8800 Introduction and Planning Guide*, GC27-2297.

3.3.2 Zoning considerations

Follow these best practice zoning considerations:

- Plan to have two fabrics to prevent outages.
- Use an individual zone for each single initiator.
- Respect the zoneset and zone naming conventions.
- Use soft zones instead of hard zones.
- Implement a policy to back up the Switch and ZoneSet configuration.
- Consider implementing core / edge architecture.
- Keep the firmware updated.
- Plan a baseline for SAN structure.
- Always check the compatibility matrix.

3.4 Host Bus Adapter (HBA) configuration

In this section, we first describe the general requirements for SAN Boot and then continue with with the setup procedures. As explained in Chapter 1, "Introduction to SAN Boot" on page 1, the basic steps to have a SAN Boot in place are as follows:

- Check if your host, cards, and storage can support the SAN Boot.
- Update your server and cards using the supported firmware version.
- Create a LUN to be used as a bootable disk.
- Install your operating system following the procedure outlined in this chapter and/or vendor recommendation.

3.4.1 General requirements for attaching a host

Before you attach a host to the DS8000, review this list of general requirements for all hosts. Then, review the specific host requirements described in the section for each host.

Perform the following steps before you attach any host system to DS8000:

1. Go to the System Storage Interoperation Center (SSIC) website for the most current information about supported hosts, operating systems, adapters, and switches.
2. Obtain a list of supported host bus adapters (HBAs), firmware, and device driver information for your host system and host adapter on the System Storage Interoperation Center (SSIC) website.
3. Ensure that you can reference the following documentation:
   - The *IBM System Storage DS8000 User's Guide* from the System Storage CD that you receive with the storage unit
   - The IBM System Information Center link http://publib.boulder.ibm.com/eserver/
   - The IBM System Storage DS8000 Information Center (SSIC) link as shown in Figure 3-9. http://publib.boulder.ibm.com/infocenter/ds8000ic/index.jsp

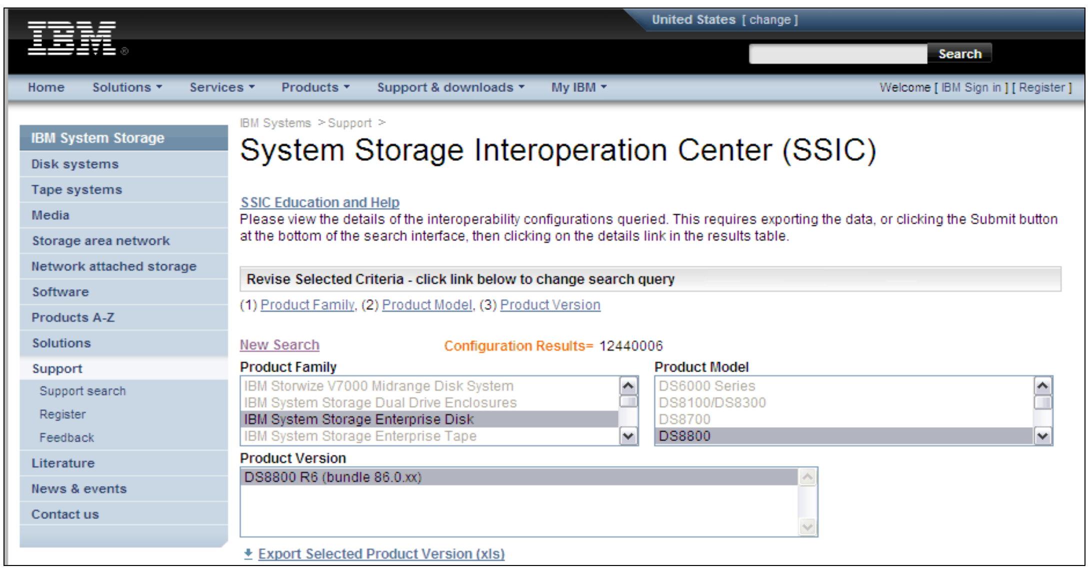


Figure 3-9 SSIC web page

4. Have an IBM service support representative install the storage unit.
5. Check the LUN limitations for your host system and verify that there are enough adapters installed in the server to manage the total LUNs that you want to attach.
6. Review the "I/O adapter features" section of the *IBM System Storage DS8000 Introduction and Planning Guide* for detailed host adapter configuration rules and for information about host adapter and cable feature codes, host adapter balancing for optimal performance, and supported cabling distances.
7. Use your preferred DS8000 Management tool to define the host and I/O port configurations. Make sure that you define the worldwide port names for Fibre Channel ports.
8. Install the adapter driver with your host adapter or use the steps defined in "Downloading and installing a host adapter driver" to download and install an updated adapter driver.

3.4.2 Downloading and installing a host adapter driver

Complete this task to download and install host adapter drivers to use with the DS8000.

Attention: You need to download the relevant vendor documentation for the driver that you select in order to correctly install and configure the host adapter.

Installation procedure

Follow this procedure to install the latest host adapter driver:

1. Go to the System Storage Interoperation Center (SSIC) website and verify that the host adapter you want to use is compatible with the DS8000 series, your host, and your host operating system.
2. Obtain a list of supported host bus adapters (HBAs), firmware, and device driver information for your host system and host adapter on the System Storage Interoperation Center (SSIC) website.
3. Download the adapter driver from the adapter vendor website. Table 3-1 provides vendor webpages and specific download steps. For the most current steps, see the vendor website download page.

Table 3-1 Host Adapter driver download steps by vendor

Host adapter vendor	Website	Steps to locate download page
AMCC/JNI	http://www.amcc.com	1. Click **Downloads ? FC HBA/OEM Files**. 2. Click **IBM**. 3. Locate your adapter. 4. Download the appropriate files for your
Emulex	http://www.emulex.com/downloads/ibm/oneconnect-software-kits.html	1. Click the tab for the adapter type that you are looking for. 2. Locate your adapter. 3. Download the appropriate files for your adapter.
Hewlett-Packard	http://www.hp.com	1. Enter the name of the host adapter in the **Search** field at the bottom of the Hewlett-Packard home page. 2. In the search results, click the link for the driver for your operating system. 3. Click **Download** to download the adapter driver. 4. Return to the search results and review the links to documentation, such as installation requirements and release notes.
IBM (Netfinity®)	http://www.ibm.com/servers/storage/support/disk	1. Select your storage unit from the Enterprise Storage Servers list. 2. Click the **Plan/Upgrade** tab. 3. Click **HBA interoperability search tool**. 4. Select the appropriate options for your product and operating system and click **Submit**. 5. Locate the section for the current version of the driver and firmware that you want, and click **View Details**. 6. Click the **Driver Level** that you want. 7. Click **Continue** to leave the IBM System Storagesupport website. 8. Click one of the **Download** options for the feature code that you want.

Host adapter vendor	Website	Steps to locate download page
QLogic	http://support.qlogic.com/support/oem_ibm.asp	1. Click the DS8000 link. 2. Locate your adapter. 3. Download the appropriate files for your adapter.
Sun	http://www.sun.com/storage/san	If you are using Solaris 8 or 9, perform the following steps. If you are using Solaris 10, you can skip these steps because the SAN software is included in the operating system. 1. Scroll down to the **Get the Software** section. 2. Locate the current driver in the list and click the appropriate link. 3. Type your **Username** and **Password** in the fields, and click **Log in and Continue**. If you do not have a user name and password[a], complete the registration process and return to this page when you have received them. 4. Click **Accept** to agree to the license agreement (required). 5. Download the appropriate files for your adapter.

a. Sun only grants user names and passwords to customers who have purchased maintenance contracts.

6. Follow the installation instructions from the vendor to install the host adapter driver. Some installation instructions might be included in the README files that are included in the compressed download, but some installation documentation might need to be downloaded separately. Return to the vendor website that is listed in Table 3-1 on page 132 to locate the installation and configuration instructions.
7. Locate the chapter for your host in *IBM System Storage DS8000 Host Attachment and Interoperability*, SG24-8887, and follow any additional driver configuration instructions.

Platform and operating system vendor pages

The platform and operating system vendors also provide much support information for their clients. Refer to this information for general guidance about connecting their systems to SAN-attached storage. However, be aware that in some cases you cannot find information to help you with third-party vendors. Always check with IBM about interoperability and support from IBM in regard to these products. It is beyond the scope of this book to list all the vendor websites.

3.4.3 BIOS configuration

The first step in configuring the host is to disable the local boot media, such as IDE, SAS, and SATA drives, in the server's BIOS settings. Perform the following tasks:

1. Power on the server and interrupt the boot sequence to enter the system BIOS by pressing the **F1** key.
2. Select **Devices and I/O ports** from the main menu.
3. Select **IDE configuration** from the next menu and disable the Planar SAS, as shown Figure 3-10.

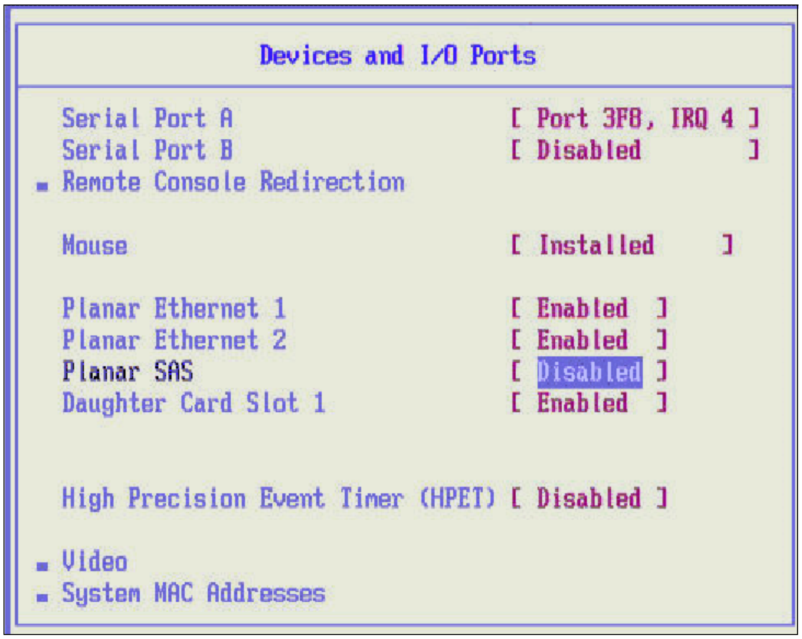


Figure 3-10 Disable Planar SAS

4. Press Esc to exit out of this menu and save the changes.

The software configuration parameters in the BIOS of a Fibre Channel (FC) HBA allows the server to identify and map the boot disk on the SAN. In the following sections, we walk you through use of HBA configuration utilities for SAN Boot solutions.

3.4.4 Finding the QLogic WWPN

To find the WWPN of the QLogic adapter, follow this procedure:

1. Remove the cables attached in your QLogic adapter.
2. Reboot the system with the QLogic HBA.
3. Press Ctrl-Q or Alt-Q and the QLogic HBA banner appears, as shown in Figure 3-11.

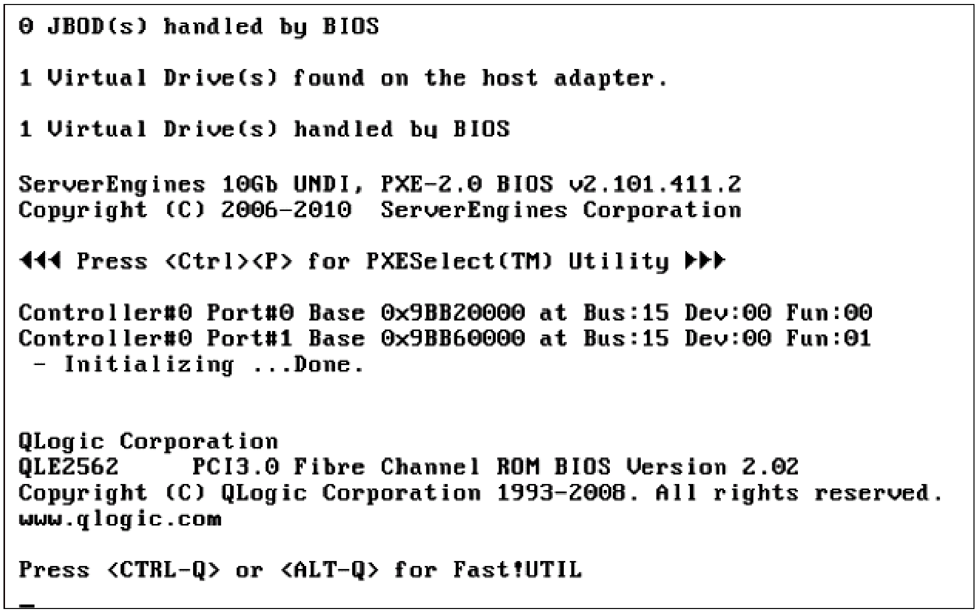


Figure 3-11 Enter QLogic' Fast!Util

4. Select the adapter port to be zoned to storage, as shown in Figure 3-12.

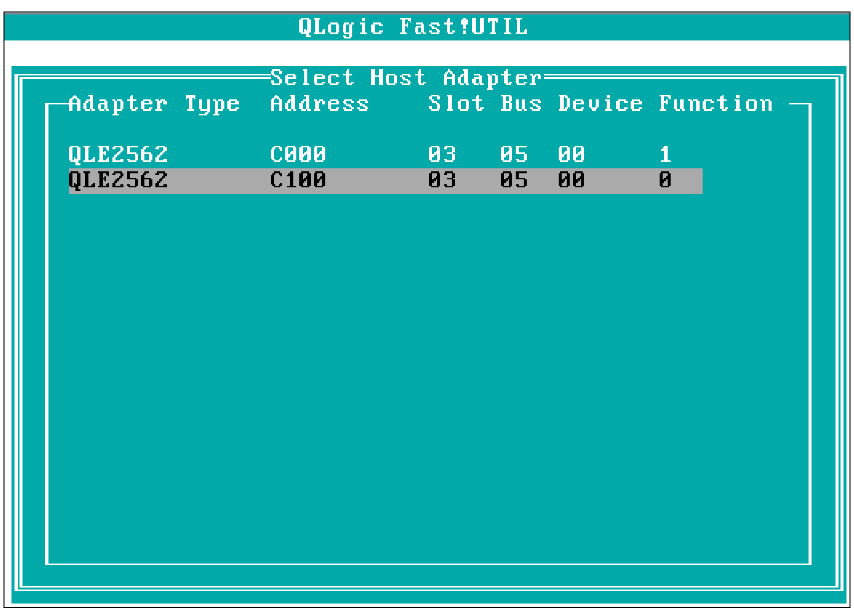


Figure 3-12 Select QLogic Adapter

5. Select **Configuration Settings**, as shown in Figure 3-13.

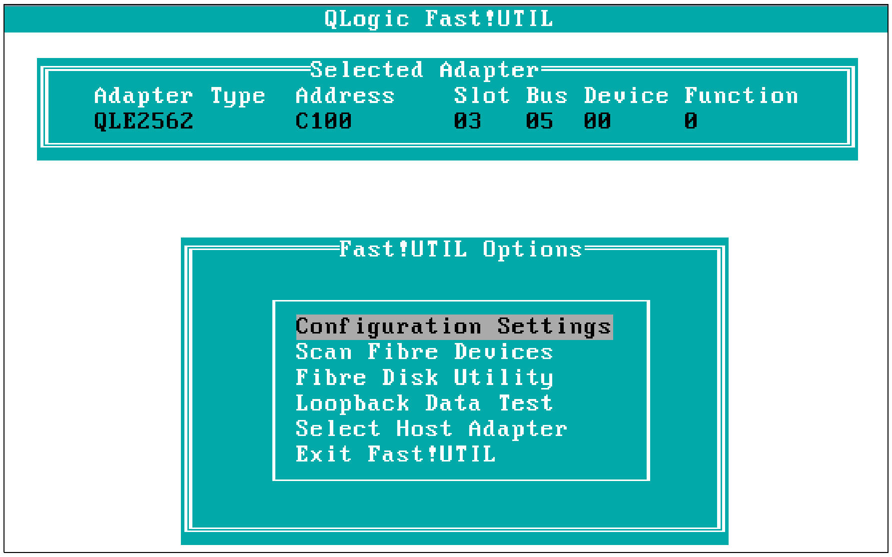


Figure 3-13 Configuration Settings

6. Select **Host Adapter Settings**, as shown in Figure 3-14.

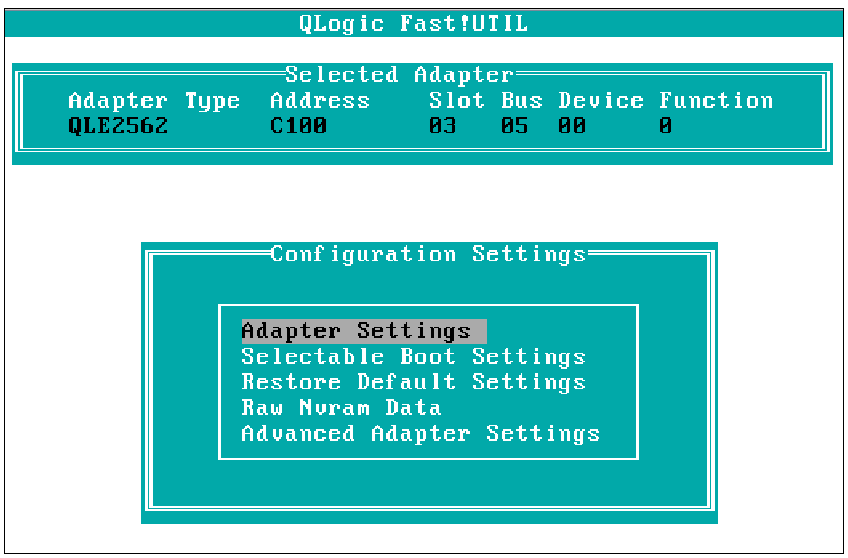


Figure 3-14 Adapter Settings

7. Make a note of the Adapter Worldwide Port Name (WWPN), as shown in Figure 3-15.

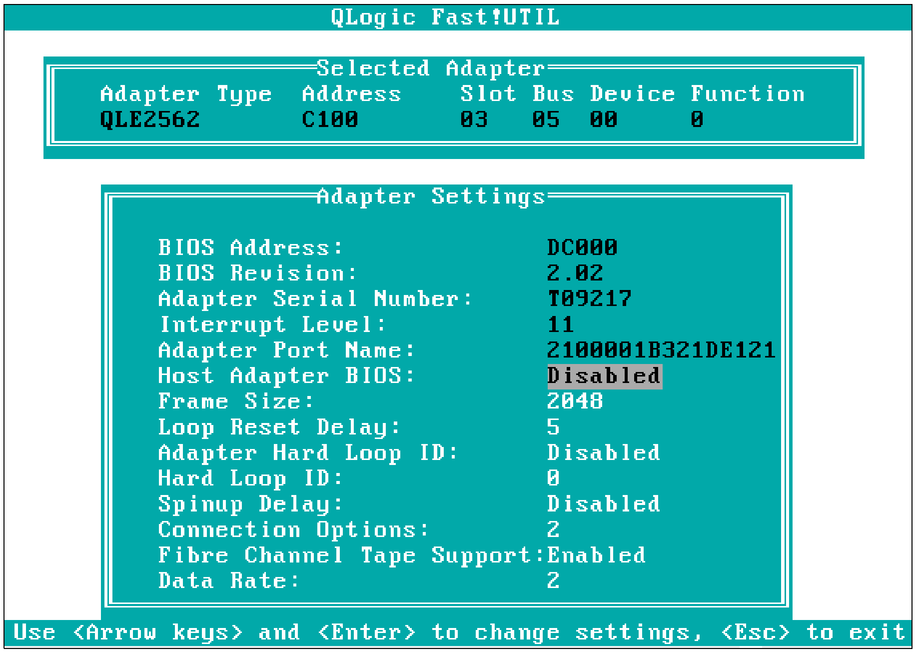


Figure 3-15 Note WWPN

Now you have the WWPN for the HBA port to be zoned to the host port on storage.

8. Assuming that you have created hosts (as described in 3.5.2, “Configuring the LUN for SAN Boot”), you can find the WWPN of the host port on the IBM DS8000 Subsystem in the Management GUI by following the steps in 3.8, “Installation of SUSE Linux Enterprise Server 11 SP 1 for SAN Boot”.

3.4.5 Configuring QLogic adapter for SAN Boot

Follow this procedure for configuring the QLogic adapters:

1. Power up the server containing the HBA.
2. Follow steps 1 through 4 of the section 3.4.4, “Finding the QLogic WWPN” on page 135.
3. Attach the cables and configure the zones.
4. Configure your host with the WWPN associated with your card, as described in 3.6, “DS8000 Subsystem host creation” on page 155.
5. Create and assign the LUN for your host, as described in 3.5, “DS8000 storage provisioning” on page 150.

6. From the Configuration Settings panel (Figure 3-13 on page 136) of your host, select **Scan Fibre Devices**. You should see the DS8000 Subsystem host-port WWPN, as shown in Figure 3-21 on page 140.

   However, if you are not able to find the LUN, as shown here in Figure 3-16, recheck your cables, zones, and the configuration for your host on the DS8000 management interface.

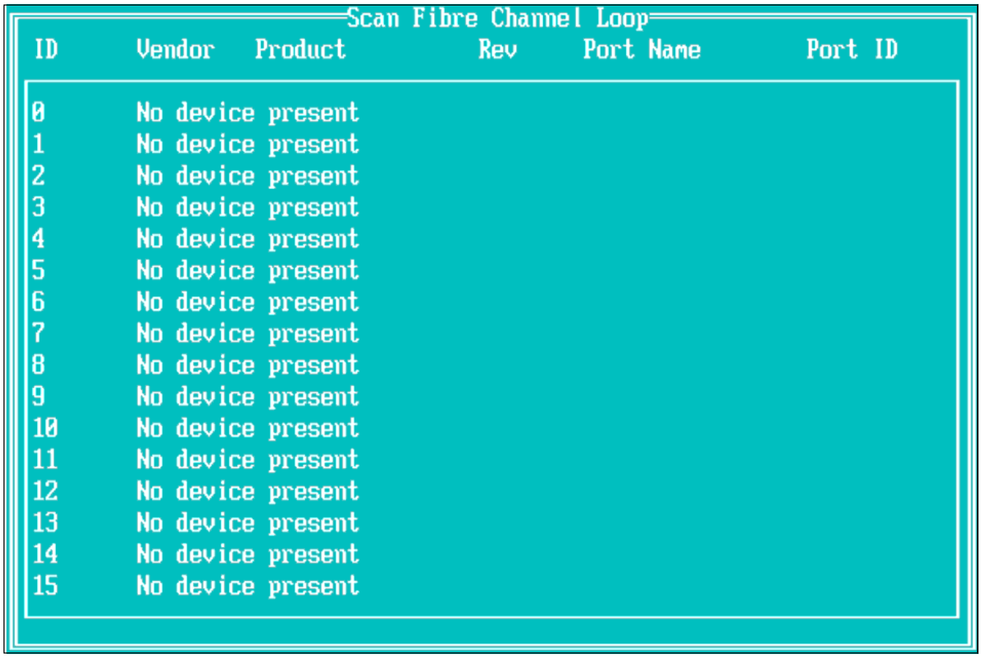


Figure 3-16 No device present yet

Tip: If you do not see any devices upon scanning the fibre devices, verify your switch zoning and hardware connectivity, then rescan.

7. When the storage host port's WWPN is found, press <Esc> to exit out to the Fast!Util Options page.
8. Select **Configuration Settings**, then **Adapter Settings**, as shown in Figure 3-17.

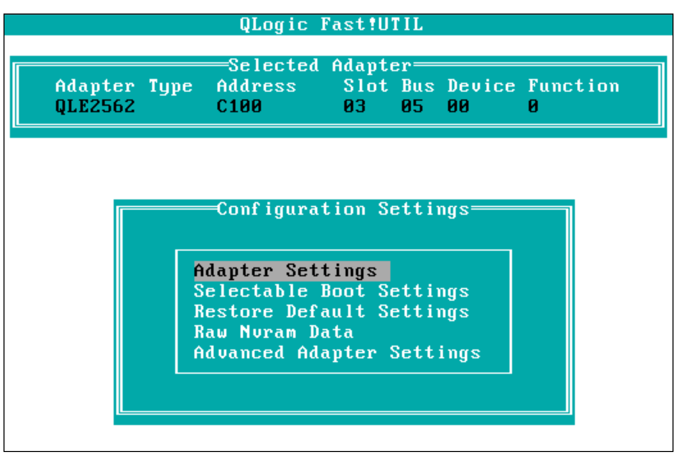


Figure 3-17 Adapter Settings

9. Select **Host Adapter BIOS** and enable it, shown in Figure 3-18.

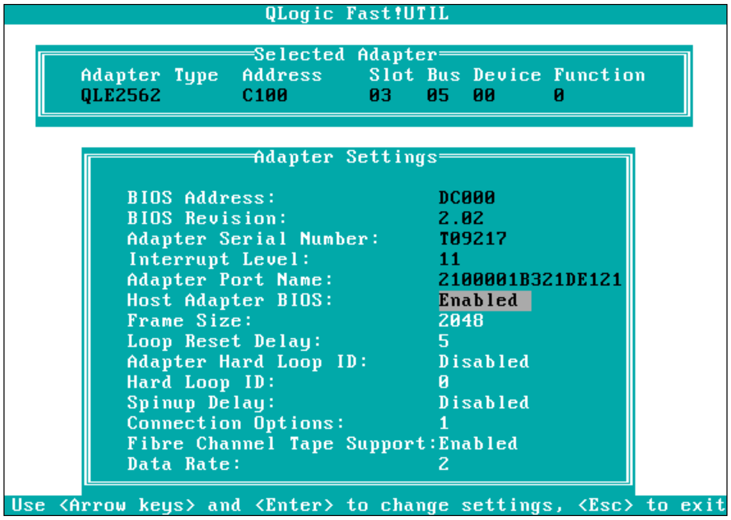


Figure 3-18 Enable BIOS

10. Press the <Esc> key to back up one menu. Next, select **Selectable Boot Settings**, as shown in Figure 3-19.

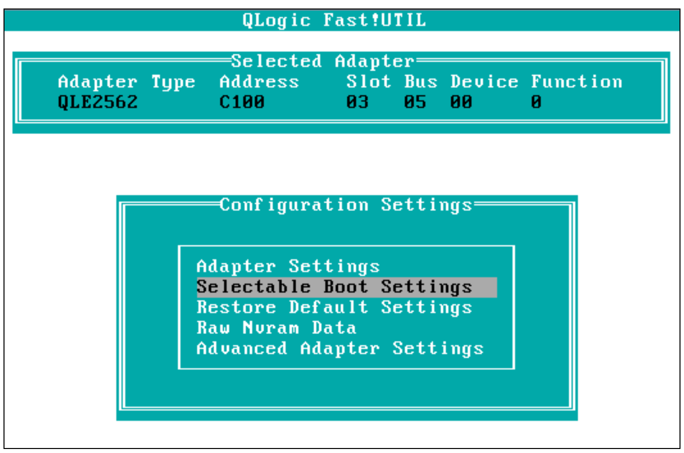


Figure 3-19 Selectable Boot Settings

11. Change Selectable Boot Settings to **Enabled**, as shown in Figure 3-20.

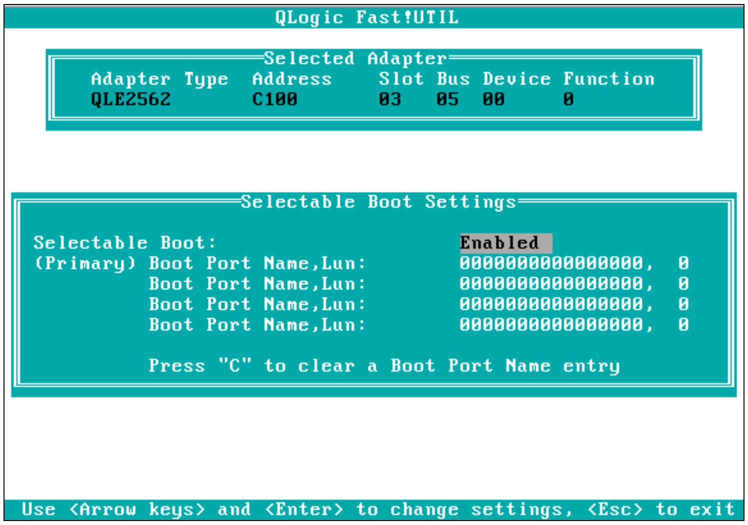


Figure 3-20 Enable Boot

12. Now click the (Primary) Boot Port Name to select the source LUN for SAN Boot, as shown in Figure 3-21. Here, IBM 2107-900 means that it is a DS8000 Subsystem.

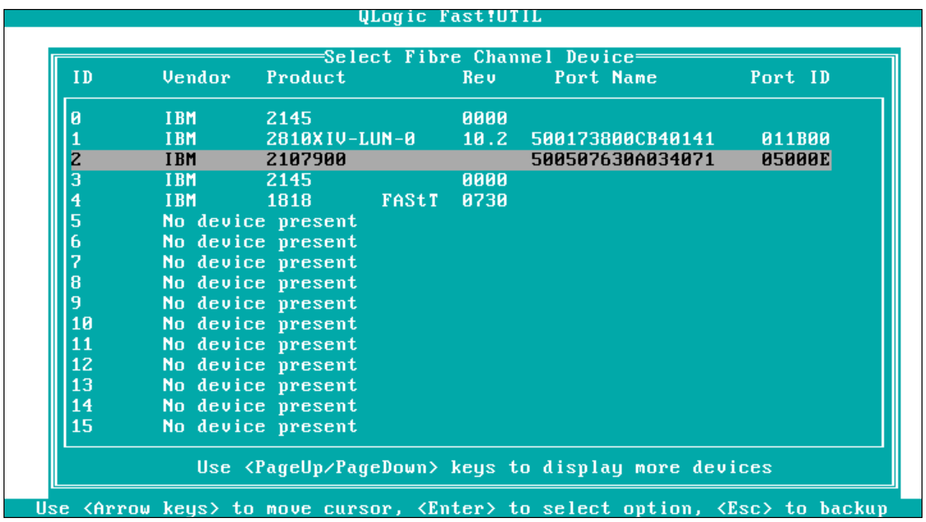


Figure 3-21 Select Boot Volume

Tip: IBM DS8000 Subsystem WWPNs are based on: 50:05:07:63:0A:**YY:Y**x:xx
where x:xx is unique for each DS8000 and the **Y** value is taken from the port position.

13.After selecting the appropriate storage Boot Volume, you come back to the previous panel. It will now have information under (Primary) Boot Port Name, LUN for the device you selected in the previous step, as shown in Figure 3-22.

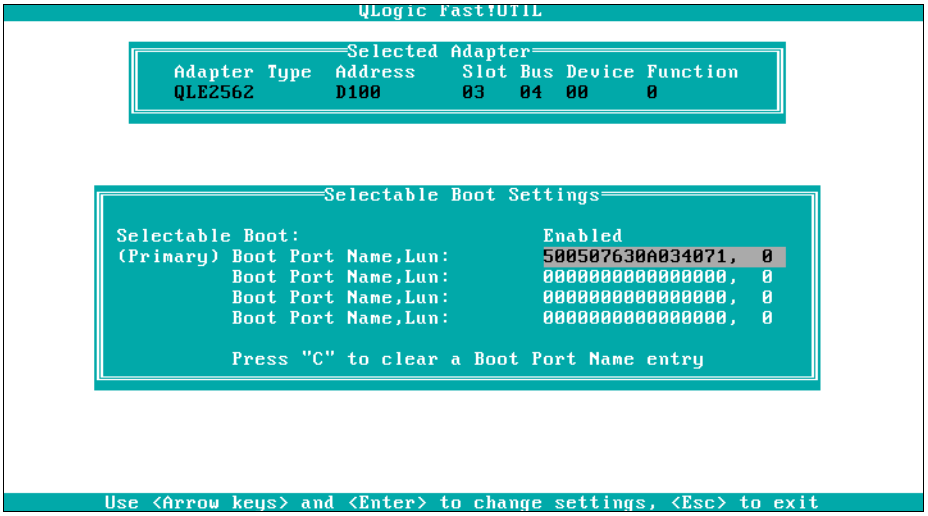


Figure 3-22 Boot LUN

14.Most common operating systems in the market have already the driver to start the installation using the multipath. If your choice matches this requirement, you can repeat the process for the second card (go back to 3.4.5, “Configuring QLogic adapter for SAN Boot” on page 137 ”). Configure both cards to have access to the same LUN and thus avoid the downtime if you lose any of these paths.

15.Press the <Esc> key twice to exit out of the menus and select **Save Changes** to save the changes, as shown in Figure 3-23.

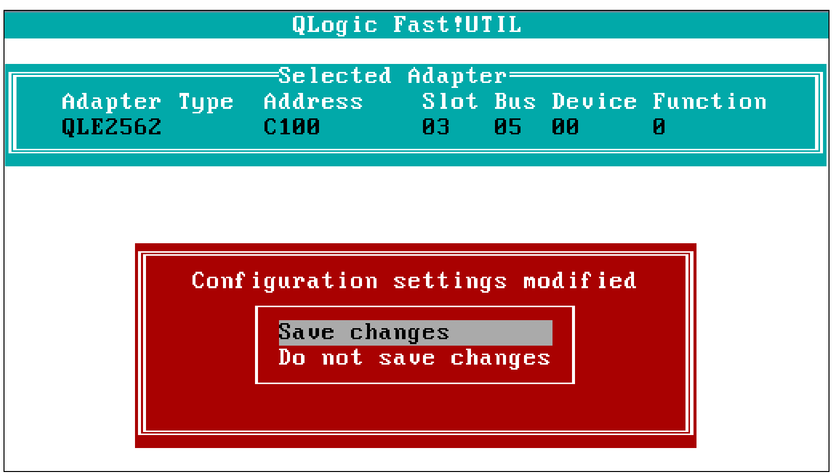


Figure 3-23 Save config. changes

16.Exit **QLogic Fast!UTIL** and reboot the server.

17. During POST, you see QLogic reporting the Boot LUN, as shown in Figure 3-24.

```
HA -0 (Bus 3 Dev 0) ServeRAID M1015 SAS/SATA Controller
FW package: 20.10.1-0022


0 JBOD(s) found on the host adapter
0 JBOD(s) handled by BIOS

0 Virtual Drive(s) found on the host adapter.

0 Virtual Drive(s) handled by BIOS


QLogic Corporation
QLE2562     PCI Fibre Channel ROM BIOS Version 2.02
Copyright (C) QLogic Corporation 1993-2008. All rights reserved.
www.qlogic.com

Press <CTRL-Q> or <ALT-Q> for Fast!UTIL
Firmware Version 4.03.01

Device Device   Adapter Port  Lun    Vendor    Product          Product
Number Type     Number   ID   Number   ID        ID             Revision
       Disk       0    05000E  0     IBM       2107900
ROM BIOS Installed
```

Figure 3-24 QLogic reporting boot LUN from Storage

18. Press F12 to select the boot device, then select the CD/DVD ROM (or appropriate source) to install the operating system, as shown in Figure 3-25 and Figure 3-26.

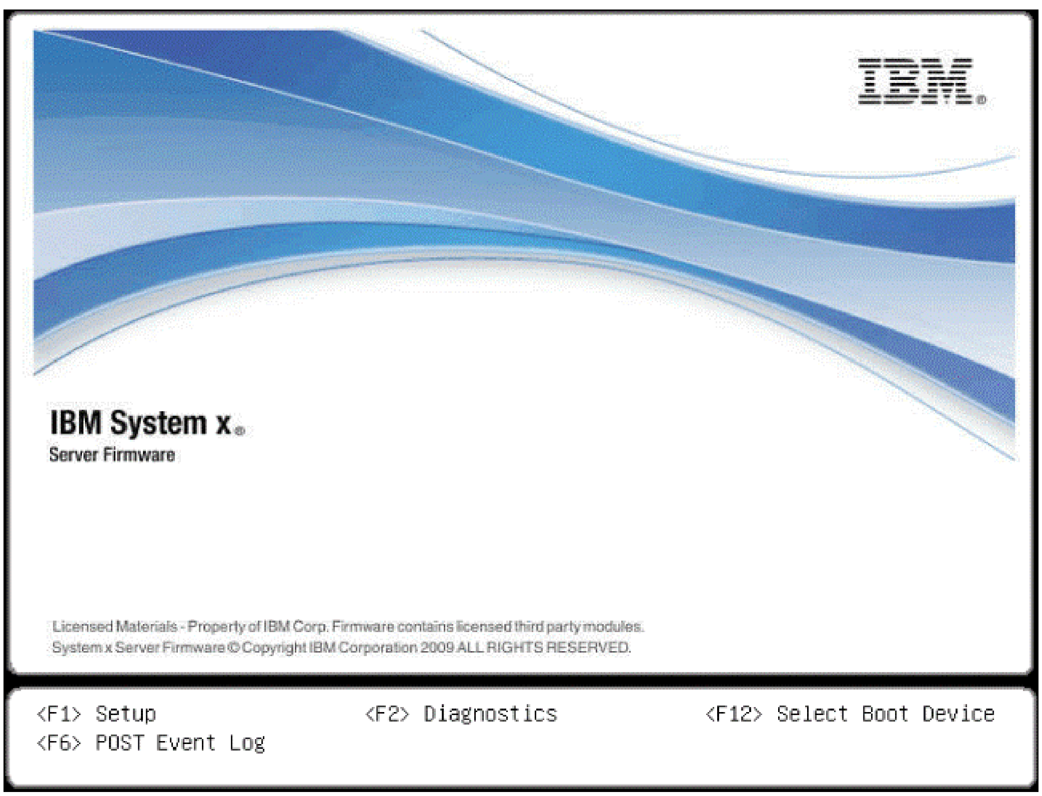


Figure 3-25 F12 to select Boot Device

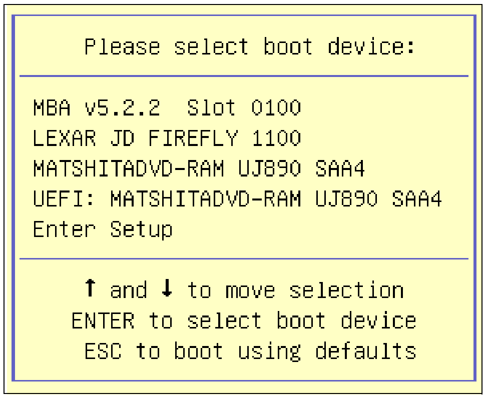


Figure 3-26 Select DVD to install OS

19. Assuming that this installation is from DVD, insert the media and follow the installation instructions outlined here:
    - Windows Installation, in 3.7, “Installation of Microsoft Windows 2008 using SAN Boot” on page 156,
    - SUSE Linux Enterprise Server 11, in 3.8, “Installation of SUSE Linux Enterprise Server 11 SP 1 for SAN Boot” on page 162,
    - Red Hat Linux 5.5, in 3.9, “Installation of Red Hat Enterprise Server 5.5 for SAN Boot” on page 173
    - ESX vSphere 4.1, in 3.11, “Installation of VMware ESX 4.1 for SAN Boot” on page 196

3.4.6 Configuring Emulex adapter for SAN Boot

To configure Emulex adapters for SAN Boot, perform the following steps:

1. Power up the server containing the HBA.
2. Attach the cables and configure the zones.
3. Configure your host with the WWPN associated with your card, as described in 3.6, “DS8000 Subsystem host creation” on page 155.
4. Create and assign the LUN for your host, as described in 3.5, “DS8000 storage provisioning” on page 150.

5. Press Ctrl-E or Alt-E to access the Emulex configuration utility, as shown in Figure 3-27.

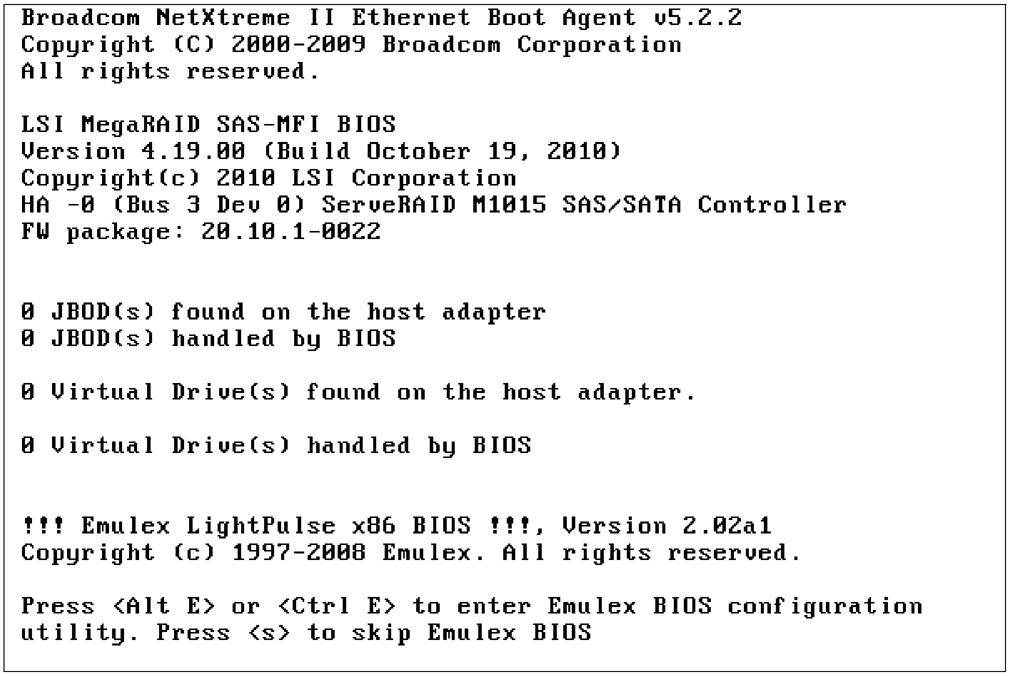


Figure 3-27 Alt E for Emulex HBA Util

6. Select the HBA port for Boot LUN settings, as shown in Figure 3-28.

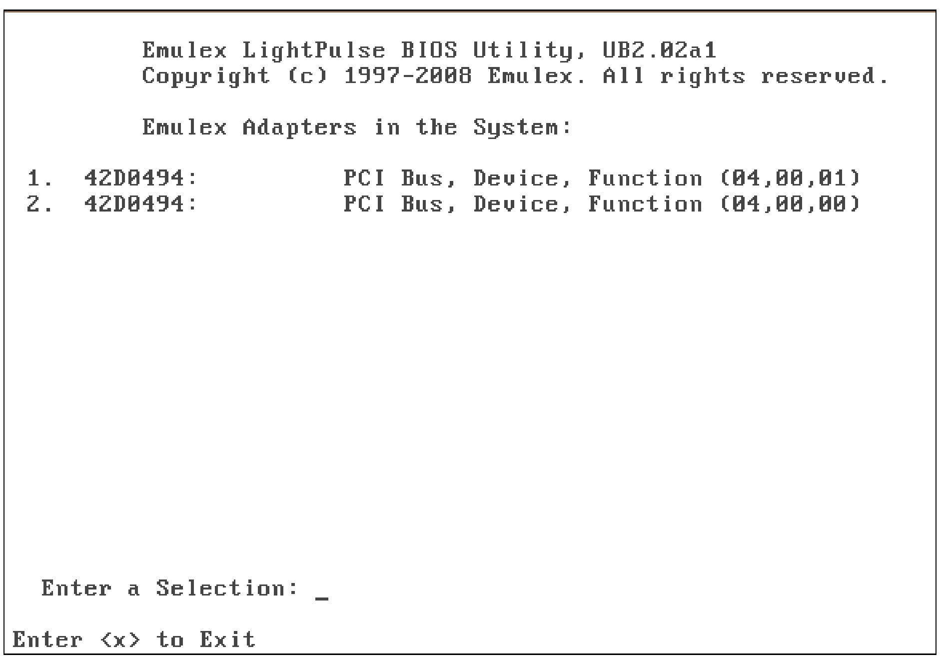


Figure 3-28 Select HBA Port

7. Select **Configure This Adapter's Parameters**, as shown in Figure 3-29.

```
    Adapter 01:           PCI Bus, Device, Function (04,00,01)

    42D0494:       Mem Base: FCA80000  Firmware Version: US1.00A12
    Port Name: 10000000 C98E438B   Node Name: 20000000 C98E438B
    Topology: Auto Topology: Loop First (Default)
    The BIOS for this adapter is Enabled

1.  Configure Boot Devices
2.  Configure This Adapter's Parameters

  Enter a Selection: _

Enter <x> to Exit    <d> to Default Values     <Esc> to Previous Menu
```

Figure 3-29 Configure HBA

8. Select Option **1** to enable BIOS for this HBA, as shown in Figure 3-30.

```
    Adapter 01:           PCI Bus, Device, Function (04,00,01)

    42D0494:       Mem Base: FCA80000  Firmware Version: US1.00A12
    Port Name: 10000000 C98E438B   Node Name: 20000000 C98E438B
    Topology: Auto Topology: Loop First (Default)
    The BIOS for this adapter is Enabled

1.  Enable or Disable BIOS
2.  Change Default ALPA of this Adapter
3.  Change PLOGI Retry Timer (+Advanced Option+)
4.  Topology Selection (+Advanced Option+)
5.  Enable or Disable Spinup Delay (+Advanced Option+)
6.  Auto Scan Setting (+Advanced Option+)
7.  Enable or Disable EDD 3.0 (+Advanced Option+)
8.  Enable or Disable Start Unit Command (+Advanced Option+)
9.  Enable or Disable Environment Variable (+Advanced Option+)
10. Enable or Disable Auto Boot Sector (+Advanced Option+)
11. Link Speed Selection (+Advanced Option+)

  Enter a Selection:

Enter <x> to Exit          <Esc> to Previous Menu
```

Figure 3-30 Enable BIOS

Figure 3-31 shows the option to enable or disable BIOS by pressing 1 or 2 appropriately. Here we enable the BIOS.

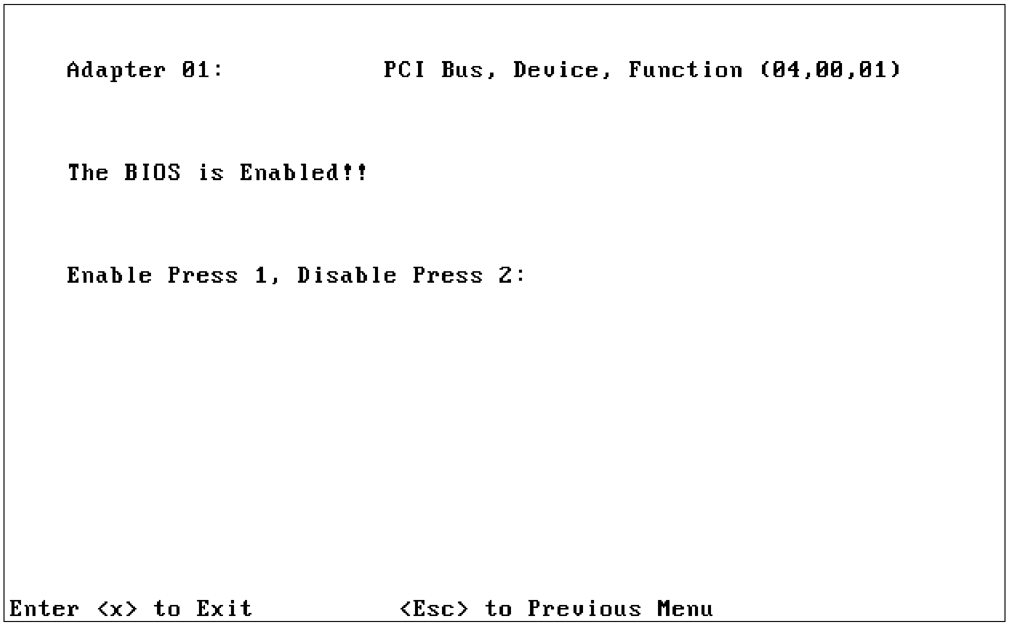


Figure 3-31 BIOS enabled

9. Press <Esc> to go back. Then select **Configure Boot Devices**.
10. Enter Boot Entry number of the first Unused boot device (Primary Boot), from the List of Saved Boot Devices, as shown in Figure 3-32.

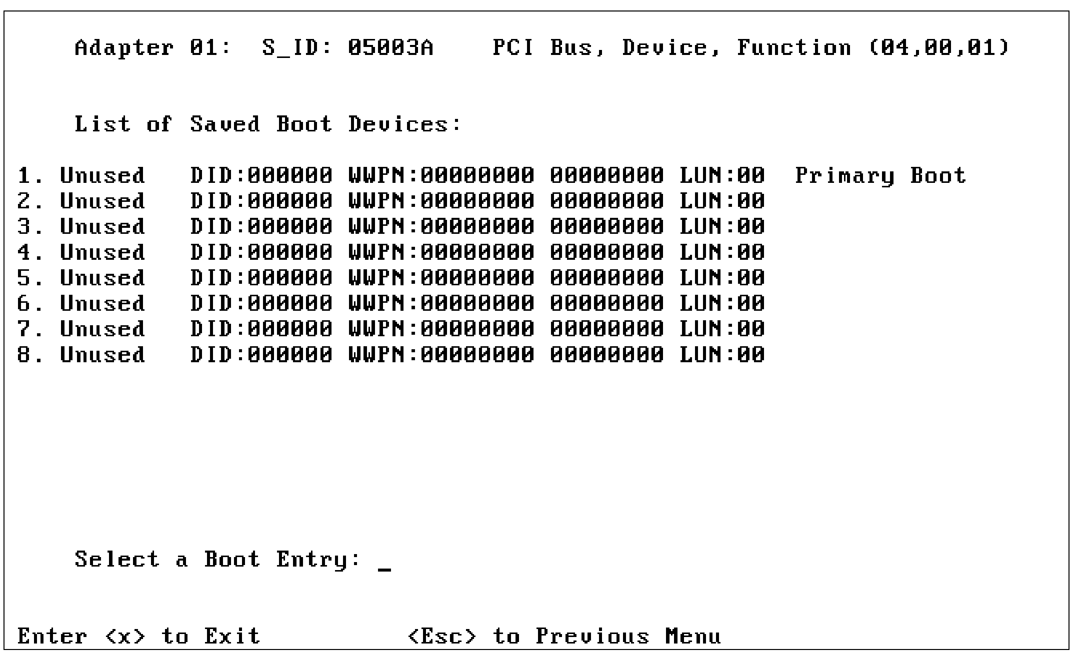


Figure 3-32 Saved Boot Devices

11. The HBA will scan the devices it can see. If you added this host port's WWPN to DS8000 Subsystem host mappings, it will show the WWPN of the hostport on DS8000 Subsystem, as shown in Figure 3-33. If you do not see the DS8000, as shown in Figure 3-33, recheck your cables, zones, and the configuration for your host on the DS8000 management interface.

```
    Adapter 02:  S_ID: 050037    PCI Bus, Device, Function (04,00,00)

00. Clear selected boot entry!!
01. DID:05000E WWPN:50050763 0A034071 LUN:00    IBM     2107900
02. DID:050500 WWPN:201600A0 B8473932 LUN:1F    IBM     Universal Xport 0730




  Select The Two Digit Number of The Desired Boot Device:_

Enter <x> to Exit         <Esc> to Previous Menu       <PageDn> to Next Page
```

Figure 3-33 Scan Fibre Devices

12. Select **01** for "Select The Two Digit Number Of The Desired Boot Device" (**01** in our case). Then it will prompt "Enter two digits of starting LUN (Hex). In our case, it is **00**, as shown in Figure 3-34.

```
    Adapter 02:  S_ID: 050037    PCI Bus, Device, Function (04,00,00)

00. Clear selected boot entry!!
01. DID:05000E WWPN:50050763 0A034071 LUN:00    IBM     2107900
02. DID:050500 WWPN:201600A0 B8473932 LUN:1F    IBM     Universal Xport 0730

                DID:05000E WWPN:50050763 0A034071

             Enter two digits of starting LUN (Hex):00

                     <Esc> to Previous Menu

  Select The Two Digit Number of The Desired Boot Device:01

Enter <x> to Exit           <Esc> to Previous Menu     <PageDn> to Next Page
```

Figure 3-34 Enter digits of starting LUN

13.A new panel will ask you to select the LUN and the boot mode as shown in Figure 3-35. The options for booting the device are by WWPN, or by DID. We choose **1** to boot by WWPN, as shown in Figure 3-35.

```
    Adapter 02:  S_ID: 050037    PCI Bus, Device, Function (04,00,00)

    DID:05000E WWPN:50050763 0A034071

01.      LUN:00              IBM     2107900




                    DID:05000E WWPN:50050763 0A034071 LUN:00

                          1. Boot this device via WWPN
                          2. Boot this device via DID

                            <Esc> to Previous Menu
                              Enter a Selection:  1

  Enter a Selection: 01
  B#W: Boot number via WWPN. B#D: Boot number via DID
Enter <x> to Exit          <Esc> to Previous Menu
```

Figure 3-35 Emulex boot mode selection.

14.Now the LUN, on DS8000 Subsystem, shows up in the List of Saved Boot Devices, as shown in Figure 3-36.

```
    Adapter 02:  S_ID: 050037    PCI Bus, Device, Function (04,00,00)


    List of Saved Boot Devices:

1. Used     DID:000000 WWPN:50050763 0A034071 LUN:00  Primary Boot
2. Unused   DID:000000 WWPN:00000000 00000000 LUN:00
3. Unused   DID:000000 WWPN:00000000 00000000 LUN:00
4. Unused   DID:000000 WWPN:00000000 00000000 LUN:00
5. Unused   DID:000000 WWPN:00000000 00000000 LUN:00
6. Unused   DID:000000 WWPN:00000000 00000000 LUN:00
7. Unused   DID:000000 WWPN:00000000 00000000 LUN:00
8. Unused   DID:000000 WWPN:00000000 00000000 LUN:00






    Select a Boot Entry: _


Enter <x> to Exit          <Esc> to Previous Menu
```

Figure 3-36 Boot LUN's WWPN

15.Most common operating systems (OS) in the market already have the driver to start the installation using multipath. If you are using such an OS, you can repeat the process for the second card (starting with Step 6 on page 144). You configure both the cards to have access to the same LUN; thus you can avoid the downtime if you lose any of these paths.

16.Enter <x> to exit and save changes, which will reboot the server.

17. During post, after the Emulex banner, you will see the DS8000 Subsystem being seen by the HBA, as shown in Figure 3-37.

```
0 JBOD(s) found on the host adapter
0 JBOD(s) handled by BIOS

0 Virtual Drive(s) found on the host adapter.

0 Virtual Drive(s) handled by BIOS


!!! Emulex LightPulse x86 BIOS !!!, Version 2.11a0
Copyright (c) 1997-2008 Emulex. All rights reserved.

Press <Alt E> or <Ctrl E> to enter Emulex BIOS configuration
utility. Press <s> to skip Emulex BIOS

Installing Emulex BIOS ......
Bringing the Link up, Please wait...
Bringing the Link up, Please wait...
--Adapter 1 42D0494:        S_ID:050038  PCI Bus, Device, Function (04,00,01)
D_ID:   05001A  LUN: 00 IBM     2107900         .432
--Adapter 2 42D0494:        S_ID:050037  PCI Bus, Device, Function (04,00,00)
D_ID:   05000E  LUN: 00 IBM     2107900         .432

Emulex BIOS is installed successfully!!!
```

Figure 3-37 Emulex seeing SAN Boot volume

18. Press F12 to select the boot device, then select the CD/DVD ROM (or appropriate source) to install the operating system, as shown in Figure 3-38 and Figure 3-39.

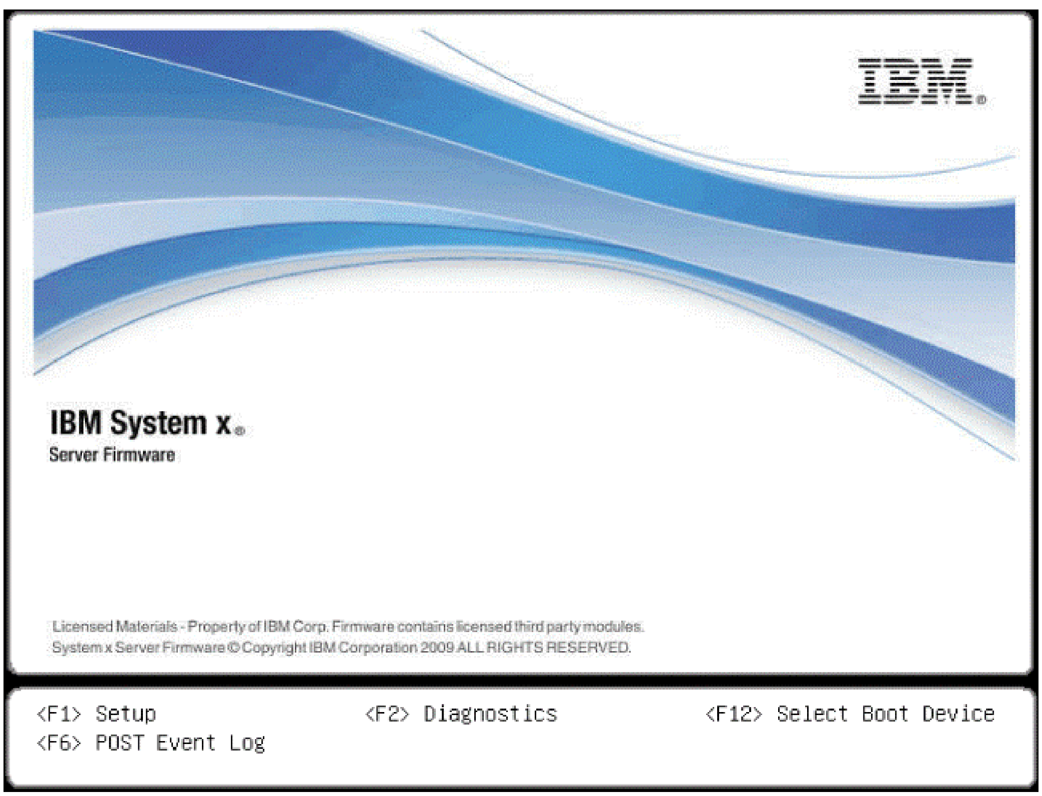


Figure 3-38 F12 to select Boot Device

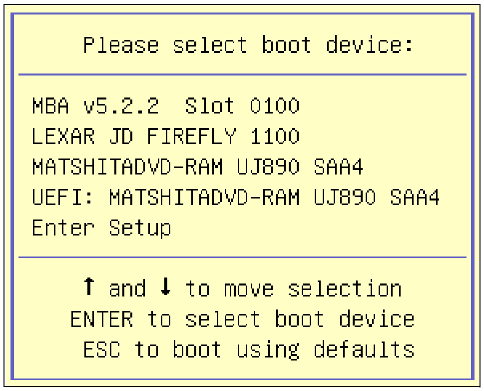


Figure 3-39 Select DVD to install OS

19. Assuming this installation is from DVD, insert the media and follow the installation instructions outlined here:
   - Windows installation, in 3.7, “Installation of Microsoft Windows 2008 using SAN Boot” on page 156
   - SUSE Linux Enterprise Server 11, in 3.8, “Installation of SUSE Linux Enterprise Server 11 SP 1 for SAN Boot” on page 162
   - Red Hat Linux 5.5, in 3.9, “Installation of Red Hat Enterprise Server 5.5 for SAN Boot” on page 173
   - ESX vSphere 4.1, in 3.11, “Installation of VMware ESX 4.1 for SAN Boot” on page 196

3.5 DS8000 storage provisioning

Here we provide an overview of DS8000 storage subsystem terminology, then describe the DS8000 storage provisioning procedures. For detailed instructions for all the topics described here, see *IBM System Storage DS8800: Architecture and Implementation*, SG24-8886.

3.5.1 Overview

Here we describe the various DS8000 Storage provisioning features.

Storage complex

A storage complex is a group of DS8000s managed by a single S-HMC (Storage Hardware Management Console). It can consist of a single DS8000 Storage unit. A storage complex is sometimes referred to as a storage-plex.

Storage facility

Storage facility refers to a single DS8000 unit (including the base frame and the optional expansion frames). A storage facility is also referred to as a storage unit. As an example, if your organization has one DS8000, then you have a single storage complex that contains a single storage unit.

Processor complex

Processor complex refers to one of the Power5 servers which runs, controls, and delivers all the services of the DS8000. There are two processor complexes in one storage facility: processor complex 0 and processor complex 1. Each processor complex can support one or more LPARs concurrently.

Logical partition (LPAR)

An LPAR uses software and firmware to logically partition the resources on a system. An LPAR consists of processors, memory, and I/O slots available in one processor complex.

Array site

An array site is a group of 8 DDMs selected by the DS8000 server algorithm in a storage facility image. An array site is managed by one storage facility image.

Array

Each array site can be individually formatted by the user to a specific RAID format. A formatted array site is called an array. The supported RAID formats are RAID-5 and RAID-10. The process of selecting the RAID format for an array is also called defining an array.

Rank

A rank is defined by the user. The user selects an array and defines the storage format for the rank, which is either Count Key Data (CKD) or Fixed Block (FB) data. One rank will be assigned to one extent pool by the user.

Extents

The available space on each rank is divided into extents. The extents are the building blocks of the logical volumes. The characteristic of the extent is its size, which depends on the specified device type when defining a rank:

- For fixed block format, the extent size is 1 GB.
- For CKD format, the extent size is .94 GB for model 1.

Extent pools

An extent pool refers to a logical construct to manage a set of extents. The user defines extent pools by selecting one to N ranks managed by one storage facility image. The user defines which storage facility image server (Server 0 or Server 1) will manage the extent pool. All extents in an extent pool must be of the same storage type (CKD or FB). Extents in an extent pool can come from ranks defined with arrays of different RAID formats, but the same RAID configuration within an extent pool is recommended. The minimum number of extent pools in a storage facility image is two (each storage facility image server manages a minimum of one extent pool).

Rank groups

Ranks are organized in two rank groups:

- Rank group 0 is controlled by server 0.
- Rank group 1 is controlled by server 1.

Logical volume

A logical volume is composed of a set of extents from one extent pool.

- A logical volume composed of fixed block extents is called a LUN.
- A logical volume composed of CKD extents is referred to as a CKD volume or logical device.

Logical subsystem

A logical subsystem (LSS) is a logical construct grouping logical volumes. One LSS can group up to 256 logical volumes from extent pools. The user can define up to 255 LSSs in a storage facility image with the following restriction: the logical volumes in one LSS must be of extent pools with identical extent types and from the same rank pool in one storage facility image. As a result, LSSs are either CKD or FB and have affinity with one storage facility image server. Up to 128 LSSs can be managed by Server 0 and up to 127 LSSs can be managed by Server 1 (one LSS address is reserved).

Address group

An address group refers to a group of LSSs. Up to 16 LSSs can be grouped into one address group. All LSSs in an address group must be of the same format (CKD or FB). The address groups are defined by the user. A storage facility image can manage up to 16 address groups.

Host attachment

One host attachment is a named group of World Wide Port Names (WWPNs) defined by the user. The definition of host attachment is necessary to manage the LUN masking. One WWPN can be defined in only one host attachment. The user assigns one host attachment to one volume group. Each WWPN in the host attachment will get access to all of the LUNs defined in the volume group.

Volume group

The user gathers LUNs into volume groups. The definition of volume groups is necessary to manage the LUN masking. One LUN can be defined in several volume groups. One volume group can be assigned to several host attachments.

3.5.2 Configuring the LUN for SAN Boot

In this section, we explain the process to create the LUNs on DS8000 storage subsystem.

> **Tip:** We are not explaining the whole process to install, configure, or format the DS8000, only the final process to create and assign the LUN to be used for SAN Boot process.
>
> For complete procedures, see *IBM System Storage DS8800: Architecture and Implementation*, SG24-8886.

Follow these steps:

1. First, access the storage using your preferred management tool, with admin or a user with administrative rights. In our case, we use the DSCLI interface, as shown in Figure 3-40.

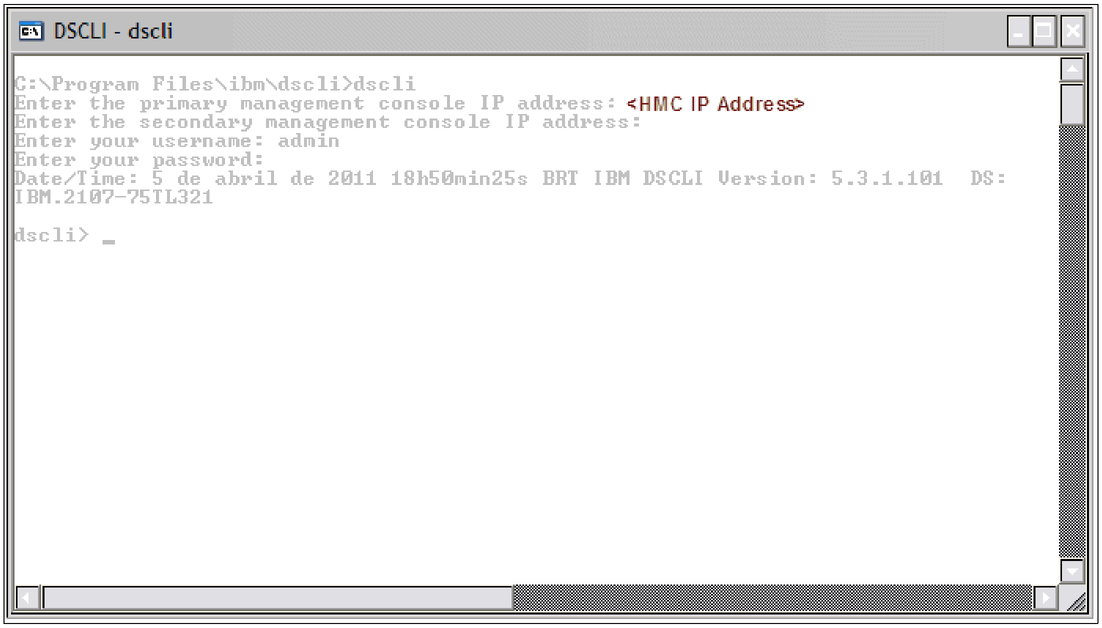


Figure 3-40 DSCLI interface to access DS8000

2. Select the pool with required space to be used for SAN Boot LUNs as shown in Figure 3-41.

```
dscli> lsextpool -dev IBM.2107-75TL321
Date/Time: 5 de abril de 2011 18h57min27s BRT IBM DSCLI Version: 5.3.1.101 DS:
IBM.2107-75TL321
Name              ID stgtype rankgrp status availstor (2^30B) %allocated
available reserved numvols
===============================================================================
====================
svc_mdisk_sanboot P0 fb            0  below            22277          4
22277         0       4
svc_vdisk_io      P1 fb            1  below            19369         17
19369         0     228
host_sanboot      P2 fb            0  below             9431          6
9431         0      19
host_io_lun       P3 fb            1  below            19565          2
19565         0      91
ckd_sanboot       P4 ckd           0  below             3130          0
3553         0       0
dscli>
```

Figure 3-41 Select the pool to create the SAN Boot LUNs

3. Create the Volume Group to receive the LUN for Hosts using the **mkvolgrp** command, as shown in Figure 3-42.

```
dscli> mkvolgrp -dev IBM.2107-75TL321 RB_ST-1L05-R01_VG1
Date/Time: 5 de abril de 2011 20h6min9s BRT IBM DSCLI Version: 5.3.1.101 DS:
IBM.2107-75TL321
CMUC00030I mkvolgrp: Volume group V43 successfully created.
dscli>
```

Figure 3-42 Command to create the Volume Group

4. Check for the next free Volume_ID to be used for SAN Boot LUN ID to be created using the **lsfbvol** command, as shown in Figure 3-43.

```
dscli> lsfbvol -dev IBM.2107-75TL321 -extpool p2
Date/Time: 5 de abril de 2011 19h18min16s BRT IBM DSCLI Version: 5.3.1.101 DS:
IBM.2107-75TL321
Name            ID   accstate datastate configstate deviceMTM datatype extpool
cap (2^30B) cap (10^9B) cap (blocks)
===============================================================================
====================================
g01hlda8_2107c  2006 Online   Normal    Normal      2107-900  FB 512   P2
80.0           -    167772160
dscli>
```

Figure 3-43 Command to list the next available ID for your pool

5. Create the next LUN ID (0010) available for your Extent Pool assigning directly for the Volume Group already created for the server as shown in Figure 3-44.

```
dscli> mkfbvol -dev IBM.2107-75TL321 -extpool p2 -cap 40 -name RB_ST1L05R01_01
-type ds -volgrp v43 0010
Date/Time: 5 de abril de 2011 19h22min45s BRT IBM DSCLI Version: 5.3.1.101 DS:
IBM.2107-75TL321
CMUC00025I mkfbvol: FB volume 0010 successfully created.
dscli>
```

Figure 3-44 Command to create the LUN

Now you are ready to create the host connectivity at DS8000 Subsystem, as described in 3.6, “DS8000 Subsystem host creation” on page 155.

Tip: For more detailed information about commands and additional parameters, access the IBM Systems Information Centers at this website:

http://publib.boulder.ibm.com/eserver/

3.6 DS8000 Subsystem host creation

Initially, you need to check if your zones were created as expected, or collect the WorldWide Port Name (WWPN) to be used for DS8000 Subsystem host configuration.

1. Log into the DS8000 storage subsystem preferred interface, in our case, it was the DSCLI.
2. Check if the I/O ports type and topology were configured in the DS8000 Subsystem, as shown in Figure 3-45.

> **Tip:** For open systems, we can use the followed configurations:
>
> - Types:
>   - Fibre Channel-SW: (SW stands for short wave)
>   - Fibre Channel-LW: (LW stands for long wave, 10 KM)
> - Topologies:
>   - FC-AL
>   - SCSI-FCP

```
dscli> lsioport -dev IBM.2107-75TL321
Date/Time: 7 de Abril de 2011 15h36min40s BRT IBM DSCLI Version: 6.6.0.305 DS:
IBM.2107-75TL321
ID    WWPN             State  Type             topo     portgrp
===============================================================
I0000 500507630A000071 Online Fibre Channel-SW SCSI-FCP 0
I0001 500507630A004071 Online Fibre Channel-SW SCSI-FCP 0
I0002 500507630A008071 Online Fibre Channel-SW SCSI-FCP 0
```

Figure 3-45 Output from ioport showing Protocol Mode used

3. Change the configuration at the I/O ports if necessary to use the correct type and topology as shown in Figure 3-46.

```
dscli>setioport -dev IBM.2107-75TL321 -topology scsi-fcp I0111
Date/Time: 7 de Abril de 2011 15h36min40s BRT IBM DSCLI Version: 6.6.0.305 DS:
IBM.2107-75TL321I/O Port I0111 successfully configured.
I/O Port I0111 successfully configured.
dscli>
```

Figure 3-46 Changing the I/O port configuration

4. Start your machine and check if it had the HBA started.
5. Collect the WWPNs in the server to be configured.
6. Configure the new server and interfaces in the DS8000 Subsystem. See Figure 3-47.

```
dscli> mkhostconnect -dev IBM.2107-75TL321 -volgrp v43 -wwname 2100001B329D1682
-profile "IBM pSeries - AIX" P700_RBH_port_1
Date/Time: 7 de Abril de 2011 16h7min6s BRT IBM DSCLI Version: 6.6.0.305 DS:
IBM.2107-75TL321
CMUC00012I mkhostconnect: Host connection 005C successfully created.
dscli> lshostconnect -dev IBM.2107-75TL321 005C
Date/Time: 7 de Abril de 2011 17h15min6s BRT IBM DSCLI Version: 6.6.0.305 DS:
IBM.2107-75TL321
Name            ID   WWPN             HostType Profile           portgrp
volgrpID ESSIOport
================================================================================
============
P700_RBH_port_1 005C 2100001B329D1682 -        IBM pSeries - AIX       0 V43
all
dscli>
```

Figure 3-47 Command to create and list the new hosts

Now you can check if the server finds the new LUNs to be used to install your host.

Tip: For more detailed information about commands and additional parameters, access the IBM Systems Information Centers at this website:

http://publib.boulder.ibm.com/eserver/

3.7 Installation of Microsoft Windows 2008 using SAN Boot

Following are the general steps to set up a Windows host using SAN Boot:

1. Configure the DS8000 Subsystem so that only the boot volume is mapped to the host.
2. Configure the zones on SAN so the host only sees one DS8000 Subsystem system node port.

Support: Multiple paths during installation are *not supported*.

3. Configure and enable the HBA BIOS, as shown in 3.4.5, “Configuring QLogic adapter for SAN Boot” on page 137 and/or 3.4.6, “Configuring Emulex adapter for SAN Boot” on page 143.
4. Install the operating system, selecting the volume configured for SAN Boot (see 3.5.2, “Configuring the LUN for SAN Boot” on page 152) as the partition on which to install.
5. For our Windows installation, we used the Trial Version of Windows Server 2008 R2 with SP1, from this website:

   http://www.microsoft.com/windowsserver2008/en/us/

6. Follow the instructions by Microsoft to install Windows 2008 R2 with SP1, until it reaches the dialog box, asking **Where do you want to install Windows?** as shown in Figure 3-48.

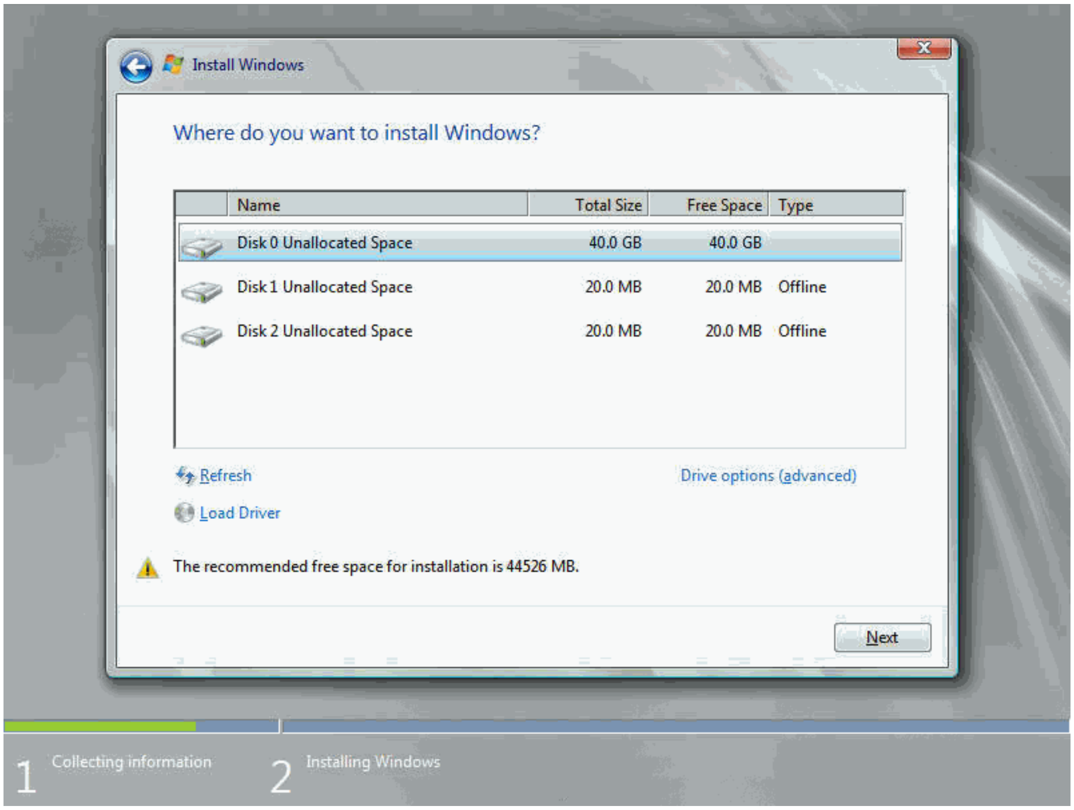


Figure 3-48 Selecting the SAN Volume

7. If you do not see the Volume (size) in the list, it means that Windows did not have the drivers for the HBA. You need to add it manually by clicking the **Load Driver** button at the bottom-left side of the dialogue box, or you probably have issues with the zones.

   Select the appropriate volume and click **Next** to create a new partition, format the drive, and continue with the installation of Windows 2008, as shown in Figure 3-49.

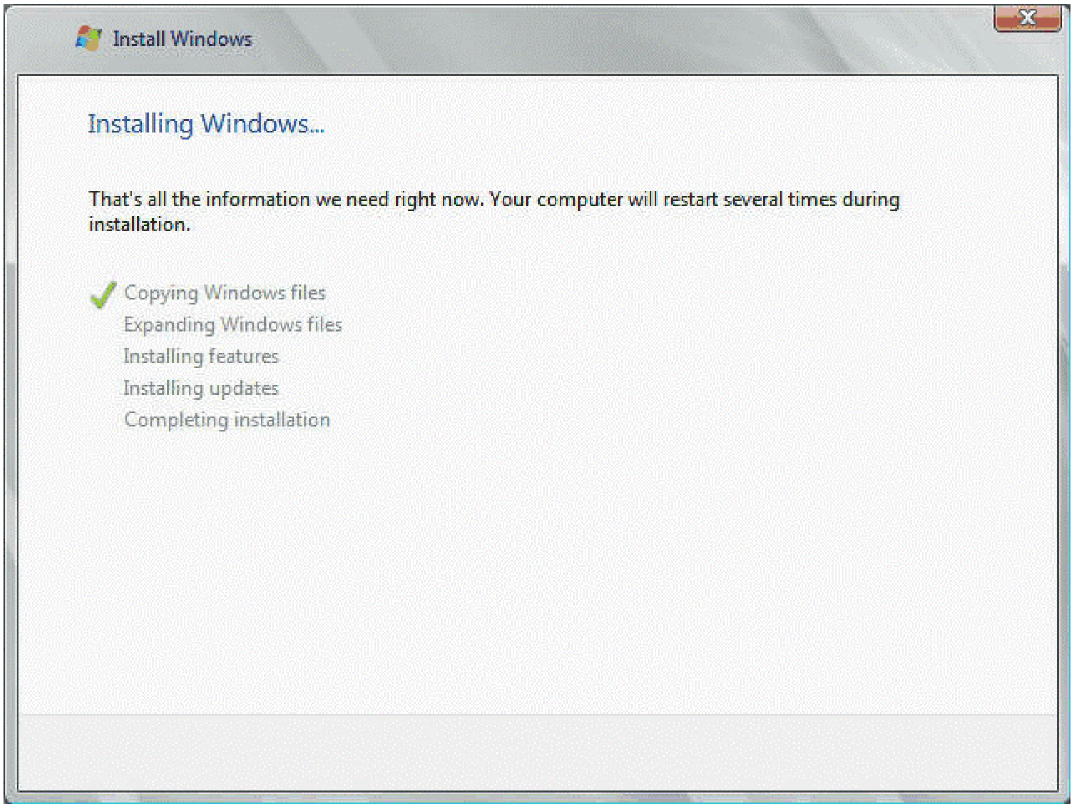


Figure 3-49 Installing Windows 2008

8. After finishing the installation, you can see the attributes of the DS8000 LUN assigned to the Windows server from the device manager, as shown in Figure 3-50.

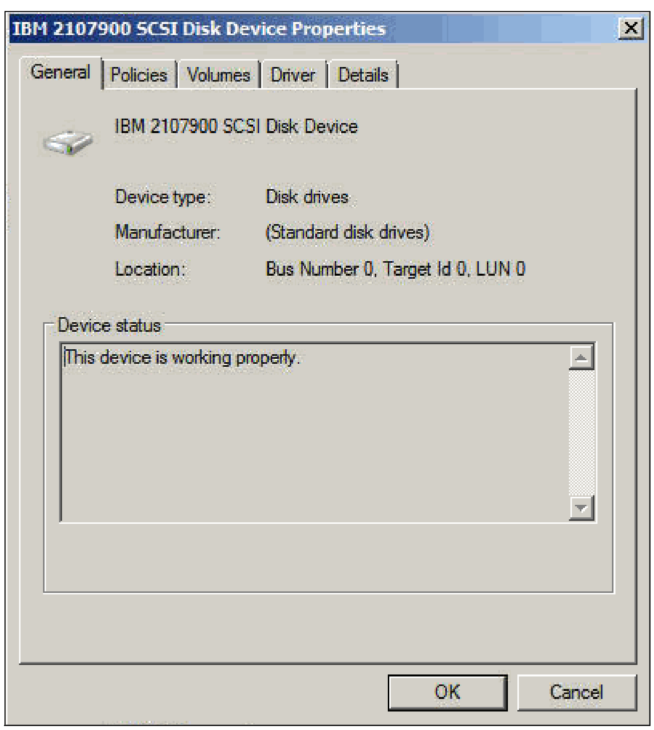


Figure 3-50 Standard Disk

9. After the operating system is installed, check if the Microsoft Multi-path I/O driver is installed or enabled.

> **Important:** Otherwise, the Windows Server 2008 server can lose the ability to access existing data on that device.

10. After ensuring that there is a single connection from your host to your storage, perform the following steps to install SDDDSM on your system from external media:

   a. Obtain the latest Subsystem Device Driver Device Specific Module (SDDDSM) installation package for the Windows Server 2008 operating system from this website:

      http://www-01.ibm.com/support/docview.wss?rs=540&context=ST52G7&dc=D430&uid=ssg1S4000350

   b. Click the link, **SDDDSM Package for IBM DS8000**. Download the appropriate package for your host.

   c. Extract the zip file for the SDDDSM package and run the setup.exe program.

   d. Follow the instructions, as shown in Figure 3-51.

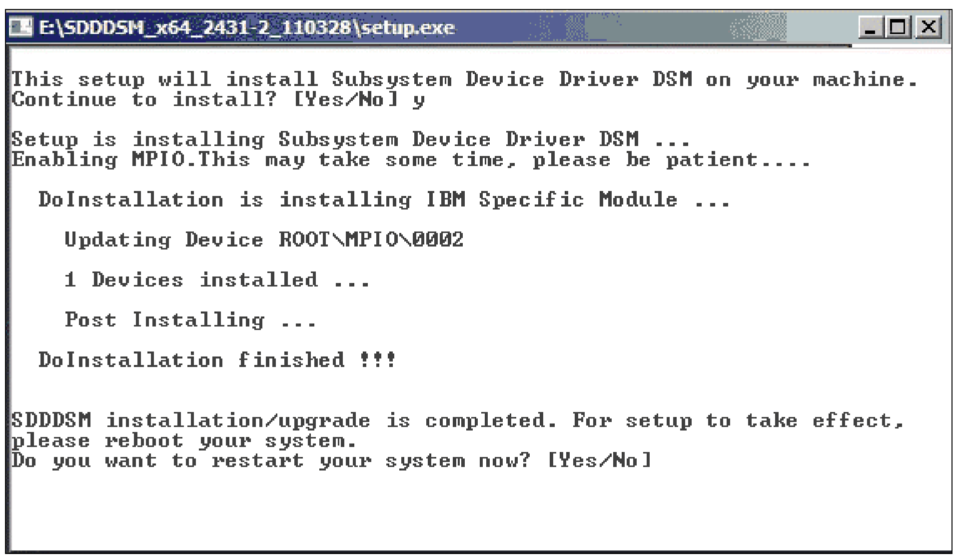


Figure 3-51 Installing SDDDSM

   iii. Configure the second path for your host and reboot the machine.

e. After rebooting the server, check the Disk proprieties, as shown in Figure 3-52.

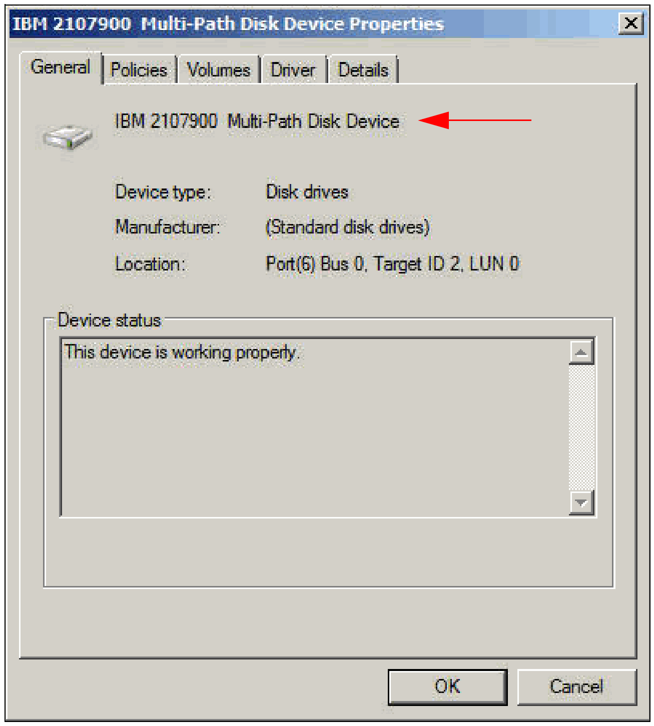


Figure 3-52 Disk with multipath installed

iv. Another way to check the multipath is to use the SDDDSM commands from the command line. Open the prompt and execute **datapath query adapter** and/or **datapath query device** to check the paths as shown in Figure 3-53.

```
C:\Program Files\IBM\SDDDSM>datapath query adapter

Active Adapters :2

Adpt#        Name        State    Mode      Select    Errors   Paths   Active
    0  Scsi Port6 Bus0   NORMAL   ACTIVE     14788         0       1        1
    1  Scsi Port7 Bus0   NORMAL   ACTIVE      1551       591       1        1

C:\Program Files\IBM\SDDDSM>
```

Figure 3-53 SDDDSM multipath verification

3.7.1 SAN Boot restriction

When booting from the FC storage systems, the following restrictions apply:

- With Windows Server 2000, do not use the same HBA as both the FC boot device and the clustering adapter, because the usage of SCSI bus resets MSCS and breaks up disk reservations during quorum arbitration. Because a bus reset cancels all pending I/O operations to all FC disks visible to the host through that port, an MSCS-initiated bus reset can cause operations on the C:\ drive to fail.
- With Windows Server 2003 and 2008, MSCS uses target resets. See the Microsoft technical article, *Microsoft Windows Clustering: Storage Area Networks*, which can be found at the following website:

  http://www.microsoft.com/windowsserver2003/techinfo/overview/san.mspx

- Windows Server 2003 and 2008 allow boot disk and the cluster server disks to be hosted on the same bus. However, you need to use Storport miniport HBA drivers for this functionality to work. It is *not* a supported configuration in combination with drivers of other types (for example, SCSI port miniport or full port drivers).
- If you reboot a system with adapters while the primary path is in a failed state, you must manually disable the BIOS on the first adapter and manually enable the BIOS on the second adapter. You cannot enable the BIOS for both adapters at the same time. If the BIOS for both adapters is enabled at the same time and there is a path failure on the primary adapter, the system stops with an INACCESSIBLE_BOOT_DEVICE error upon reboot.

3.7.2 Best practices for Windows SAN Boot implementation

This section provides some best practice guidelines for Windows SAN Boot implementation.

Booting from a SAN: For information about Microsoft support for booting from a SAN, see the following article:

http://support.microsoft.com/kb/305547

Paging disk

A pagefile is a reserved portion of the hard disk that is used to expand the amount of virtual memory available to applications. Paging is the process of temporarily swapping out the inactive contents of system physical memory to hard disk until those contents are needed again. Because the operating system must have unrestricted access to the pagefile, the pagefile is commonly placed on the same drive as system files. Thus, the C: drive normally includes boot, system, and paging files.

While there is negligible contention between the boot reads and paging writes, there can be considerable resource contention between systems on the SAN. This contention occurs when they are all trying to do paging I/O, or when many systems attempt to boot simultaneously from the same storage port.

One way to lessen resource contention is to separate non-data I/O (such as paging, registry updates, and other boot-related information) from data I/O sources (such as SQL or Exchange). Different scenarios to separate non-data I/O and data I/O are shown in Table 3-2.

Table 3-2 Possible combination for locations for non-data I/O files

Files	Scenario 1	Scenatio 2	Scenario 3
Boot	SAN	SAN	SAN
System	SAN	SAN	Local
Pagefile	SAN	Local	Local

Redundancy: Avoiding a single point of failure

One of the major benefits of SAN adoption is high availability. Table 3-3 outlines some of the SAN inter-components that can be configured redundantly to avoid any single point of failure.

Table 3-3 SAN inter-component redundant configuration

Storage controller redundancy	Configure storage arrays with multiple controllers to provide redundant array ports and avoid any single point of failure at the array controller level.
Disk redundancy	Configure the array using different RAID groups as required to provide redundancy at the disk level.
Path redundancy	Configure SAN infrastructure (switches, HBA ports) to provide redundancy and avoid any point of path failures.

3.8 Installation of SUSE Linux Enterprise Server 11 SP 1 for SAN Boot

In this section, we describe installation and troubleshooting for SUSE Linux with SAN Boot.

3.8.1 Setting up a SUSE Linux host using SAN Boot

Following are the general steps to set up a SUSE Linux host using SAN Boot:

1. Configure the DS8000 Subsystem so that only the boot volume is mapped to the host, preferentially.
2. Configure the SAN Zones to the host that sees the DS8000 Subsystem system node ports. SUSE Linux already has the multipath to accept multiple paths to the SAN Boot LUN, so you can configure multiple paths for your host.
3. Configure and enable the HBA BIOS, as shown in 3.4.5, “Configuring QLogic adapter for SAN Boot” on page 137 and/or 3.4.6, “Configuring Emulex adapter for SAN Boot” on page 143.
4. Install the operating system, selecting the volume configured for SAN Boot (3.5.2, “Configuring the LUN for SAN Boot” on page 152) as the partition on which to install.
5. For our SUSE Linux installation, we used the version of SUSE Linux Enterprise Server version 11 SP1.
6. Collect additional information about SUSE installation process by accessing the Novell documentation site:

   http://www.novell.com/documentation/

7. The first panel asks you to select the installation process option, as shown in Figure 3-54.

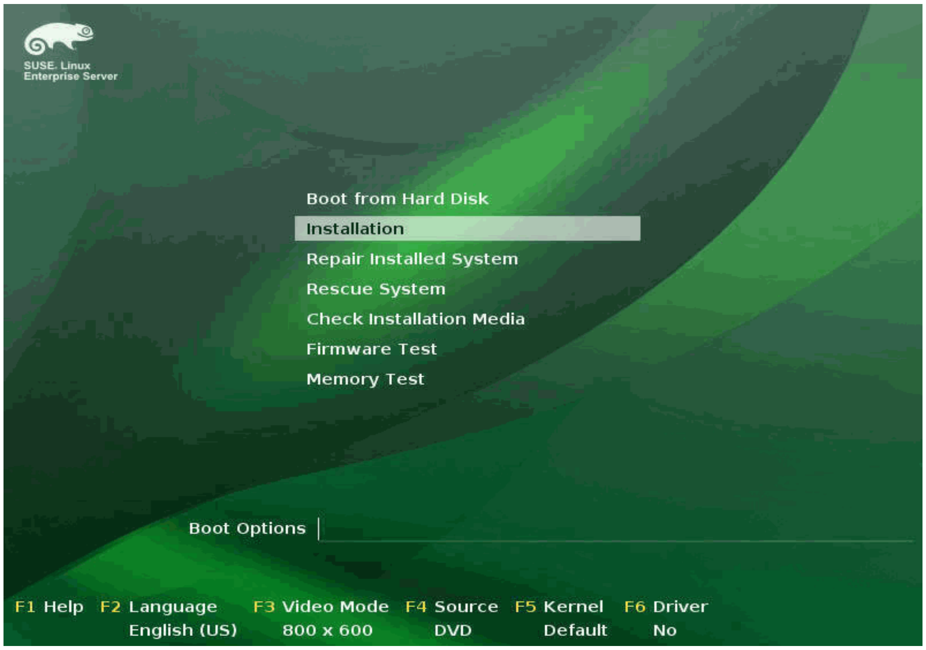


Figure 3-54 Selecting the Installation process

8. Read and accept the License Terms and click **Next**.
9. Click **Next** at Media Check panel.
10. Select the **New Installation** option to proceed with the installation for SAN Boot, as shown in Figure 3-55.

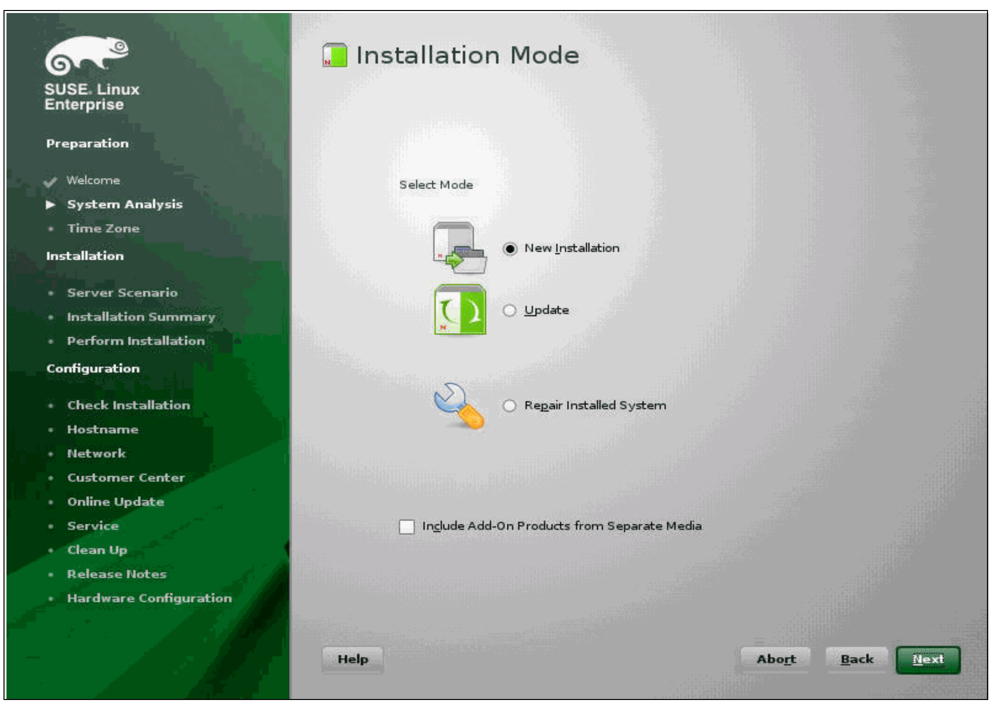


Figure 3-55 SUSE New installation option

11. Set the clock and Time Zone and click **Next**.

12. Choose the Server Base Scenario which better matches your configuration, as shown Figure 3-56.

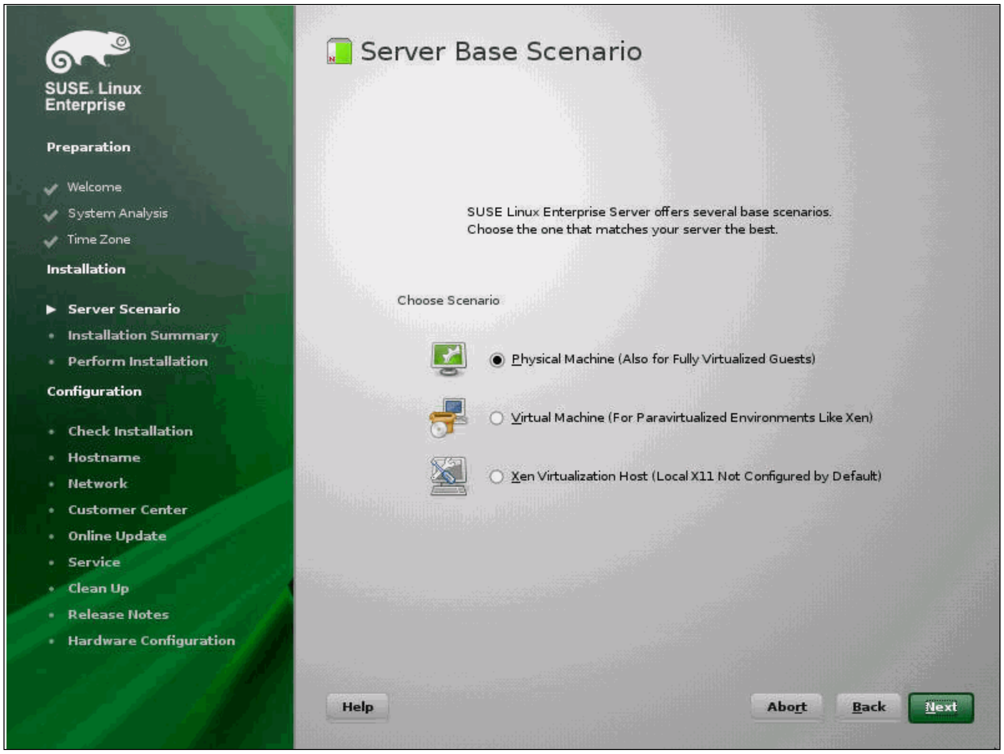


Figure 3-56 Selecting the Scenario option

13.SUSE Linux already has the drivers and multipath tool to support a DS8000 Subsystem with multiple accesses to the LUN, so you can configure the installation process to use multiple paths in this case, during the installation process. To proceed using multiple paths, you need to use the *Expert Installation* setting and change the Partitioning and Boot options, as shown in Figure 3-57.

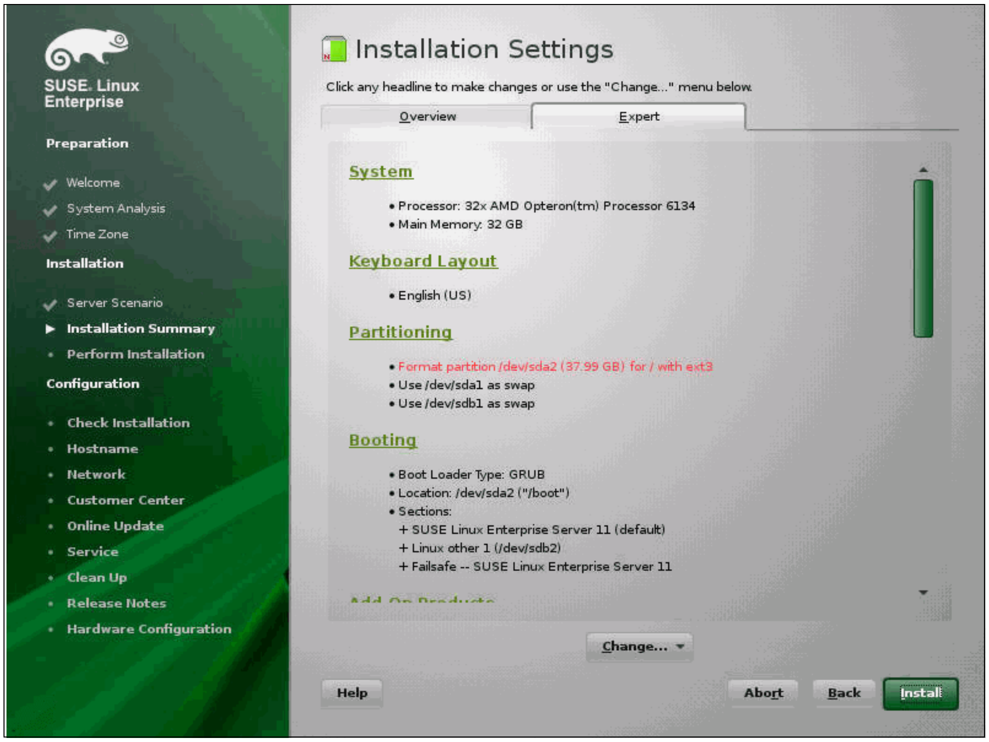


Figure 3-57 Installing using the Expert Configuration

14.Click **Partitioning** and then **Custom Partitioning**, as shown in Figure 3-58 and click **Accept**.

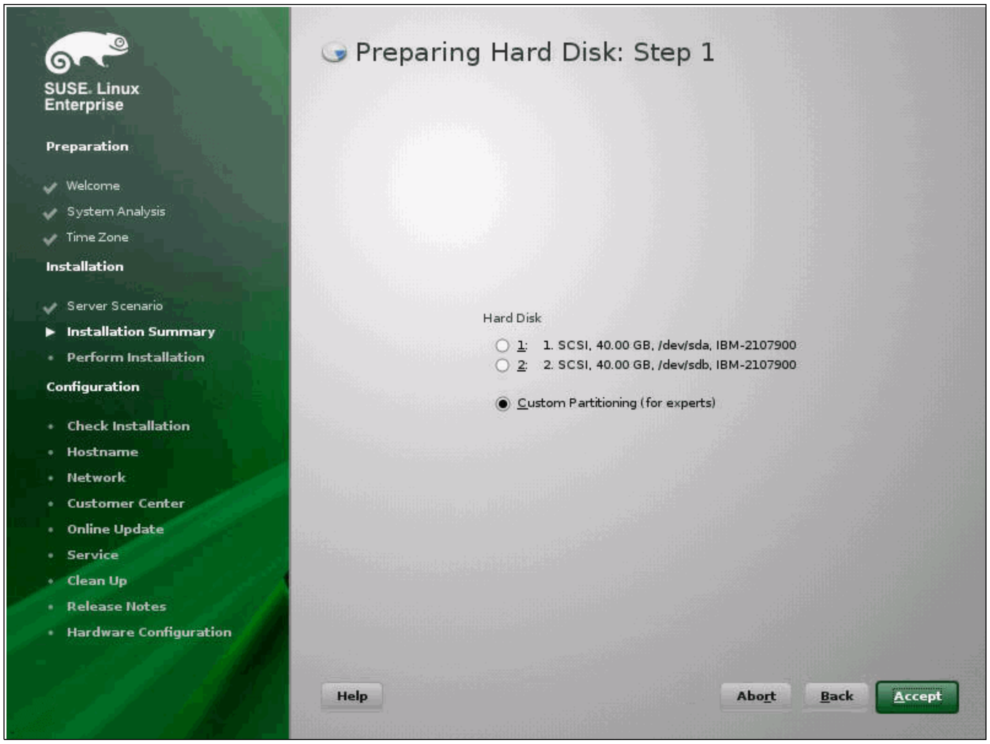


Figure 3-58 Partitioning the LUN for Experts

15.Click the **Hard Disks**, then click **Configure** and **Configure Multipath**, as shown in Figure 3-59. Afterwards, click **Yes** to Activate multipath.

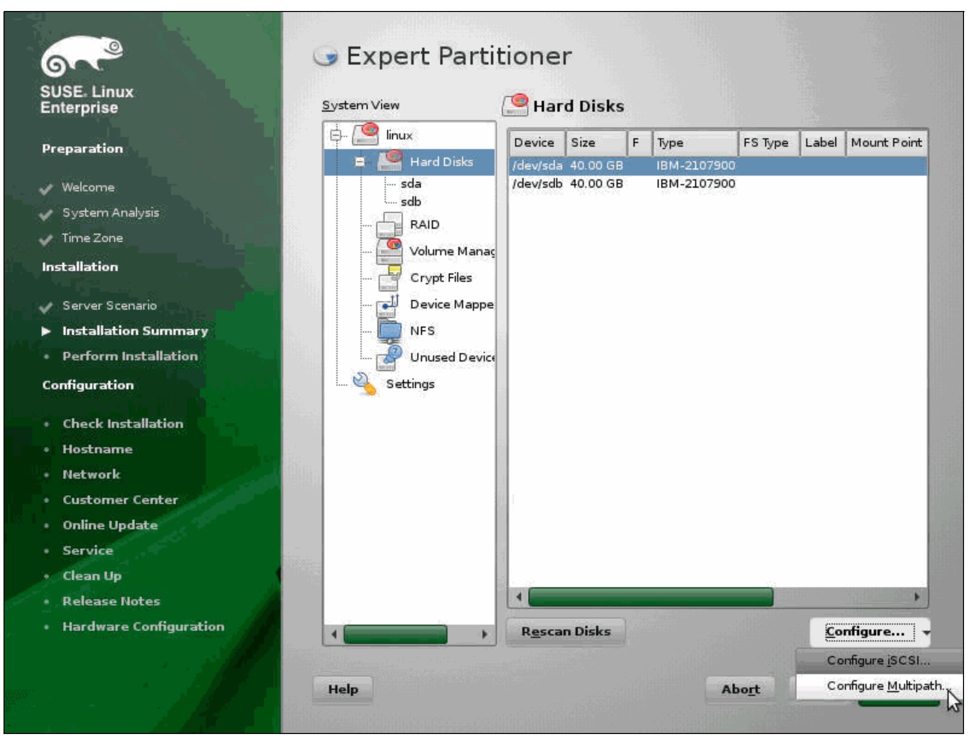


Figure 3-59 Configuring multipath

16. This process automatically creates the multipath disk that will be used to create the SWAP partition and the Root partition. You can change the partition here or you can use the standard as shown in Figure 3-60, Figure 3-61, Figure 3-62, and Figure 3-63.

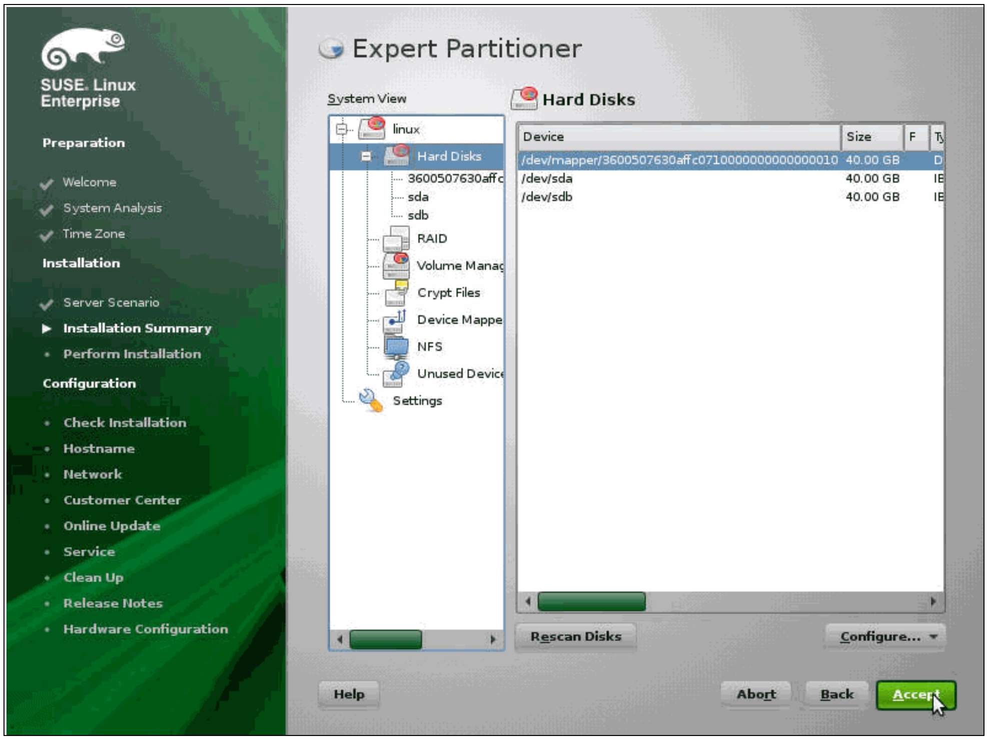


Figure 3-60 Multipath LUN created

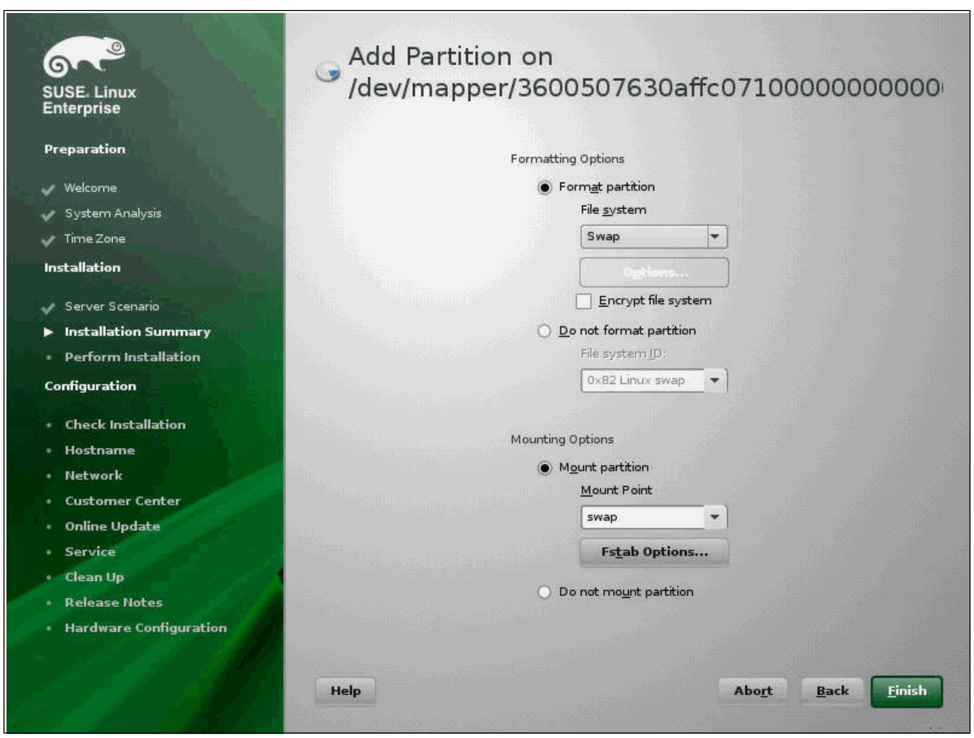


Figure 3-61 Creating SWAP partition

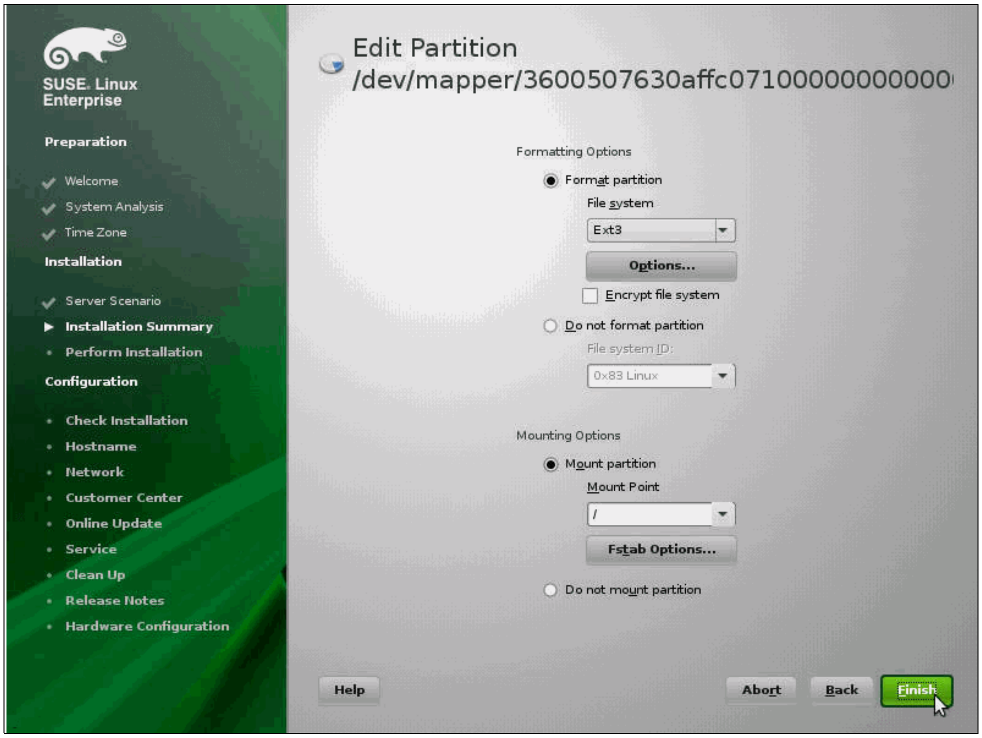


Figure 3-62 Creating root partition

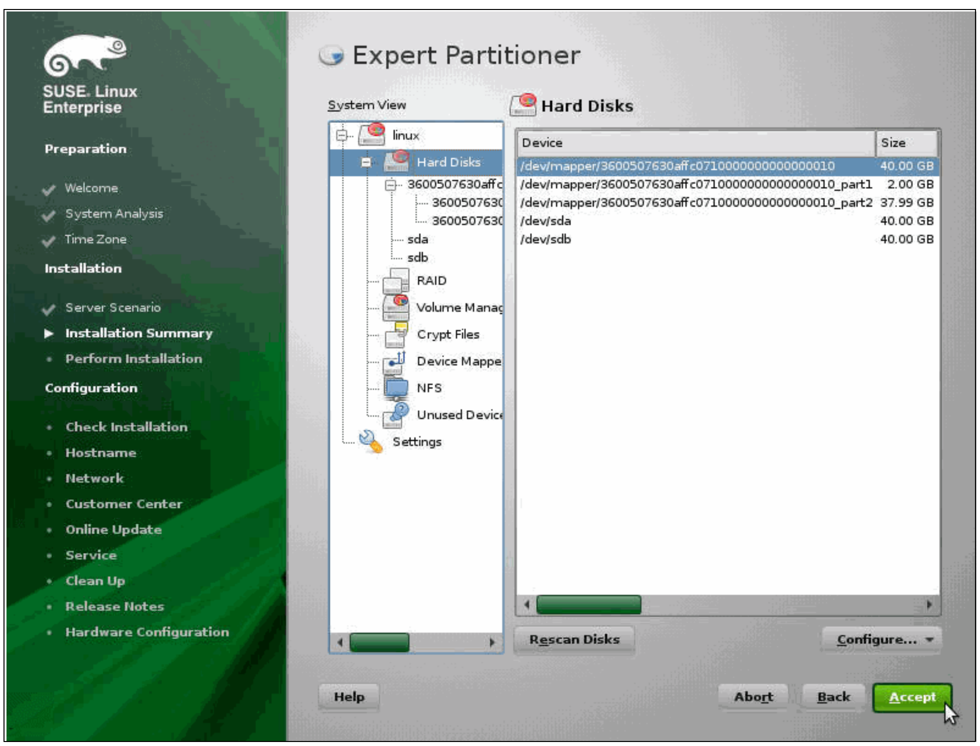


Figure 3-63 Accepting partitions to install

17. Change the Boot loader to reflect the multipath partitioning changes. Click the **Booting** option, **Boot Loader Installation**, as shown in Figure 3-64.

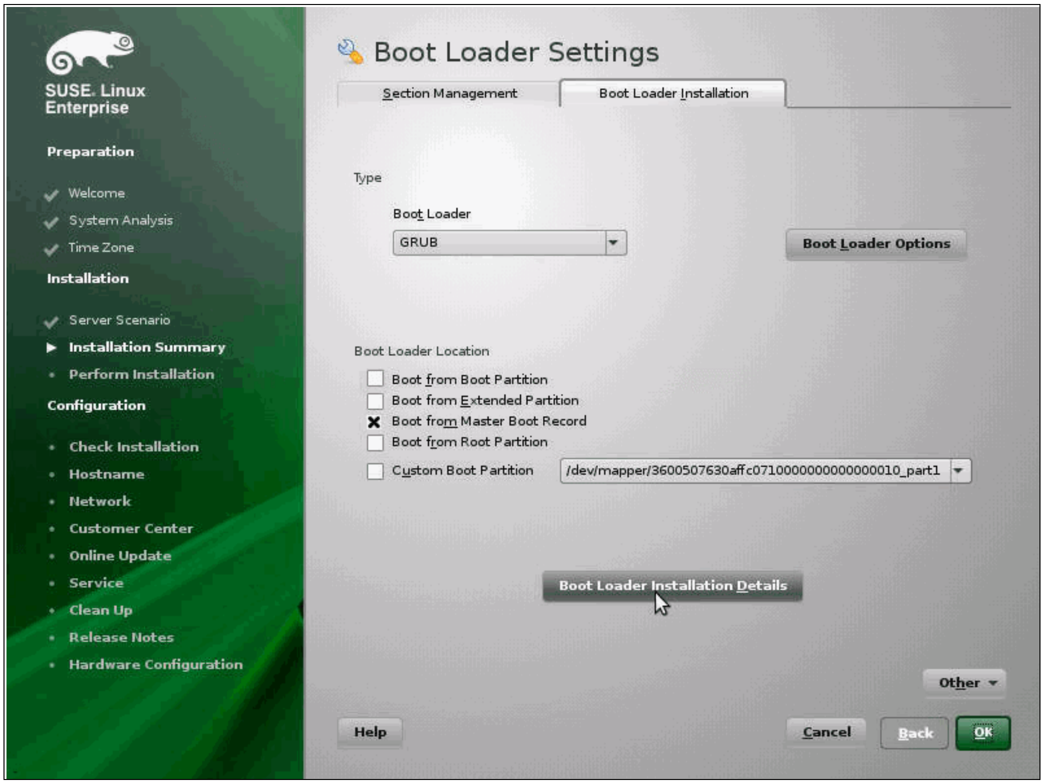


Figure 3-64 Boot loader installation details

18. Click **Boot Loader Installation Details**, as shown in Figure 3-65. Accept the multipath as a boot loader by clicking **OK** as shown in Figure 3-65.

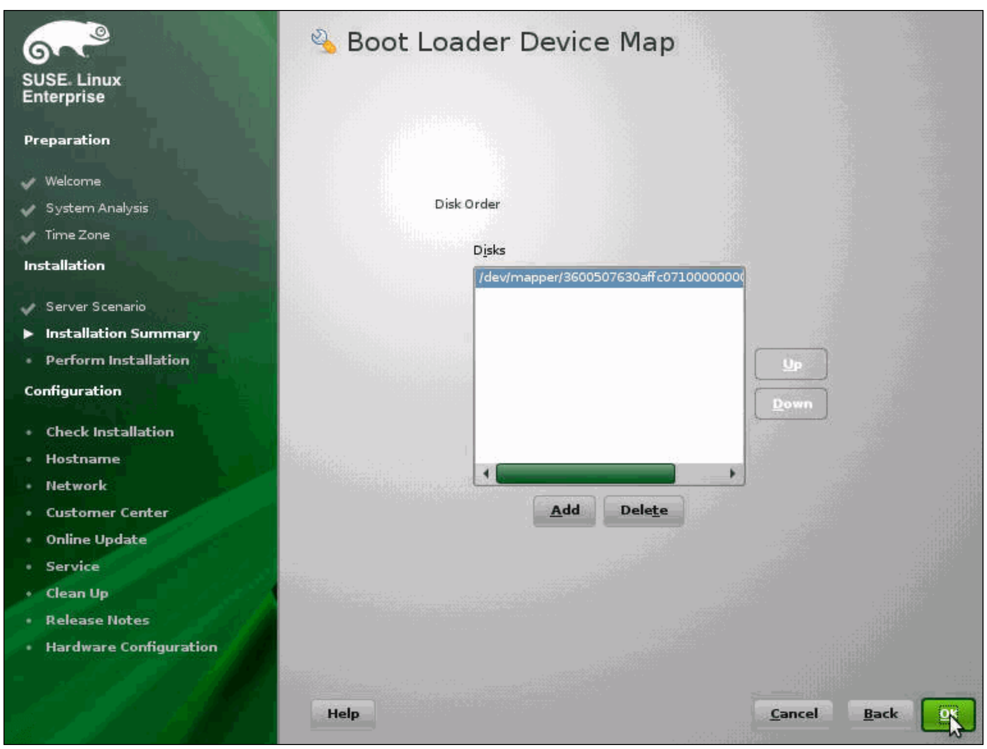


Figure 3-65 Changing the Boot Loader option

19. Click the **Install** button to initiate, accept the license term and start your installation. This process will take around 30 minutes to finish.
20. After the installation, the machine will reboot, and then you can start setting the administrator (root) password and click **Next**.
21. Set the hostname and click **Next**.
22. Set the network configuration and click **Next**. You will have the option to test the network access here.
23. Select the authentication method to access the server and click **Next**.
24. After finishing this configuration, you should be able to add/change the /etc/multipath.conf file including the DS8000 storage subsystem information and reboot your machine, as shown Example 3-1.

Example 3-1 Changing the /etc/multipath.conf

```
defaults {
    polling_interval    30
    failback            immediate
    no_path_retry       5
    rr_min_io           100
    path_checker        tur
    user_friendly_names yes
}

devices {
# These are the default settings for 2107 (IBM DS8000)
# Uncomment them if needed on this system
    device {
        vendor                  "IBM"
        product                 "2107900"
        path_grouping_policy    group_by_serial
    }
}
```

25. Reboot the server and check that the multiple paths are showing as in Example 3-2.

Example 3-2 Checking the multipath.

```
rb1sanboot:~ # multipath -ll
3600507630affc071000000000000000f dm-0 IBM,2107900
[size=40G][features=1 queue_if_no_path][hwhandler=0][rw]
\_ round-robin 0 [prio=2][active]
 \_ 6:0:2:0 sda 8:0   [active][ready]
 \_ 7:0:2:0 sdb 8:16  [active][ready]
rb1sanboot:~ #
```

For additional information about how to set your host to access the DS8000 storage subsystem, see *IBM System Storage DS8000 Host Attachment and Interoperability*, SG24-8887.

3.8.2 Best practices and troubleshooting for SUSE 11 with SAN Boot

In some cases, the /etc/multipath.conf file will not be created during the installation process. We recommend that you create the file by following the structure described in Example 3-1 on page 170, to better reflect the parameters required for IBM Storage.

In other cases, after rebooting, if your server started without the multipath, you need to configure the multipath tool to identify and work with the LUNs from the DS8000 Subsystem.

Follow these steps to configure the multipath tool:

1. Check that the multipath packages were installed as expected. Use the **rpm -qa** command as shown Example 3-3.

Example 3-3 Checking the multipath packages

```
rb1sanboot:~ # rpm -qa|grep mapper
device-mapper-1.02.27-8.6
device-mapper-32bit-1.02.27-8.6
rb1sanboot:~ # rpm -qa|grep multi
multipath-tools-0.4.8-40.1
```

2. Change or create the /etc/multipath.conf file using the IBM recommendation for DS8000 as shown Example 3-4.

Example 3-4 Creating the multipath.conf file.

```
rb1sanboot:/etc # vi /etc/multipath.conf
devices {
lling_interval    30
    failback           immediate
    no_path_retry      5
    rr_min_io          100
    path_checker       tur
    user_friendly_names yes
}

devices {
# These are the default settings for 2107 (IBM DS8000)
# Uncomment them if needed on this system
    device {
        vendor                  "IBM"
        product                 "2107900"
        path_grouping_policy    group_by_serial
    }
}
:wq
rb1sanboot:/etc #
```

3. Edit the /etc/sysconfig/kernel, adding the dm-multipath as shown Example 3-5.

Example 3-5 Editing the /etc/sysconfig/kernel

```
rb1sanboot:~ # vi /etc/sysconfig/kernel
## Description:
## Type:        string
## Command:     /sbin/mkinitrd
#
# This variable contains the list of modules to be added to the initial
# ramdisk by calling the script "mkinitrd"
# (like drivers for scsi-controllers, for lvm or reiserfs)
#
INITRD_MODULES="processor thermal ahci fan jbd ext3 edd dm-multipath scsi_dh
scsi_dh_alua"
:wq
rb1sanboot:
```

4. Check the multipath configuration as shown Example 3-6.

Example 3-6 Checking the multipath daemons

```
rb1sanboot:~ # chkconfig multipathd on
rb1sanboot:~ # chkconfig boot.multipath on
rb1sanboot:~ # chkconfig multipathd
multipathd  on
rb1sanboot:~ # chkconfig boot.multipath
boot.multipath  on
rb1sanboot:~ #
```

5. Reboot the server and check the multiple paths again as shown Example 3-7.

Example 3-7 Checking the multipaths

```
rb1sanboot:~ # multipath -ll
mpatha (3600507630affc071000000000000000f) dm-0 IBM,2107900
[size=40G][features=1 queue_if_no_path][hwhandler=0][rw]
\_ round-robin 0 [prio=2][active]
 \_ 6:0:2:0 sda 8:0   [active][ready]
 \_ 7:0:2:0 sdb 8:16  [active][ready]
rb1sanboot:~ #
```

3.9 Installation of Red Hat Enterprise Server 5.5 for SAN Boot

In this section, we describe installation and troubleshooting for Red Hat Linux with SAN Boot.

3.9.1 Setting up a Red Hat Linux 5.5 host using SAN Boot

Following are the general steps to set up a Red Hat Linux 5.5 host using SAN Boot:

1. Configure the DS8000 Subsystem so that only the boot volume is mapped to the host, preferentially.
2. Configure the SAN Zones so the host sees the DS8000 Subsystem system node ports. Red Hat already has the multipath to accept multiple paths to the SAN Boot LUN, so you can configure multiple path zones for your host.
3. Configure and enable the HBA BIOS, as shown in 3.4.5, “Configuring QLogic adapter for SAN Boot” on page 137 and/or 3.4.6, “Configuring Emulex adapter for SAN Boot” on page 143.
4. Install the operating system selecting the volume configured for SAN Boot (see 3.5.2, “Configuring the LUN for SAN Boot” on page 152) as the partition on which to install.
5. For our Red Hat installation, we used Red Hat Enterprise Server version 5.5.
6. Collect additional information about the Red Hat installation process by accessing the Red Hat documentation site:

   `http://docs.redhat.com/docs/en-US/index.html`

7. The first panel will request you to select the installation process option, as shown in Figure 3-66. If you are configuring multiple paths to access your SAN Boot LUN, type **F2** and then type **`linux mpath`**. Red Hat is able to identify the DS8000 LUNs, then install and configure the multipath tools without any additional tasks.

Figure 3-66 Selecting the Installation process

8. Choose the language and click **OK**.
9. Choose the keyboard type and click **OK**.
10. Select the **CD-ROM** for the installation media and click **OK**. After this selection, Red Hat installation will start the graphical interface. Click **Next** to proceed with the installation as shown in Figure 3-67.

Figure 3-67 Red Hat new installation

11. Enter the installation number to validate your license.
12. When performing the installation for the first time, Red Hat will ask you to format your disk to prepare that for multipath. That is the last opportunity to check if you are using the correct LUNs. After confirmation. the LUN will be formatted as shown in Figure 3-68.

Figure 3-68 Formatting the LUN

13. Click **Next** to create the partitions, as shown in Figure 3-69.

Figure 3-69 Formation the partitions to install

14. Set the network parameters such as IPs, DNS, and gateway, and then click **Next**.
15. Select your TimeZone and click **Next**.
16. Set the password for the root user and click **Next**.

17. Decide if you want to include additional packages to be installed or if you want to customize them later, as shown in Figure 3-70. Click **Next** to proceed.

Figure 3-70 Selecting the additional Software to be installed

18. Click **Next** to start the installation, as shown in Figure 3-71.

Figure 3-71 Starting the installation

19. After the installation, the Red Hat will request you to reboot the server, as shown in Figure 3-72. Click **Reboot** and proceed to the next step.

Figure 3-72 Rebooting the server

20. After the server reboots, you can proceed with the rest of the configuration starting with the welcome panel as shown in Figure 3-73. After the completing the steps here, we proceed with the multipath configuration.

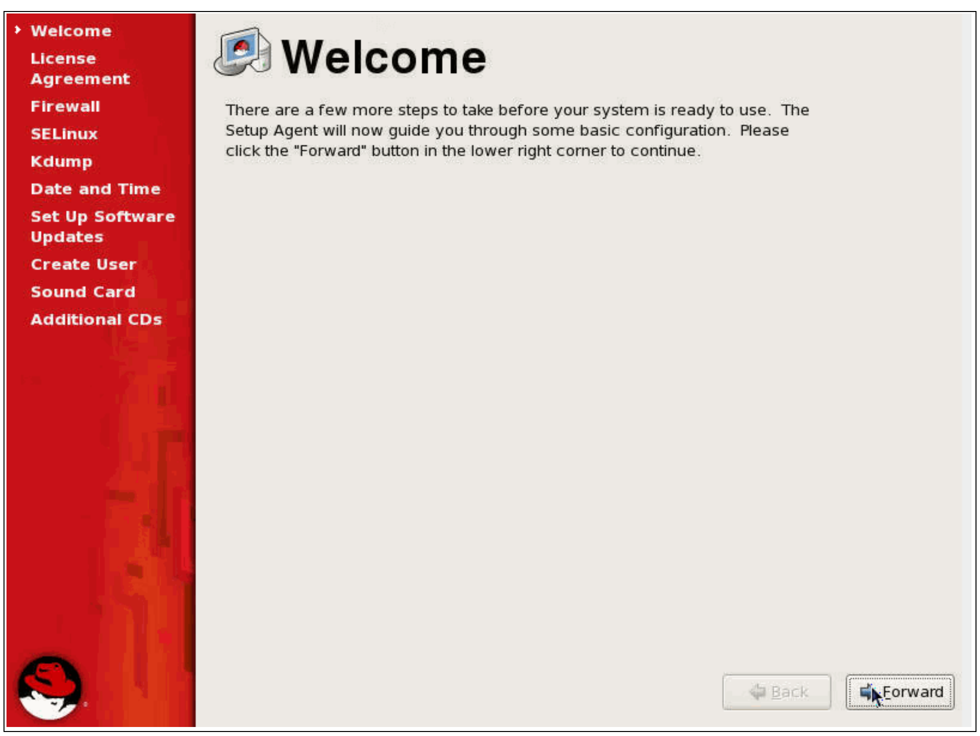


Figure 3-73 Configuring the last basic steps

21. After finishing the Red Hat installation and configuration, we can check if the multipath grouped the disks and was configured as planned. Example 3-8 shows an example of the multipath configuration.

Example 3-8 Multipath.conf and checking the multpath drivers

```
[root@rb1sanboot ~]# cat /etc/multipath.conf
defaults {
        user_friendly_names yes
}

blacklist {
        devnode "^(ram|raw|loop|fd|md|dm-|sr|scd|st)[0-9]*"
        devnode "^(hd|xvd|vd)[a-z]*"
        wwid "*"
}

# Make sure our multipath devices are enabled.

blacklist_exceptions {
        wwid "3600507630affc071000000000000000e"
}

[root@rb1sanboot ~]# multipath -ll
mpath0 (3600507630affc071000000000000000e) dm-0 IBM,2107900
[size=40G][features=1 queue_if_no_path][hwhandler=0][rw]
\_ round-robin 0 [prio=2][active]
 \_ 0:0:1:0 sda 8:0   [active][ready]
 \_ 1:0:1:0 sdb 8:16  [active][ready]
[root@rb1sanboot ~]#
```

See *IBM System Storage DS8000 Host Attachment and Interoperability*, SG24-8887 for additional information about how to set your host to access the DS8000 storage subsystem.

3.9.2 Troubleshooting tips for Red Hat Linux 5.5 with SAN Boot

Here we provide troubleshooting tips for SAN Boot implementation of Red Hat Linux 5.5:

- After the installation is complete, if your system does not boot up, verify these items:
  - Verify that the server supports booting from Fibre Channel connected storage by looking into the firmware boot options.
  - Verify that you are using the appropriate LUN shown as a boot device by your HBA.
- If the installation menu does not show the devices in mapper/mpath0 format, verify the following items:
  - Verify that you provided the mpath in the command line booting up to install.
  - Verify that your hardware is configured appropriately, such that the storage devices are seen through multiple paths.
- Tips for LVM on multipath devices:

  When using LVM on dm-multipath devices, it is preferred to turn off LVM scanning on the underlying SCSI devices. It can be done by changing the filter parameter in the /etc/lvm/lvm.conf file to be filter = ["a/dev/mapper/.*/", "r/dev/sd.*/"].

  If your root device is also a multipath LVM device, then make the foregoing change before you create a new initrd image.

- In some cases, the /etc/multipath.conf file will be not created during the installation process. If so, we recommend that you create the file following the structure described in Example 3-1 on page 170, to better reflect the parameters required for IBM storage.
- In other cases, the server can start without the multipath after rebooting. In this case, you need to configure the Multipath tool to identify and work with the LUNs from DS8000 Subsystem. Follow these steps to configure the multipath tool:

  a. To check that the multipath packages were installed correctly, use the **rpm -qa** command as shown in Example 3-9.

Example 3-9 Checking the multipath packages

```
rb1sanboot:~ # rpm -qa|grep mapper
device-mapper-1.02.27-8.6
device-mapper-32bit-1.02.27-8.6
rb1sanboot:~ # rpm -qa|grep multi
multipath-tools-0.4.8-40.1
```

  b. Change or create the /etc/multipath.conf using the IBM recommendation for DS8000 as shown in Example 3-10.

Example 3-10 Creating the multipath.conf file

```
rb1sanboot:/etc # vi /etc/multipath.conf
devices {
lling_interval    30
    failback           immediate
    no_path_retry      5
    rr_min_io          100
    path_checker       tur
    user_friendly_names yes
}
devices {
# These are the default settings for 2107 (IBM DS8000)
# Uncomment them if needed on this system
    device {
        vendor                  "IBM"
        product                 "2107900"
        path_grouping_policy    group_by_serial
    }
}
:wq
rb1sanboot:/etc #
```

  c. Edit the /etc/sysconfig/kernel, adding the dm-multipath as shown in Example 3-11.

Example 3-11 Editing the /etc/sysconfig/kernel

```
rb1sanboot:~ # vi /etc/sysconfig/kernel
## Description:
## Type:        string
## Command:     /sbin/mkinitrd
#
# This variable contains the list of modules to be added to the initial
# ramdisk by calling the script "mkinitrd"
# (like drivers for scsi-controllers, for lvm or reiserfs)
#
```

```
INITRD_MODULES="processor thermal ahci fan jbd ext3 edd dm-multipath scsi_dh
scsi_dh_alua"
:wq
rb1sanboot:
```

d. Build a new initramfs using the *mkinitrd* command as shown in Example 3-12.

Example 3-12 Including the multipathd for boot

```
rb1sanboot:~ # mkinitrd -f multipath
Kernel image:   /boot/vmlinuz-2.6.27.19-5-default
Initrd image:   /boot/initrd-2.6.27.19-5-default
Root device:    /dev/disk/by-id/scsi-mpatha-part2 (/dev/dm-2) (mounted on / as
ext3)
Resume device:  /dev/disk/by-id/scsi-3600507630affc071000000000000000f-part1
(/dev/sdb1)
Kernel Modules: hwmon thermal_sys processor thermal dock scsi_mod libata ahci fan
jbd mbcache ext3 edd scsi_dh dm-mod dm-multipath dm-snapshot scsi_tgt
scsi_transport_fc qla2xxx crc-t10dif sd_mod usbcore ohci-hcd ehci-hcd uhci-hcd
ff-memless hid usbhid dm-round-robin scsi_dh_emc scsi_dh_hp_sw scsi_dh_rdac
scsi_dh_alua
Features:       dm block usb multipath kpartx resume.userspace resume.kernel
Bootsplash:     SLES (800x600)
28920 blocks
rb1sanboot:~ #
```

e. Check the multipath configuration as shown in Example 3-13.

Example 3-13 Checking the multipath daemons

```
rb1sanboot:~ # chkconfig multipathd on
rb1sanboot:~ # chkconfig boot.multipath on
rb1sanboot:~ # chkconfig multipathd
multipathd  on
rb1sanboot:~ # chkconfig boot.multipath
boot.multipath  on
rb1sanboot:~ #
```

f. Reboot the server and check the multipaths again as shown in Example 3-14.

Example 3-14 Checking the multipaths

```
rb1sanboot:~ # multipath -ll
mpatha (3600507630affc071000000000000000f) dm-0 IBM,2107900
[size=40G][features=1 queue_if_no_path][hwhandler=0][rw]
\_ round-robin 0 [prio=2][active]
 \_ 6:0:2:0 sda 8:0   [active][ready]
 \_ 7:0:2:0 sdb 8:16  [active][ready]
rb1sanboot:~ #
```

3.10 Installation of AIX 6.1 for SAN Boot

This section describes the processes available to install or migrate the AIX operating system using a SAN Boot in the DS8000 storage subsystem.

The subsections cover the following topics:

- 3.10.1, “AIX requirements for installation”
- 3.10.2, “Installing AIX 6.1 using a pre-existing installation”
- 3.10.3, “New installation process for SAN Boot”
- 3.10.4, “New AIX installation with Network Installation Manager”
- 3.10.5, “Identifying the logical drive to hdisk”
- 3.10.6, “AIX post installation tasks with DS8000 Subsystems best practices”

3.10.1 AIX requirements for installation

In this section, we describe considerations regarding the AIX installation.

Prerequisites

Before proceeding with the AIX installation, ensure that the following requirements are met:

- Plan your OS installation while keeping in mind what OS version, hardware, storage subsystems, and SAN architecture, as well as the multipaths, are required.
- Check the status of your Hardware Management Console (HMC) and hardware.
- Verify that the microcode for your server and firmware for your HBA have been updated.
- Collect all installation media to be used before starting the process.
- Check the license agreements.
- Verify the LUN assignments and host connectivity were properly done.
- Verify the zoning for the host.
- Check for special requirement and conduct your installation following the recommendations.

This section describes the possibilities of installing AIX in the SAN Boot environment.

Support: At the time of writing this book, only the FC configuration is supported on AIX for SAN Boot and newest Power Servers on the DS8000 Subsystem.

Implementation possibilities

Implementations of SAN Boot with AIX include the following possibilities:

- To implement SAN Boot on a system with an already installed AIX operating system:
  - You can use the alt_disk_install system utility.
  - You can mirror an existing SAN Installation to several other LUNs using Logical Volume Copy.
- To implement SAN Boot on a new system:
  - You can start the AIX installation from a bootable AIX CD install package.
  - You can use the Network Installation Manager (NIM).

The methods known as alt_disk_install or mirroring are simpler to implement than using the Network Installation Manager (NIM).

You can find more information about the differences between each process by accessing the Information Center and searching for the topic “Installation,” using the following link:

http://publib.boulder.ibm.com/infocenter/aix/v6r1/index.jsp?topic=/com.ibm.aix.baseadmn/doc/

3.10.2 Installing AIX 6.1 using a pre-existing installation

Follow this procedure for installing AIX 6.1 using a pre-existing installation.

Creating a boot disk with alt_disk_install

The following procedure is based on the **alt_disk_install** command, cloning the operating system located on an internal SCSI disk to the storage subsystem SAN disk.

> **Tip:** To use the **alt_disk_install** command, the following file sets must be installed:
>
> - bos.alt_disk_install.boot_images
> - bos.alt_disk_install.rte (for rootvg cloning)

It is necessary to have an AIX system up and running and installed on an existing SCSI disk. Otherwise, you must first install a basic AIX operating system on the internal SCSI disk and then follow this procedure:

1. Create a logical drive on the storage subsystem big enough to contain all rootvg data, and make an appropriate zoning to see the storage subsystem from the AIX system.
2. Assuming that the source internal SCSI disk is hdisk0 and the target disk on the storage subsystem is hdisk3, you can use one of the following commands to clone hdisk0 to hdisk3:

```
/usr/sbin/alt_disk_install –C –B -P all hdisk3
/usr/sbin/alt_disk_copy -O -B -d hdisk3 /AIX 6.1/
smitty alt_clone /appropriate input required/
```

> **Important:** The target disk must have the same or greater capacity than the source disk.

3. In Figure 3-74, you can see the result of executing **alt_disk_install** on our test AIX 5L V6.1 system.

```
9.11.218.138 - PuTTY
Creating logical volume alt_hd10opt
Creating /alt_inst/ file system.
Creating /alt_inst/home file system.
Creating /alt_inst/opt file system.
Creating /alt_inst/tmp file system.
Creating /alt_inst/usr file system.
Creating /alt_inst/var file system.
Generating a list of files
for backup and restore into the alternate file system...
Backing-up the rootvg files and restoring them to the
alternate file system...
Modifying ODM on cloned disk.
Building boot image on cloned disk.
forced unmount of /alt_inst/var
forced unmount of /alt_inst/usr
forced unmount of /alt_inst/tmp
forced unmount of /alt_inst/opt
forced unmount of /alt_inst/home
forced unmount of /alt_inst
Changing logical volume names in volume group descriptor area.
Fixing LV control blocks...
Fixing file system superblocks...
[root@b50-ds5k]:
>
```

Figure 3-74 Executing alt_disk_install

> **Considerations:** If the rootvg is mirrored, remove the second copy before running the **alt_disk_install** command, or alternatively, use the other *mirroring* method, which is described in 3.10.3, “New installation process for SAN Boot” on page 185.

4. The **alt_disk_install** command can run for several minutes. It creates a new volume group on the storage subsystem disk called altinst_rootvg and creates and copies all logical volumes from rootvg. You can verify the result with the **lspv** command, activating the volume group with **varyonvg altinst_rootvg** and **lsvg -l altinst_rootvg** commands as shown in Figure 3-75.

```
9.11.218.138 - PuTTY
[root@b50-ds5k]:
> lspv
hdisk0          0004362acb666243                    rootvg          active
hdisk1          0004362a29493cb0                    altinst_rootvg
hdisk2          0007041a2648506b                    dddd
hdisk3          0007041a26485446                    dddd
hdisk4          0007041a2648557e                    dddd
hdisk5          0007041a264856b5                    dddd
[root@b50-ds5k]:
> lsvg -l altinst_rootvg
0516-010 : Volume group must be varied on; use varyonvg command.
[root@b50-ds5k]:
> varyonvg altinst_rootvg
[root@b50-ds5k]:
> lsvg -l altinst_rootvg
altinst_rootvg:
LV NAME             TYPE       LPs     PPs     PVs  LV STATE      MOUNT POINT
hd5                 boot       1       1       1    closed/syncd  N/A
hd6                 paging     8       8       1    closed/syncd  N/A
hd8                 jfslog     1       1       1    closed/syncd  N/A
hd4                 jfs        33      33      1    closed/syncd  /
hd2                 jfs        29      29      1    closed/syncd  /usr
hd9var              jfs        1       1       1    closed/syncd  /var
hd3                 jfs        9       9       1    closed/syncd  /tmp
hd1                 jfs        33      33      1    closed/syncd  /home
hd10opt             jfs        2       2       1    closed/syncd  /opt
[root@b50-ds5k]:
>
```

Figure 3-75 Checking the results of alt_disk_install

5. Clean up the alternate disk volume group and make hdisk0 a system disk using either of the following commands:
   - **alt_disk_install -X**
   - **alt_rootvg_op -X / AIX 6.1/**
6. Ensure the proper boot logical setting for the target disk:

   /usr/lpp/bosinst/blvset -d /dev/hdisk3 -g level
7. If you find that the boot logical volume settings on the cloned disk are blank, update the cloned boot logical volume manually with the following command:

   echo y | /usr/lpp/bosinst/blvset -d /dev/hdisk3 -plevel
8. Change the AIX boot list to force AIX to start on hdisk3 at the next reboot with a fallback or secondary boot disk of hdisk0 if there is any problem with hdisk3:

   bootlist -m normal hdisk3 hdisk0

   > **Attention:** This procedure requires a reboot, so run **shutdown -Fr**.

   The system comes back up, booting from the storage subsystem rootvg disk.
9. Check which disk was used to boot by using the following command:

   bootinfo -b

For more information about the use of the **alt_disk_install** command, see *IBM eServer Certification Study Guide - AIX 5L Installation and System Recovery*, SG24-6183.

Alternatively, the cloned disk (hdisk3 in the previous example) can be remapped on the storage subsystem and used to boot up another AIX machine (or LPAR), a method that is significantly quicker than installing from CD-ROM a second machine. An added advantage is that any customized configuration on the source AIX system such as user IDs and tuning parameters are duplicated on the new AIX system or LPAR.

Be careful using this method on the new machine, because any network adapters will carry duplicated IP addresses causing network problems. This problem can be easily overcome by ensuring that the new machine does not have any physical network connectivity when it is initially booted using the cloned disk.

Mirroring ROOTVG for SAN Boot

Mirroring process is a basic concept and can be used to migrate the current internal disk for a SAN Boot disk. You can find this basic commands in the Redbooks publication, *IBM eServer Certification Study Guide - AIX 5L Installation and System Recovery*, SG24-6183

Here are the basic steps covered in the aforementioned documentation:

1. Configure the AIX server to recognize the new SAN Boot disk.
2. Proceed with HBA configuration for AIX.
3. Add the SAN Boot disk to the current ROOTVG.
4. Using LVM Mirror, create and synchronize the system operation Logical Volumes.
5. Remove the LV copy from the internal disk using LVM commands.
6. Remove the internal disk from the ROOTVG.
7. Configure the bootlist and bosboot to reflect the new configuration.

3.10.3 New installation process for SAN Boot

To install AIX on storage subsystem disks, make the following preparations:

1. Update the system/service processor microcode to the latest level. For the downloadable versions, see the following website:

   http://techsupport.services.ibm.com/server/mdownload/

2. Update the FC adapter (HBA) microcode to the latest supported level. For a detailed description of how to update the microcode, see 3.4.2, “Downloading and installing a host adapter driver” on page 132.
3. Make sure that you have an appropriate SAN configuration. Check that the host is properly connected to the SAN, the zoning configuration is updated, and at least one LUN is mapped to the host.

> **Tips:**
>
> - If the system cannot see the SAN fabric at login, you can configure the HBAs at the server prompt, `Open Firmware`.
> - By default, AIX will install using native drivers and the MPIO package. Using this default, the server can be configured with multiple paths before starting the installation.

Installation procedure

Complete the Base Operating System (BOS) Installation using the following procedure. In this procedure, we perform a new and complete base operating system installation on a logical partition using the partition's media device. This procedure assumes that there is an HMC attached to the managed system.

Prerequisites before starting the installation process

> **AIX:** The information in this how-to scenario was tested using specific versions of AIX 6.1. The results you obtain might vary significantly depending on your version and level of AIX.

Here are few prerequisites to consider before starting the installation process:

- For the installation method that you choose, ensure that you follow the sequence of steps as shown. Within each procedure, you must use AIX to complete some installation steps, while other steps are completed using the HMC interface.
- Before you begin this procedure, you should have already used the HMC to create a partition and partition profile for the client. Assign the SCSI bus controller attached to the media device, a network adapter, and enough disk space for the AIX operating system to the partition. Set the boot mode for this partition to SMS mode. After you have successfully created the partition and partition profile, leave the partition in the Ready state. For instructions about how to create a logical partition and partition profile, see “Creating logical partitions and partition profiles” in *IBM PowerVM Virtualization Introduction and Configuration*, SG24-7940.

Activating and installing the partition

Perform the following steps on the HMC interface to activate and install the partition:

1. Activate the partition, as follows:
   a. Insert the AIX Volume 1 installation media into the media device of the managed system.
   b. Check if the DVD is already assigned to the current partition.
   c. Right-click the partition to open the menu.
   d. Select **Activate**. The Activate Partition menu opens with a selection of partition profiles. Be sure the correct profile is highlighted.
   e. Select **Open a terminal window or console session** at the bottom of the menu to open a virtual terminal (vterm) window.
   f. Select **Advanced** to open the Advanced options menu.
   g. For the Boot mode, select **SMS**.
   h. Select **OK** to close the Advanced options menu.
   i. Select **OK**. A vterm window opens for the partition.
2. In the SMS menu on the vterm, do the following steps:
   a. Press the 5 key and press Enter to select **5. Select Boot Options** as shown in Figure 3-76.

```
 Version EL340_075
 SMS 1.7 (c) Copyright IBM Corp. 2000,2008 All rights reserved.
-------------------------------------------------------------------------------
 Main Menu
 1.   Select Language
 2.   Setup Remote IPL (Initial Program Load)
 3.   Change SCSI Settings
 4.   Select Console
 5.   Select Boot Options

-------------------------------------------------------------------------------
 Navigation Keys:
X = eXit System Management Services

-------------------------------------------------------------------------------
 Type menu item number and press Enter or select Navigation key:5
```

Figure 3-76 SMS Basic Menu

   b. Press the 1 key and press Enter to select **1. Select Install/Boot Device**.
   c. Press the 3 key and press Enter to select **3. CD/DVD**.
   d. Select the media type that corresponds to the media device and press Enter.
   e. Select the device number that corresponds to the media device and press Enter. The media device is now the first device in the Current Boot Sequence list.
   f. Press the 2 key and press Enter to select **2. Normal Mode Boot**.
   g. Press the 1 key and press Enter to confirm your boot option.

   After you complete these steps, the system will boot your machine using the CD/DVD and load the basic Kernel to install the machine.

3. Boot from the AIX Volume 1, as follows:
   a. Select console and press Enter.
   b. Select language for BOS Installation menus, and press Enter to open the Welcome to Base Operating System Installation and Maintenance menu.
   c. Type 2 to select **Change/Show Installation Settings and Install** (as shown in Figure 3-77) in the Choice field and press Enter.

```
                         Welcome to Base Operating System
                           Installation and Maintenance

Type the number of your choice and press Enter.  Choice is indicated by >>>.

>>> 1 Start Install Now with Default Settings

    2 Change/Show Installation Settings and Install

    3 Start Maintenance Mode for System Recovery

    4 Configure Network Disks (iSCSI)


    88  Help ?
    99  Previous Menu

>>> Choice [1]: 2
```

Figure 3-77 Base Installation menu

4. Verify or change BOS Installation Settings, as follows:
   a. Type 2 in the Choice field to select the **Change/Show Installation Settings option**.
   b. Type 1 for System Setting to select **New and Complete Overwrite** in the Choice field and press Enter.
   c. When the Change Disk(s) panel opens, you can change the destination disk for the installation. If the default shown is correct, type 0 in the Choice field and press Enter. To change the destination disk, do the following steps:
      i. Type the number for each disk that you choose in the Choice field and press Enter. Do not press Enter a final time until you have finished selecting all disks. If you must deselect a disk, type its number a second time and press Enter.
      ii. By typing 77 two times, you can check if the Storage WWPN and LUN number associated with your machine is correct; see Example 3-15.

Example 3-15 Checking Storage WWPN and LUN ID.

```
Change Disk(s) Where You Want to Install

Type one or more numbers for the disk(s) to be used for installation and press
Enter. To cancel a choice, type the corresponding number and Press Enter.
At least one bootable disk must be selected. The current choice is indicated
by >>>.
Name      Device Adapter Connection Location
                  or Physical Location Code
    1  hdisk1    U789C.001.DQDU764-P2-D5
```

```
>>>  2  hdisk17   ...-W50050768014052D1-LA000000000000
     3  hdisk6    ...-W5001738000D00142-L1000000000000
     4  hdisk0    U789C.001.DQDU764-P2-D8
     5  hdisk2    U789C.001.DQDU764-P2-D4
    06  MORE CHOICES...

>>>  0   Continue with choices indicated above
    55  More Disk Options
    66  Devices not known to Base Operating System Installation
    77  Display More Disk Information
    88  Help ?
    99  Previous Menu

>>> Choice [0]:
```

iii. When you have selected the disks, type 0 in the Choice field and press Enter. The Installation and Settings panel opens with the selected disks listed under System Settings.

Important: Be sure that you have made the correct selection for root volume group, because the existing data in the destination root volume group will be destroyed during Base Operating System (BOS) installation.

d. If needed, change the primary language environment. Use the following steps to change the primary language used by this installation to select the language and cultural convention you want to use.

Tip: Changes to the primary language environment do not take effect until after the Base Operating System Installation has completed and your system is rebooted.

i. Type 2 in the Choice field on the Installation and Settings panel to select the Primary Language Environment Settings option.

ii. Select the appropriate set of cultural convention, language, and keyboard options. Most of the options are a predefined combination, however, you can define your own combination of options.

e. After you have made all of your selections, verify that the selections are correct. Press Enter to confirm your selections and to begin the BOS Installation. The system automatically reboots after installation is complete.

5. Select **Open terminal window** to open a virtual terminal (vterm) window.
   a. Type the model of your terminal as the terminal type.
   b. In the License Agreement menu, select Accept License Agreements.
   c. Select **Yes** to ACCEPT Installed License Agreements.
   d. Press F10 (or Esc+0) to exit the License Agreement menu.
   e. In the Installation Assistant main menu, select **Set Date and Time** as shown in Figure 3-78.

```
                              Installation Assistant

Move cursor to desired item and press Enter.

  Set Date and Time
  Set root Password
  Configure Network Communications
  Install Software Applications
  Using SMIT (information only)


F1=Help             F2=Refresh          F3=Cancel           F8=Image
F9=Shell            F10=Exit            Enter=Do
```

Figure 3-78 Post installation menu - Installation Assistant

f. Set the correct date, time, and time zone. Press the F3 (or Esc+3) key to return to the Installation Assistant main menu.

g. Select **Set root Password** and set a root password for the partition.

h. Select **Configure Network Communications**. Select **TCP/IP Startup**. Select from the Available Network Interfaces and press Enter. Enter the appropriate network information in the Minimum Configuration and Startup menu and press Enter. Use the F3 (or Esc+3) key to return to the Installation Assistant main menu.

i. Exit the Installation Assistant by typing F10 (or Esc+0).

The vterm window will display a login prompt at this time.

Switch the partition to Normal Mode

When possible, switch the partition to Normal Mode from the HMC, as follows:

1. Right-click the partition profile to open the menu. Be sure the correct partition profile is highlighted.
2. Select **Properties** and then select the **Settings** tab.
3. For the Boot Mode, select **Normal** and select **OK** to close the Properties menu.
4. Right-click the partition and select **Restart Partition**.
5. Select **Immediate** for the Restart Options and confirm that you want to restart the partition.

3.10.4 New AIX installation with Network Installation Manager

Network Installation Manager (NIM) is a client server infrastructure and service that allows remote installation of the operating system, manages software updates, and can be configured to install and update third-party applications. Although both the NIM server and client file sets are part of the operating system, a separate NIM server must be configured that will keep the configuration data and the installable product file sets.

NIM preparations

We assume that the following preparations have been completed:

- The NIM environment is deployed and all of the necessary configuration on the NIM master is already done.

- The NIM server is properly configured as the NIM master, and the basic NIM resources have been defined.
- The Fibre Channel Adapters are already installed in the machine onto which AIX is to be installed.
- The Fibre Channel Adapters are connected to a SAN and, on the storage subsystem, have the designated logical drive defined and mapped to the host or NIM client.
- The target machine (NIM client) currently has no operating system installed and is configured to boot from the NIM server.

For more information about how to configure a NIM server, see *NIM from A to Z in AIX 5L*, SG24-7296.

NIM installation procedure

Prior the installation, you can modify the bosinst.data file (for a more automated install), where the installation control is stored. Insert your appropriate values at the following stanza:

`SAN_DISKID`

This stanza specifies the World Wide Port Name and a Logical Unit ID for Fibre Channel attached disks. The World Wide Port Name and Logical Unit ID are each in the format returned by the **lsattr** command, that is, “0x” followed by 1-16 hexadecimal digits. The ww_name and lun_id are separated by two slashes (//):

`SAN_DISKID = <worldwide_portname//lun_id>`

Example 3-16 LUN ID example

```
SAN_DISKID = 0x0123456789FEDCBA//0x2000000000000

You can specify PVID (the example uses an internal disk):
target_disk_data:
PVID = 000c224a004a07fa
SAN_DISKID =
CONNECTION = scsi0//10,0
LOCATION = 10-60-00-10,0
SIZE_MB = 34715
HDISKNAME = hdisk0
```

To create a BOS installation client using the NIM, follow these steps:

1. Enter the command:

   `# smit nim_bosinst`

2. Select the **lpp_source** resource for the BOS installation.
3. Select the **SPOT** resource for the BOS installation.
4. Select the **BOSINST_DATA** to use during installation option, and select a bosinst_data resource that is capable of performing a unprompted BOS installation.
5. Select the **RESOLV_CONF** or define the host to use for network configuration option, and select a **resolv_conf** resource.
6. Select the **Accept New License Agreements** option, and select Yes. Accept the default values for the remaining menu options.
7. Press Enter to confirm and begin the NIM client installation.

8. To check the status of the NIM client installation, for example, where client name is “va09,” enter the command:

   # lsnim -l va09

9. After checking that the NIM Server and Client are ready, proceed using the SMS menu to boot and configure the LPAR to access the NIM Master Server.

10. Access the partition console and select **1** to access the SMS Menu as shown in Figure 3-79.

```
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM

1 = SMS Menu 5 = Default Boot List
8 = Open Firmware Prompt 6 = Stored Boot List
```

Figure 3-79 Boot menu

11.; Select **2** for IPL configuration Menu as shown in Figure 3-80.

```
PowerPC Firmware
Version SF220_001
SMS 1.5 (c) Copyright IBM Corp. 2000, 2003  All rights reserved.
----------------------------------------------------------------------------
Main Menu

1. Select Language
2. Setup Remote IPL (Initial Program Load)
3. Change SCSI Settings
4. Select Console
5. Select Boot Options



----------------------------------------------------------------------------
Navigation Keys:

X = eXit System Management Services
----------------------------------------------------------------------------
Type the number of the menu item and press Enter or select Navigation Key:
```

Figure 3-80 SMS first menu

12.Be sure that the partition already was cabled to have access to the same LAN where the NIM Master is.

13.Select the NIC card cabled already to access the NIM Master. We select **1** in our case, as shown in Figure 3-81.

```
PowerPC Firmware
Version SF220_001
SMS 1.5 (c) Copyright IBM Corp. 2000, 2003  All rights reserved.
-----------------------------------------------------------------------------
NIC Adapters
Device Slot Hardware Address
1. 10/100 Mbps Ethernet PCI Adapt Integ:U0.1-P1/E2 0002554f5c46
2. IBM 10/100/1000 Base-TX PCI-X 4:U0.1-P2-I4/E1 00145eb7f39d
-----------------------------------------------------------------------------
Navigation keys:
M = return to Main Menu
ESC key = return to previous screen X = eXit System Management Services
-----------------------------------------------------------------------------
Type the number of the menu item and press Enter or select Navigation Key: 1
```

Figure 3-81 NIC card list for remote IPL

14.Select **1** to configure IPs and Network Masks, as shown in Figure 3-82.

```
PowerPC Firmware
Version SF220_001
SMS 1.5 (c) Copyright IBM Corp. 2000, 2003  All rights reserved.
-----------------------------------------------------------------------------
Network Parameters
IBM 10/100/1000 Base-TX PCI-X Adapter: U0.1-P2-I4/E1
1. IP Parameters
2. Adapter Configuration
3. Ping Test
-----------------------------------------------------------------------------
Navigation keys:
M = return to Main Menu
ESC key = return to previous screen X = eXit System Management Services
-----------------------------------------------------------------------------
Type the number of the menu item and press Enter or select Navigation Key: 1
```

Figure 3-82 Setting the IP Parameters panel

15.Configure the IP defined for this partition in the /etc/hosts of NIM Master, IP of the NIM Master server, the gateway, and the mask, as shown in Figure 3-83.

```
PowerPC Firmware
Version SF220_001
SMS 1.5 (c) Copyright IBM Corp. 2000, 2003  All rights reserved.
-----------------------------------------------------------------------------
IP Parameters
IBM 10/100/1000 Base-TX PCI-X Adapter: U0.1-P2-I4/E1
1. Client IP Address [9.3.58.217]
2. Server IP Address [9.3.58.194]
3. Gateway IP Address [9.3.58.194]
4. Subnet Mask [255.255.255.000]
-----------------------------------------------------------------------------
Navigation keys:
M = return to Main Menu
ESC key = return to previous screen X = eXit System Management Services
-----------------------------------------------------------------------------
Type the number of the menu item and press Enter or select Navigation Key:
```

Figure 3-83 Menu to configure the network to access NIM

16. After the correct IP parameters have been entered, return to the main menu to select which boot device you want to use.
17. In the main menu, select **5** (Select boot option)
18. Select **6** (Network) to define your boot by Network
19. Select the right NIC card to boot. We use **1** as described in the steps before.
20. Select **2** for Normal Mode Boot.
21. Select **Yes** to proceed with the normal installation process as described in the documentation for using the NIM Master.

3.10.5 Identifying the logical drive to hdisk

Because a SAN allows access to a large number of devices, identifying the logical drive to hdisk for installation of AIX onto it can be difficult. Use the following method to facilitate the discovery of the lun_id to hdisk correlation:

1. For a new POWER machine that does not have AIX installed, map only the logical drive that you want to install AIX on. After AIX has been installed, then map all other logical drives required.
2. If more than one disk is assigned to an installed AIX host, you need to make sure you are using the right hdisk and its associated logical drive mapped on the storage subsystem:
   - Identify the hdisk by comparing the unique ID of the hdisk with the logical drive ID when viewing its properties on the storage subsystem. On AIX, use the **lsattr** command to identify the unique ID, for example, on hdisk9, which correlates to “logical drive ID” when you view its properties on the storage subsystem:

     ```
     lsattr -El hdisk9 | grep unique_id | awk '{print (substr($2,6,32))}'
     600A0B80004777D800006FC14AA6F708
     ```

   - The simplest method of checking an AIX system already running, where native AIX MPIO is being used, is to run the following command:

     mpio_get_config -Av

     The output of this command, as shown in Example 3-17, clearly details the hdisk - logical drive relationship.

Example 3-17 mpio_get_config -Av output

```
Frame id 0:
    Storage Subsystem worldwide name: 608e50017b5bc00004a955e3b
      Controller count: 2
      Partition count: 1
      Partition 0:
      Storage Subsystem Name = 'ITSO_5020'
          hdisk      LUN #   Ownership          User Label
          hdisk2         2   A (preferred)      Secured1
          hdisk3         3   A (preferred)      AIX_Boot
          hdisk4         4   A (preferred)      test3
          hdisk5        40   B (preferred)      p630_Test_40_64KB
          hdisk6        41   A (preferred)      p630_Test_41_8KB
          hdisk7        42   B (preferred)      p630_Test_42_512KB
          hdisk8        43   A (preferred)      p630_Test_43_128KB
```

– For AIX systems using SDDPCM with AIX MPIO, then issuing the equivalent command to the **mpio_get_config** will give the output in the same format as before:

sddpcm_get_config -Av

3.10.6 AIX post installation tasks with DS8000 Subsystems best practices

After installing AIX, some tasks must be done to preserve the environment, especially when using SAN Boot.

HBA additional configuration

List the Fibre Channel devices and attributes as shown in Figure 3-84.

```
# lsdev -Cc adapter|grep fcs
fcs0    Available 03-00 4Gb FC PCI Express Adapter (df1000fe)
# lsattr -El fscsi0
attach       switch       How this adapter is CONNECTED        False
dyntrk       no           Dynamic Tracking of FC Devices       True
fc_err_recov delayed_fail FC Fabric Event Error RECOVERY Policy True
scsi_id      0xc0700      Adapter SCSI ID                      False
sw_fc_class  3            FC Class for Fabric                  True
#
```

Figure 3-84 List of devices and default attributes for HBA0/FCS0

As a best practice, the recommendation for HBA will be to use the following settings:

- DYNTRK=YES (AIX supports dynamic tracking of Fibre Channel (FC) devices.)

  Previous releases of AIX required the user to unconfigure FC storage device and adapter device instances before making changes on the system area network (SAN) that might result in an N_Port ID (SCSI ID) change of any remote storage ports.

  If dynamic tracking of FC devices is enabled, the FC adapter driver detects when the Fibre Channel N_Port ID of a device changes. The FC adapter driver then reroutes traffic destined for that device to the new address while the devices are still online. Events that can cause an N_Port ID to change include moving a cable between a switch and storage device from one switch port to another, connecting two separate switches using an inter-switch link (ISL), and possibly rebooting a switch.

Dynamic tracking of FC devices is controlled by a new fscsi device attribute, dyntrk. The default setting for this attribute is dyntrk=no. To enable dynamic tracking of FC devices, set this attribute to dyntrk=yes, as shown Example 3-18.

Example 3-18 changing dyntrk

```
# chdev -l fscsi0 -a dyntrk=yes -a fc_err_recov=fast_fail -P
fscsi0 changed
```

- FC_ERR_RECOV=FAST_FAIL

  If the driver receives an RSCN from the switch, it can indicate a link loss between a remote storage port and switch. After an initial 15 second delay, the FC drivers query to see if the device is on the fabric. If not, I/Os are flushed back by the adapter. Future retries or new I/Os fail immediately if the device is still not on the fabric. If the FC drivers detect that the device is on the fabric but the SCSI ID has changed, the FC device drivers do not recover, and the I/Os fail with PERM errors. The default setting for this attribute is delayed_fail. To change the recovery model of FC devices, set this attribute to fc_err_recov=fast_fail, as shown in Example 3-19.

Example 3-19 changing recovery mode

```
# chdev -l fscsi0 -a dyntrk=yes -a fc_err_recov=fast_fail -P
fscsi0 changed
```

Installing the drivers and multipathing subsystem

By default, the AIX will be installed using native MPIO drivers and they must be updated with the right driver to access 2107 (DS8000 Subsystems). AIX 6.1 and DS8800 support MPIO/SDDPCM or SDD drivers; only one of them can be used.

1. Initially, go to the IBM web page and download the compatible version and guidance to install the drivers:
   - Host Attachment for SDDPCM on AIX:

     http://www-01.ibm.com/support/docview.wss?uid=ssg1S4000203

     The following packages must be installed:

     - devices.fcp.disk.ibm.mpio.rte
     - devices.sddpcm.61.2.6.0.3.bff
     - devices.sddpcm.61.rte
   - Host Attachments for SDD on AIX

     http://www-01.ibm.com/support/docview.wss?uid=ssg1S4000106

     The following packages must be installed:

     - devices.sdd.61.rte
     - devices.fcp.disk.ibm.rte

   > **Tip:** Only one package, SDD or SDDPCM, can be installed, The other one must be removed.

2. Log in to the new installed server, check the space available on the server, and transfer the package for the selected/created directory.
3. Unpackage (untar) the package to be installed.
4. Install the packages using the **`installp`** or **`smitty installp`** option and reboot the server.

5. Check the configuration as shown in Example 3-20.

Example 3-20 Checking the paths status

```
# pcmpath query device 6

DEV#:   6  DEVICE NAME: hdisk6  TYPE: 2107900  ALGORITHM:  Load Balance
SERIAL: 13031810702
==========================================================================
Path#      Adapter/Path Name          State     Mode     Select     Errors
    0           fscsi0/path0           OPEN   NORMAL       7666          0
    1           fscsi2/path1           OPEN   NORMAL       9117          0
```

Check if SDDPCM fixed the **reserve_policy** and **queue_depth** parameters for the SANBoot disk as shown in Figure 3-85.

```
# lsattr -El hdisk6
...
hcheck_interval 60                                Health Check Interval
...
queue_depth     20                                Queue DEPTH
reserve_policy  no_reserve                        Reserve Policy
...
#
```

Figure 3-85 Checking policy and queue_depth parameters

3.11 Installation of VMware ESX 4.1 for SAN Boot

Following are the general steps to set up a VMware ESXi host using SAN Boot:

1. Configure the DS8000 Subsystem so that only the boot volume is mapped to the host.
2. Configure the SAN Zones to the host to see the DS8000 Subsystem system node ports. VMware already has the multipath to accept multiple paths to the SAN Boot LUN, so you can configure multiple path zones for your host.
3. Configure and enable the HBA BIOS, as shown in 3.4.5, “Configuring QLogic adapter for SAN Boot” on page 137 and/or 3.4.6, “Configuring Emulex adapter for SAN Boot” on page 143.
4. Install the operating system selecting the volume configured for SAN Boot (see 3.5.2, “Configuring the LUN for SAN Boot” on page 152) as the partition on which to install.
5. For our ESX installation, we used VMware ESXi vSphere version 4.1.
6. Collect additional information about VMware installation process by accessing the VMware documentation site:

   http://www.vmware.com/support/pubs/vs_pubs.html

7. After booting from the DVD, select the ESXi Installer, as shown in Figure 3-86, to start the installation. ESXi installation already has the drivers required. It is able to identify the DS8000 LUNs, then install and configure the multipath tools without any additional tasks.

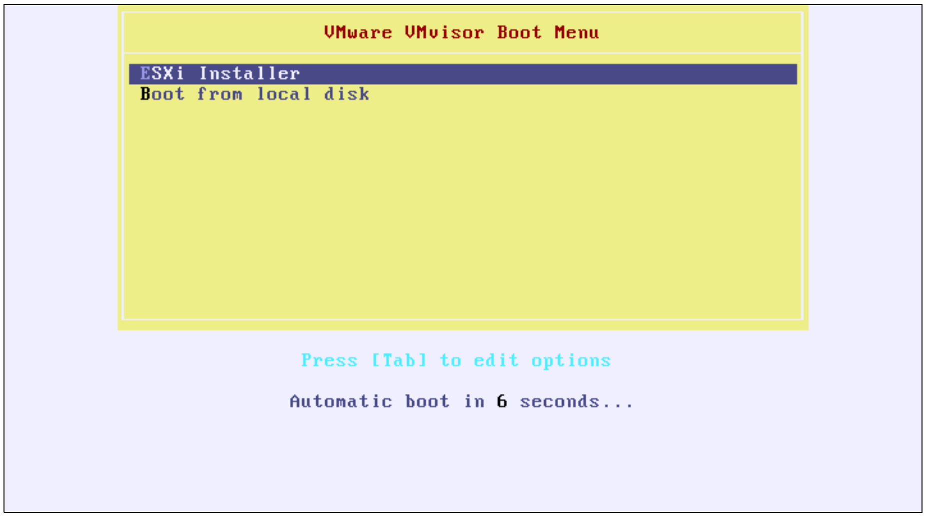


Figure 3-86 Selecting the Installation process

8. After loading the basic drivers and kernel, press the Enter key to proceed with the installation as shown in Figure 3-87.

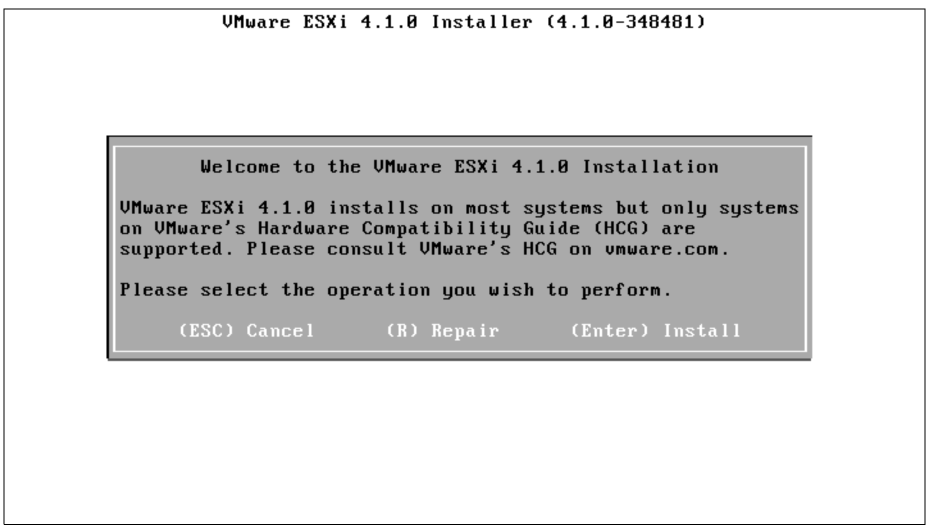


Figure 3-87 Welcome Installation page

9. Accept the License Agreement Terms by pressing the F11 key.
10. Select the **IBM 2107900 Disk** to install the ESXi and press the Enter key.

11.Confirm the Install options by pressing the F11 key, as shown in Figure 3-88.

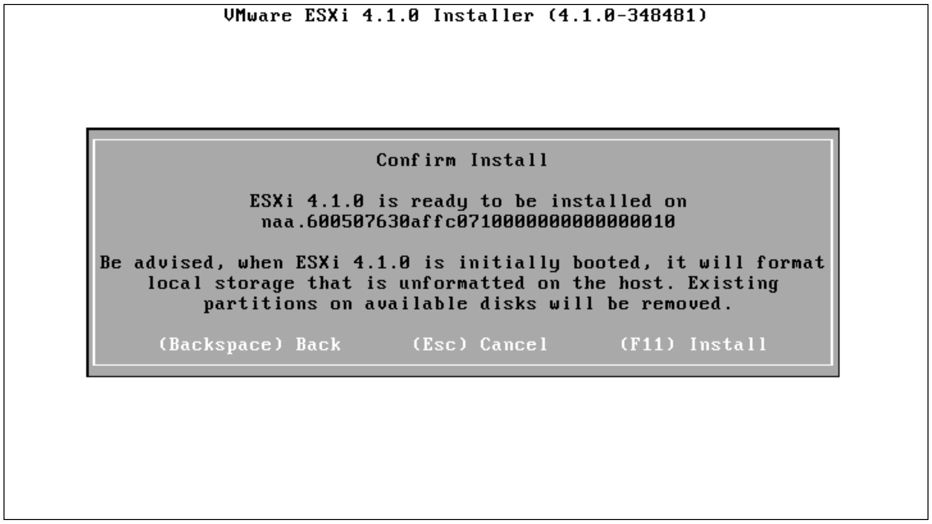


Figure 3-88 Confirm the disks to install ESXi

12.After confirmation, VMware ESXi installation will continue on the SAN Boot LUN.

13.After finishing the installation, reboot the host by pressing the Enter key, as shown Figure 3-89.

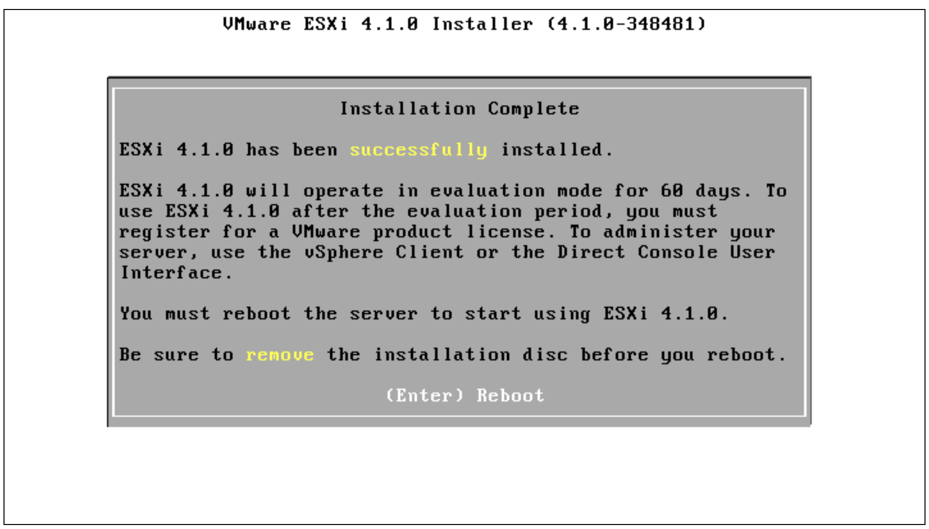


Figure 3-89 Installation Complete

14. By pressing the **F2** key, you can configure the basic information for your host, such as IPs, gateways, hostname, and other items, as shown in Figure 3-90.

Figure 3-90 Additional configuration steps

15. After finishing these steps, your server is ready to be accessed by the VMware administration interface, such as the VMware vSphere client or CLI interface. We use the GUI interface for access, as shown in Figure 3-91.

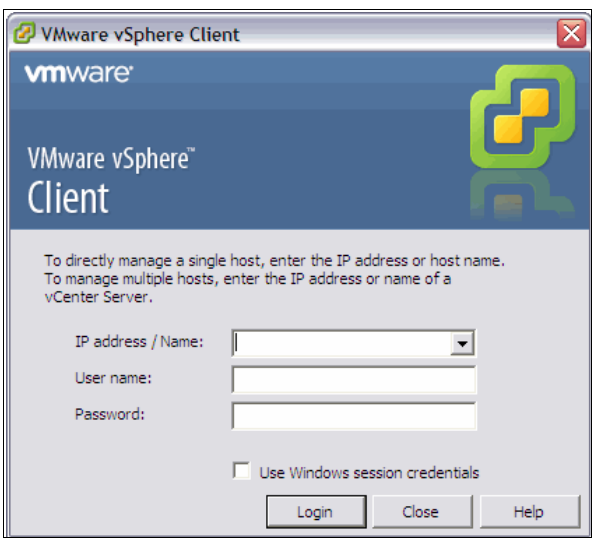


Figure 3-91 VMware vSphere Client Interface

16. To check the multipaths and the DS8000 LUNs associated with this server, click **Summary** and then select the data storage created, as shown in Figure 3-92.

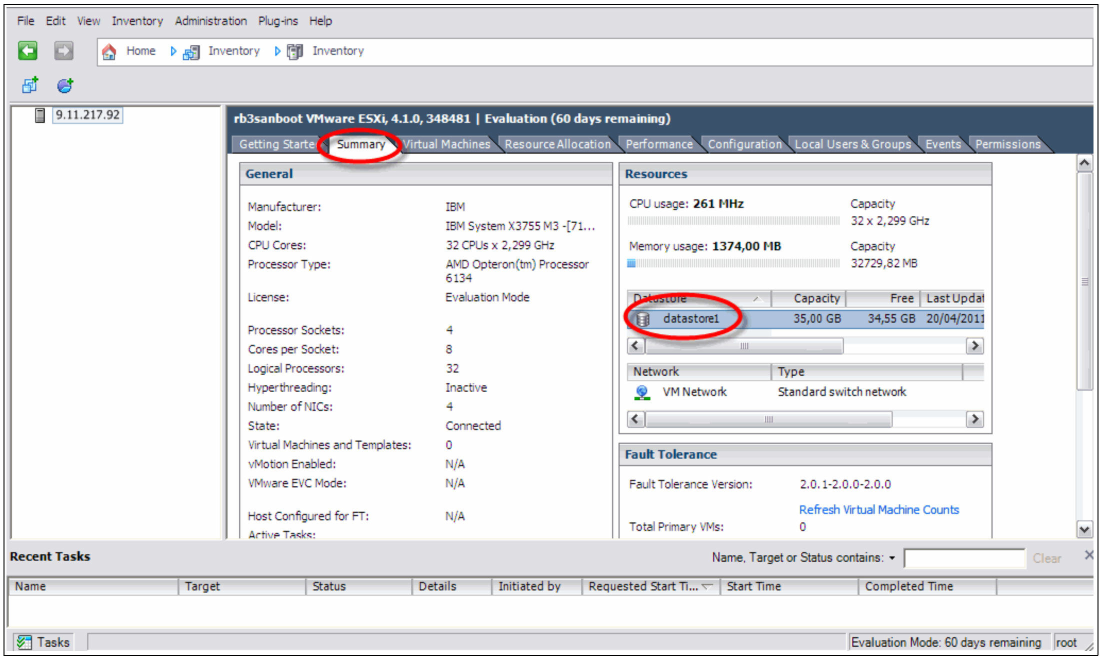


Figure 3-92 Checking the DataStore created

17. Right-click the appropriate datastore and select properties.

18. Select your Fibre Channel disk and then click **Manage Paths**, as shown in Figure 3-93.

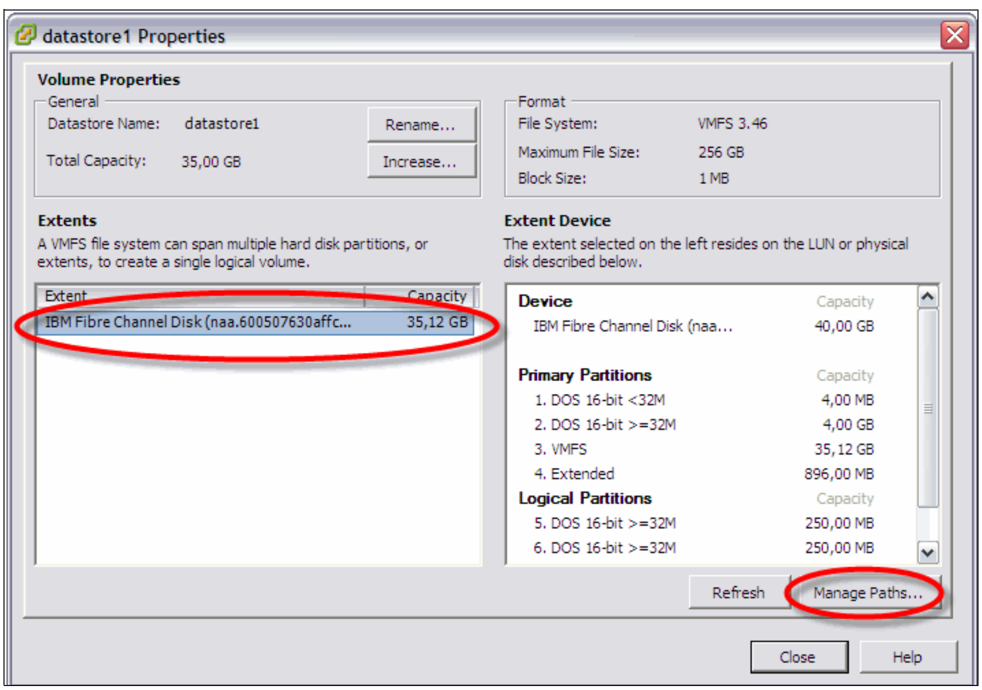


Figure 3-93 Checking multiple paths for SAN Boot LUNs

19. Here you see the details of the LUN 0 that has been accessed by two different HBAs zoned to two different DS8000 ports, as shown Figure 3-94.

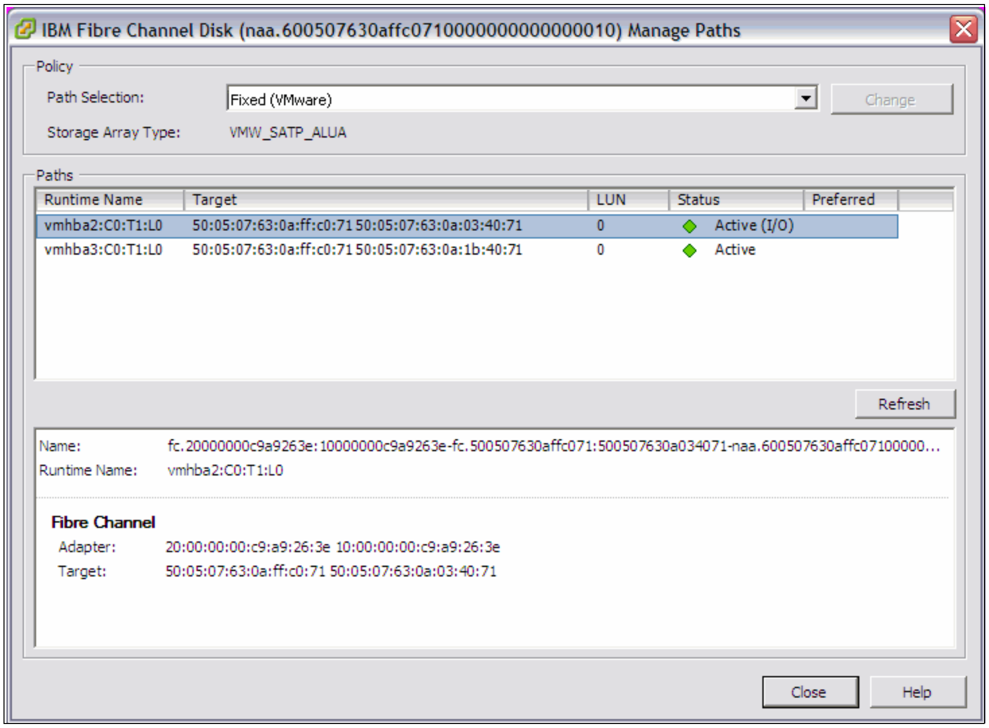


Figure 3-94 ESX multipath volumes

For additional information about how to set your host to access the DS8000 storage subsystem, see *IBM System Storage DS8000 Host Attachment and Interoperability*, SG24-8887.

3.11.1 Best practices and troubleshooting for VMware ESX 4.1 with SAN Boot

To ensure a trouble-free installation of ESX/ESXi, here are some useful tips:

1. Make sure that your hardware is compliant on the Hardware Compatibility Guide, including the following areas:
   - System compatibility
   - I/O compatibility (Network and HBA cards)
   - Storage compatibility
   - Backup software compatibility
2. Note that VMware ESX/ESXi 4.1 only installs and runs on servers with 64 bit x86 CPUs. 32 bit systems are no longer supported.
3. Make sure Intel VT is enabled in the host's BIOS.
4. If you are installing ESX/ESXi on the SAN disks and if an internal disk attached is connected to the ESX/ESXi host, then detach or disable them before proceeding with the installation. If you do not disconnect it, you might inadvertently choose an internal disk as the primary boot partition, which can result in the loss of data on the LUNs attached to the internal adapter.
5. The /, swap, and all the optional partitions are stored on a virtual disk called esxconsole-<UUID>.vmdk. Set a size minimum of 8 GB for this virtual disk.

> **Tip:** For /var/log, a separate partition is recommended to prevent unexpected disk space constraints due to extensive logging.

3.11.2 Persistent binding

Persistent binding is highly recommended in a SAN environment especially when several storage subsystems are available to the VMware ESX server.

Persistent binding means that a device has the same identification to the operating system after it restarts and after other devices are added to the operating system.

In SAN, the storage SCSI ID is dictated by the order in which the VMware ESX server discovers the subsystems at system start time. In some instances, this order might differ because of a device addition or removal. It can cause VMware ESX server to assign a different target ID and change the device addressing (vmhba <C> : <T> : <D>). It can cause the virtual machines to report their storage as missing when the system is starting. Persistent binding forces the VMware ESX server always to use the same target ID for a storage unit, no matter what the discovery order.

To enable persistent binding on all adapters, use the command: **`pbind.pl -A`**.

> **Tip:** You do not need to use this command if you are using v3.x.

See the *IBM System Storage DS8000 Host Attachment and Interoperability*, SG24-8887 for additional informations in how to set your host to access the DS8000 storage subsystem.

4

SAN Boot implementation with IBM XIV System Storage

In this chapter, we describe how to implement a host SAN Boot from XIV using Fibre Channel (FC) protocol connectivity. We provide an overview of XIV and host configuration steps necessary to perform a SAN Boot operating system installation.

At the time of writing, the ability to do a SAN Boot by the iSCSI protocol is not supported on XIV, even if an iSCSI HBA is used. Due to this current restriction, only SAN Boot by FC is described.

The term *host* in this chapter refers to a server running a supported operating system. For our purposes, we describe the following operating systems:

- AIX 6.1
- VMware ESXi 4.1
- Windows Server 2008 R2
- Linux (Red Hat Enterprise Linux 5.5, SUSE Linux Enterprise Server 11)

This chapter covers tasks that pertain to SAN Boot implementation. For operating system specific information regarding general host attachment, see the IBM Redbooks publication *XIV Storage System: Host Attachment and Interoperability*, SG24-7904.

For the latest information, see the XIV Hosts attachment kit publications at this website:

http://publib.boulder.ibm.com/infocenter/ibmxiv/r2/index.jsp

© Copyright IBM Corp. 2012. All rights reserved.

4.1 Overview

The XIV Storage System can be attached to various host platforms using the following methods:

- Fibre Channel adapters for support with the Fibre Channel Protocol (FCP)
- iSCSI software initiator for support with the iSCSI protocol

Support: At the time of writing, it is not supported to boot with iSCSI, even if an iSCSI HBA is used. Therefore, the description of SAN Boot for XIV is limited to boot with the FC protocol only.

The XIV Storage System has up to six interface modules, depending on the rack configuration. Figure 4-1summarizes the number of active interface modules as well as the FC ports for different rack configurations.

Rack configuration No. of modules	6	9	10	11	12	13	14	15
Module 9 state	NA	Disabled	Disabled	Enabled	Enabled	Enabled	Enabled	Enabled
Module 8 state	NA	Enabled	Enabled	Enabled	Enabled	Enabled	Enabled	Enabled
Module 7 state	NA	Enabled	Enabled	Enabled	Enabled	Enabled	Enabled	Enabled
Module 6 state	Disabled	Disabled	Disabled	Disabled	Disabled	Enabled	Enabled	Enabled
Module 5 state	Enabled	Enabled	Enabled	Enabled	Enabled	Enabled	Enabled	Enabled
Module 4 state	Enabled	Enabled	Enabled	Enabled	Enabled	Enabled	Enabled	Enabled
Number of dedicated Data modules	4	5	6	6	7	7	8	9
Number of Active Interface modules	**2**	**4**	**4**	**5**	**5**	**6**	**6**	**6**
FC ports	8	16	16	20	20	24	24	24

Figure 4-1 XIV hardware configuration matrix

Hosts attach to the FC ports through an FC switch.

Restriction: Direct attachment between hosts and the XIV Storage System is currently not supported.

With XIV, all interface modules and all ports can be used concurrently to access any logical volume in the system. The only affinity is the mapping of logical volumes to host. Although it simplifies storage management, balancing traffic and zoning (for adequate performance and redundancy) is more critical, although not more complex, than with traditional storage systems.

Important: Host traffic can be served through any of the interface modules. However, I/Os are not automatically balanced by the system. It is the storage administrator's responsibility to ensure that host connections avoid single points of failure and that the host workload is adequately balanced across the connections and interface modules. Balancing should be reviewed periodically or when traffic patterns change.

4.2 SAN configuration planning

Prior to installing a SAN Boot operating system, the planning and implementation of FC connectivity must be addressed. This section outlines basic topics and considerations related to FC SAN implementation.

4.2.1 Fibre Channel (FC) connectivity

Before you can attach an FC host to the XIV Storage System, there are a number of steps that you must complete.

Preparation

Here is a list of general steps that pertain to all hosts. However, in addition to these steps, review any procedures that pertain to your specific hardware and/or operating system:

1. Ensure that your HBA is supported. Information about supported HBAs and the recommended or required firmware and device driver levels is available at the IBM System Storage Interoperability Center (SSIC) at this website:

   http://www.ibm.com/systems/support/storage/config/ssic/index.jsp

   For each query, select the XIV Storage System, a host server model, an operating system, and an HBA vendor. Each query shows a list of all supported HBAs. Unless otherwise noted in SSIC, you can use any supported driver and firmware by the HBA vendors (the latest versions are always preferred).

   Also review any documentation that comes from the HBA vendor and ensure that any additional conditions are met.

2. Check the LUN limitations for your host operating system and verify that there are enough adapters installed on the host server to manage the total number of LUNs that you want to attach.
3. Check the optimum number of paths that must be defined. It will help in determining the zoning requirements.
4. Install the latest supported HBA firmware and driver. If these are not the ones that came with your HBA, they must be downloaded.

Platform and operating system vendor pages

The platform and operating system vendors also provide much support information for their clients. Refer to this information for general guidance about connecting their systems to SAN-attached storage. However, be aware that in some cases you cannot find information to help you with third-party vendors. Always check with IBM about interoperability and support from IBM in regard to these products. It is beyond the scope of this book to list all the vendors websites.

4.2.2 Host FC HBA configuration

For SAN Boot, you need to go into the HBA configuration mode, set the HBA BIOS to be *Enabled*, select at least one XIV target port, and select a LUN to boot from. In practice, you will typically configure 2-4 XIV ports as targets and you might need to enable the BIOS on two HBAs; however, it depends on the HBA, driver, and operating system. See the documentation that comes with your HBA and operating system.

The information supplied in this section outlines both QLogic and Emulex FC HBA configurations that are needed to allow SAN Boot functionality when installing Windows 2008 R2, Red Hat Enterprise Server 5.5, SUSE Linux Enterprise Server 11, and VMware ESXi 4.1.

SAN Boot for AIX 6.1 is separately addressed in 4.9, "Installation of AIX 6.1 for SAN Boot" on page 258.

QLogic FC HBA configuration

The procedures for setting up your server and HBA to SAN Boot can vary. The procedures in this section are for a QLogic QLE2562 HBA:

1. Boot your server. During the boot process, press **CTRL-Q** when prompted to load the configuration utility and display the *Select Host Adapter* menu. See Figure 4-2

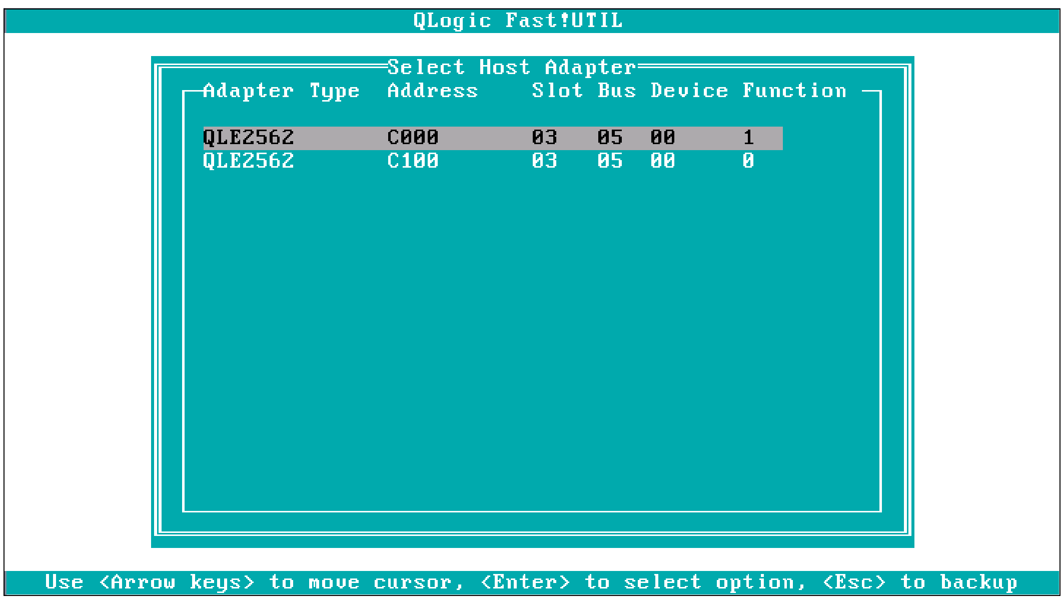


Figure 4-2 QLogic configuration panel: Select Host Adapter

2. You normally see one or more ports. Select a port and press Enter. It takes you to a panel as shown in Figure 4-3. If you will only be enabling the BIOS on one port, then make sure to select the correct port.
3. Select **Configuration Settings**.

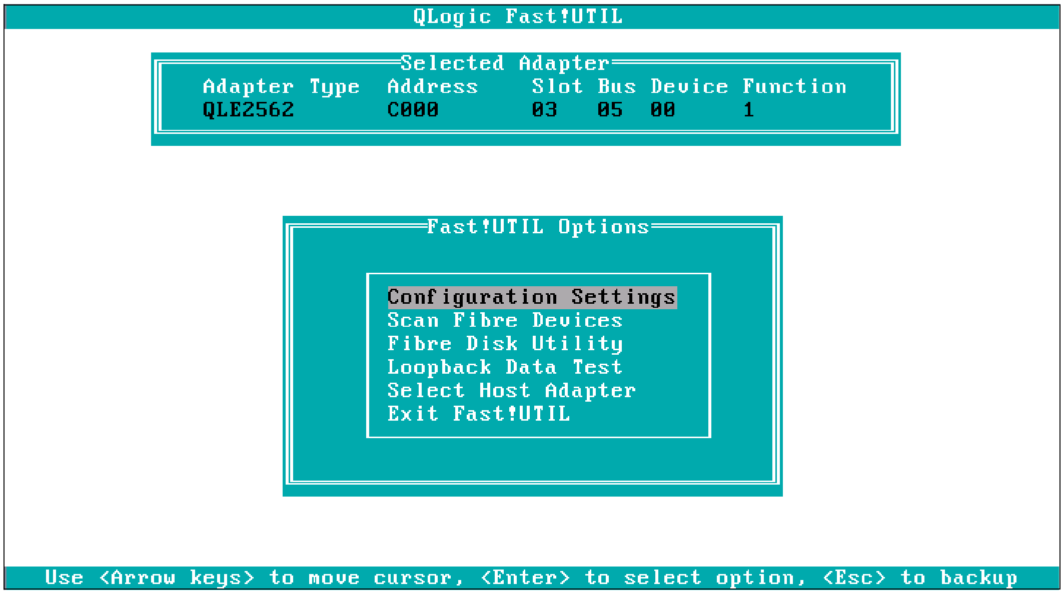


Figure 4-3 QLogic configuration panel: Fast!UTIL Options

4. In the panel shown in Figure 4-4, select **Adapter Settings**.

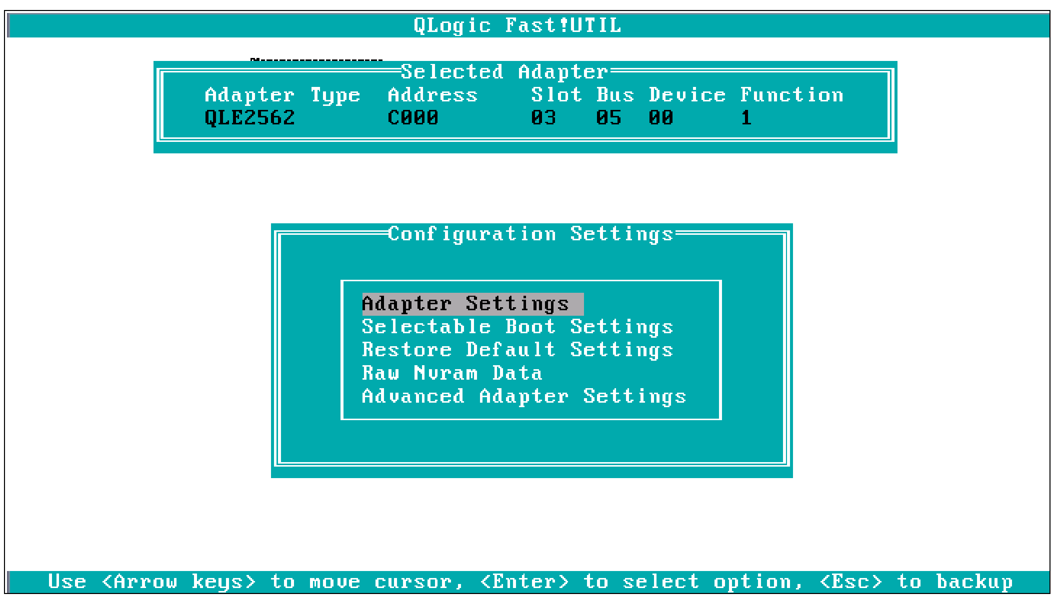


Figure 4-4 QLogic configuration panel: Configuration Settings

5. The *Adapter Settings* menu is displayed as shown in Figure 4-5.

Attention: The value associated with the *Adapter Port Name* in the Adapter Settings menu represents that port's World Wide Port Name (WWPN). Notice and record this value for zone planning as described in 4.2.5, "Zoning" on page 219.

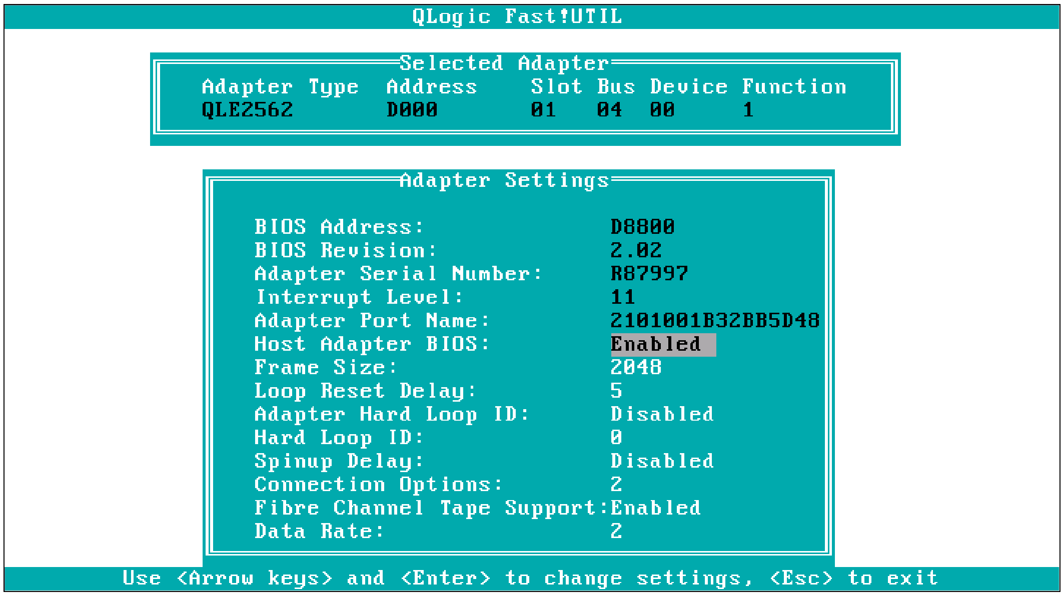


Figure 4-5 QLogic configuration panel: Adapter Settings

6. On the *Adapter Settings* panel, change the *Host Adapter BIOS* setting to **Enabled**, then press Esc to exit and go back to the *Configuration Settings* menu.
7. From the *Configuration Settings* menu, select **Selectable Boot Settings**, to get to the panel shown in Figure 4-6.

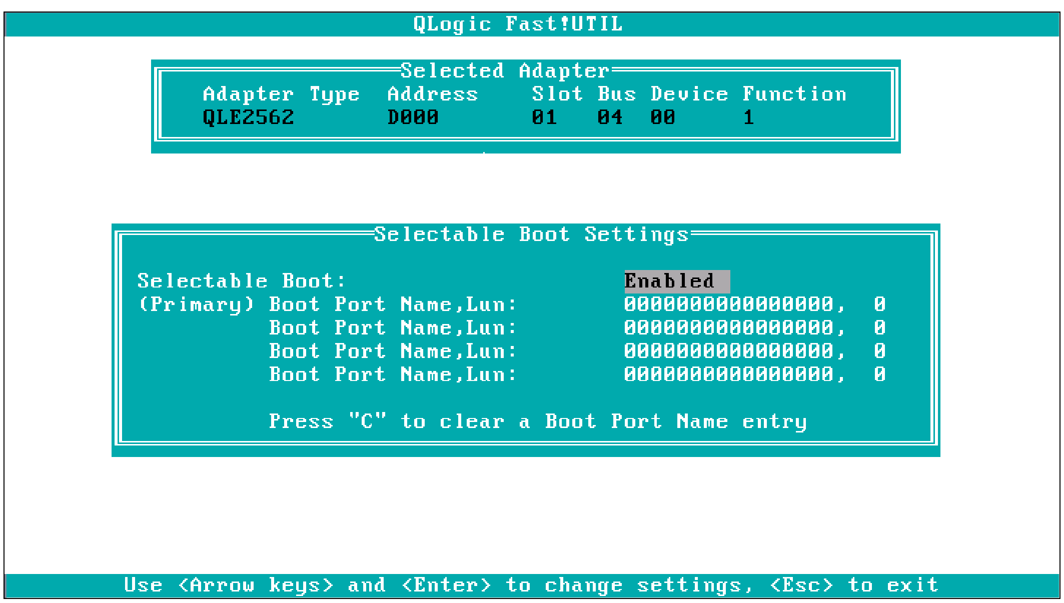


Figure 4-6 QLogic configuration panel: Selectable Boot Settings

8. Change the *Selectable Boot* option to **Enabled**. Select **Boot Port Name, Lun:** and then press Enter to get the *Select Fibre Channel Device* menu.
9. From the *Select Fibre Channel Device* panel in Figure 4-7, select the **IBM 2810XIV** device, and press Enter.

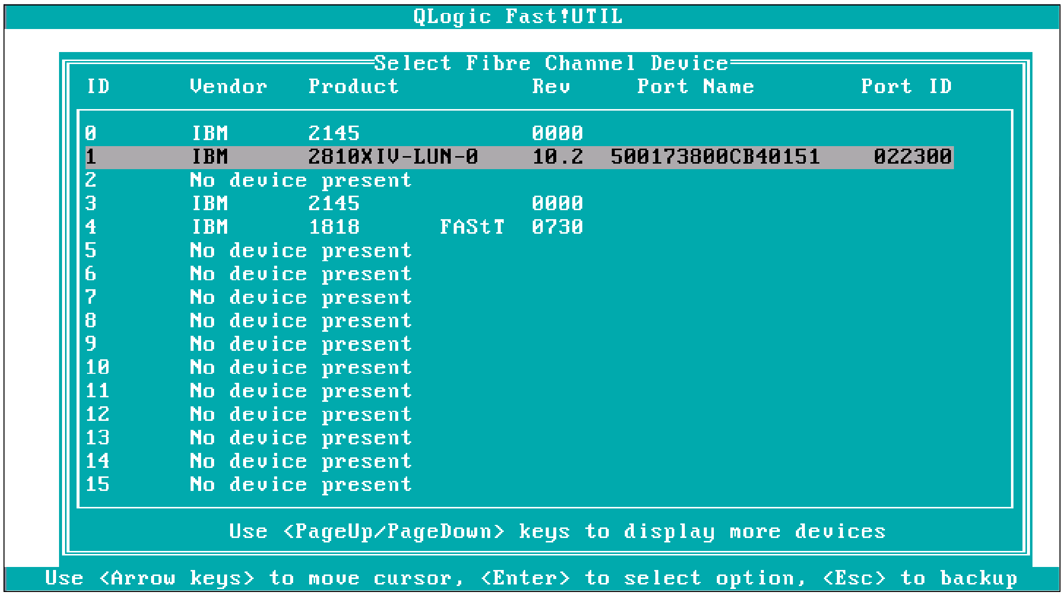


Figure 4-7 QLogic configuration panel: Select Fibre Channel Device

10. From the *Select LUN* panel in Figure 4-8, select **LUN 1** and then press Enter.

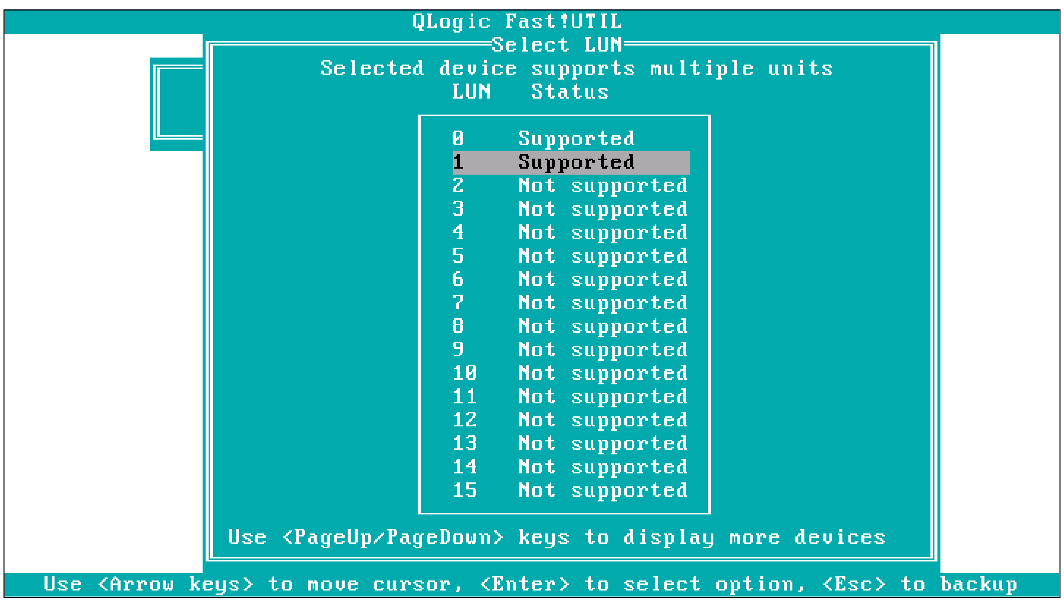


Figure 4-8 QLogic configuration panel: Select LUN

11. Select the boot LUN (in our case it is **LUN 1).** You are taken back to the *Selectable Boot Setting* menu and boot port with the boot LUN is displayed as illustrated in Figure 4-9.

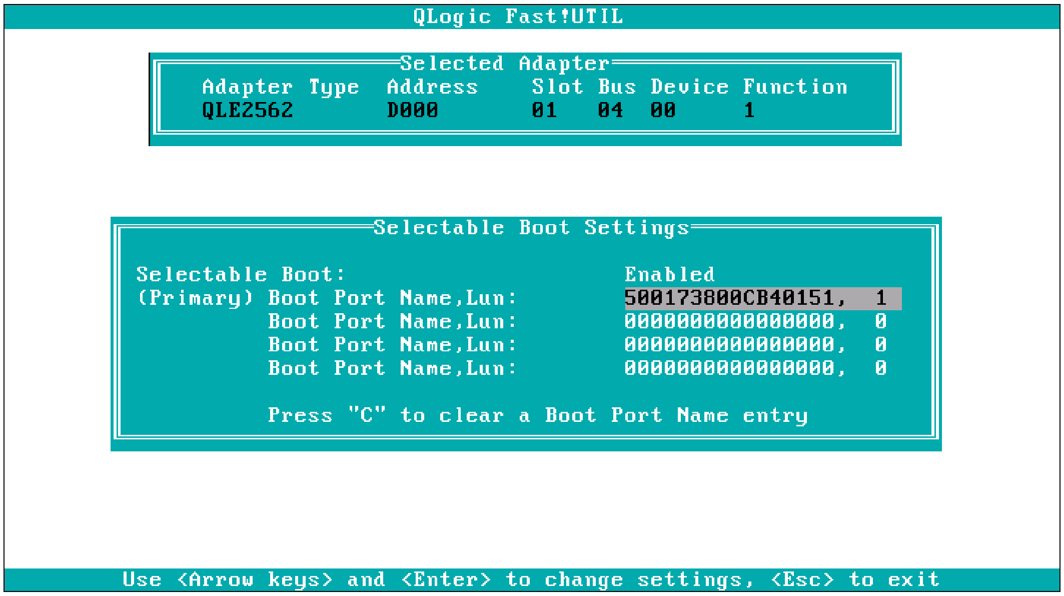


Figure 4-9 QLogic configuration panel: Selectable Boot Settings

12. When all the controllers are added, press Esc to exit (Configuration Setting panel). Press Esc again to get the **Save changes** option, as shown in Figure 4-10.

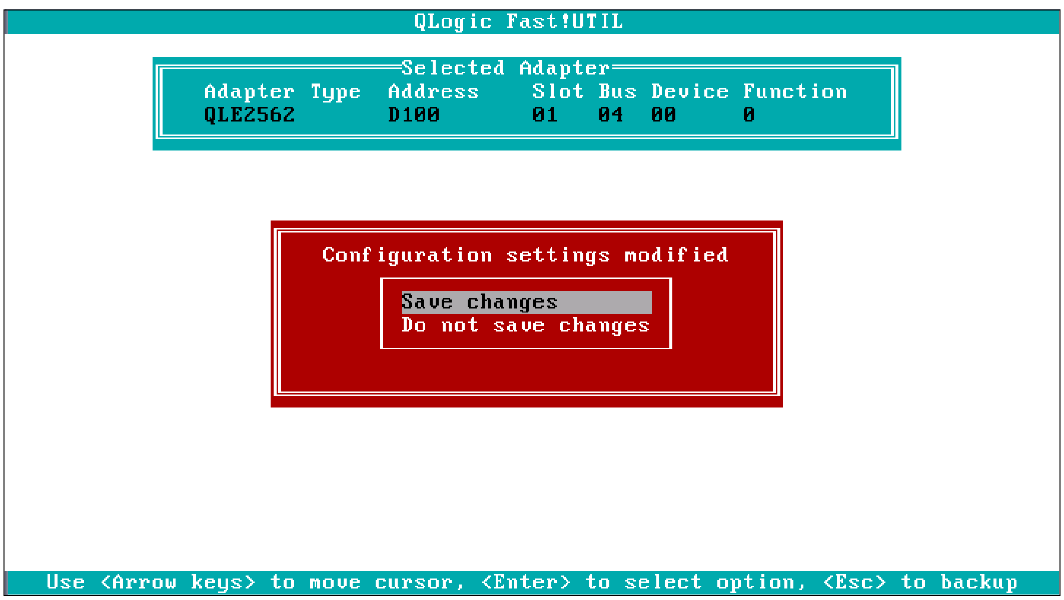


Figure 4-10 QLogic configuration panel: Configuration settings modified

13. Select **Save changes**. It takes you back to the Fast!UTIL option panel.

14. Repeat the steps 2-13 to add additional controllers. Note that any additional controllers must be zoned so that they point to the same boot LUN.

15. When all controllers have been configured, from the **Fast!UTIL Options** menu (Figure 4-3 on page 207), select **Exit Fast!UTIL.** The Exit Fast!UTIL menu is displayed as shown in Figure 4-11.

Figure 4-11 QLogic configuration panel: Exit Fast!UTIL

> **Important:** Depending on your operating system and multipath drivers, you might need to configure multiple ports as "SAN Boot" ports. See your operating system documentation for more information.

Emulex FC HBA configuration

The procedures for setting up your server and HBA to SAN Boot can vary. The procedures in this section are for an Emulex LPe12002 FC HBA (IBM model 42D0494):

1. Boot your server. During the boot process, press **CTRL-E** when prompted to load the configuration utility and display the *Select Host Adapter* menu (see Figure 4-12).

```
        Emulex LightPulse BIOS Utility, UB2.11a0
        Copyright (c) 1997-2008 Emulex. All rights reserved.

        Emulex Adapters in the System:

 1.  42D0494:          PCI Bus, Device, Function (04,00,01)
 2.  42D0494:          PCI Bus, Device, Function (04,00,00)




  Enter a Selection:

Enter <x> to Exit
```

Figure 4-12 Emulex configuration panel: Select Host Adapter

2. You normally see one or more ports. Select a port and press Enter. It takes you to a panel as shown in Figure 4-13. If you will only be enabling the BIOS on one port, then make sure to select the correct port.

> **Tip:** The value associated with the *Port Name* in Figure 4-13 represents that port's World Wide Port Name (WWPN). Notice and record this value for zone planning as described in 4.2.5, "Zoning" on page 219.

3. Select **Configure This Adapter's Parameters**.

```
    Adapter 01:           PCI Bus, Device, Function (04,00,01)

    42D0494:       Mem Base: FE980000  Firmware Version: US1.11A5
    Port Name: 10000000 C9A9263F   Node Name: 20000000 C9A9263F
    Topology: Auto Topology: Loop First (Default)
    The BIOS for this adapter is Disabled

1.  Configure Boot Devices
2.  Configure This Adapter's Parameters




  Enter a Selection: _

Enter <x> to Exit        <d> to Default Values       <Esc> to Previous Menu
```

Figure 4-13 Emulex configuration panel: Host Adapter

4. The *Adapter's Parameters* menu is displayed as shown in Figure 4-14.

```
    Adapter 01:           PCI Bus, Device, Function (04,00,01)

    42D0494:       Mem Base: FE980000  Firmware Version: US1.11A5
    Port Name: 10000000 C9A9263F   Node Name: 20000000 C9A9263F
    Topology: Auto Topology: Loop First (Default)
    The BIOS for this adapter is Disabled

1.  Enable or Disable BIOS
2.  Change Default ALPA of this Adapter
3.  Change PLOGI Retry Timer (+Advanced Option+)
4.  Topology Selection (+Advanced Option+)
5.  Enable or Disable Spinup Delay (+Advanced Option+)
6.  Auto Scan Setting (+Advanced Option+)
7.  Enable or Disable EDD 3.0 (+Advanced Option+)
8.  Enable or Disable Start Unit Command (+Advanced Option+)
9.  Enable or Disable Environment Variable (+Advanced Option+)
10. Enable or Disable Auto Boot Sector (+Advanced Option+)
11. Link Speed Selection (+Advanced Option+)

  Enter a Selection:

Enter <x> to Exit          <Esc> to Previous Menu
```

Figure 4-14 Emulex configuration panel: Adapter's Parameters

5. On the *Adapter's Parameters* panel, select **Enable or Disable BIOS**.
6. On the Enable or Disable BIOS panel, as seen in Figure 4-15, enter **1** to enable the HBA BIOS.

```
    Adapter 02:           PCI Bus, Device, Function (04,00,00)


    The BIOS is Enabled!!


    Enable Press 1, Disable Press 2:1_


Enter <x> to Exit          <Esc> to Previous Menu
```

Figure 4-15 Emulex configuration panel: Enable or Disable Bios

7. After the BIOS has been enabled, press **x** to exit the Adapter's Parameters panel and return to the Host Adapter panel.
8. From the Host Adapter panel, as seen in Figure 4-16, select **Configure Boot Devices**.

```
    Adapter 02:           PCI Bus, Device, Function (04,00,00)

    42D0494:       Mem Base: FE984000  Firmware Version: US1.11A5
    Port Name: 10000000 C9A9263E   Node Name: 20000000 C9A9263E
    Topology: Auto Topology: Loop First (Default)
    The BIOS for this adapter is Enabled

1.  Configure Boot Devices
2.  Configure This Adapter's Parameters









  Enter a Selection:

Enter <x> to Exit        <d> to Default Values       <Esc> to Previous Menu
```

Figure 4-16 Emulex configuration panel: Host Adapter

9. From the Configure Boot Devices panel, as seen in Figure 4-17, select **1** to configure the Primary Boot entry and press Enter.

```
    Adapter 02:  S_ID: 050037    PCI Bus, Device, Function (04,00,00)


    List of Saved Boot Devices:

1. Unused   DID:000000 WWPN:00000000 00000000 LUN:00  Primary Boot
2. Unused   DID:000000 WWPN:00000000 00000000 LUN:00
3. Unused   DID:000000 WWPN:00000000 00000000 LUN:00
4. Unused   DID:000000 WWPN:00000000 00000000 LUN:00
5. Unused   DID:000000 WWPN:00000000 00000000 LUN:00
6. Unused   DID:000000 WWPN:00000000 00000000 LUN:00
7. Unused   DID:000000 WWPN:00000000 00000000 LUN:00
8. Unused   DID:000000 WWPN:00000000 00000000 LUN:00






    Select a Boot Entry: _


Enter <x> to Exit          <Esc> to Previous Menu
```

Figure 4-17 Emulex configuration panel: Configure Boot Devices

10. From the Boot Entry panel, as Figure 4-18, identify the XIV LUN to be used as the SAN Boot LUN. Enter the two-digit number of the device and press Enter.

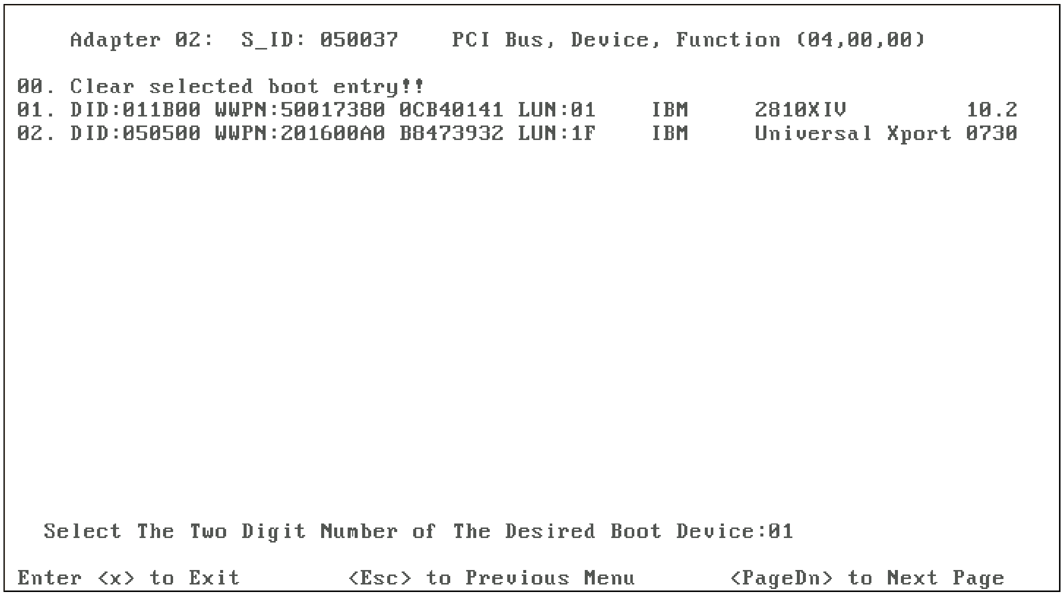


Figure 4-18 Emulex configuration panel: Boot Entry

11. From the Boot Entry panel, enter the two digit hex value for the LUN, as seen in Figure 4-19, to complete the Boot Entry process.

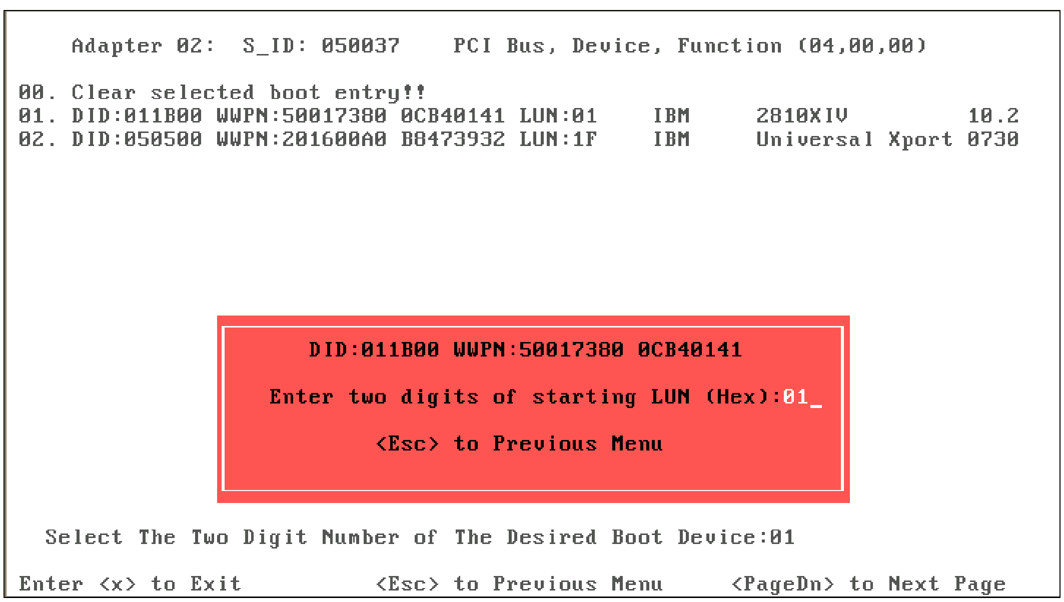


Figure 4-19 Emulex configuration panel: Boot Entry confirmation

12. After the Boot Entry confirmation panel, the Boot number panel, as seen in Figure 4-20, is displayed. Enter the two-digit identifier for the LUN, in this case **01**, and press Enter.

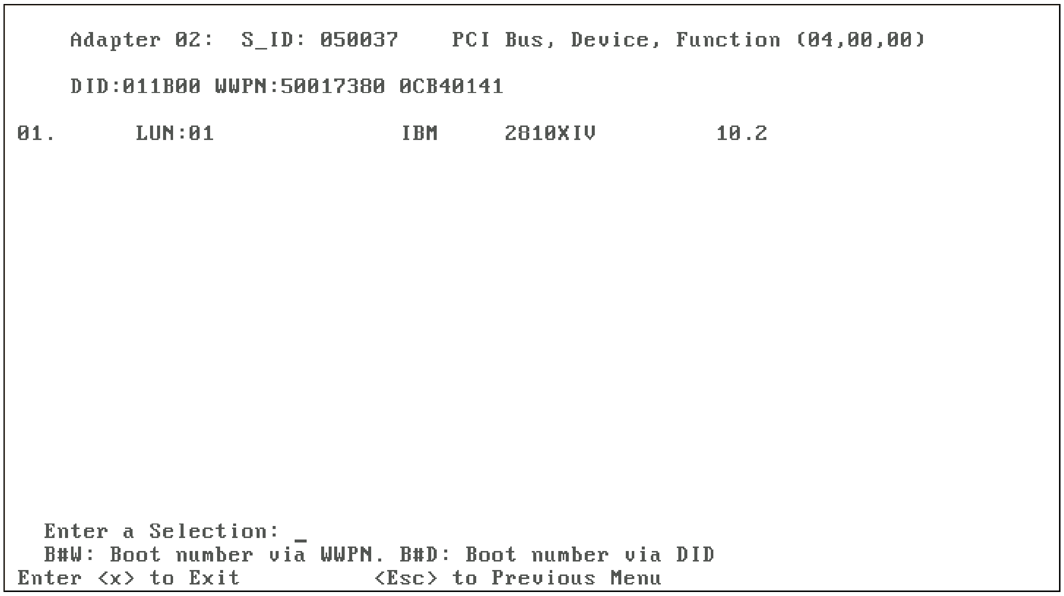


Figure 4-20 Emulex configuration panel: Boot number

13. A panel, as seen in Figure 4-21, is displayed to specify how to boot the device. Select **1** (Boot this device by WWPN) and press Enter.

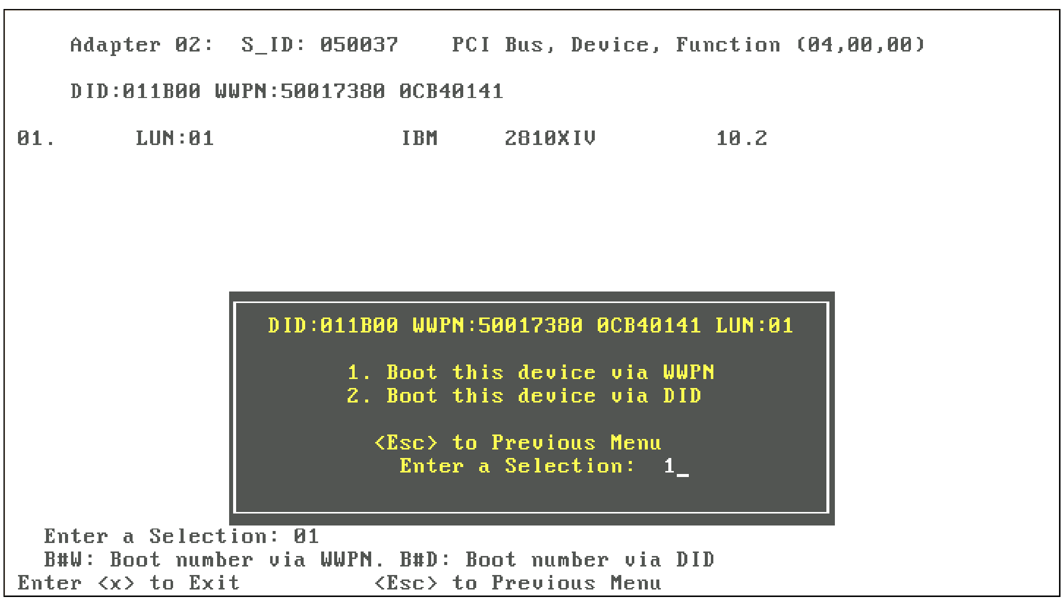


Figure 4-21 Emulex configuration panel: Boot device id

14. Repeat the steps 2-13 to configure additional controllers. Note that any additional controllers must be zoned so that they point to the same boot LUN.

15. After all Emulex controllers have been configured, use **x** to Exit the Emulex panels. As seen in Figure 4-22, when the Emulex BIOS has been exited, press **Y** to reboot the system.

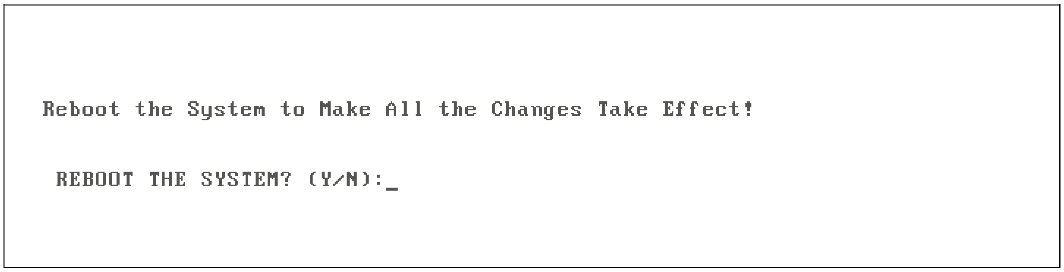


Figure 4-22 Emulex configuration panel: Reboot request

4.2.3 Identification of XIV FC ports

Identification of a port is required for setting up the zoning, to aid with any modifications that might be required, or to assist with problem diagnosis. The unique name that identifies an FC port is called the World Wide Port Name (WWPN).

As seen in Example 4-1, the easiest way to get a record of all the WWPNs on the XIV is to use the XIV Command Line Interface (XCLI). Access to the XIV management tools is described in 4.3.1, “XIV Management Tools” on page 221

Example 4-1 Displaying XIV FC ports using XCLI

```
>>fc_port_list
Component ID   Status   Currently Functioning   WWPN               Port ID   Role
1:FC_Port:4:1  OK       yes                     500173800CB40140   00030500  Target
1:FC_Port:4:2  OK       yes                     500173800CB40141   00011B00  Target
1:FC_Port:4:3  OK       yes                     500173800CB40142   00B0D500  Target
1:FC_Port:4:4  OK       yes                     500173800CB40143   00000000  Initiator
1:FC_Port:5:1  OK       yes                     500173800CB40150   00040300  Target
1:FC_Port:5:2  OK       yes                     500173800CB40151   00022300  Target
1:FC_Port:5:3  OK       yes                     500173800CB40152   00B0DF00  Target
1:FC_Port:5:4  OK       yes                     500173800CB40153   00000000  Initiator
```

However, as seen in Figure 4-23, it is also possible to display the XIV FC ports by clicking the **Backview** arrow that is displayed at the bottom and to the right of the XIV appliance display. The XIV Patch Panel is displayed. Port information can be seen by pointing to individual ports.

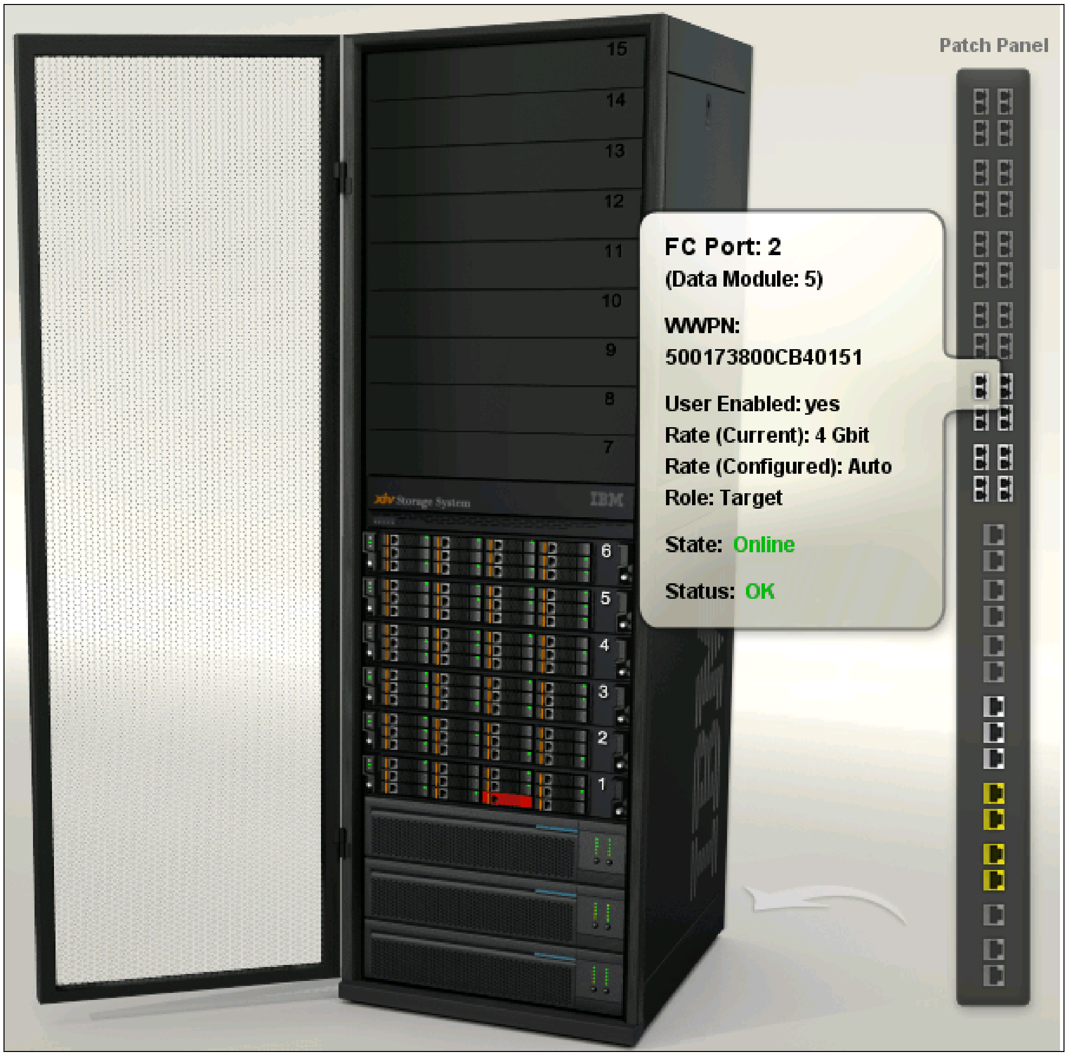


Figure 4-23 XIVGUI Patch Panel displayed after clicking the Backview arrow

4.2.4 FC fabric configurations

Several Fibre Channel fabric configurations are technically possible, and they vary in terms of their cost and the degree of flexibility, performance, and reliability that they provide.

Redundant (high availability) configurations

Production environments must always have a redundant (high availability) configuration. There must be no single points of failure. Hosts must have as many HBAs as needed to support the operating system, application, and overall performance requirements.

As an example of one possible FC fabric configuration, we provide the following redundant configuration for reference.

Redundant configuration example

The fully redundant configuration is illustrated in Figure 4-24.

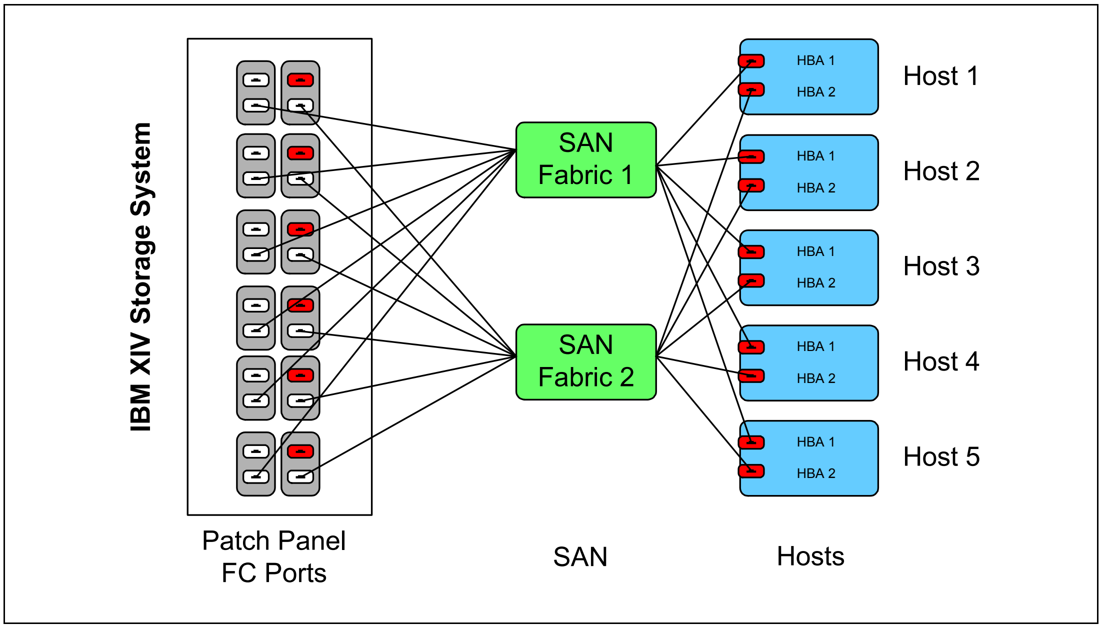


Figure 4-24 FC fully redundant configuration

This configuration has the following features:

- Each host is equipped with dual HBAs. Each HBA (or HBA port) is connected to one of two FC switches.
- Each of the FC switches has a connection to a separate FC port of each of the six interface modules.
- Each LUN has 12 paths.

4.2.5 Zoning

Zoning is mandatory when connecting FC hosts to an XIV Storage System. Zoning is configured on the SAN switch and is a boundary whose purpose is to isolate and restrict FC traffic to only those HBAs within a given zone.

Hard zones and soft zones

A zone can be either a hard zone or a soft zone. Hard zones group HBAs depending on the physical ports they are connected to on the SAN switches. Soft zones group HBAs depending on the World Wide Port Names (WWPNs) of the HBA. Each method has its merits and you need to determine which is right for your environment.

Correct zoning helps avoid many problems and makes it easier to trace cause of errors. Here are some examples of why correct zoning is important:

- An error from an HBA that affects the zone or zone traffic will be isolated.
- Disk and tape traffic must be in separate zones, as they have different characteristics. Being in the same zone can cause performance problems or have other adverse effects.
- Any change in the SAN fabric, such as a change caused by a server restarting or a new product being added to the SAN, triggers a *Registered State Change Notification (RSCN)*. An RSCN requires that any device that can “see” the affected or new device to acknowledge the change, interrupting its own traffic flow.

Zoning guidelines

There are many factors that affect zoning, such as host type, number of HBAs, HBA driver, operating system, and applications. Therefore, it is not possible to provide a solution to cover every situation. The following list gives some guidelines, which can help you to avoid reliability or performance problems. However, you also need to review documentation regarding your hardware and software configuration for any specific factors that must be considered:

- Each zone (excluding those for SAN Volume Controller (SVC)) must have one initiator HBA (the host) and multiple target HBAs (the XIV Storage System).
- Each host (excluding SVC) must have two paths per HBA unless there are other factors dictating otherwise.
- Each host must connect to ports from at least two interface modules.
- Do not mix disk and tape traffic on the same HBA or in the same zone.

For more in-depth information about SAN zoning, see the IBM Redbooks publication, *Introduction to Storage Area Networks*, SG24-5470.

An example of soft zoning using the “single initiator - multiple targets” method is illustrated in Figure 4-25.

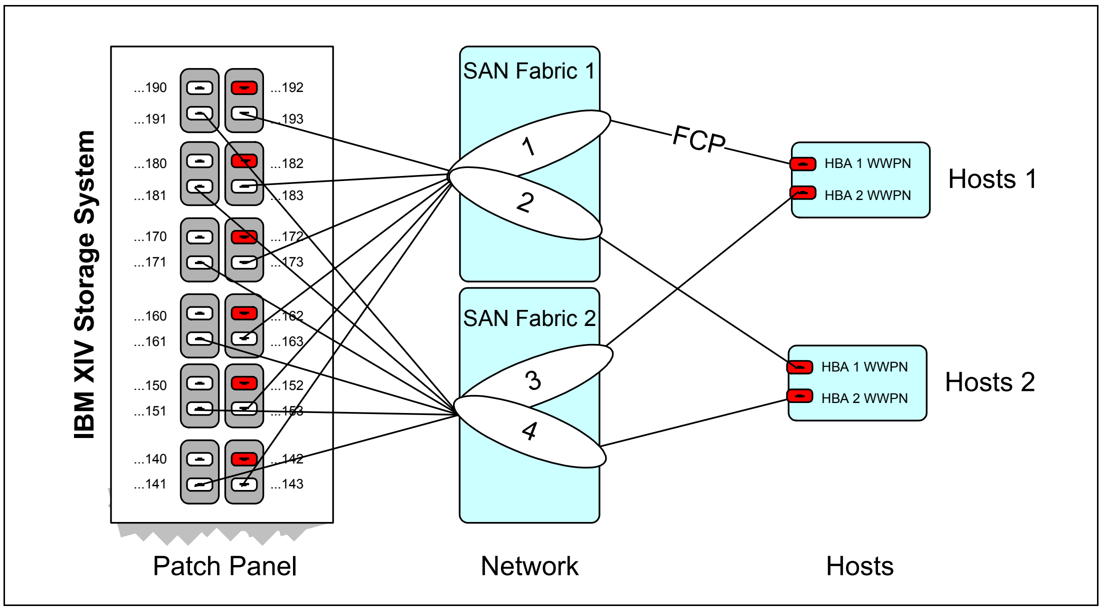


Figure 4-25 FC SAN zoning: single initiator - multiple target

> **Attention:** Use a single initiator and multiple target zoning scheme. Do not share a host HBA for disk and tape access.

Zoning considerations also include spreading the I/O workload evenly between the different interfaces. For example, for a host equipped with two single port HBAs, you must connect one HBA port to one port on modules 4,5,6, and the second HBA port to one port on modules 7,8,9.

4.3 XIV Management Tools and Support Tools

The following management and support tools are available for XIV systems.

4.3.1 XIV Management Tools

To facilitate the configuration and monitoring of one or more XIV systems, IBM provides XIV Management Tools.

IBM XIV Management Tools include the following options:

- IBM XIV Graphical User Interface (XIVGUI)
- IBM XIV Command Line Interface (XCLI)
- IBM XIV Online Monitoring Tool (XIVTop)

XIV Management Tools can be downloaded from the following website:

http://www.ibm.com/support/search.wss?q=ssg1*&tc=STJTAG+HW3E0&rs=1319&dc=D400&dtm

4.3.2 Host Attachment Kits

Starting with version 10.1.x of the XIV system software, IBM also provides updates to all of the Host Attachment Kits (version 1.1 or later). It is *mandatory* to install the Host Attachment Kit to be able to get support from IBM, even if it might not be technically necessary for some operating systems. Host Attachment Kits (HAKs) are built on a Python framework with the intention of providing a consistent look and feel across various operating system (OS) platforms. The following features are included:

- Provides backwards compatibility with versions 10.0.x of the XIV system software
- Validates patch and driver versions
- Sets up multipathing
- Adjusts system tunable parameters (if required) for performance
- Provides installation wizard
- Includes management utilities
- Includes support and troubleshooting utilities

Host Attachment Kits can be downloaded from the following website:

http://www.ibm.com/support/search.wss?q=ssg1*&tc=STJTAG+HW3E0&rs=1319&dc=D400&dtm

4.4 Logical configuration for host connectivity

This section describes the tasks required to define a volume (LUN) and assign it to a host:

1. Create a storage pool.
2. Create a volume within the storage pool.
3. Define a host.
4. Map the volume to the host.
5. Add FC ports to the host.

Important: For the host system to effectively see and use the LUN, additional and operating system specific configuration tasks are required. The tasks are described in subsequent chapters of this book according to the operating system of the host that is being configured.

4.4.1 Creating an XIV storage pool

In order to perform a SAN Boot installation, an XIV volume is required. However, prior to creating an XIV volume, an XIV storage pool must exist, to which the volume will be associated.

Using the XCLI interface, existing XIV storage pools can be displayed by executing the **pool_list** command. However, as can be seen in Figure 4-26, the existing XIV storage pools can also be displayed in the XIV GUI by referencing the pull-down shown.

Figure 4-26 XIVGUI Pull-down menu to display storage pools

If a new XIV storage pool needs to be created, it can be done using either the XCLI or XIVGUI interfaces. Example_1-2 shows use of the **pool_create** XCLI command to create a 1013 GB storage pool with a 103 GB snapshot size.

Example 4-2 Adding a storage pool using the XCLI pool_create command

```
>>pool_create pool=RB_Storage_Pool size=1013 snapshot_size=103
Command executed successfully.
```

For additional information regarding XIV XCLI commands, refer to the IBM XIV Storage System Information Center:

http://publib.boulder.ibm.com/infocenter/ibmxiv/r2/index.jsp

Figure 4-27 shows how to initiate the creation of the same XIV storage pool using the GUI **Add Pool** from the Storage Pools panel.

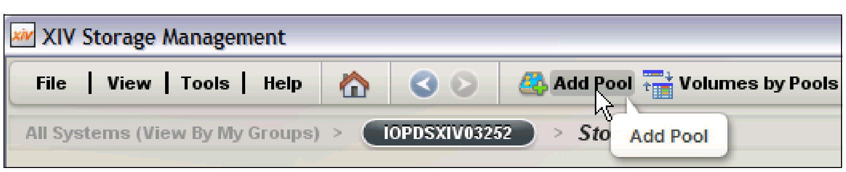


Figure 4-27 Adding a storage pool using the XIVGUI menu option

Figure 4-28 shows the XIVGUI panel that is presented to specify the attributes of the storage pool to be created.

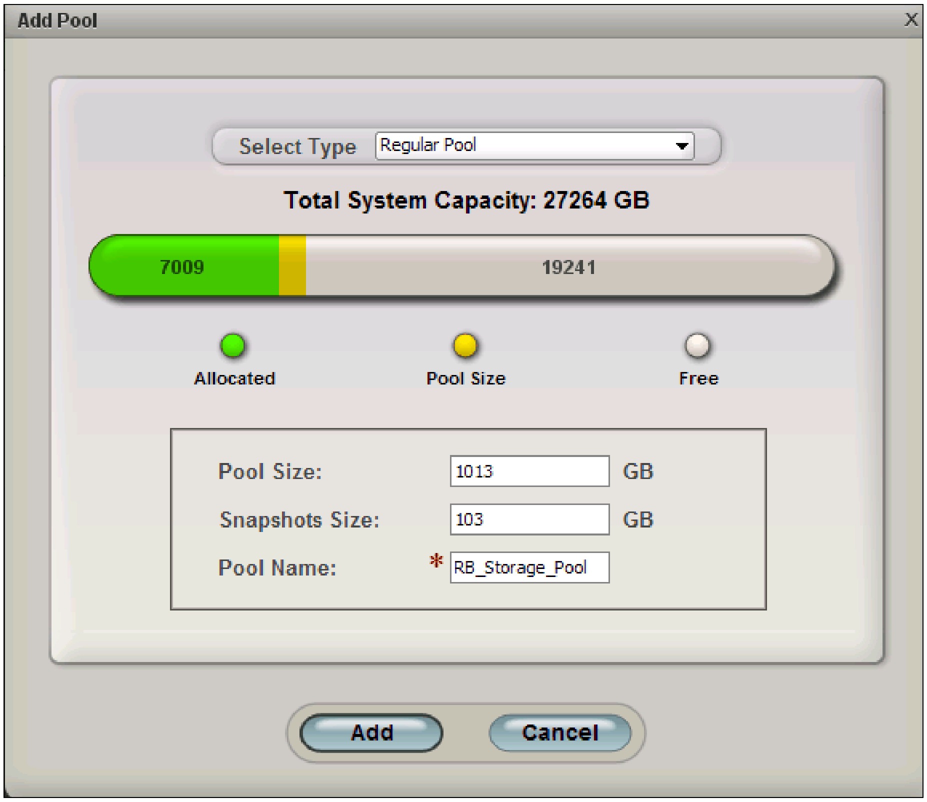


Figure 4-28 XIVGUI Add Pool panel

4.4.2 Creating an XIV Volume

After an existing or newly created XIV storage pool has been identified for an XIV Volume to be associated with, the next step is to identify an XIV Volume to be used for SAN Boot.

Using the XCLI interface, existing XIV Volumes can be displayed by executing the **vol_list** command. However, as can be seen in Figure 4-29, the existing XIV Volumes can also be displayed in the XIVGUI by referencing the pull-down shown.

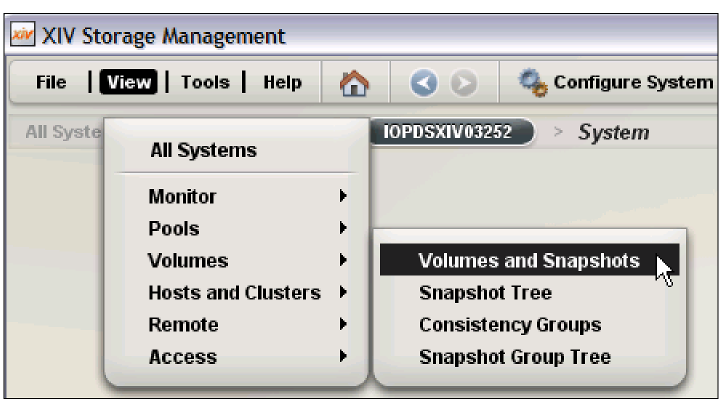


Figure 4-29 XIVGUI Pull-down menu to display Volumes

If a new volume is to be created, Example 4-3 shows how it can be performed using the **vol_create** XCLI command.

Example 4-3 Creating an XIV Volume using the XCLI vol_create command

```
vol_create size=51 pool=RB_Storage_Pool vol=SAN_Boot_1
```

Figure 4-30 shows how to initiate the creation of the same XIV storage pool using the XIVGUI **Add Volumes** from the Storage Pools panel.

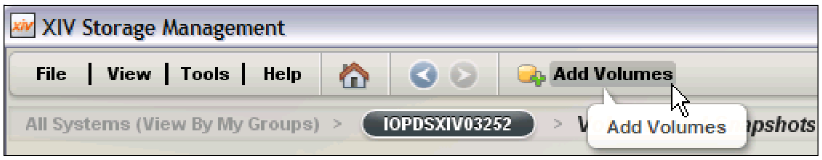


Figure 4-30 XIVGUI Add Volumes button

Figure 4-31 shows the XIVGUI panel that is presented to specify the attributes of the volume to be created.

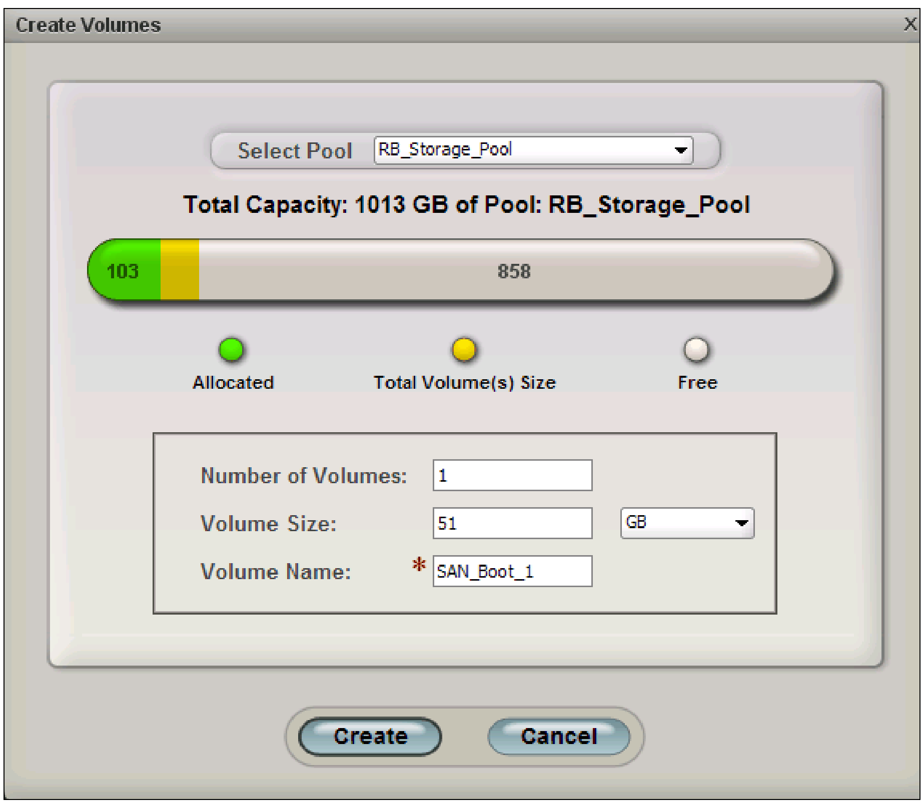


Figure 4-31 XIVGUI Add Volume panel

4.4.3 Defining an XIV Host

The following section outlines the steps necessary to create an XIV Host definition to represent the host that will be performing the SAN Boot installation process.

Using the XCLI interface, an XIV Host can be defined by executing the **host_define** command as seen in Example 4-4.

Example 4-4 Creating an XIV Host using the XCLI host_define command

```
host_define host=SAN_Boot_Host_1
```

If a new Host is to be defined, it can also be performed using the XIVGUI interface. Figure 4-32 on page 225 shows how the pull-down menu can be used to access the **Hosts and Clusters** display.

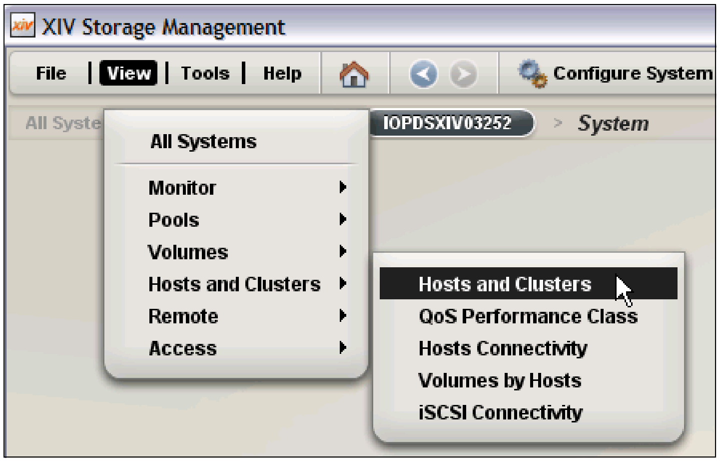


Figure 4-32 XIVGUI Pull-down menu to display Hosts

Figure 4-33 shows how to initiate the creation of the same XIV storage pool using the GUI **Add Host** from the Storage Pools panel.

Figure 4-33 XIVGUI Add Host button

Figure 4-34 shows the XIVGUI panel that is presented to specify the attributes of the host to be created.

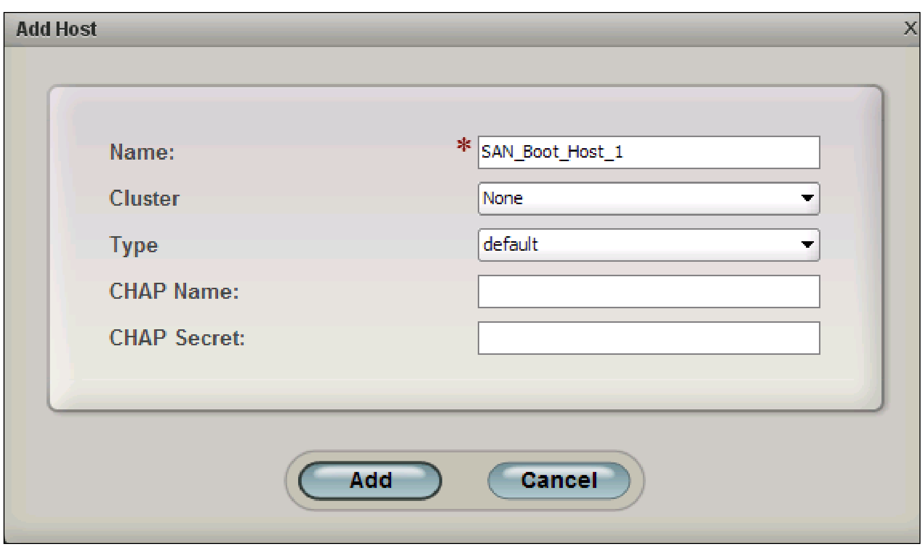


Figure 4-34 XIVGUI Add Host panel

4.4.4 Assigning an XIV Volume to an XIV Host

The following section shows how to associate a previously created XIV Volume to an existing XIV Host.

Using the XCLI interface, an XIV Host can be defined by executing the **map_vol** command as seen in Example 4-5.

Example 4-5 Assigning an XIV Volume to a host using the XCLI map_vol command

```
map_vol host=SAN_Boot_Host_1 vol=SAN_Boot_1 lun=1
```

Figure 4-35 shows how the pull-down menu can be used to access the **Hosts and Clusters** display.

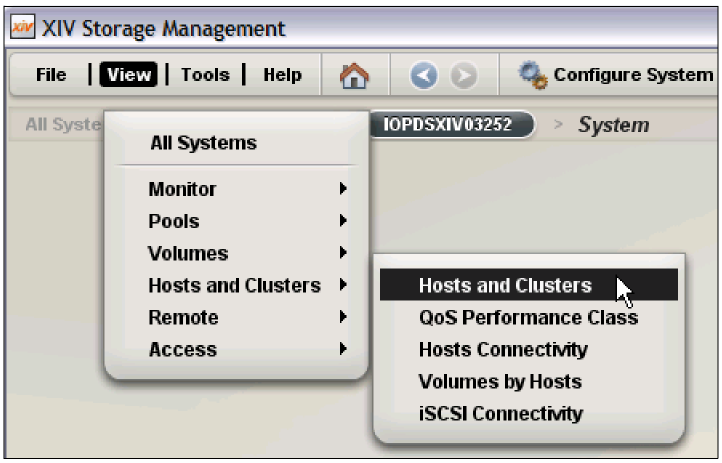


Figure 4-35 XIVGUI Pull-down menu to display Hosts

To associate an XIV Volume to an XIV Host, after clicking the desired XIV Host, a right-click will display the menu seen in Figure 4-36. Select **Modify LUN Mapping** from the menu.

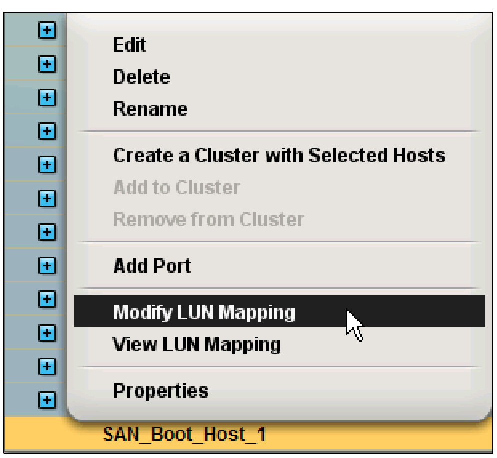


Figure 4-36 XIVGUI Host menu options with right click

Figure 4-37 shows the **LUN Mapping** panel. To complete the mapping process, select the appropriate volume from the left side of the panel and click the **Map** button. It is recommended that the SAN Boot volume be assigned to LUN 1.

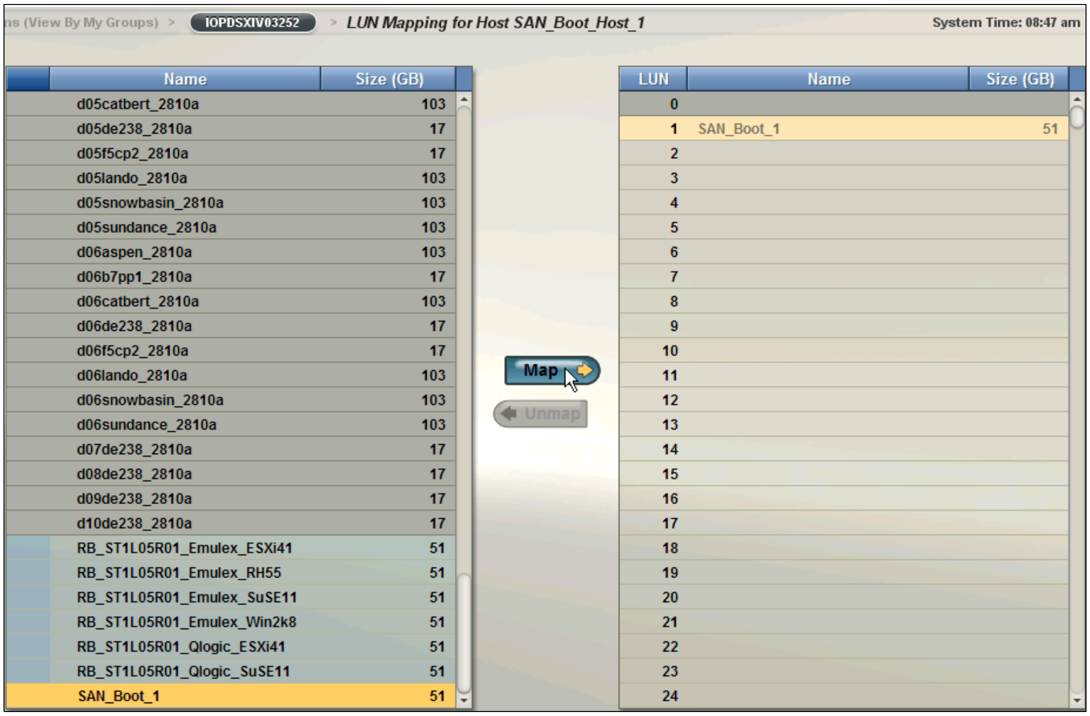


Figure 4-37 XIVGUI LUN Mapping panel

4.4.5 Adding FC ports to the XIV Host

The final XIV configuration step that must be taken is to assign the FC WWPNs for the HBA ports in the host to the XIV Host definition. The host's HBA WWPNs can be identified by various processes. See 4.2.2, "Host FC HBA configuration" on page 206 for an example of how the WWPNs can be identified from within the HBA BIOS.

Using the XCLI interface, an XIV Host can have an FC WWPN mapped to it using the **`host_add_port`** command as seen in Example 4-6.

Example 4-6 Creating an XIV Volume using the XCLI vol_create command

```
host_add_port host=SAN_Boot_Host_1 fcaddress=10000000C98E437E
```

To perform the same operation from the XIVGUI, Figure 4-38 shows how the pull-down menu can be used to access the **Hosts and Clusters** display.

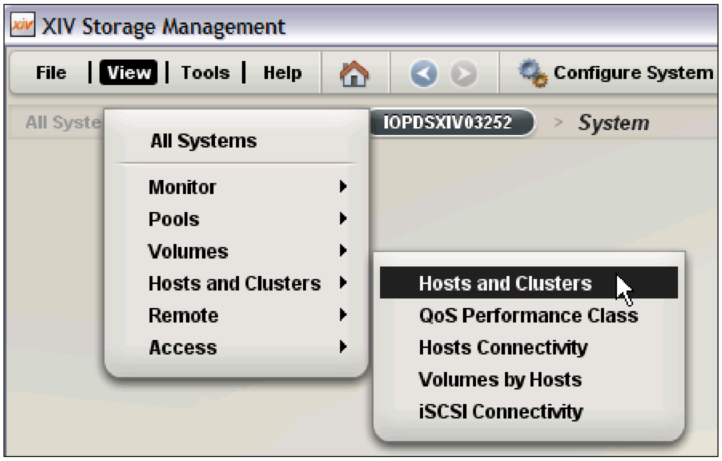


Figure 4-38 XIVGUI Pull-down menu to display Hosts

To add an FC WWPN address to an XIV Host, after clicking the desired XIV Host, a right-click will display the menu seen in Figure 4-39. Select **Add Port** from the menu.

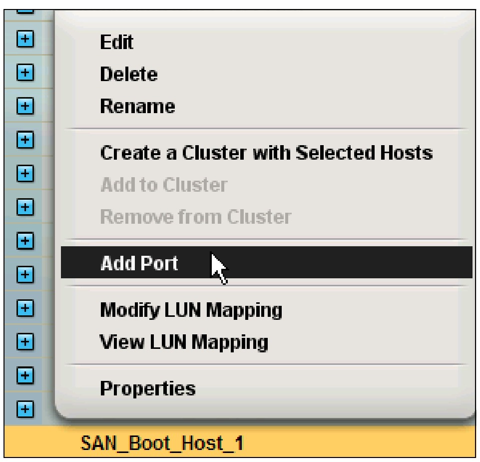


Figure 4-39 XIVGUI Host menu options with right-click

Figure 4-40 shows the XIVGUI panel that is presented to specify the port type of FC and to assign the WWPN. The port name can be entered manually, or if the host HBA has successfully logged into the SAN that includes the XIV, the WWPN should be selectable by the pull-down.

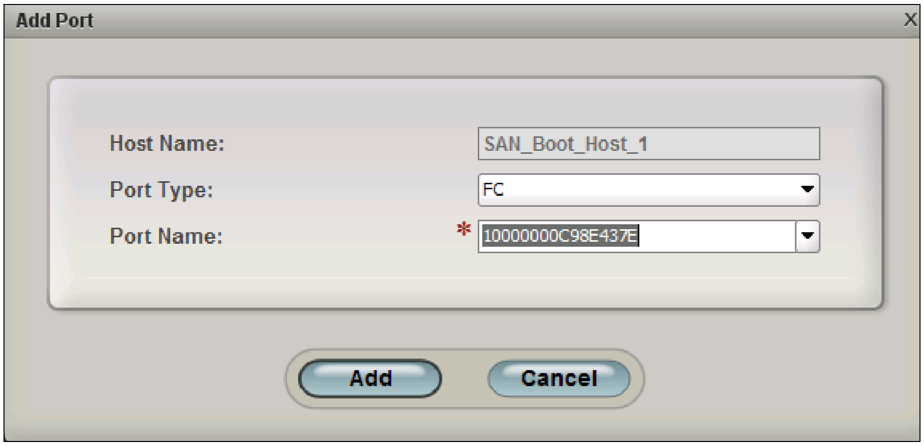


Figure 4-40 XIVGUI Add Port panel

4.5 Installation of Microsoft Windows 2008 R2 for SAN Boot

This section provides instructions for doing a SAN Boot installation of the Microsoft Windows 2008 R2 operating system using Fibre Channel connectivity to XIV.

4.5.1 Overview

This section covers prerequisites and restrictions that apply to the installation.

Microsoft support for booting from a SAN: Refer to the following article for information about Microsoft support for booting from a SAN:

`http://support.microsoft.com/kb/305547`

Prerequisites

To successfully SAN Boot a Windows host from XIV, a number of prerequisites must be met. Here is a generic list. However, your environment might have additional requirements:

- Review 4.2, “SAN configuration planning” on page 205
- Review 4.3, “XIV Management Tools and Support Tools” on page 221
- Review 4.4, “Logical configuration for host connectivity” on page 221

Installation restrictions

For the installation of Windows, the host should only be presented with a single path to the XIV volume that Windows will be installed on. It is done to prevent issues that can arise when Windows encounters duplicate drive devices that are not managed by multipathing software (which is installed and configured after the initial operating system installation).

Depending on the implemented SAN configuration, the host can be presented a single path to the XIV volume by temporarily updating the SAN zoning and XIV Host port configurations to present only a single path to the XIV volume through a single host HBA port.

Attention: For the installation of Windows, the host should only be presented with a single path to the XIV volume that Windows will be installed on.

4.5.2 Operating system installation considerations

This section describes any specific portions of the Windows 2008 R2 installation process that relate to XIV or SAN Boot considerations.

Installing HBA drivers

Windows 2008 includes drivers for many HBAs, however, it is likely that they will not be the latest version for your HBA. You should install the latest available driver that is supported. HBAs drivers are available from IBM, Emulex, and QLogic websites. They all come with instructions that should be followed to complete the installation.

Newer versions of HBA drivers

Figure 4-41 shows the Windows installation panel that allows newer versions of the HBA drivers to be loaded as well as the ability to select the disk device to use during the installation.

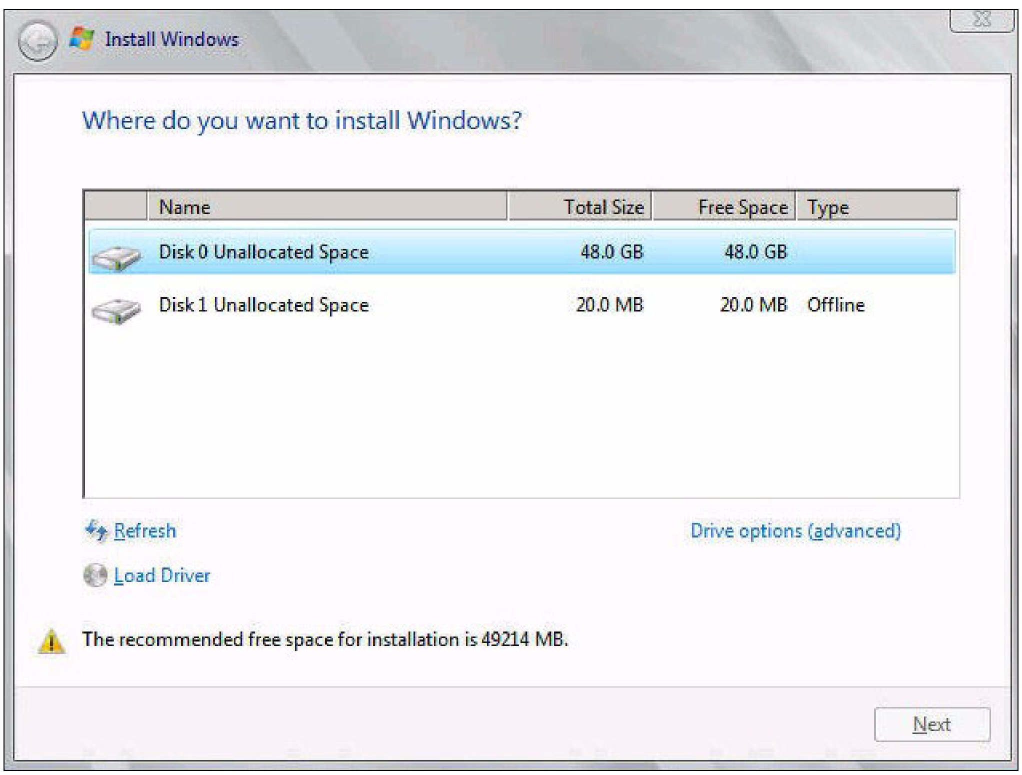


Figure 4-41 Windows 2008 R2 installation panel

After the XIV volume has been detected and selected, there are no other special SAN Boot considerations for the remainder of the installation process. After the installation completes, the next step is to install the XIV Host Attachment Kit.

4.5.3 Post operating system installation considerations

The Windows 2008 Host Attachment Kit (HAK) must be installed to gain access to XIV storage. Note that there are different versions of the Host Attachment Kit for different versions of Windows, and it is further sub-divided into 32-bit and 64-bit versions. The Host Attachment Kit can be downloaded from the following website:

http://www.ibm.com/support/search.wss?q=ssg1*&tc=STJTAG+HW3E0&rs=1319&dc=D400&dtm

Support: The installation of the XIV HAK is a mandatory IBM support prerequisite.

The XIV HAK utilizes Microsoft's native multipath functionality for path redundancy.

Microsoft provides a multi-path framework and development kit called the Microsoft Multi-path I/O (MPIO). The driver development kit allows storage vendors to create Device Specific Modules (DSMs) for MPIO and to build interoperable multi-path solutions that integrate tightly with the Microsoft Windows family of products.

MPIO allows the host HBAs to establish multiple sessions with the same target LUN but present it to Windows as a single LUN. The Windows MPIO drivers enables a true active/active path policy allowing I/O over multiple paths simultaneously.

Further information about Microsoft MPIO support is available at this website:

http://download.microsoft.com/download/3/0/4/304083f1-11e7-44d9-92b9-2f3cdbf01048/mpio.doc

Starting with version 1.5.0, the Host Attachment Kit for Windows automatically adds the Multipath I/O feature. Prior to version 1.5.0 of the HAK, you were required to manually add this feature.

Also starting with version 1.5.0 of the HAK, required hot-fixes for the English versions of Windows are also automatically applied. Additional information regarding Windows hot-fixes that relate to the XIV HAK is available within the related XIV HAK Release Notes document.

The following instructions are based on the installation performed at the time of writing. See the instructions in the HAK *Host Attachment Guide* documentation for any changes. The instructions included here show the GUI installation; for command line instructions, see the *Host Attachment Guide*.

Follow these steps:

1. Run the **XIV_host_attach-1.5.3-windows-x64.exe** file. When the setup file is run, it first determines if the python engine *(xpyv)* is required. If required, it will be automatically installed when you click **Install** as shown in Figure 4-42. Proceed with the installation following the installation wizard instructions.

2. After xpyv is installed, the XIV HAK installation wizard is launched. Follow the installation wizard instructions and select the complete installation option. See Figure 4-42.

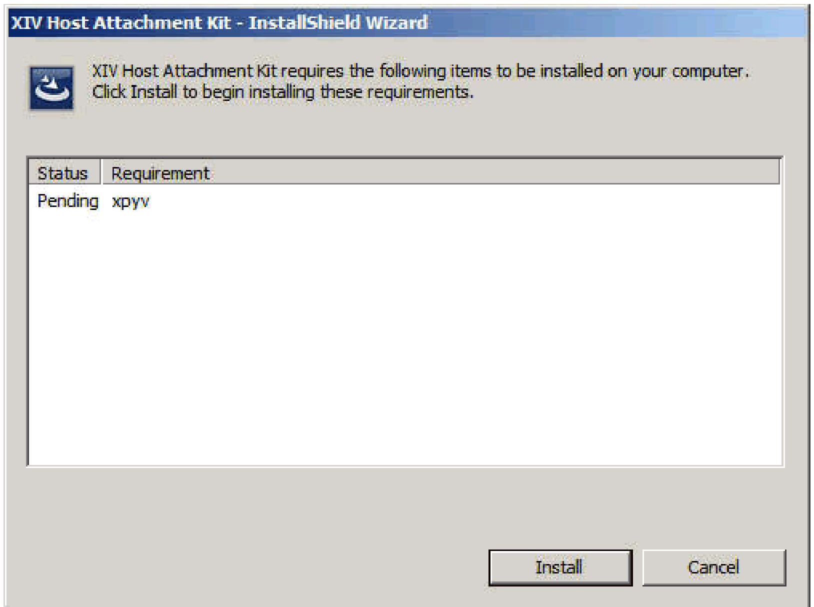


Figure 4-42 Windows HAK: Installation of xpyv

3. Next, you need to run the XIV Host Attachment Wizard as shown in. Click **Finish** to proceed. See Figure 4-43.

Figure 4-43 Windows HAK: Installation of the Host Attachment Kit

4. Answer questions from the XIV Host Attachment wizard as indicated Example 4-7. At the end, you need to reboot your host.

Example 4-7 First run of the XIV Host Attachment wizard

```
Welcome to the XIV Host Attachment wizard, version 1.5.3.
This wizard will assist you to attach this host to the XIV system.
```

```
The wizard will now validate host configuration for the XIV system.
Press [ENTER] to proceed.

-------------------------------------------------------------------------------
Please choose a connectivity type, [f]c / [i]scsi : f
-------------------------------------------------------------------------------
Please wait while the wizard validates your existing configuration...
The wizard needs to configure the host for the XIV system.
Do you want to proceed? [default: yes ]: yes
Please wait while the host is being configured...
-------------------------------------------------------------------------------
A reboot is required in order to continue.
Please reboot the machine and restart the wizard

Press [ENTER] to exit.
```

5. After rebooting, run the XIV Host Attachment Wizard again (from the **Start** button on your desktop, select **All programs**, then select **XIV** and click **XIV Host Attachment Wizard**. Answer the questions prompted by the wizard as indicated in Example 4-8.

Example 4-8 Attaching host over FC to XIV using the XIV Host Attachment Wizard

```
C:\Users\Administrator>xiv_attach
-------------------------------------------------------------------------------
Welcome to the XIV Host Attachment wizard, version 1.5.3.
This wizard will assist you to attach this host to the XIV system.

The wizard will now validate host configuration for the XIV system.
Press [ENTER] to proceed.

-------------------------------------------------------------------------------
Please choose a connectivity type, [f]c / [i]scsi : f
-------------------------------------------------------------------------------
Please wait while the wizard validates your existing configuration...
This host is already configured for the XIV system
-------------------------------------------------------------------------------
Please zone this host and add its WWPNs with the XIV storage system:
10:00:00:00:c9:a9:26:3e: [IBM 42D0494 8Gb 2-Port PCIe FC HBA for System x]: 42D0494
10:00:00:00:c9:a9:26:3f: [IBM 42D0494 8Gb 2-Port PCIe FC HBA for System x]: 42D0494
Press [ENTER] to proceed.

Would you like to rescan for new storage devices now? [default: yes ]: yes
Please wait while rescanning for storage devices...
-------------------------------------------------------------------------------
The host is connected to the following XIV storage arrays:
Serial    Ver    Host Defined  Ports Defined  Protocol  Host Name(s)
7803252   10.2   Yes           All            FC        RB_xseries_01_emulex
This host is defined on all FC-attached XIV storage arrays

Press [ENTER] to proceed.

-------------------------------------------------------------------------------
The XIV host attachment wizard successfully configured this host

Press [ENTER] to exit.
```

At this point, your Windows host should have all the required software to successfully attach to the XIV Storage System. To view the XIV path configurations, execute the **xiv_devlist** command, as shown in Example 4-9.

Example 4-9 View the XIV path configurations

```
C:\Users\Administrator>xiv_devlist
XIV Devices
-------------------------------------------------------------------------------------------------
Device              Size    Paths  Vol Name              Vol Id   XIV Id   XIV Host
-------------------------------------------------------------------------------------------------
\\.\PHYSICALDRIVE2  51.5GB  2/2    RB_ST1L05R01_Emulex_  120      7803252  RB_xseries_01_emulex
                                   Win2k8
-------------------------------------------------------------------------------------------------


Non-XIV Devices
---------------------------------
Device              Size    Paths
---------------------------------
\\.\PHYSICALDRIVE0  21.0MB  N/A
---------------------------------
\\.\PHYSICALDRIVE1  21.0MB  N/A
---------------------------------
```

For additional XIV considerations, see 4.10.3, “Windows 2008 R2 best practices” on page 267.

4.6 Installation of ESXi 4.x for SAN Boot

This section provides instructions for doing a SAN Boot installation of the VMware ESXi 4.1 operating system using Fibre Channel connectivity to XIV.

4.6.1 Prerequisites

To successfully SAN Boot a Windows host from XIV, a number of prerequisites must be met. Here is a generic list. However, your environment might have additional requirements:

- Review 4.2, “SAN configuration planning” on page 205
- Review 4.3, “XIV Management Tools and Support Tools” on page 221
- Review 4.4, “Logical configuration for host connectivity” on page 221

4.6.2 Operating system installation considerations

During the installation of ESXi 4.1, Figure 4-44 shows the panel that relates to the disk storage to utilize for the installation. ESXi is multipath-aware during the installation process and presents a single instance of the XIV volume as a disk device. Select the XIV volume and press Enter to continue the installation process.

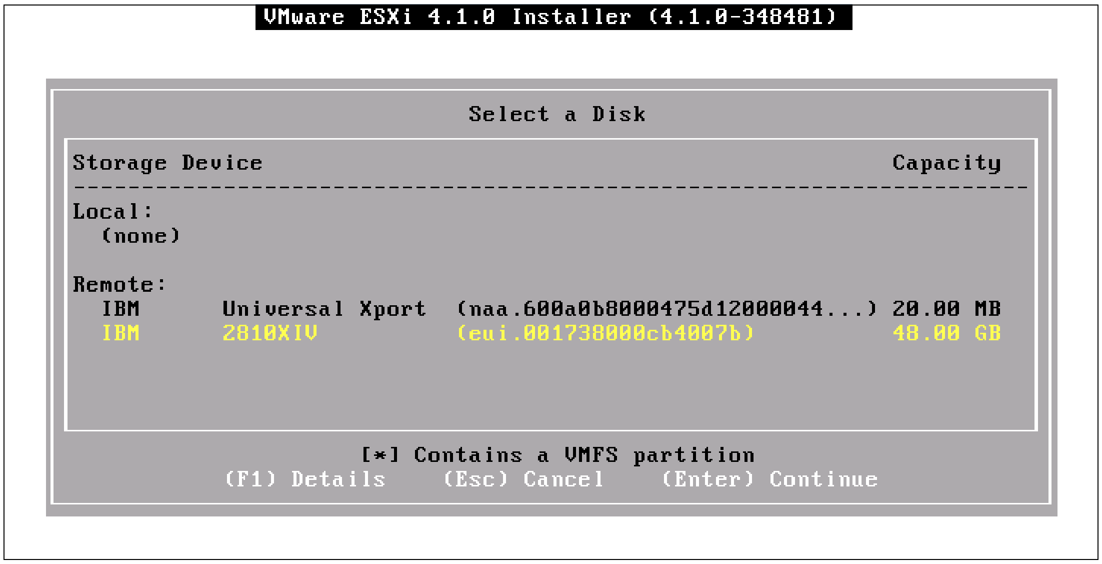


Figure 4-44 VMware ESXi 4.1: Select a Disk

4.6.3 Post operating system installation considerations

After the installation of ESXi 4.1 has completed, additional VMware post-installation steps must be performed. Although these steps are outside the scope of this SAN Boot installation description, those steps include the following actions:

- Configure an ESXi root password.
- Configure the ESXi network connectivity.
- Install VMware client software (that is, VSphere client) on an alternate host.

4.6.4 Verification of FC Multipath Connectivity

To verify the FC connections being used by ESXi, Figure 4-45 shows how it can be done from the VMware VSphere client. After the client is connected to the newly installed ESXi host, this information can be viewed from the *Summary* panel. Select the appropriate **Datastore** and right-click it to display its properties.

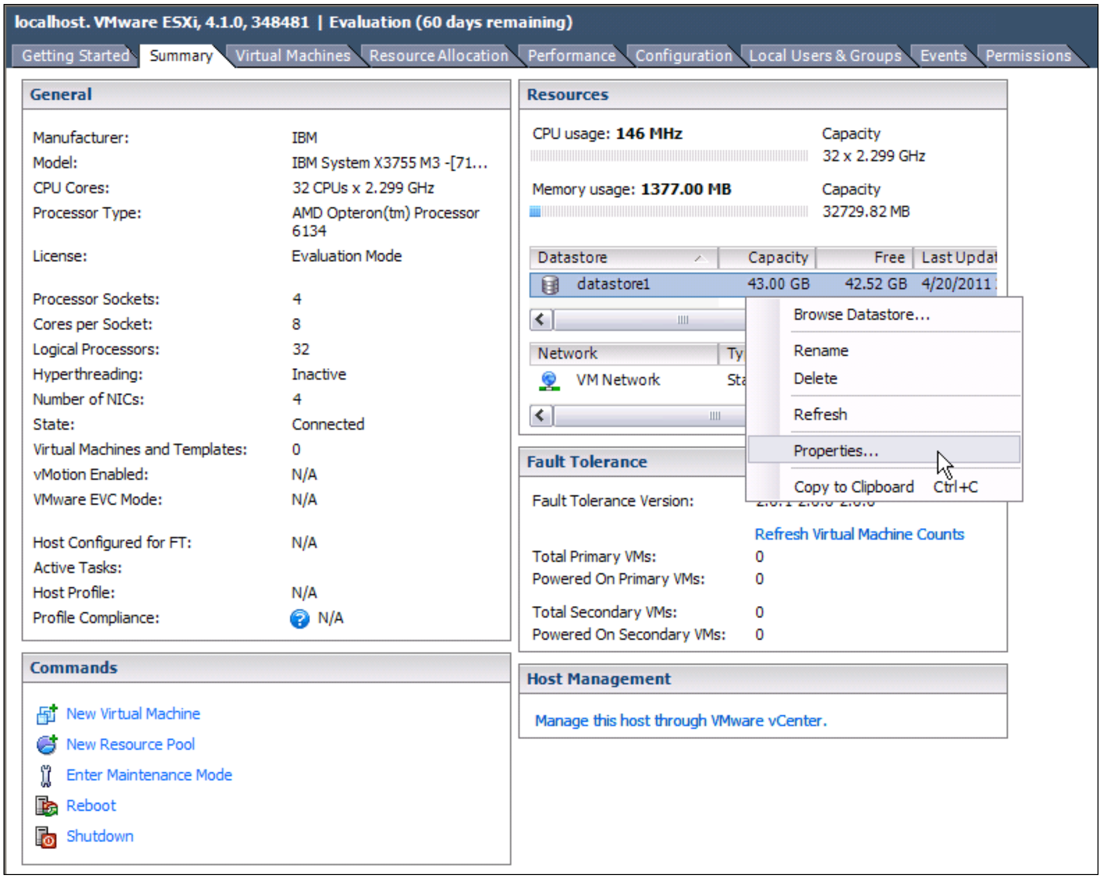


Figure 4-45 VMware VSphere client: Summary panel

Figure 4-46 displays an example of a datastore that represents the SAN Boot XIV volume that was used during the installation of ESXi 4.1. To display explicit information about the Fibre Channel (FC) paths associated with that datastore, click **Manage Paths**.

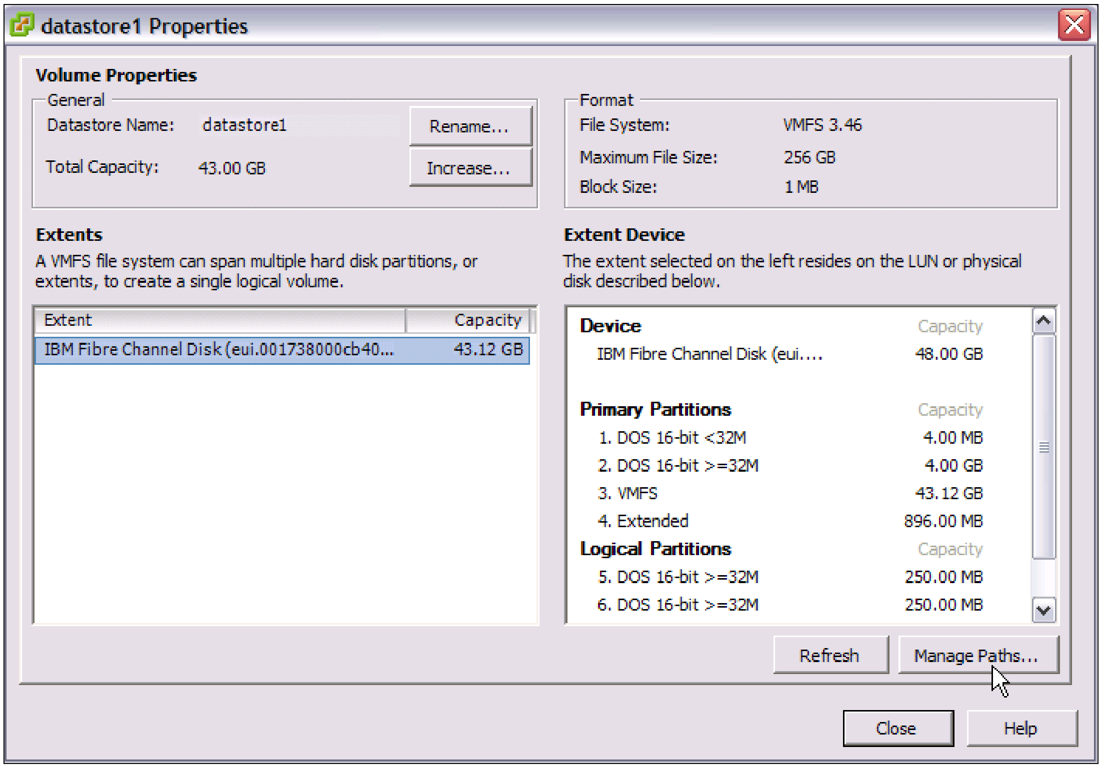


Figure 4-46 VMware VSphere client: Datastore properties

Figure 4-47 displays that two active FC paths are active and being used Round Robin.

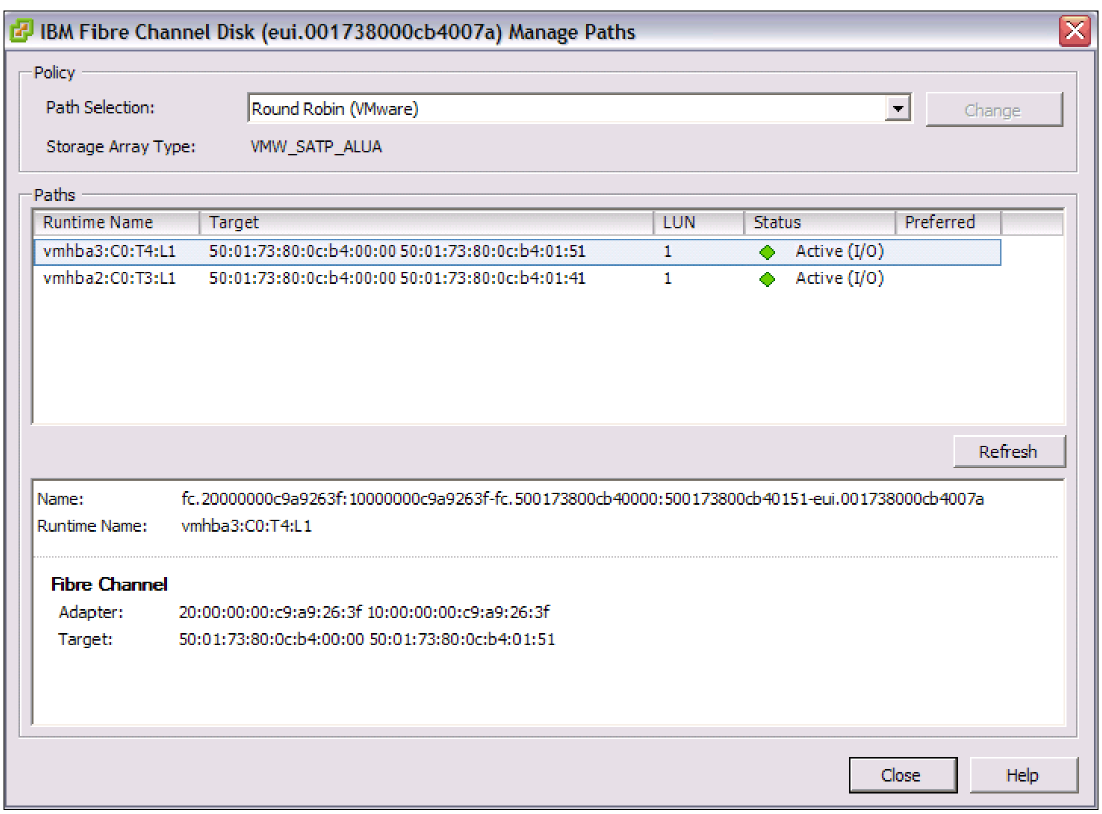


Figure 4-47 VMware VSphere client: Manage Paths

For additional XIV considerations, see 4.10.1, “VMware ESXi best practices” on page 264.

4.7 Installation of SUSE Linux Enterprise Server 11 for SAN Boot

This section provides instructions for doing a SAN Boot installation of the SUSE Linux Enterprise Server 11 operating system using Fibre Channel connectivity to XIV.

4.7.1 Prerequisites

To successfully SAN Boot a Windows host from XIV, a number of prerequisites must be met. Here is a generic list. However, your environment might have additional requirements:

- Review 4.2, “SAN configuration planning” on page 205
- Review 4.3, “XIV Management Tools and Support Tools” on page 221
- Review 4.4, “Logical configuration for host connectivity” on page 221

4.7.2 Operating system installation considerations

During the installation of SUSE Linux Enterprise Server 11, Figure 4-48 shows the panel that relates to the disk storage and boot configuration for the installation. By default, the SUSE Linux Enterprise Server 11 installer does not enable multipathing functionality. To configure the installation to be multipath aware, click the **Partitioning** link when the *Installation Settings* (Expert) panel is displayed.

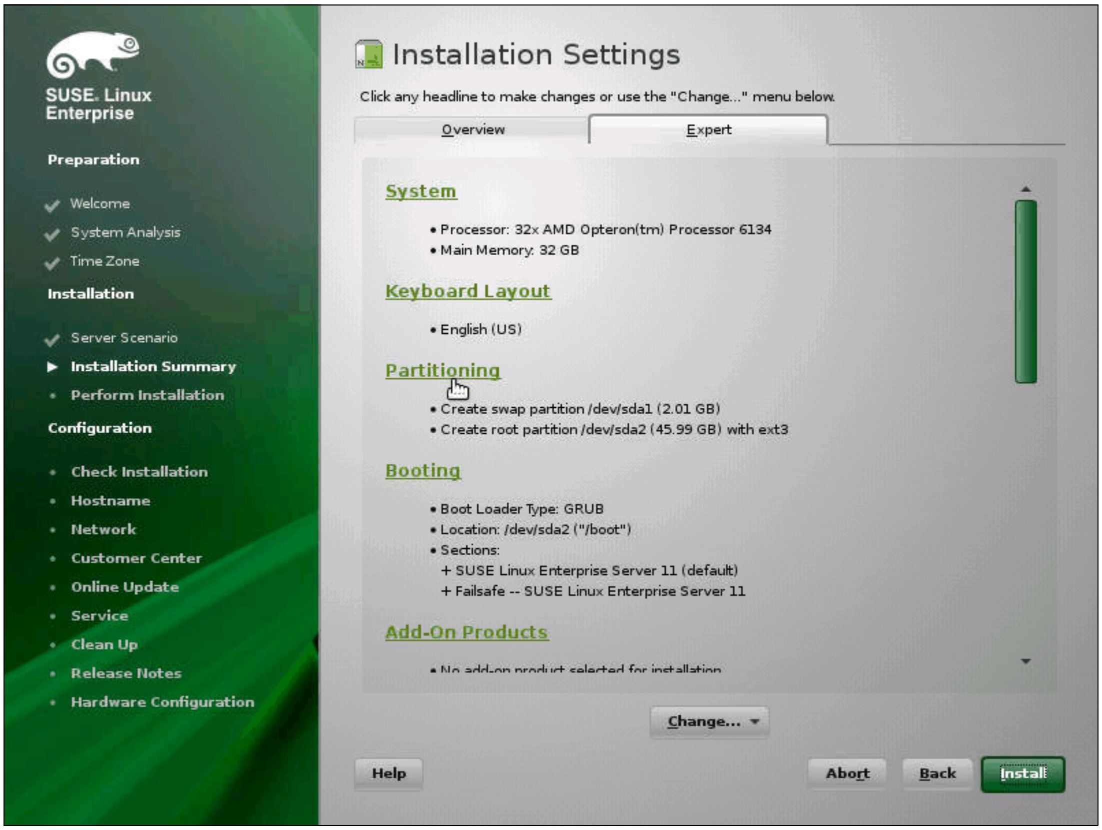


Figure 4-48 SUSE Linux Enterprise Server 11: Installation Summary / Installation Settings (expert)

Figure 4-49 illustrates the “Preparing Hard Disk: Step 1” panel that is displayed next. In this example, because there are two FC paths to the same XIV volume, that volume is displayed as two separate disk devices (sda and sdb). To enable multipathing, make sure that the **Custom Partitioning (for experts)** button is selected and click **Next**.

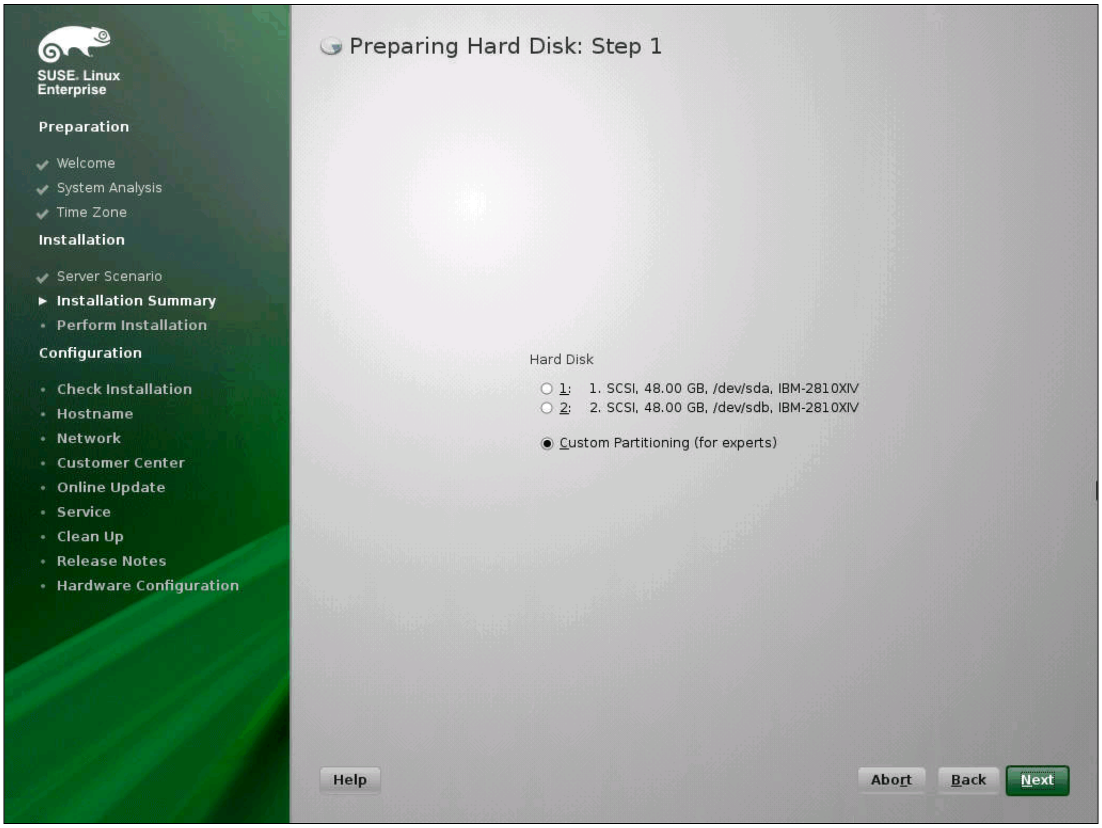


Figure 4-49 SUSE Linux Enterprise Server 11: Preparing Hard Disk: Step 1

As seen in Figure 4-50, the next panel you see is the “Expert Partitioner”. Here you enable multipathing. After selecting **Hard disks** in the navigation section on the left side, the tool offers the **Configure** button in the bottom right corner of the main panel. Click it and select **Configure Multipath**.

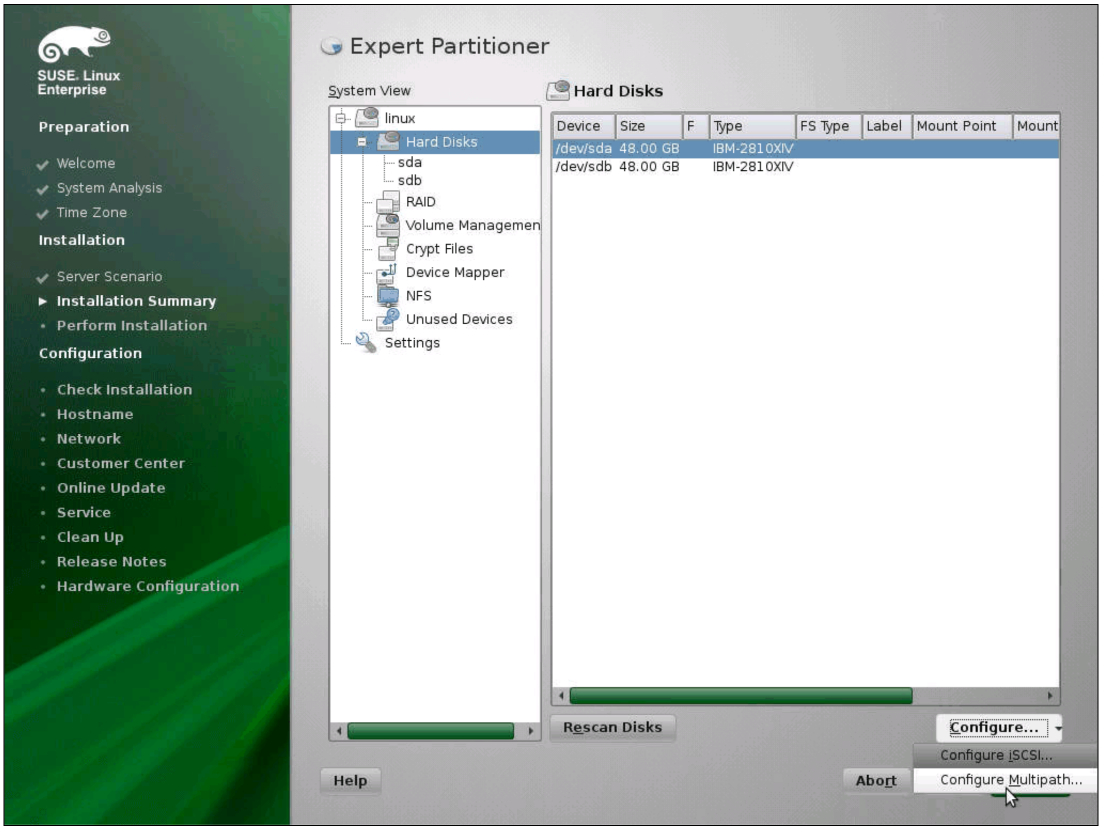


Figure 4-50 SUSE Linux Enterprise Server 11: Expert Partitioner panel before Configure Multipathing

The tool asks for confirmation and then rescans the disk devices. When finished, it presents an updated list of hard disks, which also shows the multipath devices it has created, as can be seen in Figure 4-51.

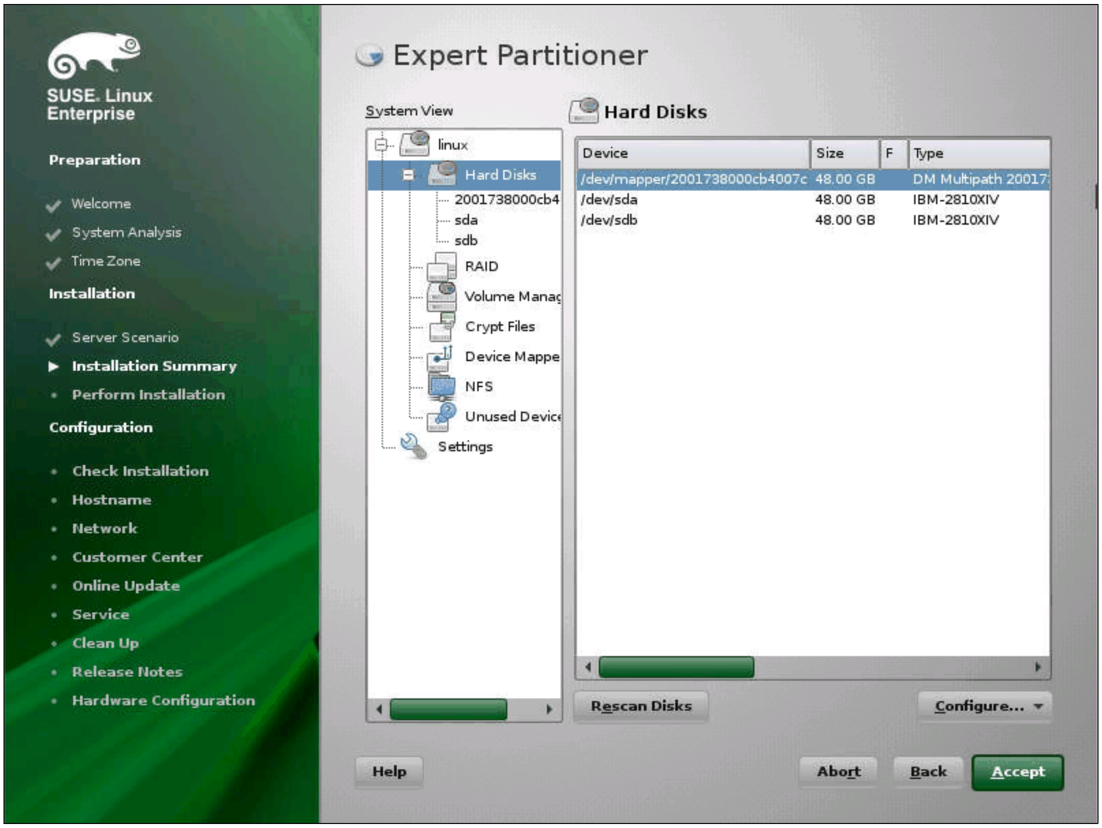


Figure 4-51 SUSE Linux Enterprise Server 11: Expert Partitioner panel after Configure Multipathing

With multipathing enabled and the multipath device displayed, the next step is to configure the desired partitions for the SUSE Linux Enterprise Server 11 installation.

Figure 4-52 provides an example of how the XIV volume can be partitioned for a SAN Boot configuration. The configuration steps involved:

1. On the left side of the panel under *System View*, select the multipath device under *Hard Disks* by left-clicking it.
2. At the bottom of the *Partitions* tab, click **Add**. In this example, a 2.01 GB primary partition was created for swap space.
3. At the bottom of the *Partitions* tab, click **Add**. In this example, a 45.98 GB primary partition was created for the root (/) file system.

After the desired partitions have been created, click **Accept**.

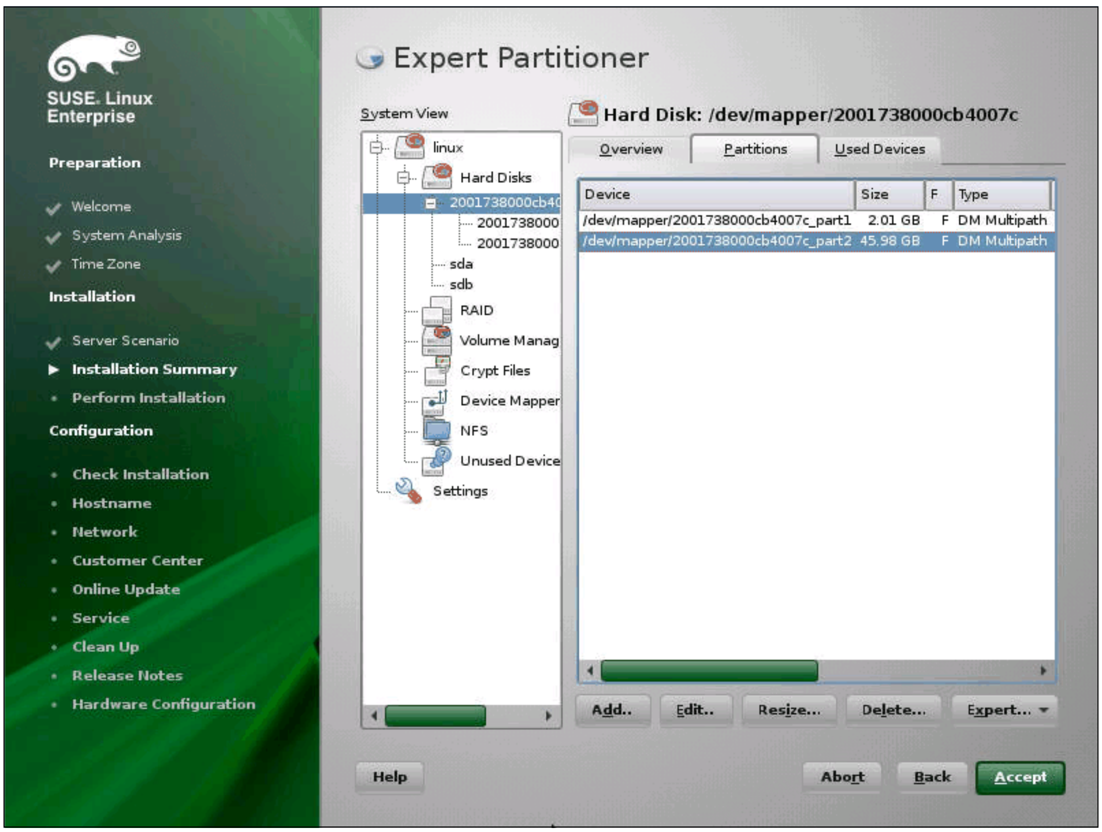


Figure 4-52 SUSE Linux Enterprise Server 11: Expert Partitioner with multipath-based partitions

After the Expert Partitioner panel has been accepted, the SUSE Yast2 utility detects that a configuration change has occurred that impacts the boot loader. As illustrated in Figure 4-53, a pop-up will be displayed asking if the default boot loader location should be set. Click **Yes**.

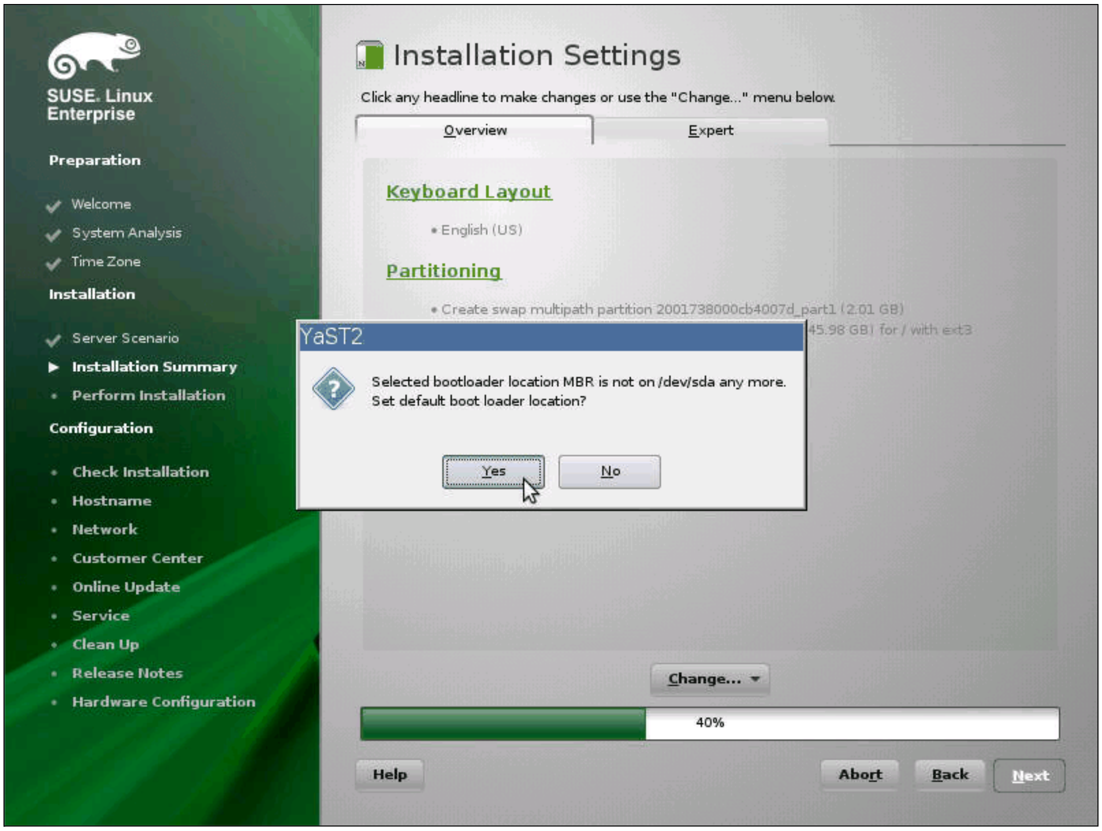


Figure 4-53 SUSE Linux Enterprise Server 11: Set default boot loader location pop-up

To ensure that the boot loader is configured to use a multipath device, the *Booting* section of Installation Settings must be reviewed. It is accomplished by selecting the **Booting** option as seen in Figure 4-54.

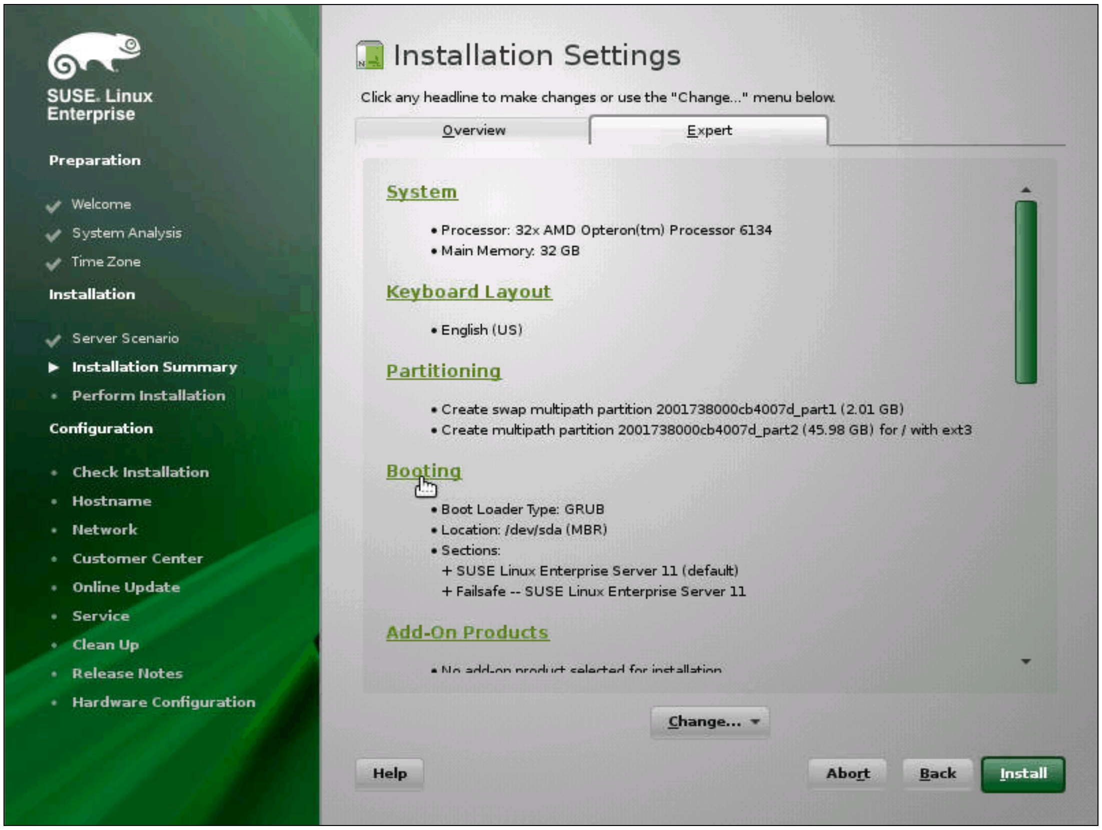


Figure 4-54 SUSE Linux Enterprise Server 11: Installation Summary / Installation Settings (expert)

Figure 4-55 displays the default Boot Loader Settings that are configured by default during this SUSE Linux Enterprise Server installation. To verify the disk device to be used by the boot loader, click **Boot Loader Installation Details**.

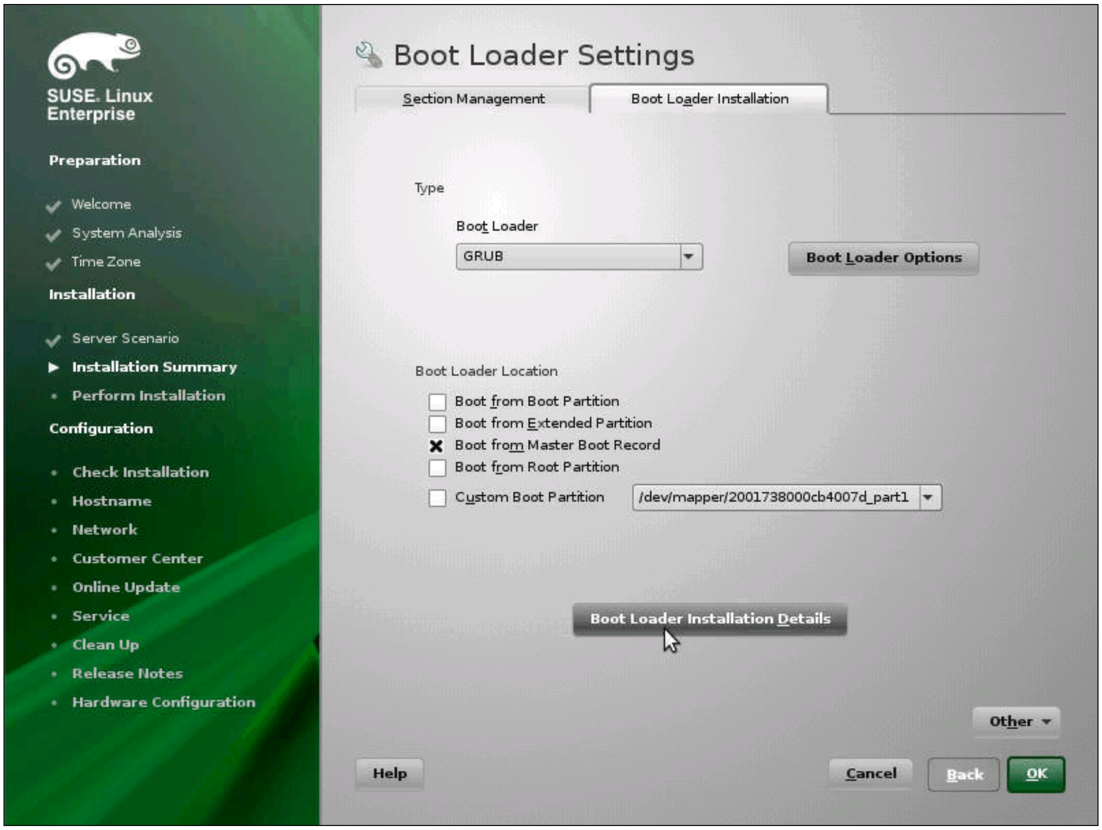


Figure 4-55 SUSE Linux Enterprise Server 11: Boot Loader Settings panel

Figure 4-56 illustrates the *Boot Loader Device Map* panel that is displayed. Notice that the disk referenced for the boot loader is a multipath device. No additional configuration is necessary in this example. If multiple multipath devices are displayed, ensure that the multipath device associated with the SAN Boot XIV volume is moved to the top of the list. Select **OK** to return to the high level Installation Settings panel.

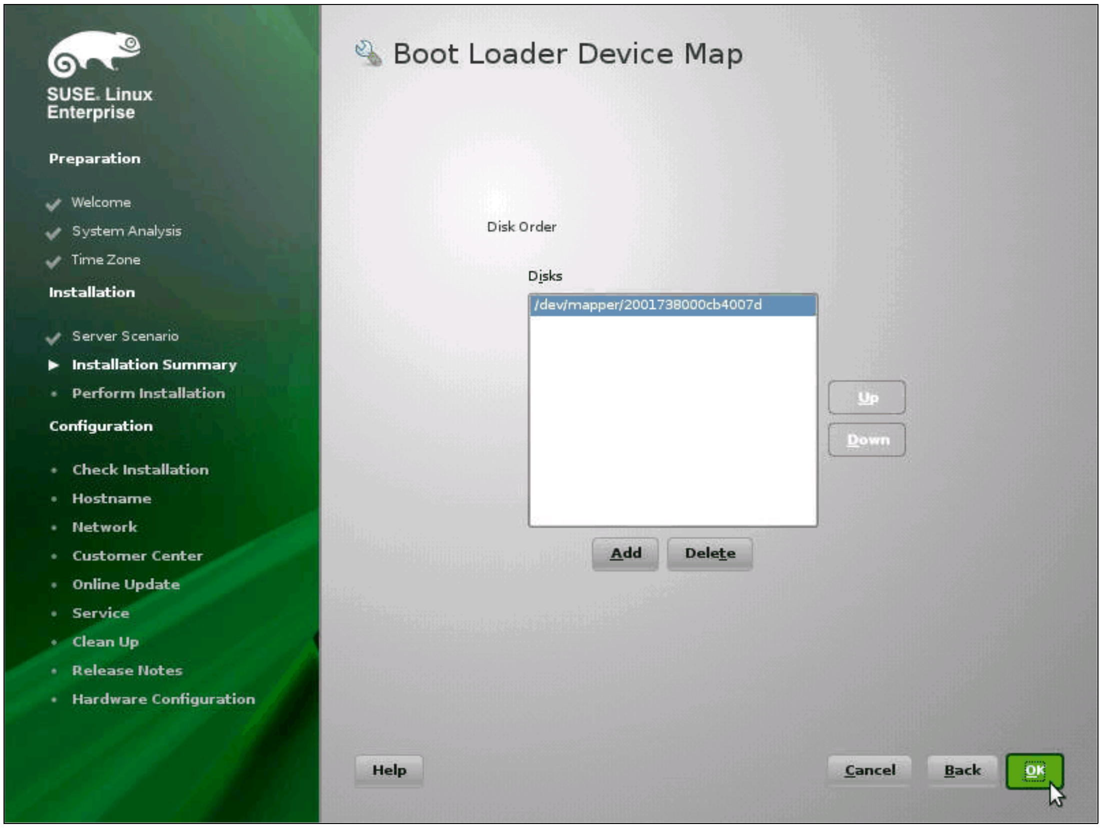


Figure 4-56 SUSE Linux Enterprise Server 11: Boot Loader Device Map panel

Figure 4-57 shows the SUSE Linux Enterprise Server 11 *Installation Summary* after the following actions have taken place:

1. Multipathing was enabled and multipath devices were created.
2. Partitions were created on the SAN Boot XIV volume using the multipath device.
3. The boot loader was updated to reference the SAN Boot multipath device.

At this point, you need to click the **Install** button to complete the SUSE Linux Enterprise Server 11 installation. No other changes to the installation configuration are required for SAN Boot.

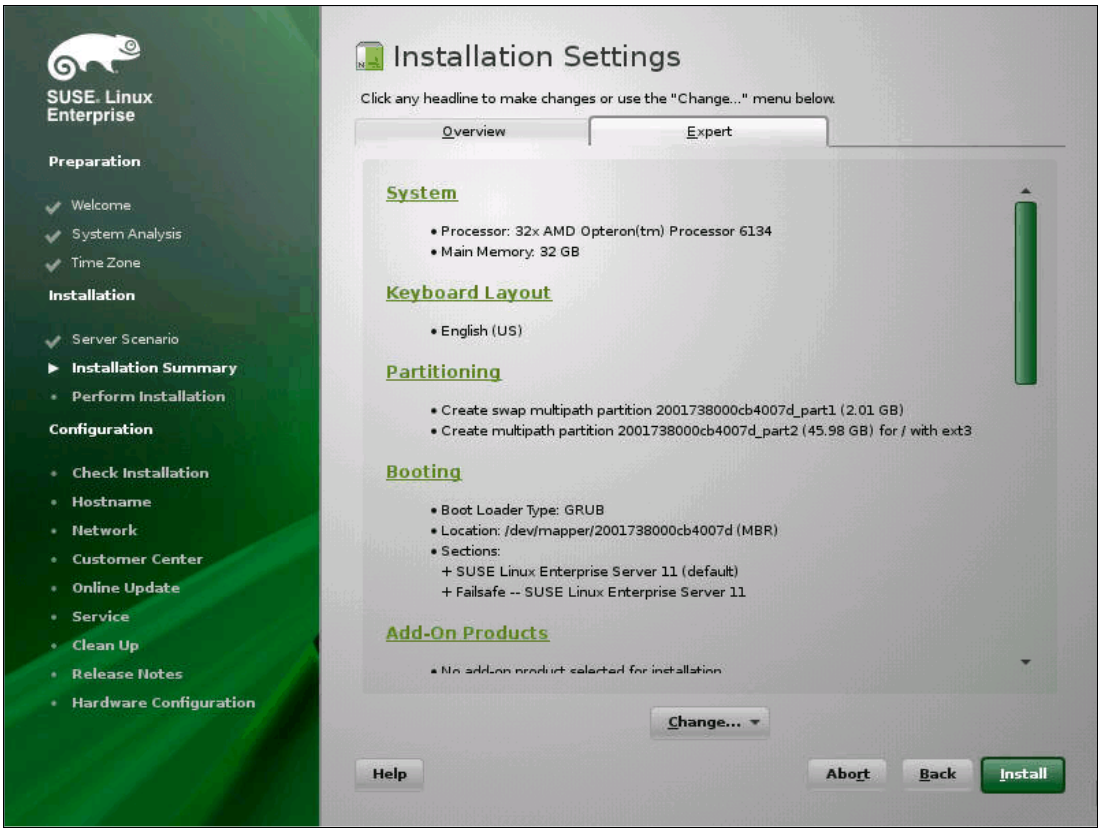


Figure 4-57 SUSE Linux Enterprise Server 11: Multipath-enabled Installation Settings panel

4.7.3 Post operating system installation considerations

After the installation of SUSE Linux Enterprise Server 11 has completed, the XIV Host Attachment Kit (HAK) must be installed to complete the multipathing configuration process. Because the HAK has OS package prerequisites associated with the HAK, a review of the associated HAK release notes, for the version to be installed, is highly recommended prior to the start of the installation. Example 4-10 shows the execution of the installation script.

Example 4-10 Installation of the XIV HAK for SUSE Linux Enterprise Server 11

```
# tar -xzvf XIV_host_attach-1.5.2-sles11-x64.tar.gz
# cd XIV_host_attach-1.5.2-SLES11-x64
# /bin/sh ./install.sh
Welcome to the XIV Host Attachment Kit installer.
```

```
Required OS packages:
+--------------------------+-----------------+-----------------+
| RHEL                     | SLES 11         | SLES 10         |
+--------------------------+-----------------+-----------------+
| device-mapper-multipath  | multipath-tools | multipath-tools |
| sg3_utils                | sg3_utils       | scsi            |
+--------------------------+-----------------+-----------------+
NOTE: Installation would FAIL if any of the above packages is missing.


Optional OS packages (iSCSI support):
+--------------------------+-----------------+-----------------+
| RHEL                     | SLES 11         | SLES 10         |
+--------------------------+-----------------+-----------------+
| iscsi-initiator-utils    | open-iscsi      | open-iscsi      |
+--------------------------+-----------------+-----------------+


Would you like to proceed and install the Host Attachment Kit? [Y/n]:
Y
Please wait while the installer validates your existing configuration...
---------------------------------------------------------------
Please wait, the Host Attachment Kit is being installed...
---------------------------------------------------------------
Installation successful.
Please refer to the Host Attachment Guide for information on how to configure this
host.
```

After the installation of the HAK, the **xiv_attach** command can be used to finish the XIV multipathing configuration. Example 4-11shows the process based upon Fibre Channel (FC) connectivity to the XIV and a rescan of the FC connectivity.

Example 4-11 Execution of the xiv_attach command

```
# xiv_attach
-------------------------------------------------------------------------------
Welcome to the XIV Host Attachment wizard, version 1.5.2.
This wizard will assist you to attach this host to the XIV system.

The wizard will now validate host configuration for the XIV system.
Press [ENTER] to proceed.

-------------------------------------------------------------------------------
iSCSI software was not detected. Refer to the guide for more info.
Only fibre-channel is supported on this host.
Would you like to set up an FC attachment? [default: yes ]:yes
-------------------------------------------------------------------------------
Please wait while the wizard validates your existing configuration...
The wizard needs to configure the host for the XIV system.
Do you want to proceed? [default: yes ]:yes
Please wait while the host is being configured...
The host is now being configured for the XIV system
-------------------------------------------------------------------------------
Please zone this host and add its WWPNs with the XIV storage system:
21:00:00:1b:32:9b:5d:48: [QLOGIC]: N/A
21:01:00:1b:32:bb:5d:48: [QLOGIC]: N/A
```

```
Press [ENTER] to proceed.

Would you like to rescan for new storage devices now? [default: yes ]:yes
Please wait while rescanning for storage devices...
-------------------------------------------------------------------------------
The host is connected to the following XIV storage arrays:
Serial    Ver    Host Defined  Ports Defined  Protocol  Host Name(s)
7803252   10.2   Yes           All            FC        RB_xseries_01_qlogic
This host is defined on all FC-attached XIV storage arrays

Press [ENTER] to proceed.

-------------------------------------------------------------------------------
The XIV host attachment wizard successfully configured this host

Press [ENTER] to exit.
```

With the completion of the **xiv_attach** command, the **xiv_devlist** command can be executed to review XIV volumes discovered, the multipath devices created for them, and the number of paths configured. As seen in Example 4-12, multipath device **mpatha** was created and has two paths to that XIV volume.

Example 4-12 Output of the xiv_devlist command

```
# xiv_devlist
XIV Devices
-----------------------------------------------------------------------------------------------------
Device             Size    Paths  Vol Name                   Vol Id   XIV Id   XIV Host
-----------------------------------------------------------------------------------------------------
/dev/mapper/mpatha  51.5GB  2/2    RB_ST1L05R01_Qlogic_SuSE11  125      7803252  RB_xseries_01_qlogic
-----------------------------------------------------------------------------------------------------


Non-XIV Devices
---------------------
Device   Size   Paths
---------------------
```

As seen in Example 4-13, to get additional information about the multipathing configuration, the **multipath -ll** command can be executed.

Example 4-13 Output of the multipath-ll command

```
# multipath -ll
mpatha (2001738000cb4007d) dm-0 IBM,2810XIV
[size=48G][features=1 queue_if_no_path][hwhandler=0][rw]
\_ round-robin 0 [prio=1][active]
 \_ 6:0:1:1 sda 8:0   [active][ready]
\_ round-robin 0 [prio=1][enabled]
 \_ 7:0:1:1 sdb 8:16  [active][ready]
```

Tip: Though the foregoing example shows that multiple paths exist to the same XIV volume, the multipathing configuration is not optimal for performance based on the path group configuration. See 4.10.2, "Linux best practices" on page 264 for additional information.

4.8 Installation of Red Hat Enterprise Linux 5.5

This section provides instructions for doing a SAN Boot installation of the Red Hat Enterprise Linux (RHEL) 5.5 server operating system using Fibre Channel connectivity to XIV.

4.8.1 Prerequisites

To successfully SAN Boot a Windows host from XIV, a number of prerequisites must be met. Here is a generic list. However, your environment might have additional requirements:

- Review 4.2, “SAN configuration planning” on page 205
- Review 4.3, “XIV Management Tools and Support Tools” on page 221
- Review 4.4, “Logical configuration for host connectivity” on page 221

4.8.2 Operating system installation considerations

To perform an installation of RHEL 5.5 that uses multipath devices, the installation process must be started with the **`mpatha`** keyword. Figure 4-58 shows the initial Red Hat installation panel in which **`linux mpath`** has been specified start the installation process with multipathing enabled.

Figure 4-58 RHEL 5.5: Installation panel with linux mpath specified

Figure 4-59 illustrates how the multipath drive device is referenced for install partitioning.

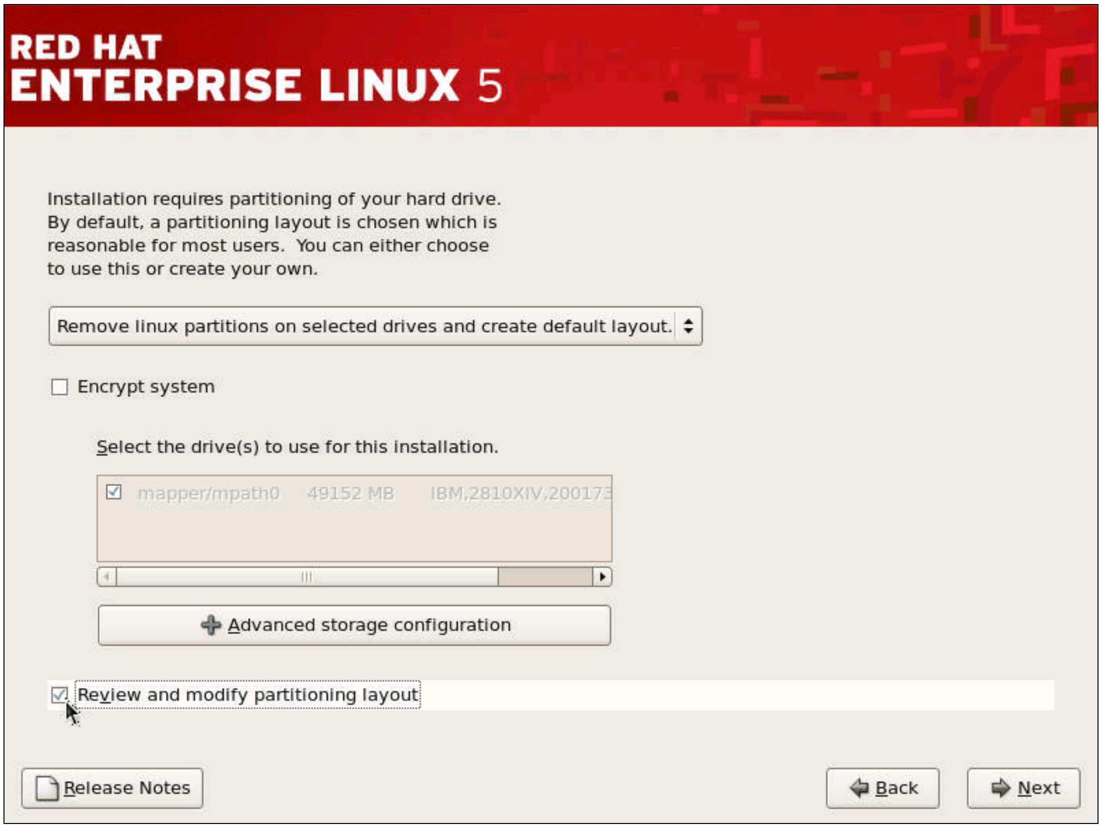


Figure 4-59 RHEL 5.5: Partition overview panel with review and modify layout specified

After the **Next** button is clicked, Figure 4-60 displays additional details regarding the default RHEL partitioning and file system layout and the use of the related multipath devices. Click **Next** to continue the installation process.

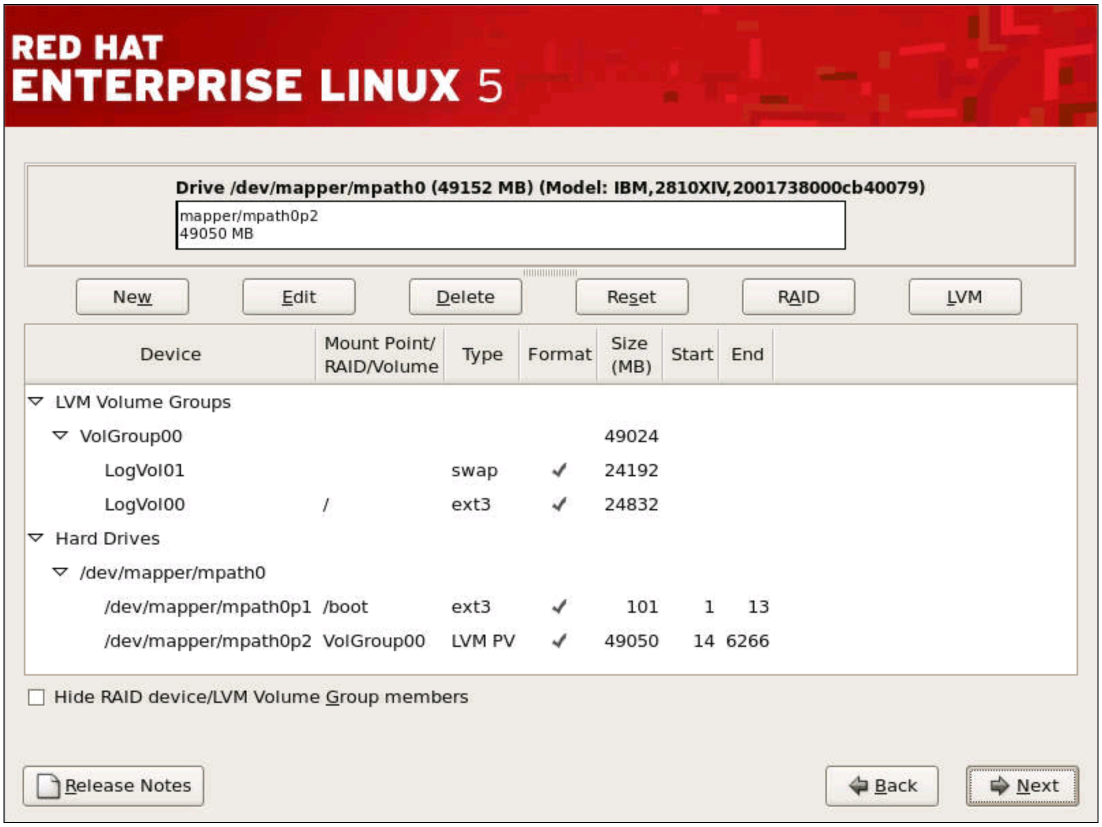


Figure 4-60 RHEL 5.5: Review of default install partitioning

Figure 4-61 illustrates the boot loader options during the installation. By default, RHEL 5.5 will install the GRUB boot loader. Notice that, like the partition and file system layout, the multipathing device will also be used for the boot loader installation. Click **Next** to continue the installation process.

Figure 4-61 RHEL 5.5: Boot loader panel

Lastly, during the RHEL 5.5 installation, you need to verify that the boot loader will also utilize a multipath device. Figure 4-62 shows the installation panel that is used to configure the boot loader installation. Notice that, in this example, the /dev/mapper/mpath0 device is used.

Figure 4-62 RHEL 5.5: Boot loader installation panel

4.8.3 Post operating system installation considerations

After the installation of RHEL 5.5 has completed, the XIV Host Attachment Kit (HAK) must be installed to complete the multipathing configuration process. Because the HAK has OS package prerequisites associated with the HAK, a review of the associated HAK release notes, for the version to be installed, is highly recommended prior to the start of the installation. Example 4-14 shows the execution of the installation script.

Example 4-14 Installation of the XIV HAK for RHEL 5.5

```
# tar -xzvf XIV_host_attach-1.5.2-rhel5-x64.tar.gz
# cd XIV_host_attach-1.5.2-rhel5-x64
# /bin/sh ./install.sh
Welcome to the XIV Host Attachment Kit installer.

Required OS packages:
+--------------------------+-----------------+-----------------+
| RHEL                     | SLES 11         | SLES 10         |
+--------------------------+-----------------+-----------------+
| device-mapper-multipath  | multipath-tools | multipath-tools |
| sg3_utils                | sg3_utils       | scsi            |
+--------------------------+-----------------+-----------------+
NOTE: Installation would FAIL if any of the above packages is missing.
```

```
Optional OS packages (iSCSI support):
+--------------------------+-----------------+-----------------+
| RHEL                     | SLES 11         | SLES 10         |
+--------------------------+-----------------+-----------------+
| iscsi-initiator-utils    | open-iscsi      | open-iscsi      |
+--------------------------+-----------------+-----------------+


Would you like to proceed and install the Host Attachment Kit? [Y/n]:
Y
Please wait while the installer validates your existing configuration...
---------------------------------------------------------------
Please wait, the Host Attachment Kit is being installed...
---------------------------------------------------------------
Installation successful.
Please refer to the Host Attachment Guide for information on how to configure this
host.
```

After the installation of the HAK, the **xiv_attach** command can be used to finish the XIV multipathing configuration. Example 4-15 shows the process based upon Fibre Channel (FC) connectivity to the XIV and a rescan of the FC connectivity.

Example 4-15 Execution of the xiv_attach command

```
# xiv_attach
-------------------------------------------------------------------------------
Welcome to the XIV Host Attachment wizard, version 1.5.2.
This wizard will assist you to attach this host to the XIV system.

The wizard will now validate host configuration for the XIV system.
Press [ENTER] to proceed.

-------------------------------------------------------------------------------
iSCSI software was not detected. Refer to the guide for more info.
Only fibre-channel is supported on this host.
Would you like to set up an FC attachment? [default: yes ]: yes
-------------------------------------------------------------------------------
Please wait while the wizard validates your existing configuration...
The wizard needs to configure the host for the XIV system.
Do you want to proceed? [default: yes ]:
Please wait while the host is being configured...
The host is now being configured for the XIV system
-------------------------------------------------------------------------------
Please zone this host and add its WWPNs with the XIV storage system:
21:00:00:1b:32:9b:5d:48: [QLOGIC]: N/A
21:01:00:1b:32:bb:5d:48: [QLOGIC]: N/A
Press [ENTER] to proceed.

Would you like to rescan for new storage devices now? [default: yes ]: yes
Please wait while rescanning for storage devices...
-------------------------------------------------------------------------------
The host is connected to the following XIV storage arrays:
Serial    Ver    Host Defined  Ports Defined  Protocol  Host Name(s)
7803252   10.2   Yes           All            FC        RB_xseries_01_qlogic
```

```
This host is defined on all FC-attached XIV storage arrays

Press [ENTER] to proceed.

-------------------------------------------------------------------------------
The XIV host attachment wizard successfully configured this host

Press [ENTER] to exit.
```

With the completion of the **xiv_attach** command, the **xiv_devlist** command can be executed to review XIV volumes discovered, the multipath devices created for them, and the number of paths configured. As seen in Example 4-16, multipath device **mpatha** been created and has two paths to that XIV volume.

Example 4-16 Output of the xiv_devlist command

```
# xiv_devlist
XIV Devices
----------------------------------------------------------------------------------------------------
Device              Size    Paths  Vol Name                  Vol Id   XIV Id   XIV Host
----------------------------------------------------------------------------------------------------
/dev/mapper/mpath0  51.5GB  2/2    RB_ST1L05R01_Qlogic_RH55  127      7803252  RB_xseries_01_qlogic
----------------------------------------------------------------------------------------------------


Non-XIV Devices
---------------------
Device   Size   Paths
---------------------
```

As seen in Example 4-17, to get additional information about the multipathing configuration, the **multipath -ll** command can be executed.

Example 4-17 Output of the multipath -ll command

```
# multipath -ll
mpath0 (2001738000cb4007f) dm-0 IBM,2810XIV
[size=48G][features=1 queue_if_no_path][hwhandler=0][rw]
\_ round-robin 0 [prio=2][active]
 \_ 0:0:1:1 sda 8:0   [active][ready]
 \_ 1:0:1:1 sdb 8:16  [active][ready]
```

In the above example, the output of the multipath command shows that a single mpath device *(mpath0)* has been created and will distribute I/O requests across two paths (devices *sda* and *sdb*).

For additional information about the information displayed by the **multipath** command and how it relates to I/O optimization, see 4.10.2, “Linux best practices” on page 264.

4.9 Installation of AIX 6.1 for SAN Boot

This section provides instructions for doing a SAN Boot installation of the AIX 6.1 operating system using Fibre Channel connectivity to XIV.

4.9.1 Prerequisites

To successfully SAN Boot a Windows host from XIV, a number of prerequisites must be met. Here is a generic list. However, your environment might have additional requirements:

- Review 4.2, “SAN configuration planning” on page 205
- Review 4.3, “XIV Management Tools and Support Tools” on page 221
- Review 4.4, “Logical configuration for host connectivity” on page 221

4.9.2 Activating and installing the partition (HMC interface)

1. Activate the partition, as follows:
   a. Insert the AIX Volume 1 installation media into the media device of the managed system.
   b. Check if the DVD is already assigned to the current partition.
   c. Right-click the partition to open the menu.
   d. Select **Activate**. The Activate Partition menu opens with a selection of partition profiles. Be sure the correct profile is highlighted.
   e. Select **Open** a terminal window or console session at the bottom of the menu to open a virtual terminal (vterm) window.
   f. Select **Advanced** to open the Advanced options menu.
   g. For the Boot mode, select **SMS**.
   h. Select **OK** to close the Advanced options menu.
   i. Select **OK**. A vterm window opens for the partition.
2. In the SMS menu on the vterm, do the following steps:
   a. Press the 5 key and press Enter to select **5. Select Boot Options**. See Figure 4-63.

```
 Version EL340_075
 SMS 1.7 (c) Copyright IBM Corp. 2000,2008 All rights reserved.
-------------------------------------------------------------------------------
 Main Menu
 1.   Select Language
 2.   Setup Remote IPL (Initial Program Load)
 3.   Change SCSI Settings
 4.   Select Console
 5.   Select Boot Options

-------------------------------------------------------------------------------
 Navigation Keys:
X = eXit System Management Services

-------------------------------------------------------------------------------
 Type menu item number and press Enter or select Navigation key:5
```

Figure 4-63 SMS Basic Menu

b. Press the 1 key and press Enter to select **1. Select Install/Boot Device**.
c. Press the 3 key and press Enter to select **3. CD/DVD**.
d. Select the media type that corresponds to the media device and press Enter.
e. Select the device number that corresponds to the media device and press Enter. The media device is now the first device in the Current Boot Sequence list.
f. Press the 2 key and press Enter to select **2. Normal Mode Boot**.
g. Press the 1 key and press Enter to **confirm your boot option**.

Afterwards, the system will boot your machine using the CD/DVD and load the basic Kernel to install the machine.

3. Boot from the AIX Volume 1, as follows:
   a. Select console and press Enter.
   b. Select language for BOS Installation menus, and press Enter to open the **Welcome to Base Operating System Installation and Maintenance** menu.
   c. Type 2 to select **Change/Show Installation Settings and Install**, as in Figure 4-64, in the Choice field and press Enter.

```
                         Welcome to Base Operating System
                          Installation and Maintenance

Type the number of your choice and press Enter.  Choice is indicated by >>>.

>>> 1 Start Install Now with Default Settings

    2 Change/Show Installation Settings and Install

    3 Start Maintenance Mode for System Recovery

    4 Configure Network Disks (iSCSI)


    88  Help ?
    99  Previous Menu

>>> Choice [1]: 2
```

Figure 4-64 Base Installation menu

4. Verify or Change BOS Installation Settings, as follows:
   a. Type 2 in the Choice field to select the **Change/Show Installation Settings** option.
   b. Type 1 for System Setting to select **New and Complete Overwrite** in the Choice field and press Enter.
   c. When the Change Disk(s) panel opens, you can change the destination disk for the installation. If the default shown is correct, type 0 in the Choice field and press Enter. To change the destination disk, do the following steps:
      i. Type the number for each disk you choose in the Choice field and press Enter. Do not press Enter a final time until you have finished selecting all disks. If you must deselect a disk, type its number a second time and press Enter.

ii. By typing 77 two times, you can check if the Storage WWPN and LUN number associated with your machine is correct, see Example 4-18.

Example 4-18 Checking Storage WWPN and LUN ID

```
Change Disk(s) Where You Want to Install

Type one or more numbers for the disk(s) to be used for installation and
press Enter. To cancel a choice, type the corresponding number and Press
Enter. At least one bootable disk must be selected. The current choice is
indicated by >>>.
Name      Device Adapter Connection Location or Physical Location Code
     1  hdisk1    U789C.001.DQDU764-P2-D5
>>>  2  hdisk16 ...-W50050768014052D1-LA000000000000
     3  hdisk6    ...-W5001738000D00142-L1000000000000
     4  hdisk0    U789C.001.DQDU764-P2-D8
     5  hdisk2    U789C.001.DQDU764-P2-D4
    06  MORE CHOICES...

>>>  0   Continue with choices indicated above
    55  More Disk Options
    66  Devices not known to Base Operating System Installation
    77  Display More Disk Information
    88  Help ?
    99  Previous Menu

>>> Choice [0]:
```

iii. When you have selected the disks, type 0 in the Choice field and press Enter. The **Installation and Settings** panel opens with the selected disks listed under System Settings.

Important: Be sure that you have made the correct selection for root volume group, because the existing data in the destination root volume group will be destroyed during Base Operating System (BOS) installation.

d. If needed, change the primary language environment. Use the following steps to change the primary language used by this installation to select the language and cultural convention you want to use.

Tip: Changes to the primary language environment do not take effect until after the Base Operating System Installation has completed and your system is rebooted.

i. Type 2 in the Choice field on the **Installation and Settings** panel to select the Primary Language Environment Settings option.

ii. Select the appropriate set of cultural convention, language, and keyboard options. Most of the options are a predefined combination, however, you can define your own combination of options.

e. After you have made all of your selections, verify that the selections are correct. Press Enter to confirm your selections and to begin the BOS Installation. The system automatically reboots after installation is complete.

5. Select **Open terminal window** to open a virtual terminal (vterm) window.
   a. Type the model of your terminal as the terminal type.
   b. In the License Agreement menu, select **Accept License Agreements**.
   c. Select **Yes** to ACCEPT Installed License Agreements.
   d. Press F10 (or Esc+0) to exit the License Agreement menu.
   e. In the Installation Assistant main menu, select **Set Date and Time**. See Figure 4-65.

```
                           Installation Assistant

Move cursor to desired item and press Enter.

  Set Date and Time
  Set root Password
  Configure Network Communications
  Install Software Applications
  Using SMIT (information only)


F1=Help             F2=Refresh          F3=Cancel           F8=Image
F9=Shell            F10=Exit            Enter=Do
```

Figure 4-65 Post installation menu - Installation Assistant

   f. Set the correct date, time, and time zone. Press the F3 (or Esc+3) key to return to the Installation Assistant main menu.
   g. Select **Set root Password**. Set a root password for the partition.
   h. Select **Configure Network Communications**. Select **TCP/IP Startup**. Select from the Available Network Interfaces and press Enter. Enter the appropriate network information in the Minimum Configuration and Startup menu and press Enter. Use the F3 (or Esc+3) key to return to the Installation Assistant main menu.
   i. Exit the Installation Assistant by typing F10 (or Esc+0).

   The vterm window will display a login prompt at this time.

6. When needed, switch the partition to Normal Mode, as follows:
   a. Right-click the partition profile to open the menu. Be sure the correct partition profile is highlighted.
   b. Select **Properties**.
   c. Select the **Settings** tab.
   d. For the Boot Mode, select **Normal**.
   e. Select **OK** to close the Properties menu.
   f. Right-click the partition to open the menu.
   g. Select **Restart Partition**.
   h. Select **Immediate** for the Restart Options.
   i. Confirm that you want to restart the partition.

When the partition has restarted, right-click the partition to open the menu.

4.9.3 Post operating system installation considerations

After the installation of AIX 6.1 has completed, the XIV Host Attachment Kit (HAK) must be installed to complete the multipathing configuration process. Because the HAK has OS package prerequisites associated with the HAK, a review of the associated HAK release notes, for the version to be installed, is highly recommended prior to the start of the installation. Example 4-19 shows the execution of the installation script.

Example 4-19 Installation of the XIV HAK for AIX 6.1

```
# /bin/sh ./install.sh
Welcome to the XIV Host Attachment Kit installer.

NOTE: This installation defaults to round robin multipathing,
if you would like to work in fail-over mode, please set the environment
variables before running this installation.


Would you like to proceed and install the Host Attachment Kit? [Y/n]:
Y
Please wait while the installer validates your existing configuration...
---------------------------------------------------------------
Please wait, the Host Attachment Kit is being installed...
---------------------------------------------------------------
Installation successful.
Please refer to the Host Attachment Guide for information on how to configure this
host.
```

After the installation of the HAK, the **xiv_attach** command can be used to finish the XIV multipathing configuration. Example 4-20 shows the process based upon Fibre Channel (FC) connectivity to the XIV and a rescan of the FC connectivity.

Example 4-20 Execution of the xiv_attach command

```
# xiv_attach
-------------------------------------------------------------------------------
Welcome to the XIV Host Attachment wizard, version 1.5.2.
This wizard will assist you to attach this host to the XIV system.

The wizard will now validate host configuration for the XIV system.
Press [ENTER] to proceed.

-------------------------------------------------------------------------------
Please choose a connectivity type, [f]c / [i]scsi : f
-------------------------------------------------------------------------------
Please wait while the wizard validates your existing configuration...
This host is already configured for the XIV system
-------------------------------------------------------------------------------
Please zone this host and add its WWPNs with the XIV storage system:
10:00:00:00:C9:8D:90:7E: fcs0: [IBM]: N/A
10:00:00:00:C9:8D:90:7F: fcs1: [IBM]: N/A
10:00:00:00:C9:8D:98:3A: fcs2: [IBM]: N/A
10:00:00:00:C9:8D:98:3B: fcs3: [IBM]: N/A
Press [ENTER] to proceed.

Would you like to rescan for new storage devices now? [default: yes ]:
```

```
Please wait while rescanning for storage devices...
Warnings and errors given by the cfgmgr command:
Method error (/usr/lib/methods/cfgefscsi -l fscsi1 ):
        0514-061 Cannot find a child device.
Warnings and errors given by the cfgmgr command:
Method error (/usr/lib/methods/cfgefscsi -l fscsi3 ):
        0514-061 Cannot find a child device.
-------------------------------------------------------------------------------
No XIV LUN0 devices were detected
Press [ENTER] to proceed.

-------------------------------------------------------------------------------
The XIV host attachment wizard successfully configured this host

Press [ENTER] to exit.
```

With the completion of the **xiv_attach** command, the **xiv_devlist** command can be executed to review XIV volumes discovered, the multipath devices created for them, and the number of paths configured. As seen in Example 4-21, the disk device /dev/hdisk16 has two paths to that XIV volume.

Example 4-21 Execution of the xiv_devlist command

```
# xiv_devlist
XIV Devices
-----------------------------------------------------------------------
Device        Size    Paths  Vol Name      Vol Id   XIV Id   XIV Host
-----------------------------------------------------------------------
/dev/hdisk16  34.4GB  2/2    RB_XIV_AIX61  6573     1300208  RB_AIX_61
-----------------------------------------------------------------------
```

Example 4-22 illustrates how additional path information can be viewed using the **lspath** command.

Example 4-22 Execution of the lspath command

```
# lspath -H -l hdisk16
status  name    parent

Enabled hdisk16 fscsi2
Enabled hdisk16 fscsi0
```

For additional XIV considerations, see 4.10.4, “AIX 6.1 best practices” on page 268.

4.10 Best practices

Due to the distributed data features of the XIV Storage System, high performance is achieved by parallelism. With the grid architecture and massive parallelism inherent to XIV system, the recommended approach is to maximize the utilization of all the XIV resources at all times.

In the case of host connectivity to the XIV, this maximizing of XIV resources translates into distributing host I/O across multiple XIV interface modules.

The following sections outline some best practices that should be reviewed post-installation. These best practices, in general, focus on multipathing configuration considerations that can impact performance.

4.10.1 VMware ESXi best practices

The following section outlines VMware ESXI best practices that relate to how multipathing settings can be reviewed.

Optimizing multipathing policies

For VMware ESXi, the *Round Robin* multipathing policy can be used to effectively distribute I/O across XIV interface modules and their FC ports. Current ESXi multipathing policies can be viewed from a vSphere client as previously shown in Figure 4-47 on page 237.

More information about VMware ESXi multipathing

For more information about VMware ESXi multipathing policies, see the following link:

http://kb.vmware.com/selfservice/microsites/search.do?language=en_US&cmd=displayKC&externalId=1011340

4.10.2 Linux best practices

The following section outlines Linux best practices (Red Hat RHEL 5.5 and SUSE Linux Enterprise Server 11) that relate to how multipathing settings can be reviewed.

Updating the Linux Boot Image

When the Linux operating system is booted, Linux uses an initial ramdisk (initrd) image to preload modules that allow the root file system to be accessed by the Linux kernel.

When Linux is SAN Booted and multipathing is used for the XIV connectivity, it is important to ensure that the initrd used by the host during its boot process has incorporated in it the current multipath configuration. In the case of XIV, it involves the following steps:

1. Installing XIV Host Attachment Kit
2. Executing the XIV **`xiv_attach`** script to configure the multipath environment
3. Executing the Linux mkinitrd utility to build a new initrd
4. Verifying that the boot loader is configured to use the new initrd
5. Rebooting the host

Linux Boot installation examples

The following examples outline the previous steps based upon what can be observed during a SUSE Linux Enterprise Server 11 SAN Boot installation.

Example 4-23 illustrates the multipath configuration after a SUSE Linux Enterprise Server 11 SAN Boot installation.

Example 4-23 Output of multipath -l after SAN Boot installation of SUSE Linux Enterprise Server 11

```
# multipath -l
2001738000cb4007c dm-0 IBM,2810XIV
[size=48G][features=0][hwhandler=0][rw]
\_ round-robin 0 [prio=-1][active]
 \_ 0:0:4:1 sda 8:0   [active][undef]
\_ round-robin 0 [prio=-1][enabled]
 \_ 1:0:4:1 sdb 8:16  [active][undef]
```

The previous example shows that two paths (*sda* and *sdb*) exist to the XIV volume, but the example also shows that two *path groups* have also been created (shown in **bold**). Though this configuration provides path failure protection, the creation of two path groups does not allow I/O requests to be distributed across both paths. In this case, all I/O requests will handled solely by the sda path. Though the sdb path will be used if the sda path fails, this configuration does not provide the performance benefit of distributing I/O across multiple XIV FC ports.

To update the multipath configuration used by this host at boot time, the initrd must be updated. However, prior to updating the initrd, the initrd currently being used by the boot loader must be identified. Example 4-24 shows the contents of the menu.lst file that is being used by the GRUB boot loader. The initrd used by default boot entry is highlighted in **bold**.

Example 4-24 Contents of GRUB's menu.lst file

```
# cat /boot/grub/menu.lst
# Modified by YaST2. Last modification on Tue Apr 26 16:00:20 UTC 2011
default 0
timeout 8
gfxmenu (hd0,1)/boot/message
##YaST - activate

###Don't change this comment - YaST2 identifier: Original name: linux###
title SUSE Linux Enterprise Server 11 - 2.6.27.19-5
    root (hd0,1)
    kernel /boot/vmlinuz-2.6.27.19-5-default
root=/dev/disk/by-id/scsi-2001738000cb4007c-part2
resume=/dev/disk/by-id/scsi-2001738000cb4007c-part1 splash=silent
crashkernel=256M-:128M@16M showopts vga=0x314
    initrd /boot/initrd-2.6.27.19-5-default

###Don't change this comment - YaST2 identifier: Original name: failsafe###
title Failsafe -- SUSE Linux Enterprise Server 11 - 2.6.27.19-5
    root (hd0,1)
    kernel /boot/vmlinuz-2.6.27.19-5-default
root=/dev/disk/by-id/scsi-2001738000cb4007c-part2 showopts ide=nodma apm=off
noresume edd=off powersaved=off nohz=off highres=off processor.max_cstate=1
x11failsafe vga=0x314
    initrd /boot/initrd-2.6.27.19-5-default
```

To create an updated initrd image, the Linux **mkinitrd** command is executed. This process will incorporate the multipathing modules and multipathing configuration information (based on the definitions within **/etc/multipath.conf**) into the initrd.

Example 4-25 shows the output when **mkinitrd** is run in SUSE Linux Enterprise Server 11 environment. Note that the initrd image created use the same name and directory structure as the initrd being referenced by GRUB's menu.lst file.

Example 4-25 Execution of SUSE Linux Enterprise Server 11 mkinitrd command

```
# mkinitrd

Kernel image:   /boot/vmlinuz-2.6.27.19-5-default
Initrd image:   /boot/initrd-2.6.27.19-5-default
Root device:    /dev/disk/by-id/scsi-mpatha-part2 (/dev/dm-2) (mounted on / as
ext3)
Resume device:  /dev/disk/by-id/scsi-2001738000cb4007c-part1 (/dev/sdb1)
Kernel Modules: scsi_mod scsi_tgt scsi_transport_fc lpfc hwmon thermal_sys
processor thermal dock libata ahci fan jbd mbcache ext3 edd scsi_dh dm-mod
dm-multipath dm-snapshot crc-t10dif sd_mod usbcore ohci-hcd ehci-hcd uhci-hcd
ff-memless hid usbhid dm-round-robin scsi_dh_emc scsi_dh_hp_sw scsi_dh_rdac
scsi_dh_alua
Features:       dm block usb multipath kpartx resume.userspace resume.kernel
Bootsplash:     SLES (800x600)
29341 blocks
```

When the **mkinitrd** process has completed, reboot the host to utilize the new initrd.

Example 4-26 shows the output of the multipath command after the reboot. Note that the XIV volume still has two paths associated with it, but we now only have a single policy group (shown in **bold**).

Example 4-26 Output of multipath -l after xiv_attach, mkinitrd, and reboot

```
# multipath -l
mpatha (2001738000cb4007c) dm-0 IBM,2810XIV
[size=48G][features=1 queue_if_no_path][hwhandler=0][rw]
\_ round-robin 0 [prio=-2][active]
 \_ 0:0:4:1 sda 8:0    [active][undef]
 \_ 1:0:4:1 sdb 8:16   [active][undef]
```

This example reflects that the path is now configured based on the /etc/multipath.conf file that was created from the installation of the XIV HAK and the initial XIV **xiv_attach** command execution. In this particular example, the change to the policy group observed is based upon the use of the **path_grouping_policy multibus** option configured in /etc/multipth.conf.

For additional information regarding the output of the **multipath -l** command and how it relates to the /etc/multipath.conf configuration file, see the following link.

http://sourceware.org/lvm2/wiki/MultipathUsageGuide

4.10.3 Windows 2008 R2 best practices

The following section outlines Windows 2008 R2 best practices related to multipathing settings.

Optimizing multipathing policies

For Windows 2008 R2, the use of the *Round Robin* multipathing policy can be used to effectively distribute I/O. To verify current multipath policy settings, the Windows *Disk Management* utility can be used.

To start Disk Management, the command **diskmgmt.msc** can be executed from a command line. To display multipathing information for a given disk device, right-click the device and select **Properties**. It is outlined in Figure 4-66.

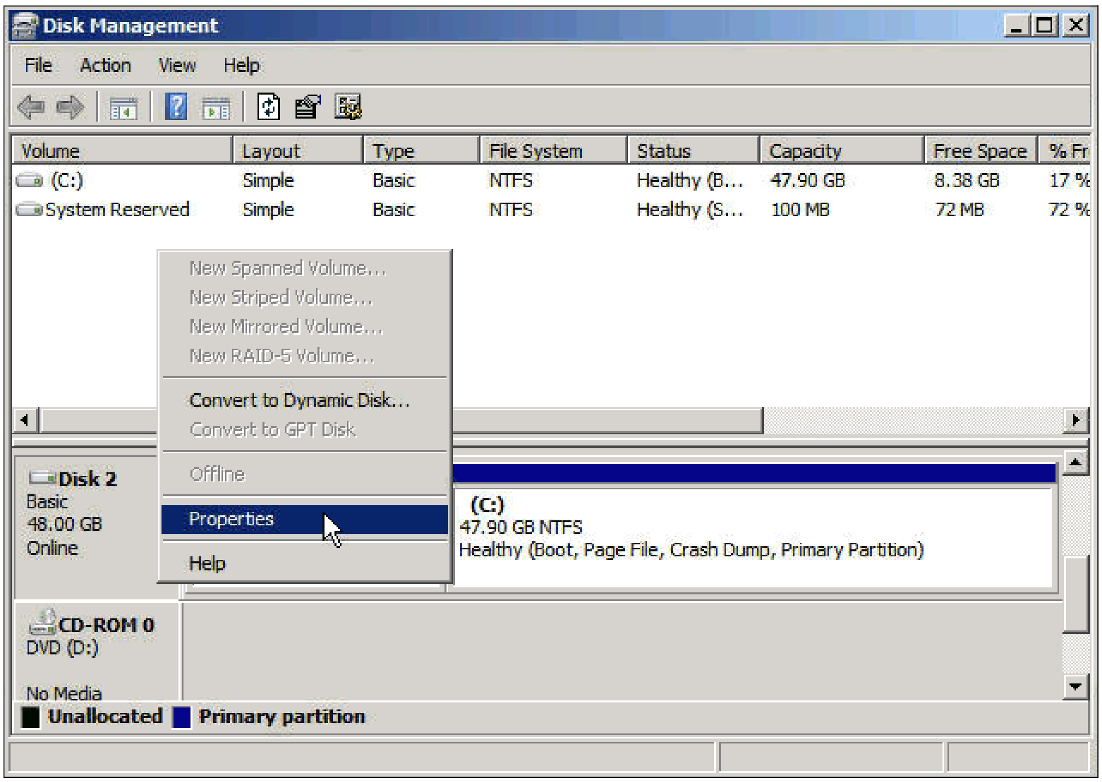


Figure 4-66 Displaying disk properties in Windows 2008 R2 Disk Management

From within the properties panel, the *MPIO* tab displays information about the current multipathing configuration. In the example shown in Figure 4-67, we can see that XIV I/O will be distributed across two active paths by the use of the *Round Robin* policy.

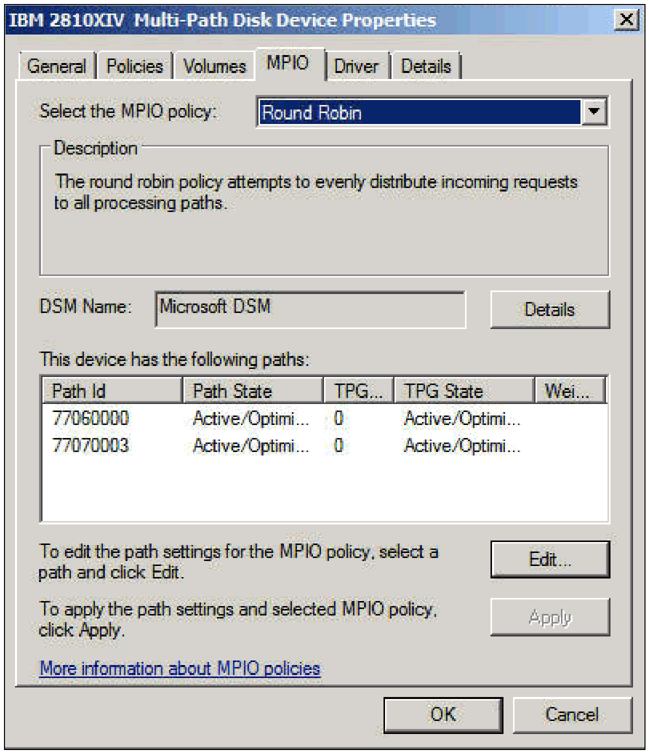


Figure 4-67 MPIO Policy - Round Robin policy

More information about Windows multipathing

For more information about Windows multipathing policies, see the following link:

http://download.microsoft.com/download/3/0/4/304083f1-11e7-44d9-92b9-2f3cdbf01048/mpio.doc

4.10.4 AIX 6.1 best practices

The following section outlines AIX 6.1best practices related to multipathing settings.

Optimizing multipathing policies

For AIX 6.1, the use of the *round_robin* multipathing policy can be used to effectively distribute I/O. To verify current multipath policy settings, the lsattr utility can be used against the multipath device.

Example 4-27 shows the use of the **xiv_devlist** command can to identify the multipathing devices.

Example 4-27 Use of the xiv_devlist command

```
# xiv_devlist
XIV Devices
----------------------------------------------------------------------
Device        Size    Paths  Vol Name      Vol Id   XIV Id   XIV Host
----------------------------------------------------------------------
/dev/hdisk16  34.4GB  2/2    RB_XIV_AIX61  6573     1300208  RB_AIX_61
----------------------------------------------------------------------
```

To display additional information about a given multipathing device, the **lsattr** command can be used. Figure 4-68 shows that for this multipathing device, the *round_robin* algorithm is being used to distribute I/O across multiple paths.

```
# lsattr -El hdisk16
PCM             PCM/friend/fcpother                     Path Control Module              False
algorithm       round_robin                             Algorithm                        True
clr_q           no                                      Device CLEARS its Queue on error True
dist_err_pcnt   0                                       Distributed Error Percentage     True
dist_tw_width   50                                      Distributed Error Sample Time    True
hcheck_cmd      inquiry                                 Health Check Command             True
hcheck_interval 60                                      Health Check Interval            True
hcheck_mode     nonactive                               Health Check Mode                True
location                                                Location Label                   True
lun_id          0x1000000000000                         Logical Unit Number ID           False
lun_reset_spt   yes                                     LUN Reset Supported              True
max_retry_delay 60                                      Maximum Quiesce Time             True
max_transfer    0x40000                                 Maximum TRANSFER Size            True
node_name       0x5001738000d00000                      FC Node Name                     False
pvid            00cd5bc44aa6664e0000000000000000        Physical volume identifier       False
q_err           yes                                     Use QERR bit                     True
q_type          simple                                  Queuing TYPE                     True
queue_depth     40                                      Queue DEPTH                      True
reassign_to     120                                     REASSIGN time out value          True
reserve_policy  no_reserve                              Reserve Policy                   True
rw_timeout      30                                      READ/WRITE time out value        True
scsi_id         0xf0300                                 SCSI ID                          False
start_timeout   60                                      START unit time out value        True
unique_id       261120017380000D019AD072810XIV03IBMfcp  Unique device identifier         False
ww_name         0x5001738000d00140                      FC World Wide Name               False
```

Figure 4-68 lsattr command

More information about AIX multipathing

For more information about AIX multipathing, see the following link:

http://publib.boulder.ibm.com/infocenter/aix/v6r1/topic/com.ibm.aix.baseadmn/doc/baseadmndita/devmpio.htm

4.11 Troubleshooting

Though it is difficult to anticipate what specific problems might be experienced during a SAN Boot installation, following are some common resources that can be referenced to assist with the diagnostic process:

- Review the installation configuration based upon *XIV Storage System: Host Attachment and Interoperability*, SG24-7904.
- Review the installation based upon the *XIV Host Attachment Guide* for the XIV Host Attachment Kit installed:

  http://www-947.ibm.com/support/entry/portal/Downloads_and_fixes/Hardware/System_Storage/Disk_systems/Enterprise_Storage_Servers/XIV_Storage_System_%282810,_2812%29
- Contact IBM support.

5

SAN Boot implementation with IBM SAN Volume Controller and IBM Storwize v7000

Storage area networks enable you to share homogeneous storage resources across the enterprise. For many companies, however, information resources are spread over various locations and storage environments with products from different vendors. The best solution takes advantage of the investment already made and provides growth when needed. Storage virtualization resolves this situation by combining multiple storage devices into a single, logical resource, with a single view for easy management.

IBM has almost 40 years experience in virtualization technologies. Among IBM storage virtualization products, IBM SAN Volume Controller (SAN Volume Controller) has established itself in the Enterprise Arena as a credible and powerful solution that incorporates maximum benefits of Storage Virtualization. An exciting latest addition to the IBM virtualization family is IBM Storwize v7000, which is a new midrange virtualized disk system with built-in storage efficiency capabilities.

In this chapter, we describe the SAN Boot implementation and best practices using IBM SAN Volume Controller and IBM Storwize v7000. Because of the commonality of the implementation procedures for SAN Boot with SAN Volume Controller and Storwize V7000, we have provided only the instructions for Storwize V7000. In other words, the information given mentioning Storwize V7000 can apply equally to the SAN Volume Controller.

© Copyright IBM Corp. 2012. All rights reserved.

5.1 What storage virtualization is

In a SAN environment, with the increase of data loads, the complexity and capacity of your storage environment also increases. But if you choose to virtualize your storage, you can reduce storage network complexity. *Storage virtualization* is an intelligent “layer” or abstraction that pools storage from multiple storage devices into a common storage pool. Virtualized storage appears as one device to the server-operating systems and can be centrally managed and provisioned from a single view, thus offering simplified management services across heterogeneous devices.

The term *storage virtualization* describes the shift from managing physical volumes of data to managing logical volumes of data. It aims to “mask” SAN complexity by aggregating multiple storage devices into a common managed “virtual” storage pool and by isolating servers from the physical storage.

Figure 5-1 shows an overview of Storage Virtualization, where a storage pool is being created from storage subsystems offered by various vendors. The Virtualization engine presents virtual disks as volumes to hosts.

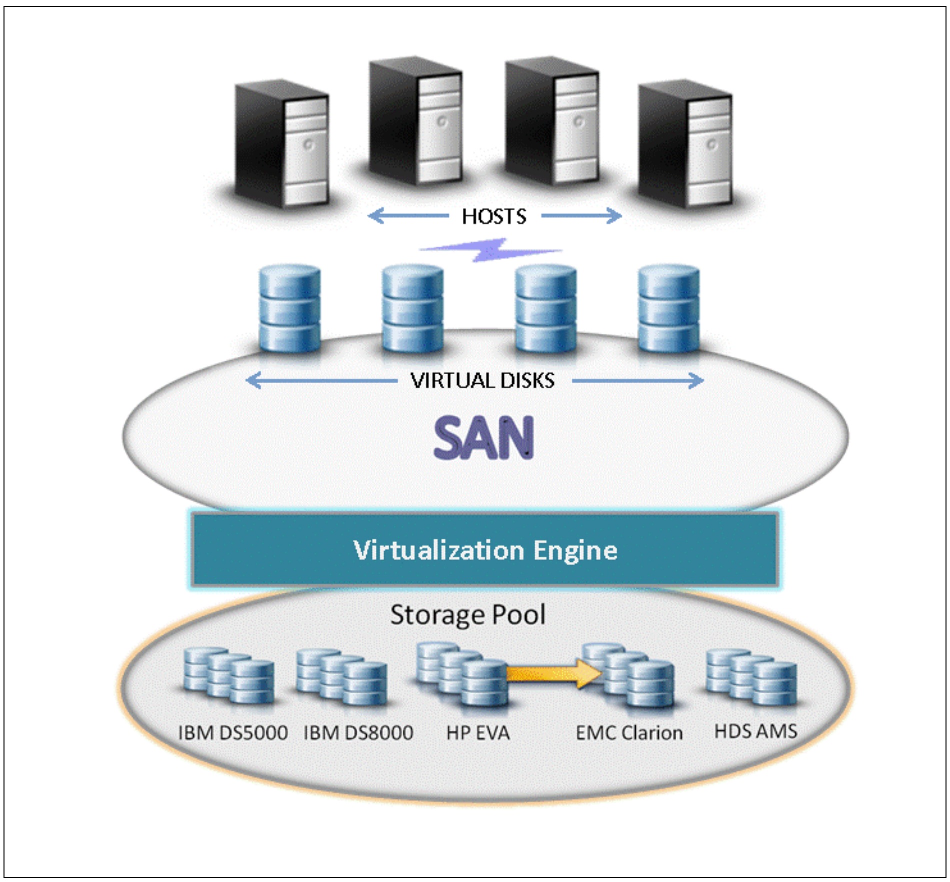


Figure 5-1 Storage Virtualization

Storage Virtualization delivers a multitude of advantages for the virtualized enterprise. It improves application and information availability, makes it easier to manage stored information, and increases productivity by simplifying management of the storage infrastructure.

To a business executive, storage virtualization presents reduced IT cost and complexity while helping to improve the business responsiveness, storage administrator productivity, and maximum storage utilization.

To a Storage Administrator, storage virtualization offers the following benefits:

- Simplified day-to-day management:

  Virtualization helps simplify management of SAN comprised of various storage devices from different vendors by eliminating the need to manage multiple-device interfaces, commands and service-level requirements across your SAN. It helps reduce device-specific training time and costs.

- Minimum Server/Network reconfigurations:

  You can remove and change storage-volume definitions and alter assignments from one storage device to another, or even add a new storage device, without requiring server and/or network reconfiguration.

- Reduced service outages:

  Virtualization helps bring advanced storage-management functionality, like advanced copy services, to the entire SAN and not just to individual devices. Backup and restore operations, as well as data migration and hardware and software upgrades can be handled non disruptively, improving data availability.

- Dynamic resource allocation:

  Virtualization enables dynamic allocation and management of storage resources either manually or by automated, customizable policies. It allows for smooth growth in changing workload conditions.

  For example, virtualization allows a server requiring additional storage to find unused space on another storage device. Conversely, a server requiring less storage reallocates the space back into the storage pool.

As a leader in both storage technology and virtualization solutions, IBM offers hardware, software and a full line of services designed to help meet all modern virtualization needs.

For details on the IBM current storage virtualization offerings, see this website:

http://www.IBM.com/systems/storage/virtualization/

For more information and helpful tutorials on Virtualization, see the 'Storage Networking Industry Association' (SNIA) website:

http://www.snia.org

The next two sections provide an overview of the major IBM virtualization solution offerings for enterprise and midrange markets, IBM San Volume Controller, and IBM Storwize V7000.

5.1.1 IBM SAN Volume Controller

IBM SAN Volume Controller is a proven IBM offering that has been delivering benefits to customers for over 6 years. It is designed to pool storage volumes into a reservoir of capacity for centralized management. SAN Volume Controller helps to hide the boundaries among disk systems, which simplifies management and enables customers to focus on managing storage as a resource to meet business requirements and not as a set of boxes.

SAN Volume Controller demonstrates scalability with the fastest disk SPC-1 benchmark results. It can virtualize IBM and non-IBM storage (over 200 systems from IBM, EMC, HP, Dell, NetApp, and so on.)

SAN Volume Controller offers these benefits:

- Overall, SAN Volume Controller helps reduce the cost of managing SAN by enabling tiers of storage, which allows clients to purchase cheaper disks for their data that does not need the high performance disks.
- Helps improve storage utilization by making better use of existing storage and control growth. By combining the capacity of multiple disk arrays into a single pool of storage, SAN Volume Controller enables clients to manage storage as a business resource available for deployment as required, and not as set of boxes.
- Enables clients to implement multi-vendor strategies because it can manage so many different disk arrays.
- Designed to improve application availability by allowing changes to storage and move data without without disrupting your applications.
- Offers greater efficiency and productivity for storage management staff. Storage administrators can manage their single pool of storage from a central point and they only need to learn a single interface.
- Offers network-based replication, which enables greater choice when buying storage
- Helps improve Business Continuity by supporting continuous operations and enabling copy services to be applied across all the managed storage. In combination with IBM TotalStorage Productivity Center, SAN Volume Controller also supports dynamically adding storage to applications automatically and on demand.

Figure 5-2 shows an overview of SAN Volume Controller, virtualizing source storage from various vendors.

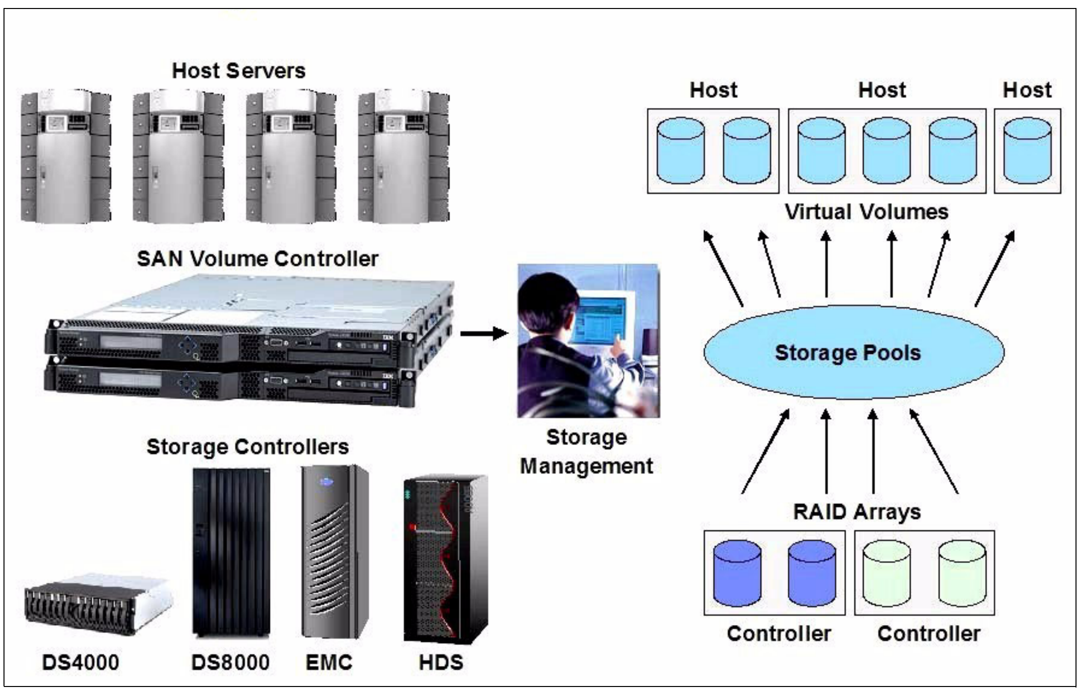


Figure 5-2 SAN Volume Controller overview

The SAN Volume Controller is highly scalable, providing an easy growth path to two-n nodes (grow in a pair of nodes). Presently, SAN Volume Controller supports FC and iSCSI SAN interfaces. It is external storage RAID controller-independent, providing a continual and ongoing process to qualify additional types of controllers. SAN Volume Controller is able to utilize disks internally located within the nodes (solid state disks) as well as the disks locally attached to the nodes (SAS drives).

On the SAN storage that is provided by the disk subsystems, the SAN Volume Controller manages multiple tiers of storage, provides block-level virtualization (logical unit virtualization), and provides advanced functions to the entire SAN, as follows:

- Large scalable cache
- Advanced Copy Services
- FlashCopy (point-in-time copy)
- Metro Mirror and Global Mirror (remote copy, synchronous/asynchronous)
- Non-disruptive and current data migration between storage tiers.

All communication between SAN Volume Controller nodes is performed through the SAN. All SAN Volume Controller configuration and service commands are sent to the system through an Ethernet network.

For detailed information on SAN Volume Controller components, see *Implementing the IBM System Storage SAN Volume Controller V6.1*, SG24-7933.

The latest details regarding SAN Volume Controller supported hardware list, device driver, firmware and recommended software levels are available at this website:

http://www-01.ibm.com/support/docview.wss?uid=ssg1S1003697

IBM SAN Volume Controller models

The IBM System Storage SAN Volume Controller hardware configuration consists of a cluster of IBM 2145 servers, called SAN Volume Controller Storage Engines. Each SAN Volume Controller Storage Engine is a node; two nodes make up a node pair (I/O group); up to four I/O groups constitute a SAN Volume Controller cluster. The cluster is scalable from one to four node pairs (I/O groups), and is supported by dual read/write cache. This design is intended to avoid single points of failure and support concurrent maintenance.

SAN Volume Controller 2145-CF8 Storage Engine

This component is the latest SAN Volume Controller model, based on the IBM System x3550M2 server running an Intel Xeon 5500 2.4-GHz quad-core processor. It has four 8-Gbps FC ports and up to 24 GB of cache per engine. It also provides scale-out high performance solid-state device support, with the ability to have up to four solid-state devices installed directly into each SAN Volume Controller hardware engine.

Figure 5-3 shows the front view of the SAN Volume Controller 2145-CF8.

Figure 5-3 2145-CF8

Tip: New engines can be intermixed in pairs with other engines in SAN Volume Controller clusters. Mixing engine types in a cluster results in VDisk throughput characteristics of the engine type in that I/O group.

SAN Volume Controller 2145-8A4 Storage Engine (for entry edition)

Based on the IBM System x3250 server running an Intel Xeon E3110 3.0GHz dual-Core processor. It has four 4-Gbps FC ports and 8 GB of cache.

Figure 5-4 shows SAN Volume Controller 2145-8A4.

Figure 5-4 2145-8A4

The above-mentioned are new offerings of SAN Volume Controller models. However, SAN Volume Controller firmware 6.1 is also supported on IBM System Storage SAN Volume Controller 2145-CF8, 2145-8A4, 2145-8G4, 2145-8F4, or 2145-8F2 storage engines with dual uninterruptible power supplies.

Support: SAN Volume Controller V6.1.0 software is not supported on clusters containing 2145-4F2 nodes.

5.1.2 IBM Storwize v7000: Overview

IBM Storwize V7000 introduces a new era in midrange storage by offering a virtualized storage system that complements virtualized server environments. It provides unmatched performance, availability, advanced functions, and highly scalable capacity never seen before in midrange disk systems.

IBM Storwize v7000 benefits

The IBM Storwize v7000 solution provides a modular storage system that includes the capability to consolidate external virtualized SAN-attached storage and its own internal storage and present it as a centralized pool of storage. It is built upon the IBM SAN Volume Controller (SAN Volume Controller) technology base and offers enterprise capabilities from the IBM DS8000 family technology. In the following sections, we describe the features of the IBM Storwize V7000 solution.

Enterprise capability at a midrange price that grows as your needs grow

This capability provides the following benefits:

- Storage virtualization, thin provisioning, disaster recovery, solid-state drive support, sophisticated copy functions
- Modular and affordable, to add capacity and capability as required
- Improve application availability and service levels through high-performance, near-instant backup and restore capabilities that help reduce downtime

High storage efficiency

This capability provides the following benefits:

- Virtualization features increases disk utilization by up to 30%
- Thin provisioning dramatically reduces disk storage needs
- Easy Tier® in combination with solid-state drives provide the performance required at lower cost

Breakthrough ease of use increases storage administrator efficiency

This capability provides the following benefits:

- Clients can grow capacity without adding more administrators
- Key tasks are simplified: storage configuration, provisioning, storage tiering (automate with Easy Tier)
- Consolidated server and storage management with IBM Systems Director

Control of costs

This capability provides the following benefits:

- Consolidate your heterogeneous environment with a single point of control
- Use efficiency features to maximize utilization of existing assets
- Data migration capability speeds time to value, minimizes application downtime

Figure 5-5 shows a high-level view of IBM Storwize V7000 virtualization.

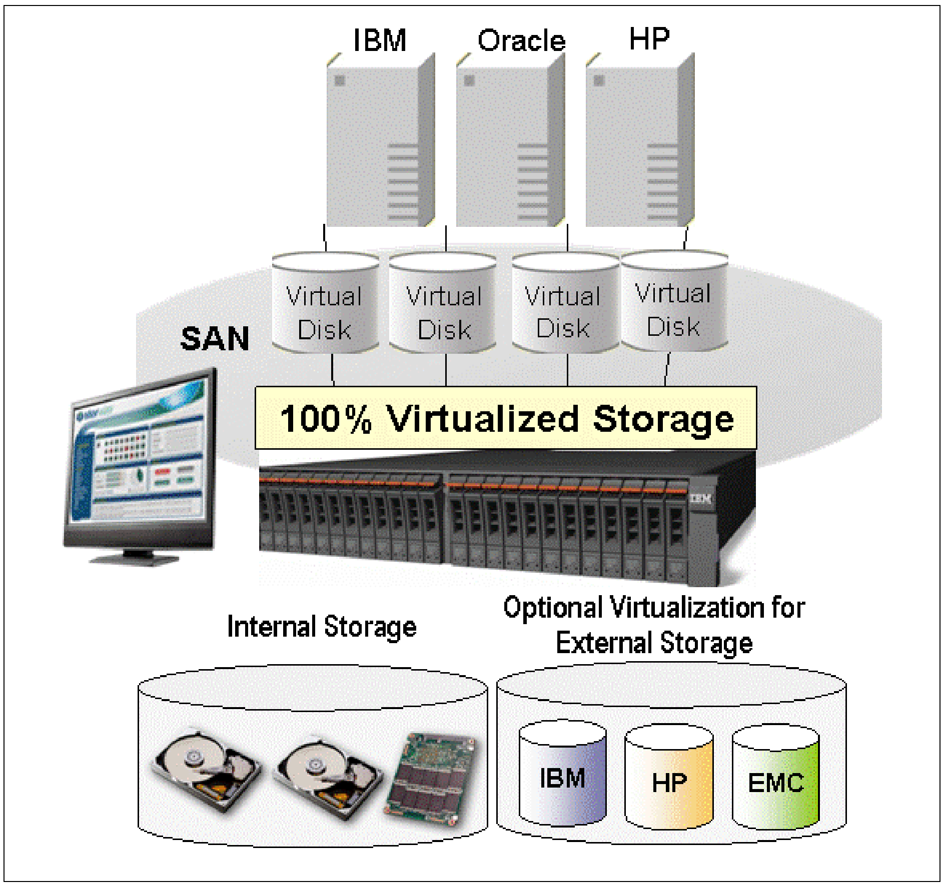


Figure 5-5 IBM Storwize V7000 virtualization

IBM Storwize v7000 models

The IBM Storwize v7000 is a modular storage solution composed of Control Enclosures and Expansion Enclosures. There are two models of the IBM Storwize V7000 control enclosures and two expansion enclosures.

Enclosures with 12 slots for 3.5-inch drives

The following options are available:

- 2076-112 - Control Enclosure
- 2076-212 - Expansion Enclosure

Enclosures with 24 slots for 2.5-inch drives

The following options are available:

- 2076-124 - Control Enclosure
- 2076-224 - Expansion Enclosure

All the models have internal disks slots for Serial Attached SCSI (SAS) drives only. Figure 5-6 shows the front view of the IBM Storwize V7000 models, with 12-bay enclosure (top) and 24-bay enclosure (bottom),

Figure 5-6 IBM Storwize V7000 Front View

> **Tip:** It is possible to mix 12 disk slots and 24 disk slots models.

Each IBM Storwize v7000 system will have one control enclosure, which contains two node canisters. The two nodes act as a single processing unit and form an I/O group that is attached to the fabric using the Fibre Channel or iSCSI ports. The two nodes provide high availability and fault tolerance, because if one node fails, the surviving node can automatically take over.

Figure 5-7 shows the rear view of a Control Enclosure (top) and an Expansion Enclosure (bottom).

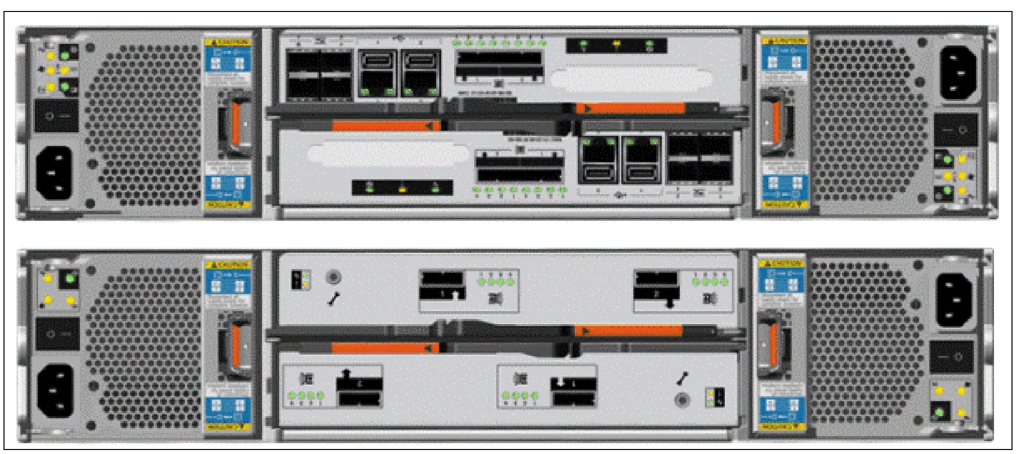


Figure 5-7 IBM Storwize V7000 rear view

The IBM Storwize V7000 hardware and software components, along with advanced features, are described at length in *Implementing the IBM Storwize V7000*, SG24-7938.

Drive features

IBM Storwize V7000 supports a maximum 240 drives. At the time of writing, the following drives are supported:

- 3.5-inch disk drives:
  - 2 TB 3.5" 7.2k Near-Line SAS disk
- 2.5-inch disk drives:
  - 300 GB 2.5" 10k SAS disk
  - 450 GB 2.5" 10k SAS disk
  - 600 GB 2.5" 10k SAS disk
  - 300 GB 2.5" E-MLC (enterprise-grade multilevel cell) SSD

> **Tip:** As more drives continue to be tested with Storwize V7000, check the latest supported drives at this website:
>
> https://www-03.IBM.com/systems/storage/disk/storwize_v7000/specifications.html

IBM Storwize v7000 connectivity

Before installing the IBM Storwize v7000 hardware, review the product manuals, including the *Environmental and Safety notices*, and the *IBM Storwize v7000 Quick Installation Guide,* at this website:

http://www.IBM.com/support/docview.wss?uid=ssg1S7003318

Host connections to the IBM Storwize V7000 are either SCSI over the Fibre Channel SAN or iSCSI over an Ethernet network:

- External Storage can be attached by Fibre Channel connections.
- IBM Storwize v7000 Expansion enclosures are connected by SAS network.

SAS Network

In Control Enclosure of IBM Storwize v7000, there are two Node Canisters. Each node canister has two SAS ports:

- Up to nine Expansion Enclosures can connect to the Control Enclosure by SAS network.
- Each Control Enclosure can have two SAS chains.
- The Control Enclosure is part of *chain 1*.
- The first Expansion Enclosure to be added must be connected to port 2. That is, it will be the start of *chain 2*.
- Subsequent additional Expansion Enclosures must be added such that the length of each chain is the same or only one different.

SAS cabling rules

The following rules apply:

- No more than five expansion enclosures can be chained to port 2 on the node canisters.
- No more than four expansion enclosures can be chained to port 1 on the node canisters.
- The list of enclosures attached to each node canister port 1 must match (same rule for port 2).
- The list of enclosures attached to any given port must be connected in the same order.

Figure 5-8 shows the recommended cabling scheme, with all 9 expansion enclosures attached to the control unit.

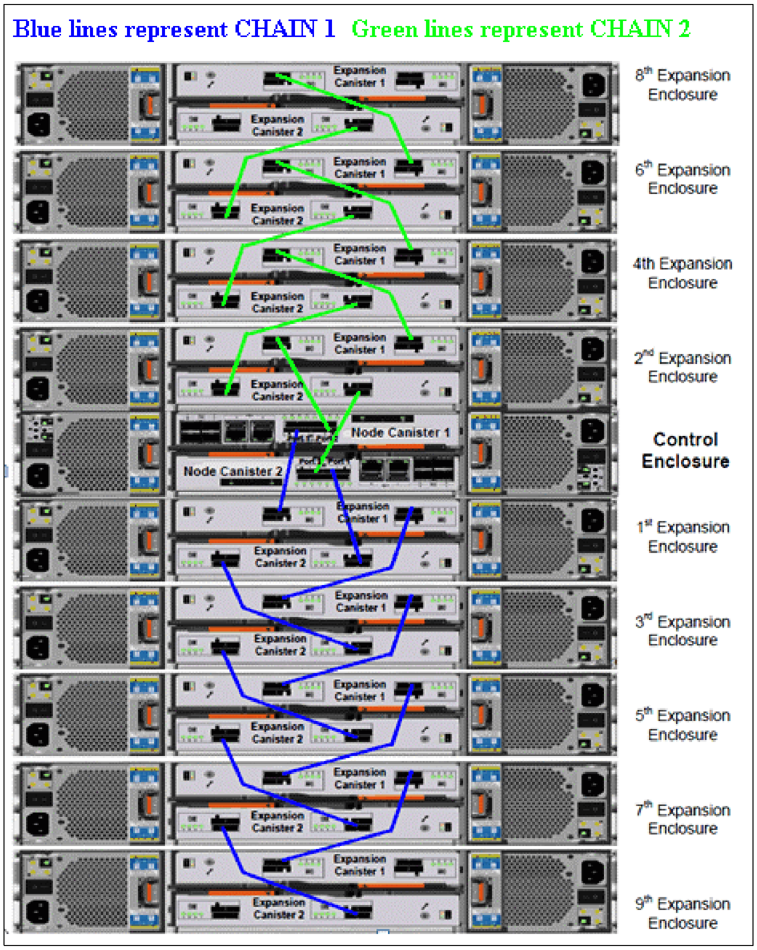


Figure 5-8 IBM Storwize v7000 cabling

IBM Storwize v7000 can connect to hosts by 8 Gb Fibre Channel (FC) and/or 1 Gb iSCSI.

5.2 Management and support tools

Simplified management is one of the key features of IBM SAN Volume Controller v6.1 and IBM Storwize v7000. It provides a leading-edge and intuitive web user interface, which has a number of preset configuration options and automated wizards to assist in resolving any events that might occur.

Presently, the following methods can be used to manage an IBM SAN Volume Controller and Storwize v7000 system:

- By using the Management Graphical User Interface (GUI)
- By using the SSH protocol Command Line Interface (CLI)
- Through the IBM Systems Storage Productivity Center (SSPC)
- Through an existing SAN Volume Controller Master Console

The primary tool for managing both IBM SAN Volume Controller and Storwize V7000 is the Graphical User Interface (GUI), which is a powerful browser-based interface used to configure, manage, and troubleshoot the IBM SAN Volume Controller and Storwize v7000 systems.

The CLI interface can be utilized for advanced customization, however, due to the debug-level strength of CLI commands, we suggest using it with the assistance of IBM Support.

5.2.1 Management GUI (Graphical User Interface)

The Management GUI provides a web user interface with a number of preset configuration options to reduce user input, and automated wizards to assist in resolving any events that might occur.

The new graphical user interface is designed to simplify storage management. It is used primarily to configure RAID arrays and logical drives, assign logical drives to hosts, replace and rebuild failed disk drives, and expand the logical drives. It allows for troubleshooting and management tasks, such as checking the status of the storage server components, updating the firmware, and managing the storage server. Finally, it offers advanced functions, such as FlashCopy, Disk Mirroring, and Enhanced Remote Mirroring.

The new GUI installs with the Cluster code (6.x). It is Web 2.0 based, giving the user the ability to connect directly to the cluster management IP address. It means that to use the GUI, no dedicated SAN Volume Controller GUI server, or SSPC, or Master Console is needed.

To log in to an IBM SAN Volume Controller or Storwize V7000, use the system management IP address in a browser. The system will ask you for *User Name* and *Password*. Figure 5-9 shows login panels for SAN Volume Controller and Storwize V7000.

> **Tip:** The default system *superuser* password is *passw0rd (with a zero)*. The password and the IP address are used to connect to the management GUI to finalize creating the system.

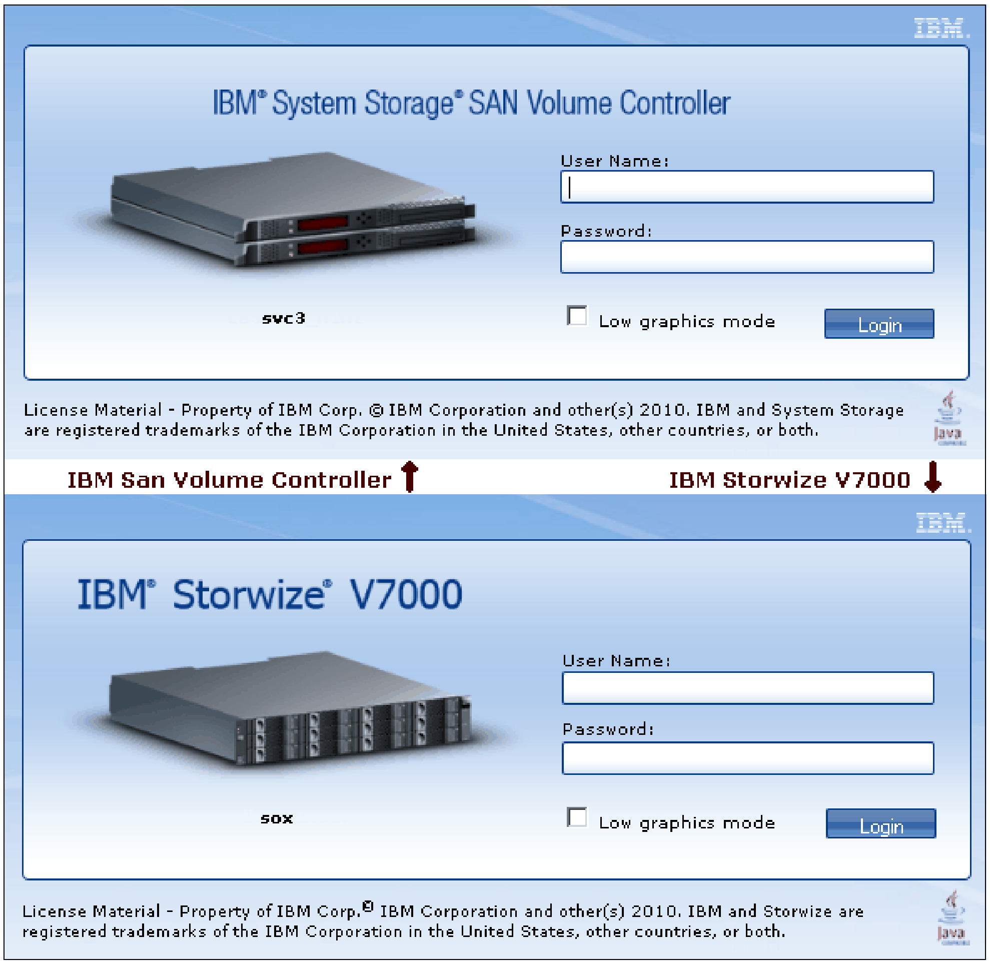


Figure 5-9 Login panels for SAN Volume Controller and Storwize V7000

> **Tip:** You might need to enable JavaScript in your browser. Additionally, if using Firefox, under Advanced JavaScript Settings, you need to disable or replace context menus, and allow cookies.

After logging into the Management GUI, the first view is *Getting Started*. This view is customized to your storage subsystem. Figure 5-10 shows the view for IBM Storwize V7000.

> **Tip:** SAN Volume Controller always shows 0 (zero) internal drives.

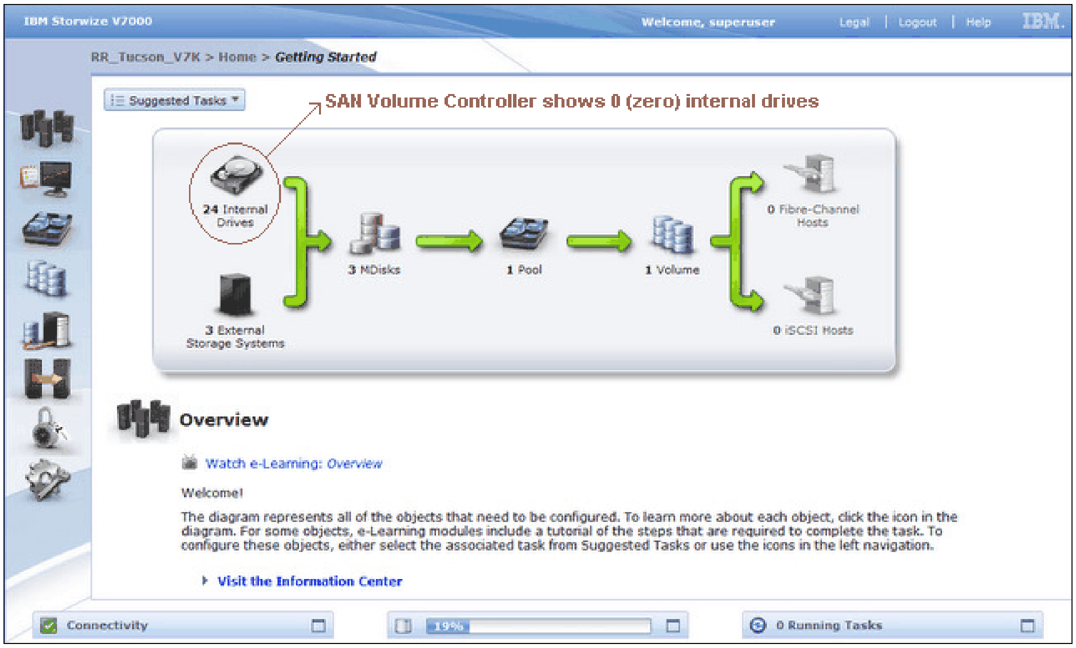


Figure 5-10   IBM Storwize V7000

This view shows the menu, **Suggested Tasks**, which intelligently predicts logical next actions. It contains a link to **Watch e-Learning: Overview** that shows eight videos explaining various management functions in the GUI. Another very helpful link is **Visit the Information Center,** which takes you to the IBM Storwize v7000 Information Center website.

On the far left hand side of the panel are eight function Icons. Figure 5-11 describes the icons and their associated functions.

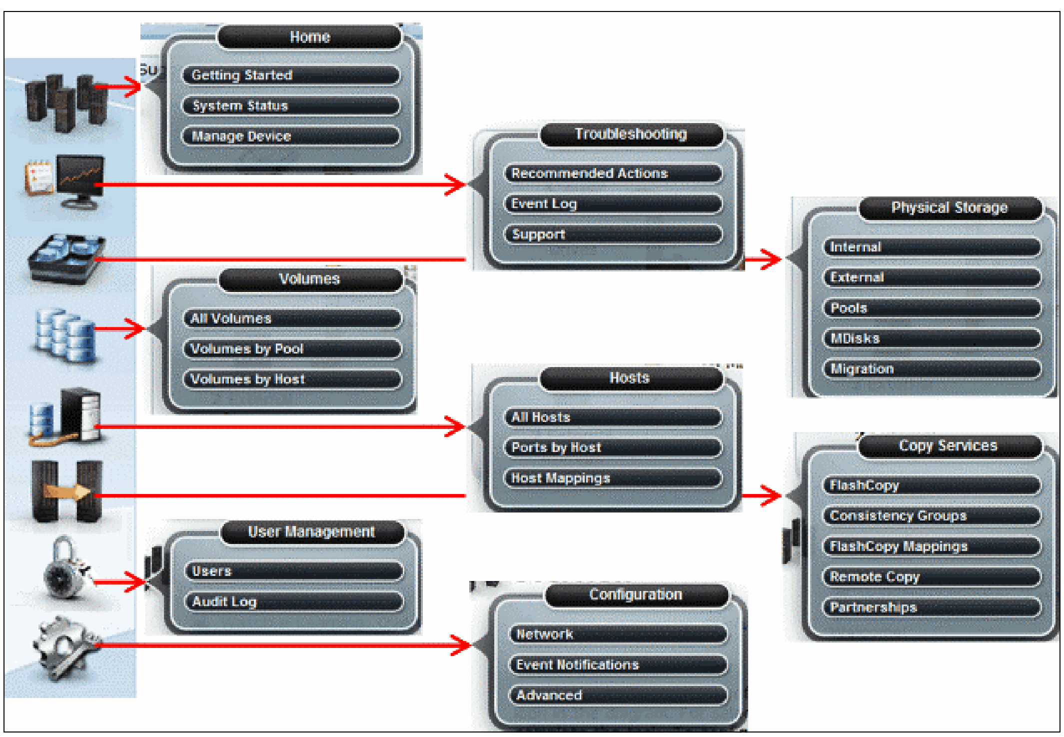


Figure 5-11   Navigation

The functions used in the SAN Boot process are described in 5.3, “SAN Boot with IBM Storwize v7000 and SAN Volume Controller” on page 285.

Detailed information regarding all functions and features of the Management GUI is presented in *Implementing the IBM Storwize V7000*, SG24-7938.

5.2.2 Command Line Interface (CLI) using SSH

Secure Shell (SSH) is a client/server network application that provides a secure environment to connect to a remote machine.

SSH provides secure data flow between the SSH server (client) and IBM Storwize V7000 (server) using the principles of public and private keys for authentication. A public key and a private key are generated together as a pair. The public key is uploaded to the SSH server, while the private key identifies the client and is checked against the public key during the connection.

Tip: The “admin” account is always used for SSH login and is independent from the created user.

You can find more information about using SSH in *Implementing the IBM Storwize V7000*, SG24-7938.

Detailed CLI information is also available in “IBM System Storage SAN Volume Controller and Storwize V7000 Command-Line Interface User's Guide”, available at this website:

http://www-01.ibm.com/support/docview.wss?uid=ssg1S7003318

5.2.3 Other supported management tools

With the introduction of SAN Volume Controller version 6.1 and Storwize V7000, the GUI now resides on the system. Therefore, neither SSPC nor master console is required to run the new Management GUI.

Tip: For SAN Volume Controller version 4.2.1 and earlier, the *master console* provided a single point from which to manage the SAN Volume Controller nodes. Beginning with SAN Volume Controller version 4.3.0, the *IBM System Storage Productivity Center* (SSPC) was introduced as an integrated hardware and software solution that provided a single point of entry for managing SAN Volume Controller clusters.

IBM System Storage Productivity Center (SSPC)

SSPC is a hardware appliance with preinstalled software that can help you improve and centralize the management of your storage environment through the integration of products. It provides a single point of management by integrating the functionality of the IBM Tivoli Storage Productivity Center with the storage devices element managers in an easy-to-use user interface for management. Consolidating several hardware and software components on a single tested system helps support consistent interoperability.

Detailed information about SSPC can be found in *IBM System Storage Productivity Center Deployment Guide*, SG24-7560.

Master Console

There is no longer a Master Console with version 6.x of the SAN Volume Controller. The master console provides a platform on which the subsystem configuration tools can be run. It also provides a platform for remote service, which allows the desktop to be shared with remote IBM service personnel. Master console gives access to the *SAN Volume Controller Console* through a web browser, and gives access to the command-line interface through a Secure Shell (SSH) session.

More information about the Master Console is available at the following website:

http://publib.boulder.ibm.com/infocenter/svc/ic/index.jsp?topic=/com.ibm.storage.svc510.console.doc/svc_overviewnodemgmt_224cwl.html

Although multiple master console servers can access a single cluster, you cannot concurrently perform configuration and service tasks if multiple servers are accessing one cluster. Due to this limitation as well as others, the preferred method for managing IBM Storwize V7000 or SAN Volume Controller is using the Management GUI.

5.3 SAN Boot with IBM Storwize v7000 and SAN Volume Controller

Booting servers from a storage area network (SAN) is being used increasingly in today's complex data center environments due to its significant benefits over traditional method of booting from local disks. SAN Boot enables organizations to maximize consolidation of their IT resources, minimize their equipment costs, and realize the considerable management benefits of centralizing the boot process.

In SAN Boot, you can deploy diskless servers in an environment where the boot disk is located on (often RAID-capable) storage connected to the SAN. The server (*initiator*) communicates with the storage device (*target*) through the SAN using the Fibre Channel host bus adapter (HBA).

The system downtime is greatly minimized in case a critical component such as a processor, memory, or host bus adapter fails and needs to be replaced. The system administrator needs to swap only the hardware and reconfigure the *HBA's BIOS*, *switch zoning*, and *host-port definitions* on the storage server. The system image still exists on the logical drive, therefore the server is fully operational after the hardware swap and configuration change is completed.

IBM SAN Volume Controller supports SAN Boot for all major operating systems. For the latest supported SAN Boot combinations, check the following link:

http://www-01.ibm.com/support/docview.wss?uid=ssg1S1003697

For information regarding SAN Boot supported configurations on IBM Storwize v7000, see this website:

http://www-01.IBM.com/support/docview.wss?uid=ssg1S1003703

5.3.1 Considerations before setting up SAN Boot

For all physical components in the SAN, you must meet the minimum hardware and software requirements and ensure that other operating environment criteria are met:

- To obtain software installation and configuration information, planning information and host attachment information, see the following references:
  - For IBM SAN Volume Controller:

    http://publib.boulder.ibm.com/infocenter/svc/ic/index.jsp
  - For IBM Storwize V7000:

    http://publib.boulder.ibm.com/infocenter/storwize/ic/index.jsp
- Verify that the SAN switches and/or directors to which the IBM Storwize v7000 will connect, and the host bus adapters (HBAs), are at supported firmware levels:
  - For IBM SAN Volume Controller:

    http://www-01.ibm.com/support/docview.wss?rs=591&uid=ssg1S1003697
  - For IBM Storwize V7000:

    http://www-01.IBM.com/support/docview.wss?uid=ssg1S1003703
- Check the latest firmware's configuration limits and restrictions:
  - For IBM SAN Volume Controller:

    http://www-01.ibm.com/support/docview.wss?uid=ssg1S1003704
  - For IBM Storwize V7000:

    http://www-01.ibm.com/support/docview.wss?uid=ssg1S1003702
- Check the status of the hardware in the SAN environment, including switches, hosts, HBAs, Hardware Management Console (HMC) and storage enclosures.
- Verify the firmware levels for the server's BIOS, HBA, switches, and so on, before starting the operating system installation process.
- Collect and validate all installation media to be used in the installation process.
- Check the license requirements and agreements.
- Review any special requirements recommended by the hardware vendors regarding your SAN equipment, and follow best practice guidelines.
- Consider the TCP/IP address requirements of the storage.
- Verify the availability and validity of the host bus adapter's (HBA) driver packages.
- Make sure that the storage device and the switch are both powered up and completely initialized before making any configuration changes.
- Verify that a unique Domain ID and IP address have been configured on all the switches.
- Take a backup of the switch fabric configuration before making any configuration changes.
- Review the restrictions and limitations regarding SAN Boot specified in the latest operating system documentation.

5.3.2 SAN connectivity

For Fibre Channel connections, the IBM Storwize V7000 nodes must always be connected to SAN switches only. Each node must be connected to each of the counterpart SANs that are in the redundant fabric.

Attention: All Fibre Channel devices *must* be connected through SAN fabrics and *must not* use direct connections.

The recommended SAN configuration comprises a minimum of two fabrics. Configuring your SAN with at least two independent switches, or networks of switches, ensures a redundant fabric with no single point of failure. If one of the two SAN fabrics fails, the configuration is in a degraded mode, but is still valid. A SAN with only one fabric is a valid configuration but risks loss of access to data if the fabric fails. SANs with only one fabric are exposed to a single point of failure. A simple example of recommended SAN is shown in Figure 5-12.

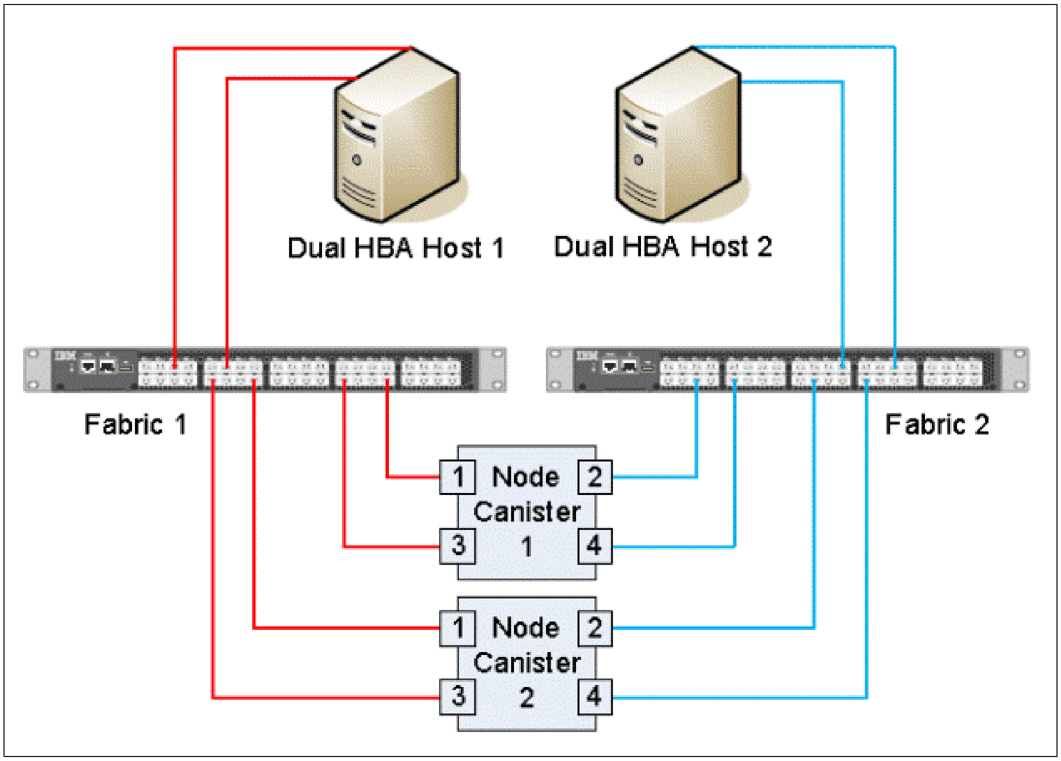


Figure 5-12 Recommended SAN

Support: Configurations with more than four SANs are not supported.

5.3.3 Zoning Considerations for Storwize V7000

A SAN is configured into a number of zones. A device using the SAN can communicate only with devices that are included in the same zones that it, itself, is in.

An IBM Storwize V7000 requires several distinct types of zones: a cluster zone, host zones, and disk zones. The intercluster zone is optional. The cluster zone contains all ports from all nodes in the cluster, unless you are using a dual-core fabric design. In a disk zone, the IBM Storwize V7000 nodes identify the storage systems. In brief, there must be three distinct zones in the fabric.

- *Cluster zone*: Create one zone per fabric with all of the Storwize V7000 ports cabled to this fabric to allow internode communication.

- *Host zone*: Create a host zone for each server accessing storage from the Storwize V7000 cluster.
- *Storage zone*: Create one storage zone for each storage subsystem that is virtualized by the Storwize V7000.

Each node has four ports and each I/O group has two nodes. Therefore, without any zoning in a dual SAN environment, the number of paths to a volume is four (4) multiplied by the number of host-ports. This rule exists to limit the number of paths that must be resolved by the multipathing device driver.

For optimal performance, limit a host with two Fibre Channel ports to only four paths: one path to each node on each SAN.

> **Support:** The number of paths through the network from the IBM Storwize V7000 nodes to a host must not exceed eight. Configurations in which this number is exceeded are not supported.

If you want to restrict the number of paths to a host, zone the switches so that each host bus adapter (HBA) port is zoned with one IBM Storwize V7000 port for each node in the cluster. If a host has multiple HBA ports, zone each port to a different set of IBM Storwize V7000 ports to maximize performance and redundancy.

IBM Storwize V7000 requires single-initiator zoning for all large configurations that contain more than 64 host objects. Each server Fibre Channel port must be in its own zone, which contains the Fibre Channel port and IBM Storwize V7000 ports. In configurations of less than 64 hosts, you can have up to 40 Fibre Channel ports in a host zone if the zone contains similar HBAs and operating systems.

> **Tip:** 512 logins are allowed per node Fibre Channel port. They include logins from server HBAs, disk controller ports, node ports within the same cluster, and node ports from remote clusters.

Regarding connectivity with external storage, in an environment where you have a fabric with multiple speed switches, best practice is to connect the Storwize V7000 and the disk subsystem to the switch operating at the highest speed.

For SAN Configuration and Zoning Rules, see the following website:

http://publib.boulder.ibm.com/infocenter/storwize/ic/index.jsp?topic=/com.ibm.storwize.v7000.doc/svc_configrulessummary_02171530.html

5.3.4 Best practices in SAN Boot implementations

The following general guidelines are intended to help employ best practices with SAN Boot implementations. These are by no means conclusive, and you must adhere to the best practice guidelines specified by the hardware and software vendors of all components in your SAN environment.

Avoiding a single point-of-failure

One of the major benefits of SAN adoption is high availability. Configure the following SAN components redundantly to avoid any single point of failure.

Storage controller redundancy

Configure storage arrays with multiple controllers to provide redundant array ports and avoid any single point of failures at the array controller level.

Disk redundancy

Configure the array using different RAID groups as required to provide redundancy at the disk level.

Path redundancy

Configure SAN infrastructure (switches, HBA ports) to provide redundancy and avoid any point of path failures.

Replacing hardware components in existing SAN Boot environment:

A hardware component carrying a WWPN, for example, an HBA, switch, node in IBM Storwize V7000 control enclosure, might need replacement at some point in SAN Boot environment due to failure, compatibility-issues, or simply for upgrade reasons. In such scenarios, you need to perform the following reconfigurations/verifications:

- The host must have the correct HBA WWPN in its definition. See 5.4.6, "Creating Fiber Channel hosts using the GUI" on page 302.
- The HBA and IBM Storwize V7000 host-port WWPNs must be correct in the working zones for SAN Boot.
- The HBA BIOS settings must be updated in the HBA's configuration utility (see the QLogic Fast!UTIL, reference in 5.5.2, "Configuring a QLogic adapter for SAN Boot" on page 315).

5.4 Provisioning storage for SAN Boot

IBM Storwize V7000 and IBM SAN Volume Controller are deployed as *Clusters*. A cluster is a group of nodes that presents a single configuration or management and service interface to the user. While referring to Storwize V7000 and SAN Volume Controller, the term *Cluster* is often used interchangeably with *System*.

Because the configuration, monitoring, and service tasks are performed at the cluster level, you can take advantage of their virtualization and advanced features only after configuring your cluster.

5.4.1 Creating a cluster using IBM SAN Volume Controller

Each node in the SAN Volume Controller is an individual server in a cluster, on which the SAN Volume Controller software runs. A system (cluster) is managed over an Ethernet connection by the use of a *management GUI* or *command-line* session. Management IP addresses are assigned to the cluster and all configuration settings are replicated across all nodes in the cluster.

Initial cluster configuration can be done on the SAN Volume Controller using the *front panel*. After you have installed all nodes, use the front panel of one of the SAN Volume Controller nodes to initiate the creation of the clustered system. The front panel display is arranged as shown in Figure 5-13.

Attention: To create a system (cluster), do not repeat these instructions on more than one node. After completing the steps for initiating system creation from the *front panel*, you can use the management GUI to create the system and add additional nodes to complete system configuration.

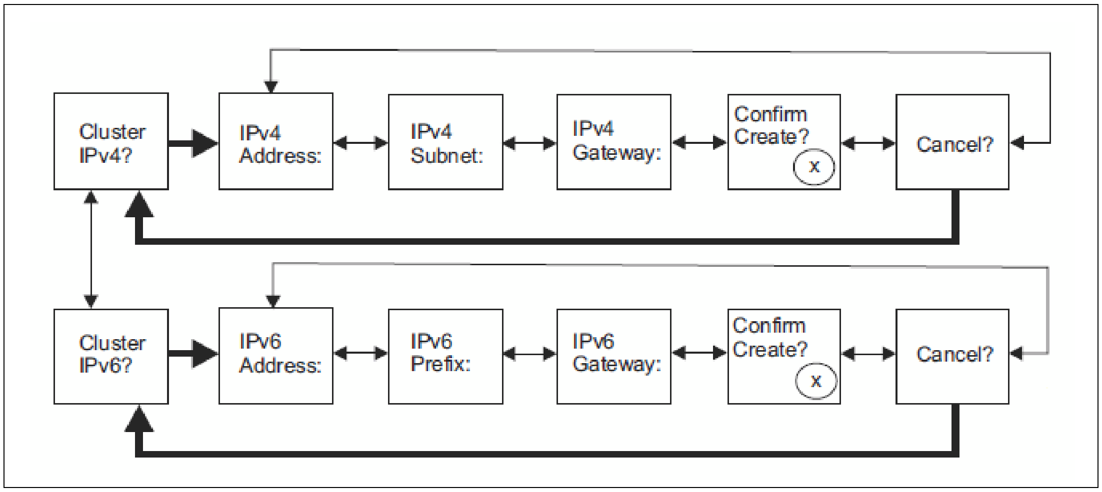


Figure 5-13 Front Panel Options

When you create the system, you must specify either an IPv4 or an IPv6 system address for port 1. Use the front panel and follow these steps to create and configure the system:

1. Choose a node that you want to make a member of the system that you are creating.
2. Press and release the **up** or **down** button until *Action* is displayed. Press and release the **select** button.
3. Depending on whether you are creating a system with an IPv4 address or an IPv6 address, press and release the **up** or **down** button until either *New Cluster IP4* or *New Cluster IP6* is displayed. Press and release the **select** button.
4. Press and release the **left** or **right** button until either *IP4 Address:* or *IP6 Address:* is displayed.
5. Use the following steps to create a cluster with an IPv4 address:
   a. The first IPv4 address number is shown. Set the desired IPv4 Address.

      Press the **select** button to enter *edit* mode. After editing, press **select** button again to leave the *edit* mode

      Press the **up** button to increase value and the **down** button to decrease it. Press the **right** or **left** buttons to move to the number field that you want to update.

   b. Press the **right** button to move to the next stage. *IP4 Subnet:* is displayed.
   c. Enter edit mode by pressing **select** button. Set the subnet to desired value, then leave edit by pressing the **select** button again.
   d. Press the **right** button to move to the next stage. *IP4 Gateway:* is displayed.
   e. Press and release the **right** button until *Confirm Create* is displayed. Press the **select** button to complete this task.

Now, on the service display panel, *Cluster:* must be displayed on line 1, and a temporary, system-assigned clustered system name displayed on line 2.

6. After you have created the clustered system on the front panel with the correct IP address format, you can finish the system configuration by accessing the management GUI.

To access the management GUI, point your supported browser to the management IP address. You can add nodes, rename objects and perform management tasks using the management GUI. The section, "First time setup," in *Implementing the IBM Storwize V7000*, SG24-7938 details the next steps.

For step-by-step instructions for creating a cluster in SAN Volume Controller, see "Setting up the SAN Volume Controller cluster" in *Implementing the IBM System Storage SAN Volume Controller V6.1*, SG24-7933.

5.4.2 Creating a cluster on IBM Storwize V7000

The IBM Storwize V7000 provides an easy-to-use initial setup contained within a USB key. The USB key is delivered with each storage system and contains the initialization application called "InitTool.exe". A system management ip address, the subnet mask and the network gateway address are required. The initialization application creates a configuration file on the USB key.

The IBM Storwize V7000 will start the initial setup as soon as you plug in the USB key with the newly created file in the storage system. If the system is configured not to autorun the USB keys, open the USB key from *My Computer* (in Windows) and double-click the InitTool.exe file. Follow the intuitive wizard to create a cluster.

For detailed, step-by-step instructions for initial configuration of IBM Storwize V7000, see "First time setup" in *Implementing the IBM Storwize V7000*, SG24-7938

5.4.3 IBM Storwize V7000 software components

Following are some terms that you need to know in order to understand SAN Volume Controller and Storwize V7000 implementation.

Node

A node is a single processing unit that provides virtualization, cache, and copy services for the cluster. SAN Volume Controller nodes are deployed in pairs called I/O groups. One node in the cluster is designated the configuration node.

Each node has two Ethernet ports that can be used for management. Ethernet port 1 must be configured with a management IP address and must be connected on all nodes in the system. The use of Ethernet port 2 is optional.

Configuration node

At any point in time, only one node in the system can operate as the focal point for configuration and monitoring requests. This node is called the configuration node. It is the only node that active cluster IP addresses, and is the only node that receives cluster management request. You can use one or more of these addresses to access the system through the management GUI or the command-line interface (CLI).

Cluster characteristics:

- In SAN Volume Controller, a cluster can consist of between two and eight SAN Volume Controller nodes.
- In Storwize V7000, nodes of a cluster are the node canisters inside a control enclosure. Two nodes are allowed per cluster in Storwize V7000.

I/O group

The nodes are always installed in pairs. Each pair of nodes is called an I/O group. The I/O groups take the storage that is presented to the SAN by the storage systems as MDisks and translates the storage into local disks, known as Volumes, which are used by applications on the hosts.

Tip: The IBM Storwize V7000 can have one I/O group, while the SAN Volume Controller can have up to 4 I/O groups in a cluster.

I/O traffic for a particular volume is, at any one time, managed exclusively by the nodes in a single I/O group. All I/O operations that are managed by the nodes in an I/O group are cached on both nodes.

When a write operation is performed to a volume, the node that processes the I/O duplicates the data onto the partner node that is in the I/O group. After the data is protected on the partner node, the write operation to the host application is completed. The data is physically written to disk later.

Read I/O is processed by referencing the cache in the node that receives the I/O. If the data is not found, it is read from the disk into the cache. The read cache can provide better performance if the same node is chosen to service I/O for a particular volume.

Tip: Although a clustered system in IBM SAN Volume Controller can have eight nodes within it, the nodes manage I/O in independent pairs. It means that the I/O capability of the SAN Volume Controller scales well, because additional throughput can be obtained by adding additional I/O groups.

Managed disk (MDisk)

An MDisk is a SCSI disk that is presented by a RAID controller and that is managed by the cluster. The MDisk is not visible to host systems on the SAN.

Storage pool

A storage pool is a collection of storage capacity, made up of MDisks, which provides the pool of storage capacity for a specific set of volumes. Each MDisk in a storage pool is broken up logically into a number of extents. The size of these pools can be changed (expanded or shrunk) at run time by adding or removing MDisks, without taking the storage pool or the volumes offline.

Tip: A single cluster can manage up to 128 storage pools. A storage pool can be a collection of up to 128 MDisks.

Extent

An extent is a number of contiguous logical blocks, a fixed-size unit of data, which is used to manage the mapping of data between MDisks and Volumes.

Volume

A volume is a logical device that appears to host systems attached to the SAN as a SCSI disk. Each volume is associated with exactly one I/O group. It will have a preferred node within the I/O group.

For detailed information regarding SAN Volume Controller and Storwize V7000 components, see the following references:

- "SAN Volume Controller Components" in *Implementing the IBM System Storage SAN Volume Controller V6.1*, SG24-7933
- "IBM Storwize V7000 Components" in *Implementing the IBM Storwize V7000*, SG24-7938

As described before, the IBM Storwize V7000 takes basic storage units called managed disks and use them to make one or more storage pools. These storage pools then provide the capacity to create volumes for use by hosts. Figure 5-14 shows the functions to create storage pool and then volumes.

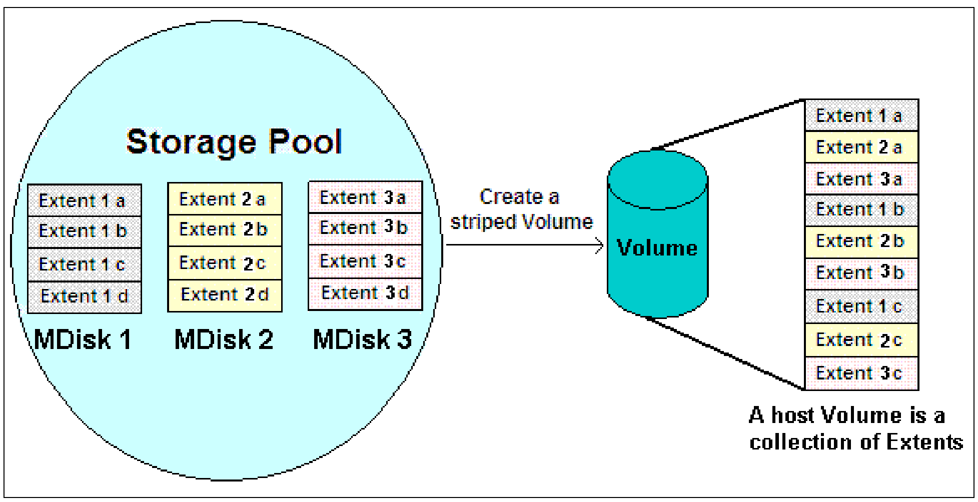


Figure 5-14 Storage pools and volumes

IBM Storwize V7000 RAID configuration presets

For internal drives in IBM Storwize V7000, RAID configuration presets are used to configure drives based on recommended values for the RAID level and drive class. Each preset has a specific goal for the number of drives per array, the number of spare drives to maintain redundancy, and whether the drives in the array are balanced across enclosure chains, thus protecting the array from enclosure failures. Table 5-1 describes the RAID presets that are used for hard disk drives for the IBM Storwize V7000 storage system.

Table 5-1 HDD RAID Presets

Preset	Purpose	RAID level	Drives per array goal	Spare goal	Chain balance
Basic RAID 5	Protects against a single drive failure. Data and one strip of parity are striped across all array members.	5	8	1	All drives in the array are from the same chain wherever possible.
Basic RAID 6	Protects against two drive failures. Data and two strips of parity are striped across all array members.	6	12	1	All drives in the array are from the same chain wherever possible.
Basic RAID 10	Protects against at least one drive failure. All data is mirrored on two array members.	10	8	1	All drives in the array are from the same chain wherever possible.
Balanced RAID 10	Protects against at least one drive or enclosure failure. All data is mirrored on two array members. The mirrors are balanced across the two enclosure chains.	10	8	1	Exactly half of the drives are from each chain.
RAID 0	Provides no protection against drive failures.	0	8	0	All drives in the array are from the same chain wherever possible.

For details on presets, see *Implementing the IBM Storwize V7000*, SG24-7938.

5.4.4 Creating storage pools

IBM Storwize V7000 discovers LUNs from the internal or external storage as one or more MDisks. SAN Volume Controller discovers LUNs from external storage as MDisks. The concept and steps for managing external storage system are the same as for internal storage. These MDisks will ultimately be added to a storage pool in which volumes will be created and mapped to hosts as needed.

This section provides a brief overview of steps to create a storage pool. At the GUI interface, go to the **Physical Storage** menu, as shown in Figure 5-15, and select **External** (with Storwize V7000, you can also select **Internal**).

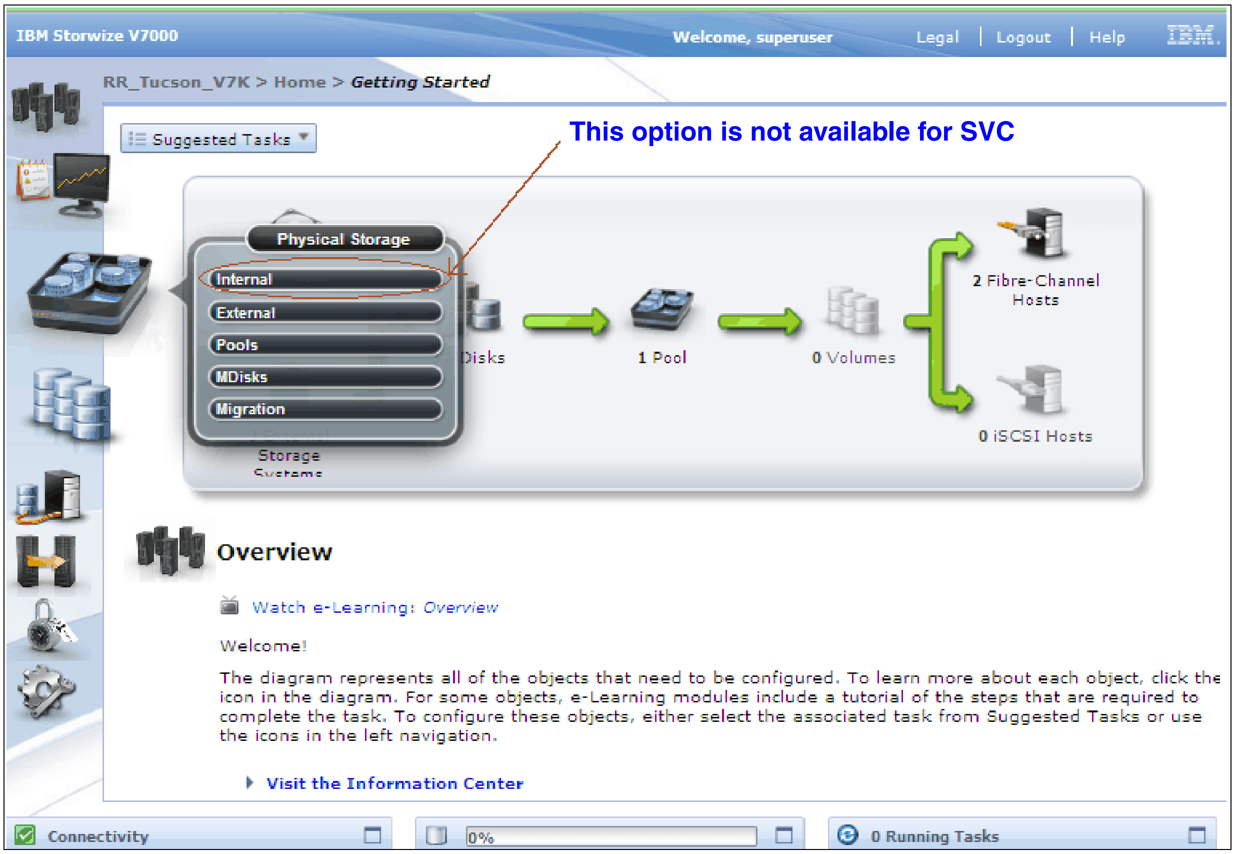


Figure 5-15 Physical Storage

The *External* panel, as shown in Figure 5-16, gives an overview of all your drives. On the left of the panel there is a catalog of the internal drives. In this catalog you can find out how many different types of disks are in this Storwize V7000 storage system. Select any type on the left, and the member MDisks will be displayed on the right.

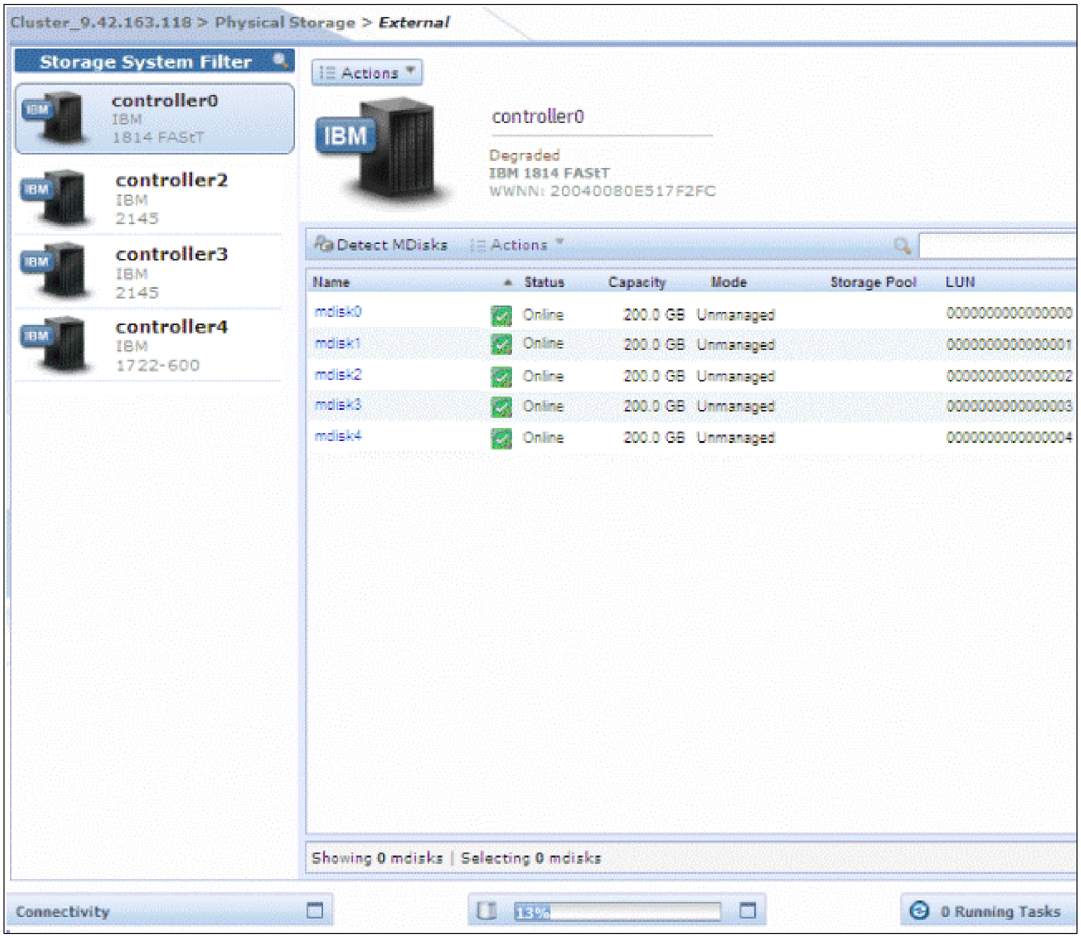


Figure 5-16 Available MDisks

Tip: Avoid splitting arrays into multiple LUNs at the external storage system level. When possible, create a single LUN for each array for mapping to IBM Storwize V7000.

Click **Physical Storage** → **Storage Pools**, as shown in Figure 5-15 on page 295, then click **new Pool**. It will present a dialogue box as shown here in Figure 5-17.

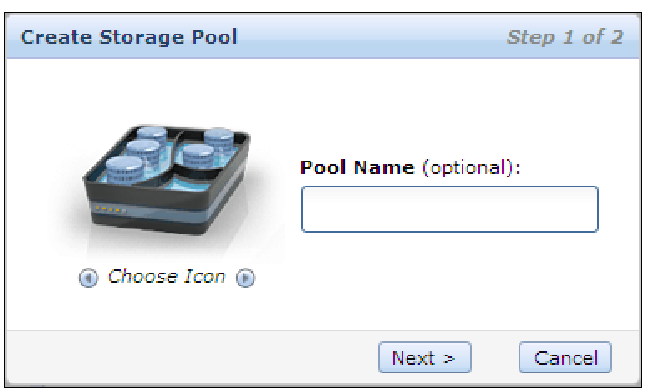


Figure 5-17 New storage pool

Next, click **Physical Storage** → **Storage Pools**, as shown in Figure 5-15 on page 295, then select **MDisks**. It will present a view of all *Managed* and *Unmanaged* MDisks, as shown here in Figure 5-18.

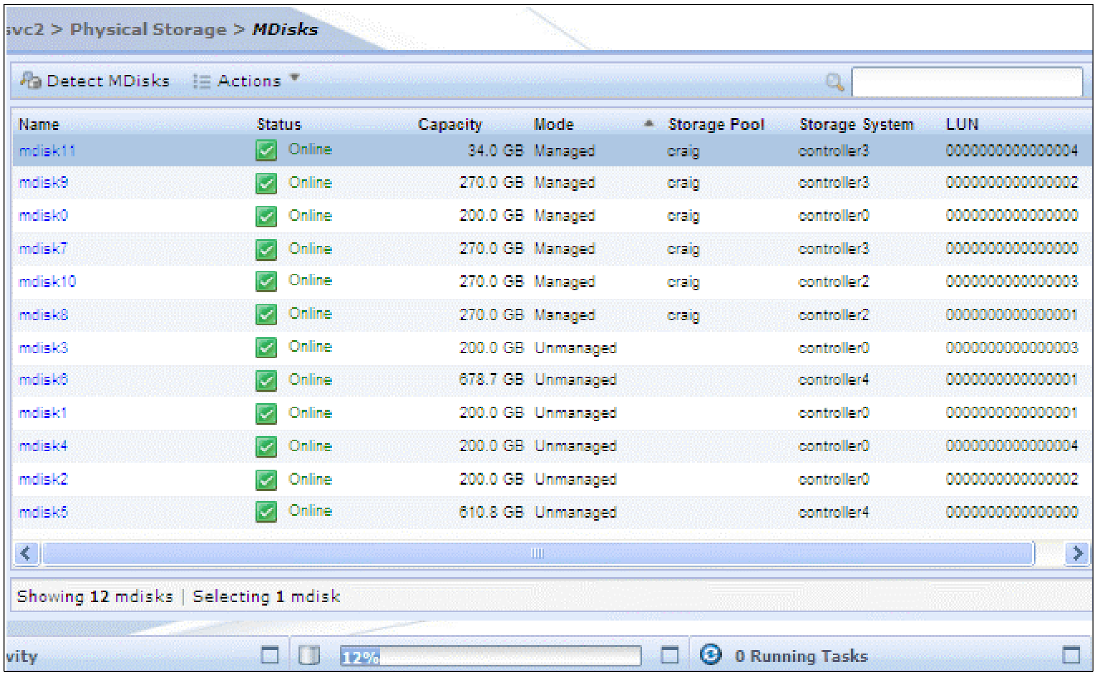


Figure 5-18 MDisks View

Tip: *Managed* means that the MDisk is assigned to a storage pool and provides extents that volumes can use. Whereas, *Unmanaged* means that the MDisk is not a member of any storage pools, in other words, it is not being used by the IBM Storwize V7000.

The LUNs presented by external storage to IBM Storwize V7000 will be discovered as unmanaged MDisks.

You can *right-click* an unmanaged disk and click **Add to Pool**. In the following dialogue box, select the Pool to which you want to add the MDisk, then click the button **Add to Pool**, as shown in Figure 5-19.

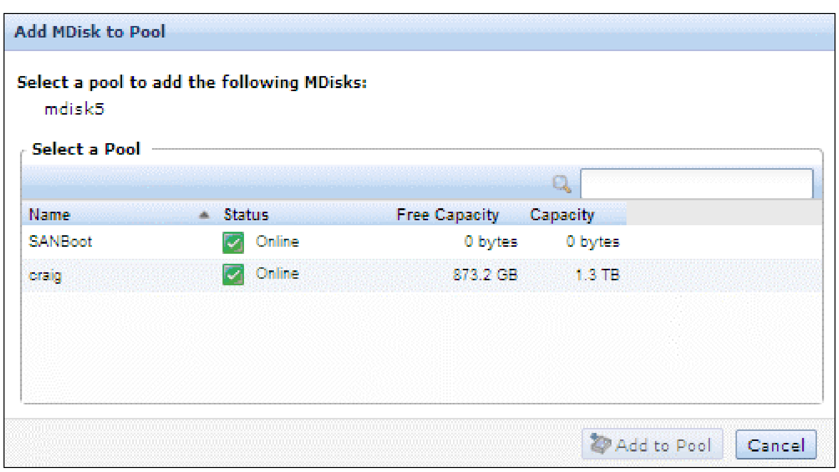


Figure 5-19 Add MDisk to Pool

After creating storage pools, the following steps are required to complete the basic setup of your environment:

- Create new volumes (5.4.5, "Creating a volume for SAN Boot" on page 298)
- Map volumes to the host (5.4.7, "Mapping SAN Boot volume to the host" on page 306)
- Discover the volumes from the host and specify multipath settings (5.5, "Host configuration" on page 312)

5.4.5 Creating a volume for SAN Boot

For this step, we assume that you have created storage pools (5.4.4, "Creating storage pools" on page 295)

Creating a new volume

To start creating a new volume, open the **All Volumes** section of the Management GUI as shown in Figure 5-20.

Figure 5-20 All Volumes

Click **New Volume** and a new window is presented, as shown in Figure 5-21.

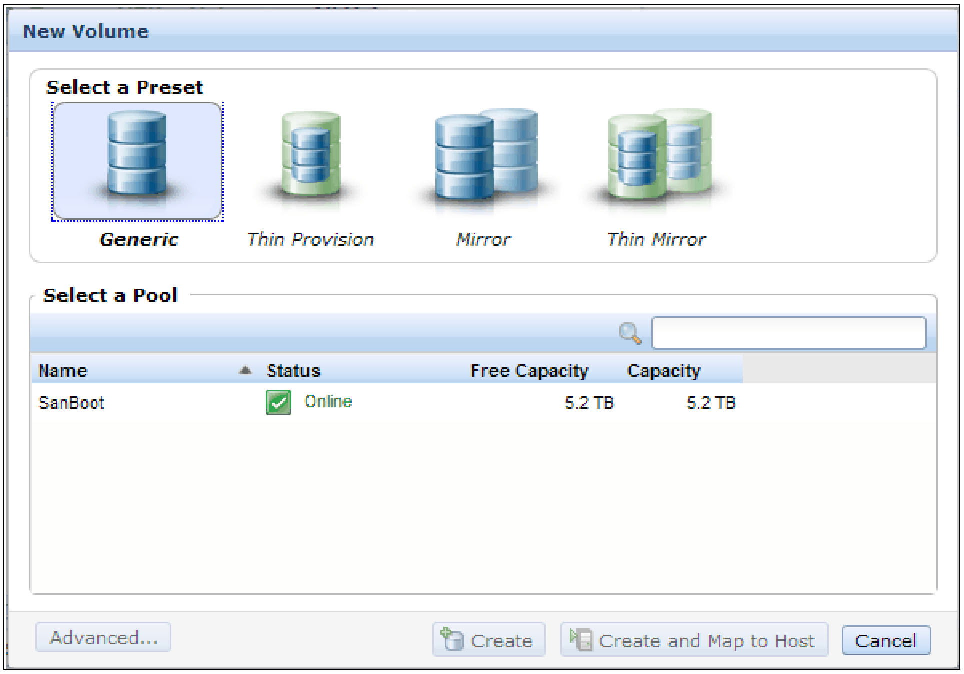


Figure 5-21 Select Preset

By default, all volumes that you create are striped across all available MDisks in one storage pool. The GUI provides the presets: Generic, Thin-provision, Mirror, and Thin-mirror.

For the purpose of this book, we selected Generic Volume type. If you want to create a thin provisioned, mirrored, or thin mirrored volume, see *Implementing the IBM Storwize V7000*, SG24-7938.

Creating a generic volume

We choose a generic volume as shown in Figure 5-22. Afterwards, we select the pool in which the volume should be created. Select the pool by clicking it.

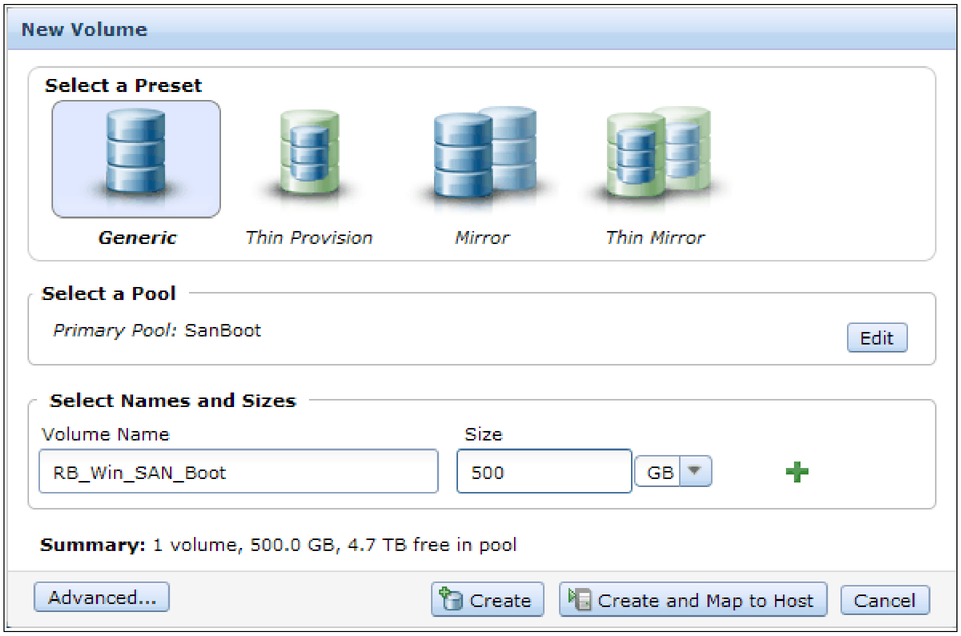


Figure 5-22 Generic Volume

Tip: The maximum volume size SAN Volume Controller is 1 PB (petabyte), while for Storwize V7000 it is 256 TB.

Figure 5-23 shows the new volume, which is also identified by UID (unique identifier).

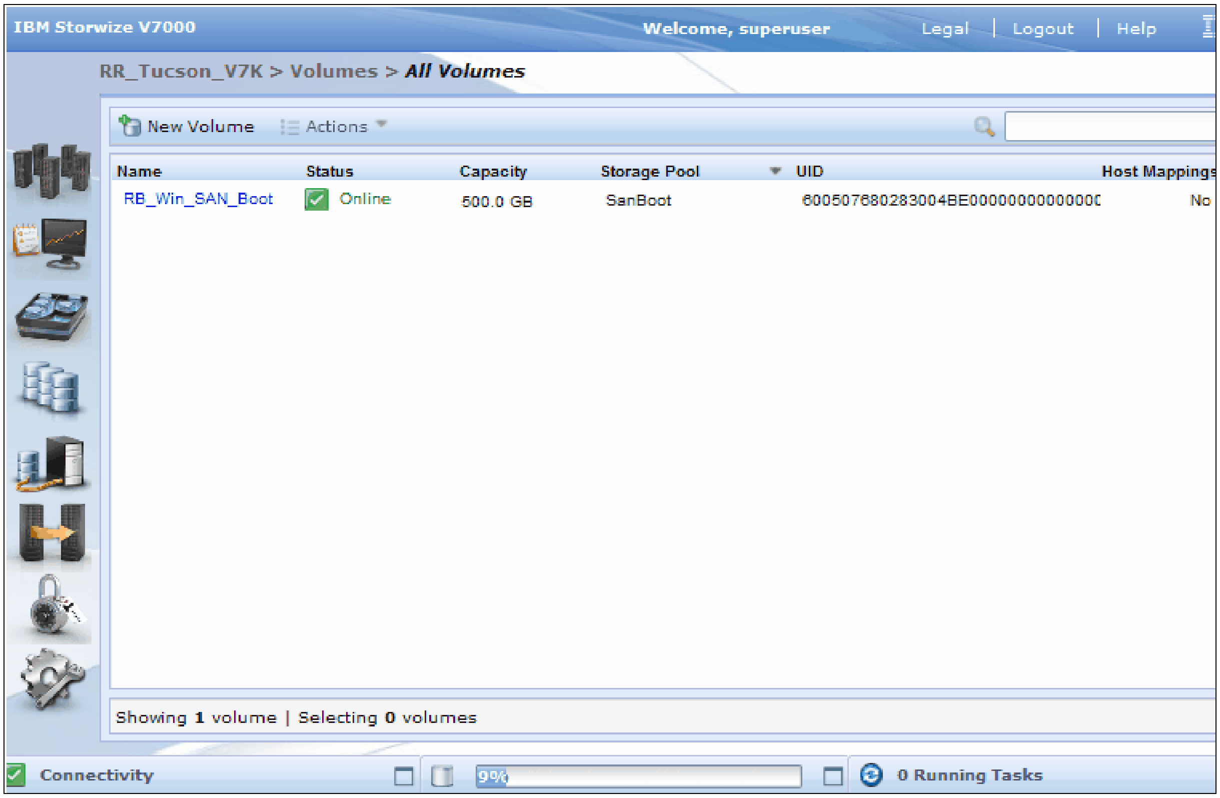


Figure 5-23 Volume Created

5.4.6 Creating Fiber Channel hosts using the GUI

This section describes how to create Fibre Channel hosts using the IBM Storwize v7000 GUI.

Preparation

For this procedure, you need the worldwide port names of the HBA ports that are going to connect to the Storwize v7000 host ports:

- To determine the QLogic HBA WWPNs, see 5.5.1, “Finding WWPNs of QLogic FC HBAs” on page 313.
- For the Emulex HBA WWPNs, see 5.5.3, “Finding WWPNs of Emulex FC HBAs” on page 321.

In the Management GUI, open the host configuration by selecting **All Hosts** (Figure 5-24).

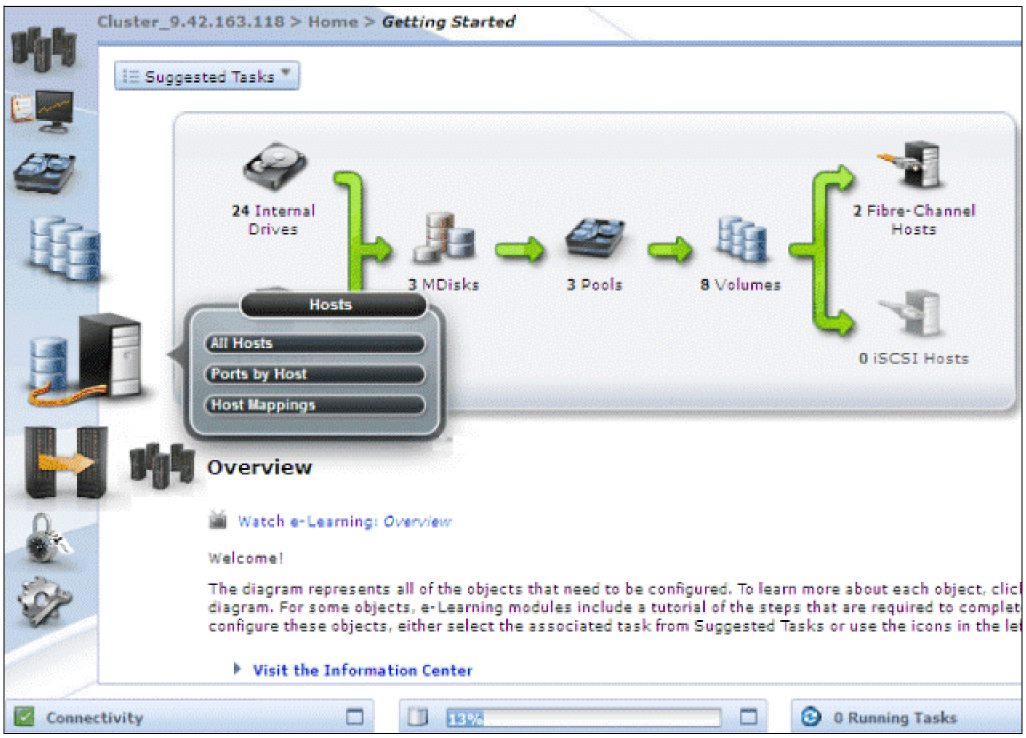


Figure 5-24 All Hosts

Create a new host by clicking **New Host** or **Actions** → **New Host**. That will start the wizard as shown in Figure 5-25.

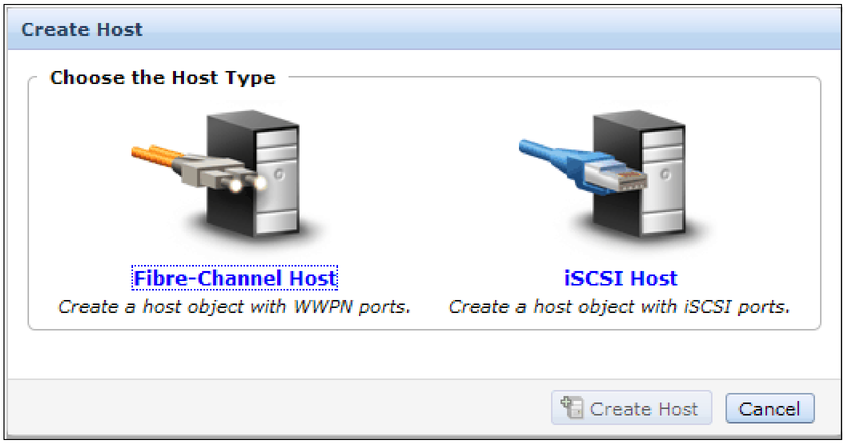


Figure 5-25 Create Host

Creating the Fiber Channel hosts

Follow these steps:

1. Click **Fibre Channel Host** as shown in Figure 5-25. The resulting window will be displayed as shown in Figure 5-26.

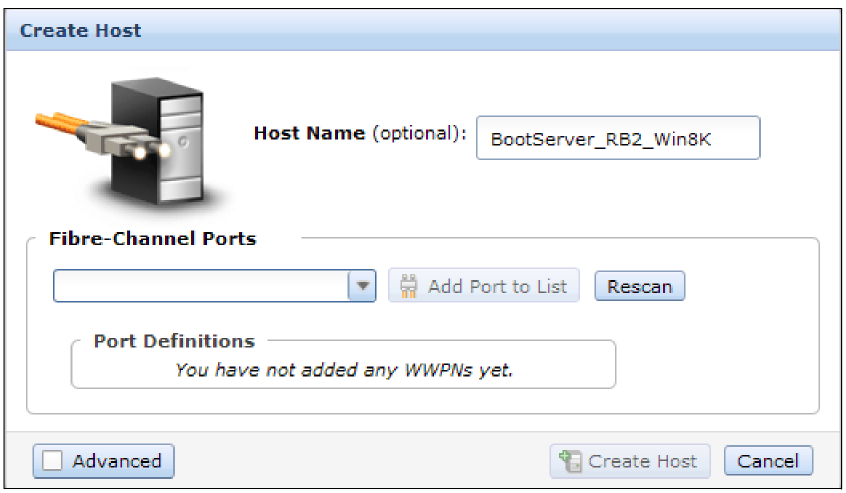


Figure 5-26 Create FC Host

2. Enter a Host Name and click the Fibre Channel Ports box to get a list of all known worldwide port names (WWPNs) as shown in Figure 5-27.

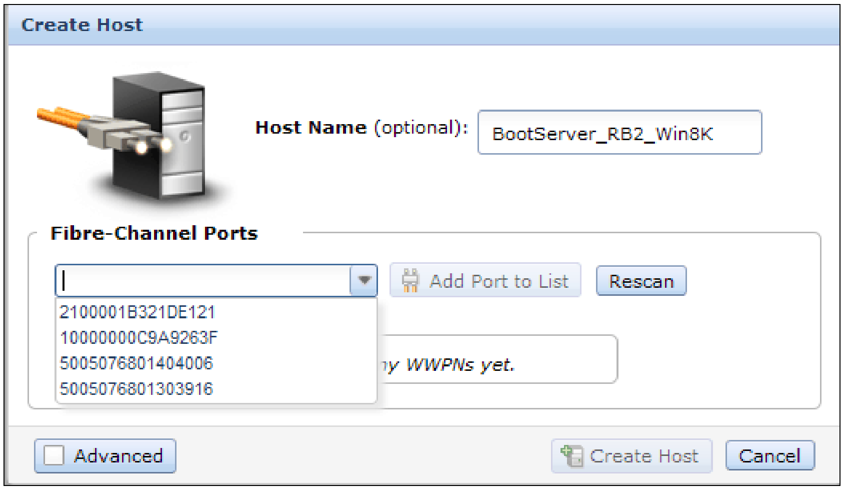


Figure 5-27 Select host-port

The drop-down shows list of the worldwide port names (WWPN) of the host bus adapters (HBAs) that are visible to the Storwize V7000's host-ports. If you do not see the desired HBA WWPN in the list, check the path between the HBA and storage's host-ports.

3. Click **Add Port to List** and then **Create Host**, as shown in Figure 5-28 and Figure 5-29.

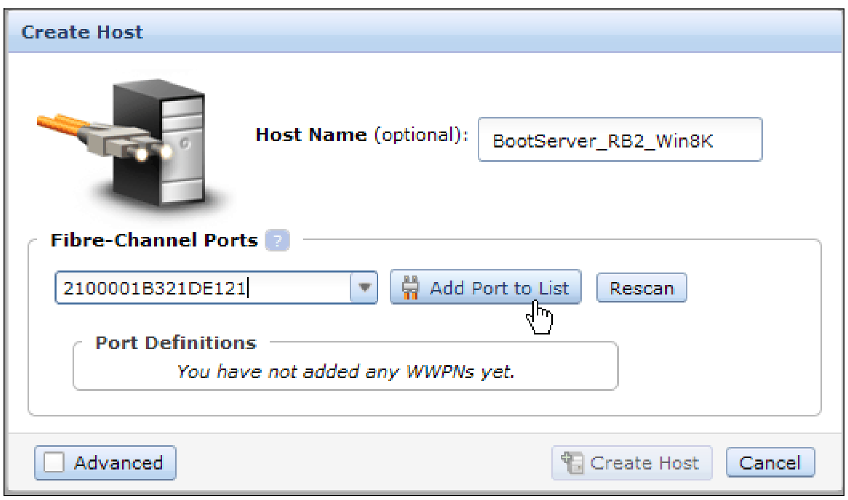


Figure 5-28 Add host-port

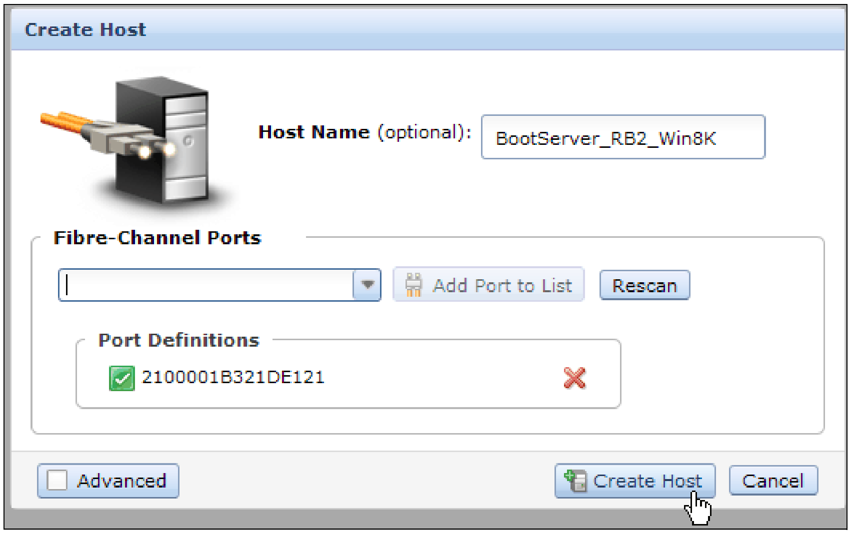


Figure 5-29 Create Host

> **Tip:** If you want to create hosts that are offline, or not connected at the moment but will be added later, it is also possible to enter the WWPNs manually. Just type them into the Fibre Channel Ports Box and add to List.

The new host is now displayed in the **All Hosts** view, as shown in Figure 5-30.

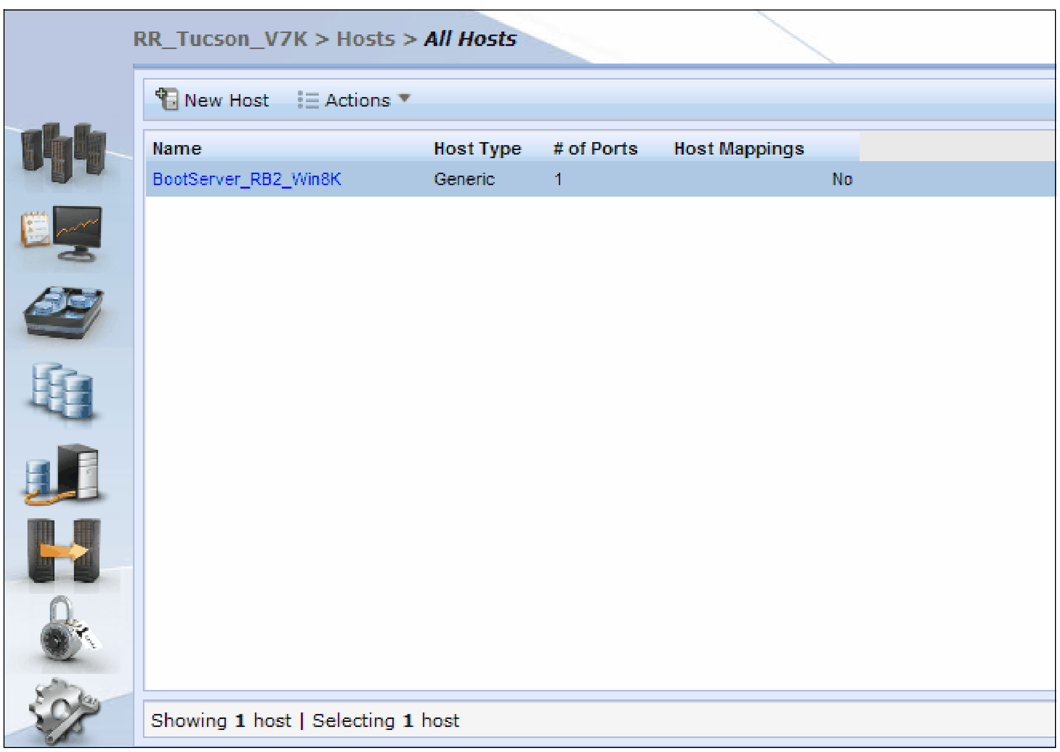


Figure 5-30 Host without Mappings

5.4.7 Mapping SAN Boot volume to the host

This section describes how to map the volume that we use for SAN Boot, to a host. At this point, we continue to map the volume we have created under “Creating a generic volume” on page 300.

As the first step of the mapping process, you need to select a host to which the new volume should be attached.

1. Click the icon **Hosts** → **All Hosts**, as shown in Figure 5-31.

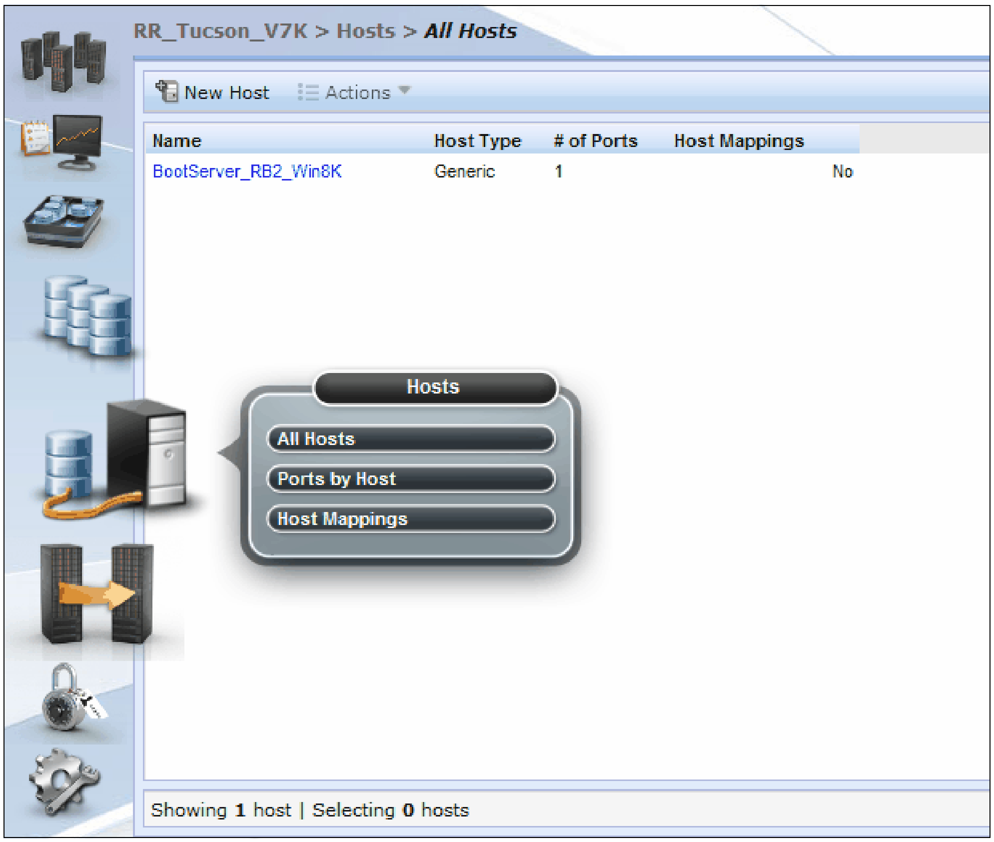


Figure 5-31 Selecting Host

2. Right-click the host and select **Modify Mappings**, as shown in Figure 5-32.

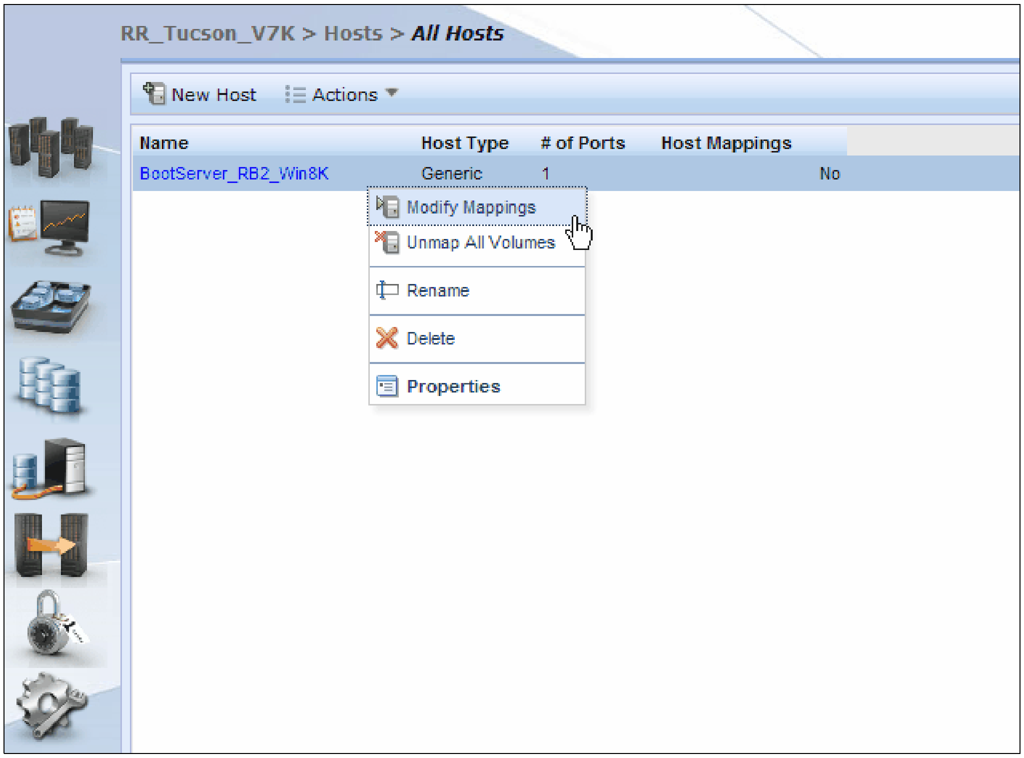


Figure 5-32 Modify Mappings

Now the wizard opens the **Modify Mappings** section, as shown in Figure 5-33.

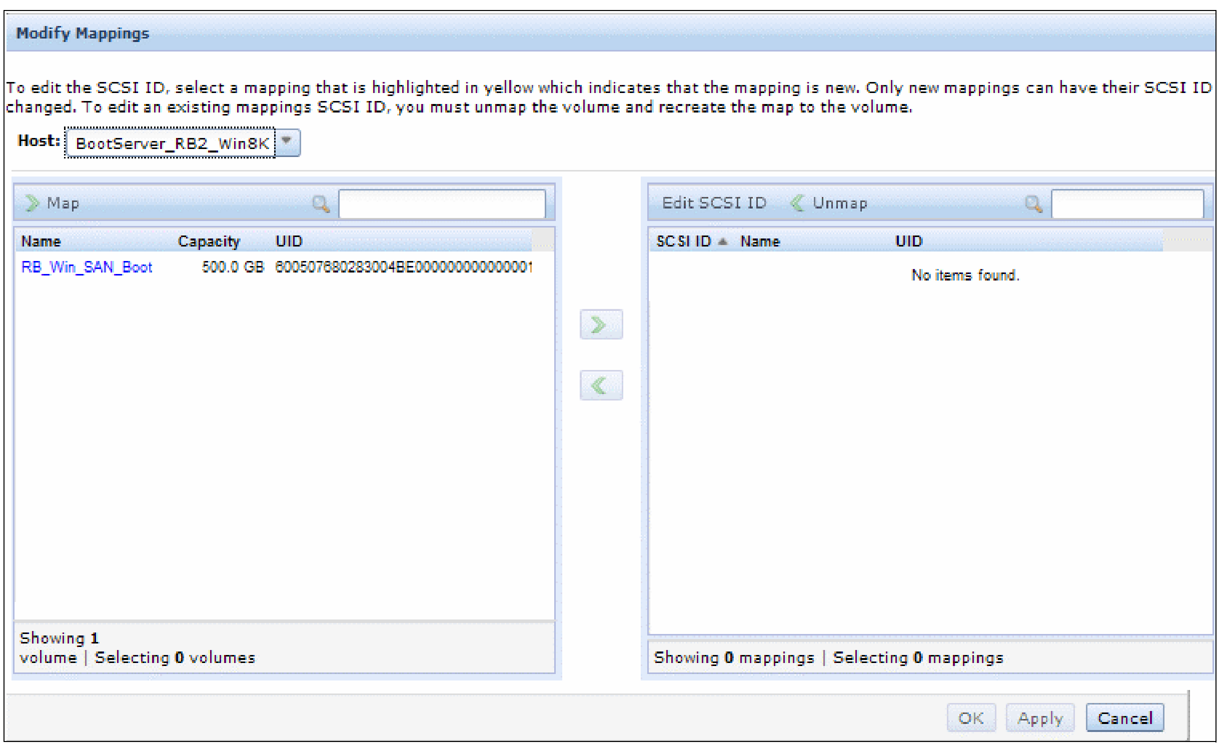


Figure 5-33 Create SCSI ID

3. Add the desired WWPN to the right pane by clicking the right-arrow icon. Click **Apply**, as shown in Figure 5-34.

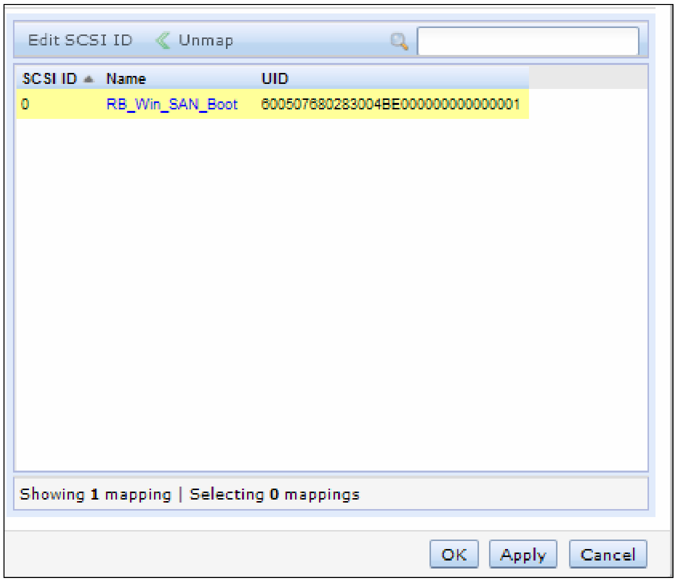


Figure 5-34 Add WWPN

The host is now able to access the volumes, as shown in Figure 5-35.

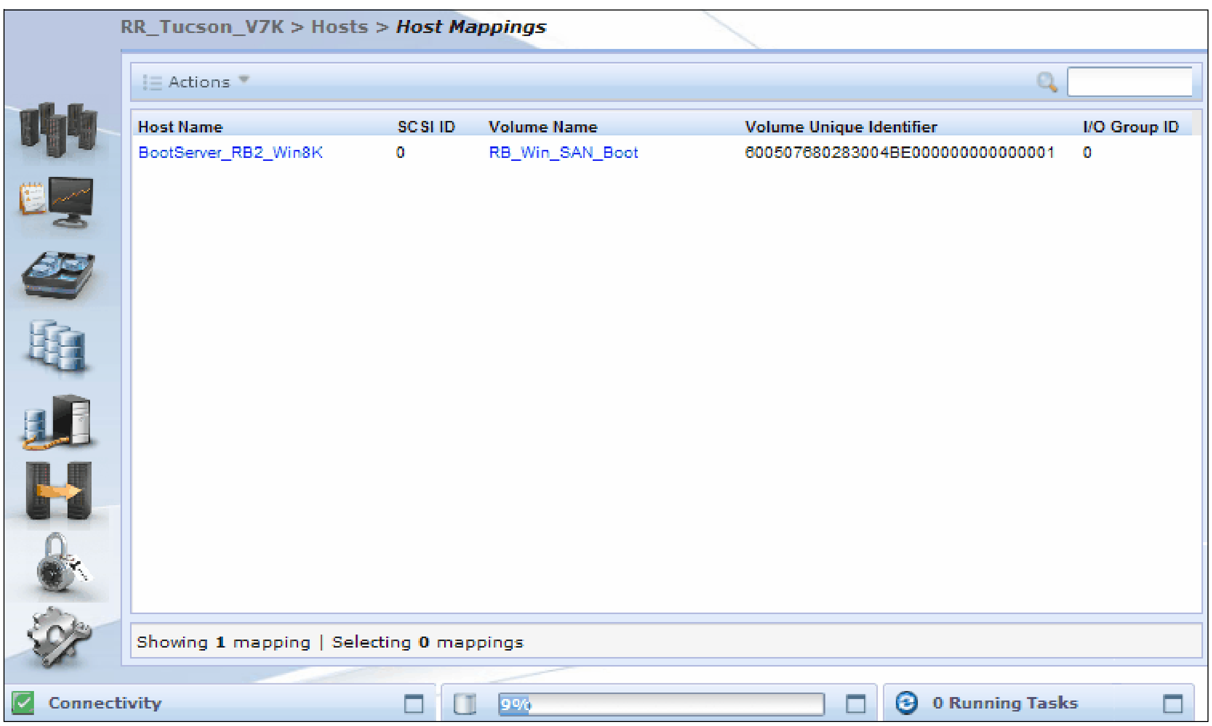


Figure 5-35 Mapped to Volume

4. The next step is to discover the volumes on the host by scanning the HBA, as described for QLogic and Emulex in 5.5, "Host configuration" on page 312.

For further details about creation of volumes, hosts, and host mappings, see *Implementing the IBM Storwize V7000*, SG24-7938.

5.4.8 Finding host port WWPNs on IBM Storwize V7000 nodes

To zone the HBA and Storage ports together, you need the WWPNs of the HBA ports and the host-ports on Storage.

To view the WWPN of the QLogic HBAs, see 5.5.1, “Finding WWPNs of QLogic FC HBAs” on page 313. To view the WWPN of the Emulex HBAs, see 5.5.3, “Finding WWPNs of Emulex FC HBAs” on page 321.

To view the host-port WWPNs on IBM Storwize V7000, perform the following steps:

1. To see the WWPN of the IBM Storwize V7000 host-ports. Go to the **Configuration** icon, in the left pane, and click **Network**, as shown in Figure 5-36.

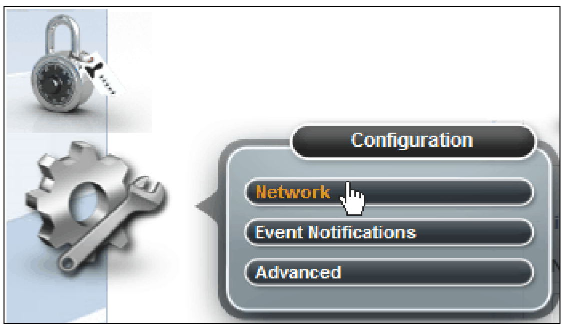


Figure 5-36 Network Configuration

2. The next view displays connectivity between nodes and other storage subsystems and the hosts that are attached through the FC network, as shown in Figure 5-37.

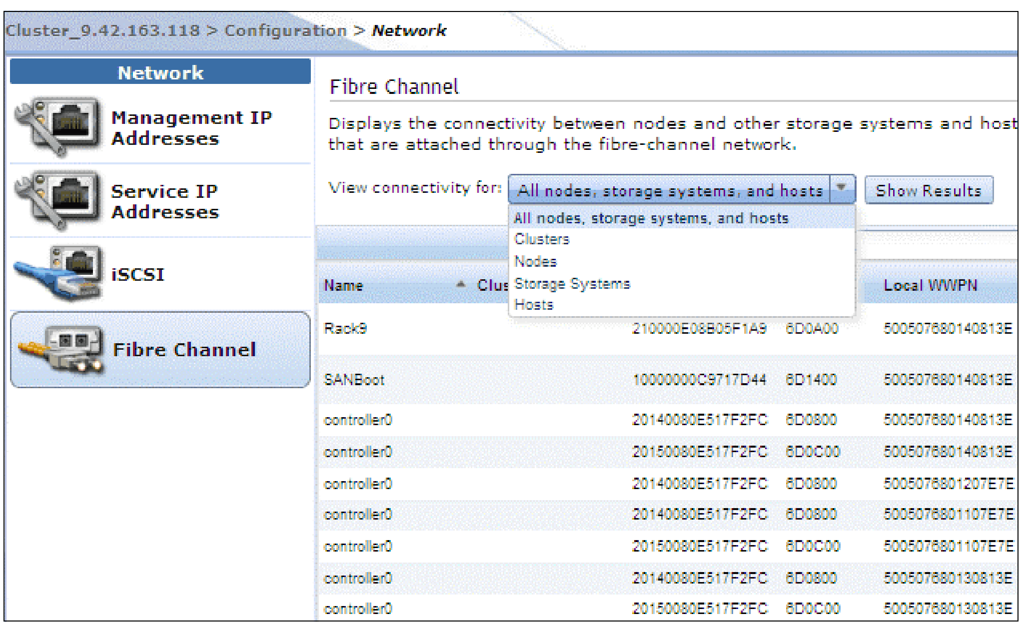
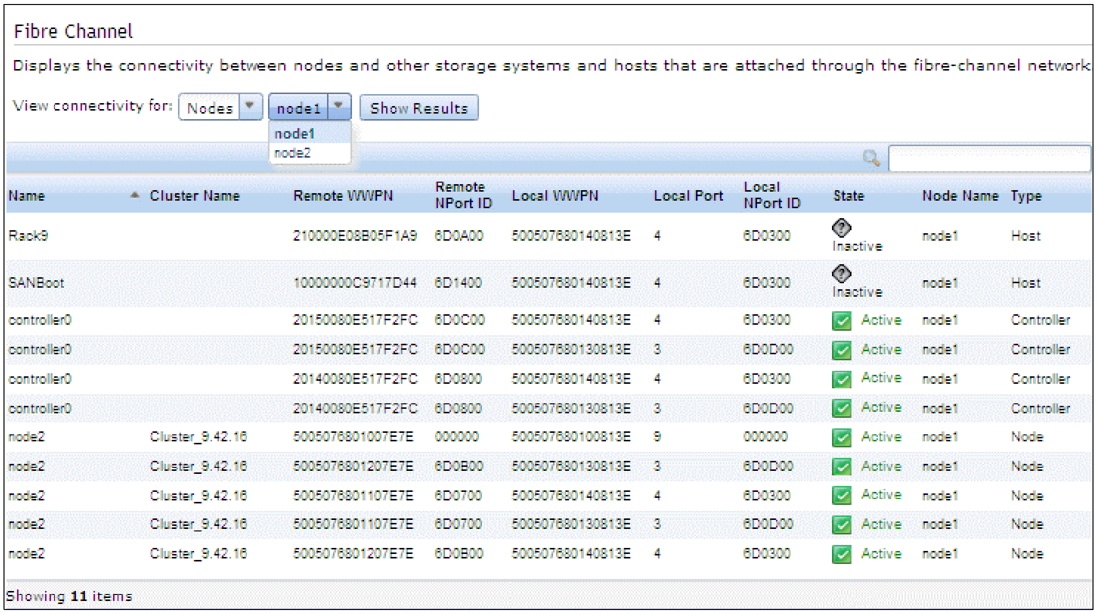


Figure 5-37 IBM Storwize V7000 FC Network - all WWPNs

3. Click the drop-down menu and select **Nodes**. In the adjacent drop-down menu, select the desired node, then click **Show Results**. All fibre connectivity to and from the node will be displayed, as shown in Figure 5-38.



Fibre Channel

Displays the connectivity between nodes and other storage systems and hosts that are attached through the fibre-channel network

View connectivity for: Nodes node1 Show Results

Name	Cluster Name	Remote WWPN	Remote NPort ID	Local WWPN	Local Port	Local NPort ID	State	Node Name	Type
Rack9		210000E08B05F1A9	6D0A00	500507680140813E	4	6D0300	Inactive	node1	Host
SANBoot		10000000C9717D44	6D1400	500507680140813E	4	6D0300	Inactive	node1	Host
controller0		20150080E517F2FC	6D0C00	500507680140813E	4	6D0300	Active	node1	Controller
controller0		20150080E517F2FC	6D0C00	500507680130813E	3	6D0D00	Active	node1	Controller
controller0		20140080E517F2FC	6D0800	500507680140813E	4	6D0300	Active	node1	Controller
controller0		20140080E517F2FC	6D0800	500507680130813E	3	6D0D00	Active	node1	Controller
node2	Cluster_9.42.16	500507680100 7E7E	000000	500507680100813E	9	000000	Active	node1	Node
node2	Cluster_9.42.16	500507680120 7E7E	6D0B00	500507680130813E	3	6D0D00	Active	node1	Node
node2	Cluster_9.42.16	500507680110 7E7E	6D0700	500507680140813E	4	6D0300	Active	node1	Node
node2	Cluster_9.42.16	500507680110 7E7E	6D0700	500507680130813E	3	6D0D00	Active	node1	Node
node2	Cluster_9.42.16	500507680120 7E7E	6D0B00	500507680140813E	4	6D0300	Active	node1	Node

Showing 11 items

Figure 5-38 Node connectivity

Verify that the connectivity between local (host-port) and remote (HBA) WWPN shows *Active*, as shown in the node's FC connectivity summary in Figure 5-38.

Alternatively, on some systems you can also view the node's host-port WWPN in the *Home* section of the GUI.

- In the Management GUI, select **Home** → **Manage Device**, as shown in Figure 5-39.

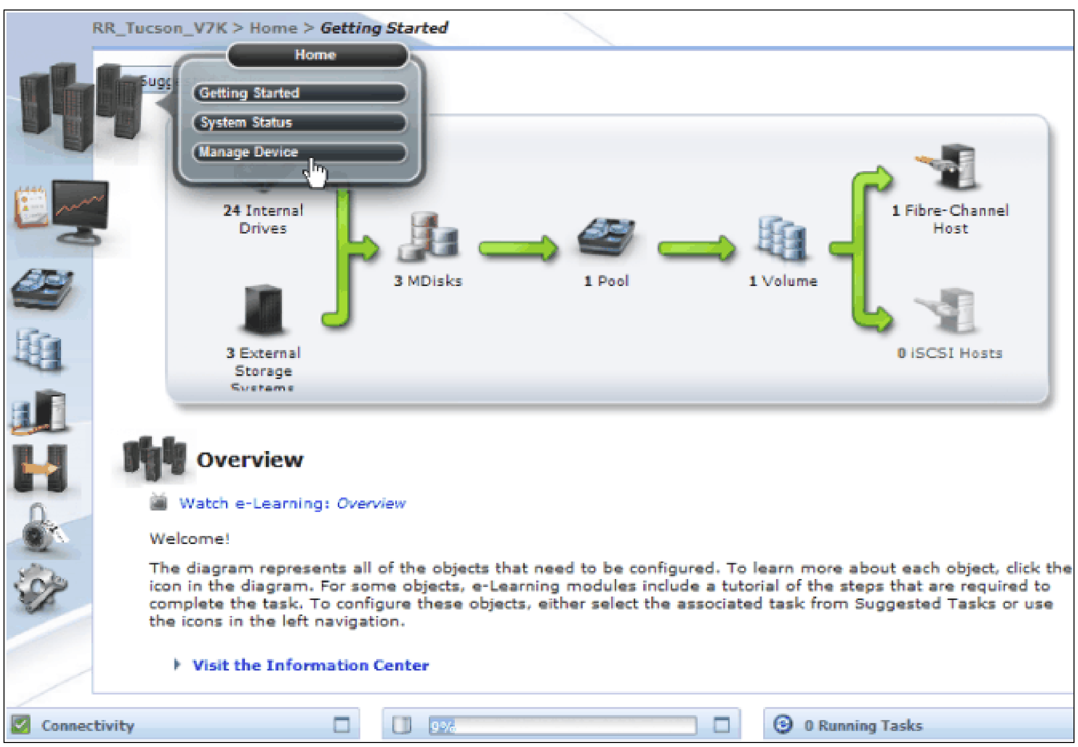


Figure 5-39 Manage Device

► Click the **Canister** to find out its associated WWPNs, as shown in Figure 5-40

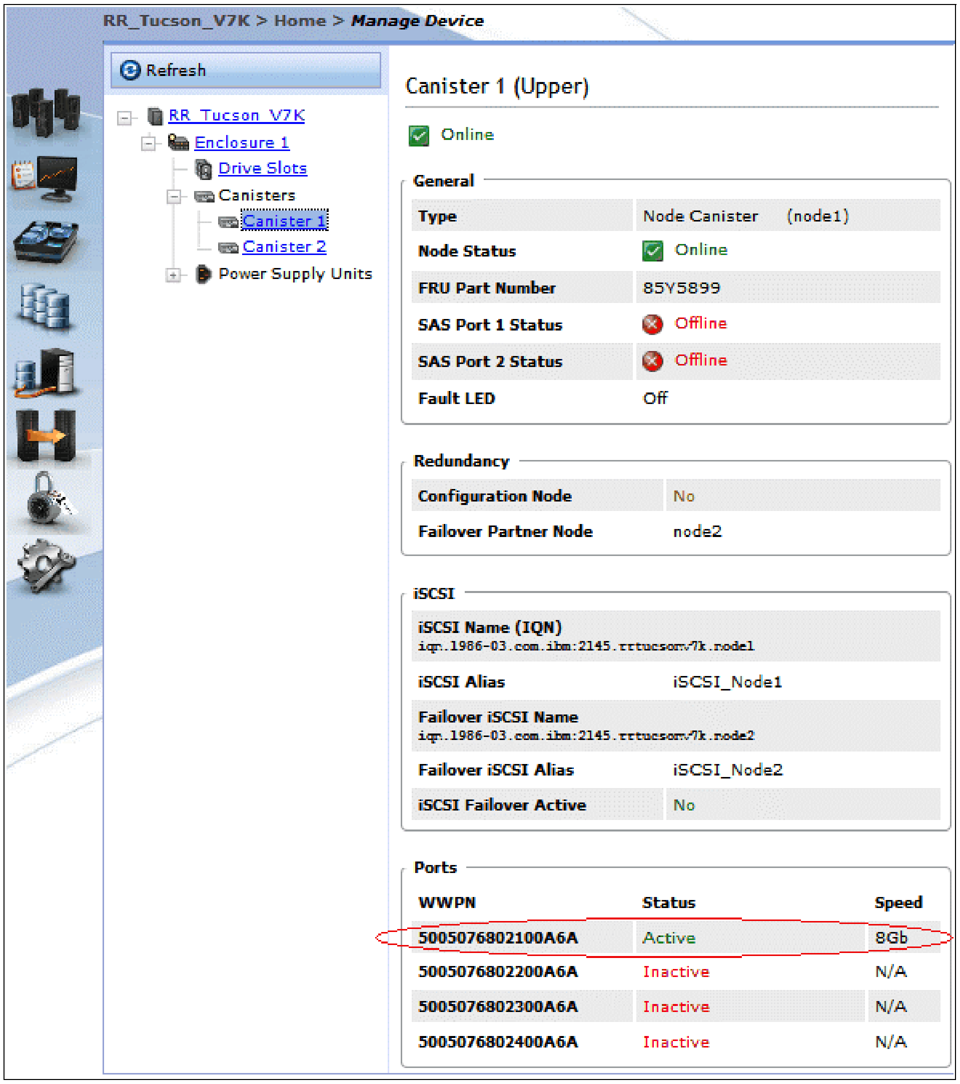


Figure 5-40 Canister WWPNs

5.5 Host configuration

In this section, we describe host configuration for SAN Boot implementation. For this purpose, we used IBM System x servers with QLogic and Emulex Fibre Channel host bus adapters (HBAs).

The first step in configuring a host for SAN Boot is to disable the local boot media, such as IDE, SAS, and SATA drives, in the Server's BIOS settings:

1. Power on the server and interrupt the boot sequence by pressing the **F1** key, to enter the system's BIOS Setup.
2. Select **Devices and I/O ports** from the main menu.
3. Select **IDE configuration** from the next menu. Disable the **Planar SAS**, as shown Figure 5-41.

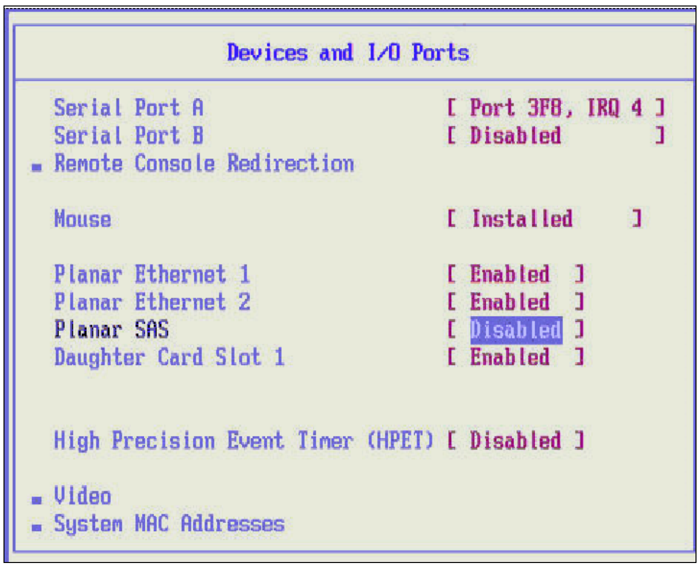


Figure 5-41 Disable Planar SAS

4. Press Esc to exit out of this menu and save the changes.
5. Next, use the HBA's configuration utilities for SAN Boot solutions. The software configuration parameters in the BIOS of a Fibre Channel (FC) HBA allows the server to identify and map the boot disk on the SAN.

The first step is to find the HBA's worldwide port name (WWPN) so it can be identified in the storage for mapping the boot volume to the host. In the following pages, you can find information regarding the use of QLogic and Emulex HBA configuration utilities for SAN Boot implementation.

5.5.1 Finding WWPNs of QLogic FC HBAs

Perform the following steps to obtain WWPNs of QLogic HBAs:

1. Boot the server after verifying that the QLogic HBA is installed properly.
2. Press **Ctrl-Q** when the QLogic HBA banner appears, as shown in Figure 5-42.

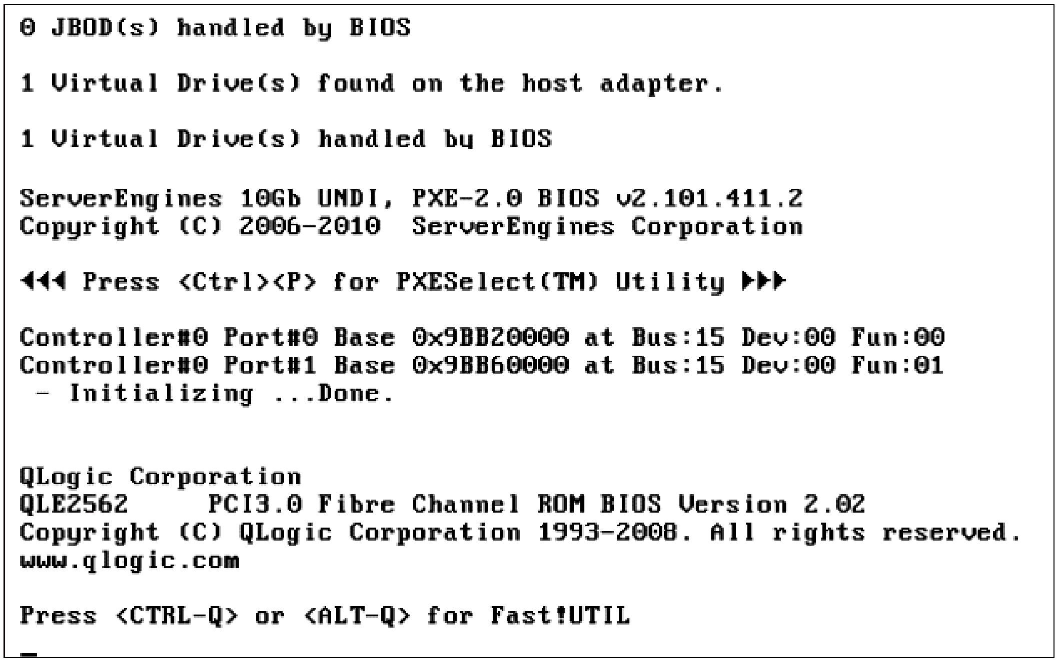


Figure 5-42 Enter QLogic Fast!UTIL

3. Select the adapter's port that will be zoned to storage, as shown in Figure 5-43.

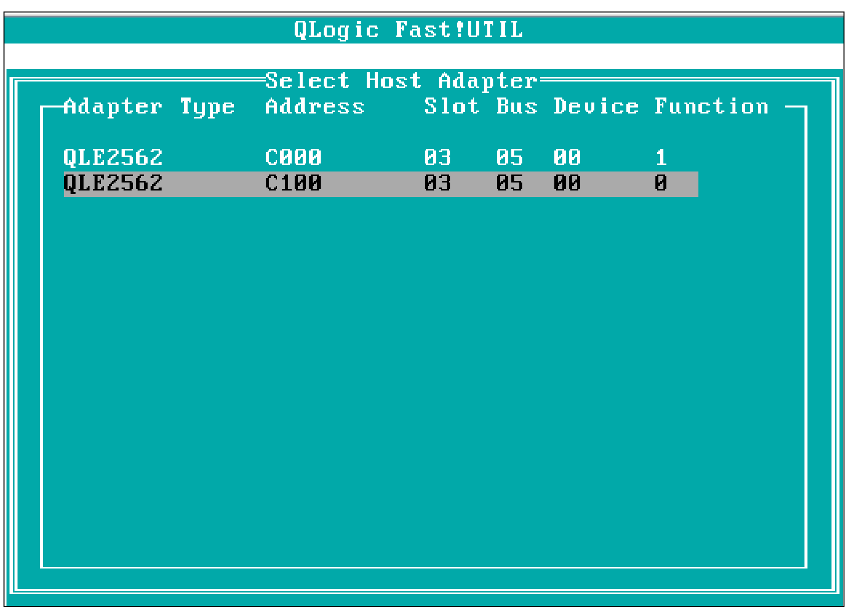


Figure 5-43 Select QLogic adapter

4. Select **Configuration Settings**, as shown in Figure 5-44.

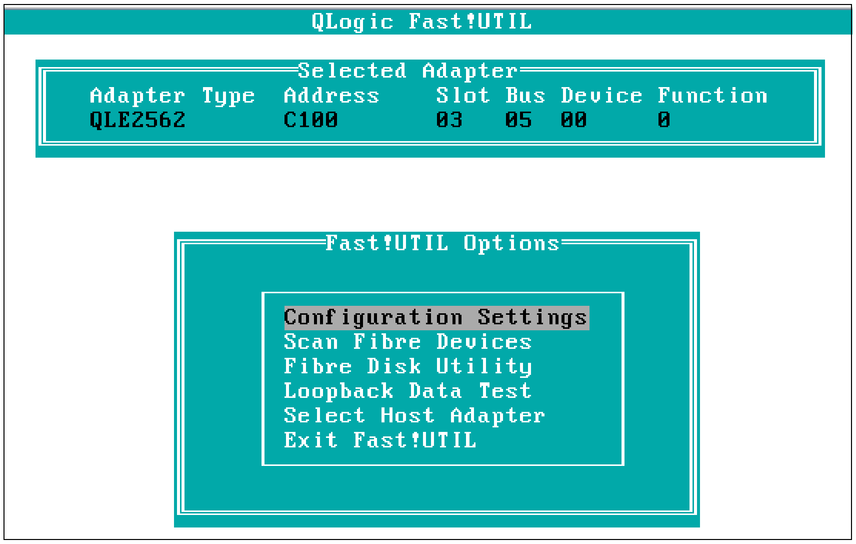


Figure 5-44 Configuration Settings

5. Select **Host Adapter Settings**, as shown in Figure 5-45.

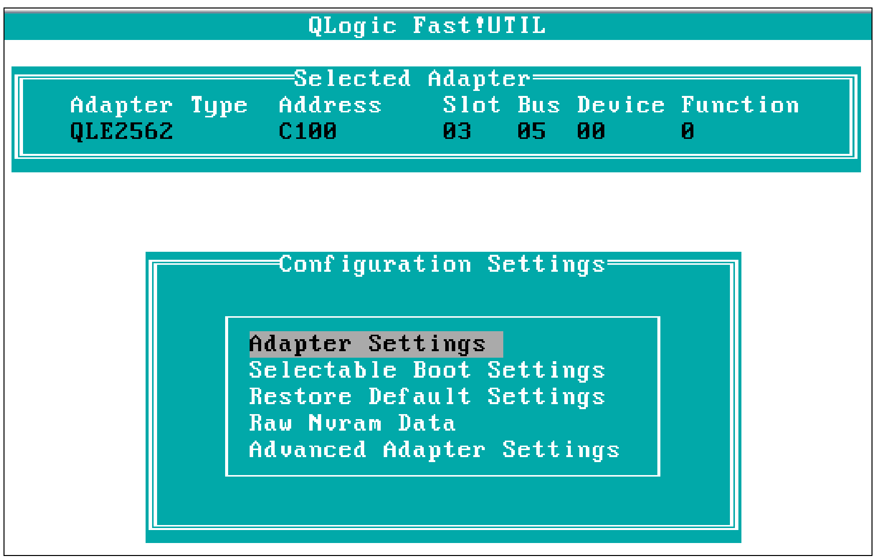


Figure 5-45 Adapter Settings

6. Make a note of the adapter's Worldwide Port Name (WWPN), as shown in Figure 5-46.

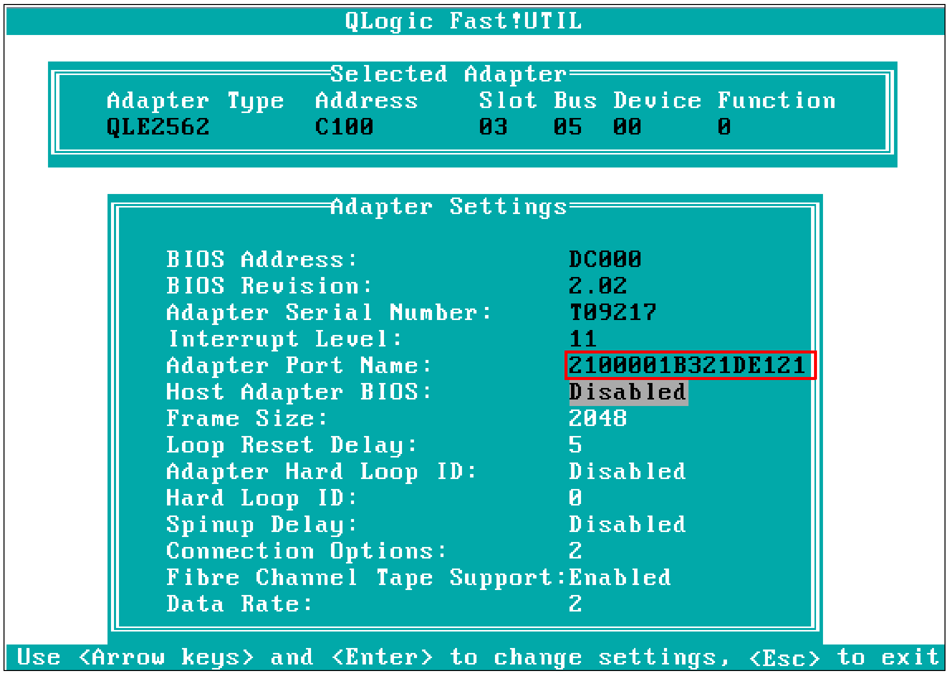


Figure 5-46 Note WWPN

In the *Zone Administration* utility of your **Switch**, use the HBA port's WWPN to Zone the HBA port to a host-port on Storage.

Tip: If you have created hosts (see 5.4.6, "Creating Fiber Channel hosts using the GUI" on page 302), you can find the WWPN of the host-port on IBM Storwize v7000 as described in 5.4.8, "Finding host port WWPNs on IBM Storwize V7000 nodes" on page 309.

5.5.2 Configuring a QLogic adapter for SAN Boot

To configuring the QLogic adapter for SAN Boot, perform the following setup:

1. Power up Server containing the HBA.
2. Follow steps **1** through **4** given in 5.5.1, "Finding WWPNs of QLogic FC HBAs" on page 313.
3. Select **Scan Fibre Devices** from the menu, as shown in Figure 5-44 on page 314.
4. You can expect to see the IBM Storwize V7000 host-port's WWPN, as shown in Figure 5-47.

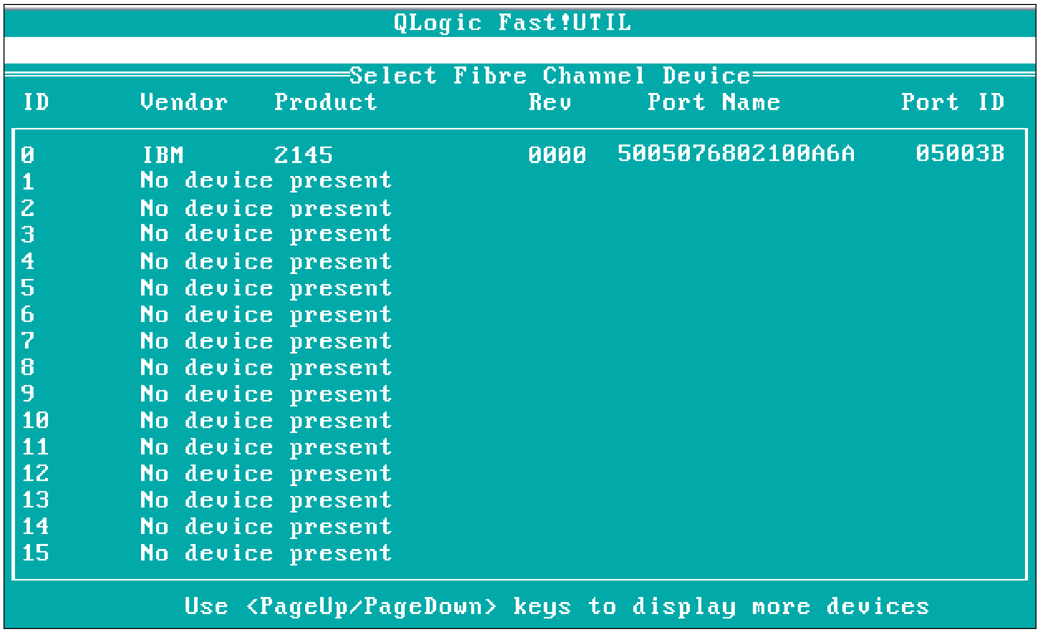


Figure 5-47 Scan Fibre Devices

> **Tip:** If you do not see any devices on scanning the fibre devices, verify your switch zoning and hardware connectivity, then rescan.

5. After the storage host-port's WWPN is found, press <Esc> to exit out to the *Fast!Util Options* page. We add the storage LUN in step 8 on page 317.

> **Tip:** Both SAN Volume Controller and IBM Storwize V7000 are seen as *Product 2145*. The difference is in WWPNs. IBM Storwize V7000 has 50:05:07:*68:02*:yx:xx:xx, while SAN Volume Controller has 50:05:07:*68:01*:yx:xx:xx.

6. Select **Configuration Settings**, then **Adapter Settings**, as shown in Figure 5-48.

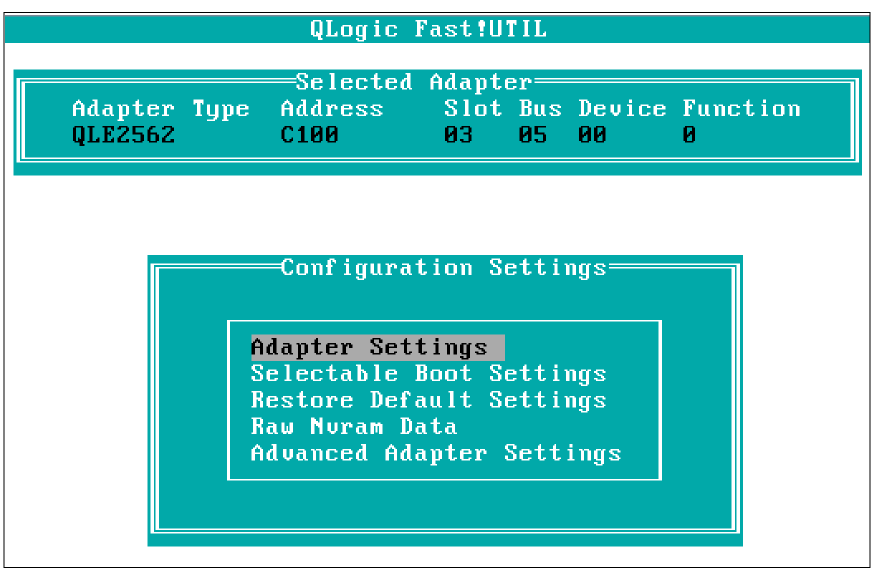


Figure 5-48 Adapter Settings

7. Select **Host Adapter BIOS** and enable it, as shown in Figure 5-49.

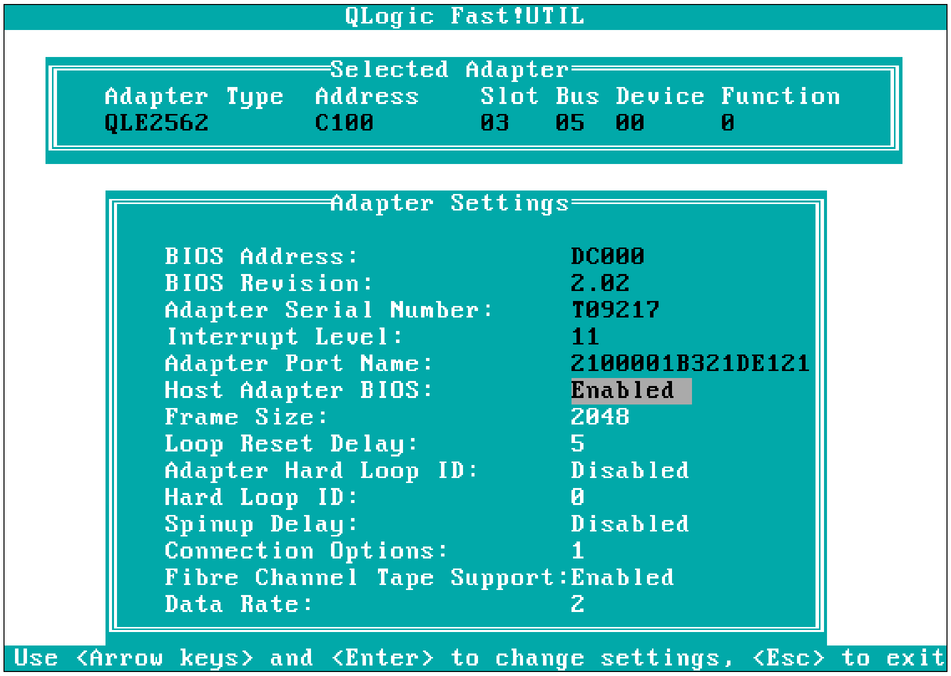


Figure 5-49 Enable BIOS

8. Press the <Esc> key to back up one menu. Next, select **Selectable Boot Settings**, as shown in Figure 5-50.

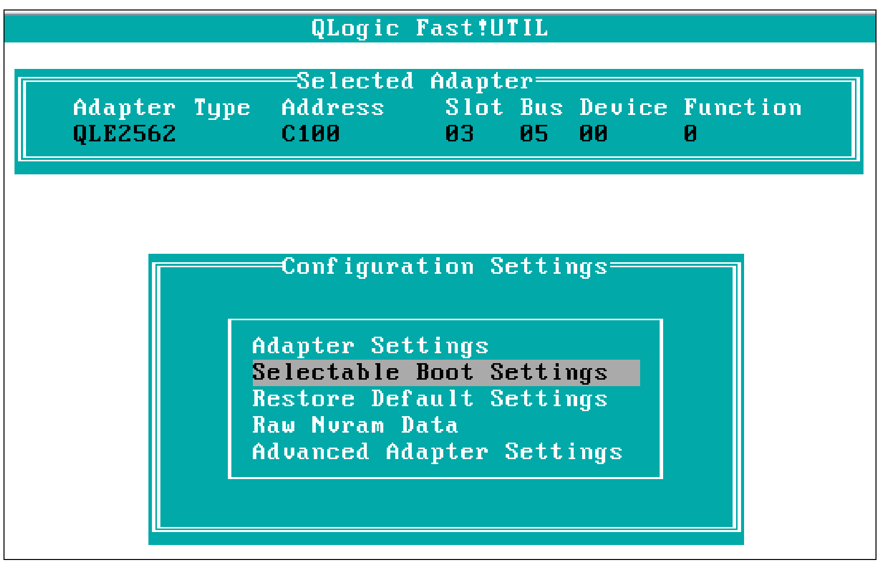


Figure 5-50 Selectable Boot Settings

9. Change *Selectable Boot Settings* to **Enabled**, as shown in Figure 5-51.

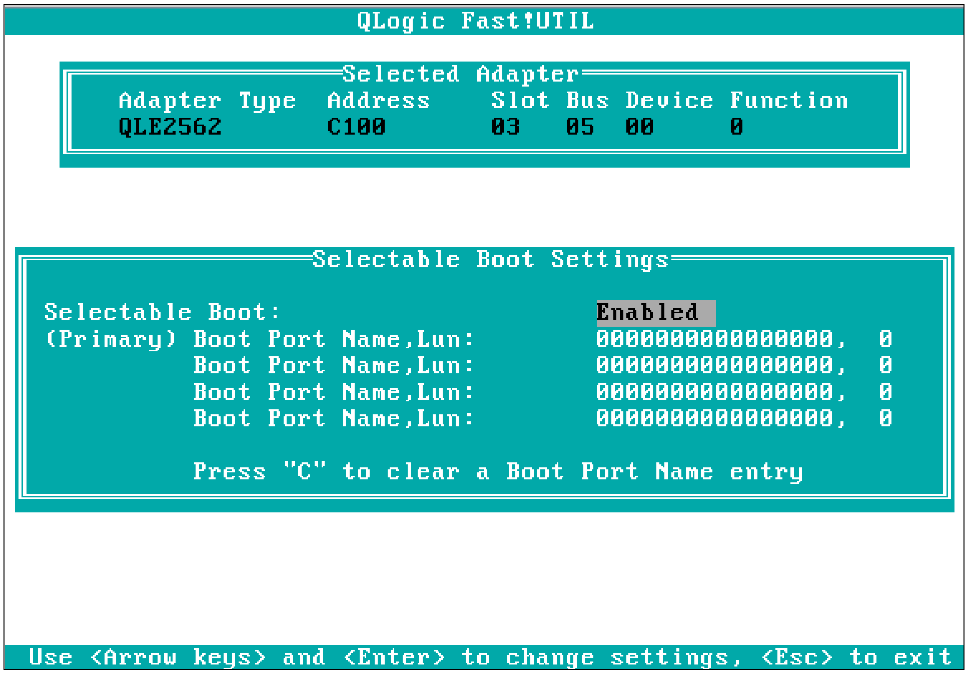


Figure 5-51 Enable Boot

10. Now click the **(Primary) Boot Port Name, Lun** to select the source LUN for SAN Boot, as shown in Figure 5-52.

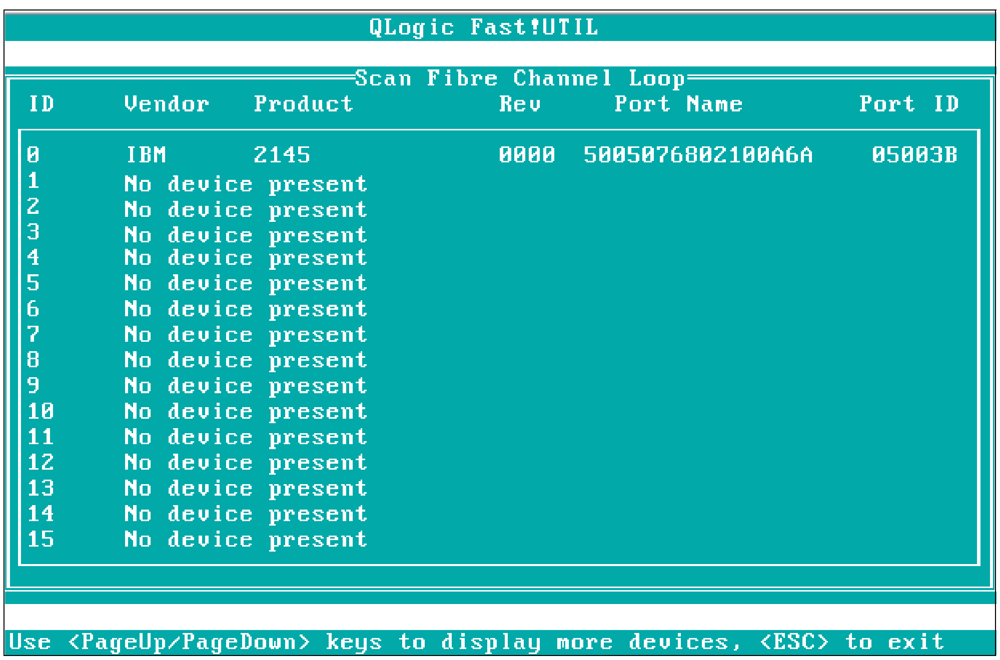


Figure 5-52 Select Boot Volume

Tip: IBM Storwize v7000 WWPNs are based on: 50:05:07:68:02:**Y**x:xx:xx, where x:xx:xx is unique for each node canister and the **Y** value is taken from the port position.

11. After selecting the appropriate storage Boot Volume, you come back to the previous panel, which now shows device information under *(Primary) Boot Port Name, LUN*, as shown in Figure 5-53.

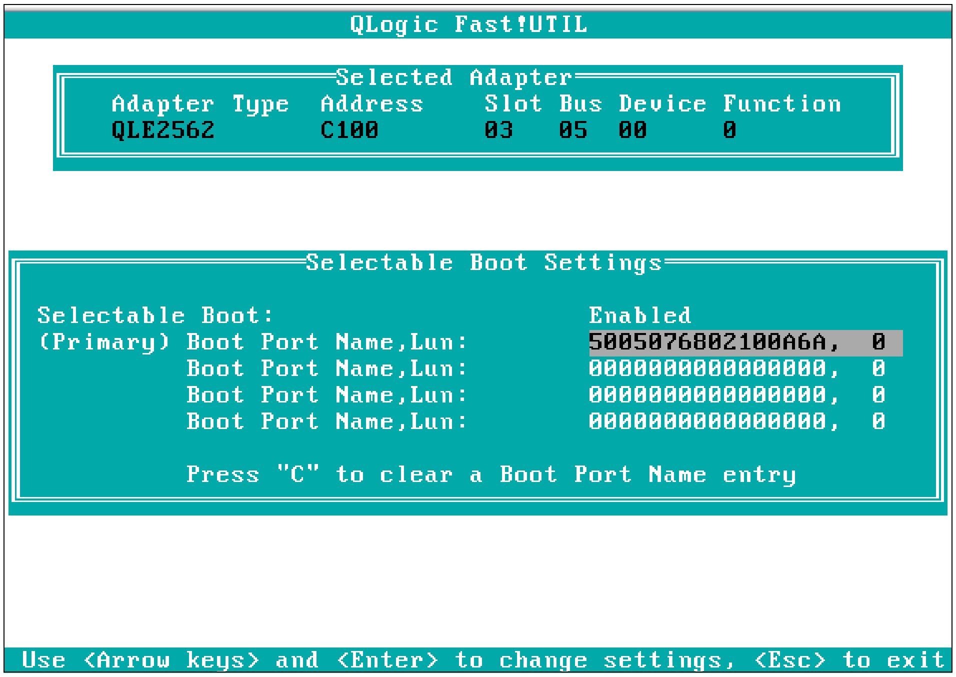


Figure 5-53 Boot LUN

12. Press the <Esc> key twice to exit out of the menus and select **Save Changes** to save the changes, as shown in Figure 5-54.

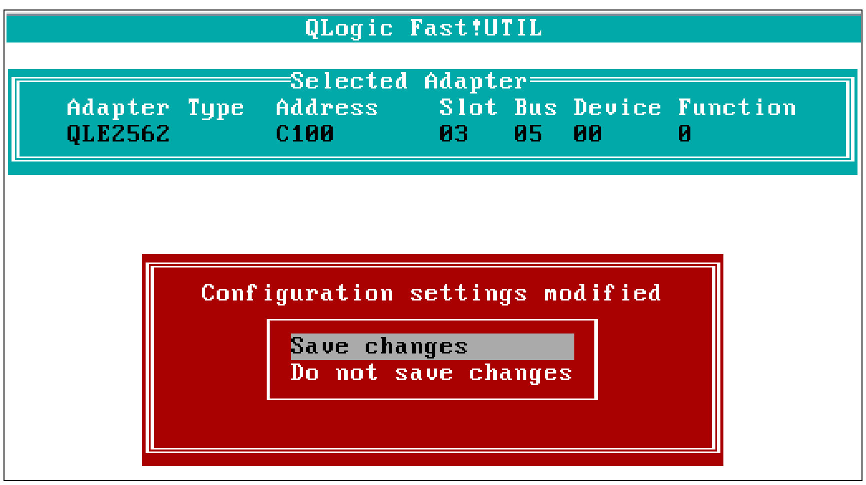


Figure 5-54 Save configuration changes

13. Exit **QLogic Fast!UTIL** and reboot server.

During POST, you can see QLogic reporting the Boot LUN, as shown in Figure 5-55.

```
0 Virtual Drive(s) handled by BIOS


!!! Emulex LightPulse x86 BIOS !!!, Version 2.02a1
Copyright (c) 1997-2008 Emulex. All rights reserved.

Press <Alt E> or <Ctrl E> to enter Emulex BIOS configuration
utility. Press <s> to skip Emulex BIOS
Emulex BIOS is Disabled on Adapter   1
Emulex BIOS is Disabled on Adapter   2

QLogic Corporation
QLE2562     PCI Fibre Channel ROM BIOS Version 2.02
Copyright (C) QLogic Corporation 1993-2008. All rights reserved.
www.qlogic.com

Press <CTRL-Q> or <ALT-Q> for Fast!UTIL
Firmware Version 4.03.01

Device Device  Adapter Port  Lun   Vendor    Product          Product
Number Type    Number   ID  Number   ID        ID             Revision
       Disk      0    05003B  0     IBM      2145              0000
ROM BIOS Installed
```

Figure 5-55   QLogic reporting boot LUN from Storage

14. Press F12 to select the boot device, then select the CD/DVD ROM (or appropriate source) to install the operating system, as shown in Figure 5-56 and Figure 5-57.

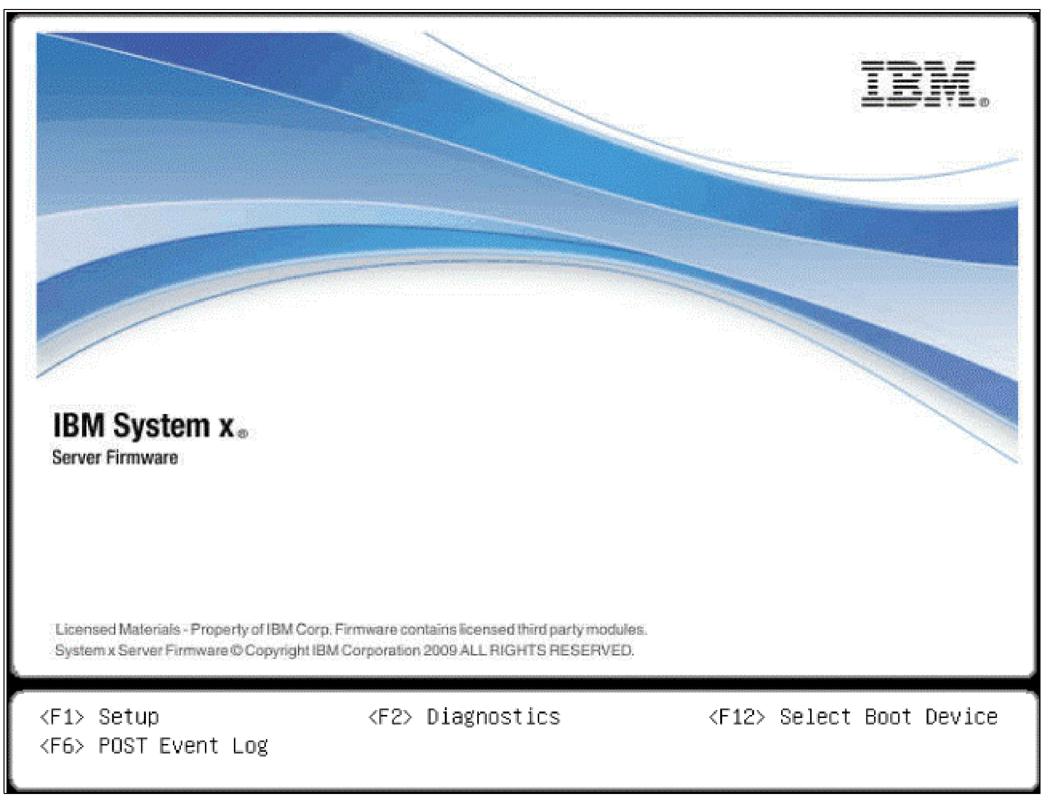


Figure 5-56   F12 to select Boot Device

```
Please select boot device:
--------------------------------
MBA v5.2.2  Slot 0100
LEXAR JD FIREFLY 1100
MATSHITADVD-RAM UJ890 SAA4
UEFI: MATSHITADVD-RAM UJ890 SAA4
Enter Setup
--------------------------------
↑ and ↓ to move selection
ENTER to select boot device
ESC to boot using defaults
```

Figure 5-57 Select DVD to install OS

15. (Assuming this installation is from DVD), insert the media and follow the installation instructions, detailed in 5.6, “Operating systems configuration” on page 329.

5.5.3 Finding WWPNs of Emulex FC HBAs

For Emulex adapters, perform the following steps:

1. Press **Alt E** to go to the Emulex BIOS Utilities, as shown in Figure 5-58.

```
Broadcom NetXtreme II Ethernet Boot Agent v5.2.2
Copyright (C) 2000-2009 Broadcom Corporation
All rights reserved.

LSI MegaRAID SAS-MFI BIOS
Version 4.19.00 (Build October 19, 2010)
Copyright(c) 2010 LSI Corporation
HA -0 (Bus 3 Dev 0) ServeRAID M1015 SAS/SATA Controller
FW package: 20.10.1-0022


0 JBOD(s) found on the host adapter
0 JBOD(s) handled by BIOS

0 Virtual Drive(s) found on the host adapter.

0 Virtual Drive(s) handled by BIOS


!!! Emulex LightPulse x86 BIOS !!!, Version 2.02a1
Copyright (c) 1997-2008 Emulex. All rights reserved.

Press <Alt E> or <Ctrl E> to enter Emulex BIOS configuration
utility. Press <s> to skip Emulex BIOS
```

Figure 5-58 Alt + E for Emulex BIOS Configuration Utility

2. Select the adapter that you want from the list, as shown in Figure 5-59.

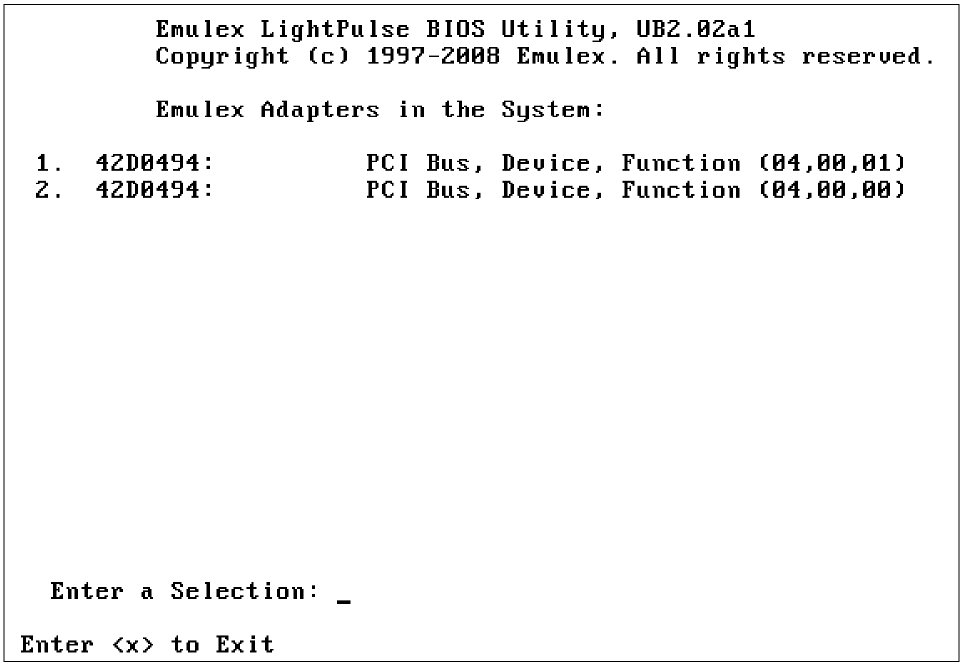


Figure 5-59 Select HBA Port

3. Select **Configure This Adapter's Parameters**, as shown in Figure 5-60.

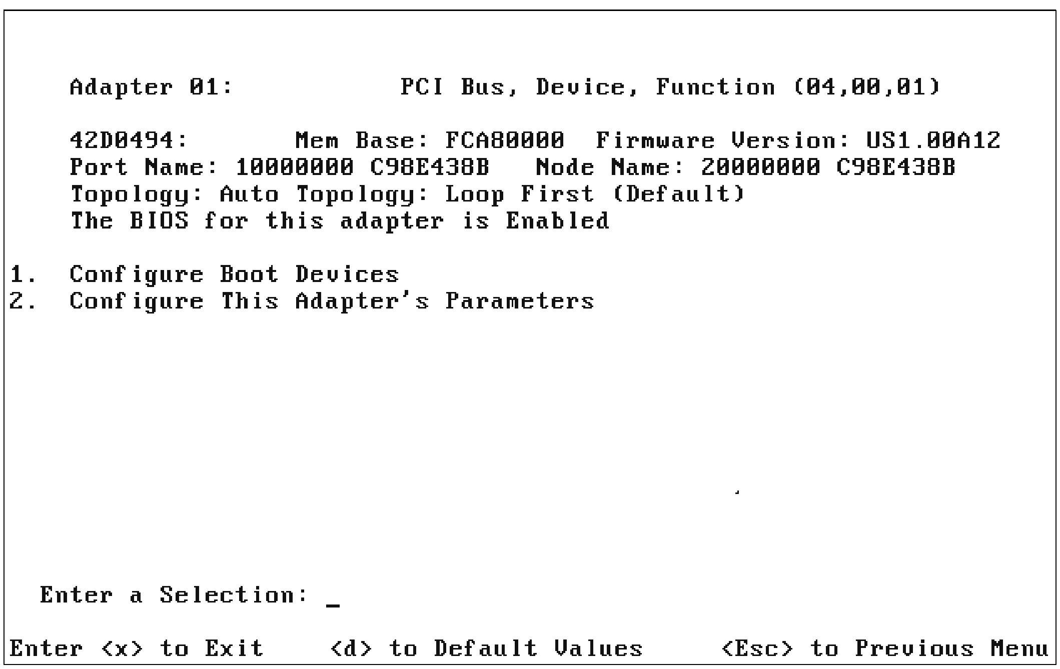


Figure 5-60 Adapter Settings

4. The Adapter's WWPN can be seen near the top of the panel, in Figure 5-61.

```
    Adapter 01:           PCI Bus, Device, Function (04,00,01)

    42D0494:       Mem Base: FCA80000  Firmware Version: US1.00A12
    Port Name: 10000000 C98E438B  Node Name: 20000000 C98E438B
    Topology: Auto Topology: Loop First (Default)
    The BIOS for this adapter is Enabled

1.  Enable or Disable BIOS
2.  Change Default ALPA of this Adapter
3.  Change PLOGI Retry Timer (+Advanced Option+)
4.  Topology Selection (+Advanced Option+)
5.  Enable or Disable Spinup Delay (+Advanced Option+)
6.  Auto Scan Setting (+Advanced Option+)
7.  Enable or Disable EDD 3.0 (+Advanced Option+)
8.  Enable or Disable Start Unit Command (+Advanced Option+)
9.  Enable or Disable Environment Variable (+Advanced Option+)
10. Enable or Disable Auto Boot Sector (+Advanced Option+)
11. Link Speed Selection (+Advanced Option+)

  Enter a Selection:

Enter <x> to Exit          <Esc> to Previous Menu
```

Figure 5-61 Emulex parameters

5. Make a note of the WWPN of the HBA's port.

To obtain WWPN of the alternate HBA port, repeat steps 2 on page 322 to 5 on page 323.

5.5.4 Configuring an Emulex adapter for SAN Boot

To configure the Emulex adapter for SAN Boot, perform the following steps:

1. Boot the server after verifying that the Emulex HBA is installed properly.
2. Follow steps **1** through *4* from 5.5.3, “Finding WWPNs of Emulex FC HBAs” on page 321.
3. Select option **1** to *Enable BIOS* for this HBA, as shown in Figure 5-62.

```
    Adapter 01:           PCI Bus, Device, Function (04,00,01)

    42D0494:       Mem Base: FCA80000  Firmware Version: US1.00A12
    Port Name: 10000000 C98E438B   Node Name: 20000000 C98E438B
    Topology: Auto Topology: Loop First (Default)
    The BIOS for this adapter is Enabled

1.  Enable or Disable BIOS
2.  Change Default ALPA of this Adapter
3.  Change PLOGI Retry Timer (+Advanced Option+)
4.  Topology Selection (+Advanced Option+)
5.  Enable or Disable Spinup Delay (+Advanced Option+)
6.  Auto Scan Setting (+Advanced Option+)
7.  Enable or Disable EDD 3.0 (+Advanced Option+)
8.  Enable or Disable Start Unit Command (+Advanced Option+)
9.  Enable or Disable Environment Variable (+Advanced Option+)
10. Enable or Disable Auto Boot Sector (+Advanced Option+)
11. Link Speed Selection (+Advanced Option+)

  Enter a Selection:

Enter <x> to Exit          <Esc> to Previous Menu
```

Figure 5-62 Enable BIOS

4. You can receive acknowledgement that the BIOS is enabled, as shown in Figure 5-63.

```
    Adapter 01:           PCI Bus, Device, Function (04,00,01)


    The BIOS is Enabled!!


    Enable Press 1, Disable Press 2:









Enter <x> to Exit          <Esc> to Previous Menu
```

Figure 5-63 BIOS enabled

5. Press <Esc> twice to exit back to the adapter's main menu. Now select option **1** to **Configure Boot Devices**, as shown in Figure 5-64.

```
    Adapter 01:           PCI Bus, Device, Function (04,00,01)

    42D0494:       Mem Base: FCA80000  Firmware Version: US1.00A12
    Port Name: 10000000 C98E438B   Node Name: 20000000 C98E438B
    Topology: Auto Topology: Loop First (Default)
    The BIOS for this adapter is Enabled

1.  Configure Boot Devices
2.  Configure This Adapter's Parameters









  Enter a Selection: _

Enter <x> to Exit    <d> to Default Values     <Esc> to Previous Menu
```

Figure 5-64 Adapter (HBA Port) settings

6. Enter **Boot Entry number** of the first *Unused boot device (Primary Boot)* from the *List of Saved Boot Devices*, as shown in Figure 5-65.

```
    Adapter 01:  S_ID: 05003A    PCI Bus, Device, Function (04,00,01)


    List of Saved Boot Devices:

1. Unused   DID:000000 WWPN:00000000 00000000 LUN:00  Primary Boot
2. Unused   DID:000000 WWPN:00000000 00000000 LUN:00
3. Unused   DID:000000 WWPN:00000000 00000000 LUN:00
4. Unused   DID:000000 WWPN:00000000 00000000 LUN:00
5. Unused   DID:000000 WWPN:00000000 00000000 LUN:00
6. Unused   DID:000000 WWPN:00000000 00000000 LUN:00
7. Unused   DID:000000 WWPN:00000000 00000000 LUN:00
8. Unused   DID:000000 WWPN:00000000 00000000 LUN:00






    Select a Boot Entry: _

Enter <x> to Exit          <Esc> to Previous Menu
```

Figure 5-65 Saved Boot Devices

7. The HBA will scan for Fibre Channel devices. If you added this HBA port's WWPN to IBM Storwize V7000's host mappings (5.4.7, "Mapping SAN Boot volume to the host" on page 306, the list will be populated with device information from Storwize (host-port WWPNs), as shown in Figure 5-66.

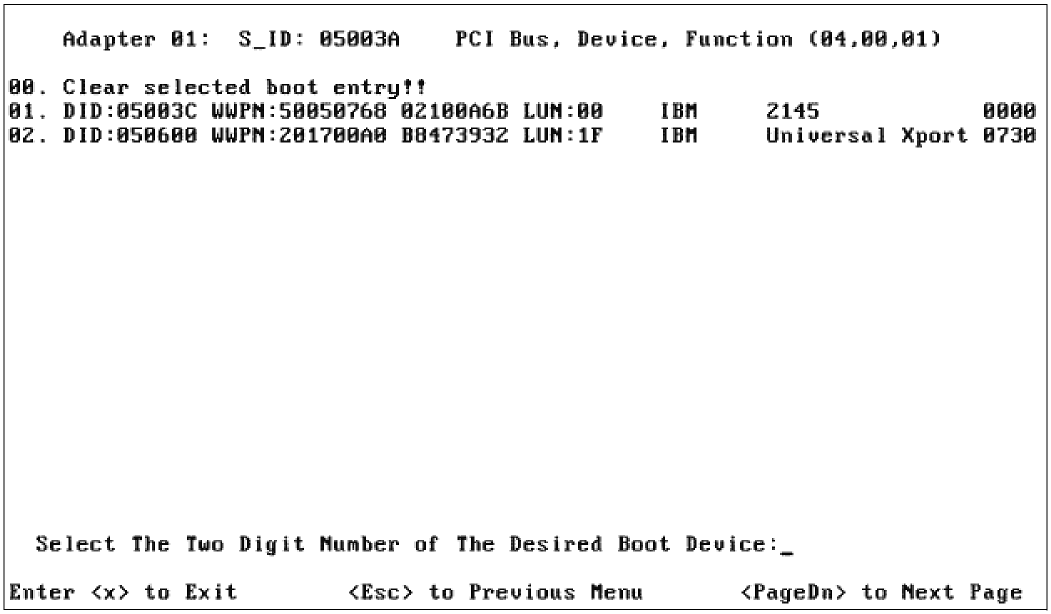


Figure 5-66 Scan Fibre Devices

Tip: Both SAN Volume Controller and IBM Storwize V7000 are seen as Product 2145. The difference is in WWPNs. IBM Storwize V7000 has 50:05:07:68:02:yx:xx:xx, while SAN Volume Controller has 50:05:07:68:01:yx:xx:xx

8. Type **01** (in our case) to **Select The Two Digit Number Of The Desired Boot Device**. Next, it prompts for the **Two Digits of the Starting LUN**, which is **00**. The next prompt is to select **Boot this device via WWPN**, or **Boot this device via DID**. See Figure 5-67.

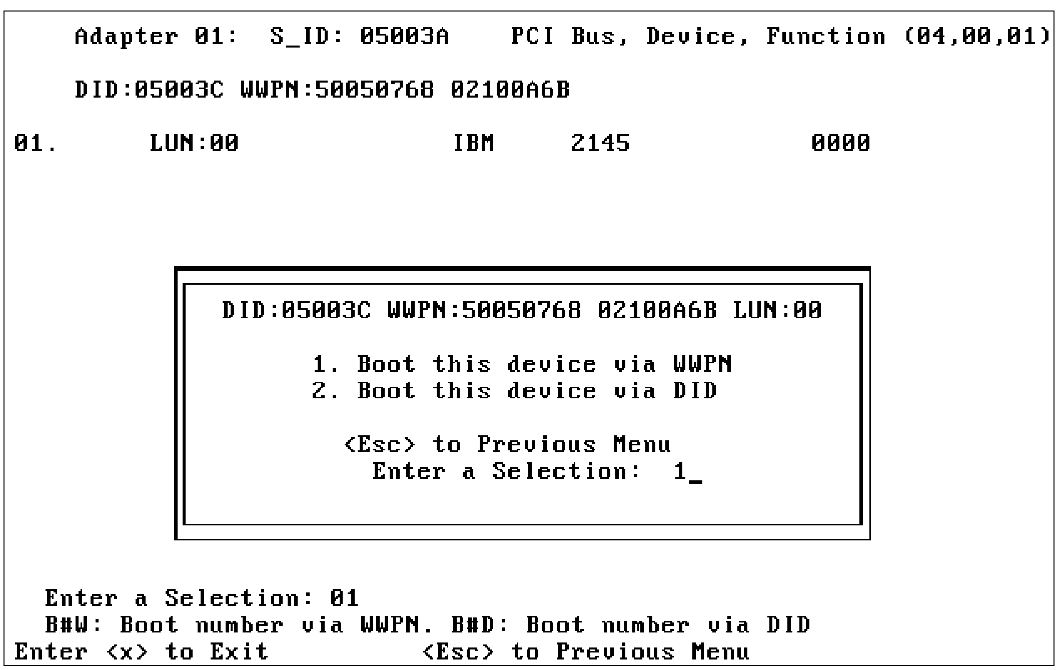


Figure 5-67 Enter digits of starting LUN

9. Enter the *Boot number* to boot from (by WWPN), as shown in Figure 5-68.

```
    Adapter 01:  S_ID: 05003A    PCI Bus, Device, Function (04,00,01)

    DID:05003C WWPN:50050768 02100A6B

01.      LUN:00              IBM     2145            0000









  Enter a Selection:
  B#W: Boot number via WWPN. B#D: Boot number via DID
Enter <x> to Exit          <Esc> to Previous Menu
```

Figure 5-68 Boot by WWPN

10. Now the LUN, on IBM Storwize V7000, shows up in the **List of Saved Boot Devices**, as shown in Figure 5-69.

```
    Adapter 01:  S_ID: 05003A    PCI Bus, Device, Function (04,00,01)


    List of Saved Boot Devices:

1. Used      DID:000000 WWPN:50050768 02100A6B LUN:00  Primary Boot
2. Unused    DID:000000 WWPN:00000000 00000000 LUN:00
3. Unused    DID:000000 WWPN:00000000 00000000 LUN:00
4. Unused    DID:000000 WWPN:00000000 00000000 LUN:00
5. Unused    DID:000000 WWPN:00000000 00000000 LUN:00
6. Unused    DID:000000 WWPN:00000000 00000000 LUN:00
7. Unused    DID:000000 WWPN:00000000 00000000 LUN:00
8. Unused    DID:000000 WWPN:00000000 00000000 LUN:00




    Select a Boot Entry: _


Enter <x> to Exit          <Esc> to Previous Menu
```

Figure 5-69 Boot LUN's WWPN

11. Enter <x> to exit and save changes, which will reboot the server.

12. During post, after Emulex banner, you will see the IBM Storwize V7000 being seen by the HBA, as shown in Figure 5-70.

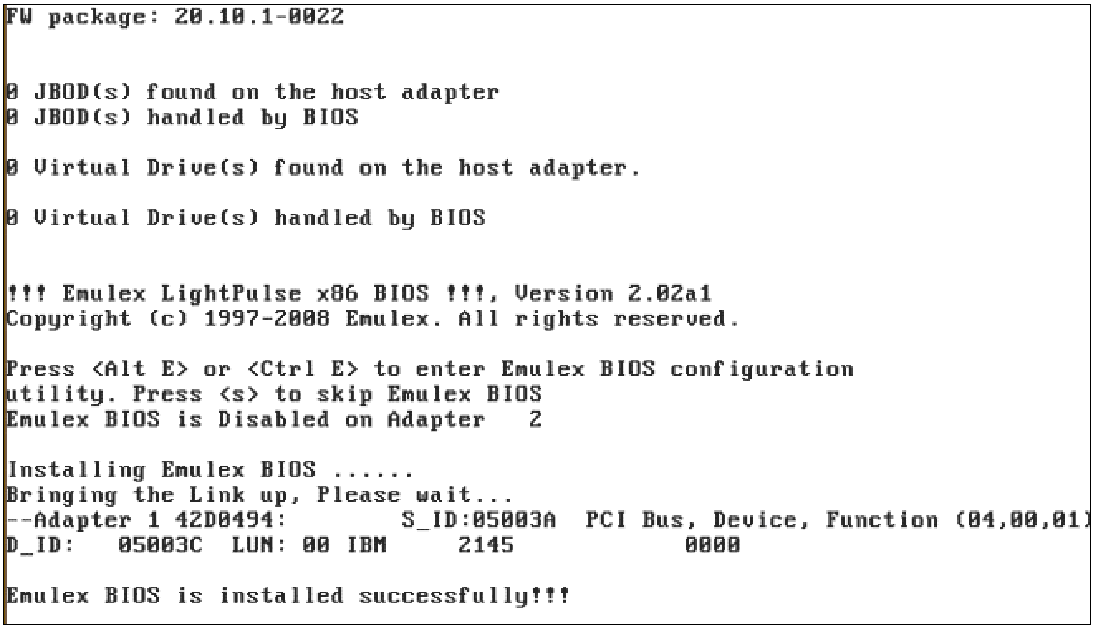


Figure 5-70 Emulex seeing SAN Boot volume

13. Press F12 to select boot device, then select the CD/DVD ROM (or appropriate source) to install the operating system, as shown in Figure 5-71 and Figure 5-72.

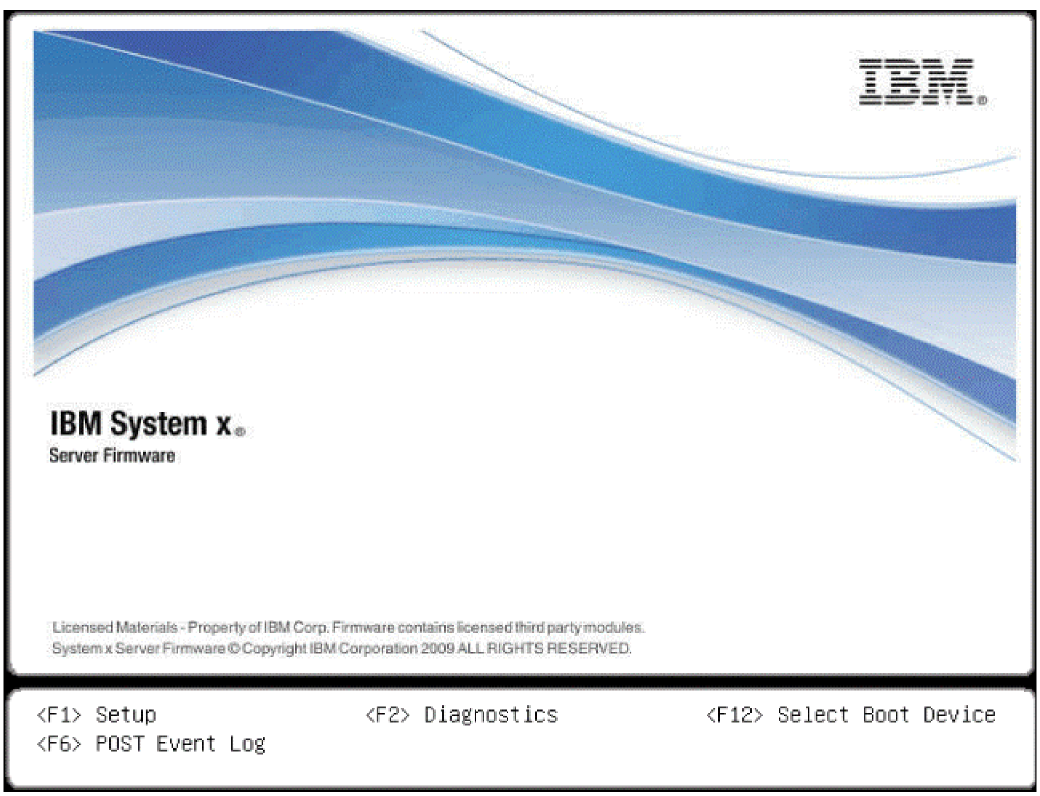


Figure 5-71 F12 to select Boot Device

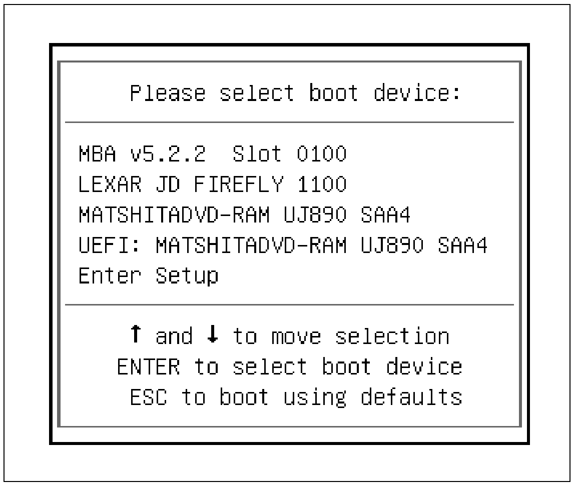


Figure 5-72 Select DVD to install OS

14. Assuming this installation is from DVD, insert the media and follow the installation instructions in section 5.6, “Operating systems configuration” on page 329.

5.6 Operating systems configuration

IBM Storwize v7000 supports SAN Boot for Windows, VMware, Linux, AIX, and many other operating systems. SAN Boot support can change from time to time, therefore we strongly recommend checking the IBM Storwize v7000 interoperability matrix before setting up SAN Boot, available at this website:

http://www-03.IBM.com/systems/storage/disk/storwize_v7000/interop.html

Support for multipathing software with SAN Boot is also updated at the foregoing interop link.

In the following sub-sections, we give step-by-step instructions for SAN Boot implementation using Microsoft Windows 2008 R2, VMware ESXi 4.1, SUSE Linux Enterprise Server 11 SP1, Red Hat Enterprise Linux 5.5 (RHEL), and IBM AIX version 6.1 operating systems.

5.7 SAN Boot for Windows

Microsoft Windows operating systems have been using SAN Boot implementations since Windows NT, due to its significant advantages in terms of the Redundancy, Availability, and Serviceability of fabric connectivity.

> **Microsoft support for booting from a SAN:** See the article at the following website for information about Microsoft support for booting from a SAN:
>
> http://support.microsoft.com/kb/305547

In the following pages, we explain the steps to implement SAN Boot using **Microsoft Windows Server 2008 R2 with SP1** on a boot volume residing at an IBM Storwize V7000 storage system.

Tip: Before setting up SAN Boot for Windows, review the information in 5.3.1, "Considerations before setting up SAN Boot" on page 286.

Following are the steps to set up a host (server) to boot from SAN using Microsoft Windows 2008 R2 operating system:

1. Configure the IBM Storwize V7000 so that only the boot volume is mapped to the host.
2. Configure the Fibre Channel SAN (switch) so that the host only sees one IBM Storwize V7000 system node port. It is done to prevent issues that can arise after rebooting the server, when Windows encounters duplicate drive devices that are not managed by multipathing software

Support: Multiple paths during installation are *not supported* with Microsoft Windows 2008 operating systems.

3. Configure and enable the HBA BIOS, as shown in 5.5.2, "Configuring a QLogic adapter for SAN Boot" on page 315 and/or 5.5.4, "Configuring an Emulex adapter for SAN Boot" on page 324.
4. Install the operating system selecting the volume configured for SAN Boot (see section 5.4.5, "Creating a volume for SAN Boot" on page 298) as the partition on which to install.
5. Follow the instructions by Microsoft to install Windows 2008 R2 with SP1, until it reaches the dialog box with the highlighted question, **Where do you want to install Windows?**, as shown in Figure 5-73.

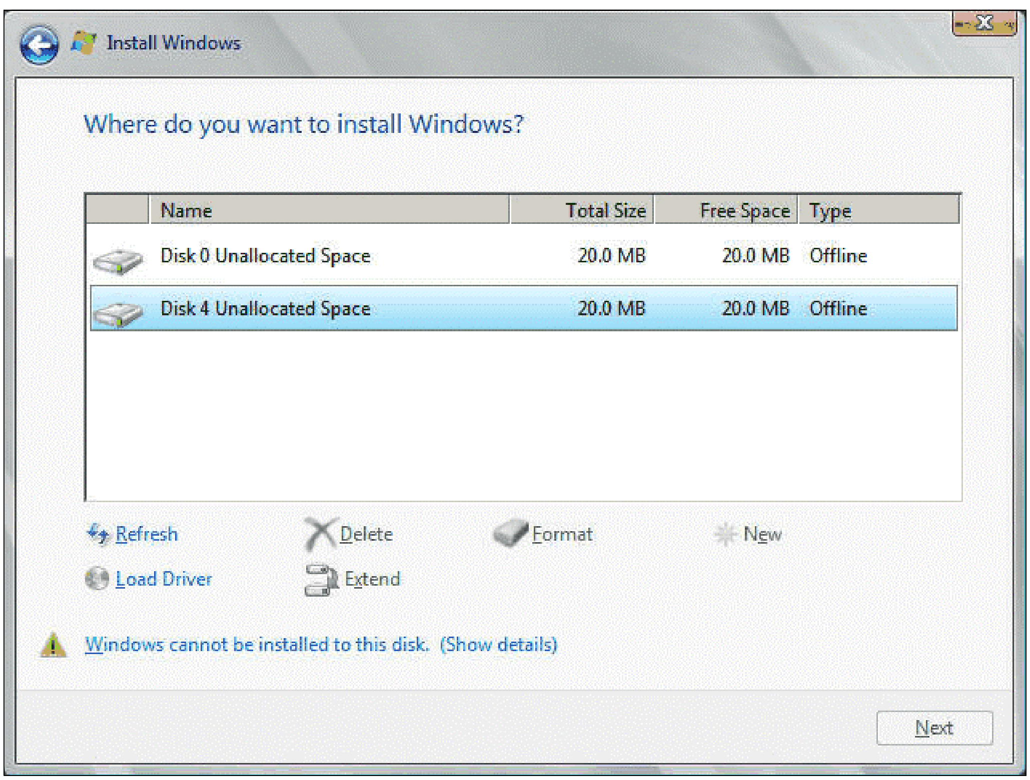


Figure 5-73 Without SAN Volume

6. If you do not see the Volume (by size) in the list, it means that Windows did not have the drivers for the HBA. You need to add it manually by clicking the **Load Driver** button at the bottom-left of the window, as shown in Figure 5-73. The resulting load driver window is shown in Figure 5-74.

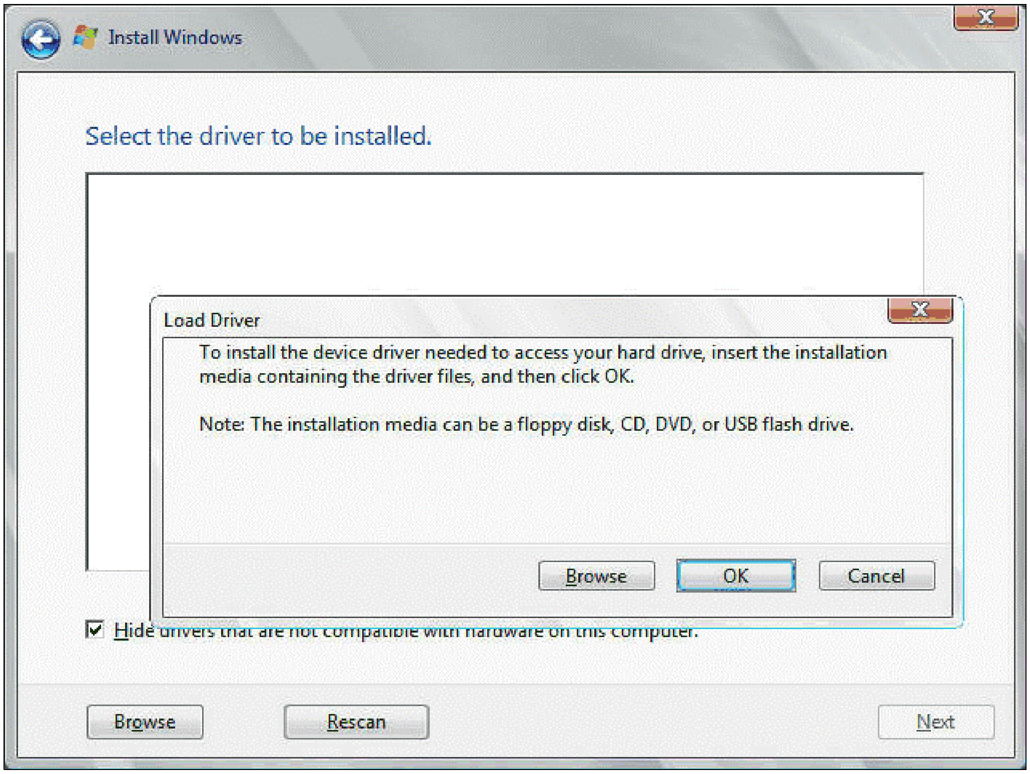


Figure 5-74 Loading HBA driver

7. Browse to the source (media) of the HBA driver package. Notice the check box that would hide the drivers which it deems not compatible, as shown in Figure 5-75.

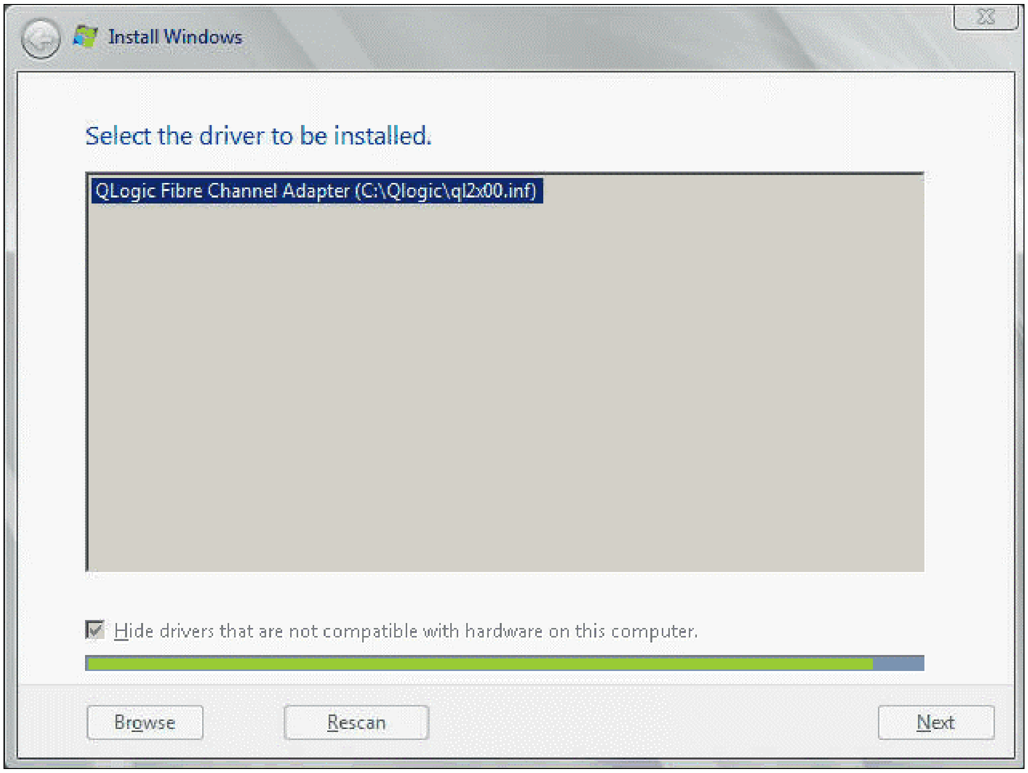


Figure 5-75 Install QLogic Driver

8. Verify that the boot volume allocated on storage is visible by the HBA (5.5, “Host configuration” on page 312). The disk information will be updated after Rescan, as shown in Figure 5-76.

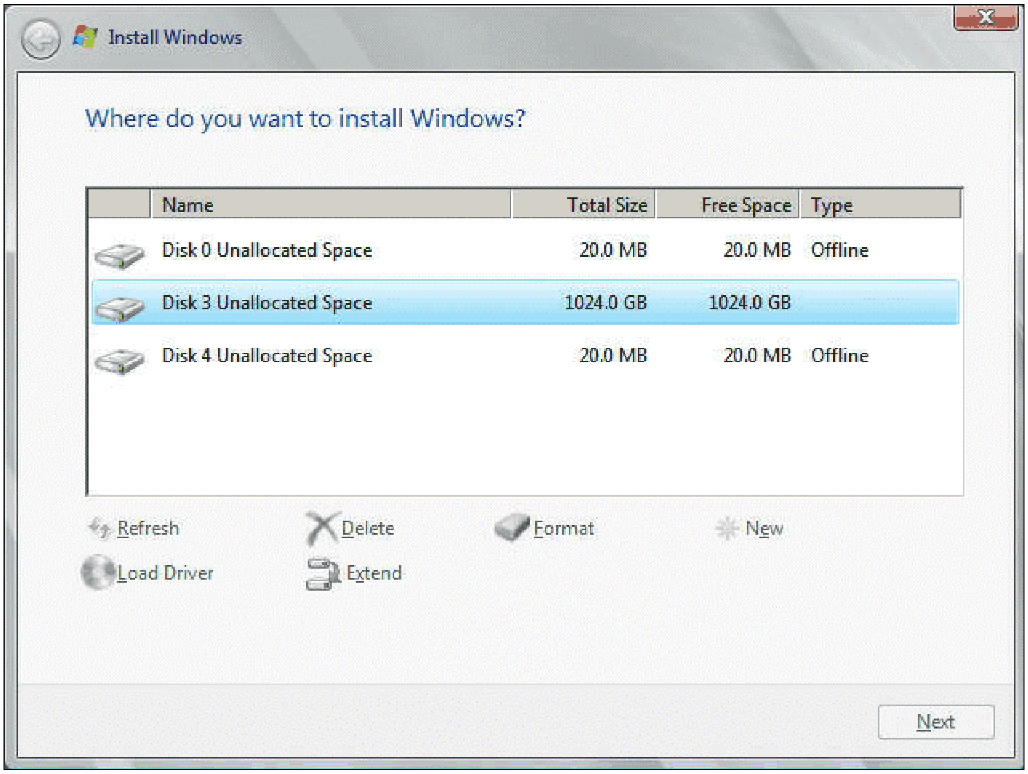


Figure 5-76 Boot Volume on Storage

9. Create **New** partition. Next, **Format** the drive and continue with the installation of Windows 2008, as shown in Figure 5-77.

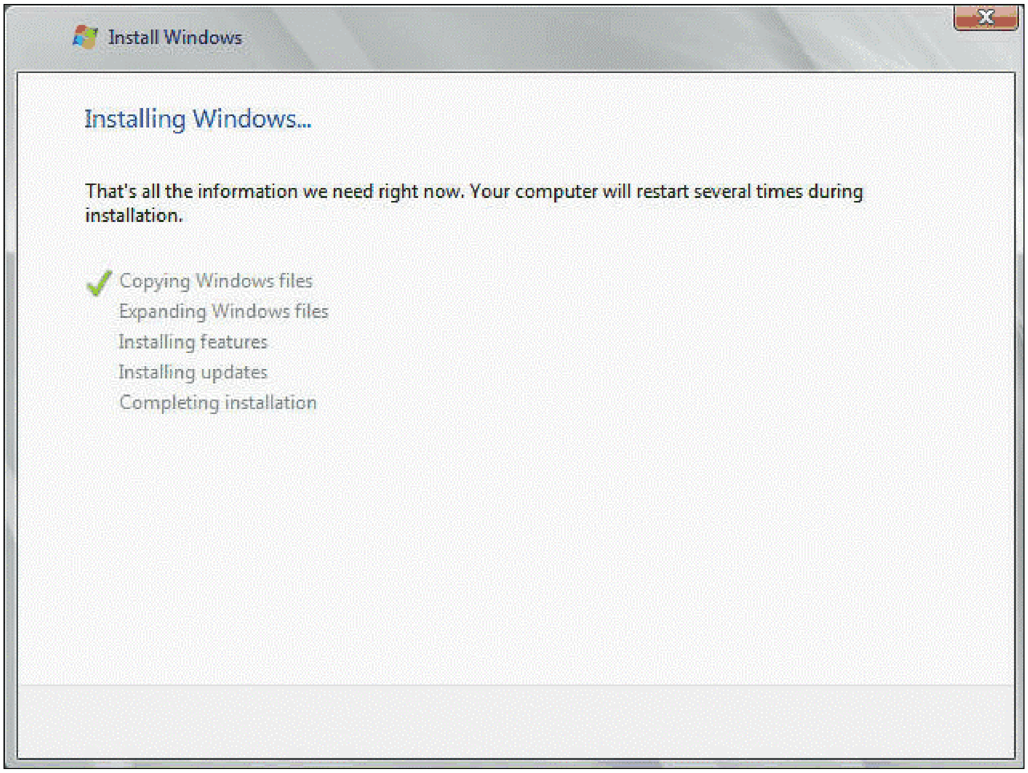


Figure 5-77 Installing Windows 2008

5.7.1 Multipath I/O

After the operating system completes installation process, the multipath driver needs to be installed to provide redundant paths between host and the storage. The ability to failover to alternate path becomes even more critical with SAN Boot because of OS residing on it.

Multiple Path I/O (MPIO): With MPIO, a device can be uniquely detected through one or more physical connections, or paths. In other words, the MPIO feature can be used to define alternate paths to a device for failover purposes.

Failover: Failover is a path-management algorithm that improves the reliability and availability of a device because the system automatically detects when one I/O path fails and re-routes I/O through an alternate path.

MPIO and SDDDSM dynamic pathing

Microsoft Multi Path Input Output (MPIO) solutions are designed to work in conjunction with device-specific modules (DSMs) written by vendors. IBM Subsystem Device Driver DSM (SDDDSM) is the IBM multipath I/O solution that is based on Microsoft MPIO technology, It is a device-specific module specifically designed to support IBM storage devices on Windows hosts.

MPIO supports dynamic pathing when you add more paths to an existing volume and when you present a new volume to the host. No user intervention is required, other than the typical new device discovery on a Windows operating system. SDDDSM uses a load-balancing policy that attempts to equalize the load across all preferred paths. If preferred paths are available, SDDDSM uses the path that has the least I/O at the time. If SDDDSM finds no available preferred paths, it tries to balance the load across all the paths it does find and uses the least active non-preferred path.

Path probing and reclamation is provided by MPIO and SDDDSM. For SDDDSM, the interval is set to 60 seconds. You can change this by modifying the following Windows system registry key: HKLM\SYSTEM\CurrentControlSet\Services\mpio\Parameters\PathVerificationPeriod

The latest supported drivers for multipathing in a SAN Boot Implementation can be found at:

http://www-03.IBM.com/systems/storage/disk/storwize_v7000/interop.html

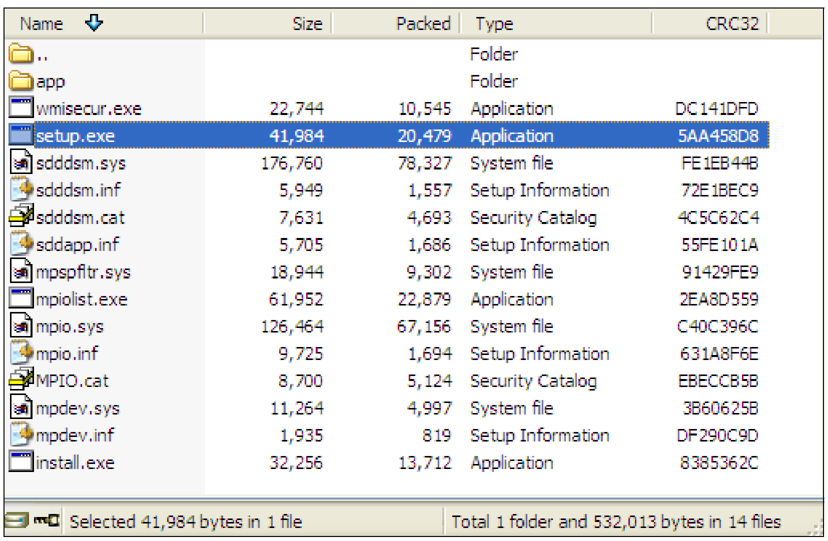
To ensure proper multipathing with IBM Storwize V7000 and SAN Volume Controller, SDDDSM must be installed on all Windows 2008 hosts:

1. Check the SDDDSM download matrix to determine the correct level of SDDDSM to install for Windows 2008 and download the package at this website:

   http://www-01.ibm.com/support/docview.wss?rs=540&uid=ssg1S7001350#WindowsSDDDSM

   > **Tip:** Using SDDDSM for multipathing, you must utilize the Storport Miniport drivers for Emulex HBAs and the STOR Miniport driver for QLogic HBAs.

2. Extract the package and start the setup application, as shown in Figure 5-78.



Name	Size	Packed	Type	CRC32
..			Folder	
app			Folder	
wmisecur.exe	22,744	10,545	Application	DC141DFD
setup.exe	41,984	20,479	Application	5AA458D8
sdddsm.sys	176,760	78,327	System file	FE1EB44B
sdddsm.inf	5,949	1,557	Setup Information	72E1BEC9
sdddsm.cat	7,631	4,693	Security Catalog	4C5C62C4
sddapp.inf	5,705	1,686	Setup Information	55FE101A
mpspfltr.sys	18,944	9,302	System file	91429FE9
mpiolist.exe	61,952	22,879	Application	2EA8D559
mpio.sys	126,464	67,156	System file	C40C396C
mpio.inf	9,725	1,694	Setup Information	631A8F6E
MPIO.cat	8,700	5,124	Security Catalog	EBECCB5B
mpdev.sys	11,264	4,997	System file	3B60625B
mpdev.inf	1,935	819	Setup Information	DF290C9D
install.exe	32,256	13,712	Application	8385362C

Selected 41,984 bytes in 1 file | Total 1 folder and 532,013 bytes in 14 files

Figure 5-78 SDDSM 2.4.3.1-2

3. The setup CLI appears, type **yes** to install the SDDDSM and press Enter, as shown in Figure 5-79.

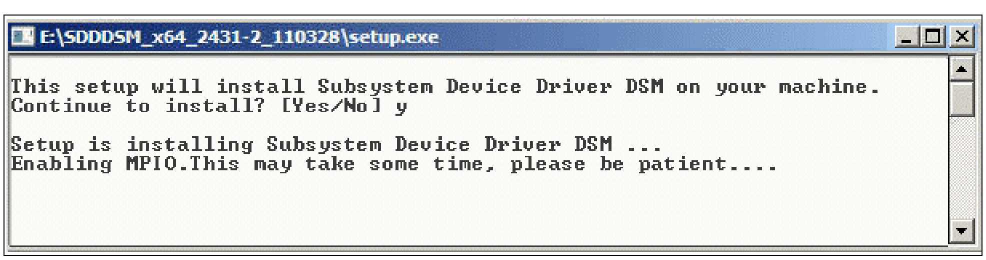


Figure 5-79 Setup CLI

4. After the setup completes, you are asked to restart the system. Confirm this by typing **yes** and press Enter, as shown in Figure 5-80.

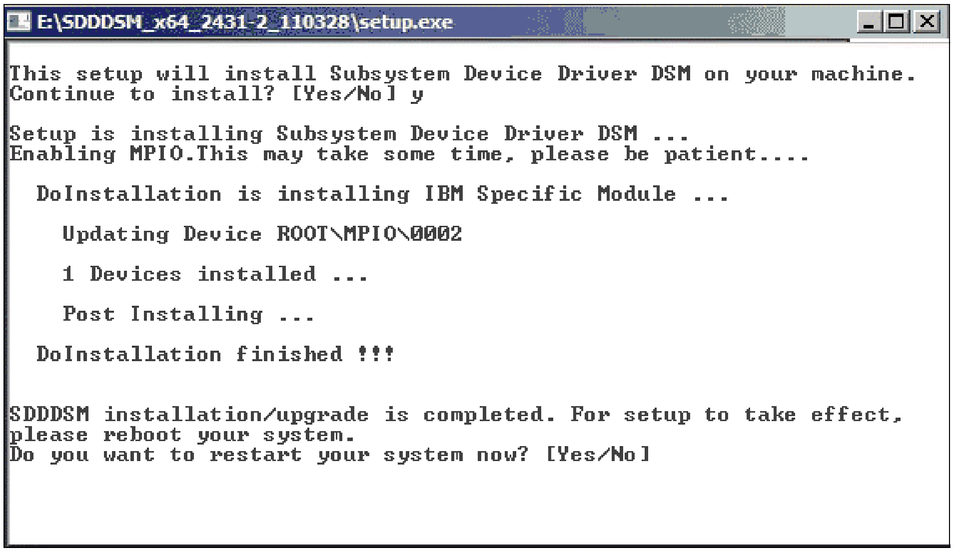


Figure 5-80 SDDDSM complete

Next, we add the second path between host and the storage.

5. Go to IBM Storwize V7000 management GUI and add the second HBA port (for failover) to the host definition, as shown in Figure 5-81

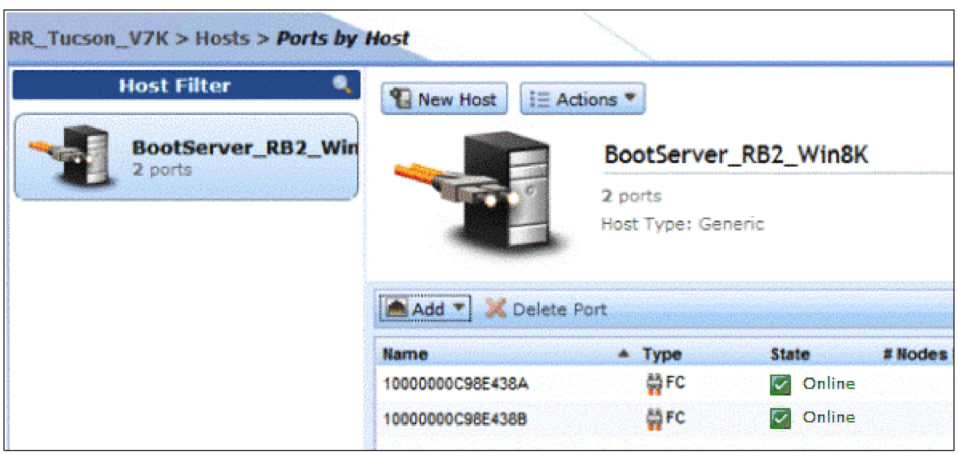


Figure 5-81 Adding Redundant path

6. Reboot server and go into HBA's BIOS Utility to enable SAN Boot options for the second path, between redundant HBA's port and Storwize V7000 node's host-port. For instructions, refer to section 5.5.2, "Configuring a QLogic adapter for SAN Boot" on page 315 and/or 5.5.4, "Configuring an Emulex adapter for SAN Boot" on page 324.
7. At the SAN switch, make necessary zoning configuration changes to add the second path between host and storage.

Now that you have SDDDSM installed and second path verified, log on to Windows OS.

Verifying the SDDDSM installation

To verify a successful SDDDSM setup, you have several possibilities:

- You can use the **SDDDSM** entry in the *My Programs* startup menu, as shown in Figure 5-82. The SDDDSM tools contain Subsystem Device Driver DSM, SDDDSM Technical Support website, README, and Notes.

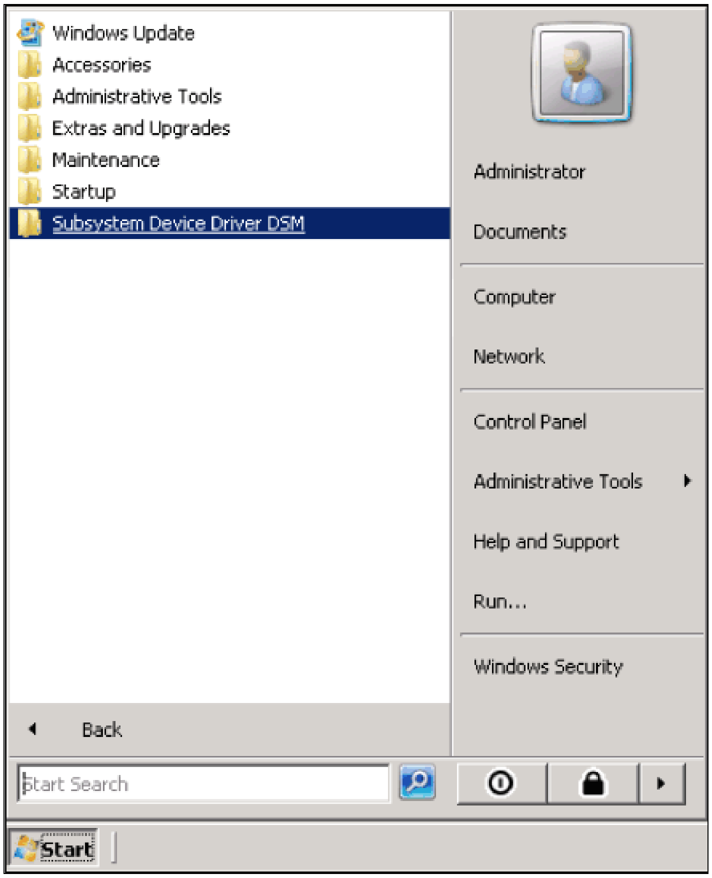


Figure 5-82 SDDDSM in Startup

- You can also verify that SDDDSM has been successfully, as follows:
  - Click **Start** → **Programs** → **Administrative Tools** → **Computer Management**.
  - Double-click **Device Manager**, then expand **Disk drives** in the right pane. You will see the Multipath Disk Device installed there, as shown in Figure 5-83.

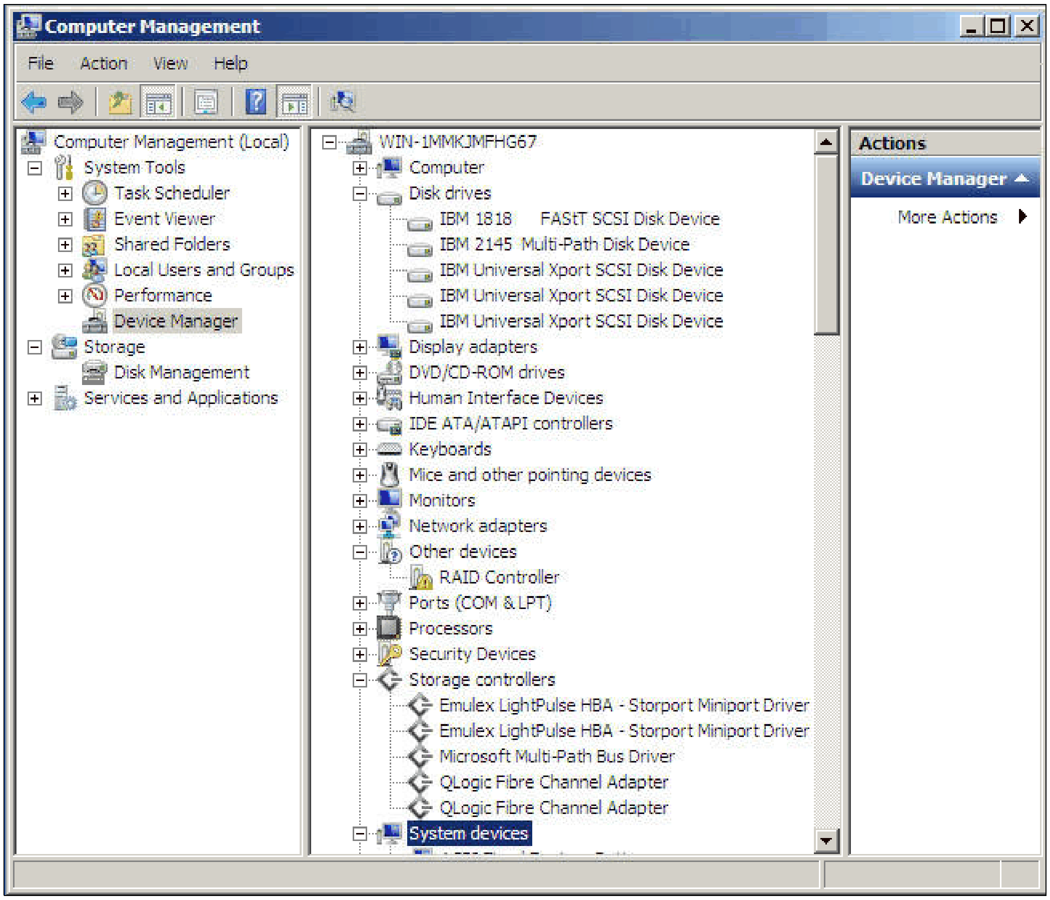


Figure 5-83 SDDDSM in Windows

Tip: When using SAN Boot feature with Microsoft clusters, to prevent inappropriate failover, set the Port Down Retry Timer (in HBA BIOS settings) to 15 seconds.

- A useful tool offered by SDDDSM is the **`datapath`** command. One common use of *datapath* is using it to query the HBA ports by issuing the **`datapath query device`** command, as shown in Figure 5-84.

```
C:\Program Files\IBM\SDDDSM>datapath query adapter

Active Adapters :2

Adpt#         Name       State    Mode      Select    Errors   Paths    Active
    0   Scsi Port6 Bus0  NORMAL   ACTIVE     14788         0       1         1
    1   Scsi Port7 Bus0  NORMAL   ACTIVE      1551       591       1         1
```

Figure 5-84 datapath query adapter

- You can also use the *datapath* tool to determine the WWPNs of the host. Type **`datapath query wwpn`** as shown in Example 5-1.

Example 5-1 Datapath query wwpn

```
C:\Program Files\IBM\SDDDSM>datapath query wwpn
        Adapter Name    PortWWN
        Scsi Port3:     100000051EAF08A0
        Scsi Port4:     100000051EAF08A1
        Scsi Port5:     100000051E0C1CCC
        Scsi Port6:     100000051E0C1CCD
```

Multipathing configuration maximums

When you configure multipathing on your hosts, you must consider the maximum supported configuration limits. Table 5-2 provides the configuration maximums for hosts running the Windows Server 2008 operating system.

Table 5-2 Maximum configurations

Object	Maximum	Description
Volume	512	The maximum number of volumes that can be supported by the IBM Storwize V7000 and SAN Volume Controller system for a host running a Windows operating system (per host object)
Paths per Volume	8	The maximum number of paths to each volume. The recommended number of paths is four.

It is important to check your operating system and HBA documentation for limitations that might be imposed by other driver software.

Tip: SDDDSM for Windows supports 16 paths per volume, however, IBM Storwize v7000 system supports only a maximum of eight paths, to support reasonable path-failover time.

For detailed information about SDDDSM, refer to *Multipath Subsystem Device Driver User's Guide,* GC52-1309-02.

5.7.2 Migrating existing SAN Boot images in Windows 2008

If you have a host running the Windows Server 2008 operating system, having existing SAN Boot images controlled by storage controllers, you can migrate these images to image-mode volumes that are controlled by the IBM Storwize V7000 system.

Perform the following steps to migrate your existing SAN Boot images:

1. Shut down the host.
2. Perform the following configuration changes on the storage controller:
   a. Remove all the image-to-host mappings from the storage controller.
   b. Map the existing SAN Boot image and any other disks to the IBM Storwize V7000 system.
3. Zone one port of each host bus adapter (HBA) to one of the IBM Storwize V7000 system ports that is associated with the I/O group for the target image-mode volume.
4. Perform the following configuration changes on the IBM Storwize V7000 system:
   a. Create an image mode volume for the managed disk (MDisk) that contains the SAN Boot image. Use the MDisk unique identifier to specify the correct MDisk.
   b. Create a host object and assign it to the HBA port that is zoned to the IBM Storwize V7000 host-port.
   c. Map the image mode volume to the host. For example, you might map the boot disk to the host with SCSI LUN ID 0.
   d. Map the swap disk to the host, if required. For example, you might map the swap disk to the host with SCSI LUN ID 1.
5. Change the boot address of the host by performing the following steps:
   a. Restart the host and open the BIOS utility of the host during the booting process.
   b. Set the BIOS settings on the host to find the boot image at the worldwide port name (WWPN) of the node that is zoned to the HBA port.
6. Boot the host in single-path mode.
7. Uninstall any multipath driver that is not supported for IBM Storwize V7000 system hosts that run the applicable Windows Server operating system.
8. Install a supported multipath driver.
9. Restart the host in single-path mode to ensure that the supported multipath driver was properly installed.
10. Zone each HBA port to one port on each IBM Storwize V7000 system node.
11. Add HBA ports to the host object that you created in Step 5 b.
12. Configure the HBA settings on the host by using the following steps:
    a. Restart the host and open the host's BIOS utility during the booting process.
    b. Ensure that all HBA ports are boot-enabled and can see both nodes in the I/O group that contains the SAN Boot image. Configure the HBA ports for redundant paths.
    c. Exit the BIOS utility and finish booting the host.
13. Map any additional volumes to the host as required.

Attention: In Microsoft Failover Clustering Windows Server 2008 allows boot disk and the cluster server disks to be hosted on the same bus. However, you need to use Storport miniport HBA drivers for this functionality to work. It is *not* a supported configuration in combination with drivers of other types (for example, SCSI port miniport or Full port drivers).

5.8 SAN Boot for VMware

VMware ESX Server systems are using SAN Boot implementations increasingly, reaping its benefits like cost saving in diskless environments, better disaster recovery, redundancy, efficient use of storage space and ease of management, etc.

In the following pages, we explain the steps to implement SAN Boot using *VMware ESXi version 4.1* on a boot volume residing at an IBM Storwize V7000 storage system.

> **Tip:** Before setting up SAN Boot for VMware, review the information in 5.3.1, "Considerations before setting up SAN Boot" on page 286.

Consider the following restrictions before starting the installation process:

1. For SAN Boot, each ESXi host must use a dedicated volume.
2. At the host system's BIOS configuration, enable **Legacy Boot** under the *Boot Menu*. It will allow for discovery of Fibre Channel attached devices on VMware ESXi servers.
3. When installing on a host with **UEFI BIOS**, move the *Legacy* option to the top of the *Boot Options* to allow Fibre Channel attached devices be discovered after reboot of VMware ESXi servers.

Perform the following to install ESXi 4.1.0:

1. Start the installation from *VMware VMvisor Boot Menu*, as shown in Figure 5-85.

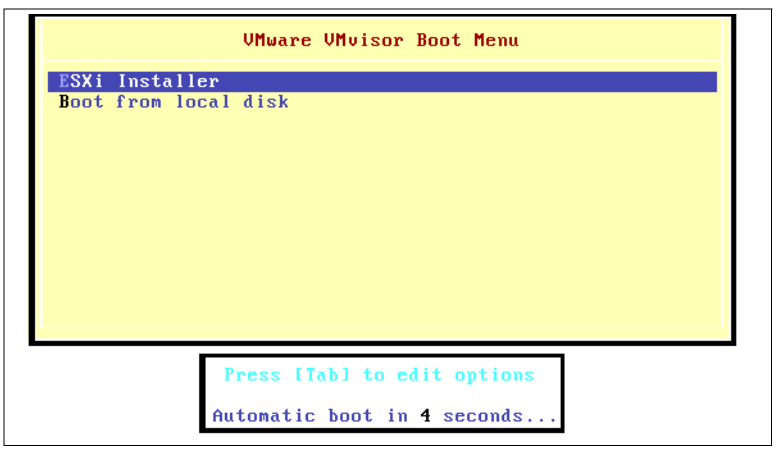


Figure 5-85 Boot Menu

2. Press Enter to install VMware ESXi 4.1.0, as shown in Figure 5-86.

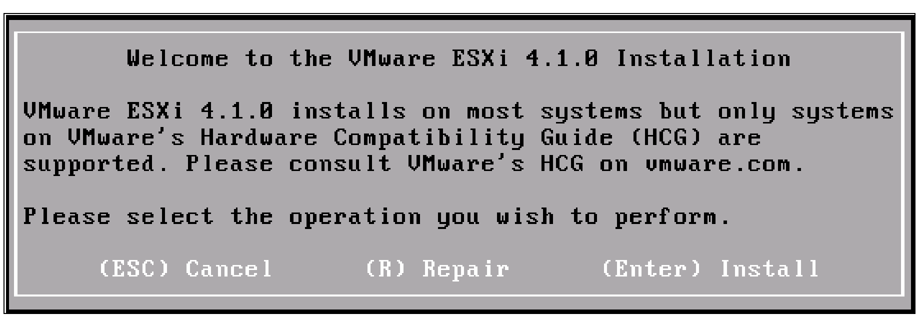


Figure 5-86 Installation panel

3. Select the **Remote** volume (*2145* for Storwize V7000 and SAN Volume Controller (SVC)) as the target for installation, as shown in Figure 5-87.

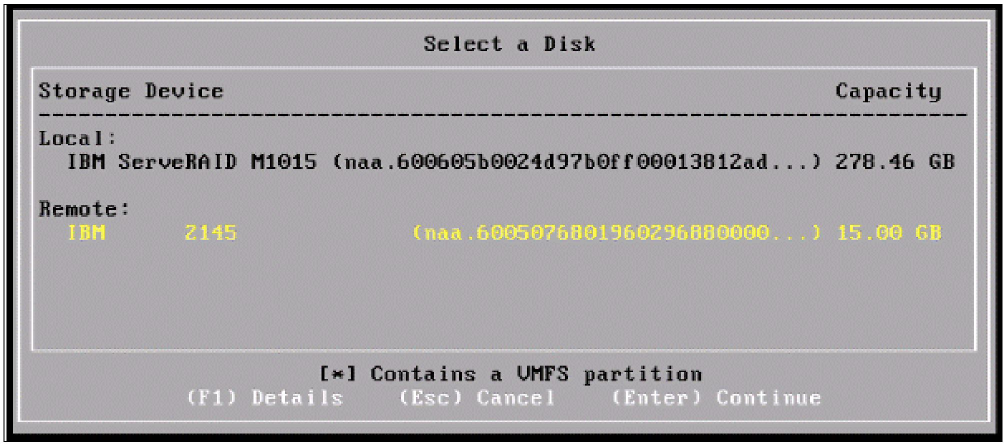


Figure 5-87 SAN Volume Controller or Storwize V7000 LUN as target for boot

4. Press F11 to confirm installation on the Boot LUN from the storage, as shown in Figure 5-88.

Figure 5-88 Confirm Install

After the installation process is complete, the system will report successful installation, as shown in Figure 5-89.

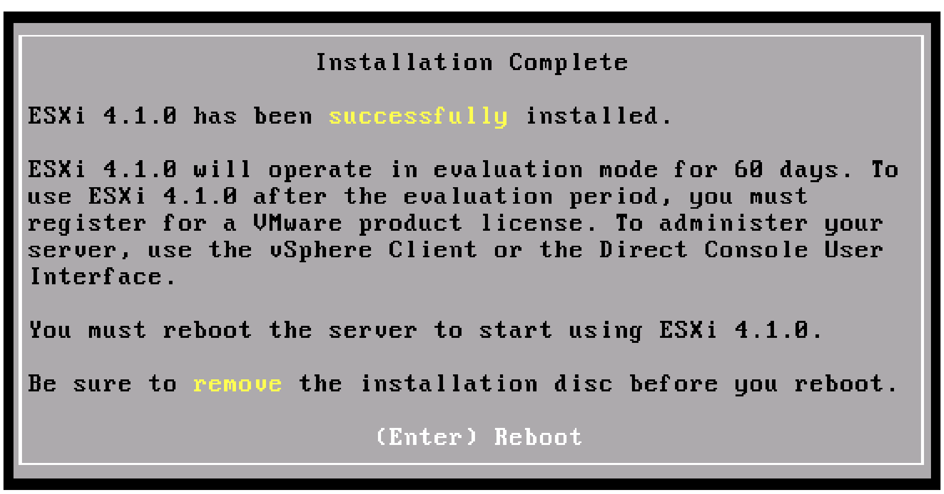


Figure 5-89 Installation Complete

5. After reboot, the host should boot from the volume designated on Storwize V7000 storage, as shown in Figure 5-90.

> **Tip:** For diskless ESXi hosts that boot from SAN, multiple ESXi host systems can share one diagnostic partition on a SAN volume.

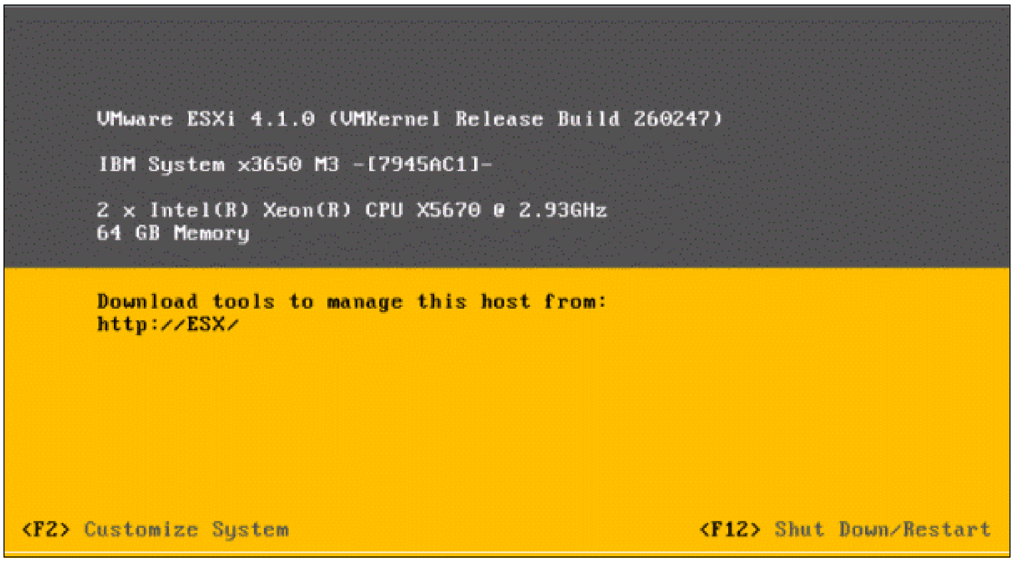


Figure 5-90 Booting VMware ESXi 4.1.0

To manage ESXi, we need a Virtual Infrastructure called vSphere Client. The VMware vSphere 4 can be downloaded from the following website:

http://downloads.vmware.com/d/info/datacenter_downloads/vmware_vsphere_4/4_0

6. Log in to the vSphere Client, as shown in Figure 5-91.

Figure 5-91 vSphere Login

5.8.1 Creating datastores and installing a guest operating system

A datastore is storage for virtual machines. After verifying volume access, configure the VMware datastores, then install your guest operating systems and applications in accordance with your virtualization strategy.

5.8.2 Multipath I/O

VMware ESXi provides multipathing support natively in one of three forms, *most recently used (MRU)*, *fixed path*, and *round robin*.

IBM Storwize v7000 is an active/active virtualized storage subsystem, and therefore *round robin* is the preferred multipathing policy.

Tip: The default policy for new IBM Storwize v7000 volumes discovered by ESXi is *fixed path*.

Using *round robin* reduces the complexity and management overhead required in maintaining *MRU* or *fixed path* policies.

To change the preferred path management policy, perform the following procedure:

1. In **Configuration** tab, select the Fibre Channel adapter. Next, right-click a *storage device* to select **Manage Paths**, as shown in Figure 5-92.

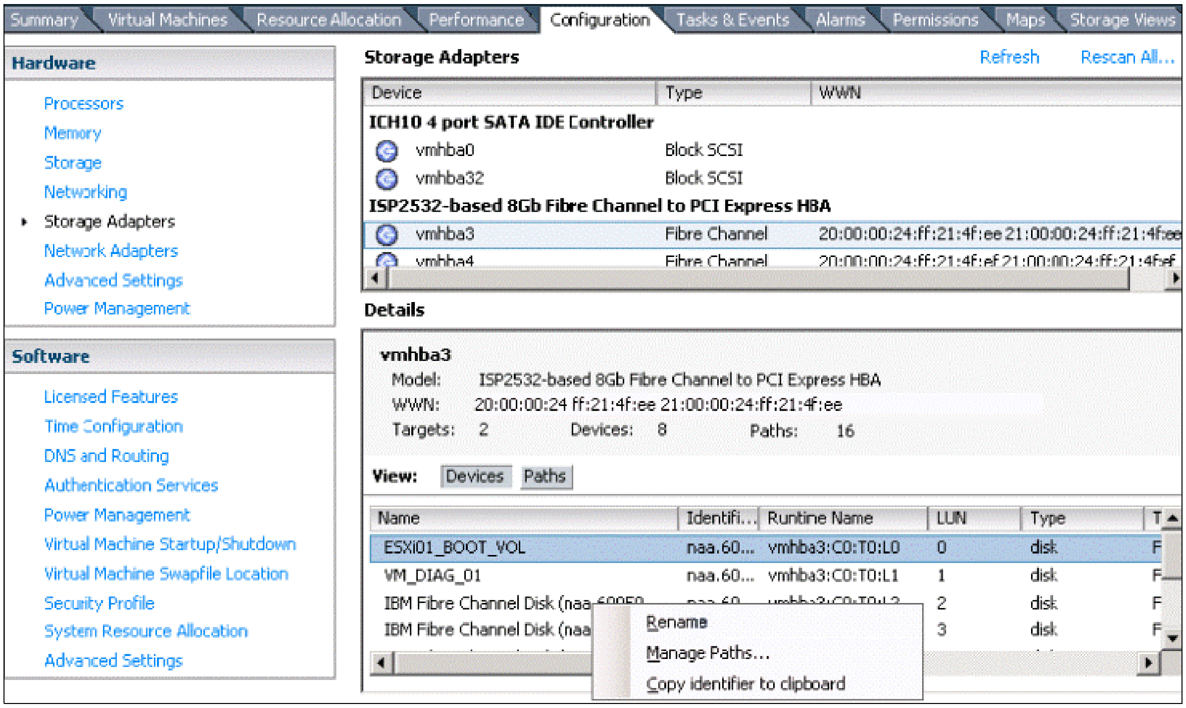


Figure 5-92 Manage HBA Paths

2. In the **Manage Paths** dialogue window, select **round robin** from the *path selection* drop-down menu. Click the **Change** button, as shown in Figure 5-93.

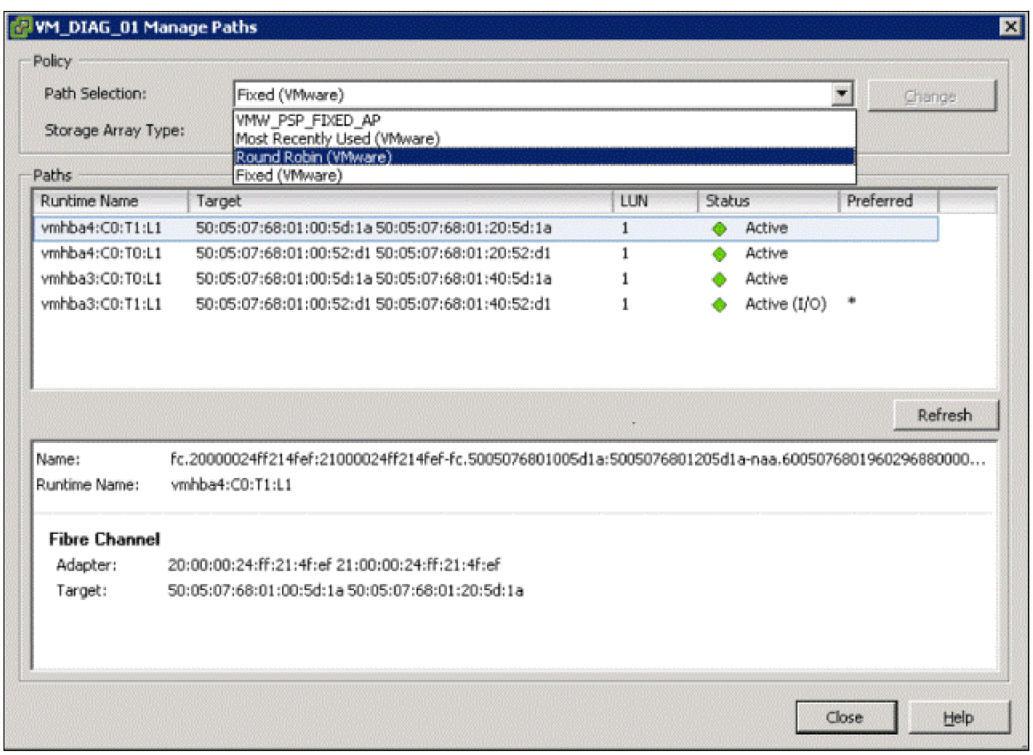


Figure 5-93 Path Policy Selection

3. You can verify the multipathing for the datastore by going to the **Summary** tab in the vSphere Client, as shown in Figure 5-94.

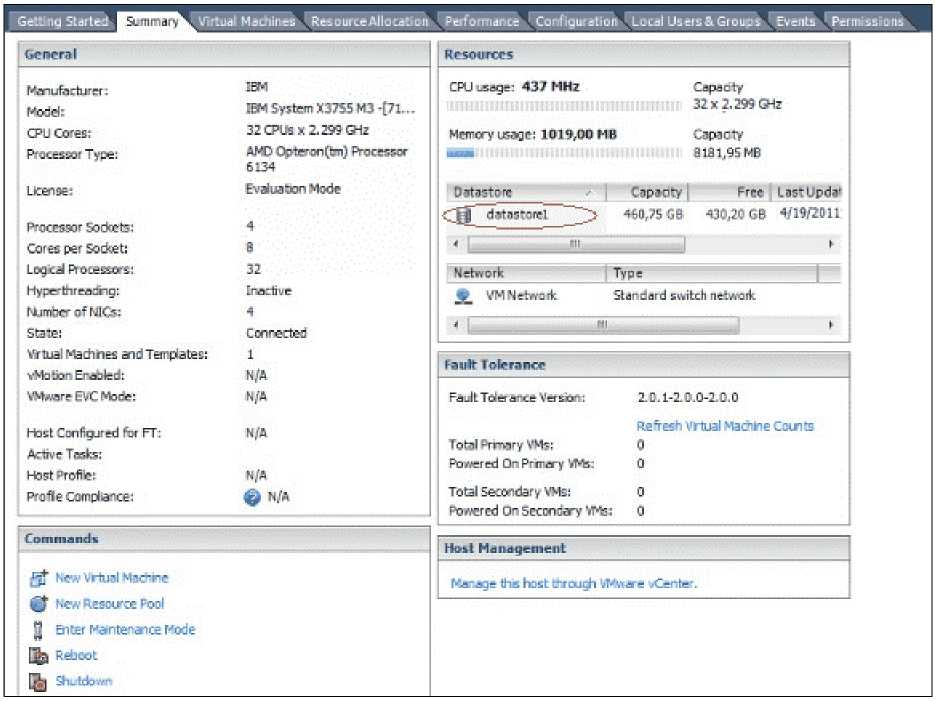


Figure 5-94 Summary in vSphere Client

4. Right-click the datastore, and select **Properties** → **Manage Paths**, as shown in Figure 5-95.

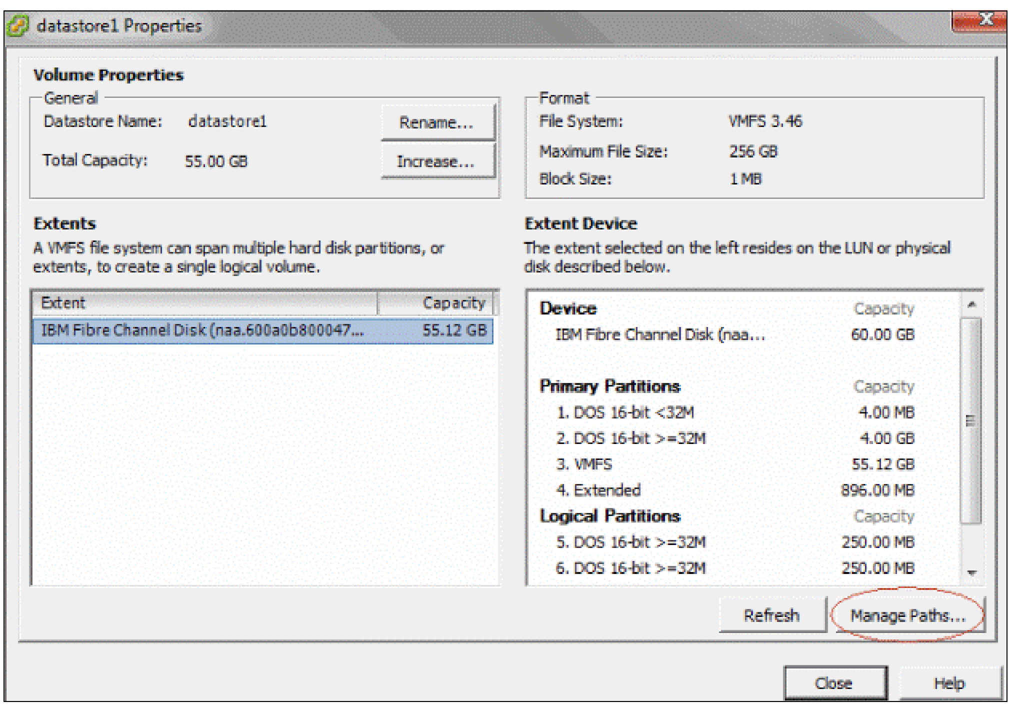


Figure 5-95 Manage Paths for datastore

5. You should see both paths in the Manage Paths dialogue window, as shown in Figure 5-96.

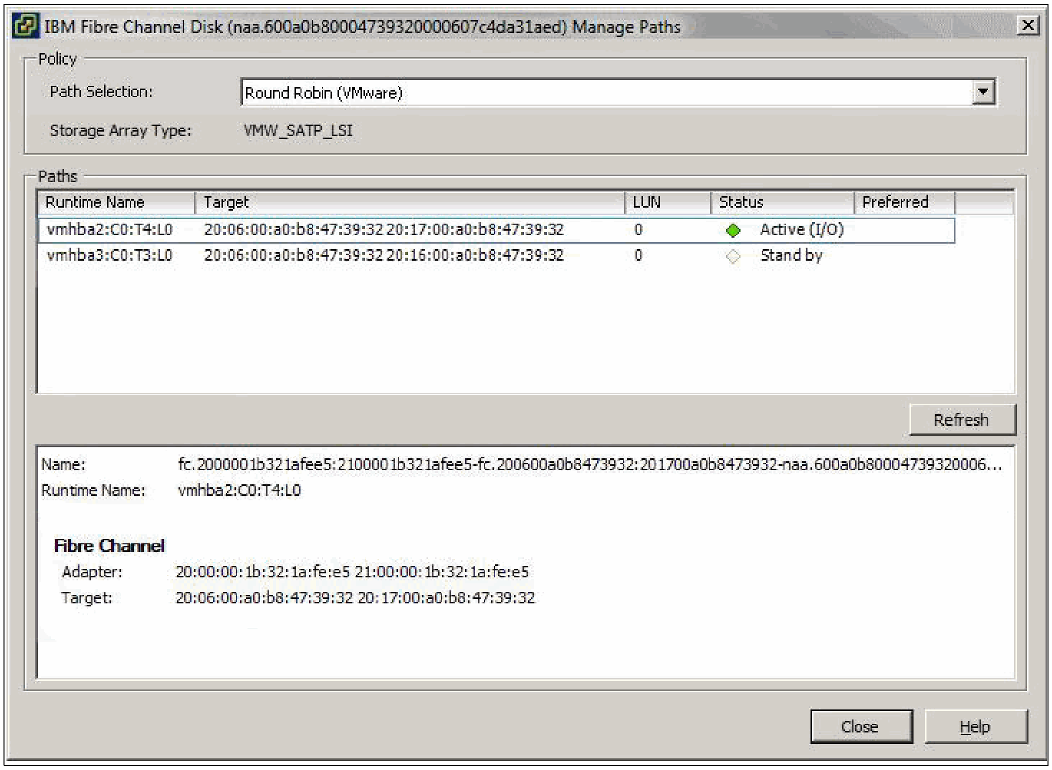


Figure 5-96 Manage Path Policy

For more information regarding SAN Boot for VMware, see the VMware (4.1) Fibre Channel SAN Configuration Guide document, at this website:

http://www.vmware.com/pdf/vsphere4/r41/vsp_41_san_cfg.pdf

5.9 SAN Boot for Linux

WIth the increasing trend towards implementing SAN Boot solutions, various distributions of Linux are being installed on external storage for SAN Boot environments.

In the following sections, we describe two of the most widely used distributions of Linux, *SUSE Linux Enterprise Server 11 SP 2* and *Red Hat Enterprise Linux 5.5 (RHEL).*

> **Tip:** Before setting up SAN Boot for Linux, review the information in 5.3.1, “Considerations before setting up SAN Boot” on page 286.

5.9.1 SAN Boot using SUSE Linux Enterprise Server 11 SP2

Following are the general steps to set up a SUSE Linux host using SAN Boot:

1. Configure the Storwize V7000 Subsystem in a such way that only the boot volume you want is mapped to the host.

> **Tip:** Unlike other operating systems, with SUSE Linux Enterprise Server 11 SP2, it is recommended to have multiple paths configured between host and the SAN Volumes before installation of the operating system.

2. Configure and enable the HBA BIOS, as shown in 5.5.1, “Finding WWPNs of QLogic FC HBAs” on page 313 and/or 5.5.3, “Finding WWPNs of Emulex FC HBAs” on page 321.
3. Selecting the volume configured for SAN Boot (see section 5.4.5, “Creating a volume for SAN Boot” on page 298) and start the installation process.
4. After initializing the installation process, you come to the first window that gives various startup options. Choose **Installation** at the *Welcome* panel, as shown in Figure 5-97.

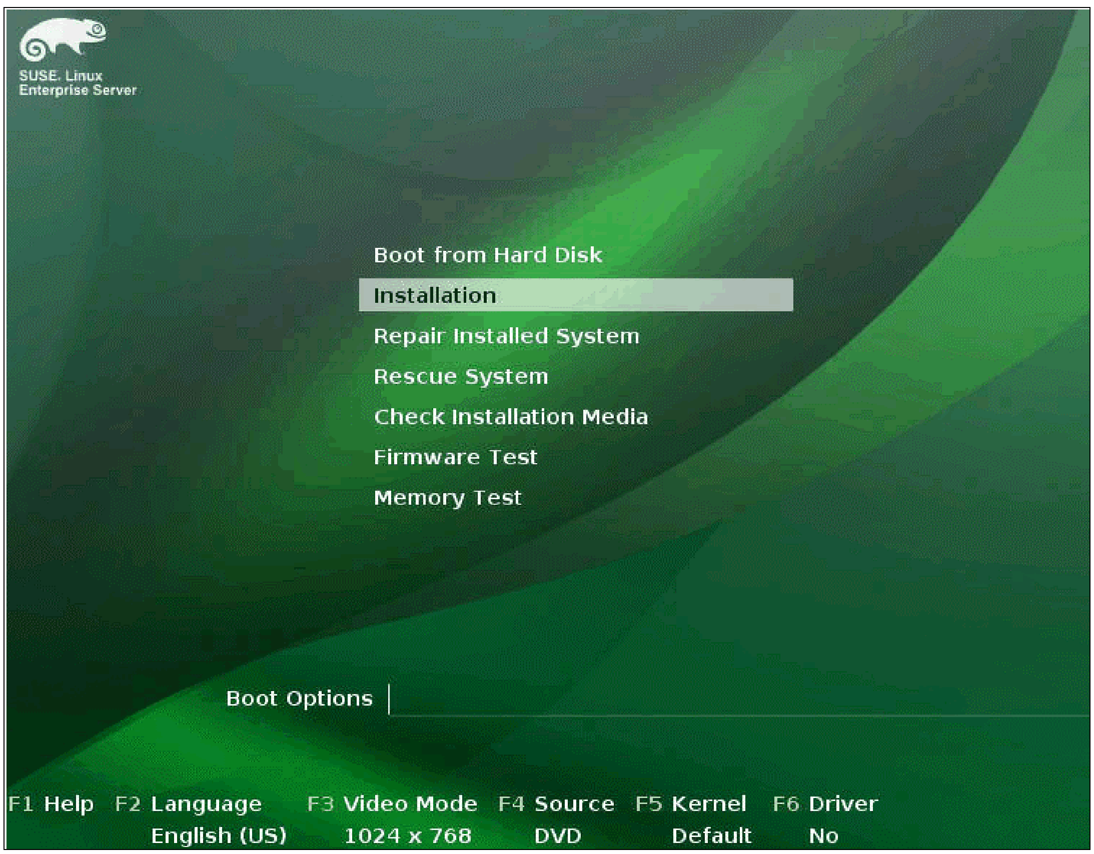


Figure 5-97 Welcome to SUSE Linux Enterprise Server 11 installation

5. Select **New Installation** as the installation mode, as shown in Figure 5-98.

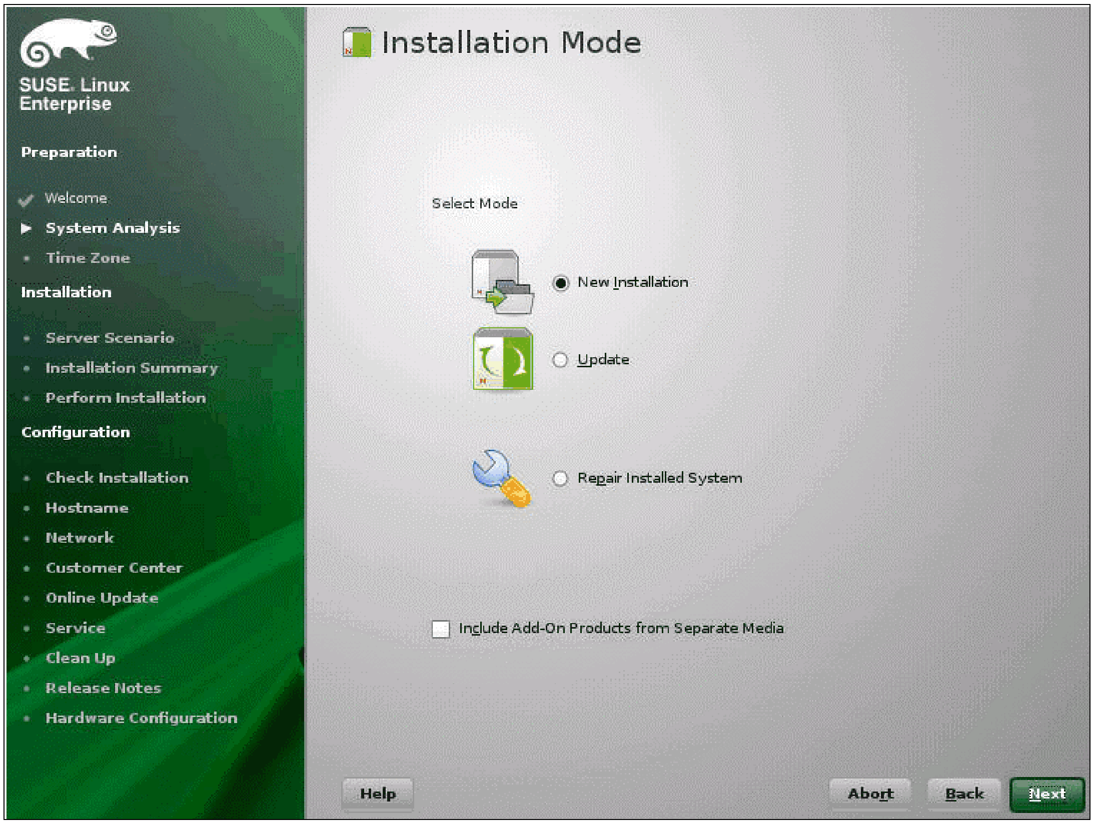


Figure 5-98 Installation Mode

6. Set the clock and time zone. Next, at the *Server Base Scenario* window, choose **Physical Machine**, as shown in Figure 5-99.

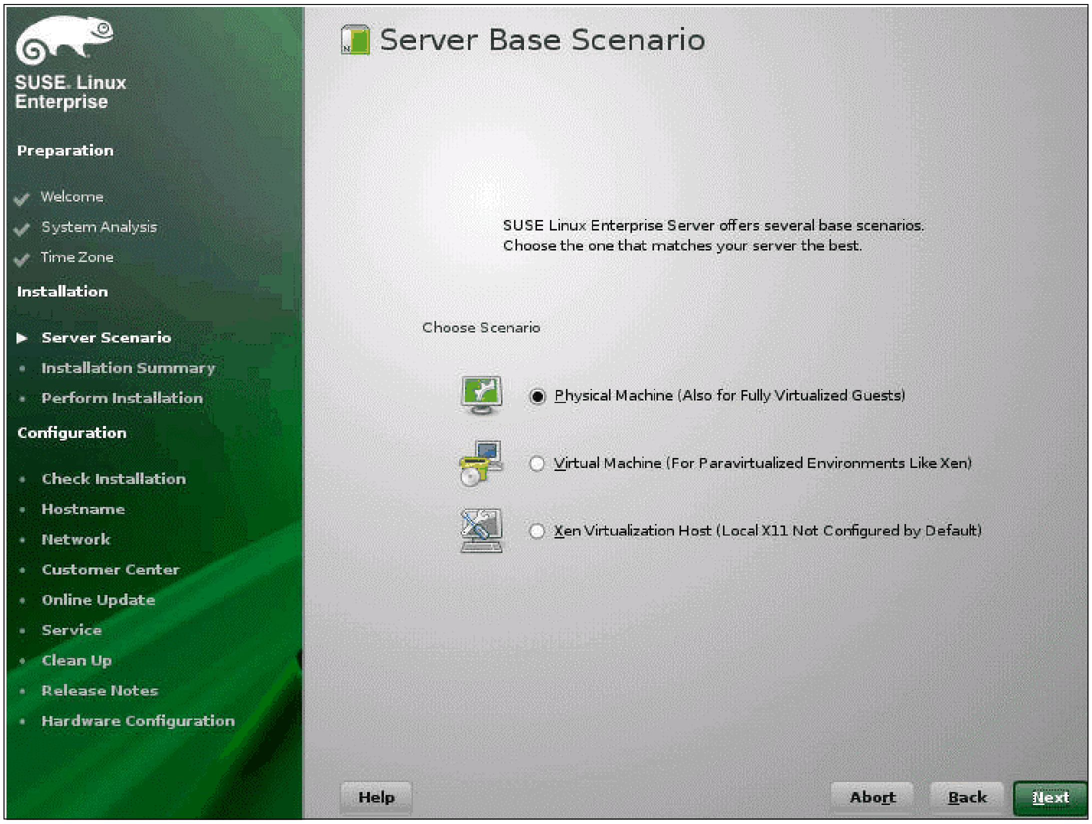


Figure 5-99 Server base scenario

7. At the *Installation Settings* windows, you can configure multiple paths during the installation process. To proceed using multiple paths, choose the **Expert** tab, as shown in Figure 5-100.

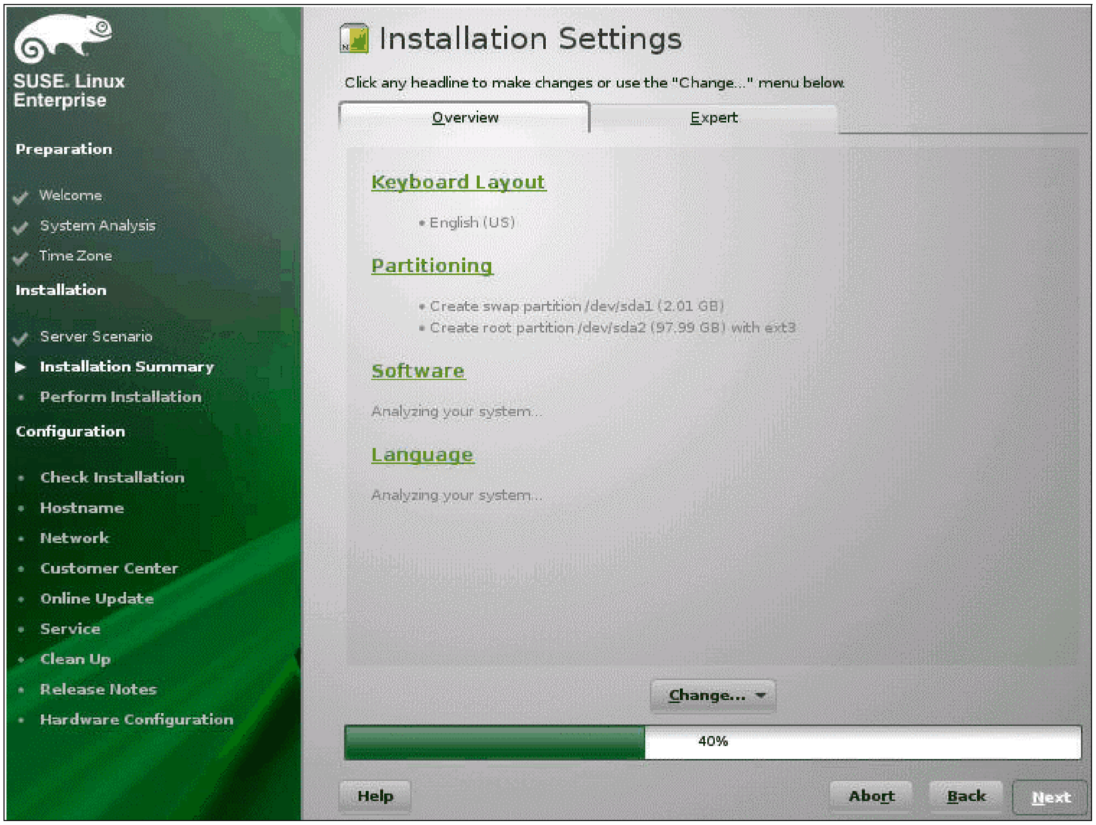


Figure 5-100 Installation Settings

8. Click the **Partitioning** link to change the partition information from default drive to the volume allocated on Storwize V7000 for SAN Boot, as shown in Figure 5-101.

Figure 5-101 Select boot LUN from SVC

9. Click the **Hard Disks**, then click **Configure** and **Configure Multipath**, as shown in Figure 5-102.

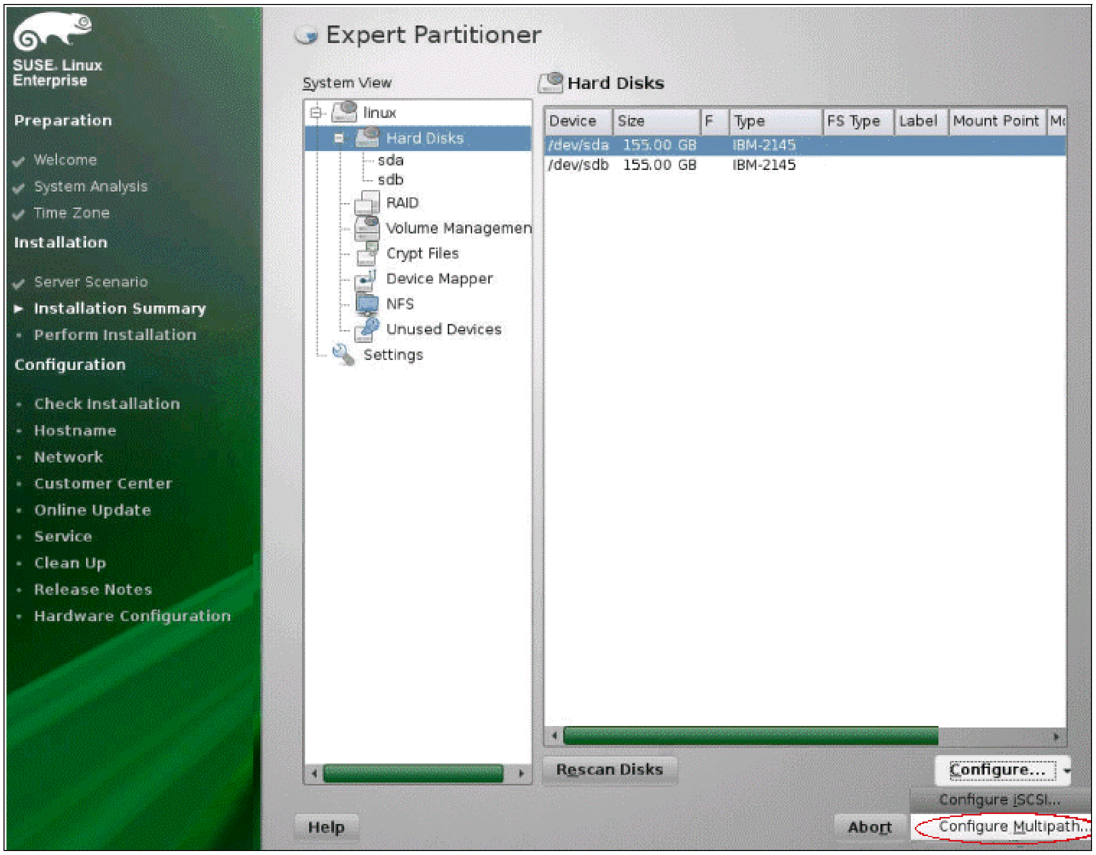


Figure 5-102 Configure Multipath

10. You see a confirmation dialogue box for *Activation of Multipath*. On clicking **Yes,** it will rescan devices and display the updated device information, as shown in Figure 5-103.

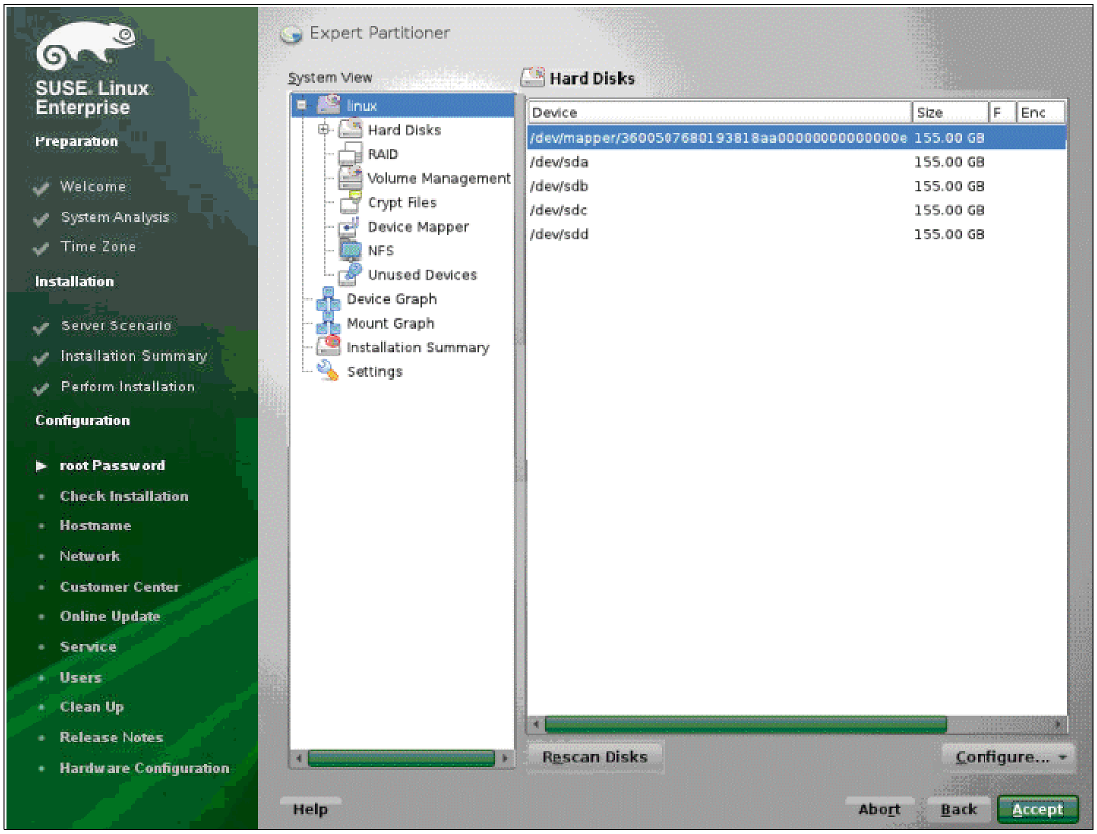


Figure 5-103 Multipathing Selected

11. Now we configure the partitions for installation. In the left pane, under **System View**, select the multipath device under **Hard Disks** by left-clicking it.

12. At the bottom of the **Partitions** tab, click **Add**. In this example, a 10.00 GB primary partition was created for swap space. Click **Add** again to create a 60.99 GB primary partition for the root filesystem, as shown in Figure 5-104.

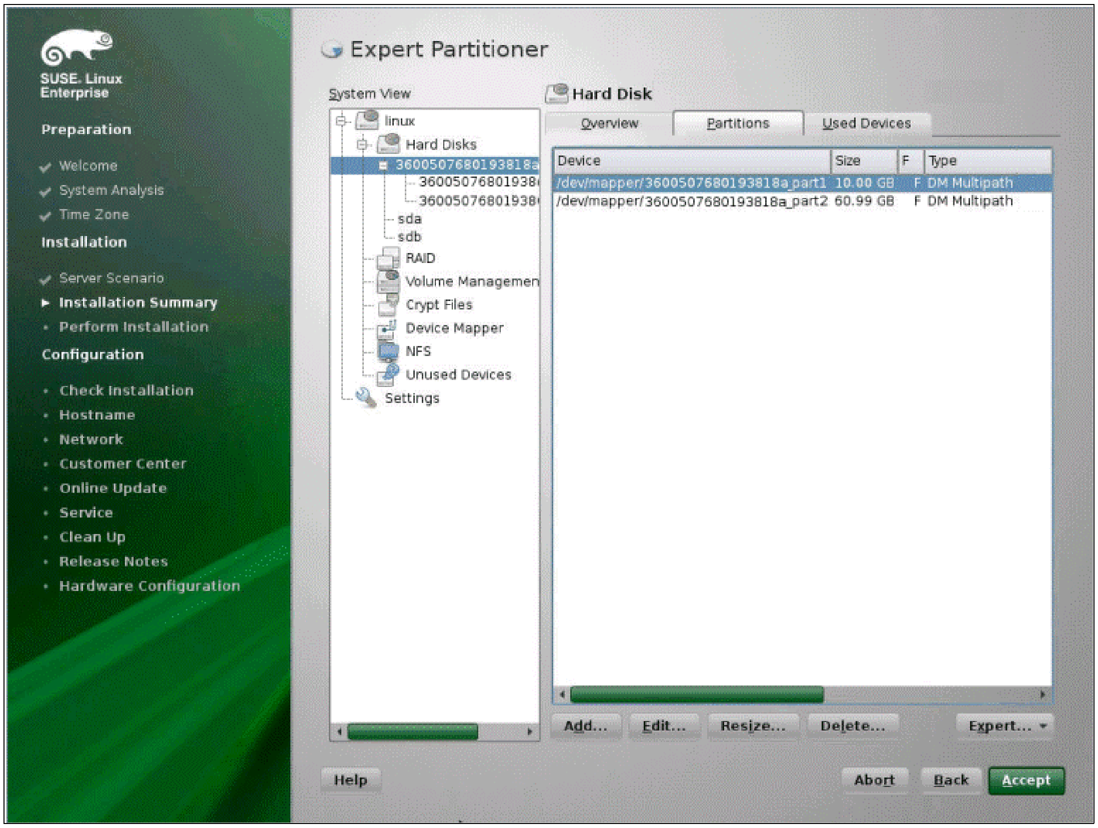


Figure 5-104 Multipath added

13. After the desired partitions have been created, click **Accept**.

14. Change the Boot loader to reflect the multipath partitioning changes. Click the **Booting** option, **Boot Loader Installation**, and then click **Details**, as shown in Figure 5-105.

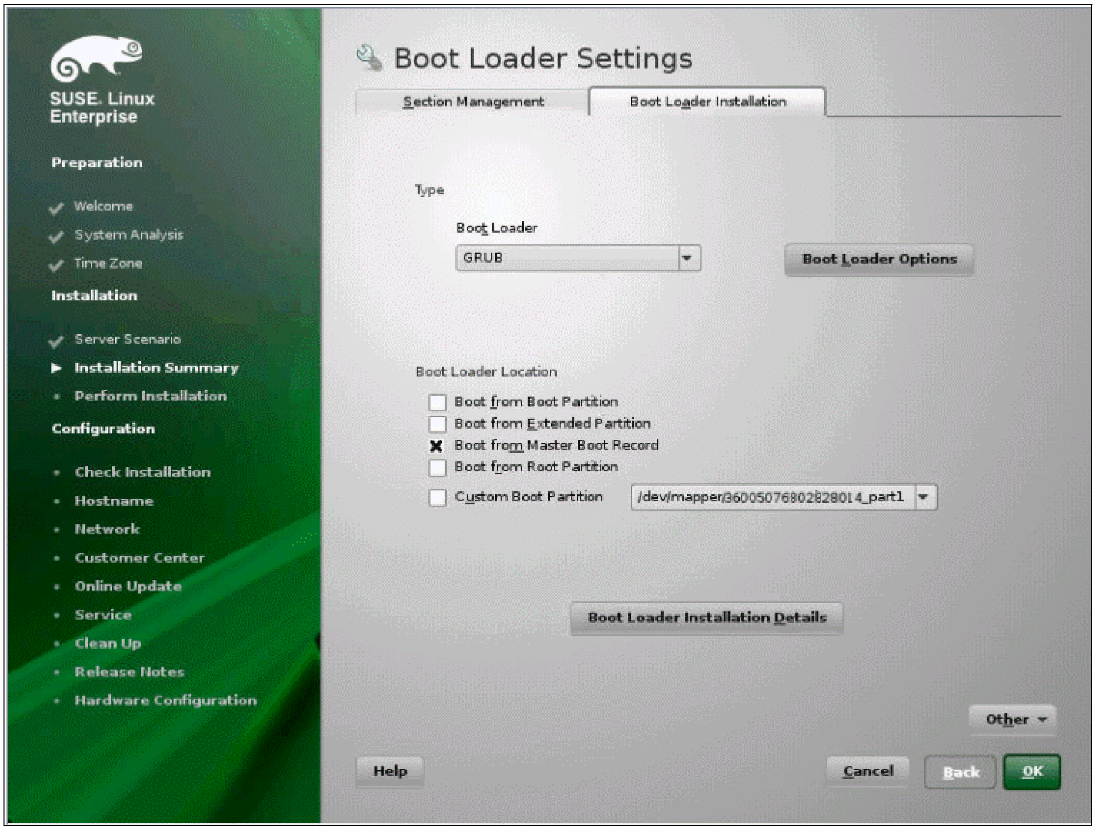


Figure 5-105 Boot from MBR

15. Click the button, **Boot Loader Installation Details**, as shown in Figure 5-106.

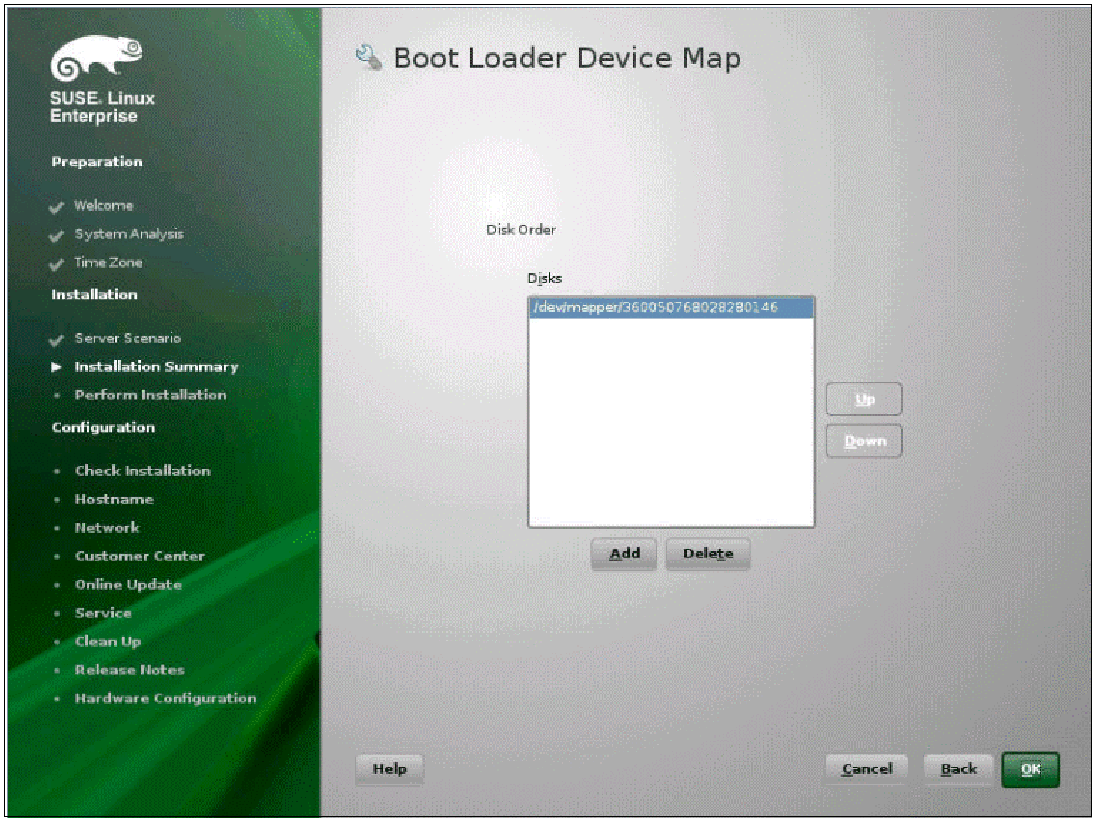


Figure 5-106 Boot Loader Device Map

16.Click **Ok**. Now the installer will show the final settings before installation, as shown in Figure 5-107.

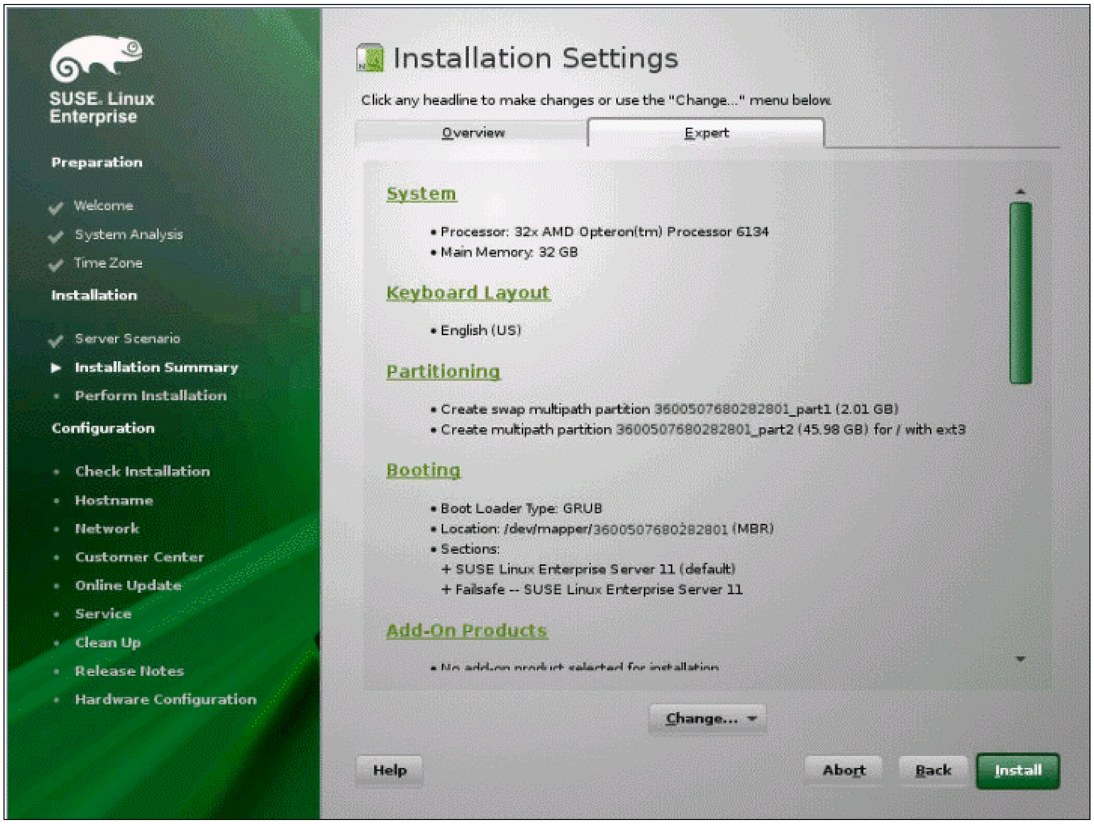


Figure 5-107 Installation Settings

17.Click **Install** to start installing the operating system on the boot LUN. At the end of installation, the system will reboot and take you to the *Login* panel. Here you type in the administrator (root) password and start using SUSE Linux Enterprise Server 11 SP2 operating system.

5.9.2 Installation of Red Hat Enterprise Server 5.5 for SAN Boot

Perform the following steps to install RHEL 5.5 on external storage (Storwize V7000):

1. Configure the Storwize V7000 Subsystem so that only the boot volume is mapped to the host.
2. Set up multipath (both paths) between host and the storage.

Tip: Starting with RHEL 5.1, Anaconda has the capability to detect, create, and install dm-multipath devices:

- *Anaconda* is the installation program used by Fedora, Red Hat Enterprise Linux, and some other Linux distributions.
- *DM-Multipath* stands for *Device mapper multipath*, which is the native multipath driver in Red Hat Enterprise Linux.

3. Start the installation program. At the installation mode's selection panel, press **F2** for *Options*. Type `linux mpath` at the *kernel boot line* and press Enter, as shown in Figure 5-108.

Figure 5-108 RHEL installation - select mode

4. After selecting the *language* and *keyboard type*, select the installation media. In our case, we select *CDROM* as the installation media, as shown in Figure 5-109.

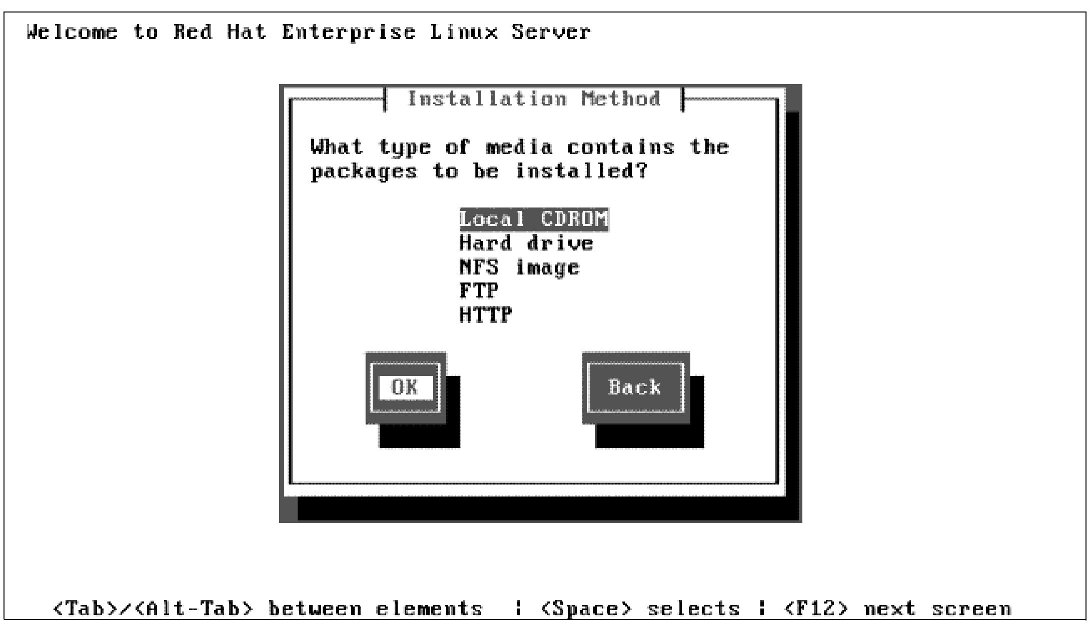


Figure 5-109 Select Installation Media

5. Click **Next** to proceed with the installation, as shown in Figure 5-110.

Figure 5-110 RHEL installation

6. You see a dialogue box asking for an *Installation Number* to validate your license. The next panel shows both paths between host and storage, as shown in Figure 5-111. Choose the drive for OS installation and click **Next**.

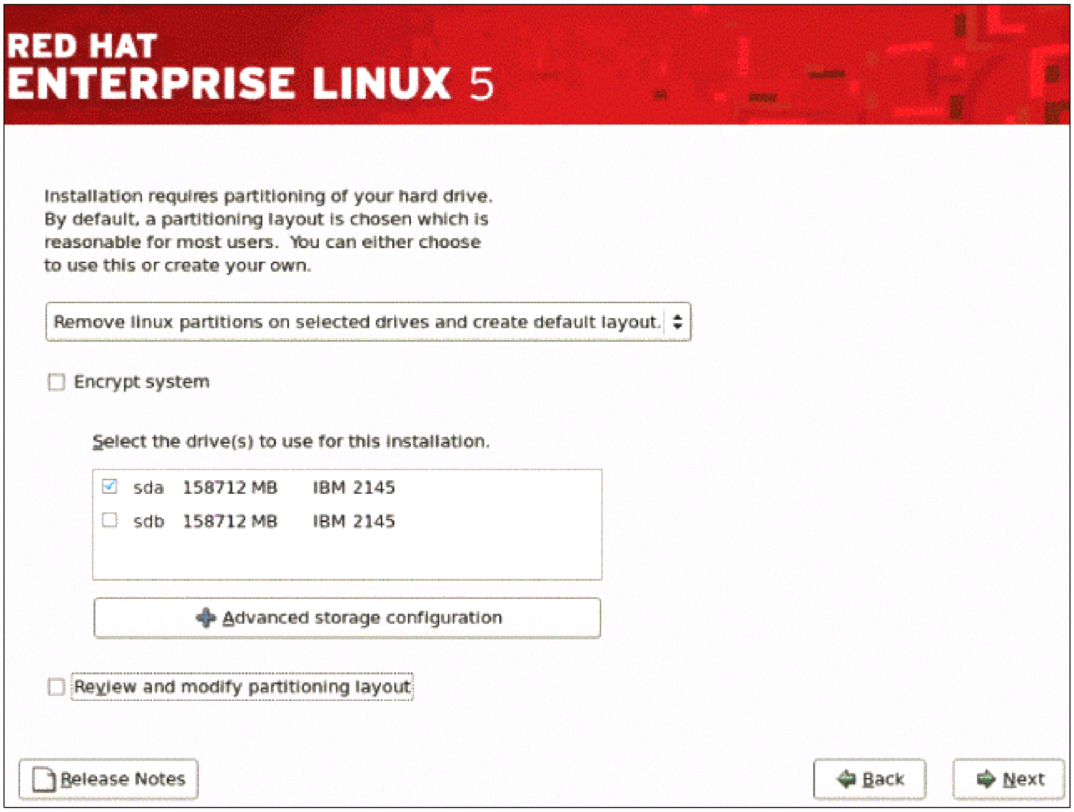


Figure 5-111 Select drive for SAN Boot installation

The next window displays additional details regarding the default Red Hat partitioning and filesystem layout and the use of the related multipath devices. Verify that the settings are as desired, then click **Next**, as shown in Figure 5-112.

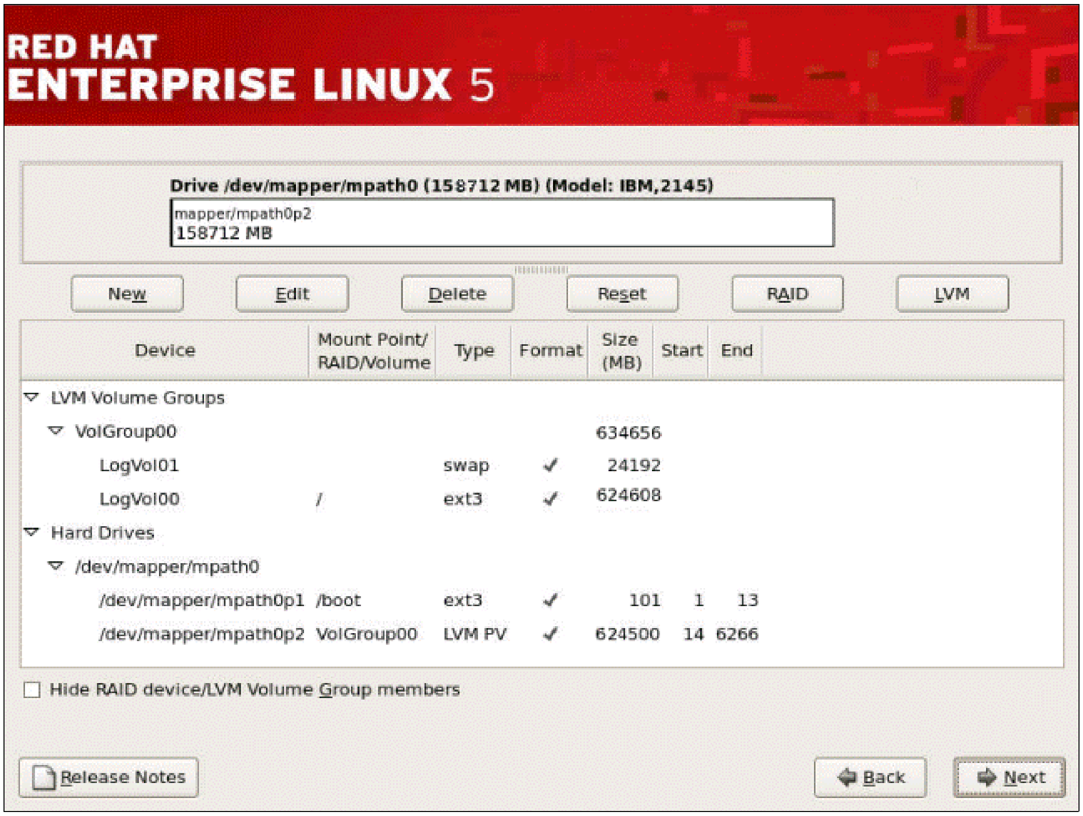


Figure 5-112 Partition Details

7. Due to the fresh installation on the SAN Boot volume, the installer prompts to create new partitions.

> **Tip:** If the installation menu does not show the devices in `mapper/mpath0` format, verify that you provided `mpath` in the command line at the first installation panel. Also verify that your hardware is configured appropriately, such that the storage devices are seen through multiple paths.

Because the partitions will be formatted and initialized, re-verify the disk information before proceeding, as shown in Figure 5-113.

Figure 5-113 Formatting unreadable partition

8. Select **Review** and modify partitioning layout, then click **Next**, as shown in Figure 5-114.

Figure 5-114 Review and modify partition

9. The next window gives you the option to configure the network devices, as shown in Figure 5-115. Choose your settings, then click **Next**.

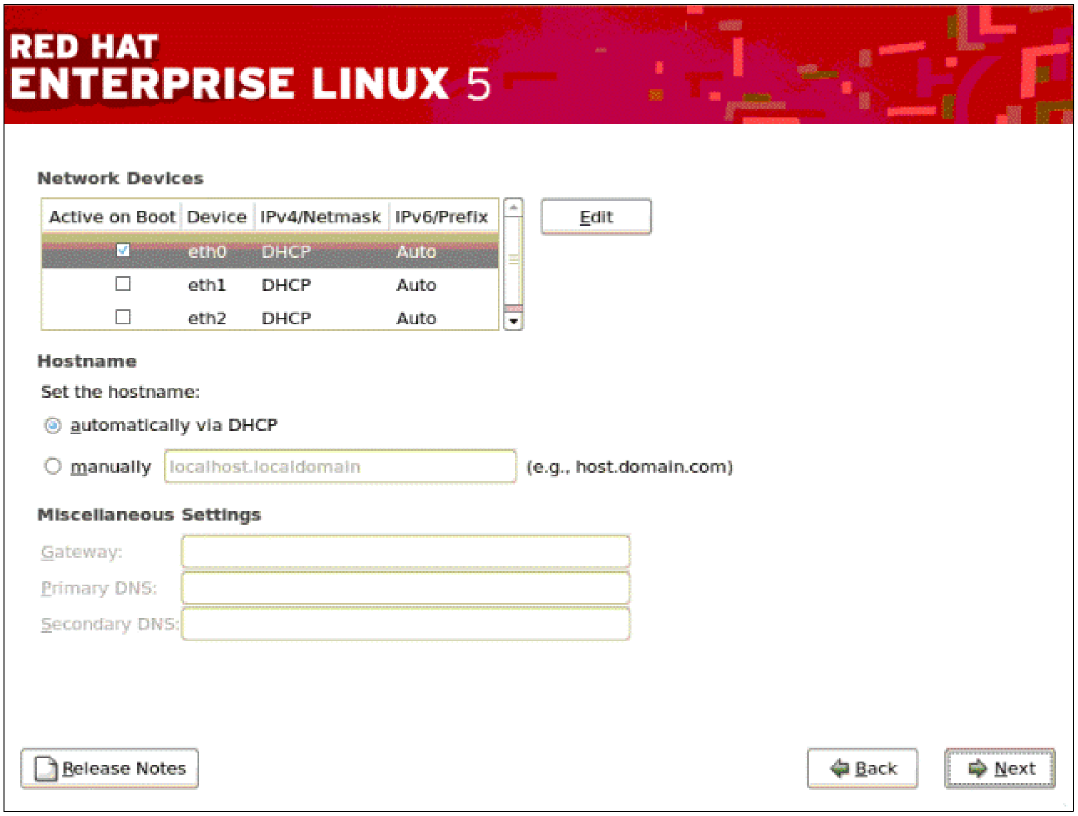


Figure 5-115 NiC settings

10. Select **TimeZone** and click **Next**. Set the *password* for the *root* user and click **Next**, as shown in Figure 5-116.

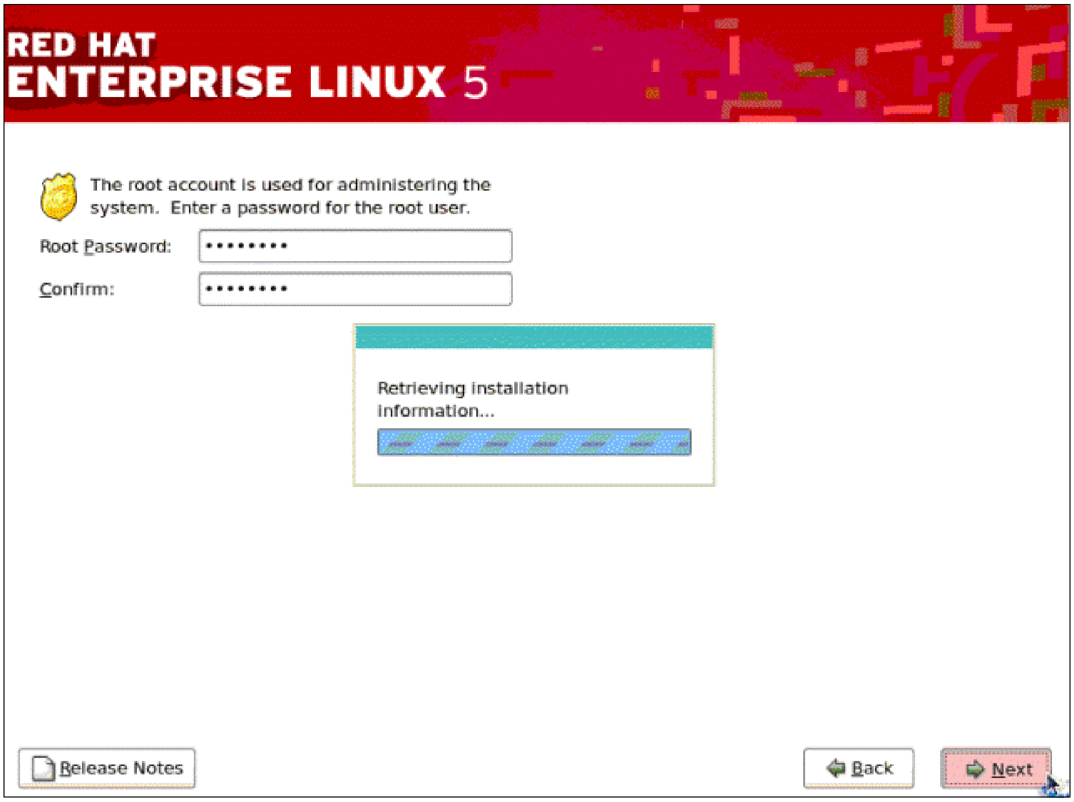


Figure 5-116 Set password

11. You can include additional packages to be installed or you can customize later, as shown in Figure 5-117.

Figure 5-117 Select (optional) packages

12. Click **Next** to start the installation on the SAN Boot volume, as shown in Figure 5-118.

Figure 5-118 Install on boot volume

13. After completing installation, you are prompted to reboot the server, as shown in Figure 5-119.

Figure 5-119 Installation complete

14. The server will boot off of the SAN. You can proceed configuring the rest of OS installation steps, as shown in Figure 5-120.

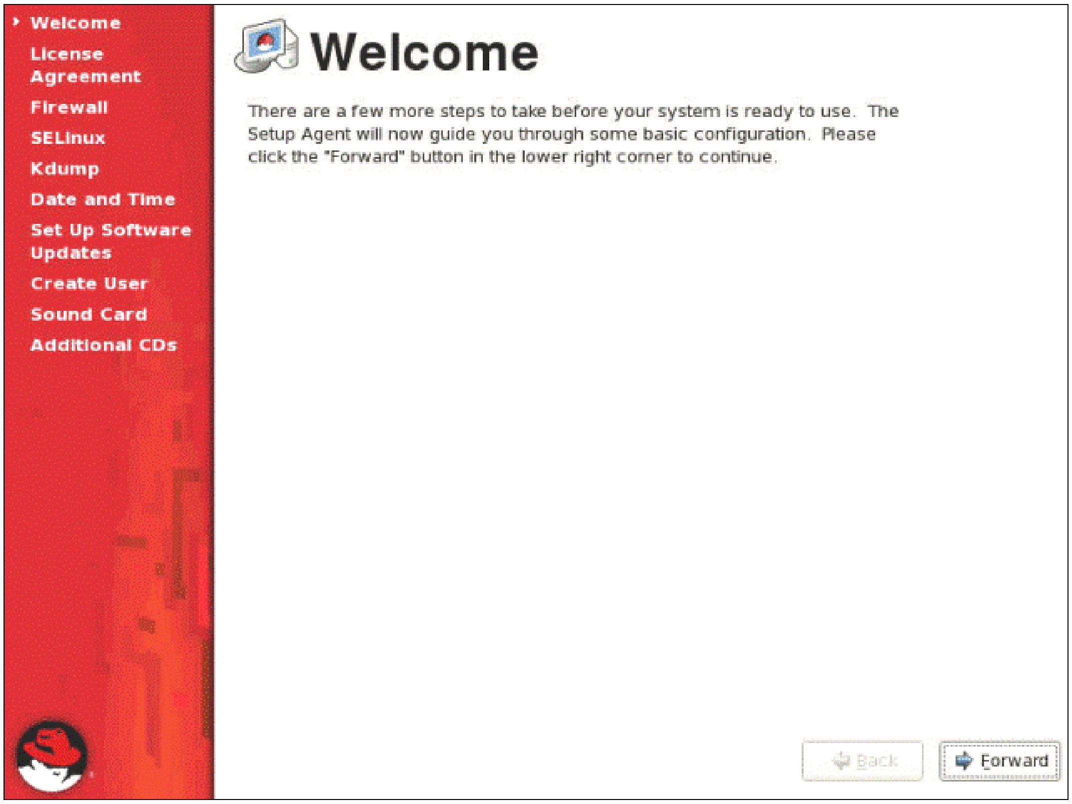


Figure 5-120 Welcome panel

15. After completing installation, to check if the multipath grouped the disks and was configured as planned, you can use the **cat** command to concatenate and display the `/etc/multipath.conf` file, as shown in Example 5-2.

Example 5-2 Multipath.conf and checking the multipath drivers.

```
[root@rb1SAN Boot ~]# cat /etc/multipath.conf
defaults {
        user_friendly_names yes
}

blacklist {
        devnode "^(ram|raw|loop|fd|md|dm-|sr|scd|st)[0-9]*"
        devnode "^(hd|xvd|vd)[a-z]*"
        wwid "*"
}

# Make sure our multipath devices are enabled.

blacklist_exceptions {
        wwid "500507680140813e1000000000000000e"
}
```

```
[root@rb1SAN Boot ~]# multipath -ll
mpath0 (500507680120813e1000000000000000e) dm-0 IBM,2145
[size=40G][features=1 queue_if_no_path][hwhandler=0][rw]
\_ round-robin 0 [prio=2][active]
 \_ 0:0:1:0 sda 8:0   [active][ready]
 \_ 1:0:1:0 sdb 8:16  [active][ready]
[root@rb1SAN Boot ~]#
```

5.9.3 Multipath I/O

Starting with SUSE Linux Enterprise Server 10 and RHEL5, IBM has transitioned Linux multipath driver support from SDD to the use of Linux-native Device Mapper Multipath module (DM-MP / DM-MPIO).

Linux Clustering is not supported, and Linux OS does not use the legacy reserve function, therefore, there are no persistent reserves used in Linux.

Device Mapper Multipath (DM-MPIO)

Device mapper is a block subsystem that provides layering mechanism for block devices. Device mapper can be written to provide a specific functionality on top of a block device. It also provides a consistent user interface for storage devices provided by multiple vendors.

There is only one block device (/dev/mapper/XXX) for a LUN and it is the device created by device mapper. Paths are grouped into priority groups, and one of the priority group will be used for I/O, and is called active. A path selector selects a path in the priority group to be used for an I/O based on some load balancing algorithm (for example round-robin).

When an I/O fails in a path, that path gets disabled and the I/O is re-tried in a different path in the same priority group. If all paths in a priority group fails, a different priority group which is enabled will be selected to send I/O.

Device Mapper consists of four components:

1. **dm-multipath:** This module is responsible for the multipathing decisions in event of a failure.
2. **multipath:** This command is used to detect multiple devices for failover, to list the available paths and to remove virtual devices.
3. **multipathd daemon:** This daemon constantly checks the paths to mark them as failed when it finds a path is faulty and switches the I/O to the next enable priority group. It keeps checking the failed path, after the failed path comes alive, based on the failback policy, it can activate the path.
4. **multipathd command:** When used with the **-k** parameter, it is possible to run a command line interface as shown in Example 5-3.

Example 5-3 multipathd command line

```
# multipathd -k
multipathd> help
multipath-tools v0.4.8 (08/02, 2007)
CLI commands reference:
list|show paths
list|show paths format $format
list|show status
list|show maps|multipaths
list|show maps|multipaths status
list|show maps|multipaths stats
list|show maps|multipaths topology
list|show topology
list|show map|multipath $map topology
Linux Config.fm Draft Document for Review March 28, 2011 12:24 pm
636 IBM System Storage DS3500: Introduction and Implementation Guide
list|show config
list|show blacklist
list|show devices
list|show wildcards
add path $path
remove|del path $path
add map|multipath $map
remove|del map|multipath $map
switch|switchgroup map|multipath $map group $group
reconfigure
suspend map|multipath $map
resume map|multipath $map
resize map|multipath $map
reload map|multipath $map
reinstate path $path
fail path $path
multipathd>
```

Installing the DMM multipath driver

Follow these steps to install the DMM multipath driver:

1. Set HBA timeout by modifying the `/etc/modprobe.conf.local` file. Add the following lines to it:

   `options lpfc_nodev_tmo 10`

   `options qla2xxx qlport_down_retry=10`

2. Stop the multipathing daemon with the service **`multipathd stop`** command if the daemon is already running.

3. Setup `multipath.conf`. It is very important to set up this file as it is consulted by the multipath command for the configuration of multipath devices. It is essentially the configuration file for devices virtualized by device mapper.

   The file has four sections:

   a. *Defaults:* Specifies system level default override.

   b. *Blacklist*: Specifies a list of devices to not be multipathed.

   c. *Devices:* Specifies the list of devices that the setting will be applied to.

   d. *Multipaths:* Specifies configuration settings for tuning

Create the multipath configuration file, `/etc/multipath.conf`. Add entry for the Storwize V7000 system in the *devices* section, as shown in Example 5-4.

Example 5-4 multipath.conf file sample

```
multipaths {
           multipath {
           wwid                   360080e50001b0c90000007504c6e060a
           alias                  mpath0
           path_grouping_policy multibus
           path_checker readsector0
           path_selector "round-robin 0"
           failback      "5"
           rr_weight priorities
           no_path_retry "5"
           }
}
devices {
        device {
        vendor "IBM"
        product "2076"
        path_grouping_policy group_by_prio
        prio_callout "/sbin/mpath_prio_alua/dev/%n"
        }
}
```

Tip: Comment out the *default blacklist* if it is listed in the `multipath.conf` file.

4. Save the file.
5. Restart the multipathing daemon with the **service multipathd start** command.

Additional multipath commands

The following **`multipath`** commands are helpful when working with multipath topologies and configurations.

multipath –ll

This command shows the current multipath topology from all available information, such as sysfs, the device mapper, or path checkers. When a new device is added to a system that is running multipathing, run the multipath **-v2** command so that the device can be monitored by the multipathing software.

multipathd –k

Puts your session into interactive mode. Commands can be used to list the current configuration or to change the configuration. For example, `<show config>` lists the current configuration, and `<reconfigure>` gets the latest configuration from the configuration file (`etc/multipath.conf`).

5.10 SAN Boot for AIX

You can SAN Boot the operating system over Fibre Channel from a IBM Storwize v7000 system volume if your hosts use AIX operating system version 5.2 or later. IBM System p hosts take significant advantage of SAN Boot in achieving the following capabilities:

- Better I/O performance due to caching and striping across multiple spindles,
- Ability to redeploy disk space when a server is retired from service
- Option to use FlashCopy to capture a rootvg backup
- Option to move an AIX image from one physical server/LPAR to another.

In our example, we have used the *IBM AIX version 6.1* operating system. The following pages explain steps to boot IBM System p hosts from a boot volume on IBM Storwize V7000 storage system.

> **Tip:** Before setting up SAN Boot for AIX, review the information in 5.3.1, "Considerations before setting up SAN Boot" on page 286.

To implement SAN Boot on a new system, you can:

- Use the Network Installation Manager (NIM).
- Start the AIX installation from a bootable AIX CD install package.

5.10.1 AIX SAN installation with NIM

Network Installation Manager (NIM) is a client server infrastructure and service that allows remote installation of the operating system, manages software updates, and can be configured to install and update third-party applications. Although both the NIM server and client file sets are part of the operating system, a separate NIM server must be configured that will keep the configuration data and the installable product file sets.

We assume that the following preparations have been completed:

- The NIM environment is deployed and all of the necessary configuration on the NIM master is already done.
- The NIM server is properly configured as the NIM master, and the basic NIM resources have been defined.
- The target machine (NIM client) currently has no operating system installed and is configured to boot from the NIM server.

For more information about how to configure a NIM server, see *NIM from A to Z in AIX 5L*, SG24-7296.

You can modify the `bosinst.data` file prior to starting the installation. `bosinst.data` is where the installation control is stored. Insert the appropriate values at the following line:

```
SAN_DISKID = <worldwide_portname//lun_id>
```

The (World Wide Port Name) and (Logical Unit ID) are each in the format returned by the **`lsattr`** command, that is, "0x" followed by 1-16 hexadecimal digits. For example:

```
SAN_DISKID = 0x0123456789FEDCBA//0x2000000000000
```

Perform the following procedure to install the NIM:

1. Enter the command:

   # smit nim_bosinst

2. Select the **lpp_source** and **SPOT** resource for the BOS installation.
3. Select the **BOSINST_DATA** to use during installation option. Next, select a **bosinst_data** resource that is capable of performing a unprompted BOS installation.
4. Select the **RESOLV_CONF** to use for network configuration option, and select a **resolv_conf** resource.
5. Accept the default values for the remaining menu options. Press Enter to confirm and begin the NIM client installation.
6. To check the status of the NIM client installation, for example, where client name is “va_v7k”, enter the command:

   # lsnim -l va_v7k

7. Proceed using the **SMS** menu to boot and configure your LPAR to access the NIM Master Server. Choose **1** to go to SMS main menu, as shown in Figure 5-121.

```
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM

1 = SMS Menu 5 = Default Boot List
8 = Open Firmware Prompt 6 = Stored Boot List
§
```

Figure 5-121 SMS main menu

8. Select 2 for IPL Configuration Menu as shown in Figure 5-122.

```
PowerPC Firmware
Version EL340_075
SMS 1.7 (c) Copyright IBM Corp. 2000,2008 All rights reserved.
-------------------------------------------------------------------------------
Main Menu
1.   Select Language
2.   Setup Remote IPL (Initial Program Load)
3.   Change SCSI Settings
4.   Select Console
5.   Select Boot Options








-------------------------------------------------------------------------------
Navigation Keys:

                                             X = eXit System Management Services
-------------------------------------------------------------------------------
Type menu item number and press Enter or select Navigation key:5
```

Figure 5-122 Setup Remote IPL

9. Select the NIC card cabled already to access the NIM Master. Select **1**, as shown in Figure 5-123.

```
PowerPC Firmware
Version SF220_001
SMS 1.5 (c) Copyright IBM Corp. 2000, 2003  All rights reserved.
-------------------------------------------------------------------------------
NIC Adapters
Device Slot Hardware Address
1. 10/100 Mbps Ethernet PCI Adapt Integ:U0.1-P1/E2 0002554f5c46
2. IBM 10/100/1000 Base-TX PCI-X 4:U0.1-P2-I4/E1 00145eb7f39d
-------------------------------------------------------------------------------
Navigation keys:
M = return to Main Menu
ESC key = return to previous screen X = eXit System Management Services
-------------------------------------------------------------------------------
Type the number of the menu item and press Enter or select Navigation Key: 1§
```

Figure 5-123 NIC card list for remote IPL

10.Select **1** to configure IP parameters, as shown in Figure 5-124.

```
PowerPC Firmware
Version SF220_001
SMS 1.5 (c) Copyright IBM Corp. 2000, 2003  All rights reserved.
-----------------------------------------------------------------------------
Network Parameters
IBM 10/100/1000 Base-TX PCI-X Adapter: U0.1-P2-I4/E1
1. IP Parameters
2. Adapter Configuration
3. Ping Test
-----------------------------------------------------------------------------
Navigation keys:
M = return to Main Menu
ESC key = return to previous screen X = eXit System Management Services
-----------------------------------------------------------------------------
Type the number of the menu item and press Enter or select Navigation Key: 1§
```

Figure 5-124 Configure IP parameters

11.Configure the IP defined for this partition in the */etc/hosts* of NIM Master, IP of the NIM Master server, the gateway, and the subnet mask, as shown in Figure 5-125.

```
PowerPC Firmware
Version SF220_001
SMS 1.5 (c) Copyright IBM Corp. 2000, 2003  All rights reserved.
-----------------------------------------------------------------------------
IP Parameters
IBM 10/100/1000 Base-TX PCI-X Adapter: U0.1-P2-I4/E1
1. Client IP Address [9.3.58.217]
2. Server IP Address [9.3.58.194]
3. Gateway IP Address [9.3.58.194]
4. Subnet Mask [255.255.255.000]
-----------------------------------------------------------------------------
Navigation keys:
M = return to Main Menu
ESC key = return to previous screen X = eXit System Management Services
-----------------------------------------------------------------------------
Type the number of the menu item and press Enter or select Navigation Key:§
```

Figure 5-125 IP parameters

12. After the correct IP parameters have been entered, return to the *Main Menu* by pressing **M**, as shown in Figure 5-126.

```
PowerPC Firmware
Version EL340_075
SMS 1.7 (c) Copyright IBM Corp. 2000,2008 All rights reserved.
-------------------------------------------------------------------------------
Main Menu
1.   Select Language
2.   Setup Remote IPL (Initial Program Load)
3.   Change SCSI Settings
4.   Select Console
5.   Select Boot Options










-------------------------------------------------------------------------------
Navigation Keys:

                                          X = eXit System Management Services
-------------------------------------------------------------------------------
Type menu item number and press Enter or select Navigation key:5
```

Figure 5-126 Main Menu

13. In the main menu, select **5** (*Select Boot Option*), as shown in Figure 5-127.

```
PowerPC Firmware
Version EL340_075
SMS 1.7 (c) Copyright IBM Corp. 2000,2008 All rights reserved.
-------------------------------------------------------------------------------
Multiboot
1.   Select Install/Boot Device
2.   Configure Boot Device Order
3.   Multiboot Startup <OFF>











-------------------------------------------------------------------------------
Navigation keys:
M = return to Main Menu
ESC key = return to previous screen         X = eXit System Management Services
-------------------------------------------------------------------------------
Type menu item number and press Enter or select Navigation key:1
```

Figure 5-127 Multiboot menu

14. Select **6** (*Network*) to define your boot by Network, as shown in Figure 5-128.

```
PowerPC Firmware
Version EL340_075
SMS 1.7 (c) Copyright IBM Corp. 2000,2008 All rights reserved.
-------------------------------------------------------------------------------
Select Device Type
1.   Diskette
2.   Tape
3.   CD/DVD
4.   IDE
5.   Hard Drive
6.   Network
7.   List all Devices







-------------------------------------------------------------------------------
Navigation keys:
M = return to Main Menu
ESC key = return to previous screen         X = eXit System Management Services
-------------------------------------------------------------------------------
Type menu item number and press Enter or select Navigation key:6
```

Figure 5-128 Select Network

15. Select the NIC card to boot from. Press **2**, as shown in Figure 5-129.

```
PowerPC Firmware
Version EL340_075
SMS 1.7 (c) Copyright IBM Corp. 2000,2008 All rights reserved.
-------------------------------------------------------------------------------
Select Device
Device  Current  Device
Number  Position  Name
1.        -       PORT - 1 IBM Host Ethernet Adapter
       ( loc=U789C.001.DQDU764-P1-C7-T1    )
2.        -       Port 1 - IBM 2 PORT 10/100/1000 Base-TX PCI-X Adapter
       ( loc=U789C.001.DQDU764-P1-C4-T1 )
3.        -       Port 2 - IBM 2 PORT 10/100/1000 Base-TX PCI-X Adapter
       ( loc=U789C.001.DQDU764-P1-C4-T2 )
4.        -       SAS 136 GB Harddisk, part=2 (AIX 6.1.0)
       ( loc=U789C.001.DQDU764-P2-D5 )
5.        -       SAS Tape
       ( loc=U789C.001.DQDU764-P2-D1 )
6.        -       SATA CD-ROM
       ( loc=U789C.001.DQDU764-P2-D2 )
-------------------------------------------------------------------------------
Navigation keys:
M = return to Main Menu      N = Next page of list
ESC key = return to previous screen         X = eXit System Management Services
-------------------------------------------------------------------------------
Type menu item number and press Enter or select Navigation key:2
```

Figure 5-129 Select NIC

16.Select **2** for *Normal Mode Boot*, as shown in Figure 5-130.

```
PowerPC Firmware
Version EL340_075
SMS 1.7 (c) Copyright IBM Corp. 2000,2008 All rights reserved.
-------------------------------------------------------------------------------
Select Task

Port 1 - IBM 2 PORT 10/100/1000 Base-TX PCI-X Adapter
    ( loc=U789C.001.DQDU764-P1-C4-T1 )

1.   Information
2.   Normal Mode Boot
3.   Service Mode Boot









-------------------------------------------------------------------------------
Navigation keys:
M = return to Main Menu
ESC key = return to previous screen         X = eXit System Management Services
-------------------------------------------------------------------------------
Type menu item number and press Enter or select Navigation key:2
```

Figure 5-130 Boot Mode

17.Select **Yes** to proceed with the *normal installation* process, as shown in Figure 5-131.

```
PowerPC Firmware
Version EL340_075
SMS 1.7 (c) Copyright IBM Corp. 2000,2008 All rights reserved.
-------------------------------------------------------------------------------
Are you sure you want to exit System Management Services?
1.   Yes
2.   No











-------------------------------------------------------------------------------
Navigation Keys:

                                            X = eXit System Management Services
-------------------------------------------------------------------------------
Type menu item number and press Enter or select Navigation key:1
```

Figure 5-131 exit SMS for normal mode installation

18.Follow the steps in the Normal Mode installation process, which is detailed in 5.10.3, “Normal Mode installation process (common in NIM and HMC methods)” on page 385.

5.10.2 Installing AIX 6.1 using HMC interface

Verify that the HMC is attached to the managed system, then perform the following steps:

1. Using the HMC, create a partition and partition profile for the client.
2. Assign the SCSI bus controller attached to the media device, a network adapter, and enough disk space for the AIX operating system to the partition.
3. Set the boot mode for the partition to be SMS mode.
4. After you have successfully created the partition and partition profile, leave the partition in the Ready state.

   The following steps are for performing a new base operating system installation on a logical partition using the partition's media device.

5. Activate the (new) partition allocated for SAN Boot, by performing these steps:
   a. Insert the AIX Volume 1 installation media into the media device of the managed system.
   b. Verify that the DVD is assigned to the current partition.
   c. Right-click the partition to open the menu.
   d. Select **Activate**. The *Activate Partition* menu opens with a selection of partition profiles. Highlight the desired Profile.
6. Select *Open Terminal* window or *Console Session* at the bottom of the menu to open a virtual terminal (vterm) window.
7. Select **Advanced** to open the *Advanced Options* menu.
8. For the *Boot* mode, select **SMS (**system management services).
9. Select **OK** to close the *Advanced Options* menu. Select **OK** again to open a *vterm* window for the partition. Figure 5-132 shows the Main Menu.

```
PowerPC Firmware
Version EL340_075
SMS 1.7 (c) Copyright IBM Corp. 2000,2008 All rights reserved.
-------------------------------------------------------------------------------
Main Menu
1.   Select Language
2.   Setup Remote IPL (Initial Program Load)
3.   Change SCSI Settings
4.   Select Console
5.   Select Boot Options









-------------------------------------------------------------------------------
Navigation Keys:

                                            X = eXit System Management Services
-------------------------------------------------------------------------------
Type menu item number and press Enter or select Navigation key:5
```

Figure 5-132 Main Menu

10. Press **5** key to *Select Boot Options*. Press **Enter**. The Multiboot menu is shown, as in Figure 5-133.

```
PowerPC Firmware
Version EL340_075
SMS 1.7 (c) Copyright IBM Corp. 2000,2008 All rights reserved.
-------------------------------------------------------------------------------
Multiboot
1.   Select Install/Boot Device
2.   Configure Boot Device Order
3.   Multiboot Startup <OFF>

-------------------------------------------------------------------------------
Navigation keys:
M = return to Main Menu
ESC key = return to previous screen         X = eXit System Management Services
-------------------------------------------------------------------------------
Type menu item number and press Enter or select Navigation key:1
```

Figure 5-133 Select Boot Device

11. Press **1** key to *Select Install/Boot Device*. Press **Enter**. The Select Device Type menu is shown, as in Figure 5-134.

```
PowerPC Firmware
Version EL340_075
SMS 1.7 (c) Copyright IBM Corp. 2000,2008 All rights reserved.
-------------------------------------------------------------------------------
Select Device Type
1.   Diskette
2.   Tape
3.   CD/DVD
4.   IDE
5.   Hard Drive
6.   Network
7.   List all Devices

-------------------------------------------------------------------------------
Navigation keys:
M = return to Main Menu
ESC key = return to previous screen         X = eXit System Management Services
-------------------------------------------------------------------------------
Type menu item number and press Enter or select Navigation key:3
```

Figure 5-134 Select Device

12.Now specify the target. Select **7** to *List all Devices*, as shown in Figure 5-135.

```
scan /pci@800000020000200/pci1014,02BD@1/sata/disk
check /pci@800000020000200/pci1014,02BD@1/sata/disk@30000
scan /pci@800000020000204/fibre-channel@0/fp
scan /pci@800000020000204/fibre-channel@0/disk
check /pci@800000020000204/fibre-channel@0/disk@5001738000d00142,1000000000000
check /pci@800000020000204/fibre-channel@0/disk@5005076801405d1a
check /pci@800000020000204/fibre-channel@0/disk@5005076801405d1a,1000000000000
check /pci@800000020000204/fibre-channel@0/disk@5005076801405d1a,2000000000000
check /pci@800000020000204/fibre-channel@0/disk@5005076801405d1a,3000000000000
check /pci@800000020000204/fibre-channel@0/disk@5005076801405d1a,4000000000000
check /pci@800000020000204/fibre-channel@0/disk@5005076801405d1a,5000000000000
check /pci@800000020000204/fibre-channel@0/disk@5005076801405d1a,6000000000000
check /pci@800000020000204/fibre-channel@0/disk@5005076801405d1a,7000000000000
```

Figure 5-135 Scan Devices

13.Select SATA CD-ROM from the list. In our case, press **6**, as shown in Figure 5-136.

```
PowerPC Firmware
Version EL340_075
SMS 1.7 (c) Copyright IBM Corp. 2000,2008 All rights reserved.
-------------------------------------------------------------------------------
Select Device
Device  Current  Device
Number  Position  Name
1.        -       PORT - 1 IBM Host Ethernet Adapter
       ( loc=U789C.001.DQDU764-P1-C7-T1    )
2.        -       Port 1 - IBM 2 PORT 10/100/1000 Base-TX PCI-X Adapter
       ( loc=U789C.001.DQDU764-P1-C4-T1 )
3.        -       Port 2 - IBM 2 PORT 10/100/1000 Base-TX PCI-X Adapter
       ( loc=U789C.001.DQDU764-P1-C4-T2 )
4.        -       SAS 136 GB Harddisk, part=2 (AIX 6.1.0)
       ( loc=U789C.001.DQDU764-P2-D5 )
5.        -       SAS Tape
       ( loc=U789C.001.DQDU764-P2-D1 )
6.        -       SATA CD-ROM
       ( loc=U789C.001.DQDU764-P2-D2 )
-------------------------------------------------------------------------------
Navigation keys:
M = return to Main Menu       N = Next page of list
ESC key = return to previous screen         X = eXit System Management Services
-------------------------------------------------------------------------------
Type menu item number and press Enter or select Navigation key:6
```

Figure 5-136 Select CD-ROM

14. Proceed with *Normal Boot*, as shown in Figure 5-137.

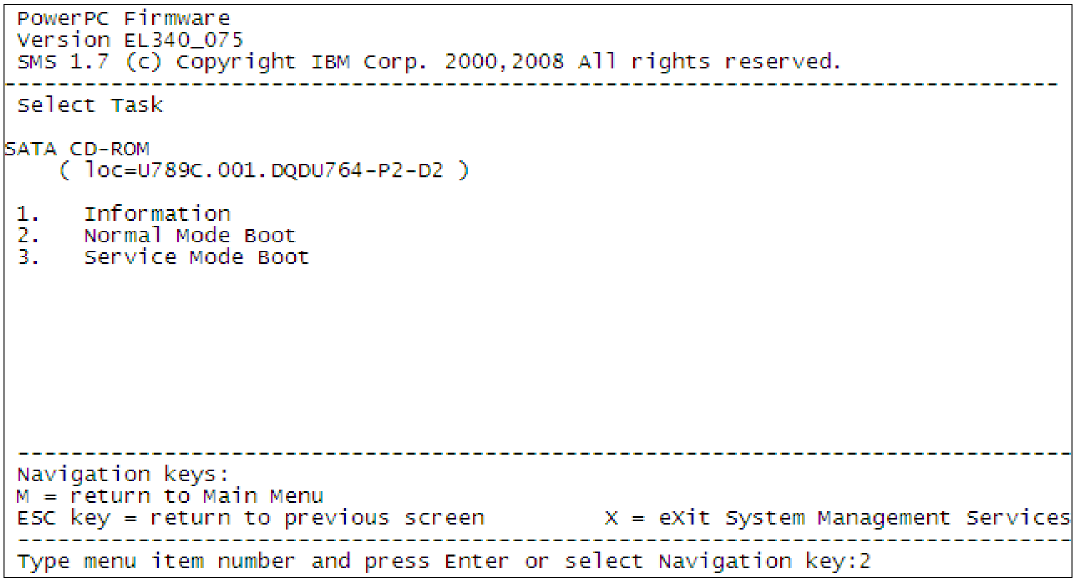


Figure 5-137 Normal Boot with CD-ROM

15. Select **Yes** to proceed with the normal installation process, as shown in Figure 5-138.

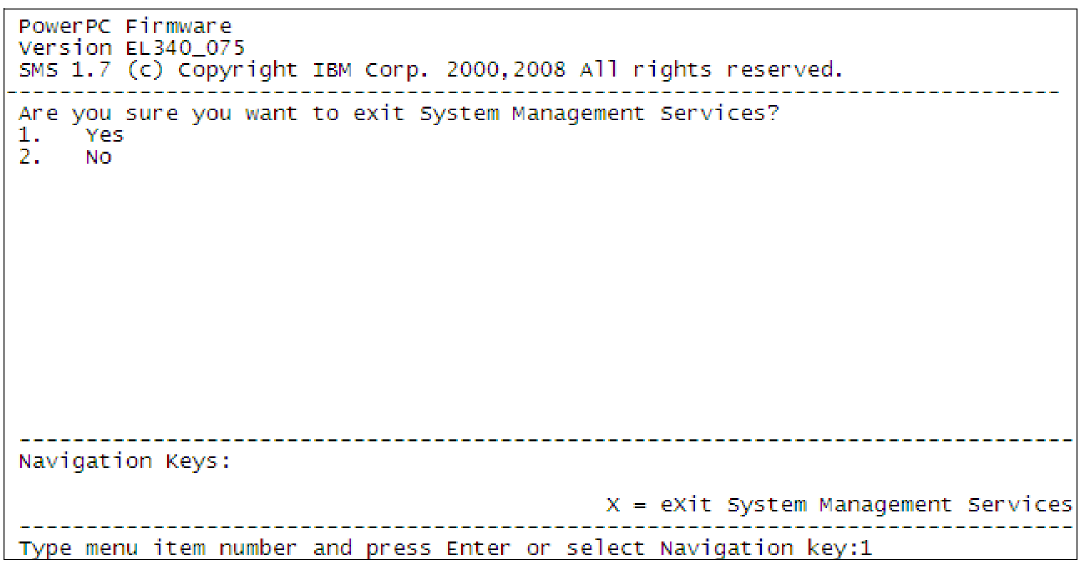


Figure 5-138 Exit SMS for installation

16. Follow the steps in Normal mode installation process, detailed at 5.10.3, “Normal Mode installation process (common in NIM and HMC methods)” on page 385

5.10.3 Normal Mode installation process (common in NIM and HMC methods)

Follow these steps to perform a Normal Mode installation:

1. After choosing *Normal Mode Boot* from the *System Management Services (SMS)*, the AIX software will load, as shown in Example 5-139.

```
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM                             IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM     STARTING SOFTWARE       IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM        PLEASE WAIT...       IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM                             IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM
IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM IBM

Elapsed time since release of system processors: 196201 mins 2 secs
```

Figure 5-139   Starting software installation

2. Press **1** to choose a terminal as System Console, as shown in Figure 5-140.

```
-------------------------------------------------------------------------------
                                Welcome to AIX.
                       boot image timestamp: 21:10 05/20
                 The current time and date: 23:45:31 04/14/2011
       processor count: 4;  memory size: 31360MB;  kernel size: 26369758
boot device:
/pci@800000020000200/pci1014,02BD@1/sata/disk@30000:\ppc\chrp\bootfile.exe
                       kernel debugger setting: enabled
-------------------------------------------------------------------------------

AIX Version 6.1
Starting NODE#000 physical CPU#001 as logical CPU#001... done.
Starting NODE#000 physical CPU#002 as logical CPU#002... done.
Starting NODE#000 physical CPU#003 as logical CPU#003... done.
Starting NODE#001 physical CPU#004 as logical CPU#004... done.
Starting NODE#001 physical CPU#005 as logical CPU#005... done.
Starting NODE#001 physical CPU#006 as logical CPU#006... done.
Starting NODE#001 physical CPU#007 as logical CPU#007... done.
Preserving 17285 bytes of symbol table [/usr/lib/drivers/pci/efcdd]
Preserving 31888 bytes of symbol table [/etc/drivers/pci/efcddpin]
Preserving 21033 bytes of symbol table [/usr/lib/drivers/efscsidd]
Preserving 22508 bytes of symbol table [/etc/drivers/efscsiddpin]
Preserving 25668 bytes of symbol table [/usr/lib/drivers/scsidisk]
Preserving 37820 bytes of symbol table [/etc/drivers/scsidiskpin]
Preserving 126407 bytes of symbol table [/usr/lib/drivers/hd_pin]
Preserving 200724 bytes of symbol table [/usr/lib/drivers/hd_pin_bot]


 ******* Please define the System Console. *******

Type a 1 and press Enter to use this terminal as the system console.
```

Figure 5-140   Use terminal as system console

3. Press **2** to *Change/Show Installation Settings and Install*, as shown in Figure 5-141.

```
                      Welcome to Base Operating System
                      Installation and Maintenance

Type the number of your choice and press Enter.  Choice is indicated by >>>.

>>> 1 Start Install Now with Default Settings

    2 Change/Show Installation Settings and Install

    3 Start Maintenance Mode for System Recovery

    4 Configure Network Disks (iSCSI)





    88  Help ?
    99  Previous Menu

>>> Choice [1]:      2
```

Figure 5-141 Install with default settings

4. The next panel (Figure 5-142) shows the current *System Settings*. Press **1** to change the settings for SAN Boot.

```
                          Installation and Settings

Either type 0 and press Enter to install with current settings, or type the
number of the setting you want to change and press Enter.

    1   System Settings:
          Method of Installation.............Preservation
          Disk Where You Want to Install.....hdisk0

    2   Primary Language Environment Settings (AFTER Install):
          Cultural Convention................English (United States)
          Language ..........................English (United States)
          Keyboard ..........................English (United States)
          Keyboard Type......................Default
    3   Security Model.......................Default
    4   More Options  (Software install options)

>>> 0   Install with the current settings listed above.

                       +-----------------------------------------------------
    88  Help ?         |    WARNING: Base Operating System Installation will
    99  Previous Menu  |    destroy or impair recovery of SOME data on the
                       |    destination disk hdisk0.
>>> Choice [0]:      1
```

Figure 5-142 Current System Settings

5. Select **1** for new installation, as shown in Figure 5-143.

```
                        Change Method of Installation

  Type the number of the installation method and press Enter.

    1 New and Complete Overwrite
      Overwrites EVERYTHING on the disk selected for installation.
      Warning: Only use this method if the disk is totally empty or if there
      is nothing on the disk you want to preserve.

>>> 2 Preservation Install
      Preserves SOME of the existing data on the disk selected for
      installation.  Warning: This method overwrites the usr (/usr),
      variable (/var), temporary (/tmp), and root (/) file systems.  Other
      product (applications) files and configuration data will be destroyed.




    88  Help ?
    99  Previous Menu

>>> Choice [2]:     1
```

Figure 5-143 New installation

6. Now you can change the target device for installation. For SAN Boot, you choose the volume allocated on the storage subsystem (refer to 5.4.5, "Creating a volume for SAN Boot" on page 298). Figure 5-144 shows the disk selection menu.

> **Tip:** On AIX 6.1 platforms, AIX host imposes the size limitation of 2 TB on disk volume sizes.

```
                 Change Disk(s) Where You Want to Install

Type one or more numbers for the disk(s) to be used for installation and press
Enter. To cancel a choice, type the corresponding number and Press Enter.
At least one bootable disk must be selected. The current choice is indicated
by >>>.

        Name      Physical Volume Identifier

     1  hdisk0   00cd5bc41a85f714
     2  hdisk18  00cd5bc44aa6664e
     3  hdisk1   00cd5bc48c49f4a8
     4  hdisk2   00cd5bc48c49f1d7
     5  hdisk3   00cd5bc48c49f346
    06  MORE CHOICES...

>>>  0   Continue with choices indicated above
    55  More Disk Options
    66  Devices not known to Base Operating System Installation
    77  Display More Disk Information
    88  Help ?
    99  Previous Menu

>>> Choice [0]:     6
```

Figure 5-144 Change Disk

7. If you do not see the disk that you allocated on SAN, select **6** for more disk choices until you see the hdisk matching the size of the disk allocated on the storage for SAN Boot. In our case, it is *hdisk17*, as shown in Figure 5-145.

```
                 Change Disk(s) Where You Want to Install

Type one or more numbers for the disk(s) to be used for installation and press
Enter. To cancel a choice, type the corresponding number and Press Enter.
At least one bootable disk must be selected. The current choice is indicated
by >>>.

        Name      Location Code    Size(MB)  VG Status   Bootable

>>> 13  ...PREVIOUS CHOICES
    14  hdisk9   03-00-01          10240   none            Yes    No
    15  hdisk10  03-00-01          10240   none            Yes    No
>>> 16  hdisk17  03-00-01          40960   none            Yes    No
    17  hdisk11  03-00-01          20480   other vg        Yes    No
    18  MORE CHOICES...

     0   Continue with choices indicated above
    55  More Disk Options
    66  Devices not known to Base Operating System Installation
    77  Display More Disk Information
    88  Help ?
    99  Previous Menu

>>> Choice [13]:      7 7
```

Figure 5-145 Choose boot LUN (defined on SAN)

8. Type **77** to get more details on the hdisk. Pressing **77** again will display the Identifier of the hdisk which correlates with the *Volume ID* of the volume defined in Storwize v7000 GUI, as shown in Figure 5-146.

```
                 Change Disk(s) Where You Want to Install

Type one or more numbers for the disk(s) to be used for installation and press
Enter. To cancel a choice, type the corresponding number and Press Enter.
At least one bootable disk must be selected. The current choice is indicated
by >>>.

        Name      Device Adapter Connection Location
                  or Physical Location Code
>>> 13  ...PREVIOUS CHOICES
    14  hdisk9    ...-W50050768014052D1-L2000000000000
    15  hdisk10   ...-W50050768014052D1-L3000000000000
>>> 16  hdisk17   ...-W50050768014052D1-LA000000000000
    17  hdisk11   ...-W50050768014052D1-L4000000000000
    18  MORE CHOICES...

     0   Continue with choices indicated above
    55  More Disk Options
    66  Devices not known to Base Operating System Installation
    77  Display More Disk Information
    88  Help ?
    99  Previous Menu

>>> Choice [13]:      77
```

Figure 5-146 Corresponds to Volume ID in Storwize V7000 GUI

9. Press **0** to go with the selected disk. It will again display the current System setting. Press **1** to continue installing on the SAN disk, as shown in Figure 5-147.

```
                            Installation and Settings

Either type 0 and press Enter to install with current settings, or type the
number of the setting you want to change and press Enter.

    1  System Settings:
         Method of Installation.............New and Complete Overwrite
         Disk Where You Want to Install.....hdisk17

    2  Primary Language Environment Settings (AFTER Install):
         Cultural Convention................English (United States)
         Language ..........................English (United States)
         Keyboard ..........................English (United States)
         Keyboard Type......................Default
    3  Security Model.......................Default
    4  More Options  (Software install options)

>>> 0  Install with the current settings listed above.

                       +-----------------------------------------------------
    88  Help ?         |    WARNING: Base Operating System Installation will
    99  Previous Menu  |    destroy or impair recovery of ALL data on the
                       |    destination disk hdisk17.
>>> Choice [0]:      1
```

Figure 5-147   Settings with volume for SAN Boot selected

10. The next panel shows an installation summary, as shown in Figure 5-148.

```
                            Overwrite Installation Summary

Disks:  hdisk17
Cultural Convention:  en_US
Language:  en_US
Keyboard:  en_US
JFS2 File Systems Created:  Yes
Graphics Software:  Yes
System Management Client Software:  Yes
Enable System Backups to install any system:  Yes

Optional Software being installed:




>>> 1   Continue with Install
                       +-----------------------------------------------------
    88  Help ?         |    WARNING: Base Operating System Installation will
    99  Previous Menu  |    destroy or impair recovery of ALL data on the
                       |    destination disk hdisk17.
>>> Choice [1]:      1
```

Figure 5-148   Installation Summary

11. Press **1** to start the installation, as shown in Figure 5-149.

```
                          Installing Base Operating System



        Please wait...










        Approximate     Elapsed time
     % tasks complete   (in minutes)


                                                                0                0
                                                                             0
     0       Preparing target disks.

                                                                2                0
Making paging logical volumes.

                                                                3                1
Making logical volumes.

                                                                3                1
Making logical volumes.

                                                                3                1
Making logical volumes.
```

Figure 5-149   Installing Base OS

12. After going through automated installation phases, it will show the copyright information of the AIX software, then reboot the system, as shown in Figure 5-150.

```
 Licensed Materials - Property of IBM

 5765G6200
   Copyright International Business Machines Corp. 1985, 2008.
   Copyright AT&T 1984, 1985, 1986, 1987, 1988, 1989.
   Copyright Regents of the University of California 1980, 1982, 1983, 1985, 1986,
1987, 1988, 1989.
   Copyright BULL 1993, 2008.
   Copyright Digi International Inc. 1988-1993.
   Copyright Interactive Systems Corporation 1985, 1991.
   Copyright ISQUARE, Inc. 1990.
   Copyright Innovative Security Systems, Inc. 2001-2006.
   Copyright Mentat Inc. 1990, 1991.
   Copyright Open Software Foundation, Inc. 1989, 1994.
   Copyright Sun Microsystems, Inc. 1984, 1985, 1986, 1987, 1988, 1991.

 All rights reserved.
 US Government Users Restricted Rights - Use, duplication or disclosure
 restricted by GSA ADP Schedule Contract with IBM Corp.
Rebooting . . .
```

Figure 5-150   Installation Complete

13. After reboot, the host should boot from SAN by a single path between the host and the storage.

5.10.4 Multiple path I/O

A multipath driver is needed to prevent a single point-of-failure caused by the host bus adapter, Fibre Channel cable, or host-port(s) on the supported storage. In SAN Boot environment, having a proper multipathing setup becomes even more critical, as the OS resides on the external storage.

> **Path-control module (PCM):** A path-control module (PCM) provides the path management functions. A PCM can support one or more specific devices. An MPIO-capable device driver can control more than one type of target device, therefore, one device driver can be interfaced to multiple PCMs that control the I/O across the paths to each of the target devices.
>
> **Subsystem Device Driver Path Control Module (SDDPCM):** SDDPCM is a loadable path control module designed to support the multipath configuration environment. It provides a health check daemon to employ an automated method of reclaiming failed paths to a closed device.

After initial installation of AIX SAN Boot completes on a single-path SAN Boot environment, the supported storage boot device is configured as an MPIO device with AIX default PCM. You need to install SDDPCM driver for multipathing, as all supported storage MPIO SAN Boot devices are now configured with SDDPCM.

To enable SAN Boot with multipathing, perform the following procedure:

1. Verify that two paths exists between host and storage (cabling, zoning, and so on.)
2. In AIX, type the following command on the command line:

   ```
   lspath -l hdiskX
   ```

   Here, *X* is the logical number of the newly configured device. The command output should display one path for each adapter you installed and the status of each.
3. After verifying two paths from host to the storage, you need to install the supported SDDPCM packages. Install *SDDPCM Package for IBM Storwize V7000* from the following IBM site:

   http://www-01.ibm.com/support/docview.wss?uid=ssg1S4000201
4. Log in into the new installed server, check the space available on server and transfer the package for the selected/created directory.
5. Unpackage (untar) the package to be installed, then install the packages using the **installp** or **smitty installp** option.
6. Restart the system.

> **Tip:** If NIM installation is used and the SDDPCM software is included in the NIM repository environment, then the second restart is not required.

7. After the system finishes rebooting completely, check the path status, as shown in Example 5-5.

Example 5-5 Checking the paths status

```
# pcmpath query device 6

DEV#:   6  DEVICE NAME: hdisk6  TYPE: 2107900  ALGORITHM:  Load Balance
SERIAL: 13031810702
==========================================================================
Path#      Adapter/Path Name          State     Mode     Select     Errors
    0           fscsi0/path0           OPEN   NORMAL       7666          0
    1           fscsi2/path1           OPEN   NORMAL       9117          0
```

5.10.5 SDDPCM

The SDDPCM provides enhanced utilities (**pcmpath** commands) to show mappings from adapters, paths, devices, as well as performance and error statistics than can be useful in SAN management for those supported devices. It allows for dynamic selection of path selection algorithm options and provides for the ability to add paths to a device dynamically, dynamically remove or replace physical adapters, and support AIX and VIOS Fibre Channel dynamic device tracking.

The commands shown in Table 5-3 can help you manage MPIO paths:

Table 5-3 Commands to manage MPIO paths

Command	Description
mkpath	Adds a path to a target device
rmpath	Removes a path to a target device
chpath	Changes an attribute or the operational status of a path to a target device
lspath	Displays information about paths to a target device

Tip: The SDDPCM is an add-on software entity and has its own update strategy and process for obtaining fixes. Therefore, the customer must manage coexistence levels between both the mix of devices and operating system levels.

The SDDPCM default reserve policy is *no_reserve* and the default path selection algorithm is *load_balance*, which means that no reservation methodology for the device has been applied. With *no_reserve* policy, the device might be accessed by other initiators, and these initiators might be on other host systems.

You can check reservation policy and other parameters for the SAN Boot disk by running the **lsattr** command. See Figure 5-151.

```
# lsattr -El hdisk6
...
hcheck_interval 60                              Health Check Interval
...
queue_depth     20                              Queue DEPTH
reserve_policy  no_reserve                      Reserve Policy
...
#
```

Figure 5-151 checking policy and queue_depth parameters

If the SAN Boot devices are configured with the default reserve policy and path selection algorithm, you must not configure these SAN Boot devices on other AIX servers, because no scsi-2 reserve or scsi-3 persistent reserve is implemented on these SAN Boot devices to prevent access from another AIX server.

To find out more about AIX OS installation procedures in multipath SDD SAN Boot environment, refer to the *SDD User's Guide,* at this website:

http://www-01.IBM.com/support/docview.wss?uid=ssg1S1002675

5.11 More information

> **Tip:** For instructions about how to create a logical partition and partition profile, see the article, "Creating logical partitions and partition profiles" in *PowerVM Virtualization on IBM System p: Introduction and Configuration Fourth Edition*, SG24-7940.

For more information about the IBM Storwize v7000, see *Implementing the IBM Storwize V7000*, SG24-7938, and the following websites:

- Product site:

  http://www.IBM.com/systems/storage/disk/storwize_v7000/index.html

- Information center:

  http://publib.boulder.IBM.com/infocenter/storwize/ic/index.jsp

- Configuration limits and restrictions:

  http://www.IBM.com/support/docview.wss?uid=ssg1S1003702

- Techdocs library:

  http://www.IBM.com/support/techdocs/atsmastr.nsf/Web/Search

- Product manuals:

  http://www.IBM.com/support/docview.wss?uid=ssg1S7003318

6

iSCSI SAN Boot implementation with IBM System Storage DS5000

Storage area networks (SANs) have rapidly gained popularity over traditional client-server configurations due to their ability to efficiently handle data. Typically, SANs require Fibre Channel, a specialized "fabric" or communication backbone. However, Fibre Channel is much more expensive than the standard Internet Protocol (IP) backbone that exists in nearly every enterprise. *iSCSI*, a new standard for storage networks, encapsulates standard SCSI commands to communicate with disks and transmits them over the existing IP network.

Not only is iSCSI seen as a lower cost alternative to Fibre Channel, but it is also expected to outperform Fibre Channel as Ethernet speeds jump from 100 Mb/sec to 1 Gb and 10 Gb in the near future. In this chapter, we describe SAN Boot implementation using the IBM DS5000 series of storage subsystems with iSCSI.

© Copyright IBM Corp. 2012. All rights reserved.

6.1 iSCSI overview

The Internet Small Computer System Interface (iSCSI) is an IP-based standard for transferring data that supports host access by carrying SCSI commands over IP networks.

iSCSI is a block-level protocol that encapsulates SCSI commands into TCP/IP packets and, thereby, uses an existing IP network instead of requiring FC HBAs and SAN fabric infrastructure.

By carrying SCSI commands over IP networks, iSCSI is used to facilitate data transfers over intranets and to manage storage over long distances. iSCSI can be used to transmit data over local area networks (LANs), wide area networks (WANs), or the Internet, and can enable location-independent data storage and retrieval.

An iSCSI session establishes a TCP relationship between an iSCSI initiator node port and an iSCSI target node port. After being established, iSCSI control, data, and status messages are communicated over the session.

iSCSI uses TCP/IP (typically TCP ports 860 and 3260). In essence, iSCSI simply allows two hosts to negotiate and then exchange SCSI commands using IP networks. Thus, iSCSI takes a popular high-performance local storage bus and emulates it over wide area networks (WANs), creating a storage area network (SAN).

Unlike some SAN protocols, iSCSI requires no dedicated cabling; it can be run over existing switching and IP infrastructure. However, the performance of an iSCSI SAN deployment can be severely degraded if not operated on a dedicated network or subnet (LAN or VLAN). As a result, iSCSI is often seen as a low-cost alternative to Fibre Channel, which requires dedicated infrastructure except in its Fibre Channel over Ethernet (FCoE) form.

An iSCSI system is composed of (at least) the following components:

- iSCSI initiator or iSCSI host bus initiator (HBA)
- iSCSI target (disk)
- One or more servers
- Ethernet network

An iSCSI client (*initiator)* sends SCSI commands over an IP network to an iSCSI Storage Resource (*target)*.

Tip: A single iSCSI initiator or iSCSI target is referred to as an *iSCSI node*.

Figure 6-1 shows an overview of an iSCSI implementation with IBM System Storage SAN Volume Controller (SVC).

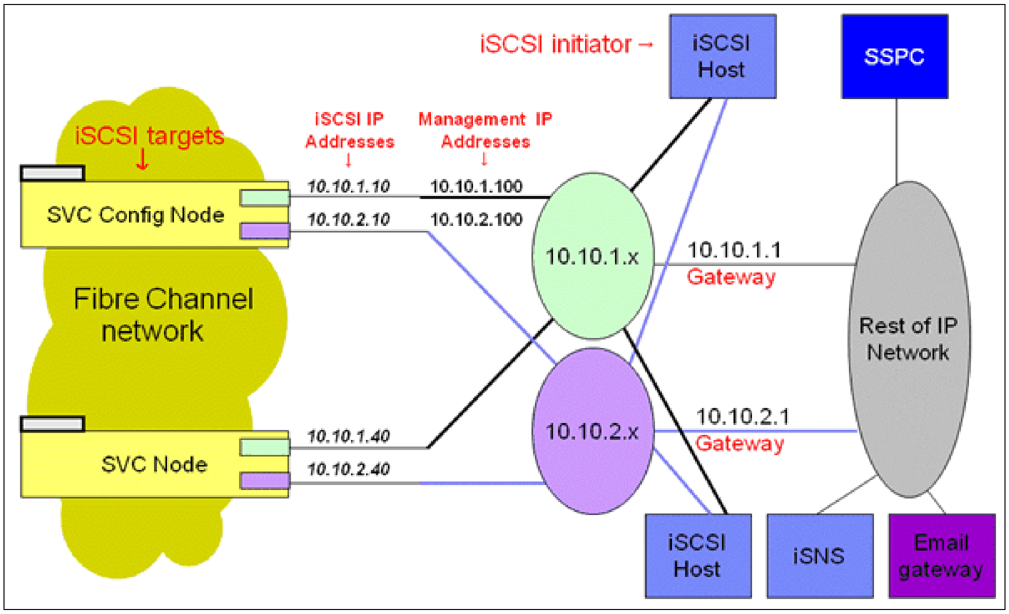


Figure 6-1 iSCSI implementation with SAN Volume Controller

6.1.1 iSCSI initiator

An iSCSI initiator functions as iSCSI client, and sends SCSI (CDB) commands over an IP network. An iSCSI initiator can be either an iSCSI HBA inside a host server, or you can define a software iSCSI initiator by using an iSCSI stack and an Ethernet network adapter.

Tip: *CDB* stands for *SCSI Command Descriptor Block*, which is a block of information that describes the commands sent from SCSI initiators to targets.

iSCSI initiators are of two types:

1. Software Initiator
2. Hardware initiator

Software initiator

A software initiator uses code to implement iSCSI. Typically, this happens in a kernel-resident device driver that uses the existing network card (NIC) and network stack to emulate SCSI devices by speaking the iSCSI protocol. An example of an iSCSI software initiator is the *Microsoft iSCSI Software Initiator*, which runs on Windows Server 2003 and Windows Server 2008.

Check currently supported iSCSI software initiators for various operating systems on the IBM interoperability website:

http://www.ibm.com/systems/support/storage/config/ssic

Hardware initiator

A hardware initiator uses dedicated hardware, typically in combination with software (firmware) running on that hardware, to implement iSCSI. An iSCSI host bus adapter (HBA) implements a hardware initiator.

Tip: An iSCSI HBA can include PCI option ROM to allow booting from an iSCSI target.

A hardware initiator mitigates the overhead of iSCSI and TCP processing and Ethernet interrupts, and therefore improves the performance of servers that use iSCSI. An iSCSI HBA off-loads the IP, TCP, AND iSCSI processes onto special chipsets, thereby reducing the amount of CPU needed by the host.

Using hardware initiators offers the following key benefits:

- It is possible to implement iSCSI boot from SAN with hardware initiators.
- They do not interfere with data networking traffic if the network topology is designed to segregate storage traffic such as by using a separate set of switches to connect servers to iSCSI storage.
- The traffic that passes through them does not load the server's CPU to the same extent that might be the case if the storage traffic passed through the standard IP stack. It results in increased performance.

As some iSCSI hardware initiators might be supported for attachment to DS5000 in the future, check the currently supported types and versions at the IBM interoperability website:

http://www.ibm.com/systems/support/storage/config/ssic

6.1.2 iSCSI target

An iSCSI target usually represents hard disk storage that works over the IP or Ethernet networks. Other types of peripheral devices, like tape drives and medium changers, can act as iSCSI targets as well.

IBM offers a wide range of storage systems that can act as iSCSI targets, including the IBM DS5000, DS4000, and DS3000 series of system storage systems, as well as the virtualized systems such as IBM Storwize V7000 and IBM SAN Volume Controller.

6.1.3 Nodes

There are one or more iSCSI nodes within a network entity. The iSCSI node is accessible by one or more *network portals*. A network portal is a component of a network entity that has a TCP/IP network address and that can be used by an iSCSI node.

6.1.4 Session

An active iSCSI connection is called a session, and a session is required for actual data transfer to occur. The iSCSI targets will listen for connection requests using TCP port 3260 (you can use another port if necessary). You might create other Ethernet ports to handle subsequent communications after a connection between an initiator and target is established.

6.1.5 iSCSI qualified name

For the initiator to communicate with the target, you must assign a unique name to both, so that this communication layer can be established. The *iSCSI qualified name, or IQN*, is used for that purpose on the IBM DS5000 series of storage subsystems.

Tip: All iSCSI HBAs and software initiators configured on the same machine must share the same IQN name.

Example

Figure 6-2 shows the example of a QLogic iSCSI HBA (initiator).

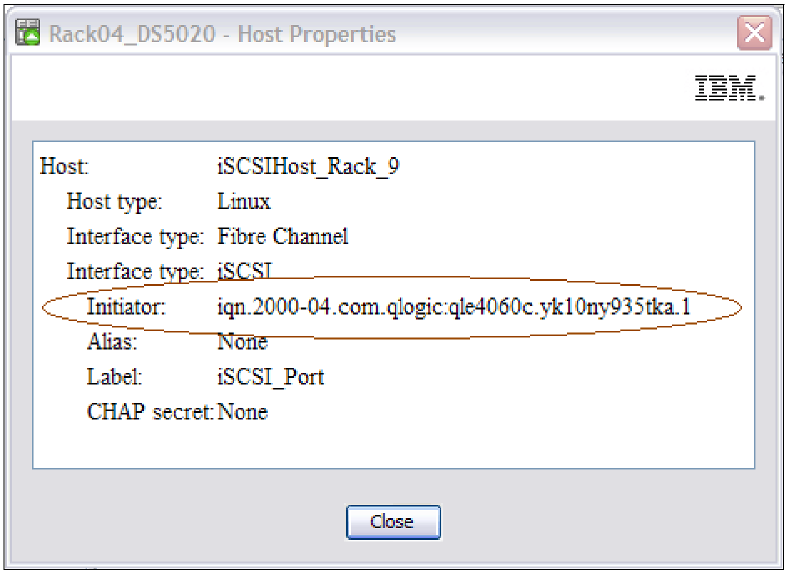


Figure 6-2 iSCSI Initiator

Security

Unlike FC SANs or direct SAS connections, Ethernet networks can be more open, so in order to provide additional security, you can configure the following additional authentication protocols on the DS5000 storage subsystems:

- The Internet Storage Name Service (iSNS) protocol allows for automated discovery, management, and configuration of iSCSI devices on a TCP/IP network. The iSNS servers offer additional security services through explicitly defined initiator-to-target mappings and simplified asset locators, similar to that provided by DNS and WINS for IP address lookup facilities
- Currently, IBM DS Storage Manager 10 has only one iSCSI authentication method called *Challenge Handshake Authentication Protocol (CHAP)*. CHAP is an authentication scheme, used by several types of servers, to validate the identity of remote clients. This authentication occurs when you establish the initial link. The authentication is based on a shared secret and is bi-directional.

CHAP: An initiator-target authentication protocol that uses a challenge to verify that systems have access to each other, either one way or both ways. CHAP occurs constantly without user interaction.

6.2 iSCSI and DS5000 storage subsystems

All the DS5000 series models support intermixed iSCSI and FC host interfaces. Using both interfaces provides a greater degree of flexibility in configuration and deployment of a consolidated storage solution to individual servers or hosts attached to a SAN.

6.2.1 iSCSI host interface connectivity

The DS5000 storage subsystems support 1 Gbps or 10 Gbps iSCSI connectivity. The iSCSI ports support IPv4 and IPv6 TCP/IP addresses, CHAP, and iSNS.

Tip: Use either Cat5E or Cat6 Ethernet cable types for iSCSI port connections. A Cat6 Ethernet cable provides optimal performance.

6.2.2 iSCSI host interface rear view

Figure 6-3 shows the rear view of an IBM DS5020 with mixed Fibre Channel and iSCSI host ports.

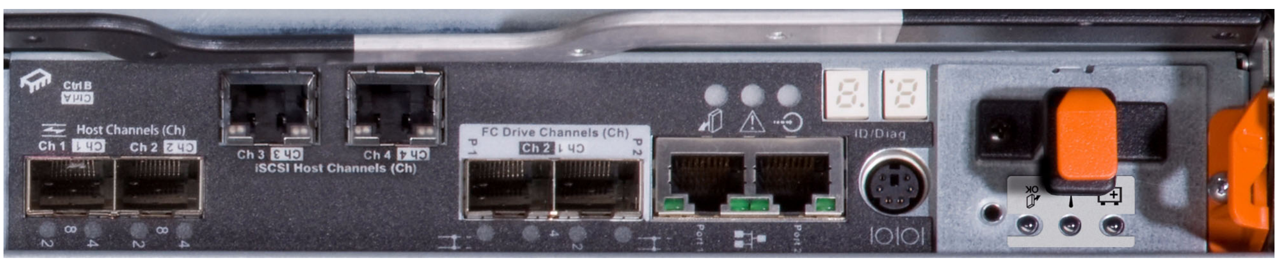


Figure 6-3 DS5020 Controller

6.3 iSCSI boot from SAN

iSCSI boot lets the host server boot from a remote operating system image located on a SAN. It uses SCSI firmware image that makes the volume on remote storage resemble a local bootable drive. The server is configured to boot from the iSCSI target after connecting to the SAN on the network and downloading the operating system (OS) image off of it.

6.3.1 Configuration procedure overview for iSCSI SAN Boot

The following section provides an overview of the configuration procedure. See 6.4, “Step-by-step iSCSI SAN Boot implementation” on page 402 for a step-by-step procedure.

Host configuration

Host configuration includes the following tasks:

1. Disable the local drives (such as IDE, SAS, and SATA) in the server BIOS.
2. Configure the parameters on the first iSCSI port of the hardware initiator from the *QLogic Fast!UTIL* menu. Set the IP address and subnet mask on both iSCSI HBA ports.
3. Ensure that only one iSCSI HBA path exists between the storage and the host system. Disconnect/disable the second path for now, which will be connected later for multipathing.
4. After the network configuration is complete, validate the IP connectivity from the host to the target by pinging the IP address of the device interfaces.

Storage configuration

Here are the necessary requirements to configure your storage system:

- Provision the boot logical unit number (LUN) which is a regular DS5000 LUN.
- Define the host:
  - Define the host OS type.
  - Define the host (initiator) IQN name.
- Map the LUN to the host.

Setting iSCSI boot settings for the primary path

Set the following iSCSI boot settings:

- Select the desired Adapter Boot Mode.
- Set the Primary Boot Device Settings.
- Scan and select the boot volume on SAN as the Primary Target Device.

Operating system installation

Perform the following tasks:

- Select the CDROM as the first boot device from the BIOS.
- Install the operating system.
- Verify that the server successfully boots from the logical drive on the primary path with a power off/on or by restarting the server.

Adding secondary iSCSI path for redundancy

Perform the following tasks:

- Add another cable from the secondary iSCSI HBA path of the server to the network switch.
- Define the host iSCSI port (for the second HBA path) under the same host used for the primary path configuration on the DS5000. The host type must be the same as for the first host iSCSI HBA.
- Verify that both controllers and paths are visible from the host.
- Test path redundancy by performing/simulating failover.

6.3.2 Limitations of iSCSI SAN Boot

When the server is configured with storage access and path redundancy, you can observe the following limitations:

- If there is a path failure and the host is generating I/O, the boot drive moves to the other path. However, while this transition is occurring, the system appears to halt for up to 30 seconds.
- By booting from the DS5000 storage subsystem, most of the online diagnostic strategies are effectively canceled, and path problem determination must be done from the diagnostics panel, which can be accessed by pressing Ctrl+Q.
- The IDE disk devices must not be re-enabled.

For iSCSI there is no specific zoning configuration to be done switches. It is highly recommended to isolate the iSCSI traffic from the regular Ethernet traffic by using different physical switches or by configuring VLANs.

6.4 Step-by-step iSCSI SAN Boot implementation

This section describes the step-by-step implementation tasks needed to set up the iSCSI SAN Boot installation.

In our sample configuration, we use the following components:

- IBM System x3550 M3 server
- QLogic iSCSI HBA hardware initiator card - QLE4060C
- IBM DS5020 storage subsystem with v7.60.40 controller firmware
- IBM DS Storage Manager v10.70

> **Tip:** The implementation steps can vary, depending on the particular host type and operating system that are involved.

6.4.1 Host configuration

This section illustrates how to configure the QLogic iSCSI Card for iSCSI boot and assumes that an iSCSI target has not previously been configured on the HBA.

Before you begin:

- Ensure that the latest BIOS and firmware is installed on the HBA.
- Ensure that the storage device and Ethernet switch (if used) are both powered up and completely initialized before proceeding further.
- Ensure that only the primary iSCSI host HBA path is plugged into the network switch at this time. Leave the secondary iSCSI host HBA path cable unplugged for now.

Server BIOS configuration

To configure the server BIOS, perform the following tasks:

1. Power on the server and press the F1 key to interrupt the boot sequence and enter the server system BIOS.
2. Select **Devices and I/O Ports** from the main menu, as shown in Figure 6-4.

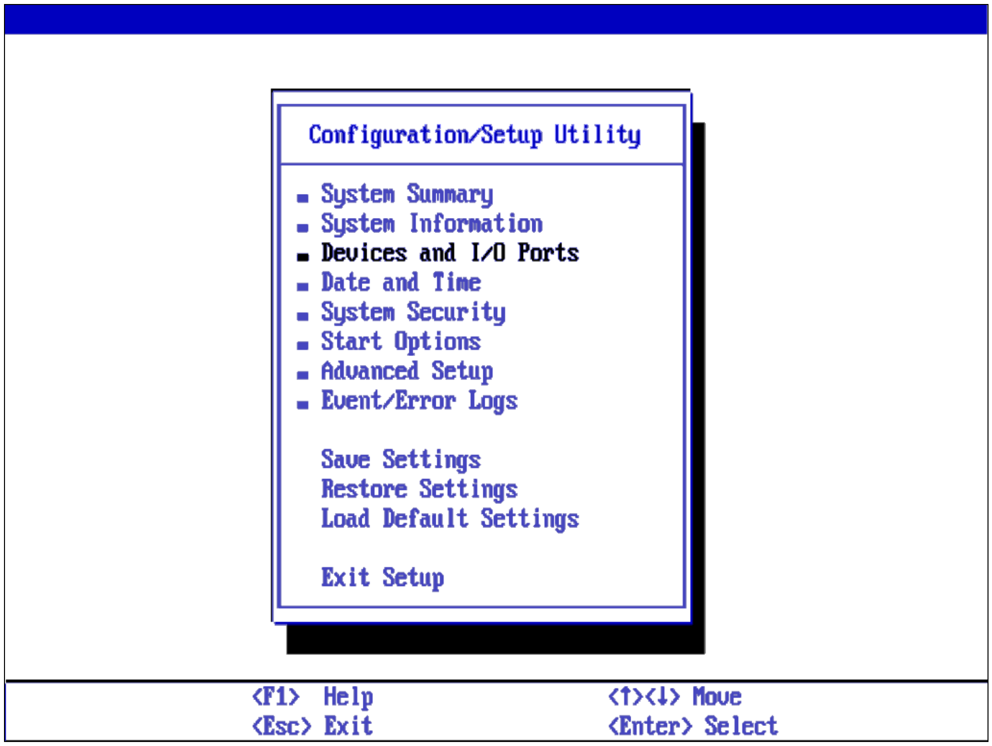


Figure 6-4 Devices and I/O Ports

3. Disable any onboard or internal hard drives. In our example, we disable the **Planar SAS** for internal drives, as shown in Figure 6-5.

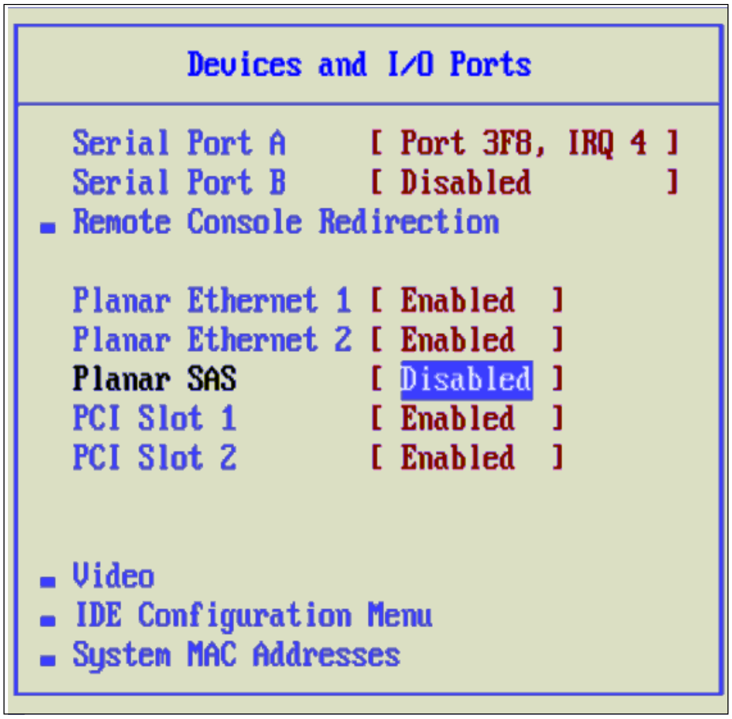


Figure 6-5 Disabling Planar SAS

4. Press Esc to exit out of this menu and save the changes.
5. Restart the server.

iSCSI HBA configuration

To configure the iSCSI HBA for SAN Boot, perform the following tasks:

1. On server startup and when you are prompted to launch the QLogic Corporation iSCSI BIOS, press the Ctrl+Q keys, as shown in Figure 6-6.

```
QLogic Corporation
QLA2300 PCI Fibre Channel ROM BIOS Version 1.16
Copyright (C) QLogic Corporation 1993-2001. All rights reserved.
www.qlogic.com

Press <CTRL-Q> for Fast!UTIL


BIOS for Adapter 0 is disabled
ROM BIOS NOT INSTALLED

QLogic Corporation
QLE4060C iSCSI ROM BIOS Version 1.15  Subsystem Vendor ID 1014
Copyright (C) QLogic Corporation 1993-2008. All rights reserved.
www.qlogic.com

Press <CTRL-Q> for Fast!UTIL

<CTRL-Q> Detected, Initialization in progress, Please wait...
```

Figure 6-6 Control-Q for QLogic FastUtil!

2. In the QLogic FastUtil menu, select the desired HBA, as shown in Figure 6-7.

```
                         QLogic Fast!UTIL

                       Select Host Adapter
 Adapter  Boot Mode Address  Slot Bus Device Function  MAC Address

 QLE4060C Manual    6100     00   08  01      1        00-14-5E-F1-E7-3F
 QLE4060C Manual    4100     00   14  01      1        00-14-5E-F1-E6-8D
```

Figure 6-7 Select HBA

3. From the Fast!UTIL Options window, select **Configuration Settings** and press Enter, as shown in Figure 6-8.

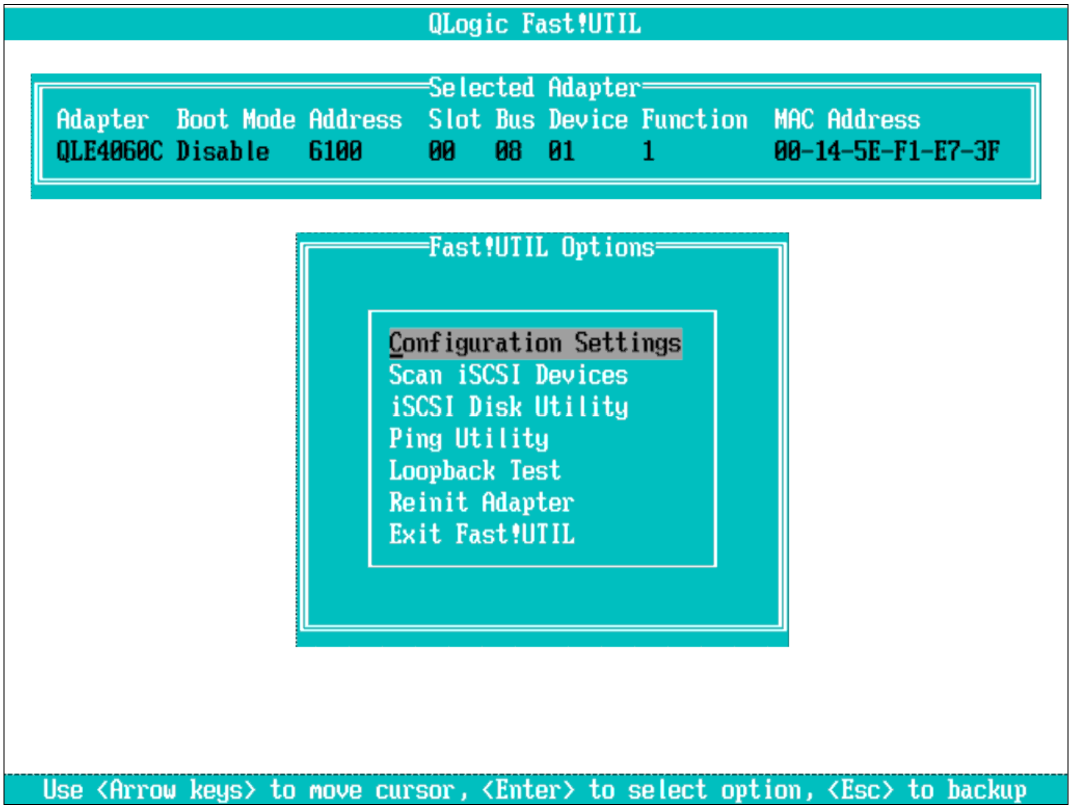


Figure 6-8 Configuration Settings

4. Select **Host Adapter Settings** and press Enter to input the HBA configuration parameters, as shown in Figure 6-9.

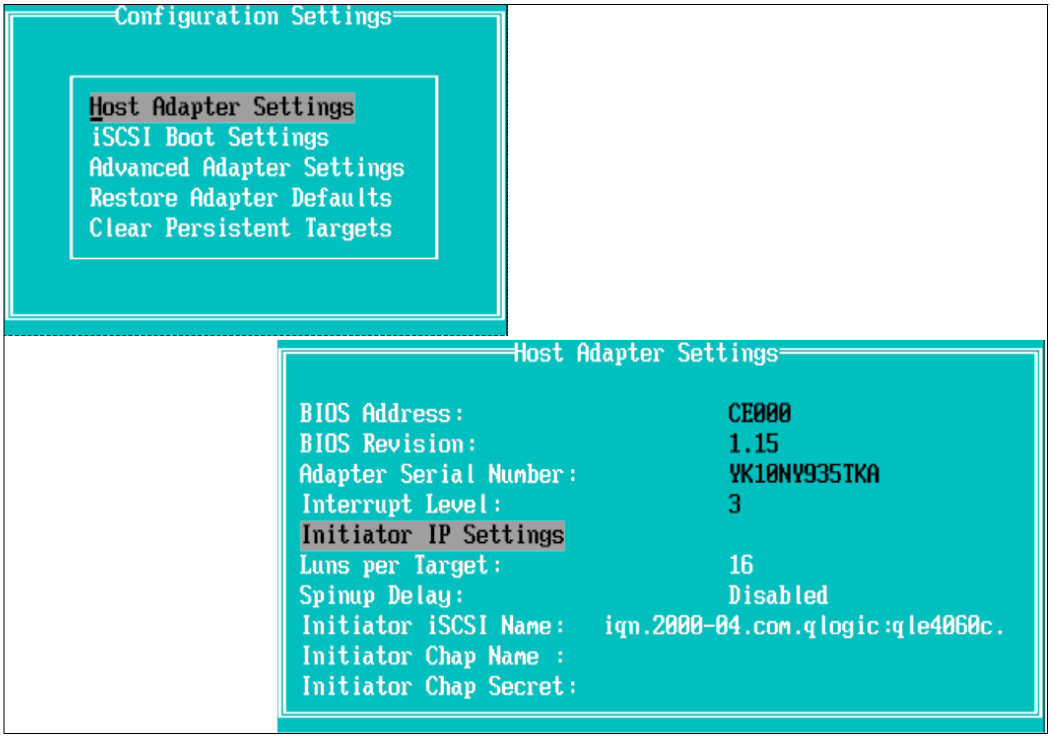


Figure 6-9 Host Adapter Settings window

By pressing Enter on the **Initiator IP Settings** selection as shown in Figure 6-9, a new window is displayed where all IP settings can be changed.

5. Enable either the IPv4 or IPv6 options, depending on whether you are booting from an IPv4 or IPv6 target (Figure 6-10). If using DHCP to obtain these parameters from a DHCP server, you need only enable the Initiator IP Address through DHCP.

Tip: It is preferred that the host and target ports reside in the same VLAN/Subnet.

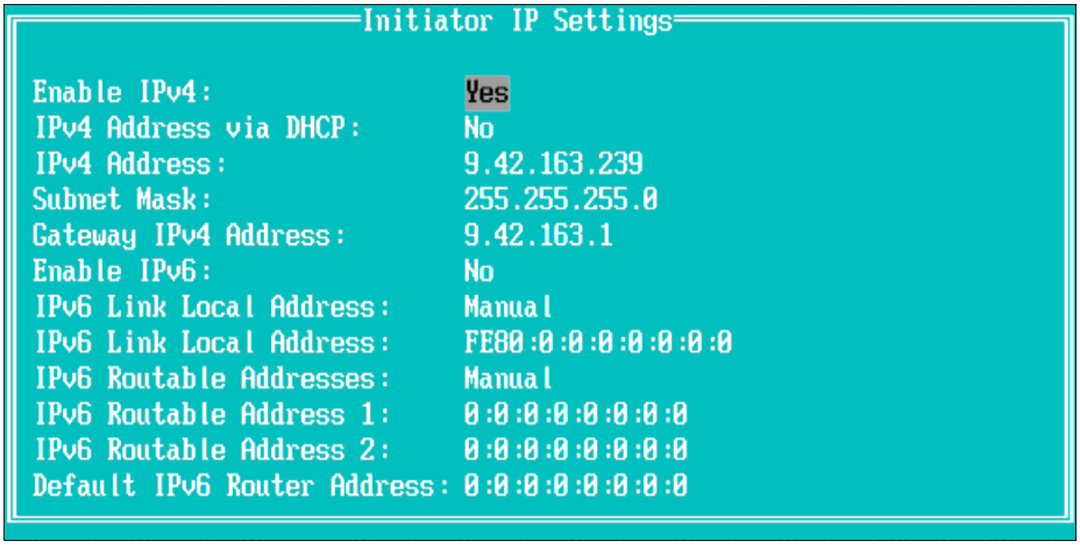


Figure 6-10 Initiator IP Settings

6. Press Esc to come back to Host Adapter Settings. Here, a default **Initiator iSCSI Name** is provided. Optionally, you can change this, however, the new name must be in the form of a correctly formatted RFC 3720 IQN (iSCSI Qualified name).

   Best practice is to set authentication to archive the right security level for your boot environment. You can also use CHAP authentication with the target by specifying the **Initiator Chap Name** and **Initiator Chap Secret** on this window, as shown in Figure 6-11.

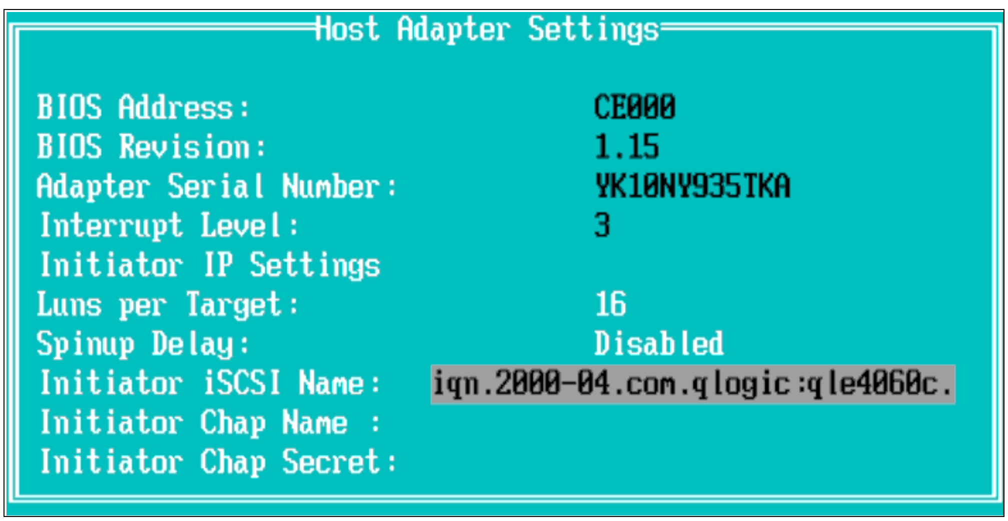


Figure 6-11 IQN

   Press Esc again to go to the **Configuration Settings** menu. Proceed to “Configuring CHAP authentication” on page 411 to complete the CHAP authentication setup on the DS5000 storage subsystem before continuing with this section.

7. From the **Configuration Settings** menu, select **iSCSI Boot Settings**., as shown in Figure 6-12.

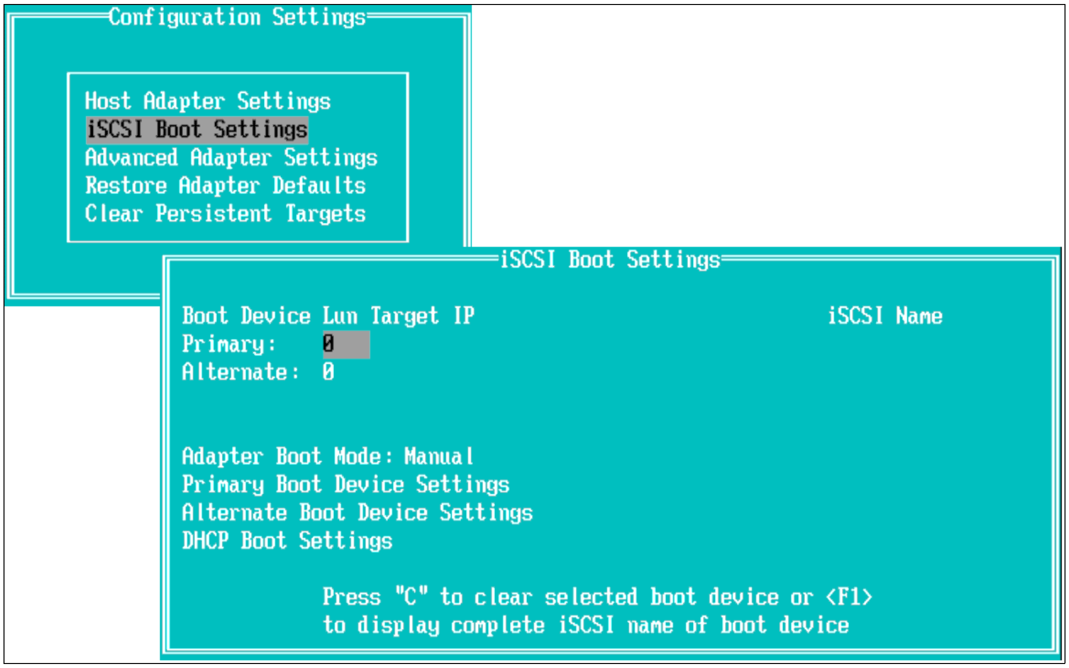


Figure 6-12 iSCSI Boot Settings

8. Select **Primary Boot Device Settings** and specify the IP address of the controller which owns the boot LUN, as shown in Figure 6-13.

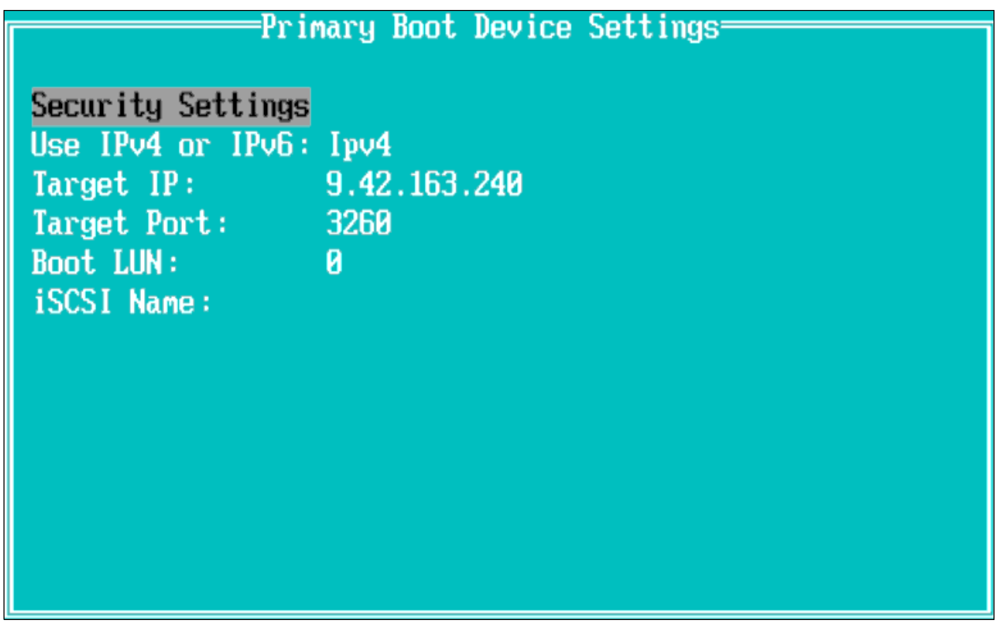


Figure 6-13 Primary Boot Device Target

> **Tip:** To define the host ports (IQN name) for the host definition, it is helpful to let the HBA log in to the storage, assuming that it is set up already.

9. Change the *Adapter Boot Mode* to **Manual** to enable the BIOS on this port. The resulting panel looks similar to Figure 6-14.

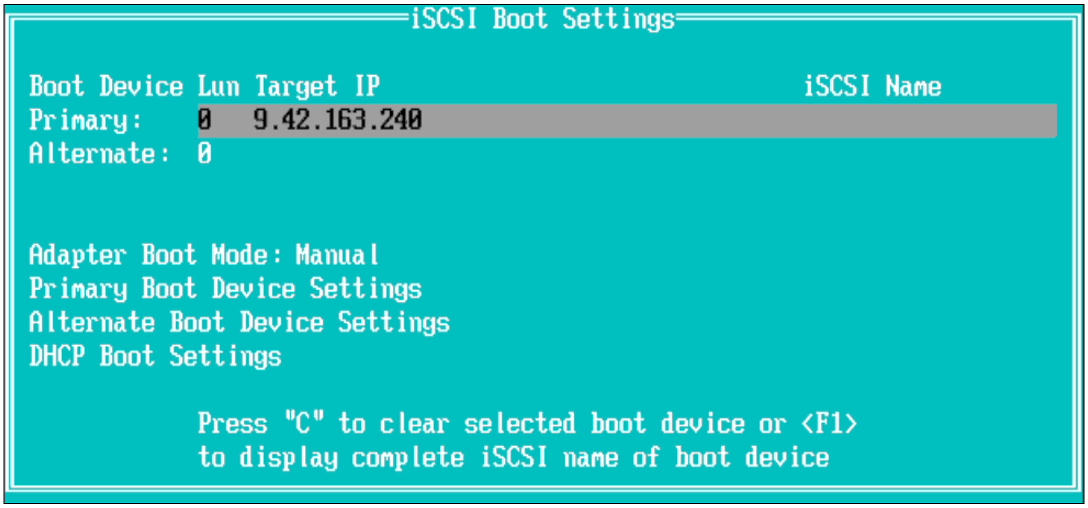


Figure 6-14 HBA Target Set

10. Press Esc to go back. As configuration settings have been modified, it will prompt you to **Save changes**. Press Enter to save, as shown in Figure 6-15.

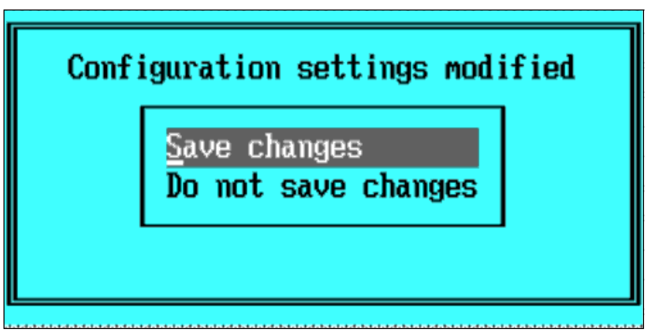


Figure 6-15 Save configuration setting changes

11. Next, in the QLogic Fast!UTIL menu, press Enter on **Reinit Adapter**, as shown in Figure 6-16. It will initialize the HBA and establish a connection between the iSCSI HBA and the DS5000 Controller.

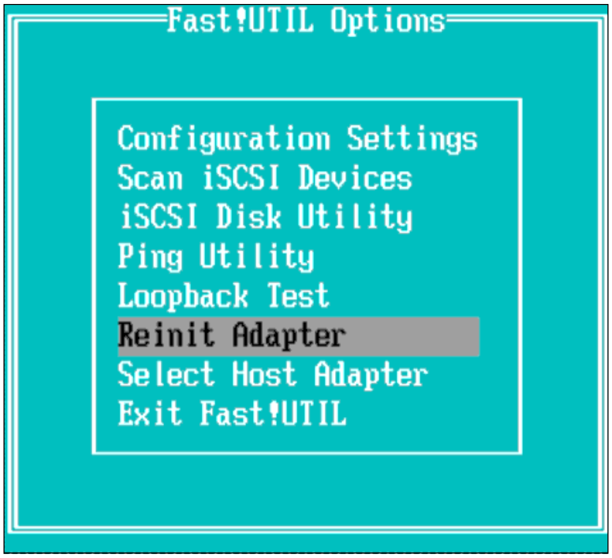


Figure 6-16 Initialize the HBA

Stay in this window, because you need to return to the HBA BIOS in order to select the boot LUN after the storage provisioning.

6.4.2 DS5000 storage provisioning

This section covers the storage configuration necessary to prepare the iSCSI target and boot device. For detailed instructions regarding iSCSI setup on DS5000, see the *IBM Midrange System Storage Hardware Guide*, SG24-7676 and the *IBM Midrange System Storage Implementation and Best Practices Guide*, SG24-6363.

We assume that you have IBM System Storage DS Storage Manager installed, and the DS5000 storage subsystem, in optimal state, is being managed by it. For the purpose of this implementation, we used Storage Manager version 10.70.

Configuring CHAP authentication

We now need to configure CHAP authentication on the target DS5000 storage subsystem to provide a secure connection to the host. We use the CHAP secret that was defined in “iSCSI HBA configuration” on page 405.

Perform the following tasks to configure CHAP authentication using the Initiator Chap Secret:

1. In the Storage Manager’s Subsystem Management window, go to the **Setup** tab.
2. Click **Manage iSCSI Setting**.

3. You are presented with a panel for CHAP Authentication as shown in Figure 6-17.

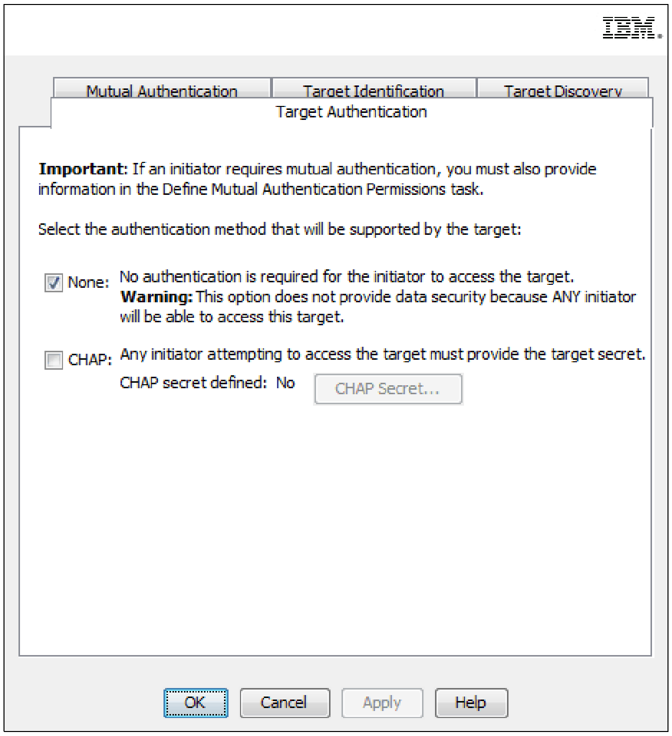


Figure 6-17 CHAP Authentication

4. Click the **Mutual Authentication** tab as shown in Figure 6-18.

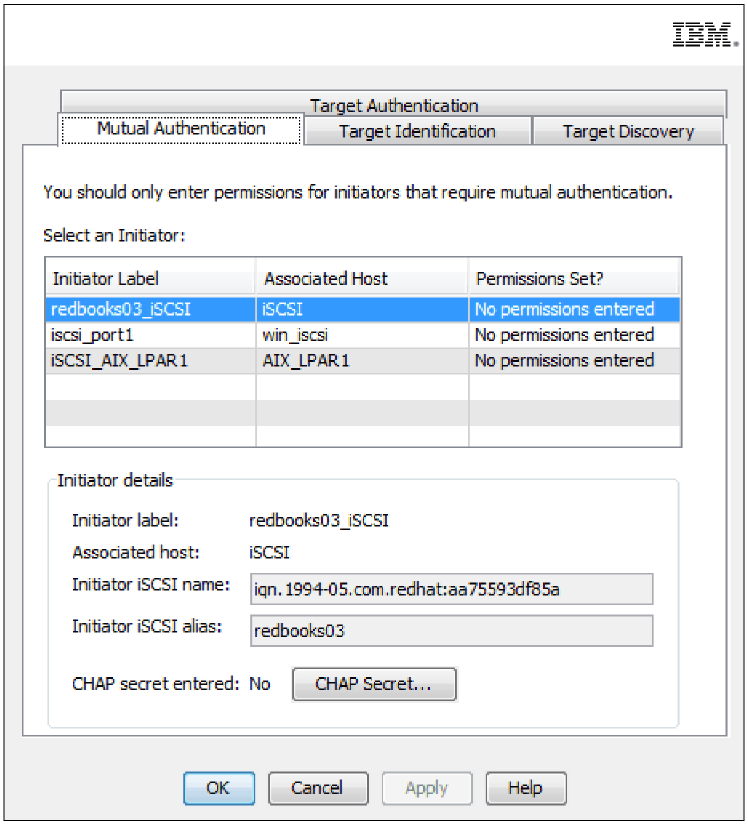


Figure 6-18 CHAP Mutual Authentication

5. Select the host initiator server in question and click the **CHAP Secret...** button to enter the Initiator CHAP Secret. Enter the Initiator Chap Secret that was defined in “iSCSI HBA configuration” on page 405, into the provided fields as shown in Figure 6-19. Ensure that you confirm the Initiator CHAP Secret correctly.

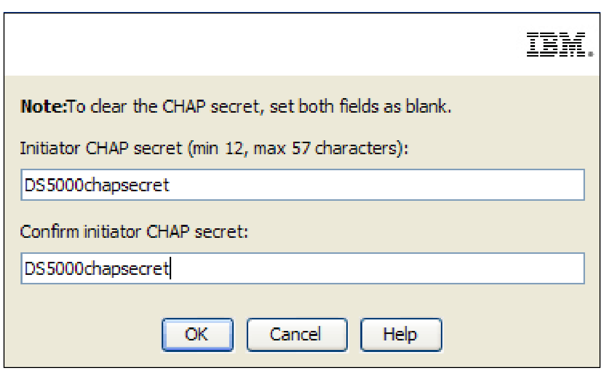


Figure 6-19 Initiator CHAP Secret

6. You are presented with a confirmation panel stating that the CHAP secret has been set successfully. Click **OK**. You will now be returned to the Mutual Authentication panel as shown in Figure 6-20. We can now see that our host redbooks03_iSCSI has CHAP authentication defined as Permission Set.

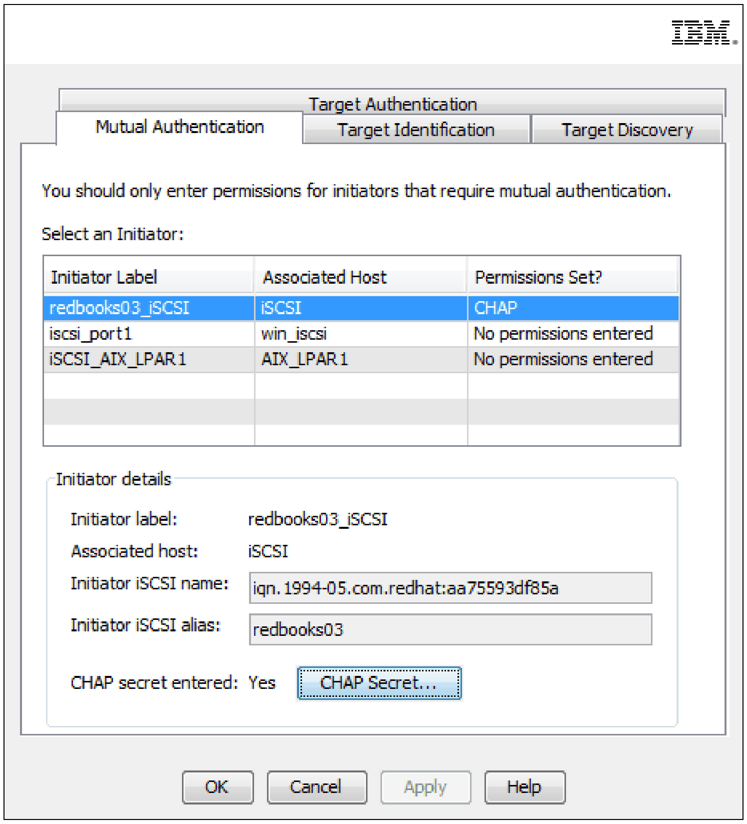


Figure 6-20 CHAP Mutual Authentication

CHAP authentication using the Initiator Chap Secret has now been set up. You should be able to establish a secure connection between the DS5000 storage subsystem and the host server.

Perform the following tasks to configure the boot LUN for your operating system partition:

1. Create a logical volume (LUN) of size that fits your OS system partition requirements:
   a. In the Storage Manager's Subsystem Management window, go to the **Logical** tab.
   b. Right-click **Free Capacity** → **Create Logical Drive...**, as shown in Figure 6-21.

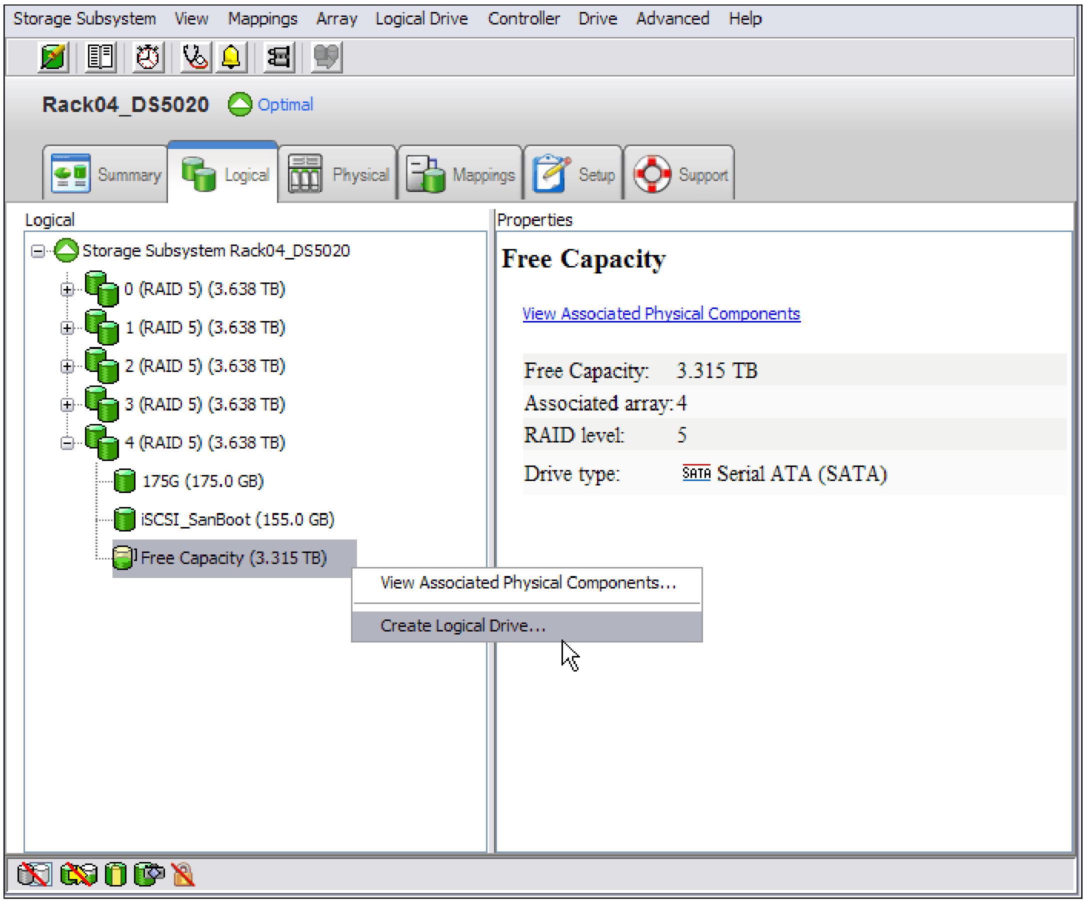


Figure 6-21 Create Logical Drive

c. A wizard will guide you through the steps of creating the logical drive, as shown in Figure 6-22 and Figure 6-23.

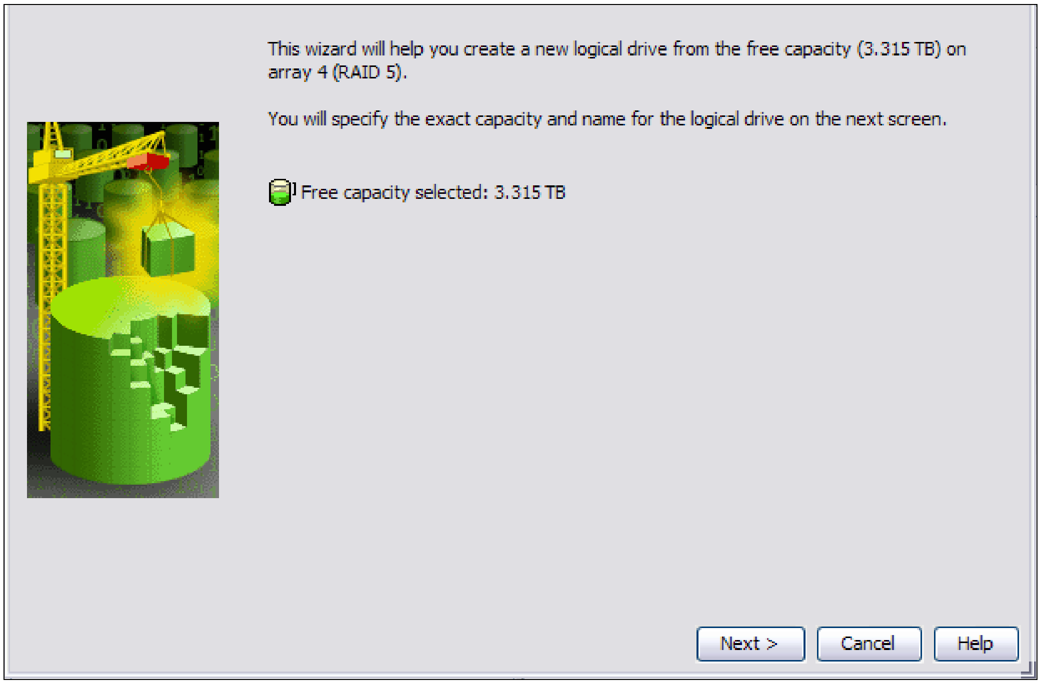


Figure 6-22 Logical Drive Creation Wizard

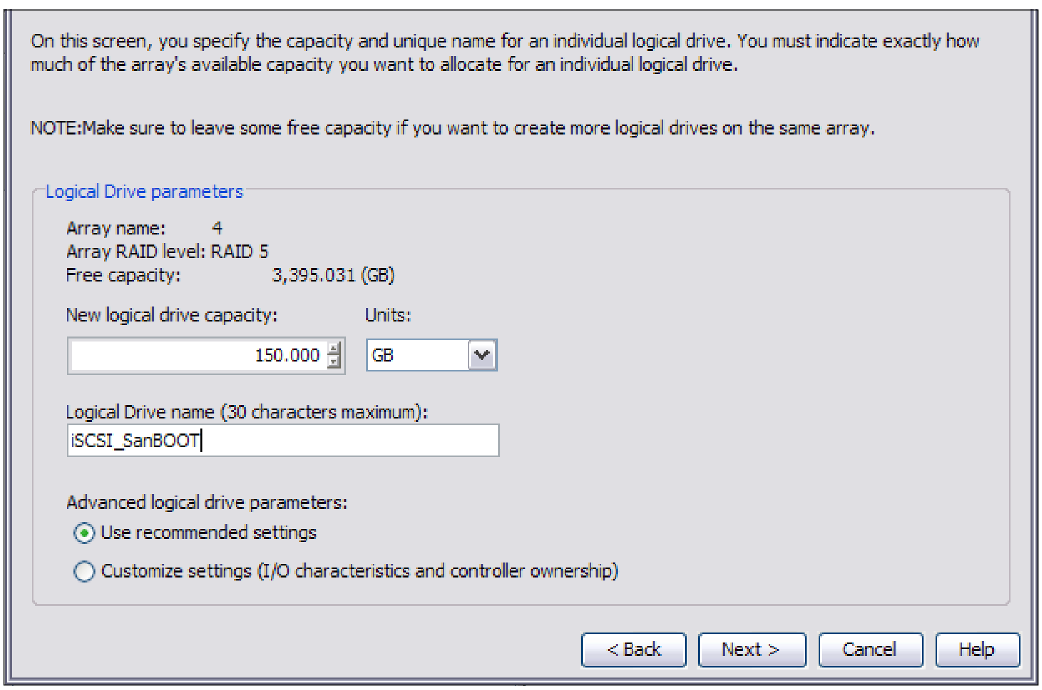


Figure 6-23 Defining Logical Drive parameters

d. In the next window, you can map the LUN to the SAN Boot host if it is defined already, otherwise, you can choose to map it later, as shown in Figure 6-24.

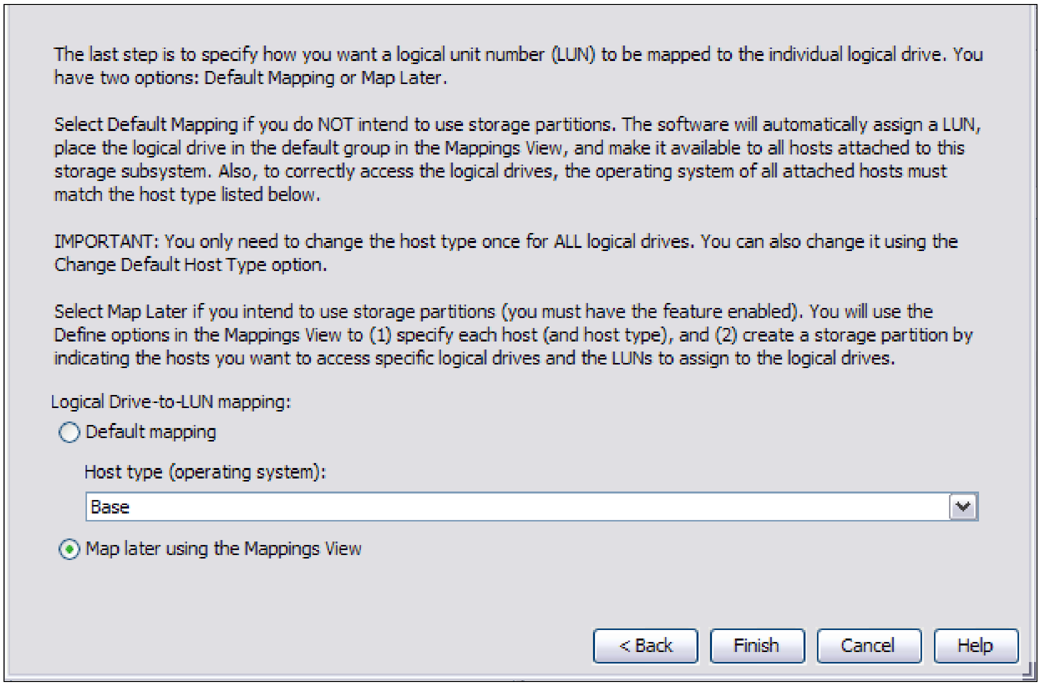


Figure 6-24 Logical drive to Lun Mapping

You get a confirmation and next-step instructions, as shown in Figure 6-25.

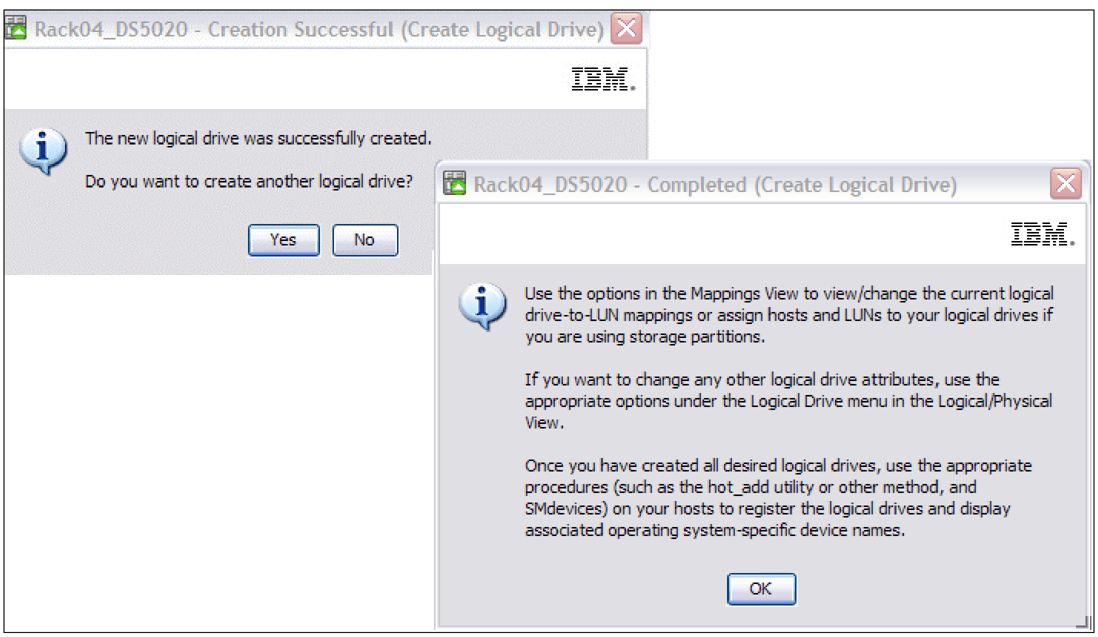


Figure 6-25 Logical Drive Created

2. Next, we create the host definition that will identify the host HBA with the DS5000 storage subsystem:

   a. In the Storage Manager Subsystem Management window, go to the **Mappings** tab. In the **Topology** frame, right-click the storage subsystem name (top item), then **Define** → **Host**..., as shown in Figure 6-26.

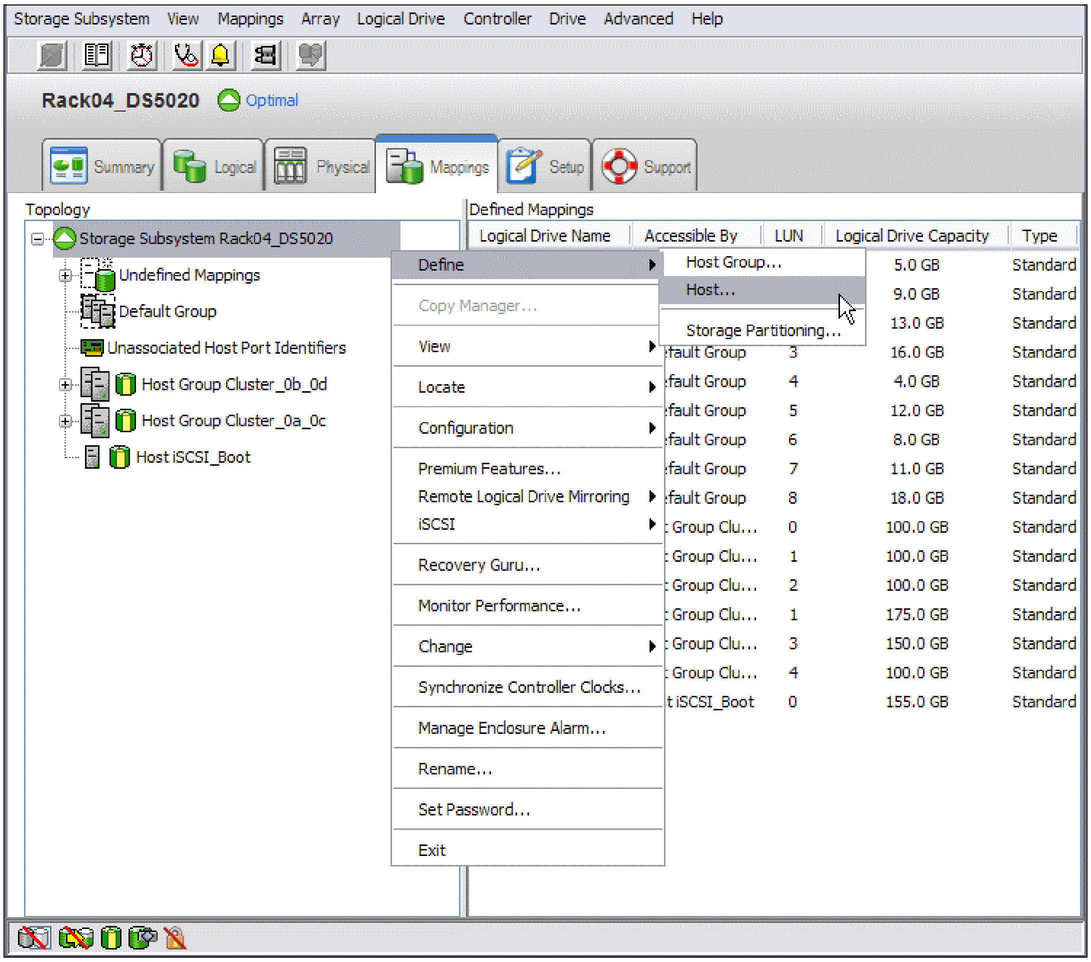


Figure 6-26 Define Host

b. A wizard window opens, to guide you through the steps of defining hosts. Provide the desired host name in the **Host name** text box, as shown in Figure 6-27.

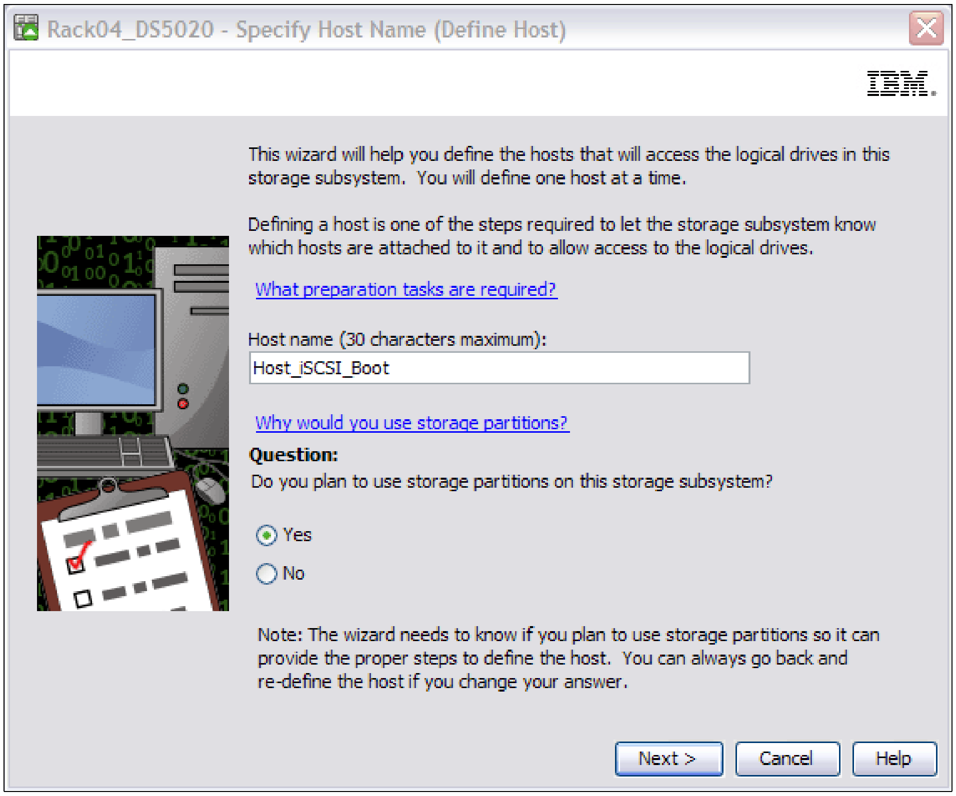


Figure 6-27 Specify Host Name

c. In the next window, we specify the host port identifier of the HBA. From the *host interface type* drop-down, choose **iSCSI**. If you have a network path between storage and the HBA, the *Known unassociated host port identifier* will automatically populate with the HBA's IQN. Otherwise, you need to type it manually under the **New host port identifier** textbox. Figure 6-28 shows the interface.

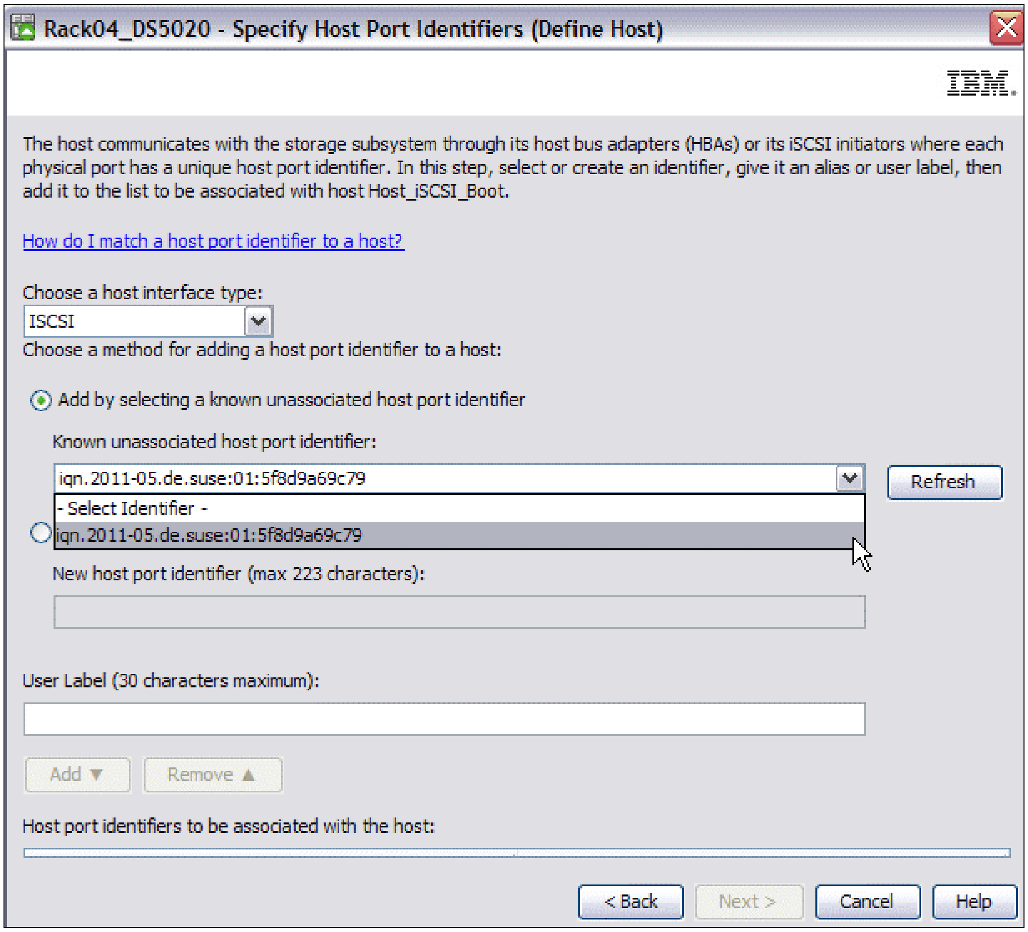


Figure 6-28 Specify host port identifier

Tip: It is useful to have the server iSCSI HBA network setup done before creating the host partition on the DS5000, which forces the HBA to do a SCSI login to the DS5000. Subsequently, the DS5000 will remember the IQN names of the host.

d. Give a **User Label** to the identifier and click **Add**, as shown in Figure 6-29.

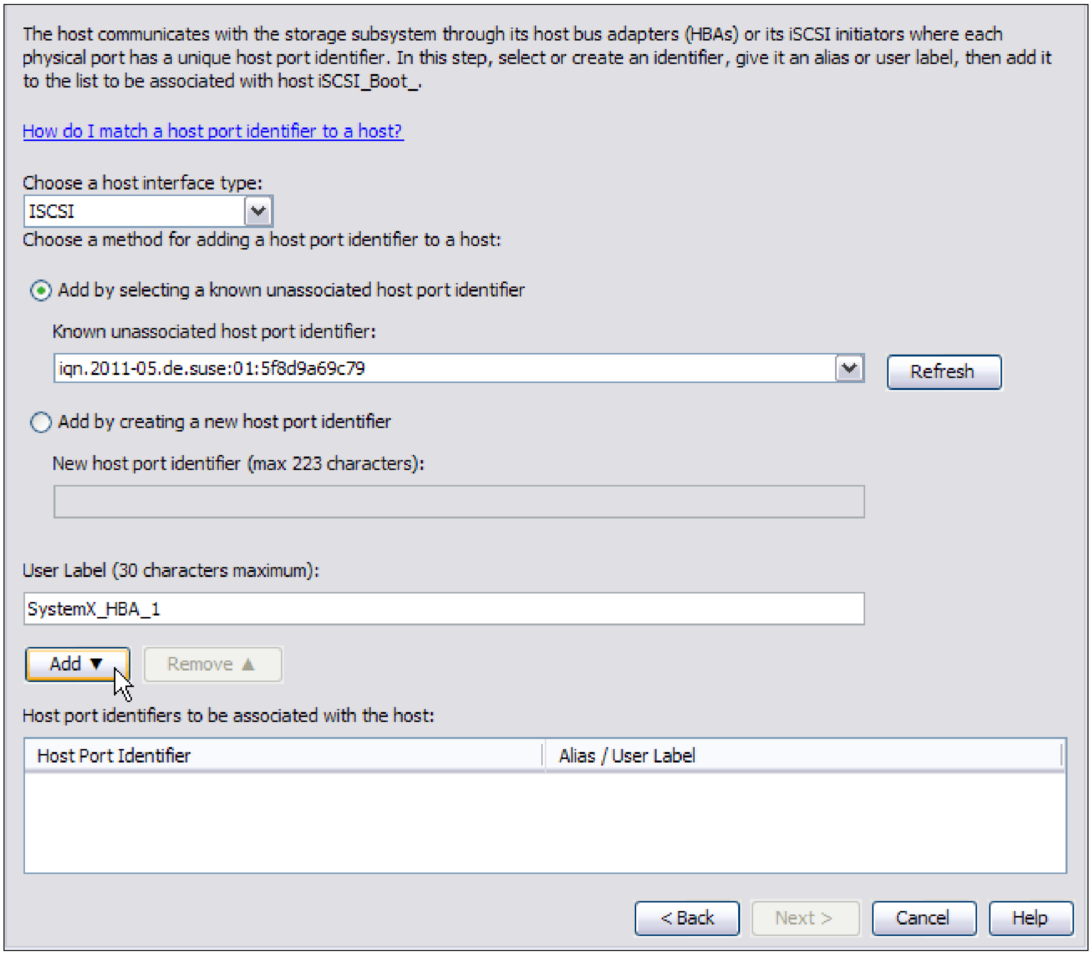


Figure 6-29 Associate host port identifier with the host

e. Click **Next** to go to the window that allows to specify the host type, as shown in Figure 6-30. In our case, we chose **Linux** from the drop-down menu.

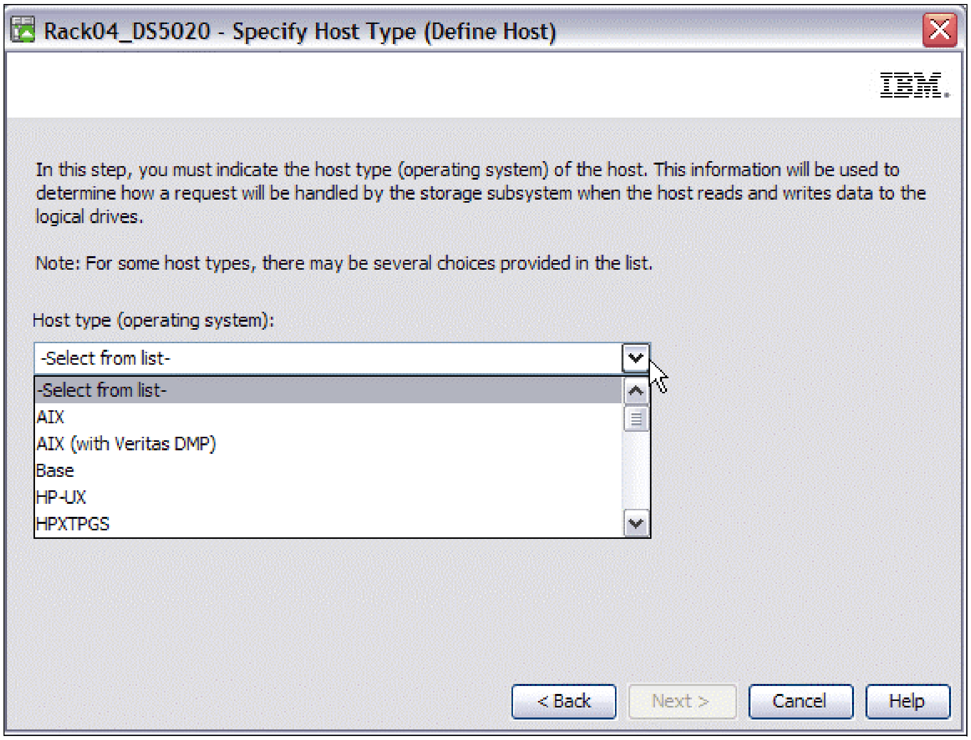


Figure 6-30 Select Host type

f. Finally, the wizard gives a preview of your newly created host's definition, as shown in Figure 6-31. Click **Finish** to complete the process.

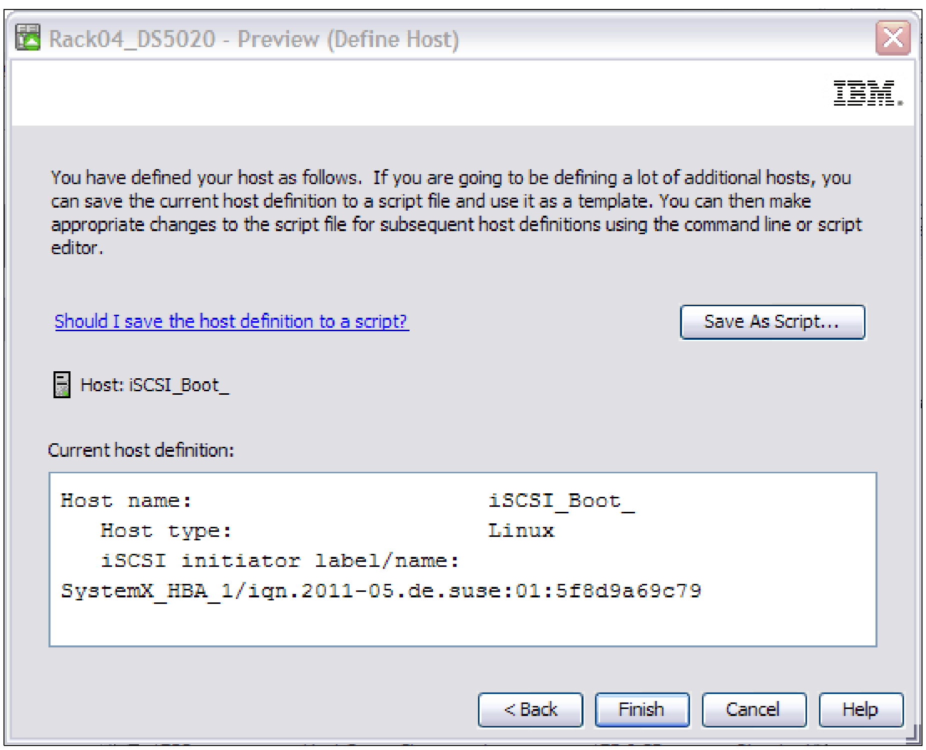


Figure 6-31 Preview after host definition process

The new host is shown under Default Mapping. In the next step we map the boot LUN to it, which will consequently create a storage partition.

3. Storage partition: We map the host to the boot volume on SAN:

   a. In the Storage Manager Subsystem Management window, go to the **Mappings** tab. Under Undefined Mappings, you can find the newly created LUN. Right-click it to **Define Additional Mapping...**, as shown in Figure 6-32.

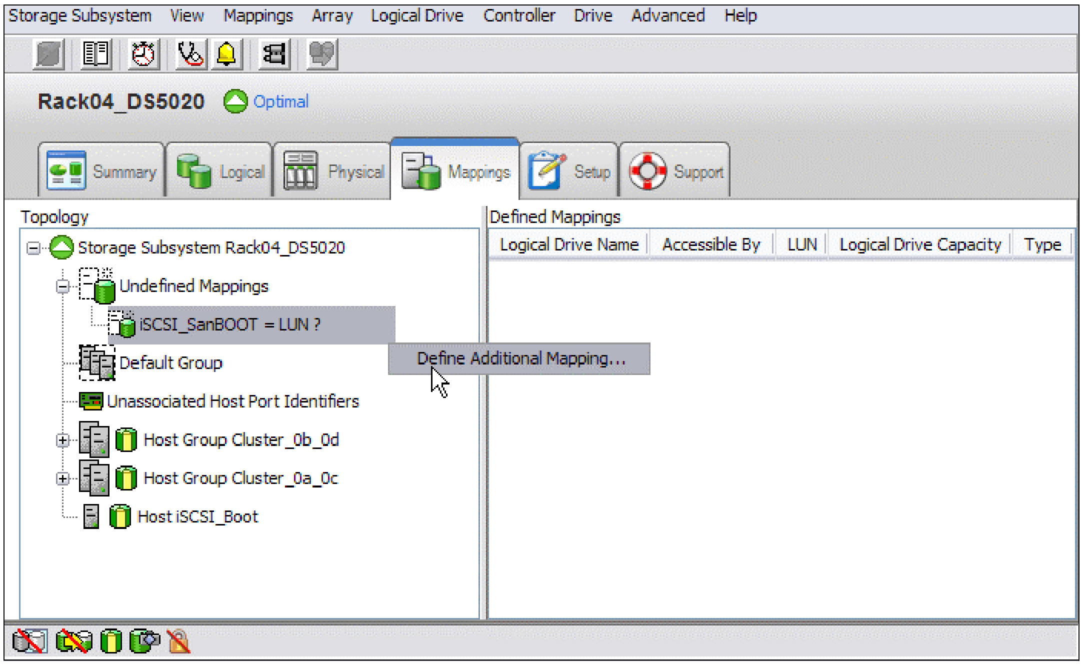


Figure 6-32 Define Mappings

b. In the Define Additional Mapping window, choose the desired host from the drop-down menu. Next, for the boot LUN, choose LUN 0 (zero), as shown in Figure 6-33.

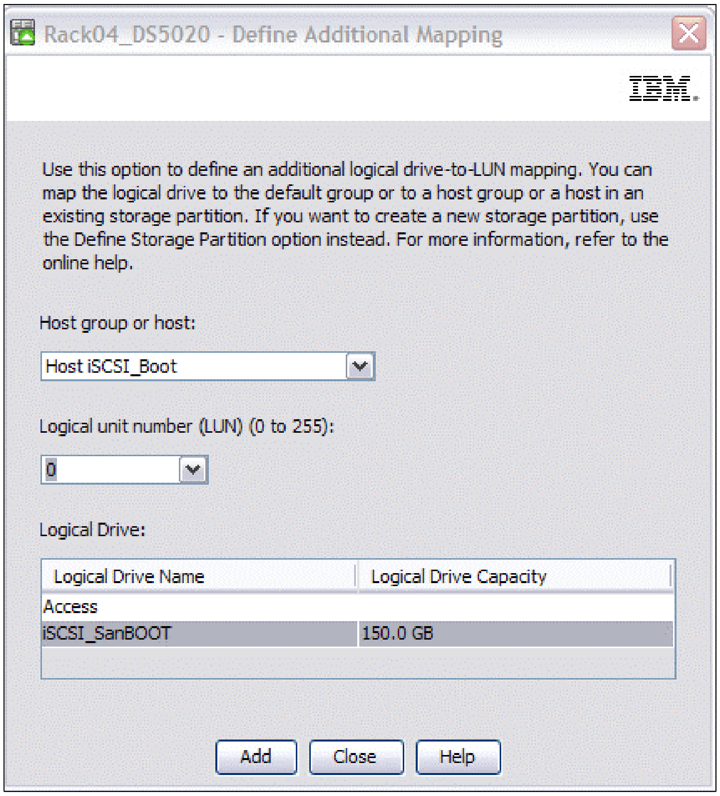


Figure 6-33 Assign Host

> **Important:** The boot logical volume must have LUN ID 0 to be able to boot the operating system,

c. In the Mappings Window, you should be able to see the newly created Boot LUN in the Defined Mappings frame, as shown in Figure 6-34.

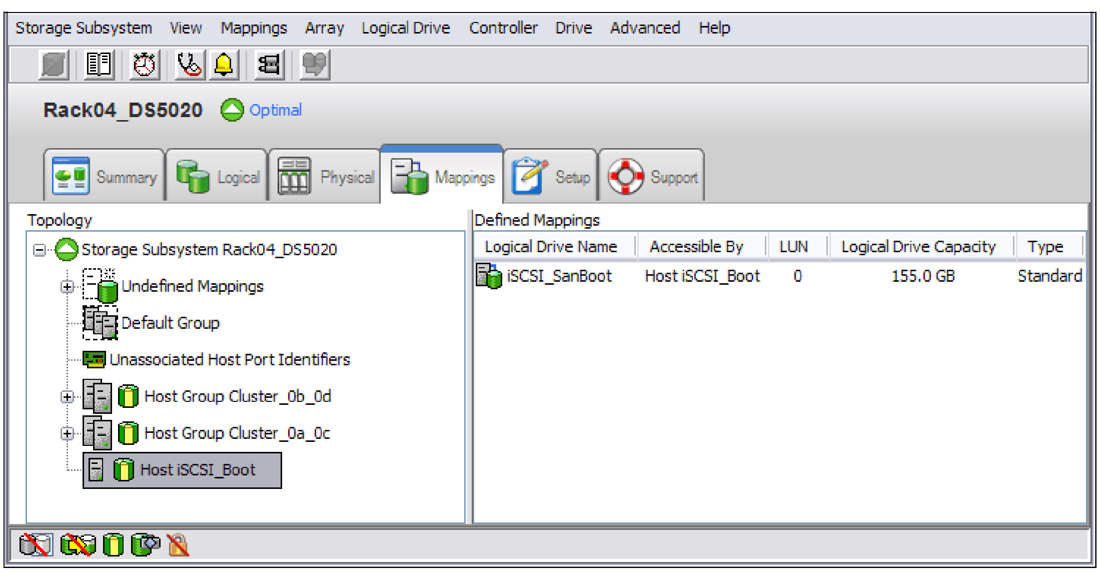


Figure 6-34 Boot LUN mapped to the host

Verify the properties of the LUN, then proceed to 6.4.3, “Additional host configuration” on page 427

> **Tip:** If you have multiple LUNs mapped to this host, *delete all mappings* and make sure that the boot disk has LUN number 0 and is not mapped to other hosts. Also, the access LUN must *not* be mapped to the host at this stage because it will be seen by the host as storage.

After the OS and failover driver installation, you might need to enter the mapping again in order to map additional data disks to your host.

Configuring the secondary path for failover

We now proceed with setting up the secondary path to the storage that will provide failover in the event of a failure of the primary path.

Initial preparation

Perform the following tasks:

1. Connect the second iSCSI cable to the second iSCSI port on the server.
2. Enter the BIOS of the second port as shown in and set the IP settings for the second HBA path as were performed for the primary path.

Configuring the secondary path to the DS5000 storage subsystem

To configure the secondary path to the DS5000 storage subsystem, perform these tasks:

1. In Storage Manager mapping view, select the host and define the second host port of the server to be part of the storage partition as seen in Figure 6-35.

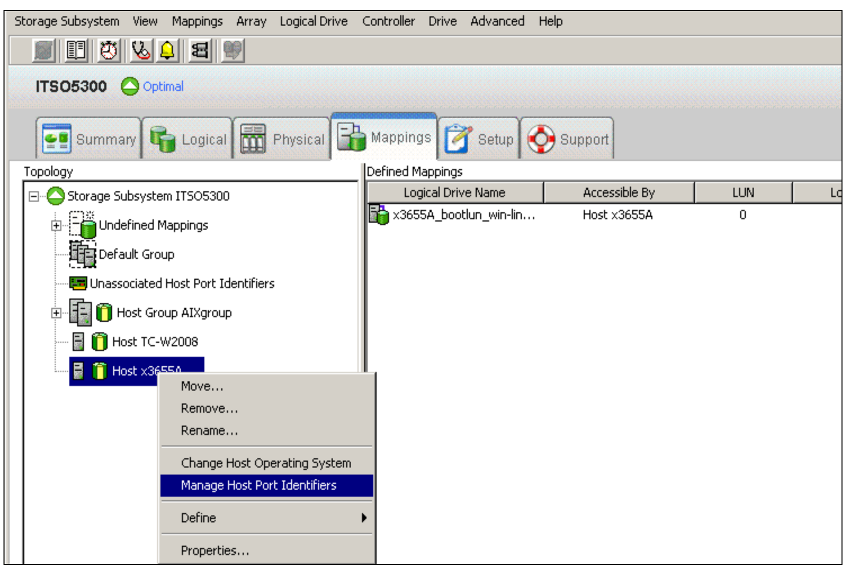


Figure 6-35 Storage Manager - Manage Host Port Identifiers

2. From the Add Host Port Identifier window, select the interface, known IQN or WWPN from the Known unassociated host port identifier drop-down list and define a host port name (as shown in Figure 6-36).

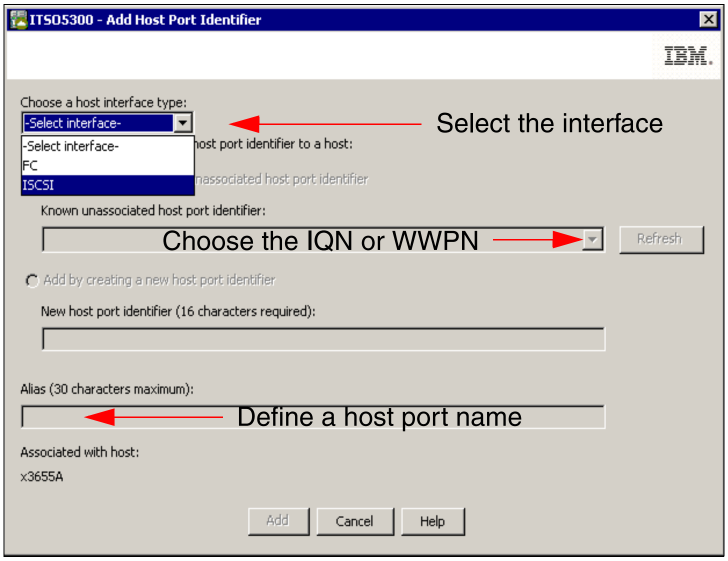


Figure 6-36 Add Host Port Identifier

The host can now see the secondary path to the DS5000 storage subsystem. You can also configure any additional host storage if needed.

6.4.3 Additional host configuration

To specify the boot LUN to the host HBA, we need to perform a few more steps in HBA BIOS utility. Perform the following tasks:

1. Reboot the host server. On the prompt to launch the QLogic Corporation iSCSI BIOS, press the Ctrl+Q keys, as shown in Figure 6-37.

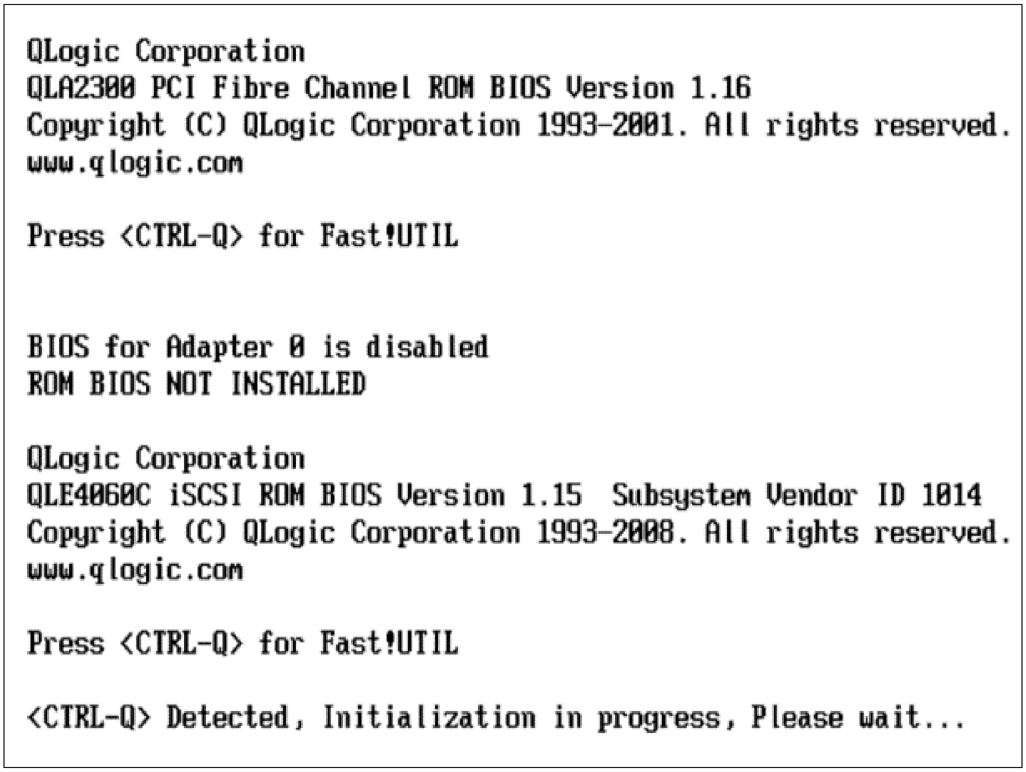


Figure 6-37 Control-Q for QLogic Fast Util!

2. In the QLogic FastUtil menu, select the desired HBA, as shown in Figure 6-38.

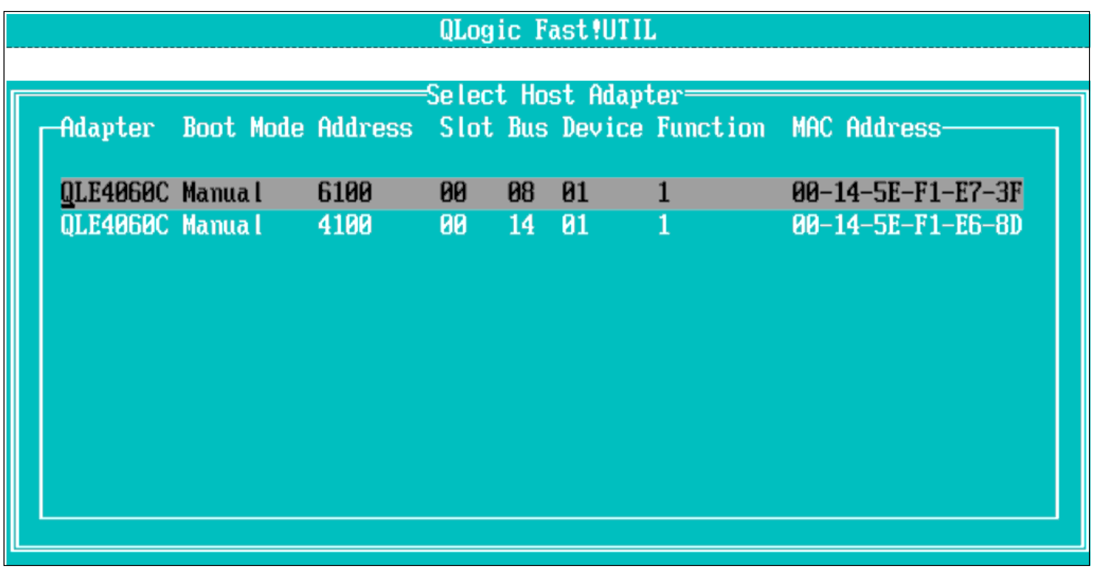


Figure 6-38 Select HBA

3. From the Fast!UTIL Options window, select **Scan iSCSI Devices** and press Enter. The HBA will scan for devices connected to it, as shown in Figure 6-39.

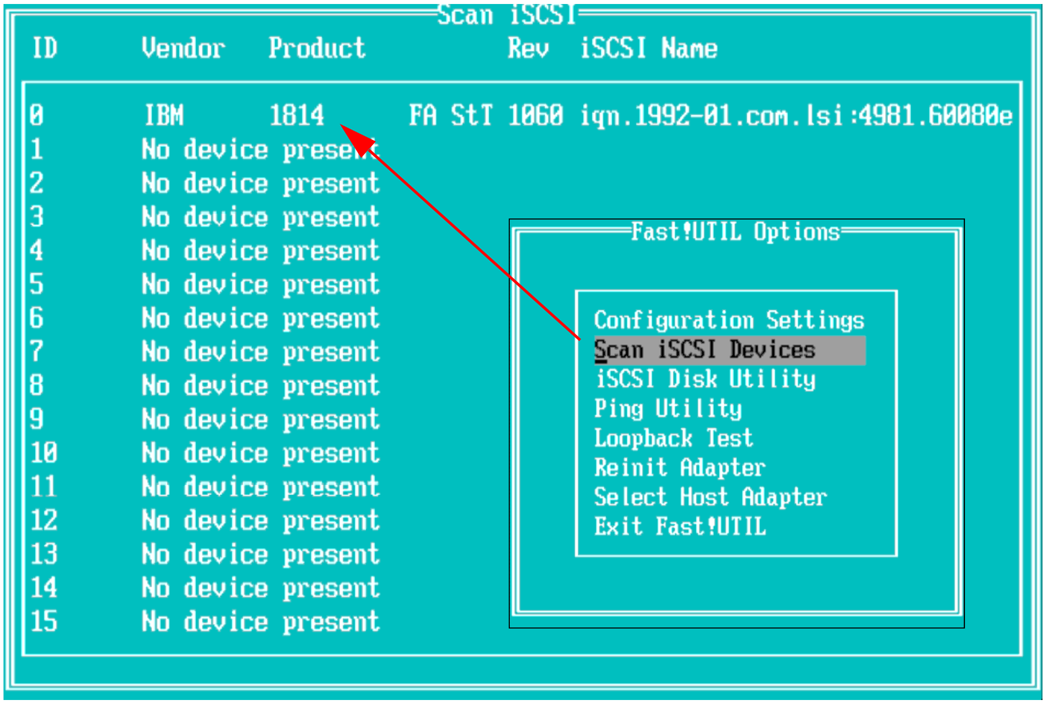


Figure 6-39 Scan iSCSI Devices

> **Tip:** At this time, the iSCSI initiator SCSI protocol is able to login to the DS5000. The DS5000 can recognize the new IQN to be used in the storage partitioning.

4. Re-enter into the **Configuration Settings** → **iSCSI Boot Settings**. You see a window that represents the new primary boot target IP address but without an IQN name, as shown in Figure 6-40.

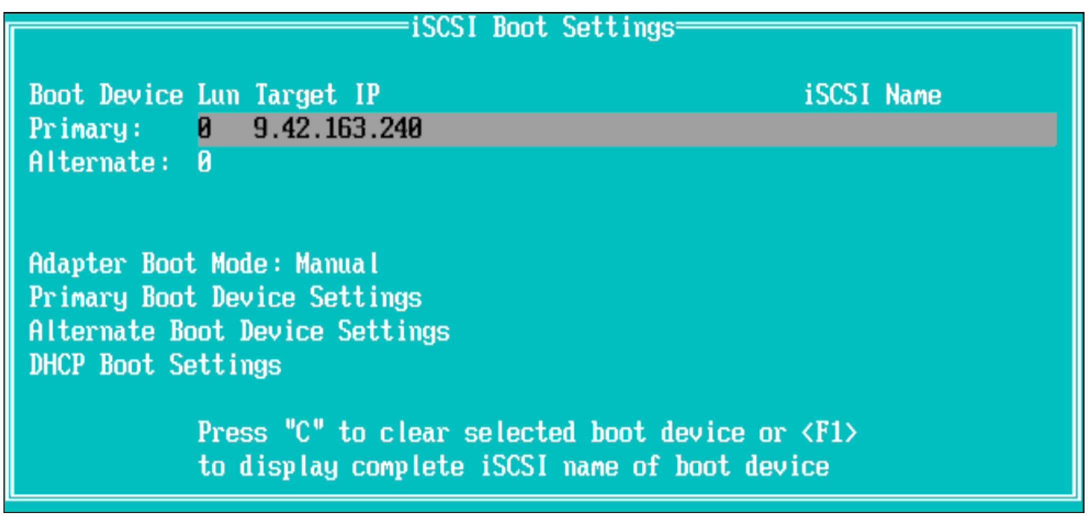


Figure 6-40 iSCSI Boot Settings

5. Press Enter on **Primary:** to scan and select the boot LUN, as shown in Select Boot LUN (Figure 6-41). Press Enter again to add it to the definition as shown in Figure 6-42.

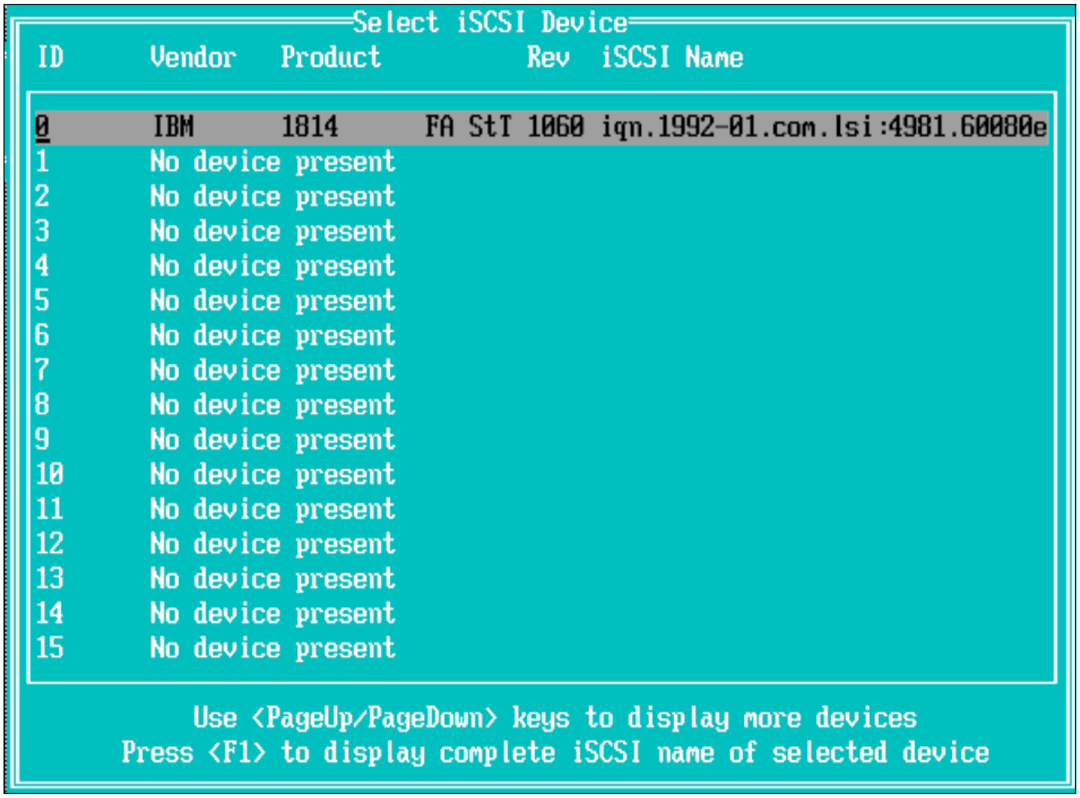


Figure 6-41 Select Boot LUN

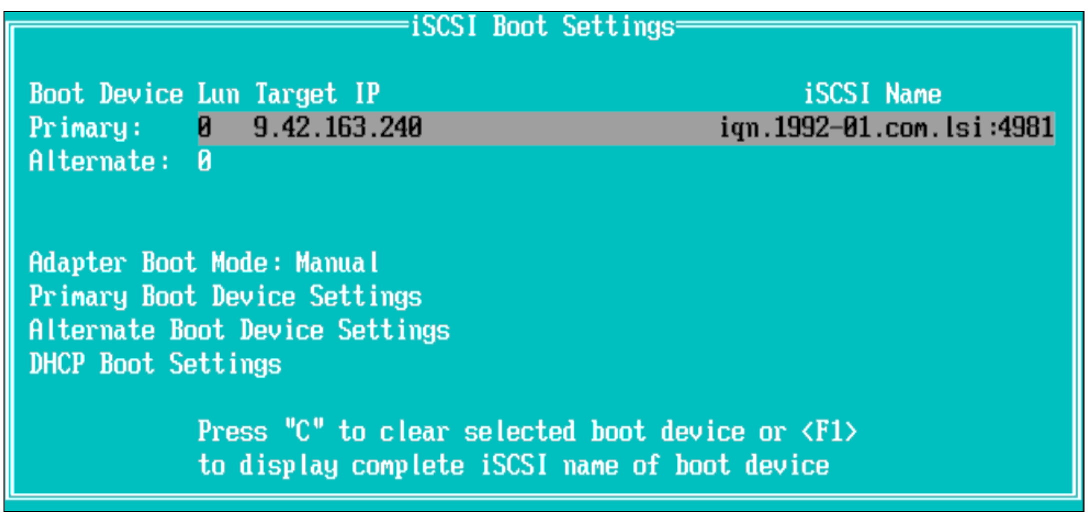


Figure 6-42 Primary Boot device

6. To save changes to the HBA, press Esc. When prompted, select **Save changes**. The iSCSI HBA is now configured to allow for iSCSI boot.

The host and storage configuration process necessary for the primary path is now complete. The operating system installation can now be performed (6.5, “Operating system installation” on page 430). For a trouble-free installation, ensure that there is only one path (primary path) available to the operating system during the installation.

6.5 Operating system installation

At the time of writing of this book, SAN Boot using iSCSI hardware initiator (HBA) is supported for *Red Hat Enterprise Linux v5.4* and *SUSE Linux Enterprise Server v10 SP3*. In the following pages, we give instructions for both operating systems.

The installation process is similar to local disk installation. Start by inserting the installation media of the operating system installer into the server's CD/DVD drive and boot the server from CD/DVD.

The installation process explained here assumes that the server does not have any special hardware (iSCSI card) that requires a specific Linux driver not included on the installer CD.

6.5.1 Implementing iSCSI SAN Boot using Red Hat Enterprise Linux v5.4

Ensure that there are multiple paths to the LUNS before starting the installation process. If there is only one path to the LUNS, Anaconda installs the operating system on the SCSI device even if mpath is specified as a kernel parameter. Enabling multipath features for the installation partitions on iSCSI SAN devices is supported in RHEL 5.3 and later.

Before starting the installation tasks, ensure that both paths between the host and boot LUN exist. In RHEL 5.3 and above, enabling mutipathing is supported at the time of install.

Perform the following tasks:

1. Begin the installation program.
2. A welcome panel is displayed, as shown in Figure 6-43. To enable the multipathing feature at the time of install, we type **`linux mpath`** at the kernel boot line.

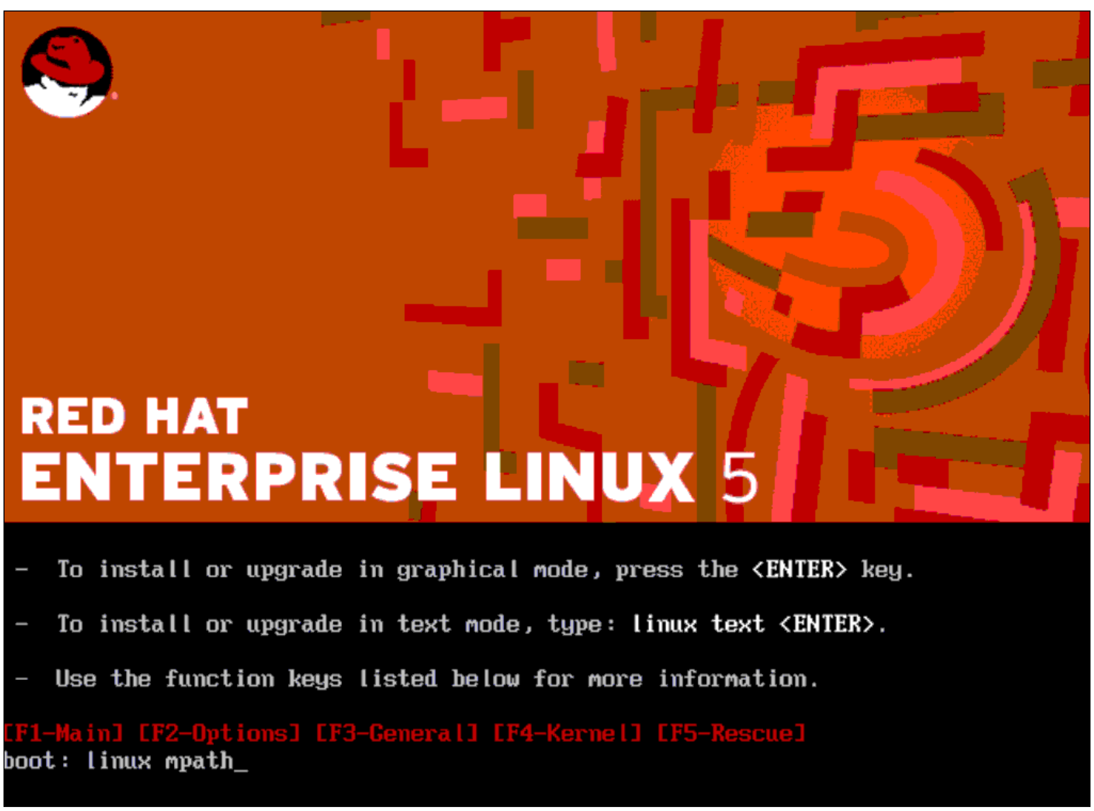


Figure 6-43 Welcome panel

3. In the next window, you should see the multipath device that you can create partitions on, as shown in Figure 6-44. Check the **Review and modify partitioning layout** check box.

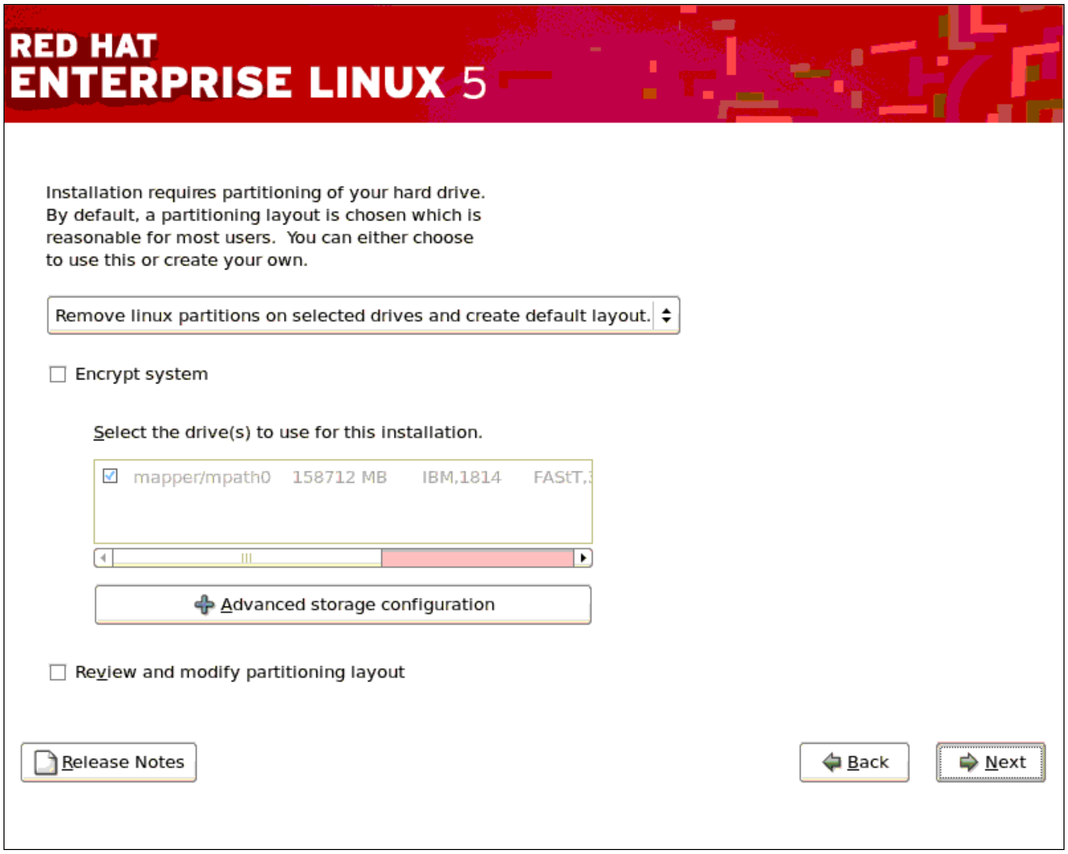


Figure 6-44 multipath device

4. The next window displays additional details regarding the default RHEL partitioning and filesystem layout and the use of the related multipath devices, as shown in Figure 6-45. The multipathing feature will be enabled for all the operating system partitions created on a multipath device during installation.

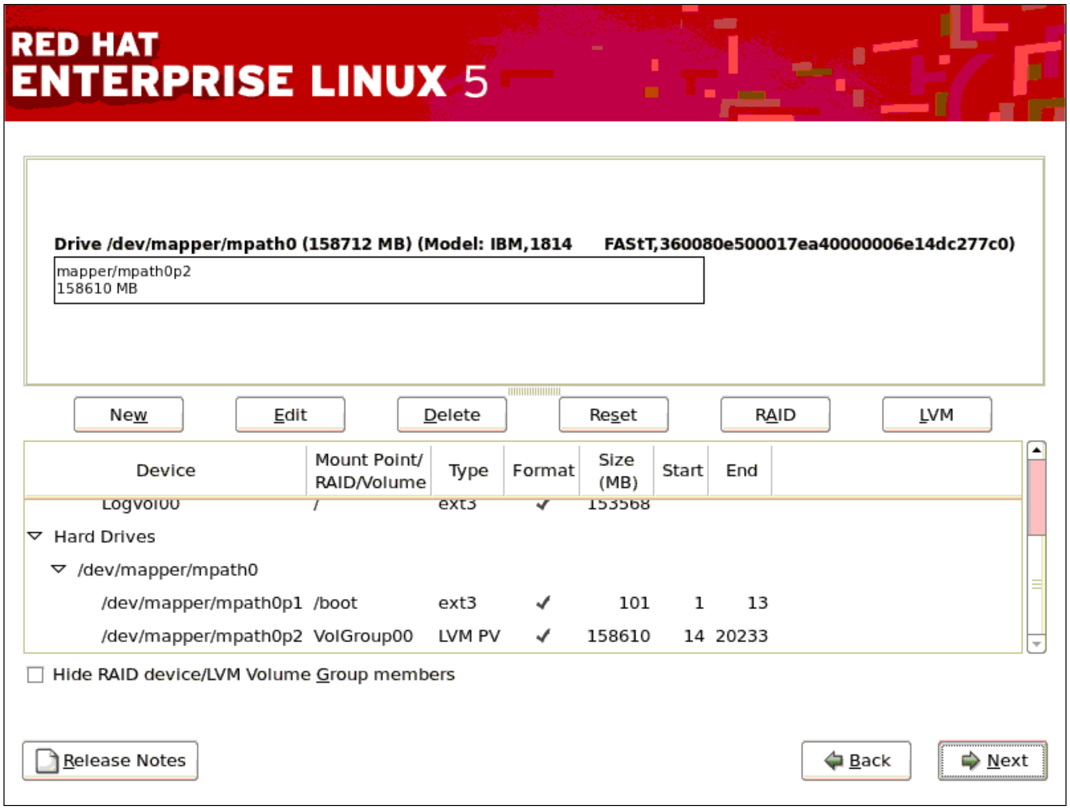


Figure 6-45 Device mapper

5. You can also verify the boot loader and the Master Boot Record it'll be installed on by clicking the **Advanced Storage Configuration** button in Figure 6-44 on page 431. The options are shown in Figure 6-46 and Figure 6-47.

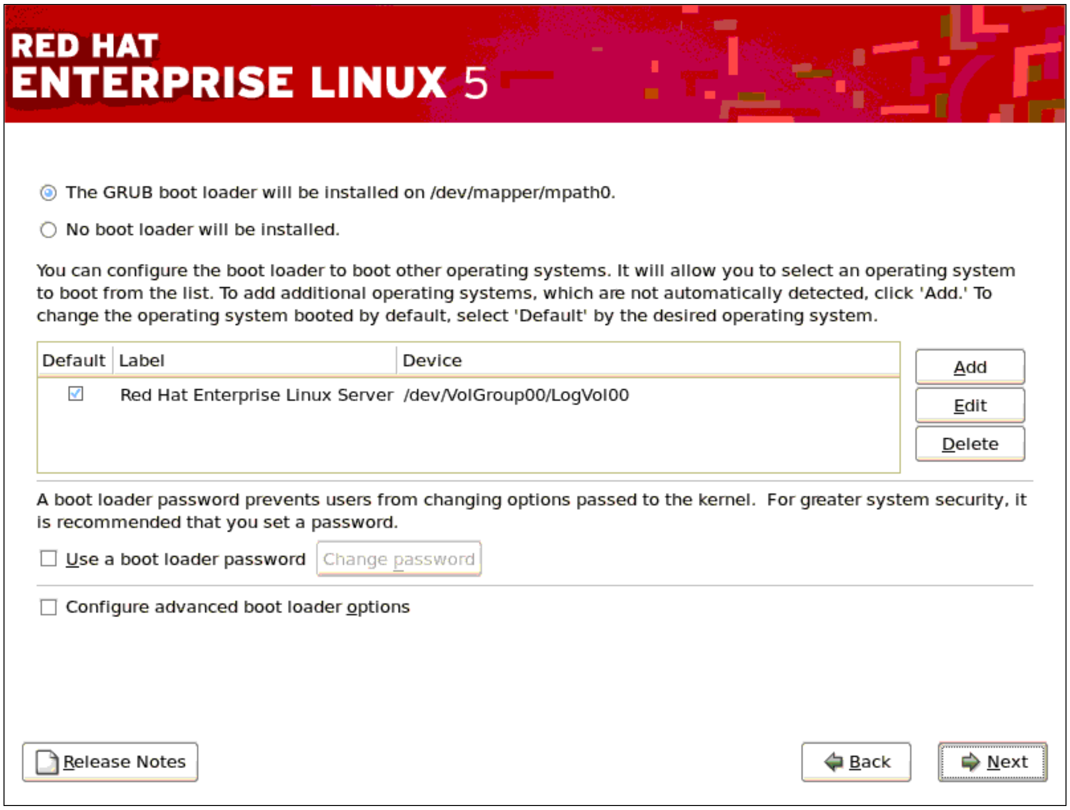


Figure 6-46 Boot Loader

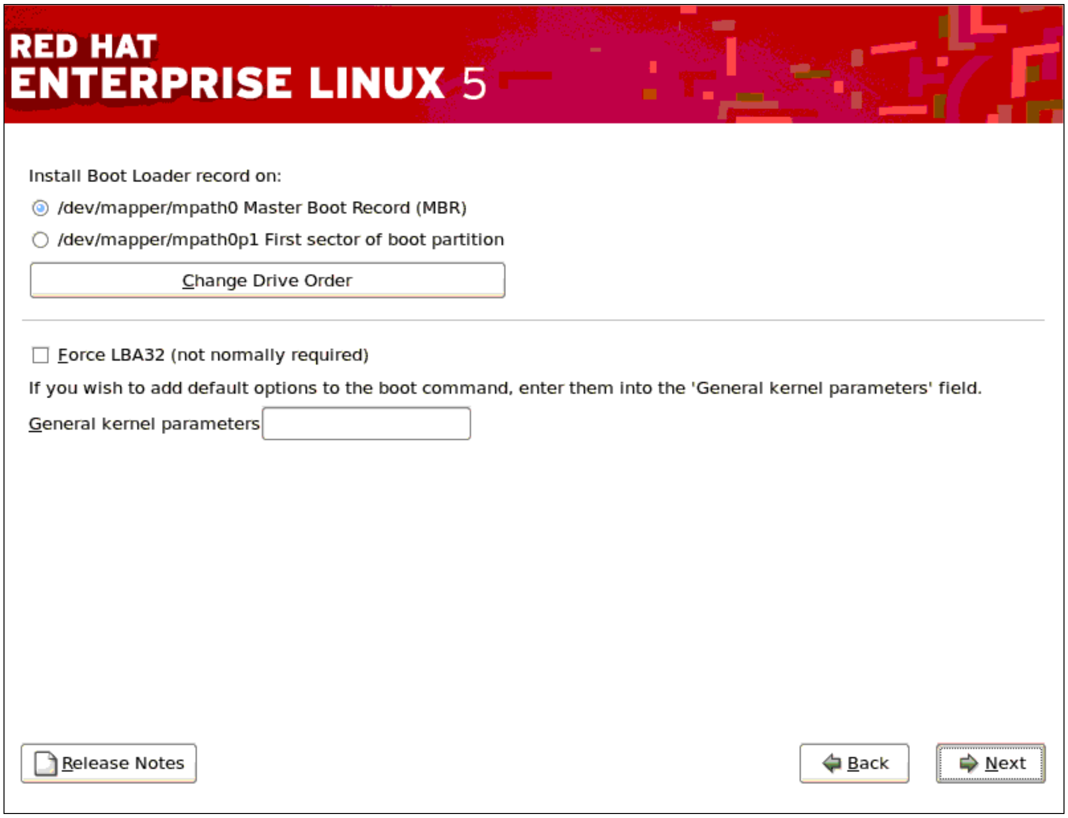


Figure 6-47 Master Boot Record (MBR)

6. The installer now tries to initialize the disk. Because it is a newly created logical drive that has no partition table yet, we get a warning, as shown in Figure 6-48, to initialize the disk. Click **Yes** to erase all existing data on the logical volume and to proceed with the installation, as shown in Figure 6-48.

> **Tip:** If you receive a message reporting an I/O error when initializing the disk, review all steps from the beginning. A possibility is that the logical disk is currently owned by the other controller in the DS5000 storage subsystem, or cabling or zoning are not properly configured.

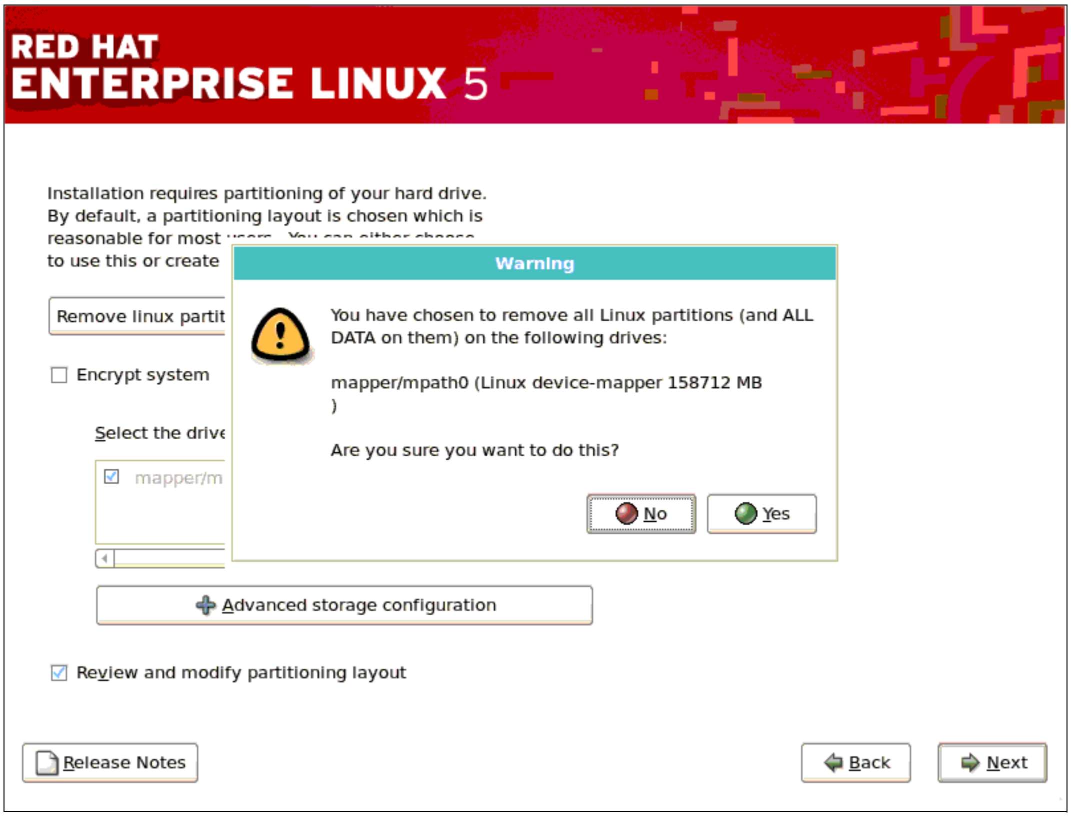


Figure 6-48 Partitioning Layout

7. Follow the installation wizard until you reach the Package Group Selection. Install the Software Development Tools package (Figure 6-49), as it contains tools to compile certain third party drivers that might be used later.

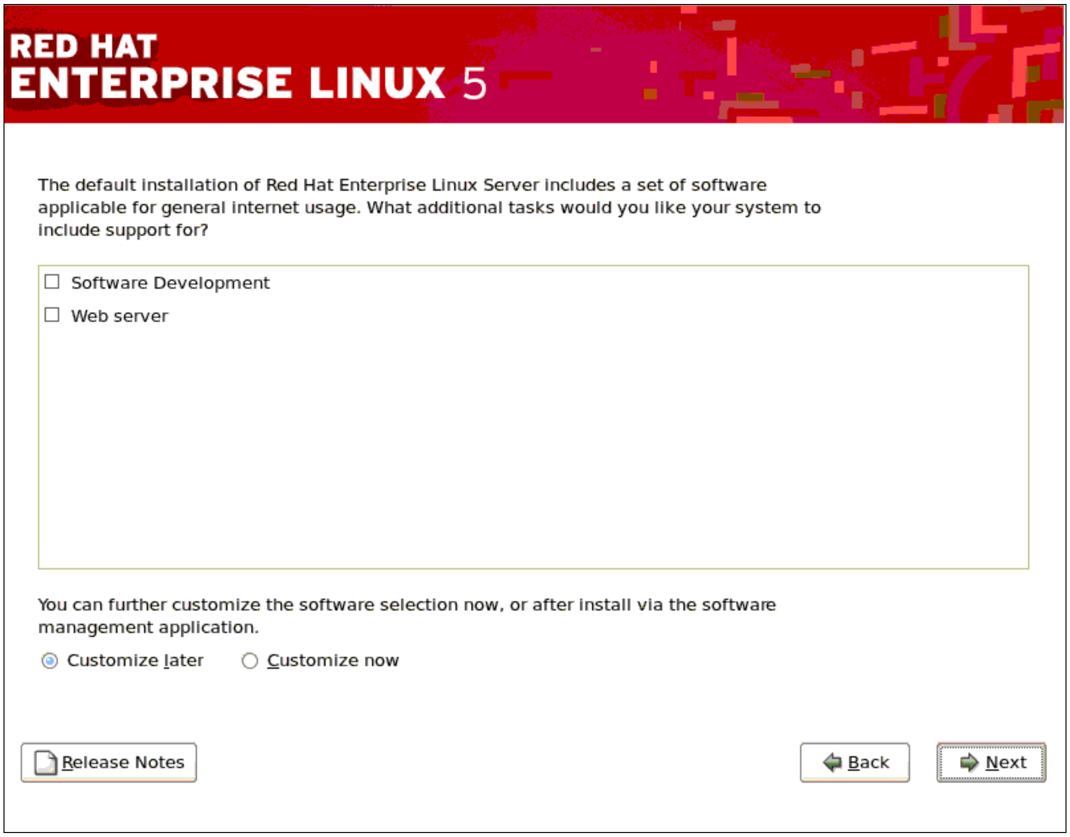


Figure 6-49 Optional Packages for Installation

8. After you have completed the package group selection installation, proceed with the remaining steps of the installation wizard until finished and reboot the server (Figure 6-50). If there was no error during the installation process, the server can boot and load the Linux OS you just installed on the DS5000 storage subsystem disk.

Figure 6-50 Installation Complete

6.5.2 Implementing iSCSI SAN Boot using SUSE Enterprise Linux Server 11 SP 3

It is recommended to have a single path to the boot volume while installing SUSE Linux Enterprise Server 10 SP 3.

Perform the following steps to install the operating system:

1. For iSCSI SAN Boot, Provide the parameters **“`withiscsi=1 netsetup=1`”** to the kernel boot line, as shown in Figure 6-51.

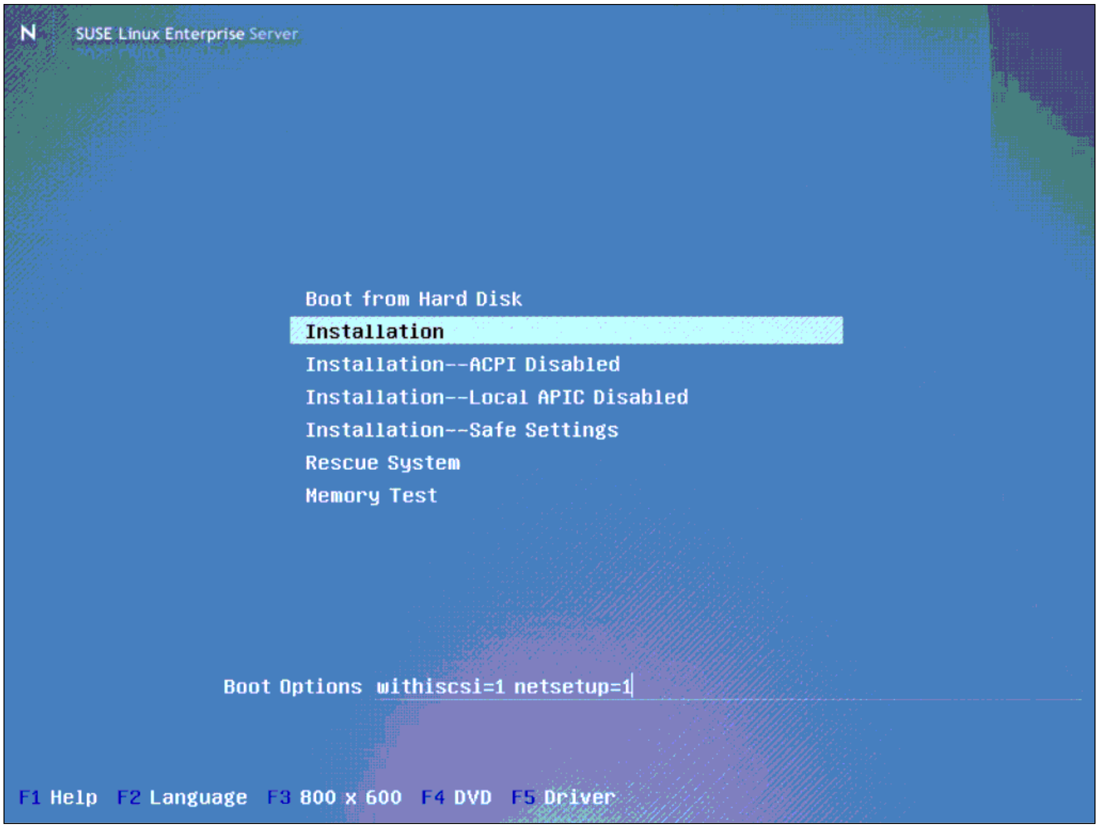


Figure 6-51 SUSE Linux Enterprise Server - Welcome panel

2. Select the device connected to the iSCSI SAN, as shown in Figure 6-52.

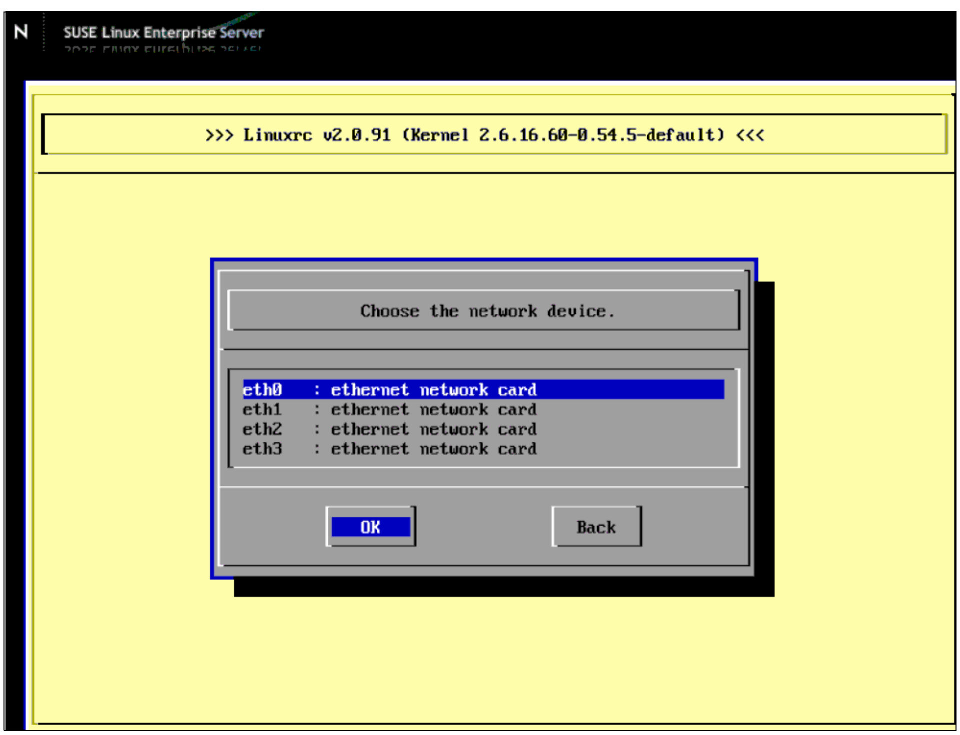


Figure 6-52 Select Ethernet card

3. Based on your network, select **Dhcp** or **Static IP**, as shown in Figure 6-53.

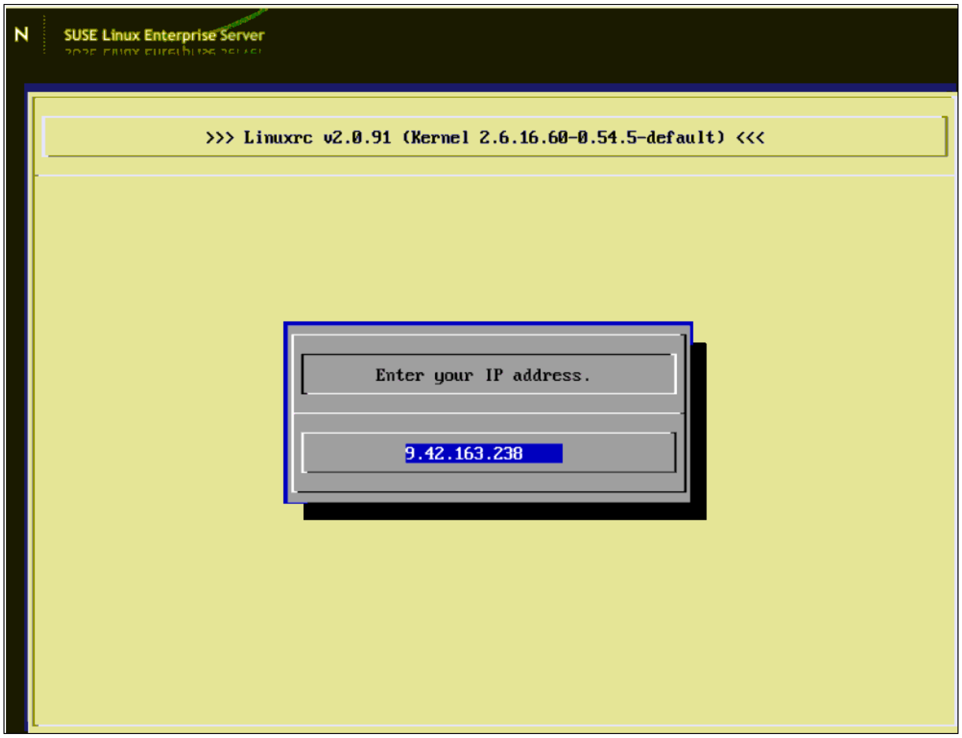


Figure 6-53 Static IP

4. Set the network mask, gateway and nameserver information. On the next panel you can choose Language, then review and accept the License Agreement.

5. The next panel displays the iSCSI initiator information, as shown in Figure 6-54.

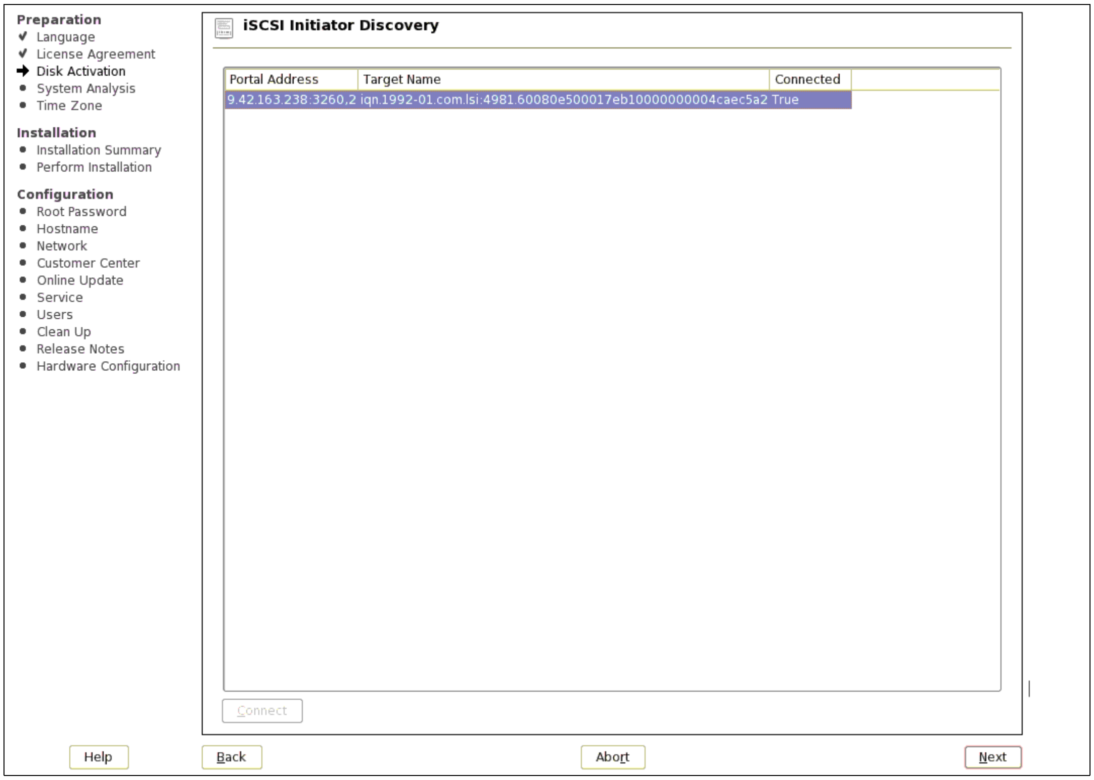


Figure 6-54 iSCSI initiator Discovery

Tip: Ensure that only one target IP is connected and the Start-Up is set to automatic.

6. Select Installation mode as **New Installation**, as shown in Figure 6-55.

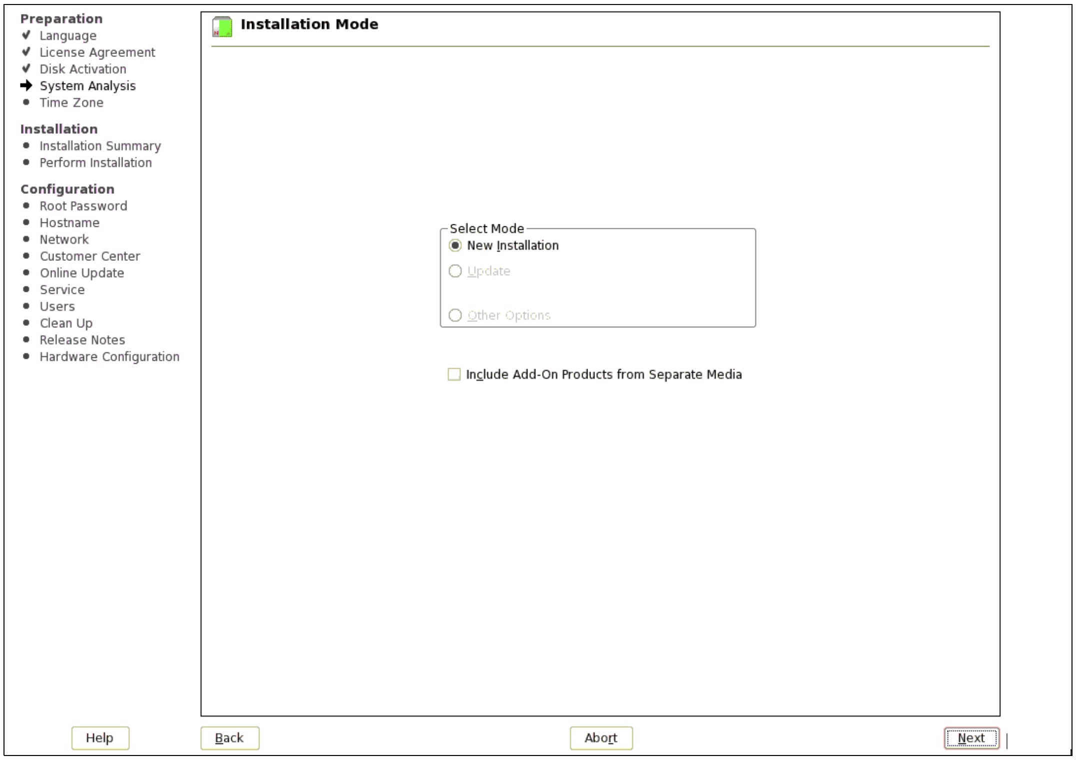


Figure 6-55 New Installation

7. The next window shows the clock and zone information. After that comes the Installation Settings window. It gives an overview of the partitions where installation will be performed (Figure 6-56). You can change the installation settings at this stage. Click **Accept** after verifying the settings that you want.

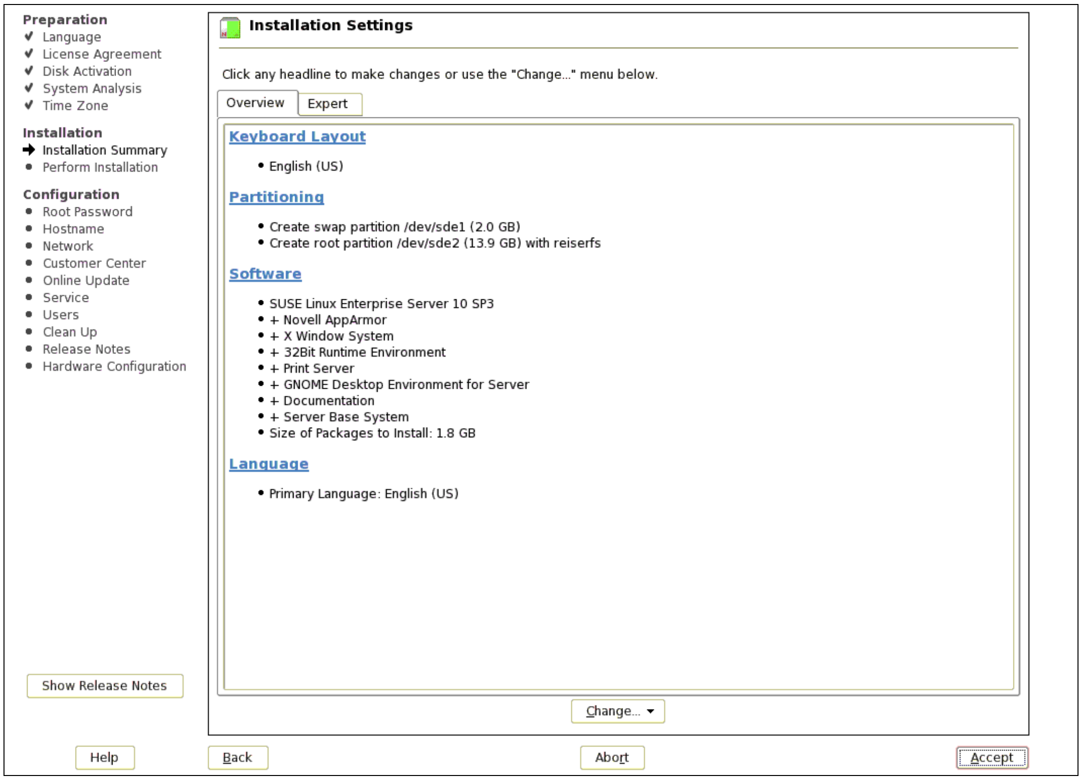


Figure 6-56 Installation Settings

8. The basic installation process will take place, as shown in Figure 6-57. On completion, system will reboot.

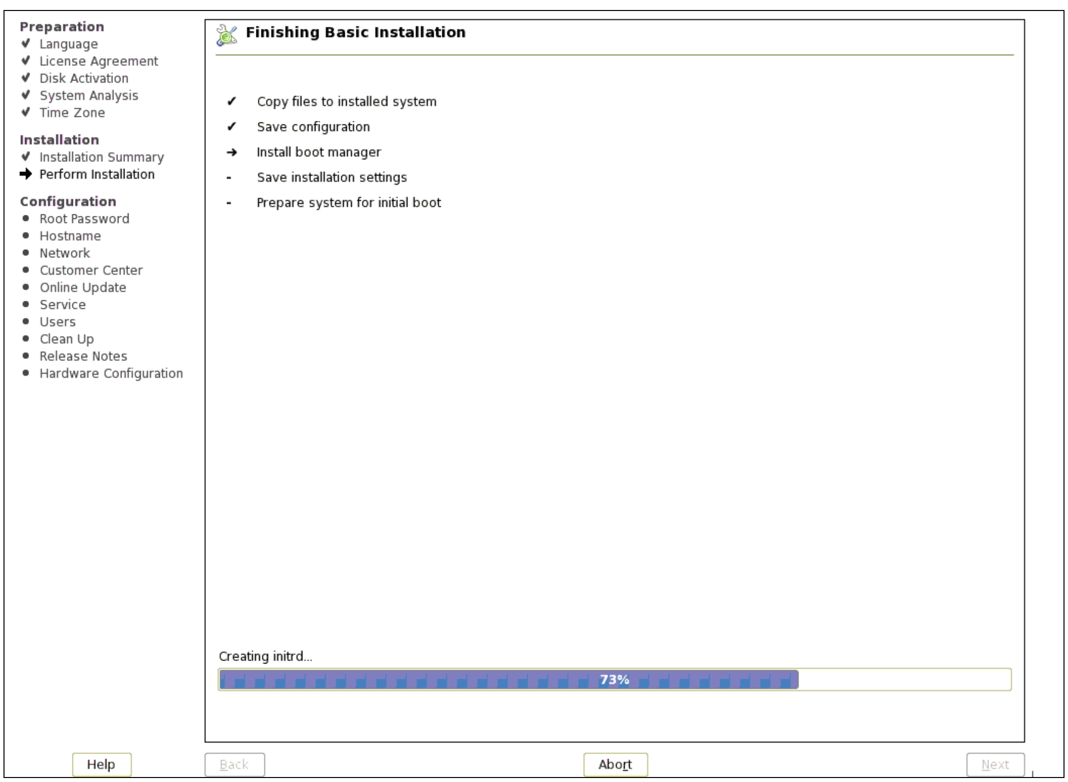


Figure 6-57 Basic Installation

9. Enter the password for the root user. Next, update the Hostname and Domain Name as appropriate. Click **Next**.
10. Verify the Network Configuration. You might want to disable the firewall if you are testing in a secure lab environment. You can configure other NICs in the system at this time as well. Click **Next**.
11. If you want to test the Internet Connection, make sure you select the NIC that is connected to your LAN that is connected to the Internet. Click **Next**.
12. Change the Installation Settings as required. Next, you can select the User Authentication Method as appropriate. Click **Next**.
13. Enter (local) user information. Click **Next**. Browse the release notes. Click **Next**.
14. Test the configuration to be sure the video mode is compatible. Click **Next**.
15. The installation is now complete, and we proceed with the multipath configuration.

6.5.3 Configuring multiple paths for redundancy

We use Device Mapper Multipath (DMM) to configure multipaths in SUSE Linux Enterprise Server 10 SP3 for DS5000 Subsystems.

Device mapper is a block subsystem that provides layering mechanism for block devices. It allows you to configure multiple I/O paths between server nodes and storage arrays creating a new virtual device that manages the different paths in order to provide failover in an active/passive storage subsystem.

DMM resolves all the issues that arise in accessing a multipathed device in Linux. It also provides a consistent user interface for storage devices provided by multiple vendors. There is only one block device (/dev/mapper/XXX) for a LUN and it is the device created by device mapper. Paths are grouped into priority groups, and one of the priority group will be used for I/O, and is called active. A path selector selects a path in the priority group to be used for an I/O based on some load balancing algorithm (for example round-robin).

When a I/O fails in a path, that path gets disabled and the I/O is re-tried in a different path in the same priority group. If all paths in a priority group fails, a different priority group which is enabled will be selected to send I/O.

DMM consists of four components:

1. **`dm-multipath:`** This module is responsible for the multipathing decisions in event of a failure.
2. **`multipath:`** This command is used to detect multiple devices for failover, to list the available paths and to remove virtual devices.
3. **`multipathd daemon:`** This daemon constantly checks the paths to mark them as failed when it finds a path is faulty and switches the I/O to the next enable priority group. It keeps checking the failed path, after the failed path comes alive, based on the failback policy, it can activate the path.
4. **`multipathd command:`** When used with the **`-k`** parameter, it is possible to run a command line interface as shown in Example 6-1.

Example 6-1 multipathd command line

```
# multipathd -k
multipathd> help
multipath-tools v0.4.8 (08/02, 2007)
CLI commands reference:

list|show paths
 list|show paths format $format
 list|show status
 list|show maps|multipaths
 list|show maps|multipaths status
 list|show maps|multipaths stats
 list|show maps|multipaths topology
 list|show topology
 list|show map|multipath $map topology
 list|show config
 list|show blacklist
 list|show devices
 list|show wildcards
 add path $path
 remove|del path $path
 add map|multipath $map
 remove|del map|multipath $map
 switch|switchgroup map|multipath $map group $group
 reconfigure
 suspend map|multipath $map
 resume map|multipath $map
 resize map|multipath $map
 reload map|multipath $map
 reinstate path $path
 fail path $path
multipathd>
```

Installing the DMM multipath driver

Follow these steps to install the DMM multipath driver:

Setting HBA timeouts

Modify the /etc/modprobe.conf.local file and add the following lines to this file:

```
options lpfc_nodev_tmo 10
options qla2xxx qlport_down_retry=10
```

Setting up multipath.conf

It is very important to set up this file as it is consulted by the multipath command for the configuration of multipath devices. It is essentially the configuration file for devices virtualized by device mapper.

The file has four sections:

- Defaults: Specifies system level default override
- Blacklist: Specifies a list of devices to not be multipathed
- Devices: Specifies the list of devices that the setting will be applied to
- Multipaths: Specifies configuration settings for tuning

Create the /etc/multipath.conf file as shown in Example 6-2.

Example 6-2 Multipath.conf file sample

```
multipaths {
        multipath {
        wwid                    360080e50001b0c90000007504c6e060a
        alias                   mpath0
        path_grouping_policy    multibus
        path_checker            readsector0
        path_selector           "round-robin 0"
        failback                "5"
        rr_weight               priorities
        no_path_retry           "5"
        }
}
devices {
      device {
      vendor "IBM"
      product "1746"
      hardware_handler "1 rdac"
      path_checker rdac
      failback 0
      path_grouping_policy multibus
      prio_callout "/sbin/mpath_prio_tpc /dev/%n"
      }
}
```

The vendor field must be the same as in the */sys/block/sdb/device/vendor* file and the product field must be the same as in the */sys/block/sdb/device/model* file

Tip: In this case, /dev/sdb is one of the paths to the logical drive

It is highly recommended to use alias for naming the devices instead of using the uid of the device. In this case we choose to use the name mpath. To choose the name, consider that the devices will be created as /dev/mapper/mpath[n]. This name will be created under the /dev/mapper directory.

Configuring services

Configure the multipathd service to start on boot with the following command:

```
chkconfig multipathd on
```

6.5.4 Scan and manage the storage logical drive

To scan for new devices run the following command

```
# rescan-scsi-bus.sh
# multipath
```

To list all the available paths use the following command (output is shown in Example 6-3):

```
#multipath -ll
```

This command is used to list all of the “virtual devices” created by device mapper, and shows the physical paths associated with each.

Example 6-3 Output of multipath -ll command

```
# multipath -ll
mpath0 (360080e50001b0c90000007504c6e060a) dm-0 IBM,1746      FAStT
[size=15G][features=1 queue_if_no_path][hwhandler=1 rdac][rw]
\_ round-robin 0 [prio=4][active]
 \_ 4:0:0:0 sdb 8:16  [active][ready]
 \_ 5:0:0:0 sdc 8:32  [active][ghost]
 \_ 5:0:1:0 sdd 8:48  [active][ghost]
 \_ 4:0:1:0 sde 8:64  [active][ready]
```

Example 6-3 shows that the disks, in this case paths, /dev/sdb and /dev/sdc are ready. The reason is that those paths are connected to the controller that is owner of the logical drive in an active/passive array storage.

The other two paths (/dev/sdc and /dev/sdd) are in standby mode (ghost). In the event when both actives paths fail, device mapper will failover to the active ghost paths. Afterwards, the Storage subsystem will change the Logical Drive ownership to the secondary Storage controller and there will be a message in Storage Recovery Guru saying that Logical Drive is not in the preferred path.

Partitioning the virtual device

According to the Example 6-3, we have four paths to the logical drive and we can create our partition on one of the paths as shown in Example 6-4.

Example 6-4 Creating a 100MB partition with fdisk

```
Linux:/dev/mapper # fdisk /dev/sdb
Device contains neither a valid DOS partition table, nor Sun, SGI or OSF disklabel
Building a new DOS disklabel with disk identifier 0xebba465c.
Changes will remain in memory only, until you decide to write them.
After that, of course, the previous content won't be recoverable.
The number of cylinders for this disk is set to 15360.
There is nothing wrong with that, but this is larger than 1024,
and could in certain setups cause problems with:
1) software that runs at boot time (e.g., old versions of LILO)
2) booting and partitioning software from other OSs
   (e.g., DOS FDISK, OS/2 FDISK)
Warning: invalid flag 0x0000 of partition table 4 will be corrected by w(rite)

Command (m for help): n
Command action
   e   extended
   p   primary partition (1-4)
p
Partition number (1-4): 1
First cylinder (1-15360, default 1):
Using default value 1
Last cylinder, +cylinders or +size{K,M,G} (1-15360, default 15360): +100M

Command (m for help): w
The partition table has been altered!

Calling ioctl() to re-read partition table.
Syncing disks.
```

We now need to create a map from the previous partition table. In order to do that, run the following command:

```
# kpartx -a /dev/mapper/mpath0
```

List the /dev/mapper directory and the partitions on /dev/dm-0 to verify the correct partitioning as shown in Example 6-5.

Example 6-5 very the partitioning

```
# ls /dev/mapper/
control  mpath0  mpath0p1

# fdisk -l /dev/dm-0
Disk /dev/dm-0: 16.1 GB, 16106127360 bytes
64 heads, 32 sectors/track, 15360 cylinders
Units = cylinders of 2048 * 512 = 1048576 bytes
Disk identifier: 0xebba465c

     Device Boot      Start         End      Blocks   Id  System
/dev/dm-0p1               1         101      103408   83  Linux
```

The last step is to format the partition and mount it with the following commands:

```
# mkfs.ext3 /dev/mapper/mpath0p1
# mount /dev/mapper/mpath0p1 /mnt
```

Important: You must make sure you are running I/O to the virtual devices created by device mapper, and not the physical device paths. If you are running I/O to the physical paths, device mapper will not be able to manage a failover and the I/O will fail.

Removing virtualized devices

The following multipath commands will “flush” all unused device mappings and it is not possible to run this command if the device is mounted.

```
multipath -F
```

After flushing the paths, use the *multipath* command with no options to add back mappings for the devices.

Tip: It is not possible to flush the mappings while the device is being used by the system. Unmount any filesystems for /dev/dm-* devices before running this command.

The same command with the -f parameter can be used to remove a single device mapping as shown in Example 6-6. Run *multipath -ll* and find the “dm-*” device you want to remove. Find the scsi id of the device to select the device.

Example 6-6 Remove a single device mapping

```
# multipath –f 360080e50001b0c90000007504c6e060a
```

6.5.5 Failover/failback verification test

To verify that the server can successfully access the secondary path in case of a primary path failure while Linux is booted up completely, perform the following tasks:

1. From the Storage Manager Subsystem Management window (Logical view), check which controller currently has the drive ownership of the boot disk for the server. In optimal condition, the preferred and current controller must be the same (Figure 6-58).

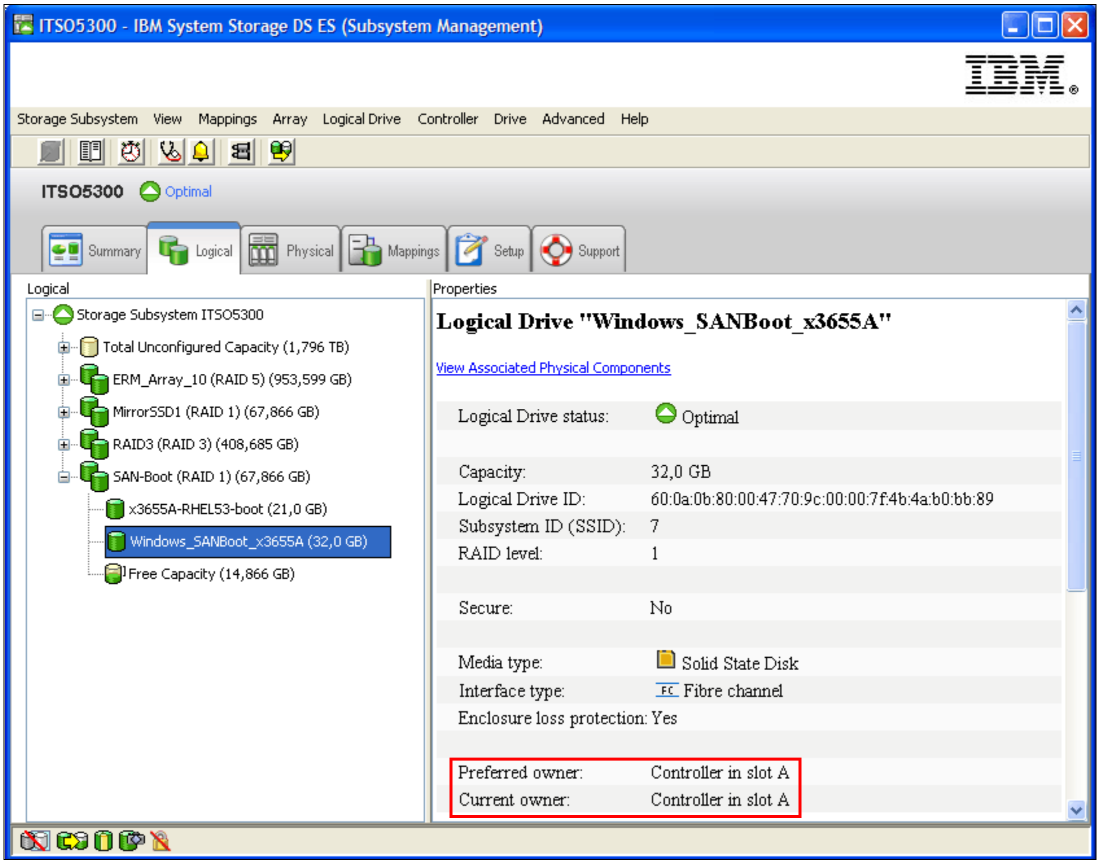


Figure 6-58 Check controller ownership

2. Now disconnect the cable on the primary server host port 1.
3. Within 120 seconds, the multipath driver will cause a failover to the secondary path, which is attached to controller B on the DS5000 storage subsystem (Figure 6-59).

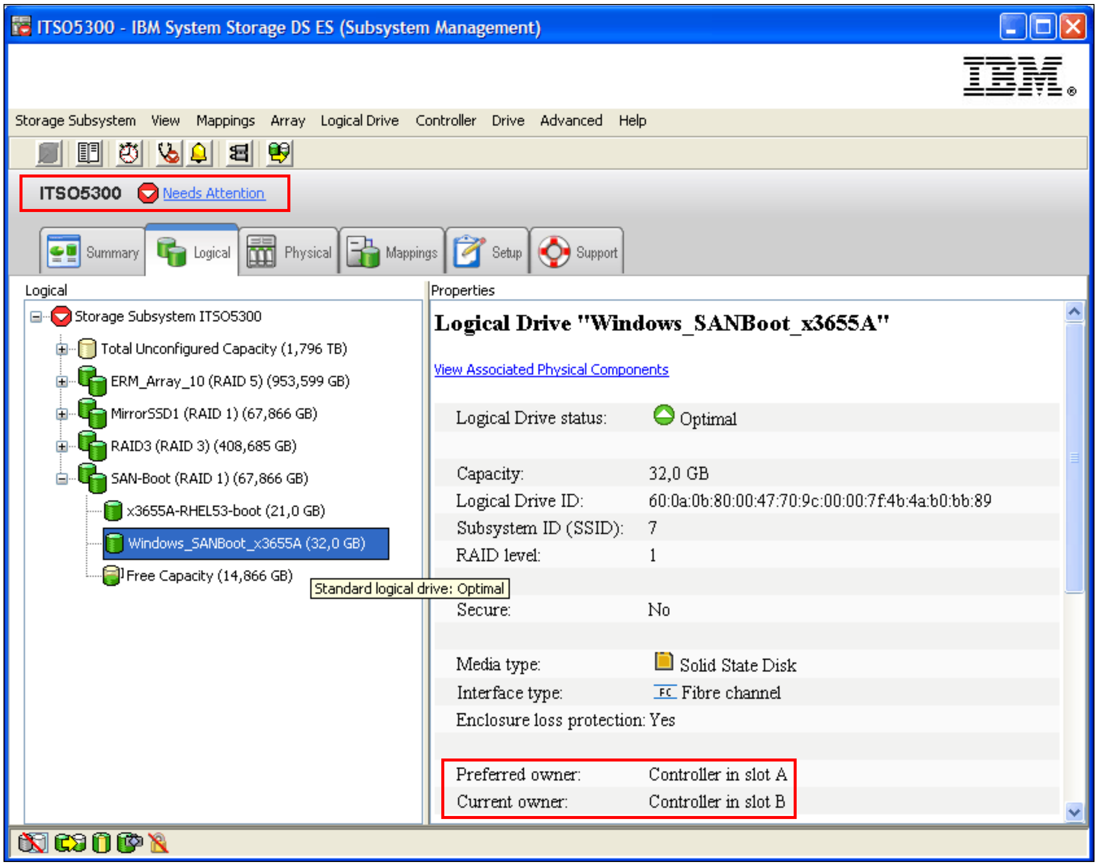


Figure 6-59 Controller B takes ownership

4. After the failover is done, connect the cable back into the server HBA port. You can see the path moving back to its primary controller within the next 120 seconds automatically.

The configuration and verification process for the Linux iSCSI boot from SAN is now complete.

If you reboot your server while the LUN ownership is not on the preferred controller, you need to enter Fast!Util during the boot sequence (press Ctrl+q when prompted), disable the BIOS of the first HBA path, and then enable the BIOS of the second HBA path and configure the boot device from the second HBA path.

> **Best practice:** Enable the BIOS on only one HBA path to avoid confusion. In Linux, you can only boot from SAN if the controller that is visible from the BIOS-enabled HBA is the owner of the boot LUN.

6.5.6 Troubleshooting

Here we present a few common problems and ways to troubleshoot.

rescan_scsi_bus.sh is not showing any new device

If the server did not recognize any new mapped LUNs after running the **`rescan_scsi_bus.sh`** command a few times, consider rebooting the server.

multipath –ll output does not show all the LUNs

In this case, most probably the Device Mapper is not seeing all the LUNs mapped to it. This condition can occur due to various reasons,

If mappings are edited with the host connected to the array. The host in some case might fail to monitor for newly added LUNs and there will not be a scsi block device created for these newly added LUNs. To fix this issue rescan for the scsi devices using the command "rescan_scsi_bus". This should create a scsi device for the newly added LUNs and run "multipath" to add the scsi devices to device mapper table.

multipath does not show paths equal to the number of LUNs mapped

The first thing that you should check is the 'lsscsi' output to see if the SCSI midlayer is seeing (# of paths x # of LUNs) number of block devices.

If not, issue a device rescan by **`rescan-scsi-bus.sh`**. It should fix the issue, else you need to reboot the host.

If "lsscsi" is seeing all the devices, then do the following steps:

- Issue **`multipath`** so the device mapper can add the devices seen by lsscsi.
- Flush all the mappings and re-add them (**`multipath –F ; multipath`**).
- Restart multipath daemon (**`service multipathd restart`**).

multipath -ll output is showing the paths as [failed][faulty]

This condition can occur if the physical path to a controller fails or the controller itself fails and failover happens:

- Verify the disk uid and parameters in multipath.conf file.
- Verify cables and zoning configuration.
- If this happens after an error in test and you want to make the configuration optimal, save the logs first, then flush and re-add multipath devices, restart multipath daemon. If the issue still persists, reboot the system.

6.6 More information

As the support for iSCSI SAN Boot is increasing on IBM products, refer to the latest interoperability information at the IBM website: *System Storage Interoperation Center (SSIC)*, at this website:

http://www-03.ibm.com/systems/support/storage/ssic/interoperability.wss

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this book.

IBM Redbooks publications

The following IBM Redbooks publications provide additional information about the topic in this document. Note that some publications referenced in this list might be available in softcopy only.

- *IBM System Storage DS4000 and Storage Manager V10.30*, SG24-7010
- *IBM Midrange System Storage Hardware Guide*, SG24-7676
- *IBM Midrange System Storage Implementation and Best Practices Guide*, SG24-6363
- *Implementing the IBM Storwize V7000*, SG24-7938
- *SAN Volume Controller Best Practices and Performance Guidelines*, SG24-7521
- *Implementing the IBM System Storage SAN Volume Controller V6.1*, SG24-7933
- *IBM System Storage DS8700 Architecture and Implementation*, SG24-8786
- *IBM System Storage DS8000 Host Attachment and Interoperability*, SG24-8887
- *IBM System Storage DS8800: Architecture and Implementation*, SG24-8886
- *IBM XIV Storage System: Architecture, Implementation, and Usage*, SG24-7659
- *IBM XIV Storage System: Copy Services and Migration*, SG24-7759
- *XIV Storage System: Host Attachment and Interoperabilty*, SG24-7904
- *IBM System Storage b-type Multiprotocol Routing: An Introduction and Implementation*, SG24-7544
- *Implementing an IBM b-type SAN with 8 Gbps Directors and Switches*, SG24-6116
- *Implementing an IBM/Cisco SAN*, SG24-7545

You can search for, view, download or order these documents and other Redbooks, Redpapers, Web Docs, draft and additional materials, at the following website:

ibm.com/redbooks

Other publications

These publications are also relevant as further information sources:

- *IBM System Storage DS: Command-Line Interface User's Guide*, GC53-1127.
- *IBM System Storage DS8700 and DS8800 Introduction and Planning Guide*, GC27-2297
- *Multipath Subsystem Device Driver User's Guide*, GC52-1309-02.

© Copyright IBM Corp. 2012. All rights reserved.

Online resources

These websites are also relevant as further information sources:

- IBM support website for HBA support on System x:

  http://www-03.ibm.com/systems/xbc/cog

- System Storage Interoperation Center (SSIC):

  http://www-03.ibm.com/systems/support/storage/config/ssic/displayesssearchwithoutjs.wss?start_over=yes

- LSI RDAC Driver Packages:

  http://www.lsi.com/rdac

- VMware documentation:

  http://www.vmware.com/support/pubs/vs_pubs.html

- VMware vSphere 4.1 Documentation:

  http://www.vmware.com/support/pubs/vs_pages/vsp_pubs_esxi41_i_vc41.html

- XIV Management Tools:

  http://www.ibm.com/support/search.wss?q=ssg1*&tc=STJTAG+HW3E0&rs=1319&dc=D400&dtm

- Storage Networking Industry Association (SNIA):

  http://www.snia.org

Help from IBM

IBM Support and downloads:

ibm.com/support

IBM Global Services:

ibm.com/services

Index

© Copyright IBM Corp. 2012. All rights reserved.

E

F

G

H

I

J

K

L

S

SAN Boot Implementation and Best Practices Guide for IBM System Storage

(1.0” spine)
0.875”<->1.498”
460 <-> 788 pages

SAN Boot Implementation and Best Practices Guide for IBM System Storage

Implementation examples using DS5000, DS8000, XIV, Storwize V7000, and SVC

Windows, Red Hat Linux, SUSE Linux, and VMware implementations

Best Practice guidelines

Booting servers from a storage area network (SAN) is being used increasingly in complex data center environments today, due to its significant benefits over the traditional method of booting from local disks. SAN Boot enables organizations to maximize consolidation of their IT resources, minimize their equipment costs, and realize the considerable management benefits of centralizing the boot process.

In SAN Boot, you can deploy diskless servers in an environment where the boot disk is located on (often RAID-capable) storage connected to the SAN. The server (initiator) communicates with the storage device (target) through the SAN using the Fibre Channel host bus adapter (HBA).

The system downtime is greatly minimized in case a critical component such as a processor, memory, or host bus adapter fails and needs to be replaced. The system administrator needs to swap only the hardware and reconfigure the HBA's BIOS, switch zoning, and host-port definitions on the storage server. The system image still exists on the logical drive, therefore the server is fully operational after the hardware swap and configuration change is completed.

This IBM Redbooks publication can help you with the SAN Boot implementation. We present various SAN Boot scenarios using IBM System Storage products that include DS5000, DS8000, XIV, and SVC. The operating systems that are covered include Windows 2008, Red Hat Linux, SUSE Linux, and VMware.

INTERNATIONAL TECHNICAL SUPPORT ORGANIZATION

BUILDING TECHNICAL INFORMATION BASED ON PRACTICAL EXPERIENCE

IBM Redbooks are developed by the IBM International Technical Support Organization. Experts from IBM, Customers and Partners from around the world create timely technical information based on realistic scenarios. Specific recommendations are provided to help you implement IT solutions more effectively in your environment.

For more information:
ibm.com/redbooks

SG24-7958-00 ISBN 0738437069