Practical Guide for SAN with pSeries

Hints and tips for pSeries and AIX

Comprehensive planning information

Structured problem determination guide

Daesung Chung
Jim Blue
Andrew Carr
Yonghee Kim

Redbooks

ibm.com/redbooks

International Technical Support Organization

Practical Guide for SAN with pSeries

March 2002

SG24-6050-00

Take Note! Before using this information and the product it supports, be sure to read the general information in “Special notices” on page 233.

First Edition (March 2002)

This edition applies to both AIX 4.3.3 and 5.1.

Comments may be addressed to:
IBM Corporation, International Technical Support Organization
Dept. JN9B Building 003 Internal Zip 2834
11400 Burnet Road
Austin, Texas 78758-3493

When you send information to IBM, you grant IBM a non-exclusive right to use or distribute the information in any way it believes appropriate without incurring any obligation to you.

Contents

Figures

Tables

Preface

This book provides a single source of the information you will need to build your own SAN fabric, with easy-to-follow instructions and ample illustrations. To allow quick access to the information readers may need, the contents of this redbook have been arranged in three parts: planning, implementation, and problem determination. With this handbook format, both Planners and Engineers in the SAN area will find this book a useful addition to their technical library. This redbook was written with a specific aim. A few SAN Redbooks have been published in the past, but the information specific to IBM pSeries servers has been scattered and hard to find. In this book the authors compile the information that is critical for building a SAN fabric with IBM pSeries servers.

The team that wrote this redbook

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

Daesung Chung is working at ITSO, Austin Center, and is in charge of e-business solutions on IBM @server pSeries and RS/6000. He writes extensively and teaches IBM classes worldwide on e-business solutions and AIX and Linux. His areas of study are implementation, capacity planning, and performance tuning. He also has years of experience in UNIX system management. Before joining ITSO, he was a Consulting IT Specialist at IBM Korea.

Jim Blue is a Software Engineer in Research Triangle Park, North Carolina, and is currently a member of the SAN Central Support Group. He has 18 years of experience in the areas of IT and networking. He holds a BS degree in computer science from North Carolina State University. His areas of expertise include SCSI and other networking protocols.

Andrew Carr is a Senior IT Specialist working as a Techline Storage Specialist in the UK. He has 18 years of experience in the computer industry. He has worked at IBM for five years and previously worked for Storagetek. His areas of expertise include both IBM and other vendor storage systems.

Yonghee Kim is a Senior IT Specialist with the Storage Techline in Korea. She has 14 years of experience in the Information Technology field. Her areas of expertise include VAX, iSeries, pSeries, and storage hardware and software. She has practical experience with SAN planning and design from working with many customers in her country.

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

International Technical Support Organization, Austin Center
Julie Czubik

IBM Austin
John Dieringer

IBM Austin
Rich Avery

IBM SAN Central, Research Triangle Park, North Carolina
Brian Tesar

IBM SAN Central, Research Triangle Park, North Carolina
Jerry Poole

Special notice

This publication is intended to help customers, business partners, and IBM professionals plan and implement a SAN fabric with IBM @server pSeries. The information in this publication is not intended as the specification of any programming interfaces that are provided by AIX. See the PUBLICATIONS section of the IBM Programming Announcement for AIX for more information about what publications are considered to be product documentation.

IBM trademarks

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

3890™	AIX®
AIX 5L™	DB2®
e (logo)®	Enterprise Storage Server™
ESCON®	FICON™
FlashCopy™	HACMP/6000™
IBM ®	iSeries™
Language Environment®	Magstar®
MORE™	Notes®
Perform™	pSeries™
Redbooks™	Redbooks Logo
RS/6000®	S/390®
Seascape®	SP™
StorWatch™	TotalStorage™
Wave®	

Comments welcome

Your comments are important to us!

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

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

  `ibm.com/redbooks`

- Send your comments in an Internet note to:

  `redbook@us.ibm.com`

- Mail your comments to the address on page ii.

1

Planning guide

In this chapter we discuss how to plan a SAN network, what points need to be considered beforehand, and how to select the ideal components to meet the requirements. We also discuss product-specific considerations and zoning considerations.

The following topics will be covered in this chapter:

- Understanding requirements
- Choosing SAN components
- Considerations for zoning
- Considerations for cables
- Product-specific considerations
- FICON interoperability

1.1 Understanding requirements

To develop a solid implementation plan for a SAN environment, we recommend that you study the following points and prepare answers to them:

- Study the specifications and characteristics of the servers and storage devices that will be connected through the SAN fabric. Particularly, pay attention to the following points:
  - Server: Manufacturer and model number
  - Operating system: Name and release level
  - HBA (Host Bus Adapter): Manufacturer, type of adapter, and supported microcode level
  - Storage devices: Manufacturer, model number, and microcode level
  - Expected peak and average throughput of the SAN device measured in MB/sec

  > **Note:** The theoretical bandwidth of Fibre Channel is 100 MB/sec full duplex. However, to develop a realistic performance expectation, we recommend that you base any decisions during the planning stage on the assumption that the throughput rate of the Fibre Channel would be around 60 MB/sec per port. This approach can help users avoid unrealistic performance expectations.

- Study the required level of availability or redundancy. Remember that the more redundancy you have, the higher the cost of your design will become. For instance, we have seen a case in which 20-30 percent of the switch ports were used for redundant inter-switch link (ISL) connections.
- Scalability is another important factor to take into consideration. What is the planned growth rate of the demand for storage devices and/or server machines?
- Identify potential bottlenecks. Which servers or storage devices will get the highest data traffic? In a large SAN implementation, you would want the devices with the highest performance requirements to have the smallest number of inter-switch hops possible. For some products, such as INRANGE 9000, optimum performance can be achieved by using specific adjacent ports.
- Examine potential security issues on servers and SAN devices. Systems that might have security exposure will have to be isolated. This will affect the zoning requirements.

- Estimate the number of servers and storage devices that will be connected to the SAN network.
- Decide whether the operating systems of the servers will be homogeneous or heterogeneous (multiple types of operating systems coexisting in the SAN fabric).

> **Note:** This point has important implications for your plan. Compatibility issues may occur if heterogeneous platforms share a SAN fabric. This restriction may force you to split the fabric into multiple zones that are specifically linked to a particular server or group of servers sharing the same operating system and storage devices. Resources are not shared between zones, so your performance and availability requirements need to be applied to each zone separately, which might change the number of ports or adapters required in your SAN.

From the questions above you now have a starting point for building your plan. We are now ready to study the key points to be considered when choosing the individual components for your SAN.

1.2 Choosing SAN components

The next step in designing a SAN network is to review the technical requirements in Section 1.5, "Product-specific considerations" on page 10. As a rule of thumb, the following items should be checked:

- Confirm that the SAN components you are considering will work with your SAN environment. This information can be found for IBM products at:

  http://www.storage.ibm.com/ibmsan/products/sanfabric.htm

- Determine whether you need to use the Fibre Channel Arbitrated Loop (FC-AL). If you do, you may need additional SAN components, such as FC-AL supported hubs.
- Determine whether to use switches or directors. By switches we mean the eight or sixteen port switches that can be linked together to build a fabric, and by directors we mean the high availability core products with large numbers of ports on one box managed by proprietary software. The larger a SAN fabric is, the more attractive directors become, considering the cost per usable port and ease of management and maintenance. On the other hand, switches are simple and flexible components. They can be interconnected for high availability and managed by a professional SAN management software, like Tivoli Storage Network Manager (TSNM).

1.3 Considerations for zoning

Zoning is where you hide parts of the SAN from some Host Bus Adapters (HBAs) and control which ports switches or targets can be seen by these HBAs. You might do this for security reasons for the purpose of isolating confidential systems or data. The more common reason is to isolate different HBAs or systems for which the interpretability has not been tested.

1.3.1 SNIA guideline

Some useful guidelines for zoning can be found in *"Open SAN Supported Solutions,"* which is a white paper published by the Storage Networking Industry Association (SNIA), and can be found at:

http://www.snia.org/

According to the white paper, the following restrictions should be considered to build an open interoperable SAN:

- Homogeneous fabric: Use switches from the same manufacturer with the same level of microcode or firmware.
- Heterogeneous HBAs may coexist on the same SAN, but should be placed in separate data zones. HBAs from different manufacturers must be zoned.
- HBAs within a host should be homogeneous. HBAs from different manufacturers, such as Emulex and JNI, should not be mixed on the same host.
- Be aware that HBAs use vendor-specific drivers and firmwares. Confirm that the firmwares and microcodes shipped on the HBAs are supported by the SAN devices that you are going to connect.
- Heterogeneous operating systems on the same SAN should be in separate data zones. Different operating systems sharing the same SAN network may cause conflicts.
- Heterogeneous storage targets on the same SAN should be in separate data zones. An example of this would be having an IBM 2105 ESS on the same SAN as an EMC Symmetrix. Any supported host can have access to either or both, but the storage systems themselves should not be share the same zone.

Note: The goal of the SNIA for the near future is to deliver a supported open SAN solution incorporating a data zone with heterogeneous servers simultaneously accessing storage systems from each of the participating storage vendors. This future enhancement will demonstrate the following characteristics:

- Homogeneous fabric
- Heterogeneous system initiators exist on the same SAN in the same data zone
- HBAs within a host are homogeneous
- HBAs will use specified drivers and firmware
- Heterogeneous operating systems exist on the same SAN in the same data zone
- Heterogeneous storage targets exist on the same SAN in the same data zone

The status of this SNIA project can be checked at:
http://www.snia.org/index.html

1.3.2 A case study of heterogeneous SAN design

In this section, as an exercise, we create a plan for zoning. Given the following scenario, you are to figure out how many zones are required in the SAN network if we faithfully apply the guidelines given in the SNIA white paper.

Given scenario

A customer has been using SUN, HP, and NT systems, one of each directly connected to an EMC Symmetrix storage system by using point-to-point Fibre Channel connections. The customer plans to expand his SAN environment, but at this time he puts forward the following conditions: First, the existing point-to-point connections should be replaced by a full-scale Fibre Channel fabric; and second, the new SAN fabric should accommodate new servers and storage devices. The customer is looking to add two IBM pSeries servers running AIX, an IBM 2105 Disk Storage System, and an IBM 3584 Tape Library. He is now wondering what would be an appropriate zone configuration.

A sample solution

1. Analyze the requirements.

   We have analyzed the interoperability requirements. According to the given scenario, the heterogeneous systems listed in Table 1-1 should be interconnected to the customer's SAN fabric.

Table 1-1 Sample audit information

Machine type	O/S	HBA
Two 7026 H80s	AIX 4.3.3	IBM Gigabit FC Adapter
One PC server	NT 4.0	Qlogic QLA2200F Adapter
One SUN Enterprise 4500	Solaris 7	Qlogic QLA2200F Adapter
One HP 9000 V	HP UX 11.0	HP A5158A Adapter

2. Confirm that the newly added components are supported.

   The next step is to check whether each SAN storage device is properly supported in the given scenario. Conduct a detailed search of support issues by each operating system and each HBA type. We are adding one ESS system and one 3584 system to the existing SAN network. Remember that EMC Symmetrix should also be included in the consideration.

   At the time of writing this redbook, the necessary information can be found at the following URLs:

   `http://www.storage.ibm.com/hardsoft/products/ess/supserver`

   `http://www.storage.ibm.com/hardsoft/tape/3584/3584opn.html`

   `http://www.EMC.com/techlib/abstract.jsp?id=65`

3. Develop a plan for zoning.

   We have defined the components that should be connected to our SAN. The next step is to create a plan for the zones. We will apply the SNIA guidelines described in Section 1.3, “Considerations for zoning” on page 4. The case study team came up with the following plan.

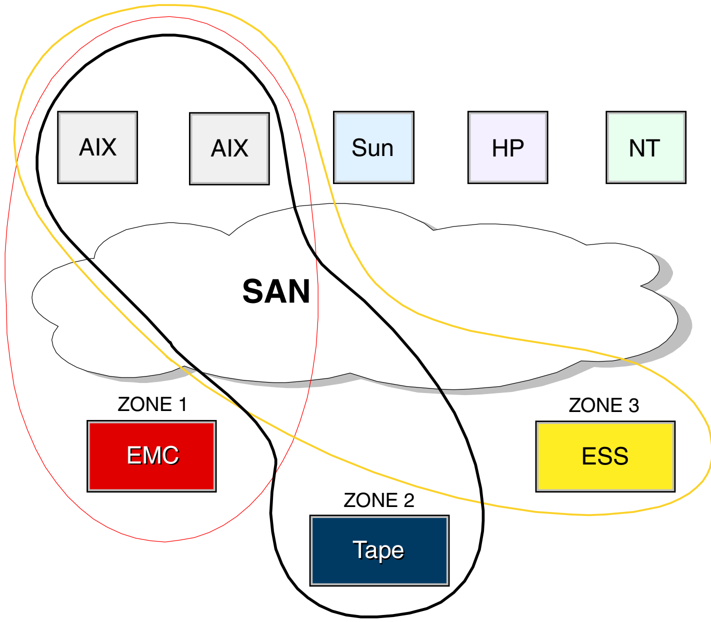


Figure 1-1 Multiple data zones for heterogeneous storage devices

Three separate zones have been defined for the two pSeries servers, as shown in Figure 1-1. The two pSeries servers are grouped together because they have the same HBAs and the same operating system. The three storage devices are assigned to separate data zones because they are heterogeneous. This allows the AIX systems to access all of the storage devices, but prevents the storage systems from interacting with each other. The same three data zones would have to be created for the other servers because either the operating system or the HBAs on these servers are heterogeneous. A total of twelve zones are required in this case. In order to keep the picture as clear and concise as possible, all of the zones are not shown in this figure.

1.4 Considerations for cables

In this section we briefly discuss the types of fiber cables that are normally used with IBM SAN products. Be aware that customers or other manufacturers often use different names and descriptions to indicate the same types of cable. If you need further information on IBM Fiber Transport Services please refer to:

http://www-1.ibm.com/services/its/us/fts.html

1.4.1 Types of cables

Types of optical fiber cables are often distinguished by the diameter of the core and cladding, measured in microns. An optical fiber cable having a core diameter of 50 microns and a cladding of 125 microns is designated as 50/125 microns. Other fibers commonly used are 62.5/125 and 9/125. These fiber cables come in two distinct types, which are linked to their sizes: Multi-Mode Fiber (MMF) for short distances (up to 2 km), and Single-Mode Fiber (SMF) for longer distances (up to 10 km, or for the IBM 2032 up to 100 km). Table 1-2 shows the cable specifications and the IBM supported distances for these cables and modes.

Table 1-2 Cables and modes

Diameter (microns)	**Mode**	**Laser type**	**Distance**
9	Single-mode	Long wave	<= 10 km
50	Multi-mode	Short wave	<= 500 m
62.5	Multi-mode	Short wave	<= 175 m

IBM pSeries Gigabit Fibre Channel PCI Adapters (#6227 and #6228) support 50 and 62.5 micron multi-mode connections. For 9 micron single-mode connections to these adapters you must use a suitable SAN device (switch or hub) as a link between the two modes and cable types.

Other types of cables you may need to be aware of are:

Jumper cables An optical fiber cable, physically, is a pair of optical fibers that provide two unidirectional serial bit transmission lines. Commonly used for short fiber optical links between servers and storage devices within the same room.

Multi-jumper cables Contain more than one pair of fiber cables. These are commonly used for fiber optical links within the same building. Multi-jumpers are used in large DP-Centers to interconnect more than one pair of Fibre Channel ports.

Trunk cables Typically contain 12 to 144 fibers. These have a strength member and an outer jacket. Each Fibre Channel link requires two trunk fibers. The physical trunk cable configuration varies and depends on the user requirements, environmental conditions, and the type of installation (for example, above ground or underground).

1.4.2 Long wave cable example

Figure 1-2 shows an example of two SAN switches that are separated by a distance longer than 300 m but shorter than 9 km. This would be the case when a customer has built two sites, each of which has its own SAN network connecting to pSeries servers. The customer wants to link the two sites with a cable that supports a distance of up to 9 km.

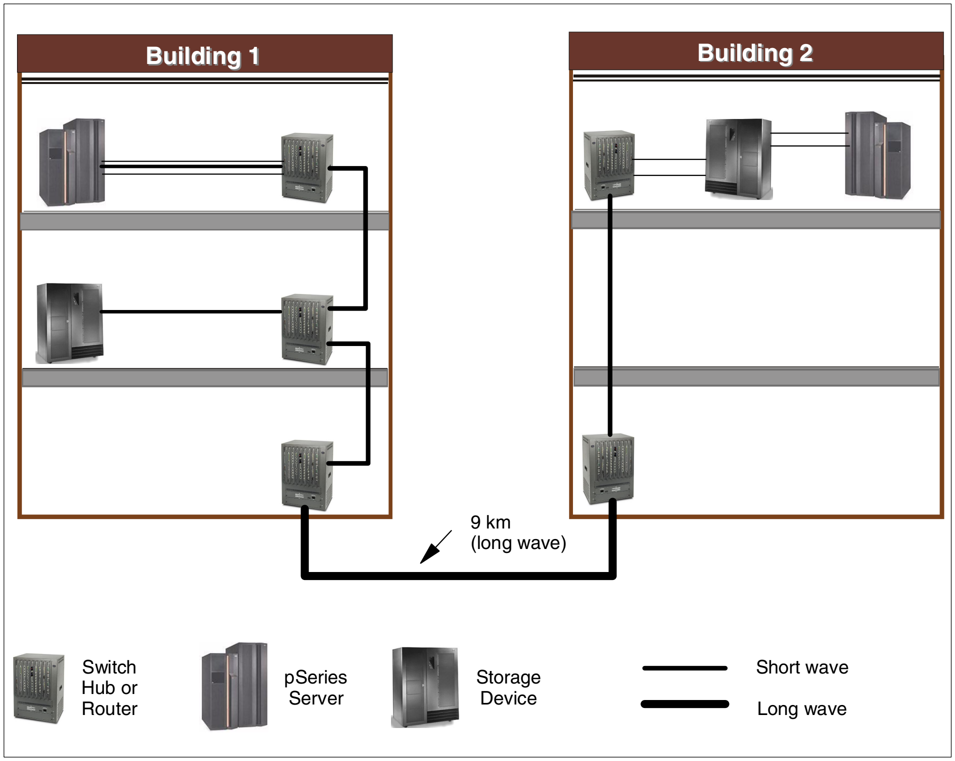


Figure 1-2 Long wave and short wave

In Section 1.4.1, “Types of cables” on page 8, we learned that a long wave cable can be used in this case. The point we need to pay attention to is that normally the ports on the SAN switches cannot accommodate long wave cables. To address this issue, the customer needs to prepare either long wave GBIC upgrades for the existing switches or new switches that have long wave ports.

We have not given any details of the switches used in our example, as there are different choices for switch products. Table 1-3 lists the feature codes for the long wave GBICs available for IBM SAN switches.

Table 1-3 Long wave feature codes

Model	Feature code for long wave
2109 S16 and S08	Feature code 2020
2109 F16	Feature code 2220
2031 016 and 032 (McDATA)	Feature code 3010
2032 064 (McData)	Feature code 6011 Card with four LW ports on each card Feature code 6012 Card with three SW ports and one LW
2042 001 (INRANGE)	Feature code 2020 GBICs added in groups of eight Feature code 2030 Extended length LW allows Director-to-Director connection up to 80 Km

1.5 Product-specific considerations

The IBM SAN initiative is an IBM corporate commitment to provide the complete range of services, infrastructure, and technology required for businesses to successfully implement a SAN. The SAN value-add solutions layer covers the solutions that provide the true customer value of SAN technology. Software, servers, fabric, and storage are the building blocks used to construct, use, and manage the SAN. The services and system integration components are critical to the success of the SAN initiative because of the complexity of heterogeneous, multi-platform environments.

In the following sections we cover these topics:

- Section 1.5.1, SAN servers - pSeries
- Section 1.5.2, SAN switches
- Section 1.5.3, SAN storage

1.5.1 SAN servers - pSeries

IBM currently has a wide range of SAN-ready RS/6000 and pSeries servers: our entry 170, 270, F50, and F80 servers; our rack-mounted H70, H80, M80, and new pSeries 640 servers; our family of high-end Enterprise SMP servers, including the new 24-way SMP, the pSeries 680; and our Large Scale SP system, including the newest POWER3 SMP High Node.

In order to connect a pSeries server to a Fibre Channel fabric, the pSeries server should have at least one Fibre Channel adapter. The Gigabit Fibre Channel PCI Adapter, IBM's connectivity solution for pSeries, is an RS/6000 I/O adapter that provides Fibre Channel connectivity for AIX system hosts. Each adapter, generally known as a Host Bus Adapter (HBA), provides one port for connection into the configuration. 7017-S70 supports only FC #6227 to provide Fibre Channel attachment. Except for the S70 server, a wide range of pSeries servers provide Fibre Channel attachment capabilities through the following solutions:

- Gigabit Fibre Channel Adapter for 64-bit PCI Bus (FC #6228)

  This Fibre Channel adapter is a 64-bit address/data short form factor PCI adapter with LC-type external fiber connectors that provides single or dual initiator capability over an optical fiber link or loop running up to 100 MBps. With the use of appropriate optical fiber cabling, this adapter provides the capability for a network of high speed local and remotely located storage. It supports up to 500-meter distances. When used with the IBM Fibre Channel storage hub, distances of up to 10 km are supported. If we are attaching a device or switch with a SC-type fiber connector, an LC-SC Fibre Channel conversion cable is required. The minimum OS level required is AIX Version 4.3.3 with AIX Update CD-ROM LCD4-0995-12 or higher, or APAR IY20367 from FIXDIST which can be found at the following URL:

  http://techsupport.services.ibm.com/rs6k/fixes.html

- Gigabit Fibre Channel Adapter for PCI Bus (FC #6227)

  This Fibre Channel adapter provides single or dual initiator capability over an optical fiber link or loop running up to 100 MBps. With the use of optical fiber cabling, this adapter provides the capability for a network of high-speed local and remotely located storage. It supports up to 500-meter distances. The minimum required OS level is AIX 4.3.3 or higher, or AIX 5.1 or higher.

> **Attention:** As of 12/3/2001, the Gigabit Fibre Channel Adapter for PCI Bus (FC #6227) is no longer available, and has been replaced by FC #6228.

Refer to the following IBM storage subsystem Web page for additional supported server attachment information for IBM devices:

http://www.storage.ibm.com/hardsoft/disk/products.htm

1.5.2 SAN switches

Considerations for high availability

- Selecting proper SAN switches

  The SAN infrastructure itself is typically one of the highest availability components of storage networks. To provide high availability the switches should have features such as hot-pluggable and redundant power supplies, and cooling and hot-pluggable gigabit interface converter (GBIC) modules.

  – Availability rate: Look into the availability rate of switch products you are considering. The availability rate of a switch is represented by the Mean Time Between Failures (MTBF) number of a tested system, including redundant components.

  – Switch upgrade: The ability to upgrade switches seamlessly without stopping services may be another important factor to consider. What you need is a network of switches that can provide alternate paths within a SAN fabric. With such an ability administrators can upgrade switches in the network without interrupting operations.

  – Firmware download: After an administrator has tested a new firmware code, the tested firmware code should be easily downloaded to other portions of the SAN fabric. The ability to upgrade only the selected parts of the network can be a key advantage over single monolithic switch designs.

- Design considerations

  Basic fabric design should also have built-in redundancy to allow for alternate paths to end node devices. A single switch fabric will have no alternate paths between devices should the switch fail. You can avoid the single-point-of-failure by simply configuring two switch fabrics along with redundant elements on host systems and storage nodes. Adding redundancy in your fabric allows routing through an alternate switch.

  A number of factors should be considered when designing a SAN. There is no single answer regarding which one is the best, but a properly designed SAN fabric might have taken the following points into account:

  – Hosts with key applications should have redundant paths to the storage nodes via the fabric. For this purpose, adding multiple HBAs per host should be considered so that a single HBA failure will not result in an entire loss of access to data.

  – Also consider high availability of the storage nodes. This topic will be discussed in a greater detail in Section 1.5.3, “SAN storage” on page 21. A critical storage node can be mirrored locally within a fabric or mirrored across an extended fabric link. Mirroring may slow performance a little, but this effect can be compensated for by using proper techniques such as buffer credits, which allow the sending device to continue sending data up to the credit limit without having to wait for acknowledgment.

– The extending of a SAN fabric to a remote location can be a viable option to enhance availability. You can use a link extender that allows from 30 to 120 km of distance between switch elements. However, there is a latency penalty for extended links that needs to be considered where performance is a primary concern. The shorter the links are, the lower the latency is. Use a rough estimate of 100 microseconds of delay per 10 km of distance for round-trip traffic.

Considerations for performance

- Inter-switch links (ISL)

  Inter-switch link means connection between switches. Multiple links can operate concurrently between any two switches in the fabric, allowing multiple redundant paths to be defined. In the event of a link failure, the traffic one link was carrying will be automatically and immediately transferred to another link. However, increasing the number of ISLs may require a longer time of propagating updated routing and zoning information across all ISLs. Changing ISL configurations necessitates the recalculation of routes within the fabric. This task places a burden on the switches in the fabric, so frequent changes should be avoided.

- Zoning

  The total number of zones and zoned devices should also be weighed when designing a large switch fabric. Zoning information also needs to be passed among switches when the fabric is reconfigured. Each switch may have to consume CPU cycles to accept, resolve, and reset zone data.

  – Software zoning: Software zoning is implemented within the Simple Name Server (SNS) running inside the fabric switch. When using software zoning, the members of the zone can be defined with node world wide name (WWN) and port world wide name. This option provides more flexibility in moving devices in the fabric; the zoning information follows the WWN no matter what port the device is plugged in to the fabric. However, CPU cycles are used to translate the WWN to a domain/port ID, an extra task imposing an additional processing load on the switches in the fabric.

  – Hardware zoning: Hardware zoning is based on the physical fabric port number. The members of a zone are physical ports on the fabric switch. This technique does not require CPU translation, but it decreases flexibility. For example, if you move a cable to a different port, the zoning information must be reconfigured to reflect the change on the switch hardware.

- Switch hop count

  Fewer hops are better. Hops are allowed between any source and destination using switches. However, this is likely to change over time. The hop count limit is set by the fabric OS and is used to derive a frame hold-time value per switch. The hold-time value is the maximum amount of time a frame can be held in a switch before it is dropped or returned. A frame would be held if its destination port is not available.

IBM SAN FC switches

IBM offers several different types of IBM SAN Fibre Channel switches, as follows:

- IBM SAN Fibre Channel switch Model 2109-S08, which has 8 ports
- IBM SAN Fibre Channel switch Model 2109-S16, which has 16 ports
- IBM SAN Fibre Channel switch Model 2109-F16, which has 16 ports of 2-gigabit bandwidth.
- IBM SAN switch S16 Integrated Fabric, which is an RPQ product and has 64 ports

These switches consist of a system board with up to 8/16/64 ports, and a SAN fabric operating system required for creating and managing a SAN fabric.

All Fibre Channel cable connections are made to the front panel of the switches using short wavelength (SWL) fiber optic and long wavelength (LWL) fiber optic with dual standard connector (SC) plugs. The IBM SAN Fibre Channel switches provide a 10/100BaseT Ethernet Port that can be used for connecting switch management consoles. This feature allows access to the switch’s internal SNMP agent, and also allows remote Telnet and Web access, enabling remote monitoring and testing. A serial port is provided on IBM SAN Fibre Channel switch Model 2109-S08 for recovery from loss of password information, and debugging purposes by recovering factory settings and for the initial configuration of the IP address for the switch. It is not used during normal operation.

Considerations for IBM FC switches

- Switch hop count

  The maximum number of switches allowed in the fabric is 239. The other limitation is that only seven hops are allowed between any source and destination using IBM 2109 switches. However, this is likely to change over time. Whatever the number of hops allowed, fewer hops are better. There is a maximum of three inter-switch hops for any path, and a maximum of two hops recommended for normal operation with the third hop reserved for back-up paths.

- Remote Switch Activation (FC #7302)

  Select this feature to obtain the unique activation key necessary to enable the remote switch capability. The switch should have a firmware level of 2.2.1 or higher. The remote switch function is used on two FC switches that are interconnected with a pair of CNTs Open System Gateways, providing the Fibre Channel to asynchronous transfer mode (ATM) conversion, across a wide area network.

- Extended Fabric Activation (FC #7303)

  Select this feature to obtain the unique activation key necessary to enable the remote switch capability. The switch should have a firmware level of 2.2.1 or higher. Extended fabric function provides extensions within the internal switch buffers for maintaining performance with distances greater than 10 km, and up to 70 km, by maximizing buffering between the selected switch interconnect links.

- F16 Performance Bundle Activation (FC #7421)

  Select this to obtain the unique activation key necessary to enable the Performance Bundle function within the SAN switch Model F16 Firmware. The Performance Bundle provides both inter-switch link (ISL) trunking and performance monitoring capabilities. The trunking enables Fibre Channel packets to be distributed across up to four inter-switch connections (links) between two Model F16 SAN switches providing up to eight gigabits per second of bandwidth and preserving in-order delivery. Both SAN switches must have this feature activated.

McDATA overview

- ED-6064 Director

  The McData ED-6064 Director, IBM 2032-064, provides high performance port connections to end devices such as servers, mass storage devices, and other peripherals in a Fibre Channel switch network. Up to 64 Fibre Channel connections are provided and its management software configures and supports any-to-any port connectivity.

- ED-5000 Director

  The McDATA ED-5000 Fibre Channel Director offers up to a 32-port switching capability. Each port delivers 100 MBps, full-duplex data transfer. 3,200 MBps transmission bandwidth supports full non-blocking 32-port switch performance. It offers excellent redundancy and maintenance capabilities.

- ES-3016, ES-3012 switch

  The McDATA ES-3016 switch, IBM 2031-016, provides 16 Fibre Channel generic ports for attachment to device ports or director expansion ports through fiber optic links. This switch provides full duplex, bidirectional data transfer at 1.0625 Gbps for all ports.The ES-3012 switch, IBM 2031-032, provides 32 ports with the same characteristics.

- ES-1000 switch

  The ES-1000 switch acts as a loop-switching hub and fabric-attach switch. This switch provides connectivity between attached Fibre Channel Arbitrated Loop (FC-AL) devices such as IBM Magstar Tape and IBM LTO and a fabric. The switch also incorporates a bridging function that provides dynamic connectivity between FC-AL devices and McDATA Directors participating in a switched fabric.

Considerations for McDATA switches

- EFC Management Software (FC #3750)

  EFC Management Software is required for each EFC Server PC (FC #3755), and can support up to 36 individual McDATA Directors and switches that have EFC Product Managers. The EFC Product Manager applications provide the user interface for each McDATA Director or switch.

  – For ED-6064 and ED-5000: The EFC Product Manager is required for each McDATA Director. Its feature code is #6300 for ED-6064 and #3775 for ED-5000.

  – For ES-3032 and ES-3016: Each of these switches provides an embedded Web-based management function suitable for small entry configurations. Yet the EFC Product Manager can also be used to manage these products. The feature code of the EFC Product Manager on ES-3032 is #3773, and the feature code for ES-3016 is #3772.

  – For Loop Switch, ES-1000: The EFC Product Manager is included as part of the base configuration.

- Open System Management Software (FC #6000)

  This feature is used to support third party SAN management software from vendors such as Tivoli, Veritas, or BMC. This feature is enabled through the EFC Manager and extends the management capabilities to include in-band (via Fibre Channel) management by the open system host-based SAN management application.

- Ethernet Hub (FC #5030)

  This feature provides a 24-port Ethernet Hub for attachment of additional McDATA Fabric Switches and Directors.

- Fiber optic cables

  Cables are typically customer-supplied, and details depend on the specific configuration being implemented. For simple configurations, there are two cable features available with the ED-6064:

  – FC #1811, which provides a 10 meter 50 micron/125 micron multi-mode cable with the new smaller form factor duplex LC connectors at both ends.

  – FC #1810, which provides a 10 meter 50 micron/125 micron multi-mode cable with the new smaller form factor duplex LC connector on one end and a duplex SC connector at the other end. A cable coupler is available to connect the LC/SC cable (FC#1811) to existing SC cables.

- Tape systems connected to ES-1000 that require the FC-AL attachment feature

  – IBM Magstar Tape 3590 FC-AL Drive

  – IBM LTO 3584 UltraScalable Tape Library FC-AL Drive (FC #1456)

  Note: IBM LTO 3583 Scalable Tape Library with integrated SDG Module (FC #8005) does not require using a McDATA ES-1000.

INRANGE switches

The INRANGE FC/9000 Fibre Channel Director provides enterprise-class availability and Fibre Channel fabric connectivity for IBM servers and storage supporting the FICON protocol. Each FC/9000 Director can be configured with up to 64 ports (2042 Model-001) or 128 ports (2042 Model-128). The Director also provides fully-redundant Fibre Channel fabric switches when configured with a high-availability option. Each of the Fibre Channel ports supports either long wave or short wave GBICs. Each port is self-configuring, automatically recognizing the type of connections to a fabric, a loop, or an inter-switch connection. The INRANGE-Virtual Storage Networking (IN-VSN) Enterprise Manager software is designed for centralized management control of multiple FC/9000 Directors and provides functions such as zoning, monitoring, diagnostics, name server, SNMP alerts, and Call-Home.

Considerations for INRANGE switches

- FIO Module (FC #5010)

  INRANGE FC/9000 Fibre Channel Directors can be configured with either short wave or long wave GBICs for distances up to 10 km, or up to 80 km with extended long wave GBICs.

  Any mixture of short wave and long wave GBICs are supported, but GBICs must be ordered in increments of eight, meaning either eight short wave GBICs or eight long wave GBICs, or a mix of short wave and long wave GBICs.

  – For Model-001: The base Director contains three FIO Modules with 24 ports on this model. We can add from one to five additional 8-port FIO Modules (FC #5010) for a maximum of 64 ports.

  – For Model-128: The base Director contains six FIO Modules with 48 ports on the Model 128. We can add from one to ten additional 8-port FIO Modules (FC #5010) for a maximum of 128 ports.

- Host Adapter (HA) Option (FC #5920)

  The HA Option provides a second Fabric Switch Module, a second Fibre Controller Module, and a second Power Module for redundant power supplies.

- Server/Client Module

  – Additional IN-VSN Client Software (FC #7201): For additional Windows NT or Windows 2000 workstation interfaces to the IN-VSN Management Server

  – IN-VSN Enterprise Manager Server Bundle (FC #7600): A desktop PC running Windows 2000, including a 17-inch monitor, internal and external modems for call-home and remote diagnostics, and an 8-port Ethernet Hub with cables for private LAN connections to up to seven FC/9000 FC Directors Control Modules

  – Eight-port Ethernet Hub (FC #7603): To support the attachment of multiple FC/9000 Directors to a private LAN

- Interconnect Kit for 128-port Capability (FC #4128)

  Provides the hardware and firmware required for interconnecting two 64-port FC/9000 Directors installed in the same cabinet to operate as a 128-port Director. What are included in this feature are interconnect cables, a FC/9000 Firmware upgraded to 128-port capability, and an IN-VSN Management Server upgraded to 128-port capability. Two 64-port FC/9000 Directors interconnected with this feature are functionally equivalent to the Model-128, the 128-port FC/9000 Director.

Support matrix

Refer to Table 1-4 to find the support matrix of IBM SAN Directors and switches and pSeries servers with their associated host bus adapters, storage devices, and firmware levels.

Table 1-4 Support matrix for IBM SAN

SAN element	SAN component
Software	Operating systems: IBM pSeries servers & RS/6000: AIX Version 5.1 or AIX Version 4.3.3 update (4330-08 recommended maintenance package) RS/6000 SP servers: AIX Version 4.3.3 - PSSP Version 3.1.1 or PSSP 3.2 HACMP: Version 4.4.0 with APARS IY11563, IY11565, IY11480, IY12022, and IY12057 HACMP: Version 4.3.1 with APARS IY11564, IY11560, IY12021, IY 12056, IY 03438 (CRM/ES CRM), IY11110 (F10/F20)
Server	IBM pSeries servers: 620 7025-6F0,620 7025-6F1, 640 7026-B80, 660 7026-6H0, 660 7026-6H1, 680 7017-S85 IBM RS/6000 servers: 7013-S70,S7A, 7015-S70,S7A, 7017-S70, S7A, S80, 7025-F50,F80, H70, 7026-H50,H70, 7026-H80,M80, 7043-270, 7044-170,270 RS/6000 SP servers: All PCI nodes

SAN element	SAN component
SAN interconnects	HBA FC# 6227: Gigabit Fibre Channel Adapter for PCI Bus with HBA Firmware Version 3.22A0 or higher HBA FC# 6228: Gigabit Fibre Channel Adapter for 64-bit PCI Bus with HBA Firmware Version 3.82A0 or higher

Storage	Disk systems:
	Enterprise Storage Server (2105-Ex0, 2105-Fx0) with Fibre Channel Host Adapter FC# 3022
	
	Refer to the following URL for additional details and prerequisites:
	`http://www.storage.ibm.com/hardsoft/products/ess/supserver.htm`
	
	Tape systems:
	Refer to the following URL for additional details and prerequisites:
	`http://www.storage.ibm.com/hardsoft/tape/index.html`
	
	Network-attached storage:
	Refer to the following URL for additional details and prerequisites:
	`http://www.storage.ibm.com/snetwork/nas/index.html`

Topology	Direct attach: One host server per FC port on the switch FC Hub Attach: IBM Fibre Channel Storage Hub (2103-H07) is supported for distance solutions beyond 500 m and up to 10 km. Hub is used for conversion between short wave and long wave optical only. Multiple initiators or targets on the same loop are not supported Cascaded Switches (2109) are supported in configurations up to a 2 x 4 array. Maximum of 3 inter-switch hops for any path. Maximum of 2 hops recommended for normal operation with the 3rd hop reserved for back-up paths

Since the matrix may change from time to time, you must check for the latest information when you create an implementation plan.

More information can be found at the following Web sites:

For IBM SAN FC switches:

http://www.storage.ibm.com/ibmsan/products/2109/supserver.htm

For IBM McDATA Director and switches:

http://www.storage.ibm.com/ibmsan/products/2032/prod_data/supserver-064.html

For INRANGE switches:

http://www.storage.ibm.com/ibmsan/products/directors/prod_data/supserver-042.html

1.5.3 SAN storage

Here we discuss aspects of SAN storage.

IBM Enterprise Storage Server

Disk storage is the centerpiece for any storage infrastructure. The IBM Enterprise Storage Server (ESS) is a second-generation Seascape Disk Storage System that provides industry-leading availability, performance, manageability, and scalability. Virtually all types of servers can concurrently be attached to Enterprise Storage Servers.

Product overview

When planning a SAN fabric with an ESS attached, it is useful to understand the architecture of the ESS. Figure 1-3 on page 22 illustrates the architecture of the ESS, showing the physical level data flow.

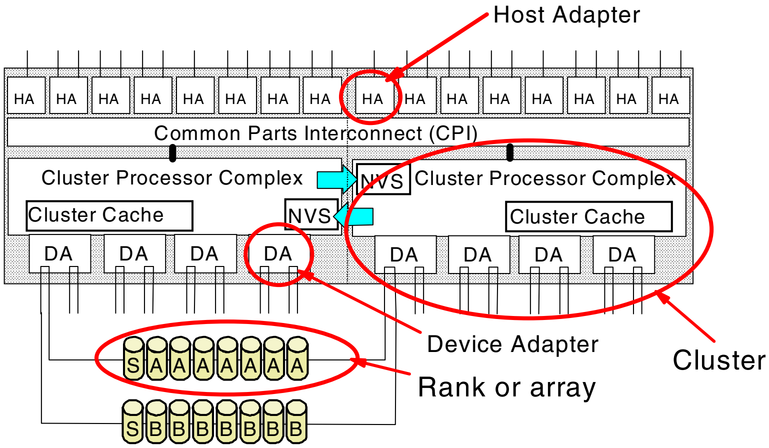


Figure 1-3 ESS architecture

The host adapter (HA) card attaches the ESS to a host system. HA cards can be installed to support any combination of ESCON, SCSI, or Fibre Channel connections. The ESCON and SCSI cards provide two ports. The Fibre Channel HA card provides a single port interface. The HA card can access either of two clusters installed in independent bays using the common parts interconnect (CPI).

The clusters use a four-way symmetric multiprocessor (SMP) to coordinate data transfer and perform advanced functions between the host adapters and device adapters. Each cluster includes cache and nonvolatile storage (NVS). The cache stores frequently accessed data and makes it quickly available. Similarly, NVS together with cache storage improves performance of write operations and provides data integrity in the event of an unplanned loss of power to the cluster. The Serial Storage Architecture (SSA) device adapters (DA) connect the disk drive modules (DDMs) to the clusters in a serial loop of up to six arrays. An ESS contains eight device adapters, which are installed in pairs, one in each bay.

Each DA pair supports one or two serial loops. Arrays are composed of a group of eight DDMs that store data within the ESS. DDMs are installed in 8-pack pairs, both of which must be attached to the same loop. Within a DDM 8-pack all of the disk drives are the same capacity and can contain 9, 18, or 36 GB (gigabyte) DDMs.

Considerations for performance

In general, we can maximize an application's performance by spreading the I/O load across the clusters, disk arrays, and device adapters in an ESS. When attempting to balance the load within an ESS, placement of application data is one of the crucial determining factors. The following resources are the most important in balancing an I/O load:

- Cluster

  Each cluster has a number of peripheral component interconnect (PCI) buses that are used for moving data into and out of the cache memory. The ESS Ex0 models have three buses per cluster and the Fx0 models have four buses per cluster. Two each of the Ex0 and Fx0 models are used for transferring data between the host and the ESS cache. If you choose to use RAID arrays for a critical application, spread the redundant array of independent disks (RAID) arrays across separate clusters. Since each cluster has separate memory buses and a cache memory, this technique maximizes the use of those resources.

- Device adapter

  Adapters in host bays 1 and 3 share the same internal bus in each cluster, and host bays 2 and 4 share the other internal bus in the cluster. It is important to avoid a situation where all the activity for a cluster comes from host bays 1 and 3 or from bays 2 and 4. If that were to happen, all of the host activity would be directed to only one of the two PCI buses used for host transfers. For most applications, this is not likely to be a concern. Random workloads do not typically place very heavy usage on internal buses. However, highly optimized applications that balance the workload very uniformly across internal components and that need high bandwidth must consider this. An example would be decision support or business intelligence applications with DB2 UDB. When selecting RAID arrays within a cluster for a critical application, spread them across separate SSA device adapters.

- Subsystem Device Driver (SDD)

  Use the IBM Subsystem Device Driver (SDD) or similar software for other platforms to balance I/O activity across SCSI or Fibre Channel ports.

- Adapter bays

  There are four host bays (1, 2, 3, & 4) for host attachment. When selecting ports to assign to a given server, spread them across different adapter bays. Spreading across adapter bays pairs 1 and 4 or 2 and 3 will also balance CPI activity. Avoid the bay 1/3 or bay 2/4 combination. For fiber-attached hosts that go through a switch, avoid zoning strategies that force activity to these combinations.

- RAID

  Use as many RAID arrays as possible for the critical applications. Most performance bottlenecks occur because a few disks are saturated. Spreading an application across multiple RAID arrays ensures that as many disk drives as possible are available (for open systems environments where cache-hit ratios are usually low). The I/O intensity guidelines shown in Table 1-5 ensure that disk drives and RAID arrays are not overloaded and provide good application response time. This is based on a conservative set of workload assumptions. Also, the guideline is intended to represent a planning estimate to avoid bottlenecks, not the maximum achievable throughput. The guideline applies to a random access workload using record sizes of 16 KB or less. When configuring an ESS, define enough RAID arrays so that the range of I/O operations per second per array does not exceed these values. Use Table 1-5 to decide the number of arrays to use when configuring logical unit numbers (LUNs) for a given application.

Table 1-5 Maximum I/O operations per array

Read ratio (%)	I/O per second per array
100	400 - 500
70	275 - 325
50	200 - 300

- Disk drive capacity

  Selecting a disk capacity for your configuration can have a significant effect on the performance of the ESS. The following are some general guidelines:

Table 1-6 Selecting disk drive capacity DDMs

DDM capacity	Considerations
18 GB DDMs	Most likely the correct choice for most applications. Provides the proper balance of performance and cost.

DDM capacity	Considerations
9 GB DDMs	Should only be considered in the following circumstances: Applications with very demanding performance requirements. Workloads with very random access patterns or very high write content ("cache-hostile" workloads, usually where there is high disk activity). Workloads that have very high access rates relative to the amount of total storage.
36 GB DDMs	Should only be considered in the following circumstances: Applications that *do not* have demanding performance requirements. Applications that make very good use of cache, with very high cache-hit ratios (>95 percent). Applications with very low access rates relative to the amount of storage (<1 access per second per GB of storage).

- Number of Fibre Channel ports

  The maximum bandwidth capability of the Fibre Channel architecture is 100 MB/sec full duplex for a total of 200 MBps. In addition, the throughputs will vary based on the type of application, specific adapter, and host attachment.

  As a general guideline, we should make planning decisions assuming a throughput capability of around 60 MB/sec per Fibre Channel port configured to an ESS. This guideline can be used if the performance requirements of the application or applications are understood.

  For example, with an application that requires a 100 MB/sec of aggregate throughput, you would most likely need two Fibre Channel ports to support this requirement (2 x 60 MBps > 100 MBps). Sometimes there is little or no information about the exact performance requirements of the application. In these cases you must estimate the number of ports required. The following guidelines can help with those estimates:

  – For an ESS configured entirely for pSeries and other open systems attachment, four Fibre Channel ports are usually adequate for handling most workloads and applications. Most online transaction processing, file serving, and standard applications fall into this category. This assumes that the workload is well-balanced across the ports. Naturally, if some of the ports have very little activity while others are very busy, then you need to plan for more ports.

– For some high-bandwidth applications in open-system environments, more ports could be necessary. For example, business intelligence applications that rely on massive parallel database scans that are likely to need all available bandwidth could need as many as eight Fibre Channel ports. In this case you should install two adapters in each of the four I/O bays.

> **Tip:** An ESS is attached to an RS/6000 and pSeries through a Fibre Channel switch. The ESS port must be defined as point-to-point. It is detected automatically by Fibre Channel adapters so we do not need to configure the Fibre Channel adapter of the pSeries system.
>
> To verify the status of the fscsi devices we can use the following command:
>
> `# lsattr -EHl fscsi-number`
>
> If the output of this commands show the `attach mode - al`, we should change to point-to-point and use a switch. The Fibre Channel adapter definition must be removed from the Object Data Manager (ODM) and then rerun **cfgmgr** with the adapter attached to the switch. The commands are following:
>
> ```
> # rmdev -dl fsci0 -R
> # rmdev -dl fcs0
> # cfgmgr
> ```

Considerations for high availability

For high availability, we recommend that you configure *at least* two separate Fibre Channel interfaces all the way from the Fibre Channel host to the ESS. Therefore, you must prepare at least two Fibre Channel adapters on each pSeries server.

In point-to-point configuration, one of the Fibre Channel adapters on the pSeries system should be plugged into Cluster 1 and the other plugged into Cluster 2 of the ESS. We also must configure each LUN that is to be accessed via multipathing to each path's host Fibre Channel adapter.

In order for the host to take advantage of both paths, we should use an IBM SDD to manage the paths. Traditionally, the task of balancing an application I/O load across physical paths to a disk drive subsystem was the responsibility of a person, such as a system administrator. This can lead to problems, such as wrong assumptions regarding the I/O load that various applications impose.

The two major reasons to install SDDs on your machines are:

- Load balancing: The SDD automatically adjusts data routing for optimum performance.
- Path failover and online recovery: The the SDD automatically and non-disruptively redirects data to an alternate data path.

Figure 1-4 shows the concept of multipathing provided by the SDD.

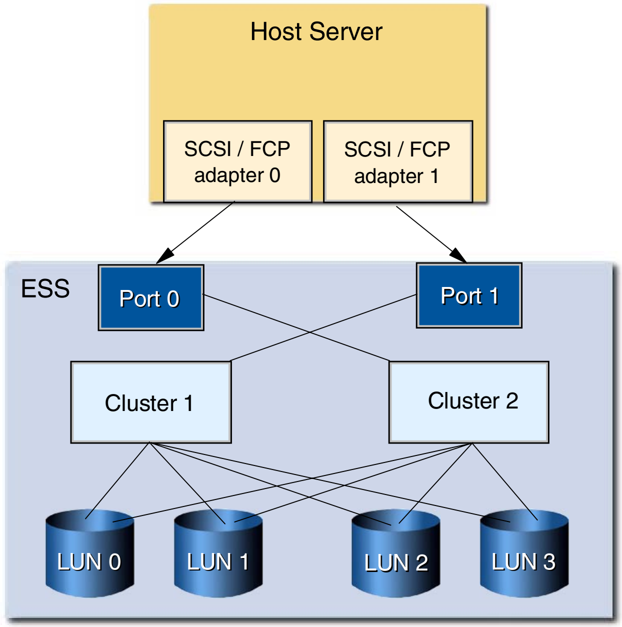


Figure 1-4 Multipath connections between a host server and ESS LUNs

The SDD runs in the host and will, in the event of failure in one path, automatically switch to the other path. This software supports up to 32 separate paths. We can therefore configure eight physical paths from our pSeries system to a fabric, then four physical paths from the fabric to the ESS. This configuration provides a total of 32 different paths by which the host can communicate with the ESS. The IBM Subsystem Device Driver can use all 32 of these paths. In addition to providing failover to an alternate path, the IBM Subsystem Device Driver also performs dynamic load balancing across all configured paths. It will spread the I/O across the paths in such a way that utilization is maximized. This improves our total performance because it will either eliminate or drastically reduce the amount of time that any I/O must wait for a path to become available before it can execute.

Considerations for data sharing

Data sharing means that multiple hosts can access the same LUNs. These configurations require a certain software on the hosts' side that is needed to manage concurrent access to the LUNs. This requirement acts as a restriction, implying that the participating hosts must have the same operating system and the same type of file system, or the same database management system (DBMS). In ESS you must configure each LUN that is to be accessed via multiple hosts to each host Fibre Channel adapter.

EMC Symmetrix

EMC has three families: the Symmetrix 3000 Series, the Symmetrix 5000 Series, and the recently announced 8000 Series. The ESN Manager is EMC software for managing Fibre Channel storage networks. With the ESN Manger Volume Logix component we can define paths from each host to its storage device. With the ESN Manager, path creation can select and access Symmetrix volumes. When creating paths, we can choose the number of paths on each host. There are two path selection choices, one to the HBA of the server paths and one to the Fibre adapter of Symmetrix paths.

Considerations for EMC Symmetrix

The following factors should be taken into consideration:

- At the time of the writing of this redbook, the supported OS level is AIX 4.3.x, AIX 5.1. Note that AIX 5.1 does not support PowerPath.
- The supported AIX Fibre Channel Adapters are FC #6227 and FC #6228. For more detailed information on supported HBAs, refer to:

  `http://www-1.ibm.com/servers/eserver/pseries/library/hardware_docs/sa38/380516.pdf`

  Three types of host bus adapters are currently supported:

  – IBM FC #6227

  – IBM FC #6228

  – EMC CKIT-E70-AIX

  For IBM 6227 and 6228 adapters, you can mix these two types of HBAs on the same server. You can also mix FC-AL connections with switched Fibre Channel connections on same server.

  For an EMC adapter, mixing FC-AL and switched Fibre Channel connections are supported on the same server. Note that device drivers for EMC HBAs are available exclusively as EMC Fibre Channel Interface Version 2.0 for AIX Platforms. Download the latest PTF packages from the EMC FTP server site:

  `ftp://ftp.emc.com/pub/symm3000/aix/EMC-FC-Kit`

- In a HACMP environment, we should check supported PTF packages and EMC ODM filesets at the following site:

  www-1.ibm.com/servers/eserver/pseries/library/hardware_docs/sa38/380516.pdf

- The booting of a pSeries server from an EMC Fibre Channel device is not supported.
- In the case of the Fibre Channel connectivity using a hub, fan-in up to four is supported. *Fan-in* is the number of hosts connecting to a single storage port, such as four hosts connecting to a single storage device port.
- In the case of the Fibre Channel connectivity using a switch, fan-in up to 10 is supported. Hops per zone has a maximum of three, though some switch connections support only two.

  To verify the interoperability, refer to the following EMC Web site:

  http://www.emc.com/horizontal/interoperability/matrices/EMCSupportMatrix.pdf

IBM Magstar Tape 3590

The Magstar 3590 B11 and 3590 E11 are frame-mounted or rack-mounted and incorporate a 10-cartridge Automated Cartridge Facility (ACF) Autoloader which supports high capacity (up to 1.2 TB for Model E11s) and unattended operation. The Magstar 3590 B1A and 3590 E1A are designed to be used in a Magstar 3494 Tape Library, Magstar Virtual Tape Server, and Magstar 3590 Silo Compatible Tape Subsystem. These models do not have an ACF.

- 3590 Model Bxx Tape Drives

  The 3590 Model B drives have an uncompressed data rate of 9 MBps. It can read or write 128-track format cartridges. 3590 B drives can be upgraded to 3590 Model E Drives.

- 3590 Model Exx Tape Drives

  The 3590 E Models may be ordered with either UltraSCSI or Fibre Channel interfaces. Each drive has two interfaces for availability reasons. A Magstar 3590 Model E Tape Drive provides an uncompressed data rate of up to 14 MBps, and can write a total of 256 tracks across the tape, doubling the cartridge capacity of B Models. 3590Es may read 128-track or 256-track cartridges, and write 256-track cartridges.

Considerations for 3590 Fibre Channel Attachments

For a 3590 native Fibre Channel Attachment, the following features need to be ordered:

- FC #9510 (Fibre Channel Attachment - Plant) can be added to all new orders of 3590 Model E11 or E1A Tape Drives. If this feature is not included then the tape drive will have the standard Ultra SCSI Attachment.
- FC #3510 (Fibre Channel Attachment - Field) can be added to any installed 3590 Model E11 or E1A Tape Drive to change the tape drive attachment from the standard Ultra SCSI Attachment to a Fibre Channel attachment.

The 3590 E Model with the Fibre Channel Attachment feature is supported for attachment to the Gigabit Fibre Channel Adapter for PCI Bus (#6227) or Gigabit Fibre Channel Adapter for 64-bit PCI Bus (#6228) on pSeries systems.

> **Note:** The Gigabit Fibre Channel Adapter for 64-bit PCI Bus (#6228) has an LC connector, while 3590 Tape Drive Fibre Channel Cables have SC connectors. Therefore, to attach 3590 Tape Drives to the 64-bit Gigabit Fibre Channel Adapter, LC-SC Fibre Channel Converter Cables (pSeries FC #2456) should also be used. Updating the 3590 Tape Drive microcode to the latest level is also required. The update microcode is a Drive Microcode Update feature (FC #0500) that customers can order.

The 3590 Model B11 or E11 requires a 7202, 7014-T00, 7014-T42, 7015-R00 rack, or space in an existing customer rack. The 7014-T00 or 7014-T42 racks should be ordered without the front door. (Refer to the IBM 7014 Model T00 or Model T42 sales manuals for details.) They can also be installed in the 7014 Model S00 and the S/70 Advanced Rack if the rack door is removed.

Considerations for high availability

AIX High Availability Failover support is available for SAN connections to Fibre Channel adapters on a pSeries system running AIX 4.3.3. A current level of the open system device driver is required for this support.

This solution is designed to allow users to configure SAN with redundant paths so that component failures are overcome in a way completely transparent to the application. This has been made possible on pSeries servers using a multipath-aware 3590 device driver. With the multipath information known to the device driver, when a command fails, the device driver knows to look down another path. Furthermore, the device driver maintains position information so that it can restore positioning and reissue the command to send information down the other path. The advantage of this approach is that the failover is done without any application involvement so any application currently running gets the failover capability.

AIX High Availability SAN Failover is a built-in feature of the IBM AIX Enhanced Tape and Medium Changer Device Driver, that is Atape. The Atape device driver alternate pathing support is designed to provide the following two basic functions:

- The Atape driver provides the ability to configure multiple physical paths, one primary and up to fifteen alternate paths to a 3590 Fibre Channel Tape Drive. The physical path can be any combination of HBAs, switches, or drive ports.
- The Atape driver also provides automatic error recovery and failover to an alternate physical path when a permanent error occurs, transparent to the running application.

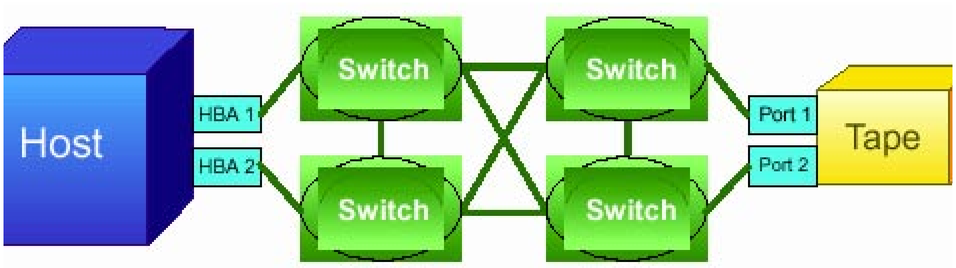


Figure 1-5 Atape configuration example

In Figure 1-5 the pSeries has two host bus adapters and the tape system has two Fibre Channel ports, one per drive. AIX configures all tape drives that are recognized through each available path. The number of tape drives recognized on the AIX side is determined by multiplying the number of host bus adapters by the number of Fibre Channel ports. In the above example, four tape devices will be created.

IBM Magstar 3494 Tape Library

The Magstar 3494 Tape Library, a tape automation system, consists of individual frame units for modular expansion that provide a wide range of configurations. The basic building block of the Magstar 3494 is the Lxx Control Unit Frame, which contains a library manager, a cartridge accessor, up to two tape drives, and slots for the storage of tape cartridges. To the Lxx frame, you can add drive frames, storage unit frames, the Magstar Virtual Tape Server, and a high-availability model to create a maximum configuration of 16 frames and two service bays. The Magstar 3494 supports both Magstar 3590 Model B1A/E1A and 3490E Model F1A Tape Drives within the same library (but in separate frames).

The main components of the Magstar 3494 Tape Library are:

- Library control unit frame
- Drive unit frame
- Storage unit frame
- High availability model
- Magstar Virtual Tape Server

We will go over each component one-by-one.

- Library control unit frame

  The library control unit frame is a central and important component of the Magstar 3494 Enterprise Tape Library Subsystem. Each control unit frame contains a tape subsystem (Magstar 3590 or IBM 3490E Tape Drives), a library manager, a cartridge accessor with single or optional high-performance dual gripper, an optional convenience I/O station, and cartridge storage cells.

  There are three control unit models:

  – Model L10: Can contain a 3490E Model C1A or C2A, or up to two 3490E Model F1A Tape Drives

  – Model L12: Can contain up to two Magstar 3590 Tape Drives

  – Model L14: Can contain up to two Magstar 3590 Tape Drives, with one 3590 Model A50 or A00 Controller

- Drive unit frame

  The drive unit contains both additional tape drives and tape cartridge slots for the Magstar 3494 Tape Library, and must be attached to a Tape Library Base Frame Model L10, L12 or L14.

  The three drive models D10, D12, and D14, differ in terms of cartridge storage cell capacity, tape subsystem supported, and host platform attachment capability.

  – Model D10

    This model supports the IBM 3490E Magnetic Subsystem Models C1A, C2A, and F1A. It can contain zero or one 3490E Tape Drive Unit (Models C1A or C2A), or up to two 3490E Model F1A Tape Drives and 400 or 300 tape cartridge storage cells.

– Model D12

  This model supports the IBM Magstar 3590 Tape Subsystem Model E1A or B1A. The base configuration of no tape drive unit and 400 tape cartridge storage cells can be expanded with up to a total of six tape drive units and 250 tape storage cells. The Model D12 Frame with three, four, five, or six Magstar 3590 Tape Drives, can be integrated with a Virtual Tape Server (VTS) to provide an ESCON-attached hierarchical storage management system.

– Model D14

  This model can contain one 3590 Model A60, A50, or A00 Controller and up to four Magstar 3590 Model E1A or Model B1A Tape Drives. The base configuration of no tape drive unit and 400 tape cartridge storage cells can be expanded with up to a total of four tape drive units and 305 tape storage cells.

- Storage unit frame (Model S10)

  The storage unit frame is used to attach additional tape cartridge storage cells to the tape library and the cartridges can be a mix of 3490E or Magstar 3590 tape cartridges. Up to fifteen storage frames can attach to any library control unit frame (Lxx) and each S10 provides storage for up to 400 tape cartridges.

- High availability model

  The high availability unit (HA1) ensures access to the data even if the library manager or cartridge accessor fails. The HA1 consists of two service bay units, one placed at each end of the library, plus a second library manager and accessor. The HA1 will operate in hot-standby mode if there is no dual active accessor option. In the event of a cartridge accessor or library manager failure, the standby accessor and/or library manager will take control of all operations in the library. The model HA1 eliminates the possibility of single-point-of-failure in the library manager and the tape cartridge.

  The Dual Active Accessor (FC #5050) is an optional feature. It activates the library's second cartridge accessor, enabling both accessors to operate simultaneously. It can improve tape mount performance and cartridge inventory time.

- Magstar Virtual Tape Server

  The VTS combines the speed of disks, the advantages of virtual tape drives, the capacity and reliability of IBM 3590 tape technology, and the automation of the 3494 to create an self-contained hierarchical storage management subsystem. VTS provides improved performance and capacity for tape processing. It includes all the basic tape backup capabilities (for example, peer-to-peer copy) and provides a variety of host attachment options, disk storage capacities, tape drive attachments, configurations, and management options.

An Enterprise Tape Library can be configured with up to two VTSs. There are two VTS models:

– Model B10: The Model B10 VTS provides host connection of up to four Extended Performance ESCON Channels or up to eight SCSI Bus Attachments.

– Model B20: The Model B20 VTS provides host connection of up to 16 Extended Performance ESCON Channels or up to eight SCSI Bus Attachments.

Considerations for Fibre Channel attachment

The 3590 Models B1A and E1A are the building blocks of the IBM TotalStorage Enterprise Automated Tape Library 3494, IBM TotalStorage Virtual Tape Server, and StorageTek Silos. The 3590 Tape Drives can be attached to a SCSI interface or Fibre Channel interface and to ESCON and FICON channels.

To get the Fibre Channel attachment capability, add one of the following features to each tape drive:

- Fibre Channel Attachment - Plant (FC #9510)

  This feature can be added to all new orders of 3590 Model E1A Tape Drives. If this feature is not included, the tape drive will have the standard Ultra SCSI Attachment.

- Fibre Channel Attachment - Field (FC #3510)

  This feature can be added to any installed 3590 Model E1A Tape Drive to change the tape drive attachment from the standard Ultra SCSI Attachment to the Fibre Channel Attachment.

The Fibre Channel Drive Option (FC #3511) must be added to the 3590 Model C12 Frame, or the 3494 Tape Library Model L12 or D12 Frame, to provide the hardware support for each 3590 Model E1A Tape Drive with Fibre Channel Attachment features installed in those frames.

A 3590 Model A60 Controller (3494-A60) is required for the attachment of these drives to FICON channels.

- FICON Attachment - Long Wave (FC #3432)

  This feature provides one long wave FICON adapter, with an SC Duplex Connector, for the attachment of 3590 Tape Drives through the Model A60 to a FICON Host System Long Wave Channel utilizing a nine micron single-mode fiber cable. The total cable length cannot exceed 10 km (up to 20 km if ordered as an RPQ).

- FICON Attachment - Short Wave (FC #3433)

  This feature provides one short wavelength FICON adapter, with an SC Duplex connector, for the attachment of 3590 Tape Drives through the Model A60 to a FICON host system short wavelength channel utilizing a 50 micron multimode fiber cable. The total cable length cannot exceed 500 m.

- Fiber Drive Attachment (FC #3463)

  This feature permits the attachment of 3590 E Model Tape Drives with a Fibre Channel Attachment (FC #3510 or #9510) to a 3590 Model A60 Controller. It replaces the standard Ultra SCSI Tape Drive Attachment with two Fibre Channel adapters, and includes the required microcode and installation instructions for the Model A60.

Note: The prerequisites for a Fibre Drive Attachment (FC #3463) are:

- All attaching 3590 Tape Drives must have either FC #3510 or FC #9510.
- An IBM 2109 Model S16 SAN Fibre Channel switch must be installed in the rack or frame containing the Model A60 Controller for attaching Fibre Channel tape drives. Each 2109 switch comes standard with four short wavelength GBICs, enough to connect to the two Model A60 Fibre Channel Attachments and two 3590 Fibre Channel Tape Drives.

Additional orderable features are:

- Short Wave GBIC (FC #2010), which provides one additional GBIC; therefore, the quantity of this feature should at least equal the number of tape drives attached to the Model A60 minus two.
- Additional Power Supply (FC #6103), which provides an additional redundant power supply that enables dual-power source configurations to minimize power outages. Non-Rack Install (FC #9205) should not be specified for this feature.

- Fiber Drive Attached Rack (FC #3465)

  This feature supplies the required hardware to support the attachment of 3590 Tape Drives with a Fibre Channel Attachment (FCs #3510 or #9510) to a 3590 Model A60 Controller in a rack. It includes the mounting hardware and instructions for installing the IBM 2109 Model S16 SAN Fibre Channel switch in the rack, including the associated Ethernet Hub and cabling between the Model A60 and the hub and switch. Fibre Channel cables from the 3590 Tape Drives to the switch in the rack with the Model A60 are included by specifying FC #9059 (one for each tape drive). For multiframe attachment of 3590 Tape Drives in other racks to the switch, the cables should be ordered with the drives.

Note: The prerequisites for a Fibre Drive Attached Rack (FC #3465) are:

- Fibre Drive Attachment (FC #3463) is mutually exclusive with FC #4065 (Multiframe SCSI Rack Attach).
- Cables: Fibre Channel cables from the 3590 Tape Drives to the 2109 switch in the rack with the Model A60 are specified with FC #9059 on the Model A60. Fibre Channel cables from the 3590 Tape Drives in other racks should be ordered with FCs #5813, #5825, or #5861 on the tape drives.

Linear Tape Open (LTO)

The Linear Tape Open (LTO) program is a joint initiative by IBM, HP, and Seagate. It is aimed at creating a new tape standard for open systems. The formats were developed to serve multiple requirements and now the standards are supported by multiple suppliers. LTO technology has developed into two open tape format specifications: Accelis and Ultrium. Not all users require the same features and functions. Some applications are *read-intensive* and require the fastest access to data possible. Some applications are *write-intensive* and fill tape cartridges very quickly, requiring the highest single-tape-cartridge capacity possible. IBM, HP, and Seagate overcame the potential trade-offs implied by a single-format solution by specifying both a fast-access open tape format specification (the Accelis format) and a high-capacity open tape format specification (the Ultrium format).

Now the IBM Ultrium family of tapes and libraries comprises four different product offerings: Stand-alone Drive Unit LTO3580, Autoloader LTO 3581, Scalable Library LTO3583, and Highly-scalable Automated Library LTO3584.

All four LTO Ultrium Tape Systems are available with both SCSI LVD and HVD interfaces. The 3584 LTO Tape Library is also available with a direct FC-AL Attachment.

The following are ways of attaching a Ultrium Tape Library into a SAN environment:

- Direct attachment

  Using a native Fibre Channel interface is an option available on the 3584. This enables you to attach each single FC-AL to the LTO Tape Drive, via the Fibre Channel Patch Panel inside the 3584. The Fibre Channel Patch Panel (FC #1462) is required with the first LTO Ultrium FC-AL Drive ordered in each frame. This feature provides a patch panel for the Fibre Channel host cables. Host server connections are made to the Fibre Channel Patch Panel.

The LTO Ultrium FC-AL Drive Canister (FC #1456) provides a canister containing one IBM LTO Ultrium Tape Drive with a Fibre Channel Arbitrated Loop (FC-AL) interface. Also included is a power supply, fans, and a drive-to-patch panel cable.

- LVD/HVD Interface

  For drives with LVD interfaces, attachment can be through the SAN Data Gateway Router, 2108-R03. For drives with HVD interfaces, attachment will be through either the SAN Data Gateway Router, 2108-R03, or the SAN Data Gateway, 2108-G07. In the case of 3580-L11, 3581-L17, an FC # 2840 is needed on the gateway. For 3580-H11, 3581-H17, an FC #2830 needed on the gateway. LTO 3583 can also connect to SAN with these interfaces.

- Using SAN Data Gateway Model in LTO3583

  There is also another way to connect LTO3583 to the SAN. We can use the internal SAN Data Gateway Module (FC# 8005) in the LTO3583 Library. It provides attachment support for Fibre Channel interfaces using a Short Wave 2 GB/sec Gigabit Interface Convertor (GBIC) with SC connectors.

Considerations for performance

The IBM LTO Ultrium tape drive has a native streaming data rate of 15MBs/sec, 30MBs/sec with 2:1 compression. *Streaming rate* is the rate at which a tape drive can read/write, not including any start/stop operation. Most uses of tape do include some start/stop, which slows down the sustained rate at which the drive operates. For example, when writing to a tape drive, normally the drive returns control to the application when the data is in the tape drives buffer, but before the data has actually been written to tape. This mode of operation provides all tape drives a significant performance improvement. If the application wants to be absolutely sure the write makes it to tape, the application needs to flush the buffer. Flushing the buffer causes the tape drive to backhitch (start/stop). The Tivoli Storage Manager (TSM) parameters TXNByte and TXNFile control how frequently TSM issues this buffer flush command.

Considerations for capacity

The 3583 Scalable Tape Library can be configured with one to six IBM Ultrium tape drives. Multiple drives offer enhanced functions, such as faster transfer of data, simultaneous backup, concurrent read/write operations, and fault tolerance. The I/O station can be defined as I/O slots or data storage, and the definable affects the number of cartridges available for data storage.

Please refer to Table 1-7 on page 38 for the various combinations and resulting data capacities for LTO libraries.

Table 1-7 LTO 3583 slots

Model	Slots w/o FC 8012 (12-cartridge I/O station)	Slots w/ FC 8012 defined as storage	Slots w/ FC 8012 defined as I/O slots
L18	Storage: 18 I/O: 1	Storage: 36 I/O: 0	Storage: 24 I/O: 12
L18 with one FC 8007 (18-slot opt.)	Storage: 6 I/O: 1	Storage: 54 I/O: 0	Storage: 42 I/O: 2
L18 with two FC 8007 (18-slot opt.)	Storage: 4 I/O: 1	Storage: 72 I/O: 0	Storage: 60 I/O: 12
L36	Storage: 36 I/O: 1	Storage: 54 I/O: 0	Storage: 42 I/O: 12
L36 with one FC 8007 (18-slot opt.)	Storage: 54 I/O: 1	Storage: 72 I/O: 0	Storage: 60 I/O: 12
L72	-	Storage: 72 I/O: 0	Storage: 60 I/O: 12

The quantity of slots for LTO 3584 depends on how many drives there are in the frame.

Table 1-8 Slots of LTO 3584 Model L32

Drives in Model L32 Frame	Slots in Model L32 Frame (w/o capacity expansion feature)	Slots in Model L32 Frame (w/ capacity expansion feature)
0-4	141	281
5-8	113	253
9-12	87	227

The IBM 3584 Model D32 Expansion Frames may be added to the base frame (Model L32) in order to add storage and/or drive capacity. Up to five D32 Expansion Frames may be added to each base frame, L32 Model. Also, the quantity of slots for D32 depends on how many drives are in the frame.

Table 1-9 Slots of LTO3584 Model D32

Drives in Model D32 Frame	Slots in Model D32 Frame
0	440
1-4	423
5-8	409
9-12	396

SCSI Disk/Tape

Industry-standard Fibre Channel technology is rapidly replacing SCSI Channel Attachment between open system servers and disk and tape storage systems. However, many disk and tape storage systems do not provide Fibre Channel attachment. To bridge the gap between Fibre Channel server adapters and SCSI-attached disk systems and tape storages, the following Data Gateways provide a solution for attaching SCSI devices to a SAN fabric:

- For SCSI disk connections:

  SAN Data Gateway for SCSI Disk, Model G07

  SAN Data Gateway for Serial Disk, Model S20

- For SCSI tape connections:

  SAN Data Gateway for SCSI Tape, Model G07

  SAN Data Gateway Router for SCSI Tape, Model R03

For more details for the support matrix, visit the following Web site:

`http://www.storage.ibm.com/hardsoft/products/sangateway/supserver.htm`

1.6 FICON interoperability

FICON is a new high performance I/O interface based on Fibre Channel technology providing Fibre Channel connectivity to the S/390 platform. Native FICON provides major improvements beyond ESCON in bandwidth, efficiency, distance, addressability, flexibility, scale, and full duplexing capabilities.

1.6.1 Native FICON support of ESS

Let us now discuss native FICON support of ESS.

Configuring ESS for FICON attachment

FICON is supported on the ESS F Model, but not on the E Model. When attaching an ESS F Model to a FICON interface, an ESS host adapter FC #3021 or FC #3023 is used.

FC #3021 is the long wave laser adapter that includes a 31 m 9 micron cable with SC duplex connectors. This adapter is also supported with either the 50 or the 62.5 micron multimode cable (ESCON cable) when mode conditioning cables are used at each end.

FC #3023 is the short wave laser adapter that includes a 31 m 50 micron cable with SC duplex connectors, and is also supported with the 62.5 micron multimode cable.

The FICON Host Attachment to Subsystem (FC #9909) must be selected. This feature is for administrative purposes only.

Since each of the FICON adapters takes up one of the four slots in one of the four I/O bays, we can have a maximum of 16 of these adapters in the ESS. This allows for a maximum of 16 FICON interface attachments. If these attachments are all point-to-point, then we can attach directly to 16 FICON channels; if some of these attachments are to switches, then the FICON zones must be in a single director because FICON does not support cascading directors. ESS allows 256 logical paths per FICON link (compared to just 64 for ESCON), 128 logical paths per Logical SubSystem (LSS), and 2048 logical paths per ESS.

- Configure Operating Mode option

  There is a Configure Operating Mode option that allows selection between S/390 and Open Systems modes. S/390 allows Port List views similar to existing ESCON Directors. Open systems devices can still share the same director, but it is limited to a single director fabric.

  Open Systems mode allows for a multi-director fabric and FICON and Fibre Channel connections, providing all FICON zones are kept within a single director.

System attachment considerations

There are a few important considerations that need to be understood in order to properly configure an S/390 system with FICON interfaces.

These are:

- S/390 FICON features

  When ESS is attached directly to FICON channels, then the maximum number of FICON attachments is 16, since that is the maximum number of host adapters that can be configured in ESS. The FICON adapters used in the IBM G5/G6 systems have FC #2314 and FC #2316 and are packaged with one port per adapter. FC #2314 is the long wave laser adapter and FC #2316 is the short wave laser adapter. The FICON adapters used in the new z900 system have FC #2315 and #2318 and are packaged two ports per adapter. FC #2315 is called FICON LX and is the long wave laser adapter; FC #2318 is called FICON SX and is the short wave laser adapter.

- The number of equivalent ESCON links

  A single FICON channel supports up to 60 MB/sec sequential bandwidth, which allows customers to collapse existing ESCON channels into FICON channels approximately 4:1 (more in some cases if ESCON channel utilizations are relatively low). And of course, for availability we will always want at least two FICON links. So a well-configured ESS typically needs no more than eight FICON channel interfaces in order to fully exploit its bandwidth. Greater simplicity of configuration is also achieved because FICON architecture allows more devices per channel (up to 16,000) to match the increased channel bandwidth.

- Multipathing

  With ESCON we would want to have at least four paths in the path group just to reach adequate performance. This is because most of the ESCON control units implemented channel command execution that at least partially synchronized the lower Direct Access Storage Device (DASD) interface with the upper channel interface. With FICON control units there is no longer any synchronization between the lower DASD interface and the upper channel interface. Therefore, the number of paths in the path group only depends on throughput requirement. Of course the number of FICON adapters must be a minimum of two and cannot exceed the architectural maximum of eight.

- The number and type of control units

  Attaching multiple control units to a channel is commonly referred to as "daisy-chaining." This term came from the S/360 OEMI parallel interface, where the interface cables would run from the channel to the first control unit, and then "daisy-chain" from the first control unit to the second, and from the second to the third, and so on. With serial interfaces like ESCON and FICON we no longer "daisy-chain" the control units.

However, when multiple control units are connected to a channel via a switch, is created the logical equivalent of the parallel "daisy-chain." With the OEMI parallel interface and with the ESCON interface, whenever the channel and control unit communicate with each other, they form a private connection. None of the other "daisy-chained" control units could communicate with the channel while this private connection was in place. The private connection could be supporting significant I/O between the channel and the control unit, or it could be running slowly, depending upon many factors affecting the control unit and the device. FICON changes the rules in this area. It does not support the private connection concept used in the previous interfaces. Instead, it performs frame (or packet) multiplexing. What this means is that in the configuration with the "daisy-chained" control units, the channel can be communicating with all of the control units concurrently. It can multiplex its I/O operations across all of the control units at the same time. Therefore, no interface cycles need be wasted due to a private connection being in place. But tape generally performs much larger I/O operations at any instant in time. Therefore, even with FICON, when we have Tape I/O running, we can temporarily "lockout" some DASD I/O. Hence, it is still better not to put tape and DASD on the same FICON channel.

1.6.2 Tape support for FICON attachment

IBM 9672 Enterprise G5 or G6 and Freeway servers provide for attaching the FICON Support Tape System with a FICON Channel Card Long Wave Laser (FC #2314) or Short Wave Laser (FC #2316).

IBM Magstar 3590 Model A60 Tape Controller with FICON Adapter and 3494 Tape Library is supported. IBM Magstar 3590 A60 Tape Subsystem with FC 3432 (FICON Attachment - Long Wave) in a 3494 Tape Library, in a 3590 Model A14 Frame or C14 Silo Compatible Frame is supported.

Up to two FICON Attachments (FC #3432 or #3433) can be specified on each Model A60, with each feature providing one FICON adapter for attachment to a FICON host channel.

1.6.3 Tips for FICON/FCP interoperability

Note: For mixed FICON and Fibre Channel configurations on the same McDATA Director, contact McDATA technical support personnel to validate your specific mixed FICON/Fibre Channel configuration. Intermixing FICON with Fibre Channel on INRANGE switches is not yet available at the time of the writing of this redbook.

The following is a summary of tips for intermixing FICON links with native Fibre Channel links:

- Set up zones with the FCP devices and the FICON hosts

  When the ESS is attached to FICON channels through one or more switches, then the maximum number of FICON attachments is 128 per ESS adapter. We are qualifying McDATA directors for use with the ESS for FICON attachment. The McDATA Directors support an in-band switch management function similar to what has been used with the ESCON Directors. In mixed configurations the FCP hosts should communicate only with the FCP devices and the FICON hosts should communicate only with the FICON devices. Although it is not required, it may be prudent to set up zones in the directors to guarantee that none of the FCP hosts or devices can affect the FICON traffic in any way.

- Cascading consideration

  When attaching FICON products to switches or directors, it is important to note that switch cascading is not allowed. That is, we cannot configure a fabric of multiple interconnected directors and then have a FICON channel attached to one director communicate to a FICON Control Unit attached to another director. The current FICON architecture prohibits this capability. The reason for this restriction is because the base S/390 I/O architecture uses a single byte for addressing the I/O devices. This 1-byte I/O address is not compatible with Fibre Channel's 3-byte port address. The FICON solution to this problem is to disallow switch cascading.

2

Installing SAN switches

Typically, a SAN offers a high-speed, dedicated Fibre Channel network fabric for the movement of data between servers and storage devices. Whatever its size, a SAN is built with various components ranging from switches to storages devices. A SAN fabric is created by establishing connections between different devices.

In this chapter we describe the actions needed to implement a SAN using pSeries servers. We start with the installation of the host bus adapter (HBA) in the server, then discuss the creation of the SAN fabric using switches from different manufacturers. Since there are still incompatibility issues due to the implementation differences by each switch vendor, we will limit the discussion to the point of view of using switches from the same manufacturer only. The following topics will be covered in this chapter:

- Preparing pSeries
- Configuring IBM 2109-S08/S16
- Configuring McDATA ED-6064
- Configuring INRANGE FC/9000

2.1 Preparing pSeries

Here we will discuss preparing a pSeries server for SAN switch installation.

2.1.1 Installing the device driver software on a pSeries system

In order to make the pSeries server properly configure the Gigabit Fibre Channel adapter, you need to install required device driver software either before or after attaching the adapter to the system. The device driver software is shipped in an AIX CD set. This section explains how to install the device driver software for IBM Gigabit Fibre Channel Adapter on a pSeries machine.

> **Note:** If the Fibre Channel adapter was installed on the machine prior to the installation of the AIX operating system, you may skip the procedure described in this section. The device driver for the Fibre Channel adapter will be automatically installed with the AIX. Proceed to Section 2.1.2, "Verifying the installation" on page 47.

If the adapter is to be installed after the AIX operating system has been installed, proceed with the software installation and perform the following steps:

1. From the AIX command line, enter the following SMIT fastpath:

   ```
   # smitty devinst
   ```

2. Install the following filesets from the appropriate AIX CD. If you have installed the 32-bit Fibre Channel Adapter (FC 6227,Type 4-S), install the devices.pci.df1000f7 fileset. If you have installed 64-bit Fibre Channel Adapter (FC 6228, Type 4-W), install the pci.df1000f9 fileset.

   ```
   devices.pci.df1000f7 or devices.pci.df1000f9
   devices.pci.common.IBM.fc
   devices.fcp.disk
   bos.diag.com
   bos.diag.rte
   ```

3. To confirm that the software is installed, type the following commands one line at a time, and press Enter after each line:

   ```
   # lslpp -l devices.pci.df1000f7 (for FC 6227)
   # lslpp -l devices.pci.df1000f9 (for FC 6228)
   # lslpp -l devices.common.IBM.fc
   # lslpp -l devices.fc.disk
   # lslpp -l bos.diag.com
   # lslpp -l bos.diag.rte
   ```

If no device driver information is displayed, or some information is missing, refer to Section 4.4.1, “HBA verification” on page 132 to determine the cause of the problem.

4. Run the **cfgmgr** command or reboot the system. This command configures the Fibre Channel adapter and is run automatically when the system is rebooted.

```
# cfgmgr -v
```

> **Tip:** There is an easier way to install the required device driver. After installing a Fibre Channel adapter, execute
>
> ```
> # cfgmgr -i /dev/cd0
> ```
>
> This command will automatically install missing device drivers from the CD ROM device. This technique can used for any type of adapter.

2.1.2 Verifying the installation

To verify whether the installation was successful, follow the steps below:

1. If the Gigabit Fibre Channel PCI adapter has been properly installed and configured, the output from the **lsdev** command will be like the following. Make it sure that the status of the adapter is “Available.”

```
# lsdev -C | grep fcs
fcs0 Available 20-60
```

If no adapter information is displayed, or if the status of the adapter appears as “Defined,” follow the problem determination procedure that is described in Section 4.4.1, “HBA verification” on page 132 to identify the cause of the problem.

2. The next step is to gather information on the adapter by entering the **lscfg** command. This should result in the display of more detailed information on the adapter.

Example 2-1 Sample output of lscfg command

```
# lscfg -vl fcs0
fcs0              14-08             FC Adapter
Part Number.................09P1162
      EC Level....................D
      Serial Number...............KT10801909
      Manufacturer................0010
      FRU Number..................09P1173
      Network Address.............10000000C9273872
      ROS Level and ID............02903290
      Device Specific.(Z0)........4002206D
      Device Specific.(Z1)........10020193
```

```
Device Specific.(Z2)........3001506D
Device Specific.(Z3)........02000909
Device Specific.(Z4)........FF101450
Device Specific.(Z5)........02903290
Device Specific.(Z6)........06113290
Device Specific.(Z7)........07113290
Device Specific.(Z8)........20000000C9273872
Device Specific.(Z9)........SS3.22A0
Device Specific.(ZA)........S1F3.22A0
Device Specific.(ZB)........S2F3.22A0
Device Specific.(YL)........U0.1-P1-I2/Q1
```

Make a note of the network address and the adapter location code (in this case it is 14-08) as this information will be useful when we come to configuring the SAN fabric. The information may also be useful when you do problem determination.

3. When a Fibre Channel adapter is configured, child fscsi devices associated with the adapters are created and activated. These child devices process the SCSI I/O protocol, which is used to control Fibre Channel-attached storage devices. Note that the device names assigned to the fscsi devices do not necessarily follow the address of the associated Fibre Channel adapter. Example 2-2 illustrates such a case. As you can see in the example, fcs0 has a child device of fscsi4.

Example 2-2 Check child devices of fcs0

```
   # lscfg | grep 14-08
   + fcs0              14-08              FC Adapter
* fscsi4            14-08-01          FC SCSI I/O Controller
```

You can confirm the status of the fscsi devices by:

```
# lsdev -C | grep fscsi
fscsi4     Available 14-08-01      FC SCSI I/O Controller Protocol Device
```

4. Finally, check the status of the fscsi devices with the **lsattr** command. Take a look at the `How this adapter is CONNECTED` line. The value of this field shows whether the adapter has been connected to a switched fabric or to an arbitrated loop.

```
# lsattr -EHl fscsi4
attribute   value    description                  user_settable
scsi_id     0x261013 Adapter SCSI ID              False
attach      switch   How this adapter is CONNECTED False
sw_fc_class 3        FC Class for Fabric          True
```

Note: If a Fibre Channel adapter is initially configured by the **cfgmgr** command with the wrap plug installed, or the adapter is directly attached to an arbitrated loop device such as an EMC Symmetrix, the value of the attach field is automatically set to “al” indicating that the adapter has been attached to an arbitrated loop.

If the display is `none` then this would indicate that the port is empty or the cable installed has no light from a switch or device. If the Attach Mode has been set to `none`, check the cable connection and run **cfgmgr** again. The system will change the value of Attach Mode to `al` or `switched` depending on the given condition.

If the Attach Mode is already set to `al` or `switched` and you want to toggle the current setting, the adapter must be logically removed with the **rmdev -dl fcsX -R** command; then re-run **cfgmgr** command.

You can also use:

```
# chdev -l fscsi4 -a attach=switch
```

The following shows an example of an arbitrated loop connection:

```
# lsattr -EHI fscsi4
attribute   value description                   user_settable
scsi_id     0x1   Adapter SCSI ID               False
attach      al    How this adapter is CONNECTED False
sw_fc_class 3     FC Class for Fabric           True
```

2.2 Configuring IBM 2109-S08/S16

One of the key building blocks of a fabric is the Fibre Channel switch. Like traditional network switches, Fibre Channel switches provide an intelligent, interconnected system between servers and storage devices. Larger Fibre Channel fabrics are created by using multiple switches that are interconnected to each other. Multiple switch schemes not only connect more devices, but also improve the bandwidth and reliability of the fabric. In the following sections we describe the process of creating a simple fabric and its expansion by using the IBM SAN Fibre Channel switch, 2109-S08 and 2109-S16.

2.2.1 Initial installation and configuration of the IBM 2109

A major SAN building step is the basic installation and configuration of the 2109 switch as the SAN fabric. This section lists the steps needed to implement the 2109. For additional details about specific actions, refer to the *IBM SAN Fibre Channel Switch 2109 Model S16 Installation and Service Guide,* SG26-7352. The actions are:

1. Physically install the switch (either table top or rack mount).
2. Install the optional second power supply. Verify that a power outlet for the second power cable is available.
3. Install the desired type of GBICs in any port of the switch. The GBIC modules supported are the short wavelength (SWL) and long wavelength (LWL) fiber optics, and Copper (Cu) versions.
4. Connect the switch's 10/100BaseT Ethernet port to the customer's LAN for switch management console interface.
5. Power the switch up and wait for approximately 2 minutes while the switch completes POST tests.
6. Set the IP address for the 10/100BaseT Ethernet port. The IP address can be set from the front panel on the 2109-S16, while the 2109-S08 uses command line access via a serial port. Both models also provide access over the customer's LAN using a factory default address of 10.77.77.77. Verify and change, if needed, the subnetmask. If connecting over the customer's LAN, the switch will prompt for a user logon
7. Verify connectivity to the switch over the customer's LAN, and continue with configuring the switch.

With the installation and LAN connectivity completed, there are a number of configuration parameter changes to consider. For additional information on configuring the 2109 switch, refer to the *IBM SAN Fibre Channel Switch 2109 Model S16 User's Guide,* SC26-7351. Some of the parameters for consideration are:

- The switch's date and time setting with the **`date`** command
- The switch's name with the **`switchName`** command
- The passwords for the user log ins with the **`password`** command
- If used in a multi-switch environment, the switch's domain ID can be modified with the **`configure`** command
- Implement zoning in the SAN fabric (see Section 3.2.2)

The 2109 switch configuration process can be accomplished with a command line interface from a Telnet session to the switch, or by means of Graphical User Interface (GUI) with the IBM StorWatch Specialist.

The steps to establish a Telnet session with the new switch and login are:

1. From the command line prompt on a pSeries server, use the command:

   ```
   >telnet <IP address>
   ```

   If the IP address of the new switch was changed, use the new IP address. Otherwise, the switch uses 10.77.77.77 as the default IP address.

2. The switch prompts for a user name. The default administrator name of the IBM 2109 switch is *admin*.

3. The switch next prompts for the password. The default administrator password is *password*.

In the command line interface, use the **help** command for assistance for a list of commands and their purposes. The IBM StorWatch Specialist view can be accessed through one of the following Java-enabled Web browsers:

For UNIX:

- Netscape 4.51 or higher

For Windows 95/98 or Windows NT:

- Internet Explorer 4.0 or higher
- Netscape 4.51 or higher

In addition to the above, Java Plug-In 1.2.2, or higher, is also required. The current version of the Java Plug-In can be downloaded from the following URL:

http://java.sun.com/products/

To launch the IBM StorWatch Specialist:

Start the Web browser if it is not already active. Enter a switch name or IP address in the Location/Address field.

There are a number of views offered by the IBM StorWatch Specialist that display various aspects of your SAN. These screens, or views, are organized into the following:

Fabric view Shows the number of switches in the fabric with world wide name, domain ID, switch name, and network IP information. The Fabric view is the first view displayed, and serves as the launch point for the

Fabric Topology view, Name Server Table view, Zone Administration Interface view, and Switch view.

Figure 2-1 shows a sample screen of the Fabric view.

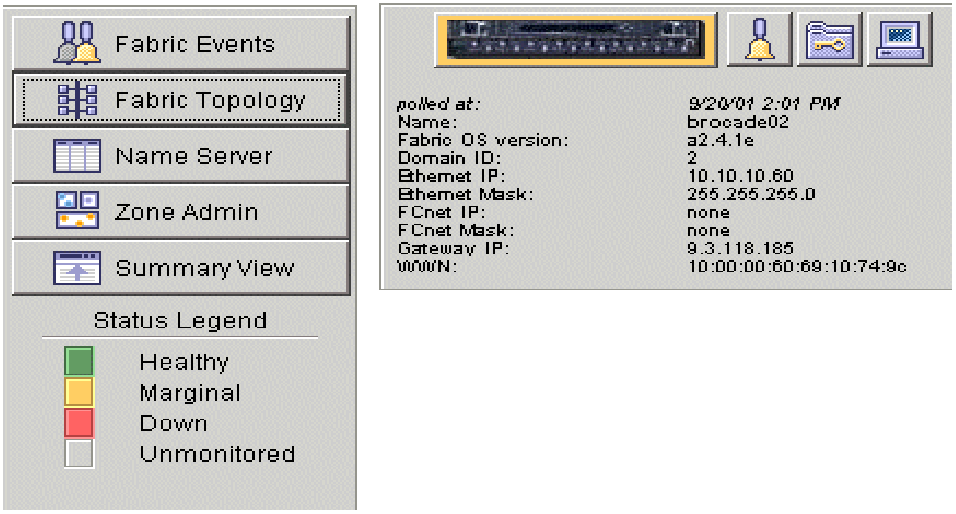


Figure 2-1 Fabric view

Fabric Topology view Displays the physical configuration, including active domains, paths, and routing information.

Name Server Table view Displays the Name Server table for the fabric. Use to view information about the devices attached to the fabric.

Zone Administration Interface view Provides an interface for configuring zoning: zone alias settings, zone settings, and zone configuration settings.

Switch view Displays the switch information. Provides a real-time view of overall switch status. The Switch view is the launch point for the following:

- Port Statistics view
- Performance view
- Administrative Interface
- Telnet interface

Figure 2-2 on page 69 shows a sample screen of the Switch view.

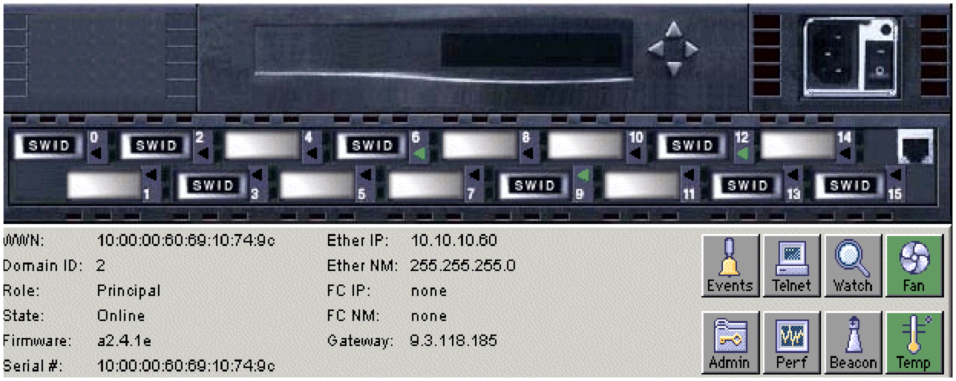


Figure 2-2 Switch view

Port Statistics view Displays the statistics, general information, and status for a specific port.

Performance view Graphically portrays the real-time data throughput for each port and the total switch bandwidth utilization.

Administrative Interface view Provides an interface for performing functions such as upgrading firmware versions or reconfiguring a switch.

Telnet Interface view Provides an interface for using Telnet commands for switch diagnostics, troubleshooting, and SAN management.

2.2.2 Implementation of zoning

Zoning is a fabric management utility that allows the segregation of devices within a SAN and enables resource partitioning for management and access control. Zoning can be used as a barrier between different environments. Thus, only devices in the same zone can communicate with each other. Some of the reasons for zoning are:

- Separation of different operating systems, such as Windows NT/2000-based systems segmented from UNIX-based systems.
- Separation by functional areas, such as keeping test areas partitioned from marketing.
- Isolating HBAs from different manufacturers.
- Securing Fabric areas by providing another level of access control.

- Separation by storage device type, such as tape systems and disks.

Zoning is typically based on either a port-by-port basis or by the Worldwide Name of the connecting devices. For the 2109 switch, either method for zone implementation can be used. Every zone has a member list that defines what devices, or the devices connected to specific ports, are allowed to see each other. At the same time, however, a device may belong to multiple zones. Zone members may be specified by one of the following notations:

- Node world wide name
- Port world wide name
- Port number, which includes the switch number (domain ID)

The 2109 switch has several components, called zone objects, that are used to define a zone. The zone objects and their purpose are:

Aliases Zone aliases simplify the repetitive use of port number entries or world wide names. For example, the name `Shark11` could be used as an alias for `10:00:00:c9:21:23:f5`; while `iod95` could be used as an alias for `1,8` (port 8 on switch 1).

Members Members are individual devices used to create zones. Members can be specified by a world wide name, port number, an alias, or a combination of any designation.

Zones A zone is set of two or more, members that are allowed to communicate with each other.

Configurations A zone configuration is a set of one or more zones. When a zone configuration is in effect, all zones that are members of the configuration are able to see each other.

There can be multiple zone aliases, members, and configurations that are stored in the 2109's memory. However, only one zone configuration can be active at any given time. Zone configurations are disabled and enabled with the **`cfgDisable`** and **`cfgEnable`** commands. An example zone configuration is shown below.

Example 2-3 IBM 2109 switch zone configuration

```
brocade02:admin> cfgShow
Defined configuration:
 cfg:   Prod_zone
                iod95_to_shark; nowhere_togo;
 cfg:   Just4show
                nowhere_togo
 zone:  iod95_to_shark
                mickey; tiod95
```

```
 zone:  nowhere_togo
                nowhere; places
 alias: nowhere 2,3
 alias: places  2,9
 alias: tiod95  2,12
 alias: Mickey  2,6

Effective configuration:
 cfg:   Prod_zone
 zone:  iod95_to_shark
                2,6
                2,12
 zone:  nowhere_togo
                2,3
                2,9
```

We defined the following configuration of two zones for use on the 2109 switch. For a more complete discussion on zone implementation in the 2109 switch, please refer to the *IBM SAN Fibre Channel Switch 2109 Model S16 User's Guide,* SC26-7351.

Table 2-1 2109 configuration

Zone name	Description
iod95_to_shark	A two port zone connecting pSeries server (tiod95 on port 12) to a 2105 ESS (Mickey on port 6)
nowhere_togo	A two port zone with no devices attached consisting of switch ports 3 and 9

2.2.3 Scaling the fabric

This section describes procedures for scaling, or enlarging, the fabric. Simply stated, the fabric is expanded by cascading additional switches. These extra switches can be brand new out of the box, or switches that are already part of another SAN. Many of the switch' parameters will be automatically configured when a new switch is added to the existing fabric. However, there are several items that need to be manually configured.

When adding switches to an existing fabric, the new switch IP address and domain ID must be unique to allow the cascading of switches. The following fabric configuration parameters must be identical for cascading to occur:

- BB_Credit
- R_A_TOV
- E_D_TOV

- Data Field Size
- Device Probing
- VC Encoded Address Mode
- Translative Mode
- Per-Frame Route Priority

Use the **configShow** command to display these parameters, and the configure command to change them. Refer to the *IBM SAN Fibre Channel Switch 2109 Model S16 User's Guide,* SC26-7351, for further information about these commands.

When adding switches to an existing fabric, make sure that the switches entering the fabric follow the same rules that presently exist in the fabric. If the new switch has zone configuration information that is already defined, then the alias names, zone names, and zone configurations on the new switch must be either unique (completely different) or exactly match the existing fabrics zone configurations. If the joining switch has zone configuration information, disable the zoning configuration on the joining switch using the **cfgDisable** command. If the joining switch has zoning configuration information that does not need to be imported into the existing fabric, then perform the **cfgClear** and **cfgSave** commands. This action will delete the zone information on the new switch before cascading it into the existing fabric.

To add a new switch with no zoning configuration information to an existing fabric, the steps are:

1. Check the device count of the existing fabric. The **fabricShow** command lists the switches currently in the fabric.
2. Power on the new switch and verify that the domain ID of the new switch does not conflict with the domain IDs of existing switches in the fabric. The **fabricShow** command lists the domain IDs currently in use for the existing fabric. The **switchShow** or **configShow** commands display the domain ID of the new switch. If the domain ID of the new switch is already being used by another switch in the fabric, the domain ID of the new switch must be changed. To make the domain ID of the new switch different, use the **configure** command.
3. Connect the new switch to the existing fabric using one InterSwitch Link (ISL) at a time. The first ISL connection cascades the new switch into the fabric. Additional ISLs can be added for load balancing or redundant paths.
4. Issue the **fabricShow** command to verify that the new switch is in the fabric.
5. Verify that none of the existing switches have dropped from the fabric.
6. Proceed with the attachment of device connections to the new switch.

If two fabrics with zoning information are joined, the fabrics will attempt to merge the two zone configurations. If the zoning information in both fabrics is the same, the two fabrics will merge into a single, larger fabric. If the fabrics have different zone configuration data, then the fabrics will automatically attempt to merge the differences. If the merger is not possible, the two fabrics remain as separate entities. The zone merger does not take place under the following conditions:

- A zone configuration is enabled in both fabrics and is different in both fabrics.
- The name of a zone object in one fabric is used as the name of a different type of zone object in the other fabric.
- The definition of a zone object in one fabric is different from its definition in the other fabric.

If one of these conditions is detected by the switch then error messages are presented in a Telnet session and displayed on the Liquid Crystal Display (LCD) of the 2109-S16 switch. For these reasons, during fabric segmentation the ISL's port light blinks green.

2.2.4 Attaching devices to the fabric

With the SAN fabric established, the next step is to attach devices to the fabric, such as pSeries servers. This section describes the actions for connecting and verifying connectivity of a device to the SAN fabric. The steps are:

1. Use the proper cable to connect the server to the desired 2109 switch port of the SAN fabric.
2. Power on the server, waiting for the device to complete start-up, and perform fabric login. Use the **switchShow** command (see Example 2-4) to verify that the server is logged in. Alternately, the Web-based GUI tool can be used to confirm this information by viewing the name server .

Example 2-4 Results from switchShow command

```
brocade02:admin> switchShow
switchName:     brocade02
switchType:     2.4
switchState:    Online
switchRole:     Principal
switchDomain:   2
switchId:       fffc02
switchWwn:      10:00:00:60:69:10:74:9c
switchBeacon:   OFF
port  0: id  No_Light
port  1: --  No_Module
port  2: id  No_Light
port  3: id  No_Light
port  4: --  No_Module
```

```
port  5: --  No_Module
port  6: id  Online       F-Port  10:00:00:00:c9:21:23:f5
port  7: --  No_Module
port  8: --  No_Module
port  9: id  No_Light
port 10: --  No_Module
port 11: --  No_Module
port 12: id  Online       F-Port  10:00:00:00:c9:27:38:72
port 13: id  No_Light
port 14: --  No_Module
port 15: id  No_Light
```

Note that the 2109 switch port with the connection to the pSeries server should appear as a F-port connection.

3. Verify that the server has successively registered with the Name Server function of the fabric with the **nsShow** command. This informational item can also be confirmed using the Web-based GUI interface.

Example 2-5 Contents of 2109 name server table

```
brocade02:admin> nsShow
The Local Name Server has 2 entries {
 Type Pid    COS     PortName NodeName           TTL(sec)
 N    021600;    2,3;10:00:00:00:c9:21:23:f5;50:05:07:63:00:c0:02:ed; na
    FC4s: FCP [IBM     2105F20         .191]
    Fabric Port Name: 20:06:00:60:69:10:74:9c
 N    021c00;    2,3;10:00:00:00:c9:27:38:72;20:00:00:00:c9:27:38:72; na
    Fabric Port Name: 20:0c:00:60:69:10:74:9c
```

These steps for individual servers should be repeated for all devices that are to become a part of the SAN. If multiple HBAs are incorporated into individual devices, then each connection should be verified as well.

2.3 Configuring McDATA ED-6064

The ED-6064 product has two basic components: a director box and an Enterprise Fabric Connectivity (EFC) Manager. The director box provides dynamic switched connections for servers and SAN devices. ED-6064 supports connection to both mainframe and open systems. Through Inter-Switch Links (ISLs), the director box can also connect to one or more additional directors to form a Fibre Channel multi-switch fabric, managed by a single EFC manager.

The EFC manager consists of an EFC server and a GUI client. The server part consists of an Enterprise Fabric Connectivity (EFC) Manager, ED-6064 Product Manager, and Fabric Manager applications. The Java-based GUI enables the user to manage the director. The EFC server runs on any type of PC, including notebook computers.

2.3.1 Useful tips before you begin

In this section we talk about a few useful tips for installation. More detailed information can be found in *McDATA ED-6064 Director Planning Manual,* 620-000106.

- Make it sure that the following items have been completed.
  - Site plan
  - Configuration planning
  - Evaluation of consideration
  - Planning checklists
- Verify the requirements for connecting the SAN fabric and the devices. Fill out the related planning worksheet. Refer to the *IBM ED-6064 Director Planning Manual* for more information. If you are not sure whether the planned configuration is supported, contact your local Techline team or ATS team.
- Ensure that technical support personnel will be available during the installation process. You may need assistance or cooperation from the customer network administrator, and representatives from other manufacturers or maintainers of equipment attached to the director.
- Check whether the required fiber optic cables (multi-mode or single mode) have been delivered. Verify that the quantity and the types are correct. Ensure that the cables are the correct length and have compatible connectors.

  > **Note:** The fiber optic transceivers used in the 6064 use duplex LC connectors.

2.3.2 Changing Director network configuration

The McDATA ED-6064 Director is delivered with the following default network configuration:

- MAC address: The Message Authentication Code (MAC) address is programmed in FLASH memory on the CTP card at the time of manufacture. The MAC address is unique for each director, and should not be changed. The address is in xx.xx.xx.xx.xx.xx format, where xx is a hexadecimal pair.
- IP address: The factory preset IP is `10.1.1.10`. This address must be changed to a unique value on each director.
- Subnet mask: The subnet mask is `255.0.0.0`. Change this value according to what your network environment requires.
- Gateway address: The gateway address is `0.0.0.0`. If the director is installed on a public LAN (for example, corporate intranet), the gateway address must be changed to the address of the corporate router on the intranet.

Connect an ASCII terminal to the director using an asynchronous RS-232 null modem cable. When you get the > prompt on the screen, type the user-level password. The default password is `password`. The director will respond with an `SSP0>` prompt at the top of the window.

At the `SSP0>` prompt, enter the **`ipconfig`** command. The director will display configuration information listed below:

- MAC address
- IP address (factory preset value is 10.1.1.10)
- Subnet mask (default is 255.0.0.0)
- Gateway address (default is 0.0.0.0)

To change the director network configuration, enter the following at the `SSP0>` prompt:

```
ipconfig <director_ip_addr> <subnet_mask> <gateway_addr>
```

2.3.3 Configuring EFC server

Check the site’s network configuration with the network administrator. If the EFC server is to be installed on an dedicated LAN, we do not need to modify the network configuration of the EFC server.

If the EFC server is to be installed on a public LAN segment, the following network information must be changed to conform to the customer's LAN addressing scheme:

- IP address of EFC server
- Subnet mask
- Gateway address

  The gateway address is the address of the local router for the corporate intranet
- Host name of Domain Name System (DNS)
- DNS domain name

The IP address of the EFC server can be either a static address or a dynamic address obtained from a Dynamic Host Configuration Protocol (DHCP) server.

2.3.4 Configure the Director to the EFC Manager

A newly installed director must be identified to the EFC Manager.

To identify the new director:

1. Start the EFC Connectivity Manager GUI on the PC where the EFC Manager has been installed.
2. You will be prompted to enter a user name and a password. The default user name is `Administrator` and the default password is `password`. You will also enter the IP address or the DNS name of the system where the EFC Manager was installed.
3. Click the Configure icon at the navigation control panel and select New product from the pop-up menu.
4. Type the IP address or DNS host name of the director.
5. Select ED-6064 from the Product Type field and click the OK button. A new director icon will be shown at the Product view as in Figure 2-3 on page 62.

.

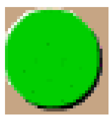
Figure 2-3 Product view

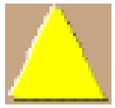
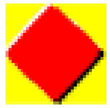
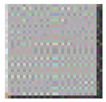
2.3.5 Verify director-to-EFC server communication

Communication must be verified between the director and the EFC server. To verify director-to-server communication:

1. On the Windows NT desktop, open the McDATA Enterprise Fabric Connectivity Manager window.
2. In the Product view, inspect the shape and color of the symbol behind the director icon. Table 2-2 explains operational states and associated symbols.

Table 2-2 Director symbols

Status	Graphic symbol
Operational Director-to-server communication is established and the director is operational. No failures are indicated.	

Status	Graphic symbol
Degraded Director-to-server communication is established, but the director is operating in degraded mode and requires service. This condition is typical if a port or redundant FRU fails.	
Failed Director-to-server communication is established, but the director failed and requires immediate service.	
Status Unknown Director status is unknown because of a network communication failure between the director and EFC server.	

3. Inspect director status at the Hardware view and perform one of the following steps:
   - If director operation appears degraded or a director failure is indicated (FRU alert symbols and a yellow triangle or red diamond at the alert panel), conduct a problem determination procedure as instructed in *McDATA ED-6064 Director Installation and Service Manual,* P/N 620-000108-200 to isolate the problem.
   - If the director appears operational (no FRU alert symbols and a green circle on the alert panel), you are now ready to continue with the installation specific to you configuration, such as:
     - Configure the Call-Home feature
     - Assign user names and passwords
     - Configure the ED-6064 Product Manager Application
     - Cable Fibre Channel Ports

For further information on the remaining installation tasks please refer to the *McDATA ED-6064 Director Installation and Service Manual,* P/N 620-000108-200.

2.4 Configuring INRANGE FC/9000

In this section we describe the procedure required for installing the INRANGE FC/6000 switch. We recommend backing up the configuration information of the current settings before making any changes to an existing INRANGE 9000. Here are the steps for creating a backup of the configuration information:

1. Launch the IN-VSN Manager GUI.
2. From the server window, click on the File drop down menu.
3. Drag the mouse pointer down to the Backup command and click on it.
4. Type in the name of the backup file and then click on the back up button to complete the process. When the file has been successfully backed up, you will receive a confirmation message.

> **Note:** Do not attempt to restore a database without first consulting an qualified INRANGE service person.

2.4.1 Installation

We will assume that the INRANGE 9000 and the PC that will run the IN-VSN mangment tools have been installed successfully using the INRANGE FC/9000-64 Port Fibre Channel switch installation instructions. For more detailed information, refer to the *FC/9000 Enterprise Manager Software Installation and Operation Manual,* 9110509-203, and the *FC/9000 Fibre Channel Switch Planning Guide,* 9106677-101.

The INRANGE Director is delivered with the current supported level of firmware and a default TCP/IP setting of 10.1.1.51 and 10.1.1.52, and a subnet mask setting of 255.255.255.0. There is also a default chassis ID, switch ID and fabric ID. The Chassis ID is used in each domain ID defined. It is the responsibility of an IBM Customer Engineer (CE) to reset all default addresses to reflect the environment that it is being installed into. For added security, the TCP/IP address can only be set or reset by the CE, using an RS232 connection and entering the CE user ID and a corresponding password. The new TCP/IP address will be displayed in a LED panel located on the FCM module. The PC that will run the IN-VSN server needs to be connected to the FCM module using an Ethernet port.

2.4.2 Confirm setup

Once the Ethernet cables have ben installed, it is good practice to do a **ping** test from the IN-VSN server to determine whether the new TCP/IP addresses and cabling have been set up correctly. Once communication between the server and the FCM has been established, the FCM will synchronize the time set on the IN-VSN server and the time set on the FCM modules to ensure that they use consistent timings for event logs.

2.4.3 IN-VSN installation

There are a number of different ways in which the IN-VSN software can be delivered. If no additional software feature codes were specified when ordering an INRANGE Director, you will receive a CD that contains the IN-VSN server code and the client code. The license information contained on the CD permits you to run only one copy of the server and one copy of the client. The server code and the client code can be installed on any PC, as long as it meets with the following minimum requirements:

- Pentium III or compatible PC
- Clock speed 500 MHz
- 128 MB SDRAM
- 4 MB Video RAM
- 13 GB hard drive
- CD-ROM
- 1 Ethernet 10Base-T/100 Base-TX
- 2 external serial communication ports (debug serial cable, external modem)
- Windows NT 4.0 with Service Pack 6a or Windows 2000 Professional

Installing IN-VSN

If you run the setup.exe contained on the CD, both the server and the client will be installed. Note that both the server and client software must be kept at the same level. The server software needs to be started prior to the client. To start either the server or the client software, double click on the corresponding icon. You will be prompted to enter a user name and the corresponding password.

Functions of IN-VSN management client

The GUI-based software operates on a Windows NT platform using a Java Web browser and offers the following capabilities:

- Defining module and port configurations

- Defining zoning parameters
- Monitoring alarms and system performance
- Invoking system diagnostics
- Displaying a graphical overview of port and system configuration
- Displaying port and fabric utilization statistics
- Displaying the status of individual modules
- Performing system diagnostics through the error and event logs

2.4.4 Zoning with INRANGE

All zone configuration changes for an INRANGE Director are performed using the Java-based IN-VSN management software. There are three types of zones that INRANGE Directors support:

- Hard zoning
- Name server zoning
- Broadcast zoning

Hard zoning

Hard zoning follows physical boundaries within a single-stage switch chassis, and limits the communication of a port to only the other ports in the same hard zone. Hard zoning, in certain circumstances, is the only way to provide the required additional level of security, but careful consideration should be applied prior to activating any hard zones, as wrong hard zoning definition can isolate connected devices.

Name server zoning

Name server zoning, or software zoning, allows the division of the fabric (one or more switch chassis) into 256 possible zones. All name sever zoning is defined using an associated IN-VSN client workstation. Soft zones can be defined by using World Wide Port Name (WWPN) addresses.

Broadcast zoning

Broadcast zoning allows the division of a SAN fabric into subareas that broadcast can reach. A particular port may be placed in one or more of these broadcast zones. A port will broadcast to all ports in the same broadcast zone (or zones) in which the port is defined. If hard zoning has also been enabled, then the broadcast zones may not cross the defined hard zone boundaries. Broadcast zoning is typically used to permit the reassignment of a TCP/IP address of a failed server to another server in a clustered configuration.

Port-positioning for performance

As the FC/9000 switch is a full non-blocking device, any switch port can be used. However, to achieve optimum performance, adjacent ports should be used as this offers a shorter network latency (less than 0.6 micro seconds). Examples of adjacent ports are pairs of 1 and 2, 3 and 4, 5 and 6, and 7 and 8 only. The examples are illustrated in Figure 2-4. With other combination of ports the latency value is less than 2.7 microseconds.

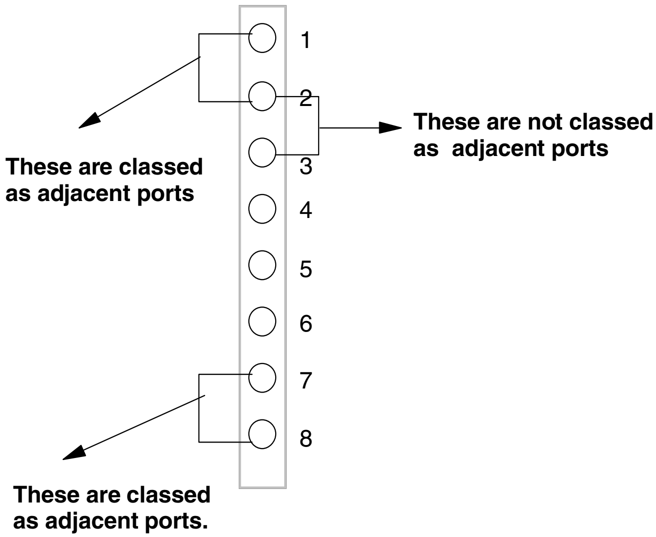


Figure 2-4 Adjacent ports

3

Installing SAN storages

In this chapter we describe the detailed steps required to attach the SAN devices listed below. We begin with a discussion of IBM ESS. We have reorganized the information from a pSeries' point of view. Besides ESS, some other SAN-attachable devices are reviewed. We have put emphasis on points to be considered when implementing those devices. The following topics will be covered in this chapter:

- Attaching an ESS to a pSeries server
- Attaching 3590
- Attaching EMC Symmetrix
- Attaching IBM TotalStorage™ SAN Controller 160
- Attaching SSG 3553 (FAStT500)

3.1 Attaching an ESS to a pSeries server

In the following sections we cover the following topics:

- Section 3.1.1 Required PTFs
- Section 3.1.2 Volume assignment using StorWatch Specialist
- Section 3.1.3 Implementation steps
- Section 3.1.4 Configuring SDD
- Section 3.1.5 Using FlashCopy

3.1.1 Required PTFs

To attach an ESS to pSeries servers, either you or an IBM Service Support Representative (SSR) should get the latest AIX software fixes, firmware updates, and serviceability tools to keep your products updated to the latest supported levels.

The required levels of software and PTFs are summarized in Table 3-1. This information may change, so you should check for the latest information on supported levels and required PTFs to ensure that the installs and operates successfully.

Table 3-1 Supported level for pSeries (as of 9/04/2001)

Software	Level	APARs / PTFs
AIX	4.3.3.0 - 08 5.1.0.0	No APARs No APARs SDD not supported
HACMP	4.3.1 4.4.0 4.4.1	Yes Yes Yes
FC Adapter	3.22A0(FC 6227) 3.82A0(FC 6228)	Yes Yes
ESS	QBLD0517	
IBM 2109 switch	a2.4.1c	
McDATA ED5000 Director	3.2	2032-001 2032-064
INRANGE 9000	2.1.3	2042-001(FC9000)

For more details about the release level for your operating system, see the following Web site:

http://www.storage.ibm.com/hardsoft/products/ess/supserver.htm

3.1.2 Volume assignment using StorWatch Specialist

The IBM ESS comes with a network-based tool that allows for managing and monitoring the ESS. This product is called ESS Specialist. It is a Web-based utility for configuration and administration, using a secure Internet connection (normally using a PC). You need a browser (such as Netscape Navigator, Microsoft Internet Explorer, or Sun Hot Java) that supports Java 1.1.6 or higher.

The ESS Specialist provides the customer with the ability to:

- Monitor error logs
- View the status of the ESS
- View and update the configuration
- Add, delete, or modify host systems
- Configure host ports (both SCSI and Fibre)
- Create RAID or JBOD ranks
- Add volumes and reassign volumes between different hosts
- View communication resource settings, such as TCP/IP configuration and users
- View cluster LIC levels
- Select an authorization level for each user

In this section we will walk through the steps for assigning physical volumes to a host. To do this, direct the Web browser to the IP address of the ESS system.

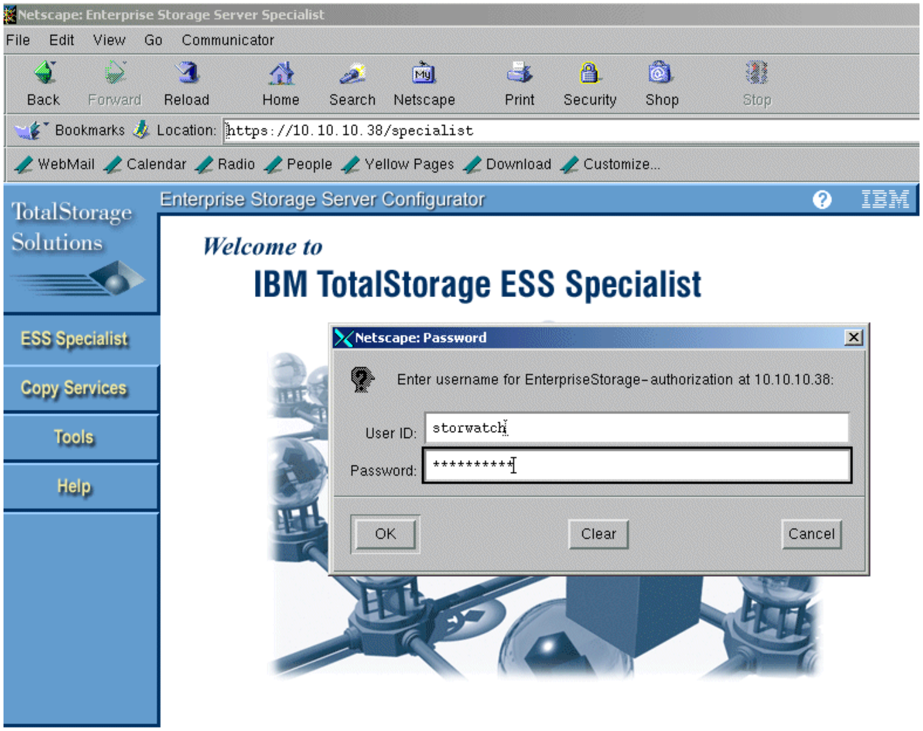


Figure 3-1 StorWatch Specialist login window

To the left is a panel from where we can access the StorWatch Specialist functions. After pressing the ESS Specialist button to start the configuration process. The default user name is `StorWatch` and the password is `specialist` in the above panel. The user name and password characters are all lowercase characters. After a successful login we can move on to configuring the ESS.

Use the Open Systems Storage panel and its buttons to configure new host system attachments and storage, or to modify existing configurations for the open storage type of AIX servers. Figure 3-2 on page 73 shows the menus available on the Open Systems Storage panel.

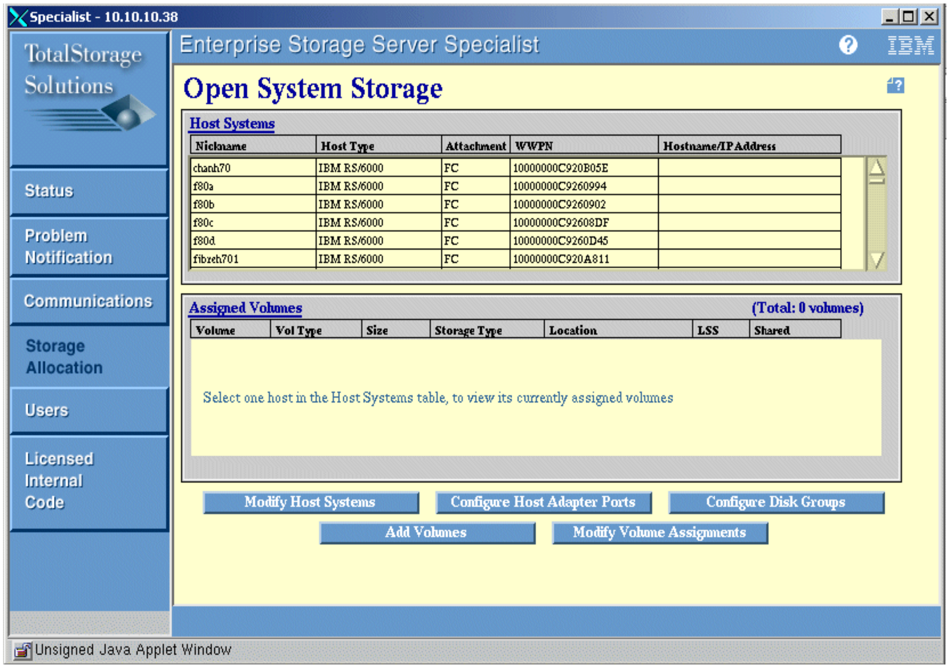


Figure 3-2 Storage Allocation panel

An overview of some of the uses are listed below.

- Modify Host System: Use this panel to identify the host systems that you are attaching to the ESS.
- Configure Host Adapter Ports: Use this panel to set the topology for Fibre Channel ports and to see the Fibre Channel access modes of the ESS.
- Configure Disk Groups: This button links you to the Fixed Block Storage panel. Use this panel to define fixed-block disk groups as RAID or non-RAID and to prepare them for volume assignment. This panel displays all disk groups that are undefined or defined as a fixed block. Once you have selected and defined the disk groups the ESS Specialist displays the group on the Storage Allocation panel.
- Add Volume: This panel is for adding logical volumes. The size of a logical volume can range from 0.5GB to 224GB (full rank size).
- Modify Volume Assignments: To remove or change the assignment of volumes

The next step is to define the host that will have access to the ESS system. Figure 3-3 on page 74 shows the user interface used for this purpose.

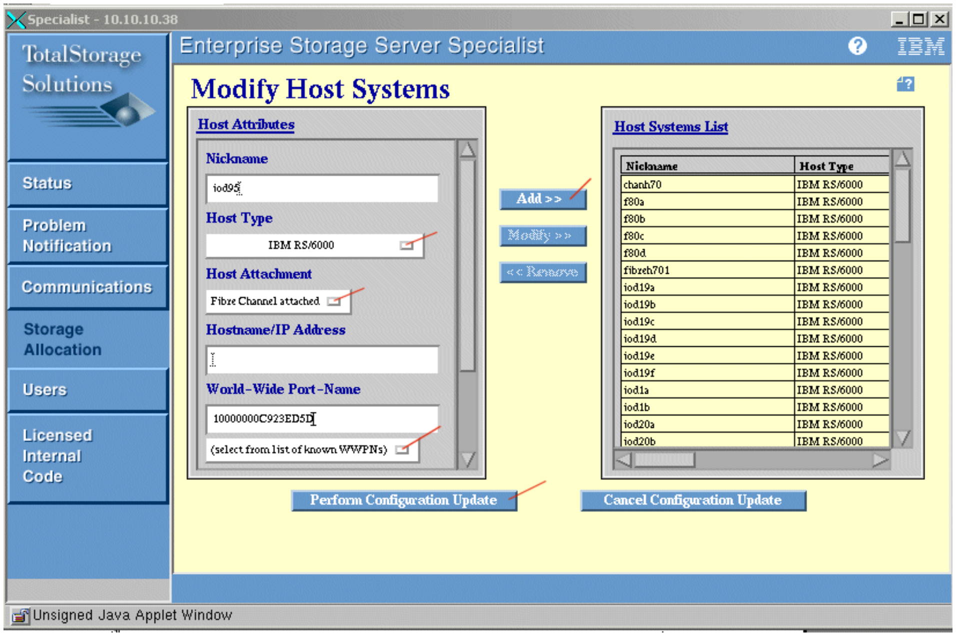


Figure 3-3 Define the host system, WWN (world wide name)

In this step we inform the ESS system of which hosts are connected to the ESS. To do this we select **Storage Allocation** -> **Modify Host Systems**. We are now presented with the Modify Host Systems window, which is shown in Figure 3-3. The procedure is:

1. In the Nickname field type in the host's nickname.
2. In the Host Type field choose **IBM RS/6000** from the drop-down list.
3. In the Host Attachment field, choose **Fibre Channel attached** from the drop down list.
4. You may fill out the Hostname/IP Address field, but this is optional.
5. In the World-Wide Port-Name field, type in or select from the pull-down menu the WWPN of the host bus adapter. Select the World-Wide Port-Name of the Fibre Channel adapter on the pSeries.
6. After completing all of the fields, select **Add>>**. The nickname of the host that you just created should appear on the Host Systems List panel.

7. Repeat steps one to six to create nicknames for each Fibre Channel adapter on the pSeries system.
8. Commit the new configuration information by clicking the Perform Configuration Update button.

Note: If there is more than one adapter on a pSeries system or an RS/6000 system, we recommend that the nickname consists of the host name and some extension to distinguish between them. This helps later when we have to identify devices; for instance, when we setup zoning.

For our example, we have configured two nicknames to represent two Fibre Channel adapters on the pSeries server. The two Fibre Channel interfaces are shown in Figure 3-4.

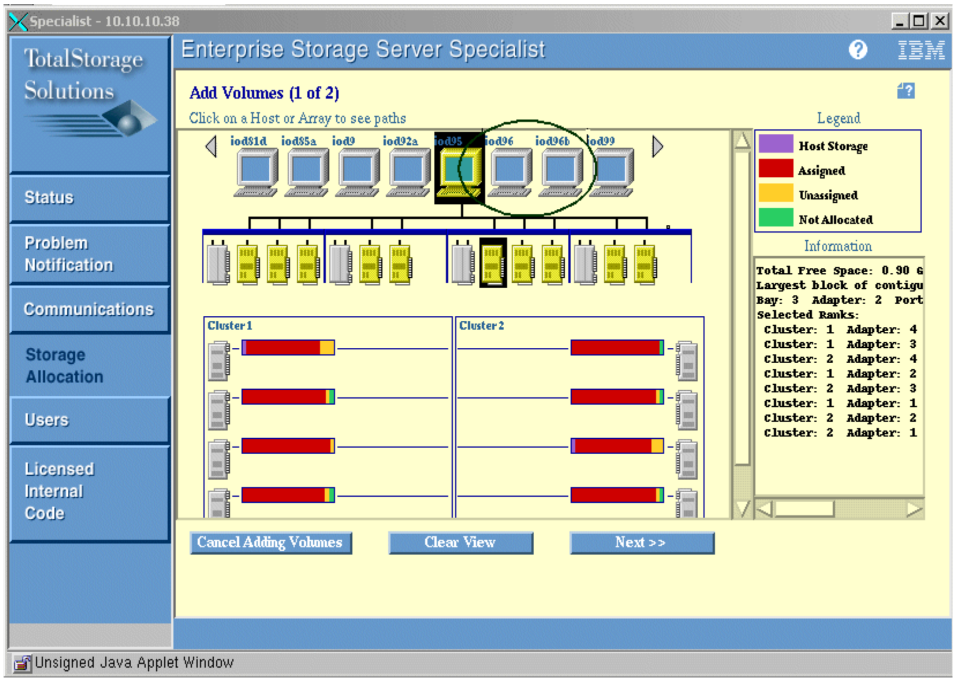


Figure 3-4 Example panel of two nicknames of one pSeries

The next step is to assign a volume to the server. We will assume that volumes were created prior to this step.

What you need to do is choose the volumes that you want to make available to the pSeries server and make the necessary changes on the ESS system.

1. From the Main Menu, go to Storage Allocation.
2. Click Open System Storage.
3. Click Modify Volume Assignments.
4. Highlight the volumes that you want to assign to the host. If multiple contiguous volumes are to be selected, highlight the first, press the Shift key, then highlight the last volume. If non-contiguous, use the Ctrl key, and highlight each volume. Refer to Figure 3-5 for an example.
5. Once done, click **Perform Configuration Update**.
6. Once the progress bar ends, click OK. The volumes you assigned are now available to the host.

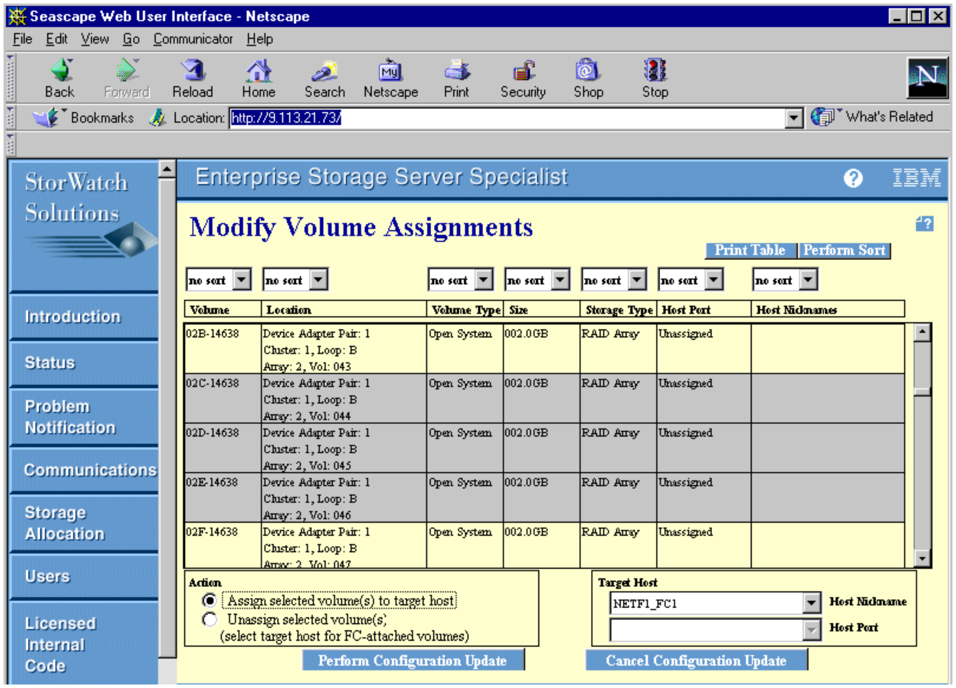


Figure 3-5 Select volumes to assign to the host

You can verify the list of assigned volumes by doing the following:

1. Go back to the Storage Allocation screen.
2. Click **Open System Storage**.

3. In the Host Systems pane search for the nickname of your host and click on the host. A list of volumes assigned to the host will appear in the Assigned Volumes pane.

Figure 3-6 on page 77 shows an example of the list of volumes assigned to our test server.

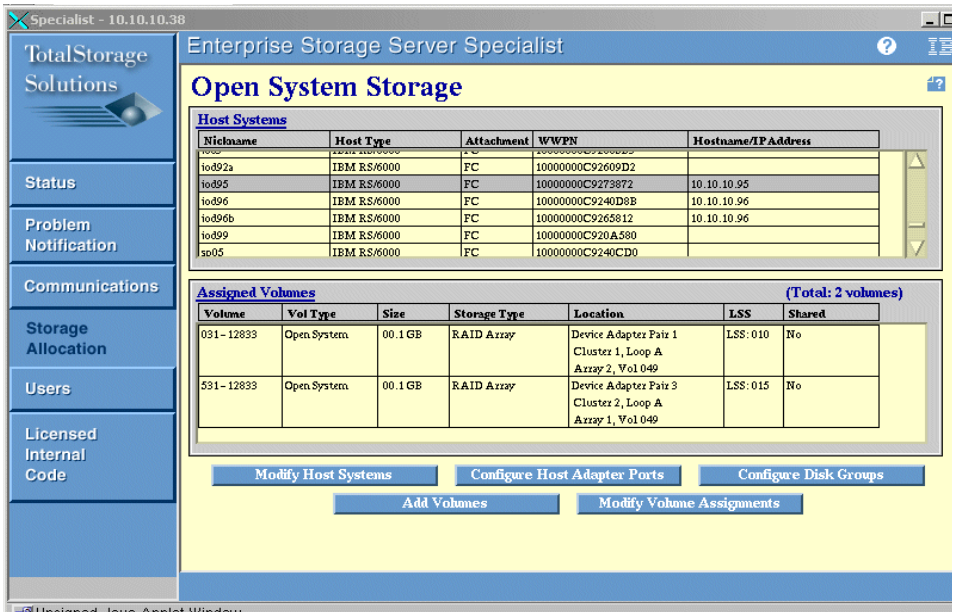


Figure 3-6 Assigned Volume List view

For more about StorWatch Specialist and how to use it refer to the *IBM Enterprise Storage Server Web Interface Users Guide for the ESS Specialist and ESS Copy Services,* SC26-7346.

3.1.3 Implementation steps

In order to make sure that your ESS disks function properly, you will need to install all of the filesets listed in Table 3-2.

The base level filesets for the “devices” filesets in the table are on a CD-ROM named "Gigabit Fiber Channel PCI Adapter Software Upgrade." This CD-ROM comes with your Fiber Channel PCI adapter (or your system that contains the adapter). The latest updates for the filesets can be obtained from the following AIX service FTP site:

ftp://service.software.ibm.com/aix/fixes/v4/devices/

Table 3-2 AIX filesets required for ESS

Fileset	Current Version	Source
devices.fcp.disk.rte	4.3.3.75	Gigabit Fiber Channel PCI Adapter Software Upgrade CD-ROM
ibm2105.rte	32.6.100.7	http://ssddom01.storage.ibm.com/techsup/swtechsup.nsf/support/sddupdates

The steps for configuring ESS drives on an AIX system is as follows:

1. Install the 2105 host attachment fileset (ibm2105.rte).

   From the AIX command line, execute **smitty install_update** to go directly to the installation panel. Install both filesets listed in Table 3-2 on page 78.

2. Configure the Fibre Channel-attached ESS disks.

   The newly installed devices must be configured before they can be used. There are two ways to configure these devices.

   a. Use the **cfgmgr -v** command.

   b. Use the **shutdown -Fr** command to restart the system.

3. Verify the IBM disk drive connectivity.

   To verify the configuration of the ESS on the AIX host system, execute the following command:

   ```
   # lsdev -Ccdisk
   ```

   A list of all IBM ESS disk devices should be displayed as shown below.

Example 3-1 Example of a list of devices

```
# lsdev -Cc disk | grep 2105
hdisk1  Available 14-08-01     IBM FC 2105F20
hdisk2  Available 14-08-01     IBM FC 2105F20
hdisk27 Available 14-08-01     IBM FC 2105E20
...
...
```

If the disk devices appear as Other FC SCSI Disk Drive (as shown in Example 3-2 on page 94, it means that the configuration was not successful.

Example 3-2 Example of a list of devices

```
hdisk0  Available 40-60-00-4,0 16 Bit LVD SCSI Disk Drive
hdisk1  Available 14-08-01     Other FC SCSI Disk Drive
```

```
hdisk2  Available 14-08-01     Other FC SCSI Disk Drive
hdisk3  Available 2D-08-01     Other FC SCSI Disk Drive
...
```

Check whether the proper filesets have been installed on the system. Delete the hdisk devices with the **rmdev** command and reconfigure the disk devices. For more information on the Fibre Channel attachment, see Section 2.1, “Preparing pSeries” on page 46.

3.1.4 Configuring the Subsystem Device Driver

Here we sill discuss configuring the Subsystem Device Driver (SDD).

Introduction to the SDD

The SDD is a software tool that provides load balancing and enhanced data availability capability in configurations with more than one I/O path between the host server and the ESS. The SDD is provided with the ESS at no additional charge. Both SCSI and Fibre Channel (SCSI-FCP) attachment configurations are supported in the AIX, HP-UX, Windows NT, Windows 2000, and Solaris environments.

The IBM Subsystem Device Driver (SDD) resides in the pSeries server with the native disk device drivers for the IBM 2105 ESS. The SDD uses redundant connections between the server and disk storage in an ESS to provide enhanced performance and data availability. In the event of a failure in one path between the host and the ESS, the SDD automatically switches to a different path, if available. This automatic switching is called *failover,* and the SDD minimizes the impact to applications using disk storage on the ESS.

Load balancing can reduce or eliminate I/O bottlenecks that occur when many I/O operations are directed to common devices via the same I/O path. The SDD also helps eliminate a potential single point of failure by automatically rerouting I/O operations when a path failure occurs, thereby supporting enhanced data availability capability.

Each SDD device represents a unique physical device on the ESS. The SDD acts as a pseudo device driver. All I/O operations sent to the SDD are passed to the AIX disk driver after path selection. Refer to Figure 3-7 on page 80 for the layers of device drivers.

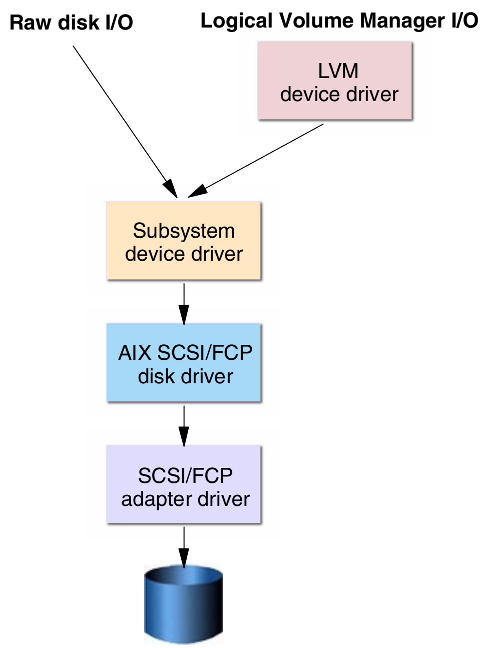


Figure 3-7 Location of SDD layer in communications path with disks

The Subsystem Device Driver provides the following functions:

- Enhanced data availability
- Automatic path failover and recovery to an alternate path
- Dynamic load balancing of multiple paths
- Path selection policies for the AIX operating system
- Concurrent download of licensed internal code

The SDD may operate under different modes and configurations.

- Concurrent Data Access mode

  A total system configuration where simultaneous access to data on a common LUN by more than one host is controlled by system application software, such as Oracle Parallel Server, or file access software that has the ability to deal with access conflicts. The LUN is not involved in access resolution.

- Non-concurrent Data Access mode

  A system configuration where there is no inherent system software control of simultaneous access to the data on a common LUN by more than one host. Therefore, access conflicts must be controlled at the LUN level by a hardware locking facility such as SCSI Reserve/Release.

  The terms single path and multi-path, when used in the path configuration between the ESS and a host system as viewed from the ESS, are defined below:

  S-Path This means that there is a single path from all hosts and adapters to the ESS. That is, there is only one channel path from all the adapters in each host server, which connects to the LUN, such as a single SCSI channel. While this configuration does not provide full redundancy and adapter failover, SDD will still work in this environment to provide ESS concurrent microcode loads for uptime while maintenance is being applied.

  M-Path This means that there are multiple I/O-paths to each LUN in the ESS. That is, there is more than one I/O channel path to each LUN. This would typically be used for adapter/channel redundancy and fail-over purposes. SDD works in this environment to provide adapter failover as well as concurrent ESS microcode loads for uptime while maintenance is being applied.

We can run the Subsystem Device Driver in concurrent and non-concurrent multi-host environments in which more than one host is attached to the same LUNs on an ESS. If you are using pSeries servers running HACMP/6000 check for a support matrix for the SDD on the following Web site:

http://ftp.software.ibm.com/storage/subsystem/tools/f2asdd00.htm#HDRINSTALL

Installing SDD

The installation steps for the SDD are as follows:

1. Check the prerequisites for the AIX host.

   To successfully install the SDD you must have AIX 4.2.1, 4.3.2, or 4.3.3 installed on your host system along with the fixes shown in Figure 3-3 on page 82.

AIX level	PTF number	Component name	Component level
4.2.1	IX62304		
	U451711	perfagent.tools	2.2.1.4
	U453402	bos.rte.libc	4.2.1.9
	U453481	bos.adt.prof	4.2.1.11
	U458416	bos.mp	4.2.1.15
	U458478	bos.rte.tty	4.2.1.14
	U458496	bos.up	4.2.1.15
	U458505	bos.net.tcp.client	4.2.1.19
	U462492	bos.rte.lvm	4.2.1.16
4.3.2	U461953	bos.rte.lvm	4.3.2.4
4.3.3	Maintenance level 4		

Table 3-3 AIX PTF required fixes (as of June/2001)

You must check for the latest information on APARs, maintenance level fixes, and microcode updates at the following downloadable Web site:

http://ssddom01.storage.ibm.com/techsup/swtechsup.nsf/support/sddupdates/

Also you can check the SDD's support for 32-bit and 64-bit applications on AIX machines at the above site.

In order to check to see if the required PTFs are installed, use the **instfix** command as shown in Example 3-3.

Example 3-3 Check for APARs installed

```
# instfix -i | grep IY10201
    All filesets for IY10201 were found.
```

2. Install the Subsystem Device Driver.

   In order to make the SDD work, you will need to install either ibmSdd_432.rte or ibmSdd_433.rte on an AIX 4.3.3 system. For AIX 5.1 systems install ibmSdd_510.rte.

Note: Use the ibmSdd_432.rte fileset for SDD 1.3.0.x when running HACMP with AIX 4.3.3 in concurrent mode. Use the ibmSdd_433.rte fileset for SDD 1.3.0.x when running HACMP with AIX 4.3.3 in non-concurrent mode.

The ibmSdd_*.rte and the ibm2105.rte filesets can be obtained from the following Web site:

http://ssddom01.storage.ibm.com/techsup/swtechsup.nsf/support/sddupdates

From your desktop window, type **smitty install_update** to go directly to the installation panel. Install the ibmSsd_432.rte fileset or the appropriate filesets depending on your operating system level and HACMP requirement.

To verify the installation of the SDD on an AIX system, type the following command:

```
# lslpp -l ibmSdd_*.rte
```

If you have successfully installed the ibmSdd_432.rte package, the output from the command looks like this:

```
Path: /usr/lib/objrepos
  ibmSdd_432.rte            1.3.0.1 COMMITTED  IBM Subsystem Device Driver
                                                for AIX V432 and V433
                                                 without HACMP
Path: /etc/objrepos
  ibmSdd_432.rte            1.3.0.1 COMMITTED  IBM Subsystem Device Driver
                                                for AIX V432 and V433
                                                without HACMP
```

3. Configure the Subsystem Device Driver.

   Type **smitty device** from your desktop window. Select **Data Path Devices**. The Data Path Device panel is displayed.

```
                              Data Path Devices

Move cursor to desired item and press Enter.

  Display Data Path Device Configuration
  Display Data Path Device Status
  Display Data Path Device Adapter Status
  Define and Configure all Data Path Devices
  Configure a Defined Data Path Device
  Remove a Data Path Device




F1=Help             F2=Refresh          F3=Cancel           Esc+8=Image
Esc+9=Shell         Esc+0=Exit          Enter=Do
```

Figure 3-8 Data Path Devices screen

   Highlight Define and Configure all Data Path Devices and press Enter. The configuration process begins.

4. Run the AIX configuration manager to recognize all new hdisk devices for each relevant FCP adapter. Run the **cfgmgr** command *n* times. (*n* is the number of Fibre Channel adapters.)

5. Run the **smitty lvm** command. The screen shown in Example 3-9 on page 101 is displayed. Then select Add a Volume Group with Data Path Devices to create a new volume group. In our example we have created a volume group named datavg1.

```
                              Volume Groups

Move cursor to desired item and press Enter.

[TOP]
  List All Volume Groups
  Add a Volume Group
  Add a Volume Group with Data Path Devices
  Add a Data Path Volume to a Volume Group
  Set Characteristics of a Volume Group
  List Contents of a Volume Group
  Remove a Volume Group
  Activate a Volume Group
  Deactivate a Volume Group
  Import a Volume Group
  Export a Volume Group
  Mirror a Volume Group
  Unmirror a Volume Group
  Synchronize LVM Mirrors
  Back Up a Volume Group
  Back Up a Volume Group with Data Path Devices
[MORE...4]

F1=Help             F2=Refresh          F3=Cancel           Esc+8=Image
Esc+9=Shell         Esc+0=Exit          Enter=Do
```

Figure 3-9 LVM panel for data path devices

6. Check the status of the disk drives.

   If you want to use a command-line interface to verify the configuration, type the **lsvpcfg** command. You should see output similar to Example 3-4 on page 101.

Example 3-4 Output panel of the lsvpcfg command

```
...
vpath222      000657af2b52d19a    datavg1
hdisk248      none                None
hdisk223      none                None
...
```

   In the above example note that hdisk248 and hdisk223 have no physical volume identifier (PVID) value. The PVIDs for these hdisks were deleted from the ODM database when the SMIT interface was executed. The SDD has created the vpath222 device that has a PVID. This behavior is normal, and the reason for doing so is to make the vpath device take over in a failover situation. Remember that AIX recognizes hard disks only through PVIDs.

Therefore vpath222, not hdisk devices, is recognized as a physical volume by the operating system. The operating system then makes I/O calls to the vpath device. A function of the SDD takes care of paths to the actual hdisk devices. In the previous example, the SDD is aware that there are two paths. In the case that one of the paths becomes unavailable, the SDD tries the other path and vice versa.

To query the multiple path information, execute:

```
# datapath query device
```

See the following output for an example. Note that vpath222 has two paths consisting of fscsi2 and fscsi0.

```
DEV#: 222  DEVICE NAME: vpath222  TYPE: 2105F20   SERIAL: 42B12833
==================================================================
Path# Adapter/Hard Disk      State     Mode        Select     Errors
 0     fscsi2/hdisk248       OPEN      NORMAL        71         0
 1     fscsi0/hdisk223       OPEN      NORMAL        73         0
```

7. Change the path-selection policy.

   The SDD provides various path-selection policies that can be used to increase the performance of multi-path configured ESSs and make path failures transparent to applications.

   Load balancing (lb) The path to use for an I/O operation is chosen by estimating the load on the adapter to which each path is attached. The load is a function of the number of I/O operations currently in process. If multiple paths have the same load, a path is chosen at random from those paths.

   Round robin (rr) The path to use for each I/O operation is chosen at random from those paths not used for the last I/O operation. If a device has only two paths, the SDD alternates between the two.

   Failover only (fo) All I/O operations for the device are sent to the same preferred path until the path fails because of I/O errors. Then an alternate path is chosen for subsequent I/O operation.

   You can determine the active attributes of a vapth device. Enter the **lsattr -El vpath*N*** command. The output should look similar to Example 3-5 on page 102.

Example 3-5 Displaying SDD path-selection policy

```
000657af2b52d19a0000000000000000 Data Path Optimizer Parent False
policy df Scheduling Policy         True
active_hdisk hdisk248/42B12833                Active hdisk              False
active_hdisk hdisk223/42B12833
```

The default policy is default load balancing. In order to change SDD path-selection policy, enter:

```
# chdev -l vpathN -a policy=[rr/fo/lb/df]
```

The output should look similar to this:

Example 3-6 Changing the SDD path-selection policy

```
# chdev -l vpath222 -a policy=rr
vpath1 changed
# lsattr -El vpath1
pvid         000657af2b52d19a0000000000000000 Data Path Optimizer Parent False
policy       df                               Scheduling Policy         True
active_hdisk hdisk248/42B12833                Active hdisk              False
active_hdisk hdisk223/42B12833
```

SDD commands

If vpath devices have only one path to hdisk devices, or PVIDs appear on the hdisk devices that are associated with our vpath devices, then use the following SDD commands to fix the problem.

addpaths
Used to dynamically add more paths to SDD devices while they are in Available state. This command will open a new path (or multiple paths) automatically if the vpath is in OPEN state, and the number of paths of the vpath devices will be increased.

hd2vp,vp2hd
Converts hdisk devices into SDD vpath devices. The **vp2hd** script does the opposite action.

dpovgfix
Recovers mixed volume groups. Mixed volume groups happen when a SDD volume group is inactivated (varied off), and certain AIX commands to the hdisk put the PVID attribute of hdisk back into the ODM database.

lsvpcfg
Displays the configuration status of vpath devices.

mkvg4vp
Creates a SDD volume group.

extendvg4vp
Extends an existing SDD volume group.

For more information regarding the usage of SDD commands, refer to the following Web site:

http://ssddom01.storage.ibm.com/techsup/swtechsup.nsf/support/sddupdates

3.1.5 Using FlashCopy

Today organizations require their applications to be available 24 x 365 (24 hours x 365 days). They require high availability, minimal application downtime for maintenance, and the ability to perform data backups with the shortest possible application outage. FlashCopy is a fast point-in-time copy technique designed to replicate all data from an entire volume onto another volume. FlashCopy creates a physical point-in-time copy of data and makes it possible to access both the source and target copies immediately. By creating an instant copy, FlashCopy enables applications using either the source or the target to operate with only a minimal interruption to initiate the copy operation.

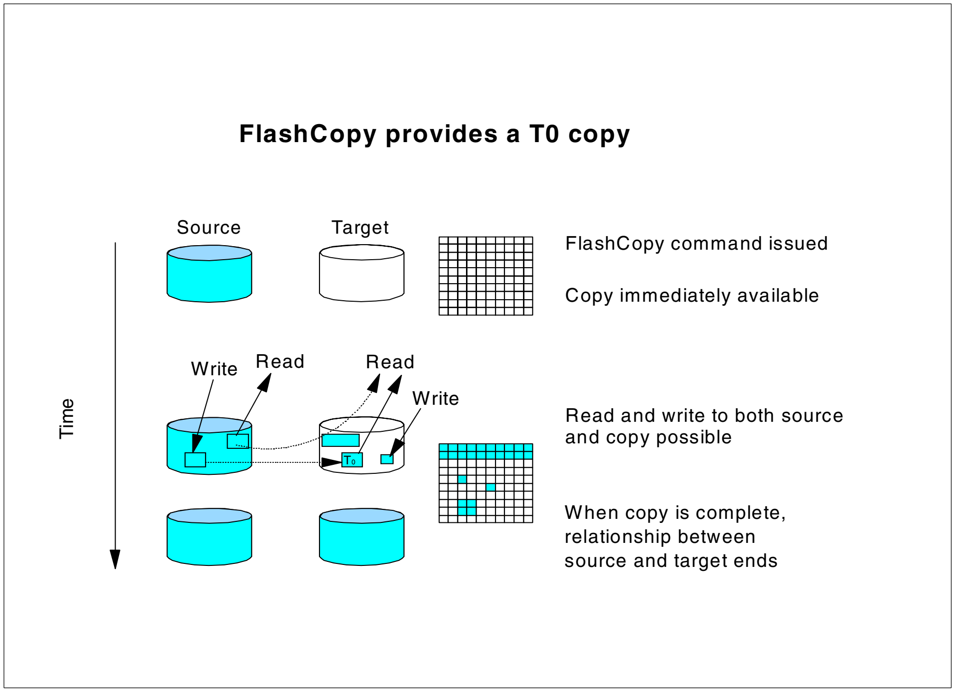


Figure 3-10 FlashCopy operations

As shown in Figure 3-10 on page 88, when a FlashCopy is invoked the command returns to the operating system as soon as the FlashCopy pair has been established and the necessary control bitmaps have been created. This process takes only a few seconds to complete. Thereafter, you have access to a T_0 copy of the source volume. As soon as the pair has been established you can read and write to both the source and the target volumes.

Note: We recommend that you unmount the target volume from all host systems before using the FlashCopy process. Be aware that the FlashCopy process is a destructive operation to the target and will overwrite the data on the target volume.

FlashCopy functions using StorWatch Specialist

We can access the FlashCopy services from the StorWatch Specialist panel. Press the Copy Services button to start the ESS FlashCopy services process. The following is the available menu options on the Copy Services screen.

Volumes To establish a FlashCopy pair. Select the LSS within which you want to perform the FlashCopy. This can be done either in the source pane or target pane appearing in the Volumes panel.

Logical subsystems You can get more detailed information about a single logical subsystem by selecting the LSS and clicking the Information Panel button.

Paths A path is used to send data between the source and the target of PPRC pairs. For more information on the Peer-to Peer Remote Copy (PPRC) functions and implementation steps, see the *Implementing ESS Copy Services on UNIX and Windows NT/2000,* SG24-5757.

Administration To Identify the Copy Services Server and view and clear the Copy Services log file, administrate users for the host command authorization.

The steps for using FlashCopy are:

1. Establish a FlashCopy pair.

   Use the Volumes panel to establish a FlashCopy pair. Select the LSS within which you want to perform the FlashCopy. This can be done in either the source or target area of the Volumes panel. You always need to have two components to establish a FlashCopy pair: a source and a target. With a left-click you select the source volume and with a right-click the target. If you do not know the source volume ID, you can find the Specialist panel, and you can click the Find button at the bottom of the above panel. Figure 3-11 on page 90 shows an example of using the Find function to locate the source volume.

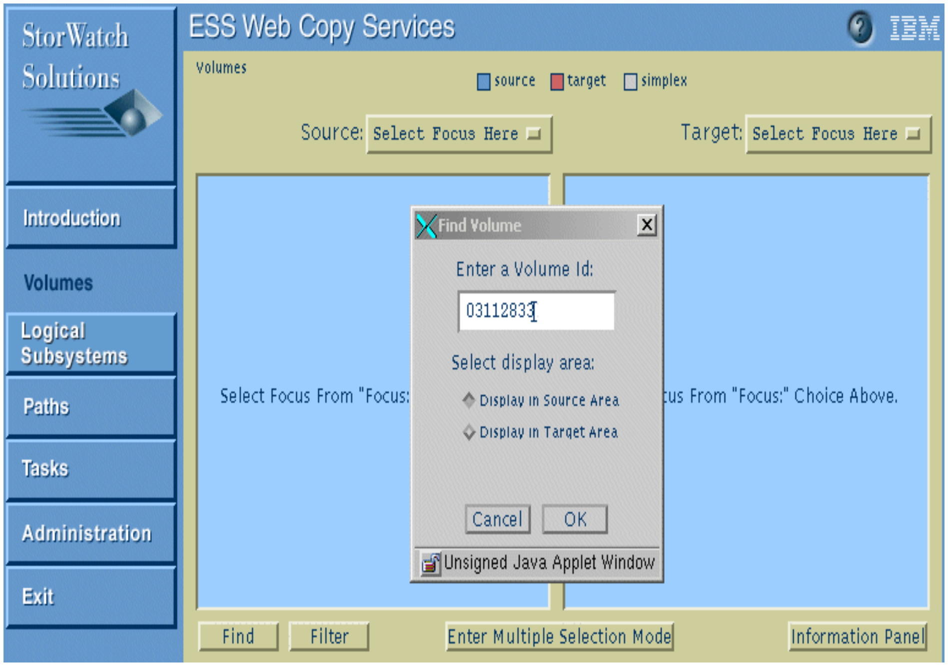


Figure 3-11 Select source volume

Figure 3-12 on page 91 shows an example of using the Find function to locate the target volume.

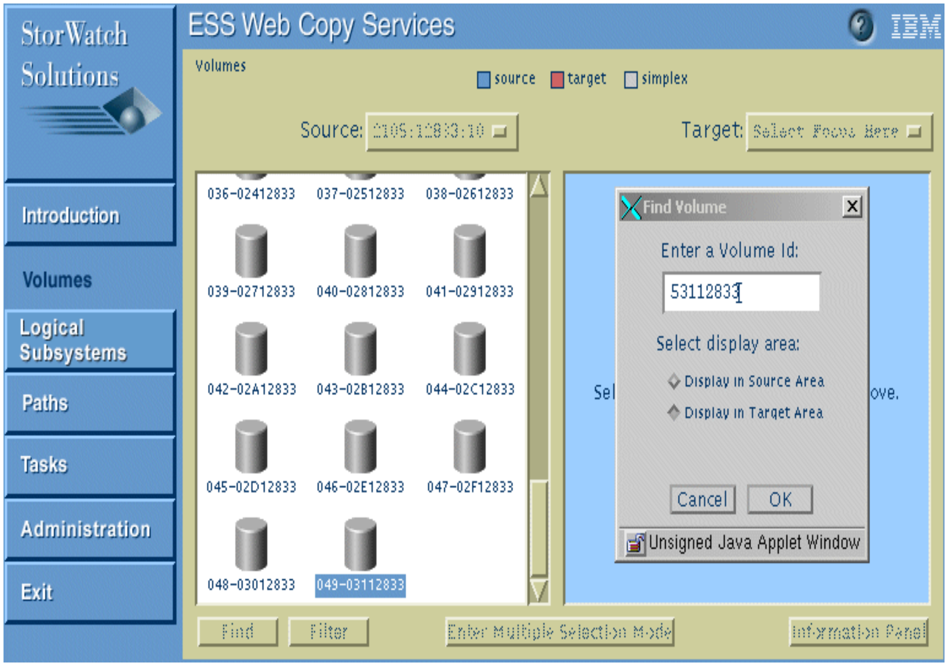


Figure 3-12 Select target volume

During the short period where the copy pair is created, only one of the triangles is filled. Once the relationship has been established successfully, both of the triangles will be solid-colored. Once the source and target volumes have been selected, a screen similar to Figure 3-13 on page 92 will appear.

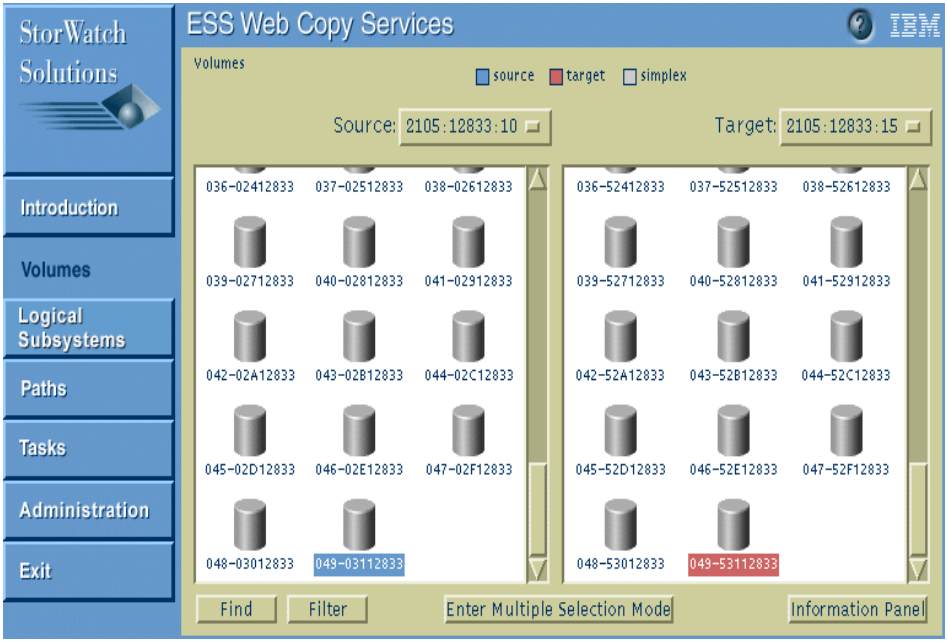


Figure 3-13 Source and target volumes have been selected

2. Select multiple volumes for a FlashCopy task.

   In some cases you may want to establish or withdraw multiple FlashCopy pairs at the same time. Click the Enter multiple selection mode button at the bottom of the Volumes panel. Select one pair at a time beginning with a left-click for the source and a right-click for the target. Once you are finished with the selection, right-click again on the last target volume, which will start the Task Wizard.

Accessing the FlashCopy target volume with recreatevg

In this section we describe a method that can be used to access the FlashCopy target volume on a single AIX system while the source is available and active on the same system.

FlashCopying a source volume's contents causes all of the data structures and identifiers used by AIX's Logical Volume Manager to be duplicated to the target volume. The duplicate definitions cause conflicts within the Logical Volume Manager (LVM). This problem is solved now with a new AIX command, **`/usr/sbin/recreatevg`**. This command is used to recreate an AIX Volume Group (VG) on a set of disks that are copied from another set of disks belonging to a specific volume group. The command will allocate new PVIDs for the member disks and a new volume group identifier (VGID) to the volume group. The command also provides options for renaming the logical volumes with a prefix that you specify, and options to rename labels to specify different mount points for file systems.

The **recreatevg** command is packaged as a PTF for AIX 4.3.3 in APAR IY10456 and higher. APAR IY10456 is officially available in:

- AIX Version 4.3.3 Recommended Maintenance Level 05 (RML05) or higher
- AIX 5L Version 5.1

The following steps are required for making the target volumes created by FlashCopy available to AIX:

1. Stop all applications that access the FlashCopy source volumes.
2. Unmount all related file systems.
3. Establish the FlashCopy pairs with the option of Do not perform background copy.
4. Mount all related file systems.
5. Restart the applications that access the FlashCopy source volumes.
6. The target volumes will now have the same volume group data structures as the source volumes. Clear the PVIDs from the target hdisks to allow a new volume group to be made.

   # chdev -l target_volume_name_1 -a pv=clear
7. Execute the **chdev** command for each target volumes.
8. Create the target volume group and prefix all file system path names with **/backup** and prefix all AIX logical volumes with **bkup**:

   ```
   # recreatevg -y fc_target_vg -L /backup -Y bkup target_volume_name1
   target_volume_name2 ...
   ```

   You must specify the hdisk names of all disk volumes participating in the volume group.
9. Mount the new file systems belonging to the target to make them accessible.

Accessing a FlashCopy target on another pSeries system

The following procedure makes the data on the FlashCopy target volume available to AIX.

1. If the target volume (hdisk) is new to AIX, run the Configuration Manager. If not, the device should first be removed using the **rmdev** command, and then the Configuration Manager should be run:

   ```
   # cfgmgr -v
   ```

2. Check which hdisk is your FlashCopy target:

```
# lsdev -Cc disk | grep 2105
```

3. Import the volume group:

```
# importvg -y <volume group name> <hdisk#>
```

4. Vary on the volume group:

```
# varyonvg <volume group name>
```

5. Verify consistency of all file systems on the FlashCopy target:

```
# fsck -y <file system name>
```

6. Mount the file system:

```
# mount <file system name>
```

FlashCopy Command Line Interface (CLI)

Copy services CLI can be used for creating scripts to automate certain procedures. Using the CLI, we are able to communicate with the ESS Copy Services Server from the AIX command line. An example usage of CLI would be automating tasks, such as doing a FlashCopy by invoking the Copy Services commands within a customized script.

The required AIX level is AIX 4.3.3 or higher with a Java JDK level of 1.1.8. The name of the fileset to be installed is **ibm2105cli.rte** and the destination of the files is set to **/usr/lpp/ibm2105cli**.

We can verify that the Command Line Interface fileset was installed properly with the following command:

```
# lslpp -L ibm2105cli.rte
```

The CLI code is closely tied to the Enterprise Storage Server microcode level. When ESS microcode updates occur, it is required that Copy Services CLI are also updated.

For more information about usage of the CLI commands, see the *Implementing ESS Copy Services on UNIX and Windows NT/2000,* SG24-5757.

3.2 Attaching 3590

Before we begin please check that you have the correct levels of code and drivers. Table 3-4 on page 95 presents sample information.

The latest information can be found at:

http://www.storage.ibm.com/hardsoft/tape/3590/3590attach

Table 3-4 3590 code and driver prerequisites

Component	Level	File name	Web site location
3590 FC Microcode Level	DOIE_32E		obtained via FMR tape or Feature Code 500
Gigabit Fibre Channel PCI Adapter - FC 6227	3.21A0 AIX 4.3.3 3.22A0 AIX5.1	df1000f7.bin	http://www.rs6000.ibm.com/support/micro/download.html
Gigabit Fibre Channel 64Bit PCI Adapter - FC6228	3.82A0	df1000f9.bin	http://www.rs6000.ibm.com/support/micro/download.html APAR IY16132 is also required for AIX 4.3.3
Atape Driver Level	6.1.4.0	Atape.6.1.4.0.bin	ftp://ftp.software.ibm.com/storage/devdrvr/AIX/
AIX Fixes	4.3.3		http://service.software.ibm.com/rs6k/fixes.html APAR IY17356
AIX Fixes	5L		http://service.software.ibm.com/rs6k/fixes.html Version 5.1
Switches			IBM 2109 9 (Note 6) and IBM 2031 (Note 4)

Component	Level	File name	Web site location
Notes			4.IBM 2031 McData ES-1000 Loop Switch requires AIX 4.3.3 with maintenance package 4330-08 or later on pSeries (RS6000) models. 3590 FC Microcode Level D0IE_32E is required. 6.If a 2109 switch is used, Code Level V2.2.1A is the minimum level recommended

The basic steps that you will need to perform to install the IBM 3590 are:

1. Install the FC Adapter Card and driver. Refer to Section 2.1, “Preparing pSeries” on page 46.
2. Check the microcode level on your Fibre Channel adapter is correct, by using the **lscfg -vl fcs0** command and reading the Z9 field as in Example 3-7.

Example 3-7 Sample output of lscfg

```
DEVICE            LOCATION          DESCRIPTION
fcs0              14-08             FC Adapter

      Part Number.................09P1162
      EC Level....................D
      Serial Number...............KT10801909
      Manufacturer................0010
      FRU Number..................09P1173
      Network Address.............10000000C9273872
      ROS Level and ID............02903290
      Device Specific.(Z0)........4002206D
      Device Specific.(Z1)........10020193
      Device Specific.(Z2)........3001506D
      Device Specific.(Z3)........02000909
      Device Specific.(Z4)........FF101450
      Device Specific.(Z5)........02903290
      Device Specific.(Z6)........06113290
```

```
Device Specific.(Z7)........07113290
Device Specific.(Z8)........20000000C9273872
Device Specific.(Z9)........SS3.22A0
Device Specific.(ZA)........S1F3.22A0
Device Specific.(ZB)........S2F3.22A0
Device Specific.(YL)........U0.1-P1-I2/Q1
```

Note: FC 6227 requires at least level SF321.A0 on AIX 4.3.3. FC 6227 requires at least level S1F3.22A0 on AIX 5.1. FC 6228 requires level C1D3.82A0 for AIX 4.3.3 and AIX 5.1. APAR IY16132 is also required for AIX 4.3.3.

3. Install the Atape device driver (Atape.n.n.n.n.bin) for Magstar and Ultrium products.
4. If necessary, update the microcode. It can be downloaded from:

   http://www.rs6000.ibm.com/support/micro/download.html#adapter
5. Connect your storage devices to the appropriate switch or director port.
6. Connect the Fibre Channel cable from the host FC Adapter to the selected switch/Director port
7. Power-on, in the following sequence:
   a. The tape drives
   b. The SAN switch, or enable the ports
   c. Run **cfgmgr -v**

At the end of these steps, to see that the AIX recognizes the tape devices, enter:

```
# lsdev -Cctape
rmt0 Available 2D-08-01 IBM 3590 Tape Drive and Medium Changer (FCP)
rmt1 Available 2D-08-01 IBM 3590 Tape Drive and Medium Changer (FCP)
```

We can check whether the 3590 drives are communicating with the AXI server by examining the messages appearing on the operator panel of the 3590 devices.

Example messages would be:

```
PORT0 ID=02 14 26
PORT1 ID:aa bb cd
```

= indicates that a hard address is being used.

: indicates that a soft address is being used.

Where:

aa = Fibre domain address or switch ID (on some switches)

bb = Fibre Area Address or switch port (on some switches)

cd = arbitrated loop

For more information regarding problem determination procedures for 3590, refer to Section 4.6.2, “IBM 3590 Tape Systems” on page 199.

Note: For performance reasons, it is strongly recommended that you have a separate Fibre Channel host bus adapter and connection for the tape path from each server to the IBM 2109 Fibre Channel switch. This is because I/Os to tape are typically quite long and can conflict with the large number of short I/Os to disk normally associated with transactional processing.

3.3 Attaching EMC Symmetrix

This section gives an overview of the tasks that need to be done on the pSeries’s side when you are attaching an EMC Symmetrix box to a SAN fabric. It is important to keep in mind that the Symmetrix box has been configured properly by an EMC support representative. Therefore, be sure to arrange for EMC technical support before you begin the installation.

3.3.1 Updating AIX ODM files for EMC Symmetrix

To install the latest AIX ODM definition files for EMC Symmetrix, please go to the following FTP site:

ftp://ftp.emc.com/pub/symm3000/aix/

Then select ODM_DEFINITIONS folder. You will find some Tape Archive (TAR) files. Download the correct file for your level of AIX. At the time of writing this book, two files were available:

EMC_AIX_ODM.4.3.tar.Z for AIX 4.3.3
EMC_AIX_ODM.5.1.tar.Z for AIX 5.1

Once the files are down loaded, uncompress them with the **uncompress** command.

```
# uncompress EMC_AIX_ODM.4.3.tar.Z
#tar -tvf EMC_AIX_ODM.20001016.tar
-rwxr-xr-x   0 0    21348 Nov 07 11:05:07 2001 README.433
-rw-r--r--   0 0  1126400 Nov 07 11:04:43 2001 EMC.AIX.4.3.3.3
```

Read the README file first. It gives detailed installation instructions. You can use the normal smitty install method to install the filesets. There are four filesets in the EMC.AIX.4.3.3.3 file:

```
Symmetrix.fcscsi.rte 4.3.3
Symmetrix.aix.rte 4.3.3.3
Symmetrix.ha.rte 4.3.3.3
Symmetrix.fcp.rte 4.3.3.3
```

In most cases you may only need to install Symmetrix.aix.rte 4.3.3.3 and Symmetrix.fcp.rte 4.3.3.3. Symmetrix.fcscsi.rte 4.3.3 requires the EMC Fibre kit and Symmetrix.ha.rte 4.3.3.3 requires HACMP Concurrent Logical Volume Manager (CLVM). If you select all of them, you will see the following error message:

```
Symmetrix.ha.rte          4.3.3.2         Requisite failure
Symmetrix.fcscsi.rte      4.3.3.2         Requisite failure
```

3.3.2 Validating installation

Once the files are loaded, do the following steps:

1. Check that the fiber connections from the host to the Symmetrix are correct.
2. Check with your EMC support representative as to whether volumes have been properly created on the EMC box. This task is normally carried by an EMC support representative with EMC’s Volume Logix software.
3. Run **cfgmgr.**
4. Enter **lsdev -Ccdisk** to display the configured EMC disks.

```
hdisk8 Available 2D-08-01     EMC Symmetrix FCP Raid1
hdisk9 Available 2D-08-01     EMC Symmetrix FCP Raid1
hdisk10 Available 2D-08-01     EMC Symmetrix FCP Raid1
hdisk11 Available 2D-08-01     EMC Symmetrix FCP Raid1
```

In case that disks were not properly configured, check the following things:

1. Check that the drivers are installed by entering the **lslpp -h | grep -p Symmetrix** command.

```
Symmetrix.aix.rte 4.3.3.2   COMMIT       COMPLETE     09/17/01     11:02:10
Symmetrix.fcp.rte 4.3.3.2   COMMIT       COMPLETE     09/17/01     11:02:10
```

2. Check your cabling.
3. More importantly, get the EMC representative to check the bin (configuration) file on the Symmetrix, or the configuration created on the EMC Volume Logix software.

3.3.3 Limitations

- AIX must not be configured to see the same disks through multiple paths, unless PowerPath is installed.
- For loop configurations the following applies:
  - Maximum of four (4) hosts per loop
  - Maximum of one (1) adapter per host per loop
  - Maximum of one (1) hub per loop
- For Direct Attach configurations the following applies:
  - 1-16 adapters depending on the host used

3.4 Attaching IBM TotalStorage™ SAN Controller 160

The IBM TotalStorage™ SAN Controller 160 is a high-function serial disk controller that provides Fibre Channel connectivity by converting Fibre Channel protocol to SSA protocol. It provides the IBM Serial Disk Storage 7133 with improved scalability, availability, and tape backup capabilities.

We will assume that you have successfully installed and connected the 160 Controller to the SSA disks using the *7140 TotalStorage SAN Control Model 160 Router Installation and User's Guide,* GC26-7433 and completed the diagnostics successfully.

3.4.1 Supported Adapters Servers SAN products

The latest information on supported configurations can be found at:

`http://www.storage.ibm.com/hardsoft/products/san160/160_support_matrix.html`

At the time of writing this redbook, the following prerequisites are required to attach a 7140-160 device:

For a Gigabit Fibre Channel Adapter for 64-bit PCI Bus (FC #6228):

- Adapter Firmware Version 02C03890 (3.82A0) or higher
- devices.pci.df1000f9.rte 4.3.3.1 or higher

- devices.pci.df1000f7.com 4.3.3.50
- devices.pci.df1000f9.diag 4.3.3.0
- devices.common.IBM.fc.rtc 4.3.3.51
- devices.fcp.disk.rte 4.3.3.50

For a Gigabit Fibre Channel Adapter for PCI Bus (FC #6227):

- Adapter firmware version 02903290 (3.22A0) or higher
- devices.pci.df1000f7.rte 4.3.3.0
- devices.pci.df1000f7.com 4.3.3.29
- devices.pci.df1000f7.diag 4.3.3.27
- devices.common.IBM.fc.rtc 4.3.3.12
- devices.fcp.disk.rte 4.3.3.27

A tar archive of these files named 7140_aix.tar can be downloaded at:

http://www.storage.ibm.com/hardsoft/products/san160/san160.html

3.5 Attaching SSG 3553 (FAStT500)

In this section we will briefly describe IBM's SSG 3552, which is also known as FAStT500.

3.5.1 Introduction

The Fibre Array Storage Technology (FAStT) storage disk subsystem consists of a FAStT500 storage server enclosure and from 1 to 10 EXP500 storage expansion enclosures. The FAStT500 storage server is a 4U rack-mountable enclosure, or a drawer, which has two RAID controller units, redundant power supplies and fans for high availability. Each RAID controller unit has a 300 MHz AMD K6 processor and from 256 MB to 512 MB of battery-backed cache, which is used to boost performance.

On the host side, each RAID controller is connected to two Fibre Channel loops through two mini-hubs, which each have two GBICs. A total of four mini-hubs (two GBICs per hub) can be used to connect up to eight host ports, or four redundant hosts, on the host side of the RAID controller. Mini-hubs 1 and 3 are connected to RAID controller 1 and mini-hubs 2 and 4 are connected to RAID controller 2.

On the drive side, each RAID controller is attached to four Fibre Channel loops, which are normally used as two redundant loops. Each redundant loop can be attached to up to five EXP500 disk expansion enclosures, for a total of up to ten EXP500 expansion enclosures for both RAID controllers. The FAStT500 offers the following features:

- RAID options of RAID 0, 1, 2, 3, 5, and 0+1
- Storage consolidation
  - Up to 16 host systems per FAStT500 storage server
  - Up to 8 AIX host servers per FAStT500 via SAN switch
- 256MB or optional maximum 512 MB of cache with battery backup
- Flexibility to configure short wave connections for distances of up to 500 M or long wave GBICs for distances of up to 10 Km
  - FC #2010 (PN 03K9308) one short-wave GBIC
  - FC #2020 (PN 03K9307) one long-wave GBIC
- Scalability
- High availability through automatic I/O path failover
  - Both controllers can be simultaneously active to provide seamless failover capability in case of emergency.
- Concurrent maintenance with hot-pluggable components
  - Dual RAID controller units
  - Power supply modules
  - Fans
  - Cache back-up battery

3.5.2 Requirements for installation of FAStT500

This section provides the following system-requirement information for AIX operating systems:

- Hardware requirements
- Software requirements
- AIX restrictions

Hardware requirements

Table 3-5 lists the supported versions of hardware required for attaching FASTtT500.

Table 3-5 Hardware requirements

Product name	Model	Product release and firmware version
IBM FAStT500 RAID Controller Enclosure Unit	3552–1RU	Firmware 04.01.02.21 NVSRAM NV3552R710NT008
IBM FAStT500 EXP500 Storage Expansion Unit	3560–1RU	ESM 9163 ESM9165
IBM Adapter	FC 6227	3.22A0, 3.22A1
IBM Adapter	FC 6228	3.82A1
Brocade switch	IBM 2109	2.1.4.e
McData switch	IBM 2032–001	3.2
McData switch	IBM 2031–16	1.2.2
McData switch	IBM 2031–64	1.2.2
INRANGE switch	IBM 2042	2.0.2 3.0.0

Note: Read the README file that is shipped with the product and access `http://www.ibm.com/storage/fast500` to ensure that you have the latest versions of the firmware and NVSRAM and for information on host-adapter device drivers.

Software requirements

Check the versions of the following software prerequisites:

- AIX Version 4.3.3
- IBM RDAC Driver: devices.fcp.disk.array.rte, Version 4.3.3.77 or higher
- Other required PTF filesets

Table 3-6 lists the PTF filesets required for installation of the Redundant Disk Array Controller (RDAC) Driver.

Table 3-6 PTF filesets required for RDAC driver installation

PTF filesets	Version
devices.fcp.disk.array.diag	4.3.3.50
devices.fcp.disk.array.rte	4.3.3.75*
devices.fcp.disk.rte	4.3.3.75
devices.common.IBM.fc.rte	4.3.3.75
devices.pci.df1000f7.com	4.3.3.76
devices.pci.df1000f7.rte	4.3.3.75
devices.scsi.scarray.rte	4.3.3.50
* The RDAC Driver Update Version 4.3.3.77 requires the presence of Version 4.3.3.75.	

The AIX driver files listed in the Table 3-6 on page 120 can be obtained at:

www.techsupport.services.ibm.com/rs6k/fixdb.html

AIX restrictions

Note the following restrictions:

- The F-RAID Manager is not supported.
- The FAStT500 is supported in a switched environment only.
- One to ten EXP500 Storage Expansion Units are supported (up to 100 disk drives).
- AVT/ADT must be disabled for path redundancy to be effective.
- Concurrent NVSRAM upgrades are not supported. You must stop all I/O to the controllers before you upgrade NVSRAM.
- Access volumes are not supported in this release; failure to delete them can cause system errors.
- Up to four mini-hubs are supported on both the host and the drive side of the RAID Controller.
- FAStT500 supports up to eight hosts per IBM @server pSeries.
- The maximum number of partitions per AIX host per FAStT500 is one.

3.5.3 Installing host software

The RDAC package is an IBM multi-path device driver that provides controller-failover support when a failure occurs anywhere along the Fibre Channel I/O path. All AIX hosts in your storage subsystem must have the RDAC multi-path driver installed on them. This section describes how to install that host software package. Perform the following procedure to install the RDAC software on an AIX system.

> **Note:** The RDAC must be installed on all AIX systems connected to the FASTtT500 disk subsystem.

Complete the following procedure to install the RDAC software (devices.fcp.disk.array.rte) on an AIX system. Repeat this procedure for all AIX systems connected to the FAStT500 disk subsystem.

1. Verify the version numbers of the AIX device drivers listed in Table 3-6 on page 104. You can use:

```
# lslpp -l <fileset_name>
```

2. Mount the CD at the /cdrom mount point:

```
# mount -v cdrfs -o ro /dev/cd0 /cdrom
```

3. Change to the appropriate directory:

```
# cd /cdrom/AIX/usr/sys/inst.images
```

4. Start the installation process:

```
# installp -acXgd .all
```

5. Verify that the correct version of the software was successfully installed:

```
# lslpp -ah devices.fcp.disk.array.rte
```

6. Unmount the CD with the following command, and then remove the CD from the CD-ROM drive:

```
# umount /cdrom
```

7. Shut down and restart the AIX system. This must be done for the software changes to take effect.

```
# shutdown -r
```

Installing and setting attributes of the RDAC software for AIX

The RDAC driver must be installed on all AIX hosts that will be attached to a FAStT500 disk subsystem.

The RDAC driver creates the following devices that represent the FAStT disk subsystem configuration:

darX The disk array router represents the entire array, including current and deferred paths to all LUNs (hdisks on AIX).

dac0, dac1 The disk array controller devices represent the two controllers of the array.

hdiskX These devices represent individual LUNs on the array.

When these devices are configured, the ODM is updated with default parameters. You must change some of these parameters for the AIX to operate properly, while others can be changed to enhance the operability of your configuration. Use the **lsattr -El** command to view attribute settings on an AIX system. When these devices are configured, the ODM is updated with default parameters. Example 3-8 shows the attribute settings for a dar device.

Example 3-8 Attribute settings of a sample dar device

```
#lsattr -El dar0
act_controller dac0,dac1                     Active Controllers False
all_controller dac0,dac1                     Available Controllers False
held_in_reset none                           Held-in-reset controller True
load_balancing no                            Dynamic Load Balancing True
autorecovery no                              Autorecover after failure is
                                             corrected True
lun_bitmap 00000000000000000000000000001010  LUN Ownership bitmap True
hlthchk_freq 600                             Health check frequency in seconds
                                             True
aen_freq 600                                 Polled AE frequency in seconds
                                             True
balance_freq 600                             Dynamic Load Balancing frequency
                                             in seconds True
fast_write_ok yes                            Fast Write available False
cache_size 512                               Cache size for both controllers
                                             False
```

Required attribute settings

This section lists the required attribute settings for dar and hdisk devices and shows how to set them using the **chdev -l** command. To make the attribute changes permanent in the Customized Devices object class, use the -P option.

- Attribute settings for dar devices: For multi-initiator configurations, the autorecovery attribute must be set to no.

  ```
  # chdev -l dar0 -a autorecovery=no
  ```

On single host systems, the load_balancing attribute can be set to `yes` to enhance performance.

```
# chdev -l dar0 -a load_balancing=yes
```

Note: On multi-host systems, the load_balancing attribute must be set to `no`.

- Attribute settings for hdisk devices: Setting the queue_depth attribute to the appropriate value is important for system performance. For large, multihost configurations, always set the attribute to less than 10.

  Use the following formula to determine the maximum queue depth for your system:

  512/(number-of-hosts x LUNs-per-host)

  For example, a system with four hosts, each with 32 LUNs (the maximum number of LUNs per AIX host), would have a maximum queue depth of 4:

  512/(4 x 32) = 4

  In this case, you would set the queue_depth attribute for hdiskX as follows:

```
# chdev -l hdiskX -a queue_depth=4 -P
```

Attention: If you do not set the queue depth to the proper level you might experience loss of filesystems and system panics.

The write_cache attribute does not control whether or not write caching is on; the FAStT500 controllers do. By default, write caching is on for the FAStT500 controllers.

Making the disk available to the host

After the FAStT500 disk subsystem has been set up, volumes have been assigned to the host, and the RDAC driver has been installed, use the following command to probe for the new devices:

```
# cfgmgr -v
```

Next, use the **`lsdev -Cc disk`** command to see if the device driver recognizes each FAStT500 LUN as a 3552 Disk Array Device.

```
# lsdev -Cc disk
hdisk0 Available 10-88-00-8,0 16 Bit LVD SCSI Disk Drive
hdisk1 Available 20-58-01 3552 Disk Array Device
hdisk2 Available 20-58-01 3552 Disk Array Device
hdisk3 Available 20-58-01 3552 Disk Array Device
hdisk4 Available 20-58-01 3552 Disk Array Device
```

One of the attributes listed by the **lsattr -El** command is the controller serial number (controller_SN) of that dac. In the following example, the controller has a serial number of 1T04810361 and is represented by dac1.

Example 3-9 Attributes of dac devices

```
#lsattr -El dac1
passive_control no                  passive controller      False
alt_held_reset  no                  Alternate held in reset False
controller_SN   1T04810361          Controller serial number False
ctrl_type       3552                Controller Type         False
cache_size      512                 Cache Size in MBytes    False
scsi_id         0x210513            SCSI ID                 False
lun_id          0x0                 Logical Unit Number     False
utm_lun_id      none                Logical Unit Number     False
location                            Location Label          True
ww_name         0x201200a0b807b856  World Wide Name         False
GLM_type        low                 GLM type                False
```

Identifying device names and bus names

After the operating system device names are found, those names must be correlated to the primary and secondary paths of the device, and then from each path to its associated logical drive.

Use the **lsdev**, f**get_config** and **lsattr** commands to get information about device names and bus numbers.

- Using the **lsdev** command

  This section shows how to use the **lsdev** command to get information about disk array routers (dars), disk array controllers (dacs), and hdisks. The following example uses the **lsdev** command to show the status of the dar, which represents the entire FAStT500 array. “darX” keeps track of the status of each dacX, and of the current and preferred paths to each hdisk.

  ```
  # lsdev -C |grep dar
  dar0 Available 3552 Disk Array Router
  ```

  Next execute **lsdev -C | grep dac**. The result sows the two dacs that represent the disk array controllers. The third column is the location code column. Two distinct paths are represented by the 91-08-01 and 11-08-01 values. Each AIX system has its own set of location codes that describe the internal path of that device, including bus and host-adapter locations. See the service manual for your system type to identify device locations.

```
# lsdev -C |grep dac
dac0 Available 91-08-01 3552 Disk Array Controller
dac1 Available 11-08-01 3552 Disk Array Controller
```

Then execute **lsdev -Ccdisk**, which lists available disks, including disks (LUNs) of type 3552 (FAStT500) recognized by this AIX host system. The third column shows the location code of each device. Notice that there are four disks per path, or four disks per dac (controller) in this example.

```
# lsdev -Cc disk
hdisk0 Available 40-60-00-4,0 16 Bit LVD SCSI Disk Drive
hdisk1 Available 91-08-01 3552 Disk Array Device
hdisk2 Available 91-08-01 3552 Disk Array Device
hdisk3 Available 91-08-01 3552 Disk Array Device
hdisk4 Available 11-08-01 3552 Disk Array Device
hdisk5 Available 11-08-01 3552 Disk Array Device
hdisk6 Available 11-08-01 3552 Disk Array Device
hdisk7 Available 11-08-01 3552 Disk Array Device
hdisk8 Available 91-08-01 3552 Disk Array Device
```

- Using the **fget_config** command

  The **fget_config** command displays the current owner of each hdisk. The following example shows that it is a quick way to determine which LUN (hdisk) is actively owned by a controller.

```
#fget_config -l dar0
dac0 ACTIVE dac1 ACTIVE
dac0-hdisk1
dac0-hdisk2
dac0-hdisk3
dac1-hdisk4
dac1-hdisk5
dac1-hdisk6
dac1-hdisk7
dac0-hdisk8
```

4

Problem determination guide

In any Fibre Channel environment, whether simple or complex, problems will be encountered eventually. With servers, fabric devices, and a multitude of storage components, problems can occur anywhere. Sometimes the problem is obvious, such as a total failure of a power supply. In this situation the source of the problem is easily detected and corrected. Often, however, a problem occurs that requires troubleshooting techniques of data collection from multiple sources within the SAN. Worse case situations require assistance from technical support groups with advanced isolation and resolution methods, tools, and techniques.

The purpose of this chapter is to assist the reader in finding a problem within a SAN environment and the cause of the problem. Once the problem and its cause are identified, the reader should be able to determine the appropriate corrective actions. Since the scope and nature of these corrective actions is so widely varied, it is beyond the scope of this publication to adequately cover the full range of fixes. Some corrective actions are well-defined, such as replacement of a defective piece of hardware. Also, there are a number of additional resources available to the reader for determining the appropriate action to correct a problem. Such resources are service guides for specific products and remote technical support groups.

This chapter provides basic, and some advanced, problem determination procedures for Fibre Channel SANs. Considering the potential complexity of Fibre Channel environments, no single process can guarantee complete problem isolation and resolution. We discuss a methodical approach to problem determination. First, the SAN environment needs to be documented to establish a baseline of how the SAN is supposed to work. We describe the various sources and types of data to collect for problem determination throughout the SAN, such as the servers, the SAN fabric components, and storage devices. Finally, we discuss what items of information should be readily available when contacting a technical support group for further assistance. The topics cover:

- Information to have before problems occur
- Problem determination process guide
- Problem determination flows
- Checking the pSeries server
- Checking the fabric
- Checking the storage systems
- Getting further assistance

4.1 Information to have before problems occur

Effective problem determination starts with a good understanding of the SAN and the components involved. Simply stated, the more information documenting a SAN, the better. Problem determination will be easier with the following:

- Diagrams of the SAN

  These drawings should document both the logical and the physical connections across the entire SAN. The physical diagrams need to include information about fiber cable runs between devices, particularly if long distances are involved. If patch panels are incorporated into the SAN, their locations must also be indicated in the diagrams.

- Current code levels with release notes

  Many problems are introduced and/or resolved with different levels of code and PTFs on specific devices. With this information on-hand, known issues can be identified and resolved with minimal effort.

- Configuration of equipment

  Many SAN devices, particularly SAN fabric switches, have the means to download, and upload, configurations to a server over a LAN connection. Having a back-up source of device configurations provides a means to compare the current configuration with a past configuration when looking for configuration changes.

- User manuals for each type of device used in the SAN environment

  User manuals (installation, configuration, and service guides) for each type of device within the SAN provide points of reference for commands. In addition, the manuals usually have detailed explanations of various troubleshooting actions for a particular device.

- Knowledge of the baseline performance of the SAN

  Performance-related issues need to have a reference to establish the degree of impact of the problem.

- Setting device clocks to the same time

  While not exactly an item that can be printed or stored in a file, setting the clocks on SAN devices to the same time is crucial. With data being collected from a multitude of sources, it is important to be able to correlate the different pieces of information.

Another benefit for having a SAN fully documented before problems occur is that the people with knowledge about a particular SAN are not always available. In other words, how can a SAN be fixed if its capabilities, features, functions, and normal behavior patterns are unknown?

4.2 Problem determination process guide

This section gives guidance and a method on how to troubleshoot a SAN. We can break problem determination into three stages. These stages are defined as follows:

- Stage 1: Initial installation of a device. In this stage, the device has been unpacked, powered up, and given its initial configuration. The device has not been connected to the SAN, but may have a LAN connection. Problems occurring in this stage are resolved with user documentation for the device. Failures when power is first applied are typically indicated with LEDs. Likewise, configuration issues can be identified and resolved using the specific product user documentation. Common problems at this stage are faulty devices and components, connectivity to a device over the LAN, and issues with implementing the initial configuration.
- Stage 2: Initial connection of the device to the SAN. In this stage devices are connected to the SAN for the first time with minimal traffic levels. Once the physical connections are established, the devices create logical connections with other devices, completing with normal port status and full connectivity. Common problems during this stage are SAN connectivity problems due to cabling issues, incorrect code levels causing incompatibility, and configuration conflicts with existing SAN devices.
- Stage 3: Routine traffic flows, or normal SAN operations. Problems during this stage deal with interruptions to normal SAN operation and performance issues. The most common problems are a result of changes made to the SAN.

For the purpose of this publication, we do not cover stage 1 issues, since the device is not connected to the SAN environment.

4.2.1 Define the problem

Time-efficient troubleshooting starts with an accurate description of the problem. Thus, a series of basic questions need to be answered to help direct problem determination efforts. Some of these basic questions are:

- What exactly is the problem?

  Define what aspect or portion of the SAN is not working correctly. This description should be kept simple and concise. This statement is not a recap of all symptoms. Rather, what action, function, or application does not work as intended.

- Is this SAN environment a new installation?

  If a new SAN installation, then configuration issues are typically the cause of the problem. Otherwise, some form of connectivity issue is likely.

- What has been changed recently?

  Typically, problems in an established SAN are a result of a change to one or more devices, or the addition or removal of a device. With luck, the failure is immediately noticed after the change (but that does not always happen).

- What is the scope of the problem?

  For example, you discover a server has a connectivity problem with a given storage device. Is the problem observed on other servers? Can the server connect with other storage devices? Are there any common devices within the scope of the problem?

- Are any specific protocols, functions, or applications related to the problem?

  For example, a file can be tarred from a server to tape with no problems, yet the tape backup application has abnormal ends (abends). If the problem appears to be related to a specific application or function, then configuration issues are typically the cause of the problem.

- Is the problem constant or intermittent?

  Constantly occurring problems are much easier to isolate. Intermittent issues typically involve more sophisticated troubleshooting techniques.

- Can the problem be duplicated?

  If the failure is constant, then this question is not a high priority. However, the ability to cause a problem to occur after a given set of actions can provide valuable clues about the nature and scope of the problem.

- Do any actions reduce the problem symptoms or stop it entirely?

  Actions that improve a faulty situation are additional clues about the source of the problem. A complete description of these actions can be a valuable resource of information for the problem determination process.

- Is the problem more prevalent at a given time of day or load condition?

  Try to determine if there is a pattern with the occurrences of the problem. Does it only occur with specific conditions or at specific times in the SAN? For example, does the problem only occur during times of peak traffic or only at night when the tape back-up utilities are running?

4.2.2 Tools and procedures

Tools to gather data are varied depending on the problem, but generally will include the following:

- User input

  Usually, problems in the SAN environment are observed by end users. End users are a potential source of information about the symptoms and scope of the problem.

- Visual indicators (LEDs, power indicators, etc.)

  A number of SAN products provide indications that help quickly pinpoint a hardware problem. If not sure of abnormal LED patterns, refer to the product user documentation.

- Console commands and other management interfaces from affected devices

  Typically, console commands, along with using a product's GUI management tool, are the most useful tools for troubleshooting. With a minimum of commands, the status of connections are easily determined. This information is then used to further refine problem determination actions to isolate and correct the problem.

- Error logs from affected devices

  Many SAN devices maintain an internal error log that can be viewed as a powerful resource during problem determination. The use and interpretation of these logs usually require more specialized knowledge than do console commands. However, their use for certain types of problems can be invaluable.

- Internal traces from affected devices

  Some SAN devices are capable of generating internal traces of events and certain types of Fibre Channel traffic. Like the error logs, internal traces require a more in-depth knowledge of the products. Internal traces can be very good for providing information about problems that are infrequent or can be duplicated.

- Protocol analyzers and external trace tools

  Fibre Channel protocol analyzers and trace tools are very expensive. Fortunately, many problems can be resolved without these tools. For performance issues, intermittent problems, and connectivity faults, these are the preferred tools.

- SAN management applications

  SNMP management applications can be another useful tool for the problem monitoring of a SAN. Typically, these applications communicate over a LAN connection and, thus, have no impact on Fibre Channel traffic.

4.3 Problem determination flows

The main goal of the problem determination process is to narrow the scope of troubleshooting from the general to the specific. In many instances, the problem and its solution become obvious after gathering much of the information needed to define the problem. When the problem is not easily discovered, this section describes additional ways to help with isolating and solving problems that may be encountered.

To simplify the problem determination process, most problems can be considered as either a connectivity issue (device A cannot communicate with device B) or performance issue (slow transfer of data). Connectivity problems can be further classified into hardware issues (for example, no power, bad port, or HBA), media defects, or configuration issues. By their nature, performance issues are usually more difficult to isolate because they can involve more components of a SAN.

With a description of the problem in conjunction with diagrams of the SAN, the reader will be able to determine the best starting point to begin the problem determination process.

4.3.1 Getting started

The problem determination process starts with the notification that a problem exists. The notification can come from a variety of sources. The most widely recognized source of this notification are users of the SAN resources. For example, one or more users may contact the network administrator about a server that is not responding, or that a particular application is generating numerous errors during data read or write operations. The main point of the notification is that something may not be working as designed.

Since SANs can range from the simple to the complex, we begin by describing the problem determination process for a hypothetical SAN, shown in Figure 4-1 on page 118. This generic SAN is the example used to describe the basic concepts for determining a starting point for further investigations. The example LAN consists of several pSeries servers connecting on one side of the fabric. There are at least two switches that comprise the SAN fabric. The storage resources can be either disk devices or tape units.

While defining the problem, one of the most critical pieces of information to determine is the scope of the problem. Does the problem appear to be related to just one server or several? Is the server(s) losing connectivity to storage resources? If so, how many resources (either disk or tape) appear to have been disconnected? Once the scope of the problem is better understood, a logical starting troubleshooting point within the SAN is the place that is the most common point of failure. Thus, the scope of the problem is narrowed down to one of the following situation types:

- Only one device
- One-to-one connectivity (such as server to storage or disk to tape)
- One-to-many connectivity (for example, a server cannot access any storage)
- Many-to-many connectivity (typically a result of an issue within the fabric)

Each of these scenarios is discussed further in the following sections. To help better understand each situation, we describe a problem example.

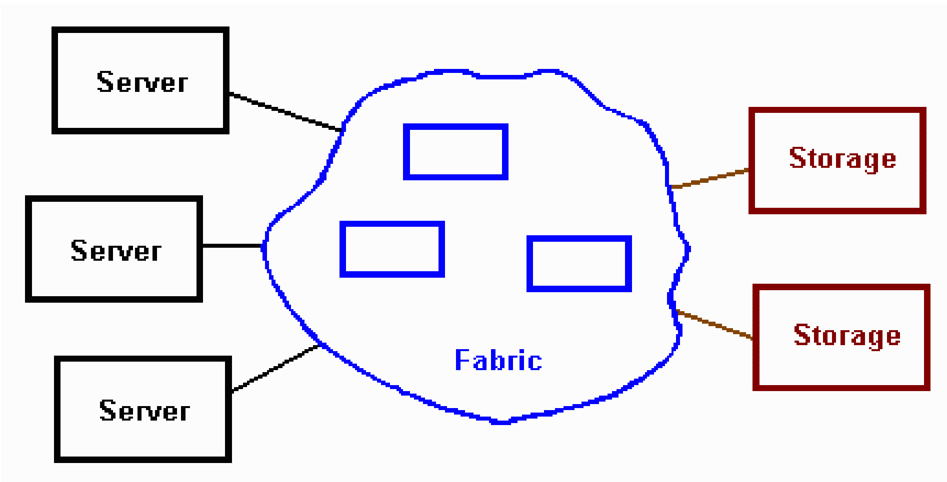


Figure 4-1 Generic example of a SAN

Only one device problem determination starting point

In the situation where only one specific device within the SAN is having problems, the place to start in-depth troubleshooting is with the affected device itself. Once additional information has been captured on the device, the scope of the problem may expand to a different scenario, such as a one-to-one issue, and thus require data collection from other points in the SAN.

With an identified starting point in the only-one-device scenario, additional information about the problem is quickly obtained. In Figure 4-2 on page 119, the one device with an issue is one of the servers, but it could just as easily be a storage device. Whatever the type of device, this problem situation will make use of the product's user documentation and/or service guide for problem determination actions.

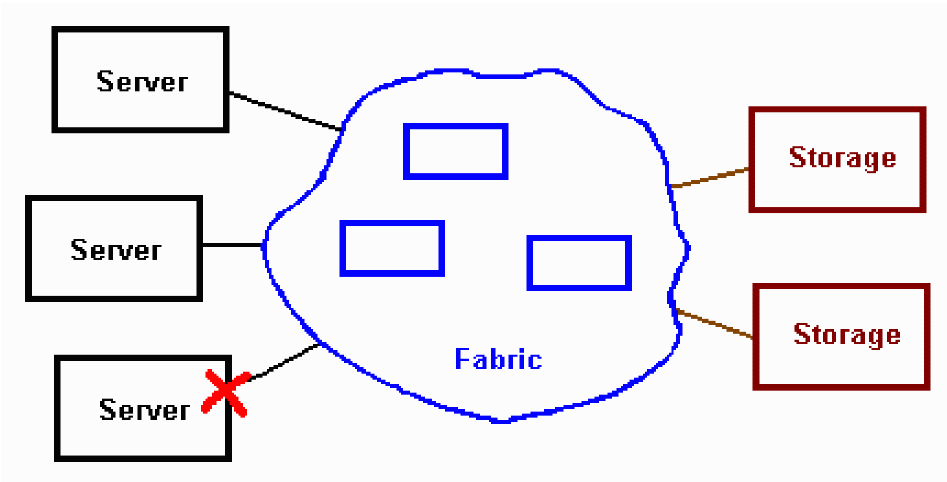


Figure 4-2 Example of a SAN with only one faulty device

One-to-one problem determination starting point

The one-to-one connectivity issue is usually encountered as part of an expansion to an existing SAN or a fault in the connection between two edge devices. This scenario is shown in our example SAN diagram in Figure 4-3 on page 120. In a one-to-one situation, the best starting point is with one of the ends of the connection. Once the start point is identified, the problem determination process is simply the verification of the various links between the two devices. A typical problem scenario is when a server cannot communicate with a storage device.

For the connectivity-type problems in a one-to-one scenario, Section 4.3.2 provides a problem determination flow for checking the connections between the two edge devices. In summary, the process starts with one edge device of the failing connection and proceeds along the physical links to the other device. If the problem is a consistent loss of communications, the source of the disconnect should be readily found and then corrected. Examples of this problem type are:

- Faulty port in a HBA or switch
- Loose connection of fiber cable into a port or patch panel
- Zoning configuration error in the fabric
- Incorrect masking of LUNs on storage device

The one-to-one scenario can also encompass situations where the problem is intermittent. In this situation, additional data collection will be used to make modifications to the troubleshooting process. Section 4.3.3 provides a flow for intermittent-type problems. As an example of this type of failure, suppose that a pSeries server can successfully back up small data files to tape storage but fails with files larger than a given size. In this example, the physical connections appear to be functional since smaller sized files are transferred through the SAN onto tape.

Continuing with this example, the physical links can be verified during the large file transfer to be absolutely sure (as outlined in Section 4.3.2, “Connectivity problem isolation flows” on page 140); and the connection between devices is okay. Some additional questions to answer are:

- Does the failure occur with a specific application? If so, does the application have the ability to perform advanced diagnostics?
- Do error logs of the edge devices provide any clues about the nature of the problem?
- Do switches within the fabric have entries in their error logs when the problem occurred?

If the collected data still does not provide sufficient information to determine the source of the problem, then refer to Section 4.7, “Further Assistance” on page 217.

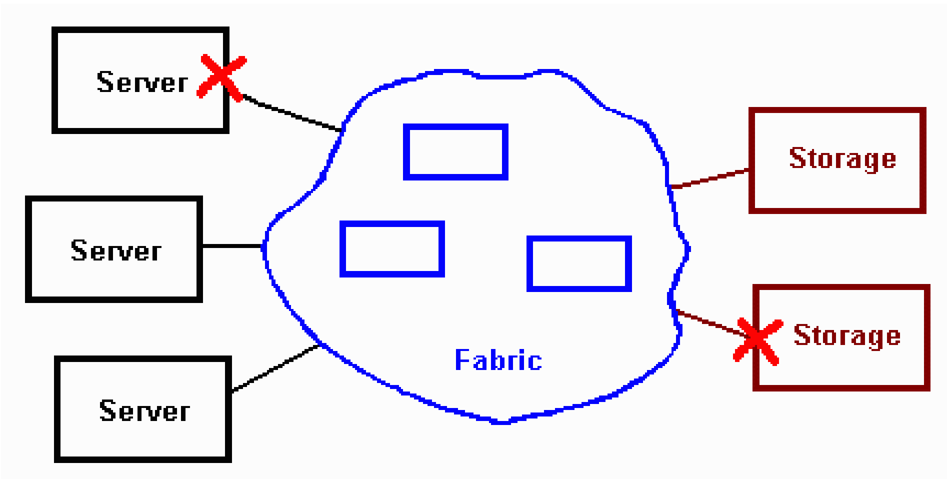


Figure 4-3 Example of a SAN with a one-to-one connectivity issue

One-to-many problem determination starting point

One-to-many connectivity problems usually occur as a result of:

- Either hardware or application configuration errors within the “one” device
- Zoning configuration conflicts within the fabric
- A failure in the physical linkage between a device on the one end and multiple devices on the other end

We provide an example of this problem type in Figure 4-4 on page 122. In this example, one server has lost connectivity to several of its storage resources. During initial troubleshooting, check other servers that communicate with the same storage resources. If other servers are not having connectivity to storage issues, then the starting point for troubleshooting is with the server. If other servers using the same storage devices are having problems, the problem is considered to be a many-to-many scenario, which is discussed in the following section.

In our example problem, the starting point is the single edge device (the server in this example). From the initial point for troubleshooting, check all physical links from the server into the SAN fabric. If this physical linkage is okay, then the fabric is the next logical place to investigate. If there are multiple switches in the fabric, start with the switch that has the physical connection back to the server in question.

If the problem does not appear to be a connectivity issue, additional data is needed to adapt the problem determination process. In this case, the problem is intermittent and not a hard failure. Section 4.3.3, “Advanced problem isolation techniques” on page 146 discusses intermittent type issues and describes a methodology for troubleshooting. As an example of this failure, we use an automated back-up application on a pSeries server, which copies data from disks to tape. In our scenario, there are random failures during the back-up process.

During the initial investigation phase, the physical links are verified (as outlined in Section 4.3.2, “Connectivity problem isolation flows” on page 140) and no problems are uncovered. Some additional questions to consider are:

- Does the backup application have advanced diagnostics and/or error logging? If so, what messages are logged? What are the results of the diagnostic routines?
- Do error logs from any of the fabric devices provide clues about the nature of the problem?
- Do the storage devices have error logs? Are there any indications of the problem reported?

If the collected data does not provide sufficient information for determining the cause of the problem, refer to Section 4.7, “Getting further assistance” on page 217.

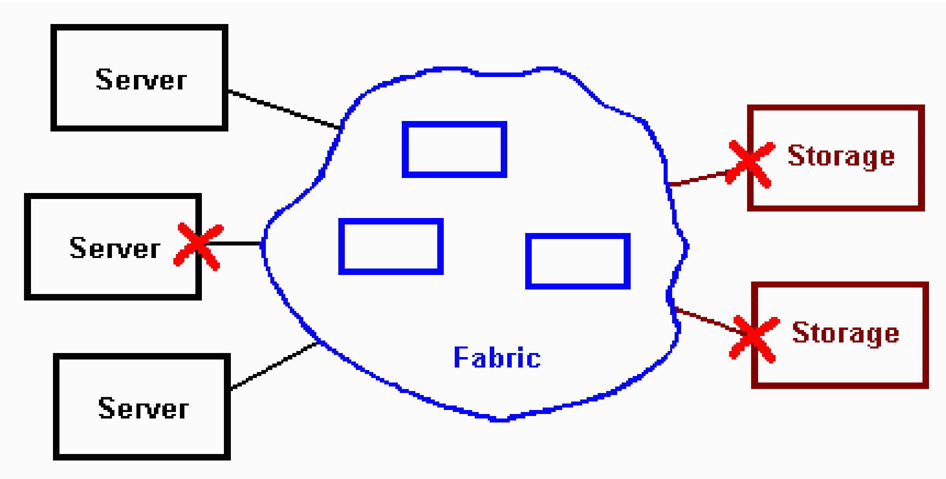


Figure 4-4 Example of a SAN with a one-to-many connectivity issue

Many-to-many problem determination starting point

Easily the most difficult of all problem situations, the cause for a many-to-many problem can be located anywhere within the SAN. With such a wide problem scope, the best starting point is within the SAN fabric itself. Some of the potential causes of this situation are:

- Changes to the SAN due to an addition or removal of devices
- Zone configuration conflicts in the SAN fabric
- The hardware failure of a device’s Fibre Channel port with minimal zoning implemented

Here is an example of a many-to-many situation: Most of the servers are not able to communicate with any of their storage resources. In this example SAN, minimal zoning has been implemented. There is at least one server that is still able to communicate with its storage resources. This server is not in the same zone as the problematic servers. A fast check of the port LEDs of the links between the problem servers and the fabric do not indicate any obvious failures. A switch has been added recently into the SAN fabric.

In our example, the SAN fabric is the first place to begin in-depth troubleshooting. When starting problem determination with the fabric, the key item to verify is the condition of the links between the various fabric components. If these connections are not at fault, then the next step is to work outward from the fabric to the affected edge devices. At the stage of expanding outwards, the main search is for common conditions between affected edge devices. Thus, a comparison for the common critical information.

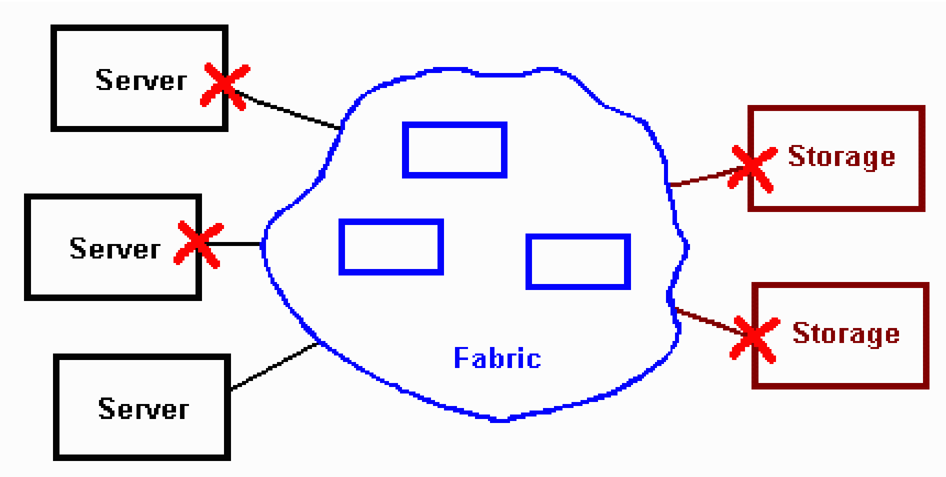


Figure 4-5 Example of a SAN with a many-to-many connectivity issue

In our example, a closer status inspection of the fabric's interconnecting ports shows that a portion of the SAN fabric has become partitioned. A review of the error logs from partitioned and non-partitioned switches confirms a zoning configuration mismatch shortly before the first calls from users. From these clues, we concluded that the newly added switch had some prior zone configuration that was incompatible with the existing fabric's zoning. Removing the switch and restoring the original configuration corrects the problem. The new switch has all zoning information removed and then is inserted into the SAN fabric again.

When the problem is not constant in a many-to-many situation, the same basic approach is used. In conjunction with checking the linkages, additional data is collected and reviewed for possible hints about the problem. Error logs from applications and devices can be a valuable source of information. In addition, advanced troubleshooting tools such as Fibre Channel trace tools can be utilized for the most difficult situations. However, when implementing such extreme measures, the problem must be defined sufficiently so that the tools are located properly within the SAN so to provide meaningful information. For most situations at this stage, additional data and technical information is usually needed. For more details on how to proceed, refer to Section 4.7, “Getting further assistance” on page 217.

4.3.2 Connectivity problem isolation flows

At this point, one or more systems within the SAN have been identified as good starting locations for the next problem determination step. This section provides two maps for checking the physical connections and logical links. We describe each step that is outlined in the maps. To further illustrate the process by example, we used the simple SAN created for this publication. Figure 4-6 on page 125 is a connections diagram of the test SAN that was created and used for this publication. As an additional reference, Table 4-1 on page 125 lists the devices by name and WWN. Units with multiple HBAs only list the ports that connects to this SAN.

The basic premise is to start with a server as the starting edge device. We then proceed to follow the connections through the SAN environment to the storage edge device on the other side of the SAN fabric. As the various steps are described, we will provide examples using our sample SAN.

If your initial investigations have determined that a fabric device or storage device is the best initial point, then the maps can be re-ordered as needed. We are using the server-to-storage device as a simplified method to explain the process. This methodology is aimed primarily at troubleshooting constantly-occurring problems. The process described in this section can also provide some valuable clues for cases with intermittent problems.

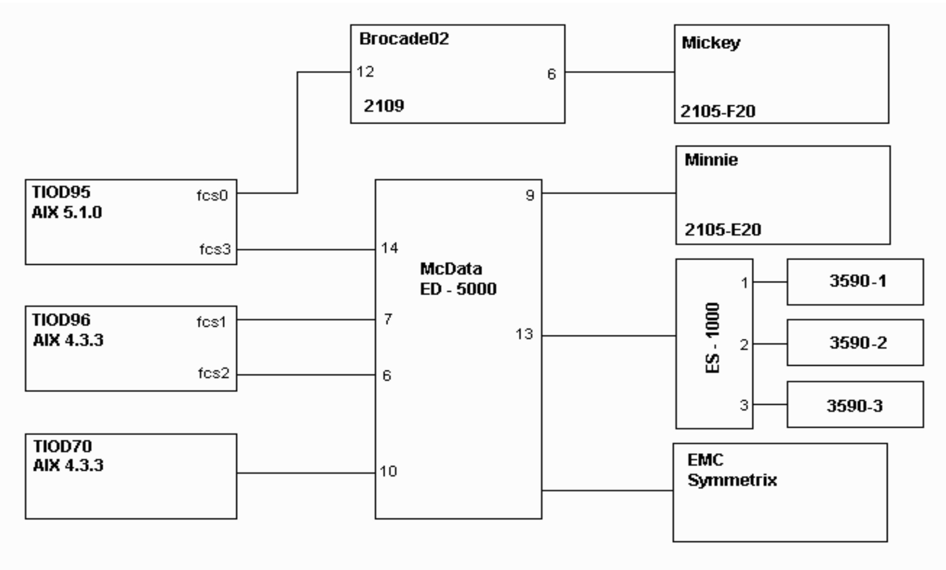


Figure 4-6 Example SAN diagram

Table 4-1 List of code levels and WWN by device name

Device name	Code level	WWN
tiod95	AIX 5.1.0	10:00:00:00:c9:27:38:72 (fcs0) 10:00:00:00:c9:23:ed:73 (fcs3)
tiod96	AIX 4.3.3	10:00:00:00:c9:26:58:12 (fcs0) 10:00:00:00:c9:24:0d:8b (fcs2)
tiod70	AIX 4.3.3	10:00:00:00:c9:26:a6:28
brocade02 (2109-S16)	2.4.1e	10:00:00:60:69:10:74:9c
McDATA ED-5000	3.0.1	10:00:08:00:88:60:6a:e5
McDATA ES-1000	1.02.0	10:00:08:00:88:60:40:e8
Mickey (2105-F20)	1.4.0.237	10:00:00:00:c9:21:23:f5
Minnie (2105-E20)	1.4.0.237	10:00:00:00:c9:23:9e:a2
EMC Symmetrix		50:06:04:82:03:3e:e5:cf
3590 - 1 (3590-E11)	1.15.9.7	50:05:07:63:00:40:3b:5e
3590 - 1 (3590-E11)	1.15.9.7	50:05:07:63:00:40:20:14
3590 - 1 (3590-E11)	1.15.9.7	50:05:07:63:00:40:36:02

The first map, Figure 4-7 on page 127, checks the physical connections from the server to a storage device. This map outlines three steps that are relatively easy to carry out with little effort and time. These first steps are nothing more than a visual inspection of the status LEDs that are now incorporated on most SAN devices. In situations with a constant problem, this initial step should quickly provide a clue as to whether the cause is due to a physical issue with the connections.

If no obvious connection problems are indicated by the LEDs on various devices, the next action in the problem determination process is as per Map 2 (Figure 4-8 on page 145). This troubleshooting map is based on the search for one or more issues in the logical connectivity between edge devices. In this map we proceed in a similar manner as in Map 1 by starting with further investigations of the pSeries server. Once started, the plan is to methodically progress from the server through the fabric to the storage resources until the problem is found.

Based on Map 2's outline, the next series of actions (Step 4) involve the verification of the basic operations within the pSeries server. This involves checking such items as:

- Correct operation of the HBA in the pSeries server
- Proper installation of device drivers for the HBA
- Proper installation of drivers for the logical SCSI I/O Controller protocol device.

Step 4 in Map 2 refers to the actions and explanations that are found in Section 4.4, "Checking the pSeries server" on page 132. If the problem is discovered during those actions, that section also contains some possible solutions. The possible solutions are based on some of the more common issues.

While most of the actions for implementing Step 4 are outlined and discussed in Section 4.4, "Checking the pSeries server" on page 132; Section 4.3.3, "Advanced problem isolation techniques" on page 129 provides additional methods that the reader can use. Please note that these additional actions are not neccessarily required to troubleshoot every problem. The reader should review the various actions and determine whether a given action will provide useful information.

If the problem is not found, and still exists, after Step 4, then the next step is to perform a basic health check of the SAN fabric. Step 5 of Map 2 is the entry point for these actions, and further explained is located in Section 4.5, "Checking the fabric" on page 147. Some of the basic items that are checked in Step 5 are:

- Confirming logical connections between all fabric devices
- Verifying basic connectivity with edge devices from the fabric

- Inspecting the fabric's zone configuration for conflicts or other errors

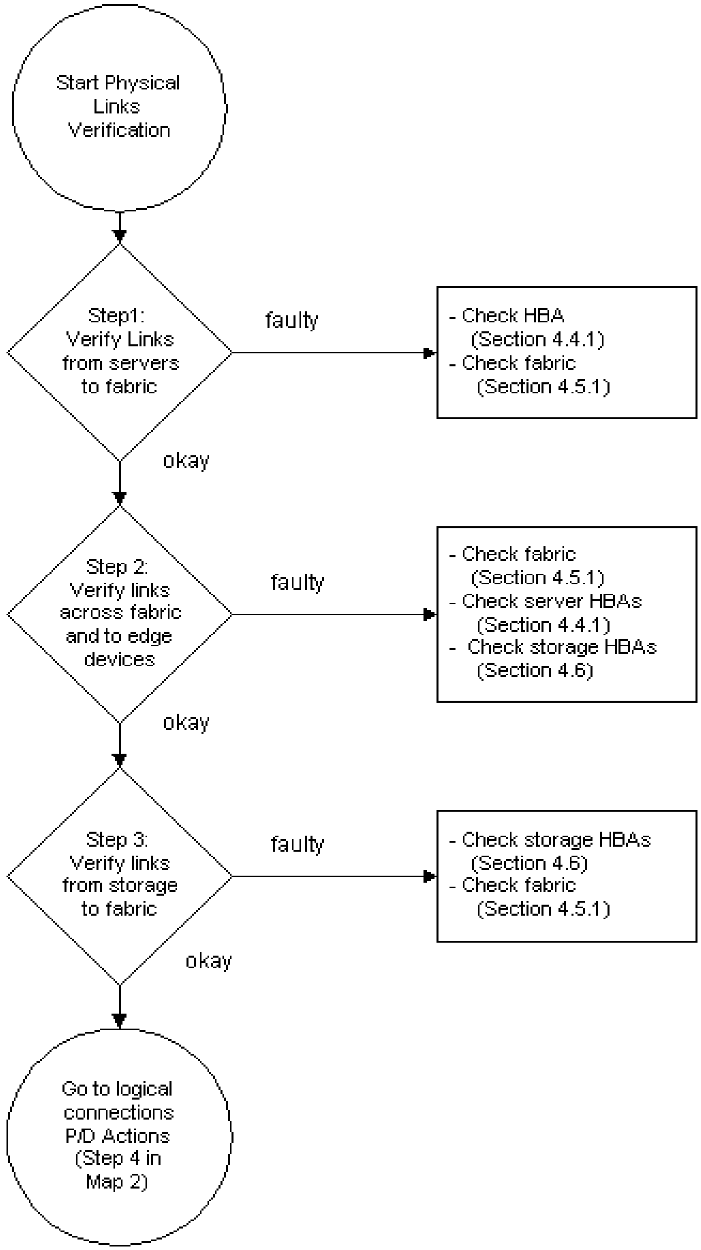


Figure 4-7 Map 1 - Troubleshooting physical links in a SAN

Outside of a given device's hard failure, a typical fabric-based problem involves consistency with the zone configuration information that is maintained is most fabric devices.

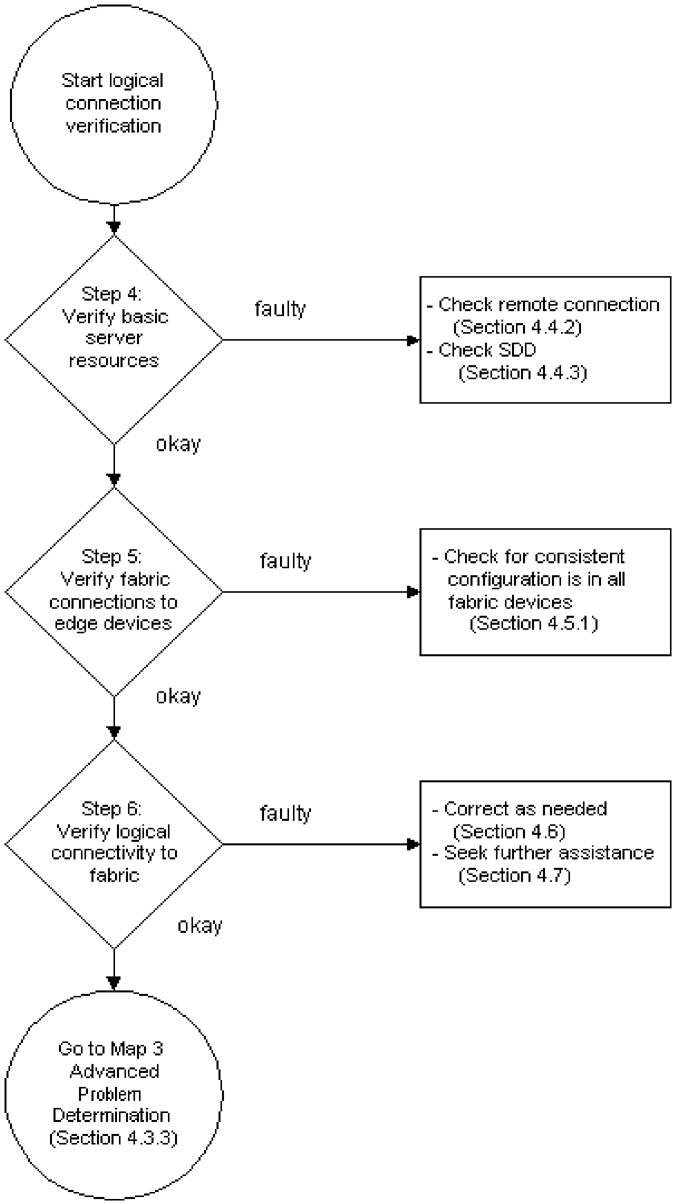


Figure 4-8 Map 2 - Troubleshooting logical links in a SAN

One method of identifying this type of issue is to manually inspect and compare the configuration of each device with all others. However, Section 4.3.3, “Advanced problem isolation techniques” on page 129 and Section 4.5, “Checking the fabric” on page 147, have additional details and enhanced troubleshooting actions for the fabric.

If the pSeries server and the SAN fabric are apparently working correctly, then the next action is Step 6. Step 6 covers troubleshooting storage devices. Further information and materials can be found in Section 4.3.3, “Advanced problem isolation techniques” on page 129, and Section 4.6, “Checking the storage systems” on page 191. Some of the basic actions for the inspection of the various storage units are:

- Verifying the logical connectivity to the SAN fabric
- Checking any masking implementation in the storage devices, such as LUN masking
- Inspecting the resource allocation to specific servers within the storage unit

The actions described in Section 4.6, “Checking the storage systems” on page 191 should be completed before implementing any of the advanced techniques.

4.3.3 Advanced problem isolation techniques

This section describes a number of advanced troubleshooting actions that can be implemented by the reader. We present these items for readers that have a knowledgeable and well-founded background working with SAN environments. A number of the actions described in this section are techniques that are typically utilized by remote technical support groups. Some of these actions can be disruptive to a SAN, and they are identified. These actions are not used in all problem situations, and the reader is cautioned.

This section will occasionally refer to some materials that will also be covered in the following sections. This advanced troubleshooting section is presented here as a natural progression of increasing complexity, with references and examples to follow.

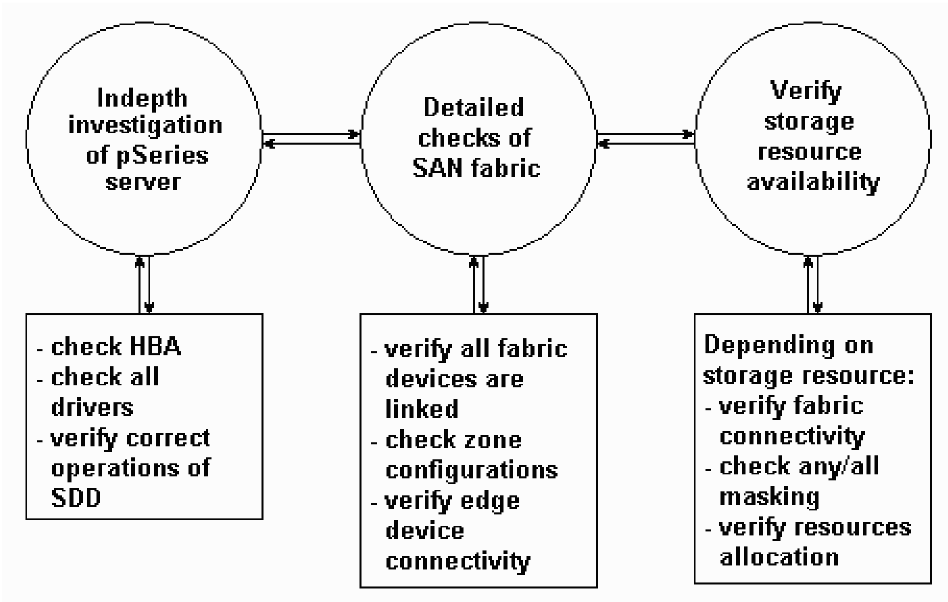


Figure 4-9 Map 3 - Advanced troubleshooting

As shown in Map 3, we do not have a set starting point for continuing the problem determination process. Also, the listing of actions under each SAN area (server, fabric, and storage) should already have been checked before considering any of the methods provided in this section. We will revisit the list of possible advanced actions with further details about each.

- Reboot the pSeries server.

  One of the most basic operations for AIX systems, this operation can resolve a number of issues. However, a server reboot does impact the SAN environment. At the same time, the reboot operation will cause the AIX operating system to reload all associated device drivers after a consistency check of the filesets, and cause the configuration manager application to be run again.

- Reload the device driver/component in question.

  If a good first rule of thumb is to reboot the server, it is closely followed by a second rule of thumb that is reloading the code component in question. This action is particularly true in new installation situations. An example of this scenario occurred during the preparations for this publication. In our case, a server had its device drivers reloaded, but only showed a subset of storage resources in a defined state. After extensive troubleshooting, we deleted all of the defined resources, reloaded the device drivers, and then ran the configuration manager application again. The problem was then resolved.

- Delete adapter instance plus any associated logical resources and rerun the configuration manager application, **cfgmgr**.

  This action is a variation of the reload device component. The intent of this action is to reinitialize the AIX operating system's knowledge of the storage resources available to it. This situation usually occurs during installation of new devices in the SAN environment. Needless to say, this action is disruptive to the server, and can impact a heavily loaded SAN when the **cfgmgr** application is run on a server with many resources available to it.

- Run a SDD trace.

  The Subsystem Device Driver supports AIX trace functions. The trace ID for the SDD is 2F8. The trace tracks routine entry, exit, and error paths of the SDD algorithm. To use the trace, manually turn on the trace function before a problem is recreated, then turn off the trace function either after the problem is seen or whenever the trace report is needed. To start the trace function, issue the **trace -a -j 2F8** command. To stop the trace function, use the **trcstop** command. To then read the report, use the **trcrpt | pg** command.

  > **Note:** To perform the AIX trace function, the bos.sysmgt.trace fileset must be installed on the AIX system.

  When the AIX trace is running, the SDD logs error conditions into the AIX errlog system. To check if the SDD generated an error log message, use the **errpt -a | grep VPATH** command. Refer to the *IBM Subsystem Device Driver Installation and User's Guide* softcopy documentation for more information about the various error log messages and their explanations.

- Generate and interpret the error log from the pSeries servers.

  A number of applications and the AIX operating system make entries into the AIX error log. This error log can contain very useful information if a problem occurs on a regular basis. Not only does a problem entry in the log provide data about the type of problem, but valuable information about the source of the problem can be gathered from the error log. This action is not disruptive to the SAN environment.

- Capture and interpret the error log from fabric devices.

  Practically all fabric devices have some mechanism to detect and record various types of events and errors. A review of this information in conjunction with data from other sources can be a valuable resource identifying the problem source. For more information on obtaining these logs from different switches, refer to Section 4.5, "Checking the fabric" on page 147.

- Run **`tapeutil`** to obtain errors reported by the drive to the host.

  Check for any messages that may have been reported back to the AIX operating system by the tape drive unit. To gather this information, run the **`tapeutil`** program. From the menu, select option 9, Error Log Analysis. For further information understanding messages, refer to Chapter 4, "Messages," of *IBM 3590 Tape Subsystem Maintenance Information*, SA37-0301.

- Run the Fibre Channel trace tool.

  The use of Fibre Channel trace tools in a SAN environment is usually implemented in very intermittent problem situations. Since the trace tool must be inserted into a link, this action is disruptive to a SAN environment. The severity of the disruption varies depending on the link to be traced. Because of the impact on a SAN, trace tools are typically used as a last resort method of problem determination.

4.4 Checking the pSeries server

This section provides information for troubleshooting a pSeries server. It presents a methodology from simple, physical inspections to detailed checks of the internal communications path from the AIX kernel to the HBA.

4.4.1 HBA verification

The first check of the HBA should be verifying that a physical connection exists between the HBA's port and another SAN device, usually a fabric port. This step is completed by visual inspection of the HBA port status LED, as in Figure 4-10 on page 133. Table 4-2 on page 133 lists the LED patterns during normal operations. If the port status is not in a normal condition, refer to the *IBM Fibre Channel Integration & Planning: User's Guide and Service Information,* SC23-4329, for additional information about the port status.

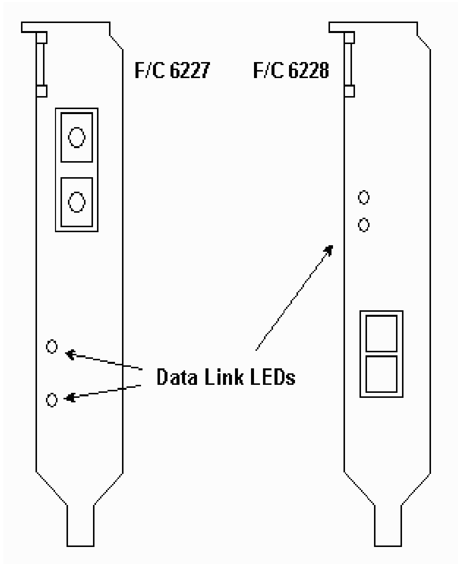


Figure 4-10 Gigabit Fibre Channel Adapter with data link LEDs

Table 4-2 LED - Normal status indicators

Green LED	Yellow LED	Port state
Slow blink (1 Hz)	Off	Normal - link down or not yet started
On	Slow blink (1 Hz)	Normal - inactive
On	Flashing (irregularly)	Normal - active
On	Fast blink (4 Hz)	Normal - busy

If the LEDs do not indicate a normal physical connection, the problem may be associated with the fiber cable or port. Possible causes of physical connectivity problems with the HBA include:

- Loose connection of fiber cable into the HBA's port

  Make sure that the fiber cable is securely plugged into the HBA port. If not, the fiber cable can be misaligned in the connector. This misalignment can significantly reduce the signal to the point that no light is detected.

- Transmit and receive lines of end ports are not correctly connected

  The transmitter at one end of a fiber cable run should be connected to the receiver at the other end. Refer to Figure 4-11 for proper orientation. A quick method for testing for this condition is by swapping the individual connectors at one end of the fiber cable. If a physical connection is not established, the problem has a different cause. Another method is to use a loopback connector on the suspect fiber cable and port.

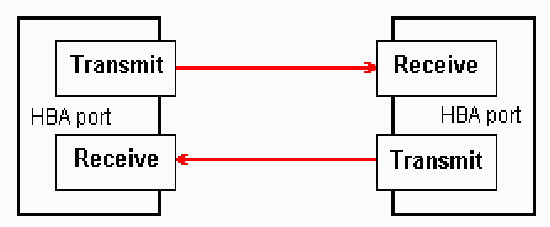


Figure 4-11 Orientation of transmit and receive connections

- High signal loss between end ports of a connection

  This situation can be prevalent in SAN environments with patch panels, or where the fiber cable is physically damaged. With some devices, use loopback tool to wrap the port on itself. This test should cause a change in the port status indicators if there is a faulty fiber cable.

- HBA port is faulty

  A faulty HBA port can be difficult to test. There are advanced diagnostics that can be implemented, but these test routines are usually disruptive to the SAN. A much easier test is to use a fiber cable from another established connection. The test simply consists of swapping the good fiber connection with the suspect fiber connection. If the suspect HBA port's status indicators do not change, refer to the service documentation for the HBA for further troubleshooting information for a possibly defective HBA port.

When the HBA port status indicators show a normal connection, the problem determination process starts with Step 1 in Map 4 (refer to Figure 4-12 on page 136). This first step determines whether the HBA is available to the AIX system. To determine an adapter's status, use the **`lsdev -C | grep fcs`** command (refer to Example 4-1 on page 135).

Example 4-1 Example listing of adapter status

```
# lsdev -C | grep fcs
fcs0      Available 10-70        FC Adapter
fcs1      Available 10-78        FC Adapter
fcs2      Available 20-58        FC Adapter
```

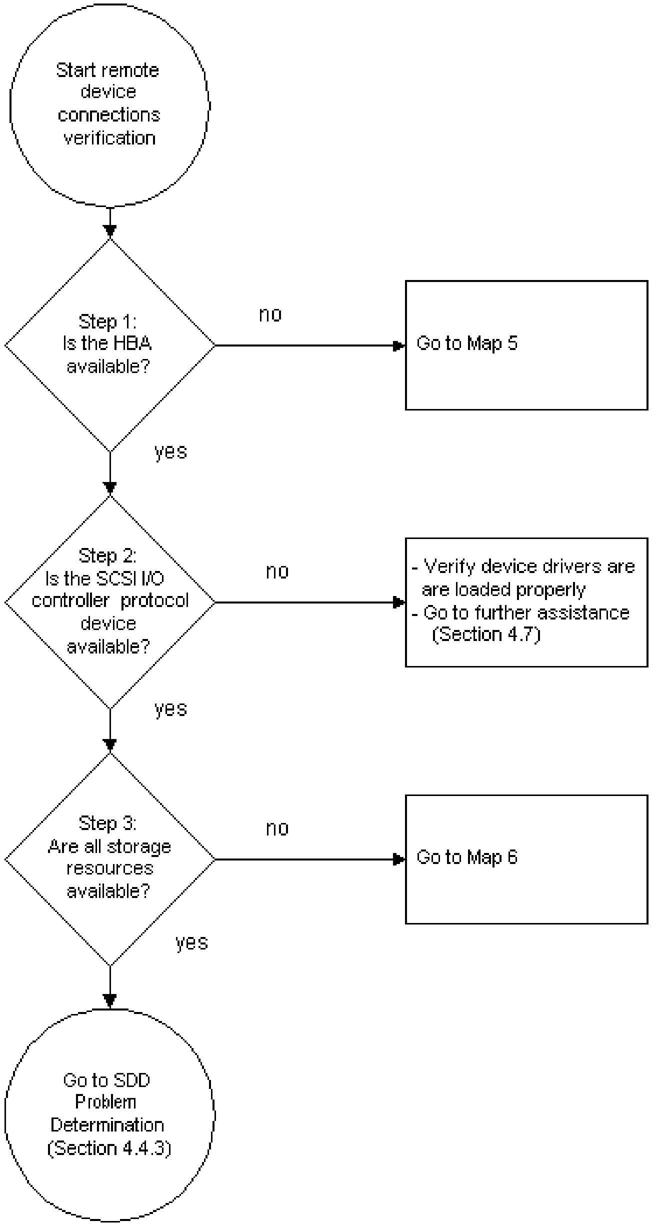


Figure 4-12 Map 4 - Advanced troubleshooting (continued)

If the adapter is available, then the various logical device and protocol drivers associated with the adapter needs further testing. Step 2 proceeds with these further tests. If the adapter is not available, then proceed with Step 4 in Map 5 (refer to Figure 4-13 on page 139).

The adapter in question is available to the AIX system. Step 2 determines whether the SCSI I/O Controller Protocol Device is available with the adapter. To check on the status, use the **lsdev -C | grep fscsi** command, shown in Example 4-2.

Example 4-2 Example listing of SCSI I/O Controller Protocol Devices

```
# lsdev -C | grep fscsi
fscsi0     Available 10-70-01      FC SCSI I/O Controller Protocol Device
fscsi1     Available 20-58-01      FC SCSI I/O Controller Protocol Device
fscsi2     Available 10-78-01      FC SCSI I/O Controller Protocol Device
```

If the SCSI I/O Controller Protocol Device is not available, then a problem may be that the device drivers for the Fibre Channel adapter are not properly installed. Depending on the type of adapter installed, run one of the following commands:

```
lslpp -l | grep df1000f7
lslpp -l | grep df1000f9
```

The df1000f7 fileset is for the adapter feature code 6227 while the df1000f9 fileset is for the 64-bit adapter feature code 6228. The result of the command should be similar to the example output shown in Example 4-3.

Example 4-3 Example of fileset for feature code 6227 HBA

```
# lslpp -l | grep df1000f7
  devices.pci.df1000f7.com  4.3.3.75  APPLIED    Common PCI FC Adapter Device
  devices.pci.df1000f7.diag
  devices.pci.df1000f7.rte  4.3.3.75  APPLIED    PCI FC Adapter Device Software
  devices.pci.df1000f7.com   4.3.3.0  COMMITTED  Common PCI FC Adapter Device
  devices.pci.df1000f7.rte   4.3.3.0  COMMITTED  PCI FC Adapter Device Software
```

If some of the above components are missing then the device drivers are not properly installed. To correct this problem refer to the *IBM Fibre Channel Integration & Planning: User’s Guide and Service Information*, SC23-4329, to reinstall the drivers. If all of the components are present, refer to Section 4.3.3, “Advanced problem isolation techniques” on page 129, or to Section 4.7, “Getting further assistance” on page 202.

If all of the components are present, then proceed to Step 3 of Map 4. Step 3 determines which, if any, storage resources are not available to the AIX system. At this point the reader may already have the answer to the question of which resources are available from earlier steps in defining the problem. If not known, then Map 6 (see Figure 4-14 on page 142) provides a method for determining the status of external storage resource availability.

Map 5 (see Figure 4-13 on page 139) outlines the methodology for troubleshooting issues where the Fibre Channel adapter is not available. The first action on this map, Step 4, is based on the assumption that the adapter is installed in the pSeries server and is known to the AIX system. If known, the adapter has a status of `defined`.

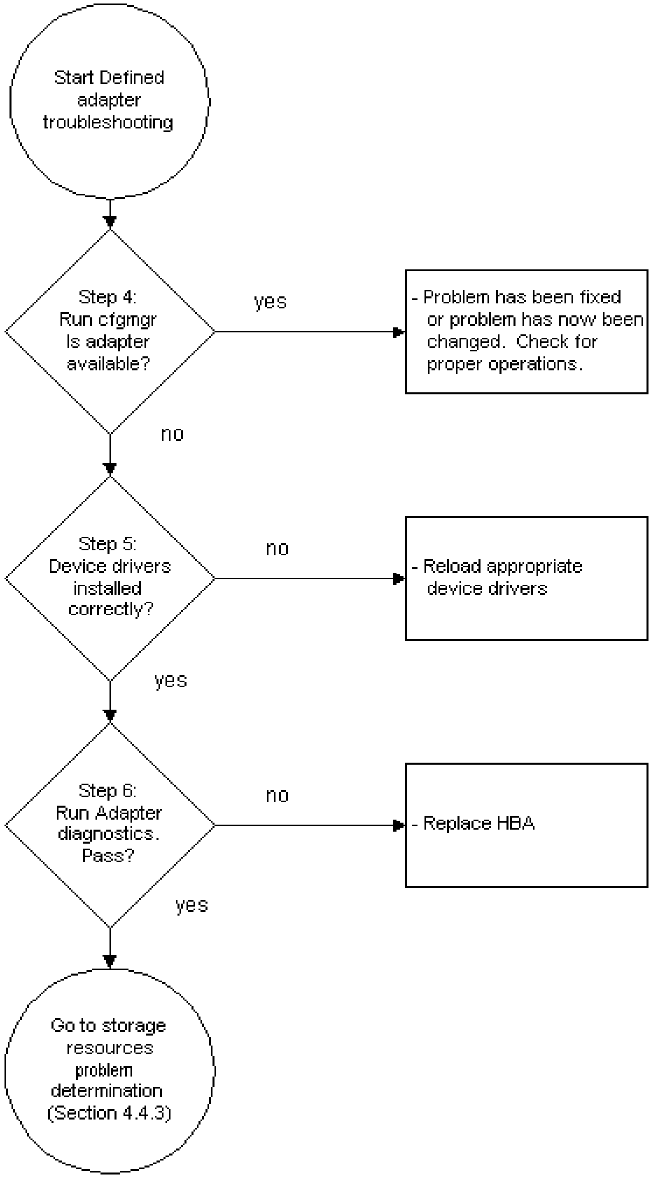


Figure 4-13 Map 5 - Advanced troubleshooting (continued)

In the defined state, attempt to configure the adapter by running the configuration manager application, **cfgmgr**. Depending on the number and type of storage resources, this command can take several minutes to complete.

After the **cfgmgr** command has been executed, check the adapter status again with the command:

lsdev -C | grep fcs

If the adapter is now in the available state, the problem has probably been resolved. Otherwise, the device drivers for the Fibre Channel adapter need to be checked. Depending on the type of adapter installed, run on of the following commands:

```
lslpp -l | grep df1000f7
lslpp -l | grep df1000f9
```

The df1000f7 fileset is for the adapter feature code 6227, while the df1000f9 fileset is for the 64-bit adapter feature code 6228. The result of the command should be similar to Example 4-4.

Example 4-4 Example of lslpp -l | grep df1000f7 command

```
# lslpp -l | grep df1000f7
  devices.pci.df1000f7.com  4.3.3.75  APPLIED    Common PCI FC Adapter Device
  devices.pci.df1000f7.diag
  devices.pci.df1000f7.rte  4.3.3.75  APPLIED    PCI FC Adapter Device Software
  devices.pci.df1000f7.com   4.3.3.0  COMMITTED  Common PCI FC Adapter Device
  devices.pci.df1000f7.rte   4.3.3.0  COMMITTED  PCI FC Adapter Device Software
```

If some of the above components are missing, then the device drivers are not properly installed. To correct this problem, refer to the *IBM Fibre Channel Integration & Planning: User's Guide and Service Information*, SC23-4329, to reinstall the drivers. If all of the components are present, then diagnostics should be run on the Fibre Channel adapter. For more information on running diagnostics, refer to the *IBM Fibre Channel Integration & Planning: User's Guide and Service Information*, SC23-4329. The adapter diagnostic routines will be disruptive to the SAN environment. If the adapter diagnostics fail, the HBA may need to be replaced. To verify this condition, we suggest using your normal technical support process for confirmation.

Important: If the HBA is replaced, there are a number of items within the SAN that will need to be checked and possibly modified. If soft zoning is used in the fabric, the zone configuration will need to be updated with the WWN of the new HBA. LUN masking on storage devices may also need similar updating. Finally, update the records of the SAN documentation of the changed values.

Another possible action at this stage is to seek remote technical support. For more information about this action, refer to Section 4.7, "Getting further assistance" on page 217.

4.4.2 Remote device connections verification

At this point, we are reasonably confident that no problems exist with the HBA, its associated drivers, and the SCSI I/O protocol devices. The next series of actions further refine the scope and nature of the problem. This is accomplished by a systematic check of all storage resources that are supposed to be available to the pSeries server, as shown in Figure 4-14 on page 142, Map 6. To be able to complete this confirmation, however, there are several assumptions that we make:

- All storage resources that should be available are known.
- The problem is not due to the addition or modification of storage resources.
- If a redundant path application (such as a SDD) is installed, the application was properly installed and previously working correctly.

To be able to determine if resources are missing, we must have a good idea of what is supposed to be available. If changes have been made due to the addition or modification of resource allocations, then it is evident that a configuration error has been introduced. In this case, troubleshooting can proceed directly to the area where the modifications were introduced. For example, a change in LUN masking in the IBM 2105 ESS will quickly be noticed and the masking changes carefully reviewed.

If a redundant pathway application, such as a SDD, is utilized and having problems, then the methodology and troubleshooting flows will naturally progress to checking that area of operations. As we have discussed earlier, the scope of the problem is determined in a methodical manner, and any lack of resources is a crucial data item to have. Even if our assumptions are not valid, using the presented methods will eventually progress to the point of checking the items that are the base of our assumptions.

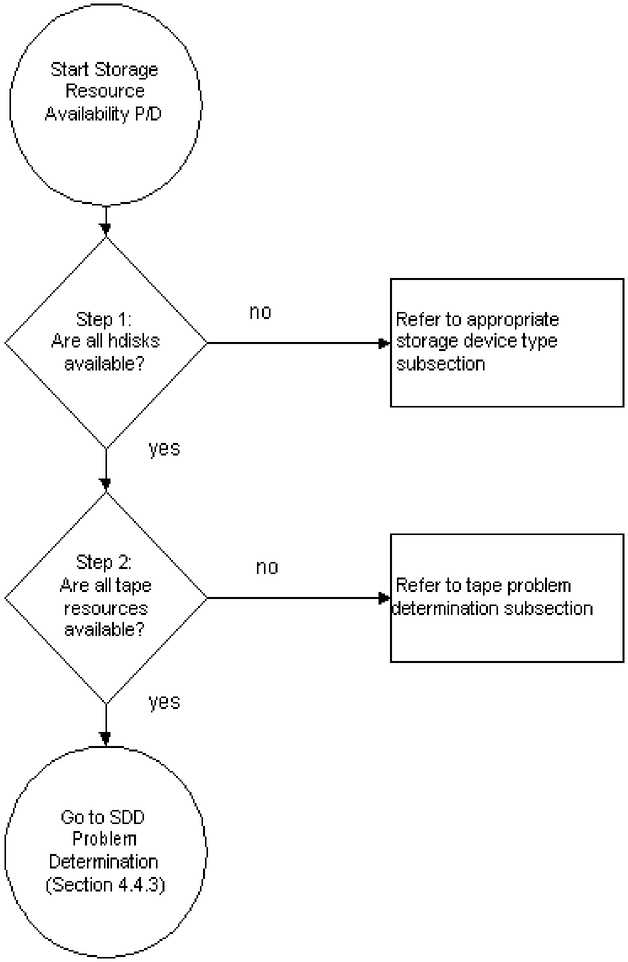


Figure 4-14 Map 6 - Advanced troubleshooting (continued)

Step 1 for checking remote device connections uses the command:

```
lsdev -C | grep hdisk
```

This command lists the logical hard disks (hdisks) that are associated with any Fibre Channel adapters in the pSeries server. If there are no disk devices in the server configuration, skip this step and proceed to Step 2. Otherwise, the result should be similar to Example 4-5. The key item to monitor is the status of the hdisk. A known hdisk has an Available status, while problems with a hdisk results in the Defined status. Care must be taken to not confuse disk subsystems that are internal to the pSeries server. These devices have the LUN assignment from the SCSI controller in the location column.

Example 4-5 Sample of available hdisks

```
# lsdev -C | grep hdisk | pg
hdisk0     Available 10-88-00-8,0  16 Bit LVD SCSI Disk Drive
hdisk1     Available 10-78-01      IBM FC 2105E20
hdisk2     Available 10-78-01      IBM FC 2105E20
hdisk3     Available 10-78-01      IBM FC 2105E20
hdisk4     Available 10-78-01      IBM FC 2105E20
hdisk5     Available 10-78-01      IBM FC 2105E20
hdisk6     Available 10-78-01      IBM FC 2105E20
hdisk236   Available 10-78-01      EMC Symmetrix FCP Raid1
hdisk237   Available 10-78-01      EMC Symmetrix FCP Raid1
```

If any disk resources are missing, or listed with a Defined status, this information can be a clue about the problem and its scope. In this situation, the next action is to determine whether the missing disk(s) reside in the same physical storage device. For any missing disks, the reader should consult the SAN documentation to determine whether the disks all reside in the same physical disk storage unit. Situations involving a missing subset of disks in a single device would indicate a configuration issue within the server or storage device itself. Thus, the scope of the problem has been localized to just two prime candidates for further investigation. In situations involving more than one physical disk subsystem with missing resources, the server requires additional troubleshooting in its various configurations. If all logical disks from a given storage subsystem are not available, then the server or fabric is the possible cause.

After checking for available hdisks to the pSeries server, we proceed to Step 2 of Map 6. If tape systems have never been assigned to the pSeries server, this step is omitted. To check installed tape devices, issue the **lsdev - Cc tape** command. Check the listing of devices to be sure that the tape units supposedly connected to the server do appear in the resulting output of the command (refer to Example 4-6).

Example 4-6 Sample listing of available tape drive units

```
# lsdev -Cc tape
rmt0 Available 10-70-01 IBM 3590 Tape Drive and Medium Changer (FCP)
rmt1 Available 10-70-01 IBM 3590 Tape Drive and Medium Changer (FCP)
rmt2 Available 10-70-01 IBM 3590 Tape Drive and Medium Changer (FCP)
rmt3 Available 10-78-01 IBM 3590 Tape Drive and Medium Changer (FCP)
rmt4 Available 10-78-01 IBM 3590 Tape Drive and Medium Changer (FCP)
rmt5 Available 10-78-01 IBM 3590 Tape Drive and Medium Changer (FCP)
```

To check the WWN of the actual attached tape unit to what is expected, issue the **tapeutil -f /dev/rmt# qrypath** command (where rmt# designates the device in question). A valid response is data indicating the device is a 3590 with its WWN.

Now that the investigations of remote device connections have been performed, the reader should have a much better idea of where to next investigate. In situations that still indicate the server as a possible cause, the next logical move is to investigate the SDD.

One main concept of the troubleshooting process is that the more information that is gathered about a problem, the better the problem and its cause are understood. With this improved understanding, the reader can start making more efficient decisions about which portions of the problem determination process can be quickly verified. At the same time, the more detailed locations for the problem will have detailed examinations. The reader should understand that we have provided a basic blueprint of sorts that can and should be flexible as more data is collected about the decision.

4.4.3 SDD verification

The SDD is an application that is included with every IBM 2105 system. It is not required that it is installed on a host server for proper I/O operations between the pSeries server and the ESS. However, the ESS must have LUN sharing for multiple paths into the ESS for each pSeries server running the SDD. For the latest code, updates, and technical documentation, refer to the following Web site:

sdddom01.storage.ibm.com/techsup/swtechsup.nsf/support/sddupdates

In this section we describe actions and a methodology for troubleshooting the SDD based on the assumption that the SDD has previously been working on a now problematic system.

There are a number of actions that should be taken when troubleshooting the SDD. These actions are:

- Verify the installation of the SDD.

  You can verify that the SDD has been successfully installed by using the **lslpp -l ibmSdd*** command. If the filesets have been installed properly, the output from the command should look like this:

```
# lslpp -l ibmSdd*
  Fileset                      Level  State      Description
----------------------------------------------------------------------------
Path: /usr/lib/objrepos
  ibmSdd_433.rte             1.3.0.1  COMMITTED  IBM Subsystem Device
                                                 Driver
                                                 AIX V433
                                                 for non-concurrent HACMP
```

- Verify the SDD configuration.

  To check the SDD configuration, use the SMIT Display Device Configuration panel. Perform the following steps on the pSeries server command line prompt:

  a. Type `smitty device` at the command prompt. The Devices menu is displayed.

  b. Highlight `Data Path Devices` and press Enter to display the Data Path Device panel.

  c. Highlight `Display Data Path Device Configuration` and press Enter. A list is displayed on the condition (either Defined or Available) of all SDD pseudo devices, as shown in Example 4-7. If any device is listed as Defined, the configuration is not correct. Refer to the *IBM Subsystem Device Driver Installation and User's Guide* documentation for more information about the configuration procedures.

Example 4-7 Verify condition of SDD pseudo devices

```
COMMAND STATUS

Command: OK            stdout: yes           stderr: no

Before command completion, additional instructions may appear below.

[TOP]
vpath0 (Avail ) 10012073 = hdisk1 (Avail pv )
vpath1 (Avail pv datavg) 10112073 = hdisk2 (Avail )
vpath2 (Avail pv datavg) 10212073 = hdisk3 (Avail )
vpath3 (Avail ) 10312073 = hdisk4 (Avail pv )
vpath4 (Avail ) 10512073 = hdisk5 (Avail pv )
vpath5 (Avail ) 10612073 = hdisk6 (Avail pv )
vpath6 (Avail ) 10712073 = hdisk7 (Avail pv )
vpath7 (Avail ) 10812073 = hdisk8 (Avail )
vpath8 (Avail ) 10912073 = hdisk9 (Avail )
vpath9 (Avail ) 10A12073 = hdisk10 (Avail )
vpath10 (Avail ) 10B12073 = hdisk11 (Avail )
vpath11 (Avail ) 10C12073 = hdisk12 (Avail )
vpath12 (Avail ) 10D12073 = hdisk13 (Avail )
[MORE...210]
```

- Verify that multiple paths are configured for each adapter connected to an ESS port.

  To check on multiple paths for each adapter, use the SMIT Display Device Configuration panel.

Perform the following steps on the pSeries server command line prompt:

a. Type `smitty device` at the command prompt. The Devices menu is displayed.

b. Highlight `Data Path Devices` and press Enter to display the Data Path Device panel.

c. Highlight `Display Data Path Device Adapter Status` and press Enter. All attached paths for each adapter are displayed, as shown in Example 4-8.

Example 4-8 Check on multiple paths for an adapter

```
COMMAND STATUS

Command: OK            stdout: yes           stderr: no

Before command completion, additional instructions may appear below.

Active Adapters :1

Adpt#     Adapter Name   State     Mode     Select     Errors  Paths  Active
    0           fscsi2  NORMAL   ACTIVE          0          0    222       0
```

- Verify the association of frame check sequence (FCS) adapters that are attached to SDD pseudo devices.

  To get a complete listing of the associations and some performance data, use the **`datapath query adaptstats`** command. This command will list information on a per-adapter basis, as shown in Example 4-9. The adapters are identified by their relative position within the pSeries server. Refer to the *IBM Subsystem Device Driver Installation and User's Guide* for more information about the data that is provided.

Example 4-9 Output of datapath command

```
# datapath query adaptstats

Adapter #:  0
=============
                Total Read  Total Write  Active Read  Active Write   Maximum
I/O:                     0            0            0             0         0
SECTOR:                  0            0            0             0         0
```

If an association discrepancy between the SDD pseudo devices and the FCS adapters is found, then verify the SDD configuration again. Also, verify that the status of each PATH is in the OPEN state.

Note: The device number and the device index number are usually the same. However, the devices can be configured out of order, thus, the numbers are not always consistent. To correlate the index number for a specific device, issue the `datapath query device` command. At the same time as when checking the correlations, also verify that the each adapter is in the NORMAL state and ACTIVE mode.

4.5 Checking the fabric

This section provides guidance and a suggested methodology for troubleshooting a SAN fabric. We will first describe a non-product specific flow for troubleshooting. This generic fabric problem determination process is then augmented with product-specific information and commands necessary to complete the process. Each of the product specific sections starts with a description of the port status indicators and normal indicator conditions are described.

Each product-specific section also describes the actions needed to complete the various steps outlined in the generic discussion. Finally, the product sections provide information about advanced testing actions and techniques that are unique for each switch.

Some of the troubleshooting steps only apply to multiple-switch environments. Thus, some of the described actions will not be applicable for all SAN designs.

4.5.1 Generic fabric checks

We first describe a methodology for troubleshooting SAN fabrics in general terms. During the descriptions of each step, we denote whether the actions of each step are applicable in single switch SANs, multi-switch environments, or both types of SAN fabrics.

Figure 4-15 on page 149 provides a view of the flow of the basic troubleshooting steps for the problem-determination process for the SAN fabric. Additional descriptions of the individual steps follow the map.

The basic intent of these procedures is to quickly verify the health of the SAN fabric. As with the previous troubleshooting processes, we determine a logical starting point and then work outwards from that starting point until either the problem source is discovered or the SAN fabric is determined to be operating properly.

The starting point for the SAN fabric in general is the links between the various devices that make up the fabric. Once the links are verified, the next step is to check on the connections from the fabric to the edge devices, such as servers and storage units. Finally, we explore the edge-to-edge logical connectivity by checking the zoning configurations.

In the product-specific sections we also provide the methods used for collecting additional information, which are used with the advanced problem-determination steps. Examples of this information include error logs, audit logs, and port performance data.

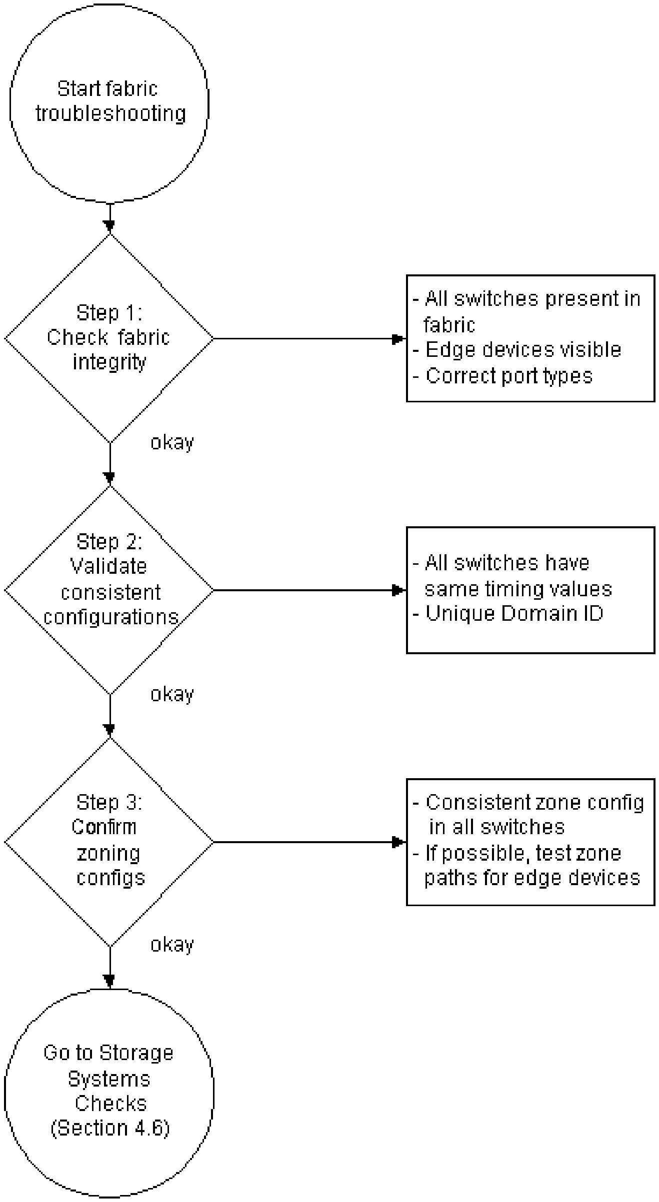


Figure 4-15 SAN fabric switch troubleshooting map

Step 1 of the SAN fabric troubleshooting map is a quick and simple integrity check for all devices that make up the SAN fabric. In the case where only a single switch is in the SAN fabric, this step can be skipped. Using the SAN diagram and the supporting information for fabric devices, verify that all of the fabric devices are known to each other. This action can be a simple command at a device prompt, or can use a graphical device management application. Either way, the devices in the fabric must know of each other's existence, and these links must be of the correct type of connection. When switches of the same type are interconnected, the end ports should be listed as E_ports (or T_Ports for INRANGE products), or expansion ports. Some switches may use different terminology, but the idea is that inter-switch links are not based on the same type of end port definitions as a port connecting an edge device.

Note: Fibre Channel standards are still being implemented at various rates of progress by the different Fibre Channel switch manufacturers. Thus, inter-switch connections between two switches of different makers may not have a sufficient level of interoperability to establish an fabric expansion link, or E_port. Refer to documentation from the various manufacturers to determine if their products are capable of establishing inter-switch links with each other's product.

If one or more fabric devices are missing from one location in the fabric, check other switches to see if the problem is consistent. If so, then areas for further checks are:

- Port status of connections to other fabric devices

  Check the port status of all connections to fabric devices. Their conditions should be operational and the proper type of connection (as in E-port for interswitch links). If the port is not operational, use the product's user's guide or service guide to determine the cause. Possible causes are cabling issues, configuration changes, or hardware failures. Most of these causes are usually found quickly.

- Error log of the switch

  Refer to the appropriate sections of the corresponding product manuals for information on gathering and making use of this information.

- Zoning configuration of the affected switches in comparison to other switches without the problem

  We describe methods to view zoning configuration information in general terms further on in this chapter.

After verification that all SAN fabric devices are know to each other, the second step is to confirm the logical connections from the fabric to the edge devices. Depending on the nature of the problem, some of these connections have already been checked from a subset of the edge devices. However, the communications links are two-way. Thus, one side may be talking while the other end of a link is not listening or able to understand the conversation. This is the reason for inspecting and verifying a link at the fabric end, even though the given link has already been checked from the edge device.

Step 2 of the fabric check map (Figure 4-15 on page 149) is a detailed examination of the links from some, or all, of the fabric devices to the attached edge devices. The main aspects of a link to investigate are that the link is established, and the link is defined as the proper type of connection. Thus, the link to an edge device should be connected to an F_port, or fabric port, at the switch's end. Hopefully, not all edge device connections will need to be checked. During initial rounds of the problem determination process, the reader has defined the scope of the problem. Only the ports with connections to edge devices identified during the initial understanding of the problem's scope need to be checked. It is a rare situation where all fabric ports to edge devices will have to be checked. However, the reader should remember that intermittent problem types or the many-to-many scenario may warrant such an extreme as testing all ports.

Step 3 of the fabric check map is a verification of the zoning configuration that is implemented in the SAN fabric. This step is based on the assumption that the various fabric port checks in Steps 1 and 2 showed that all connections are properly established. This action will verify the zoning configuration of one switch for proper connectivity through the SAN fabric to the effected edge devices. If there are multiple switches in the fabric, the confirmed zoning configuration of the first checked switch is then compared to the remaining switches. Depending on the type of switch used (IBM 2109, McDATA, or INRANGE), the comparison process may be quickly accomplished, or may take a significant amount of detailed checking.

4.5.2 IBM 2109 specific checks

When checking a switch, the first quick troubleshooting action is checking the status of ports with connections to other devices. These connections are for both edge devices and other fabric units. For the IBM 2109 switch, each port has a single LED. Refer to Figure 4-16, which shows the location of the port's LED. This port LED can be either green or amber.

Figure 4-16 Location of LED port indicators

Each port LED will be in one of following conditions:

- Constantly off
- Constantly on
- Slow blink - flashes every 2 seconds
- Fast blink - flashes twice every second
- Flicker - randomly flashes

Refer to Figure 4-3 for the meaning of the various LED colors and conditions.

Table 4-3 2109 port status indicators

LED state	Port status
No light	No light or signal carrier
Steady yellow	Light detected, but not online yet
Slow yellow	Disabled. Flashes every 2 seconds
Fast yellow	Error, fault with port. Flashes twice per second
Steady green	Online, connected to device over cable
Slow green	Online, but segmented (partitioned). Flashes every 2 seconds
Fast green	Internal diagnostic loopback. Flashes twice per second
Flickering green	Online with traffic

If one or more ports on the switch have visual indications of problems, then check the device's port condition at the other end of the connection. The nature of the problem may be better defined when the conditions at both ends of the link are known. If both ends indicate a problem, there may be a fault with the physical cabling and/or a logical connection problem. Possible causes of physical connectivity problems with the port include:

- Loose connection of fiber cable into the port

  Make sure that the fiber cable is securely plugged into the port. If not, the fiber cable can be misaligned in the connector. This misalignment can reduce the signal significantly to the point that no light is detected.

- Transmit and receive lines of end ports are not correctly connected

  The transmitter at one end of a fiber cable run should be connected to the receiver at the other end. Refer to Figure 4-11 on page 134 for proper orientation. A quick method to test for this condition is to swap the individual connectors at one end of the fiber cable. If a physical connection is not established, the problem has a different cause.

- High signal loss between end ports of a connection

  This situation can be prevalent in SAN environments with patch panels, or where the fiber cable is physically damaged. With some devices, use of a loop back tool to wrap the port on itself. This test should cause a change in the port status indicators if there is a faulty fiber cable.

- Fibre Channel port is faulty

  A faulty port can be difficult to test. There are advanced diagnostics that can be implemented, but these test routines are usually disruptive to the SAN. A much easier test is to use a fiber cable from another established connection. This test simply consists of swapping the good fiber connection with the suspect fiber connection. If the suspect port's status indicators do not change, the reader will need to proceed with further troubleshooting actions.

For problems that are not due to a physical cabling issue, the reader will need to continue with the troubleshooting methods previously outlined. The following material is presented to assist the reader with specific actions and methods for troubleshooting the IBM 2109 Fibre Channel switch.

The 2109 switch problem determination process can be accomplished with a command line interface from a Telnet session to the switch, or by means of GUI interface with the IBM StorWatch Specialist. In the following descriptions of troubleshooting actions we provide the steps needed for various problem-determination actions using both methods.

The steps for establishing a Telnet session with the new switch, and then logging in are:

1. From the command line prompt on a pSeries server, use the command:

   ```
   >telnet <IP address>
   ```

   If the IP address of the new switch was changed, use the new IP address. Otherwise, the switch uses `10.77.77.77` as the default IP address.

2. The switch prompts for a user name. The default administrator name of the IBM 2109 switch is `admin`.

3. The switch next prompts for the password. The default administrator password is `password`.

In the command line interface, use the **help** command to view a list of commands and their purpose. The IBM StorWatch Specialist view can be accessed through one of the following Java-enabled Web browsers:

For UNIX:

- Netscape 4.51 or higher

For Windows 95/98 or Windows NT:

- Internet Explorer 4.0 or higher
- Netscape 4.51 or higher

In addition to the above, Java Plug-In 1.2.2 or higher is also required. The current version of the Java Plug-In can be downloaded from the following URL:

`http://java.sun.com/products/`

To launch the IBM StorWatch Specialist:

1. Start the Web browser (if it is not already active).
2. Enter a switch name or an IP address in the Location/Address field.

Figure 4-17 shows the default view from the IBM 2109 Web interface.

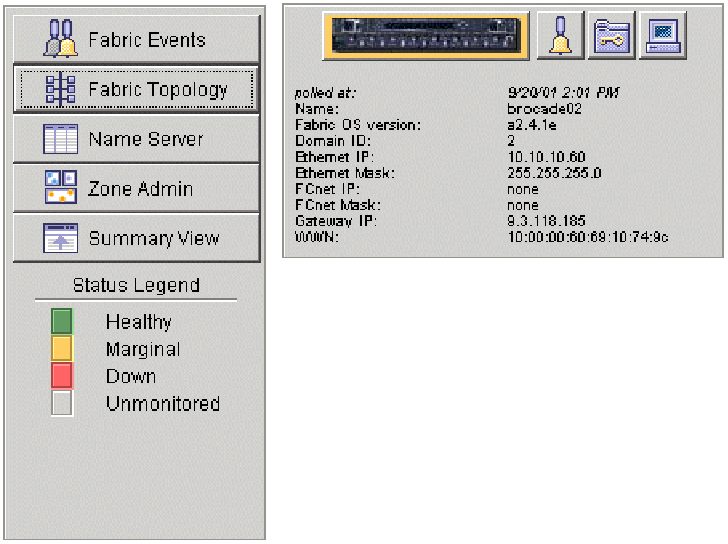


Figure 4-17 Default web interface view of IBM 2109

After visual inspection of the port LEDs, the next action is to check the port status via the Telnet session or the GUI interface. While the LED indicator may have shown an established connection, the type of connection (for example, E_Port) is not known. To see the status of all ports from the command line, use the **switchShow** command (see Example 4-10).

Example 4-10 Example output of SwitchShow command

```
brocade02:admin> switchShow
switchName:     brocade02
switchType:     2.4
switchState:    Online
switchRole:     Principal
switchDomain:   2
switchId:       fffc02
switchWwn:      10:00:00:60:69:10:74:9c
switchBeacon:   OFF
port  0: id  No_Light
port  1: --  No_Module
port  2: id  No_Light
port  3: id  No_Light
port  4: --  No_Module
port  5: --  No_Module
port  6: id  Online        F-Port  10:00:00:00:c9:21:23:f5
port  7: --  No_Module
port  8: --  No_Module
port  9: id  No_Light
port 10: --  No_Module
port 11: --  No_Module
port 12: id  Online        F-Port  10:00:00:00:c9:27:38:72
port 13: id  No_Light
port 14: --  No_Module
port 15: id  No_Light
```

From the sample output, check that:

- All ports with connections to other SAN devices have a status of On-line.
- All ports with connections to other fabric devices are present and are listed as E_Ports.
- All ports with connections to edge devices are present and are listed as F_Ports or FL_Ports. FL_Ports are for Fibre Channel arbitrated loop devices.

From the main GUI interface of the switch, the port information is displayed by clicking on the IBM 2109 icon, which displays a larger window of the IBM 2109. In this window, click on one of the ports to open the Web interface to the port status and statistics view. Figure 4-18 is an example of the GUI interface showing the status of the ports.

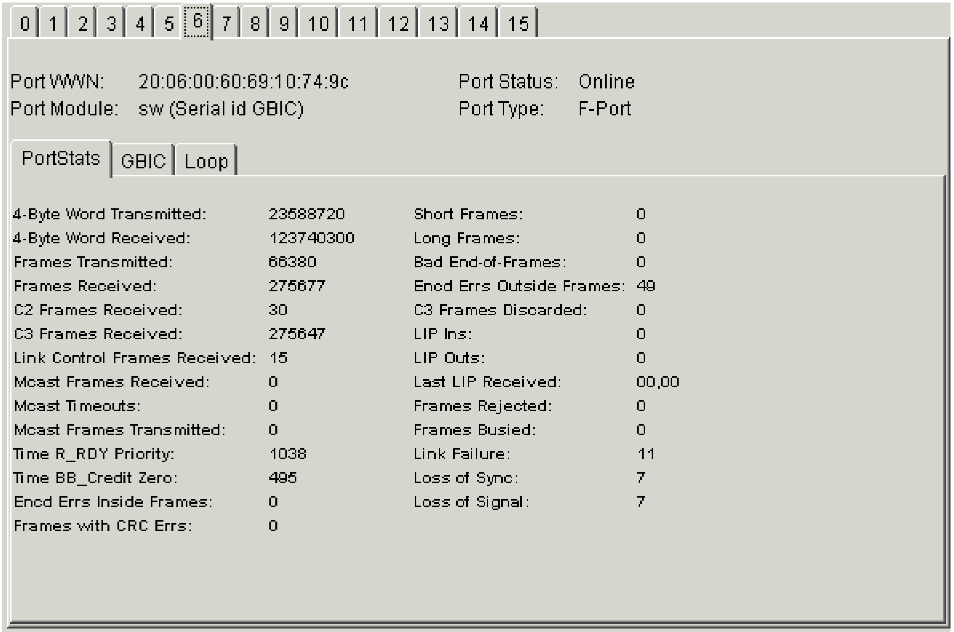


Figure 4-18 Status of ports from GUI interface of IBM 2109

If one or more ports are part of a link associated with a problem, the IBM 2109 switch has the means to check on certain types of errors. A better idea about certain types of problems can be gained by a check of the error summary for all ports. From the command line, use the **portErrShow** command. Refer to the sample output shown in Example 4-11.

Example 4-11 Sample of portErrShow command

```
brocade02:admin> porterrshow
       frames  enc  crc  too  too  bad  enc disc link loss loss frjt fbsy
      tx   rx   in  err shrt long  eof  out   c3 fail sync  sig
    ---------------------------------------------------------------------
 0:    0    0    0    0    0    0    0    0    0    0    0    1    0    0
 1:    0    0    0    0    0    0    0    0    0    0    0    0    0    0
 2:    0    0    0    0    0    0    0    0    0    0    0    1    0    0
 3:    0    0    0    0    0    0    0    0    0    0    0    1    0    0
 4:    0    0    0    0    0    0    0    0    0    0    0    0    0    0
 5:    0    0    0    0    0    0    0    0    0    0    0    0    0    0
 6:  63k 208k    0    0    0    0    0   49    0   11    6    7    0    0
 7:    0    0    0    0    0    0    0    0    0    0    0    0    0    0
 8:    0    0    0    0    0    0    0    0    0    0    0    0    0    0
 9: 2.6m 4.2k    0    0    0    0    0   27    0    1    9    9    0    0
10:    0    0    0    0    0    0    0    0    0    0    0    0    0    0
11:    0    0    0    0    0    0    0    0    0    0    0    0    0    0
```

```
12:  2.7m  60k   0    0    0    0    0  105    0   20    9   12    0    0
13:     0    0   0    0    0    0    0    0    0    0    0    2    0    0
14:     0    0   0    0    0    0    0    0    0    0    0    0    0    0
15:     0    0   0    0    0    0    0    0    0    0    0    1    0    0
```

From the main GUI interface of the switch, the port information is displayed by clicking on the IBM 2109 icon, which displays a larger window of the IBM 2109. In this window, click on one of the ports to open the Web interface to the port status and statistics view.

Some of the informational items from this port error summary are:

- Frames sent and received
- Number of frames too big or too small
- Frames with encoding errors or CRC errors
- Loss of synchronization or signal
- Frames rejected (with F_RJT) or busied (with F_BSY)

There will most likely be some specific error types due to known causes. However, this information can be used to monitor the health of connections over a period of time. An encoding error, a Cyclical Redundancy Checking (CRC) error, loss of synchronization, or loss of signal can be results of cabling issues or a port failing. The main point is that this information can be reviewed periodically to see if the errors continue to increase over time. If they do, then this can be an indication of the problems in cases of randomly occurring issues.

At this point if the problem is not yet identified, there may still be some suspicion of the SAN fabric. Even though the ports with fabric interconnects were found to be operational and have the proper setting, there are other functions within the fabric that bear checking. The first of these is the fabric controller function. The quickest check of this function is to view the controller's listing of switches in the fabric. From the Telnet session, enter the **fabricShow** command to view the listing, as shown in Example 4-12.

Example 4-12 Example output of fabricShow command

```
brocade02:admin> fabricshow
Switch ID   World wide Name          Enet IP Addr    FC IP Addr      Name
-------------------------------------------------------------------------
1: fffc01 10:00:00:60:69:10:4e:78  10.10.10.60     0.0.0.0 >"brocade01"
2: fffc02 10:00:00:60:69:10:74:9c  10.10.10.60     0.0.0.0  "brocade02"
```

All of the switches that make up the fabric should appear in this list, including the switch running the command. From the main GUI interface of the switch, the fabric controller information is displayed by clicking on the Fabric Events icon in the upper left region of the screen. Refer to Figure 4-17 on page 154.

For similar reasons as with the controller function, the fabric's Name Server function needs to be checked. From the Telnet session's command line, there are two commands that are available. The difference is the scope of the listing. If the **nsAllShow** command is used, the resulting output display is the 24-bit Fibre Channel port IDs of all of the devices in all of the switches in the fabric. It does not list any WWN information. Refer to Example 4-13 for an example of this command.

Example 4-13 Example output of nsAllShow command

```
brocade02:admin> nsAllShow
2 Nx_Ports in the Fabric {
  011200 0118e8 021900 021c00
}
```

For the output from the local switch, use the **nsShow** command. The advantage of the **nsShow** command is the inclusion of WWN information of attached devices in the display. Example 4-14 displays the output of this command.

Example 4-14 Example output of nsShow command

```
brocade02:admin> nsShow
The Local Name Server has 2 entries {
 Type Pid    COS     PortName NodeName          TTL(sec)
 N    021600;    2,3;10:00:00:00:c9:21:23:f5;50:05:07:63:00:c0:02:ed; na
    FC4s: FCP [IBM     2105F20         .191]
    Fabric Port Name: 20:06:00:60:69:10:74:9c
 N    021c00;    2,3;10:00:00:00:c9:27:38:72;20:00:00:00:c9:27:38:72; na
    Fabric Port Name: 20:0c:00:60:69:10:74:9c
```

When attempting to trace possible zoning configuration errors in a SAN, having simple, concise information about the current configuration is extremely helpful. This is especially handy when having to compare the configuration across multiple switches in the SAN fabric. Example 4-15 on page 159 displays the output of this command.

Example 4-15 Sample output of cfgShow

```
brocade02:admin> cfgShow
Defined configuration:
 cfg:   Prod_zone
                iod95_to_shark; nowhere_togo;
 cfg:   Just4show
                nowhere_togo
 zone:  iod95_to_shark
                mickey; tiod95
 zone:  nowhere_togo
                nowhere; places
 alias: nowhere 2,3
 alias: places  2,9
 alias: tiod95  2,12
 alias: Mickey  2,6

Effective configuration:
 cfg:   Prod_zone
 zone:  iod95_to_shark
                2,6
                2,12
 zone:  nowhere_togo
                2,3
                2,9
```

The 2109 switch provides this information with the **cfgShow** command. In multiple-switch SAN fabrics using the IBM 2109 switch, extreme care must be taken when comparing the zone configuration information of one switch to another. All items must compare equally, such as the alias names, zone names, zoned ports, and active configuration. If there are any conflicts with the zoning configurations between fabric switches, the fabric will become segmented into two or more separate fabrics. This check is crucial and can be painstaking with larger fabrics. If a conflict is found, determine which configuration is not correct and then make the appropriate changes. For additional information on making changes to a zone configuration, refer to *IBM SAN Fibre Channel Switch 2109 Model S16 User's Guide,* SC26-7351.

Tip: When making zoning configuration changes, attempt to make the changes during periods of low activity of the SAN. This is because multiple switch fabrics will have a burst of activity while the new zoning information is propagated across the SAN fabric. Once the changes are made, perform connectivity tests between the servers and their related storage resources. If there are a larger number of servers involved and any impacts are to be kept to a minimum, only check the servers with the most critical applications. If this quick check proves okay, remember to issue the **cfgSave** command. This command will store the new zoning configuration into flash memory (as in permanent memory) so that it is not lost in the event of a power failure or system reset.

With these checks completed, it is a reasonable assumption to say that the fabric is not the problem source in cases of consistent problems. However, for the random type of problem, the SAN fabric with IBM 2109 may still be an issue. There are still a number of items that can be checked, but these are considered to be advanced troubleshooting techniques. We will describe some of these methods for troubleshooting purposes, as well as how to have the information readily available if further assistance must sought from a technical support group. Some of the advanced methods are:

- Check the current firmware code level for known issues.

  There are times when certain problems are found to exist with specific levels of code that are corrected with later revisions. We describe a method for determining if this scenario is applicable to the current problem. First, the current code level of the IBM 2109 switch must be known. To determine the firmware code level from the command line, use the **version** command. Example 4-16 shows the output of this command.

Example 4-16 Output from the version command

```
brocade02:admin> version
Kernel:     5.3.1
Fabric OS:  a2.4.1e
Made on:    Fri Jul 20 10:29:34 PDT 2001
Flash:      Fri Jul 20 10:30:39 PDT 2001
BootProm:   Thu Jun 17 15:20:39 PDT 1999
```

  The code level is listed as the Fabric OS. To determine if there is a later code level than the switch is currently using, go to the Web site:

  http://www.storage.ibm.com/ibmsan/products/2109/index.html

Follow the links to the downloads Web page and determine whether the switch is running at the most current level of code. If not, download the code package (stored in a zip file format). Extract the various files from the download and look for a set of release notes for the newer level of code. If there is a listing for the problem that is being experienced, then install the new code in the fabric switches to resolve the issue.

Important: In SAN fabrics with multiple switches, always use the same firmware code revision on all switches. This action will prevent random problems from occurring within the SAN fabric.

- Perform a firmware upgrade on the switch.

  While not exactly an advanced troubleshooting method, the ability to perform code upgrades to the IBM 2109 can be crucial for the continued, smooth operations of a SAN environment. While the actual downloading of a new code is not disruptive, the activation of the new code will impact the SAN environment.

  Firmware downloads can be used in conjunction with either a UNIX host or a Windows 95/NT server. If using a Windows-based server, there are two additional utility programs that are required. This server should be in the same LAN segment as the fabric switches.

  The code upgrade process is comprised of three steps:

  a. Download the code from the IBM Web site to the local server.

     On Windows platforms, remote shell daemon (RSH) support must be provided. There are two utility programs available from the Web site:

     `http://www.ibm.com/storage/fcswitch`

     From this site, download the utilities:

     - cat.exe
     - rshd.exe

     These utility programs are self-extracting files. The files should unzip into a directory in a searchable path along with the firmware code file.

  b. Confirm that the Remote Procedure and TFTP deamons are active.

     On UNIX platforms, verify that the appropriate daemons (rshd and tftpd) are active.

     For Windows-based servers, start the remote procedure daemon with the **`rshd`** command.

c. At the switch, download the code from the local server over the LAN.

To initiate the download process on the switch, use the **firmwareDownload** command. This command can have the parameters included on the command line with the command, but they are not required. If not supplied, the switch will prompt for the following information:

- Host name or IP address
- User name for login
- Code file name to download
- FTP password (optional)

The inclusion of the fourth parameter will cause the download process to use the FTP protocol. Lacking this parameter causes the download process to invoke a remote shell using RCP commands.

For more information about this command and the download process, refer to *IBM SAN Fibre Channel Switch 2109 Model S16 User's Guide*, SC26-7351.

- Capture the current configuration and status.

  If remote technical support assistance is required, extensive information about the IBM 2109 switch may be requested. The switch has a single user command, **supportShow**, which will provide the output of a combination of 22 commands. This information includes configuration data for the switch and zoning, plus various data for error counters and error logs.

  When this information is requested, the best method for capturing it is setting up a screen capture function with the Telnet application that is running to connect to the switch. Once invoked, issue the **supportShow** command and allow the output to completely finish. Then stop the screen capture function and send the resulting file to the appropriate destination.

- Gather and review the error log.

  In many situations, the switch can detect problem conditions. In these cases, the switch will record the condition, with a date and time stamp, into an error log. The error log holds a maximum of 64 entries of the latest error conditions. To view the switch's error log, use the **errShow** command. Example 4-17 shows a sample of the output.

Example 4-17 Sample output of the errShow command

```
brocade02:admin> errShow

Error 64
--------
0x10e688f0 (tThad): Oct 12 09:32:09
    Error FW-BELOW, 3, fopportState006 (FOP Port State Changes 6) is below low
```

```
oundary. current value : 0 Change(s)/minute. (normal)

Type <CR> to continue, Q<CR> to stop:

Error 63
--------
0x10e688f0 (tThad): Oct 12 09:31:07
    Error FW-ABOVE, 3, fopportState006 (FOP Port State Changes 6) is above high
boundary. current value : 3 Change(s)/minute. (faulty)

Type <CR> to continue, Q<CR> to stop:
```

The various items of information for each error message is beyond the body of this publication. However, Example 4-17 on page 179 is from port 6 having a fiber cable attached and a device establishing a connection with the switch. In most cases the timestamp is critical when relating the entries in the error log to occurrences of the problem. This error log is part of the combined output of the **supportShow** command. This information will typically be provided to remote technical support personnel for intensive problem determination.

- Capture and review the port log.

  The IBM 2109 switch also maintains a log of switch activity associated with fabric services (such as link services and fabric logins) communicating with attached devices. One major advantage of this log is the arguments column. This information represents different values depending on the task and event. Typically, the arguments information is the first portion of the actual payload from the Fibre Channel frame. Thus, it can be an invaluable resource when performing problem determination actions.

  To view this log, use the **portLogShow** command. Example 4-18 is a portion of this log during the time when port 6 had a connection established on it from Example 4-17 on page 162. Additional information for understanding the data in this log can be found in the *IBM SAN Fibre Channel Switch 2109 Model S16 User's Guide*, SC26-7351.

Example 4-18 A portion of the portLogShow command output

```
brocade02:admin> portlogshow
Oct  5        task       event port  cmd  args
----------------------------------------------
09:31:04.516  tReceive   loopscn  6  LIP  8002
09:31:04.849  tReceive   loopscn  6  TMO  2
09:31:04.849  tReceive   pstate   6  LF2        * 2
09:31:04.849  tReceive   pstate   6  OL2
09:31:04.849  tReceive   pstate   6  LR3
09:31:04.849  tReceive   pstate   6   AC
09:31:05.349  interrupt  scn      6    1
```

```
09:31:05.366  tFabric    ioctl    6   90  102fcad0,0
09:31:05.366  tFabric    Tx       6  164  02fffffd,00fffffd,0311ffff,10000000
09:31:06.433  tReceive   Rx2      6  116  22fffffe,00000000,001fffff,04000000
09:31:06.433  tTransmit  Tx2      6    0  c0000000,00fffffe,001f0312
09:31:06.433  tTransmit  scn      6    6
09:31:06.433  tFspf      ioctl    6   ac  0,0
09:31:06.449  tFspf      ioctl    6   aa  ffffff,9      * 3
09:31:06.449  tFspf      ioctl    9   aa  ffffff,6
09:31:06.449  tFspf      ioctl   12   aa  ffffff,6
09:31:06.449  tFspf      ioctl   16   aa  ffffff,6
09:31:06.449  tFspf      ioctl    6   ad  0,0
09:31:06.449  tSwitch    ioctl    6   92  10a7e83c,0
09:31:06.449  tSwitch    Tx2      6  116  23021600,00fffffe,001f0312,02000000
```

4.5.3 McDATA specific checks

The principle method of interfacing with McDATA switches is provided from an Ethernet-connected host running the Enterprise Fabric Connectivity (EFC) client application. The client application communicates with the EFC server. The EFC server's responsibility is to provide an interface between EFC clients and the switches. The EFC client can reside on the same host as the server, or be installed from a separate host system.

The EFC server can work with most McDATA products, including:

- ED-5000 Enterprise Director
- ED-6064 Enterprise Director
- ES-3032 Fabric switch
- ES-3016 Fabric switch
- ES-1000 Workgroup switch

As a result of this wide range of managed products, there are consistent methods, or pathways, for gathering information from different switch models. The main difference in methods is the graphics used to display information from the various models. Even though the displays are different, depending on the switch model being queried, the same icons and buttons are used in the EFC GUI to view the same information. Because of this similarity of functionality, we provide example figures from several switch models during the following discussions.

For the purposes of our following descriptions, we assume that an EFC server is already running and communicating with all switches in the SAN fabric via LAN connections. The following methods are being used on an EFC client that is communicating with the server application.

When the EFC client is first started, we select **View -> Hardware** to start the problem determination process. This view presents icons of all the switches that have been defined to the server. One of the first indicators of possible problems is the status color around the switch icon. If the circle is not green, then some type of problem exists in relation to that switch. To see the Hardware view of a given switch, simply click on its icon.

When checking a switch, the first quick troubleshooting action is to check the status of the ports with connections to other devices. These connections are for both edge devices and other fabric units. For the McDATA ED-5000 switch, each port has two LEDs, green and amber. Refer to Figure 4-19 for the location of the LEDs in relation to ports on the McDATA ED-5000.

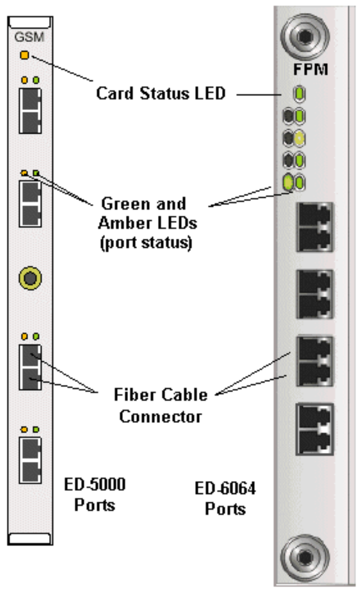


Figure 4-19 McDATA ED-5000 port status LED locations

Refer to Table 4-4 for the correlation of the LED condition and port status for ports on the ED-5000. On the ED-6064, the green LED signals that a link has been established, while the amber LED denotes traffic movement.

Table 4-4 McDATA ED-5000 port status indicators

Green LED	Amber LED	Port status
Off	Off	No light detected
On	Off	Connection established to a device
Flicker	Off	Traffic is flowing on port

Green LED	Amber LED	Port status
Off	On	Port failure
Off	Flicker	Running diagnostic

If one or more ports on the switch have visual indications of problems, then check the device's port condition at the other end of the connection. The nature of the problem may be better defined when the conditions at both ends of the link are known. If both ends indicate a problem, there may be a fault with the physical cabling and/or a logical connection problem. Possible causes of physical connectivity problems with the port include:

- Loose connection of fiber cable into the port

  Make sure that the fiber cable is securely plugged into the port. If it is not, the fiber cable can be misaligned in the connector. This misalignment can reduce the signal significantly to the point that no light is detected.

- Transmit and receive lines of end ports are not correctly connected

  The transmitter at one end of a fiber cable run should be connected to the receiver at the other end. Refer to Figure 4-11 on page 134 for proper orientation. A quick method to test for this condition is to swap the individual connectors at one end of the fiber cable. If a physical connection is not established, then the problem has a different cause.

- High signal loss between end ports of a connection

  This situation can be prevalent in SAN environments with patch panels, or where the fiber cable is physically damaged. With some devices, use a loopback tool to wrap the port on itself. This test should cause a change in the port status indicators if there is a faulty fiber cable.

- Fibre Channel port is faulty

  A faulty port can be difficult to test. There are advanced diagnostics that can be implemented, but these test routines are usually disruptive to the SAN. A much easier test is to use a fiber cable from another established connection. The test simply consists of swapping the good fiber connection with the suspect fiber connection. If the suspect port's status indicators do not change, the reader will need to proceed with further troubleshooting actions.

For problems that are not due to a physical cabling issue, the reader will need to continue with troubleshooting methods previously outlined. The following material is presented to assist the reader with specific actions and methods for McDATA switches.

After visual inspection of the port LEDs, check the port status by using the client's GUI interface. While the LED indicator may indicate an established connection, the type of connection (for example, E_Port or F_Port) is not known. There are several navigation paths through the client's GUI that will let you see the status of ports and additional information.

> **Note:** The following procedures will make use of the McDATA EFC application. We assume the client is able to communicate with the server while the server is communicating with any McDATA switches in the SAN fabric.

From the EFC client's Hardware view, we are presented with the front and the back of the selected switch. These are interactive views where you can query the status of individual modules by clicking on them. Once clicked, more information about the module is displayed. Also included with the expanded view of a module is information describing its operational status. In our troubleshooting method, we selected a Port I/O module with four ports. By clicking on one of the ports, another window is displayed with detailed port information. Figure 4-20 on page 168 shows a Port Card view with detailed port information. This informational window is called the Port Properties panel.

The Port Properties panel includes a number of data items of interest such as:

- Current operational status of the port
- WWN of any active, attached device
- Type of port connection (such as port 20 being a F_Port in our example)
- Link incident listing last identified problem with the port

This detailed information provides the data for confirming that another SAN device is properly connected to the switch's port.

There may be an additional icon on this panel in the shape of a triangle in the Link Incident field. If a triangular icon is present, then the switch has detected one or more errors on a given port. By clicking on the icon, a new informational panel will appear with additional data and descriptions about the detected error.

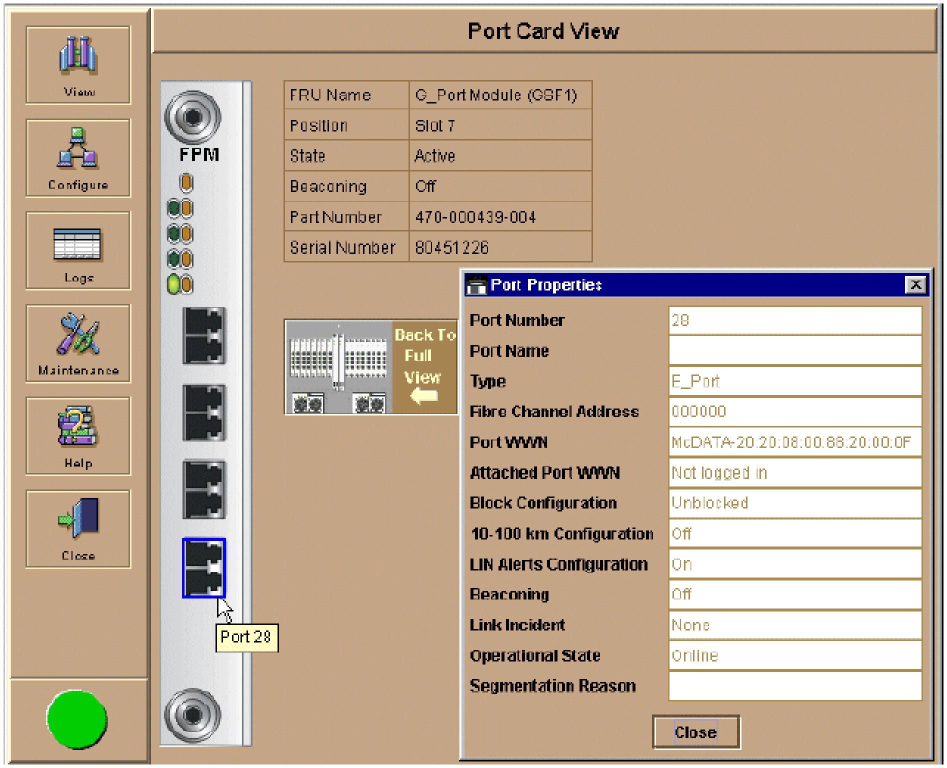


Figure 4-20 Verification of port state and previous errors (McDATA ED-6064)

Another method for quickly viewing the status of all ports is by selecting the Port List view. This view provides the following information for all ports of the switch:

- User-assigned name of port (if any)
- Block configuration
- Port status (as in No Light, Not Installed, Online)
- Type of port (as in G_Port or F_port)
- An triangular icon, if any alerts or events are associated with a specific port

One major drawback of the Port List view is that the WWN of any attached devices is not listed in the data. However, the WWN can be found by clicking on any entry that has an attached device. Using these panels, we can quickly determine whether all SAN devices attached to the switch ports are operating properly.

If there are indications of problems with ports on the switch, there is a log of link incidents. A link incident is a problem on a link that the switch has detected. An example of a Link Incident Log is shown in Figure 4-21. To view this log, select **Logs -> Link Incident Log**.

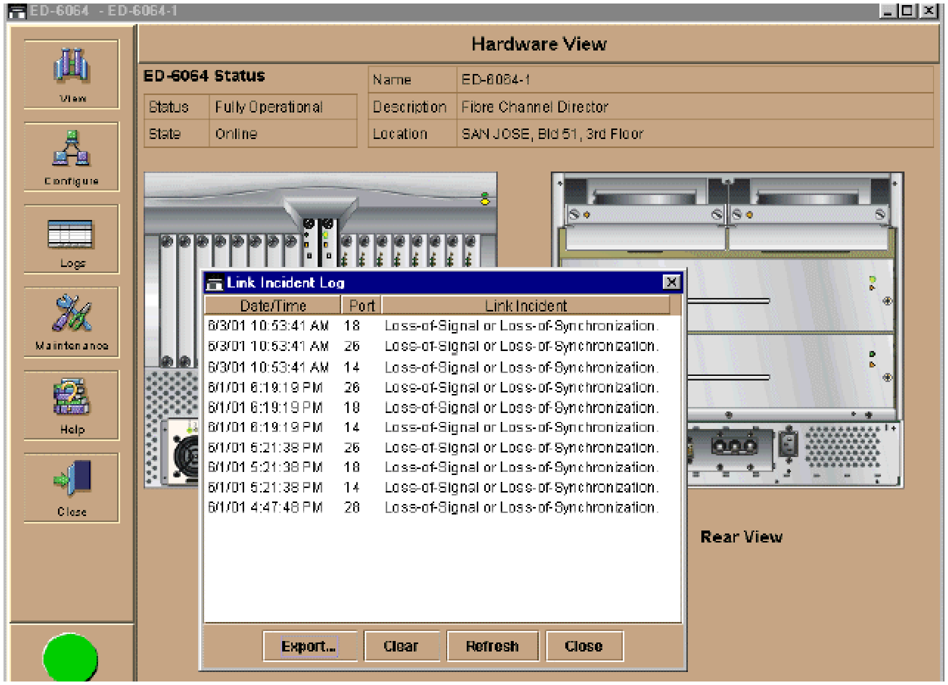


Figure 4-21 Example of link log showing error conditions

Since we have accessed the Port List view, all ports of the switch are displayed at one time. With this convenience during the problem determination process, we do not have to separate the verification process of fabric and edge device connections into two separate actions. However, the SAN must be documented properly so that the reader knows what the expected settings for the connected ports should be. For example, we use an example where the status of a given link is a N_Port when it should be an F_Port. In this scenario the switch may show no problems for the port. However, since the port is linked to another switch but treated as an edge device, communications between real edge devices may not be working at all.

If at this point the problem has not yet been identified, there may still be some suspicion of the SAN fabric. Even though the ports with fabric interconnects were found to be operational and at the proper setting, there are other functions within the fabric that bear checking. The first of these is the interconnections with other switches within the fabric. The quickest check of this is by displaying a graphical listing of all switches in the fabric.

The EFC client GUI provides a very quick graphical view of inter-switch connections. From the main display, select **View -> Fabric**. This path will then display all fabrics that have been defined to the EFC server. If there are serious problems within a SAN fabric, the icon visually displays an indication of issues. If the server has more than one SAN fabric that it is monitoring, select the desired fabric to be investigated by clicking on the appropriate icon. This action results in the fabric Topology view, shown in Figure 4-22 on page 170.

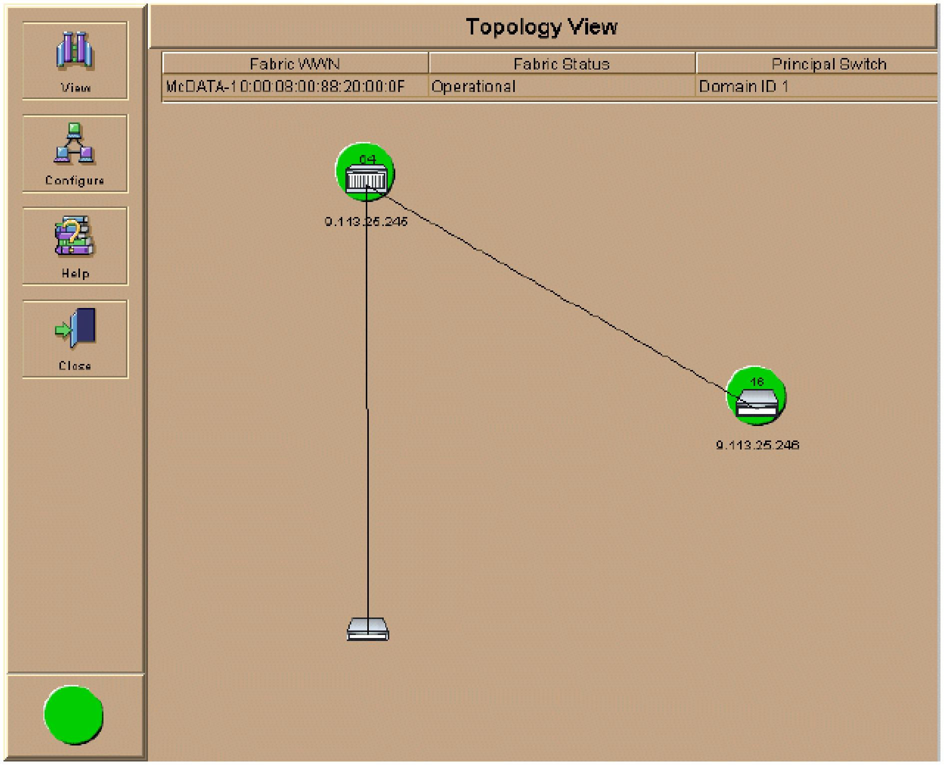


Figure 4-22 Checking of fabric links and devices with EFC

With this high-level view of the SAN fabric and the fabric documentation, any missing fabric devices can be quickly identified for further troubleshooting actions.

By using the EFC application, many potential zoning configuration issues in a multiple-switch SAN fabric are eliminated. This is because the EFC server loads a consistent zone configuration into all defined fabric switches. Thus, conflicting zone configuration issues between two or more switches are eliminated. Further, the EFC server will scan any changes for obvious configuration errors before allowing the changes to the existing zone configurations.

However, the EFC application cannot determine whether there are user-induced mistakes in a given zone configuration, such as an invalid WWN, a port wrongly specified, or storage resources incorrectly allocated. In cases such as these, the EFC manager may determine that a zone configuration is okay, while the improper zone configuration blocks all communications between certain SAN devices.

As stated earlier in this chapter, proper documentation is crucial to effective problem-determination success. By displaying the current zone configuration and comparing with the documentation, this type of problem can be quickly found and corrected.

To see the current zoning configuration of a SAN fabric, we start from the Topology view on the EFC client. From here, select **View -> Zoning**. This action will display the Zoning view, as shown in Figure 4-23. When troubleshooting potential zoning configuration issues, consult the SAN environment's documentation to verify the configuration.

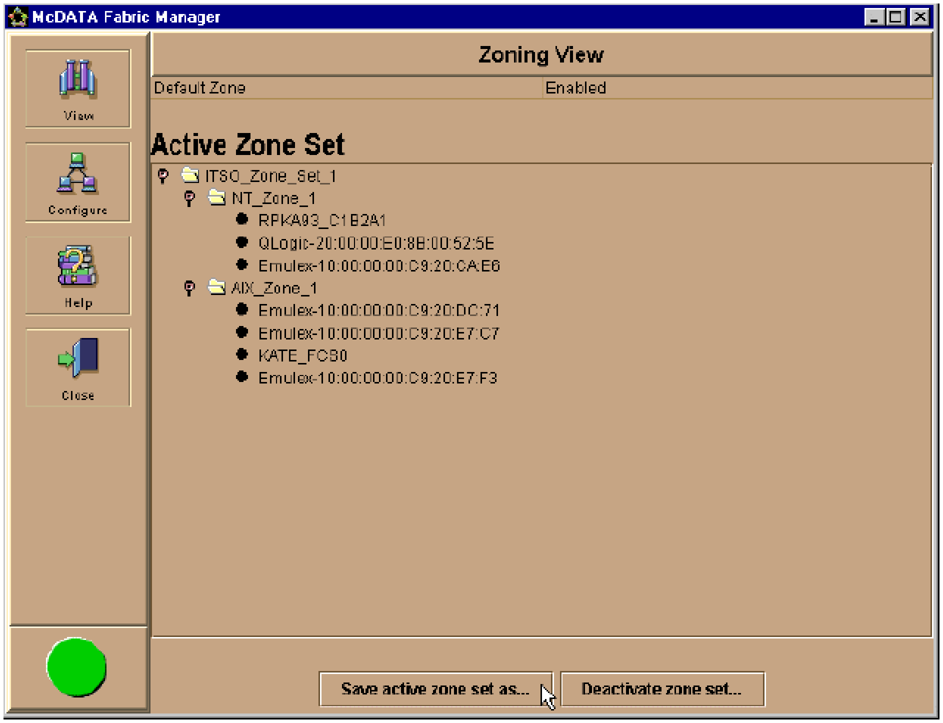


Figure 4-23 Example view of current zone configuration in McDATA switch

With these checks completed, it is a reasonable assumption to say that the fabric is not the problem source in cases of consistent problems. However, for the random type of problem, the SAN fabric with the McDATA ED-5000 may still be an issue. There are still a number of items that can be checked, but these are considered to be advanced troubleshooting techniques. We will go on to describe some of the methods for troubleshooting, as well as how to have the information readily available in case further assistance must be sought from a technical support group. Some of the advanced troubleshooting methods are:

- Determine the current code level.

  From the Product view, click on the switch to be inspected. This action displays the Hardware view. From the navigation control panel, select **Maintenance -> Firmware Library**. See Figure 4-24. In the lower left corner of the window is the current code level.

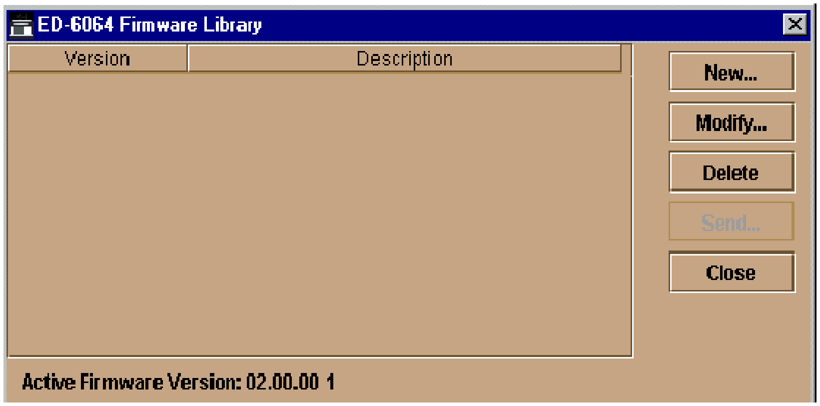


Figure 4-24 The EFC window with code level information

If there are questions or concerns about known code bugs in the current version, contact McDATA technical support. To contact the McDATA Solution Center for additional remote technical support:

– 800.752.4572 (North America customers)

– 720.566.3910 (all other customer geographies)

– 720.566.3851 (fax)

– support@mcdata.com (e-mail)

Note: The relationship between IBM and McDATA is a reseller agreement. In this arrangement, McDATA typically provides technical support for its switches, regardless of who sold the product. Depending on the type of maintenance support agreement in place, you should use your normal, local technical support process with the vendor from whom you received the product.

- Download the code to the switch.

  While not exactly an advanced troubleshooting method, the ability to perform code upgrades to the McDATA switch can be crucial to smooth operations in a SAN environment. While the actual downloading of new code is not disruptive, the activation of the new code will impact the SAN environment.

Note: For the description of the code upgrade process, we assume that the new code is available on the EFC server host platform, and is known to the EFC application. To reach this condition, please refer to the Release Notes or EC instructions that are part of the code package for the specific actions.

Firmware downloads are initiated from the EFC server application. From Product view, select the icon representing the switch that is to receive the download. This action opens the Hardware view. At the navigation control panel, select **`Maintenance`** -> **`Firmware Library`**. Select by highlighting the firmware version to be downloaded. Click the Send button. Once the EFC has verified that the switch is available and ready to receive the new code, it prompts the user for confirmation. Click the *Y*es button to start the code transfer.

While the code is downloading, a dialog box displays the progress of the file transfer. Once complete, the switch will reset for the new code to take effect. After the reset, the new code is loaded into the backup CTP module, if installed in the switch. Wait for this action to complete before closing out the Firmware Library box.

Important: The release notes or EC instructions that are part of the code package should always be reviewed before starting the download process. Use the procedural actions from the code documentation if different from the general procedure we have provided.

- Back up the current switch configuration.

  To back up the director configuration file to the EFC server from the Product view display. Select the icon representing the switch for which the configuration file will be backed up. The Hardware view for the selected director is displayed. At the navigation control panel, select **Maintenance -> Backup and Restore Configuration**. The Backup and Restore Configuration dialog box is displayed. See Figure 4-25.

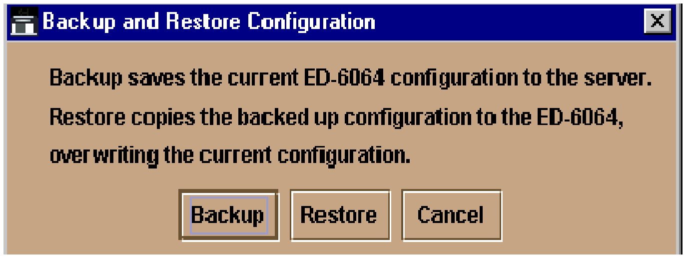


Figure 4-25 Backup and Restore Configuration window

  Click the *Backup* button. When the back-up process finishes, the Backup Complete dialog box is displayed. Click the OK button to close the dialog box and return to the Hardware view.

- Capture the maintenance data for further technical assistance.

  When the switch detects a critical error or module failure, it automatically copies the contents of its dynamic random access memory (DRAM) to a dump area in the FLASH memory on the active CTP module. If possible, the switch then transfers the captured dump file to the EFC server over its LAN connection.

  Perform the maintenance data collection procedure after a firmware fault is corrected or a failed module is replaced to capture the data for analysis by third-level support personnel. Maintenance data includes the dump file, hardware log, audit log, and an engineering log viewable only by support personnel.

  From the Product view displays, select the icon representing the director for which the data collection procedure will be performed. The Hardware view for the selected director is displayed. At the navigation control panel, select **Maintenance -> Data Collection**. The Save Data Collection dialog box is displayed. At the Save Data Collection dialog box, select the appropriate directory to store the file from the Look in drop-down menu, then type in a descriptive name for the collected maintenance data in the File name field. Ensure that the file name has a .zip extension, then click the *S*ave button.

- Capture and review the switch event log.

The switch event log displays a history of events for the switch, such as system events, degraded operations, module failures, module removals and replacements, port problems, Fibre Channel link incidents, and EFC server-to-director communication problems. All detected software and hardware failures are recorded in the Event Log. This information is extremely valuable for troubleshooting and repair verification.

To open the ED-6064 Event Log from the Hardware view, Port Card view, Port List view, FRU List view, Node List view, or Performance view, from the navigation control panel, select **Logs -> Event Log**. The log contains the following information:

– The date and time of the event.

– A three-digit code associated with the event. Refer to the specific product's service manual for an explanation of the event codes.

– A brief description of the event.

– The level of severity of the event (informational, minor, major, or severe).

– The switch module that is associate with the event.

– Supplementary event data may also be available for each event. This data can contain up to 32 bytes of hexadecimal data. Refer to the specific product's service manual for an explanation of the supplementary event data.

4.5.4 INRANGE 9000 specific checks

The primary method of interfacing with the INRANGE 9000 is through an Ethernet-connected host running the INRANGE Virtual Storage Network Enterprise Manager (IN-VSN). The IN-VSN is a browser-based application used to manage and control one or more INRANGE Directors. The IN-VSN suite consists of two components:

- Server software
- Client software

The server software communicates with the INRANGE Director, while the IN-VSN client software communicates with the IN-VSN server. All user interface is performed by the client software.

The IN-VSN management software capabilities include:

- Defining module and port configurations
- Defining zoning parameters
- Monitoring alarms and system performance

- Invoking system diagnostics

Another method for working with the director is via serial debug port on the FCM board. However, many of the available options are only for qualified Field Service Technicians. With this limitation in mind, we do not describe or make use of this access method. We will limit our material to working with the IN-VSN application.

When checking a switch, the first quick troubleshooting action is to take is to check the status of ports with connections to other devices. These connections are for both edge devices and other fabric units. For the INRANGE 9000 Enterprise Director, each port has two LEDs: green and amber. Refer to Figure 4-26 on page 178 for the locations of the LEDs for each port. Table 4-5 on page 178 lists the LED indicators for normal port operations.

> **Note:** A number of the initial installation parameters are accessible to a Customer Engineer only. Based on this limitation, there are several restrictions as to the type and degree of troubleshooting that you can utilize on the INRANGE 9000 Fibre Channel Director. We account for these restrictions with the following descriptions and problem determination methods. Further, we use the assumption that the IN-VSN server and client applications are running correctly and communicating with each other and the Director properly over the LAN.

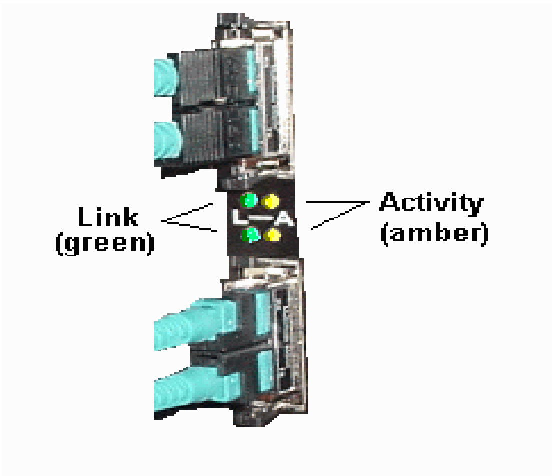


Figure 4-26 INRANGE 9000 port status LEDs

Table 4-5 Normal port indicator conditions for INRANGE 9000

Green LED	Amber LED	Port status
On	Off	Link established, no traffic
On	On/Flicker	Link established, traffic passing

If one or more ports on the switch have visual indications of problems, then check the device's port condition at the other end of the connection. The nature of the problem may be better defined when the conditions at both ends of the link are known. If both ends indicate a problem, then there may be a fault with the physical cabling and/or a logical connection problem.

Possible causes of physical connectivity problems with the port include:

- Loose connection of fiber cable into the port

  Make sure that the fiber cable is securely plugged into the port. If it is not, the fiber cable can be misaligned in the connector. This misalignment can reduce the signal significantly to the point that no light is detected.

- Transmit and receive lines of end ports are not correctly connected.

  The transmitter at one end of a fiber cable run should be connected to the receiver at the other end. Refer to Figure 4-11 on page 134 for proper orientation. A quick method for testing for this condition is to swap the individual connectors at one end of the fiber cable. If a physical connection is not established, then the problem has a different cause.

- High signal loss between end ports of a connection

  This situation can be prevalent in SAN environments with patch panels, or where the fiber cable is physically damaged. With some devices use a loopback tool to wrap the port on itself. This test should cause a change in the port status indicators if there is a faulty fiber cable.

- Fibre Channel port is faulty

  A faulty port can be difficult to test. There are advanced diagnostics that can be implemented, but these test routines are usually disruptive to the SAN. A much easier test is to use a fiber cable from another established connection. The test simply consists of swapping the good fiber connection with the suspect fiber connection. If the suspect port's status indicators do not change, then the reader will need to proceed with further troubleshooting actions.

For problems that are not due to a physical cabling issue, the reader will need to continue with troubleshooting methods previously outlined. The following material is presented to assist the reader with specific actions and methods for the INRANGE FC/9000 Fibre Channel Director.

> **Note:** The following procedures will be making use of the IN-VSN application. We assume that the client is able to communicate with the server while the server is communicating with any INRANGE switches in the SAN fabric.

After visual inspection of the port LEDs, check the port status by using the client's GUI interface. While the LED indicator may indicate an established connection, the type of connection (for example, T_Port or F_Port) is not known. To view the status of the ports, from the main window of the IN-VSN client, select Fabric from the Navigation Tree. This action displays a list of the user-defined fabrics. Select the desired fabric name in the Navigation Tree to display the directors within that fabric. Next, select the director for the problem determination process. This action will open the general view of the director. This display is the default view, and is shown in Figure 4-27.

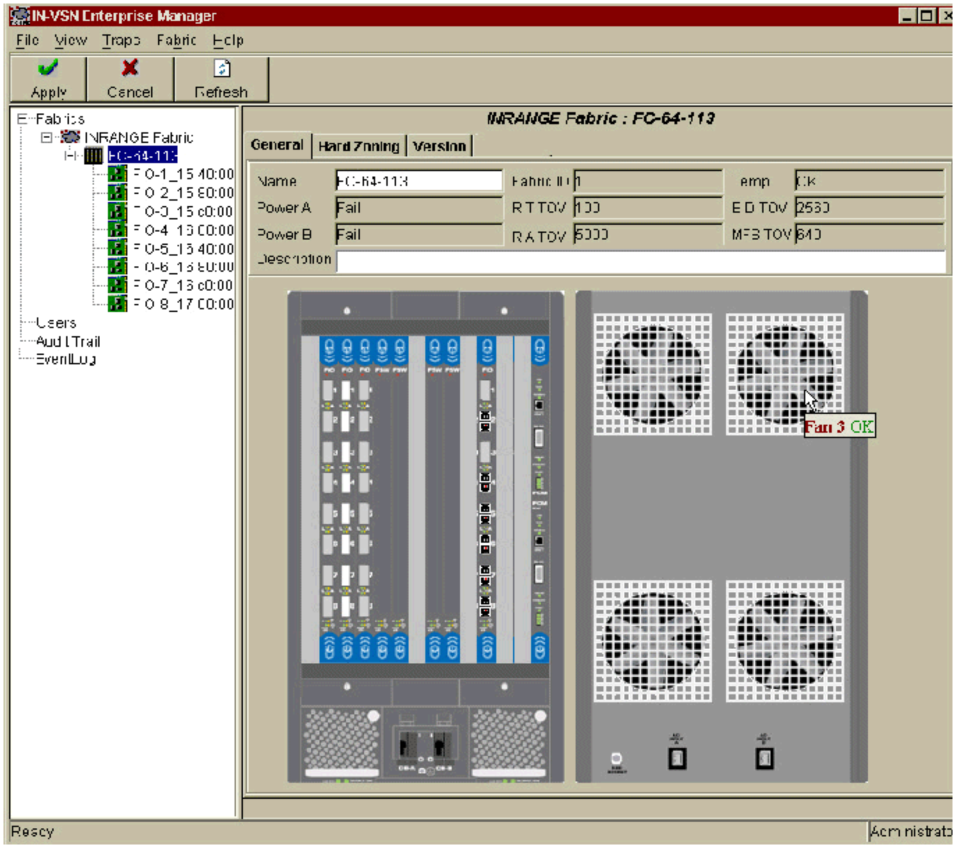


Figure 4-27 Default general view of a director

A method to quickly view the status of all ports is to select the Port tab (refer to Figure 4-28 on page 181). This view provides the following information for all ports of the switch:

- Port status (as in Offline or Online)
- WWN of any attached device and its alias
- Information about various zones (Name Server, Broadcast, or Hardware) that contain the ports

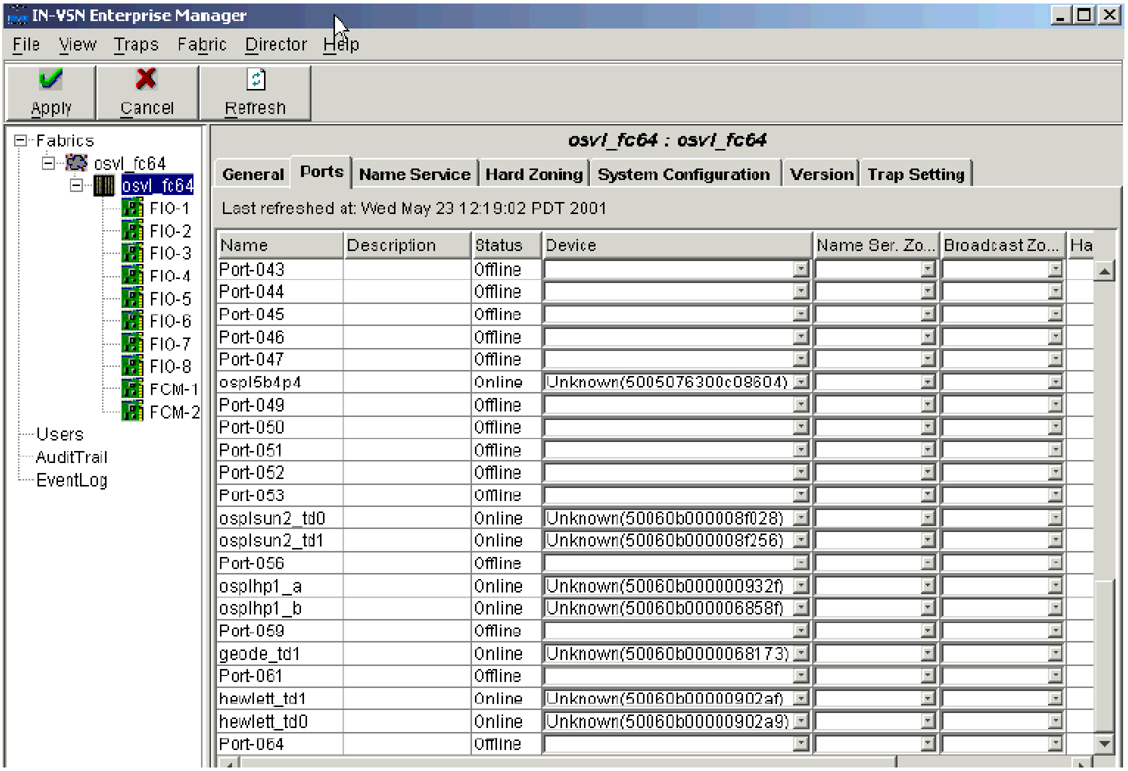


Figure 4-28 Example view of checking port states and connected devices

One major drawback of this view is the lack of information about the type of connection (T_Port or F_Port) for the ports. However, this information can be viewed by using the Navigation Tree. Select the appropriate Fibre Channel I/O module (FIO) in the director and then select the desired port for a detailed information display. Refer to Figure 4-29 on page 182.

The information that is displayed from this panel includes:

- Current connection type of port
- Administrative and operational status
- Fibre Channel address
- Statistical information about the traffic in tabular and graphical formats

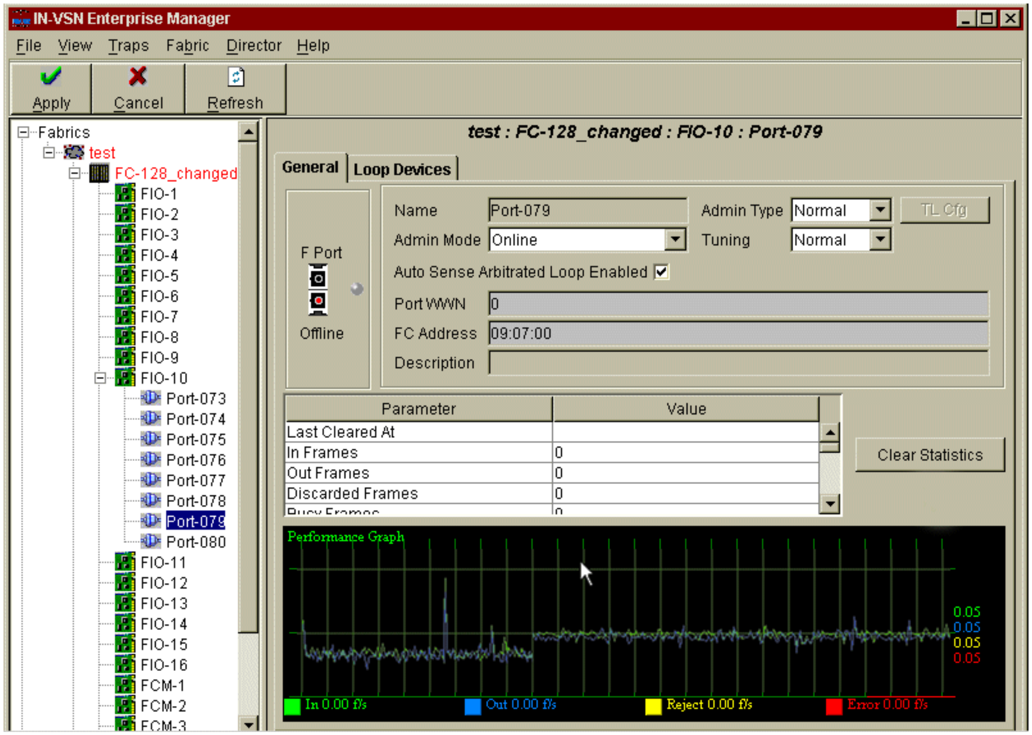


Figure 4-29 Checking status of a specific port with performance charting

Using the Ports view and additional data about specific ports, you can confirm the connectivity from the director to any SAN device attached to the director.

At this point, if the problem has not yet been identified, there may still be some suspicion of the SAN fabric. Even though the ports with fabric interconnects were found to be operational and have the proper setting, there are other functions within the fabric that should be checked. The first of these is the interconnections with other switches within the fabric. To check this item display a graphical listing of all of the devices in the fabric.

The IN-VSN client GUI provides a very quick graphical view of inter-switch connections. From the main start point on the Navigation Tree, select Fabrics. This action displays all fabrics that have been defined to the IN-VSN server. Select the desired fabric for further troubleshooting. This path will display the fabric devices and the connecting links (refer to Figure 4-30 on page 183). If there are serious problems within a SAN fabric, the Fabric Topology view visually indicates where the issue has been detected.

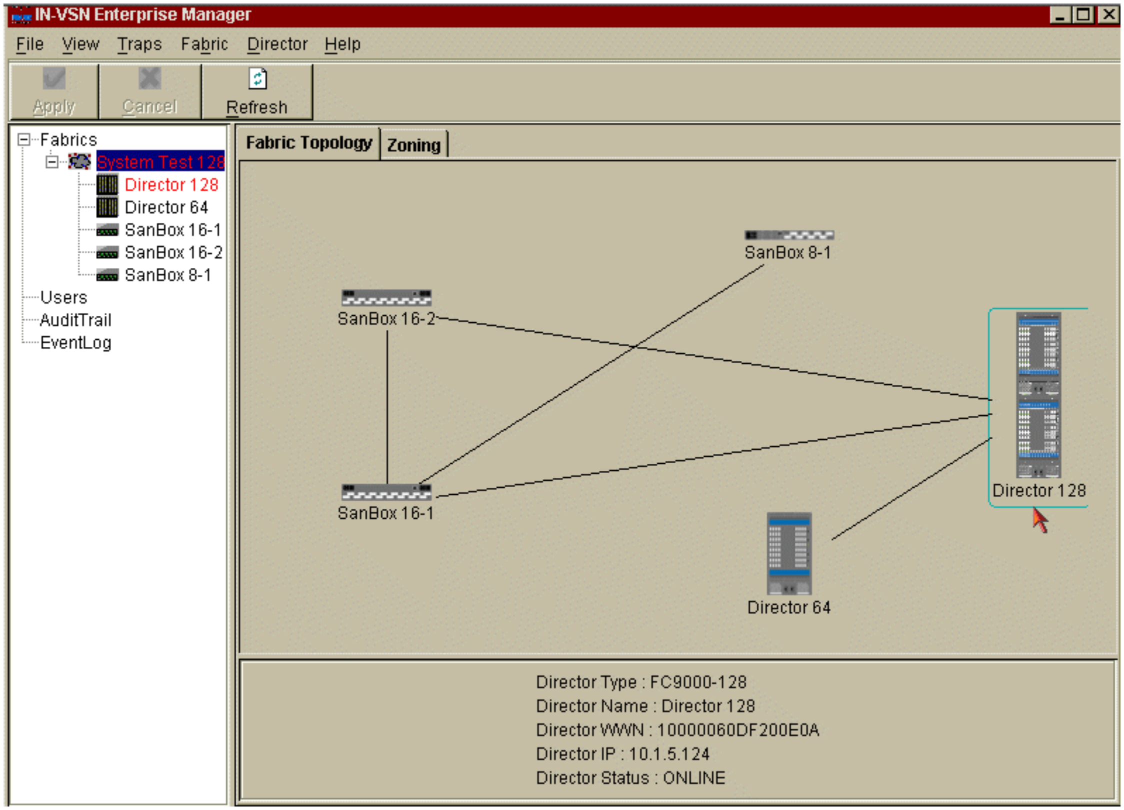


Figure 4-30 Checking fabric health with IN-VSN Enterprise Manager

With this high-level view of the SAN fabric and the fabric documentation, any missing fabric devices can be quickly identified for further, in-depth troubleshooting actions.

By using the IN-VSN application, many potential zoning configuration issues in a multiple switch SAN fabric are eliminated. This is because the IN-VSN server loads a consistent zone configuration into all defined directors in a fabric. Thus, conflicting zone configuration issues between two or more directors are eliminated. Further, the IN-VSN server will scan any changes for obvious configuration errors before allowing the changes to the existing zone configurations.

However, the fabric management application cannot determine if there are user-induced mistakes in a given zone configuration, such as the wrong port specified or storage resources incorrectly allocated. In cases such as these, the IN-VSN manager may determine that a zone configuration is okay, while the improper zone configuration blocks all communications between certain SAN devices.

As stated earlier in this chapter, proper documentation is crucial to effective problem determination success. By displaying the current zone configuration and comparing it with the documentation, this type of problem can be quickly found and corrected.

For purposes of zoning, the ports of the directors are divided into clusters with up to four ports per cluster. You have the option of putting each cluster into a hard zone. Each hard zone is color-coded, as shown in Figure 4-31 on page 185, and may be named (up to 24 characters). Ports within zones may only connect with other ports within that particular zone. By using the SAN fabric documentation, you can determine whether a device is connected to a port that has been assigned to the proper hard zone cluster.

If there are questions about the validity of the cluster configurations, use the Test option button. This feature reviews the configuration and will flag any clusters that are invalid.

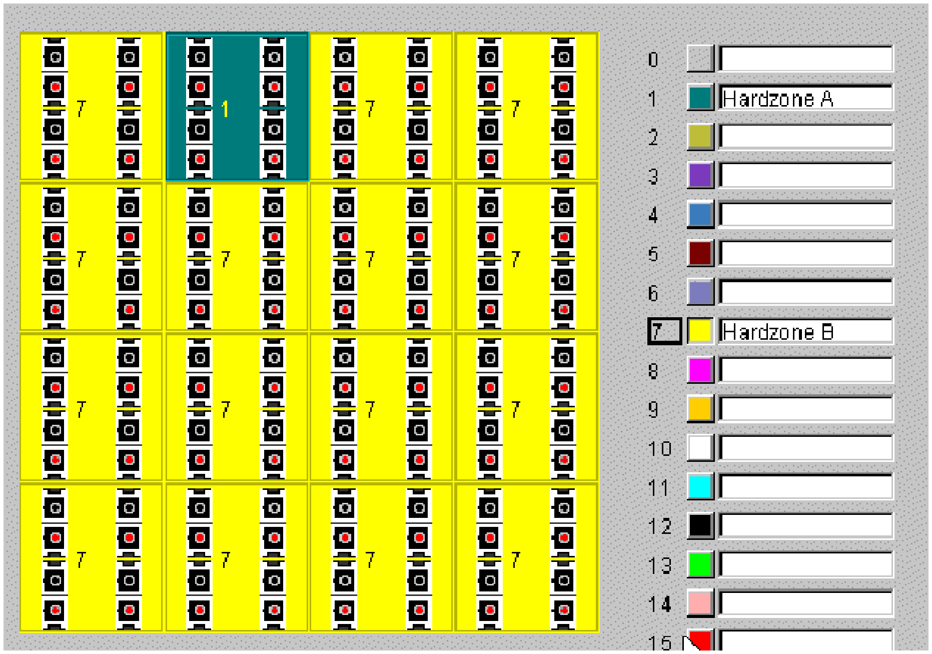


Figure 4-31 Graphical view of hard zoning INRANGE 9000 with IN-VSN Manager

Name server zones, also called soft zoning, allow the division of the fabric (one or more directors) into as many as 256 fabric-wide zones that define which ports receive name server information. A particular port may be defined in one or more of these name server zones. If hard zones are enabled, name server zones may not cross the defined hard zone boundaries.

Name server zones are definable by port. It is selected from the Fabric Topology window and applies to the entire fabric. If a zone is defined by port number, a port will receive name server information for all ports in the same name server zone (or zones) in which the port is defined.

From the Navigation Tree, select a fabric and then the Zoning tab. You can now view the name server zones (refer to Figure 4-32 on page 186). To get information about a name server zone, select it. The information for the zone now appears at the bottom of the window. If an HBA has recently been replaced within the network, this view is an effective problem-determination tool for verifying that the zoning configuration has been modified appropriately.

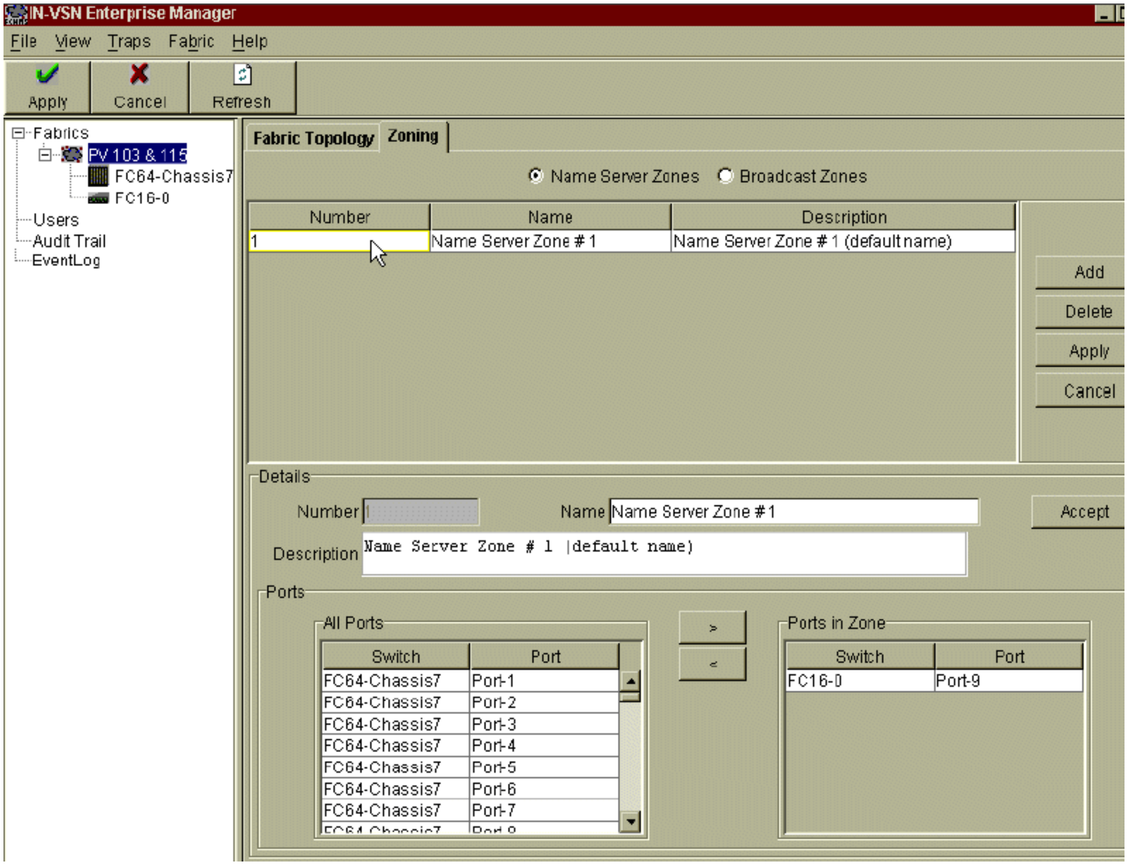


Figure 4-32 Checking fabric-wide soft zoning configuration

With these checks completed, it is reasonable to assume that the fabric is not the problem source in cases of consistent problems. However, for the random type of problem, the SAN fabric with the INRANGE F/C 9000 may still be an issue. There are still a number of items that can be checked, but these are considered to be advanced troubleshooting techniques. We will describe some of these methods for troubleshooting purposes, as well as how to have the information readily available in case further assistance is sought from a technical support group. Some of the advanced methods are:

- View and capture event log information.

  At the Navigation Tree, select on the director from which to capture the event information. Select the Event Log option at the Navigation Tree view. The event log of the director will now appear. To view a fabric-wide event log (as in a merged event log for all directors in the fabric), select a fabric instead of an individual director. An example of a fabric event log is shown in Figure 4-33. A major advantage of viewing a fabric-wide event log is seeing the relationship of how specific events are logged by different devices. However, this situation occurs only in extremely large SAN fabrics, so we describe working with a single director.

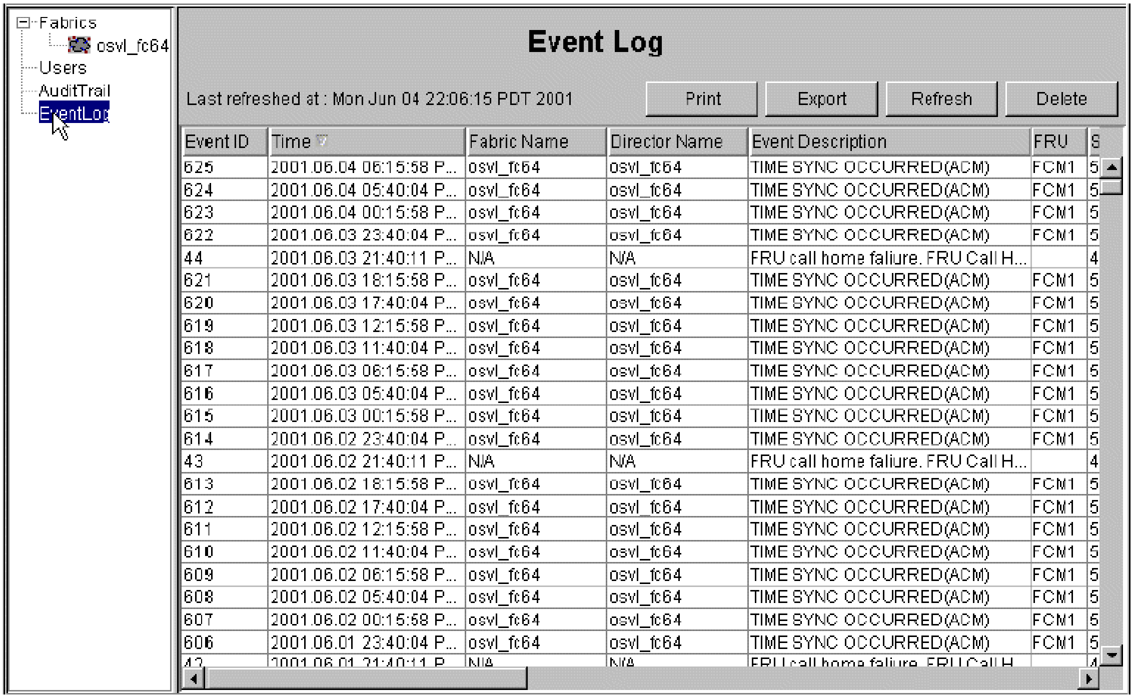
A number of informational items are provided in an event log entry. Some of the data are:

– The date and time of event

– The data to identify the director and the specific module associated with the event

– A short textual description of the event

– The event codes and extended information for support staff



Event ID	Time	Fabric Name	Director Name	Event Description	FRU	S
625	2001.06.04 06:15:58 P...	osvl_fc64	osvl_fc64	TIME SYNC OCCURRED(ACM)	FCM1	5
624	2001.06.04 05:40:04 P...	osvl_fc64	osvl_fc64	TIME SYNC OCCURRED(ACM)	FCM1	5
623	2001.06.04 00:15:58 P...	osvl_fc64	osvl_fc64	TIME SYNC OCCURRED(ACM)	FCM1	5
622	2001.06.03 23:40:04 P...	osvl_fc64	osvl_fc64	TIME SYNC OCCURRED(ACM)	FCM1	5
44	2001.06.03 21:40:11 P...	N/A	N/A	FRU call home faliure. FRU Call H...		4
621	2001.06.03 18:15:58 P...	osvl_fc64	osvl_fc64	TIME SYNC OCCURRED(ACM)	FCM1	5
620	2001.06.03 17:40:04 P...	osvl_fc64	osvl_fc64	TIME SYNC OCCURRED(ACM)	FCM1	5
619	2001.06.03 12:15:58 P...	osvl_fc64	osvl_fc64	TIME SYNC OCCURRED(ACM)	FCM1	5
618	2001.06.03 11:40:04 P...	osvl_fc64	osvl_fc64	TIME SYNC OCCURRED(ACM)	FCM1	5
617	2001.06.03 06:15:58 P...	osvl_fc64	osvl_fc64	TIME SYNC OCCURRED(ACM)	FCM1	5
616	2001.06.03 05:40:04 P...	osvl_fc64	osvl_fc64	TIME SYNC OCCURRED(ACM)	FCM1	5
615	2001.06.03 00:15:58 P...	osvl_fc64	osvl_fc64	TIME SYNC OCCURRED(ACM)	FCM1	5
614	2001.06.02 23:40:04 P...	osvl_fc64	osvl_fc64	TIME SYNC OCCURRED(ACM)	FCM1	5
43	2001.06.02 21:40:11 P...	N/A	N/A	FRU call home faliure. FRU Call H...		4
613	2001.06.02 18:15:58 P...	osvl_fc64	osvl_fc64	TIME SYNC OCCURRED(ACM)	FCM1	5
612	2001.06.02 17:40:04 P...	osvl_fc64	osvl_fc64	TIME SYNC OCCURRED(ACM)	FCM1	5
611	2001.06.02 12:15:58 P...	osvl_fc64	osvl_fc64	TIME SYNC OCCURRED(ACM)	FCM1	5
610	2001.06.02 11:40:04 P...	osvl_fc64	osvl_fc64	TIME SYNC OCCURRED(ACM)	FCM1	5
609	2001.06.02 06:15:58 P...	osvl_fc64	osvl_fc64	TIME SYNC OCCURRED(ACM)	FCM1	5
608	2001.06.02 05:40:04 P...	osvl_fc64	osvl_fc64	TIME SYNC OCCURRED(ACM)	FCM1	5
607	2001.06.02 00:15:58 P...	osvl_fc64	osvl_fc64	TIME SYNC OCCURRED(ACM)	FCM1	5
606	2001.06.01 23:40:04 P...	osvl_fc64	osvl_fc64	TIME SYNC OCCURRED(ACM)	FCM1	5

Figure 4-33 Event log from INRANGE 9000

There are two features from the Event Log panel the reader should notice. The buttons are Print and Export. If a printer is available to the client's host system, the event log can be printed as hardcopy. Using the Export button, create a plain text version of the file to store on the client's host system. The default location of the file is the directory with the application files, but the destination directory can be modified. This file may be requested by support groups if they are assisting with the problem-determination process.

When the Export option is utilized, the client software presents a new window prompting for the location of the resulting text file. Make any changes for the directory path and a file name and click Save. Once the file has been written, a pop-up message is displayed with a confirmation of the export operation.

- View and capture the audit log.

  The IN-VSN application has a special feature that can assist with answering the question "What has changed?" This information comes from the audit log, shown in Figure 4-34 on page 205. The audit log is a listing of all changes initiated by the IN-VSN server. Some of the captured information in the audit log includes:

  – The date and time of the operation

  – The type of operation and the operation's completion status

  – The director that implemented the operation

  – The user ID of the person making the change.

  To view the audit log, start in the Navigation Tree and select the Audit Log option. This action yields the summary of changes initiated from the server (refer to Figure 4-34 on page 189).

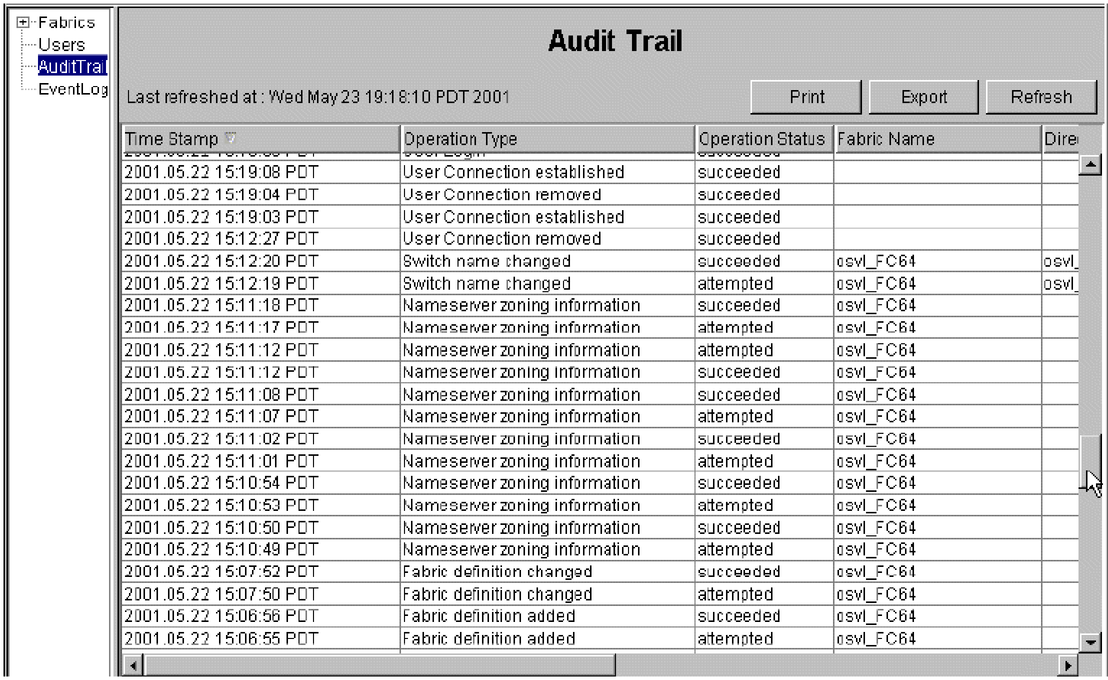


Figure 4-34 Audit trail from INRANGE 9000

The audit log information should be viewed in situations where the SAN has been stable for a period of time before a connectivity problem occurs.

- Determine the code levels on the modules.

This information will usually be captured at the request of technical support groups. We provide a method for viewing, and storing to file, this information. From the Navigation Tree, select the director to display the code information. In the display window, click the Version button. This action displays the code information for all modules in the director (refer to Figure 4-35 on page 190). This data can be stored in a plain text file using the Export button. The method for exporting the code revision data is similar to the method outlined for the Event Log.

You should note that we are not including information about upgrading code levels on the INRANGE Director. This is due in large part to the fact that INRANGE explicitly states that a number of situations must only be performed under the direction of qualified technical support. Failure to comply can void the warranty of the specific unit.

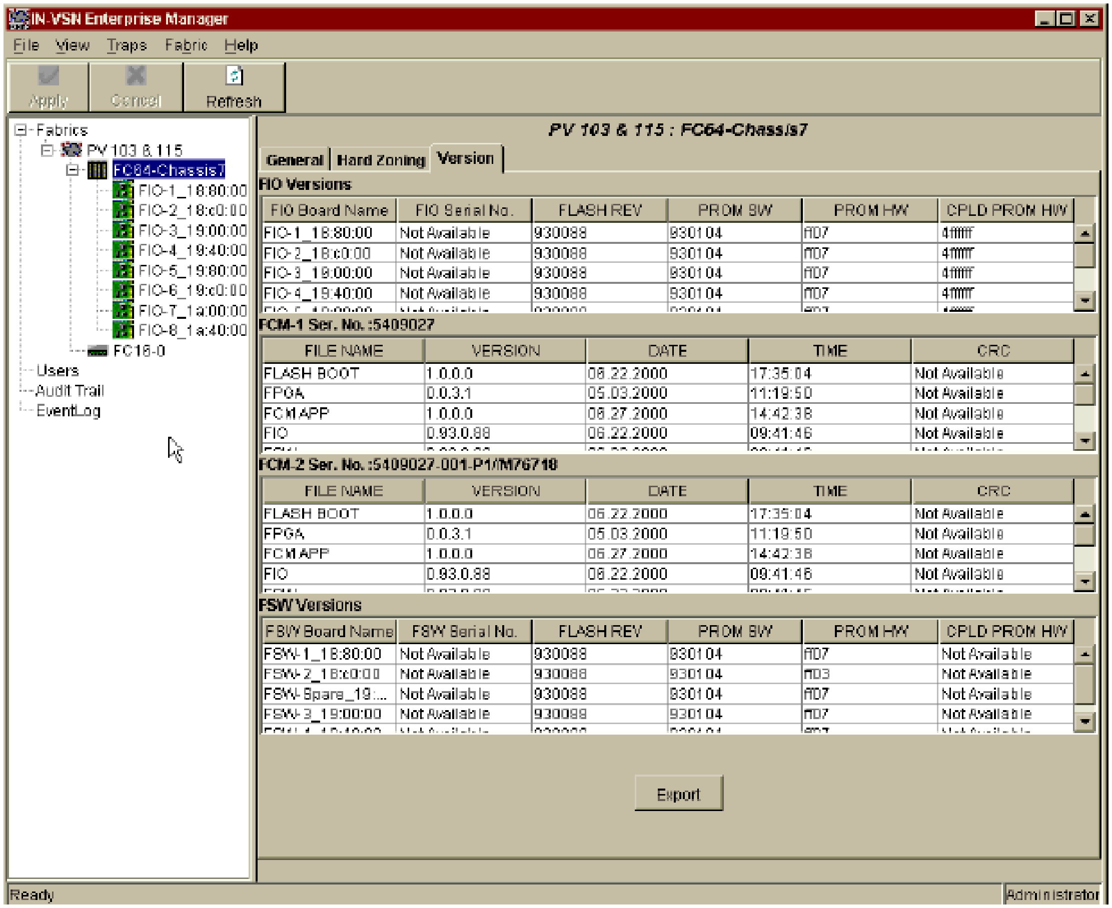


Figure 4-35 Code versions panel of INRANGE Director

- Capture, or back up, the director configuration.

  The IN-VSN application does not have the means to create an all-inclusive configuration file of the entire director. According to the *FC/9000 Fibre Channel Switch - Maintenance Manual*, DCN 9110774-307-00, the following information needs to be collected for the physical and logical configuration of a director. The data to capture includes:

  – Chassis or cabinet serial number

  – Number of FIO modules

  – Number of FSW modules

  – Number of FCMs

– Physical/slot locations of FIOs in chassis
– Physical configuration/ports attached
– List of active/inactive ports
– System log files

One item that is missing that we add is data about zoning configurations, whether hard zones or name server zones. As stated earlier, this information can be exported to files for future reference.

4.6 Checking the storage systems

This section provides some guidance for the problem determination process for several different types of storage systems. Regardless of the type of maker, the storage products are complex units and the troubleshooting options available are very limited. With such a wide range of storage types (as in disk to tape), we limited our coverage to the following products:

- IBM 2105 ESS
- IBM 3590 Tape System
- EMC Symmetrix

4.6.1 IBM 2105 ESS

The IBM 2105 ESS is an incredibly complex storage system. With its numerous features and functions for reliability and robustness, the IBM 2105 can suffer a rare failure in a subsystem, yet continue operations as a storage resource for multiple servers. However, this complexity has associated costs when troubleshooting problems potentially involving the IBM 2105. The sheer number of components, the multitude of subsystem microcodes, and the range of configuration options are topics of a small library of user guides and service manuals.

We provide a method of verification of host connectivity as the basis for the SAN problem-determination process involving the IBM 2105. To check the operations and connectivity of the ESS, the sequence of problem determination actions are:

1. Check known problems with the ESS.
2. Verify that one or more servers are known to the ESS.
3. Confirm that a connection path exists from the server to the ESS resources.
4. Verify the ESS resource allocation for the server.

During the writing of this publication, we used the IBM 2105-E20 and IBM 2105-F20 models as external disk storage resources for the pSeries servers.

The first troubleshooting action with the IBM 2105 checks for known issues with the IBM 2105. The first way to accomplish this task is to visually inspect the MESSAGES LEDs on the ESS front panel. If either of the LEDs are lit, then there is a physical problem within the ESS unit. With the richness of the internal diagnostic functionality, the ESS is very capable of detecting obvious failures and other significant degredations of its various subsystems.

We do not suggest opening the doors of the ESS to verify physical connectivity at a specific port. Because of the complexity of the ESS, there are few user-interaction options available. As a result, the rest of the troubleshooting process will make use of the StorWatch Specialist Web user interface in the ESS and an external Web browser. For a listing of Web browsers supported by the ESS Web interface, refer to the *IBM Total Storage Enterprise Storage Server Web Interface User's Guide*, SC26-7346.

To continue the problem-determination process, the ESS Web interface is utilized. Start a Microsoft Windows-based browser (such and Netscape or MS IE). Enter the name alias or IP address for either cluster controller of the ESS in question. When first accessed, the ESS Specialist Welcome panel is displayed. The Welcome panel provides the following information about the ESS:

- Machine type, in this case 2105
- Model type, for example, E20 or F20
- ESS serial number
- Worldwide node name (WWNN) of the ESS

After this window, the browser will prompt for a valid user name and password before continuing with routine operations. Once supplied, the first item to review is the ESS Specialist Status panel (refer to Figure 4-36 on page 193). The Status panel is a remote method for capturing additional information when problems are indicated by the MESSAGES LED.

To access this panel, from the navigation frame, click Status. Once the Status panel is open, select Refresh Data to display the current, not a previously cached, status.

Ensure that the Status panel shows all machine resources are in a normal state and that there are no problem records listed. If there are no problem records, continue with the problem-determination process. Otherwise, contact your ESS service provider to resolve any problems. If these problems do not resolve the SAN issue, then further investigation is needed. To gather more information about a particular problem, select Problem Log to access the Problem Log panel. Some of the information in this panel is:

- Cluster number
- Problem description, status, and age
- Possible user actions
- Probable cause

If these problems do not resolve the SAN issue, then further investigation is needed.

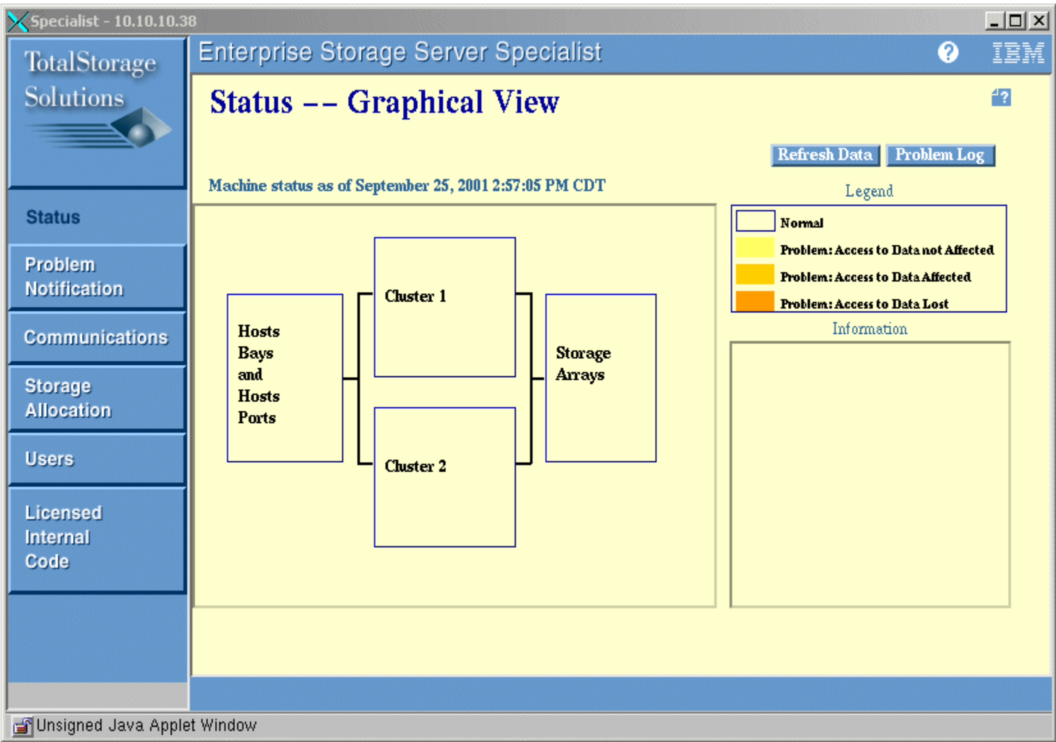


Figure 4-36 Status panel from ESS Web interface

After a check of the ESS status, verify the known information of a particular server by the ESS. To display this information, select **Storage Allocation -> Open Systems** to open the Open Systems Storage panel. From this panel, select the Modify Host Systems button. The Modify Host Systems panel (see Figure 4-37 on page 194) consists of a Host Attributes window and a Host Systems List window. Scan the Host Systems List for the particular server of interest, and click its line. This action displays the information about the server in the Host Attributes window. Review the configuration of the server in question. Check for valid settings in the following fields:

- Host type: For the proper operating system
- Host attachment: Should be Fibre Channel attached
- Hostname/IP address: Check that the correct server's data is being reviewed
- World Wide Port Name: The WWPN of the HBA in the server

Verify that the information in the Host Attributes window is correct and valid. This information should be compared with the SAN documentation, such the SAN devices table. If there are discrepancies with any information, make the corrections and then select the Perform Configuration Update button.

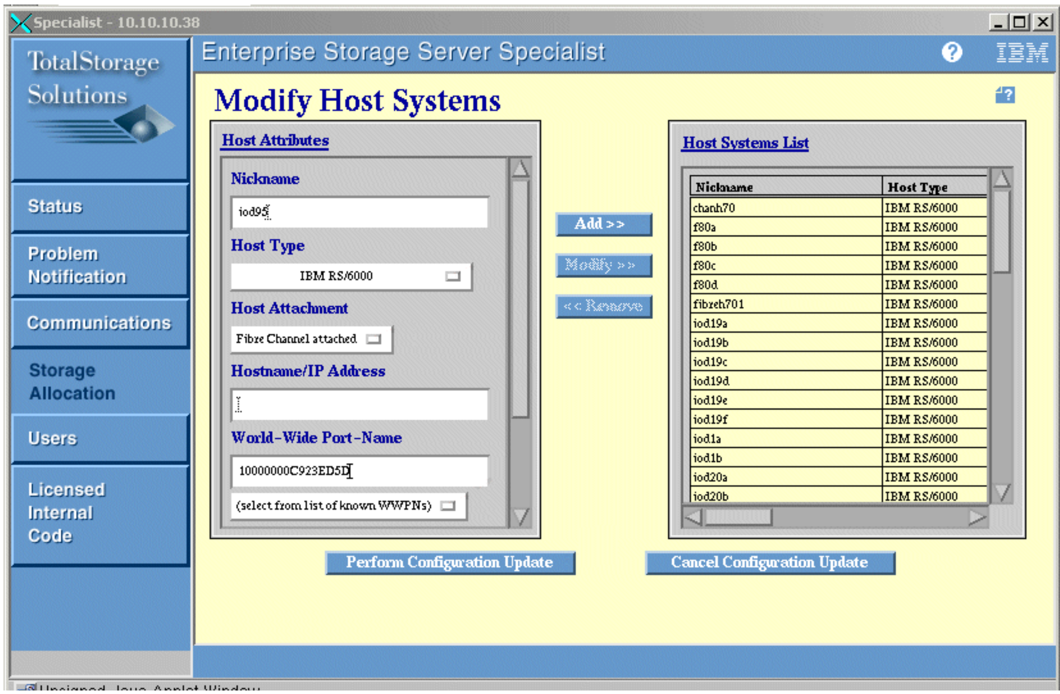


Figure 4-37 Verify host information

Once the server information is verified, check on the attachment of the server to an ESS adapter port, also known as the storage server host port. First, we determine which host adapter port is assigned to provide ESS connectivity to the server. To gather this information, we use the Add Volumes (1 of 2) panel. To access the Add Volumes panel, select **Storage Allocation -> Open Systems** to open the Open Systems Storage panel. From this panel, select the **Add** Volumes button. This action displays a graphical listing of known servers to the ESS, as shown in Figure 4-38 on page 195. By highlighting the server of interest, the panel shows which adapter port is assigned to provide connectivity for a given server HBA.

The information window provides the details about which ESS adapter port is assigned to a given server for connectivity. Refer to the SAN documentation or SAN diagram to verify that the correct adapter port is being used by the server.

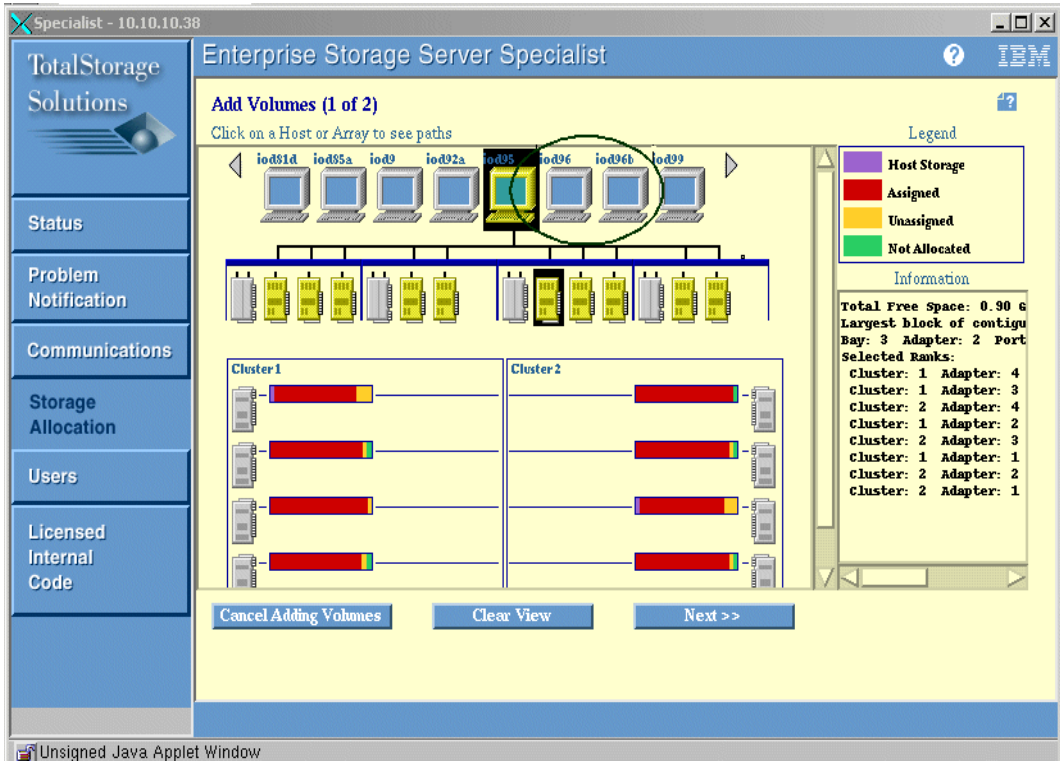


Figure 4-38 Display of adapter port assignments for servers

At this point in the troubleshooting process, one or more servers are correctly identified to the ESS system, and a valid ESS adapter port that provides connectivity to a given server has been identified. The next step is to verify the ESS adapter port configuration. To display this information, we use the Host Adapter Port panel. To access the Host Adapter Port panel, select **Storage Allocation -> Open Systems** to open the Open Systems Storage panel. From this panel, select the Configure Host Adapter Port button. This action displays the Configure Host Adapter Port panel, as shown in Figure 4-39 on page 197.

The items to verify in the Host Adapter Port panel are:

- Worldwide Port Name (WWPN)
- Fibre Channel topology
- Fibre Channel protocol

The WWPN is for the ESS adapter port that provides connectivity for certain servers in the SAN. This WWPN should appear as a name server entry within the SAN fabric. With most of today's SAN fabrics being created with switches, the topology setting is typically Point-to-Point (Switched Fabric). However, some switches have the ability to support Fibre Channel Arbitrated Loop on ports, so care must be taken when checking the topology setting. Finally, the protocol setting will usually be FCP (Open Systems).

If the ESS adapter port WWPN does not appear as a fabric name server entry, then closer investigation of the fiber cabling between the ESS and its attached fabric device is needed. Incorrect and invalid settings for topology and protocol should be corrected and the adapter port configuration updated.

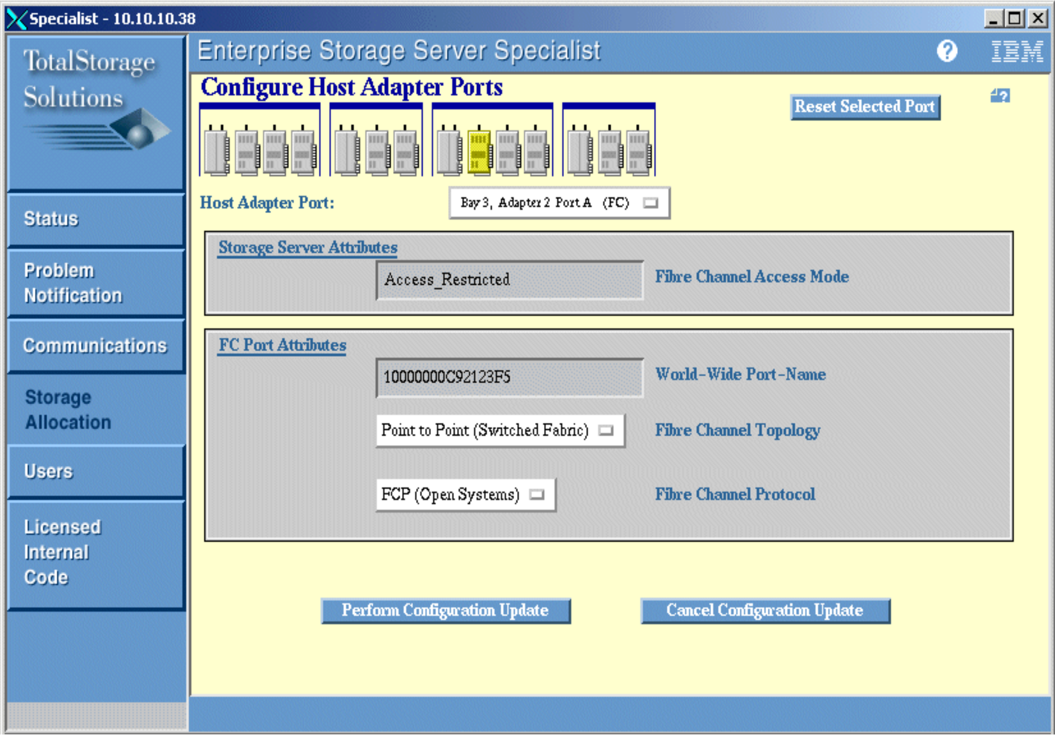


Figure 4-39 ESS adapter port setting

From the original checklist of four problem determination steps for the ESS, the first three actions have been completed. The final troubleshooting step is verifying that ESS resources have been allocated to a given server. A quick way of checking that storage has been allocated is by using the Add Volumes (1 of 2) panel. To access this panel (shown in Figure 4-40) select **Storage Allocation -> Open Systems** and click on the Add Volumes button.

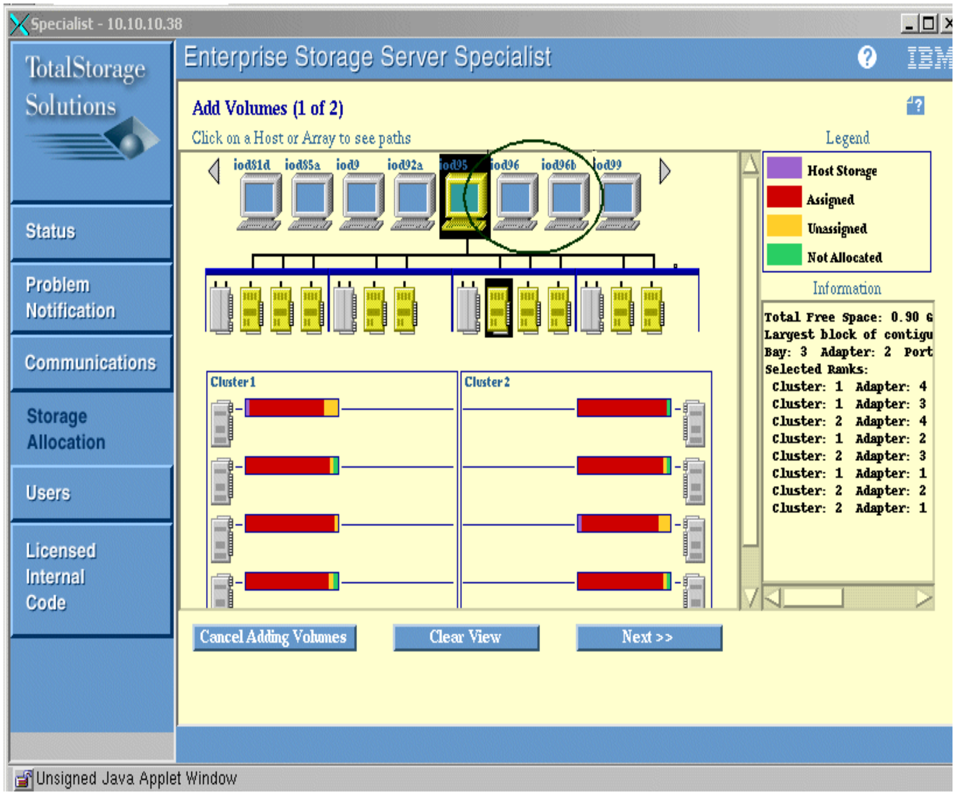


Figure 4-40 GUI display of allocated storage to a server

In the Add Volumes panel, select the server of interest with a single click on its icon at the top of the screen. The storage allocation for the selected server is then shown graphically. Thus, the fast check method confirms whether that disk space is assigned to a particular server. This graphical display does provide detailed information about the resource allocation.

For a detailed listing of storage allocation, use the Open Systems Storage panel. To open this panel, select **Storage Allocation -> Open Systems**. Scroll through the Host Systems list to highlight (single click) the server of interest. Once a server is selected, the Assigned Volumes window displays additional data about the ESS resources allocated to the server.

The main data item from the Assigned Volumes window is the number of volumes in the list and the type of volumes. The volume type should be Open System. The most important item deals with the number of volumes in the listing. The number of volumes should be equal to the number of external hdisks that the server reports having access. If there are discrepancies between the number of volumes belonging to a server by the ESS and the number of hdisks reported by the server, then the most likely cause is a configuration issue either at the server or in the ESS. For these types of configuration issues refer to the appropriate product's user guide for further assistance.

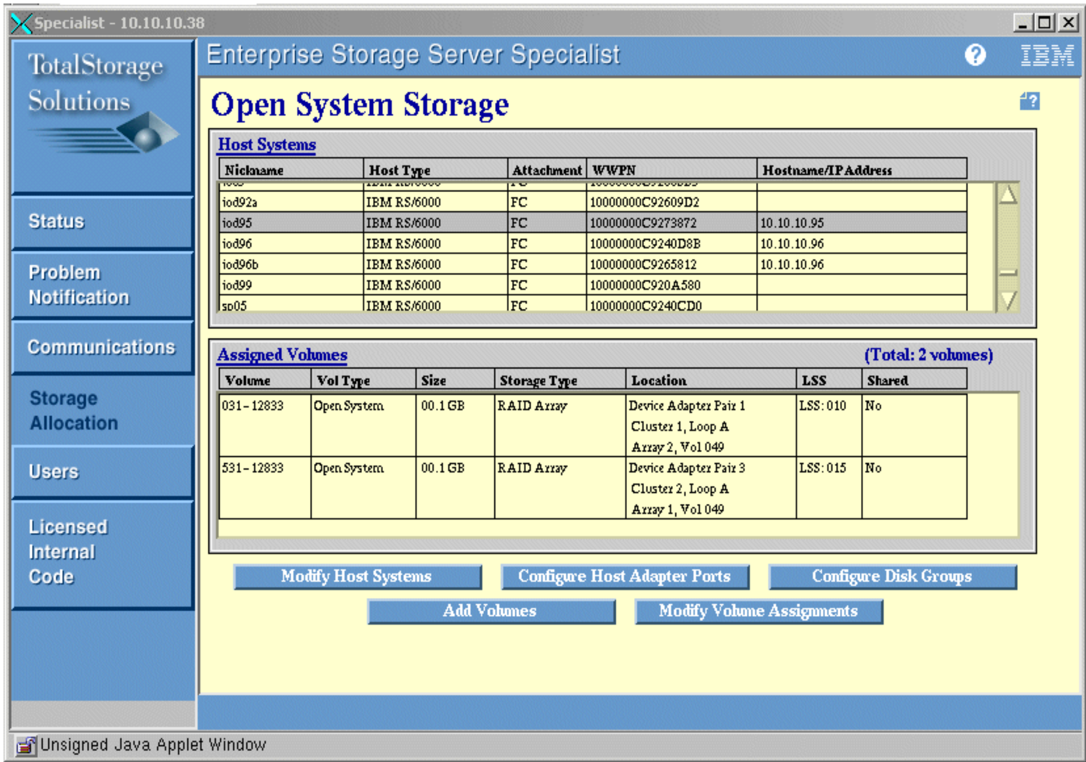


Figure 4-41 Tabular display of assigned storage to host

4.6.2 IBM 3590 Tape Systems

The tape devices used during the writing of this publication were IBM 3590-E11 Tape Systems. Thus, the main focus of this troubleshooting section will center on this particular tape system as a good example of fiber attached tape units. We acknowledge the variety of other tape products offered by IBM and other manufacturers that can be utilized in a SAN environment. In general, however, many of the concepts, methods, and data items that we describe can be implemented with other makes and models using commands specific for those units.

To further complicate the troubleshooting process, the 3590 can be configured as a standalone unit or as part of a tape library system. In the standalone configuration, the operator/CE panel is readily available. If the 3590 unit is a part of a tape library, you may not be able to utilize this panel display. To simplify our presentation of available methods and processes, we will use the standalone configuration of the 3590. The concepts and methods we use can be applied to the other variants of the 3590 as well.

3590 operator/CE panel available

One of the first checks is to verify connectivity from the tape to the SAN fabric. Figure 4-42 shows the layout of the operator/CE panel.

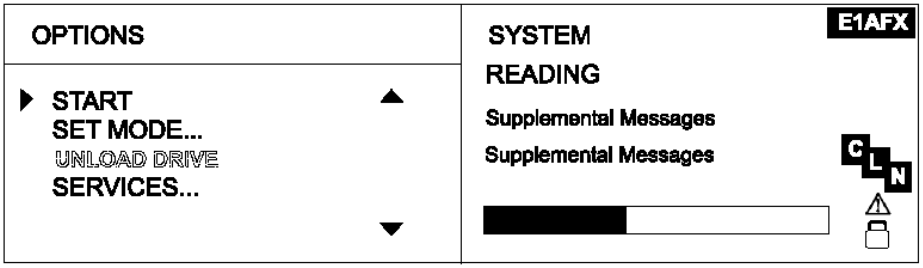


Figure 4-42 Layout of the 3590 operator/CE panel

The specific area to check is the two lines labeled Supplemental Message. The port address format is:

PORT0 ID: aa bb cc

where:

aa = fiber domain address or switch ID (on some switches)

bb = fiber area address or switch port (on some switches)

cd = arbitrated loop physical address or AL_PA

The port status is determined by the settings of aa bb cd. The settings based on port status are shown in Table 4-6.

Table 4-6 Decoding supplemental message on LCD panel

aa bb cd displayed	Meaning
`-- -- --`	No connection. No light seen on port. Soft addressing of the AL_PA.
`-- -- cd`	No connection. No light seen on port. Hard addressing of the AL_PA.

aa bb cd displayed	Meaning
?? ?? ??	Not communicating or did not establish a connection. Port does detect light. Soft addressing of the AL_PA.
?? ?? cd	Not communicating or did not establish a connection. Port does detect light. Hard addressing of the AL_PA.
cd CONFLICT	AL_PA is being used by another device on the fiber loop.
cd OFFLINE	The AL_PA is bypassed on the fiber loop.
02 14 26	Communications established.

If problems are indicated for the physical connection from the tape unit to the SAN fabric, the following items need to be investigated further:

- Ensure the tape drive power is ON.
- Ensure the drive serial number is the same as the drive serial number being used by the host application.
- Ensure the Operator Fibre Channel Address menu is set correctly.
- Check the fiber cable and plug connections from the tape unit to the fabric device. If a patch panel is in the run, those connections should be checked as well.
- Check that the maximum length of the cable does not exceed product specifications for the tape unit and the device to which the tape is attached. Use the appropriate product service guide for this step.

If the following items did not find the problem and the tape unit's port is still not communicating, then advanced diagnostic techniques are probably needed. Due to the strong possibility of excessive impact on the SAN environment, we only list further diagnostics but do not discuss them in detail. If further details about these tests are desired, then refer to the *3590 Tape Subsystem MI Models B11, B1A, E11, & E1A,* SA 37-0301.

Some of the possible advanced diagnostics are:

- Fibre Channel Wrap Test: This test is disruptive to the SAN environment and should only be implemented by qualified service personnel.
- Obtain the errors reported by the drive from the host server. For pSeries servers, run the **errpt** command with the appropriate flags.
- If there is partial connectivity to the SAN to the host, run the **tapeutil** command on the pSeries for additional error information. Refer to Section 4.4.2, “Remote device connections verification” on page 141 for more information about this test.

- Capture port error logs, if available, from the device to which the tape unit is attached. For more information, refer to the appropriate product's User or service guide.
- Run a device driver trace to capture more information. For example, if you are using the AIX tape device driver, issue the **atrc** command to start a trace.

  At the next failure, collect a 3590 microcode dump.

4.6.3 EMC Symmetrix

Other disk systems used during the writing of this publication were the EMC Symmetrix 8230 and the Symmetrix 8530. Due to time constraints, we had limited access to management utilities that can provide important data and information about the Symmetrix system and its links to the SAN. The principle management tool is the EMC ControlCenter. The ControlCenter is a series of integrated application modules that are available to users of EMC products. The EMC ControlCenter Console displays topology information, and provides alert management and a number of data representations.

From the topology view, clicking on a Symmetrix device causes the Symmetrix Manager to launch. The Symmetrix Manager application displays detailed fabric-specific and path-specific views. The application makes extensive use of the FibreAlliance MIB (V1.5 and V2.2). It actively polls the Symmetrix storage units for data on a regular, user-defined, time period.

Due to our limited exposure to the EMC management products, we cannot adequately provide troubleshooting guidance for these storage systems. We choose not to present incomplete and/or inaccurate techniques. More information about these applications and modules can be found at:

http://www.emc.com/

Click the Enterprise Storage Management Solutions link.

4.7 Getting further assistance

This section gives guidance when seeking further assistance. At this point, you have followed the previous maps and troubleshooting procedures without determining the cause of the problem. However, there is a large amount of data that you captured from the problem determination actions. Now the main question is: What is the next course of action?

There are online technical support resources from many product makers. Some of them offer search engines that can use symptoms as a criteria. Some have forums where other users and/or technical support groups can offer suggestions and ideas. The variety of offerings from these resources is as wide as the number of Web sites. There are several Web sites from IBM. For fabric and storage related issues, the IBM storage technical support Web site is at:

http://www.storage.ibm.com/techsup.htm

IBM pSeries server support can be found at:

http://techsupport.services.ibm.com/server/support?view=pSeries

If the various Web sites are not able to help with additional tips or fixes, then the next action is to engage remote technical support. We do not describe exactly who to contact, since that depends on many factors, such as world location and equipment vendor, to name a couple. The key principle here is that the problem is better defined and the scope is hopefully more narrowly defined. With these considerations, you can determine which support group is best suited to readily offer effective assistance.

When contacting a remote technical support group, there are a number of items you should have ready to provide. Some of these items are:

- SAN diagram(s)
- Code levels of affected devices
- Diagnostic data from problem-determination process to date, including error logs
- Description of problem

This list of information to have on-hand is not comprehensive, by far, since each support organization will have variations of desired data.

A

Fibre Channel boot

Booting an AIX system from a Fibre Channel-attached disk is now supported. Even though the procedure to implement this is not complicated, there are some tips you need to know. In this appendix the procedure to install an AIX operating system onto a Fibre Channel-attached disk is described briefly.

This appendix is organized into the following sequence:

- Limitations and restrictions
- Checking prerequisites
- Preparing ESS for Fibre Channel boot
- Installation procedure
- Update system microcode
- Update Fibre Channel adapter microcode
- Install AIX on Fibre Channel disk
- Gotchas

Limitations and restrictions

- A SAN storage device has a number of disks connected to a server. The fact that there are a huge number of available disks can cause a problem, or at least be a nuisance, because having too many disks makes the system administrator select the one to install the AIX to. To ease this problem, the new AIX installation menu provides the option of separating the disks by adapters. Another tip that can be used for avoiding the problem is to create a zone that contains only disks required for installing the operating system. Thereby, the system administrator can reduce the risk of loading the operating system onto a wrong disk.
- As for the INRANGE switch, the auto-loop-detect option should be disabled on the port that is connected to the Fibre Channel adapter in the pSeries machine. If the option is not disabled, the Fibre Channel adapter will fail to bring up the link.
- Booting from a Fibre Channel-attached tape is *not* supported.

Checking prerequisites

The following are required for a Fibre Channel boot:

- Proper level of system firmware
- Proper level of microcode for Fibre Channel adapters
- Proper level of operating system

Check the level of each of the above items with the following procedure.

Check the system firmware level. The method used for checking the system firmware level varies model by model. The best place to look up related information is `http://www.rs6000.ibm.com/support/micro/download.html`. The following command is provided as an example applicable to some pSeries models:

```
# lscfg -vp|grep -F .CL

        ROM Level (alterable).......CL010507
```

Check the Fibre Channel adapter microcode level.

The IBM Gigabit Fibre Channel adapter is available in two kinds, the 32-bit adapter (FC #6227) and 64-bit adapter (FC #6228).

The following command can be applied to both types in order to check the level of the installed microcode:

For the 32-bit adapter:

```
# lscfg -vl fcs1 |grep Z9

        Device Specific.(Z9)........SS3.22A0
```

For the 64-bit adapter:

```
# lscfg -vl fcs0 |grep Z9

        Device Specific.(Z9)........CS3.81A1
```

Check the AIX maintenance level.

The Fibre Channel boot is not supported in the older releases of the AIX. You should check the “Release Note” of your AIX to verify whether your release is capable of supporting the Fibre Channel boot.

On an AIX 4.3.3 system, messages similar to the following will appear:

```
# instfix -i |grep 4330
       All filesets for 4330-02_AIX_ML were found.
       All filesets for 4330-03_AIX_ML were found.
       All filesets for 4330-04_AIX_ML were found.
       All filesets for 4330-05_AIX_ML were found.
       All filesets for 4330-01_AIX_ML were found.
       All filesets for 4330-06_AIX_ML were found.
       All filesets for 4330-07_AIX_ML were found.
       All filesets for 4330-08_AIX_ML were found.
```

On an 5.1 system, messages similar to the following will appear.

```
# instfix -i|grep ML
       All filesets for 5.0.0.0_AIX_ML were found.
       All filesets for 5.1.0.0_AIX_ML were found.
       All filesets for 5.1.0.0_AIX_ML were found.
```

Preparing ESS for Fibre Channel boot

We recommend that you assign one and only one volume of the ESS to the AIX host that will boot from a Fibre Channel-attached disk. The reason is to prevent too many ESS volumes from being shown on the AIX installation menu. If too many volumes are presented on the menu, the system administrator may get confused and overwrite a volume that has important data.

Use the procedure described in Section 3.1.2, “Volume assignment using StorWatch Specialist” on page 71 to assign a volume to the AIX host. The following picture (Figure A-1 on page 223) shows that one volume has been allocated to our test system iod70b. Take note of the LUN ID of the volume. The LUN ID will be referred to later to identify the disk.

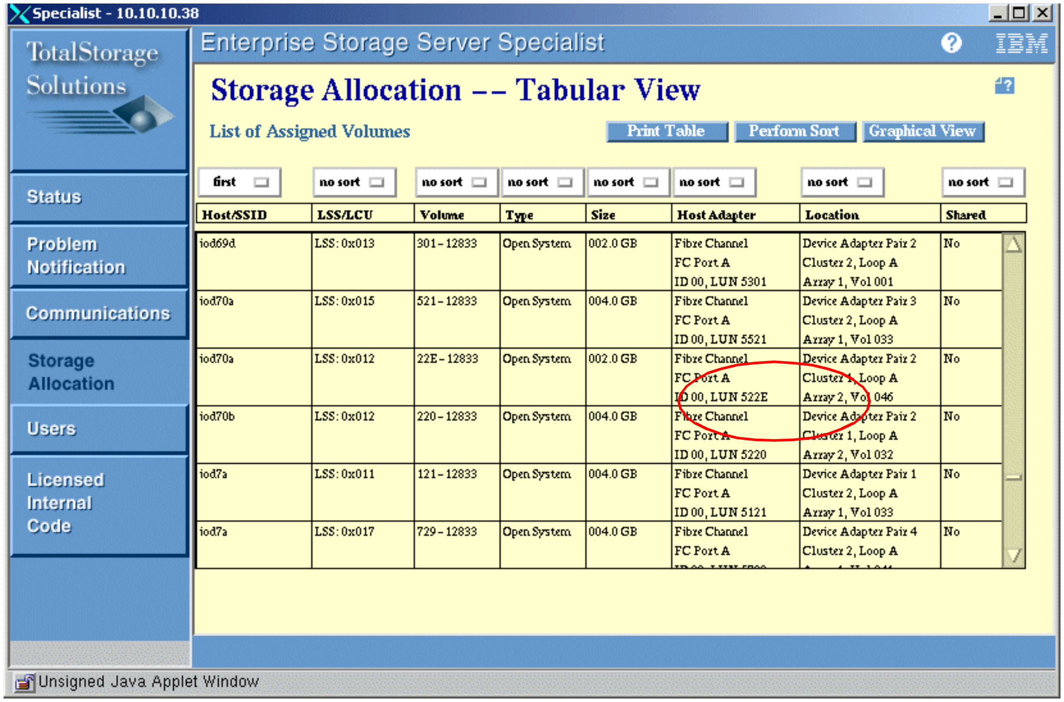


Figure A-1 ESS Specialist screen showing LUN_ID

The following steps are not mandatory, but are recommended. By doing these steps, you can check whether the Fibre Channel device can be successfully accessed.

```
# cfgmgr -v
# lsdev -Ccdisk
hdisk0 Available 11-08-00-2,0 16 Bit LVD SCSI Disk Drive
hdisk1 Available 11-08-00-4,0 16 Bit LVD SCSI Disk Drive
hdisk2 Available 37-08-01     Other FC SCSI Disk Drive
```

Note: Note that the ESS disk is shown as Other FC SCSI Disk. This is because the test system did not add the proper ODM entries for the ESS. We have configured the disk only for test purposes. You do not need to update the ODM at this step. We will do so later.

You can also check the LUN ID of the allocated volume with the following command:

```
# lsattr -Elhdisk2 | grep lun_id
lun_id        0x5220000000000000               N/A True
```

Installation procedure

Here we will describe the procedure fo rinstalling AIX on a Fibre Channel attached disk.

Update system microcode

Download the required system microcode from IBM's support site. At the time of writing this book, the latest copies of the system microcode for the pSeries and the RS/6000 were available on http://www.rs6000.ibm.com/support/micro/download.html. This Web site also provides the installation instruction for the system microcode. Since the procedure used to update the system microcode varies model by model, we strongly recommend that you to confer with the information provided on the Web site. We will not describe the update procedure in detail here. After the update, check the level of the system microcode with the following command. Note that the latest system microcode that supported the Fibre Channel boot for our test system at the time of writing this redbook was CL010913.

```
# lscfg -vp|grep -F .CL

        ROM Level (alterable).......CL010913
```

Update Fibre Channel adapter microcode

The latest microcodes for Fibre Channel adapters can also be obtained on the same Web site. The instructions for updating the adapter microcodes are also provided on the same Web site.

The following instructions tell you the procedure used to update the microcode of Gigabit Fibre Channel adapters.

1. Log in as root.
2. Please ensure there are no active I/Os, file systems, or volume groups using the adapter. Varyoff all volume groups associated with the adapter.
3. Download the device driver to the /tmp directory.
4. Enter the **# chmod 777 df1000f9.bin** command.
5. Execute the file.
6. It will prompt you for the unzip password. The password can be obtained by signing the license agreement form that is available on `http://www.rs6000.ibm.com/support/micro/downproc.html`. After you sign the agreement you will receive a password that authorizes you to unzip and install the microcode update.
7. The self-extracting zip file will unzip the Readme and the microcode files into the /etc/microcode directory under the current directory. Copy the extracted files into /etc/microcode.
8. You are now ready to FLASH the EEPROM in the adapter using the following command:

   ```
   # diag -c -d fcsX -T "download -s /etc/microcode -f -l latest"
   ```

   Where X is the number found from the **lsdev -C | grep fcs** command.
9. Varyon the volume groups associated with the adapter.
10. Repeat the above command for all of the cards that need the update.

Use the following command to check whether the update procedure was successful.

```
# lscfg -vl fcs1 |grep Z9

        Device Specific.(Z9)........SS3.22A1
# lscfg -vl fcs0 |grep Z9

        Device Specific.(Z9)........CS3.82A1
```

Note that the latest adapter microcodes that supported the Fibre Channel boot at the time of writing the redbook are:

SS3.22A1 for 32-bit adapter

CS3.82A1 for 64-bit adapter

Install AIX on Fibre Channel disk

Turn the system unit off and wait for 30 seconds, then turn the system unit on again. Insert a CD that has a bootable image of AIX 4.3.3 Maintenance Level 9 or higher to the CD-ROM drive. You need to change system's bootlist to make the system boot from the CD. The way to change the bootlist varies model by model. In most PCI-based RS/6000 models and pSeries models, this can be done by using the System Management Services (SMS) menu. Refer to the user's guide for your model. The menu presented by SMS may also vary. The screens shown in the following sections were taken from F80.

The following screen is the first page of the SMS on F80. Select 6. MultiBoot to change the primary boot device to the CD-ROM.

```
Version M2P01113
(c) Copyright IBM Corp. 2000  All rights reserved.
----------------------------------------------------------------------------
Utilities

1  Password Utilities
2  Display Error Log
3  Remote Initial Program Load Setup
4  SCSI Utilities
5  Select Console
6  MultiBoot
7  Select Language
8  OK Prompt

                                                                    .------.
                                                                    |X=Exit|
                                                                    `------'

===>6
```

Figure A-2 SMS menu for pSeries machines

In the MultiBoot menu, select 4. Select Boot Devices.

```
Version M2P01113
(c) Copyright IBM Corp. 2000  All rights reserved.
----------------------------------------------------------------------------
Multiboot

1  Select Software
2  Software Default
3  Select Install Device
4  Select Boot Devices
5  OK Prompt
6  Multiboot Startup <OFF>

                                                                   .------.
                                                                   |X=Exit|
                                                                   `------'

===>4
```

Figure A-3 MultiBoot menu

Select this option to view and change the custom boot list, which is the sequence of devices read at startup time. The boot list can contain up to five devices. Enter 3 to define the first boot device.

```
Version M2P01113
(c) Copyright IBM Corp. 2000  All rights reserved.
----------------------------------------------------------------------------
Select Boot Devices

1  Display Current Settings
2  Restore Default Settings
3  Configure 1st Boot Device
4  Configure 2nd Boot Device
5  Configure 3rd Boot Device
6  Configure 4th Boot Device
7  Configure 5th Boot Device

                                                                   .------.
                                                                   |X=Exit|
                                                                   `------'
===>3
```

Figure A-4 Select boot devices

In the following screen, enter the number for the CD-ROM device and press Enter.

```
Configure 1st Boot Device

Device  Current  Device
Number  Position  Name
1        -        Diskette
2        -        SCSI Tape ( loc=P1/Z1-A0 )
3        -        SCSI CD-ROM ( loc=P1/Z1-A1 )
4        -        SCSI 9100 MB Harddisk ( loc=P1/Z1-A2 )
5        -        Token-Ring  ( loc=P1-I3/T1 )
6        -        Ethernet  ( loc=P1/E1 )
7        -        Ethernet  ( loc=P1-I9/E1 )
8                 None

                                                                     .------.
                                                                     |X=Exit|
                                                                     `------'
===>3
```

Figure A-5 Configure 1st boot device

Now the CD-ROM has become the first boot device. Make sure that you have inserted the AIX Maintenance Level 9, or higher, CD in the CD-ROM device.

```
Version M2P01113
(c) Copyright IBM Corp. 2000  All rights reserved.
-----------------------------------------------------------------------------
Current Boot Sequence

1  SCSI CD-ROM ( loc=P1/Z1-A1 )
2    None
3    None
4    None
5    None

                                                                     .------.
                                                                     |X=Exit|
                                                                     `------'
```

Figure A-6 Current boot sequence

Keep entering x a number of times until you see the following menu. Enter x again and the system will begin to boot from the boot device.

```
Version M2P01113
(c) Copyright IBM Corp. 2000  All rights reserved.
----------------------------------------------------------------------------
Utilities

1  Password Utilities
2  Display Error Log
3  Remote Initial Program Load Setup
4  SCSI Utilities
5  Select Console
6  MultiBoot
7  Select Language
8  OK Prompt

                                                                     .------.
                                                                     |X=Exit|
                                                                     `------'

===>x
```

Figure A-7 Rebooting from SMS menu

The three-digit LED should display `c31` after several minutes. If you have more than one console, each terminal and direct-attached display device (or console) may display a screen that directs you to press a key to identify your system console. A different key is specified for each terminal displaying this screen. If this screen is displayed, then press the specified key on the device to be used as the system console.

```
 ******* Please define the System Console. *******

Type a 2 and press Enter to use this terminal as the
  system console.
Typ een 2 en druk op Enter om deze terminal als de
  systeemconsole to gebruiken.
Pour definir ce terminal comme console systeme, appuyez
  sur 2 puis sur Entree.
Taste 2 und anschliessend die Eingabetaste druecken, um
  diese Datenstation als Systemkonsole zu verwenden.
Premere il tasto 2 ed Invio per usare questo terminal
  come console.
Escriba 2 y pulse Intro para utilizar esta terminal como
  consola del sistema.
Escriviu 1 2 i premeu Intro per utilitzar aquest
  terminal com a consola del sistema.
Digite um 2 e pressione Enter para utilizar este terminal
  como console do sistema.
```

Figure A-8 Define system console

A screen is displayed that prompts you to select an installation language. Follow the directions on this screen to select the language in which the installation instructions will be displayed.

```
>>>  1 Type 1 and press Enter to have English during install.
     2 Entreu 2 i premeu Intro per veure la instal·lació en català.
     3 Entrez 3 pour effectuer l'installation en français.
     4 Für Installation in deutscher Sprache 4 eingeben
        und die Eingabetaste drücken.
     5 Immettere 5 e premere Invio per l'installazione in Italiano.
     6 Digite 6 e pressione Enter para usar Português na instalação.
     7 Escriba 7 y pulse Intro para usar
        el idioma español durante la instalación.
     8 Skriv 8 och tryck ned Enter = Svenska vid installationen.



    88  Help ?


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

Figure A-9 Define language

The Welcome to the Base Operating System Installation and Maintenance screen is displayed. You need to change the installation and system settings that have been set for this machine in order to select a Fibre Channel-attached disk as a target disk. Type `2` 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



    88  Help ?
    99  Previous Menu

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

Figure A-10 BOS Installation menu

The following screen appears. You will see `Preservation Install` on the screen. This is the default setting for a machine installed with AIX 4.3. Since we are installing the operating system on a new disk we do not have anything to preserve. Enter `1` to change the system settings.

```
 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  Install Trusted Computing Base...... No

>>> 0  Install AIX 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 A-11   Installation and Settings menu

The next menu allows you to change the method of installation. Since you are loading the operating system on a new Fibre Channel-attached disk, you do not have anything to preserve. Enter `1` to select New and Complete Overwrite (installation).

```
 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 A-12 Change Method of Installation menu

Next you will be presented with the Change (the destination) Disk screen. This allows you to change the hard disk where BOS will be installed. The location codes of the hard disks are displayed in the Location Code column of the Change Disk(s) Where You Want to Install screen. You may wish to keep a record of the location code for the destination disk. In the future, you can use this location code to identify which disk contains the root volume group in order to do system maintenance.

Type the number, but *do not* press Enter, for each disk you choose. Typing the number of a selected disk will deselect the device. You need to deselect hdisk0, otherwise the AIX will be loaded onto the hdisk0 again.

```
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

>>>  1  hdisk0   11-08-00-2,0      8678   rootvg         Yes
     2  hdisk1   11-08-00-4,0      8678   none           Yes
     3  hdisk2   37-08-01          3814   none           Yes


>>>  0   Continue with choices indicated above

    66  Disks not known to Base Operating System Installation
    77  Display More Disk Information
    88  Help ?
    99  Previous Menu

>>> Choice [0]:1
```

Figure A-13 Change Disk menu

Now select the Fibre Channel-attached disk. It is shown in the following screen that hdisk2 is the only Fibre Channel-attached disk. Remember 37-08-01 was he location code of the Fibre Channel adapter on our test system. You can select hdisk2 now, but look into more information to be sure. You will find the following steps useful when you have multiple Fibre Channel disks. Enter `77` to display more detailed disk information.

```
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

     1  hdisk0   11-08-00-2,0      8678   rootvg          Yes
     2  hdisk1   11-08-00-4,0      8678   none            Yes
     3  hdisk2   37-08-01          3814   none            Yes


>>>  0   Continue with choices indicated above

    66  Disks not known to Base Operating System Installation
    77  Display More Disk Information
    88  Help ?
    99  Previous Menu

>>> Choice [0]:77
```

Figure A-14 Deselect hdisk0

In the next screen, PVIDs are displayed. Enter `77` to display more information.

```
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   0000cbcfefb65e28
     2  hdisk1   0000000000000000
     3  hdisk2   00002328a2155c12


>>>  0   Continue with choices indicated above

    66  Disks not known to Base Operating System Installation
    77  Display More Disk Information
    88  Help ?
    99  Previous Menu

>>> Choice [0]:77
```

Figure A-15 Display more disk information

Finally the system shows us the WWPN and LUN_ID information for Fibre Channel-attached disks. Check the LUN_ID information to see if its values matches with those you previously obtained from the ESS Specialist, then enter the number representing the disk. This action will set the disk as the target disk.

```
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 World Wide Port Name//Lun ID         SCSI ID
     1  hdisk0    scsi0//2,0
     2  hdisk1    scsi0//4,0
     3  hdisk2    0x10000000c92123f5//0x5220000000000000           0x261013

>>>  0   Continue with choices indicated above

    66  Disks not known to Base Operating System Installation
    77  Display More Disk Information
    88  Help ?
    99  Previous Menu

>>> Choice [0]:2
```

Figure A-16 Display WWPN and LUN_ID

After you have selected Fibre Channel-attached disks, the Installation and Settings screen is displayed with the selected disks. Verify the installation settings in the following screen. If everything looks OK, type 0 and press Enter. Then the Installation process begins.

Attention: It is extremely important to select the correct root volume group because the existing data in the destination root volume group will be destroyed during BOS installation.

```
 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.....hdisk2

    2  Primary Language Environment Settings (AFTER Install):
         Cultural Convention................English (United States)
         Language ..........................English (United States)
         Keyboard ..........................English (United States)
         Keyboard Type......................Default

    3  Install Trusted Computing Base...... No

>>> 0  Install AIX 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 hdisk4.
>>> Choice [0]:0
```

Figure A-17   Start installation

The following screen appears and will show the progress of the installation process. Wait until the installation is complete. The system will automatically reboot.

```
Installing Base Operating System

If you used the system key to select SERVICE mode,
turn the system key to the NORMAL position any time before the
installation ends.

        Please wait...


Approximate     Elapsed time
     % tasks complete   (in minutes)


          7             0      Restoring base operating system
```

Figure A-18 During installation

When your system reboots, the following message will appear on the screen for a moment. The message tells you the address of the device from which the system is reading the boot image. As in the following example, the system is booting from the Fibre Channel-attached disk with LUN_ID 5220.

```
----------------------------------------------------------------------------
                                Welcome to AIX.
                       boot image timestamp: 19:56 09/26
                 The current time and date: 20:01:16 09/26/2001
                number of processors: 6    size of memory: 3072Mb
boot device:
/pci@fff7f0a000/pci@c,4/fibre-channel@1/disk@10000000c92123f5,52200000000000
00:2
closing stdin and stdout...
----------------------------------------------------------------------------
```

Figure A-19 Booting from a Fibre Channel-attached disk

Upon the completion of the rebooting process, log on to the system as root. Run the following commands to check if everything is OK.

```
# lspv
hdisk0        0000cbcfefb65e28    None
hdisk1        none                None
hdisk2        00002328a2155c12    rootvg
```

```
# lsdev -Ccdisk
hdisk0 Available 11-08-00-2,0 16 Bit LVD SCSI Disk Drive
hdisk1 Available 11-08-00-4,0 16 Bit LVD SCSI Disk Drive
hdisk2 Available 37-08-01     Other FC SCSI Disk Drive
```

Note that hdisk2 appears as Other FC SCSI Disk Drive. This is normal behavior because the operating system does not have proper ODM entries for the ESS. The ODM entries are not automatically updated during the initial installation process. It has to be manually updated by system administrator. To update the ODM entries, install the ibm2105.rte fileset. Rebooting the machine is also required.

Upon completion of rebooting, log on to the machine as root user and run the following command. The Fibre Channel-attached disk should appear to be an ESS device.

```
# lsdev -Ccdisk
hdisk0 Available 11-08-00-2,0 16 Bit LVD SCSI Disk Drive
hdisk1 Available 11-08-00-4,0 16 Bit LVD SCSI Disk Drive
hdisk2 Available 37-08-01     IBM FC 2105F20
```

Surprise problems

When the pSeries machine has been connected to tens of Fibre Channel-attached disks, the time required for bringing up SMS can take much longer than the time required when the machine boots from internal SCSI disks. The reason is that the system searches for each FC disk to recognize it. On the average, it takes 10-15 seconds per disk. There is no way to remedy this.

```
                        .-----------------.
                        | PLEASE WAIT.... |
                        `-----------------'


/pci@fff7f09000/pci@b,4/fibre-channel@1/disk@10000000c924635f,56700000000000
0
```

Figure A-20   A sample screen of boot dialog

If you do not update the microcode of the Gigabit Fibre Channel adapter to the latest level, the Fibre Channel-attached disk may appear as a non-bootable disk.

The system may not boot after replacing a Gigabit Fibre Channel adapter. In most cases, it is because the replaced adapter has a down-level microcode. Check the level of the microcode and update it.

If the installation menu in SMS does not show any Fibre Channel attached disks, check the following items:

- Level of the microcode of the Fibre Channel adapters
- Level of the system firmware
- Zone definitions

If you have taken one of the following actions, then go to either the SMS menu or the AIX maintenance menu to fix the problems.

- You have changed LUD ID on the SAN-attached storage.
- You have changed the WWPN of the boot disk.

In the cases of any of the following actions, no particular follow-up action is required:

- You have replaced a Fibre Channel adapter on the pSeries machine.
- You have connected the pSeries machine to a different switch port.
- You have connected the SAN-attached storage to a different switch port.

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this redbook.

IBM Redbooks

For information on ordering these publications, see "How to get IBM Redbooks" on page 230.

- *Designing an IBM Storage Area Nework,* SG24-5758
- *IBM SAN Survival Guide,* SG24-6143
- *IBM SAN Survival Guide Featuring the IBM 2109,* SG24-6127
- *IBM SAN Survival Guide Featuring the INRANGE Portfolio,* SG24-6150
- *IBM SAN Survival Guide Featuring the McDATA Portfolio,* SG24-6149
- *Implementing ESS Copy Services on UNIX and Windows NT/2000,* SG24-5757

Other resources

These publications are also relevant as further information sources:

- *3590 Tape Subsystem MI Models B11, B1A, E11, & E1A,* SA 37-0301
- *7140 TotalStorage SAN Control Model 160 Router Installation and User's Guide,* GC26-7433
- *FC/9000 Enterprise Manager Software Installation and Operation Manual,* 9110509-203
- *FC/9000 Fibre Channel Switch - Maintenance Manual*, DCN 9110774-307
- *FC/9000 Fibre Channel Switch Planning Guide,* 9106677-101
- *IBM 2105 Enterprise Storage Server Systems Assurance Product Review (SAPR) Guide,* SA99-004
- *IBM Fibre Channel Integration & Planning: User's Guide and Service Information,* SC23-4329
- *IBM SAN Fibre Channel Switch 2109 Model S16 Installation and Service Guide,* SC26-7352
- *IBM SAN Fibre Channel Switch 2109 Model S16 User's Guide,* SC26-7351

- *IBM SDD TotalStorage ESS System Device Driver User's Guide,* GC26-7442
- *IBM Subsystem Device Driver Installation,* GC26-7442
- *McDATA Director Planning Manual,* 620-000106
- *McDATA ED-6064 Director Installation and Service Manual,* P/N 620-000108-200

Referenced Web sites

These Web sites are also relevant as further information sources:

- 3590 Attachment:

  `http://www.storage.ibm.com/hardsoft/tape/3590/3590attach`
- Adapter microcode:

  `http://www.rs6000.ibm.com/support/micro/download.html#adapter`
- AIX Atape Driver Levels:

  `ftp://ftp.software.ibm.com/storage/devdrvr/AIX`
- AIX ODM definition files for EMC Symmetrix:

  `ftp://ftp.emc.com/pub/symm3000/aix/`
- AIX service FTP site:

  `ftp://service.software.ibm.com/aix/fixes/v4/devices/`
- APAR IY20367 from FIXDIST:

  `http://techsupport.services.ibm.com/rs6k/fixes.html`
- Disk Storage Systems:

  `http://www.storage.ibm.com/hardsoft/disk/products.htm`
- EMC home page:

  `http://www.emc.com/`
- EMC Symmetrix interoperability for AIX:

  `www.emc.com/horizontal/interoperability/matrices/EMCSupportMatrix.pdf`
- IBM @server pSeries & RS/6000 microcode download procedure:

  `http://www.rs6000.ibm.com/support/micro/downproc.html`
- IBM @server pSeries & RS/6000 microcode updates:

  `http://www.rs6000.ibm.com/support/micro/download.html`
- IBM Fiber Transport Services

  `http://www-1.ibm.com/services/its/us/fts.html`

- IBM LTO support:

  http://www.storage.ibm.com/hardsoft/tape/3583/3583opn.html

- IBM McDATA Director and switches:

  http://www.storage.ibm.com/ibmsan/products/2032/prod_data/supserver-064.htm
l

- IBM pSeries server support:

  http://techsupport.services.ibm.com/server/support?view=pSeries

- IBM SAN FC switches:

  http://www.storage.ibm.com/ibmsan/products/2109/supserver.htm

- IBM TotalStorage FAStT500 Storage Server:

  http://www.ibm.com/storage/fast500

- IBM TotalStorage SAN Controller 160:

  http://www.storage.ibm.com/hardsoft/products/san160/san160.html

- IBM TotalStorage SAN Controller 160 Supported servers:

  http://www.storage.ibm.com/hardsoft/products/san160/160_support_matrix.html

- IBM TotalStorage SAN switches for Core-to-Edge SAN Solutions:

  http://www.ibm.com/storage/fcswitch

- IBM TotalStorage SAN switch S08 and S16:

  http://www.storage.ibm.com/ibmsan/products/2109/index.html

- IBM zTechnical Library:

  http://www.EMC.com/techlib/abstract.jsp?id=65

- INRANGE switches:

  http://www.storage.ibm.com/ibmsan/products/directors/prod_data/supserver-04
2.html

- Java Plug-In:

  http://java.sun.com/products/

- McDATA ED-5000 IBM supported/planned Storage and server matrix:

  http://www.mcdata.com/IBM/suptdev/ed5000.html

- Network attached storage details:

  http://www.storage.ibm.com/snetwork/nas/index.html

- PTF packages:

  ftp://ftp.emc.com/pub/symm3000/aix/EMC-FC-Kit

- SAN component compatability:

  http://www.storage.ibm.com/ibmsan/products/sanfabric.htm
- SDD technical support:

  http://ssddom01.storage.ibm.com/techsup/swtechsup.nsf/support/sddupdates/
- Storage Networking Industry Association (SNIA):

  http://www.snia.org/
- Storage technical support:

  http://www.storage.ibm.com/techsup.htm
- Subsystem Device Driver documentation:

  http://ssddom01.storage.ibm.com/techsup/swtechsup.nsf/support/sddupdates
- Supported HBAs:

  http://www-1.ibm.com/servers/eserver/pseries/library/hardware_docs/sa38/380516.pdf
- Supported server list for ESS:

  http://www.storage.ibm.com/hardsoft/products/ess/supserver.htm
- Support matrix:

  ttp://www.storage.ibm.com/hardsoft/products/sangateway/supserver.htm
- Support matrix for the SDD pSeries servers running HACMP/6000:

  http://ftp.software.ibm.com/storage/subsystem/tools/f2asdd00.htm#HDRINSTALL
- Tape System support details:

  http://www.storage.ibm.com/hardsoft/tape/index.html

How to get IBM Redbooks

Search for additional Redbooks or Redpieces, or to view, download, or order a hardcopy from the Redbooks Web site, go to:

ibm.com/redbooks

Also download additional materials (code samples or diskette/CD-ROM images) from this Redbooks site.

Redpieces are Redbooks in progress; not all Redbooks become Redpieces and sometimes just a few chapters will be published this way. The intent is to get the information out much quicker than the formal publishing process allows.

IBM Redbooks collections

Redbooks are also available on CD-ROMs. Click the CD-ROMs button on the Redbooks Web site for information about all the CD-ROMs offered, as well as updates and formats.

Special notices

References in this publication to IBM products, programs or services do not imply that IBM intends to make these available in all countries in which IBM operates. Any reference to an IBM product, program, or service is not intended to state or imply that only IBM's product, program, or service may be used. Any functionally equivalent program that does not infringe any of IBM's intellectual property rights may be used instead of the IBM product, program or service.

Information in this book was developed in conjunction with use of the equipment specified, and is limited in application to those specific hardware and software products and levels.

IBM may have patents or pending patent applications covering subject matter 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 the IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785.

Licensees of this program who wish to have information about it for the purpose of enabling: (i) the exchange of information between independently created programs and other programs (including this one) and (ii) the mutual use of the information which has been exchanged, should contact IBM Corporation, Dept. 600A, Mail Drop 1329, Somers, NY 10589 USA.

Such information may be available, subject to appropriate terms and conditions, including in some cases, payment of a fee.

The information contained in this document has not been submitted to any formal IBM test and is distributed AS IS. The use of this information or the implementation of any of these techniques is a customer responsibility and depends on the customer's ability to evaluate and integrate them into the customer's operational environment. While each item may have been reviewed by IBM for accuracy in a specific situation, there is no guarantee that the same or similar results will be obtained elsewhere. Customers attempting to adapt these techniques to their own environments do so at their own risk.

Any pointers in this publication to external Web sites are provided for convenience only and do not in any manner serve as an endorsement of these Web sites.

The following terms are trademarks of other companies:

Tivoli, Manage. Anything. Anywhere.,The Power To Manage., Anything. Anywhere.,TME, NetView, Cross-Site, Tivoli Ready, Tivoli Certified, Planet Tivoli, and Tivoli Enterprise are trademarks or registered trademarks of Tivoli Systems Inc., an IBM company, in the United States, other countries, or both. In Denmark, Tivoli is a trademark licensed from Kjøbenhavns Sommer - Tivoli A/S.

C-bus is a trademark of Corollary, Inc. in the United States and/or other countries.

Java and all Java-based trademarks and logos are trademarks or registered trademarks of Sun Microsystems, Inc. in the United States and/or other countries.

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

PC Direct is a trademark of Ziff Communications Company in the United States and/or other countries and is used by IBM Corporation under license.

ActionMedia, LANDesk, MMX, Pentium and ProShare are trademarks of Intel Corporation in the United States and/or other countries.

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

SET, SET Secure Electronic Transaction, and the SET Logo are trademarks owned by SET Secure Electronic Transaction LLC.

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

Abbreviations and acronyms

ACF	Automated Cartridge Facility
ATM	Asynchronous transfer mode
CE	Customer Engineer
CLI	Command Line Interface
CLVM	Concurrent Logical Volume Manager
CPI	Common parts interconnect
CPU	Central processing unit
CRC	Cyclical Redundancy Checking
Cu	Copper
DA	Device adapter
dacs	Disk array controllers
dars	Disk array routers
DASD	Direct Access Storage Device
DBMS	Database management system
DDM	Disk drive module
DHCP	Dynamic Host Configuration Protocol
DNS	Domain Name System
DRAM	Dynamic random access memory
EFC	Enterprise Fabric Connectivity
EFC	Enterprise Fabric Connectivity
ESCON	Enterprise Systems Connection
ESN	Electronic Serial Number
ESS	Enterprise Storage Server
FAStT	Fibre Array Storage Technology
FC	Feature code
FC-AL	Fibre Channel Arbitrated Loop
FCP	Fibre Channel Protocol
FCS	Frame check sequence
FICON	Fiber Connectivity
FIO	Future I/O
FIO	Fibre Channel I/O module
FRU	Field replaceable unit
FTP	File Transfer Protocol
GB	Gigabytes
GBIC	Gigabit interface converter
GUI	Graphical User Interface
HA	Host adapter
HA1	High Availablility Unit
HBA	Host Bus Adapter
IBM	International Business Machines Corporation
IN-VSN	INRANGE Virtual Storage Network Enterprise Manager
ISL	Inter-Switch Links
ITSO	International Technical Support Organization
JBOD	Just a Bunch Of Disks
Km	Kilometer
LAN	Local area network
LCD	Liquid Crystal Display
LIC	Line interface coupler
LSS	Logical SubSystem
LTO	Linear Tape Open
LUN	Logical unit number
LVM	Logical Volume Manager
LWL	Long wavelength
m	Meter
MAC	Message Authentication Code
MMF	Multi-Mode Fiber
MTBF	Mean Time Between Failures

NVS	Nonvolatile storage
ODM	Object Data Manager
PCI	Peripheral component interconnect
PPRC	Peer-to Peer Remote Copy
PTF	Program temporary fix
PVID	physical volume identifier
RAID	Redundant array of independent disks
RPQ	Request for Price Quotation
RSH	Remote shell daemon
SAN	Storage area network
SC	Standard connector
SCSI	Small computer system interface
SDD	Subsystem Device Driver
SINA	Storage Networking Industry Association
SMF	Single-Mode Fiber
SMP	Symmetric multiprocessor
SMS	System Management Services
SNIA	Storage Networking Industry Association
SNMP	Simple Network Management Protocol
SNS	Simple Name Server
SSA	Serial Storage Architecture
SSR	Service Support Representative
SWL	Short wavelength
TAR	Tape Archive
TSM	Tivoli Storage Manager
TSNM	Trivoli Storage Network Manager
VG	Volume group
VGID	Volume group identifier
VTS	Virtual Tape Server
WWN	World wide Name

Index

IBM

Redbooks

Practical Guide for SAN with pSeries

Practical Guide for SAN with pSeries

Redbooks

Hints and tips for pSeries and AIX

Comprehensive planning information

Structured problem determination guide

Practical Guide for SAN with pSeries provides a single source of the information you will need to build your own SAN fabric, with easy-to-follow instructions and ample illustrations. To allow quick access to the information readers may need, the contents of this redbook have been arranged in three parts:

- Planning
- Implementation
- Problem determination

With this handbook format, both planners and practitioners in SAN area will find it a useful addition to their technical library. This redbook was written with a specific aim.

A few SAN redbook were already published in the past but the information specific to IBM pSeries servers has been hard to find and scattered around. Now the authors complied most of the information that would be critical when you build a SAN fabric with IBM pSeries machines.

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

SG24-6050-00 ISBN 0738423386