IBM

SAN Multiprotocol Routing
An Introduction and Implementation

- Read about the basics of the IBM approach to multiprotocol routing
- Learn about the IBM products and solutions
- Understand how to install routers

Jon Tate
Steve Fodor
Jure Arzensek

Redbooks

ibm.com/redbooks

International Technical Support Organization

SAN Multiprotocol Routing: An Introduction and Implementation

November 2006

SG24-7321-00

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

First Edition (November 2006)

This edition applies to the IBM multiprotocol routing solutions that were current at the time of writing. We clearly state the supported version of software and hardware in the respective chapters.

Contents

Notices

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

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

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

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

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

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

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

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

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

COPYRIGHT LICENSE:

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

Trademarks

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

AIX 5L™
AIX®
BladeCenter®
DS4000™
DS8000™
Enterprise Storage Server®
Enterprise Systems Architecture/390®
ESCON®
eServer™
Everyplace®
FICON®
FlashCopy®
IBM TotalStorage Proven™
ibm.com®
IBM®
PowerPC®
PR/SM™
pSeries®
Redbooks (logo) ™
Redbooks™
S/360™
S/370™
S/390®
Storage Tank™
System p™
System Storage™
System x™
System/360™
System/370™
Tivoli®
TotalStorage Proven™
TotalStorage®
WebSphere®
xSeries®
z/Architecture™
z/OS®
zSeries®

The following terms are trademarks of other companies:

Java, JRE, Solaris, Sun, Sun Microsystems, Ultra, and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

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

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

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

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

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

Preface

The rapid spread and adoption of production storage area networks (SANs) has fuelled the need for multiprotocol routers. The routers provide improved scalability, security, and manageability by enabling devices in separate SAN fabrics to communicate *without* merging fabrics into a single, large SAN fabric. This capability enables clients to initially deploy separate SAN solutions at the departmental and data center levels. Then, clients can consolidate these separate solutions into large enterprise SAN solutions as their experience and requirements grow and change.

Alternatively, multiprotocol routers can help to connect existing enterprise SANs for a variety of reasons. For instance, the introduction of Small Computer System Interface over IP (iSCSI) provides for the connection of low-end, low-cost hosts to enterprise SANs. The use of an Internet Protocol (IP) in the Fibre Channel (FC) environment provides for resource consolidation and disaster recovery planning over long distances. And the use of FC-FC routing services provides connectivity between two or more fabrics without having to merge them into a single SAN.

This IBM® Redbook targets storage network administrators, system designers, architects, and IT professionals who sell, design, or administer SANs. It introduces you to the products, concepts, and technology in the IBM System Storage™ SAN Routing portfolio. This book shows the features of each product and examples of how you can deploy and use them.

Prior to reading this book, you must be familiar with SANs. If not, we recommend that you read the following IBM Redbooks™ before you start this one:

- *IBM TotalStorage: SAN Product, Design, and Optimization Guide*, SG24-6384
- *Introduction to Storage Area Networks*, SG24-5470

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, San Jose Center.

Jon Tate is a Project Manager for IBM TotalStorage® SAN Solutions at the International Technical Support Organization, San Jose Center. Before joining the ITSO in 1999, he worked in the IBM Technical Support Center, providing Level 2 support for IBM storage products. Jon has 20 years of experience in storage software and management, services, and support, and is both an IBM Certified IT Specialist and an IBM SAN Certified Specialist.

Steve Fodor is an IT Specialist in the Data Solutions, Information Technology Services team in Canada. He has 16 years experience in the IT field, spending the last eight years focused on the SAN storage world. His areas of expertise include open systems storage solutions with Intel®-based servers. Steve has coauthored a white paper about IBM storage and Microsoft® Exchange. He is also a SNIA Certified Professional.

Jure Arzensek is an Advisory IT Specialist at IBM Slovenia, and responsible for post-sales technical support for IBM System x™, Personal Computing Division, and TotalStorage DS4000™ products in the Central and Eastern Europe (CEE) region. He has six years of experience with IBM storage products and has worked for IBM since 1995. He holds a degree in computer science from the University of Ljubljana. His other areas of expertise include the System x servers and network operating systems for the Intel platform. He has coauthored four other IBM Redbooks.

This IBM Redbook uses material that was contained in previous Redbooks and the authors of those deserve special mention:

Steve Garraway
Jim Kelly
Andy McManus
Pauli Rämö
Leos Stehlik
Marcus Thordal

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

Tom Cady
Emma Jacobs
Leslie Parham
Deanna Polm
Sangam Racherla
Sokkieng Wang
International Technical Support Organization, San Jose Center

Lisa Dorr
IBM Storage Systems Group

Jim Baldyga
Silviano Gaona
Brian Steffler
Brocade Communications Systems

Hui Chen
Dan Hersey
John McKibben
Darshak Patel
Cisco Systems

Brent Anderson
Prasad Pammidimukkala
McDATA Corporation

Tom and Jenny Chang
Garden Inn Hotel, Los Gatos, California

Become a published author

Join us for a two- to six-week residency program! Help write an IBM Redbook dealing with specific products or solutions, while getting hands-on experience with leading-edge technologies. You'll have the opportunity to team with IBM technical professionals, Business Partners, and Clients.

Your efforts will help increase product acceptance and customer satisfaction. As a bonus, you'll develop a network of contacts in IBM development labs, and increase your productivity and marketability.

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

`ibm.com/redbooks/residencies.html`

Comments welcome

Your comments are important to us!

We want our 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 e-mail to:

  redbooks@us.ibm.com

- Mail your comments to:

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

1

SAN routing

SAN routing provides the following new tools to manage:

- Departmental isolation and resource sharing
- Technology migration and integration
- Remote replication of disk systems
- Remote access to disk and tape systems
- Low-cost connection to SANs

This chapter introduces the terminology, technologies, and the value propositions for SAN routing.

1.1 SAN routing definitions

It is important that you clearly understand the terms and principles of Fibre Channel (FC) routing before you learn about designing routed networks.

For an excellent introduction, refer to *Multiprotocol Routing for SANs* written by Josh Judd from Brocade. Another book that addresses the topic is *IP SANs: An Introduction to iSCSI, iFCP, and FCIP Protocols for Storage Area Network*, written by Tom Clark from McDATA. For details about locating these books, see "Related publications" on page 547.

1.1.1 Fibre Channel

Fibre Channel is a set of standards for a serial input/output (I/O) bus developed through industry cooperation. A Fibre Channel frame consists of a header, payload, and 32-bit CRC bracketed by start of frame (SOF) and end of frame (EOF) delimiters. The header contains the control information necessary to route frames between N_PORTs and manage exchanges and sequences.

It is beyond the scope of this book to cover Fibre Channel in any great depth. For further reading, we recommend:

- *Introduction to Storage Area Networks*, SG24-5470
- *IBM TotalStorage: SAN Product, Design, and Optimization Guide*, SG24-6384

Figure 1-1 shows the layout of a Fibre Channel frame.

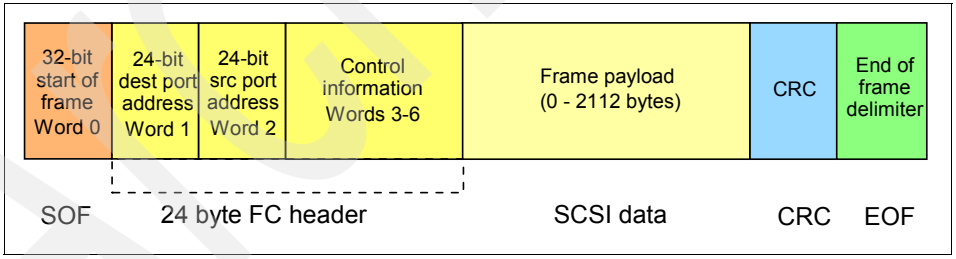


Figure 1-1 Fibre Channel frame structure

1.1.2 Fibre Channel switching

A Fibre Channel switch filters and forwards packets between Fibre Channel connections on the *same* fabric, but it cannot transmit packets between fabrics. As soon as you join two switches together, you merge the two fabrics into a single fabric with one set of fabric services, which then becomes a single point of failure.

Important: Fibre Channel switching cannot transfer packets between fabrics.

1.1.3 Fibre Channel routers

A router forwards data packets *between* two or more fabrics. Routers use headers and forwarding tables to determine the best path for forwarding the packets.

Separate fabrics each have their own addressing schemes. When they are joined by a router, there must be a way to translate the addresses between the two fabrics. This mechanism is called *network address translation* (NAT) and is inherent in the Cisco, Brocade, and McDATA multiprotocol switch/router products. It is sometimes referred to as FC-NAT to differentiate it from a similar mechanism that exists in IP routers.

Important: Fibre Channel routers forward packets between fabrics.

1.1.4 Tunneling

Tunneling is a technique that allows one network to send its data through another network's connections. Tunneling works by encapsulating a network protocol within packets carried by the second network. For example, in a Fibre Channel over Internet Protocol (FCIP) solution, Fibre Channel packets can be encapsulated inside IP packets. Tunneling raises issues of packet size, compression, out-of-order packet delivery, and congestion control.

1.1.5 Routers and gateways

When a Fibre Channel router needs to provide protocol conversion or tunneling services, it is a *gateway* rather than a router. However, it has become common usage to broaden the term *router* to include these functions. FCIP is an example of tunneling, while Small Computer System Interface over IP (iSCSI) and Internet Fibre Channel Protocol (iFCP) are examples of protocol conversion.

1.1.6 Fibre Channel routing between physical or virtual fabrics

Brocade, Cisco, and McDATA all offer FC-FC routing between separate physical fabrics. Cisco also offers Inter-VSAN Routing (IVR), which is Fibre Channel routing between separate logical (virtual) fabrics. In October 2004, the Technical Committee T11 of the International Committee for Information Technology Standards (INCITS) selected Cisco's VSAN technology for approval by the

American National Standards Institute (ANSI) as the industry standard for virtual fabrics.

1.2 Gateway protocols

The topics that follow introduce the protocols that are encountered in a “routed” environment.

1.2.1 FCIP

FCIP is a method for tunneling Fibre Channel packets through an IP network. FCIP encapsulates Fibre Channel block data and transports it over a TCP socket, or tunnel. TCP/IP services are used to establish connectivity between remote devices. The Fibre Channel packets are not altered in any way. They are simply encapsulated in IP and transmitted.

Figure 1-2 shows FCIP tunneling, assuming that the Fibre Channel packet is small enough to fit inside a single IP packet.

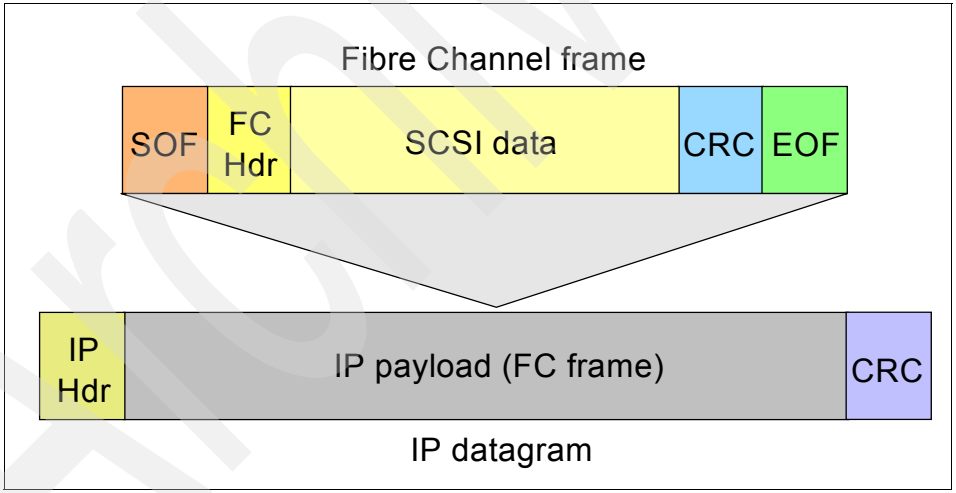


Figure 1-2 FCIP encapsulates the Fibre Channel frame into IP packets

The main advantage is that FCIP overcomes the distance limitations of native Fibre Channel. It also enables geographically distributed devices to be linked using the existing IP infrastructure, while keeping fabric services intact.

The architecture of FCIP is outlined in the Internet Engineering Task Force (IETF) Request for Comment (RFC) 3821, “Fibre Channel Over TCP/IP (FCIP),” available on the Web at:

http://ietf.org/rfc/rfc3821.txt

Merging fabrics

Because FCIP simply tunnels Fibre Channel, creating an FCIP link is like creating an inter-switch link (ISL), and the two fabrics at either end are merged into a single fabric. This creates issues in situations where you do not want to merge the two fabrics for business reasons, or where the link connection is prone to occasional fluctuations.

Many corporate IP links are robust, but it can be difficult to be sure because traditional IP-based applications tend to be retry-tolerant. Fibre Channel fabric services are not as retry-tolerant. Each time the link disappears or reappears, the switches re-negotiate and the fabric is reconfigured.

By combining FCIP with FC-FC routing, the two fabrics can be left unmerged, each with its own separate Fibre Channel services.

1.2.2 iFCP

iFCP is a gateway-to-gateway protocol. It provides Fibre Channel fabric services to Fibre Channel devices over a TCP/IP network. iFCP uses TCP to provide congestion control, error detection, and recovery. iFCP’s primary purpose allows interconnection and networking of existing Fibre Channel devices at wire speeds over a IP network.

Under iFCP, IP components and technology replace the Fibre Channel switching and routing infrastructure. iFCP was originally developed by Nishan Systems, which was acquired by McDATA in September 2003.

To learn more about the architecture and specification of iFCP, refer to the document at the following IETF Web site:

http://tools.ietf.org/wg/ips/draft-ietf-ips-ifcp/draft-ietf-ips-ifcp-14.txt

There is a popular myth that iFCP does not use encapsulation. In fact, iFCP encapsulates the Fibre Channel packet in much the same way that FCIP does. In addition, it maps the Fibre Channel header to the IP header and a TCP session, as shown in Figure 1-3.

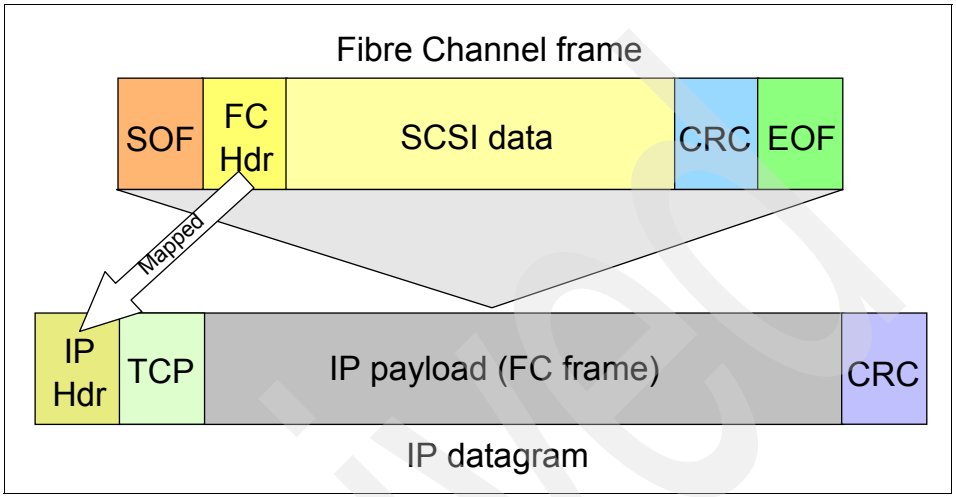


Figure 1-3 iFCP encapsulation and header mapping

iFCP uses the same Internet Storage Name Service (iSNS) mechanism that is used by iSCSI.

iFCP also allows data to fall across IP packets and share IP packets. Some FCIP implementations can achieve a similar result when running software compression, but not otherwise. FCIP typically breaks each large Fibre Channel packet into two dedicated IP packets. iFCP compression is payload compression only. Headers are not compressed to simplify diagnostics.

iFCP uses one TCP connection per fabric login (FLOGI), while FCIP typically uses one connection per router link (although more are possible). A FLOGI is the process by which an N_PORT registers its presence on the fabric, obtains fabric parameters such as classes of service supported, and receives its N_PORT address. Because under iFCP there is a separate TCP connection per N_PORT to N_PORT couple, each connection can be managed to have its own quality of service (QoS) identity. A single incidence of congestion does not need to drop the sending rate for all connections on the link.

While all iFCP traffic between a given remote and local N_PORT pair must use the same iFCP session, that iFCP session can be shared across multiple gateways or routers.

1.2.3 iSCSI

The Small Computer Systems Interface (SCSI) protocol has a client/server architecture. Clients (called *initiators*) issue SCSI commands to request services from logical units on a server known as a *target*. A SCSI *transport* maps the protocol to a specific interconnect.

The SCSI protocol has been mapped over various transports, including Parallel SCSI, Intelligent Peripheral Interface (IPI), IEEE-1394 (firewire), and Fibre Channel. All of these transports are ways to pass SCSI commands. Each transport is I/O specific and has limited distance capabilities.

The iSCSI protocol is a means of transporting SCSI packets over TCP/IP to take advantage of the existing Internet infrastructure.

A session between a iSCSI initiator and an iSCSI target is defined by a session ID that is a combination of an initiator part (ISID) and a target part (Target Portal Group Tag).

The iSCSI transfer direction is defined with respect to the initiator. Outbound or outgoing transfers are transfers from an initiator to a target. Inbound or incoming transfers are transfers from a target to an initiator.

For performance reasons, iSCSI allows a “phase-collapse.” A command and its associated data can be shipped together from initiator to target, and data and responses can be shipped together from targets.

An iSCSI name specifies a logical initiator or target. It is not tied to a port or hardware adapter. When multiple network interface cards (NICs) are used, they should generally all present the same iSCSI initiator name to the targets, because they are simply paths to the same SCSI layer. In most operating systems, the named entity is the operating system image.

The architecture of iSCSI is outlined in IETF RFC 3720, “Internet Small Computer Systems Interface (iSCSI),” available on the Web at:

`http://www.ietf.org/rfc/rfc3720.txt`

Figure 1-4 shows the format of the iSCSI packet.

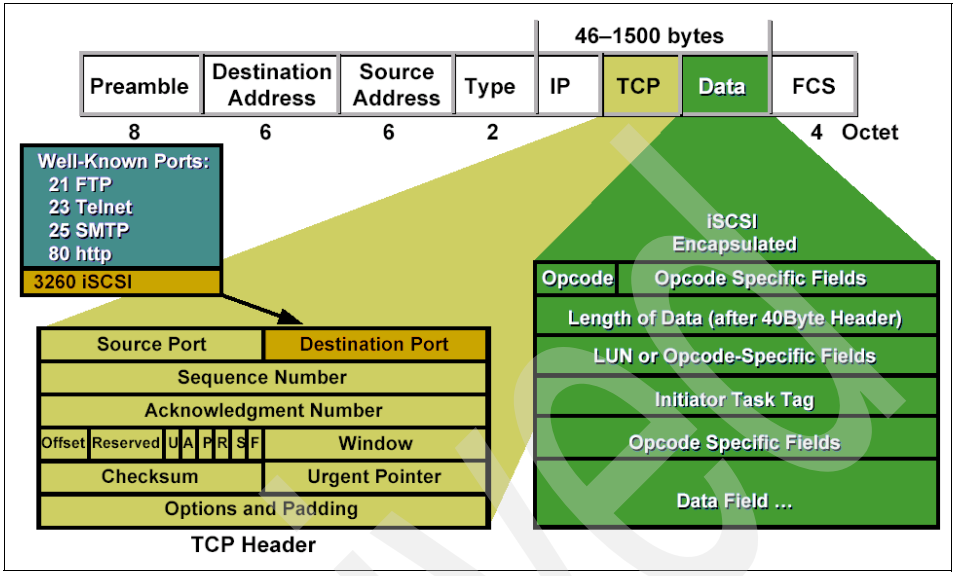


Figure 1-4 iSCSI packet format

Testing on iSCSI latency has shown a difference of up to 1 ms of additional latency for each disk I/O as compared to Fibre Channel. This does not include such factors as trying to do iSCSI I/O over a shared, congested, or long-distance IP network, all of which might be tempting for some clients. iSCSI generally uses a shared 1 Gbps network. The round trip delays in “Time of frame in transit” on page 11 also apply.

iSCSI naming and discovery

There are three ways for an iSCSI initiator to understand what devices are in the network:

- In small networks, you can use the **`sendtargets`** command.
- In larger networks, you can use the Service Location Protocol (SLP, multicast discovery).
- In large networks, we recommend that you use Internet Storage Name Service (iSNS).

Note: At time of writing, not all vendors’ have delivered iSNS.

You can find a range of drafts that cover iSCSI naming, discovery, and booting on the following Web site:

http://www.ietf.org/proceedings/02mar/220.htm

1.3 Routing issues

The topics that follow briefly describe some issues associated with a routed Fibre Channel environment.

1.3.1 Packet size

The standard size of a Fibre Channel packet is 2148 bytes, and the standard IP packet size is 1500 bytes (with a 1460 byte payload). When transporting Fibre Channel over IP, you can use jumbo IP packets to accommodate larger Fibre Channel packets. Keep in mind that jumbo IP packets must be turned on for the whole data path. In addition, a jumbo IP packet is not compatible with any devices in the network that do not have a jumbo IP packet enabled.

Alternatively, you can introduce a variety of schemes to split Fibre Channel packets across two IP packets. Some compression algorithms can allow multiple small Fibre Channel packets or packet segments to share a single IP packet.

Each technology and each vendor might implement this differently. They all try to avoid sending small, inefficient packets.

1.3.2 TCP congestion control

Sometimes standard TCP congestion mechanisms might not be suitable for tunneling storage. Standard TCP congestion control is designed to react quickly and severely to network congestion and recover slowly. This is well suited to traditional IP networks being somewhat variable and unreliable. But for storage applications, this approach is not always appropriate and might cause disruption to latency-sensitive applications.

When three duplicate unanswered packets are sent on a traditional TCP network, the sending rate backs off by 50%. When packets are successfully sent, it does a slow-start linear ramp-up again. The minimum send rate is normally set to:

```
minimum = maximum/20
```

Some vendors tweak the back-off and recovery algorithms. For example, the tweak causes the send rate to drop by 12.5% each time congestion is

encountered, and then to recover rapidly to the full sending rate by doubling each time until full rate is regained.

Other vendors take a simpler approach to achieve much the same end. Rather than introduce new algorithms, they suggest setting:

```
minimum = maximum x 0.8
```

If you are sharing your IP link between storage and other IP applications, using either of these storage friendly congestion controls might impact your other applications.

You can find the specification for TCP congestion control on the Web at:

http://www.ietf.org/rfc/rfc2581.txt

1.3.3 Round-trip delay

Round-trip link latency is the time it takes for a packet to make a round-trip across the link. The term *propagation delay* is also sometime used. Round-trip delay generally includes both inherent latency and delays due to congestion.

Fibre Channel cable has an inherent latency of approximately five microseconds per kilometer each way. Typical Fibre Channel devices, such as switches and routers, have inherent latencies of around five microseconds each way. IP routers might vary between 5 and 100 microseconds in theory, but when tested with filters applied, the results are more likely to be measured in milliseconds.

This is the essential problem with tunneling Fibre Channel over IP. Fibre Channel applications are generally designed for networks that have round-trip delays measured in microseconds. IP networks generally deliver round-trip delays measured in milliseconds or tens of milliseconds. Internet connections often have round-trip delays measured in hundreds of milliseconds.

Any round-trip delay caused by additional routers and firewalls along the network connection also needs to be added to the total delay. The total round-trip delay varies considerably depending on the models of routers or firewalls used and the traffic congestion on the link.

If you are purchasing the routers or firewalls yourself, we recommend that you include the latency of any particular product in the criteria that you use to choose the products. If you are provisioning the link from a service provider, we recommend that you include at least the maximum total round-trip latency of the link in the service level agreement (SLA).

Time of frame in transit

The time of frame in transit is the actual time that it takes for a given frame to pass through the slowest point of the link. Therefore, it depends on both the frame size and link speed.

The maximum size of the payload in a Fibre Channel frame is 2112 bytes. The Fibre Channel headers add 36 bytes to this, for a total Fibre Channel frame size of 2148 bytes. When transferring data, Fibre Channel frames at or near the full size are usually used.

If we assume that we are using jumbo frames in the Ethernet, the complete Fibre Channel frame can be sent within one Ethernet packet. The TCP and IP headers and the Ethernet medium access control (MAC) add a minimum of 54 bytes to the size of the frame, giving a total Ethernet packet size of 2202 bytes, or 17616 bits.

For smaller frames, such as the Fibre Channel acknowledgement frames, the time in transit is much shorter. The minimum possible Fibre Channel frame is one with no payload. With FCIP encapsulation, the minimum size of a packet with only the headers is 90 bytes, or 720 bits.

Table 1-1 details the transmission times of this FCIP packet over some common wide area network (WAN) link speeds.

Table 1-1 FCIP packet transmission times over different WAN links

Link type	Link speed	Large packet	Small packet
Gigabit Ethernet	1250 Mbps	14 µs	0.6 µs
OC-12	622.08 Mbps	28 µs	1.2 µs
OC-3	155.52 Mbps	113 µs	4.7 µs
T3	44.736 Mbps	394 µs	16.5 µs
E1	2.048 Mbps	8600 µs	359 µs
T1	1.544 Mbps	11 400 µs	477 µs

If we cannot use jumbo frames, each large Fibre Channel frame needs to be divided into two Ethernet packets. This doubles the amount of TCP, IP, and Ethernet MAC usage for the data transfer.

Normally, each Fibre Channel operation transfers data in only one direction. The frames going in the other direction are close to the minimum size.

1.3.4 Write acceleration

Write acceleration, or fast write as it is sometimes called, is designed to mitigate the problem of the high latency of long-distance networks. Write acceleration eliminates the time spent waiting for a target to tell the sender that it is ready to receive data. The idea is to send the data before receiving the ready signal, knowing that the ready signal will almost certainly arrive as planned. Data integrity is not jeopardized because the write is not assumed to have been successful until the final acknowledgement has been received.

Figure 1-5 shows a standard write request.

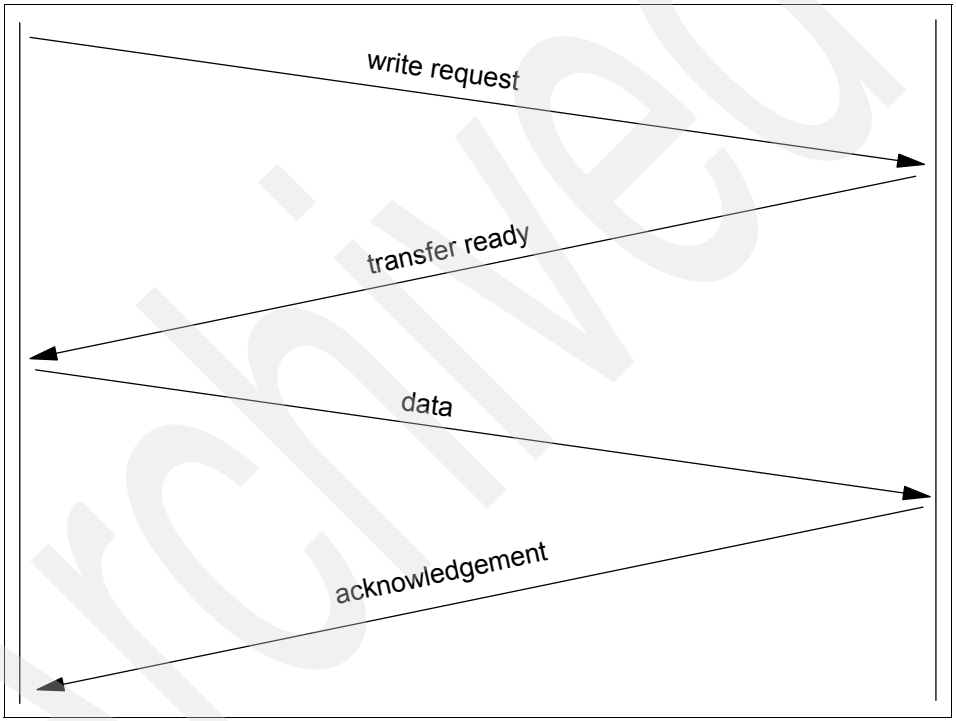


Figure 1-5 A standard write request

Figure 1-6 shows an accelerated write request.

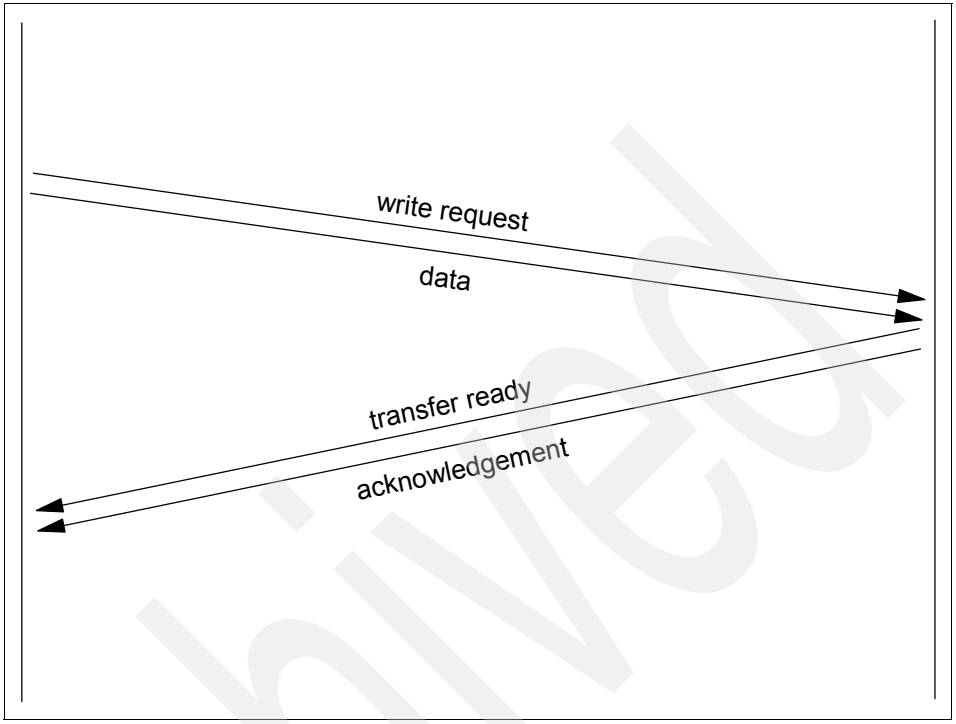


Figure 1-6 Write acceleration or fast write request

1.3.5 Tape acceleration

Tape acceleration (TA) takes write acceleration one step further by “spoofing” the transfer ready and the write acknowledgement. This gives the tape transfer a better chance of streaming rather than running stop and start. The risk here is that writes have been acknowledged but might not have completed successfully.

Without tape acceleration, a sophisticated backup/restore application, such as IBM Tivoli® Storage Manager, can recover and restart from a broken link. However, with TA, Tivoli Storage Manager believes that any write for which it has received an acknowledgement must have completed successfully. The restart point is therefore set after that last acknowledgement. With TA, that acknowledgement was spoofed so it might not reflect the real status of that write.

Tape acceleration provides faster tape writes at the cost of recoverability. Although the write acknowledgments are spoofed, the writing of the final tape mark is never spoofed. This provides some degree of integrity control when using TA.

Figure 1-7 shows how you can use tape acceleration to improve data streaming.

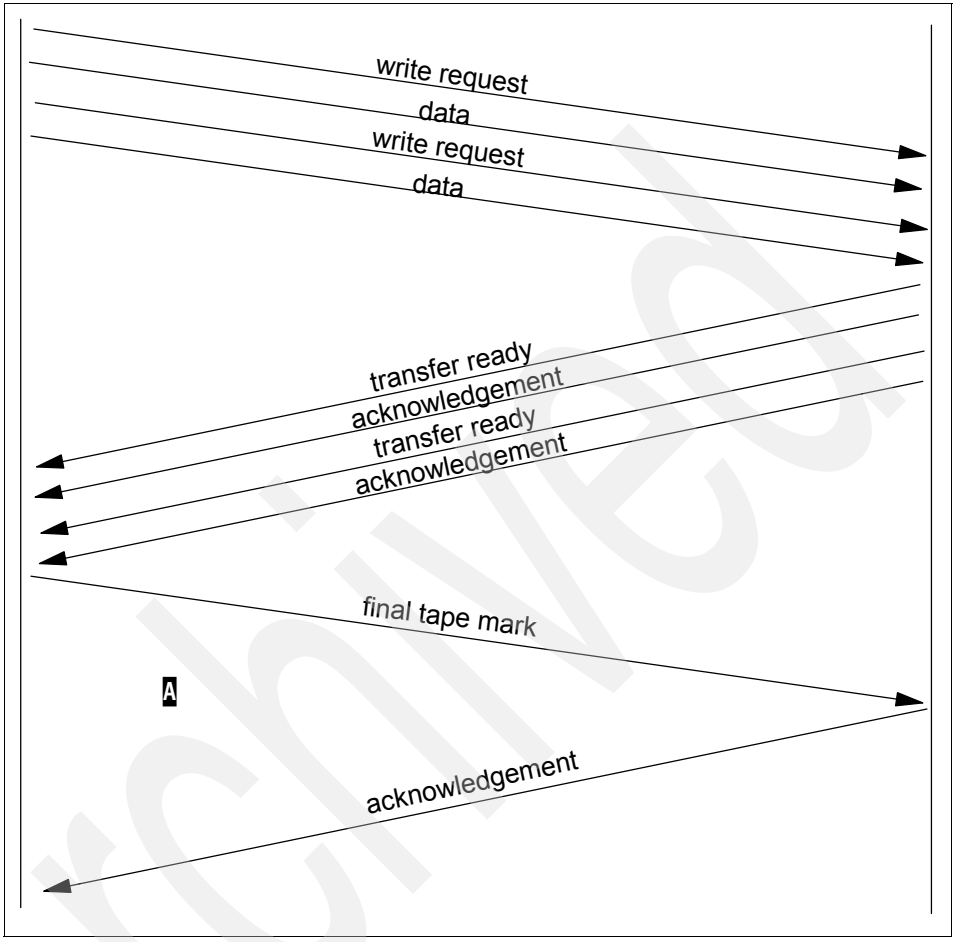


Figure 1-7 Tape acceleration example

1.4 Multiprotocol scenarios

The solution briefs in the following sections show how you can use multiprotocol routers.

1.4.1 Dividing a fabric into sub-fabrics

Suppose you have eight switches in your data center, and they are grouped into two fabrics of four switches each. Two of the switches are used to connect the

development/test environment, two are used to connect a joint-venture subsidiary company, and four are used to connect the main production environment.

The development/test environment does not follow the same change control disciplines as the production environment. Also, systems and switches can be upgraded, downgraded, or rebooted on occasion.

The joint-venture subsidiary company is up for sale. The mandate is to provide as much separation and security as possible between it and the main company, and the subsidiary. The backup/restore environment is shared between the three environments.

In summary, we have a requirement to provide a degree of isolation, and a degree of sharing. In the past, this would have been accommodated through zoning. Some fabric vendors might still recommend that approach as the simplest and most cost-effective. However, as the complexity of the environment grows, zoning can become complex. Any mistakes in setup can disrupt the entire fabric. Adding FC-FC routing to the network allows each of the three environments to run separate fabric services and provides the capability to share the tape backup environment.

In larger fabrics with many switches and separate business units, for example, in a shared services hosting environment, separation and routing are valuable in creating a larger number of simple fabrics, rather than fewer, more complex fabrics.

If using virtual fabrics as well (Cisco only at time of writing), you can apply additional separation within a physical fabric.

Note: FC-FC routing can provide departmental isolation and accommodate resource sharing.

1.4.2 Connecting a remote site over IP

Suppose you want to replicate your disk system to a remote site, perhaps 50 km away synchronously, or 500 km away asynchronously. Using FCIP tunneling or iFCP conversion, you can transmit your data to the remote disk system over a standard IP network. The router includes Fibre Channel ports to connect back-end devices or switches and IP ports to connect to a standard IP wide area network router. Standard IP networks are generally much lower in cost to provision than traditional high-quality dedicated dense wavelength division multiplexing (DWDM) networks. They also often have the advantage of being well understood by internal operational staff.

Similarly, you might want to provision storage volumes from your disk system to a remote site. You can do this by using FCIP tunneling (Brocade and Cisco) or iFCP protocol conversion (McDATA).

Note: FCIP and iFCP can provide a low-cost way to connect remote sites using familiar IP network disciplines.

1.4.3 Connecting hosts using iSCSI

Many hosts do not require high bandwidth low latency access to storage. For such hosts, iSCSI might be a more cost-effective connection method. iSCSI can be thought of as an IP SAN. There is no requirement to provide a Fibre Channel switch port for every server, nor to purchase Fibre Channel host bus adapters (HBAs), nor to lay Fibre Channel cable between storage and servers.

The iSCSI router has both Fibre Channel ports and Ethernet ports to connect to servers located either locally on the Ethernet or remotely over a standard IP wide area network connection.

Note: The iSCSI router is effectively the iSCSI target. Each iSCSI initiator, such as the server, is mapped to a worldwide name (WWN) generated by the router. As far as the Fibre Channel disk system is concerned, it sees each initiator as a separate Fibre Channel attached server. Fibre Channel logical unit numbers (LUNS) are mapped to iSCSI Qualified Names (IQNs) generated by the router. As far as the iSCSI initiator is concerned, it sees iSCSI targets.

The iSCSI connection delivers block I/O access to the server, so it is application independent. That is, an application cannot really tell the difference between direct SCSI, iSCSI, or Fibre Channel, because all three are delivery SCSI block I/Os.

Different router vendors quote different limits on the number of iSCSI connections that are supported on a single IP port.

iSCSI places a significant packetizing and depacketizing workload on the server CPU. This can be mitigated by using TCP/IP offload engine (TOE) Ethernet cards. However, because these cards can be expensive, they somewhat undermine the low-cost advantage of iSCSI.

Note: You can use iSCSI to provide low-cost connections to the SAN for servers that are not performance critical.

Part 1

IBM b-type family

In this part, we discuss OEM products from Brocade.

2

IBM TotalStorage b-type family routing products

This chapter introduces the storage area network (SAN) routing concepts in the IBM TotalStorage b-type family of SAN products and the products involved. We examine the hardware and software, SAN routing terminology, and the routing solutions.

2.1 IBM TotalStorage b-type family

At the time of writing, the IBM TotalStorage b-type family has three products that implement SAN routing:

- IBM TotalStorage SAN18B-R (2005-R18)
- M48 FC Routing Blade (for installation in IBM TotalStorage SAN256B Director)
- IBM TotalStorage SAN16B-R (2109-A16)

The SAN18B-R and the M48 FC Routing Blade are the new generation SAN router products. The SAN16B-R differs from the two newer products in several respects:

- The SAN16B-R has multiprotocol ports; all 16 ports can be configured as either FC or IP (GbE) ports. Therefore, tri-rate SFPs are required.
- The SAN18B-R and the FC Routing Blade have 16 dedicated FC ports and two dedicated IP ports.
- The two new products have all 16 FC ports and both IP ports enabled as standard. In contrast, the SAN16B-R can be purchased with only eight active ports (the remaining eight ports can be activated at a later time by purchasing the Ports On Demand license).
- SAN16B-R supports FC speeds up to 2 Gbps, while the two newer products support speeds up to 4 Gbps.
- SAN16B-R offers iSCSI support, but SAN18B-R and the FC Routing Blade do not.
- SAN16B-R uses XPath OS. Newer products use the standard Fabric OS.

We describe each SAN router product in the following sections.

2.1.1 SAN18B-R (2005-R18)

The SAN18B-R is designed to enable consolidation of SAN islands for infrastructure simplification, without the need to merge the SAN islands into one large SAN. High performance is assured by the 4 Gbps FC ports and hardware-assisted traffic processing across Gigabit Ethernet IP ports. Hardware-based compression, large window sizes, and selective acknowledgement of IP packets optimize performance of IP traffic as well.

Figure 2-1 shows the router.

Figure 2-1 SAN18B-R

The SAN router can provide metropolitan and global SAN extension for business continuity solutions, using the existing IP MAN or WAN infrastructure. Eight virtual FCIP tunnels per IP port provide scalability and efficient utilization of MAN and WAN resources.

Integrated management tools allow for simple installation, configuration, and administration.

Hardware components

The SAN18B-R offers high-speed FC and IP ports for excellent performance; in addition, hot swap redundant power supplies and fans provide high availability. We discuss these hardware components in the following sections.

Fibre Channel ports

The router has 16 FC ports, numbered 0 through 15. The ports support FC routing services at link speeds up to 4 Gbps. You can use a variety of SFP modules with these ports. A LED below each FC port indicates the operational status of the port.

The configuration and management utilities, such as WebTools, show 32 ports, numbered 0 through 31. Ports 0 through 15 are the 16 physical FC ports. Ports 16 through 31 represent the virtual ports assigned to FCIP tunnels defined across two GbE ports.

Gigabit Ethernet ports

Two Gigabit Ethernet (GbE) ports support the FCIP and FCR services at link speeds up to 1 Gbps. The ports are numbered *ge0* and *ge1*.

You can configure up to eight FCIP tunnels across each GbE port. Therefore, there are eight virtual ports for each physical GbE port. The virtual ports can be either VE_ports (fabrics connected through the FCIP tunnel merge) or VEX_ports (fabrics do not merge). Virtual ports 16 through 23 represent eight FCIP tunnels across physical port *ge0*, and virtual ports 24 through 31 correspond to eight tunnels across *ge1*.

Serial management port

The serial management port is used for initial router configuration—setting the management IP address using the terminal emulation utility (for example, HyperTerminal in Microsoft Windows®). After setting the IP address, you can use the Ethernet management port for all further configuration tasks. Apart from setting the IP address, the serial port is not intended for normal management and maintenance tasks. However, it can be used for switch recovery in case the Ethernet management access is lost.

A serial cable with RJ-45 connectors and an RJ-45-to-DB9 converter are supplied with the router.

Ethernet management port

This is a 10/100 Mbps Ethernet port intended for the management workstation attachment (either directly with a cross-over Ethernet cable or through the Ethernet LAN). The port has two LEDs, indicating the link status and speed.

You can connect to the management IP address either with Telnet (for CLI management) or with a Web browser (GUI management). We discuss both methods in the sections that follow.

Power supplies

The SAN18B-R contains two redundant hot-pluggable power supplies. The power supplies use separate power cords for increased availability. Each power supply is equipped with a power switch, a status LED, a captive screw, and a handle for easy removal and replacement.

Cooling fans

Three cooling fan assemblies provide redundant cooling. Each fan assembly actually contains two fans; the total number is therefore six fans. The fan assemblies contain a status LED, a captive screw, and a handle.

Standard and optional features

The SAN18B-R comes with the following standard features:

- All 18 ports (16 FC and 2 IP) active
- WebTools support
- FC-FC Routing Service
- Advanced Zoning

The following features are optional and require additional license:

- FCIP activation
- Advanced security
- Advanced Performance Monitoring
- Enhanced inter-switch link (ISL) trunking

WebTools

WebTools is a comprehensive set of management tools that use a Web browser interface. Simply point your browser at the management IP address and WebTools will launch. The tools enable you to install, configure, and manage the SAN18B-R router. The set includes modules for all aspects of the SAN router configuration.

FC-FC Routing Service

This service enables devices in separate SAN fabrics to communicate across routers without actually merging the fabrics. The routed network consisting of several SAN fabrics is known as a Meta SAN.

Advanced Zoning

Advanced Zoning provides hardware-enforced access control over fabric resources to prevent unauthorized storage access. Zone membership can be specified at port, AL-PA, and WWN level. It also simplifies heterogeneous storage management.

FCIP activation

FCIP provides extension of FC SANs over longer distances across IP connections. Basically, the FC traffic is encapsulated within IP packets, which are transferred across the IP network. Native FC connectivity is a lot more expensive over long distance than IP connectivity, and this is usually the most important reason for FC over IP implementation. With FCIP, you can extend your SANs across thousands of kilometers at a sensible cost.

You can configure up to eight virtual FCIP tunnels per physical IP port.

Advanced security

Interconnecting SANs over long distances has an impact on security requirements. Brocade offers Secure Fabric OS that provides policy-based security protection to fulfill these requirements. The methods include digital certificates and signatures, password protection on multiple levels, strong password encryption, PKI-based authentication, and 128-bit encryption used for digital signatures.

Advanced Performance Monitoring

The Advanced Performance Monitoring feature identifies end-to-end bandwidth usage and provides useful information for capacity planning. This comprehensive tool is based on Brocade Frame Filtering technology and a unique performance counter engine. Besides capacity planning, it can also be effectively used for troubleshooting.

Enhanced ISL trunking

This feature enables FC frames to be efficiently load balanced across multiple ISL connections, while preserving in-order delivery. Up to eight 4 Gbps links can be combined to set up an ISL connection with up to 32 Gbps throughput per trunk.

2.1.2 M48 FC Routing Blade

The FC Routing Blade can be installed in the 2109-M48 Director to provide FC routing and FCIP capabilities. At the time of writing, the M48 Director supports up to two FC Routing Blades. Figure 2-2 shows the FC Routing Blade.

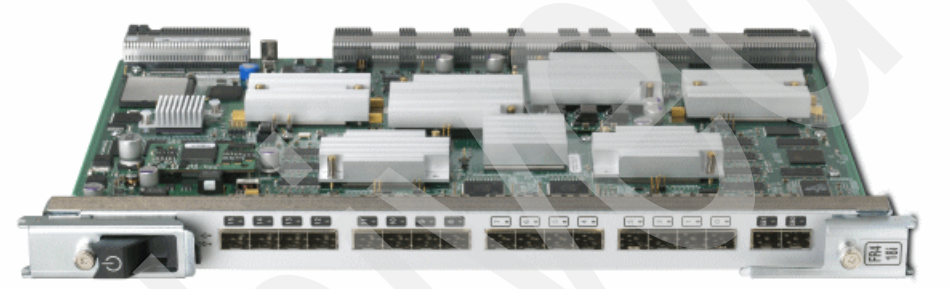


Figure 2-2 M48 FC Routing Blade

There are additional prerequisites for proper implementation of the Routing Blade:

- For adequate cooling, install one blade in slot grouping 1-4 and the other in slot grouping 7-10.
- Note that the FC Routing Blade requires that two additional power supplies are installed in the M48 Director.
- The FC Routing Blade requires Fabric OS v5.1.0b or later. You need to make sure that the M48 Director is on the appropriate Fabric OS version *before* you install any FC Routing Blades. When the Routing Blade is powered on, it will go through POST and then autolevel with the firmware version in the active CP. If the firmware version is below V5.1.0b, the Routing Blade will not work.
- Another requirement is chassisConfig mode 5. If the M48 Director is not in correct mode, use the **`chassisConfig 5`** command to change it. Be aware that this command is disruptive and requires a reboot.

The features of M48 FC Routing Blade are similar to those of SAN18B-R. The blade has 16 FC ports and two IP ports. FC ports support speeds up to 4 Gbps and the two IP ports are Gigabit Ethernet (GbE). Each GbE port can support up to eight FCIP tunnels, as is the case with SAN18B-R.

By default, the blade will power up in disabled state until the ports are configured and enabled. The configuration is not stored on the routing blade itself, but by slot location in the CP. If you move the blade to another slot, the configuration will not follow; it will be preserved in the previous slot location. If you install another routing blade to that slot, it will use the existing slot configuration.

One feature currently not supported on the FC Routing Blades is ISL trunking on EX_Ports.

2.1.3 SAN16B-R (2109-A16)

The 2109-A16 is designed to provide intelligent multiprotocol routing services to help enable Internet Protocol (IP)-based Global Mirroring business continuity solutions. It also provides Small Computer System Interface over IP (iSCSI)-based host server storage consolidation solutions. And, it provides SAN routing capabilities for infrastructure simplification solutions to selectively share resources between fabrics.

Figure 2-3 shows the 2109-A16.

Figure 2-3 IBM TotalStorage SAN16B-R

The 2109-A16 provides the following hardware features:

- Eight or sixteen active multiprotocol ports with the following features:
  - Support for 1 Gbps and 2 Gbps Fibre Channel (FC) or Gigabit Ethernet
  - Automatic negotiation to the highest common speed of all Fibre Channel devices connected to port
  - Support for short wavelength (SW), long wavelength (LW), and extended long wavelength (ELW) small form-factor pluggable (SFP) transceivers
  - Support for F_Port, FL_Port, or E_Port operation
  - Support of each active Gigabit Ethernet port for either iSCSI gateway or FC over IP (FCIP) tunneling functionality
- No SFPs included; one Tri-rate SFP for each active port required
- Two 10/100 Mbps Ethernet ports and one RS-232 serial port for management
- Redundant, hot pluggable power supplies and fans
- Two U chassis

The following standard software features are included with the 2109-A16:

- Fibre Channel switch support
- Advanced WebTools
- Advanced Zoning
- Exchange-based inter-switch link (ISL) and inter-fabric link (IFL) trunking
- Extended fabric capability
- iSCSI gateway functionality

The optional software features available for the router include:

- Fibre Channel routing
- FCIP
- FCIP and FC Routing Bundle

We describe the hardware and software components in more detail in the following sections.

2109-A16 hardware components

The 2109-A16 is a multiple board design. Below the system board in the chassis is a dc power printed circuit board (PCB) that provides the required system voltages. These voltages are derived from and regulated by the 2109-A16 redundant power supply units. This regulated power is supplied to the main system PCB. Mounted on top of the system board is a daughter board that contains a high-performance 800 MHz PowerPC® 745x reduced instruction set computer (RISC) processor core with SDRAM controller, PCI bus interface, peripheral local bus for external ROM and peripherals, direct memory access (DMA), I2C interface, and general purpose I/O.

The system uses four types of memory devices in the design: SDRAM, kernel flash, compact flash (user flash), and boot flash. The fabric application and switching section of the system board, the XPath per-port processing application-specific integrated circuit (ASIC) and memory chip sets, the XPath Fabric ASIC, and the SFP media are the key components that provide high-speed data manipulation and movement. The SFP media interface to external devices and support any combination of SW, LW, and ELW optical media.

Power supplies

The 2109-A16 power supply is a hot-swappable field replaceable unit (FRU), enabling 1+1 redundant configurations. The unit is a universal power supply that is capable of functioning worldwide without voltage jumpers or switches. The fully enclosed, self-contained unit has internal fans to provide cooling. The power supply provides three dc outputs (5V standby, 12V, and 48V), with a total maximum output power of 320W. Each power supply provides an integral on/off

switch, input filter, and power indicator, as well as a serial EEPROM device that provides identifying information.

Multiprotocol ports

The 2109-A16 has 16 multiprotocol ports (numbered 0 through 15, left to right). You can purchase it with either eight or 16 active ports. If you purchase the 2109-A16 with only eight active ports, you can activate the other eight ports by purchasing and installing the Ports on Demand license key.

Each of the 16 multiprotocol ports needs to be equipped with an SFP. The SFPs are hot swappable and use industry-standard local channel connectors. Each port is supported by its own dedicated ASIC, which contains three embedded 133 MHz ARM processors.

In Fibre Channel mode, each port provides support for E_port, F_port, and FL_port modes. The ports also support automatic negotiation of both port mode and speed. With the current firmware, each port can provide up to 255 buffer credits.

Important: Because the multiprotocol ports of the 2109-A16 support both Fibre Channel and Gigabit Ethernet, they require the use of special Tri-rate SFPs. The router only activates ports that are equipped with supported SFPs. We recommend that you always use only the SFPs delivered with the router.

Fabric ASIC

The XPath Fabric ASIC provides non-blocking connectivity between the 16 separate port ASICs.

Management ports

The 2109-A16 provides dual 10/100 BaseT Ethernet ports for management purposes. You can configure the TCP/IP address for each port through the serial port.

Serial port

An RS-232 serial port is provided on the 2109-A16. The serial port uses an RJ-45 connector. Communication parameters are fixed at 9600 baud, 8 data bits, and no parity, with flow control set to None. This connection is used for the initial IP address configuration and for recovery of the switch to its factory default settings, should flash memory contents be lost. The serial port connection is not intended for normal administration and maintenance functions.

Cooling fans

The non-port side of the 2109-A16 includes dual hot-swappable cooling fan assemblies. Each contain three fans and system status LEDs.

Software features

The XPath OS is used on the 2109-A16. It provides full Fibre Channel switch capability, FC-FC routing, FCIP tunneling, and iSCSI to FC gateway.

Fibre Channel switch support

The XPath OS includes the following Fibre Channel switch features:

- Name server support
- Zone server support
- Exchange-based ISL trunking
- Extended fabric support

FC-FC routing

The FC-FC routing service provides connectivity to devices in different fabrics without merging the fabrics. FC-FC routing enables the creation of logical storage area networks (LSANs). An LSAN can span multiple fabrics, allowing Fibre Channel zones to cross physical SAN boundaries without merging the fabrics, yet maintaining access control of the zones.

FC-FC routing also enables you to share devices, such as tape drives, across multiple fabrics without the associated administrative problems that can result from merging the fabrics, including change management, network management, scalability, and reliability, availability, and serviceability.

FCIP tunneling

The FCIP tunneling service enables tunneling of Fibre Channel frames through TCP/IP networks. It encapsulates them in TCP packets and then reconstructs them at the other end of the link.

Note: The XPath OS supports an FCIP connection only between two 2109-A16s.

iSCSI gateway

The iSCSI gateway service provides connectivity to Fibre Channel targets for servers using iSCSI. Servers use an iSCSI adapter (or an iSCSI driver and Ethernet adapter) to connect to a Fibre Channel fabric over IP.

Management capability

There are three ways to manage the 2109-A16:

- Command-line interface (CLI)
- Advanced WebTools AP Edition
- Fabric Manager

The CLI and Advanced WebTools are similar to the same features in the IBM TotalStorage b-type switches. They are included with the 2109-A16 at no extra cost.

Fabric Manager provides a Java™-based application that can simplify management of a multiple-switch fabrics. Use it to administer, configure, and maintain fabric switches and SANs. In addition, it has an easy-to-use wizard for creating LSANs. Fabric Manager is available as an optional feature of most IBM TotalStorage b-type SAN switches.

2.2 SAN routing terminology

Table 2-1 lists and explains some of the routing terminology that is used throughout this book.

Table 2-1 Routing terminology

Term	Explanation
E_Port	A port on an FC switch or router that connects to another switch or router, forming an ISL. If the devices previously formed separate fabrics, these fabrics merge, putting all fabric services into one distributed image.
VE_Port	An FCIP port on an FC switch will create a *Virtual E_Port*. This is physically an IP/Ethernet interface, but each FCIP tunnel *looks* like an FC E_Port to the rest of the fabric.
EX_Port	FC Routers use EX_Ports instead of E_Ports on routed interfaces. To connect a router to a switch, you connect its EX_Port to another switch's E_Port through an appropriate cable. Routers still use E_ or VE_Ports to form a backbone fabric.
VEX_Port	In addition to supporting virtual E_Ports, Brocade platforms allow the FCIP and FC Router features to be combined, creating a *Virtual EX_Port*.
ISL	The connection between two E_Ports is an *inter-switch link*.
IFL	The connection between an E_Port and an EX_Port is an *inter-fabric link*.

Term	Explanation
Tunnel	The FCIP connection between switches is a tunnel. The tunnel can contain one or many logical ISLs or IFLs.
Edge fabric	This is Fibre Channel fabric connected to a router through an EX_Port (IFL). This is largely the same as any standard Fibre Channel fabric. This is, for the most part, where the hosts and storage are attached.
Backbone fabric	Routers provide a backbone (BB) fabric to interconnect routers for more scalable and flexible routed SANs. Each router can have many edge fabric connections, but only one BB fabric. Routers connect to the BB fabric through E_Ports, and all N_ and NL_ port connections on a router are part of the BB fabric. With 4-bit routers, a number of hosts and storage devices can be connected to the BB fabric.
Meta SAN	The collection of edge fabrics and backbone fabrics connected together is called a *Meta SAN*.
Fabric ID	Each edge fabric has a *fabric ID (FID)* that is unique within the same Meta SAN, and configured to all EX_Ports connected to the edge fabric. The backbone fabric must also have a unique FID configured.
LSAN	Logical SANs are zones that span fabrics. They traverse at least one EX_Port or VEX_Port. LSANs are how connectivity is configured across routers.

Figure 2-4 shows a Meta SAN that consists of four edge fabrics connected to a single backbone fabric.

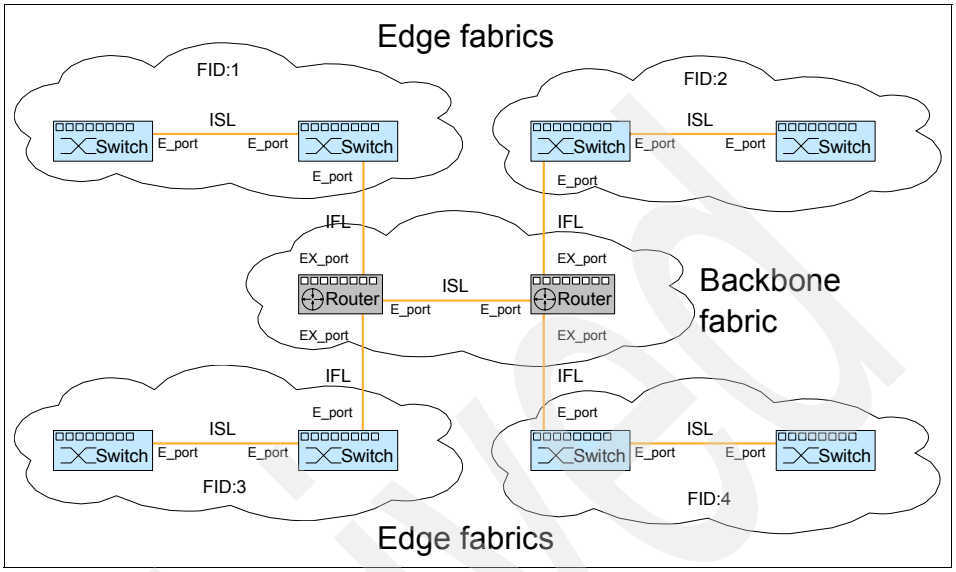


Figure 2-4 Meta SAN with four edge fabrics connected to a backbone fabric

The switches in the edge fabric treat an IFL like a normal ISL. Through the IFL, they gain access to a set of *phantom domains* that are never principal domains in the edge fabric. Two types of phantom domains are created in the edge fabric:

- Every EX_Port connected to an edge fabric is represented by a front domain (fd) that participates in the edge fabric.
- Every remote edge fabric that has at least one node exported to the local edge fabric is represented by a single translate domain (xlate).

The phantom domains can have the following connections:

- The virtual links connecting front domains to translate domains are called *phantom links*.
- The exported nodes of a remote fabric are represented by port addresses in the translate domain called proxy devices.

The translate domains are used to perform Fibre Channel network address translation (FC-NAT) between the different edge fabrics. The translate domain IDs are persistent across router reboots and can be assigned manually.

The front domains are used to provide multiple paths to the translate domains through the different IFLs that are available and allow normal Fabric Shortest Path First (FSPF) routing across the paths.

When counting the hop count in the complete Meta SAN, you need to count the ISLs and IFLs as hops. You do not need to count the phantom links, because they are not physical links and do not add any delay to the data path.

The port IDs (PIDs) of the proxy devices follow the specific format 0xAABBBB, where:

- AA is the translate domain for the remote fabric where the physical device is attached.
- BBBB is the *virtual slot* number in the range 0xf001-0xffff.

The virtual slot numbers can be assigned automatically, or you can assign them manually.

Note: Some operating systems, such as IBM AIX® 5L™ or HP-UX, assume that the PID of any device stays constant. If you have any servers running either of these operating systems, we recommend that you define both the translate domain IDs and the virtual slot numbers manually to ensure that they remain constant.

Figure 2-5 shows the logical view of an edge fabric with four IFLs having a connection to another edge fabric.

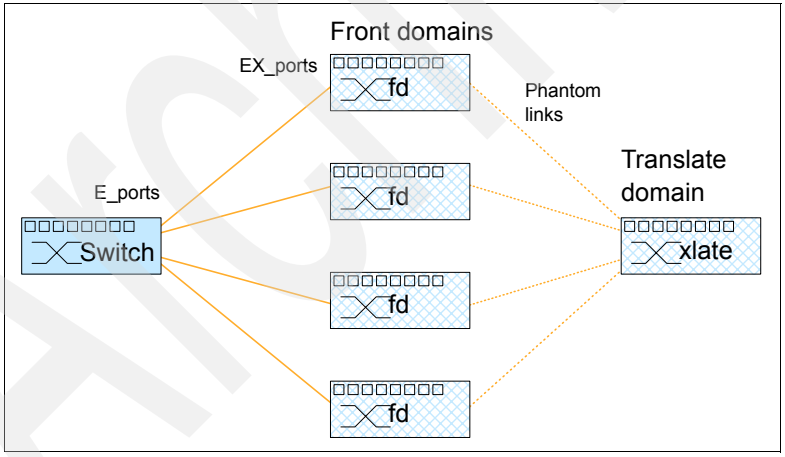


Figure 2-5 Edge fabric logical view of four EX_Ports

In addition, to support FCIP tunneling, we introduce the *VE_Port*, or rather the virtual E_Port created over the FCIP tunnel. Figure 2-6 shows an FCIP tunnel between two SAN routers.

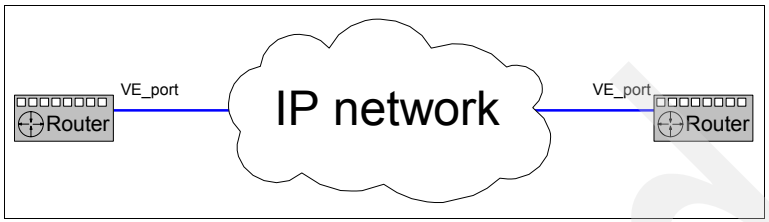


Figure 2-6 FCIP tunnel

The VE_Port functions exactly like a normal E_Port, and a normal ISL is created between the two VE_Ports.

2.3 Description of the routing solution

This example has a small Meta SAN, consisting of two edge fabrics of one switch each and a single router backbone fabric. Each edge fabric has two connections to the router for redundancy.

The fabric parameters are set as follows:

- Fabric 1
  - Core PID: 1
  - Domain IDs used: 1
  - Fabric ID: 1
- Fabric 2
  - Core PID: 0
  - Domain IDs used: 1
  - Fabric ID: 2

Because we use a different core PID format in the fabrics, we cannot merge them into a single fabric. We also use the same domain ID on both fabrics. Changing either of these parameters would be disruptive to the fabric involved.

We need to enable the server connected to fabric 1 access to a storage device that is connected to fabric 2. We implement this by creating an LSAN called LSAN_zone1.

Figure 2-7 shows the physical layout of our environment.

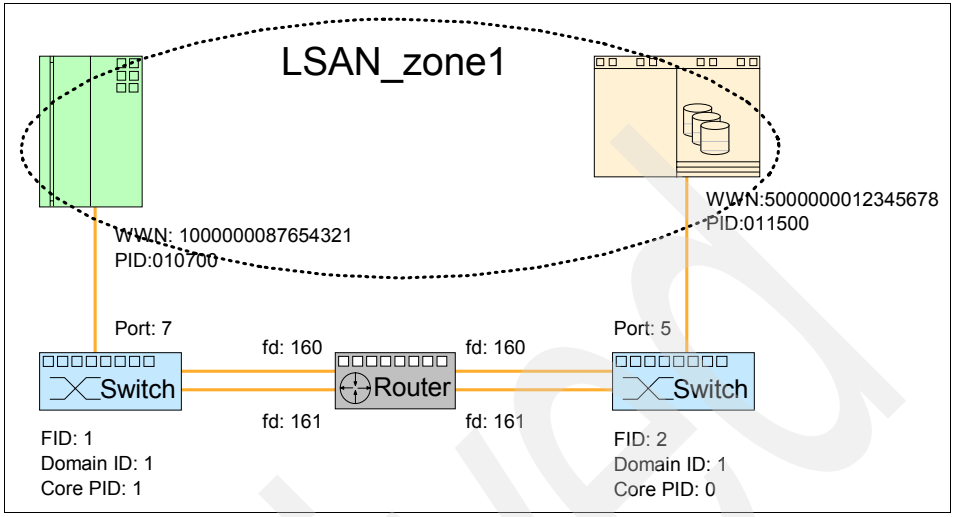


Figure 2-7 FC-FC routing physical layout

As you can see, the router automatically assigns a separate front domain for each IFL in each edge fabric, starting from a default preferred domain ID 160. The front domain IDs need to be unique only within a single edge fabric, and the same front domain ID can be assigned in several different edge fabrics.

We create the LSAN by creating a zone named LSAN_zone1 that has two members: the port worldwide names (PWWNs) 10:00:00:00:87:65:43:21 and 50:00:00:00:12:34:56:78. The router automatically intercepts any zone with a name starting with LSAN and exports any LSAN members between fabrics as required.

In our case, we assume that fabric 2 is given translate domain ID 5 in fabric 1, and fabric 1 is given translate domain ID 6 in fabric 2. Any fabric 2 members exported to fabric 1 are represented in fabric 1 by proxy devices with PIDs starting from 05f001. And any fabric 1 members exported to fabric 2 are represented in fabric 2 by proxy devices with PIDs starting from 06f001.

Figure 2-8 shows the logical view seen by the server in fabric 1.

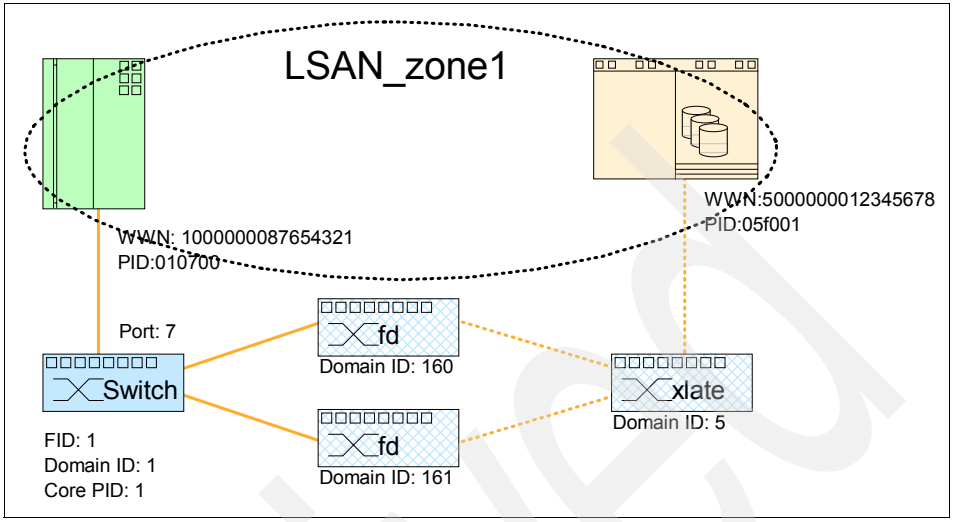


Figure 2-8 Logical view from fabric 1

Figure 2-9 shows the logical view seen by the storage device in fabric 2.

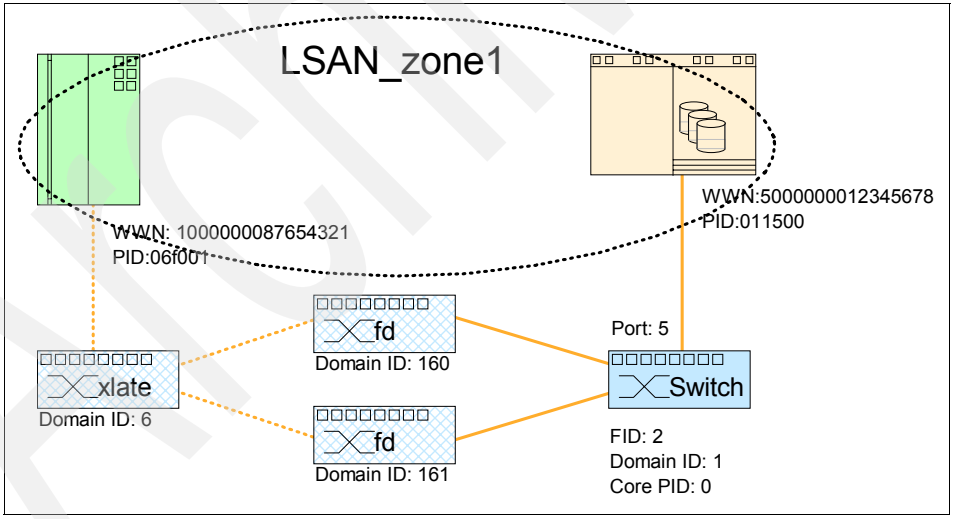


Figure 2-9 Logical view from fabric 2

The WWNs of the LSAN members are not changed by the NAT. This way, you can still use the real WWN of the server for logical unit number (LUN) masking in the storage device.

2.4 Current limitations

This section examines the specific limitations of the different functions of SAN routers. The limitations are based on the capabilities of the current hardware and operating system version, and are subject to change. We recommend that you always check the current values in the interoperability matrix for your SAN router product. You can find the matrix on the support Web site:

- Interoperability matrix for SAN18B-R and FC Routing Blade:

  http://www.ibm.com/servers/storage/support/san/2005-r18/planning.html

- Interoperability matrix for SAN16B-R:

  http://www.ibm.com/servers/storage/support/san/2109a16/planning.html

2.4.1 FC-FC routing

Because the backbone fabric can consist of multiple SAN routers and multiple FC switches, there is no practical limit on the number of edge fabrics that can be connected to the backbone fabric.

SAN18B-R and FC Routing Blade maximums

The Fabric OS v5.1 limitations of SAN18B-R and the Routing Blade are as follows:

- For each edge fabric:
  - Maximum number of local switches: 26
  - Maximum number of front domains: 10
  - Maximum number of translate domains: 17
  - Maximum number of total domains: 53
  - Maximum number of local WWNs: 1000
  - Maximum number of imported devices: 500
  - Maximum number of local and remote WWNs: 1300
- For each Meta SAN:
  - Maximum number of edge fabrics: 16
  - Maximum number of FCR switches: 10
  - Maximum number of device DB entries: 5000
  - Maximum number of LSAN zones: 1000
- For each backbone fabric:
  - Maximum number of local switches: 5

  - Maximum number of translate domains: 16
  - Maximum number of total domains: 21
  - Maximum number of local WWNs: 256
- For each backbone switch:
  - Maximum number of EX_ports: 32
  - Maximum number of EX_ports to one edge fabric: 4
- Maximum number of entries per LSAN zone: 64
- Maximum number of hops between edge switches: 12

SAN16B-R maximums

The FC-FC routing function has the following maximums when using SAN16B-R:

- Edge fabrics in interoperability mode are currently not supported.
- LSAN can only contain nodes from edge fabrics, not from the backbone fabric.
- LSAN members must be specified by port WWN.
- The maximum number of hops between switches, including routers is 12.
- For each edge fabric:
  - Maximum number of front domains: 15
  - Maximum number of translate domains: 33
  - Maximum number of physical domains (switches): 32
  - Maximum total number of domains (physical + translate + front): 80
  - Maximum number of all devices (local + proxy): 1280
  - Maximum number of proxy devices: 1000
- For the complete routed fabric (Meta SAN):
  - Maximum number of edge fabrics: 34
  - Maximum number of FC routers: 14
  - Maximum number of LSAN zones: 1000
  - Maximum number of LSAN zone members: 10 000
  - Maximum number of members/LSAN zone: 200

2.4.2 FCIP tunneling on SAN16B-R

The FCIP tunneling function has the following limitations:

- One point-to-point FCIP connection is supported for each port configured for FCIP.
- The 2109-A16 does not implement compression on FCIP links.
- There is no fast write or write acceleration feature in the 2109-A16.

2.4.3 iSCSI gateway on SAN16B-R

The iSCSI gateway function has the following limitations:

- A maximum of 12 ports of the 2109-A16 can be configured as iSCSI portals.
- Each port can support up to eight iSCSI sessions.
- iSCSI traffic cannot exit the 2109-A16 through an EX_port or a VE_port.
- The current iSCSI client software is the Microsoft Initiator v2.02.

3

IBM TotalStorage b-type family routing solutions

This chapter describes the solutions that are available using the IBM TotalStorage SAN multiprotocol routers. We address the following three solutions:

- Fibre Channel to Fibre Channel (FC-FC) routing.
- FC over Internet Protocol (FCIP) tunneling.
- Small Computer System Interface over IP (iSCSI) gateway, available on SAN16B-R only. SAN18B-R and the FC Routing Blade do not support iSCSI.

3.1 FC-FC routing

This section describes the FC-FC routing functionality and possible scenarios where the functionality can be used.

3.1.1 Local FC-FC routing

Figure 3-1 shows the local FC-FC routing solution between two storage area network (SAN) fabrics.

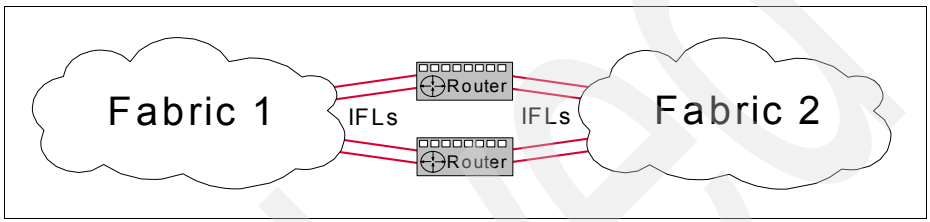


Figure 3-1 Local FC-FC routing between two SAN fabrics

This solution has two redundant SAN routers connected to two SAN fabrics. Both routers are connected to both fabrics using two inter-fabric link (IFLs). We can extend this configuration to span up to eight SAN fabrics, or up to 16 fabrics by adding two more routers.

Each edge fabric can be connected to the routers with up to four IFLs. If you have multiple switches in your fabric, we recommend that you distribute the IFL connections across them for maximum availability. If you are using a core-edge fabric, we recommend that you connect IFLs to the core switches.

Some of the solutions that can be provided by local FC-FC routing include:

- Scalability

  The Fibre Channel addressing can theoretically support up to 16 million nodes in a single fabric. However, the practical limit for the number of nodes is much lower. This is similar to TCP/IP networks, where the network address space is divided into smaller subnets of limited number of IP addresses, and traffic is routed between them. Usually, the practical limit of a fabric from both technical and management standpoint is between 250 and 1000 nodes.

  FC-FC routing enables you to divide the environment into several fabrics, while providing access to shared resources between fabrics.

- Multiple SAN administrators

  In many cases, enterprises have several small SAN islands that are managed by different SAN administrators, often as a result of mergers or acquisitions.

Using FC-FC routing between the fabrics allows each fabric to be managed separately from other fabrics. It prevents the propagation of any management errors to other fabrics.

The logical storage area network (LSAN) configuration is the only part of the fabric that needs to be coordinated among the SAN administrators. Because the LSAN zones need to be defined on all fabrics before they can route traffic, the devices in each fabric are protected against unplanned access from the other fabrics.

- Interoperability between storage vendors

  Several storage vendors offer Brocade Silkworm SAN products, either as resellers or as OEM products. Although these products are theoretically compatible with each other, each vendor usually only supports specific levels of Fabric OS, and it may be difficult to find a common supported version among multiple storage vendors. The fabric-wide parameters and recommended zoning methodologies can also differ between vendors.

  If the different vendors are each separated to their own SAN fabrics, as shown in Figure 3-1 on page 40, we avoid these problems. This solution also allows each edge fabric to be supported and even managed by the corresponding storage vendor, while enabling storage access between different SAN fabrics.

- Interoperability between old and new fabrics

  In many cases when implementing a new SAN fabric, you already have an existing fabric. The existing fabric can have some parameter settings that you want or need to set up differently in the new fabric. One good example is the core PID setting.

  By using FC-FC routing to connect the fabrics, you do not need to change the settings in the old fabric, and are free to choose the settings that you need for the new fabric. You can also use only a single Fabric OS level in any fabric, independent on the Fabric OS levels supported by the old hardware.

- Migration between old and new fabrics

  The storage hardware is usually replaced with new hardware every three to five years. When refreshing the disk hardware, it might make sense to refresh the SAN hardware as well, especially if the new disk vendor is different than the old vendor.

  FC-FC routing enables you to implement the new disk subsystems and SAN fabric in the final configuration and connect the complete new environment to the current SAN fabrics. This way, you can have simultaneous access from the servers to both old and new disk subsystems, and use server-based tools, such as Logical Volume Manager (LVM), to migrate the data from the old disks to the new disks.

After you migrate any host to new disks, you can move the Fibre Channel ports of the server to the new SAN fabric as well. Because you can do this one server at a time, the outage needed is minimized.

- Storage consolidation

  Many enterprises implement a separate SAN fabric for tape backups. Without FC-FC routing, this requires a separate Fibre Channel adapter in each server that needs to be connected to the backup devices, as well as the additional fiber cabling to support these adapters. If you set up FC-FC routing between the normal SAN fabrics and the backup fabric, you can share the tape devices across any adapters in any fabric, as required.

  Another example of storage consolidation is implementing a single IBM TotalStorage SAN Volume Controller (SVC) cluster across multiple SAN fabrics.

3.1.2 Fabric extension with FC-FC routing

Figure 3-2 shows a simple SAN fabric extension using the SAN routers.

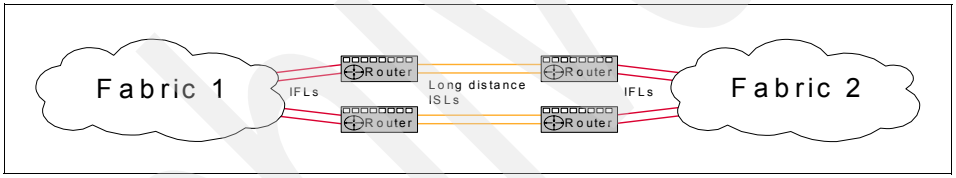


Figure 3-2 Fabric extension with FC-FC routing

This solution takes advantage of the large number of buffer credits available on each port in the router and the Extended Fabrics feature. The long-distance inter-switch links (ISLs) can be implemented with dark fiber, or with dense wavelength division multiplexing (DWDM) technology.

This solution has the advantage that the edge fabrics are separated from each other and from the long-distance ISLs. This isolates any SAN fabric failure in one site to that site only and prevents the other site from being affected. This is especially important in a disaster recovery environment. Any link failures on the long-distance links are also isolated from both edge fabrics.

3.2 FCIP tunneling

In situations where dark fiber for DWDM is not available or the distances are longer than supported by DWDM, the SAN fabrics can be connected using the FCIP tunneling functionality. Figure 3-3 shows this solution.

Figure 3-3 FCIP tunneling

We recommend that you combine FCIP tunneling with the FC-FC routing feature. This way, you can isolate any SAN fabric failure in one site to that site only and prevent the other site from being affected. Any link failures on the FCIP links are also isolated from both edge fabrics.

3.3 iSCSI gateway

The iSCSI gateway feature of the 2109-A16 enables you to connect servers with no Fibre Channel adapters to Fibre Channel-attached storage devices using the iSCSI protocol. Figure 3-4 shows this solution.

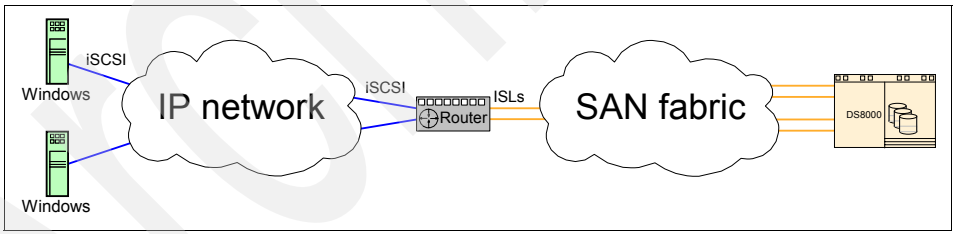


Figure 3-4 iSCSI gateway

Each 2109-A16 port configured for iSCSI supports up to eight concurrent iSCSI sessions. A total of 12 ports in a single 2109-A16 can be configured for iSCSI.

Note: In the current implementation, the iSCSI gateway traffic has to leave the 2109-A16 through an E_port. It cannot leave through an EX_Port or a VE_Port. The 2109-A16 must be a part of an edge fabric to connect to the storage devices in any edge fabric. Therefore, it is usually not feasible to share a single 2109-A16 between the iSCSI function and the other functions described in this chapter.

4

IBM TotalStorage b-type family routing best practices

This chapter describes some of the key features of the IBM TotalStorage b-type family routing solution. It also provides information about best practices when implementing such a solution.

Specifically, this chapter examines:

- Planning considerations
- Compatibility and interoperability
- Availability
- Security
- Performance
- IP network issues

4.1 Planning considerations

This section addresses the specific topics you need to consider when planning solutions with the IBM TotalStorage SAN multiprotocol router products.

4.1.1 Piloting new technology

Whenever you plan to use advanced features, such as Fibre Channel (FC) over IP (FCIP) or FC-FC routing, it always pays to implement it initially as a pilot with the understanding that experience gained in your own environment will always be slightly unique.

When implementing leading-edge technologies, many clients prefer to avoid the uncertain outcomes that a pilot implies. Instead, they secure implementation guarantees from vendors. But in fact, the outcome can never really be guaranteed, and piloting allows the solution to be tailored based on lessons learned in your own environment.

4.1.2 FC-FC routing considerations

Because logical storage area networks (LSANs) are created like any other zones, apart from the name, it is possible to define all zones as LSAN zones. However, if you do this, each SAN router has to keep track of all your zones, limiting the future scalability of the Meta SAN. Therefore, we do not recommend this approach.

We recommend that you create zones that enable traffic within a single edge fabric as normal zones. Also, create a separate LSAN zone wherever you need to access a storage device from a host in different edge fabric. This way, you also avoid any unnecessary access between devices.

If you are using FC-FC routing in an environment with AIX 5L or HP-UX servers, we recommend that you define the translate domain IDs and virtual slot numbers manually so that you can ensure that FC IDs of any devices used by the AIX 5L or HP-UX servers remain persistent.

4.1.3 FCIP tunneling considerations

When implementing an FCIP connection, the quality of the IP link is critical. When you order the IP link from a vendor, you need to have a service level agreement (SLA) that concerns the operation of the link. We recommend that you ensure that the link vendor takes into account the specific requirements of the storage environment in their design.

The SLA should include at least the following parameters:

- Guaranteed link bandwidth
- Round-trip latency
- Maximum packet loss rate
- Whether packet delivery can be out-of-order

You need an IP address and a subnet mask for both ends of the connection. Usually, these are provided by the vendor to suit the vendor's addressing scheme.

If you have a direct connection with no Ethernet routers, the IP addresses should be in the same subnet, and the subnet mask should be same for both ends. Also, you do not need to specify any default gateway information for the link. Figure 4-1 shows an example of a direct FCIP connection.

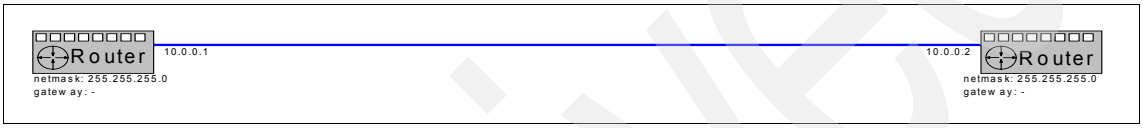


Figure 4-1 Direct FCIP connection

If you are using a routed connection, the IP addresses should be in different subnets, and you need to specify the correct default gateway address for both routers. The default gateway address is the address of the IP router in the same network with the SAN router. Figure 4-2 shows an example of a routed FCIP connection.

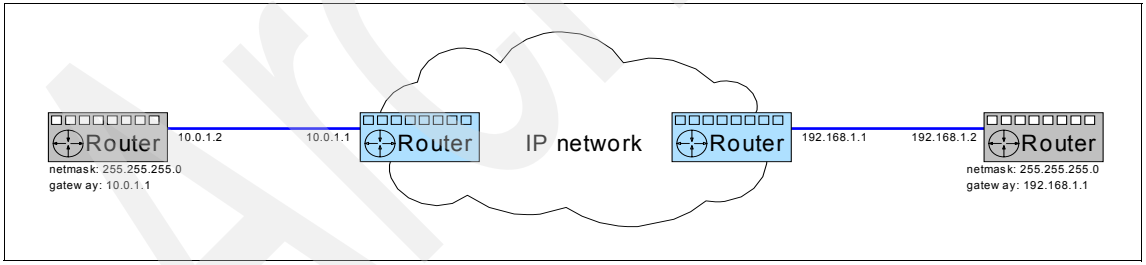


Figure 4-2 Routed FCIP connection

We also recommend that you synchronize the internal clocks of the SAN routers to an external NTP time server.

4.2 Compatibility and interoperability

An interoperability matrix is the best source of compatibility and interoperability information for a particular SAN switch or router. For the matrixes for SAN router products, refer to:

- SAN18B-R and the M48 FC Routing Blade

  ftp://service.boulder.ibm.com/storage/san/brcd/SM2005R18.pdf

- SAN16B-R

  http://service.boulder.ibm.com/storage/san/brcd/SM2109A16.pdf

We recommend that you to always check these matrixes for the most current information.

SAN18B-R and M48 FC Routing Blade interoperability

Table 4-1 lists supported switches and the required minimum OS levels at the time of writing.

Table 4-1 SAN18B-R and FC Routing Blade interoperability

Switch	Non-secure	Secure
2109-S08, 2109-S16, 3534-1RU	2.6.0 or 2.6.1	2.6.1
2109-F16, 3534-F16	3.0.2 or 3.1.2	3.1.0 or 3.1.2
2109-F32, 2109-M12	4.1	4.2
BladeCenter® SSM (3016) 2 Gbps	4.2.1a	4.2.1a
SSM 2 Gbps, 2109-F32, 2005-H08, 2005-B32, 2109-M12, 2109-M14	4.4.0	4.4.0
2005-B16, 2109-M48	5.0.1	5.0.1
BladeCenter SSM (4020) 4 Gbps	5.0.3a	5.0.3a
2109-M48 (FC3450), 2005-B64, 2005-R18	5.1.0b	5.1.0b
2109-A16	7.4.1	N/A

Note: The minimum supported Fabric OS version on SAN18B-R and M48 FC Routing Blade is 5.1.0b.

SAN16B-R interoperability

With the current firmware version, the 2109-A16 is interoperable with most current and previous IBM b-type SAN switches, as well as the Brocade SilkWorm products. Table 4-2 lists the currently supported switches and the required minimum switch Fabric OS levels.

Table 4-2 IBM TotalStorage SAN16B-R interoperability

IBM products	Brocade products	Fabric OS
2109-S08, 2109-S16	SilkWorm 2000-series	v2.6.1 or later
3534-F08, 2109-F16	SilkWorm 3000, 3200, 3600, 3800	v3.1.0 or later
2109-F32, 2109-M12	SilkWorm 3900, 12000	v4.1.0x or later
2005-H08, 2005-H16, 2109-M14	SilkWorm 3250, 3850, 24000	v4.2.0x or later
2005-B32	SilkWorm 4100	v4.4.0b or later
2005-B16, 2109-M48	SilkWorm 200E and 48000	v5.0.1a or later

The 2109-A16 supports these switches in edge fabrics in their native mode with any currently used Core PID format. The switches are also supported in the backbone fabric.

The 2109-A16 also supports interoperability with McDATA through the IBM RPQ process.

The 2109-A16 ports configured as Gigabit Ethernet are compatible with any standard Ethernet switch with a fiber connection. If you need to connect the 2109-A16 to a switch that has only copper ports, use a media converter.

The FCIP implementation of the 2109-A16 only supports tunnels to another 2109-A16.

The 2109-A16 supports the Microsoft iSCSI Initiator (current version at the time of writing is 2.02).

4.3 Availability

The SAN18B-R and SAN16B-R routers have redundant, hot-swappable power supplies and fans. The SAN16B-R currently supports hot code load, but not hot code activation. The SAN18B-R supports hot code load and hot code activation for FC ports.

For FC-FC routing and FCIP tunneling, we recommend that you install two or more SAN routers to ensure availability of the backbone fabric.

For the iSCSI gateway function, the 2109-A16 can be configured to support iSCSI failover between two routers. However, not all iSCSI clients support such a configuration.

4.4 Security

The SAN routers can be used in several different roles. The following sections describe the security features used in each role.

4.4.1 FC-FC routing security

The FC-FC routing security is based on the LSAN concept. Each LSAN is a zone that contains multiple port worldwide names (PWWNs) from separate edge fabrics. It is implemented by defining the zone in each of the edge fabrics involved. The LSAN zones are separated from all zones by having their name start with LSAN_. Note the underscore after the letters LSAN. The letters in this context are not case-sensitive. When the zones are defined, the SAN router automatically exports any shared nodes across the fabrics.

Because the LSAN zone needs to be separately defined in both edge fabrics, a zoning change in one edge fabric cannot enable unauthorized access to any node in any other edge fabric. This is especially important in cases where the edge fabrics are managed by different SAN administrators.

4.4.2 FCIP tunneling security

When you configure an FCIP tunnel between two 2109-A16 routers, you need to choose which of the routers you want to initiate the connection. The other 2109-A16 is configured to listen for connections.

The 2109-A16 configured to initiate the connection opens a connection to the FCIP well-known Transmission Control Protocol (TCP) port 3225 in the 2109-A16 that is configured to listen for the connection. If you have one or more firewalls

between the 2109-A16s, you need to allow this traffic in the firewall. We recommend that you also allow **ping** between the devices to make problem determination easier.

To enhance security, you can configure each end of the FCIP tunnel with the WWN of the SAN router in the remote end. This prevents any other router from connecting to the port.

4.4.3 iSCSI gateway security

When opening the iSCSI session, the iSCSI initiator connects to the iSCSI well-known TCP port 3260 in the router. If you have a firewall between the iSCSI client and the 2109-A16, you need to allow this traffic in the firewall.

The iSCSI gateway function of the 2109-A16 supports the optional use of Challenge Handshake Authentication Protocol (CHAP) in either one-way or two-way configurations.

In a one-way CHAP configuration, the iSCSI initiator is authenticated by the 2109-A16. The iSCSI initiator can open an iSCSI session only if the CHAP secret of the initiator matches the CHAP secret configured for the initiator in the 2109-A16. This way, the iSCSI connection is protected against any other iSCSI initiator trying to use the same iSCSI Qualified Name (IQN).

In a two-way CHAP configuration, the iSCSI initiator is first authenticated by the 2109-A16, and the 2109-A16 is then authenticated with the iSCSI initiator. The iSCSI initiator can open an iSCSI session only if the CHAP secret of the initiator matches the CHAP secret configured for the initiator in the 2109-A16, and the CHAP secret of the 2109-A16 matches the CHAP secret configured for the 2109-A16 in the iSCSI initiator. This way, the iSCSI initiator is also protected against any other device imitating the 2109-A16 iSCSI portal.

We recommend that you always use at least one-way CHAP configuration.

4.5 Performance

This section addresses the performance of SAN routers when they are used in different solutions.

4.5.1 FC-FC routing performance

The SAN18B-R and the FC Routing Blade are capable of routing traffic at a full 4 Gbps line rate on all FC ports. Therefore, in an FC-FC routing scenario, the

performance of an inter-fabric link (IFL) is practically the same as performance of a 4 Gbps inter-switch link (ISL).

The 2109-A16 is different in one respect: It supports speeds up to 2 Gbps.

Due to having a separate application-specific integrated circuit (ASIC) for each port, the 2109-A16 supports exchange-level trunking, instead of the frame-level trunking available in the switches in the IBM TotalStorage b-type SAN family.

4.5.2 FCIP tunneling performance

FCIP performance is a complex issue, because it is affected by many separate factors, such as performance of the FCIP links and latency caused by the FCIP encapsulation itself.

In general, an FCIP link has significantly lower performance and higher latency than the same link would have if it was a native Fibre Channel. Therefore, we recommend that you use FCIP tunneling only when a native Fibre Channel link cannot be implemented for technical or financial reasons.

The SAN routers implement the traffic shaping feature that enables you to limit the amount of bandwidth used for the FCIP link. This way, you can avoid overloading a wide area network (WAN) link slower than the Gigabit Ethernet interface. The routers also support jumbo frames and exchange level trunking between up to four FCIP links.

You can learn more about FCIP link performance in 4.6, “IP network issues” on page 53.

4.5.3 iSCSI performance

iSCSI gateway functionality allows small servers, which do not have high performance or availability requirements for their storage connection, access to Fibre Channel-based storage systems.

Although it is possible to share the same network interface between normal TCP/IP traffic and iSCSI traffic, we recommend that you implement a separate network for the iSCSI traffic. The iSCSI network should have low latency, less than 15 ms, and minimal packet loss. Due to the network latency requirements, we do not recommend using iSCSI over long-distance connections.

Each iSCSI portal in the 2109-A16 is capable of providing up to 450 Mbps of throughput. We recommend that, if you have iSCSI clients using 100 Mbps Ethernet connections, you separate those clients to a different iSCSI portal than the clients that are using 1 Gbps Ethernet.

4.6 IP network issues

Several factors affect the performance of an FCIP link:

- Link bandwidth
- Link latency
- TCP receive window
- Packet loss rate
- Out-of-order packet delivery

This section discusses these factors in more detail.

Ask your telecommunications company about high-quality quality of service (QoS) managed links, including information about such offerings as IP over SONET/SDH. Work closely with your telecommunications company so that they clearly understand your network quality expectations and you clearly understand the cost and management implications of any decisions you make.

4.6.1 Link bandwidth

The link bandwidth is the most obvious factor affecting performance. It is also one of the key metrics used when provisioning the link. In storage environments, the link bandwidth used should always be the *guaranteed bandwidth* from the service provider.

If the guaranteed bandwidth is anything less than one full Gigabit Ethernet, it needs to be configured into the routers in both ends of the link to use the traffic shaping features and avoid overrunning the link. We recommend that you set the maximum allowed speed of the FCIP port to 96% of the guaranteed bandwidth of the link on both ends of the link.

4.6.2 Link latency

Link latency is a metric of the round-trip time (RTT) it takes for a packet to cross the link. The key factors contributing to the link latency include:

- Distance
- Router and firewall latencies
- Time of frame in transit

Distance

The speed of light in optical fiber is approximately 208 000 km/s. Therefore, the delay caused by a fiber connection is approximately 4.8 μs/km. To calculate the round trip latency, we have to count this delay both ways.

For example, for a 100 km link, the round trip latency is approximately:

100 km x 4.8 μs/km x 2 = 960 μs

Similarly, for a 1000 km link, the round trip latency is 9600 μs, or 9.6 ms.

Router and firewall latencies

Any delay caused by routers and firewalls along the network connection needs to be added to the total latency. The latency varies a lot depending on the routers or firewalls and the traffic load. It can range from a few microseconds to several milliseconds.

You also need to remember that the traffic generally passes the same routers both ways. For round-trip latency, you need to count the one-way latency twice.

If you are purchasing the routers or firewalls yourself, we recommend that you include the latency of any particular product among the criteria you use to choose the products. If you are provisioning the link from a service provider, we recommend that you include at least the maximum total round trip latency of the link in the SLA.

Time of frame in transit

The time of frame in transit is the actual time that it takes for a given frame to pass through the slowest point of the link. Therefore, it depends on both the frame size and link speed.

The maximum size of payload in a Fibre Channel frame is 2112 bytes. The Fibre Channel headers add 36 bytes to this, for a total Fibre Channel frame size of 2148 bytes. When transferring data, Fibre Channel frames at or near the full size are usually used.

If we assume that we are using jumbo frames in the Ethernet, the complete Fibre Channel frame can be sent within one Ethernet packet. The TCP and IP headers and the Ethernet medium access control (MAC) add a minimum of 54 bytes to the size of the frame, for a total Ethernet packet size of 2202 bytes, or 17616 bits.

For smaller frames, such as the Fibre Channel acknowledgement frames, the time in transit is much shorter. The minimum possible Fibre Channel frame is one with no payload. With FCIP encapsulation, the minimum size of a packet with only the headers is 90 bytes, or 720 bits.

Table 4-3 details the transmission times of this FCIP packet over some common WAN link speeds.

Table 4-3 FCIP packet transmission times over different WAN links

Link type	Link speed	Large packet	Small packet
Gigabit Ethernet	1250 Mbps	14 µs	0.6 µs
OC-12	622.08 Mbps	28 µs	1.2 µs
OC-3	155.52 Mbps	113 µs	4.7 µs
T3	44.736 Mbps	394 µs	16.5 µs
E1	2.048 Mbps	8600 µs	359 µs
T1	1.544 Mbps	11 400 µs	477 µs

If we cannot use jumbo frames, each large Fibre Channel frame needs to be divided into two Ethernet packets. This doubles the amount of TCP, IP, and Ethernet MAC usage for the data transfer.

Normally, each Fibre Channel operation transfers data in only one direction. The frames going in the other direction are close to the minimum size.

4.6.3 TCP receive window

In addition to the line speed available, the maximum throughput available on any given TCP connection is determined by the TCP receive window of the receiving device. This is similar to the buffer-to-buffer-credit flow control used in Fibre Channel networks. To enable the full utilization of a given FCIP link, the size of TCP window allocated for the link needs to be large enough for all FCIP packets that are being transferred on the line.

4.6.4 Packet loss rate

In traditional TCP/IP networks, packet loss is a normal and accepted behavior. The built-in retransmission mechanism in the protocols handle retransmitting any dropped packet. Most protocols used in TCP/IP networks can easily handle high packet loss rates, such as 1%, without significant performance degradation.

Because FCIP uses a TCP connection for data transfer, it also uses the same mechanism. However, because latency is usually critical in storage applications, the storage networks do not cope well with retransmissions. Therefore, networks used for storage traffic need a much lower packet loss rate. Even an IP network

with a packet loss rate of 0.01% is considered a low quality network, compared to the baseline of zero frame loss in Fibre Channel networks.

4.6.5 Out-of-order packet delivery

Fibre Channel networks rely on in-order delivery of Fibre Channel frames. The SAN router receiving FCIP traffic needs to ensure that the Fibre Channel frame order is retained.

If IP packets are received out of order, as is often the case with shared IP networks using Multiprotocol Label Switching (MPLS), the SAN router needs to buffer them until it receives the complete sequence of packets.

Usually, the only effect of out-of-order packet delivery is slightly increased latency. However, in extreme cases, the receiving router can run out of buffer space and have to drop packets.

5

IBM TotalStorage b-type family real-life routing solutions

This chapter presents some real-life solutions implemented with the IBM TotalStorage b-type family routing products. We discuss the following solutions:

- Backup consolidation
- Migration to a new storage environment
- Long-distance disaster recovery over IP

Important: The solutions and sizing estimates that we discuss or make in this chapter are unique. Make no assumptions that they will be supported on or apply to each environment. We recommend that you engage IBM to discuss any proposal.

5.1 Backup consolidation

This scenario presents a solution to consolidate local area network (LAN)-free tape backups from two separate storage area network (SAN) fabrics.

5.1.1 Client environment and requirements

The client has two existing SAN fabrics and is currently using ArcServe software to back up the Microsoft Windows servers in the SAN fabrics to tape. The client also has several application servers that do not have SAN attachment. Figure 5-1 shows the client environment.

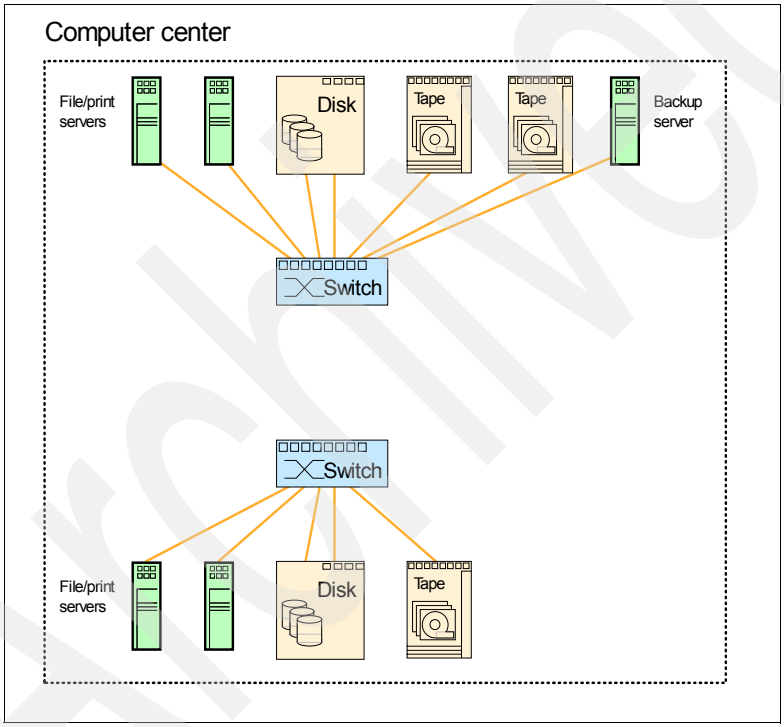


Figure 5-1 Current backup environment

The client has the following requirements for the new solution:

- Consolidate the tape backups to a single Tivoli Storage Manager environment
- Provide for LAN-free backups from both current SAN fabrics
- Implement the new backup system to a separate location from the computer center
- Leverage the existing investment in SAN hardware

In the first SAN fabric, the client currently has 160 GB of disk space. This space is projected to grow to 630 GB in the near future. In the second SAN fabric, the client has 100 GB of disk space.

5.1.2 The solution

Our solution has the following new components:

- IBM System p™ server for Tivoli Storage Manager
- IBM 3583-L72 tape library with four Fibre Channel (FC) drives
- IBM TotalStorage SAN32B-2 switch for the backup environment
- IBM TotalStorage SAN16B-R router, with FC-FC routing enabled

The following graphic shows the new backup environment (Figure 5-2).

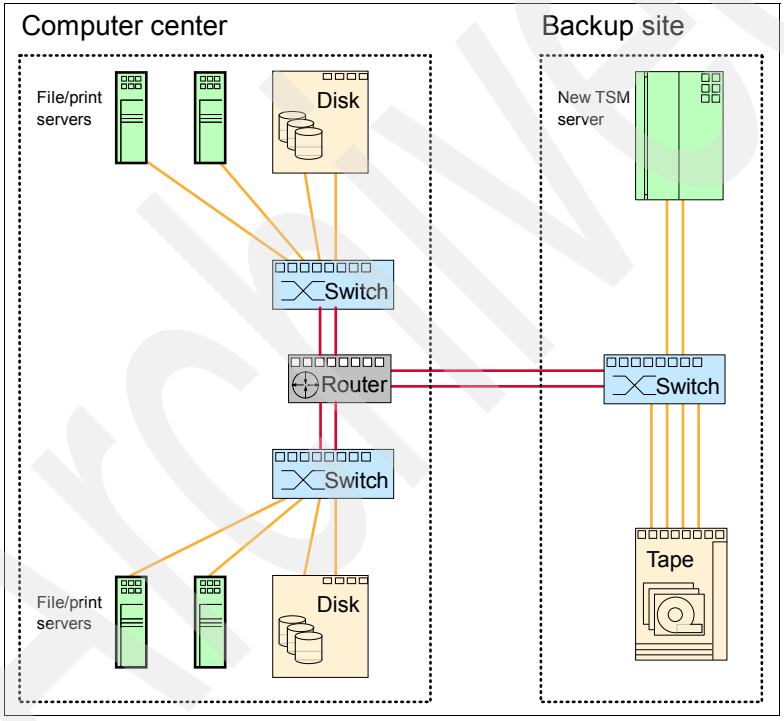


Figure 5-2 New backup environment

We locate the router in the computer center to minimize the need of fiber connections between the computer center and the backup site. All other components are located in a single rack at the backup site.

We connect each of the current fabrics, and the new backup switch, to the router with two inter-fabric links (IFLs) for redundancy. The client provides the two long wave fiber connections required between the computer center and the backup site.

The Tivoli Storage Manager server will use its internal disks for both Tivoli Storage Manager databases and disk storage pools. Therefore, it does not need any access to the existing client SAN fabrics. The tape drives are divided evenly to the two Fibre Channel adapters in the Tivoli Storage Manager server.

We create a separate logical SAN (LSAN) zone for each server in any SAN fabric that needs access to the tape drives. The LSAN will contain the port worldwide name (PWWN) of the host bus adapter (HBA) of the server and the PWWNs of all the tape drives. We also ask the client to create the same LSANs in the existing SAN fabrics.

Because we use our new environment only for daily backups, it does not have the high availability requirements that SAN fabrics use for disk access. Therefore, it is adequate to have a single backup switch and a single router in the solution.

The application servers that are not connected to any SAN fabric are backed up to the Tivoli Storage Manager server over a LAN connection.

5.1.3 Failure scenarios

This section explains how the failure of different components affects the operation of our solution:

- Power failure

  The Tivoli Storage Manager server, the tape library, and all SAN fabric components in the environment have dual redundant power supplies connected to different power circuits. Therefore, a power failure in one circuit does not have any effect on operation.

- IFL failure

  If an IFL fails, the system remains operational, but the maximum bandwidth available is reduced by 50%.

- Router failure

  If the SAN router fails, it is impossible to run LAN-free backups. In this situation, the Tivoli Storage Manager client automatically uses a LAN-based method for any backups and restores. The Tivoli Storage Manager server and the servers that are not using LAN-free backups are not affected.

- Backup switch or Tivoli Storage Manager server failure

  The failure of either the backup switch or the Tivoli Storage Manager server prevents any backup and restore activity.

5.2 Migration to a new storage environment

This scenario presents a solution to migrate the client's current storage environment to a new environment.

5.2.1 Client environment and requirements

The client has a Hewlett-Packard (HP) XP512 storage system that is shared between AIX 5L, HP-UX, and Windows servers. Due to historical reasons, each server platform has its own SAN fabrics and connections to the XP512.

Each SAN fabric consists of a single 16-port, 1 Gbps Brocade 2800 switch. Because the lease period of the environment expires within a few months, the client needs a new solution to replace the current environment. Figure 5-3 on page 62 shows the initial environment. For clarity, we show only some of the servers.

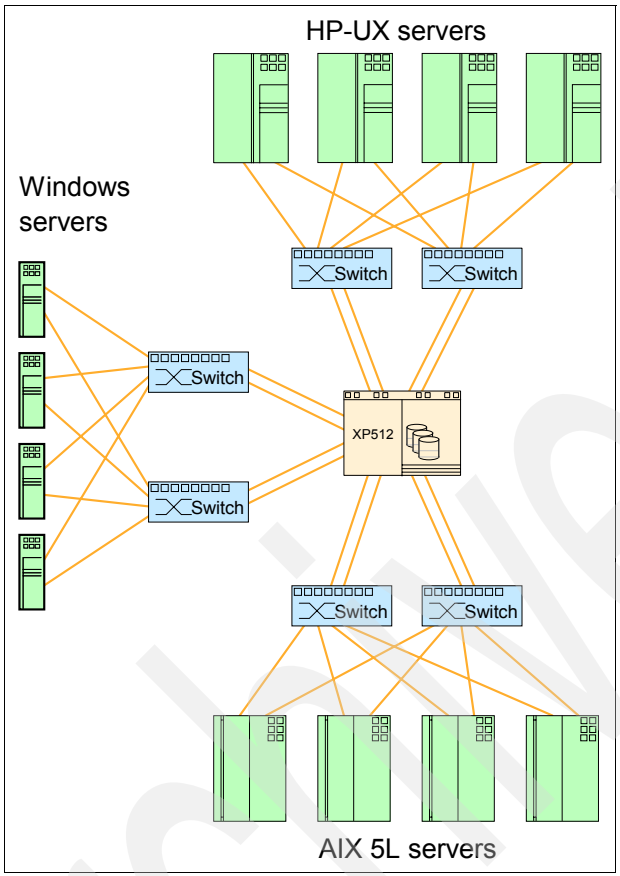


Figure 5-3 Initial storage environment

The client has the following requirements for the new solution:

- New hardware to replace the current disk system and SAN fabric
- Flexibility in allocating ports between different platforms
- Scalability to support future applications
- Minimize the amount of downtime of servers due to migration

5.2.2 The solution

Our solution has the following new components:

- IBM TotalStorage DS8100 disk subsystem
- Two IBM TotalStorage SAN256B Directors with 64 ports each
- Two IBM TotalStorage SAN18B-R routers

We install the components of the new storage environment, and connect the environment to the old environment with IFLs, as shown in Figure 5-4.

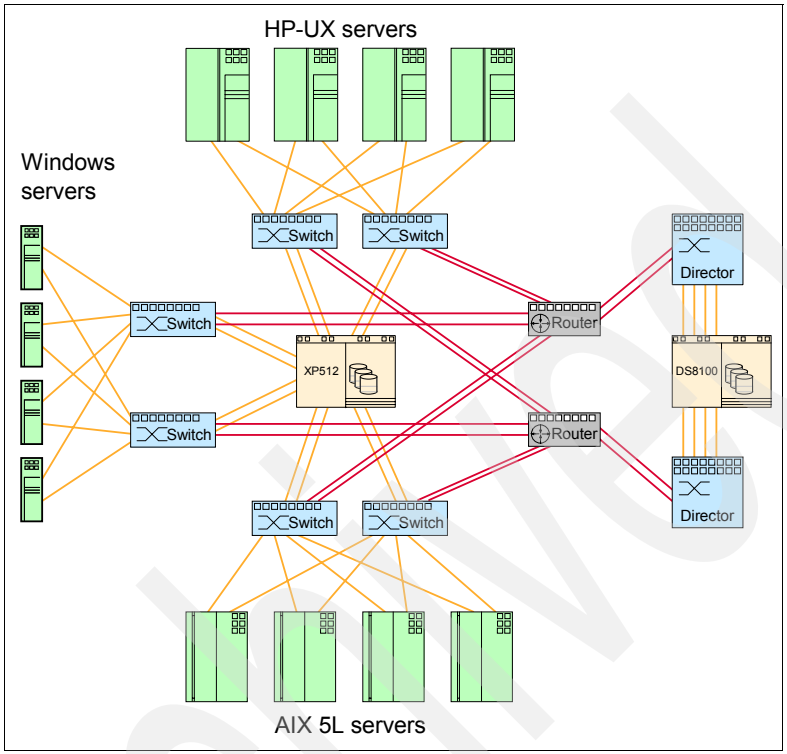


Figure 5-4 Interim environment for migration

In the new environment, all TotalStorage DS8100 ports are shared among all servers. Because we are only migrating a few servers at the same time, using a limited number of IFLs does not cause any performance degradation to the servers.

When the new storage environment is completely installed, we start migrating the servers one server or a group of servers at a time, using the following procedure:

1. Create LSANs to allow the server to access TotalStorage DS8100.
2. Install the IBM Subsystem Device Driver (SDD) package and any other TotalStorage DS8100-specific software on the server.
3. Allocate new storage in TotalStorage DS8100 to the servers.
4. Migrate all server data from the old storage to the new storage using the following operating system-based tools:
   - Native Logical Volume Manager (LVM) for AIX 5L

- PVLinks for HP-UX
- Veritas Volume Manager for Windows

5. Create non-LSAN zones to allow the server to access the storage from the new SAN fabrics.
6. Disconnect the server from the old switches and move it to the new directors.
7. Delete the LSANs created in step 1.

The only step that requires server downtime in the procedure is step 6. If the new cabling is prepared before, this step should take little time.

After the migration of all servers is complete, no servers should be connected to the old switches and the XP512 should be idle. At this time, we can remove the old storage hardware from the environment. The IBM TotalStorage SAN18B-R routers are also freed and can be used for other purposes, such as SAN extension with FC over IP (FCIP).

Figure 5-5 shows the final storage environment.

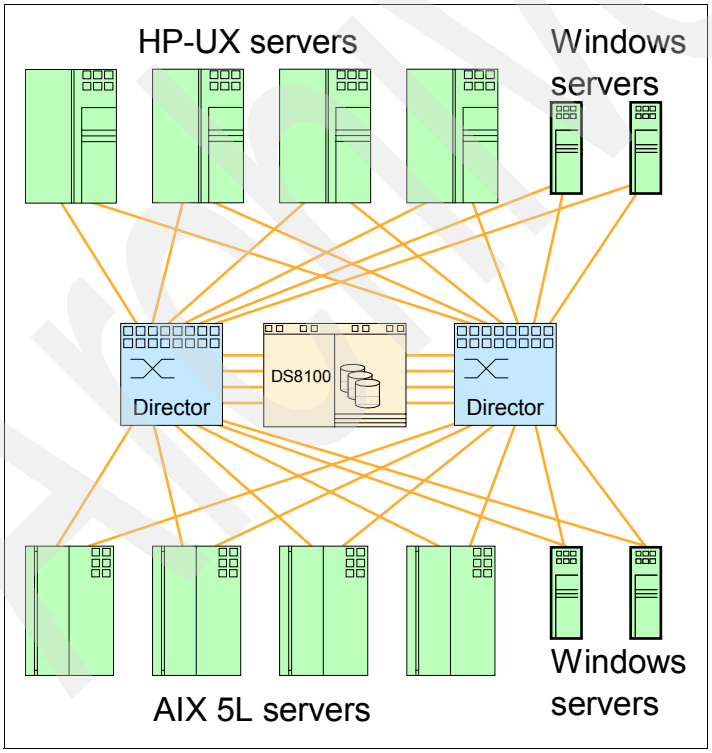


Figure 5-5 Final storage environment

5.3 Long-distance disaster recovery over IP

This scenario presents a solution that provides for long-distance disaster recovery over an IP connection.

5.3.1 Client environment and requirements

The client has three different SAN islands that need to be connected:

- Development SAN in the primary site
- Production SAN in the primary site
- Disaster recovery SAN in the disaster recovery site

The distance between the primary site and the disaster recovery site is 600 km. The amount of data in the productive environments is expected to grow to 5 TB within two years, and we expect that 3% of the data is changing in the peak hour.

The client has the following requirements for the solution:

- Provide asynchronous replication for production data from the primary site to the disaster recovery site, with a five minute recovery point objective (RPO) and a five minute recovery time objective (RTO)
- Keep the dual fabrics of each SAN both physically and logically separate
- Provide access to a point-in-time copy of productive data from the test environment at the development SAN
- Provide for LAN-free backup from the development network to the tape library in the productive network

The current environment contains the following components:

- Production environment at the primary site
  - Dual SAN fabrics, based on IBM TotalStorage SAN256B Directors
  - IBM TotalStorage DS8100 disk subsystem with eight Fibre Channel ports
  - IBM TotalStorage 3584 tape library with six IBM 3592 tape drives
  - Eight IBM eServer™ pSeries® servers with dual Fibre Channel adapters
  - Sixteen IBM eServer xSeries® servers with dual Fibre Channel adapters
- Development environment at the primary site
  - Dual SAN fabrics, based on IBM TotalStorage SAN 32B-2 switches
  - IBM TotalStorage DS6800 disk subsystem with four Fibre Channel ports
  - Eight pSeries servers with dual Fibre Channel adapters
  - Sixteen xSeries servers with dual Fibre Channel adapters

- Disaster recovery environment at the disaster recovery site
  - Dual SAN fabrics, based on IBM TotalStorage SAN256B Directors
  - IBM TotalStorage DS8100 disk subsystem with eight Fibre Channel ports
  - IBM TotalStorage 3584 tape library with six IBM 3592 tape drives
  - Eight pSeries servers with dual Fibre Channel adapters
  - Sixteen xSeries servers with dual Fibre Channel adapters

Figure 5-6 shows the environment. For clarity, we show only some of the servers and connections.

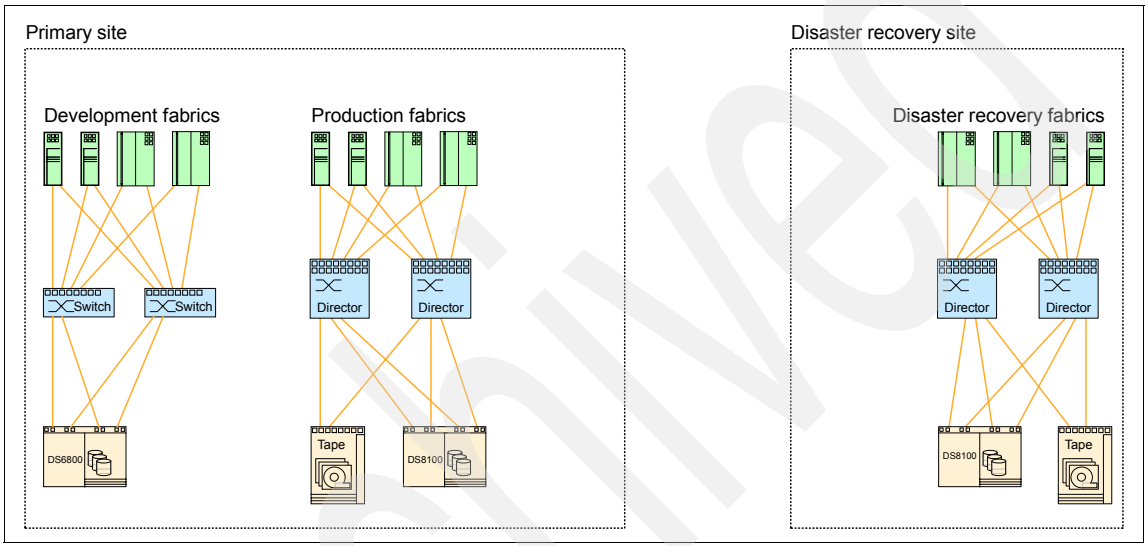


Figure 5-6 Client environment

5.3.2 The solution

Our solution has the following components:

- TotalStorage DS8100 Global Mirroring feature for asynchronous replication
- Four IBM TotalStorage SAN18B-R routers (2005-R18), with FCIP tunneling feature activated
- Four IP links between the SAN18B-R routers from the primary site to the disaster recovery site
- IBM enterprise Remote Copy Management Facility (eRCMF) software to provide automatic failover of both pSeries and xSeries servers

Figure 5-7 shows the complete solution.

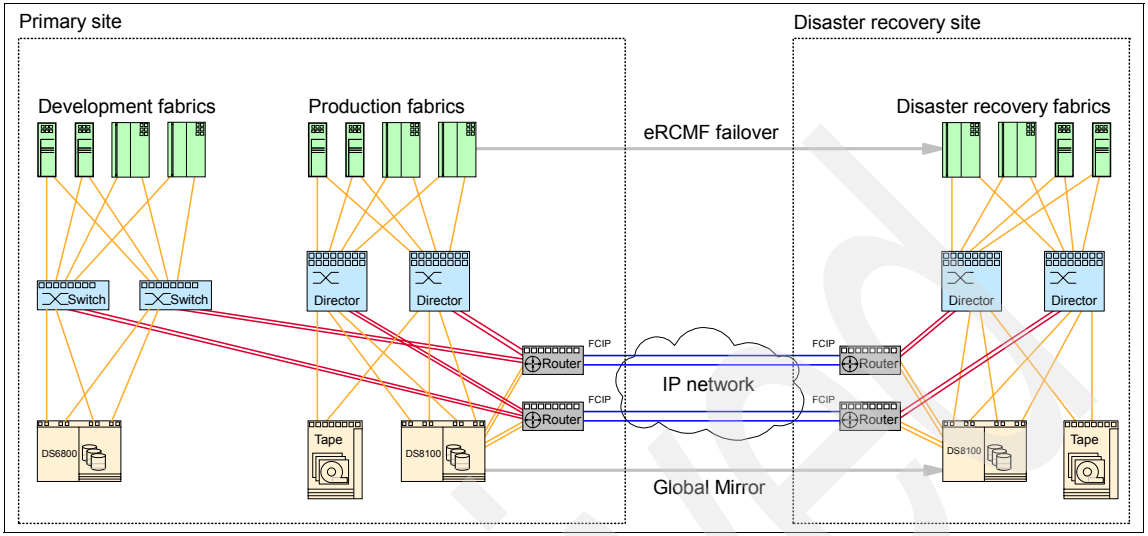


Figure 5-7 Disaster recovery solution

FCIP link sizing

Because we are using only the FCIP links for Global Mirror between the TotalStorage DS8100 systems, we only need to take into account any changes to the data when sizing the links.

Based on the client requirements, the amount of data changing during the peak hour is 3% of 5 TB, or 150 GB. If we assume that the changes are evenly divided over the hour, the changes are 2.5 GB/min, or approximately 42 MBps, or 336 Mbps. We use this number as the basis for our link sizing.

If we divide the amount evenly across four links, we get a traffic of 84 Mbps over each link. However, to allow the loss of one link or any peaks in the traffic, we divide the traffic only across three links, for 33% extra bandwidth and 112 Mbps traffic over each link. We also plan to have a maximum of 90% utilization on the link, so the minimum link speed we need is 125 Mbps.

Each link can be implemented over an OC-3 line that has the capacity of 155 Mbps. An alternative is to use a Multiprotocol Label Switching (MPLS)-based, shared connection. However, due to possible router latency issues, we prefer the private OC-3 -based connection.

The most significant part of the OC-3 link latency is the propagation time of the light within the fiber. For a 600 km connection, with 1200 km round trip, it is

1200 x 4.8 μs, or 5.8 ms. We round this to 6 ms to account for the packet transmission time over the 155 Mbps OC-3 link.

5.3.3 Normal operation

In normal operation, the production servers use only TotalStorage DS8100 disks in the primary site. The disaster recovery servers are connected to TotalStorage DS8100 in the disaster recovery site, but do not have the disk mounted or any applications running.

The development servers use TotalStorage DS6800 disks in the primary site, and some capacity from TotalStorage DS8100 in the primary site.

For TotalStorage DS8100 disk subsystems, four of the eight ports are used for host attachments and the remaining four are used for Global Mirroring. The ports used for Global Mirroring are directly connected to the routers.

In addition to normal zoning, we define the following LSANs in our environment:

- Separate LSANs for the HBAs of any server in the development fabric that needs access to TotalStorage DS8100, containing:
  - The HBA of the server
  - Both Fibre Channel ports of TotalStorage DS8100 used for host attachment in the fabric
- Separate LSANs for the HBAs of any server in the development fabric that needs access to LAN-free backup, containing:
  - The HBA of the server
  - All Fibre Channel ports of the tape drives in the primary site connected to the fabric

In addition, we define zones for each Global Mirror connection in the backbone fabric.

5.3.4 Failure scenarios

This section explains how the failure of different components affects the operation of our solution:

- Power failure

  All of the SAN fabric components in the environment have dual redundant power supplies connected to different power circuits. Therefore, a power failure in one circuit does not have any affect on operation.

- FCIP link failure

  The failure of a single FCIP link reduces the available bandwidth between the sites by 25%. However, because we assumed three available links in our sizing, the performance of the system will remain adequate.

- Development fabric switch failure

  The failure of a switch in development fabric reduces the Fibre Channel bandwidth available for development and test servers by 50%. The traffic is automatically routed through the remaining paths by the SDD. The production environment is not affected.

- Primary site router failure

  If one of the routers at the primary site fails, the capacity of the Global Mirror connection will be reduced by 50%. However, because we rounded up our link speed, we still have about 300 Mbps or about 90% of the peak hour capacity available. In addition, it reduces the Fibre Channel bandwidth available between development and test servers, and the storage in the production fabrics, by 50%.

- Primary site director failure

  Director failure at the primary site reduces the Fibre Channel bandwidth available for production servers by 50%. It also reduces the Fibre Channel bandwidth available between development and test servers, and the storage in the production fabrics, by 50%.

- Disaster recovery site router failure

  If the router at the disaster recovery site fails, the capacity of the Global Mirror connection is reduced by 50%. However, because we rounded up our link speed, we still have about 300 Mbps or about 90% of the peak hour capacity available.

- Disaster recovery site director failure

  Director failure at the disaster recovery site reduces the Fibre Channel bandwidth available for disaster recovery servers by 50%. However, in normal situations, those servers are idle, so this reduction affects the system only in the case where the production workload is already running at the disaster recovery site.

- Primary site TotalStorage DS8100 port failure

  If a port used for host access in TotalStorage DS8100 at the primary site fails, the Fibre Channel bandwidth available for host access is reduced by 25%.

  If a port that is used for Global Mirror in TotalStorage DS8100 at the primary site fails, the remaining Fibre Channel ports can sustain the full Global Mirror performance.

- Primary site TotalStorage DS8100 failure

  If TotalStorage DS8100 at the primary site fails, all hosts lose access to it. This event can be promoted to site failure, and production can resume at the disaster recovery site.

- Primary site TotalStorage DS8100 port failure

  If a port used for host access in TotalStorage DS8100 at the disaster recovery site fails, the Fibre Channel bandwidth available for host access is reduced by 25%. However, in normal operation, those servers are idle, so this reduction affects the system only in the case where the production workload is already running at the disaster recovery site.

  If a port that is used for Global Mirror in TotalStorage DS8100 at the primary site fails, the remaining Fibre Channel ports can sustain the full Global Mirror performance.

- Disaster recovery site TotalStorage DS8100 failure

  If TotalStorage DS8100 at the disaster recover site fails, the Global Mirror connections change to a *suspended* state. TotalStorage DS8100 at the primary site will accumulate changes to the data, and copy the changed data over to the disaster recovery site, when TotalStorage DS8100 becomes available.

- Primary site failure

  If the complete primary site fails, the IBM eRCMF software starts the production at the disaster recovery site automatically. Although manual failover is also possible, it is difficult to manually reach the RTO target.

6

IBM TotalStorage b-type router implementation

In this chapter, we discuss the implementation of the IBM TotalStorage b-type SAN router products. We perform the steps required to install and configure the router and perform basic management functions, including upgrading firmware and implementing it into an existing SAN infrastructure, creating a Meta SAN.

The chapter has two distinct sections:

- Installing and configuring SAN16B-R

  SAN16B-R is an older product and is configured in a different manner than the two newer products. It also uses XPath OS, while the newer products use standard Fabric OS.

- Installing and configuring SAN18B-R and the M48 FC Routing Blade

  We configure these two SAN routers in a very similar manner, so they are both discussed in one common section.

6.1 Installing the SAN16B-R (2109-A16)

During the implementation of this router, we perform the following steps:

1. Set up IP addresses through the serial port.
2. Upgrade the switch to the latest supported firmware version.
3. Attach the router to the LAN and connect to its management interface (WebTools).
4. Configure FCIP links between two routers at different sites.
5. Configure EX_Ports between the routers and their local SANs.
6. Set up LSAN zones allowing the routing of specific devices.

6.1.1 Initial setup

The first step is to physically install the 2109-A16 into a rack. This should be performed by an IBM engineer. After this is completed, connect the appropriate power cables to the router and turn on.

After the router completes its power-up sequence, we need to perform some initial configuration steps.

Power-up sequence

This should not take any more than three minutes to complete. During the POST sequence, the port LEDs will light up in sequence from left to right. If you see no change on the front of the router after three minutes, turn it off one minute, and then try turning it on again. After the POST completes successfully, the dc and ac LEDs, including the system LED, will be green.

Connecting to the serial port

To connect to the serial port, perform the following steps:

1. To connect to the router serial port, we need to use the provided console cable. It is a rolled RJ45-to-RJ45 cable with an RJ45-to-DB9 converter at one end. The Standard b-type straight-through (non-null modem) serial cable will not work. First, we connect the RJ45 end to the console port on the router, which is the right-most port of the three RJ45 ports, labeled “10101”. Connect the DB9 end to either a mobile computer or a Windows machine that is close to the router.

2. Start Hyperterm from the PC or mobile computer. Select the appropriate COM port that is associated to the serial connection (Figure 6-1).

Figure 6-1 Selecting the correct COM port

3. The settings for this connection are the same as all the other b-type family of switches: **9600**, **8**, **None**, **1**, **None**, as shown in Figure 6-2.

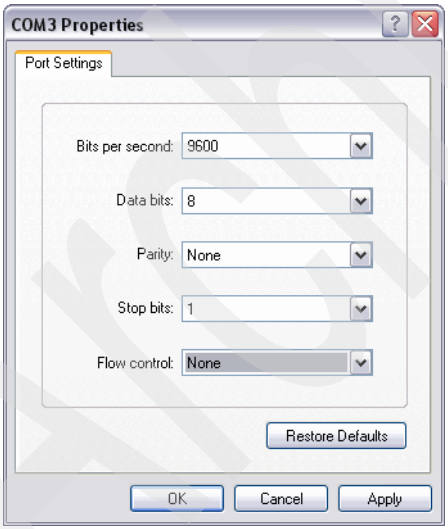


Figure 6-2 Hyperterm connection settings

4. After clicking the **OK** button, we press Enter on the keyboard to get a login prompt.
5. Log in. The default login details are also the same as all other b-type switches. The login ID is **`admin`** and default password is **`password`**.

After entering the default login and password details, we press Ctrl+c to escape from the request to change the default password. We perform this at a later date.

Note: All commands are **bold** in the output that follows. Other items are **bold** to help clarify the current discussion.

Example 6-1 Initial login

```
Login: admin
Password:
Last login: Thu Nov 3 13:25:43 2005 on ttyOO
Please change your passwords now.
Use Control-C to exit or press 'Enter' key to proceed.
^C
Password was not changed. Will prompt again at next login
until password is changed.

Please remember to config time zone for the switch

IBM_2109_A16:admin>
```

Setting the IP address

Now, we are connected and logged in to the switch. The first task is to set up the IP address so that we can add the router to our LAN and start to manage it through the GUI interface. Therefore, we need to use the **ipaddrshow** command to display our current IP settings, and use the **ipaddrset** command to change these settings.

Example 6-2 Displaying the IP address

```
IBM_2109_A16:admin> ipaddrshow
MGMT 1          Configuration           Current
IP Address      10.77.77.77             10.77.77.77
Net Mask        255.255.255.0           255.255.255.0
Gateway         10.77.77.71             10.77.77.71
MGMT 2          Configuration           Current
IP Address      0.0.0.0                 0.0.0.0
Net Mask        0.0.0.0                 0.0.0.0
Gateway         0.0.0.0                 0.0.0.0
IBM_2109_A16:admin>
```

From the output of our **ipaddrshow** command, we can see the default IP address is 10.77.77.77, the default subnet mask is 255.255.255.0, and the default gateway is 10.77.77.71.

We can also see that two Ethernet interfaces are represented in the output. The router has two physical external Ethernet connections. The connection on the right represents MGMT 1 and the port on the left represents MGMT 2. We only need to use MGMT 1. The second interface was designed to work with future application plans that might require some type of failover IP functionality.

So, now we use the **ipaddrset** command to set our own LAN settings. This command is not interactive as with the Fabric Operating System (FOS) but requires command line parameters as with any standard UNIX® flavor. The structure of the command is as follows.

Example 6-3 Changing the IP address

```
IBM_2109_A16:admin> ipaddrset
usage: ipaddrset <mgmt interface num> <-i ipAddress> <-n netmask> <-g
gateway> <-a action> [-s] [-r]
where
 mgmt interface num: 1 or 2
 -a action: "cfgnow" or "cfgafterreboot"
 -s: to set the switch virtual IP as same as the IP of the management
interface
 -r: to reset IP configuration of the management interface 2
```

Now, we can enter our IP details and specify that we want the configuration to take place immediately (using the **-a cfgnow** option) and not after the next reboot.

Example 6-4 Entering the IP details

```
IBM_2109_A16:admin> ipaddrset 1 -i 10.64.210.204 -n 255.255.240.0 -g
10.64.208.1  -a cfgnow -s
Management ethernet interface configuration for port 1 is set
The switch virtual ip is set
IBM_2109_A16:admin> ipaddrshow

MGMT 1          Configuration           Current
IP Address      10.64.210.204           10.64.210.204
Net Mask        255.255.240.0           255.255.240.0
Gateway         10.64.208.1             10.64.208.1

MGMT 2          Configuration           Current
IP Address      0.0.0.0                 0.0.0.0
```

```
Net Mask        0.0.0.0                 0.0.0.0
Gateway         0.0.0.0                 0.0.0.0
```

From the command, we can see a reference to a virtual address. This can be configured, but is not necessary. Use the **svipaddrshow** command to see the details of the virtual address.

Example 6-5 Details of the virtual address

```
IBM_2109_A16:admin> svipaddrshow
The switch virtual IP   configuration   current
IP address              10.64.210.204   10.64.210.204
Netmask                 255.255.240.0   255.255.240.0
Gateway                      -          10.64.208.1
IBM_2109_A16:admin>
```

We physically connect the MGMT 1 interface onto our LAN and point our Web browser at the router's IP address to access the WebTools management GUI.

However, at this point, we continue using the CLI to upgrade the current firmware version on this router.

Upgrading the firmware

The b-type router currently does not use the standard Brocade Fabric Operating System (FOS) but instead uses XPath OS. XPath OS is built on NetBSD UNIX.

The XPath OS splits the router's flash memory into two different banks. You can view these banks using the **firmwarestatus** command, which we show later. These banks are architected so that one is active and the other Inactive. On initiating a firmware download, the code is loaded into the inactive bank first. It can then be copied over to the active bank manually or during a reboot.

There are two parts of the XPath OS firmware of which we need to be aware: first the recovery kernel (RK) and second the base operating system. The recovery kernel does not commonly require updating with minor code level increases, for example, from 7.3.0a to 7.3.0b. However, it does require updating during a major code update, for example, from 7.3.x to 7.4.x. We are upgrading from 7.3.0b to 7.4.0b and, therefore, need to upgrade the RK.

The command we use to perform a firmware upgrade is **firmwaredownload**. We can choose whether we will just upgrade the operating system or the recovery kernel as well. When upgrading the recovery kernel, this process automatically upgrades the operating system.

Downloading the code

To download the latest firmware for the 2109-A16 router, we start at the following URL:

http://www.ibm.com/servers/storage/support/san/2109a16/downloading.html

From, here we click the **Downloadable files** link. Note that this redirects us to the IBM support section of the Brocade Web site, as shown in Figure 6-3.

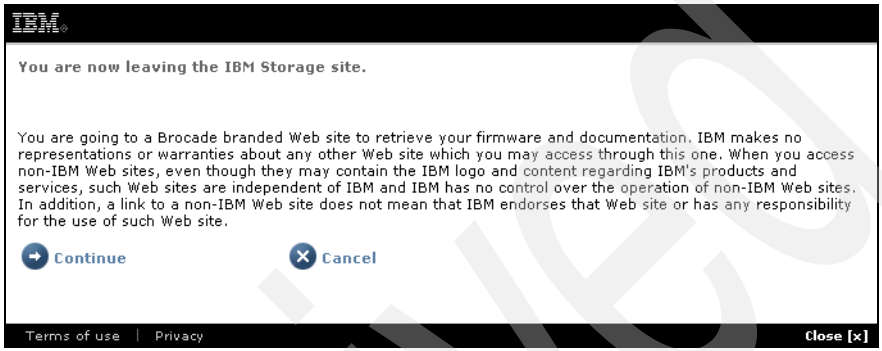


Figure 6-3 IBM redirect warning

From the Brocade Web site, we click the appropriate version of the XPath OS firmware to download. After downloading this code, we need to decide on an FTP mechanism to get the code to the router. If you use Fabric Manager 5.0 with its built-in FTP server, you can simply download code onto this machine. Otherwise, you need to copy the firmware onto a machine that has an FTP server available to which the router can contact. After downloading the code and copying it to a suitable machine, we unzipped it.

Installing the code

To install the code, perform the following steps:

1. We use the **version** command from the router CLI to display the currently running code level.

Example 6-6 Displaying the code version

```
IBM_2109_A16:admin> version
RPG file server         : 10.64.209.27
Root directory          : /MPR/xpath_os_v7.3.0b
FTP username            : akirk
FTP password            : ******
Download protocol       : ftp

==================
```

```
Installed Packages:
===================
Package Name            : xpath_os_v7.3.0b
Installed from          : bank2
Installed date          : Apr 13  2005
Administrative status   : (1)
Primary status          : up
Secondary status        : installed and running

Disk usage on root fs -  Total: 198 Mbytes, Free:  93 Mbytes.

IBM_2109_A16:admin>
```

We see version 7.3.0b installed. Therefore, we proceed with our upgrade to version 7.4.0b.

2. The first task is to back up the configuration of our router using the **configupload** command.
3. Next, we use the **firmwaredownload** command with the following syntax.

Example 6-7 firmwaredownload command

```
switch:admin> firmwaredownload <ftp server ip address> <ftp userid>
<path to firmware using forward flashes / regardless of platform>
<password for userid>
```

This fetches the appropriate firmware from the FTP server. In the following example, we use the built-in FTP server within Fabric Manager 5.0.

4. Before starting the firmware download process, it is always a good idea to alter the 10 minute telnet session timeout value, just in case. To do this, we use the **timeout** command.

Example 6-8 Setting the idle timeout value

```
IBM_2109_A16:admin> timeout
Current IDLE Timeout is 10 minutes
IBM_2109_A16:admin>
IBM_2109_A16:admin> timeout 0
IDLE Timeout Changed to 0 minutes
The modified IDLE Timeout will be in effect after NEXT login
IBM_2109_A16:admin>
```

5. With the unzipped firmware file, make a note of the path to the RK if that is what you are intending to upgrade. If you are simply upgrading the operating system itself, you will point at the base file:

IBM_2109_A16:admin> firmwaredownload 10.64.223.17 fm /Switches/7.x/v7.4.x/v7.4.0b/xpath_os_v7.4.0b/xpath_rk_v1.5.2 brocade

6. Upgrading the recovery kernel causes the switch to reboot during **firmwaredownload**. Your Telnet session will be lost as the switch reboots automatically at least once. The whole process can take up to 25 minutes.

Example 6-9 Firmware update

```
Starting to download the package......
###
Starting to install the package......
####################
Nov  3 13:51:03 reboot: rebooted by root


firmwaredownload[9386]: SERVER SIGTERM: Exiting ...
shell[8260]: SERVER SIGTERM: Exiting ...
syncing disks... done
rebooting..
```

Here it is important to note that even though the XPath OS version is 7.0.4b, the RK is at version 1.5.2.

Important: Note that the recovery kernel file name format here is "xpath_**rk**_v1.5.2". Choosing any other file name format is incorrect for a recovery kernel.

7. During the upgrade process, the router restarts on the new RK and also loads the new base OS. It restarts one final time when all the new code is in place. After booted onto the new code, we log in to the switch and use the **firmwareshow** command to check the status of the flash memory banks and to watch the switch commit the new code to the inactive bank.

Example 6-10 Issuing firmwareshow command

```
IBM_2109_A16:admin> firmwareshow
=========================== Active (Bank 1) Version
===========================
==================
Installed Packages:
==================
Package Name:      xpath_os_v7.4.0b
Install Date:      Nov  3, 2005 13:55
```

```
========================== Inactive (Bank 2) Version
==========================
Firmware download or commit in progress...
IBM_2109_A16:admin>
```

8. Here, we can see the firmware commit process, and it is currently copying the code from the active bank to the inactive bank. If we run the command again a few minutes later, the process should have completed.

Example 6-11 Issuing the firmwareshow command again

```
IBM_2109_A16:admin> firmwareshow
========================== Active (Bank 1) Version
==========================
==================
Installed Packages:
==================
Package Name:      xpath_os_v7.4.0b
Install Date:      Nov  3, 2005 13:55


========================== Inactive (Bank 2) Version
==========================
==================
Installed Packages:
==================
Package Name:      xpath_os_v7.4.0b
Install Date:      Nov  3, 2005 13:55

IBM_2109_A16:admin>
```

The firmware is now in place and copied into both flash memory banks.

It is possible to manually manipulate the code load process to enable different versions of code in each bank. There is a CLI command that enables you to choose which bank from which to boot. For further details about how to perform this operation, refer to the *XPath OS Administrator's Guide* available from the Brocade Web site for further instructions.

We use the **version** command to check the currently installed code version.

Example 6-12 Displaying the current version

```
IBM_2109_A16:admin> version
==================
Installed Packages:
```

```
==================
Package Name:      xpath_os_v7.4.0b
Install Date:      Nov  3, 2005 13:55
BM_2109_A16:admin>
```

The output clearly shows our firmware download process completed successfully. For further firmware upload/download options, refer to the *XPath OS Administrator's Guide*.

Initial router settings

Before starting the configuration process of adding this router into a Meta SAN, ensure that the time, date, and time zone are correct, and that the correct licenses are installed.

Time services

All routers within a SAN backbone fabric or Meta SAN need their time set correctly. In fact, best practices recommend that all routers and switches (where available) use the NTP service to synchronize their time to a known good external time source.

To correctly set the router's date, we use the **date** command, but note the format is different from FOS. Here, we use **date *YYMMDDHHMM***.

Example 6-13 Setting the date and time

```
IBM_2109_A16:admin> date 0511161152
Wed Nov 16 11:52:00 UTC 2005
IBM_2109_A16:admin>
```

Now, we set up the time zone and NTP services. For setting the time zone, we use the **timezoneset** command. This initiates an interactive menu system, enabling us to select the correct time zone.

Example 6-14 Setting the time zone

```
IBM_2109_A16:admin> timezoneset
Please select a continent or ocean
 1). Africa            2). America           3). Antarctica        4).
Arctic Ocean
 5). Asia              6). Atlantic Ocean    7). Australia         8).
Europe
 9). Indian Ocean     10). Pacific Ocean    11). US               12).
Canada
Enter the option #: 11
Please select a country or city
```

```
 1). Alaska            2). Aleutian          3). Arizona           4).
Central
 5). East-Indiana      6). Eastern           7). Hawaii            8).
Indiana-Stake
 9). Michigan         10). Mountain         11). Pacific          12).
Pacific-New
13). Samoa
Enter the option #: 11
time zone is set
IBM_2109_A16:admin> date
Wed Nov 16 06:38:02 PST 2005
IBM_2109_A16:admin>
```

Now, we use the **tsclockserver** command to set an NTP server address for the router to talk with. The command syntax is **tsclockserver "*ip address*"**. By default, NTP is set to local or LOCL.

Example 6-15 Setting the NTP server address

```
IBM_2109_A16:admin> tsclockserver
tsclockserver LOCL
IBM_2109_A16:admin> tsclockserver "9.149.21.143"
tsclockserver is set
IBM_2109_A16:admin> tsclockserver
tsclockserver 9.149.21.143
IBM_2109_A16:admin>
```

Licensing

The following licenses are available for a 2109-A16:

- Web
- Zoning
- Base Switch
- Trunking
- Fibre Channel Routing Services
- Ports on Demand
- FCIP

At this time, the router, by default, ships with a Web license, Zoning license, Base Switch license, and Trunking license. The optional licenses are FC Routing Services, FCIP, and an 8-15 Ports on Demand (POD) upgrade license.

To view the currently installed licenses, we use the CLI command **licenseshow**.

Example 6-16 Displaying licenses

```
IBM_2109_A16:admin> licenseshow
License Key: SyRcRydee9TzSdSv
               Web

License Key: bzddSQyQzRcOeeTt
               Zoning

License Key: cy9cQb9yzdzAdRI
               Base switch license

License Key: beQReccdcdkfRSfn
              Trunking

License Key: beQReccdcdcfTSfh
              Fibre Channel Routing Services

License Key: beQReccdcdcvRSfv
              Ports on Demand - enable all 16 ports

License Key: beQReccdcdcnRSfn
              FCIP

IBM_2109_A16:admin>
```

Licenses are added or removed using either the **licenseadd** or **licenseremove** command.

Example 6-17 Removing and adding licenses

```
IBM_2109_A16:admin> licenseremove "bzddSQyQzRcOeeTt"
license bzddSQyQzRcOeeTt is removed
Zoning license is removed, please reboot the switch
IBM_2109_A16:admin>
IBM_2109_A16:admin> licenseadd "bzddSQyQzRcOeeTt"
license bzddSQyQzRcOeeTt added
IBM_2109_A16:admin>
```

The implementation we perform requires the FC Routing Services license and the FCIP license. These are already installed on our router, so we can continue.

This completes the initial setup of a 2109-A16 router.

Because we are connecting two routers together through an FCIP link to form a backbone fabric, each router device in this fabric needs the same initial setup performed:

1. Configure an IP address.
2. Upgrade to the latest firmware.
3. Set up the time services correctly.
4. Ensure that the appropriate licenses are installed.

6.1.2 WebTools introduction

After completing the initial setup of the router, including the assignment of an IP address to the MGMT1 interface, we can connect a patch/drop cable into this port and physically connect the router to the LAN.

Check that the router is accessible on your LAN by pinging its IP address. Assuming this ping completes successfully, we can now access the GUI management interface, WebTools.

Note: Many of the screen captures seen here from WebTools were taken after the router had been added into a configuration, enabling us to see device details rather than blank forms.

6.1.3 WebTools prerequisites

WebTools is supported on three distinct platforms: Microsoft Windows, Sun™ Solaris™ and Red Hat Linux®. All of these platforms require Java Plugin 1.4.2_06. Using a Windows platform requires the use of Microsoft Internet Explorer® 6.0, and using either Solaris (2.8 or 2.9) or Red Hat requires Mozilla_1.6.

The recommend amount of video RAM is 8 MB. Also note the following system RAM requirements:

- For fabrics consisting of up to 10 switches, 128 MB of RAM is required.
- For fabrics consisting of up to 15 switches, 256 MB of RAM is required.
- For fabrics consisting of more than 15 switches, 512 MB of RAM is required.

6.1.4 Using WebTools

We start with Internet Explorer running. Type the IP address of the router into the address bar, as shown in Figure 6-4.

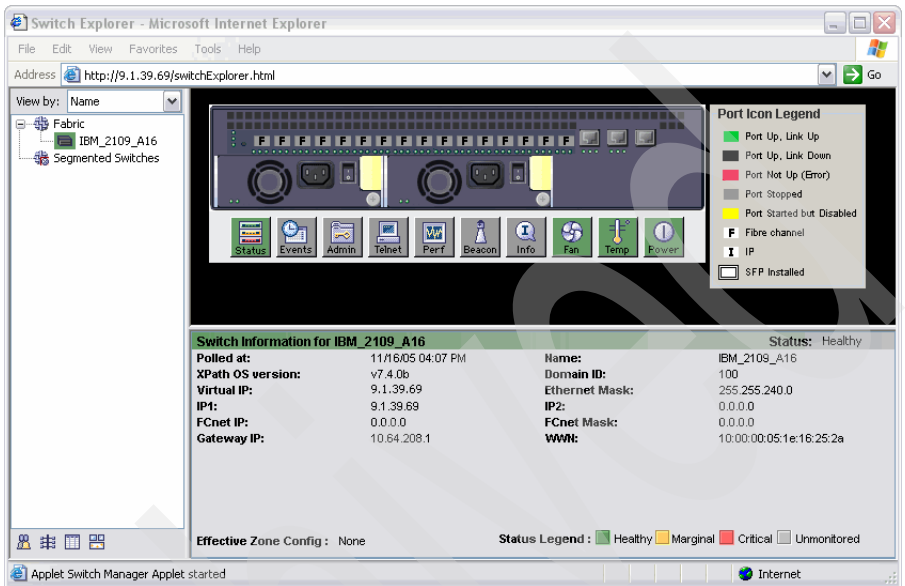


Figure 6-4 Router WebTools startup window

We can see from Figure 6-4 that the router WebTools has a very similar look and feel to that of the FOS WebTools.

Figure 6-5 shows the left pane in more detail. Here, we are presented with a list of fabrics and their associated switches. This list can be sorted by IP, WWN, or switch name.

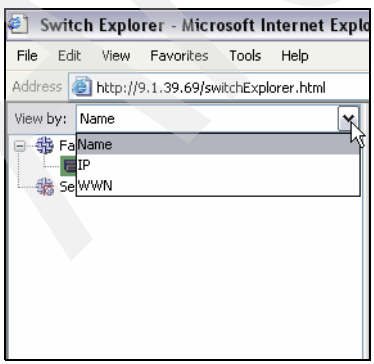


Figure 6-5 Sorting device list by device Name, IP, or WWN

At the bottom-left corner of this main window, we see a block of four buttons. In Figure 6-6, from left to right, they are Fabric Events, Topology, Name Server, and Zone Administration.

Figure 6-6 WebTools bottom left side button cluster

Clicking either the Topology, Name Server, or Zone Administration button opens the Switch Manager (SM). We cover the SM in more detail later in this chapter.

Clicking Fabric Events opens a new window showing all events across all switches in the fabric, as shown in Figure 6-7.

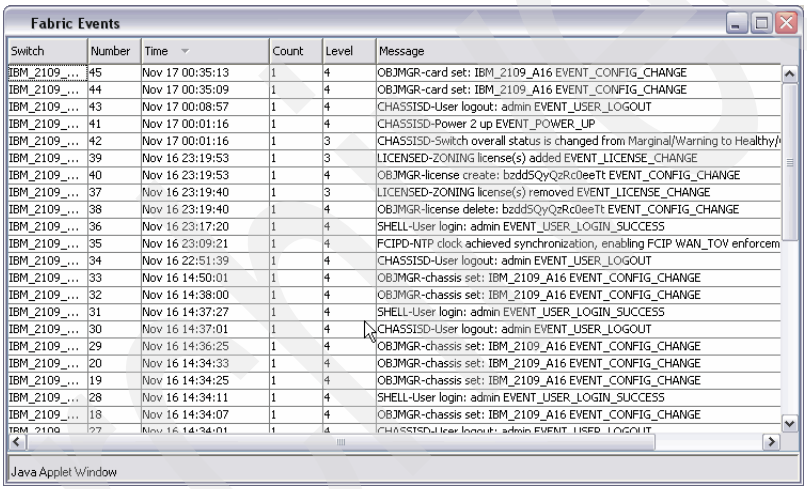


Figure 6-7 Fabric Events

Table 6-1 shows further information in relation to the fabric events data. All the possible messages you might expect to see in this log are documented fully in the *XPath OS System Error Message Reference Manual*, which you can download from the Brocade Web site.

Table 6-1 Fabric events description

Column name	Description
Switch Name	Which switch in the fabric has reported the message
Number	The particular error number from the fabric
Time	The time at which this message was reported

Column name	Description
Count	The number of times this exact message has been reported
Level	The category of error: 1=info, 2=warning, 3=critical, and 4=debug
Message	The error itself

On the right side of the main WebTools window, the lower half of the section shows a summary of information relating to this switch (see Figure 6-8). This information is also color coded: Healthy (green), Marginal (amber), Critical (red), and therefore provides a real-time status of the router.

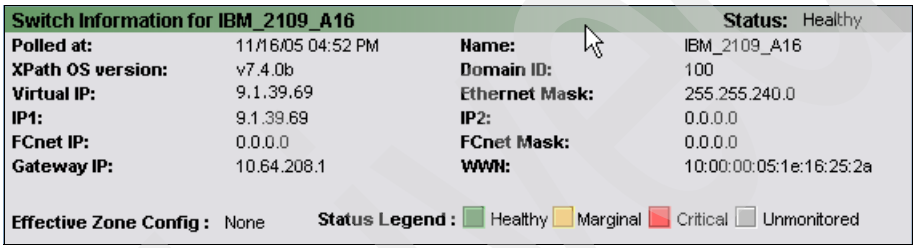


Figure 6-8 Summary and health status information

The top half of the same WebTools window shows a graphical view of the router's front (see Figure 6-9).

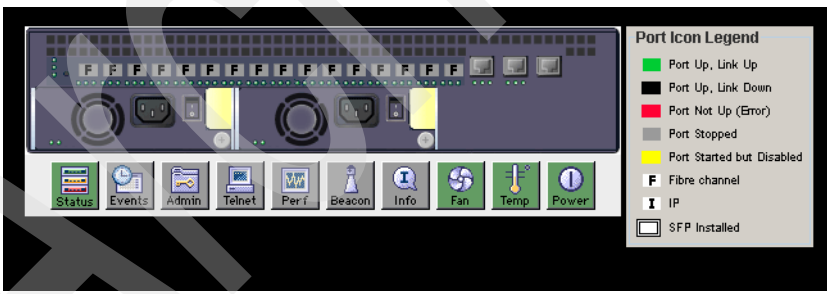


Figure 6-9 Physical view of router from WebTools with status indicators

The Events, Admin, and Perf buttons under the picture of the switch open the Switch Manager, which we discuss later. The environmental buttons, here in green, show the status of their individual sections: power, fans, temperature, and overall status.

Figure 6-10 on page 88 shows each new window opened by clicking the environmental buttons.

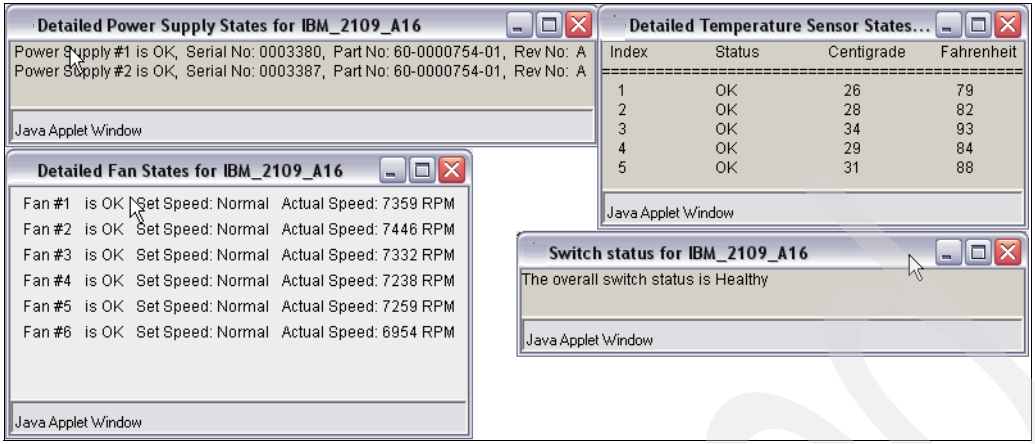


Figure 6-10 Hardware status of power, fans, temperature, and chassis

The three remaining buttons (which do not open Switch Manager) are the Info button, Beacon button, and the Telnet button. The Telnet button simply opens an external Telnet session onto the router.

After clicking the Beacon button, you initially see a confirmation box asking if you want to turn on the beaconing function. After accepting the confirmation, the system LED on the front, upper-left side blinks green and the Beacon icon changes. See Figure 6-11.

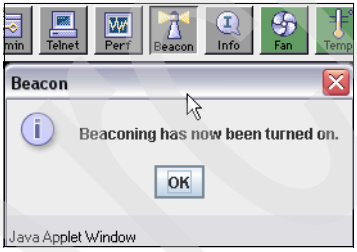


Figure 6-11 Beaconing confirmation with working lighthouse icon

The Info icon displays further chassis information from the router, as shown in Figure 6-12.

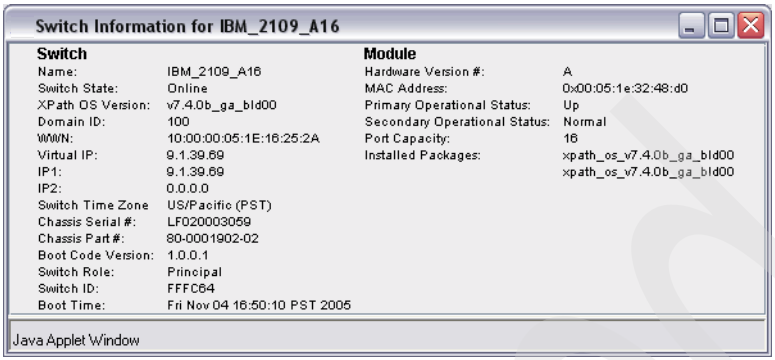


Figure 6-12 Output from Info button

6.1.5 Switch Manager

In this section, we walk through all the available functions within Switch Manager and explain roles. Switch Manager (SM) is opened from within WebTools and provides us with the majority of functions needed to implement a routing solution.

The main functions of the Switch Manager are:

- In-depth status of the hardware
- Port configurations between FC, FCIP, and iSCSI
- Access to the switch error and event log
- A Fabric and Topology view
- A detailed view of the name server
- Zoning status and administration
- Full administration of SNMP: Users, config, firmware, licenses, trunking, routing, and network functions
- Performance monitoring
- FC Routing and iSCSI Configuration windows

Clicking the Admin button in WebTools opens the SM.

The main Switch Manager window is divided into two sections: a narrow panel on the left that displays a menu selection, and a larger panel on the right where the menu selection functionality is available.

Application platform

The Application Platform Summary panel in Figure 6-13 shows a switch summary: port status with color coding, switch WWN, chassis part and serial number, firmware levels, and so on. As usual, the environmental icons are color-coded to allow easy identification of a warning or critical issue within the chassis.

This window also has links to enable beaconing or to open a Telnet session into the CLI.

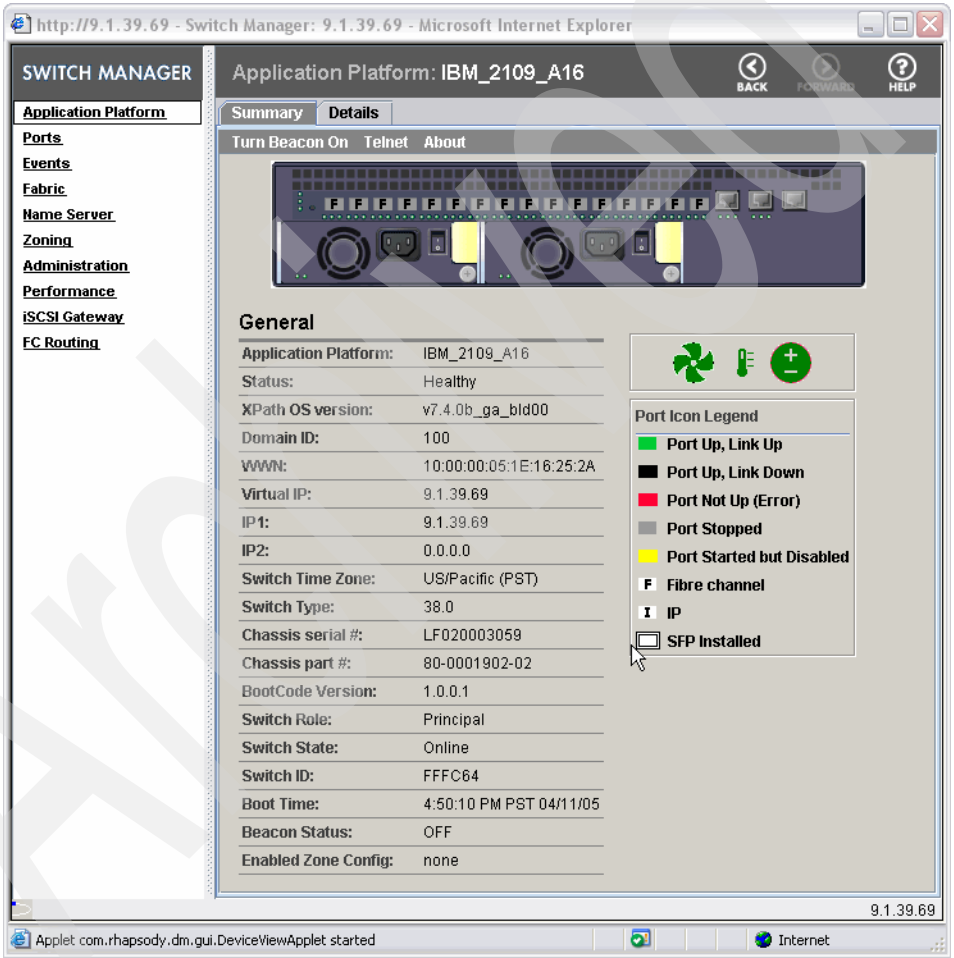


Figure 6-13 Overall status of multiprotocol router

Application platform details

We now click the Details tab, as shown in Figure 6-14.

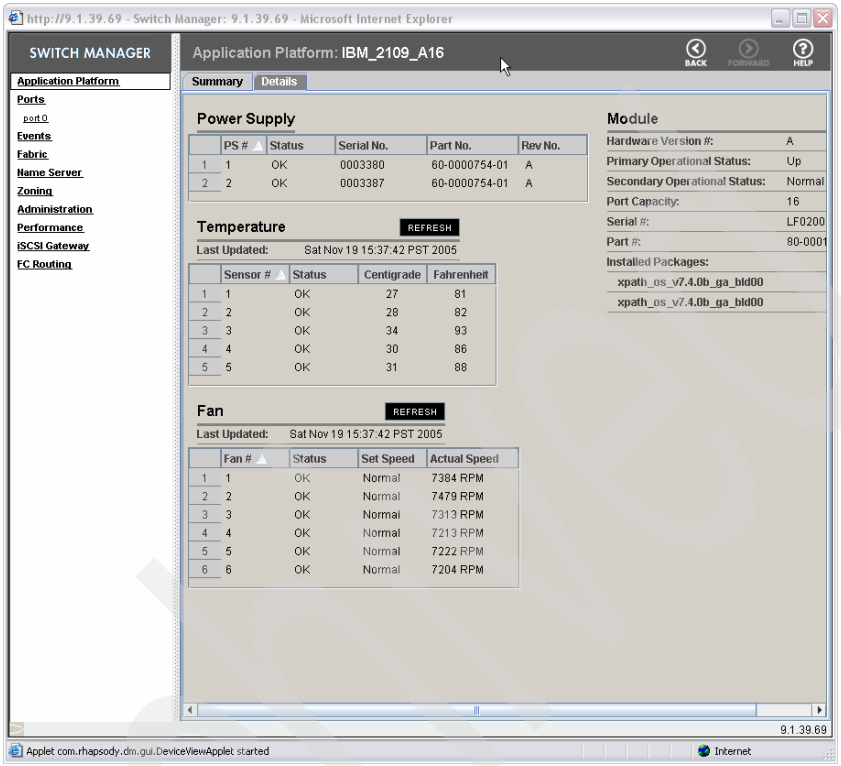


Figure 6-14 More in-depth hardware summary

This panel (Figure 6-14) shows the power supply information for each installed power supply unit (PSU), its current status, and the manufacturing part and serial numbers.

The temperature table displays the current temperature in five separate zones within the chassis, in both centigrade and fahrenheit, and whether this current temperature is within tolerances. A REFRESH button enables this temperature to be updated instantaneously.

The fan table provides us with the status of all six fans within the chassis. We can see the RPM of each fan and whether this is an expected speed.

Ports

We describe the ports in “Administration” on page 95.

Event log

From the Events panel we have access to the switch event log, as shown in Figure 6-15. Both hardware errors and Fabric Watch messages are written into this log.

All messages are assigned a severity: Info (1), Warning (2), Critical (3), and Debug (4).

The drop-down menu at the top of this panel enables us to sort these events by different time zones: Local Time Zone, Switch Time Zone, or both Local and Switch Time Zone.

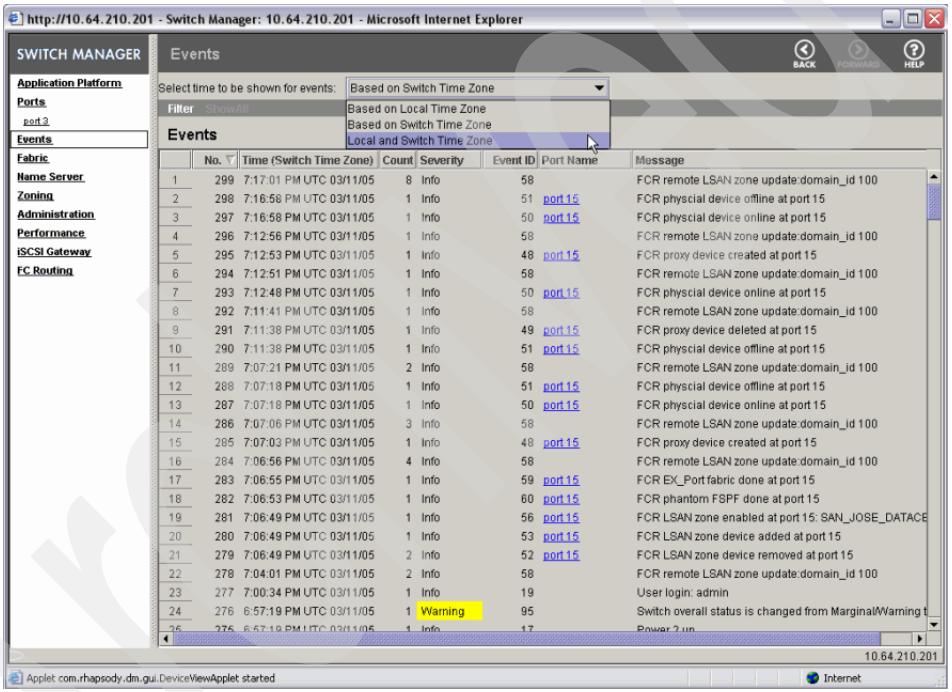


Figure 6-15 Switch event log

In Table 6-2, we describe each column in Figure 6-15.

Table 6-2 Switch events

Column name	Description
Number	The error number counter since the error log was last cleared.
Time	The time at which this error occurred, in either of the time zones: local time or switch time.

Column name	Description
Count	The number of times this particular error has been reported.
Severity	The severity of this message: informational, warning, critical, or debug.
Event ID	The event ID is a numeric tag given to each specific error message. These can be looked up in the *XPath OS System Error Message Reference Manual* for further details and can be downloaded from the Brocade Web site.
Port name	If an error is related to a specific port, that port is displayed here.
Message	The error message itself.

Fabric

The next option in the SM menu is *Fabric*, as shown in Figure 6-16. This window provides us with information regarding the fabric to which this router is connected.

In Figure 6-16, we show only one router in the fabric. Normally, all switches in this fabric will be represented here, displaying their name, domain ID, OS version, WWN, IP address, role (principal or subordinate), and whether it is a router or a standard switch.

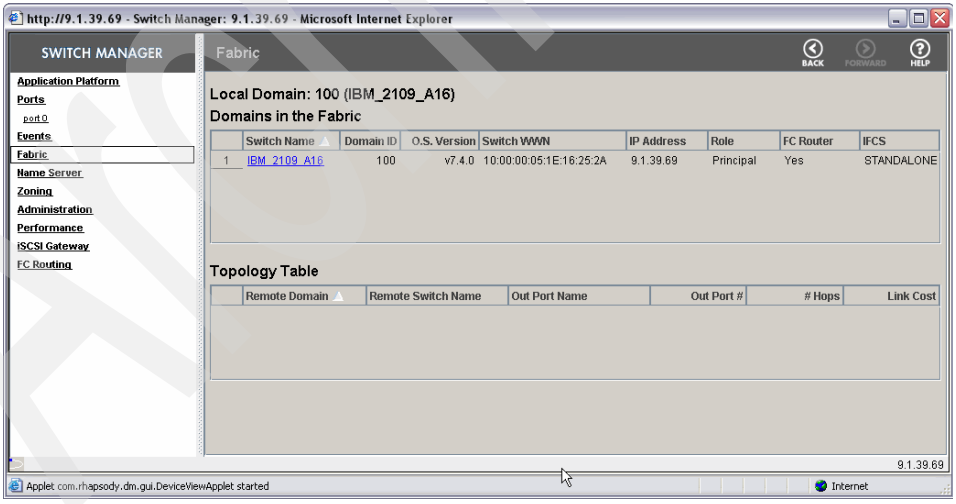


Figure 6-16 Topology and Fabric information

Name Server

When devices are directly connected to a router, the Name Server panel is populated with a variety of details from those devices: Vendor, WWN, domain ID, port ID, and so on, as shown in Figure 6-17.

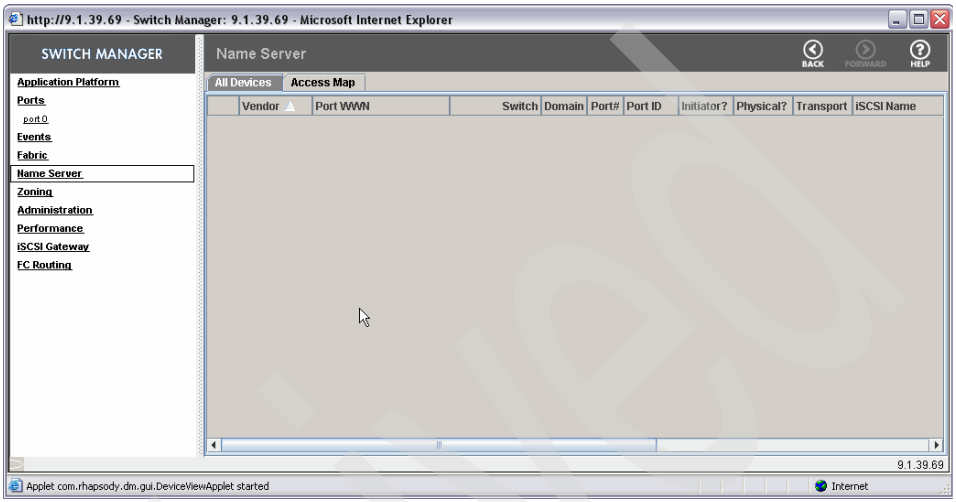


Figure 6-17 Local devices logged into the Name Server

Zoning

The *Zoning* panel shows us various information regarding local, fabric-based zoning. Here, by navigating with the use of tabs, we can view zones, aliases, and zone configurations separately. Each tab also gives us the ability to edit or create zones/aliases or a configuration.

Note that this is not router/LSAN zoning information. We discuss LSAN zones in a different panel and later in other topics in this book.

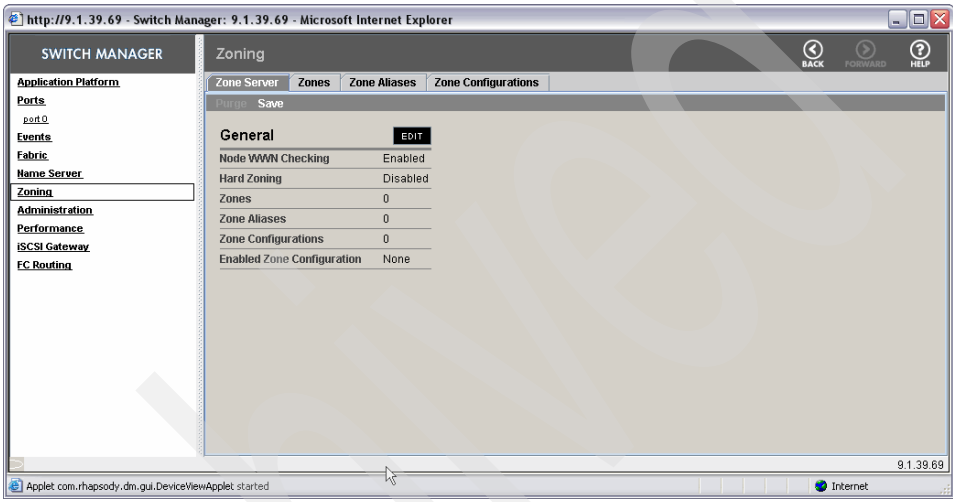


Figure 6-18 Local Zoning information

Administration

In the Administration panel, we have access to most of the configuration and monitoring functions available on this router. The information is split into 11 tabs:

- Switch Information
- Ports
- Configure Fabric
- Routing
- Network Config
- SNMP
- Users
- Config Upload/Download
- Firmware
- Licenses
- Trunking

Switch Information tab

From this tab, you can rename the router, disable the router, change the domain ID, set up the time/date, and reboot the router.

The panel itself shows us further router-based information. We can view the IP addresses on the management ports, the WWN, switch status, installed firmware, and so on, as shown in Figure 6-19.

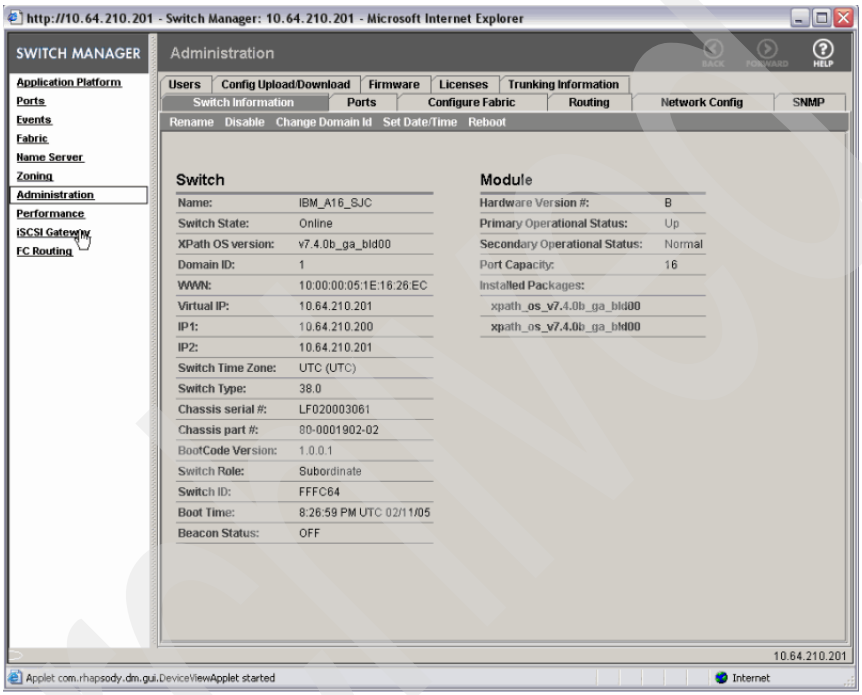


Figure 6-19 Switch configuration information

Ports tab

From the Ports tab (Figure 6-20), we can list all physical port information, including which protocol each port is running, its current status (up or down), and its port type. Additional tabs (FC Ports, FCIP Ports, and iSCSI Ports) provide further details about the physical port.

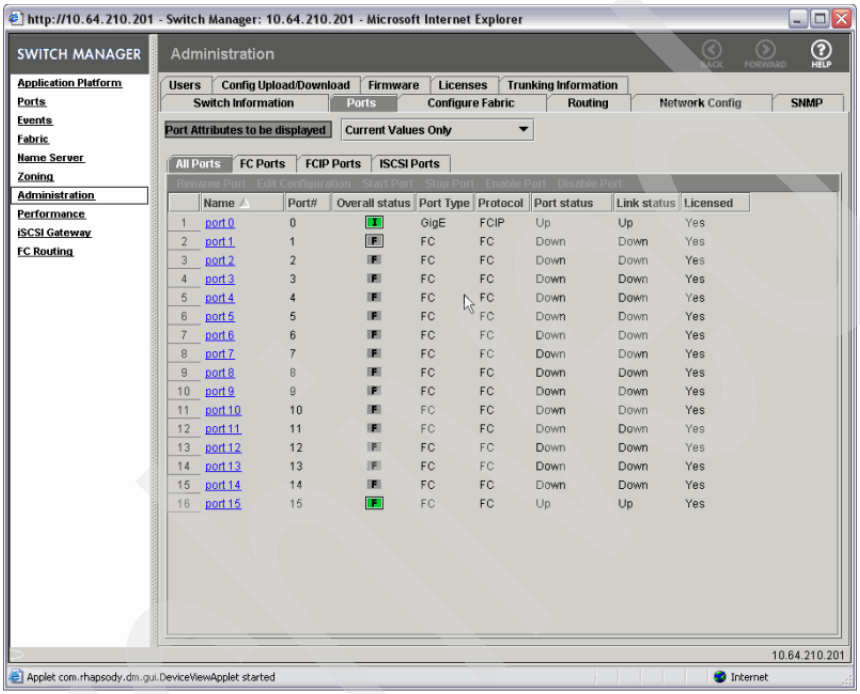


Figure 6-20 Port status overview

Figure 6-21 shows the same port view as in the previous figure but here we selected to view the port status and the currently configured values.

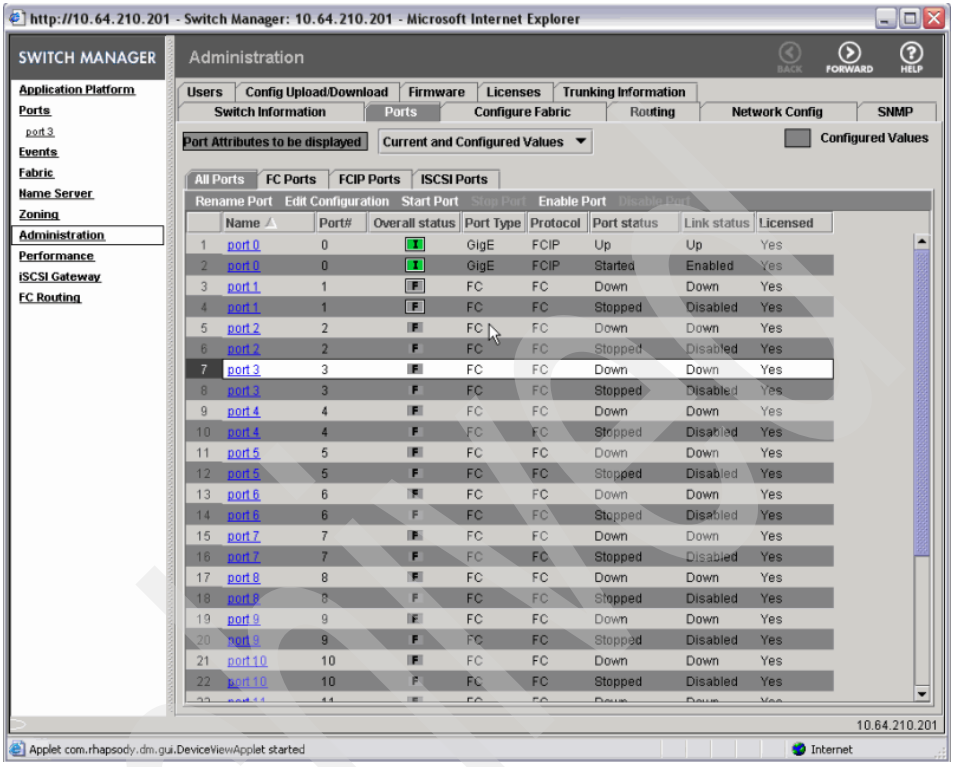


Figure 6-21 Detailed port status and configuration information

When clicking an individual port, we have access to all port details, as shown in Figure 6-22.

From this window, we can also start or stop a port. This involves the loading and unloading of code at the ASIC level per port. We can enable or disable the port and, most importantly, edit the configuration of the port.

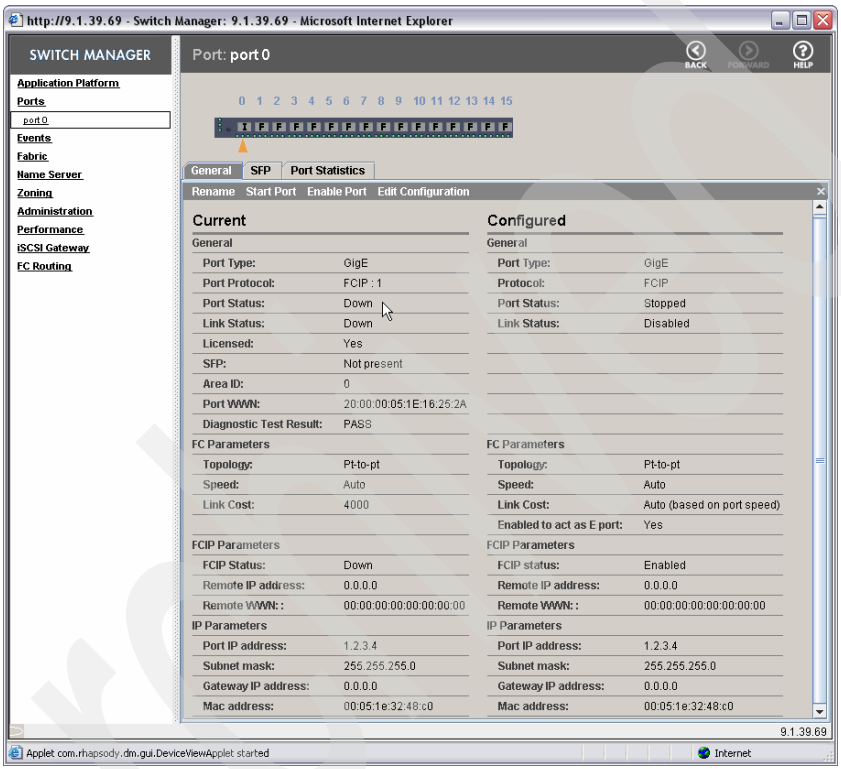


Figure 6-22 Individual port status and configuration information

The Port Statistics subtab shown in Figure 6-23 displays transmit and receive counters, overall throughput, per protocol counters, and error counters.

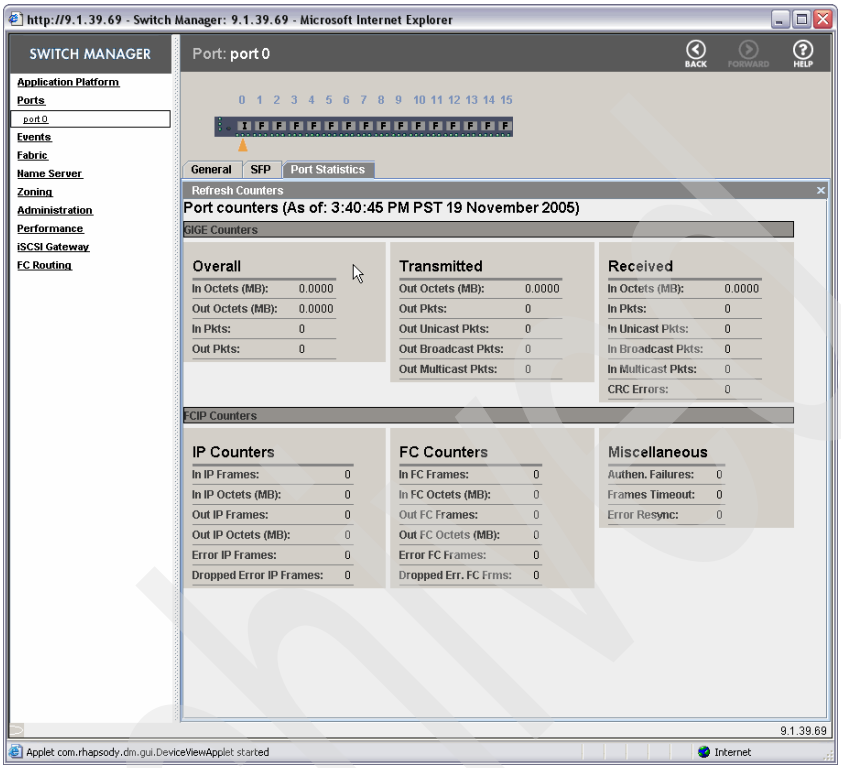


Figure 6-23 Individual port statistic counters

Configure Fabric tab

When using the Configure Fabric tab, as shown in Figure 6-24, we can adjust some familiar fabric settings, such as the default buffer-to-buffer credit settings (per port), timeout values (TOV), and PID format, to name a few. Some of these values require the router to be disabled when changing them.

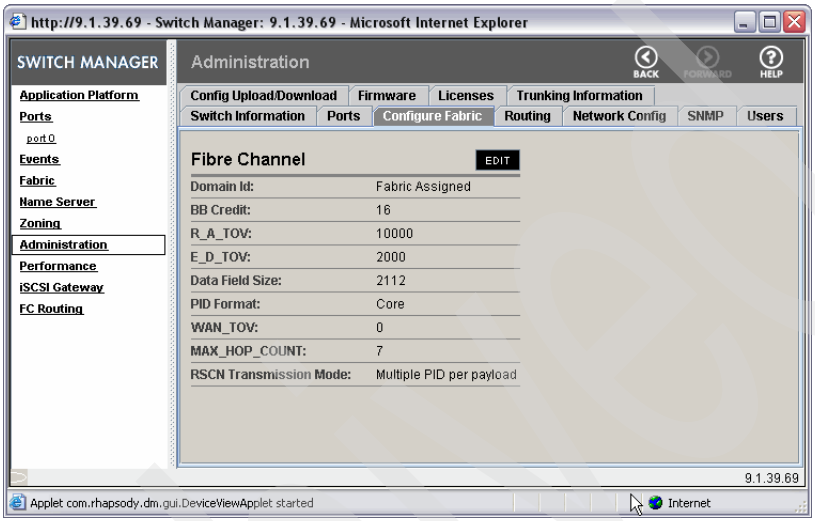


Figure 6-24 Configure fabric settings: PID, TOV, and BB Credit

Routing tab

As shown in Figure 6-25, the Routing tab is divided into four areas: IOD/DLS, FSPF Routes, Static Routes, and Link Cost.

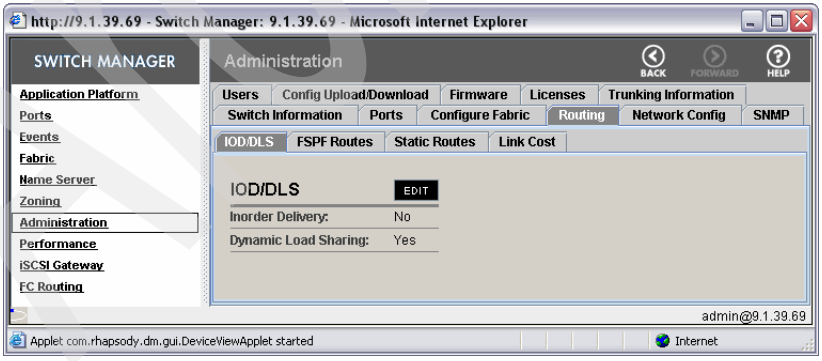


Figure 6-25 Inorder Delivery and Dynamic Load Sharing settings

From the IOD/DLS subtab, we can toggle the Inorder Delivery (IOD) setting between Yes and No. IBM recommends that IOD be set to **Yes**.

The Dynamic Load Sharing (DLS) setting can also be toggled between Yes and No. Note that IBM commonly uses different settings than the switch default settings. Refer to the router installation guide to confirm the latest supported settings.

Within the FSPF Routes table, shown in Figure 6-26, we see the preferred paths this router would use through a fabric. This calculation is based on each hop's link cost.

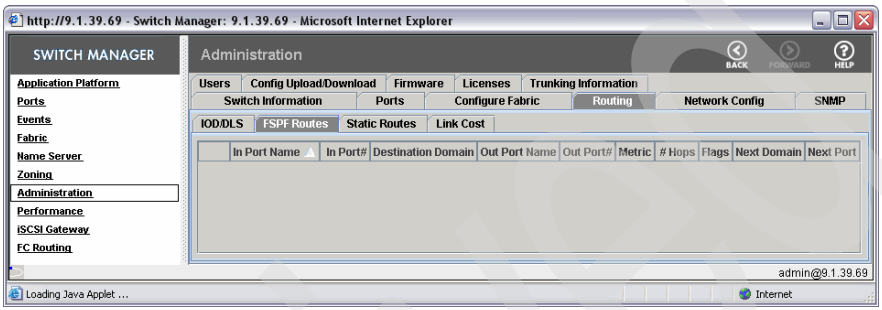


Figure 6-26 FSPF Routes table

In the Static Routes tab (Figure 6-27), we can manually add static routes to better tune the fabric for a specific scenario.

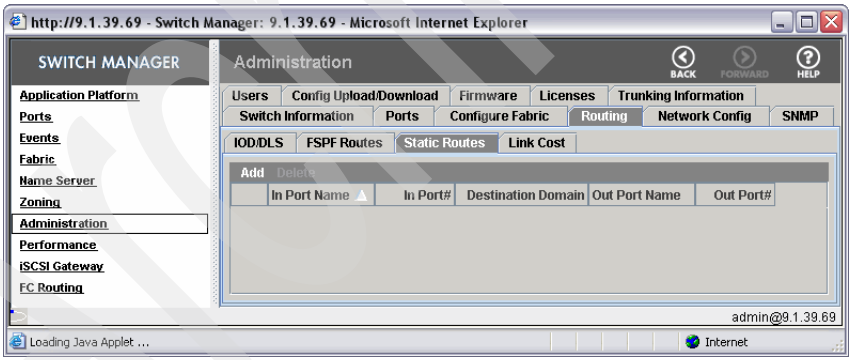


Figure 6-27 Defined Static Routes

Figure 6-28 shows Link Costs. These are commonly set by the switch.

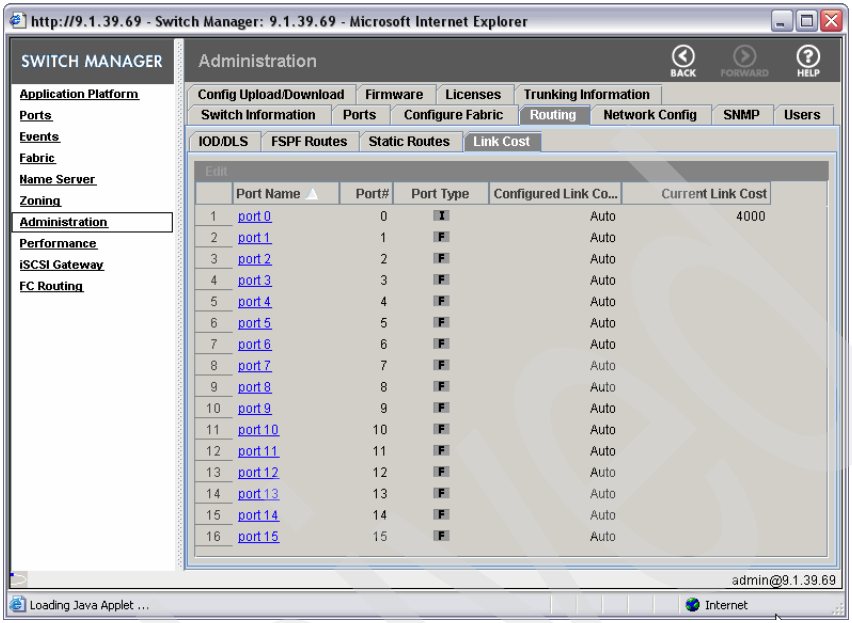


Figure 6-28 Routing Link Cost information

Network Config tab

All network configuration settings are adjustable in the Network Config tab, as shown in Figure 6-29. We can adjust the IP address, subnet mask, gateway address, speed negotiation of both management interfaces, and the virtual management port.

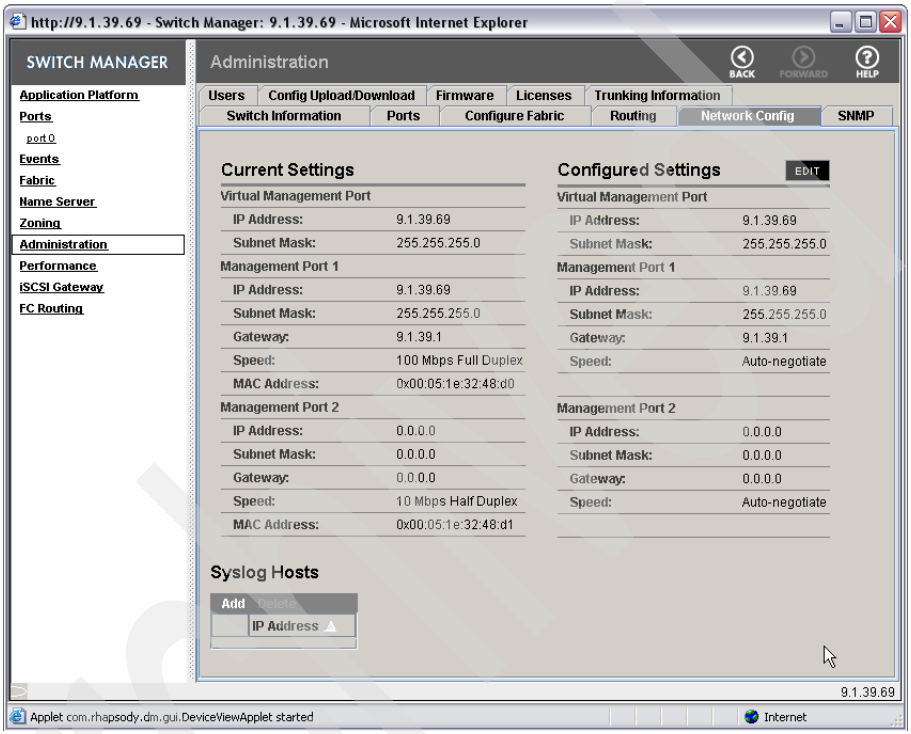


Figure 6-29 Network interface configuration settings

SNMP tab

You can view and set the SNMP settings from here (Figure 6-30).

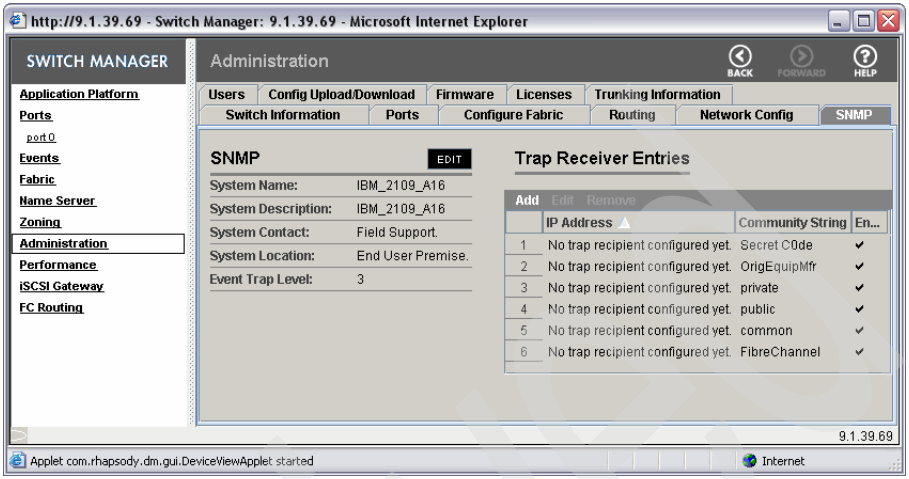


Figure 6-30 SNMP trap settings

Users tab

From this tab, shown in Figure 6-31, you can manipulate user IDs. You can delete, change the password, and adjust the role of existing users. You can also add new users here.

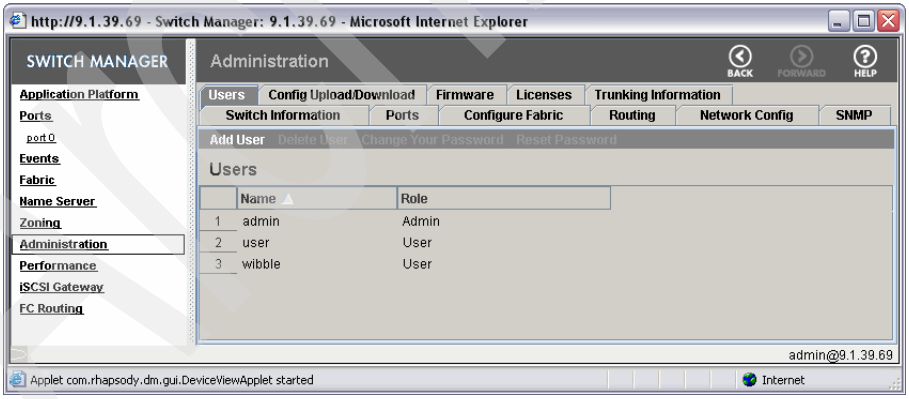


Figure 6-31 User ID settings

Config Upload/Download tab

Figure 6-32 shows the Config Upload/Download tab. Here, we set up the details of the external FTP user from or to which upload or download a configuration.

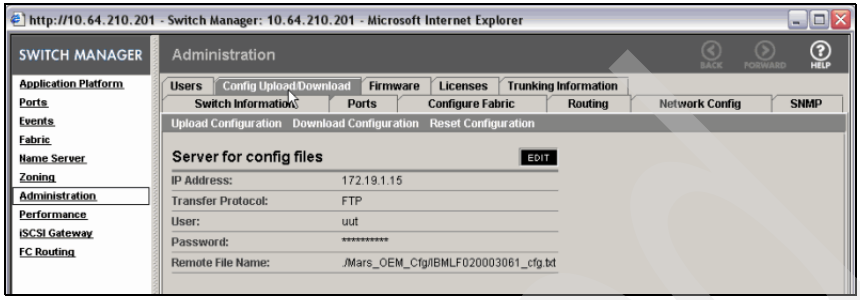


Figure 6-32 Switch Config Upload/Download tab

Firmware tab

From the Firmware tab, shown in Figure 6-33, you can upgrade or downgrade firmware from within WebTools. We can see the current version of code installed and from where it was last downloaded. An FTP server can be permanently defined from this tab for easier code loading.

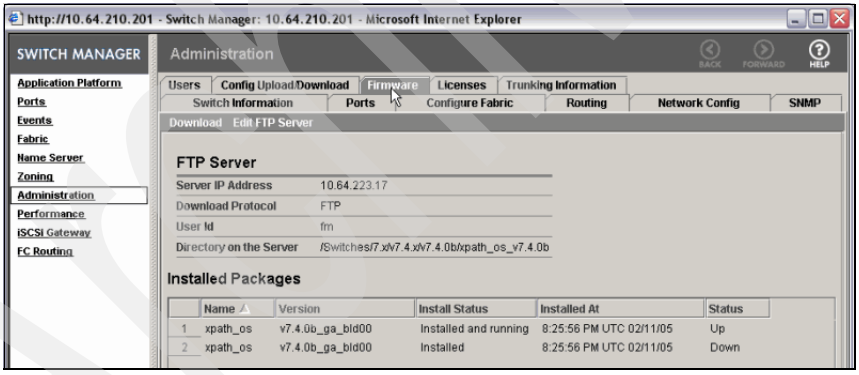


Figure 6-33 Firmware download panel

Licenses tab

From the License tab (Figure 6-34), we can perform basic license maintenance. You can add or remove licenses here or through the CLI using the **`licenseadd`** or **`licenseremove`** commands.

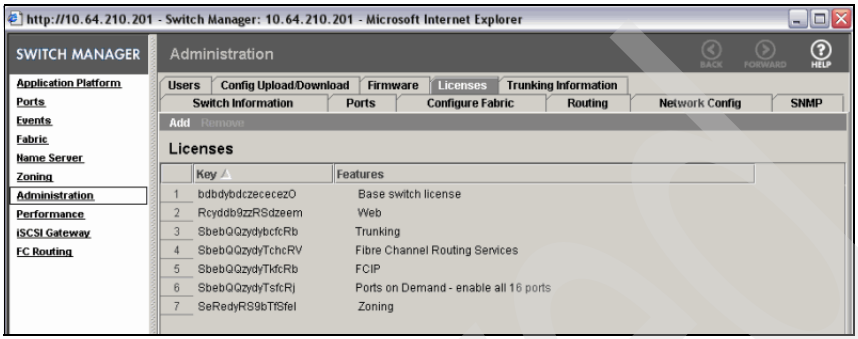


Figure 6-34 Installed Licenses

Trunking Information tab

As we can see in Figure 6-35, this tab shows some basic information regarding any trunking configuration. First, we can see if trunking is disabled or enabled. If it is enabled, we can set the option to include Exchange based trunking.

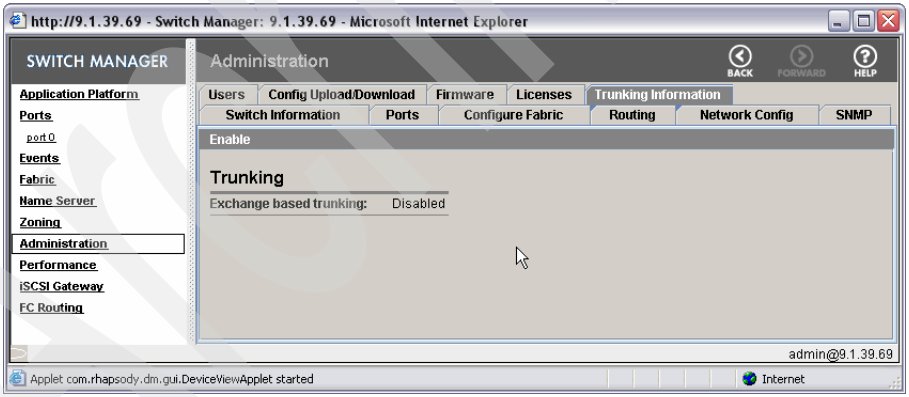


Figure 6-35 Trunking Information tab

Performance

Figure 6-36 shows performance data from the router. You can view both counter-based data and chart-based data.

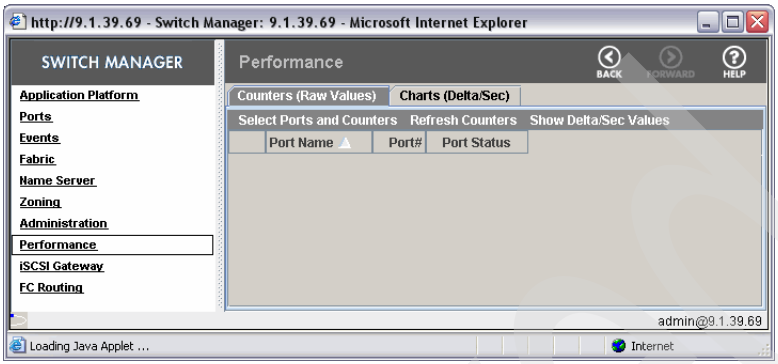


Figure 6-36 Router Performance panel

iSCSI gateway

To configure the 2109-A16 router to provide an iSCSI gateway service, we use the iSCSI Gateway panel (Figure 6-37).

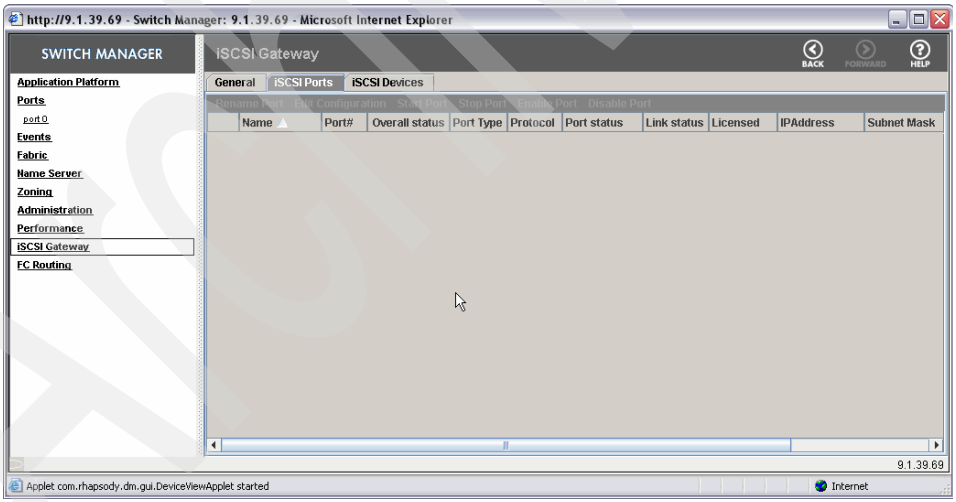


Figure 6-37 iSCSI Gateway configured port information

Figure 6-38 shows the iSCSI Devices tab, where we would see a list of currently attached iSCSI initiators to the gateway.

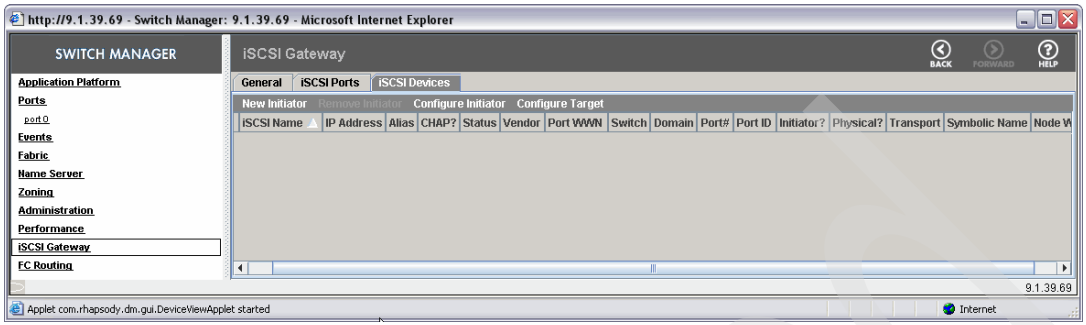


Figure 6-38 iSCSI Gateway devices

FC Routing

The final option in the Switch Manager, shown in Figure 6-39, lets us configure routing-specific information. This section is split into five subtabs: General, EX_Ports (for inter-fabric links, Edge Fabrics, LSAN Zones (for exporting devices across separated fabrics), and LSAN Devices.

EX_Ports

In Figure 6-39, we can see a single port on this router configured as an EX_Port. This shows its port number, overall status, the link status between this port and the interconnecting SAN, and the fabric ID that this port will use to route into the connecting SAN.

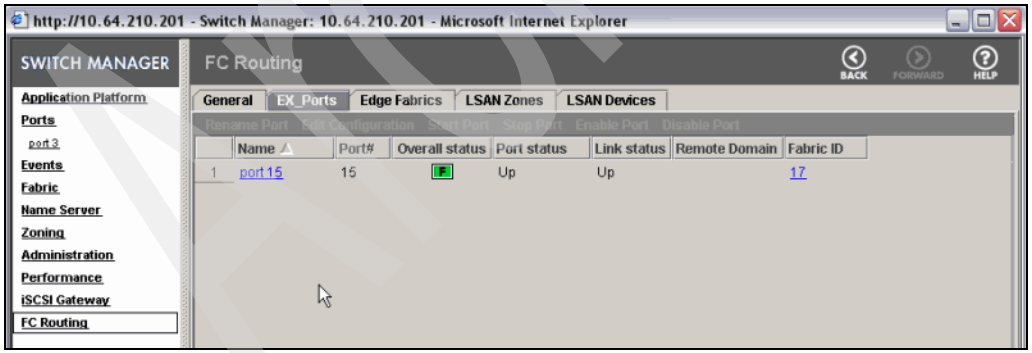


Figure 6-39 FC Routing EX_Ports

Edge Fabrics tab

The Edge Fabrics tab (Figure 6-40) shows a list of switches, each of which represents the connection point into an edge fabric. The WWN of each edge fabric switch is listed, including a reference to whether the edge fabric is connected locally or remotely (through another router).

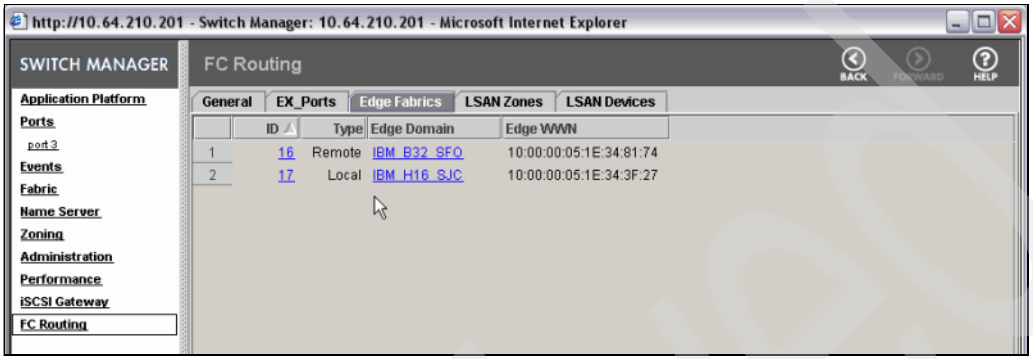


Figure 6-40 Connected Edge Fabrics

LSAN Zones tab

The LSAN Zones tab (Figure 6-41) contains a reference to each LSAN zone from both the local and remote fabrics, including a fabric ID reference.

Figure 6-41 LSAN Zones

LSAN Devices tab

This tab (Figure 6-42) is split into two sections: a Physical subtab and a Proxy subtab.

The Physical tab shows a list of all edge fabric-attached devices that are configured in one of the LSAN zones. This also displays the physical PID for each of these devices.

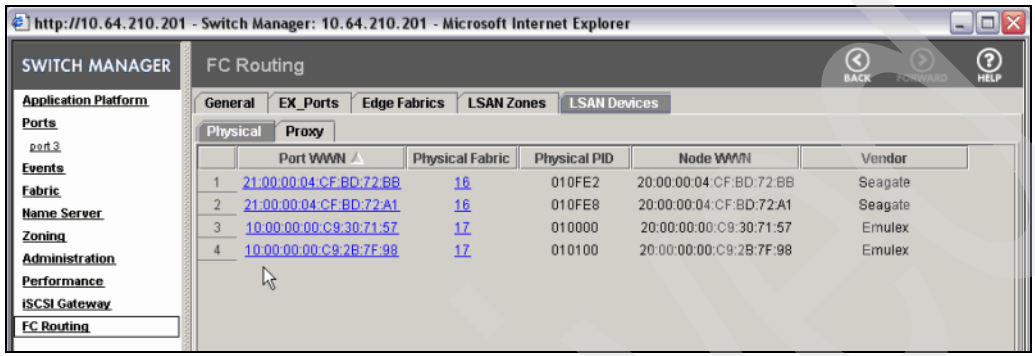


Figure 6-42 LSAN Physical devices

The Proxy subtab of LSAN Devices tab (Figure 6-43) shows all the proxy devices that will be used to assist with the communication across the backbone fabric into each edge fabric. We see both their local physical PID and proxy PIDs.

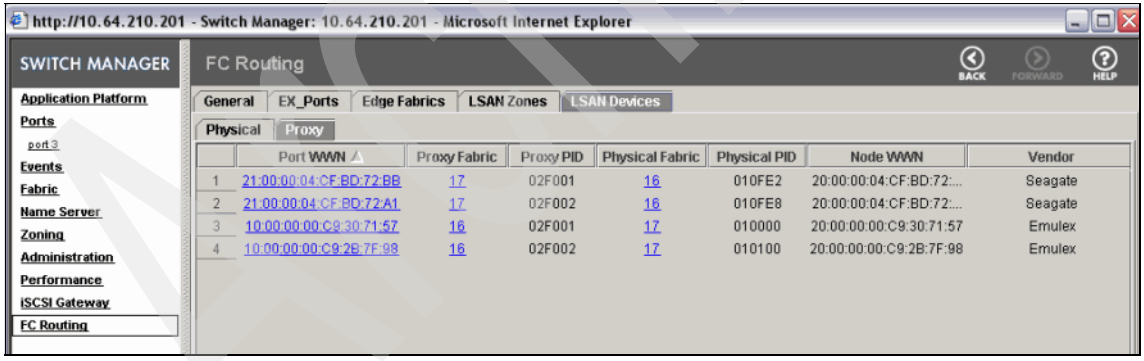


Figure 6-43 LSAN proxy devices

This concludes our walkthrough of the router WebTools functionality.

6.1.6 Configuring a routing service using FCIP

In this section, we use two existing SAN fabrics. We connect a router into each fabric, and then connect the routers to each other through an FCIP link. After we establish this backbone fabric using the routers, we can configure LSAN zones, allowing designated traffic to be routed between SANs. This is then known as a Meta SAN.

The SAN fabrics

We start with two independent SANs. The *San Francisco SAN* (SFO) contains two IBM x330 hosts, one 2005-H16 switch, and one 2005-B32 switch. The switches are connected by ISLs to form a single fabric. This fabric also has JBOD-connected disks. Some of these disks will be routed across to our other SAN. The second SAN is called the *San Jose SAN* (SJ), which also contains two IBM x330 hosts and a single 2005-H16 switch, and is connected to a JBOD disk (Figure 6-44).

Figure 6-44 The starting configuration

Establishing an FCIP link

First, we physically connect both routers together. We chose to use port 0 of each router. Next, we configure each port 0 for FCIP. Each router-to-router port will become a VE_Port after it is enabled and started. We perform all the port configuration using WebTools but you can just as easily use the CLI. For all the associated CLI commands, refer to the Brocade *XPath OS Procedures Guide*, available from the Brocade Web site.

To establish an FCIP link:

1. Starting at the main WebTools window (Figure 6-45), we click the port 0 icon.

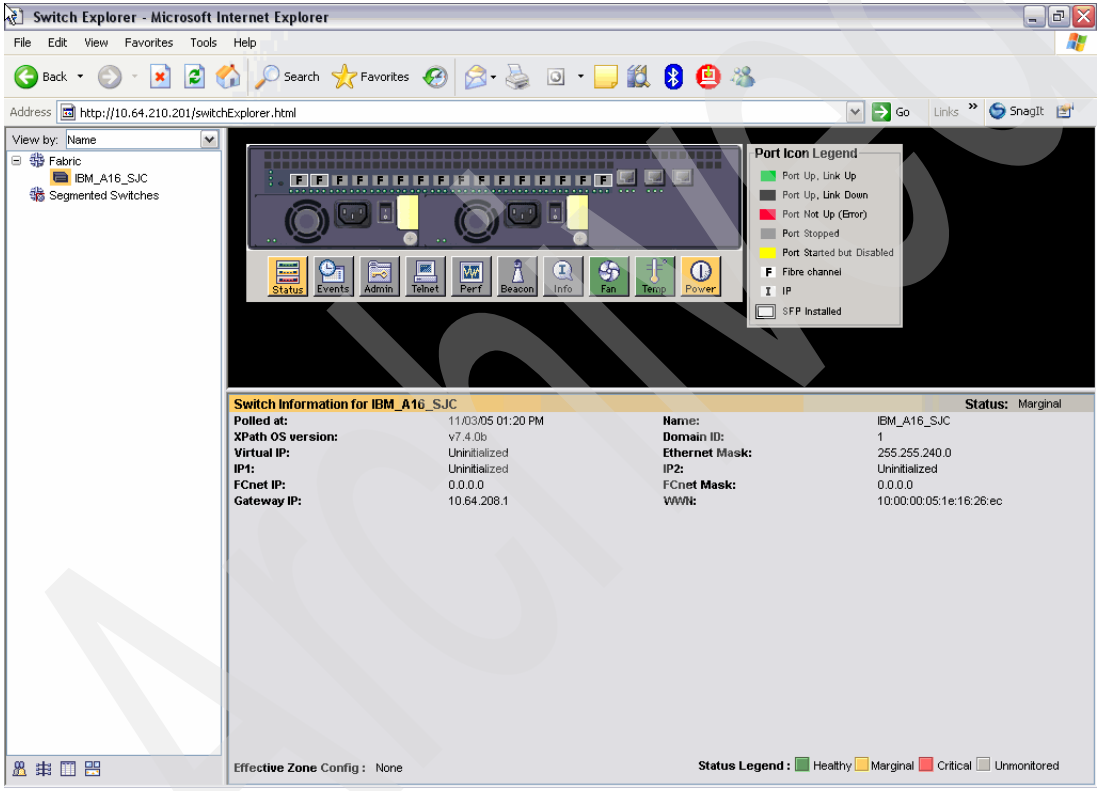


Figure 6-45 Starting point

Switch Manager opens, as shown in Figure 6-46, showing the current configuration of port 0. Here, we can see that the port is currently stopped and is down with no previous configuration.

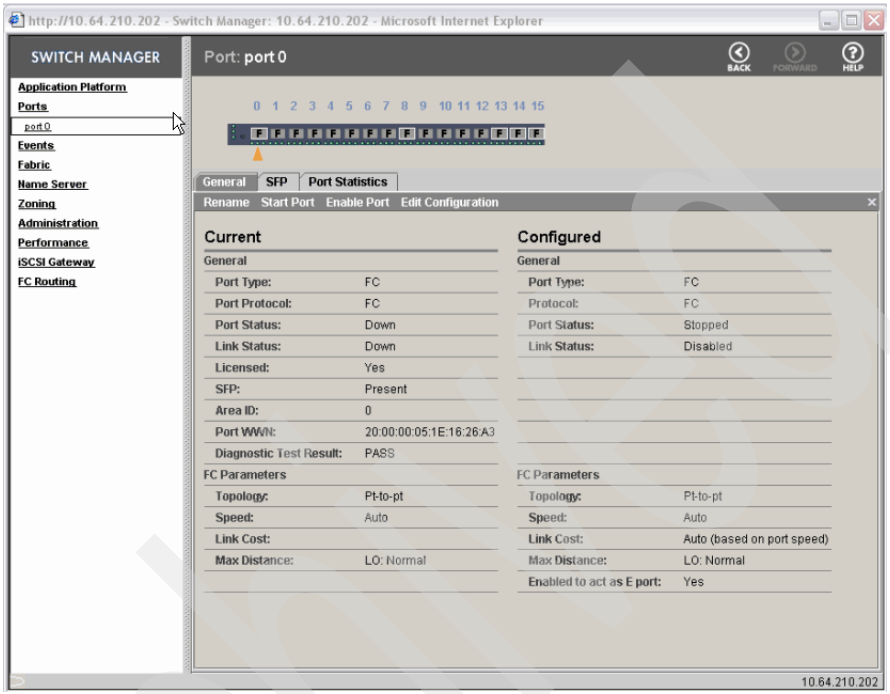


Figure 6-46 Port 0 configuration panel

2. Now, we click the **Edit Configuration** link and the Edit Port Configuration window opens, as shown in Figure 6-47.

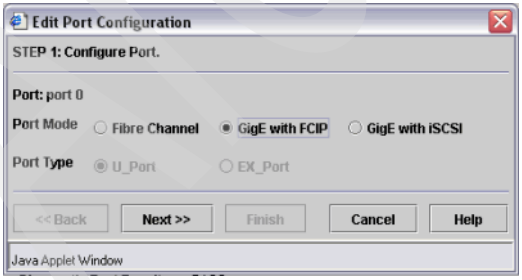


Figure 6-47 Step 1 of configuring Port 0

3. The first step is to choose the port mode, or effectively the protocol we want to use on that port. The options are:

   – Standard Fibre Channel port
   – Gigabit Ethernet (GigE) with FCIP traffic
   – Gigabit Ethernet (GigE) with iSCSI traffic

   If we select **Fibre Channel**, the port can be configured as a standard U_Port and act as any other b-type switch port. Or, we can choose **EX_Port**. We use this later, but effectively, it is the port type used when connecting a router into an existing SAN, similar to an E_Port when connecting switches.

   If we wanted to implement an iSCSI gateway solution using this product, this is where we configure the port to support that protocol.

   We select **GigE with FCIP** for this example.

4. The second configuration step is to give this port an IP address, as shown in Figure 6-48. Because FCIP is encapsulating FC traffic within standard IP packets, an IP address is needed for the source and destination. Because this will be a dedicated point-to-point connection, we do not need a gateway address and are free to choose any IP range we want.

   Each router needs to be told the IP address of its counterpart during this phase of the configuration.

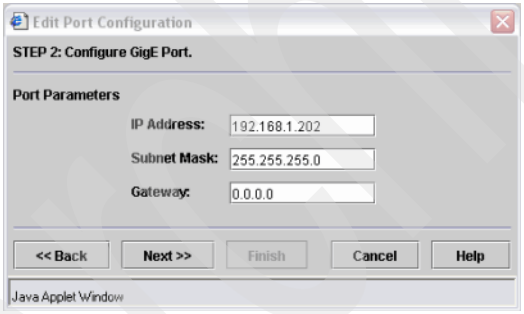


Figure 6-48 Setting an IP Address on the port for FCIP traffic

5. Here, we have two sets of configuration values (Figure 6-49). FC Values specifically relate to link cost. The link cost is used by FSPF to build the routing tables of paths through a fabric. There is no reason to manually adjust this in our configuration, so we leave the default setting of **Dynamic**.

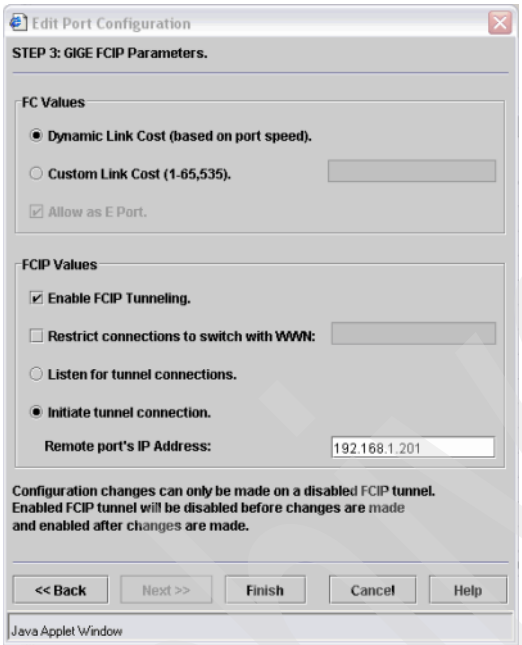


Figure 6-49 Specific FCIP Parameters

In the FCIP Values section, we have a few choices to make. First, we select **Enable FCIP Tunneling**. Because we are effectively tunneling one protocol through another, we need to select this for FCIP.

The second option here is a security feature, Restrict connections to switch with WWN. We can specifically lock this port to only negotiate with a particular switch WWN. At this point, we can add the WWN of our remote switch, thereby locking all traffic negotiation down port 0 to our remote switch only.

Finally (and probably the most important setting), we set this port to either Initiate tunnel connection or Listen for tunnel connections. One router must be set to initiate the connection and the other router to listen. We set our router to **Initiate tunnel connection**, and therefore have to make it aware of the address of the remote router.

We click **Finish** to complete our settings.

6. We can see from the details behind this final window (Figure 6-50) that the remote address has been populated. We click **Close** to close this last window.

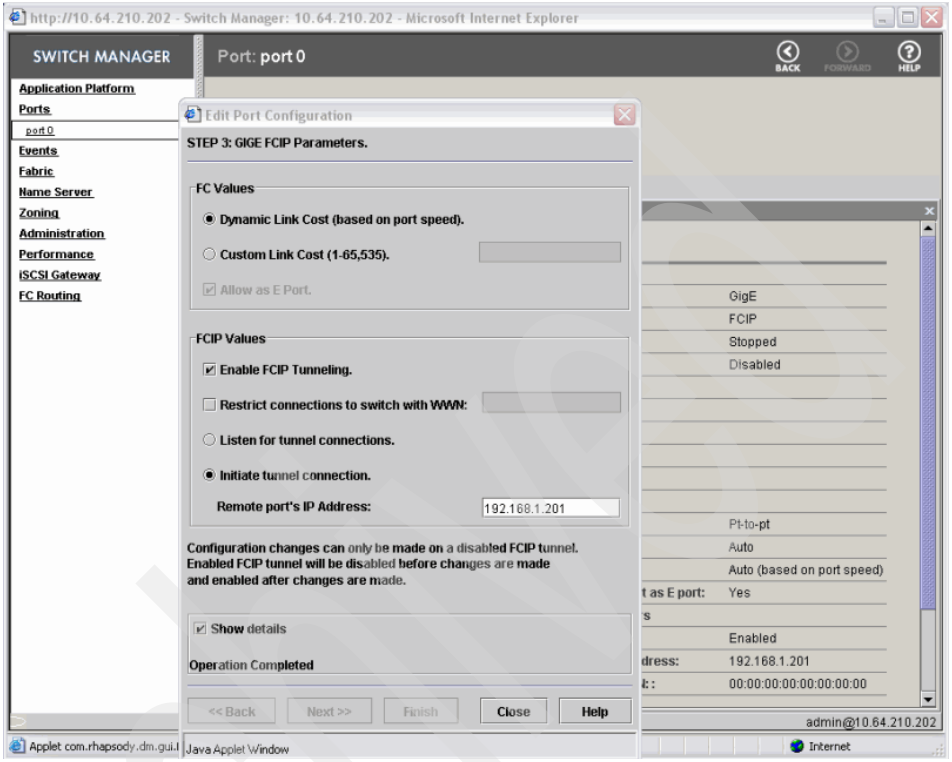


Figure 6-50 Completed port 0 configuration

7. Now, set up the remote router in the same way, but remember it must be set to **Listen for tunnel connections** and not to Initiate tunnel connection, as shown in Figure 6-51.

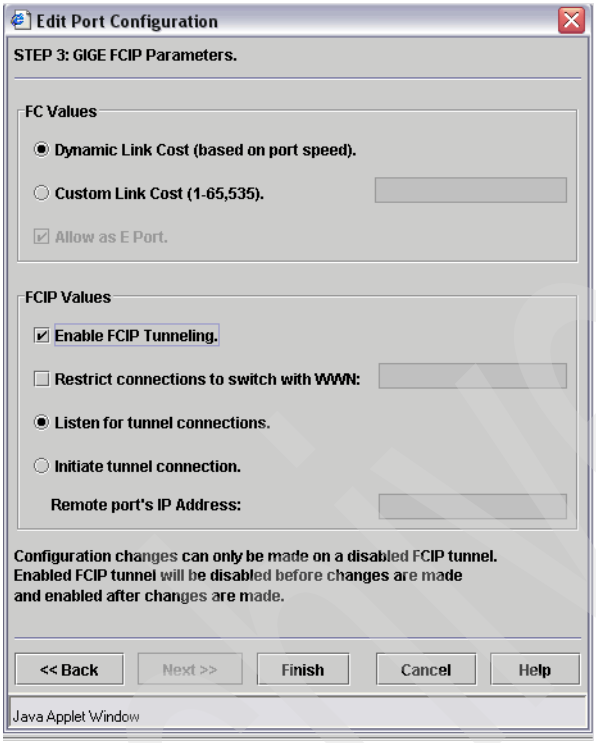


Figure 6-51 Remote router set to Listen for tunnel connections

8. At this point, both routers should have their port 0 configured correctly. Now, enable and start each port. Figure 6-52 shows this process.

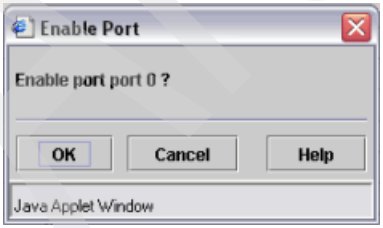


Figure 6-52 Enabling port 0

Figure 6-53 shows successfully enabling port 0.

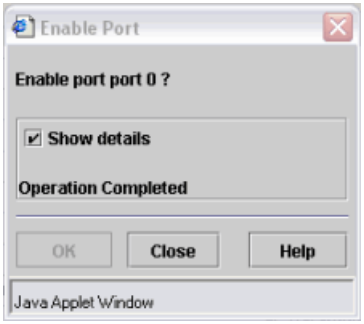


Figure 6-53 Port 0 enabled

9. Now, we start port 0, as shown in Figure 6-54.

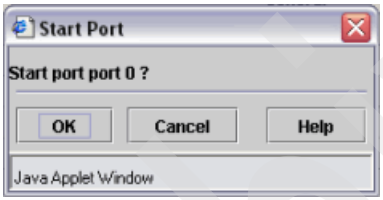


Figure 6-54 Starting port 0

Looking carefully at the information behind the Start Port window in Figure 6-55, we can see this port is enabled, started, and configured as per our needs.

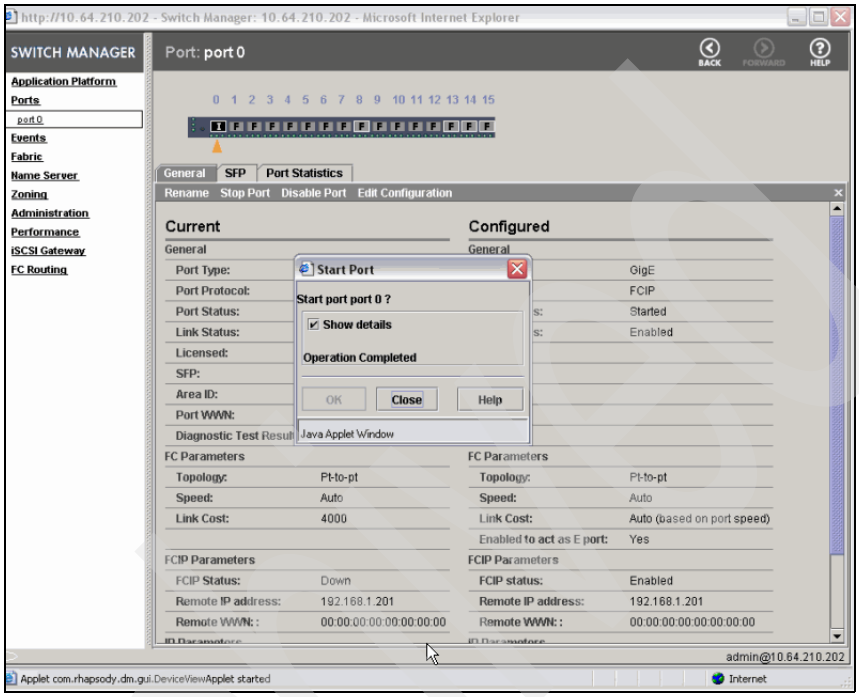


Figure 6-55 Completed startup of port 0

10. Now, we need to enable and start port 0 on the remote switch. When this completes successfully, the port color turns green in WebTools.

We see WebTools with a green port status for port 0 in Figure 6-56. We have both routers successfully talking with each other. This type of connection is called a VE_Port.

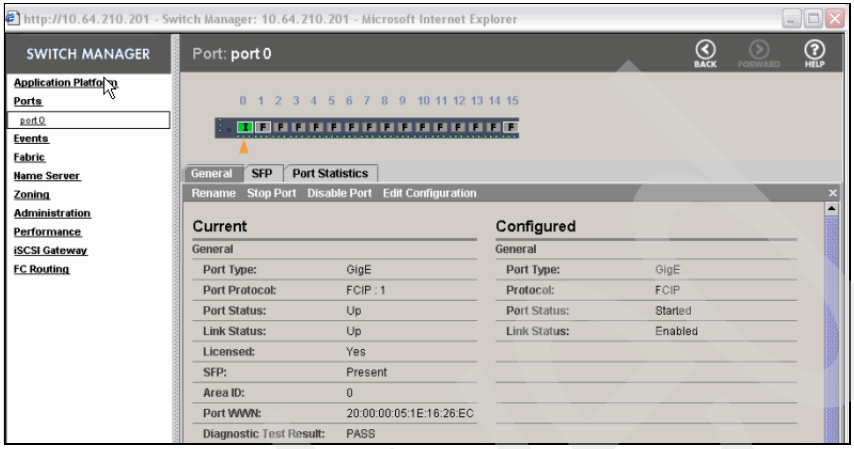


Figure 6-56 Successfully established connection between two routers

We can also use the CLI to confirm the working status of this port from both routers. Using the **switchshow** command, we see the port in VE_Port mode, online, and displaying the WWN of its remote router.

Example 6-18 Confirming the status

```
IBM_A16_SFO:admin> switchshow
Switch Name  :  IBM_A16_SFO
Switch State :  Online
Switch Type  :  38.0
Switch Role  :  Principal
Switch Domain:  100
Switch ID    :  FFFC64
Switch WWN   :  10:00:00:05:1e:16:26:a3
beacon status:  OFF
zoning       :  OFF

FC router BB Fabric ID: 1

Port Media Speed State     Info
=======================================
 0    id    1G   Online    VE_PORT  10:00:00:05:1e:16:26:ec
"IBM_A16_SJC" (downstream)
 1    --    AN   No_Module stopped
 2    --    AN   No_Module stopped
```

```
 3    --    AN   No_Module stopped
 4    --    AN   No_Module stopped
 5    --    AN   No_Module stopped
 6    --    AN   No_Module stopped
 7    --    AN   No_Module stopped
 8    id    AN   No_Light  stopped
 9    --    AN   No_Module stopped
10    --    AN   No_Module stopped
11    --    AN   No_Module stopped
12    --    AN   No_Module stopped
13    id    AN   No_Light  stopped
14    id    AN   No_Sync   stopped
15    id    AN   No_Sync   stopped
IBM_A16_SFO:admin>
```

Using the **switchshow** command from the other router, we confirm that both connections are established.

Example 6-19 Confirming the connections

```
IBM_A16_SJC:admin> switchshow
Switch Name  :  IBM_A16_SJC
Switch State :  Online
Switch Type  :  38.0
Switch Role  :  Subordinate
Switch Domain:  1
Switch ID    :  FFFC01
Switch WWN   :  10:00:00:05:1e:16:26:ec
beacon status:  OFF
zoning       :  OFF

FC router BB Fabric ID: 1

Port Media Speed State     Info
======================================
 0    id    1G   Online    VE_PORT  10:00:00:05:1e:16:26:a3
"IBM_A16_SFO" (upstream)
 1    id    AN   No_Light  stopped
 2    --    AN   No_Module stopped
 3    --    AN   No_Module stopped
 4    --    AN   No_Module stopped
 5    --    AN   No_Module stopped
 6    --    AN   No_Module stopped
 7    --    AN   No_Module stopped
 8    --    AN   No_Module stopped
```

```
 9    --    AN   No_Module stopped
10    --    AN   No_Module stopped
11    --    AN   No_Module stopped
12    --    AN   No_Module stopped
13    --    AN   No_Module stopped
14    --    AN   No_Module stopped
15    id    AN   No_Sync   stopped
IBM_A16_SJC:admin>
```

We use the **fabricshow** command, which is a simple CLI command, to verify that both routers are now within this backbone fabric.

Example 6-20 Verifying that both routers are in the backbone fabric

```
IBM_A16_SFO:admin> fabricshow
Switch ID   Worldwide Name           Enet IP Addr    Name
------------------------------------------------------------
  1: fffc01 10:00:00:05:1e:16:26:ec  10.64.210.201   "IBM_A16_SJC"
100: fffc64 10:00:00:05:1e:16:26:a3  10.64.210.202  >"IBM_A16_SFO"

The Fabric has 2 switches

IBM_A16_SFO:admin>
```

You can use three other CLI commands to verify that the connection is valid. The **fcipshow *0*** command, where ***0*** is the configured port; **portshow *0***, where ***0*** is the port number; and **rnping <remoteIP> -l <packet size>**.

Example 6-21 Verifying the connection

```
IBM_A16_SFO:admin> portshow 0
        port   0  info
                Configuration   Current
Name :          port 0
State:          STARTED         UP
Type :          GIGE            GIGE
Link Status:    ENABLED         UP
IP addr:        192.168.1.202   192.168.1.202
Net mask:       255.255.255.0   255.255.255.0
Default route:  0.0.0.0         0.0.0.0
Mac address:    00:05:1e:32:77:d0

Protocol:       fcip ver 1      fcip ver 1

Licensed        : YES
```

```
Diag result     : PASSED

IBM_A16_SFO:admin>
IBM_A16_SFO:admin> fcipshow 0
----------------- fcip protocol info(port 0) --------------------
                  Configured                    Current
State:            UP                            UP
Local  IP addr:   192.168.1.202192.168.1.202
Remote IP addr:   192.168.1.201                 192.168.1.201
Link Bandwidth:   1000                          1000
Jumbo Support:    disabled                      disabled
WAN_TOV timeout:  enabledenabled
Load balance:     exchange                      none
Remote WWN:       00:00:00:00:00:00:00:00       00:00:00:00:00:00:00:00

Time sync state:  synchronized (Since Wed Nov 16 15:09:21 2005)

in_frame_ip:      0
in_frame_fc:      0
out_frame_ip:     0
out_frame_fc:     0
in_octet_ip:      0
in_octet_fc:      0
out_octet_ip:     0
out_octet_fc:     0
error_frame_ip:   0
error_frame_fc:   0
error_resync:     0
drop_frame_fc:    0
drop_frame_ip:    0
frame_timeout:    0
authen_failure:   0
IBM_A16_SFO:admin>
IBM_A16_SFO:admin> rnping 1 192.168.1.201 -l 1024
Pinging 192.168.1.201
Reply from 192.168.1.201: bytes=1024 time<14ms TTL=255
Reply from 192.168.1.201: bytes=1024 time<14ms TTL=255
Reply from 192.168.1.201: bytes=1024 time<14ms TTL=255
Reply from 192.168.1.201: bytes=1024 time<14ms TTL=255
Reply from 192.168.1.201: bytes=1024 time<14ms TTL=255
The rnping is completed
IBM_A16_SFO:admin>
```

This concludes the setup of our FCIP link. For the next stage, we add these routers into their associated SANs using EX_Ports.

Adding a router into an existing SAN

In this section, we connect our routers into their local SANs by using functionality similar to that of E_Ports; however, in a router context, these are called EX_Ports, or inter-fabric links.

In our example, we only configure a single EX_Port from the router into its SAN, but for redundancy, a number of EX_Ports should be considered.

To add a router into an existing SAN:

1. We start at the WebTools main window, as shown in Figure 6-57.

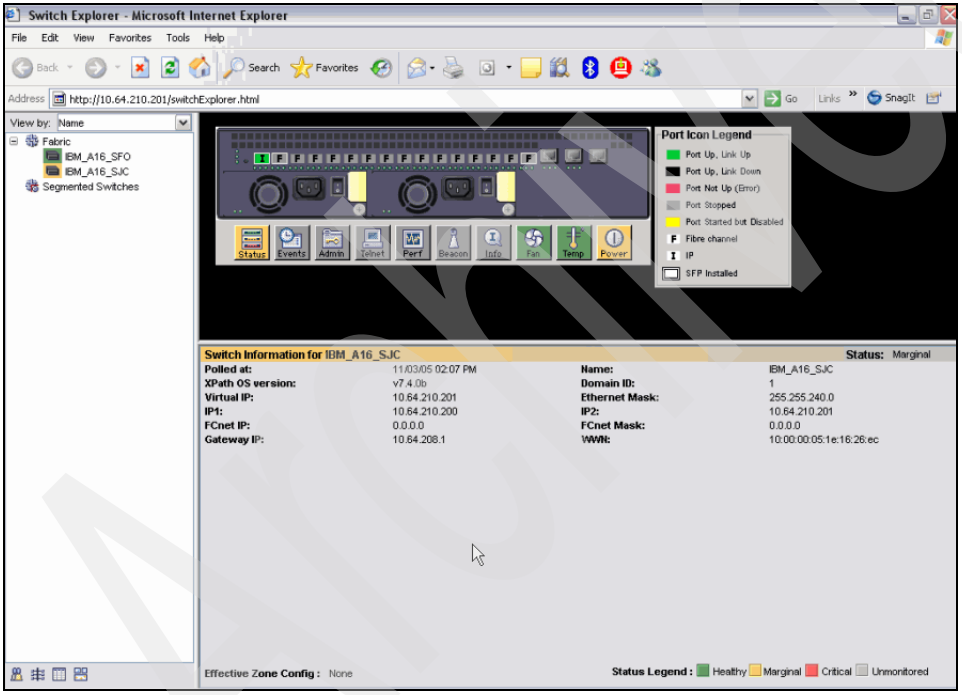


Figure 6-57 Starting point for EX_Port configuration

2. As with configuring the VE_Ports, we click the port we are going to configure. In this case, we choose port 15 as an EX_Port.

3. Here, we choose the **Fibre Channel** protocol for this port, and additionally select that we want it to be an **EX_Port** (Figure 6-58).

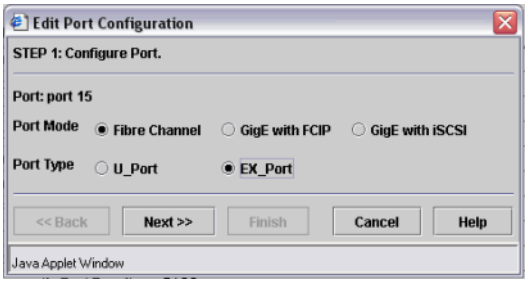


Figure 6-58 EX_Port configuration: First step

4. Figure 6-59 shows the next step of the configuration. Two configuration options are available. The FC link speed between the router and switch option provides the Auto, 1 Gbps, or 2 Gbps choices. We select **Auto**.

   The second choice is the Max Distance setting. This defines the type of long-distance mode needed for this link. The options here are L0 for normal distance, commonly up to 500 m; LE mode, up to 10k m; L0.5, up to 25 km; and LS mode, which is up to 300 km. We select **L0** mode because the switch and router are in the same rack.

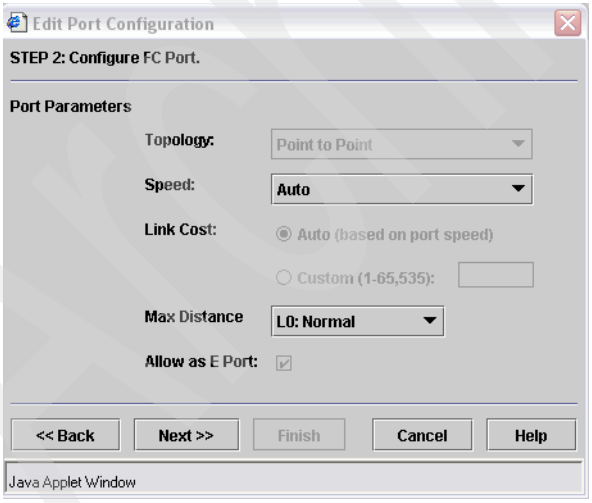


Figure 6-59 EX_Port configuration details

5. We now supply a Fabric ID, as shown in Figure 6-60. A default of 17 is specified. A fabric ID is set on every EX_Port that connects to a unique fabric. Because the router can be connected to multiple fabrics at the same time, it needs a unique ID for each fabric. For example, if we connect port 15 to fabric A and port 14 to fabric B, we need to choose a different fabric ID for ports 14 and 15. However, if both port 14 and 15 were connected to the same fabric, the fabric ID needs to be set the same on both ports.

   We are connecting a single port to one fabric, thus selecting the default is valid.

   We can also set the Interoperability Mode in this window. If we connect this port to a different vendor's switch, we have a selection of options from which to choose. Older McDATA switches, for example, 3016 or 3032, need McDATA Fabric Legacy Mode set. Newer McDATA products work in McDATA Fabric Mode.

   Because we are connecting to another Brocade product, we select **Brocade Native Mode**.

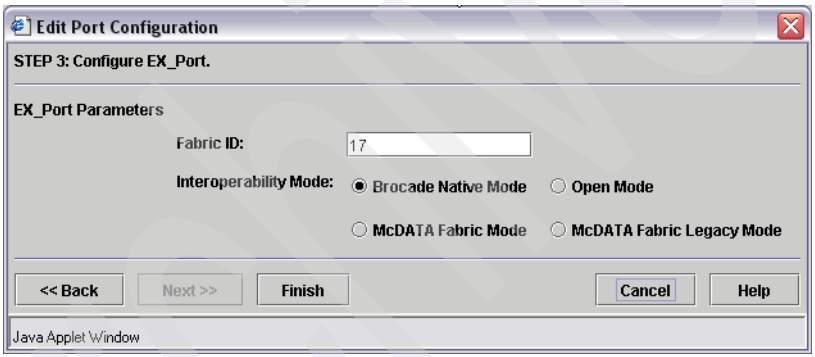


Figure 6-60 Final stage of EX_Port configuration

Figure 6-61 shows these values.

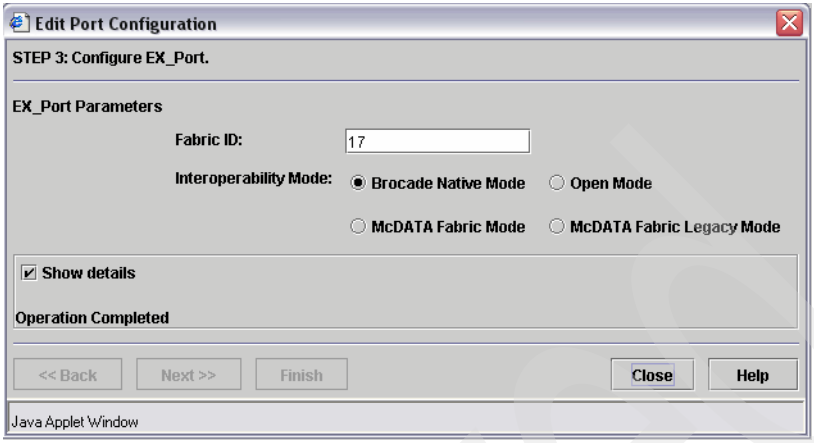


Figure 6-61 Completing the EX_Port configuration.

6. We have completed the initial configuration and now can enable and start this port, as per Figure 6-62.

Figure 6-62 Enabling the EX_Port

7. We start the port, as shown in Figure 6-63.

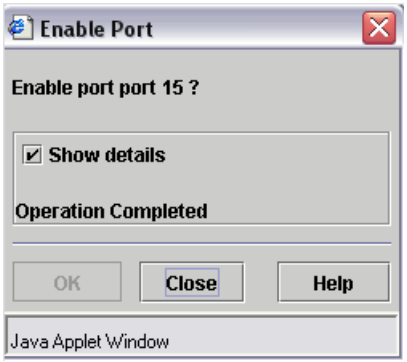


Figure 6-63 EX_Port enabled

The port is successfully started, shown here in Figure 6-64, and therefore, we should be able to see this router's connection into the existing SAN.

Figure 6-64 port-15 started

8. We can use the CLI command **switchshow** to check if the EX_Port is now online and whether we can see the WWN of the switch to which we are connected.

Example 6-22 The switchshow command

```
IBM_A16_SJC:admin> switchshow
Switch Name  :  IBM_A16_SJC
Switch State :  Online
Switch Type  :  38.0
Switch Role  :  Subordinate
Switch Domain:  1
Switch ID    :  FFFC01
Switch WWN   :  10:00:00:05:1e:16:26:ec
```

```
beacon status:  OFF
zoning       :  OFF

FC router BB Fabric ID: 1

Port Media Speed State     Info
=====================================
 0    id    1G   Online    VE_PORT  10:00:00:05:1e:16:26:a3
"IBM_A16_SFO" (upstream)
 1    id    AN   No_Light  stopped
 2    --    AN   No_Module stopped
 3    --    AN   No_Module stopped
 4    --    AN   No_Module stopped
 5    --    AN   No_Module stopped
 6    --    AN   No_Module stopped
 7    --    AN   No_Module stopped
 8    --    AN   No_Module stopped
 9    --    AN   No_Module stopped
10    --    AN   No_Module stopped
11    --    AN   No_Module stopped
12    --    AN   No_Module stopped
13    --    AN   No_Module stopped
14    --    AN   No_Module stopped
15    id    N2   Online    EX_PORT 10:00:00:05:1e:34:3f:27
"IBM_H16_SJC" (fabric id = 17)
IBM_A16_SJC:admin>
```

9. From this output, we see that both our VE_Port and EX_Port are online. The next step is to perform the same actions, adding the remote router into its local SAN.

10. After completing this, we run the **switchshow** command on that router and verify that the EX_Port connectivity is online.

Example 6-23 Running the switchshow command again

```
IBM_A16_SFO:admin> switchshow
Switch Name  :  IBM_A16_SFO
Switch State :  Online
Switch Type  :  38.0
Switch Role  :  Principal
Switch Domain:  100
Switch ID    :  FFFC64
Switch WWN   :  10:00:00:05:1e:16:26:a3
beacon status:  OFF
zoning       :  OFF
```

```
FC router BB Fabric ID: 1

Port Media Speed State     Info
======================================
 0    id    1G   Online    VE_PORT  10:00:00:05:1e:16:26:ec
"IBM_A16_SJC" (downstream)
 1    --    AN   No_Module stopped
 2    --    AN   No_Module stopped
 3    --    AN   No_Module stopped
 4    --    AN   No_Module stopped
 5    --    AN   No_Module stopped
 6    --    AN   No_Module stopped
 7    --    AN   No_Module stopped
 8    id    AN   No_Light  stopped
 9    --    AN   No_Module stopped
10    --    AN   No_Module stopped
11    --    AN   No_Module stopped
12    --    AN   No_Module stopped
13    id    AN   No_Light  stopped
14    id    N2   Online    EX_PORT 10:00:00:05:1e:34:81:74
"IBM_B32_SFO" (fabric id = 16)
15    id    N2   Online    EX_PORT 10:00:00:05:1e:34:59:a7
"IBM_H16_SFO" (fabric id = 16)
IBM_A16_SFO:admin>
```

We can see here that both EX_Ports are online. Because they both connect into the same SAN, we gave them the same fabric ID (16).

Now, the physical connectivity is complete, we define which devices are allowed to communicate with each other. We do this by using LSAN zones.

Setting up LSAN zones

Before starting to configure the LSAN zones, we verify the current zoning configuration of each fabric.

The San Francisco fabric zoning configuration shows each host as having a single HBA defined that can access two JBOD disks.

Example 6-24 San Francisco fabric zoning configuration

```
cfg
   San_FRANCISCO_DATACENTER
      server_blue_disk
      server_green_disk
```

```
zone
   server_blue_disk
      jbod_blue_1
      jbod_blue_2
      x330_blue
   server_green_disk
      x330_green
      jbod_green_2
      jbod_green_1

alias.jbod_blue_1:20:00:00:04:cf:92:74:99
alias.jbod_blue_2:20:00:00:04:cf:bd:56:3e
alias.jbod_green_1:20:00:00:04:cf:bd:71:a0
alias.jbod_green_2:20:00:00:04:cf:bd:71:a8

alias.x330_blue:20:00:00:00:c9:2b:7f:90
alias.x330_green:20:00:00:00:c9:30:71:93

alias.jbod_green_4:21:00:00:04:cf:bd:72:bb
alias.jbod_green_3:21:00:00:04:cf:bd:72:a1
alias.jbod_blue_3:20:00:00:04:cf:bd:72:a1
alias.jbod_blue_4:20:00:00:04:cf:bd:72:bb
```

In the San Jose SAN zoning configuration, we see an identical setup, where each hosts has a single HBA defined that can access two JBOD disks.

Example 6-25 San Jose SAN zoning configuration

```
cfg
   SAN_JOSE_DATACENTER
      x330_sj_bottom
      x330_sj_top
zone
   x330_sj_bottom
      x330_SJ_bottom_hba0
      seagate_sj_03
      seagate_sj_04
   x330_sj_top
      x330_SJ_top_hba0
      seagate_sj_01
      seagate_sj_02

alias.seagate_sj_01:20:00:00:20:37:42:66:bb
alias.seagate_sj_02:20:00:00:20:37:42:65:71
alias.seagate_sj_03:20:00:00:20:37:42:65:6e
```

```
alias.seagate_sj_04:20:00:00:20:37:42:66:2c

alias.x330_SJ_bottom_hba0:10:00:00:00:c9:30:71:57
alias.x330_SJ_top_hba0:10:00:00:00:c9:2b:7f:98

alias.seagate_sj_05:20:00:00:20:37:15:0b:9a
alias.seagate_sj_06:20:00:00:20:37:15:17:6b
```

Checking the Device Manager on one of the SJ hosts, we confirm the configuration because it shows two SAN-connected SCSI disk devices (Figure 6-65).

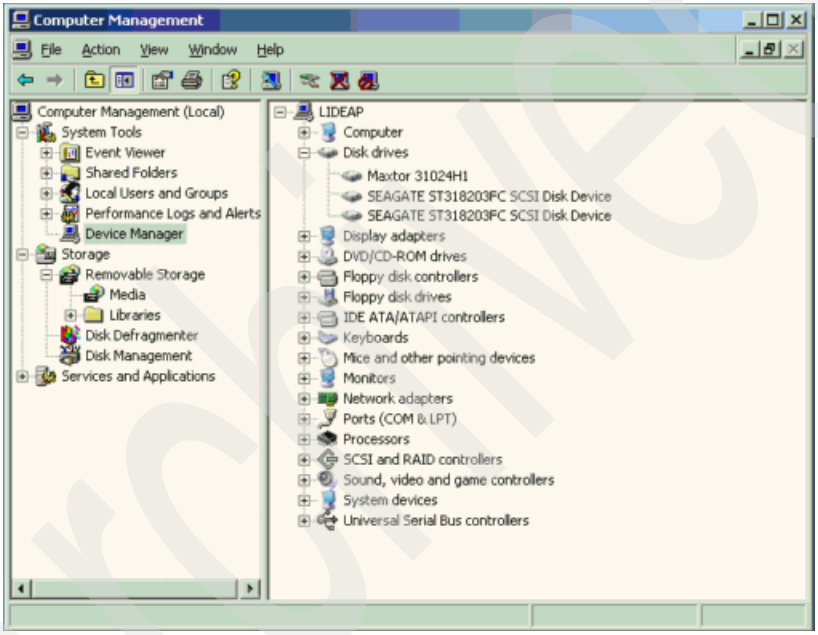


Figure 6-65 Device Manager showing the two local SAN-connected FC disks

With the b-type SAN router, it is important to remember that, for any devices we want the routers to have knowledge of, they must be put into a zone that starts with the five characters LSAN_. Note the underscore after LSAN. LSAN can be in uppercase or lowercase letters, for example, LSAN_zone_SAN-AtoB or lsan-zone_sanBtoA. “LSAN” is the tag that the router uses to acknowledge which devices it is allowed to export to the other SAN.

Here, we map a single JBOD disk from the San Francisco SAN to each of the hosts in the San Jose SAN.

To do this, we need to create LSAN zones for each SAN that contains the devices to be exported. The contents of these zones must be the same for each SAN or the devices will not be exported. However, it is not necessary for the zone names to be identical.

From the SJ SAN, we need the WWN of the HBAs. From the earlier configuration information, we know these are:

```
alias x330_SJ_bottom_hba0 10:00:00:00:c9:30:71:57
alias x330_SJ_top_hba0 10:00:00:00:c9:2b:7f:98
```

From the SFO SAN, we choose two unused disks, one for each of the previous HBAs:

```
alias jbod_blue_3 21:00:00:04:cf:bd:72:bb
alias jbod_blue_4 21:00:00:04:cf:bd:72:a1
```

One point to keep in mind is that alias names are only defined in the local SANs. So when we create LSAN zones, they will contain a mixture of local alias and remote WWNs.

In the SJ SAN we create two LSAN zones, one for each HBA:

- The first zone uses our local HBA alias and the WWN of one of the remote disks:

  ```
  x330_SJ_bottom_hba0 and 21:00:00:04:cf:bd:72:bb
  ```

- For the second zone, we use:

  ```
  x330_SJ_top_hba0 and 21:00:00:04:cf:bd:72:a1
  ```

After completing the setup in the SJ SAN, we show the following zoning information.

Example 6-26 Zoning information on SJ SAN

```
zone.
   LSAN_sj_bot_SFO_disk:
      x330_SJ_bottom_hba0;
      21:00:00:04:cf:bd:72:bb

   LSAN_sj_top_SFO_disk:
      x330_SJ_top_hba0;
      21:00:00:04:cf:bd:72:a1
```

Now, we perform the same operation in the SFO SAN, taking the WWN of the adapters in the SJ zone and using the aliases of our chosen disks to create the following LSAN zoning configuration.

Example 6-27 Zoning information on SFO SAN

```
zone.
   LSAN_sfo_disk_SJ_host_bot:
      jbod_green_4;
      10:00:00:00:c9:30:71:57

   LSAN_sfo_disk_SJ_host_top:
      jbod_green_3;
      10:00:00:00:c9:2b:7f:98
```

For simplicity, we use WebTools to create our LSAN zones. We do this on each fabric separately. However, with Fabric Manager 5.0, the LSAN zone wizard automatically propagates the same zone information into multiple SANs where necessary.

Figure 6-66 shows the final zone setup from the SJ SAN, and Figure 6-67 on page 136 shows the final configuration from the SFO SAN.

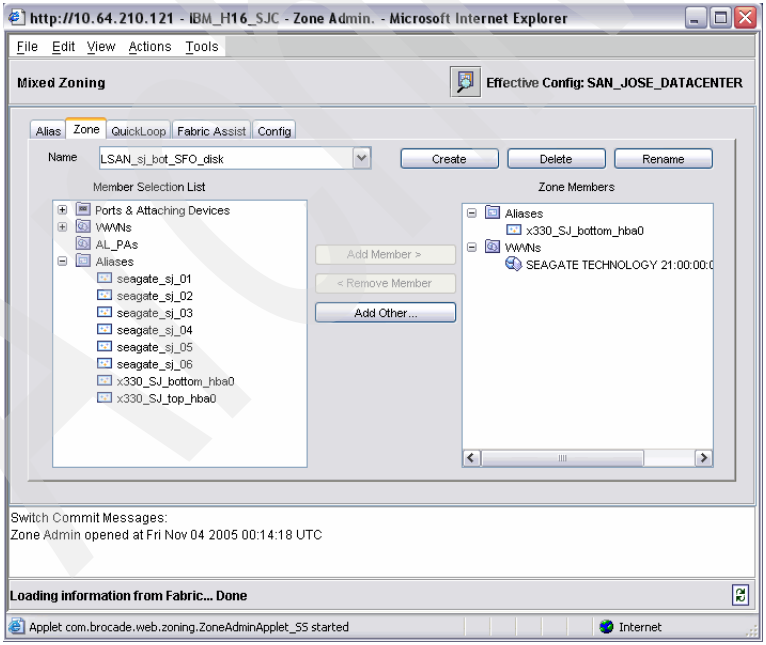


Figure 6-66 SJ final zoning configuration

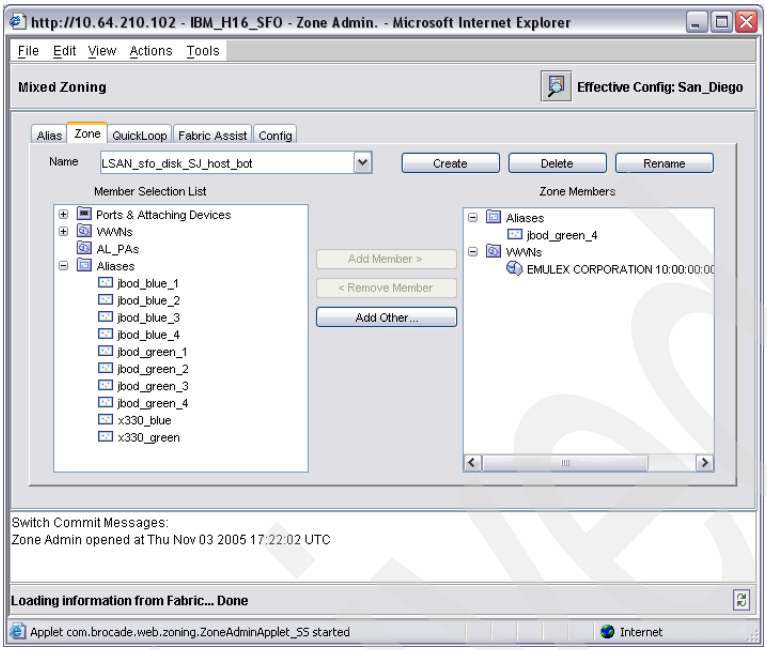


Figure 6-67 SFO final LSAN zoning configuration

After updating both SANs to include the new LSAN zones, we enable this new zone configuration.

Using the **lsanzoneshow** router CLI command, we verify that each router can see both copies of the LSAN zone information.

Example 6-28 Using the lsanzoneshow command

```
IBM_A16_SFO:admin> lsanzoneshow
Fabric ID: 16 Zone Name: LSAN_sfo_disk_SJ_host_bot
        21:00:00:04:cf:bd:72:bb
        10:00:00:00:c9:30:71:57
Fabric ID: 16 Zone Name: LSAN_sfo_disk_SJ_host_top
        21:00:00:04:cf:bd:72:a1
        10:00:00:00:c9:2b:7f:98
Fabric ID: 17 Zone Name: LSAN_sj_bot_SFO_disk
        10:00:00:00:c9:30:71:57
        21:00:00:04:cf:bd:72:bb
Fabric ID: 17 Zone Name: LSAN_sj_top_SFO_disk
        10:00:00:00:c9:2b:7f:98
        21:00:00:04:cf:bd:72:a1
IBM_A16_SFO:admin>
```

We should have the same output from the SJ router.

Example 6-29 Output from the SJ router

```
IBM_A16_SJC:admin> lsanzoneshow
Fabric ID: 16 Zone Name: LSAN_sfo_disk_SJ_host_bot
        21:00:00:04:cf:bd:72:bb
        10:00:00:00:c9:30:71:57
Fabric ID: 16 Zone Name: LSAN_sfo_disk_SJ_host_top
        21:00:00:04:cf:bd:72:a1
        10:00:00:00:c9:2b:7f:98
Fabric ID: 17 Zone Name: LSAN_sj_bot_SFO_disk
        10:00:00:00:c9:30:71:57
        21:00:00:04:cf:bd:72:bb
Fabric ID: 17 Zone Name: LSAN_sj_top_SFO_disk
        10:00:00:00:c9:2b:7f:98
        21:00:00:00:04:cf:bd:72:a1
IBM_A16_SFO:admin>
```

This output confirms that we successfully transferred information between routers, albeit zoning information. It also confirms that both routers correctly see the same zoning information.

To provide further confirmation that our devices are correctly configured for exporting, we use the **fcrphydevshow** command. This displays the routers local devices. The **fcrproxydevshow** command shows the remote devices.

For the SFO SAN, we export local disks, so these should be seen in output of the **fcrphydevshow** command. The remote devices would be SJ HBAs, and therefore, these should be displayed in the output of the **fcrproxydevshow** command. This output confirms that the router has knowledge of the two local disks it intends to export.

Example 6-30 SFO: The fcrphydevshow command

```
IBM_A16_SFO:admin> fcrphydevshow
 Device          WWN            Physical
 Exists                           PID
in Fabric
----------------------------------------
   16    21:00:00:04:cf:bd:72:a1  010fe8
   16    21:00:00:04:cf:bd:72:bb  010fe2
IBM_A16_SFO:admin>
```

The output shows that the SFO router can correctly see the remote HBA devices from the SJ SAN.

Example 6-31 SFO: The fcrproxydevshow command

```
IBM_A16_SFO:admin> fcrproxydevshow
  Proxy           WWN             Proxy      Device   Physical    State
 Created                           PID       Exists     PID
in Fabric                                   in Fabric
-------------------------------------------------------------------------
-----
    16   10:00:00:00:c9:2b:7f:98  02f002       17      010100
Imported
    16   10:00:00:00:c9:30:71:57  02f001       17      010000
Imported
IBM_A16_SFO:admin>
```

Now, we check the same is true for the SJ router. This confirms that SJ can see its local devices correctly.

Example 6-32 SJ: The fcrphydevshow command

```
IBM_A16_SJC:admin> fcrphydevshow
 Device           WWN            Physical
 Exists                           PID
in Fabric
---------------------------------------
   17    10:00:00:00:c9:2b:7f:98  010100
   17    10:00:00:00:c9:30:71:57  010000
IBM_A16_SJC:admin>
```

This confirms that SJ can see the remote devices correctly.

Example 6-33 SJ: The fcrproxydevshow command

```
IBM_A16_SFO:admin> fcrproxydevshow
  Proxy           WWN             Proxy      Device   Physical
State
 Created                           PID       Exists     PID
in Fabric                                   in Fabric
-------------------------------------------------------------------------
-----
    16   10:00:00:00:c9:2b:7f:98  02f002       17      010100
Imported
    16   10:00:00:00:c9:30:71:57  02f001       17      010000
Imported
IBM_A16_SFO:admin>
```

Now all that remains is to scan for new devices on our host at SJ.

As shown in Figure 6-68, our SJ Windows host can now see a third remote disk that was exported from our SFO SAN. Perform the same device probe on all remaining hosts that are expected to see remote disks.

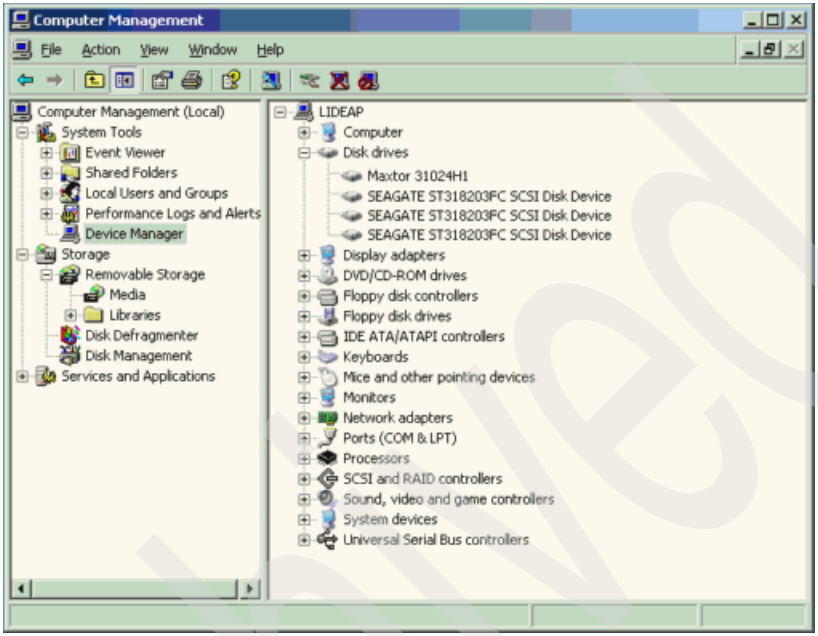


Figure 6-68 Device Manager showing a third remotely attached Seagate disk

This concludes our basic implementation of an FCIP routed Meta SAN. For further details of FCIP implementation procedures, refer to the *XPath OS Procedures Guide*, which is available for downloading from the document library at the Brocade Web site.

6.1.7 Configuring a routing service using FC to FC

Configuring an FC-to-FC routing service between two separate edge fabrics or SANs is very similar to the previous section where we configured an FC-to-FC routing service through FCIP. The same implementation process is applied, but without creating a backbone fabric over FCIP.

A single router can be used to route between separate fabrics, or a number of interconnected routers, using standard E_Port connectivity to provide a higher port count in the routing backbone fabric.

For a simple one router FC-to-FC routing fabric, we perform the following procedure:

1. Physically install the router into a rack and turn it on.
2. Set up IP connectivity using the serial port.
3. Use WebTools to perform the following tasks:
   a. Install the appropriate licenses where required.
   b. Select appropriate ports to be used as EX_Ports or inter-fabric links, and ensure that they are stopped and disabled.
   c. Configure the selected ports as EX_Ports.
   d. Physically connect the router through these configured ports to each of the separate SANs.
   e. Enable and start the selected ports.
   f. Use the FC-FC Routing panel in Switch Manager to confirm that the edge fabrics can be viewed by the router.
   g. Create a single LSAN zone on each of the fabrics, containing one of the host HBAs and target HBAs to be routed between.
   h. Use the FC-FC Routing panel in Switch Manager to ensure that the LSAN zones can be viewed by the router.
   i. Use the FC-FC Routing panel to view the exported/imported devices from each separate fabric, also known as physical versus proxy devices.
4. Use the host operating system to probe for new devices.
5. Assuming the devices are found correctly, continue to add further LSAN zones to each edge fabric.
6. Reprobe the edge fabrics from the appropriate hosts to find the new routed devices.

We describe the details for each of these steps in 6.1.6, “Configuring a routing service using FCIP” on page 112.

6.1.8 Configuring an iSCSI gateway service

You can configure SAN16B-R router as an iSCSI gateway.

The implementation of an iSCSI gateway environment is beyond the scope of this book. However, the Brocade *XPath OS Procedures Guide* covers the basic implementation steps required to set up this type of environment.

6.2 Installing the SAN18B-R and M48 FC Routing Blade

This section contains a detailed description of setting up the SAN18B-R and the FC Routing Blade. We begin with the initial setup and then continue with the FCIP and LSAN configuration. Finally, we mention the FCR administration.

Most of the procedures are exactly the same on the SAN18B-R and the Routing Blade. There are some differences in the initial setup part, because the Routing Blade is installed in the M48 Director. We perform some initial tasks, such as IP address configuration, not on the FC Routing Blade, but on the CP blade in M48 Director.

The main installation and configuration tasks are as follows:

1. Set up IP addresses through the serial port.
2. Upgrade the switch to the latest supported firmware version.
3. Attach the router to the management LAN and connect to its management interface (with CLI or WebTools).
4. Configure FCIP links between the two routers at different sites.
5. Configure EX_Ports between the routers and their local SANs.
6. Set up LSAN zones allowing the routing of specific devices.

6.2.1 Initial setup

The first step is to physically install the 2005-R18 into a rack. Next, connect the appropriate power cables to the router and turn it on.

After the router completes its power-up sequence, we need to perform some initial configuration steps.

Power-up sequence

This should not take more than five minutes to complete. During the POST sequence, the port LEDs light up in sequence from left to right. If you do not see a change on the front of the router after three minutes, turn if off for one minute, and then try to power it up again. After the POST completes successfully, the power and attention LEDs are green. The port lights flash slowly.

Connecting to the serial port

To connect to the serial port, perform the following steps:

1. To connect to the router serial port, use the provided console cable. It is a rolled RJ45 to DB9. The standard b-type straight-through (non-null modem) serial cable will not work. We first connect the RJ45 end to the console port on the router, which is the RJ45 on the left side, labeled “10101”. Connect the DB9 end to either a mobile computer or a Windows machine that is close to the router.
2. Start HyperTerminal from the PC or mobile computer. Select the appropriate COM port that is associated to the serial connection (Figure 6-69).

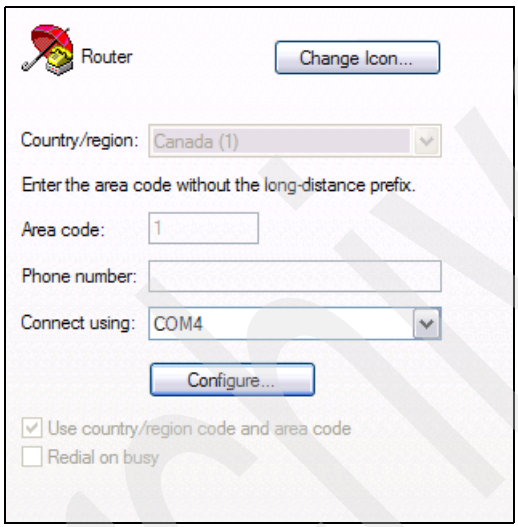


Figure 6-69 Selecting the correct COM port

3. The settings for this connection are the same as all the other b-type family of switches: **9600**, **8**, **None**, **1**, **None**, as shown in Figure 6-70.

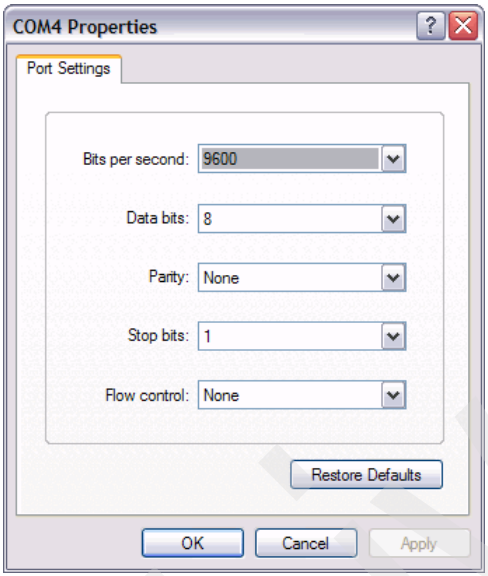


Figure 6-70 Hyperterm connection settings

4. After clicking the **OK** button, we press Enter on the keyboard to get a login prompt.
5. The default login details are also the same as all other b-type switches. The login ID is **admin** and default password is **password** (Example 6-34).

   After entering the default login and password details, we press Ctrl+C to escape from the request to change the default password. You can change these at a later date.

Example 6-34 Initial login

```
Fabric OS (IBM-2005-R18)


IBM-2005-R18 console login: admin
Password:
Please change passwords for switch default accounts now.
Use Control-C to exit or press 'Enter' key to proceed.

Password was not changed. Will prompt again at next login
until password is changed.
IBM-2005-R18:admin>
```

Setting the IP address

Now we are connected and logged into the switch. The first task is to set up the IP address so that we can add the router to our LAN and start to manage it through the GUI interface. Therefore, we need to use the **ipaddrshow** command to display our current IP settings and the **ipaddrset** command to change these settings (Example 6-35).

Example 6-35 Displaying the IP address

```
IBM-2005-R18:admin> ipaddrshow

SWITCH
Ethernet IP Address: 10.77.77.77
Ethernet Subnetmask: 255.255.255.0
Fibre Channel IP Address: 0.0.0.0
Fibre Channel Subnetmask: 0.0.0.0
Gateway Address: 0.0.0.0
```

From the output of the **ipaddrshow** command, we see the default IP address is 10.77.77.77, the default subnet mask is 255.255.255.0, and the default gateway is not defined.

Now we use **ipaddrset** to set our own LAN settings (Example 6-36). This command is interactive and prompts us for each line of input. We can accept the defaults, or input new values.

Example 6-36 Changing the IP address

```
IBM-2005-R18:admin> ipaddrset
Ethernet IP Address [10.77.77.77]: 10.64.210.253
Ethernet Subnetmask [255.255.255.0]:
Fibre Channel IP Address [0.0.0.0]:
Fibre Channel Subnetmask [0.0.0.0]:
Gateway IP Address [0.0.0.0]: 10.64.210.1
Issuing gratuitous ARP...Done.
2006/08/22-21:13:53, [WEBD-1007], 22,, INFO, IBM-2005-R18, HTTP server
will be restarted due to change of IP Address
IP address is being changed...Done.
Committing configuration...Done.
```

Now, we physically connect the Ethernet interface to our LAN and point our Web browser at the router's IP address to access the WebTools management GUI. We also need this connectivity to fetch the firmware from our FTP server for the firmware upgrade we do in the next section. So, connect your router to your network now, and ensure that there is IP connectivity.

Next, we set the switch's name using the **switchName** command (Example 6-37).

Example 6-37 Setting the switch name

```
IBM-2005-R18:admin> switchName
IBM-2005-R18
IBM-2005-R18:admin> switchname IBM_R18_SJC
Committing configuration...
Done.
IBM_R18_SJC:admin> switchname
IBM_R18_SJC
IBM_R18_SJC:admin>
```

Upgrading the firmware

To perform a firmware upgrade, use the **firmwaredownload** command. The SAN18B-R router uses the same FOS as the rest of the Brocade family.

To download the latest firmware for the 2005-R18 router, we start at the following URL:

http://www.ibm.com/servers/storage/support/san/2005-r18/downloading.html

From here, we click the **5.x firmware** link. This redirects us to the IBM support section of the Brocade Web site, as shown in Figure 6-71.

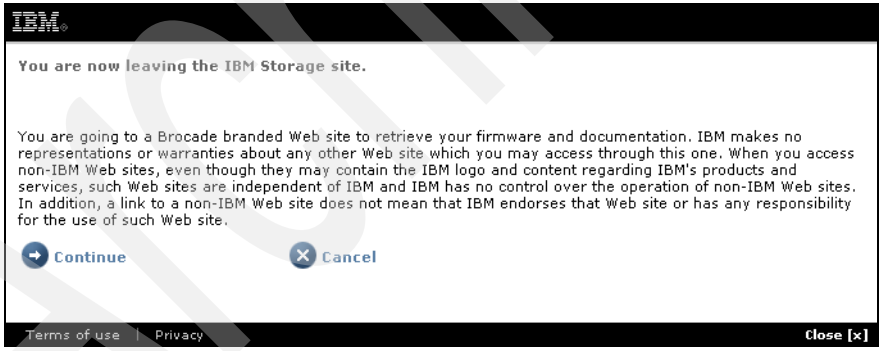


Figure 6-71 IBM re-direct warning

From the Brocade Web site, we click the appropriate version of OS firmware to download. After downloading the code, we need to decide on an FTP mechanism to get the code to the router. If you use Fabric Manager 5.x with its built-in FTP server, we can simply download code onto this machine. Otherwise, we need to copy the firmware onto a machine that has an FTP server available, which the router can contact. After downloading the code and copying it to a suitable machine, we unzipped it.

Important: We recommend that you read the release notes for the version of FOS you will be installing. For release notes, refer to the following Web page:

http://www.ibm.com/support/docview.wss?rs=1155&uid=ssg1S1002711

Installing the code

We use the **version** command from the router CLI to display the currently running code level.

Example 6-38 Displaying the code version

```
IBM-2005-R18:admin> version
Kernel:     2.4.19
Fabric OS:  v5.1.0b
Made on:    Fri May 26 23:58:16 2006
Flash:      Wed Jun 28 04:28:33 2006
BootProm:   4.5.3
```

From here, we see version 5.1.0b installed; therefore, we proceed with our upgrade to version 5.1.0c.

Tip: It is best to make a backup of the configuration prior to performing the firmware update. Use the CLI **configupload** command.

Our router is brand new and does not currently have an active configuration to backup, so we continue with our update.

Next, we use the **firmwaredownload** command.

Before starting the firmware download process, it is always a good idea to alter the 10 minute Telnet session timeout value. Setting the value to zero, sets the session to never timeout. We use the **timeout** command, and then we log out and back in as suggested.

Example 6-39 Setting the idle timeout value

```
IBM-2005-R18:admin> timeout
Current IDLE Timeout is 10 minutes
IBM-2005-R18:admin> timeout 0
IDLE Timeout Changed to 0 minutes
The modified IDLE Timeout will be in effect after NEXT login

Logout then in
```

```
IBM-2005-R18:admin> timeout
Current IDLE Timeout is 0 minutes
```

We unzipped our firmware file and set our FTP server to point to the directory. We then specify the release.plist file as the file name when requested.
Figure 6-72 shows the structure of the unzipped firmware file. We unzipped the file into the code directory on the C drive.

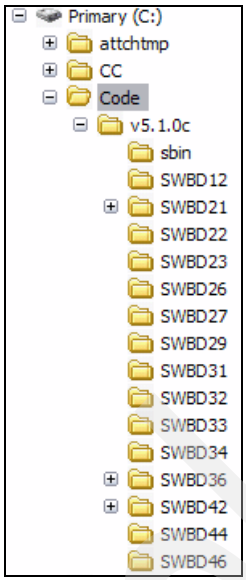


Figure 6-72 Firmware file structure

Figure 6-73 demonstrates our FTP program setup. You can see that the server's root directory is our code's root directory.

Tip: If you have issues contacting your FTP server, try disabling your firewall.

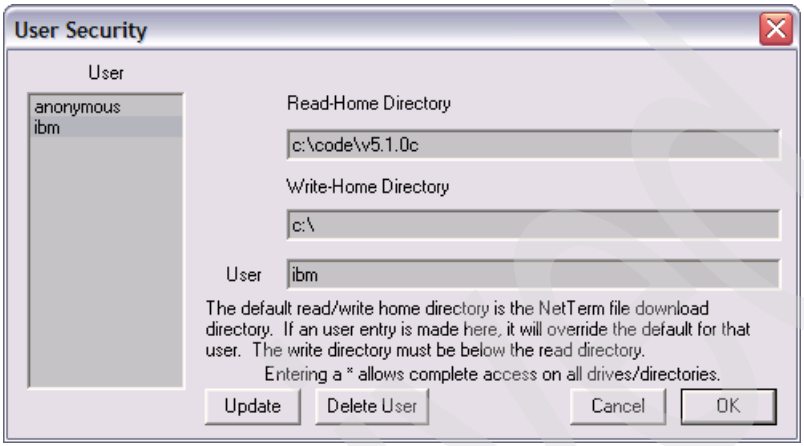


Figure 6-73 FTP server setup

Upgrading the firmware causes the switch to reboot during **firmwaredownload**. Your Telnet session might be lost because the switch reboots automatically at least once, and the whole process can take up to 30 minutes. Example 6-40 takes us through the update.

Example 6-40 Firmware update (output shortened)

```
IBM-2005-R18:admin> firmwaredownload
Server Name or IP Address: 10.64.210.100
FTP User Name: ibm
File Name: release.plist
FTP Password:
You can run firmwaredownloadstatus to get the status
of this command.

This command will cause the switch to reset and will
require that existing telnet, secure telnet or SSH
sessions be restarted.

Do you want to continue [Y]: y
Firmware is being downloaded to the switch. This step may take up to 30
minutes.
```

```
2006/08/22-22:31:54, [SULB-1001], 26,, WARNING, IBM_2005_R18,
Firmwaredownload command has started.
2006/08/22-22:31:54, [SULB-1036], 27,, INFO, IBM_2005_R18, The current
Version:
Fabric OS v5.1.0b

Checking system settings for firmwaredownload...
Start to install packages...
dir
################################################
ldconfig
################################################

....

Please avoid powering off the system during prom update.

....

Removing unneeded files, please wait ...
Finished removing unneeded files.

All packages have been downloaded successfully.
Firmware has been downloaded to the secondary partition of the switch.
2006/08/22-22:48:41, [SULB-1002], 28,, INFO, IBM_2005_R18,
Firmwaredownload command has completed successfully.
HA Rebooting ...
IBM-2005-R18:admin> Loopback backup before go standby

.....

Firmware commit completes successfully.
2006/08/22-22:53:05, [SULB-1036], 39,, INFO, IBM_2005_R18, The new
Version: Fabric OS v5.1.0c

2006/08/22-22:53:05, [SULB-1002], 40,, INFO, IBM_2005_R18,
Firmwaredownload command has completed successfully.
```

The firmware download has completed successfully, as shown in the previous message.

Verify the level with the **firmwareshow** command.

Example 6-41 Displaying current firmware level

```
IBM-2005-R18:admin> firmwareshow
Primary version:        v5.1.0c
Secondary version:      v5.1.0c
```

We can use the **version** command as well.

Example 6-42 Displaying current version

```
IBM-2005-R18:admin> version
Kernel:     2.4.19
Fabric OS:  v5.1.0c
Made on:    Thu Jun 29 22:30:10 2006
Flash:      Tue Aug 22 22:40:43 2006
BootProm:   4.5.3
```

The output clearly shows that our code load process completed successfully. For further firmware upload/download options, refer to the *Fabric OS Administrator's Guide*.

Initial router settings

Before starting the configuration process of adding this router into a Meta SAN, ensure the time, date, and time zone are correct and that the correct licenses are installed.

Time services

All routers within a SAN backbone fabric or Meta SAN need their time set correctly. In fact, best practices recommend all routers and switches (where available) use the NTP service to synchronize their time to a known, good, external time source.

To correctly set the router's date, we use the **date** command, in the following format (Example 6-43):

`date "mmddhhmmyy"`

Example 6-43 Setting the date and time

```
IBM-2005-R18:admin> date
Tue Aug 22 23:12:46 UTC 2006
IBM-2005-R18:admin> date "0822160506"
Tue Aug 22 16:05:00 UTC 2006
```

Now, we set up the time zone and NTP services. To set the time zone, we use the **tstimezone** command, in the format, **tstimezone** ***gmtoffset***. Without a parameter specified, we see the current setting, as shown in Example 6-44.

Example 6-44 Displaying the time zone

```
IBM-2005-R18:admin> tstimezone
Time Zone Hour Offset: 0
Time Zone Minute Offset: 0
```

We are in the Pacific time zone, which is -8 hours from GMT. From the message displayed, you can see that a reboot is needed for the change to take effect. We save the reboot until we are done configuring with the CLI.

Example 6-45 Setting the time zone

```
IBM-2005-R18:admin> tstimezone -8
Updating Time Zone configuration...done.
System Time Zone change will take effect at next reboot.
```

Now we use the **tsclockserver** command to set an NTP server address for the router to talk with. The command syntax is **tsclockserver "*ip address*"**. By default, NTP is set to local or LOCL.

Example 6-46 Displaying current NTP server

```
IBM-2005-R18:admin> tsclockserver
LOCL
```

Tip: We recommend that you configure a known, good NTP server. A local server is ideal, but an Internet one works just as well.

Licensing

The following licenses are available for a 2005-R18:

- Web
- Zoning
- Base Switch
- Trunking
- Fibre Channel Routing Services
- FCIP

To view the currently installed licenses, we use the CLI command **licenseshow**.

Example 6-47 Displaying licenses

```
IBM-2005-R18:admin> licenseshow
bcyczebSyycdzd0G:
    Web license
bSQ9Syb9bcTRATW:
    Zoning license
RScQQ9RdzeSTdRRb:
    Fabric license
bRSdyRRRQdcSTezb:
    Extended Fabric license
eybeczdyhzcfdd:
    Fabric Watch license
eybeczdyf7cfdj:
    FCIP license
```

Licenses are added or removed using either the **licenseadd** or **licenseremove** command.

The implementation we perform requires the FC Routing Services and FCIP. FC Routing Services is part of the base Fabric OS on the SAN18B-R and the M48 FC Routing blade. An FCIP license is required, but is already installed on our router, so we can continue.

We reboot our switch now.

This completes the initial setup of a 2005-R18 router.

We connect two routers together through an FCIP link to form a backbone to edge fabric. Each router device in this fabric needs the same initial setup performed:

1. Configure an IP address.
2. Upgrade to the latest firmware.
3. Set up the time services correctly.
4. Ensure that the appropriate licenses are installed.

In this section, we addressed all these steps.

6.2.2 Configuring FCIP and LSAN

In this section, we establish a connection between two edge fabrics across a backbone fabric, using two SAN routers and an FCIP link. We then set up LSAN zones using the devices from both edge fabrics.

You can configure the FCIP connection using the command-line interface (CLI) or with WebTools. We show both in the following sections.

The FCIP configuration steps are as follows:

1. Check if the FCIP license is applied.
2. Enable persistently disabled ports.
3. Configure the virtual GigE ports.
4. Define an IP interface on each virtual GigE port.
5. Optionally, define IP routes on GigE ports.
6. Verify IP connectivity.
7. Configure the FCIP tunnel.
8. Verify the FCIP tunnel setup.
9. Check that the VE_Port or VEX_Port is online and that the FCIP tunnel is online.

Lab configuration

Let us describe our lab configuration, which we use to set up the FCIP and LSAN (Figure 6-74).

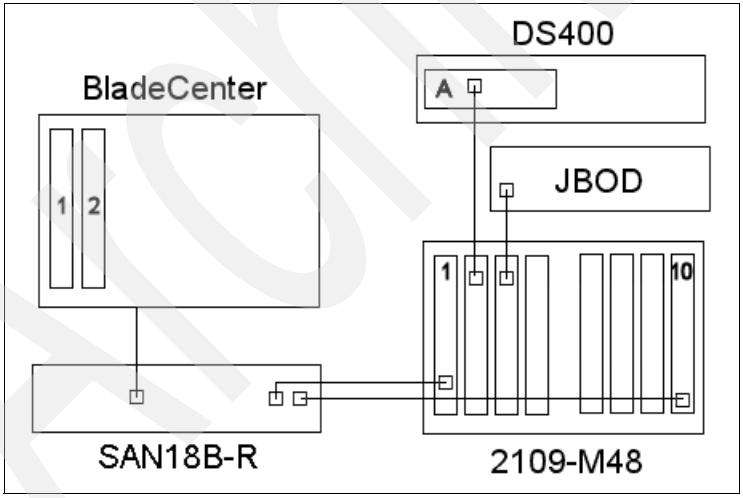


Figure 6-74 Our lab configuration

We want to establish a connection between the two Blade servers in the BladeCenter and the logical drives defined on the DS400 Storage Server across an FCIP link between two SAN routers (a SAN18B-R and an M48 FC Routing

Blade). We have two FC Routing Blades installed in the M48 Director, in slot 1 and 10.

In our first example, we show how to set up the FCIP and LSAN using the CLI. We use the SAN18B-R and the M48 FC Routing Blade in slot 1. In addition to DS400 logical drives, we also provide access to a disk drive in a separate JBOD enclosure.

Our second example demonstrates how to configure the FCIP and LSAN using GUI tools. In this case, we use the SAN18B-R and the M48 FC Routing Blade in slot 10.

FCIP setup with CLI

We configure an FCIP link between our 2005-R18 and an FC Routing Blade in slot 1 of our 2109-M48, maintaining separate, distinct SANs, through the CLI interface. We configure paths over our FCIP link using LSANs to assign disk from the backbone SAN to our server located in our edge fabric.

FCIP configuration

Launch your favorite Telnet program. We use putty, a freeware program that can easily be downloaded from the Internet.

We start our configuration on our edge fabric router, the 2005-R18, named IBM_R18_SJC. This end of the link will end up being a VE_Port.

After connecting to the switch, we need to log in to it. After logging in, we run a **switchshow** command just to verify the status of our switch and to ensure that we connected to the correct one. In addition, we explain some of the output. If this is your first time working with the Brocade router product, the port listing does need explanation. Example 6-48 shows the output.

Example 6-48 switchshow command

```
login as: admin
admin@10.64.210.253's password:

IBM_R18_SJC:admin> switchshow
switchName:     IBM_R18_SJC
switchType:     46.2
switchState:    Online
switchMode:     Native
switchRole:     Principal
switchDomain:   2
switchId:       fffc02
switchWwn:      10:00:00:05:1e:37:6d:ec
```

```
zoning:         ON (IBM_SJC_CONFIG)
switchBeacon:   OFF
FC Router:      ON
FC Router BB Fabric ID: 1

Area Port Media Speed State
==============================
  0   0   id    N4   No_Light
  1   1   --    N4   No_Module
  2   2   id    N4   No_Sync   Disabled (Persistent)
  3   3   --    N4   No_Module Disabled (Persistent)
  4   4   id    N4   No_Sync   Disabled (Persistent)
  5   5   --    N4   No_Module Disabled (Persistent)
  6   6   id    N4   No_Sync   Disabled (Persistent)
  7   7   --    N4   No_Module Disabled (Persistent)
  8   8   --    N4   No_Module Disabled (Persistent)
  9   9   --    N4   No_Module Disabled (Persistent)
 10  10   --    N4   No_Module Disabled (Persistent)
 11  11   id    N4   No_Sync   Disabled (Persistent)
 12  12   --    N4   No_Module Disabled (Persistent)
 13  13   --    N4   No_Module Disabled (Persistent)
 14  14   --    N4   No_Module Disabled (Persistent)
 15  15   --    N4   No_Module Disabled (Persistent)
 16  16   --    --   Offline   Disabled (Persistent)
 17  17   --    --   Offline   Disabled (Persistent)
 18  18   --    --   Offline   Disabled (Persistent)
 19  19   --    --   Offline   Disabled (Persistent)
 20  20   --    --   Offline   Disabled (Persistent)
 21  21   --    --   Offline   Disabled (Persistent)
 22  22   --    --   Offline   Disabled (Persistent)
 23  23   --    --   Offline   Disabled (Persistent)
 24  24   --    --   Offline   Disabled (Persistent)
 25  25   --    --   Offline   Disabled (Persistent)
 26  26   --    --   Offline   Disabled (Persistent)
 27  27   --    --   Offline   Disabled (Persistent)
 28  28   --    --   Offline   Disabled (Persistent)
 29  29   --    --   Offline   Disabled (Persistent)
 30  30   --    --   Offline   Disabled (Persistent)
 31  31   --    --   Offline   Disabled (Persistent)
     ge0  id    1G   Online
     ge1  id    1G   Online
```

The 2005-R18 has 16 FC ports and two GigE ports. These ports are shown in the list as ports 0 through 15, ge0 and ge1. The listed ports 16 through 31 are in bold, because these ports represent the 16 FCIP tunnels that can be configured

on either of the two GigE ports. The Brocade implementation of the FCIP tunnel allows for a maximum of eight tunnels per GigE interface. We have two GigE interfaces, ge0 and ge1. The first eight ports listed, ports 16 through 23, correspond to FCIP tunnels 0 through 7 on GigE port ge0. Ports 24 through 31 represent FCIP tunnels 0 through 7 on GigE port ge1.

Our configuration uses FCIP tunnel 0, which is represented by FC port 16. So, let us look at this port a little more closely (see Example 6-49).

You can see the **portState** line, which shows the port persistently disabled.

Note: Brocade currently ships all routers with all ports persistently disabled.

Example 6-49 Displaying a port

```
IBM_R18_SJC:admin> portshow 16
portName:
portHealth: OFFLINE

Authentication: None
portDisableReason: Persistently disabled port
portCFlags: 0x0
portFlags: 0x4021       PRESENT VIRTUAL U_PORT DISABLED LED
portType:  11.0
portState: Persistently Disabled
portPhys:  6    In_Sync
portScn:   2    Offline
port generation number:    14
portId:    021000
portIfId:    4302081b
portWwn:   20:10:00:05:1e:37:6d:ec
portWwn of device(s) connected:
       None
Distance:  normal
```

We need to persistently enable the port before we start our configuration. The first command, **portcfgpersistentenable**, without any parameters, shows the persistent status of all the ports. Port 16 does not say yes, so it is persistently disabled.

Example 6-50 Displaying port persistent setting

```
IBM_R18_SJC:admin> portcfgpersistentenable
Slot 0     0   1   2   3   4   5   6   7   8   9  10  11  12  13  14
15
---------+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+-
--
Enabled   YES YES  -   -   -   -   -   -   -   -   -   -   -   -   -
-

Slot 0    16  17  18  19  20  21  22  23  24  25  26  27  28  29  30
31
---------+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+-
--
Enabled    -   -   -   -   -   -   -   -   -   -   -   -   -   -   -
-

Slot 0    ge0 ge1
---------+---+---
Enabled   YES YES
```

We persistently enable the port and check the status.

Example 6-51 Changing the port persistent setting

```
IBM_R18_SJC:admin> portcfgpersistentenable 16

IBM_R18_SJC:admin> portcfgpersistentenable
Slot 0     0   1   2   3   4   5   6   7   8   9  10  11  12  13  14
15
---------+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+-
--
Enabled   YES YES  -   -   -   -   -   -   -   -   -   -   -   -   -
-

Slot 0    16  17  18  19  20  21  22  23  24  25  26  27  28  29  30
31
---------+---+---+---+---+---+---+---+---+---+---+---+---+---+---+---+-
--
Enabled   YES  -   -   -   -   -   -   -   -   -   -   -   -   -   -
-

Slot 0    ge0 ge1
---------+---+---
Enabled   YES YES
```

Now that we have persistent-enabled the port, we must disable it to continue our configuration. This is the normal port disable, as opposed to persistent port disable.

Example 6-52 Disabling a port

```
IBM_R18_SJC:admin> portdisable 16
```

Now, we need to add an IP address to our ge0 interface. First, we look at the interface with the **portshow ipif** command. As you can see, there is no IP configured on the GigE port ge0.

Example 6-53 Display IP for GigE interface

```
IBM_R18_SJC:admin> portshow ipif ge0

Port: ge0
Interface       IP Address      NetMask         MTU
-----------------------------------------------------------
```

To configure an IP address on the interface, we use the **portcfg ipif** command. The syntax is shown, as well as our configuration.

Example 6-54 Configuring IP on GigE interface

```
IBM_R18_SJC:admin> portcfg ipif
portcfg ipif
Usage:   portCfg ipif [Slot/]Port <Option> <Args>
Option:  create - create ip interface
         delete - delete ip interface
Args: Option specific args
       ip_address
       netmask
       mtu_size

IBM_R18_SJC:admin> portcfg ipif ge0 create 10.10.10.1 255.255.255.0
1500

IBM_R18_SJC:admin> portshow ipif ge0

Port: ge0
Interface       IP Address      NetMask         MTU
-----------------------------------------------------------
  0             10.10.10.1      255.255.255.0   1500
```

Important: There is no affiliation of the IP address we just configured to the FCIP tunnel we plan to use (fc16) yet. If you are creating more than one tunnel, you can create multiple interfaces at this time.

We now configure our FCIP tunnel. First, we look at our current configuration, if any, and the syntax of the **portshow fciptunnel** command.

Example 6-55 portshow fciptunnel command

```
IBM_R18_SJC:admin> portshow fciptunnel ge0 all

Port: ge0

IBM_R18_SJC:admin> portshow fciptunnel

Usage:
  portshow [<SlotNumber>/]<PortNumber>
  OR
  portshow [ipif|arp|iproute|fciptunnel] [<SlotNumber>/]ge<PortNumber>
[<Args>]

Args:  all|<tunnel_id> - Show All/<tunnel_id> FCIP tunnels on this GigE
port.
```

We currently do not have any FCIP tunnels configured on our ge0 interface. To configure our FCIP tunnel, we use the **portcfg fciptunnel** command. First, we run the command without parameters to discover the syntax, and then we configure our parameters. We further explain the command afterward.

Example 6-56 Configuring the FCIP tunnel

```
IBM_M48_SJC:admin> portcfg fciptunnel
Usage:   portCfg fciptunnel [Slot/]Port <Option> <Args> <Optional Args>
Option: create - create an FCIP tunnel
        delete - delete an FCIP tunnel
Args:   Option specific args
        tunnel_id
        remote_ip_addr
        local_ip_addr
        comm_rate                        - Committed Rate (kbps)

Optional Args:
       -n wwn               - remote switch wwn
       -m time              - minimum retransmit time
```

```
        -c                  - turn compression on
        -s                  - turn selective ACK off
        -k timeout          - keepalive_timeout
        -r retransmissions  - maximum retransmissions

IBM_R18_SJC:admin> portcfg fciptunnel ge0 create 0 10.10.10.2
10.10.10.1 0
```

Even though the command format is shown, we want to clarify the command, explaining the parameters we used:

portcfg fciptunnel The command.

ge0 The GigE port our tunnel will use for transport.

create Constructing the tunnel.

0 The tunnel ID. The value you use here directly relates to the port you want to use for your tunnel. We use port 16, which corresponds to tunnel ID 0 on interface ge0, as shown in Table 6-3.

Table 6-3 Port to tunnel ID relationship

Port	ge0 tunnel ID	Port	ge1 tunnel ID
Port 16	0	Port 24	0
Port 17	1	Port 25	1
Port 18	2	Port 26	2
Port 19	3	Port 27	3
Port 20	4	Port 28	4
Port 21	5	Port 29	5
Port 22	6	Port 30	6
Port 23	7	Port 31	7

10.10.10.2 The IP of the remote switch to which we will connect. We have not configured this yet, but we will.

10.10.10.1 Our local IP. (This is when the IP gets associated with the tunnel and port.)

0 The committed rate value. The value 0 allows the tunnel to use *all* the available bandwidth. If you are sharing the link with other traffic, you might want to limit the total bandwidth that this tunnel can consume. The units used are Kbps.

We now look at our configuration using the **portshow fciptunnel** command.

Example 6-57 Displaying the tunnel

```
IBM_R18_SJC:admin> portshow fciptunnel ge0 all

Port: ge0
-----------------------------------------
        Tunnel ID 0
        Remote IP Addr 10.10.10.2
        Local IP Addr 10.10.10.1
        Remote WWN Not Configured
        Local WWN 10:00:00:05:1e:37:6d:ec
        Compression off
        Uncommitted bandwidth, minimum of 1000 Kbps (0.001000 Gbps)
        SACK on
        Min Retransmit Time 100
        Keepalive Timeout 10
        Max Retransmissions 8
        Status : Inactive
```

Tip: We highly recommend that you implement the **-n** argument when configuring your FCIP tunnel. You can see the syntax in Example 6-56 on page 159. This option provides for a more secure connection across your WAN cloud by specifically allowing only the router with the configured WWN to establish a link. Configure this argument at both ends of your tunnel.

You can see from the output that our status is inactive. This should be obvious, because we do not have the other end of our link configured. One last thing (before we leave this switch) is to enable the port (16) so that it is online.

Example 6-58 Enabling a port

```
IBM_R18_SJC:admin> portenable 16
```

We turn our focus to the backbone side of our FCIP tunnel, the 2109-M48, specifically on the first blade in the chassis, which is an FC Routing Blade. We connect to the GigE interface 1/ge0 (slot/port). We again use FCIP tunnel 0.

Because we do not want our fabrics to merge, we need to do an extra step at this end of our link. We need to configure the port to a VEX_Port. If you wanted the fabrics to merge across the FCIP link (usually not advisable but is possible if the available bandwidth is large enough and the link is stable), skip the next step. This way, when you enable the port, the fabrics merge, provided all the merge criteria are met. We do not want our fabrics to merge, so we make the port a VEX_Port.

Open a Telnet session and connect to the switch. Our M48 is named IBM_M48_SJC.

Example 6-59 Switch login

```
login as: admin
admin@10.64.210.116's password:
IBM_M48_SJC:admin>
```

Our M48 is a fully configured chassis with blades in every slot. All the output has been concatenated as needed, so we can focus on our configuration.

We look at our switch using the **switchshow** command.

Example 6-60 switchshow command

```
IBM_M48_SJC:admin> switchshow
switchName:     IBM_M48_SJC
switchType:     42.2
switchState:    Online
switchMode:     Native
switchRole:     Principal
switchDomain:   69
switchId:       fffc45
switchWwn:      10:00:00:60:69:e4:27:52
zoning:         ON (IBM_SFO_CONFIG)
switchBeacon:   OFF
blade1 Beacon:  OFF
blade2 Beacon:  OFF
blade3 Beacon:  OFF
blade4 Beacon:  OFF
blade7 Beacon:  OFF
blade8 Beacon:  OFF
blade9 Beacon:  OFF
blade10 Beacon: OFF
FC Router:      ON
FC Router BB Fabric ID: 1
```

```
Area Slot Port Media Speed State
====================================
  0    1    0   --    N4   No_Module
  1    1    1   --    N4   No_Module
  2    1    2   --    N4   No_Module
  3    1    3   --    N4   No_Module
  4    1    4   --    N4   No_Module
  5    1    5   --    N4   No_Module
  6    1    6   --    N4   No_Module
  7    1    7   --    N4   No_Module
  8    1    8   --    N4   No_Module
  9    1    9   --    N4   No_Module
 10    1   10   --    N4   No_Module
 11    1   11   --    N4   No_Module
 12    1   12   --    N4   No_Module
 13    1   13   --    N4   No_Module Disabled (Persistent)
 14    1   14   --    N4   No_Module Disabled (Persistent)
 15    1   15   --    N4   No_Module Disabled (Persistent)
128    1   16   --    --   Offline   Disabled
129    1   17   --    --   Offline   Disabled (Persistent)
130    1   18   --    --   Offline   Disabled (Persistent)
131    1   19   --    --   Offline   Disabled (Persistent)
132    1   20   --    --   Offline   Disabled (Persistent)
133    1   21   --    --   Offline   Disabled (Persistent)
134    1   22   --    --   Offline   Disabled (Persistent)
135    1   23   --    --   Offline   Disabled (Persistent)
136    1   24   --    --   Offline   Disabled (Persistent)
137    1   25   --    --   Offline   Disabled (Persistent)
138    1   26   --    --   Offline   Disabled (Persistent)
139    1   27   --    --   Offline   Disabled (Persistent)
140    1   28   --    --   Offline   Disabled (Persistent)
141    1   29   --    --   Offline   Disabled (Persistent)
142    1   30   --    --   Offline   Disabled (Persistent)
143    1   31   --    --   Offline   Disabled (Persistent)
       1  ge0   id    1G   Online
       1  ge1   id    1G   No_Sync
```

We use port 16 (1/16), which is FCIP tunnel 0 on port ge0.

Note: The blade is logically equivalent to the 2005-R18, and behaves in the same manner, so Table 6-3 on page 160 is valid here as well.

Port 16 has already been persistently enabled and shows as disabled, which means that we are ready to start our configuration.

We start by looking at port 16 using the **portshow** command.

Example 6-61 Display a port

```
IBM_M48_SJC:admin> portshow 1/16
portName:
portHealth: OFFLINE

Authentication: None
portDisableReason: None
portCFlags: 0x0
portFlags: 0x4021        PRESENT VIRTUAL U_PORT DISABLED LED
portType:  11.0
portState: 2    Offline
portPhys:  6    In_Sync
portScn:   2    Offline
port generation number:    12
portId:    458000
portIfId:    43120029
portWwn:   20:80:00:60:69:e4:27:52
portWwn of device(s) connected:
        None
Distance:  normal
```

The current port setting shows that the port is configured as a normal FC port and can assume any port type as needed. We want to configure the port to be a VEX_Port. We do this with the **portcfgvexport** command. We run the command without parameters to see the syntax, and then configure the port.

Important: Refer to Table 2-1 on page 29 if you need clarification of the different port types and router terminology.

Example 6-62 EX port configuration

```
IBM_M48_SJC:admin> portcfgvexport

Usage: portcfgvexport [SlotNumber/]PortNumber
                [-a 1-enable 2-disable] [-f fid(1..128)]
                [-r r_a_tov] [-e e_d_tov] [-d domain]
                [-p 0-native 1-core 2-extended edge]
              [-t 1-Enable 2-Disable]

IBM_M48_SJC:admin> portcfgvexport 1/16 -a 1 -f 128 -d 128
IBM_M48_SJC:admin> portshow 1/16
portName:
```

```
portHealth: OFFLINE

Authentication: None

EX_Port Mode:   Enabled
Fabric ID:      128
Front Phantom:  state = Not OK  Pref Dom ID: 128
Fabric params:  R_A_TOV: 0      E_D_TOV: 0      PID fmt: auto

Authentication Type: None
Hash Algorithm: N/A
DH Group: N/A
Edge fabric's primary wwn: N/A
Edge fabric's version stamp: N/A

portDisableReason: None
portCFlags: 0x0
portFlags: 0x4021       PRESENT VIRTUAL U_PORT EX_PORT DISABLED LED
portType:  11.0
portState: 2    Offline
portPhys:  6    In_Sync
portScn:   2    Offline
port generation number:    12
portId:    458000
portIfId:    43120029
portWwn:   20:80:00:60:69:e4:27:52
portWwn of device(s) connected:
    None
Distance:  normal
```

From the **portshow** command, we can see that the port is now configured to be an EX_Port, which will become a VEX_Port because port 1/16 is actually an FCIP tunnel (a Virtual E_Port that is a routing port).

We look at the command **portcfgvexport 1/16 -a 1 -f 128 -d 128** a little more closely:

portcfgvexport	The command.
1/16	The port we want to configure.
-a 1	This enables the port to be an EX_Port (VEX in our case, because the port we are configuring is an FCIP tunnel, not an FC port).
-f 128	This is what we configured our fabric ID to. This value can be any value between 1 and 128.

-d 128 This is the domain ID of our backbone fabric connection, as seen from the outside world. This must be unique to your fabric, and follows the same rules as any other domain.

The rest of the configuration is basically the same as the edge fabric. Follow Example 6-63 through the configuration.

Example 6-63 Configuring IP and FCIP

```
IBM_M48_SJC:admin> portshow ipif 1/ge0

Slot: 1 Port: ge0
Interface       IP Address      NetMask         MTU
-----------------------------------------------------------


M_M48_SJC:admin> portcfg ipif
portcfg ipif
Usage:   portCfg ipif [Slot/]Port <Option> <Args>
Option:  create - create ip interface
         delete - delete ip interface
Args: Option specific args
       ip_address
       netmask
       mtu_size

IBM_M48_SJC:admin> portcfg ipif 1/ge0 create 10.10.10.2 255.255.255.0
1500
IBM_M48_SJC:admin> portshow ipif 1/ge0

Slot: 1 Port: ge0
Interface       IP Address      NetMask         MTU
-----------------------------------------------------------
  0             10.10.10.2      255.255.255.0   1500

IBM_M48_SJC:admin> portshow fciptunnel 1/ge0 all

Slot: 1 Port: ge0

IBM_M48_SJC:admin> portcfg fciptunnel 1/ge0 create 0 10.10.10.1
10.10.10.2 0

IBM_M48_SJC:admin> portshow fciptunnel 1/ge0 all

Slot: 1 Port: ge0
```

```
----------------------------------------
        Tunnel ID 0
        Remote IP Addr 10.10.10.1
        Local IP Addr 10.10.10.2
        Remote WWN Not Configured
        Local WWN 10:00:00:60:69:e4:27:52
        Compression off
        Uncommitted bandwidth, minimum of 1000 Kbps (0.001000 Gbps)
        SACK on
        Min Retransmit Time 100
        Keepalive Timeout 10
        Max Retransmissions 8
        Status : Inactive
```

Our configuration looks good so far, but our status still shows inactive because we still need to enable the port. We enable the port, and then wait one minute and check the status of our tunnel.

Example 6-64 Enable a port

```
IBM_M48_SJC:admin> portenable 1/16

IBM_M48_SJC:admin> portshow fciptunnel 1/ge0 all
Slot: 1 Port: ge0
----------------------------------------
        Tunnel ID 0
        Remote IP Addr 10.10.10.1
        Local IP Addr 10.10.10.2
        Remote WWN Not Configured
        Local WWN 10:00:00:60:69:e4:27:52
        Compression off
        Uncommitted bandwidth, minimum of 1000 Kbps (0.001000 Gbps)
        SACK on
        Min Retransmit Time 100
        Keepalive Timeout 10
        Max Retransmissions 8
        Status : Active
```

This is the active status we were hoping for.

Tip: Depending on your physical WAN link, you might need to wait longer than one minute for the link to initialize.

We use the **switchshow** command and see what type of port has been configured; it should be a VEX_Port. If it is not, our fabrics will have merged, or at least tried to. Example 6-65 shows the port status.

Example 6-65 Port display

```
 11    1   11   --    N4   No_Module
 12    1   12   --    N4   No_Module
 13    1   13   --    N4   No_Module Disabled (Persistent)
 14    1   14   --    N4   No_Module Disabled (Persistent)
 15    1   15   --    N4   No_Module Disabled (Persistent)
128    1   16   --    --   Online    VEX-Port 10:00:00:05:1e:37:6d:ec
"IBM_R18_SJC" (fabric id = 128 )
```

We can see the port is a VEX_Port, and the WWN of the switch at the far end of the link is displayed, along with the switch's name.

The **portcfgshow** command gives a summary of the port configuration. We also could have used the **portshow** command to get more information.

Example 6-66 Display port configuration

```
IBM_M48_SJC:admin> portcfgshow 1/16
Area Number:              128
Speed Level:              AUTO
Trunk Port                ON
Long Distance             OFF
VC Link Init              OFF
Locked L_Port             OFF
Locked G_Port             OFF
Disabled E_Port           OFF
ISL R_RDY Mode            OFF
RSCN Suppressed           OFF
Persistent Disable        OFF
NPIV capability           ON
EX Port                   ON
```

We log back in to our edge switch and check the status of our tunnel from that side and the type of port that has been configured.

Example 6-67 Display tunnel status

```
IBM_R18_SJC:admin> portshow fciptunnel ge0 all

Port: ge0
-----------------------------------------
        Tunnel ID 0
        Remote IP Addr 10.10.10.2
        Local IP Addr 10.10.10.1
        Remote WWN Not Configured
        Local WWN 10:00:00:05:1e:37:6d:ec
        Compression off
        Uncommitted bandwidth, minimum of 1000 Kbps (0.001000 Gbps)
        SACK on
        Min Retransmit Time 100
        Keepalive Timeout 10
        Max Retransmissions 8
        Status : Active


  9   9   --    N4   No_Module Disabled (Persistent)
 10  10   --    N4   No_Module Disabled (Persistent)
 11  11   id    N4   No_Sync   Disabled (Persistent)
 12  12   --    N4   No_Module Disabled (Persistent)
 13  13   --    N4   No_Module Disabled (Persistent)
 14  14   --    N4   No_Module Disabled (Persistent)
 15  15   --    N4   No_Module Disabled (Persistent)
 16  16   --    --   Online    VE-Port  50:06:06:9e:42:75:2e:80
"fcr_fd_128_128" (downstream)
 17  17   --    --   Offline   Disabled (Persistent)
 18  18   --    --   Offline   Disabled (Persistent)
 19  19   --    --   Offline   Disabled (Persistent)
```

This side of the tunnel also shows active. In addition, we can see that the port at this end is a VE_Port, which is what we wanted and expected.

This is the end of the FCIP tunnel configuration.

LSAN setup with CLI

We established an inter-fabric link (IFL) between the SAN18B-R and the M48 FC Routing Blade through our FCIP link. We now want to use this link to configure disk access to our server. Our two fabrics have not merged, a direct result of our

VEX_Port configuration. Using LSAN zoning, we can share devices across our link. This is called FC routing.

We have our Windows host connected to the IBM_R18_SJC switch. Our Disk is attached to the IBM_M48_SJC switch.

We start at our edge switch and run the **nsshow** command to find the WWPN of our host's HBA.

Example 6-68 Display name server on R18

```
IBM_R18_SJC:admin> nsshow
{
 Type Pid    COS     PortName                NodeName
TTL(sec)
 N    040f00;    2,3;10:00:00:00:c9:30:71:93;20:00:00:00:c9:30:71:93;
na
    FC4s: FCP
    NodeSymb: [36] "Emulex LP9002 FV3.90A7  DV5-5.10A10 "
    Fabric Port Name: 20:0f:00:05:1e:37:6d:ec
    Permanent Port Name: 10:00:00:00:c9:30:71:93
The Local Name Server has 1 entry }
```

The name server shows us the WWPN (in bold) we need to enter into our LSAN zone.

On the M48, which has our disks attached, we run the **nsshow** command as well, to acquire the WWPNs of our storage.

Example 6-69 Display name server on M48

```
IBM_M48_SJC:admin> nsshow
{
 Type Pid    COS     PortName                NodeName
TTL(sec)
 N    457300;      3;21:01:00:00:d1:26:73:1d;20:00:00:00:d1:26:73:1d;
na
    FC4s: FCP [IBM     DS400    S320   5.87]
    Fabric Port Name: 20:73:00:60:69:e4:27:52
    Permanent Port Name: 21:01:00:00:d1:26:73:1d
 NL   45cfc9;      3;21:00:00:04:cf:92:74:99;20:00:00:04:cf:92:74:99;
na
    FC4s: FCP [SEAGATE ST318452FC      0004]
    Fabric Port Name: 20:cf:00:60:69:e4:27:52
    Permanent Port Name: 21:00:00:04:cf:92:74:99
```

```
 NL   45cfca;      3;21:00:00:04:cf:bd:58:1c;20:00:00:04:cf:bd:58:1c;
na
    FC4s: FCP [SEAGATE ST318452FC      0005]
    Fabric Port Name: 20:cf:00:60:69:e4:27:52
    Permanent Port Name: 21:00:00:04:cf:bd:58:1c
 NL   45cfcb;      3;21:00:00:04:cf:bd:56:3e;20:00:00:04:cf:bd:56:3e;
na
    FC4s: FCP [SEAGATE ST318452FC      0005]
    Fabric Port Name: 20:cf:00:60:69:e4:27:52
    Permanent Port Name: 21:00:00:04:cf:bd:56:3e
 NL   45cfcc;      3;21:00:00:04:cf:bd:56:a1;20:00:00:04:cf:bd:56:a1;
na
    FC4s: FCP [SEAGATE ST318452FC      0005]
    Fabric Port Name: 20:cf:00:60:69:e4:27:52
    Permanent Port Name: 21:00:00:04:cf:bd:56:a1
 NL   45dfe4;      3;21:00:00:04:cf:bd:71:a0;20:00:00:04:cf:bd:71:a0;
na
    FC4s: FCP [SEAGATE ST318452FC      0005]
    Fabric Port Name: 20:df:00:60:69:e4:27:52
    Permanent Port Name: 21:00:00:04:cf:bd:71:a0
 NL   45dfe8;      3;21:00:00:04:cf:bd:72:a1;20:00:00:04:cf:bd:72:a1;
na
    FC4s: FCP [SEAGATE ST318452FC      0005]
    Fabric Port Name: 20:df:00:60:69:e4:27:52
    Permanent Port Name: 21:00:00:04:cf:bd:72:a1
 NL   45dfef;      3;21:00:00:04:cf:bd:71:a8;20:00:00:04:cf:bd:71:a8;
na
    FC4s: FCP [SEAGATE ST318452FC      0005]
    Fabric Port Name: 20:df:00:60:69:e4:27:52
    Permanent Port Name: 21:00:00:04:cf:bd:71:a8
The Local Name Server has 8 entries }
```

We are interested in the two bolded WWPNs at the top of the list. The first belongs to our DS400 disk array, which has our disk already carved and masked to our Windows server. The second entry is from a JBOD array and is a physical disk accessed through an FC controller.

To accomplish disk-to-server access, we create LSAN zones for each SAN that contains the devices to be exported. The contents of these zones must be the same for each SAN or the devices will not be exported; however, it is not necessary for the zone names to be identical.

With the b-type SAN router, it is important to remember that for any devices we want the routers to have knowledge of, they must be put into a zone that starts with the five characters LSAN_. Note the underscore after LSAN. “LSAN” can be

in uppercase or lowercase letters, for example, LSAN_zone_SAN-AtoB or lsan-zone_sanBtoA. “LSAN” is the tag that the router uses to acknowledge which devices it is allowed to export to the other SAN.

We name the zones lsan_zone_SJC on the R18 and lsan_zone_Guelph on the M48. Each zone has the exact same members inside them. We must remember to save our configuration. We start on the backbone fabric.

Example 6-70 Create and save the zone

```
IBM_M48_SJC:admin> zonecreate "lsan_zone_Guelph",
"21:01:00:00:d1:26:73:1d"

IBM_M48_SJC:admin> zoneadd "lsan_zone_Guelph",
"21:00:00:04:cf:92:74:99;10:00:00:00:c9:30:71:93"

IBM_M48_SJC:admin> zoneshow

zone:  lsan_zone_Guelph
               21:01:00:00:d1:26:73:1d; 21:00:00:04:cf:92:74:99;
               10:00:00:00:c9:30:71:93

IBM_M48_SJC:admin> cfgsave
You are about to save the Defined zoning configuration. This
action will only save the changes on Defined configuration.
Any changes made on the Effective configuration will not
take effect until it is re-enabled.
Do you want to save Defined zoning configuration only?  (yes, y, no,
n): [no] y
Updating flash ...
```

Now, we add this zone to the active configuration. First, we determine the name of the currently active configuration. We then add our new zone to this configuration and activate it nondisruptively.

Note: Some of the following output has been concatenated.

Example 6-71 Displaying the configuration

```
IBM_M48_SJC:admin> cfgshow
Defined configuration:
....

Effective configuration:
 cfg:   IBM_SFO_CONFIG
```

```
 zone:  Blade_1_DS400_ZONE
                20:00:00:09:6b:36:01:10
                20:00:00:00:d1:26:73:1d
 zone:  Blade_2_DS400_ZONE
                20:00:00:09:6b:36:40:14
                20:00:00:00:d1:26:73:1d
 zone:  IBM_BC_Zone_SFO
                132,15
                128,14
 zone:  IBM_xSeries_Zone_SFO
                128,15
                129,7
 zone:  LSAN_Shared_SJC_SFO_zone
                22:00:00:04:cf:3b:1c:a8
                22:00:00:20:37:15:0b:9a
                10:00:00:00:c9:2b:7f:90
```

From the output, we see that the effective configuration's name is IBM_SFO_CONFIG. We add our zone to this configuration and verify it. Notice that our zone is part of the "Defined configuration," but is not part of the "Effective configuration" yet.

Example 6-72 Add zone to the active configuration

```
IBM_M48_SJC:admin> cfgadd "IBM_SFO_CONFIG", "lsan_zone_Guelph"

IBM_M48_SJC:admin> cfgshow
Defined configuration:
 cfg:   IBM_SFO_CONFIG
                IBM_BC_Zone_SFO; IBM_xSeries_Zone_SFO;
                LSAN_Shared_SJC_SFO_zone; Blade_1_DS400_ZONE;
                Blade_2_DS400_ZONE; lsan_zone_Guelph
....

Effective configuration:
 cfg:   IBM_SFO_CONFIG
 zone:  Blade_1_DS400_ZONE
                20:00:00:09:6b:36:01:10
                20:00:00:00:d1:26:73:1d
 zone:  Blade_2_DS400_ZONE
                20:00:00:09:6b:36:40:14
                20:00:00:00:d1:26:73:1d
 zone:  IBM_BC_Zone_SFO
                132,15
                128,14
```

```
 zone:  IBM_xSeries_Zone_SFO
                128,15
                129,7
 zone:  LSAN_Shared_SJC_SFO_zone
                22:00:00:04:cf:3b:1c:a8
                22:00:00:20:37:15:0b:9a
                10:00:00:00:c9:2b:7f:90
```

We activate our configuration so that our zone can become part of the effective configuration.

Example 6-73 Enable the configuration

```
IBM_M48_SJC:admin> cfgenable "IBM_SFO_CONFIG"
You are about to enable a new zoning configuration.
This action will replace the old zoning configuration with the
current configuration selected.
Do you want to enable 'IBM_SFO_CONFIG' configuration  (yes, y, no, n):
[no] y
zone config "IBM_SFO_CONFIG" is in effect
Updating flash ...

IBM_M48_SJC:admin> cfgshow

....

Effective configuration:
 cfg:   IBM_SFO_CONFIG
 zone:  Blade_1_DS400_ZONE
                20:00:00:09:6b:36:01:10
                20:00:00:00:d1:26:73:1d
 zone:  Blade_2_DS400_ZONE
                20:00:00:09:6b:36:40:14
                20:00:00:00:d1:26:73:1d
 zone:  IBM_BC_Zone_SFO
                132,15
                128,14
 zone:  IBM_xSeries_Zone_SFO
                128,15
                129,7
 zone:  LSAN_Shared_SJC_SFO_zone
                22:00:00:04:cf:3b:1c:a8
                22:00:00:20:37:15:0b:9a
                10:00:00:00:c9:2b:7f:90
 zone:  lsan_zone_Guelph
```

```
21:01:00:00:d1:26:73:1d
21:00:00:04:cf:92:74:99
10:00:00:00:c9:30:71:93
```

Now our LSAN zone is part of our effective configuration.

Our backbone switch configuration is complete.

Before moving to our edge fabric, we look at our Windows server and launch the Disk Manager so that we have something with which to compare our results.

Figure 6-75 shows the disk configuration as seen in Disk Manager on our Windows Server® 2003 host.

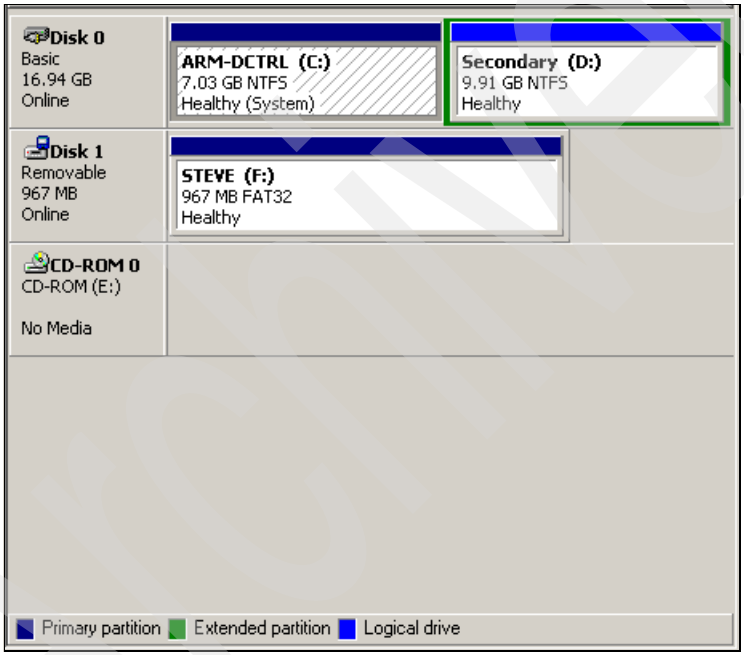


Figure 6-75 Windows disk manager 1

Now, we configure the zone on the other side of our link.

This time, we create the zone and add all three members at the same time.

Example 6-74 Create the zone

```
IBM_R18_SJC:admin> zonecreate "lsan_zone_SJC",
"21:01:00:00:d1:26:73:1d;21:00:00:04:cf:92:74:99;10:00:00:00:c9:30:71:9
3"

IBM_R18_SJC:admin> zoneshow
Defined configuration:
....

zone:  lsan_zone_SJC
                21:01:00:00:d1:26:73:1d; 21:00:00:04:cf:92:74:99;
                10:00:00:00:c9:30:71:93
....
```

We need to determine our currently active configuration, add our new zone, and enable it.

Example 6-75 Display the configuration

```
IBM_R18_SJC:admin> cfgshow

....
Defined configuration:
....
Effective configuration:
 cfg:   IBM_SFO_CONFIG
 zone:  IBM_BC_Zone_SFO
                132,15
                128,14
 zone:  IBM_xSeries_Zone_SFO
                128,15
                129,7
 zone:  LSAN_Shared_SJC_SFO_zone
                22:00:00:04:cf:3b:1c:a8
                22:00:00:20:37:15:0b:9a
                10:00:00:00:c9:2b:7f:90

IBM_R18_SJC:admin> cfgadd "IBM_SFO_CONFIG", "lsan_zone_SJC"


IBM_R18_SJC:admin> cfgshow
....
Defined configuration:
```

```
....
 cfg:   IBM_SFO_CONFIG
                IBM_BC_Zone_SFO; IBM_xSeries_Zone_SFO;
                LSAN_Shared_SJC_SFO_zone; lsan_zone_SJC
```

We must remember to save our configuration changes.

Example 6-76 Save the configuration

```
IBM_R18_SJC:admin> cfgsave
You are about to save the Defined zoning configuration. This
action will only save the changes on Defined configuration.
Any changes made on the Effective configuration will not
take effect until it is re-enabled.
Do you want to save Defined zoning configuration only?  (yes, y, no,
n): [no] y
Updating flash ...
```

Our zone is now part of the defined configuration. We enable it so that it becomes part of our effective configuration.

Example 6-77 Enable the configuration

```
IBM_R18_SJC:admin> cfgenable  IBM_SFO_CONFIG
You are about to enable a new zoning configuration.

This action will replace the old zoning configuration with the
current configuration selected.
Do you want to enable 'IBM_SFO_CONFIG' configuration  (yes, y, no, n):
[no] y
zone config "IBM_SFO_CONFIG" is in effect
Updating flash ...
```

We check our effective configuration.

Example 6-78 Display the configuration

```
IBM_R18_SJC:admin> cfgshow

....
Defined configuration:
....

Effective configuration:
 cfg:   IBM_SFO_CONFIG
 zone:  IBM_BC_Zone_SFO
```

```
                132,15
                128,14
zone:  IBM_xSeries_Zone_SFO
                128,15
                129,7
zone:  LSAN_Shared_SJC_SFO_zone
                22:00:00:04:cf:3b:1c:a8
                22:00:00:20:37:15:0b:9a
                10:00:00:00:c9:2b:7f:90
zone:  lsan_zone_SJC
                21:01:00:00:d1:26:73:1d
                21:00:00:04:cf:92:74:99
                10:00:00:00:c9:30:71:93
```

Our new LSAN zone is now active.

We go to our Windows server and see if any new disks are discovered when we rescan the HBAs, as shown in Figure 6-76 and Figure 6-77 on page 179.

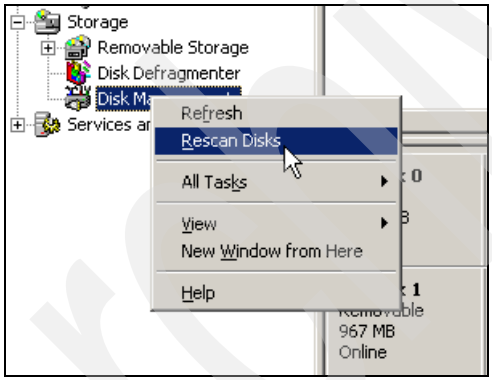


Figure 6-76 Rescan Disks

As you can see, our new disk has been discovered by our host.

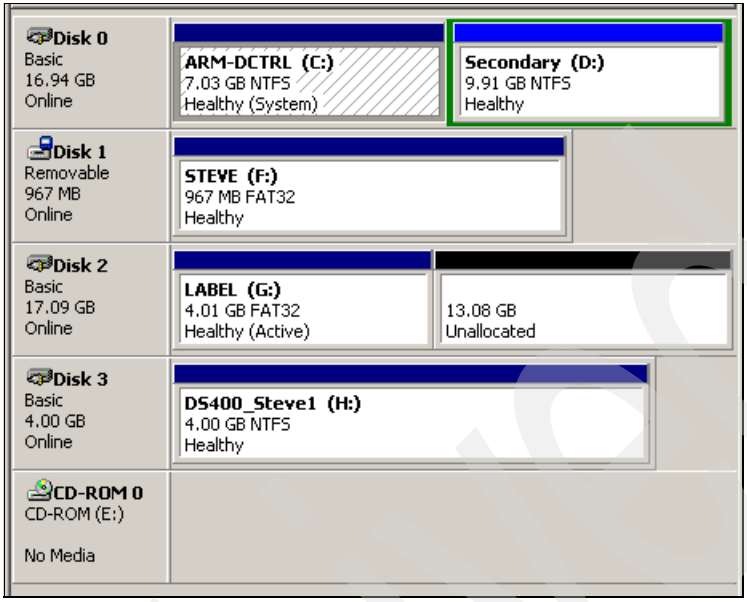


Figure 6-77 New disk

LSAN verification

There are some commands we can use to verify our configuration. Let us explore a few of them here.

The first command we describe is **`lsanZoneShow -s`**. We look at it on the M48 switch first, the VEX side of our link.

Note: We show the commands as they appear in the Brocade documentation. We found that *most* commands can be entered in all lowercase. If the command does not work, try to enter it using uppercase letters.

Example 6-79 lsanZoneShow command on M48

```
IBM_M48_SJC:admin> lsanzoneshow -s
Fabric ID: 1 Zone Name: LSAN_Shared_SJC_SFO_zone
        22:00:00:04:cf:3b:1c:a8  ABSENT
        22:00:00:20:37:15:0b:9a  ABSENT
        10:00:00:00:c9:2b:7f:90  ABSENT
Fabric ID: 1 Zone Name: lsan_zone_Guelph
        21:01:00:00:d1:26:73:1d  EXIST
        21:00:00:04:cf:92:74:99  EXIST
```

```
        10:00:00:00:c9:30:71:93  Imported
Fabric ID: 128 Zone Name: LSAN_Shared_SJC_SFO_zone
        22:00:00:04:cf:3b:1c:a8  ABSENT
        22:00:00:20:37:15:0b:9a  ABSENT
        10:00:00:00:c9:2b:7f:90  ABSENT
Fabric ID: 128 Zone Name: lsan_zone_SJC
        21:01:00:00:d1:26:73:1d  Imported
        21:00:00:04:cf:92:74:99  Imported
        10:00:00:00:c9:30:71:93  EXIST
```

As you can see, the VEX side of our FCIP tunnel (backbone), which is the routing side, knows the status of the devices and where they are physically located. Both of our LSAN zones are displayed and show the state of the devices as seen by their respective switches. Lsan_zone_Guelph shows that the two disk WWPNs are local and the server HBA WWPN is imported (from the other side of our tunnel).

Lsan_zone_SJC shows the opposite, both disk ports are imported and the server WWPN is local.

Running the same command on the edge fabric produces a slightly different output.

Example 6-80 lsanZoneShow command on R18

```
IBM_R18_SJC:admin> lsanzoneshow -s
Fabric ID: 1 Zone Name: LSAN_Shared_SJC_SFO_zone
        22:00:00:04:cf:3b:1c:a8  ABSENT
        22:00:00:20:37:15:0b:9a  ABSENT
        10:00:00:00:c9:2b:7f:90  ABSENT
Fabric ID: 1 Zone Name: lsan_zone_SJC
        21:01:00:00:d1:26:73:1d  EXIST
        21:00:00:04:cf:92:74:99  EXIST
        10:00:00:00:c9:30:71:93  EXIST
Fabric ID: 20 Zone Name: LSAN_DS400_Blade1
        21:00:00:09:6b:36:01:11  EXIST
        21:01:00:00:d1:26:73:1d  ABSENT
Fabric ID: 20 Zone Name: LSAN_DS400_Blade2
        21:00:00:09:6b:36:40:15  EXIST
        21:01:00:00:d1:26:73:1d  ABSENT
```

We can only see the local zone, and it states that all the devices simply exist. This is the normal behavior for an E_port, or VE_port in our case. This side does not know that we are routing and it believes the fabrics are merged.

The next command is **fcrPhyDevShow**. This command displays the physical device WWN, PID, and FID.

Example 6-81 displays the output from running this command on both switches.

Example 6-81 fcrPhyDevShow command

```
IBM_M48_SJC:admin> fcrphydevshow
 Device           WWN             Physical
 Exists                             PID
in Fabric
----------------------------------------
    1    21:00:00:04:cf:92:74:99  45cfc9
    1    21:01:00:00:d1:26:73:1d  457300
  128    10:00:00:00:c9:30:71:93  040f00

IBM_R18_SJC:admin> fcrphydevshow
 Device           WWN             Physical
 Exists                             PID
in Fabric
----------------------------------------
    1    10:00:00:00:c9:30:71:93  040f00
    1    21:00:00:04:cf:92:74:99  01f001
    1    21:01:00:00:d1:26:73:1d  01f003
```

The last command is **fcrProxyDevShow**. We show the output from both switches.

Example 6-82 fcrProxyDevShow command

```
IBM_M48_SJC:admin> fcrproxydevshow
  Proxy           WWN              Proxy      Device   Physical    State
 Created                            PID       Exists     PID
in Fabric                                    in Fabric
--------------------------------------------------------------------------
-----
     1   10:00:00:00:c9:30:71:93  02f002      128      040f00
Imported
   128   21:00:00:04:cf:92:74:99  01f001        1      45cfc9
Imported
   128   21:01:00:00:d1:26:73:1d  01f003        1      457300
Imported


IBM_R18_SJC:admin> fcrproxydevshow
No proxy device found
```

FCIP configuration in GUI

In our example, we establish an FCIP tunnel between a SAN18B-R router and a SAN Routing Blade. We use GigE port ge1 on both routers. SAN18B-R is attached to a BladeCenter Fibre Switch Module because we use two Blade servers as our hosts. These two Blade servers require access to the DS400 Storage Server logical drives and the DS400 is attached to the M48 Director. The M48 contains two FC Routing Blades; in our example, we use the Routing Blade in slot 10. Figure 6-74 on page 153 shows all of this.

Connection between the SAN18B-R and BladeCenter FSM

The BladeCenter Fibre Switch Module (FSM) is connected to port 12 on our SAN18B-R. Port 12 must, therefore, be configured as an EX_Port.

Perform the following steps:

1. The first step is to persistent enable port 12, as shown in Figure 6-78.

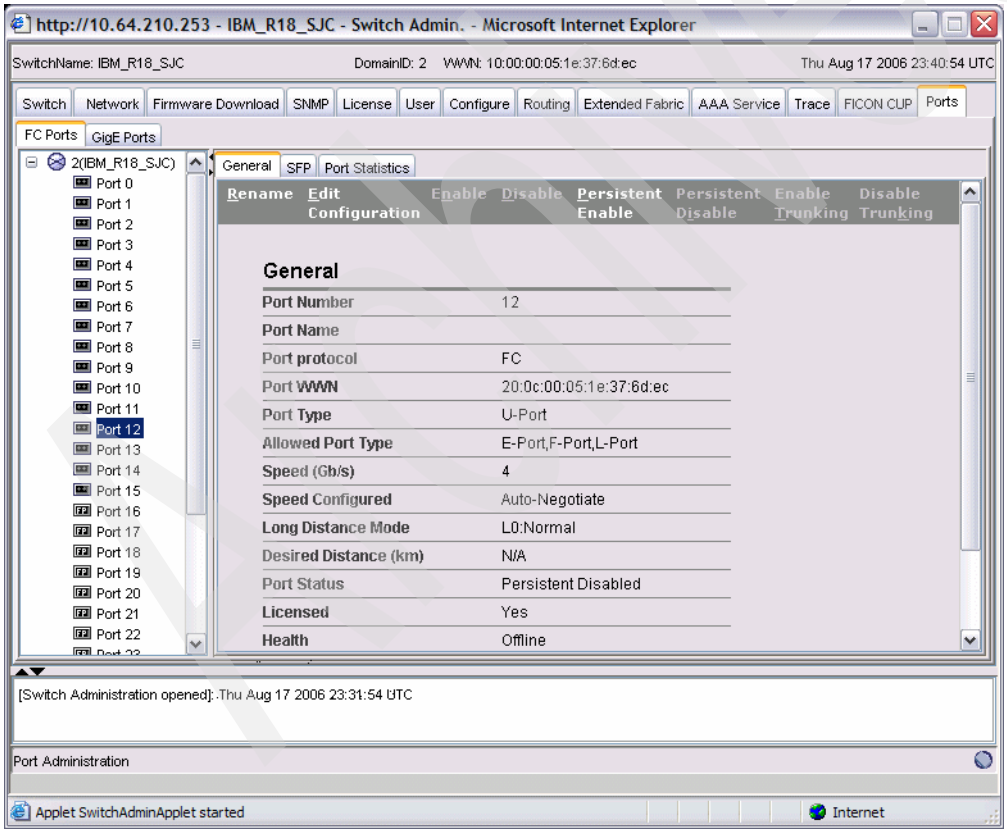


Figure 6-78 Persistent Enable Port 12

2. We now need to disable the port, because its configuration will change (Figure 6-79).

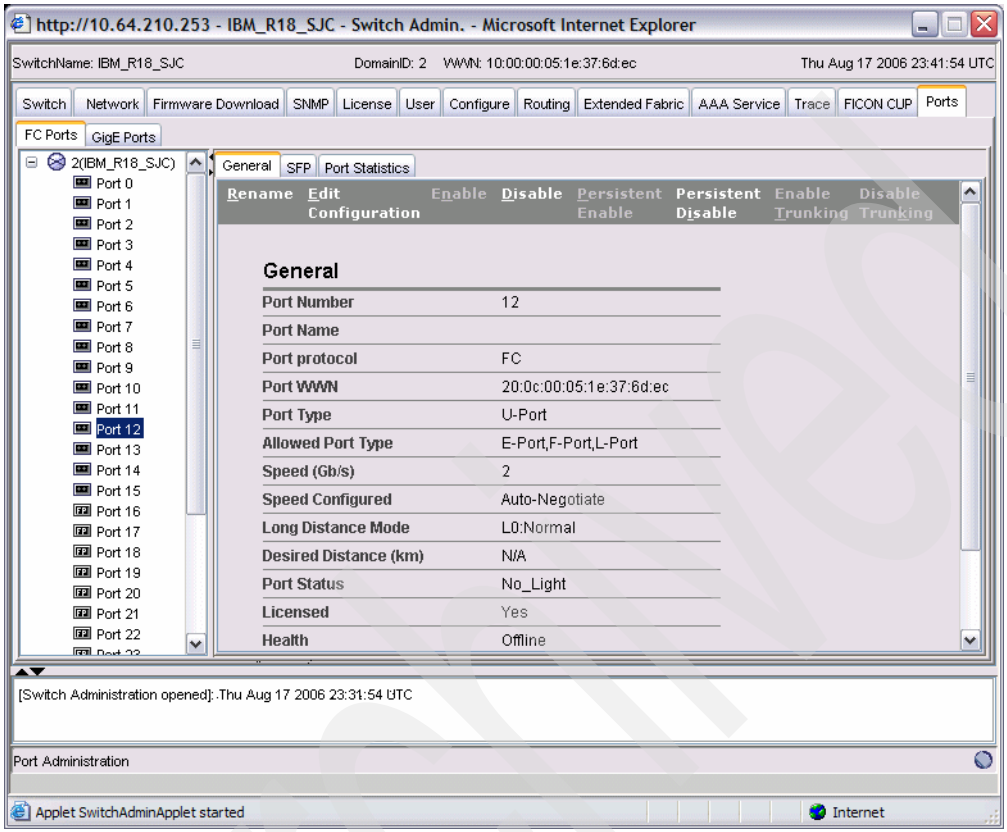


Figure 6-79 Disabling Port 12

3. Next, we set the port type to EX_Port. Click **Edit Configuration**, as shown in Figure 6-80. The FC Port Configuration wizard opens.

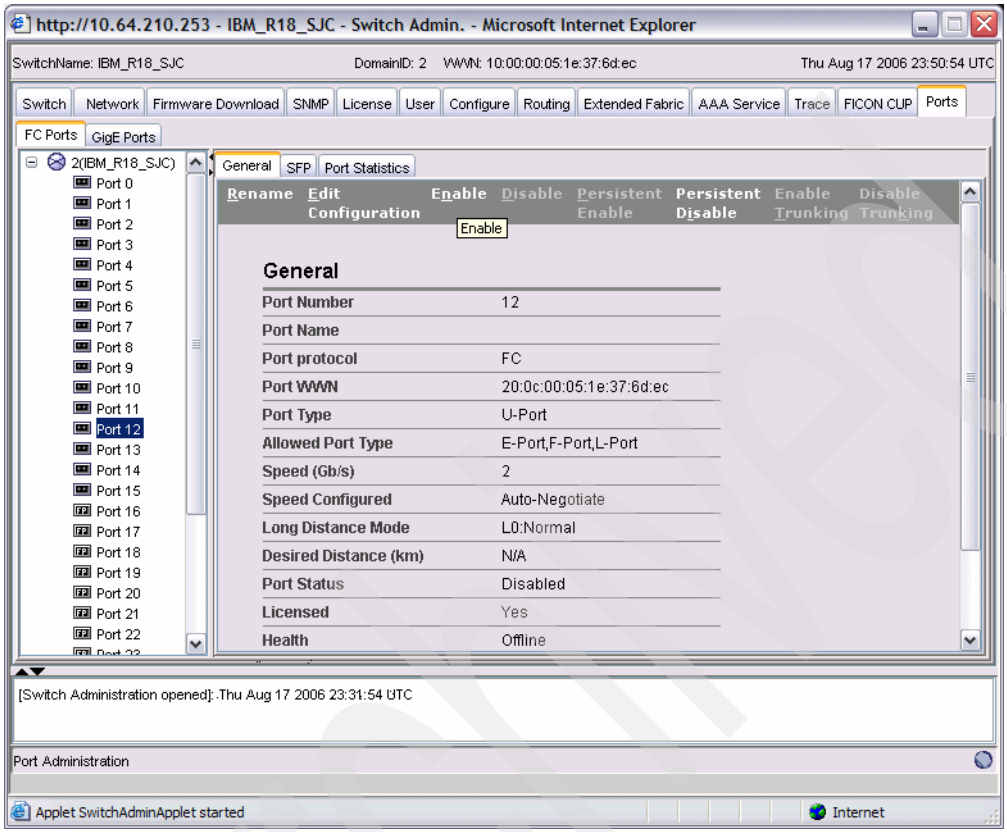


Figure 6-80 Edit Configuration Port 12

4. The EX-port type requires you to set the Fabric ID as well (see Figure 6-81).

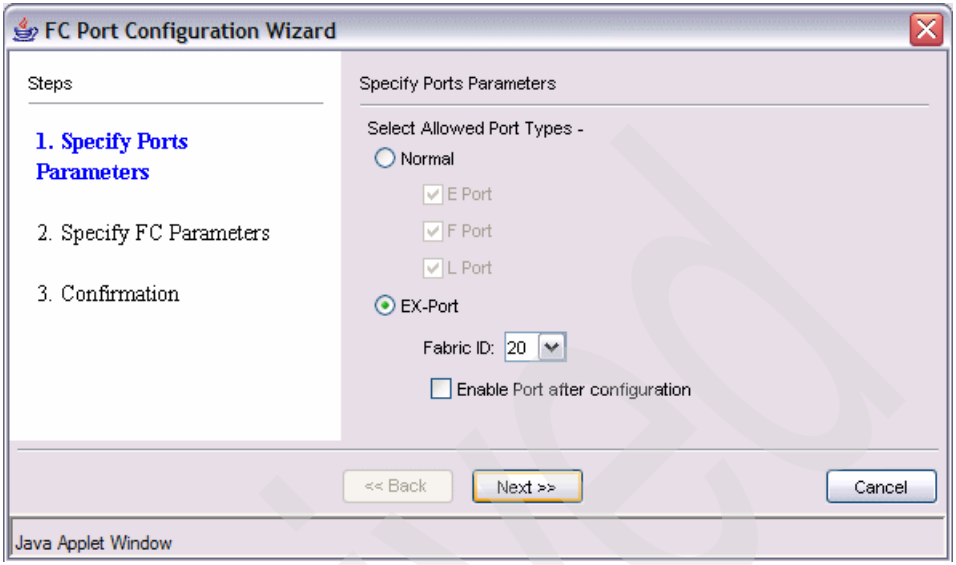


Figure 6-81 Setting the EX-port type

5. We define the FC parameters, Speed and Long Distance Mode, as shown in Figure 6-82.

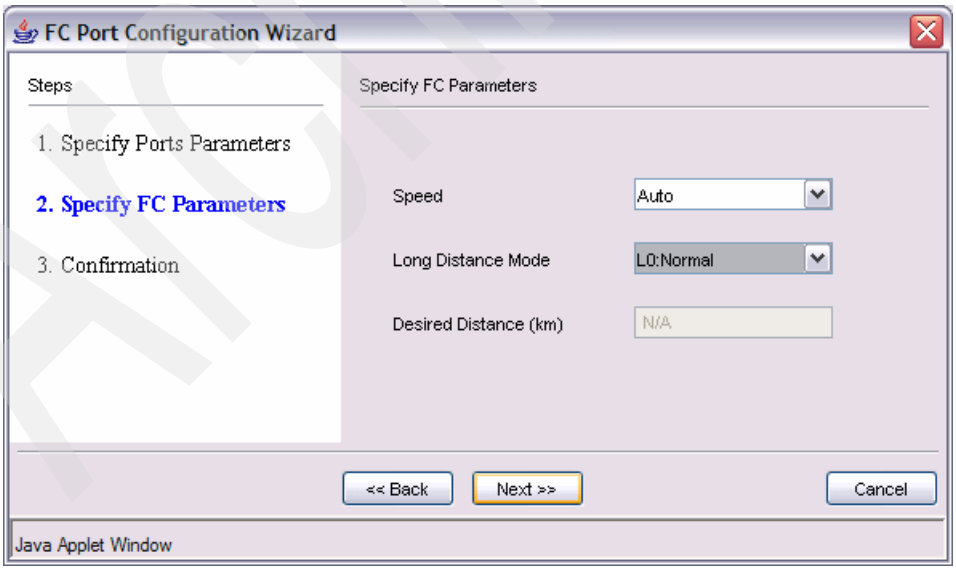


Figure 6-82 FC Parameters

6. Finally, confirm the settings and enable port 12. As you can see in Figure 6-83, the BladeCenter Fibre Switch Module and SAN18B-R router are now connected.

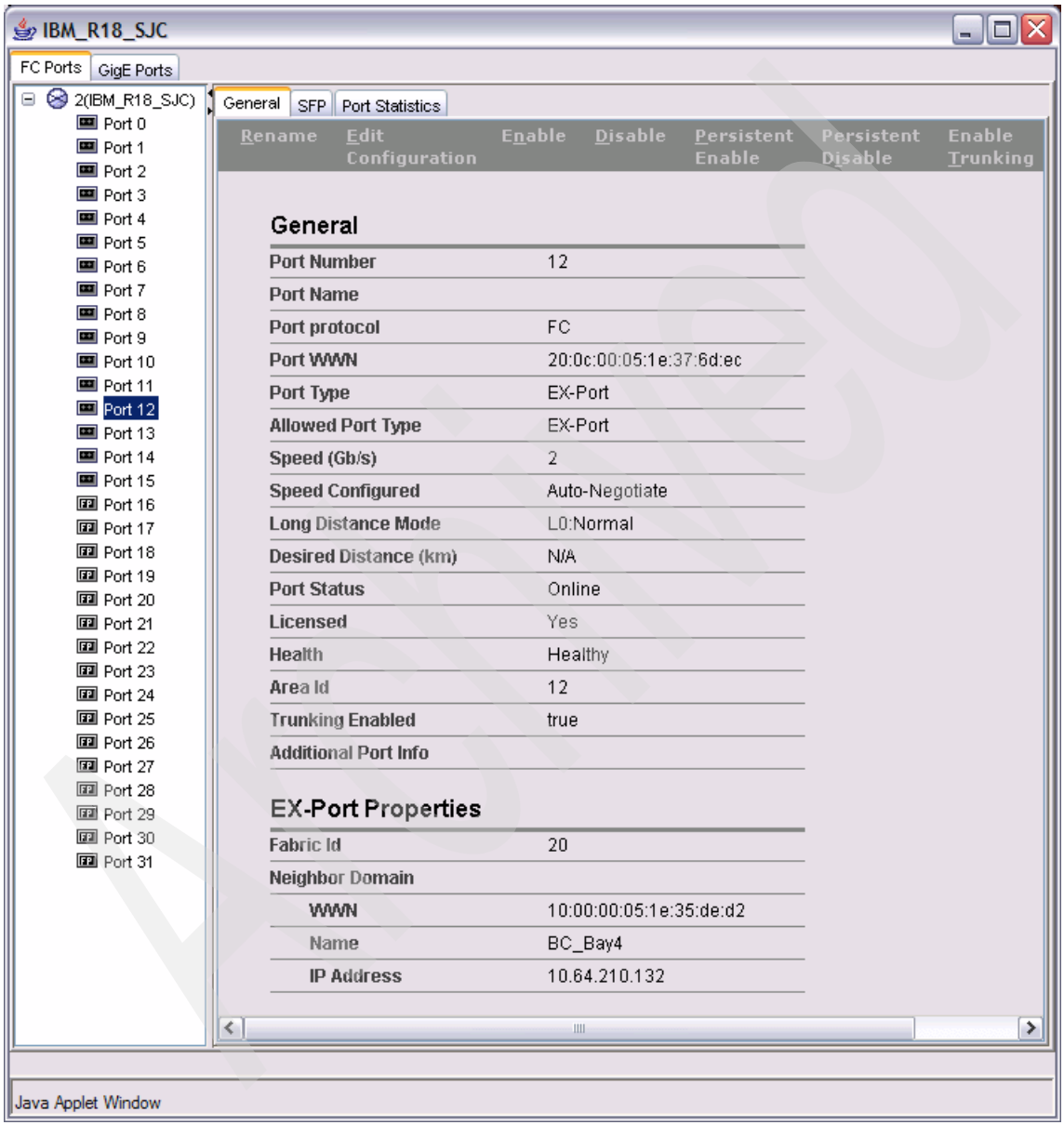


Figure 6-83 SAN Router Port 12 information

FCIP connection between the SAN18B-R and SAN Routing Blade

Now, we configure the FCIP connection between the SAN18B-R and the SAN Routing Blade. We use the ge1 GigE port; the corresponding virtual ports are 24-31. In our case, we use virtual port 24.

Perform the following steps:

1. The first step is to **Persistent Enable** port 24 (Figure 6-84).

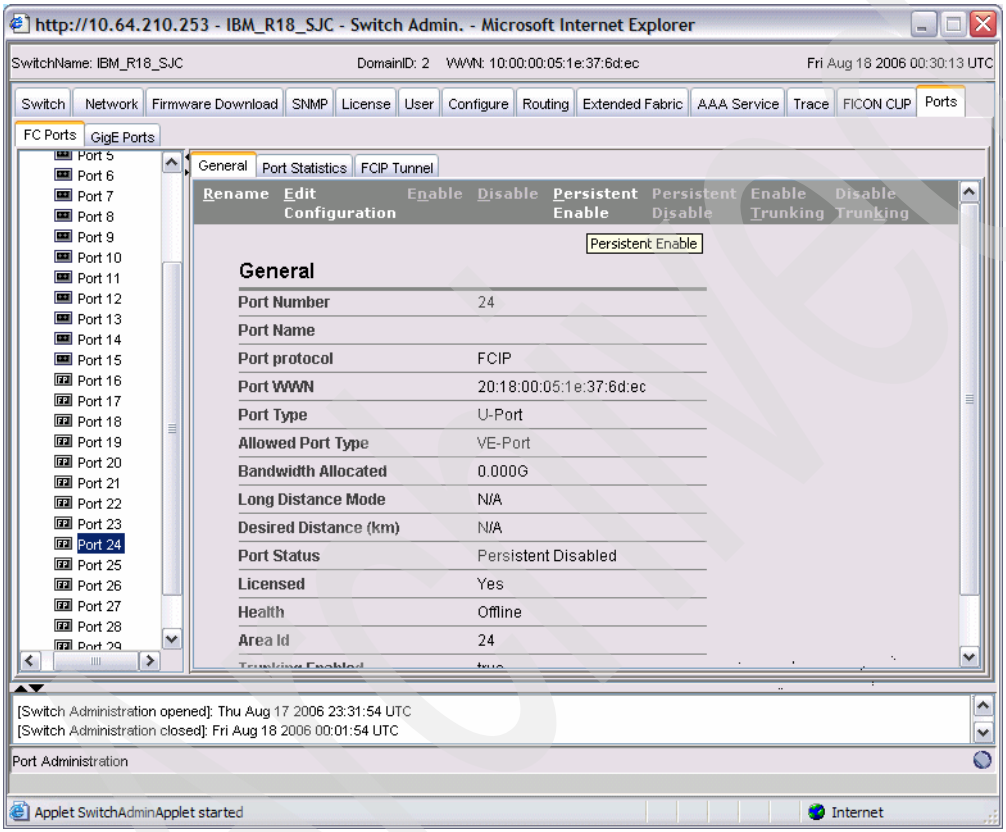


Figure 6-84 Persistent Enable Port 24

2. We now launch the GigE Port Configuration wizard. Select the **GigE Ports** tab and click **ge1**. Now, click **Edit Configuration**, as shown in Figure 6-85.

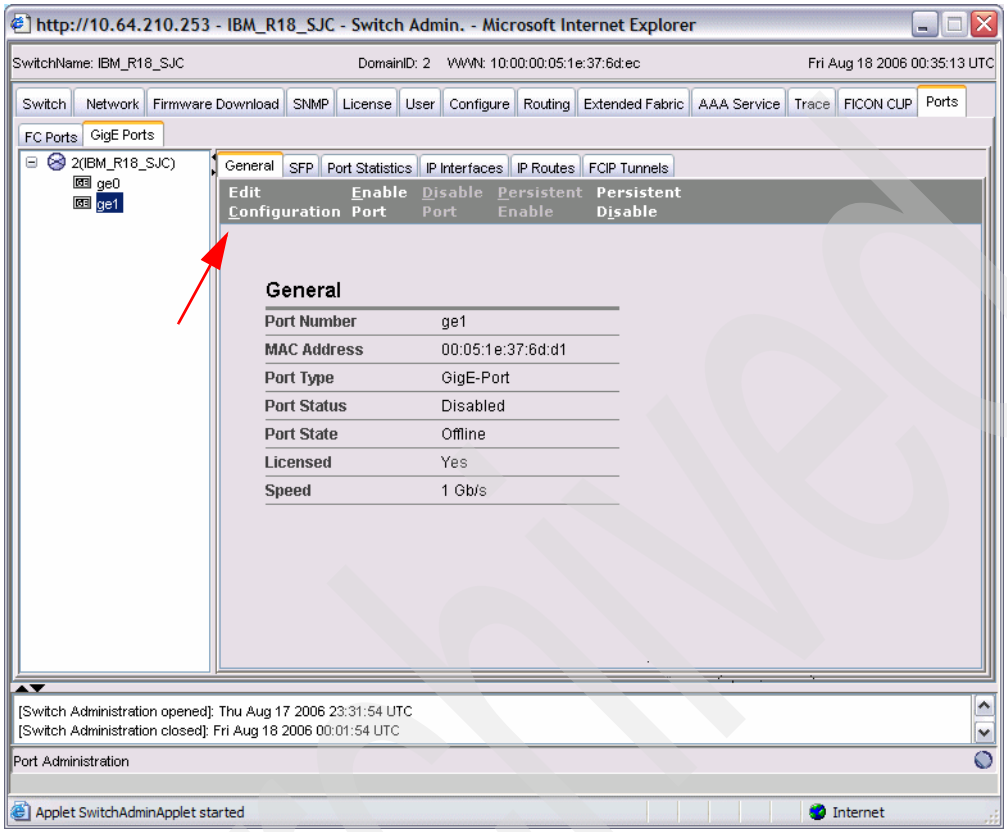


Figure 6-85 GigE ports: Edit Configuration

The wizard guides you through the FCIP configuration steps. The first step is simply an overview (Figure 6-86).

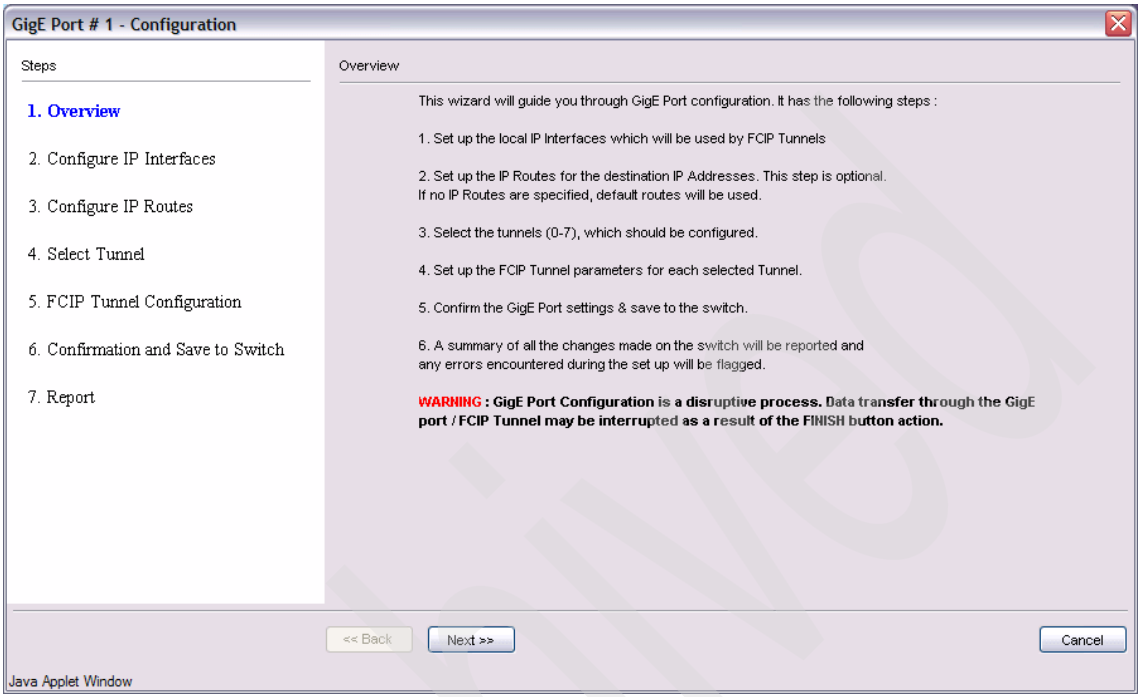


Figure 6-86 GiGE Port Configuration wizard: Overview

3. We define an IP interface. Click **Add Interface** (highlighted in Figure 6-87), and then complete the three mandatory fields: IP Address, Subnet Mask, and MTU Size. We chose the following values:
   - IP Address: `10.10.20.1`
   - Subnet Mask: `255.255.255.0`
   - MTU Size: `1500`

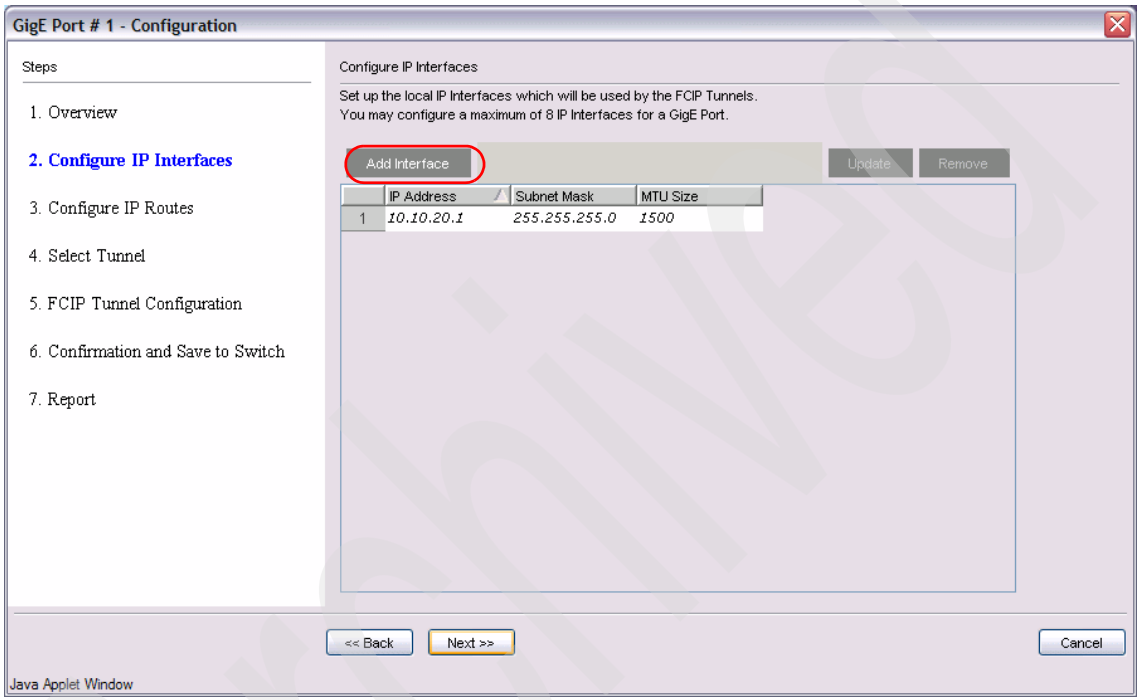


Figure 6-87 Configure IP Interfaces

4. Step number three is to configure the IP routes. This is optional. In our case, we do not need to specify any additional IP routes, so we proceed to step four.

5. We select the FCIP tunnel which will be configured. Up to eight tunnels, numbered 0-7, can be set up across each GigE port. As shown in Figure 6-88, we selected tunnel **0**.

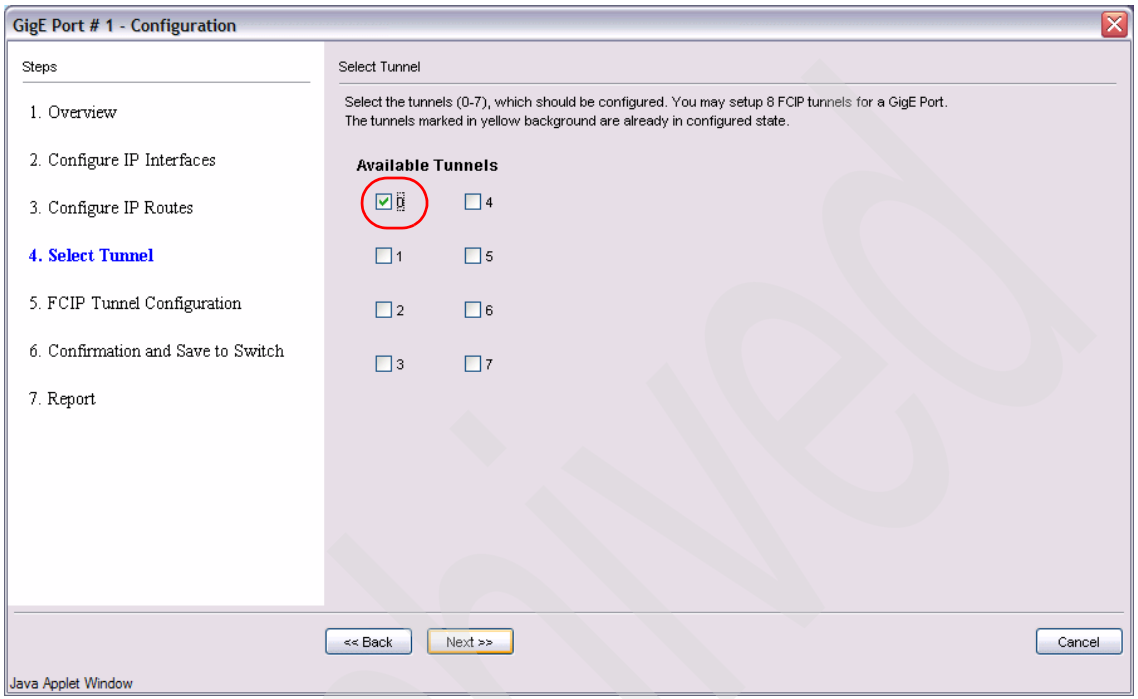


Figure 6-88 Select the FCIP Tunnel

6. Specify the tunnel settings: local and remote IP address and the commit rate. In addition, you can restrict connections to a particular switch and enable compression (Figure 6-89).

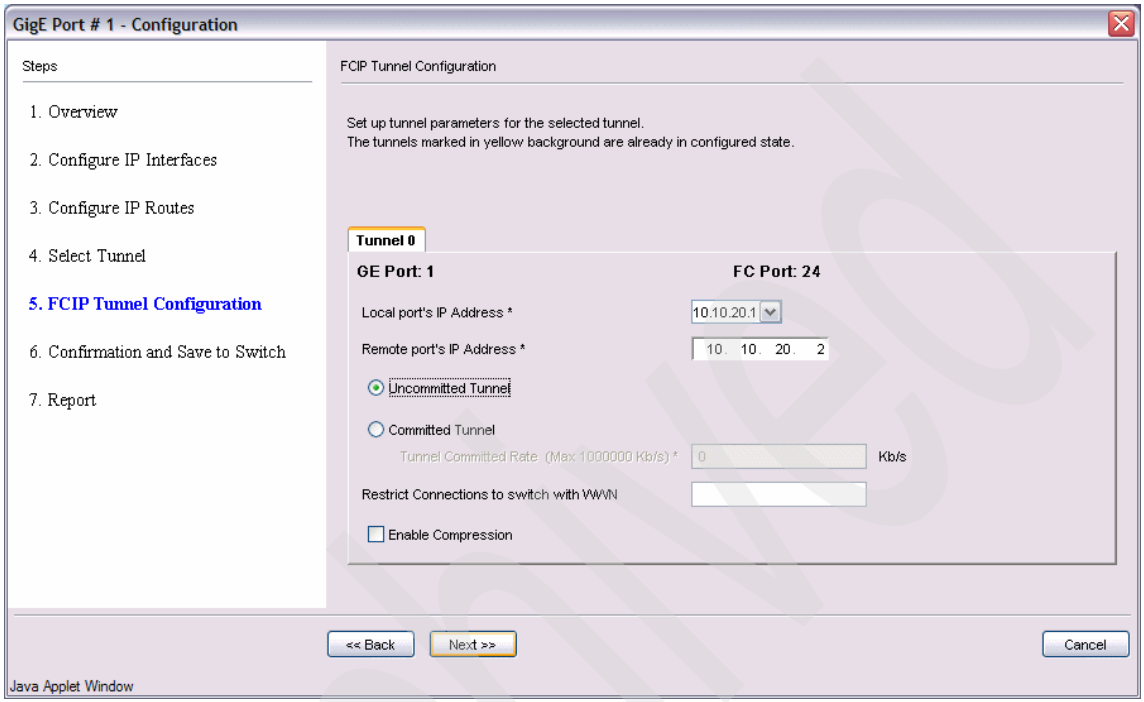


Figure 6-89 FCIP Tunnel Configuration

7. The next step is to confirm the changes and save them to the switch. You can review the configuration and go back if any corrections are necessary. Click **Finish** to save the changes to the switch (Figure 6-90).

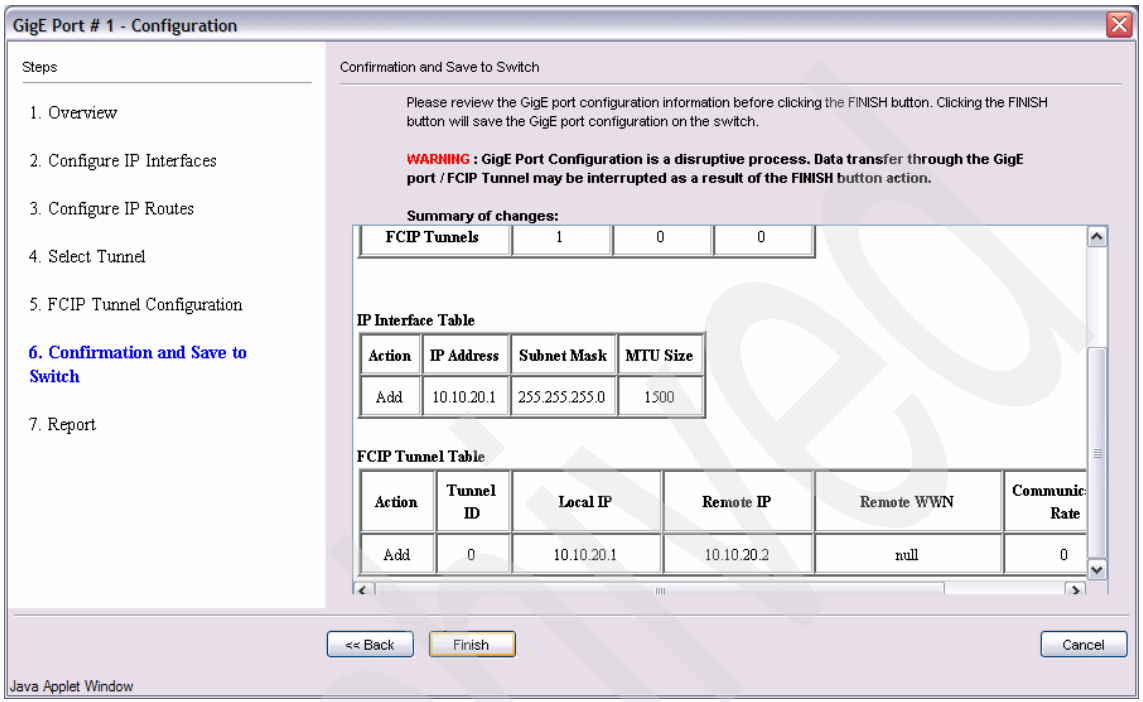


Figure 6-90 Confirm and save the configuration

8. Figure 6-91 shows the final report: The configuration changes were completed successfully. Close the GigE Port Configuration wizard.

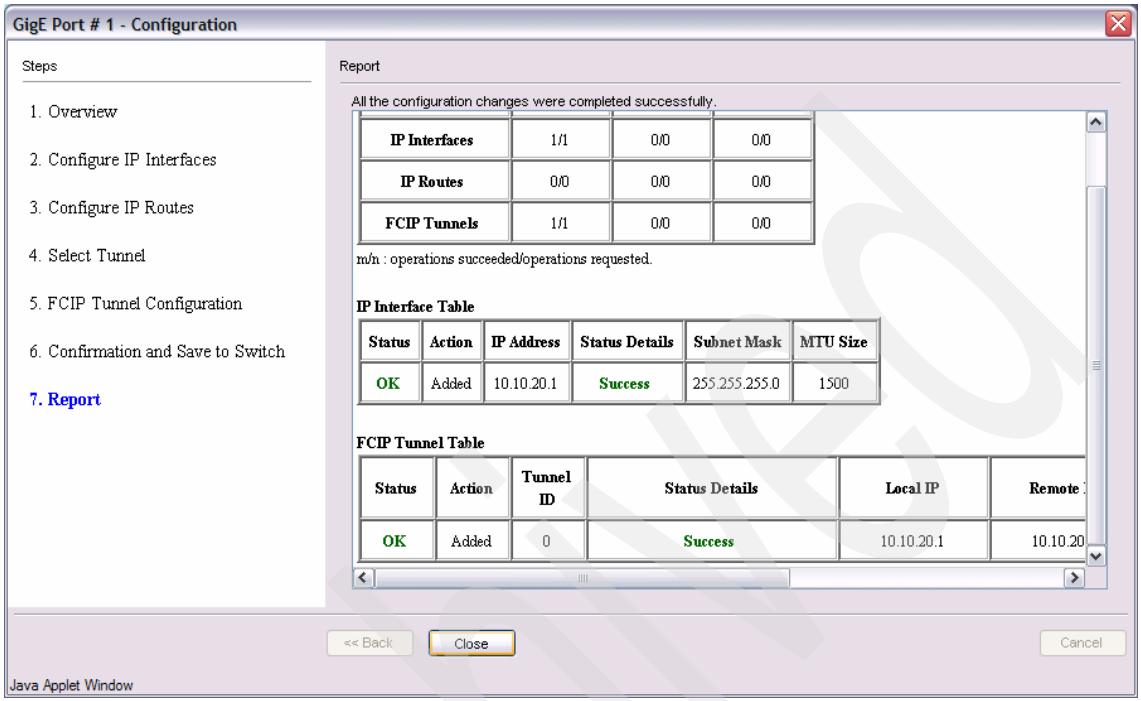


Figure 6-91 Successful completion of configuration changes

VEX_Port configuration

The FCIP connection will be established between a VE_Port and a VEX_Port. Therefore, one of the two virtual ports must be the VEX_Port. We use virtual port 24 on the SAN18B-R and the SAN Routing Blade. The port on the SAN18B-R will be set to VEX_Port type.

Perform the following steps:

1. Figure 6-92 shows the FC Port Configuration Wizard, launched for virtual port 24 on the SAN18B-R. We set the port type to **EX-Port** and set the correct Fabric ID.

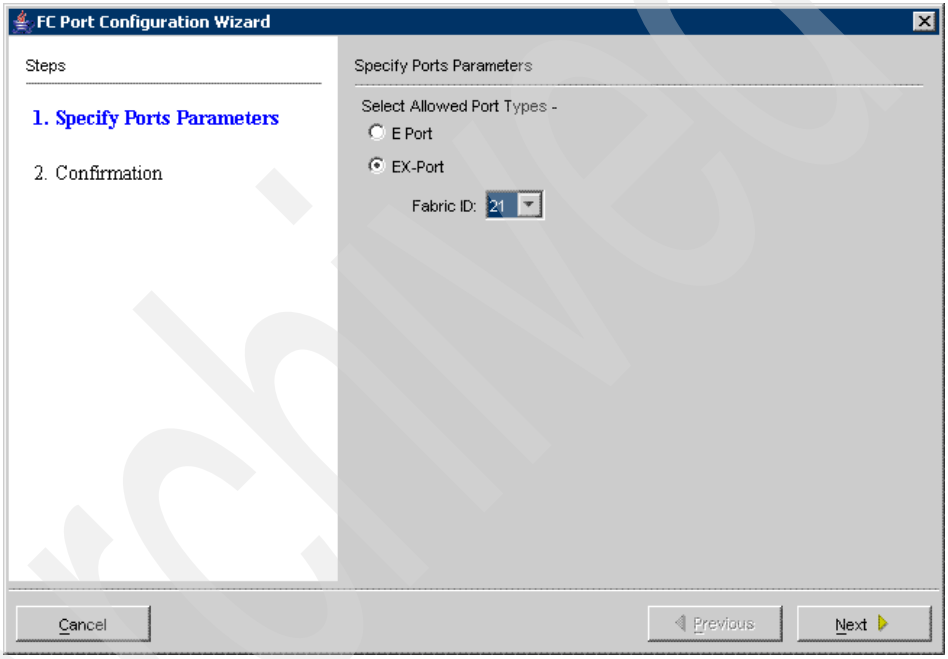


Figure 6-92 VEX-Port Configuration

2. Next, we confirm the settings. Click **Finish** to close the wizard and complete the port configuration, as shown in Figure 6-93.

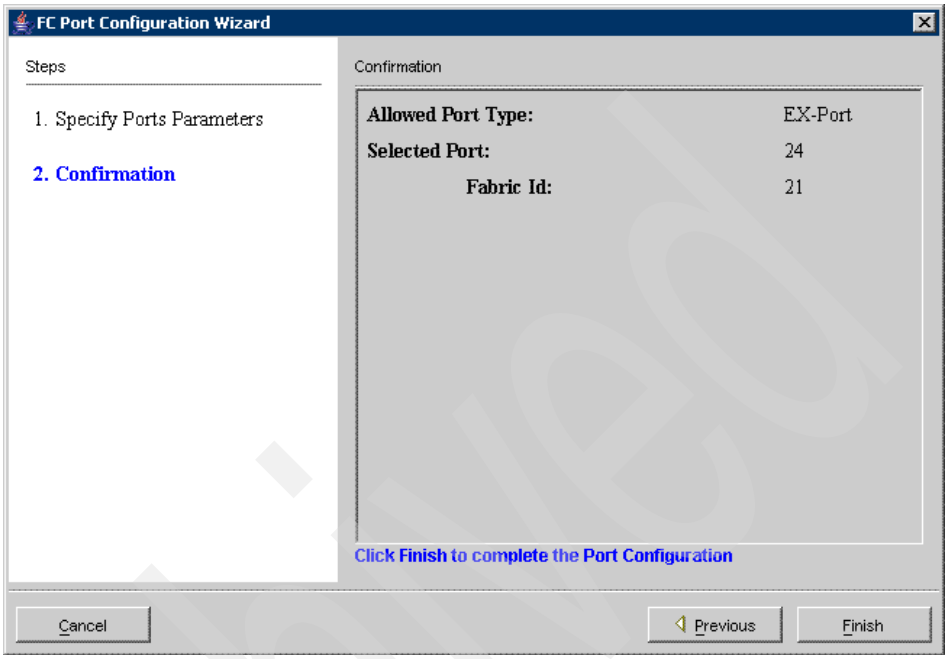


Figure 6-93 Port Configuration Confirmation

3. Enable the virtual FC ports (port 24 in our case) and GigE ports (ge1 in our case) now. If the configuration settings are correct, the FCIP tunnel becomes active (Figure 6-94).

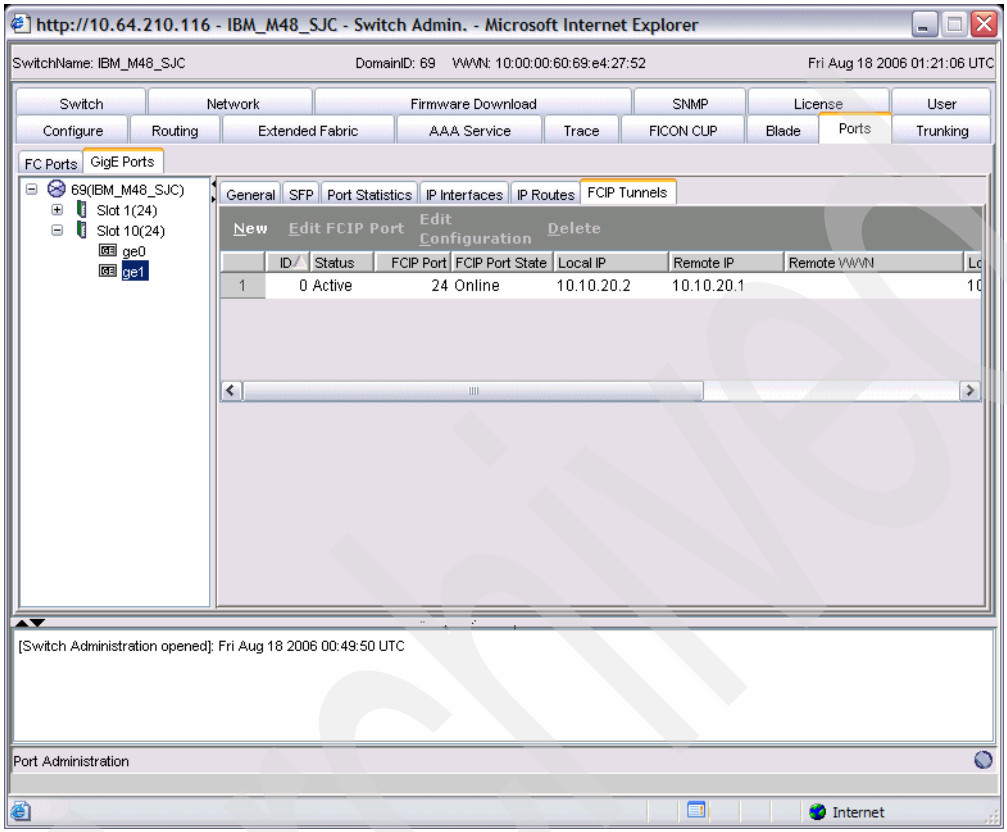


Figure 6-94 FCIP tunnel is active

LSAN configuration using the GUI

Now, we configure the LSAN zones. LSAN zones must be defined on both sides of the FCIP link. In our case, one side is the edge fabric with the BladeCenter Fibre Switch Module and SAN18B-R. The other side is the M48 Director, which contains the SAN Routing Blade in slot 10.

We want to attach the DS400 logical drives to two Blade servers. Our Blade servers are not clustered, so we need to define two separate LSAN zones, one for each Blade server. Each zone will contain a WWN of Blade server FC Expansion Card port and the WWN of the DS400 controller.

Figure 6-95 shows a definition of zone named *LSAN_DS400_Blade1*. The zone has two members:

- WWN of the Blade server 1 FC port (alias *Blade1_HBA2*)
- WWN of the DS400 controller FC port (alias *DS400_A_pt1*)

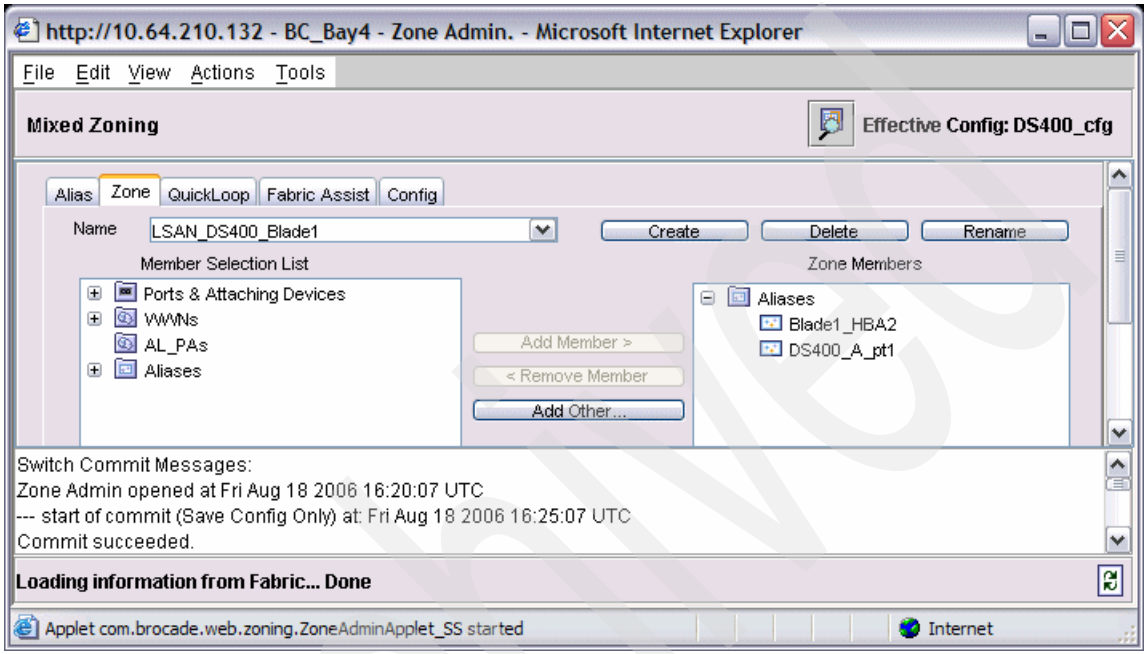


Figure 6-95 LSAN zone configuration on BladeCenter Fibre Switch Module

We need to create another zone, named *LSAN_DS400_Blade2*. This zone will contain the WWN of Blade server 2 FC port and the DS400 controller FC port.

Next, include both LSAN zones in the Fibre Switch Module configuration, as shown in Figure 6-96. The name of the configuration is *DS400_cfg*.

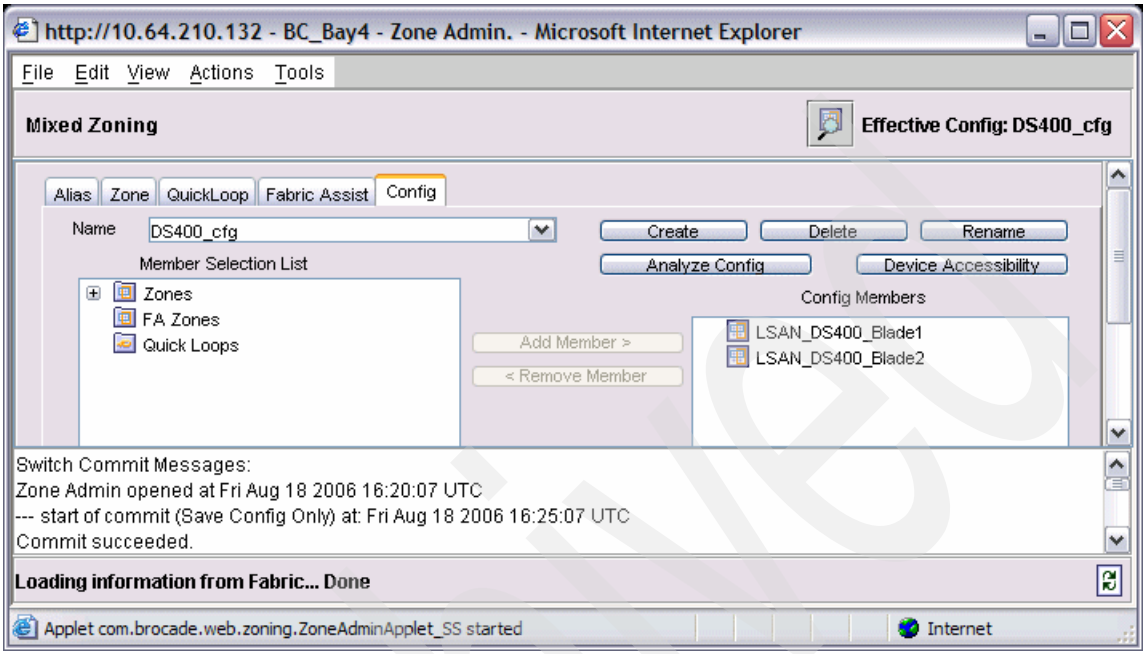


Figure 6-96 BladeCenter Fibre Switch Module configuration

In our two LSAN zones, the Blade server FC Expansion Card port is the local device (it exists in the same edge fabric) and the DS400 controller FC port is the remote device (it exists in the backbone fabric).

We now create corresponding LSAN zones on the other side as well. In our case, this is the M48 (backbone) fabric.

When this is done, the Blade servers should be able to access the DS400 logical drives across the FCIP connection.

FCR administration

The FCR administration module enables you to verify that everything is configured properly. It is accessible from the main WebTools window by clicking the FCR Administration icon.

The module has five tabs: General, EX_Ports, LSAN Fabrics, LSAN Zones, and LSAN Devices. We describe all the tabs in the following sections.

General

Figure 6-97 shows the General tab.

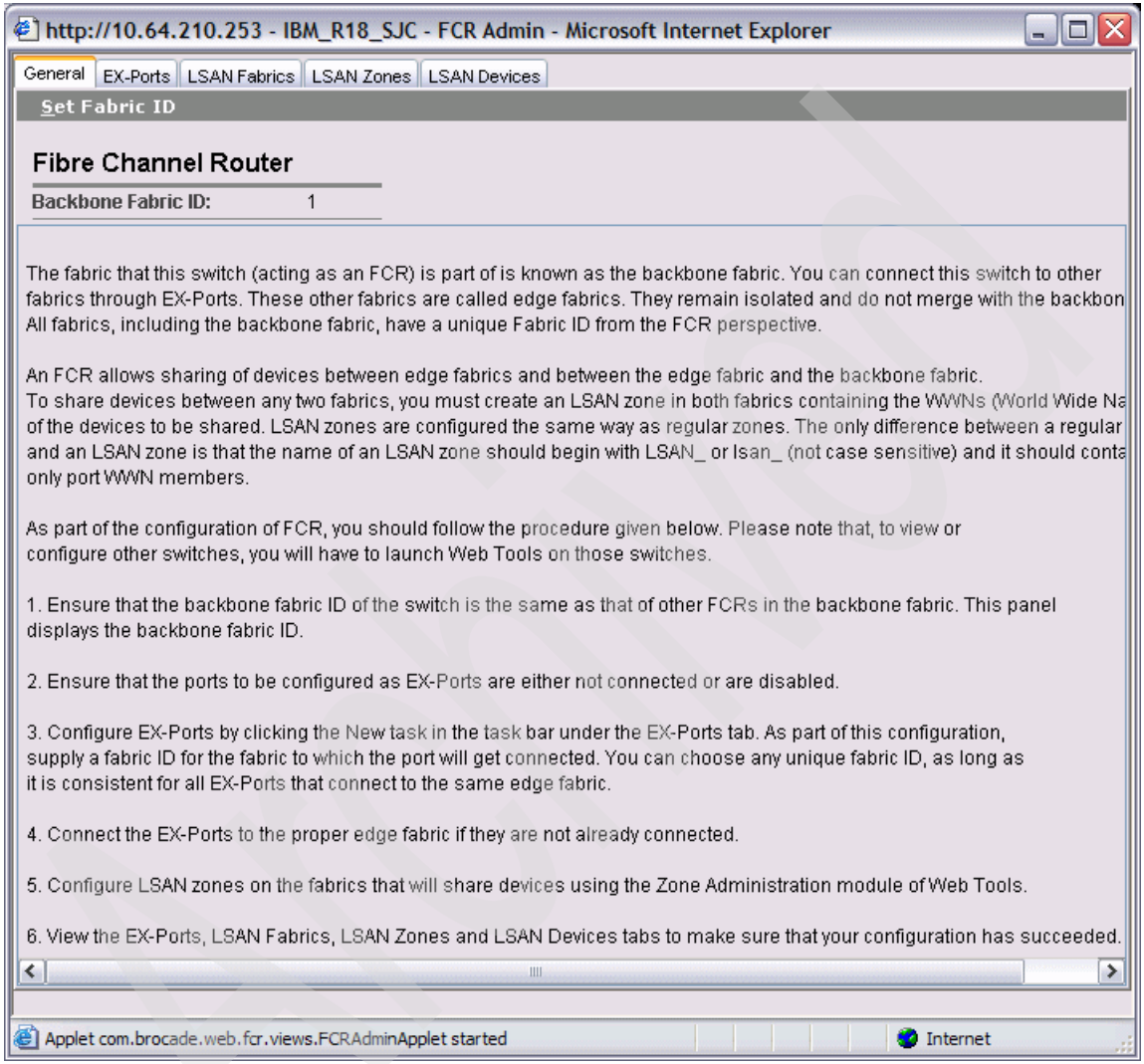


Figure 6-97 FCR Administration: General tab

The window displays basic information about FCR and LSANs. It also outlines the process to set up the FC Routing. We already completed all of the listed steps, but we still want to verify that everything is set up properly.

EX_Ports

Figure 6-98 shows the EX-Ports tab. The left pane lists the EX_Ports known in this fabric. The right pane contains detailed information about the selected EX_Port.

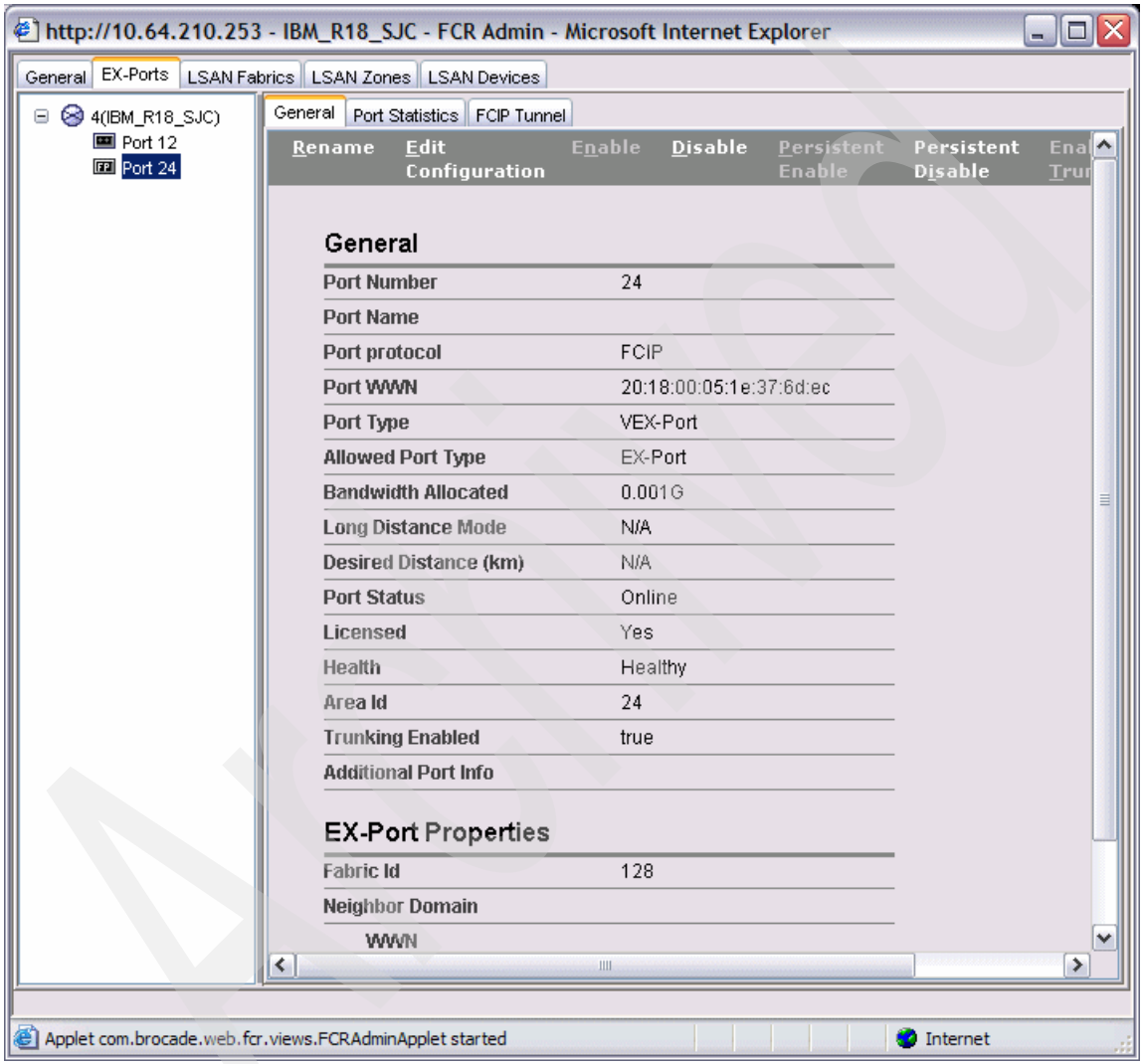


Figure 6-98 FCR Administration: EX-Ports tab

In our case, two EX_Ports are listed. Port 12 is attached to the BladeCenter Fibre Switch Module and virtual port 24 is the VEX_Port used for the FCIP connection.

LSAN Fabrics

The LSAN Fabrics tab lists the available fabrics in its left pane (Figure 6-99). The right pane contains information about the selected fabric. Among other things, it shows the list of Physical LSAN Devices (devices local to this particular fabric) and Proxy LSAN Devices (devices imported from another edge fabric).

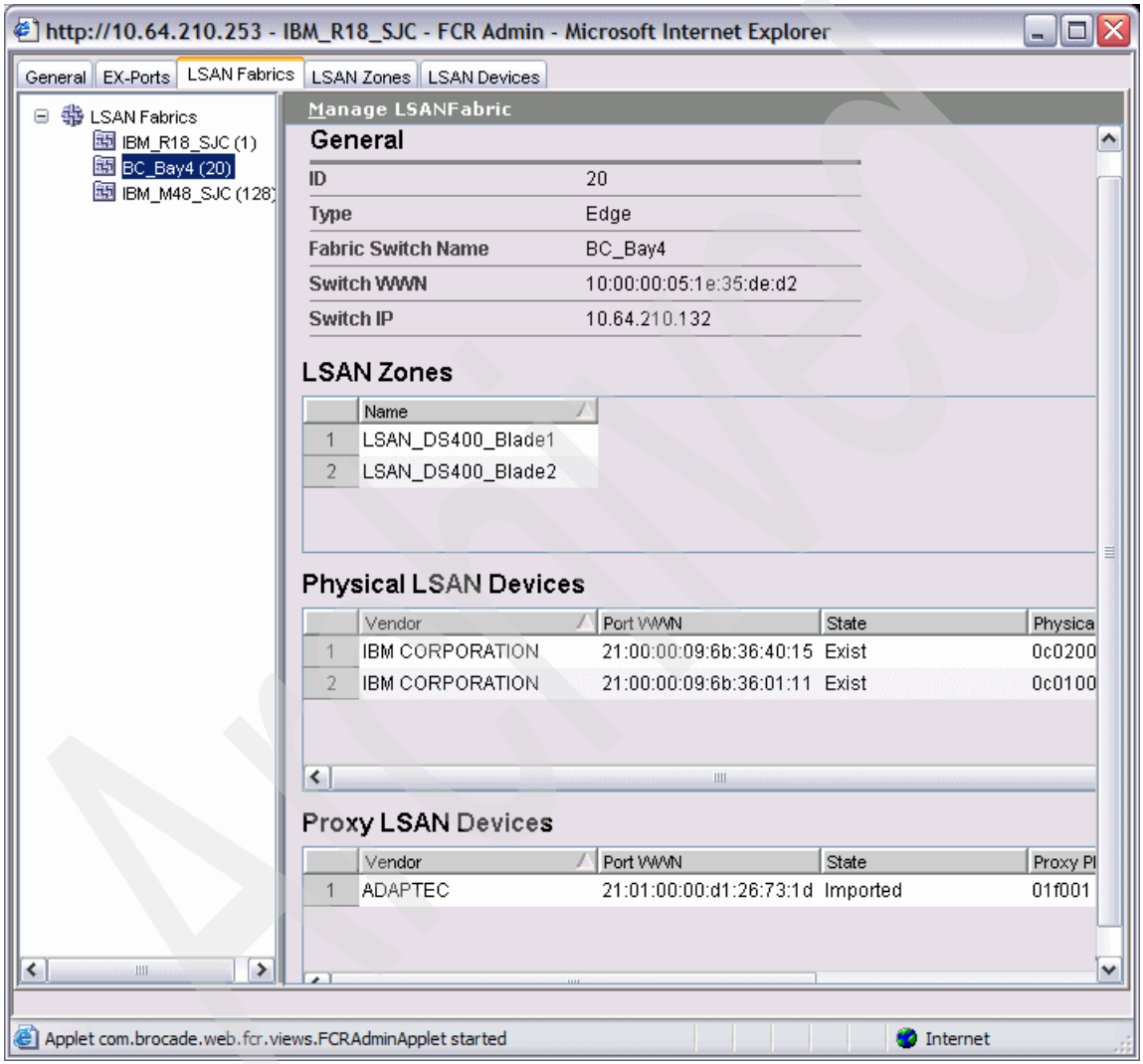


Figure 6-99 FCR Administration: LSAN Fabrics tab

In our case, the FC Expansion Card ports of both Blade servers are local (or physical) devices. The DS400 controller FC port is the imported proxy device.

LSAN Zones

This view lists all the LSAN zones defined in the fabrics (Figure 6-100). You can click any of the zones listed in the left pane. Detailed information about the particular zone appears in the right pane.

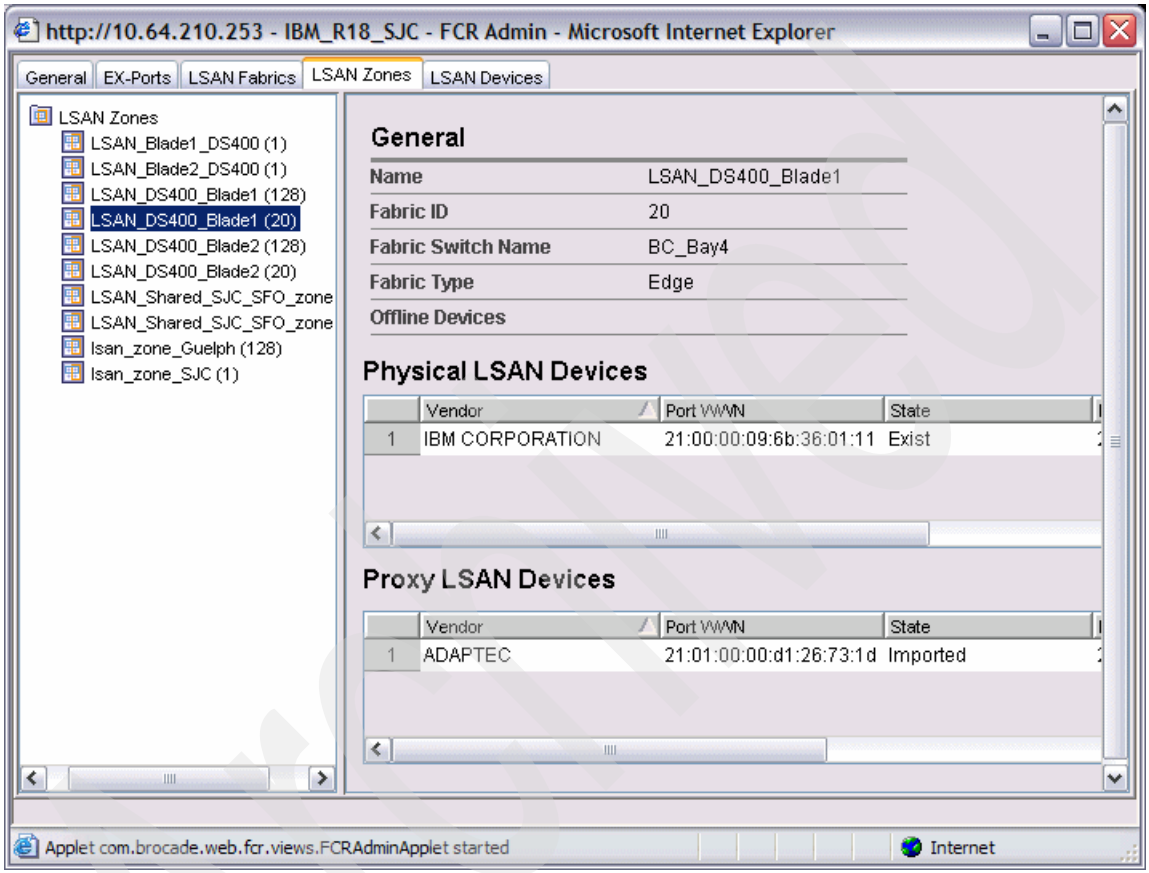


Figure 6-100 FCR Administration: LSAN Zones tab

Figure 6-100 shows information about the LSAN_DS400_Blade1 zone. The zone contains two members: Blade server 1 FC Expansion Card port (listed under Physical LSAN Devices) and DS400 controller FC port (under Proxy LSAN Devices).

LSAN Devices

The final tab is LSAN devices. It simply provides a list of all known physical and proxy devices used in the LSAN zones (Figure 6-101).

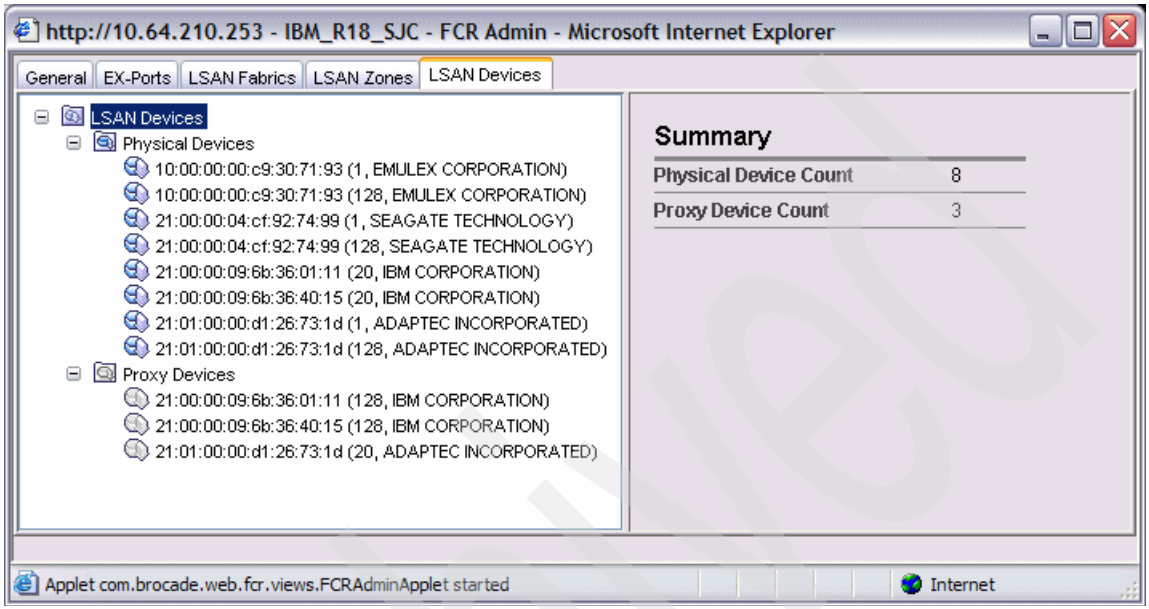


Figure 6-101 FCR Administration: LSAN Devices tab

This concludes our SAN router configuration.

Part 2

Cisco family

In this part, we discuss products associated with the Cisco reseller agreement.

7

Cisco family routing products

This chapter provides information about the routing functions of the Cisco MDS 9000 Multilayer Switches including:

- Hardware and software
- Advanced management
- Key features
- Licensing
- Interoperability

7.1 Overview of the Cisco MDS Family

The Cisco MDS Family is fundamentally different from products designed by other vendors. Routing features are inherent in every one of the Cisco MDS switches, so switch and routing functions do not need to be separated into multiple devices.

Cisco was the first to implement virtual storage area network (VSAN), which has now been adopted as an official standard by the industry. In addition, Cisco is the first to implement Inter-VSAN Routing (IVR) on every port and to support Small Computer System Interface over IP (iSCSI) and Fibre Channel (FC) over IP (FCIP) over Gigabit Ethernet.

The Cisco Systems MDS 9020 (2061-420), MDS 9120 (2061-020), MDS 9140 (2061-040), MDS 9216x (2062-D1A/D1H) Multilayer Fabric Switches, MDS 9506 (2062- D04/T04), MDS 9509 (2062-D07/T07), and the MDS9513 (2062-E11) Multilayer Directors are available from IBM and authorized IBM Business Partners.

Note: Every Fibre Channel port on a Cisco switch is an FC-to-FC routing port.

Introduction to VSAN

VSAN technology is designed to enable efficient storage area network (SAN) use by dividing a physical fabric into multiple logical fabrics. Each VSAN can be zoned as a typical SAN and maintains its own fabric services for added scalability and resilience.

VSAN is a standard feature of all Cisco MDS switches. Transmitting data between VSANs requires the IVR feature, which is included with the Enterprise license option. The Cisco MDS Family also supports routing on every port, including routing to IBM m-type and IBM b-type switches.

SAN-OS 2.1 and later provides support for network address translation (NAT). NAT allows routing between switches with the same domain identifier and routing between VSANs with the same VSAN identifier.

SAN-OS 2.1 and later also provides support for the new Storage Services Module (SSM), which implements the SANTap protocol, Fabric Application Interface Standard (FAIS), and Network Accelerated Serverless Backup (NASB). Aside from Fibre Channel fast write, the SSM requires layered applications from third-party vendors to deliver user functionality.

The following sections discuss the Cisco products that are available as part of the IBM Storage Networking solutions portfolio.

7.2 Hardware and software

This section describes the software and hardware of the Cisco MDS 9000 Multilayer Switch Family.

Important: Not all features are included with the hardware. It is likely that you need to purchase additional software licenses depending on your requirements. Consult your IBM representative for more details.

7.2.1 Cisco MDS 9120 and 9140 Multilayer Switches

The Cisco MDS 9120 Multilayer Fabric Switch (IBM 2061-020) and Cisco MDS 9140 Multilayer Fabric Switch (IBM 2061-040) are one rack-unit (RU) fabric switches that can support 20 or 40 shortwave or longwave small form-factor pluggable (SFP) fiber optic transceivers. Some of these ports operate with a 3.2 to one (3.2:1) over-subscription (fanout) and are referred to as *host optimized ports*.

The MDS 9120 has a total of 20 ports. The first group of four ports on the left side are full bandwidth ports and are identified by a white border. The remaining four groups of ports are host optimized port groups.

Cisco MDS cooling and airflow: MDS 9120 and MDS 9140 switches use what Cisco calls *front-to-rear airflow for cooling*. Be careful, because the *front* is where the Fibre Channel cables are. Only the power cables are in the back. If you install the switches with the ports facing the back for ease of server cabling, the switches will draw in hot air from the servers and might overheat.

If you mount the switches with the ports to the front, as Cisco recommends, you might need to plan cable management carefully because the cables need to connect from the front of the rack to the back of the rack where the server ports are. Alternatives include mounting the switches in a separate communications rack, or mounting the switches with the ports facing the back at the *bottom* of the server rack. However, this might be less convenient for access and does not comply with the best practice of mounting the heaviest devices at the bottom.

By way of contrast, the MDS 92xx and MDS 95xx use *right-to-left cooling*, looking from the front (which is the ports side).

Figure 7-1 shows the MDS 9120 switch.

Figure 7-1 MDS 9120 Multilayer Switch (IBM 2061-020)

The MDS 9140 has a total of 40 ports. The first eight ports on the left side are full bandwidth ports and are identified by a white border. The remaining eight groups of ports are host optimized port groups. Figure 7-2 shows the MDS 9140 switch.

Figure 7-2 MDS 9140 Multilayer Switch (IBM 2061-040)

The switches are configured with dual redundant power supplies, either of which can supply power for the whole switch. They include a hot-swappable fan tray to manage the cooling and airflow for the entire switch.

The 91*n*0 switches share a common firmware architecture with the Cisco MDS 9500 Series of Multilayer Directors, making them intelligent and flexible fabric switches.

Note: The MDS 9120 and MDS 9140 also both support optional coarse wavelength division multiplexing (CWDM) SFPs to provide aggregation of multiple links onto a single optical fiber through a passive optical mux. This is a unique Cisco feature, which allows architects to design relatively low-cost CWDM solutions around Cisco equipment.

7.2.2 MDS 9216A Multilayer Switch

The Cisco MDS 9216A Model D01 (IBM 2062-D1A) is a three RU, 2-slot fabric switch that can support from 16 to 64 shortwave or longwave SFP fiber optic transceivers. Figure 7-3 shows the MDS 9216A switch.

Figure 7-3 MDS 9216A Multilayer Switch (IBM 2062-D1A)

The chassis consists of two slots. The first slot contains the supervisor module. This provides the control and management functions for the 9216A and includes 16 full capability 2 Gbps target-optimized Fibre Channel ports. It contains 2 GB of DRAM and has one internal CompactFlash card that provides 256 MB of storage for the firmware images.

The second slot can contain any one of the modules described in 7.2.7, “Optional modules” on page 217.

The MDS 9216A also supports optional CWDM SFPs to provide aggregation of multiple links onto a single optical fiber through a passive optical mux.

7.2.3 Cisco MDS 9216i Multilayer Switch

The Cisco MDS 9216i uses the same backplane as the MDS 9216A. However, the MDS 9216i includes a fixed 14+2 supervisor module to provide 14 full capability 2 Gbps target-optimized Fibre Channel ports and two Gigabit Ethernet interfaces. The Gigabit Ethernet interfaces support iSCSI initiators connecting to Fibre Channel disk systems. The FCIP and IVR features are bundled with the Cisco MDS 9216i Switch and do not require the Enterprise package.

The second slot can contain any one of the modules described in 7.2.7, “Optional modules” on page 217.

FCIP can help to simplify data protection and business continuance strategies by enabling backup, remote replication, and other disaster recovery services over wide area network (WAN) distances using open-standard FCIP tunneling. Figure 7-4 shows the MDS 9216i.

Figure 7-4 MDS9216i (IBM 2062-D1H)

MDS 9216i can support the following features:

- Integrated IP and Fibre Channel SAN solutions.
- Simplified large storage network management and improved SAN fabric utilization helping to reduce total cost of ownership.
- Throughput of up to 4 Gbps per port (the 14 supervisor module ports are 2 Gbps) and up to 16 Fibre Channel links in a single PortChannel inter-switch link (ISL) connection.
- Scalability.
- Gigabit Ethernet ports for iSCSI or FCIP connectivity.
- Modular design with excellent availability capabilities.
- Intelligent network services that help simplify SAN management and reduce total cost.
- Assistance with security for large enterprise SANs.
- VSAN capability for creating separate logical fabrics within a single physical fabric.
- Compatibility with a broad range of IBM servers, as well as disk and tape storage devices.
- Hardware-based encryption and compression to ensure secure IP links, as well as reducing the bandwidth requirements (and thus cost) on the IP link.
- Up to 3500 buffer-to-buffer credits can be assigned to a single Fibre Channel port.

The MDS 9216i also supports optional CWDM SFPs to provide aggregation of multiple links onto a single optical fiber through a passive optical mux.

7.2.4 MDS 9506 Multilayer Director

The Cisco MDS 9506 (IBM 2062-D04) is a seven RU Fibre Channel director that can support from 12 to 192 shortwave or longwave SFP fiber optic transceivers. These ports fully support 1, 2, and 4 Gbps Fibre Channel and are auto-sensing.

The chassis has six slots, two of which are reserved for dual, redundant supervisor 2 modules, providing 1.4 Tbps of internal system bandwidth. The dual supervisor modules provide the logic control for the director. They also provide high availability and traffic load balancing capabilities across the director. Either supervisor module can control the whole director, with the standby supervisor module providing full redundancy in the event of an active supervisor failure.

The remaining four slots can contain any combination of the modules described in 7.2.7, “Optional modules” on page 217.

The director is configured with dual, redundant power supplies, either of which can supply power for the whole chassis. It also includes a hot-swappable fan tray that manages the cooling and right to left (looking from the SFP side) airflow for the entire director.

Figure 7-5 shows the MDS 9506 Multilayer Director.

Figure 7-5 MDS 9506 Multilayer Director (IBM 2062-D04)

The IBM 2062-T04 product is designed for the telecommunications industry and ships with -48 to -60V dc fed 1900W power supplies. This is the only difference when compared to the 2062-D04.

The MDS 9506 also supports optional CWDM SFPs to provide aggregation of multiple links onto a single optical fiber through a passive optical mux.

7.2.5 MDS 9509 Multilayer Director

The Cisco MDS 9509 Model D07 (IBM 2062-D07) is a 14 RU Fibre Channel director that can support from 12 to 336 shortwave or longwave SFP fiber optic transceivers. These ports fully support 1, 2, and 4 Gbps Fibre Channel and are auto-sensing.

Figure 7-6 on page 214 shows the MDS 9509 Multilayer Director. The chassis has nine slots, two of which are reserved for dual, redundant supervisor 2 modules, providing 1.4 Tbps of internal system bandwidth. The dual supervisor modules provide the logic control for the director and provide high availability and traffic load balancing capabilities across the director. Either supervisor module can control the whole director, with the standby supervisor module providing full redundancy in the event of an active supervisor failure.

Figure 7-6 MDS 9509 Multilayer Director (IBM 2062-D07)

The backplane of the 9509 provides the connectivity for two supervisor modules and up to seven switching modules. In addition to the supervisor and switching modules, the redundant power supplies and the redundant dual clock modules also plug directly into the backplane. If one clock module fails, the remaining clock module takes over operation of the director.

Note: Although there are dual redundant clock modules in the Cisco MDS 950x Directors, if one clock module needs to be replaced, a director outage is required because these modules are not hot-pluggable.

The remaining seven slots can contain any combination of the modules described in 7.2.7, “Optional modules” on page 217.

7.2.6 MDS 9513 Multilayer Director

The Cisco MDS 9513 (IBM 2064-E11) is a 14 RU Fibre Channel director that can support from 12 to 528 shortwave or longwave SFP fiber optic transceivers. These ports fully support 1, 2, and 4 Gbps Fibre Channel and are auto-sensing.

Figure 7-7 on page 215 shows the MDS 9513 Multilayer Director. The chassis has 13 slots, two of which are reserved for dual, redundant supervisor 2 modules, providing 2.2 Tbps of internal system bandwidth. The dual supervisor modules provide the logic control for the director and provide high availability and traffic load balancing capabilities across the director. Either supervisor module can control the whole director, with the standby supervisor module providing full redundancy in the event of an active supervisor failure.

Figure 7-7 MDS 9513 Multilayer Director (IBM 2064-E11)

Table 7-1 compares hardware features within the Cisco 95xx Series Multilayer Directors.

Table 7-1 Cisco MDS 95xx hardware feature comparison

Feature	MDS 9506	MDS 9509	MDS 9513
Available slots	6	9	13
Available option slots	4	7	11
Redundant Supervisor	Yes	Yes	Yes
Maximum 1/2/4 Gbps FC ports per chassis	192	336	528
Maximum 10 Gbps FC ports per chassis	16	28	44
Maximum iSCSI and FCIP ports per chassis	24	48	60
Rack units	7	14	14

Supervisor 2 modules

The supervisor 2 module is the heart of the 9500 Series Directors. It provides the control and management functions for the director, as well as an integrated crossbar switching fabric. The crossbar fabric provides up to 2.2 Tbps full duplex switching capacity in the MDS 9513, and 1.4 Tbps in the MDS 9509 and MDS 9506.

The MDS 9500 Series comes with two supervisor 2 modules as standard for redundancy and availability. In the event of a supervisor module failing, the surviving module becomes active, taking over the operation of the director.

Note: The MDS 9216x uses a different supervisor module, which integrates 16 (9216A) or 14 (9216i) target optimized ports with up to 670 Gbps full duplex switching. The function provided by the MDS 9216x supervisor is the same as the one described in this section.

Control and management

The supervisor module provides the following control and management features:

- Multiple paths avoid a single point of failure.
- A redundant central arbiter provides traffic control and access fairness.
- It performs a nondisruptive restart of a single failing process on the same supervisor. A kernel service running on the supervisor module keeps track of the high availability policy of each process and issues a restart when a process fails. The type of restart issued is based on the process's capability:
  - Warm or stateful (state is preserved).
  - Cold or stateless (state is not preserved).

  If the kernel service is unable to perform a warm restart of the process, it issues a cold restart.
- It performs a nondisruptive switchover from the active supervisor to a redundant standby without loss of traffic.

If the supervisor 2 module has to be restarted, the secondary supervisor (continuously monitoring the primary) takes over. The switchover is non-revertive. After a switchover has occurred and the failed supervisor is replaced or restarted, the operation does not switch back to the original primary supervisor, unless it is forced to switch back, or another failure occurs.

Crossbar switching fabric

The MDS 9500 supervisor 2 module provides a crossbar switching fabric that connects all the modules. A single crossbar provides 700 Gbps per switching module and 1.4 Tbps with both modules present in the MDS 9506 and the MDS 9509. In the MDS 9513, a single crossbar provides 1.1 Tbps per switching module and 2.4 Tbps with both modules present.

Figure 7-8 shows the 9500 Series supervisor 2 module.

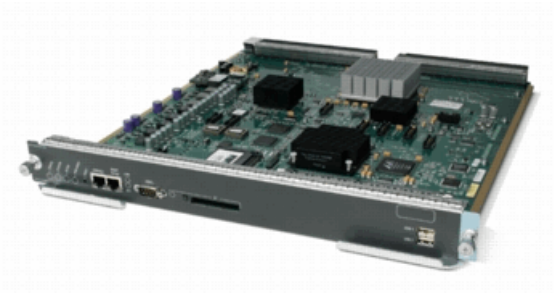


Figure 7-8 MDS 9500 Series supervisor 2 module

7.2.7 Optional modules

The MDS 9200 and 9500 Families allow optional modules to provide additional port connectivity, IP services, or storage virtualization functionality into empty expansion slots.

Refer to Table 7-1 on page 215 for option slot availability on your switch.

The 16-port Switching Module (feature code (fc) 2116)

The 16-port Switching Module provides up to 64 Gbps of continuous aggregate bandwidth. Autosensing 1 Gbps and 2 Gbps target-optimized ports deliver 200 MBps and 255 buffer credits per port.

Note: The 64-Gbps, continuous, aggregate bandwidth is based on 2 Gbps per port in full duplex mode, that is:

16 ports at 2 Gbps (or 213 MBps) in both directions = 64 Gbps

The 16-port module is designed for attaching high-performance servers and storage subsystems, and for connecting to other switches using ISL connections. This module also supports optional CWDM SFPs to provide aggregation of multiple links onto a single optical fiber through a passive optical mux.

Figure 7-9 shows the 16-port Switching Module for the Cisco MDS 9000 Family.

Figure 7-9 16-port Switching Module

The 32-port Switching Module (fc 2132)

The 32-port Switching Module is designed to deliver an optimal balance of performance and port density. This module provides high line-card port density along with 64 Gbps of total bandwidth and 12 buffer-to-buffer credits per port. Bandwidth is allocated across eight 4-port groups, with each port group sharing 2.5 Gbps, making it an aggregate bandwidth of approximately 5 Gbps full-duplex. This module provides a low-cost means to attach lower performance servers and storage subsystems to high-performance crossbar switches without requiring ISLs.

By combining 16- and 32-port Switching Modules in a single, modular chassis, administrators can configure price- and performance-optimized storage networks for a wide range of application environments.

The 32-port Switching Module also supports optional CWDM SFPs to provide aggregation of multiple links onto a single optical fiber through a passive optical mux.

Switching modules are designed to be interchanged or shared between all Cisco MDS 9200 Switches and 9500 Directors. Figure 7-10 shows the 32-port Switching Module for the Cisco MDS 9000 Family.

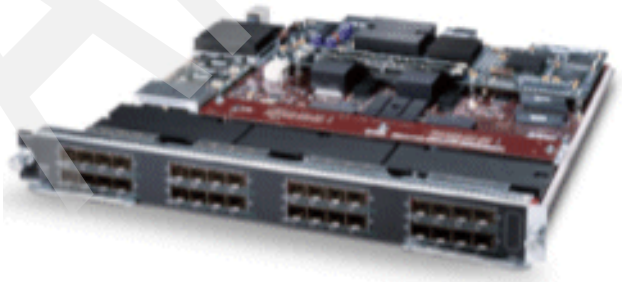


Figure 7-10 32-port Switching Module

Cisco MDS 9000 14+2 Multiprotocol Services Module (fc 2214)

The Cisco MDS 9000 14+2 Multiprotocol Services Module is designed to provide 14 Fibre Channel and two IP storage interfaces. The 14 Fibre Channel ports are based around the same full rate target optimized ports as the 16-port module, providing all the same operating modes. In addition, the 14+2 card can be configured with high buffer credits on one Fibre Channel port to support longer distance FC-to-FC connections.

The two IP storage interfaces are similar to the IP Services Module, including hardware compression and security.

Restriction: The two Ethernet ports on the 14+2 Multiprotocol Services Module *cannot* be combined into a single EtherChannel. However, PortChannel can be used.

Figure 7-11 shows the Cisco MDS 9000 14+2 Multiprotocol Services Module.

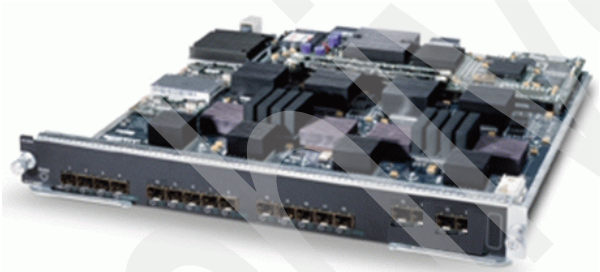


Figure 7-11 Cisco MDS 9000 14+2 Multiprotocol Services Module

This module also supports optional CWDM SFPs to provide aggregation of multiple links onto a single optical fiber through a passive optical mux.

Eight-port IP Services Module (fc 2208)

The IP Services (IPS) Module provides eight Gigabit Ethernet ports that can support iSCSI and FCIP protocols simultaneously. Because the bit rate of Gigabit Ethernet is different from the bit rate of Fibre Channel, the card requires tri-rate SFPs.

Figure 7-12 shows the 8-port IP Services Module.

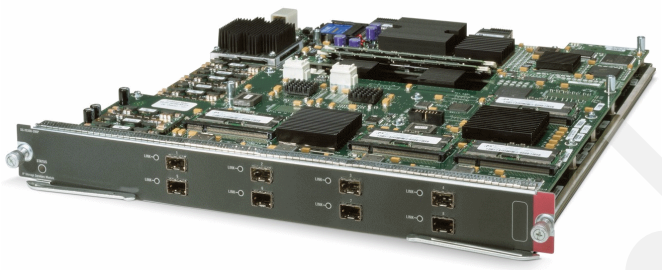


Figure 7-12 The 8-port IP Services Module

Note: Two Ethernet ports on the IPS modules *can* be combined into a single EtherChannel, but only between ports that share the same application-specific integrated circuit (ASIC). However, PortChannel can be used.

Ports configured to run FCIP

The ports configured for FCIP can support up to three, virtual ISL connections (FCIP tunnels). This way, you can transport Fibre Channel traffic transparently, except for latency, over an IP network between two FCIP-capable switches. Each virtual ISL connection acts as a normal Fibre Channel ISL or extended ISL (EISL). Advanced functionality includes FCIP compression, FCIP write compression, and FCIP tape acceleration.

To use FCIP, you need to purchase the FCIP Activation for 8-port IP Services Line Card feature for every 8-port IP line card that needs to support FCIP.

Ports configured to run iSCSI

Ports configured to run iSCSI work as a gateway between iSCSI hosts and Fibre Channel-attached targets. The module terminates iSCSI commands and issues new Fibre Channel commands to the targets.

The Cisco Fabric Manager is used to discover and display iSCSI hosts. These iSCSI hosts are bound to assigned worldwide names (WWNs) and create a static relationship that enables:

- Zoning of iSCSI initiators
- Accounting against iSCSI initiators
- Topology mapping of iSCSI initiators
- 5000 simultaneous connections per switch/director

Storage Services Module (fc 2400)

The Storage Services Module (SSM) is based on the 32-port Fibre Channel Switching Module and provides intelligent storage services in addition to 1 Gbps and 2 Gbps Fibre Channel switching. The SSM uses eight IBM PowerPC processors for SCSI data-path processing. It can be combined with the optional Cisco MDS 9000 Enterprise package to enable Fibre Channel write acceleration (FC-WA).

FC-WA can help improve the performance of remote mirroring applications over extended distances by reducing the effect of transport latency when completing a SCSI operation over distance. This supports longer distances between primary and secondary data centers and can help improve disk replication performance.

The optional Storage Systems Enabler package bundle can enable independent software vendors (ISVs) to develop intelligent fabric applications that can be hosted on the SSM through an application programming interface (API).

ISVs can use the API to offer the following applications:

- Network-accelerated storage applications, such as serverless backup
- Network-assisted appliance-based storage applications using Cisco MDS 9000 SANTap Service, such as global data replication
- Network-hosted storage applications based on proposed Fabric Application Interface Standard (FAIS) APIs offered by ISVs

Note: IBM support for these ISV applications is limited to IBM TotalStorage Proven™ solutions. For the most current IBM TotalStorage Proven information, go to:

http://www.ibm.com/storage/proven

Figure 7-13 shows the Storage Services Module.

Figure 7-13 Storage Services Module

The 12-port 4 Gbps Switching Module (fc 2412)

The 12-port 4 Gbps Switching Module is ideal for attachment to the highest performance 4 Gbps-enabled storage and for ISL connections. The 12-port 4 Gbps Switching Module can deliver up to 96 Gbps of full duplex bandwidth.

Figure 7-14 shows the 12-port Switching Module.

Figure 7-14 12-port 4 Gbps Switching Module

The 24-port 4 Gbps Switching Module (fc 2424)

The 24-port 4 Gbps Switching Module delivers an ideal balance of performance and scalability. The twenty-four 4 Gbps ports deliver up to 96 Gbps of full duplex bandwidth. Bandwidth is allocated across four 6-port, port groups, providing 24 Gbps of full-duplex bandwidth per port group. Port Bandwidth Reservation enables switching bandwidth to be dedicated to a port, providing flexibility to optimize high-demand ports such as ISLs.

Figure 7-15 shows the 24-port Switching Module.

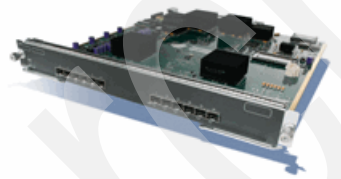


Figure 7-15 24-port 4 Gbps Switching Module

The 48-port 4 Gbps Switching Module (fc 2448)

The 48-port 4 Gbps Switching Module delivers up to 96 Gbps of total bandwidth. Bandwidth is allocated across four 12-port, port groups, providing 24 Gbps bandwidth per port group. Port Bandwidth Reservation enables switching bandwidth to be dedicated to a port, providing flexibility to optimize high-demand ports such as ISLs.

Figure 7-16 shows the 48-port 4 Gbps Switching Module.

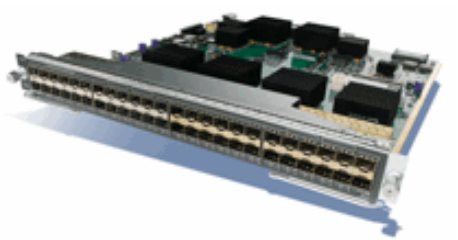


Figure 7-16 48-port 4 Gbps Switching Module

Buffer credits

Buffer credits affect the number of input/outputs (I/Os) that can be sent before an acknowledgement is received. In extended Fibre Channel networks, you need more buffer credits to keep the “pipe” filled because the latency has increased.

Each target-optimized port supports 255 buffer credits, and host-optimized ports support 12 buffer credits per port. On the 14+2 line card, up to 3500 buffer credits can be assigned to a single port if you are willing to sacrifice buffers on other ports and shut down three ports on the quad controlled by that ASIC. A maximum of 1500 buffer credits can be configured if the additional three ports are left enabled.

7.3 Advanced management

The Cisco SAN-OS is the operating system running within the MDS 9000 supervisor and modules to enable the multilayer functionality of the products. Cisco SAN-OS provides a rich suite of management tools.

Although the SAN-OS installable files are specific to each MDS 9000 platform (9100, 9200, and 9500), the standard features provided by the SAN-OS are common to all, although some features are applicable only to switches with Ethernet ports. Features include support for:

- Fibre Channel Protocol (FCP)
- iSCSI
- VSANs
- Zoning
- FCC
- IVR
- Virtual Output Queuing
- Diagnostics (SPAN, RSPAN, and so on)
- SNMPv3

- SSH
- SFTP
- RBAC
- Radius
- High availability
- PortChannels
- RMON
- Call home
- TACACS+
- FDMI
- SMI-S (XML-CIM)
- iSNS client
- iSNS
- IPS ACLs
- Fabric Manager
- Heterogeneous IVR
- WWN-based VSANs
- Zone-based quality of service (QoS)
- Auto-creation of PortChannels
- Enhanced zoning (locking)
- Cisco fabric services (lock and apply changes across the fabrics)

7.3.1 Fabric management

The Cisco MDS 9000 Family provides three modes of management:

- The MDS 9000 Family command-line interface (CLI) presents the user with a consistent, logical CLI, which adheres to the syntax of the widely known Cisco IOS CLI. This is an easy-to-use command interface with broad functionality.
- The Cisco Fabric Manager is a Java application that simplifies management across multiple switches and fabrics. It enables administrators to perform such tasks as topology discovery, fabric configuration and verification, provisioning, monitoring, and fault resolution. All functions are available through a remote management interface.
- Cisco also provides an API for integration with third-party and user-developed management tools.

Cisco MDS 9000 Fabric Manager

The Cisco Fabric Manager is included with the Cisco MDS 9000 Family of switches and is a Java and Simple Network Management Protocol (SNMP)-based network fabric and device management tool. It provides a GUI that displays real-time views of your SAN fabric and installed devices. The Cisco Fabric Manager provides three views for managing your network fabric:

- The Device View displays a continuously updated physical picture of device configuration and performance conditions for a single switch.
- The Fabric View displays a view of your network fabric, including multiple switches.
- The Summary View presents a summary view of switches, hosts, storage subsystems, and VSANs.
- The Cisco Fabric Manager provides an alternative to the CLI for most switch configuration commands.

The Cisco Fabric Manager is included with each switch in the Cisco MDS 9000 Family.

In-band management and out-of-band management

The Cisco Fabric Manager requires an out-of-band (Ethernet) connection to at least one Cisco MDS 9000 Family switch to enable it to discover and manage the fabric.

The interface used for an out-of-band management connection is a 10/100 Mbps Ethernet interface on the supervisor module, labeled mgmt0. The mgmt0 connection can be connected to a management network to access the switch through IP over Ethernet.

Ethernet connectivity is required to at least one Cisco MDS 9000 Family switch. This connection is then used to manage the other switches using in-band (Fibre Channel) connectivity. Otherwise, you need to connect the mgmt0 port on each switch to your Ethernet network.

Each supervisor module has its own Ethernet connection; however, the two connections in a redundant supervisor system operate in active or standby mode. The active supervisor module also hosts the active mgmt0 connection. When a failover event occurs to the standby supervisor module, the IP address and medium access control (MAC) address of the active Ethernet connection are moved to the standby Ethernet connection. This eliminates any need for the management stations to relearn the location of the switch.

Figure 7-17 shows an example of an out-of-band management solution.

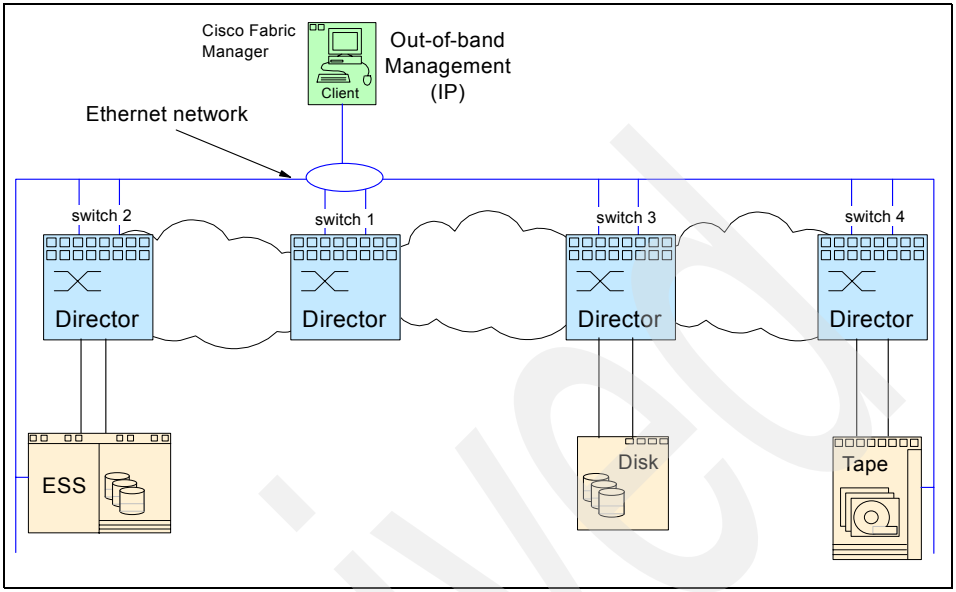


Figure 7-17 Out-of-band management connection

You can also manage switches on a Fibre Channel network using an in-band connection to the supervisor module. This in-band connection supports either management protocols over Fibre Channel or IP embedded within Fibre

Channel. The Cisco MDS 9000 Family supports RFC 2625 IP over Fibre Channel (IPFC), which allows IP to be transported between Fibre Channel devices over the Fibre Channel Protocol, as shown in Figure 7-18.

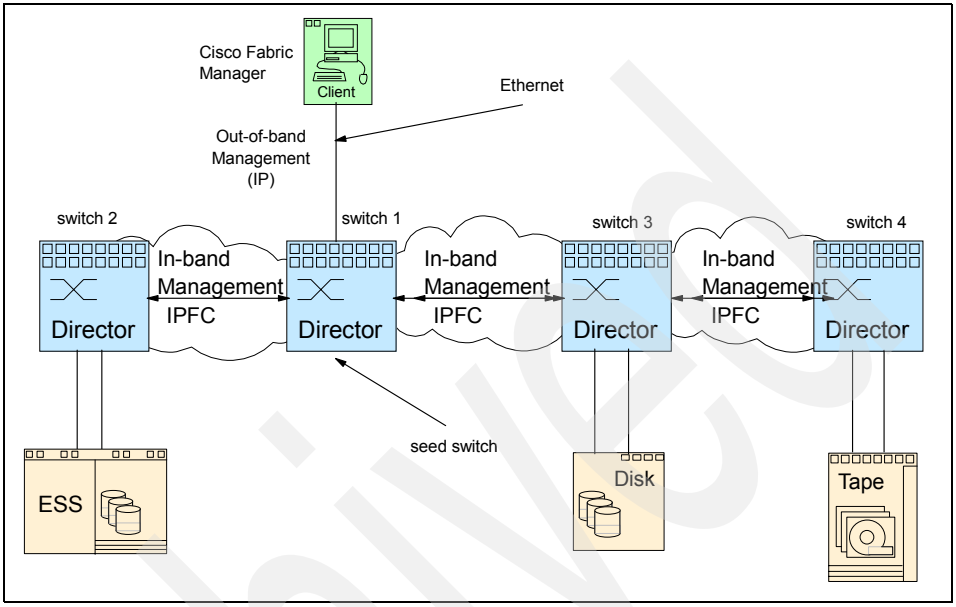


Figure 7-18 In-band management connection

IPFC encapsulates IP packets into Fibre Channel frames so that management information can cross the Fibre Channel network without requiring a dedicated Ethernet connection to each switch. IP addresses are resolved to the Fibre Channel address through the Address Resolution Protocol (ARP). With host bus adapters (HBAs) that support IP drivers, this capability allows for a completely in-band management network. The switch also uses the in-band interface to discover its own environment, including directly connected and fabric-wide elements.

Cisco now also provides the capability to assign an IP address to each VSAN and manage each VSAN in-band or out-of-band.

Note: When initially setting up a Cisco MDS Multilayer Switch, all ports are by default in VSAN1, and the speed and type are set to *autosense*.

Role-based management

The Cisco MDS 9000 Family switches support role-based management access with the CLI or the Cisco Fabric Manager. This lets you assign specific

management privileges to particular roles and then assign one or more users to each role.

Roles can also be assigned on a per-VSAN basis. For example, the administrator of one VSAN does not need to be given administrator privileges on other VSANs.

7.3.2 Optional licensed feature packages

In addition to the standard Fabric Manager features provided in the SAN-OS 3.x, there are five optional licensed feature packages to address different enterprise requirements. The five optional licensed packages are:

- FCIP Activation (one FCIP license included with each MDS 9216i):
  - SAN Extension over IP package for IPS-8 modules
  - SAN Extension over IP package for MPS-14/2 modules
- Enterprise package
- Fabric Manager Server package (FMS)
- Mainframe package
- Storage Services Enabler package

These packages have per-switch licensing, except for FCIP Activation, which is per-line-card licensing.

SAN Extension over IP package for SAN-OS 3.x

This package contains the following features and benefits:

- FCIP can be used to connect Fibre Channel across a distance using IP networks. Each Cisco MDS 9000 Family Gigabit Ethernet port is capable of managing up to three FCIP tunnels. Each 8-port IP Storage Services Module supports 24 simultaneous FCIP tunnels.
- FCIP compression in the Cisco MDS 9000 Family SAN-OS increases the effective WAN bandwidth. Gigabit Ethernet ports for IP Storage Services can theoretically achieve up to a thirty to one (30:1) compression ratio, but typical ratios of less than two to one (2:1) are more likely to be achieved in the field.
- IVR for FCIP allows selective transfer of data traffic between specific initiators and targets on different VSANs without merging VSANs into a single logical fabric. IVR can be used in conjunction with FCIP to increase the resiliency of SAN Extension over IP networks and create more efficient business continuity and disaster recovery solutions. To use IVR for Fibre Channel, the Enterprise package is also required.
- FCIP write acceleration can significantly improve application performance when storage traffic is routed over WANs using FCIP. When FCIP write acceleration is enabled, write I/O latency is decreased by minimizing the impact of WAN latencies.

- FCIP tape acceleration allows servers to transfer data across a WAN to streaming tape drives, which require a continuous flow of data to avoid write data underruns (dramatically reducing write throughput). Without FCIP tape acceleration, the effective WAN throughput for remote tape backup decreases exponentially as the WAN latency increases.
- Cisco SAN Extension Tuner helps optimize FCIP performance. The SAN Extension Tuner generates SCSI I/O commands that are directed to a specific virtual target. It reports I/Os per second and I/O latency results.

Note: The *SAN Extension over IP package* is usually referred to by IBM as *FCIP Activation*.

Enterprise package for SAN-OS 3.x

The Enterprise package optional license enables the following features:

- Enhanced security features
- Port security
- VSAN-based access control
- Fibre Channel Security Protocol (FC-SP) authentication
- Advanced traffic engineering—quality of service (QoS)
- IP Security Protocol (IPsec) for iSCSI and FCIP using the MPS-14/2 module or Cisco MDS 9216i Switch
- IKE digital certificates
- Extended credits using the MPS-14/2 module or the Cisco MDS 9216i Switch
- Enhanced VSAN routing—Inter-VSAN Routing (IVR) over FC
- IVR network address translation (NAT) over FC
- Zone-based traffic prioritizing
- Zone-based QoS
- Extended credits
- Fibre Channel write acceleration
- SCSI flow statistics
- FCIP encryption
- Fabric binding for Fibre Channel

Note: To enable FC-WA, you must also purchase the Storage Services Module and the Storage Services Enablement package.

Fabric Manager Server for SAN-OS 3.x

The standard Cisco Fabric Manager software that is included at no charge with the Cisco MDS 9000 Family Multilayer Switches provides basic switch configuration and troubleshooting capabilities. The Cisco MDS 9000 Family FMS package extends the standard Cisco Fabric Manager by providing historical performance monitoring for network traffic hotspot analysis, centralized management services, and advanced application integration.

The FMS license enables the following additional features:

- Fibre Channel Statistics Monitoring provides continuous performance statistics for Fibre Channel connections.
- Performance Thresholds enable the administrator to set two different event thresholds for each throughput statistic monitored by the Cisco FMS. Threshold values can be set with user-specified levels or with baseline values automatically calculated from performance history.
- Reporting and Graphing provides historical performance reports and graphs over daily, weekly, monthly, and yearly intervals for network hotspot analysis. Top 10 and daily summary reports for all ISLs, hosts, storage connections, and flows provide fabric-wide statistics.
- Intelligent Setup wizards are provided to quickly select information to monitor, set up flows, and estimate performance database storage requirements. Statistics are associated with host and storage devices, allowing physical connections to switches to be changed without losing historical statistics.
- Performance Database provides a compact round-robin database (RRD) maintained at a constant size by rolling up information to reduce the number of discrete samples for the oldest data points. Therefore, it requires no manual storage space maintenance.
- Web-Based Operational View provides a Web-browser interface to historical performance statistics, SAN inventory, and fabric event information needed for day-to-day operations.
- Multiple Fabrics Management allows for multiple Fibre Channel fabrics to be monitored by each management server.
- Continuous Health and Event Monitoring is enabled through SNMP traps and polling, instead of only when the application user interface is open.
- Common Discovery runs a centralized background discovery of Fibre Channel HBAs, storage devices, and switches.
- Roaming User Profiles allow user preference settings and topology map layout changes to be applied whenever the Cisco Fabric Manager client is opened. It maintains a consistent interface regardless of which computer is used for management.

- FMS Proxy Services help isolate a private IP network used for Cisco MDS management from the LAN or WAN used for remote connectivity.
- Cisco Traffic Analyzer Integration provides an easy drill down to SCSI I/O or Fibre Channel frame-level details.
- Management Server enables a server to be set up to continuously run Cisco FMA. Up to 16 remote Cisco Fabric Manager user interface clients can access this management server concurrently.

Mainframe package

The Cisco MDS 9000 Family Mainframe package is a collection of features required for using the Cisco MDS 9000 Family switches in mainframe storage networks. IBM Fibre Channel connection (FICON®) is an architecture for high-speed connectivity between mainframe systems and I/O devices. With the Mainframe package, the Cisco MDS 9000 Family has the capability to simultaneously support the FCP, iSCSI, FCIP, and FICON protocols.

Applying the Mainframe package optional license enables all FICON requirements with a single license key. The mainframe package optional license enables the following features:

- FICON Control Unit Port (CUP) for in-band management of the switch from FICON hosts
- The Fabric Binding feature to help ensure that ISLs are enabled only between switches that have been authorized in the fabric binding configuration

  This feature helps prevent unauthorized switches from joining the fabric or disrupting current fabric operations.
- The Switch Cascading feature to support FICON hosts accessing devices that are connected through ISLs
- VSAN support of FICON and FCP intermixed environments to provide separation of FCP and FICON traffic and to protect the mainframe environment from instability or excessive control traffic:
  - Qualified with IBM TotalStorage Virtual Tape Server (VTS) and IBM TotalStorage Peer-to-Peer Virtual Tape Server
  - Qualified with IBM TotalStorage Extended Remote Copy (XRC) for z/OS®
- FICON Native Mode and Native Mode Channel-to-Channel operation
- Persistent FICON Fibre Channel ID (FCID) assignment
- Port swapping for host-channel cable connections

Note: A license is required for each switch that participates in a FICON-cascaded fabric.

7.4 Key features

The following sections discuss some of the features of Cisco MDS 9000 Multilayer Switches, which are important in a routing environment.

7.4.1 Protocol support

The MDS 9000 Family supports the following protocols:

- FCP
- IP over Fibre Channel (RFC 2625)
- IPv6, IPv4, and ARP over FC (RFC 4338)
- Extensive IETF-standards based TCP/IP, SNMPv3, and remote monitoring (RMON) MIBs
- Class of Service: Class 2, Class 3, Class F
- Fibre Channel standard port types: E, F, FL, B
- Fibre Channel enhanced port types: SD, ST, TE
- IP
- IPSec
- Internet Key Exchange (IKE)

The MDS 9000 Family supports the following attachment types:

- FICON
- FCP
- FC_AL (including public and private loop support)
- FCIP over Gigabit Ethernet
- iSCSI over Gigabit Ethernet
- ISLs (attaching multiple switches or directors together)
- Interoperability (attachment to other vendors switches)

MDS 9000 Switches and Directors also support optional CWDM SFPs to provide aggregation of multiple links onto a single optical fiber through a passive optical mux.

The IP Services Modules and Multiprotocol Services Modules described in 7.2.7, “Optional modules” on page 217, provide the Gigabit Ethernet interfaces to enable iSCSI and FCIP capabilities for the 9200 and 9500 Families.

7.4.2 Supported port types

The Fibre Channel ports on all models of the MDS 9000 Family provide an auto-sensing 1, 2, or 4-Gbps SFP that use LC connectors. The operating port modes supported are described in the following sections.

Note: Not all port modules support 4 Gbps. Refer to 7.2.7, “Optional modules” on page 217.

Auto mode

Interfaces configured in the default auto mode are allowed to operate in either the fabric port (F_Port), fabric loop port (FL_Port), expansion (E_Port), or trunking E port (TE_Port) mode. The port mode is determined during interface initialization. For example, if the interface is connected to a node, server, or disk, it operates in F_Port or FL_Port mode depending on the N_Port or NL_Port mode. If the interface is attached to a third-party switch, it operates in E_Port mode. If the interface is attached to another MDS 9000 switch, it can become operational in TE_Port mode.

TL_Ports, SD_ports and ST_ports are not automatically determined during initialization and must be administratively configured.

E_Port

In E_Port mode, an interface functions as a fabric expansion port. This port can be connected to another E_Port to create an ISL between two switches. E_Ports carry frames between switches for configuration and fabric management. They serve as a conduit between switches for frames destined for remote N_Ports and NL_Ports. E_Ports support class 2, class 3, and class F services.

An E_Port connected to another switch can also be configured to form a PortChannel.

F_Port

In F_Port mode, an interface functions as a fabric port. This port can be connected to a node (server, disk, or tape) operating as an N_Port. An F_Port can be attached to only one N_Port. F_Ports support class 2 and class 3 services.

FL_Port

In FL_Port mode, an interface functions as a fabric loop port. This port can be connected to one or more NL_Ports (including FL_Ports in other switches) to form a public arbitrated loop. If more than one FL_Port is detected on the arbitrated loop during initialization, only one FL_Port becomes operational. The other FL_Ports enter a non-participating mode. FL_Ports support class 2 and class 3 services.

Fx_Port

Interfaces configured as Fx_Ports automatically negotiate operation in either F_Port or FL_Port mode. The mode is determined during interface initialization, depending on the attached N_Port or NL_Port. This administrative configuration disallows interfaces to operate in other modes, such as preventing an interface to connect to another switch.

TL_Port

In translative loop port (TL_Port) mode, an interface functions as a translative loop port. It might be connected to one or more private loop devices (NL_Ports). The TL_Port mode is specific to the Cisco MDS 9000 Family switches and has similar properties as FL_Ports. TL_Ports enable communication between private loop devices and one of the following target devices:

- A device attached to any switch on the fabric
- A device on a public loop anywhere in the fabric
- A device on a different private loop anywhere in the fabric
- A device on the same private loop

TL_Ports support class 2 and class 3 services.

TE_Port

In trunking E_Port (TE_Port) mode, an interface functions as a trunking expansion port. It connects to another TE_Port to create an EISL between two switches. TE_Ports are specific to the Cisco MDS 9000 Family. They expand the functionality of E_Ports to support these features:

- Multiple VSAN trunking
- Transport QoS parameters
- Fibre Channel trace (**`fctrace`**) feature

In TE_Port mode, all frames are transmitted in the EISL frame format, which contains VSAN information. Interconnected switches use the VSAN ID to multiplex traffic from one or more VSANs across the same physical link. This feature is referred to as *trunking* in the Cisco MDS 9000 Family.

TE_Ports support class 2, class 3, and class F services.

SD_Port

In switch port analyzer (SPAN) destination port (SD_Port) mode, an interface functions as a SPAN. The SPAN feature is specific to switches in the Cisco MDS 9000 Family. It monitors network traffic passing though a Fibre Channel interface. This monitoring is done using a standard Fibre Channel analyzer, or similar switch probe, that is attached to an SD_Port. SD_Ports do not receive frames. They merely transmit a copy of the source traffic. The SPAN feature is

nonintrusive and does not affect switching of network traffic for any SPAN source ports.

ST_Port

Interfaces configured as ST_Ports serve as an entry point port in the source switch for a Fibre Channel tunnel. ST_Ports are specific to remote SPAN (RSPAN) ports and cannot be used for normal Fibre Channel traffic.

Figure 7-19 shows an example of the port types that are available with the Cisco MDS 9000 Family of products.

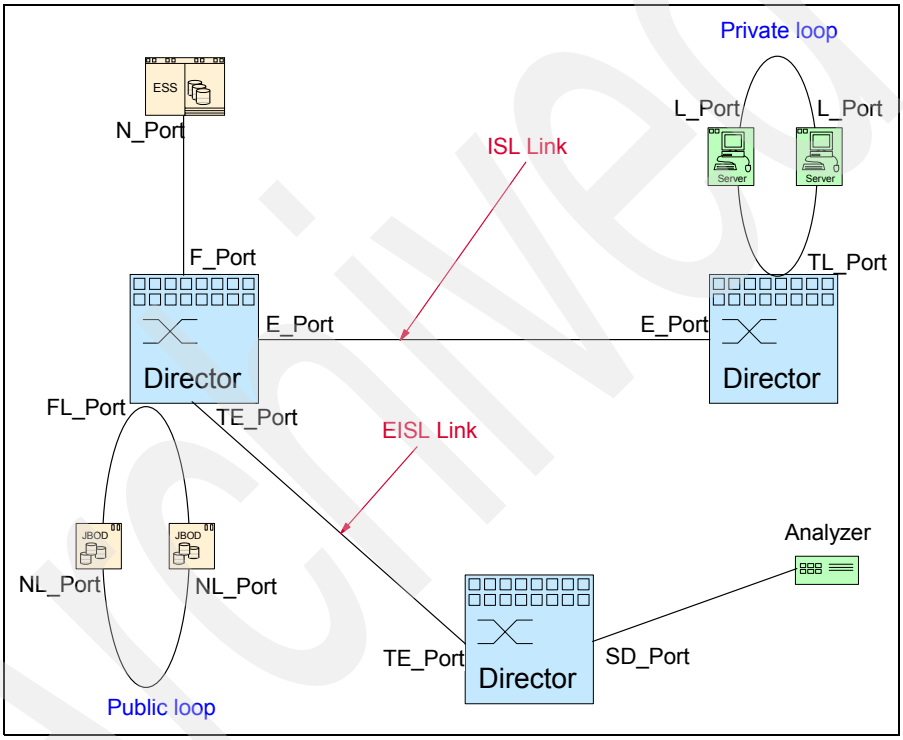


Figure 7-19 Cisco MDS 9000 Family port types

You can change the port mode for any given port through Device Manager, as shown in Figure 7-20 and Figure 7-21 on page 237.

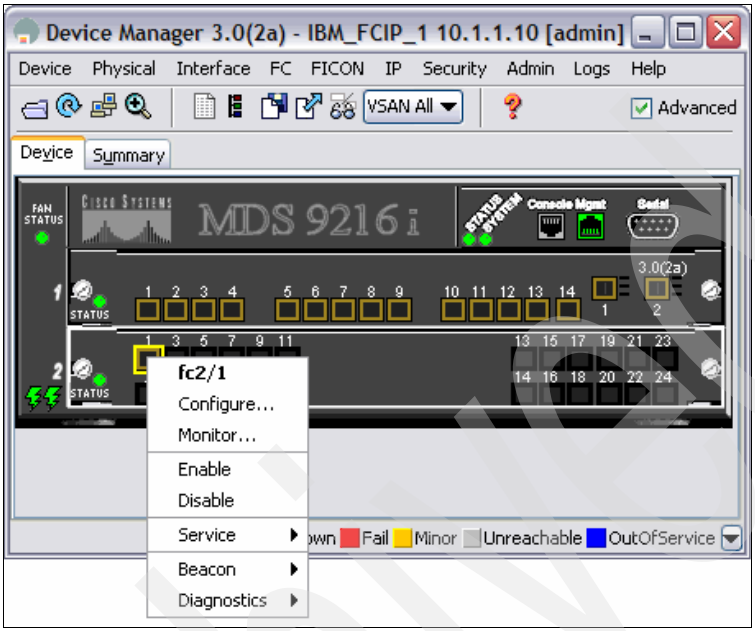


Figure 7-20 Port configuration selection

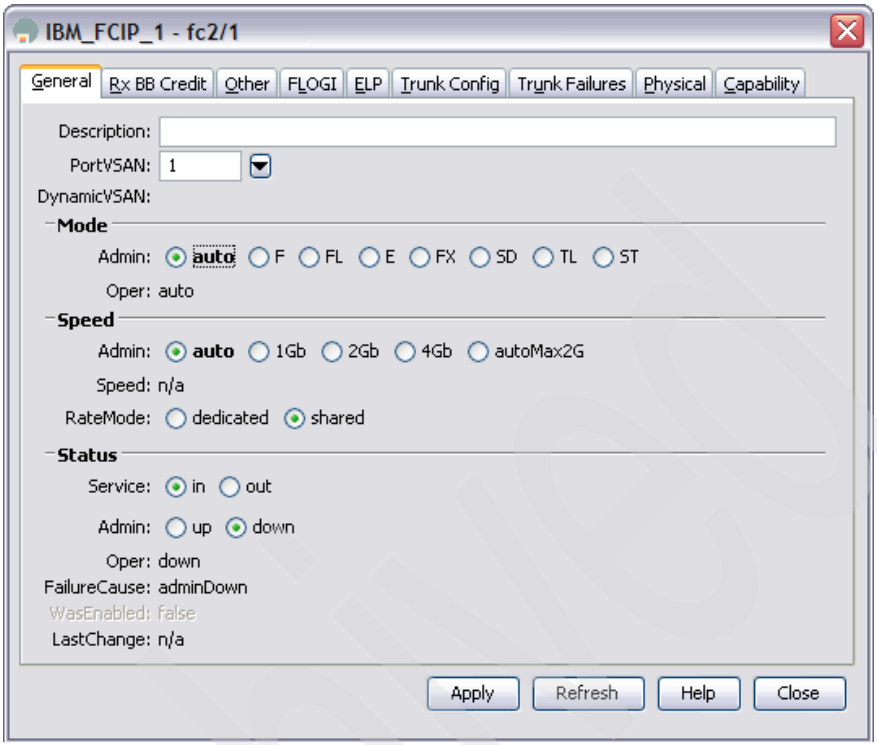


Figure 7-21 DM port configuration window

7.4.3 VSAN

VSAN technology allows virtual fabrics (enabled by a mixture of ASIC functionality and software functionality) to be overlaid on a physical fabric. Cisco's approach is to position VSAN not as a single feature, but as an architectural approach that allows flexible delivery of many features.

A given device can belong only to one VSAN. Each VSAN contains its own zoning, fabric services, and management capabilities, as though the VSAN were configured as a separate physical fabric.

VSANs offer the following features:

- Ease of configuration is enhanced because devices can be added or removed from a VSAN fabric without making any physical changes to cabling.
- Traffic is isolated to within a VSAN, unless IVR is implemented. Separate companies or divisions of a company can be segregated from each other without needing separate physical fabrics.

- Fabric services are provided separately to each VSAN. Smaller fabrics are simpler and generate fewer registered state change notifications (RSCNs) between switches. Each VSAN runs all required protocols such as Fabric Shortest Path First (FSPF), domain manager, and zoning.
- Redundancy can be configured, for example, on a dual HBA server by having each HBA in a separate VSAN. This is the same as you would typically have each HBA in a separate physical fabric if you did not have VSANs.
- Duplicate FCIDs can be accommodated on a network, provided the devices are in separate VSANs. This allows for the IVR connection of fabrics that were previously completely separate.

Figure 7-22 represents a typical SAN environment that has a number of servers, each with multiple paths to the SAN. In this case, the SAN consists of a Fibre Channel director attached to a disk and tape subsystem.

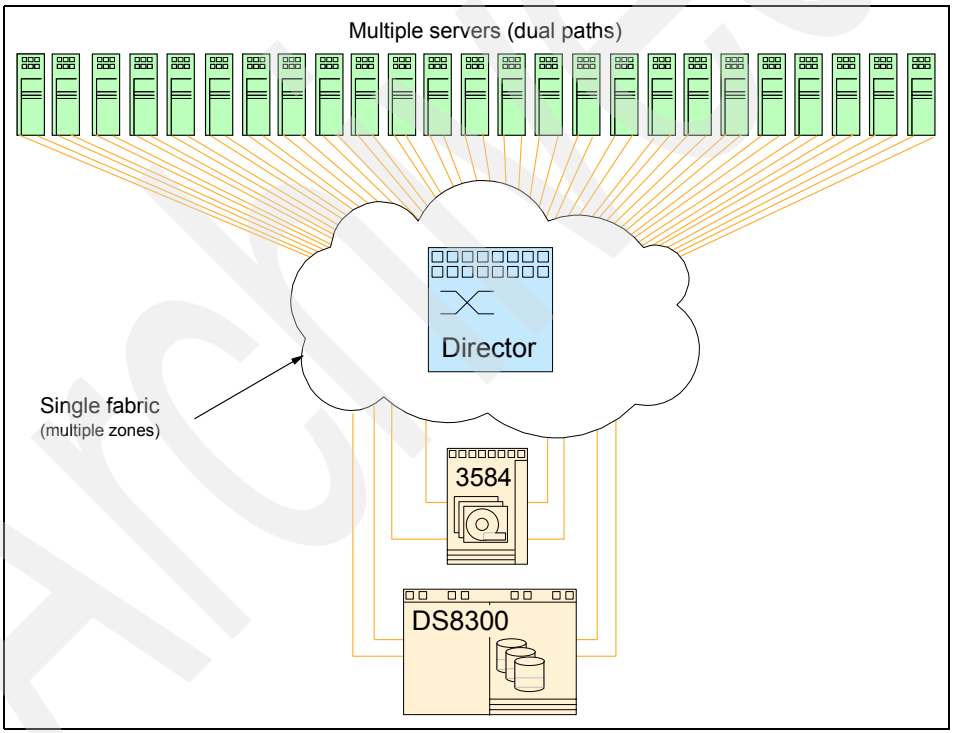


Figure 7-22 Traditional SAN

Figure 7-23 shows how the same scenario is implemented using the Cisco VSAN.

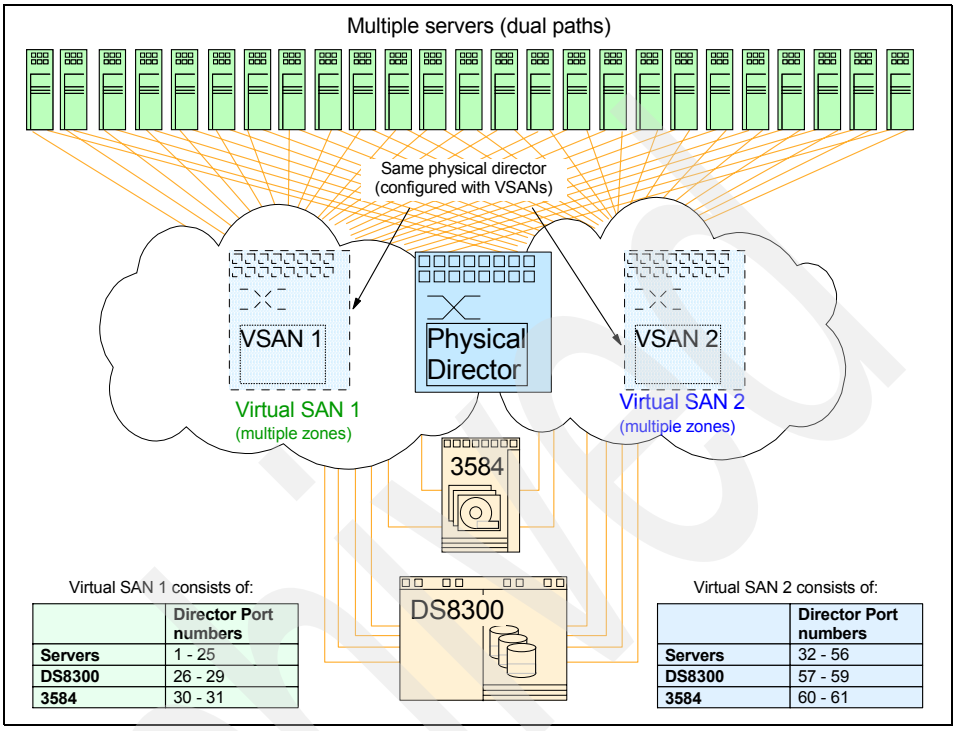


Figure 7-23 Virtual SAN

In this example, the servers are still connected to the SAN, but the SAN consists of a single 9509 attached to the same disk and tape subsystems. In this case, we configured the first 31 ports in the director into a VSAN called Virtual SAN 1, and the second 31 ports into another VSAN called Virtual SAN 2. The servers have a connection to each VSAN, thereby providing a solution that consists of multiple SAN fabrics.

The VSANs cannot communicate with each other, unless IVR is implemented. They appear to be totally separate fabrics. They have their own FSPF tables, domain manager, and zoning requirements. Any traffic disruption in one VSAN does not impact the other VSAN. A port cannot belong in multiple VSANs.

Note: A new feature in SAN-OS 2.x and later is support for WWN-based VSANs.

VSANs compared to zones

Table 7-2 shows the main differences between a zone and a VSAN.

Table 7-2 VSANs compared to zones

VSANs	Zones
A VSAN is a logical fabric with its own routing, naming, and zoning protocols.	A zone is a logical group of ports or WWNs that are allowed to talk to each other.
VSANs can contain multiple zones.	Zones are always contained within a VSAN. They cannot span a VSAN.
VSANs limit the reach of fabric services transmissions.	Zones limit the reach of I/O transmissions.
Membership is defined using a VSAN ID to Fx ports or WWN.	Membership is typically defined using WWN or port number (Fx).
HBAs can belong only to a single VSAN, the VSAN associated with the Fx port.	HBAs can belong in multiple zones.
VSANs enforce membership at each E_Port, source port, and destination port.	Zones enforce membership only at the source and destination ports.

Registered state change notifications

The registered state change notification service propagates information about a change in state of one node to all other nodes in the fabric. In the event of a device shutting down, for example, the other devices on the SAN are informed and then know not to send data to the shutdown device, avoiding timeouts and retries.

There are two types of RSCNs. Switch RSCNs (SW_RSCN) are passed from one switch to another, for example, when a new device comes online, and the local switch needs to inform the other switches. SW_RSCNs are sent at a VSAN level (within a VSAN). The second type of RSCN is issued by the switch to an end device that informs it of a change within a zone to which the end device belongs. This type of RSCN is sent to only those devices in the affected zone.

Default and isolated VSANs

Up to 4093 VSANs can be configured on a physical SAN. Of these, one is the default VSAN (VSAN 1) and another is an isolated VSAN (VSAN 4094). User-specified VSAN IDs range from 2 to 4093.

Default VSAN

The factory settings for switches in the Cisco MDS 9000 Family have only default VSAN 1 enabled. If you do not need more than one VSAN for a switch, use this

default VSAN as the implicit parameter during configuration. If no VSANs are configured, all devices in the fabric are considered part of the default VSAN. By default, all ports are assigned to the default VSAN.

Isolated VSANs

VSAN 4094 is an isolated VSAN. All non-trunking ports are transferred to this VSAN when the VSAN to which they belong is deleted. This avoids an implicit transfer of ports to the default VSAN or to another configured VSAN. All ports in the deleted VSAN are disabled.

VSAN membership

Port VSAN membership on the switch is assigned on a port-by-port basis. By default, each port belongs to the default VSAN. Trunking ports have an associated list of VSANs that are part of an allowed list.

VSAN attributes

VSANs have the following attributes:

- The *VSAN ID* identifies the VSAN as the default VSAN (VSAN 1), user-defined VSANs (VSAN 2 to 4093), and the isolated VSAN (VSAN 4094).
- The *administrative state* of a VSAN can be configured to an active (default) or suspended state. When VSANs are created, they can exist in various conditions or states.
  - The *active state* of a VSAN indicates that the VSAN is configured and enabled. By enabling a VSAN, you activate the services for that VSAN.
  - The *suspended state* of a VSAN indicates that the VSAN is configured but not enabled. If a port is configured in this VSAN, it is disabled. Use this state to deactivate a VSAN without losing the VSAN's configuration. All ports in a suspended VSAN are disabled. By suspending a VSAN, you can preconfigure all the VSAN parameters for the whole fabric and activate the VSAN immediately.
- The *VSAN name* text string identifies the VSAN for management purposes. The name can be from 1 to 32 characters long and it must be unique across all VSANs. By default, the VSAN name is a concatenation of VSAN and a four-digit string representing the VSAN ID. For example, the default name for VSAN 3 is VSAN0003.
- *Load balancing* attributes indicate the use of the source-destination ID (src-dst-id) or the originator exchange OXID (src-dst-ox-id, the default) for load balancing path selection.

Operational state of a VSAN

A VSAN is in the operational state if the VSAN is active and at least one port is up. This state indicates that traffic can pass through this VSAN. This state cannot be configured.

Deleted VSAN

When an active VSAN is deleted, all of its attributes are removed from the running configuration.

VSAN-related information is maintained by the system software:

- VSAN attributes and port membership details are maintained by VSAN manager. This feature is affected when you delete a VSAN from the configuration. When a VSAN is deleted, all the ports in that VSAN are made inactive and the ports are moved to the isolated VSAN. If the same VSAN is re-created, the ports are not automatically assigned to that VSAN. You must *explicitly* reconfigure the port VSAN membership.
- VSAN-based runtime (name server), zoning, and configuration (static route) information is removed when the VSAN is deleted.
- Configured VSAN interface information is removed when the VSAN is deleted.

7.4.4 Inter-VSAN Routing

IVR is available on the MDS9216i as a standard feature and is available on all other platforms by purchasing the Enterprise package license (minimum v1.3(2a) SAN-OS or later is required). IVR helps to allow data traffic to flow between VSANs while maintaining the VSAN segregation because no management data is passed. This proves useful, for example, when a host defined in one VSAN is required to have access to a tape drive defined in another VSAN. This feature reduces the amount of required hardware to meet the needs for multiple systems.

An IVR is defined in a similar manner to normal zoning within a VSAN. Instead of working within a VSAN and performing the zoning definitions, we work from the IVR group to create an IVR zone set, which can be activated or deactivated without affecting the VSANs.

Figure 7-24 shows how the same scenario is implemented using the Cisco IVR.

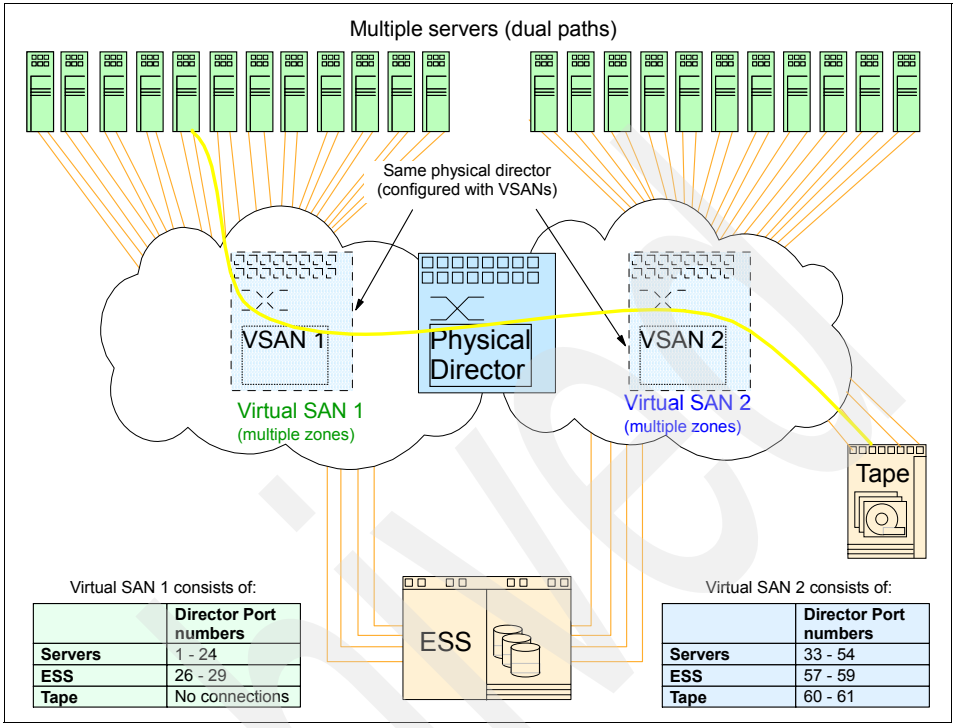


Figure 7-24 Inter-VSAN Routing

In this example, two groups of servers are connected to different VSANs within a single MDS 9509 Director. The disk has paths that are defined to both VSAN 1 and VSAN 2, although the requirement is for all servers to access all tape drives. In this case, we configured the first 29 ports in the director into VSAN 1 and the second 31 ports into VSAN 2.

The VSANs cannot communicate with each other, and they appear to be totally separate SANs. By defining an IVR zone, we can allow a data connection from a server in VSAN 1 through to a tape drive in VSAN 2. No management data is passed over this connection, and any disruptions in one VSAN do not have any impact on the other VSAN.

Tip: Use VSANs on an exception basis. For example, use them for multivendor switch interoperability, to isolate separate companies in a shared services environment, to manage QoS between test and production environments (using the VSAN-based QoS feature), and to isolate less reliable FCIP links from disrupting the main fabric.

If you have a lot of IVRs in your design, you might have too many VSANs. If you consider a LAN analogy, you would not install several routers into the middle of your corporate LAN.

7.4.5 PortChanneling

PortChanneling is the Cisco term for exchange-based load balancing across multiple ISLs. An exchange is usually a single SCSI command and the response it evokes, so it is of fairly short duration (milliseconds or seconds). However, exchanges can be longer in a FICON environment because FICON improves efficiency by retaining the exchange ID for multiple commands.

PortChanneling can also be implemented based on source ID (that is, a server HBA port) and destination ID (for example, a disk system HBA port) pairs, which give less granular load balancing but provide some traffic isolation if that is preferred.

PortChanneling does not do load balancing at a frame level. Frame-based load balancing requires additional out-of-order frame management intelligence.

With PortChannels, users can aggregate up to 16 physical ISLs into a single load-balanced bundle. Links can span any port on any module within a chassis for added scalability and resilience. These PortChannels support the following functions:

- Increase the aggregate bandwidth on an ISL or EISL by distributing traffic among all functional links in the channel
- Load balance across multiple links and maintain optimum bandwidth utilization

  Load balancing is based on a source ID (SID), destination ID (DID), and, optionally, the originator exchange ID (OXID) that identify the flow of the frame.

- Provide high availability on an ISL

  If one link fails, traffic previously carried on this link is switched to the remaining links. If a link goes down in a PortChannel, the upper protocol is not aware of it. To the upper protocol, the link is still there, although the bandwidth is diminished. The routing tables are not affected by a link failure.

Figure 7-25 shows ISLs and PortChanneling.

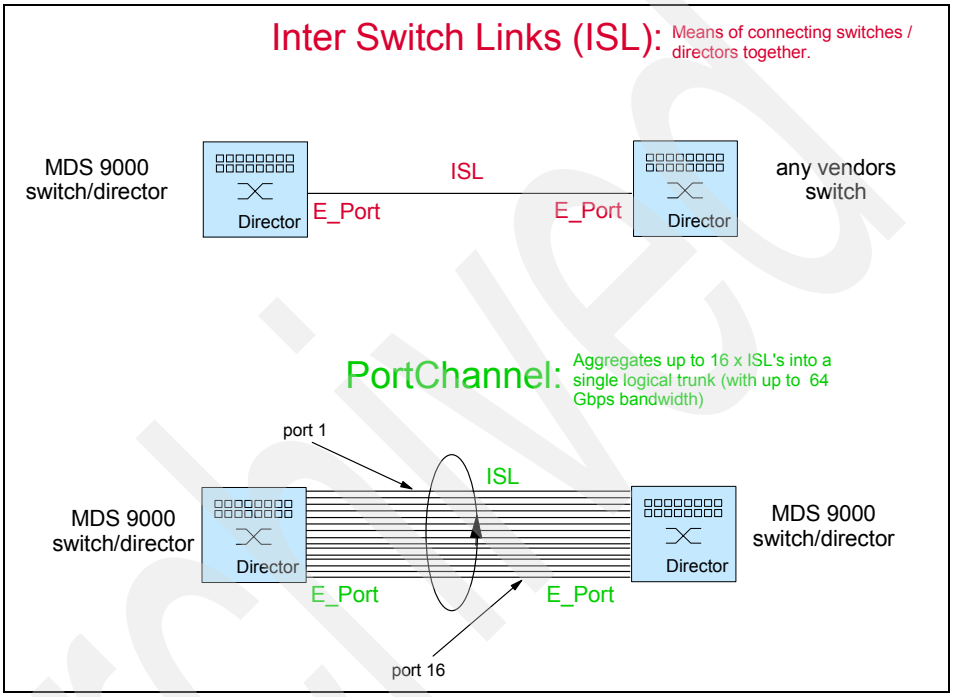


Figure 7-25 PortChannels and ISLs on the Cisco MDS 9000 switches

7.4.6 Trunking

The Cisco MDS 9000 Family uses the term *trunking* to refer to an ISL that carries one or more VSANs. Trunking ports receive and transmit EISL frames. EISL frames carry an EISL header containing the VSAN information. When EISL is enabled on an E_Port, that port becomes a TE_Port.

Trunking is also referred to as *VSAN trunking*, because it applies only to a VSAN. If a trunking-enabled E_Port is connected to another vendor's switch, the trunking protocol ensures that the port will operate as a standard E_Port.

Figure 7-26 shows a diagram of trunking. It also demonstrates how you can use a combination of PortChannels and trunking to create an aggregate bandwidth of up to 64 Gbps between switches.

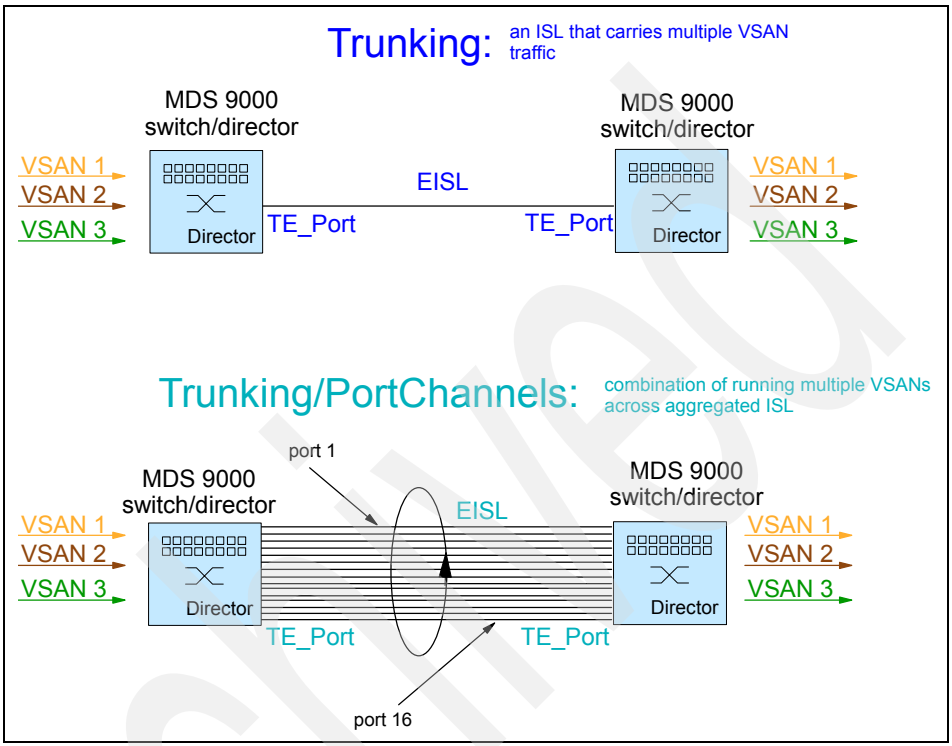


Figure 7-26 Trunking and PortChanneling

7.4.7 Quality of service

Four distinct quality of service (QoS) priority levels are available: three for Fibre Channel data traffic and one for Fibre Channel control traffic. Fibre Channel data traffic for latency-sensitive applications can be configured to receive higher priority than throughput-intensive applications using data QoS priority levels. Control traffic is assigned the highest QoS priority automatically to accelerate convergence of fabric-wide protocols such as FSPF, zone merges, and principal switch selection.

Data traffic can be classified for QoS by the VSAN identifier, zones, N-port WWN, or FCID. Zone-based QoS helps simplify configuration and administration by using the familiar zoning concept.

Note: Zone-based QoS (introduced in SAN-OS 2.1) also requires the SAN-OS Enterprise package.

QoS offers the following primary advantage:

- Prioritization of traffic

The Cisco MDS 9000 Family supports QoS for internally and externally generated control traffic. Within a switch, the control traffic is sourced to the supervisor module and is treated as a high-priority frame. A high-priority status provides absolute priority over all other traffic and is assigned in the following cases:

- Internally-generated, time-critical control traffic (generally Class F frames)
- Externally-generated, time-critical control traffic entering a switch in the MDS 9000 range from another vendor's switch

  High priority frames originating from other vendor switches retain the priority as they enter a switch in the MDS 9000 Family.

By default, the QoS feature for control traffic is enabled but can be disabled if required.

7.4.8 Fibre Channel Congestion Control

Fibre Channel Congestion Control (FCC) is a Cisco proprietary flow control mechanism that alleviates congestion on Fibre Channel networks. A switch experiencing congestion signals this condition to the upstream (source) switch, which throttles the traffic.

By default, the FCC protocol is disabled. You can enable the protocol globally for all the VSANs configured in the switch, or selectively enable or disable it for each VSAN.

Congestion control methods

With FCC enabled, there are different congestion control methods:

- *Path quench control* reduces severe congestion temporarily by slowing the source to the whole path in the fabric.
- *Edge quench control* provides feedback to the source about the rate at which frames should be entered into the network (frame intervals).

FCC process

When a node in the network detects congestion for an output port, it generates an edge or a path quench message. These frames are identified by the Fibre Channel destination ID and the source ID.

Any receiving switch in the Cisco MDS 9000 Family handles frames in one of the following ways:

- It forwards the frame.
- It limits the rate of the frame flow in the congested port.

Behavior of the flow control mechanism differs, based on the Fibre Channel DID:

- If the Fibre Channel DID is directly connected to one of the switch ports, the input rate limit is applied to that port.
- If the destination of the edge quench frame is a Cisco domain or the next hop is a Cisco MDS 9000 Family switch, the frame is forwarded.
- If neither of these conditions is true, the frame is processed in the port going toward the Fibre Channel DID.

All switches, including the edge switch, along the congested path process path quench frames. However, only the edge switch processes edge quench frames. The FCC protocol is implemented for each VSAN and can be enabled or disabled on a specified VSAN or for all VSANs at the same time.

Note: Cisco's FCC differs from standard buffer-to-buffer flow control. FCC looks at the source of congestion and passes messages upstream to report it to the nearest switch to the source so that it can apply selective buffer-to-buffer quenching as appropriate. Standard buffer-to-buffer flow control simply delivers point-to-point flow control.

If you enable FCC on one switch, be sure to enable it on all switches in the fabric.

Default QoS settings

Table 7-3 lists the default configuration for the QoS settings.

Table 7-3 Default QoS settings

Parameters	Default
FCC protocol	Disabled
QoS control traffic	Enabled
QoS data traffic	Enabled

Parameters	Default
Zone-based QoS priority	Low
Rate limit	100%

7.4.9 Switch port analyzer

The Cisco MDS 9000 Family provides a feature called the *switch port analyzer*. As mentioned in 7.4.2, “Supported port types” on page 232, the SPAN or SD_Ports enables you to monitor network traffic through the Fibre Channel interface.

Traffic through any Fibre Channel interface can be replicated to a special port called the *SPAN destination port*. Any Fibre Channel port in a switch can be configured as an SD_Port. When an interface is in SD_Port mode, it cannot be used for normal data traffic. You can attach a Fibre Channel analyzer to the SD_Port to monitor SPAN traffic.

SD_Ports do not receive frames. They only transmit a copy of the SPAN source traffic. The SPAN feature is non-intrusive and does not affect switching of network traffic for any SPAN source port.

SPAN sources

A SPAN source is the interface from which traffic can be monitored. You can also specify a VSAN as a SPAN source; in which case, all supported interfaces in the specified VSAN are included as SPAN sources. You can choose the SPAN traffic in the ingress direction, the egress direction, or both directions, for any source interface:

- Ingress source (rx)

  Traffic entering the switch fabric through this source is spanned or copied to the SD_Port, as shown in Figure 7-27 on page 250.

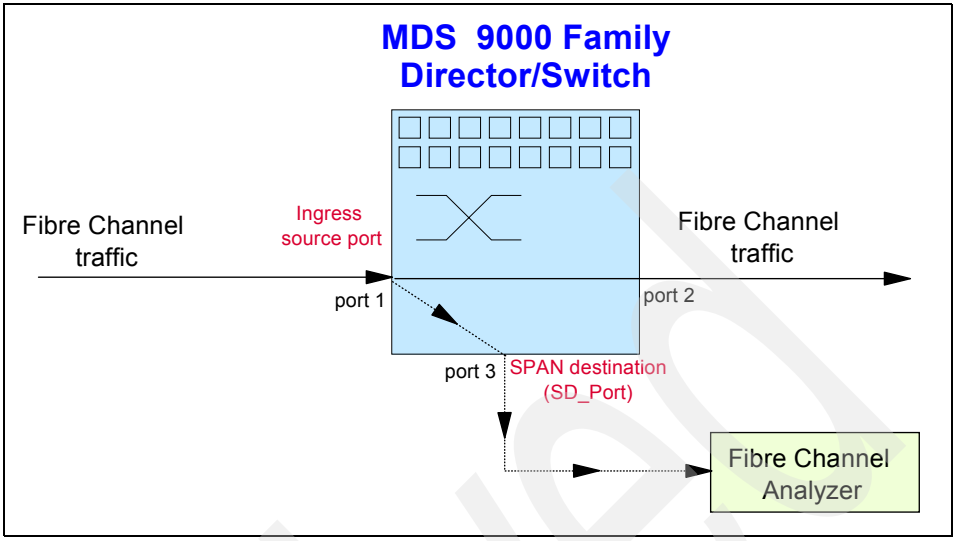


Figure 7-27 SD_Port for ingress (incoming) traffic

- Egress source (tx)

  Traffic exiting the switch fabric through this source interface is spanned or copied to the SD_Port, as shown in Figure 7-28.

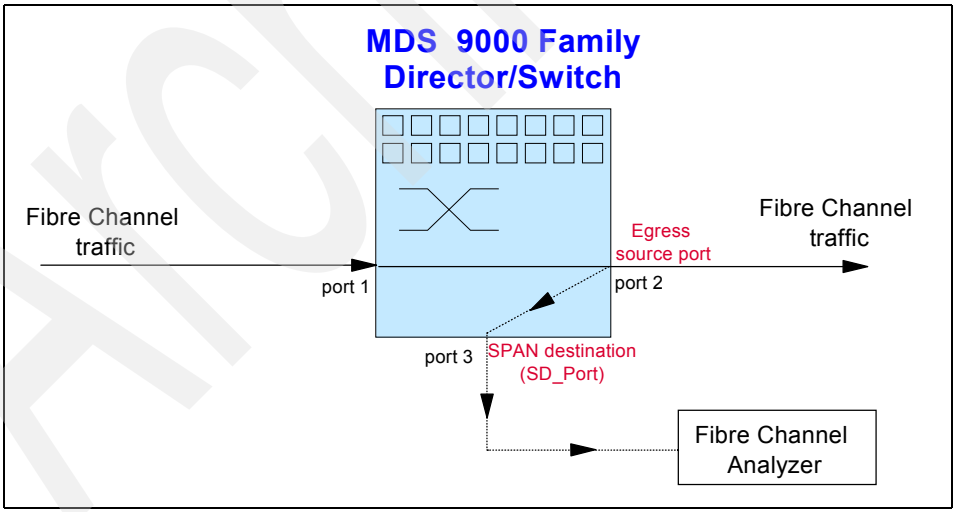


Figure 7-28 SD_Port for egress (outgoing) traffic

Allowed source interface types

The SPAN feature is available for the following interface types:

- Physical ports: F_Ports, FL_Ports, TE_Ports, E_Ports, and TL_Ports
- Interface sup-fc0 (traffic to and from the supervisor)
  - Fibre Channel traffic from the supervisor module to the switch fabric, through the sup-fc0 interface, is called *ingress traffic*. It is spanned when sup-fc0 is chosen as an ingress source port.
  - Fibre Channel traffic from the switch fabric to the supervisor module, through the sup-fc0 interface, is called *egress traffic*. It is spanned when sup-fc0 is chosen as an egress source port.
- PortChannels
  - All ports in the PortChannel are included and spanned as sources.
  - You cannot specify individual ports in a PortChannel as SPAN sources. Previously-configured, SPAN-specific interface information is discarded.

VSAN as a SPAN source

When a VSAN as a source is specified, all physical ports and PortChannels in that VSAN are included as SPAN sources. A TE_Port is included only when the port VSAN of the TE_Port matches the source VSAN. A TE_Port is excluded even if the configured allowed VSAN list might have the source VSAN, but the port VSAN is different.

Guidelines for configuring VSANs as a source

The following guidelines apply when configuring VSANs as a source:

- Traffic on all interfaces included in a source VSAN is spanned only in the ingress direction.
- When a VSAN is specified as a source, you will not be able to perform interface-level configuration on the interfaces that are included in the VSAN. Previously-configured, SPAN-specific interface information is discarded.
- If an interface in a VSAN is configured as a SPAN source, you will not be able to configure that VSAN as a source. You must first remove the existing SPAN configurations on such interfaces before configuring the VSAN as a source.
- Interfaces are only included as sources when the port VSAN matches the source VSAN.

7.5 Interoperability

The Cisco offering on interoperability includes support for Inter-VSAN Routing to Brocade, McDATA, and QLogic switch products through an RPQ.

7.5.1 Switch interoperability modes

Switch interoperability modes enable other vendors' switches to connect to the MDS Family. These modes are required because vendors often implement features on their switches that are not compatible with other manufacturers. Table 7-4 lists the modes.

Table 7-4 Cisco interoperability modes

Mode	Will interoperate with
Mode 1	Standards-based interop mode that requires all other vendors in the fabric to be in interop mode
Mode 2	Brocade native mode (Core PID 0)
Mode 3	Brocade native mode (Core PID 1)
Mode 4	McDATA native mode.

Note: VSANs are a valuable tool in managing interoperability between multivendor switches. Heterogeneous IVR was introduced in SAN-OS 2.1

The switch interoperability mode might disable a number of advanced or proprietary features, so it is worth understanding what these might affect before proceeding. Table 7-5 lists the changes required if interoperability mode is enabled on a Cisco MDS 9000 Family switch or director.

Table 7-5 Interoperability mode changes

Switch feature	Changes if interoperability mode is enabled
Domain IDs	While Cisco implements the full standard specification 239 domain IDs (switch IDs), McDATA supports only 31 domain IDs within a fabric when using midrange switches. Therefore, Interop domain IDs are restricted to the range 97 to 127 to accommodate all vendor implementations.

Switch feature	Changes if interoperability mode is enabled
Timers	All Fibre Channel timers must be set to the same value on all switches because these values are exchanged by E_Ports when establishing an ISL. The Time-Out Value timers are described in the following rows. (This is always required for connection of two switches)
F_S_TOV	Verify that the Fabric Stability Time-Out Value timers match exactly.
D_S_TOV	Verify that the Distributed Services Time-Out Value timers match exactly.
E_D_TOV	Verify that the Error Detect Time-Out Value timers match exactly.
R_A_TOV	Verify that the Resource Allocation Time-Out Value timers match exactly.
Trunking	Trunking is not supported between two different vendors' switches. This feature can be disabled on a per port basis.
Default zone	The default zone behavior of permit (all nodes can see other nodes) or deny (all nodes are isolated when not explicitly placed in a zone) might change.
Zoning attributes	Zones can be limited to the WWPN, and other proprietary zoning methods (physical port number) can be eliminated.
Zone propagation	Some vendors do not pass the full zone configuration to other switches, only the active zoneset gets passed.
VSAN	This only affects the specified VSAN.
TE_Ports and PortChannels	TE_Ports and PortChannels only apply when connecting from one MDS 9000 to another. Only E_Ports can be used to connect to non-MDS switches. TE_Ports and PortChannels can be used to connect to other MDS 9000 switches when interoperability mode is enabled.
Domain reconfiguration disruptive	This can require the entire switch to be restarted when changing the domain IDs.
Domain configuration nondisruptive	This only impacts the affected VSAN. Only the domain manager for the affected VSAN is restarted. Other VSANs are unaffected.
Name server	Verify that all vendors have the correct values in their respective name server tables.
IVR	IVR-enabled VSANs can be configured in no interop (default) mode or in any of the interop modes.

Interoperability mode in the Cisco MDS 9000 Family can be enabled disruptively or nondisruptively, but the default is to have this mode turned off.

It is still important to check with the OEM vendors involved in regard to the specific steps that must be taken.

Tip: The Brocade **msplmgmtdeactivate** command must explicitly be run prior to connecting from a Brocade switch to either Cisco MDS 9000 Family switches or to McDATA switches. This command uses Brocade proprietary frames to exchange platform information, which Cisco MDS 9000 Family switches or McDATA switches do not understand. Rejecting these frames causes the common E_Ports to become isolated.

7.5.2 Interoperability matrix

You can download the latest IBM interoperability matrixes for Cisco MDS switches from:

`ftp://service.boulder.ibm.com/storage/san/cisco/`

The interoperability matrixes include a list of servers, disk, and tape systems that have been tested and are supported with the MDS Family of switches and directors. These lists also contain supported operating system versions and links to other Web sites that document the required HBA levels.

For combinations of technologies that are not explicitly supported, contact your local IBM office or IBM Business Partner to discuss submitting an RPQ.

8

Cisco routing solutions

Routers provide access to data that is located in a different fabric. The principal uses for storage routing are:

- Storage area network (SAN) extension over Internet Protocol (IP) networks, not the 9120 and 9140
- Lowering connection costs using Small Computer System Interface over IP (iSCSI), not the 9120 and 9140
- Achieving isolation and interoperability between different business units
- Managing scalability as your SAN environment grows
- Migrating from an older storage environment to a newer one

The latter three features are inherent to the entire Cisco MDS 9000 product family. The former two uses apply to all members of the family except for the MDS 9120 and the MDS 9140 because they do not have Ethernet ports.

8.1 SAN extension with FCIP

Fibre Channel (FC) distances have traditionally been limited to either local fiber runs using 9 micron longwave Fibre Channel, or high quality wide area networks (WANs), such as SONET and SDH, in combination with coarse wavelength division multiplexing (CWDM) or dense wavelength division multiplexing (DWDM) multiplexers.

The advent of Fibre Channel over IP (FCIP) has meant that applications that can tolerate the higher latencies of IP networks can now make Fibre Channel connections across standard corporate IP WANs. The advantages of this are that it uses a widely available and well-understood infrastructure. This translates into lower cost.

We are still in a phase where people want FCIP over standard networks to be a panacea for all SAN extension applications. The inherent latencies involved are around five microseconds per kilometer (km) travelled in each direction and added latencies at every step (for example, up to 100 microseconds per router or firewall). This prevents FCIP from being used effectively for applications such as long-distance synchronous replication or online transaction processing (OLTP). For example, a high-quality network of 1000 km might have a latency of around 20 milliseconds. Given that a disk input/output (I/O) might only take 10 milliseconds, the problem with a 20 millisecond latency becomes obvious.

Because some corporate WANs provide uncertain quality of service, storage router vendors tend to be cautious about quoting distances for FCIP, and generally recommend that high quality WANs are necessary to provide services over anything more than 200 or 300 km.

In practice, the most common uses for FCIP are long-haul asynchronous replication, short-haul synchronous replication, and logical unit number (LUN) access over campus or metro distances. A client might choose to implement Fibre Channel tunneled into IP on a campus scale simply because the IP links are already in place. On a short IP network, the main problem becomes quality of service (QoS) because the latencies are not so large. The principles are the same whether running over 500 meters (m) or 5000 km. All that varies is the link latency and the service reliability and consistency.

8.1.1 Compression

FCIP compression in the Cisco MDS 9000 Family SAN-OS increases the effective WAN bandwidth. Although Gigabit Ethernet ports for IP Storage Services can theoretically achieve up to a thirty to one (30:1) compression ratio, typical ratios in the field are less than two to one (2:1).

Hardware-based compression is also available on the 14+2 line card, in addition to the MDS 9216i integrated 14+2 controller.

When software compression is turned on with the IP Services (IPS) line card and set to modes 2 or 3, the IPS runs small Fibre Channel frames together to use up space inside the IP packets rather than sending individual frames separately. This is only for use over WAN distances and is not recommended for cross-campus use.

Using jumbo packets can also improve throughput. Remember that jumbo packets need to be turned on through the entire data path.

Note: The Cisco SAN Extension Tuner helps to optimize FCIP performance. The tuner generates SCSI I/O commands that are directed to a specific virtual target. It reports I/Os per second and I/O latency results. SAN Extension Tuner is included with the FCIP Enablement license package.

8.1.2 Using Inter-VSAN Routing with FCIP

The stability of WAN links varies by geography and provider. It is usually important to separate FCIP links running over WANs from your core Fibre Channel network. Because we want to protect both ends of the core fabric from FCIP link bounce (when the network can go up and down a few times in quick succession), we create transit virtual SANs (VSANs) between the two switches and then implement Inter-VSAN Routing (IVR). Transit VSANs consist solely of the FCIP ports on the switches.

Figure 8-1 shows an example of an asynchronous replication running over IP at a 500 km distance.

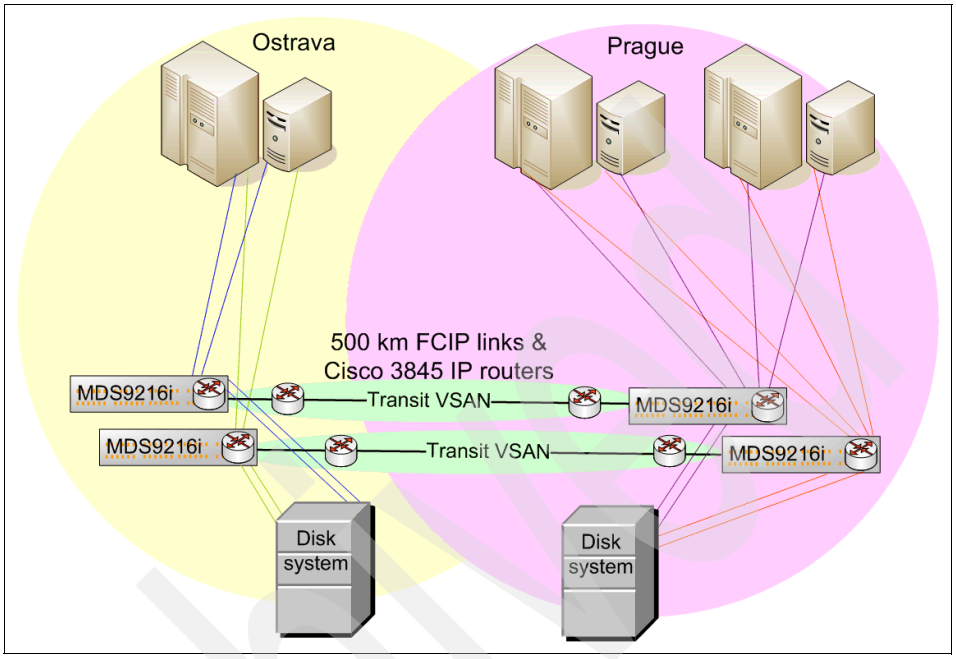


Figure 8-1 Tunneling FCIP using transit VSANs to mitigate link bounce

8.1.3 Using FCIP write acceleration

FCIP write acceleration (FCIP-WA) is an attempt to avoid the transport latency associated with long-distance SCSI I/O operations. The IPS module allows the initiator to send the data to be written before the write command is processed by the remote target and a SCSI Transfer Ready message has time to travel back to start the data transfer in the traditional way. Essentially, FCIP-WA spoofs the R_RDY.

FCIP-WA performance depends on the traffic profile of the SCSI operations being performed. The following I/O characteristics can benefit from FCIP-WA:

- Long-distance, high-latency network
- Write intensive I/O
- High number of small SCSI writes (rather than low number of large writes)
- Disk system with low write latency

This suggests that the best use for FCIP-WA is in disk system replication, but results in practice are not always worth the added complexity. Across a 100 km link, replication using FCIP-WA can be expected to deliver around 10% improvement in throughput and a similar percentage reduction in latency on a given FCIP network.

For further information about write acceleration, refer to the Cisco white paper at the following URL:

http://www.cisco.com/en/US/products/hw/ps4159/ps4358/products_white_paper0900aecd8024fd2b.shtml

Note: FCIP-WA is different from Fibre Channel write acceleration (FC-WA), which applies to FC-FC links without FCIP. FC-WA requires the Storage Services Module, but FCIP-WA does not.

8.1.4 Using Fibre Channel tape acceleration with FCIP

Fibre Channel tape acceleration is similar to FC-WA. However, where FC-WA assumes R_RDY but does not assume data is received until it receives acknowledgment from the target, tape acceleration goes one step further. Tape acceleration does not wait for R_RDY nor for acknowledgement that data has been received before sending the next packet of data. Tape acceleration leaves the target tape in an uncertain state in the event of a link failure.

If you are using a sophisticated backup tool, such as Tivoli Storage Manager, then Tivoli Storage Manager can restart a migration from a disk storage pool to the tape storage pool from the point of failure. But if you use FC-TA, Tivoli Storage Manager might think that a write has been completed when it has not. Then, any restart would be fatally flawed.

Note: Tape acceleration spoofs both the R_RDY and the ACK, but it does not spoof the final tape mark.

8.2 Low-cost connection with iSCSI

You can create low-cost connections to disk storage using one of three ways:

- Fibre Channel Arbitrated Loop (FC-AL)

  Using FC-AL does not require a switch port for each server, because up to 126 devices can share a single port. One Fibre Channel host bus adapter (HBA) is still required for each server.

- Network-attached storage (NAS) gateway

  Using an NAS gateway, you need only to provision Fibre Channel ports for the gateway device, rather than for each server. Also no Fibre Channel HBAs are required for the servers, so the primary costs are for the gateway itself and for upgrading your Ethernet network to handle the increased traffic, as well as for establishing a VLAN for this new traffic.

  Note: Some block I/O applications cannot be accessed effectively through an NAS gateway.

- iSCSI

  iSCSI is like IP SAN. Using iSCSI, you do not need to provision Fibre Channel ports for each server. Also, no Fibre Channel HBAs are required, but iSCSI imposes processing usage on each server. Therefore, in some cases, you might need a high performance Ethernet card with a TCP/IP offload engine (TOE) function. Again, look at the costs associated with upgrading your Ethernet network, such as setting up a VLAN. Because iSCSI delivers block I/O, it is likely to be compatible with all applications.

All Cisco MDS Multilayer Switches have the capability for iSCSI, except for the MDS 9020, MDS 9120, and the MDS 9140. IP ports are available in the MDS 9216i, the 14+2 Multiprotocol Services (MPS) line card, and the 8-port IPS line cards.

Note: The two Ethernet ports on the MDS 9216i and 14+2 MPS line card *cannot* be combined into a single EtherChannel. Two Ethernet ports on the IPS modules *can* be combined into a single EtherChannel, but only between ports that share the same application-specific integrated circuit (ASIC). PortChannel can be used.

Figure 8-2 shows how you can use iSCSI to provision disk storage to noncritical servers. You can also use iSCSI for critical servers. In general, you can expect lower performance and lower reliability on an Ethernet network than on a Fibre Channel network. When using iSCSI for critical servers, use iSCSI multipathing.

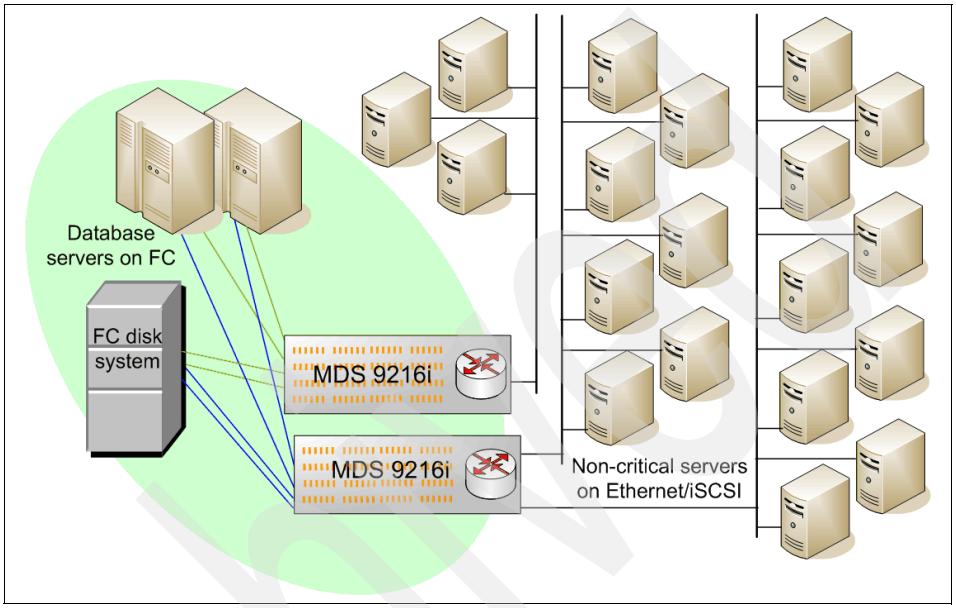


Figure 8-2 Using iSCSI routing to provision disk storage to non-critical servers

iSCSI Immediate Data

Cisco MDS 9216i and the IPS and MPS modules support iSCSI Immediate Data. iSCSI Immediate Data is similar to FC-WA. It spoofs the R-RDY and can send the initial payload with the initial write request.

8.3 Isolation and interoperability using IVR

Cisco's VSAN technology allows for the creation of separate logical fabrics on a shared physical infrastructure. There are cases where you need that isolation to be complete, but there are many other situations where you need to provide some access to that data. The following sections provide examples of ways in which you might use IVR.

8.3.1 Separating production from development

In addition to your main production environment, you might have a development or test environment that is subject to frequent reconfiguration and rebooting, or might be subject to a higher risk of failure due to less rigorous change controls. You need to isolate this from your production systems, but test systems also need occasional access to data that is stored on the production disk systems.

Figure 8-3 shows how you can achieve this fabric isolation by using a VSAN. Every Cisco MDS switch is also a router, so it can perform IVR.

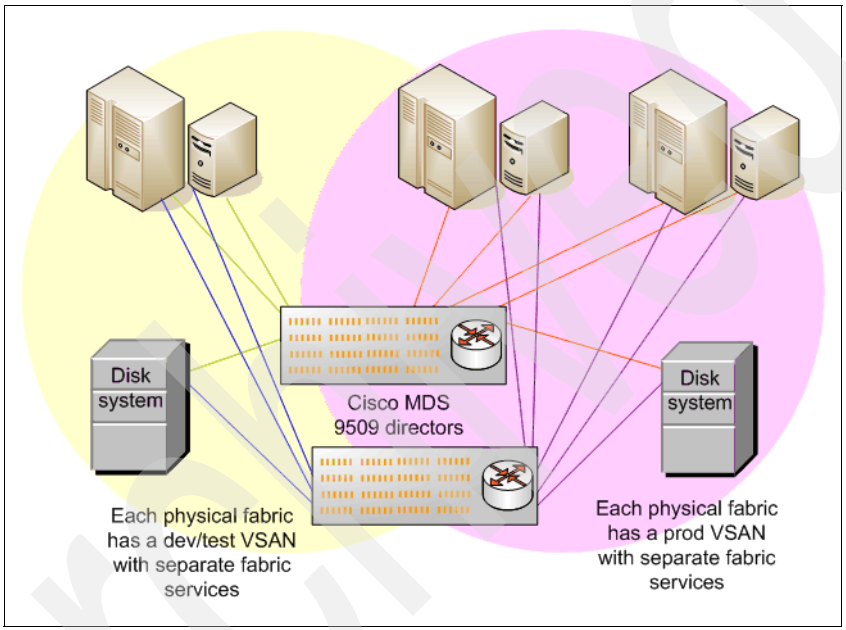


Figure 8-3 Every MDS switch is also a router: Dual director example

Figure 8-4 shows the same concept using separate switches for the two environments. Logically, this is identical to Figure 8-3 on page 262 but illustrates how you can route between two existing pairs of fabrics, so it also shows an example of scalability.

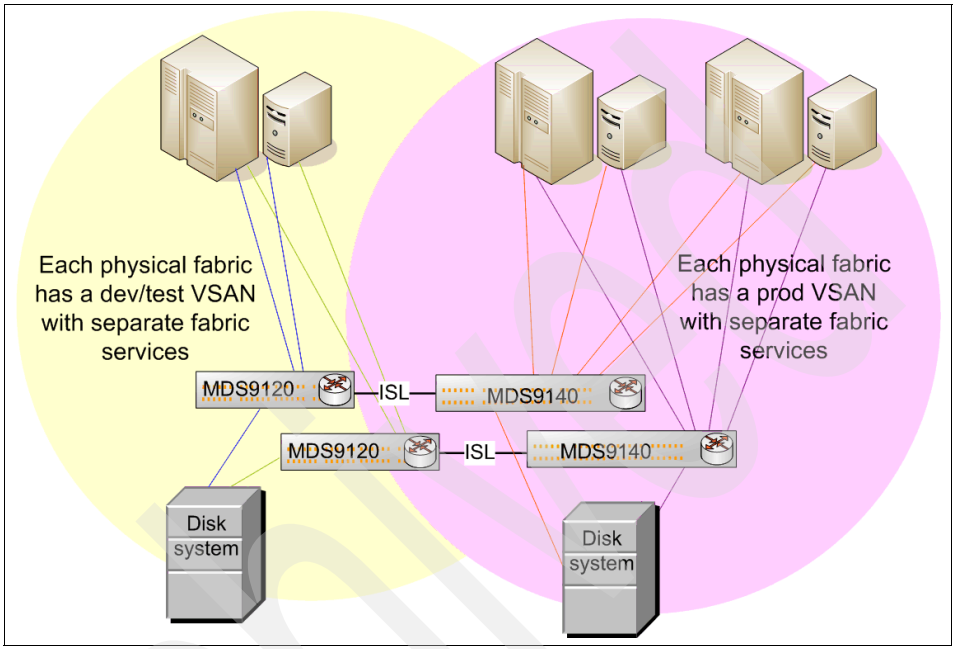


Figure 8-4 Every MDS switch is also a router: Four switch example

8.3.2 Separating corporate subsidiaries

A corporation can also choose to isolate subsidiary companies from each other while providing some shared services such as centralized backup. With Cisco MDS 9000, each VSAN can have a separate administrator with privileges granted only for that VSAN. This approach can also be used by a shared-services provider to host multiple clients on the same physical infrastructure.

Figure 8-5 shows an example where separate subsidiaries share a physical infrastructure, but live on isolated VSANs. IVR has been implemented to allow shared access to the backup infrastructure.

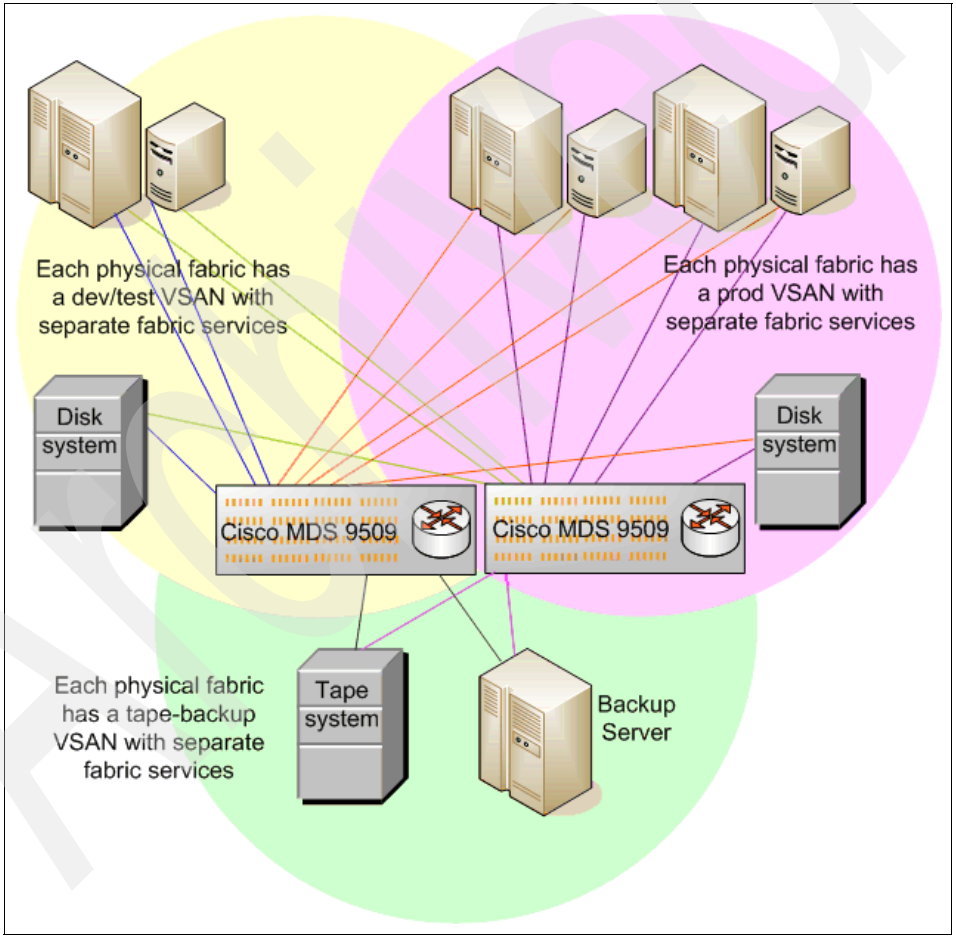


Figure 8-5 Using VSAN and IVR to isolate subsidiaries with access to shared tape

8.3.3 Isolation of multivendor switches and modes

You might have Fibre Channel switches from multiple vendors that each require different mode settings and behave slightly differently in the network. You might want to incorporate them into your network, but keep them isolated either for departmental reasons or simply keep the different modes of operation separate from each other. IVR gives the architect the confidence to combine switches from other vendors into the network, knowing that each VSAN has its own separate fabric services.

The Brocade VSAN in this case includes initiator devices attached to the Brocade switch and a single inter-switch link (ISL) port on the Cisco switch. The Cisco switch provides and manages the routing between the Brocade VSAN and the Cisco VSAN, which contains all of the other ports on the Cisco switch.

Figure 8-6 shows how switches from Brocade, McDATA, or QLogic can be incorporated into the network and yet be isolated into a separate VSAN, with IVR providing data sharing across the network.

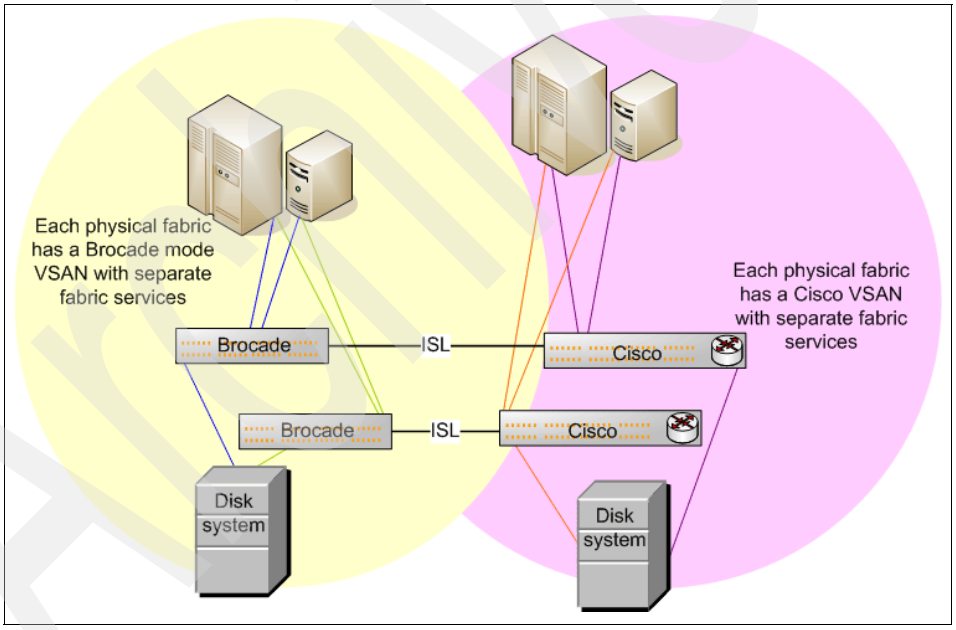


Figure 8-6 Using VSAN and IVR to provide multivendor isolation and integration

8.4 Managing scalability with IVR

It is also possible to use VSAN and IVR to prevent fabrics from growing too large because you add more ports on a modular basis. Although modern switches typically include ways to limit registered state change notifications (RSCNs), such as zone-limited RSCNs, some architects prefer to create smaller fabrics as a way to improve high availability. This is not necessarily a good idea because it introduces additional complexity. If you think about an Ethernet network, typically you do not want to introduce multiple routers into your core network. The same applies to VSAN and IVR. The guideline is to *resist the urge to route*.

You can use VSAN and IVR to manage the growth of your network and to limit the impact of a misconfiguration in a network that is subject to regular change. As you add additional switch groups, you can isolate them, as shown in Figure 8-7.

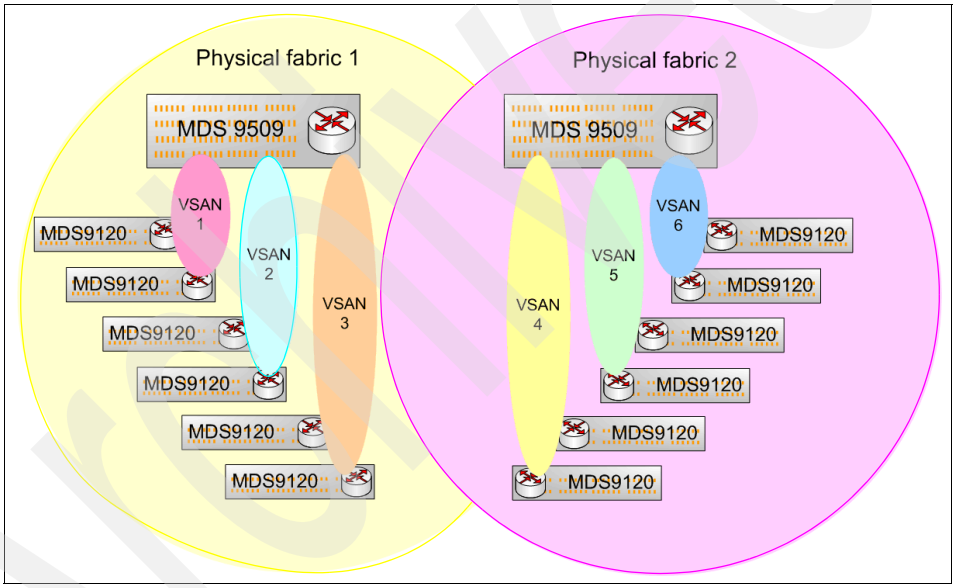


Figure 8-7 Using VSAN and IVR to manage scalability

8.5 Storage migration using IVR

The Cisco VSAN and IVR features can be valuable when considering migration from previous generation technologies. Figure 8-8 shows a site running an EMC CLARiiON FC4700 disk system and EMC Connectrix 1 Gbps switches, which are OEM from McDATA.

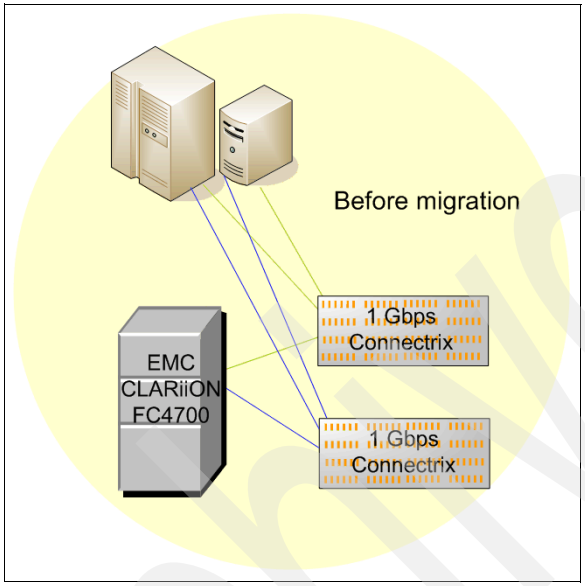


Figure 8-8 An existing SAN ready for upgrade

You can use VSAN and IVR to introduce new Cisco 2 or 4 Gbps fabric switches, and then introduce a new IBM TotalStorage disk system. In broad terms, this involves the following steps:

1. You attach the new disk systems directly to the new fabric, as shown in Figure 8-9, and can access LUNs from the servers. There is no need to have multipath access to the new disk system during the migration phase, so you can avoid issues around multipath drivers for now.

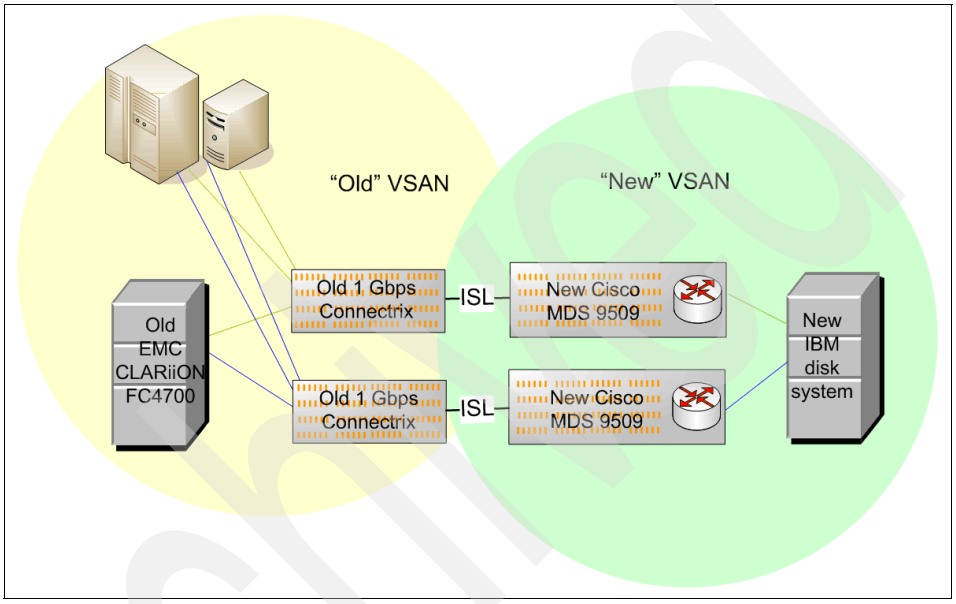


Figure 8-9 IVR allows separation during migration phase

2. Using a logical volume manager, you can create mirrors from the old LUNs on the EMC disk system to the new LUNs on the IBM disk system.
3. When the data is in sync, an outage is required on the servers so that you can directly connect them to the new fabric and install the multipathing drivers.

4. You can then remove the old fabric and disk system, as shown in Figure 8-10.

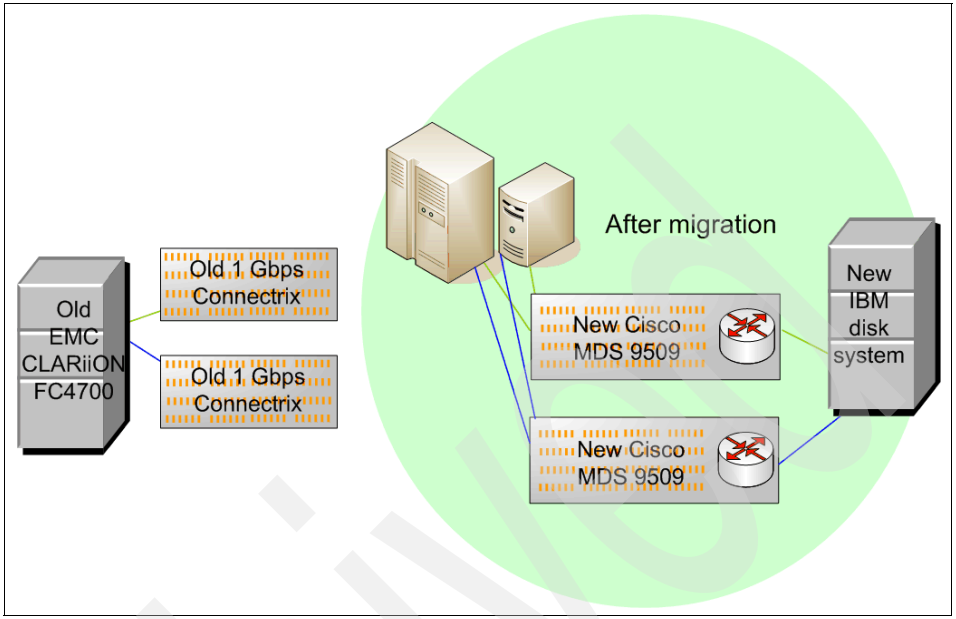


Figure 8-10 Connecting servers to the new fabric and disconnecting the old SAN

9

Cisco routing best practices

When you purchase a Cisco MDS 9000 switch, you also purchase a Fibre Channel (FC) router. The MDS 9000 is a feature-rich family. The value you gain from the products depends on which features you choose to implement and how you go about it.

9.1 To route or not to route?

That is the question. If you think about an Ethernet network, typically you do not want to introduce multiple routers into your core network, and the same applies to your Fibre Channel network. Because every port on a Cisco MDS is also a router port, it can be easy to overuse virtual storage area network (VSAN) and Inter-VSAN Routing (IVR). The guideline is to *resist the urge to route*.

Chapter 8, “Cisco routing solutions” on page 255, has examples of scenarios where you might want to route. These include the following situations:

- SAN extension over IP networks
- Lowering connection costs using Small Computer System Interface over IP (iSCSI)
- Achieving isolation and interoperability between different business units:
  - Separating production from development
  - Separating subsidiary A from subsidiary B
  - Separating multivendor switches
- Managing scalability as your SAN environment grows
- Migrating from an older storage environment to a newer one

Keep in mind that a requirement for complete isolation does not imply a requirement for routing. Routing is required only when you want isolation alongside the capability of some access between isolated environments.

Before architects and implementers deploy IVR, read the white paper *Inter-VSAN Routing with the Cisco MDS 9000 Family of Switches & Cisco SAN-OS 2.1* from Cisco on the Web at:

http://www.cisco.com/en/US/netsol/ns514/networking_solutions_white_paper0900aecd80285738.shtml

Important: Be careful with introducing unnecessary VSANs into a SAN Volume Controller (SVC) environment. Refer to “SAN Volume Controller interoperability” on page 275, for more details.

9.2 Piloting new technology

When you plan to use advanced features, such as FC over IP (FCIP), iSCSI, VSAN and IVR, compression, FCIP write acceleration (FCIP-WA), and Fibre Channel tape acceleration, it is important to implement the new technology initially as a pilot. You must do so with the understanding that the experience gained in your own environment will always be slightly unique.

When implementing leading-edge technologies, many clients prefer to avoid the uncertain outcomes that a pilot implies. Instead they secure implementation guarantees from vendors. In fact, the outcome can never really be guaranteed. Piloting allows the solution to be tailored based on lessons learned in your own environment.

9.3 iSCSI issues

Before architects and implementers deploy an iSCSI solution, read the white paper *iSCSI Design Using the MDS 9000 Family of Multilayer Switches* from Cisco on the Web at:

http://www.cisco.com/en/US/netsol/ns514/networking_solutions_white_paper09186a0080171d9e.shtml

There is a lot of enthusiasm and excitement about iSCSI. One thing that is certain is that it is slower than using Fibre Channel. In the white paper *Guide to iSCSI Performance Testing on the Cisco MDS 9000 Family*, testing was performed with an xSeries 345 (x345; dual 2.66 GHz with 1 GB RAM) running Microsoft Windows 2000, and an EMC CLARiiON CX600. This white paper offers advice about tuning and I/O block sizes, but remains inconclusive about the performance of iSCSI for specific real-world applications.

You can find this paper on the Web at:

http://www.cisco.com/en/US/products/hw/ps4159/ps4358/products_white_paper0900aecd801352e3.shtml

In Figure 9-1, the I/O response time in the bottom right graph is presumed to be in milliseconds.

The use of TCP/IP offload engine (TOE) Ethernet cards is one way to reduce the CPU usage of iSCSI processing. However, using specialized hardware also detracts from the cost and ease-of-use arguments.

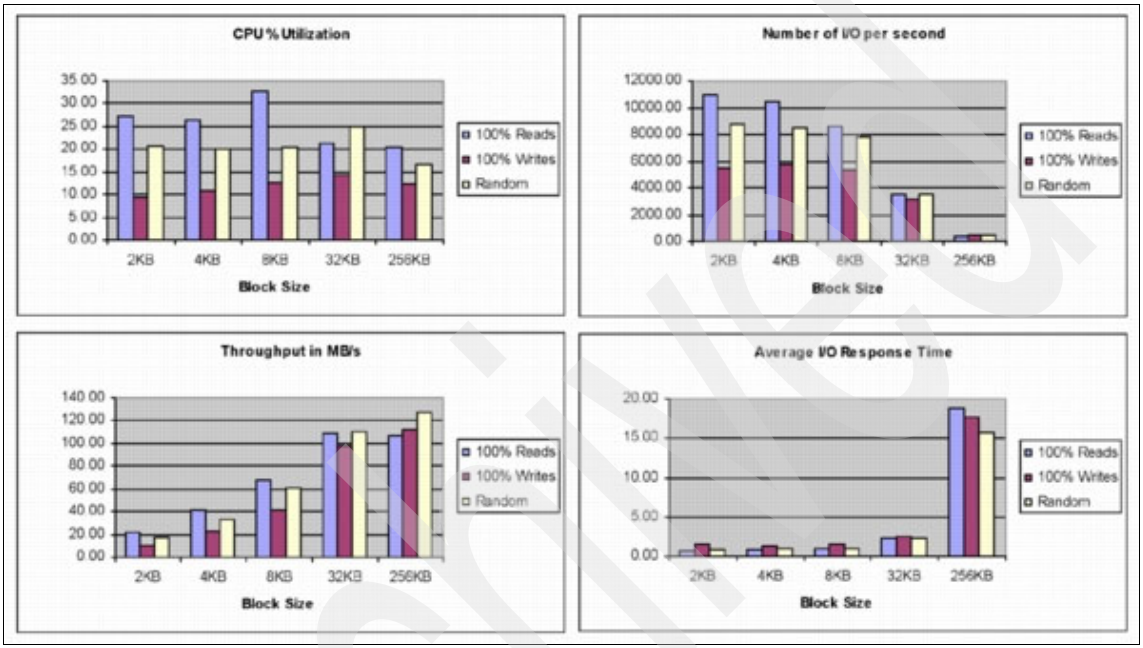


Figure 9-1 How block size affects iSCSI performance

9.4 IP network issues

Before architects and implementers deploy an FCIP solution, read the white paper *Designing FCIP SAN Extension for Cisco SAN Environments* from Cisco on the Web at:

http://www.cisco.com/application/pdf/en/us/guest/netsol/ns378/c649/cdccont_0900aecd800ed145.pdf

You can find additional white papers at the following Web site:

http://www.cisco.com/en/US/netsol/ns340/ns394/ns259/networking_solutions_packages_list.html

Also ask your telecommunications company about high quality of service (QoS) managed links, including information about offerings such as IP over SONET/SDH. Help your telecommunications company clearly understand your network quality expectations and make sure that you clearly understand the cost and management implications of any decisions that you make.

9.5 Interoperability

Cisco switches are explicitly supported for heterogeneous interconnection with:

- Brocade (including native mode)
- McDATA
- IBM BladeCenter
  - Optical pass-through modules
  - QLogic modules
  - McDATA modules
  - Brocade modules

Heterogeneous IVR was announced to the field by IBM on 24 May 2005 in the document *IBM announces Cisco MDS 9000 Intelligent Fabric features*.

Note: At time of writing, the following interoperability documents did not include mention of heterogeneous IVR.

The Cisco MDS interoperability matrixes are available by clicking the individual product details at the following Web site:

`http://www.ibm.com/servers/storage/san/c_type/`

You can find also find Cisco MDS interoperability matrixes by selecting the product information files from the following site:

`ftp://service.boulder.ibm.com/storage/san/cisco/`

The maximum hops supported is seven including one longwave ISL.

SAN Volume Controller interoperability

When deploying VSANs in an SVC environment, give careful consideration to how the different VSANs will access the SVC, given that in an SVC environment, typically all disk I/O travels through the SVC. Each SVC node has four Fibre Channel ports, so each can be part of a maximum of four fabrics or VSANs without routing through IVR.

9.6 Designing for availability

Redundancy can be used to offer some protection against hardware failure and against disruptive fabric events. The best practice with fabric design is usually to use dual redundant fabrics, which have traditionally implied redundant hardware. With the advent of the VSAN, however, the equation is not so straightforward. The topics that follow briefly discuss some areas to focus on with respect to ensuring availability.

9.6.1 Fibre Channel router hardware

All members of the Cisco Fibre Channel router/switch family include redundant hot-swappable power supplies and fans and the ability to restart a failed supervisor process.

As offered by IBM, the Cisco MDS 95xx also comes with two supervisor modules, each with its own control engine and crossbar fabric. The two crossbar fabrics operate in a load-shared active-active mode. The system does not experience any disruption or any loss of performance with the removal or failure of one supervisor module.

Even though a single Fibre Channel router/multilayer switch may be designed for high availability, two of them provide a more robust solution. Remember also that, although there are dual redundant clock modules in the Cisco MDS950n Directors, if one clock module needs to be replaced, a director outage is required because these modules are not hot-pluggable.

Common sense and the principles of balanced system design must prevail. In some cases, you might assess that deploying two 99.999% available routers or directors, such as the MDS 9509, is unnecessary if the selected midrange back-end disk system is only 99.99% available.

9.6.2 Nondisruptive software upgrade

The Cisco MDS 9500 Series of Multilayer Directors supports the ability to upgrade the supervisor module and the switching module software on the fly without disrupting traffic flowing through the switch. During this upgrade process, the standby supervisor is upgraded. Then, control automatically moves over to the upgraded supervisor (standby has become active, active has become standby), and the standby is automatically upgraded to the same level of software. Dual versions of software on the MDS are not supported.

This process allows maximum flexibility in upgrading the software while providing a path to revert back to a known level of stable software.

9.6.3 Inter-switch links

Relying on a single physical link between switches reduces the overall redundancy in the design. Redundant inter-switch links (ISLs) provide failover capacity if a link fails. Some architects are content with two separate fabrics, each with their own ISL. The issue with this is that, if an ISL fails, you are in fabric failover, because one of the fabrics is effectively broken. By having at least two ISLs per fabric, you avoid fabric failover if a link fails.

The Cisco PortChanneling feature enables you to load balance across multiple ISLs. When deploying iSCSI or FCIP networks, check that the ports are capable of using PortChanneling.

If running FCIP over an externally provisioned network, architects should also understand whether multiple links take separate physical routes through the end-to-end network.

9.6.4 VSAN and IVR

VSAN and IVR can also be used to protect against disruptive fabric events. Fabric-level events have the potential to disrupt all devices on the fabric. Mistakes made when adding a switch or changing zoning configurations can ripple through the entire connected fabric.

In large or complex networks, designing with separate VSANs helps to isolate the scope of any such events. In smaller networks, however, configuring VSANs and IVR might simply add unnecessary complexity.

9.6.5 Backup

Some members of the Cisco MDS Family offer a compact flash memory card. The settings can be backed up there, but it is generally better to back up to an external FTP server. Remember to apply sound change control procedures and retention of previous good configurations.

9.7 Designing for security

Security also impacts high availability because one of the leading causes of downtime is human error.

Role-based administration

The Cisco MDS 9000 Family of Multilayer Directors and Fabric Switches support a role-based security methodology to ensure that only authorized individuals have access to critical functions within the fabric. Each user is assigned to a role, better known as a group_ID, which is given a specific access level within each fabric. This access level dictates the commands, or more specifically, to which nodes of the command-line interface (CLI) command parser tree the particular role has access.

Note: Users can have more than one role.

Centralized management

Roles can be defined and assigned locally within a switch by using CLI commands or can be centralized in a RADIUS server for easier management. Two default roles are provided: Network Administrator (full access) and Network Operator (read-only access). Up to 64 custom roles can be defined by the user. Only a user within the Network Administrator role can create new roles.

VSAN administration

VSANs contribute to the security of a network by maintaining isolation of devices onto different fabric services even though they can share a physical fabric. Because each VSAN is a separate virtual fabric, each VSAN has its own set of role-based administrators. This adds to the security of the SAN and makes it safer to administer a shared fabric.

Encryption

You must answer these key questions:

- How do I prevent someone from viewing or modifying confidential data?
- Does connecting a Fibre Channel network to an IP network impact the integrity of my data?

The principal security mechanism on a Fibre Channel fabric is zoning. Zoning works like an access control list (ACL), so only devices that are on each others lists can talk to each other. VSANs provide an additional layer of security because they also define members. Non-members are not given access unless they are a member of another VSAN which has IVR access to the first VSAN.

Beyond ACLs, you can deploy encryption. The Cisco Multiprotocol Services (MPS) 14+2 line card and Cisco MDS 9216i switch both offer integrated hardware-based IP security protocol (IPSec) support, providing wire-rate encryption and decryption with Advanced Encryption Standard (AES) and Triple Data Encryption Standard (3DES).

Cisco provides a range of additional features that are designed to maintain the integrity of the fabric. For a discussion about some of these ideas, refer to the article “Security: Beyond Zoning” in the Cisco user’s magazine *Packet* in the second quarter 2005 edition. You can find this edition on the Web at:

http://www.cisco.com/application/pdf/en/us/guest/netsol/ns513/c666/cdccont_0900aecd802c2b74.pdf

9.8 Designing for performance

The topics that follow briefly discuss some of the items that you need to consider when designing for performance.

9.8.1 Hardware selection

At the time of writing, Cisco supports maximum port speeds up to 4 Gbps. Some disk systems, such as the DS4800, use 4 Gbps connections, but most are based on multiple 2 Gbps connections.

Place emphasis on the principles of balanced system design and promote lower throughput ports, such as iSCSI (shared 1 Gbps), and host-optimized line cards (with 3.2:1 oversubscription) as being suitable for most servers.

There are two common approaches to designing for performance. One approach is to try to understand the peak workloads, and then to project growth and allocate bandwidth accordingly. The other is to make every connection as high speed as possible, within certain cost restrictions.

When connecting back-end disk subsystems, ISLs, tape libraries, and high-throughput servers, use target-optimized ports:

- Four ports on the MDS 9120 20 port switch
- Eight ports on the MDS 9140 40 port switch
- All the ports on the 12-port 4 Gbps switch module
- Some ports on the 24-port 4 Gbps switch module (achieved using Port Bandwidth Reservation)
- Fourteen ports on the MDS 9216i and the MPS 14+2 line card

When connecting low-throughput servers, you can use host-optimized ports. Or you can use iSCSI ports or Fibre Channel Arbitrated Loop (FC-AL). Vendors with an iSCSI solution are generally enthusiastic about iSCSI, but there is still a lot of varying information in the market about performance. We recommend that you run a pilot test for iSCSI solutions before you place them into production.

Although the bandwidth on most shortwave Fibre Channel networks installed today is underused, there are applications that perform high-volume sequential I/Os, which can flood Fibre Channel ports. Remember that although a disk system might only have one, two, or four back-end 2 Gbps FC-AL loops, the system might be able to feed a lot of data onto the network from cache.

Note: You can use the Cisco SAN Extension Tuner to help understand and optimize FCIP performance. The tuner generates SCSI I/O commands that are directed to a specific virtual target. It reports I/Os per second and I/O latency results. The SAN Extension Tuner is included with the FCIP Enablement license package.

9.8.2 FCIP compression and FCIP-WA

Results can vary significantly from site to site when deploying FCIP compression and FCIP-WA.

FCIP compression

Cisco implements hardware compression in the MDS 9216i and on the MPS 14+2 line card. In IP services, module compression is implemented in software.

Software compression has one advantage. In some modes, multiple small Fibre Channel frames can be stacked into a single IP packet to make more efficient use of the payload space.

Both approaches use an LZS compression algorithm that searches for repeat data strings in the input data stream and replaces these strings with data tokens shorter in length than the original data. A compression history table is built of these string matches, pointing to previous data in the input stream. The net result is that subsequent data in the stream is compressed based on earlier data. The compression ratio is higher with repeated data and lower with greater randomness.

Because compression varies as the data varies, we recommend that you perform a pilot for the different compression modes in your environment until you arrive at the one that delivers you the best throughput.

FCIP-WA

FCIP-WA performance depends heavily on the traffic profile of the SCSI operations being performed. The following I/O characteristics are well-suited to benefit from FCIP-WA:

- Long-distance, high-latency network
- Write-intensive I/O
- High number of small SCSI writes (rather than low number of large writes)
- Disk system with low write latency

This suggests that the best use for FCIP-WA is in disk system replication. Across a 100 km link, replication using FCIP-WA can be expected to deliver between 5% and 20% throughput improvement, with 10% being a realistic expectation.

Again, the benefits of FCIP-WA vary for each client. Piloting is the only appropriate way to be sure if it is a good fit for your network.

10

Cisco routing real-life solutions

The case studies in this chapter are intended to show real-life situations. They also show the pragmatic solutions that have been developed around them using the Cisco MDS multiprotocol technology.

We present two case studies here. The first one implements inter-virtual storage area network (VSAN) routing (IVR), Fibre Channel (FC) over IP (FCIP) tunneling over a campus IP backbone, and Small Computer System Interface over IP (iSCSI). The second one implements IVR and asynchronous disk mirroring through FCIP tunneling over a long-distance IP network.

Important: The solutions and sizing estimates that we discuss or make in this chapter are unique. Make no assumptions that they will be supported or apply to each environment. We recommend that you engage IBM to discuss any proposal.

10.1 University ZYX

University ZYX, a fictional university, provides teaching and research services to 30,000 students across 10 faculties: arts, business, architecture, fine arts, education, engineering, law, medicine, science, and theology. In addition, there are 20 interdisciplinary research clusters focused on emerging fields.

The example that follows departs from best practice in some ways. Based on a history of no Fibre Channel switch failures over four years, and to save costs, the client decides that having dual physical fabrics is not important at their disaster recovery (DR) and engineering sites, especially given the capability to create VSANs each with their own fabric services. The client also decides that a single director class router or switch will provide them with adequate reliability at the core of their network.

This approach is also echoed in their willingness to use singly-attached hosts and iSCSI for non-critical servers in many cases. The client reasons that having less hardware is not so important, because they gain rich functionality with the MDS family, such as routing capability on every port of every switch.

10.1.1 Initial growth

University ZYX installed their first SAN in 2001 with an IBM TotalStorage Enterprise Storage Server® (ESS) model F20 and several 1 Gbps switches, primarily to support their PeopleSoft (now JD Edwards from Oracle) applications. In 2003, they added an IBM TotalStorage DS4500 disk system to hold Tivoli Storage Manager pools and other data that did not require tier one storage. That same year, they also added a DS4500 disk system at their cold disaster recovery site across town in the IBM data center.

In early 2004, they added a pair of SAN Volume Controller (SVC) nodes with a 32 TB license to allow virtualization of their storage volumes for increased management flexibility. Figure 10-1 shows the layout of the SAN environment after four years.

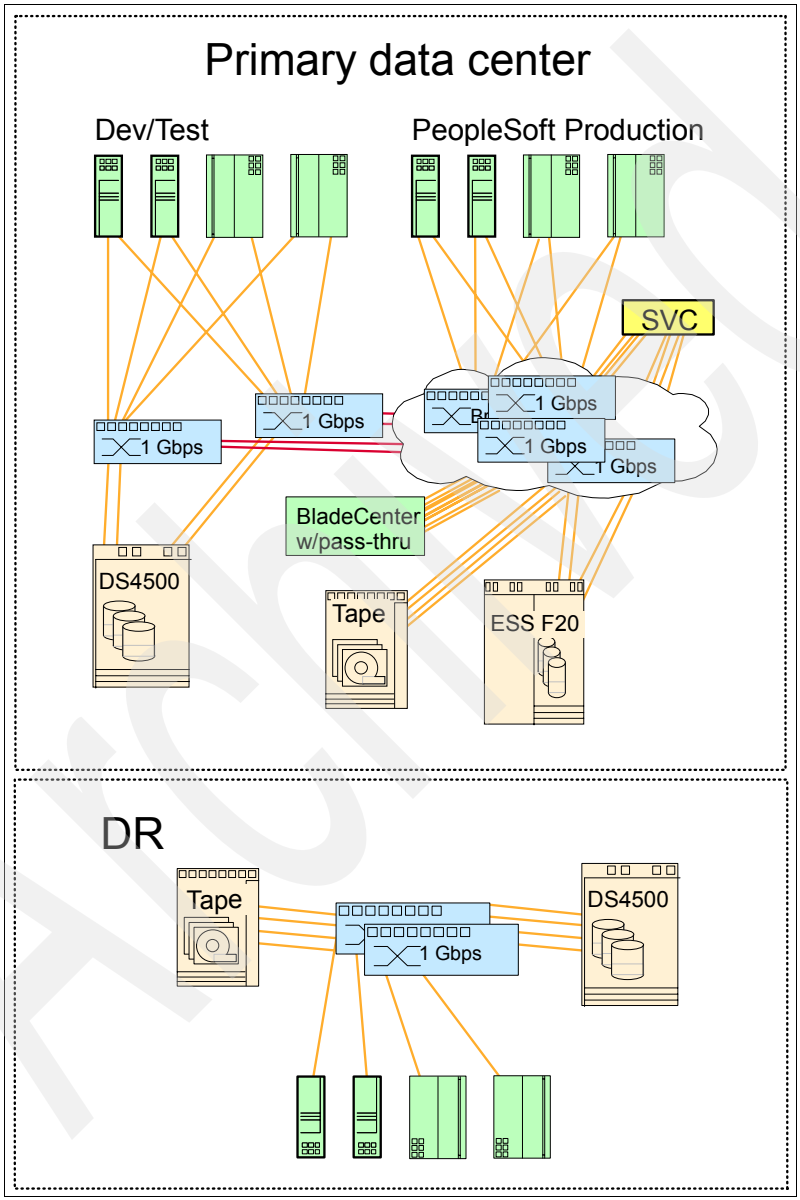


Figure 10-1 University ZYX SAN environment after four years

10.1.2 Lease expiration

At the end of 2004, University ZYX realized that the four-year lease on the ESS and the 1 Gbps Fibre Channel switches was only six months away from full term. They began planning for replacements.

As part of this, they wanted to leverage iSCSI to provide a flexible, low-cost connection to a range of small servers. They also wanted to use their existing campus IP backbone to provide some sharing capability with the School of Engineering, which was approximately 2 km down the road. Engineering wanted to purchase their own disk system as well.

At the time, the Cisco MDS was the only Fibre Channel solution with FCIP and iSCSI being offered by IBM. Because the University was a content user of Cisco IP networking devices sourced from IBM Global Services already, they decided to proceed with Cisco MDS equipment. Based on three years of fault-free operation on the ESS, they also selected IBM TotalStorage DS8100.

10.1.3 Design and purchase of new systems

Engineering went ahead with the purchase of IBM TotalStorage DS4300 and Cisco MDS 9216i. They chose to run a single physical fabric and use the VSAN feature of the Cisco MDS to create logical subfabrics. They also knew that by deploying Cisco MDS, they had options for iSCSI and FCIP available.

The IT services department decided to purchase a single MDS 9509 and initially decided on 3 x 32 port line cards and 2 x 14+2 Multiprotocol Services (MPS) card. After a review with Cisco, IBM, and the IBM Business Partner, it was decided that this configuration did not provide enough target-optimized ports. Two additional MPS 14+2 cards were substituted for one of the 32-port line cards. A decision was also made not to use CompactFlash to back up the MDS configuration, but to back it up to an external File Transfer Protocol (FTP) server instead.

Some of the 1 Gbps switches were owned and not leased. The client decided to retain some of them for an additional year and to use VSAN and IVR to isolate and connect the 1 Gbps switches. The SVC adds complexity in that all data flows through it, so it either has to be a member of every SAN, or it must be accessed thorough IVR. When all of the 1 Gbps switches are retired, this aspect of the design can be simplified.

Figure 10-2 shows the network after installing the new multiprotocol switches/routers and the new IBM TotalStorage DS8100 and DS4300.

IVR provides the test/development system with access to the tape library and to the main DS8100 when needed. Each VSAN *must* have access to the SVC.

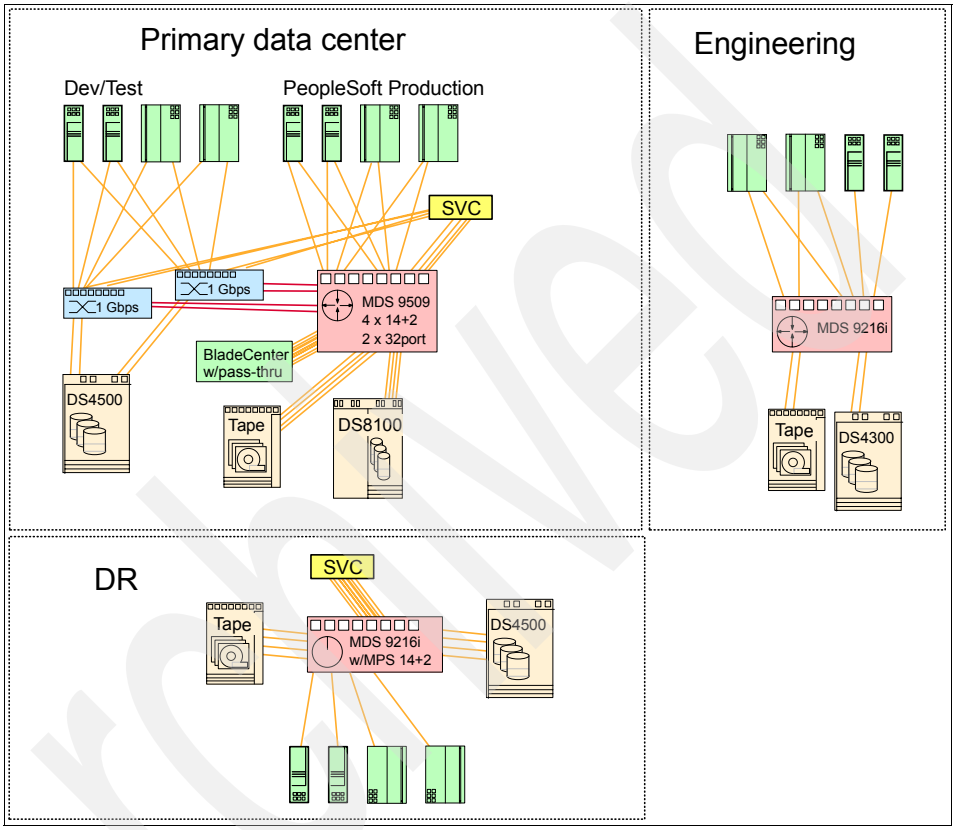


Figure 10-2 The network after installing the DS8100, DS4300, and Cisco MDS

10.1.4 Deployment of iSCSI and FCIP

iSCSI was deployed using the IP ports on the MPS 14+2 line cards. The main users are faculty file servers, which have a requirement for online or near-line storage. Additional IBM TotalStorage DS4000 EXP100s with SATA drives were purchased for the DS4500 to provision low-cost storage for the iSCSI-attached servers.

After considering a dedicated longwave Fibre Channel connection to engineering, the University opted instead to use the existing campus IP backbone and use Cisco's FCIP functionality to connect the two sites, believing

this to be a more flexible solution. The FCIP link is in a transit VSAN with IVR providing communication through the transit VSAN to the far site.

Figure 10-3 shows the physical network plan including FCIP and iSCSI.

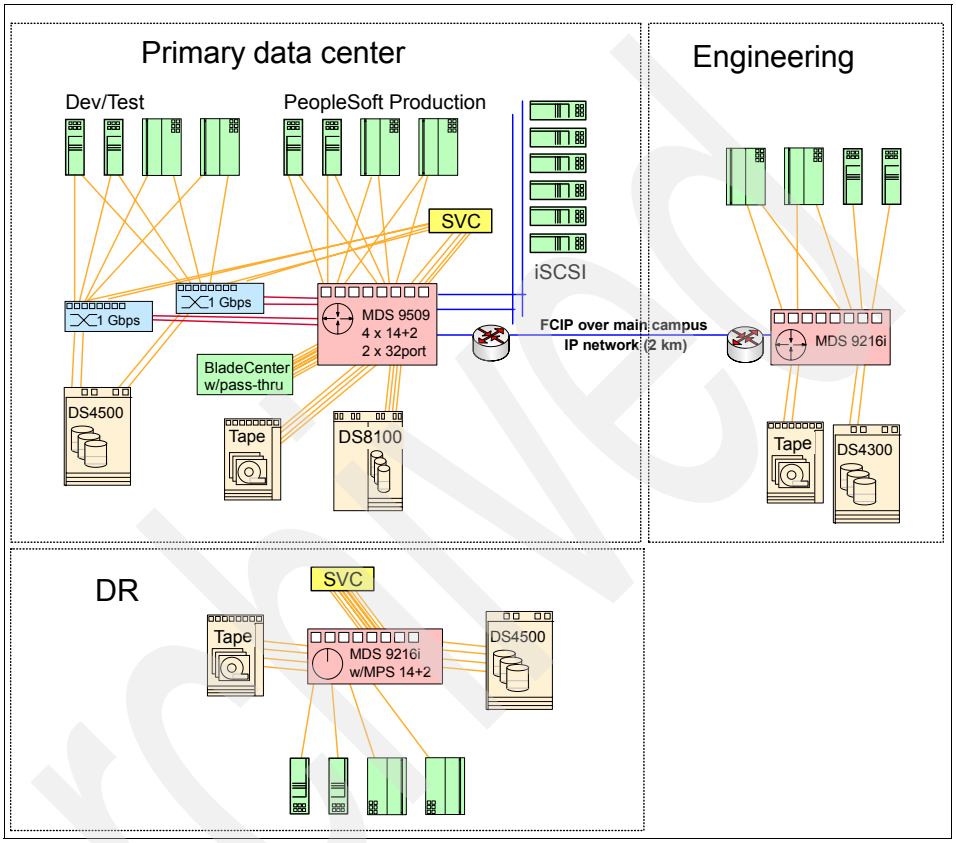


Figure 10-3 University ZYX planned network with FCIP and iSCSI

10.1.5 SVC synchronous replication for disaster recovery

The final stage in this technology refresh was the purchase of an SVC for disaster recovery and the establishment of a link to the IBM data center where the University's disaster recovery systems are housed. The advice from IBM is that a maximum 10 microsecond (ms) latency can be tolerated on SVC synchronous replication.

Estimating the latency of an IP link can be tricky. There is the 5 ms Fibre Channel cable latency per km in each direction, plus another 5 ms each way in each Fibre Channel device, plus IP router latency, which can be anything from 10 ms each

way to much higher, especially if filters are applied. Then if traffic passes through other devices, such as firewalls, they also need to be factored in.

The real question is round-trip delay, rather than latency, because delay includes any congestion that might occur. Provided the round-trip delay never goes higher than 10 ms, the replication will be successful using FCIP.

The University decided to proceed with an FCIP link. Figure 10-4 shows the physical network including SVC sync replication.

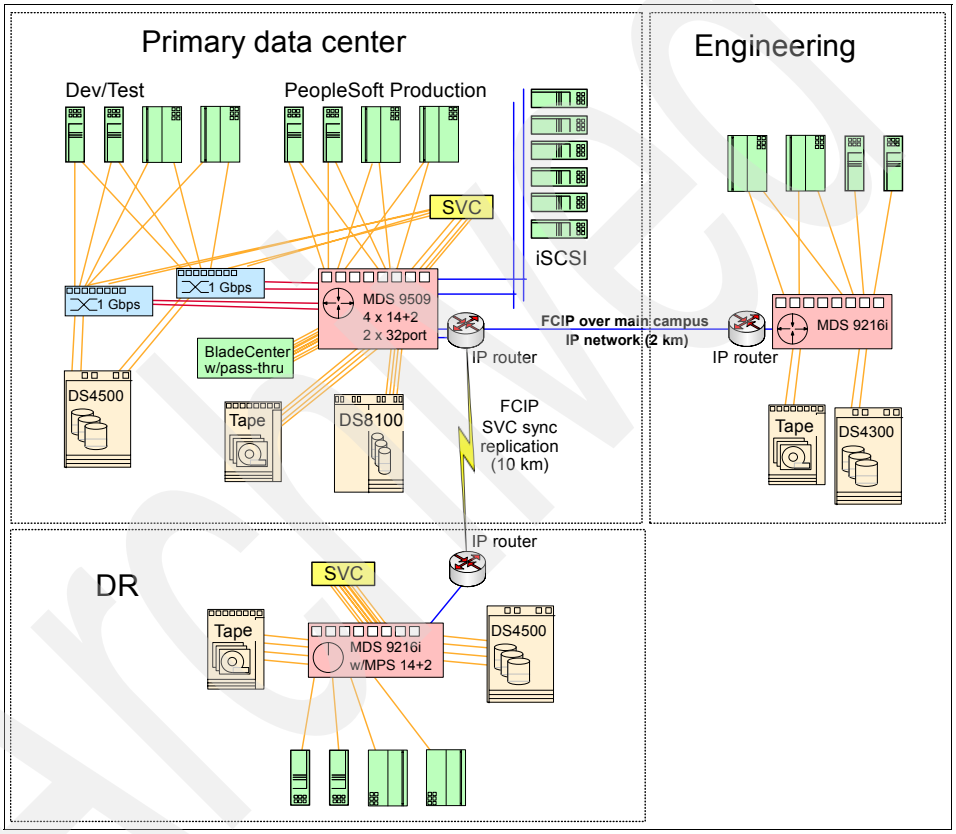


Figure 10-4 University ZYX network including SVC sync replication

10.2 Power Transmission Company ZYX

Power Transmission Company ZYX is a fictional state-owned business that is set up to own and operate the high-voltage electricity transmission grid that links generators to distribution companies and major industrial users. The company's head office is located in an area that is susceptible to earthquakes, but there is currently little provision for disaster recovery. There is a main engineering office about 600 km to the north and a customer service call center about 400 km to the south.

10.2.1 Existing systems

The company uses an asset metering application, which monitors the condition of substations, meters, transmission lines, and other assets as events occur. The data is integrated with back-office systems and analytical tools for improved decision making.

Secure data communications are provided between remote devices and back-end applications. The system gathers, filters, and communicates data on usage and status. Communications gateways are based on Arcom Controls *Director series* equipment.

The production system runs on a pSeries server with AIX 5L. Storage is shared between production and development/test on a CLARiiON CX600 disk system and an ADIC Scalar 100 tape library with two SCSI drives. Backups are done over the LAN using Tivoli Storage Manager. Current Fibre Channel switches are 1 Gbps.

Software includes IBM WebSphere® MQ Telemetry transport, IBM WebSphere Business Integration Adapters for utility industry processes and applications, and IBM WebSphere MQ Everyplace® software.

Figure 10-5 shows the existing SAN environment at Power Transmission Company ZYX.

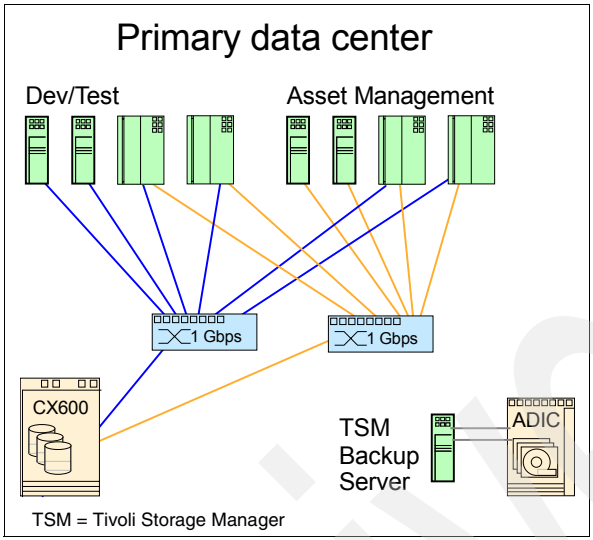


Figure 10-5 Existing SAN environment at Power Transmission Company ZYX

10.2.2 IT improvement objectives

The first main objective is to provide increased capacity and performance on the SAN in line with new application modules and increased usage of the system. The second main objective is to reduce business risk associated with IT infrastructure and site failure, including introducing:

- Dual-fabric attachment for all hosts
- Separation of development/test and production
- Improved backup/restore throughput
- Replication of data to a disaster recovery site

The company is an existing Cisco client for Ethernet switches and IP routers and has been happy with their products. The new IT infrastructure manager also has a background as a network manager. The manager has a good relationship with the Cisco account team and has some knowledge of the MDS family gained by reading the Cisco user's magazine *Packet*. It has become clear that FCIP tunneling will be the most cost-effective way to achieve remote asynchronous replication. Power Transmission Company ZYX has decided to use Cisco MDS multiprotocol switches and routers in rebuilding its SAN environment.

10.2.3 Deployment of new technology and establishment of the disaster recovery site

Power Transmission Company ZYX decided to set the 1 Gbps switches aside for use in the development/test environment. Because they plan to extend their Fibre Channel network to other servers at a later time, they elect to deploy two Cisco MDS 9509 Multiprotocol Directors at the core of their new network. The company considered using a single MDS 9216i at the disaster recovery site because it could use the VSAN feature to provide separate fabric services. In the end, the company chose to deploy two MDS 9216i multiprotocol switches at the disaster recovery site.

As part of the technology refresh, the existing tape library and several of the existing servers will be redeployed to the disaster recovery site. The disaster recovery site will initially provide cold disaster recovery only.

Figure 10-6 shows the site with separation of development/test from production and of the tape VSAN. It also shows establishment of a disaster recovery site. Fibre Channel connections are color-coded on a per-VSAN basis at each site.

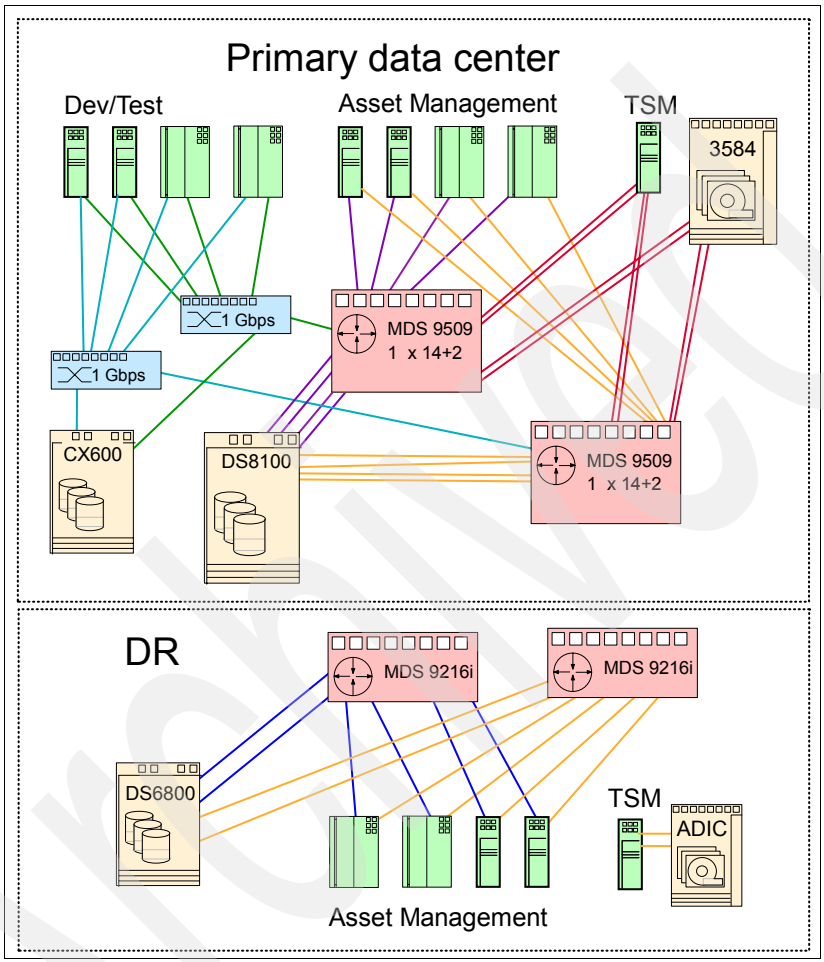


Figure 10-6 Development/test break from production; DR site established

Use of VSAN and IVR

VSANs have been deployed to provide isolation of the 1 Gbps switches. Each of these VSANs includes all of the ports connected to that 1 Gbps switch. IVR ensures that occasional access between the development/test and production areas can be accommodated.

A VSAN has also been used to create a separate fabric for the tape backup/restore solution. When other applications are added to the SAN, it is

planned that these will be separated into their own VSAN to better manage change control across different business units.

10.2.4 Global Mirroring established to the disaster recovery site

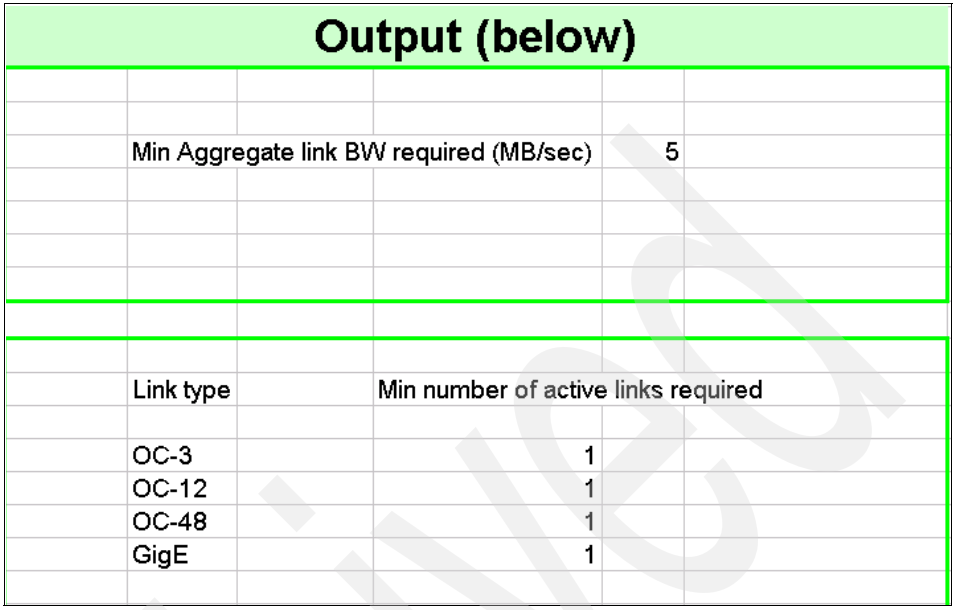
Remember that when using Global Mirroring, this is essentially the same as using Global Copy plus periodic uses of FlashCopy® to provide a safe recovery point, so extra disk space must be allowed for the FlashCopy copies.

A sizing for the FCIP link was done using the Async PPRC Bandwidth Sizing Estimator available from IBM and IBM Business Partners. Figure 10-7 shows the values entered into the fields of the Async PPRC Bandwidth Sizing Estimator.

Input (in blue box below)					
Primary Site Configuration:			**Workload (aggregate for primary site):**		
Number of ESSs		1			
Configuration per ESS:				zOS	Open
Number of HAs to host(s)		6	IO rate (SIO/sec)	0	5000
Number of HAs to 2ndary		2	Read Hit Ratio	0.92	0.6
Distance to 2ndary (miles)		500	Write Hit Ratio	1	0.87
link data compression factor		2	Read/Write Ratio	3	2.33
			Destage rate	0.5	0.5
			Seq Prestage Rate	0.4	0.4
			Avge Block size (KB)	27	4
Desired Consistency Group Interval Time (sec)	0				
Desired max drain time (sec)	60				

Figure 10-7 Input to the Async PPRC Bandwidth Sizing Estimator

Figure 10-8 shows the output of the values entered in Figure 10-7 on page 294.



Output (below)		
	Min Aggregate link BW required (MB/sec)	5
	Link type	Min number of active links required
	OC-3	1
	OC-12	1
	OC-48	1
	GigE	1

Figure 10-8 Output from the Async PPRC Bandwidth Sizing Estimator

Based on a 60 second drain time, we estimate that a bandwidth of 5 MBps (or about 40 Mbps) is required. The **ping** round-trip time (RTT) on this network is observed to be approximately 20 ms.

Note: You can use the Cisco SAN Extension Tuner to help understand and optimize FCIP performance. The tuner generates SCSI I/O commands that are directed to a specific virtual target. It reports I/Os per second and I/O latency results. The SAN Extension Tuner is included with the FCIP enablement license package.

IBM TotalStorage DS6800 was configured with 12 ranks of 73 GB 15 KRPM drives (48 drives in total) as one storage pool using Redundant Array of Independent Disks 5 (RAID5). During the design phase, using IBM Disk Magic, we checked to ensure that we had enough performance in the disaster recovery

DS6800 to process 5000 input/output processors (IOPs), plus Global Mirror, and head-room for FlashCopy copies to be created. Figure 10-9 shows the utilization statistics from IBM Disk Magic on this configuration.

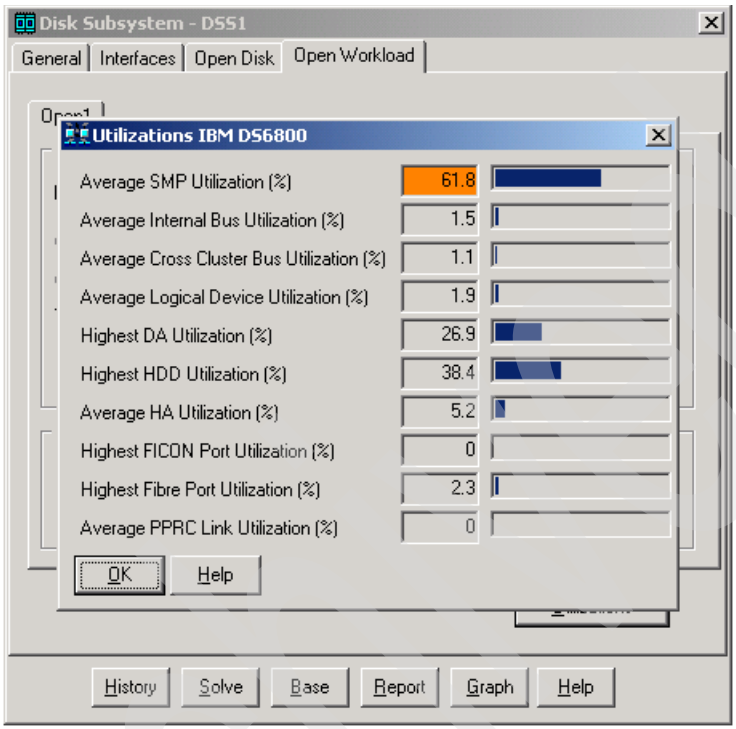


Figure 10-9 Utilization statistics for the disaster recovery DS6800 at 5000 IOPs

The company also implements Tivoli Storage Manager to stream copies of the backup both to local tape. In addition, to the remote Tivoli Storage Manager server over the IP network, the Tivoli Storage Manager traffic is transferred at night when the link is largely unused. The disaster recovery Planning module of Tivoli Storage Manager is also implemented to document the required workflow to complete a recovery.

Figure 10-10 shows that Global Mirroring has been established using FCIP tunneling. A transit VSAN was established that includes the FCIP ports on the MDS 9509 Multiprotocol Directors. This avoids fabric disruption if the wide area network (WAN) link is broken for any reason.

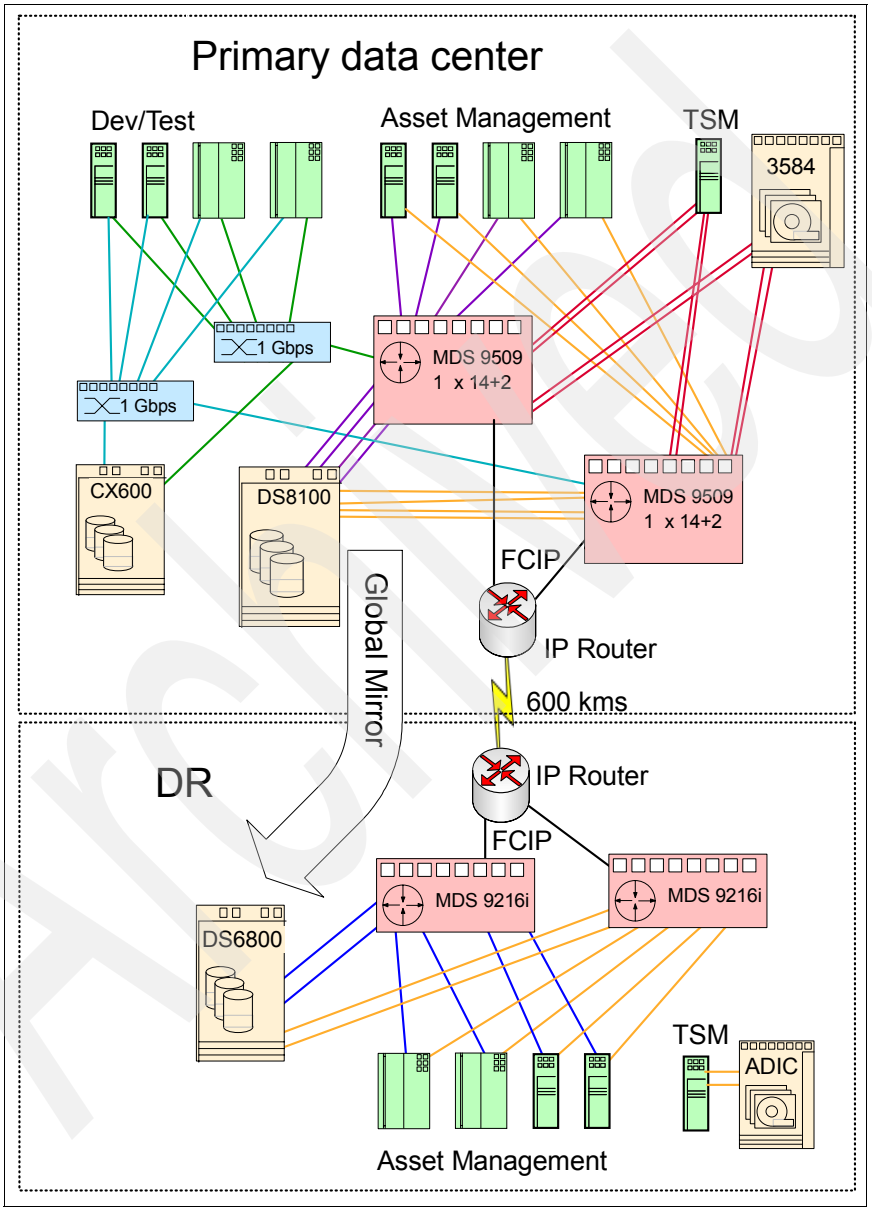


Figure 10-10 Global Mirroring established using FCIP tunneling and IVR

11

Cisco initial setup

In this chapter, we introduce the Cisco MDS 9000 Family of Fibre Channel switches and directors, focusing on the Cisco MDS 9216i. We describe the initial setups required and the process to install and activate the Cisco Fabric Manager client GUI and the Cisco Device Manager client GUI.

Note: We used Cisco Multilayer intelligent SAN operating system (SAN-OS) Version 3.0(2a) for all our testing. If your SAN-OS level is different, some of the panels might not look the same. However, the concepts introduced here should still apply.

11.1 FCP and the Cisco MDS 9000 products

In this chapter, we perform the basic switch setup and install the Fabric Manager (FM) client and Device Manager (DM) client on our workstation. If you already have the Fabric Manager client code on your workstation, and you completed the basic switch setup, you may skip 11.2, “Initial setup of the Cisco MDS 9000 products” on page 305.

11.1.1 Port addressing and port modes

The Fibre Channel ports in the Cisco MDS 9000 family are addressed with addresses in the form of fc<slot>/<port>, where <slot> is the slot number of the line card (1-9), and <port> is the port number on the line card (1-32). For example, the first port of the line card in slot 1 is fc1/1, and the seventh port of the line card in slot 3 is fc3/7.

Fibre Channel IDs and persistent FCIDs

Contrary to other switch manufacturers, there is no direct correlation between physical Fibre Channel ports and Fibre Channel IDs (FCIDs). This is necessary to allow intermixing line cards with different number of ports, while being able to use all port addresses, to allow both fabric and loop devices to coexist, and also to allow switches larger than 256 ports (currently 224 possible ports on the 9509) in the future.

The following points apply to the FCID assignment for any VSAN:

- When an N_Port or NL_Port logs into the switch, it is assigned an FCID.
- N_Ports receive the same FCID if disconnected and reconnected to any port within the same switch and within the same VSAN.
- NL_Ports receive the same FCID only if reconnected to the same port within the same switch where the port was originally connected.

If the persistent FCIDs feature is not enabled for a VSAN, the following points apply:

- The WWN of the N_Port or NL_Port and the assigned FCID are stored in a volatile cache and are not saved across switch reboots.
- The switch preserves the binding of FCID to WWN on a best-effort basis.
- The volatile cache has room for a maximum of 4000 entries, and if the cache gets full, the oldest entries are overwritten.

If the persistent FCID feature is enabled for a VSAN, the following points apply:

- The FCID to WWN mapping of the WWNs currently in use is stored to a nonvolatile database, and is saved across reboots.
- The FCID to WWN mapping of any new device connected to the switch is automatically stored into the non-volatile database.
- You can also manually configure the FCID to WWN mappings if necessary.

Note: If you attach AIX 5L or HP-UX hosts to a VSAN, you must have persistent FCIDs enabled for that VSAN. This is because these operating systems use the FCIDs in device addressing. If the FCID of a device changes, the operating system considers it to be a new device and gives it a new name.

Port modes

The Fibre Channel ports in the Cisco MDS 9000 Family can operate in several modes. Table 11-1 describes the operational modes.

Table 11-1 Fibre Channel port operational modes

Mode	Description
E_Port	An expansion port (E_Port) interconnects two Fibre Channel switches, forming an ISL between an E_Port in each switch. The ISL belongs to a single VSAN and can also be connected to third-party switches.
F_Port	A fabric port (F_Port) connects the switch to a N_Port in a host or storage device using a point-to-point link. Only one N_Port can connect to the F_Port.
FL_Port	A fabric loop port (FL_Port) connects the switch to a public FC-AL loop. Only one FL_Port can be operational in a single FC-AL loop at any given time.
TE_Port	A trunking E_Port (TE_Port) interconnects two Fibre Channel switches, forming an extended ISL (EISL) between a TE_Port in each switch. The EISL can multiplex the traffic of several VSANs. However, the EISL is currently only available in the Cisco MDS 9000 Family of switches.
TL_Port	A translative loop port (TL_Port) connects the switch to a private FC-AL loop.
SD_Port	A SPAN destination port (SD_Port) acts as a snooper port, allowing the monitoring of the switch traffic with a standard Fibre Channel analyzer.

Mode	Description
B_Port	A bridge port (B_Port) is used to connect some SAN extender devices to the switch, instead of E_Port.
Fx_Port	A Fx_Port can operate as either a F_Port or FL_Port, depending on the device connected to it. The port mode is determined during interface initialization.
Auto	A port configured as auto can operate as an E_Port, F_Port, FL_Port, or TE_Port, depending on the device connected to it. The port mode is determined during interface initialization.

11.1.2 Zoning

The Cisco MDS 9000 Family zoning can be administrated from any switch in the fabric, and all changes are automatically distributed to all of the switches.

The Cisco MDS 9000 Family supports zoning by the following criteria:

- Port world wide name (pWWN): The WWN of the Nx_Port (device) attached to the switch
- Fabric pWWN (fWWN): The WWN of the fabric port (port-based zoning)
- FCID: The FCID of the N_Port attached to the switch

To make management of zoning easier, the Cisco MDS 9000 Family supports alias names for all of these elements.

The Cisco MDS 9000 Family supports a default zone. All ports and WWNs not assigned to any zone belong to the default zone. If zoning is not activated, all devices belong to the default zone. You can control access between default zone members by a default zone policy. This is both a per-switch and per-VSAN setting.

The Cisco MDS 9000 Family supports both soft and hard zoning.

Note: Although the MDS supports soft zoning, it is not used, and there are very few conditions that can cause the MDS to “fall back” to soft zoning.

Soft zoning

In soft zoning, zoning restrictions are applied during the interaction between the name server and the end device. If an end device somehow manages to know or guess the FCID of another end device, it can access that device.

Hard zoning

In hard zoning, the zoning is enforced for each frame sent by an Nx_Port as the frame enters the switch. This prevents any unauthorized access at all times. The enforcement is done by the switch hardware at wire speed.

11.1.3 VSAN

A virtual storage area network (VSAN) is a unique feature of the Cisco MDS 9000 series that enables dividing the physical Fibre Channel fabric to virtual SAN fabrics. Each VSAN is a completely separate SAN fabric, with its own set of domain IDs, fabric services, zones, namespace, and interoperability mode.

Each port in the switch fabric belongs to exactly one of the VSANs at any given time, with the exception of trunking E_Ports (TE_Ports) that can multiplex the traffic of several VSANs over a single physical link.

Up to 256 VSANs can be configured in a single switch. The VSAN numbers can range from 1 to 4094. VSAN number 1 is called the default VSAN, and is the VSAN that initially contains all of the ports in the switch. If you do not need to divide the fabric into VSANs, you can leave all ports in the default VSAN.

The VSAN number 4094 is called the isolated VSAN, and any port configured into that VSAN is isolated from all other ports. If you delete a VSAN, all ports in it are moved to the isolated VSAN to avoid implicit transfer of the ports to the default VSAN.

Note: A good management practice is to move all unused ports out of VSAN 1 to prevent accidental usage if VSAN 1 is activated.

11.1.4 Trunking and PortChannel

In Cisco terminology, the term *trunking* is used to describe a single trunking E_Port (TE_Port) with the remit to multiplex the traffic of more than one VSAN on a single physical interface. This is in contrast to other Fibre Channel switch manufacturers who use that term (trunking) to describe the aggregation of several physical interfaces into a single logical interface. Cisco calls this latter feature *PortChannel*.

Trunking and PortChannel features are available for both Fibre Channel and Gigabit Ethernet interfaces on the Cisco MDS 9000 Family. Because the configuration rules for these features are different, we describe both of them separately.

FC trunking

Trunking, also known as VSAN trunking, enables interconnect ports to transmit and receive frames in more than one VSAN over the same physical link. In this case, the link is configured as an extended ISL (EISL) link using the EISL frame format.

Trunking is only applicable to E_Ports and used for inter-switch connections. Trunking is normally enabled for all ports in the switch but can be disabled on a port-by-port basis. If the port becomes operational as a trunking E_Port, it is referred to as a TE_Port. If a port, with trunking enabled, is connected to a third-party switch, it works as a normal E_Port.

FC PortChannel

The PortChannel feature can be used to aggregate up to 16 ISL or EISL links into a single logical link. The Fibre Channel ports can be any Fibre Channel ports in any 16-port Fibre Channel line card.

A PortChannel can consist of ports from any module within the limits of that module (that is to say, a 32-port module can only have one port of a port group participate in a PortChannel).

The PortChannel feature increases the available aggregate bandwidth of the logical link because the traffic is distributed among all functional links in the channel. It also provides high availability, because the channel remains active as long as at least one of the links forming it remains active and the traffic is transparently distributed over the remaining links.

Because PortChannel can be built on EISL links, both trunking and PortChannel are supported simultaneously.

If you need more detailed information regarding any of the mentioned topics, refer to the IBM Redbook *IBM TotalStorage: Implementing an Open IBM SAN*, SG24-6116.

11.1.5 iSCSI and FCIP support

The Cisco MDS 9000 series simultaneously supports both iSCSI and FCIP on the 8-port IP line card, the 14+2 line card, and 9216i.

iSCSI

The iSCSI support connects iSCSI-capable hosts to Fibre Channel storage devices. Support for iSCSI is included in the base price of the 4-port and 8-port IP line cards.

FCIP

The FCIP support connects separate SAN islands over an IP network. Each defined connection is a virtual E_Port (VE_Port) and can work as an E_Port or a TE_Port. Each Gigabit Ethernet interface can support up to three FCIP tunnels.

To use FCIP, depending on your particular configuration, you need to purchase a licence. The FCIP licence is a standard feature with MDS 9216i. For the rest of the switches, you need to purchase the corresponding licence feature that corresponds to your particular line card. For example, the corresponding FCIP Activation feature for the 8-port IP Services module is f/c 2210.

FCIP VE_Ports and TE_Ports can also be aggregated to form a PortChannel between Cisco MDS switches. The FCIP-based PortChannel can then be configured to carry specified VSAN traffic between switches in the same manner as that carried by conventional FC-based PortChannels.

11.2 Initial setup of the Cisco MDS 9000 products

Before you can manage the Cisco MDS 9000 series switch through the network, set up the TCP/IP parameters for the switch.

The first time the switch is turned on, it automatically runs the setup program and prompts you for the IP address and other configuration information necessary to communicate over the management Ethernet interface. You can also start the setup program with the **setup** command later if necessary.

11.2.1 Preparing to configure the switch

Before you configure the switch for the first time, you need to gather the following information:

- New administrator password
- Switch name
- IP address for the management Ethernet
- Subnet mask for the management Ethernet
- Default gateway IP address (optional)
- DNS server IP address (optional)
- NTP server IP address (optional)
- SNMP v3 secret key (optional)

11.2.2 Connecting to the switch through the serial port

To connect to the switch through the serial port, perform the following steps:

1. Connect the serial cable provided with the switch to the RJ-45 socket in the switch, using the console port in the:
   - Interface module in MDS 9216
   - Supervisor module in slot 5/6 in the MDS 9506, MDS 9509, and MDS 9513
2. Connect the other end of the serial cable to an RS-232 serial port on the workstation.
3. Disable any serial communication programs running on the workstation.
4. Open a terminal emulation application (such as HyperTerminal on a PC), and configure it as follows:

```
Bits per second: 9600
Data bits: 8
Parity: none
Stop bits: 1
Flow control: none
```

Figure 11-1 shows an example of the HyperTerminal serial port properties window.

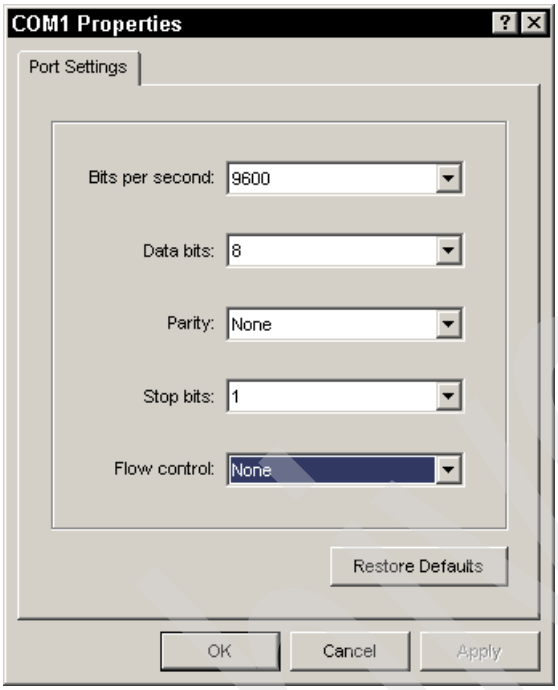


Figure 11-1 HyperTerminal serial Port Properties window

11.2.3 Setting up the initial parameters with the setup program

We assume that you are already connected to the console serial port of the switch and that the switch is still turned off. In the following example, we connect to a Cisco MDS 9216i.

Note: The steps you take might be different depending on which features you want to activate. However, the prompts of the setup program should be self-explanatory.

Note: Output is in **bold** for emphasis.

To set up the initial parameters:

1. Turn on the switch, and at the Admin account setup window, enter your admin password twice.

Example 11-1 Setup of initial parameters

```
Auto booting bootflash:/m9200-ek9-kickstart-mz.3.0.2a.bin
bootflash:/m9200-ek9-
mz.3.0.2a.bin...
Booting kickstart image:
bootflash:/m9200-ek9-kickstart-mz.3.0.2a.bin....
.....................................ImageverificationOK

Starting kernel...
INIT: version 2.78 booting
Checking all filesystems..... done.
Loading system software
Uncompressing system image: bootflash:/m9200-ek9-mz.3.0.2a.bin
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
CCCCCCCCC
CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
Uncompressing linecard components
INIT: Entering runlevel: 3

               ---- System Admin Account Setup ----


  Enter the password for "admin":
Confirm the password for "admin":

         ---- Basic System Configuration Dialog ----

This setup utility will guide you through the basic configuration of
the system. Setup configures only enough connectivity for management
of the system.

Please register Cisco MDS 9000 Family devices promptly with your
supplier. Failure to register may affect response times for initial
service calls. MDS devices must be registered to receive entitled
support services.

Press Enter at anytime to skip a dialog. Use ctrl-c at anytime
to skip the remaining dialogs.
```

2. Enter **yes** to enter setup mode:

   ```
   Would you like to enter the basic configuration dialog (yes/no): yes
   ```

3. Take the defaults for the next three questions, or enter the data if needed:

```
Create another login account (yes/no) [n]:

Configure read-only SNMP community string (yes/no) [n]:

Configure read-write SNMP community string (yes/no) [n]:
```

4. Enter the switch name:

```
Enter the switch name : IBM_FCIP_1
```

5. Take the default yes value and enter the IP, subnet, and gateway address:

```
Continue with Out-of-band (mgmt0) management configuration? (yes/no)
[y]:

Mgmt0 IPv4 address : 10.1.1.10

Mgmt0 IPv4 netmask : 255.255.255.0

Configure the default gateway? (yes/no) [y]:

IPv4 address of the default gateway : 10.1.1.1
```

6. Select the options you want to configure, or take the defaults as shown here:

```
Configure advanced IP options? (yes/no) [n]:

Enable the telnet service? (yes/no) [y]:

Enable the ssh service? (yes/no) [n]:

Configure the ntp server? (yes/no) [n]:

Configure default switchport interface state (shut/noshut) [shut]:

Configure default switchport trunk mode (on/off/auto) [on]:

Configure default zone policy (permit/deny) [deny]:

Enable full zoneset distribution? (yes/no) [n]:
```

7. When shown the summary, if it is all correct, select no:

```
The following configuration will be applied:
  switchname IBM_FCIP_1
  interface mgmt0
    ip address 10.1.1.10 255.255.255.0
    no shutdown
  ip default-gateway 10.1.1.1
  telnet server enable
  no ssh server enable
  system default switchport shutdown
  system default switchport trunk mode on
  no system default zone default-zone permit
  no system default zone distribute full

Would you like to edit the configuration? (yes/no) [n]:
```

8. Select yes to save and run this configuration:

```
Use this configuration and save it? (yes/no) [y]:
```

Note: If you do not save the configuration at this point, none of your changes are updated and the next time the switch is rebooted, you have to enter all the information again. Ensure that you type yes here to save the new configuration.

9. Wait until the configuration has been saved and the command prompt displays your newly updated switch name as an additional confirmation that your updates have been successfully applied:

```
####################################] 100%

MDS Switch
IBM_FCIP_1 login:
```

10. Your basic configuration is now finished, and you can install the Cisco Fabric Manager and Device Manager.

11.2.4 Installing the Cisco Fabric Manager and Device Manager

To install the Cisco Fabric Manager on your workstation, you need Java Runtime Environment (JRE™) 1.4 or later and Java Web Start. If you do not have these, you will be prompted to download them during the installation. Our workstation already had the required versions, so the additional downloads were not required.

Tip: Before you start, check IP connectivity to the switch by pinging it from your workstation using the new IP address that you just configured.

Start your Web browser and type the newly configured switch IP address in the URL bar. This takes you directly to the Fabric Manager installation page, as shown in Figure 11-2.

Figure 11-2 The installation page for Cisco Fabric Manager software

We recommend that you install the JRE and Java Web Start using the link provided by the installation page if you have access to the Internet from this server. If you do not, you must exit from this session and install the requested level of Java Web Start and JRE before returning to the Fabric Manager installation.

11.2.5 Installing Fabric Manager

To install Fabric Manager, complete the following steps:

1. Start the installation of Fabric Manager from the Web browser by clicking the **Fabric Manager** link.

   If this is a first time installation of the Cisco applications you might get the security warning shown in Figure 11-3. Click the **Start** button to proceed.

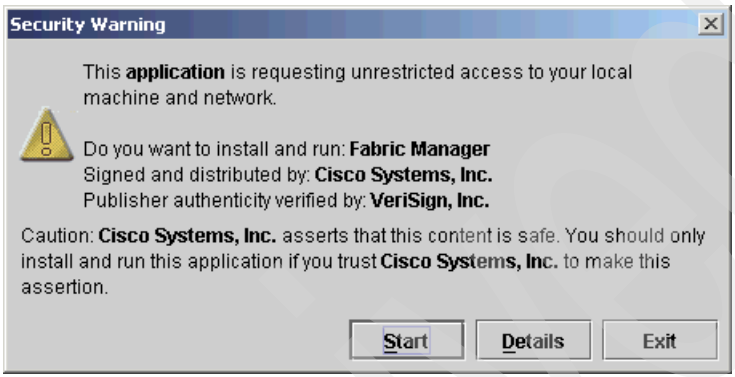


Figure 11-3 Fabric Manager Security Warning

2. Change the defaults if you want. Click **Finish (**Figure 11-4).

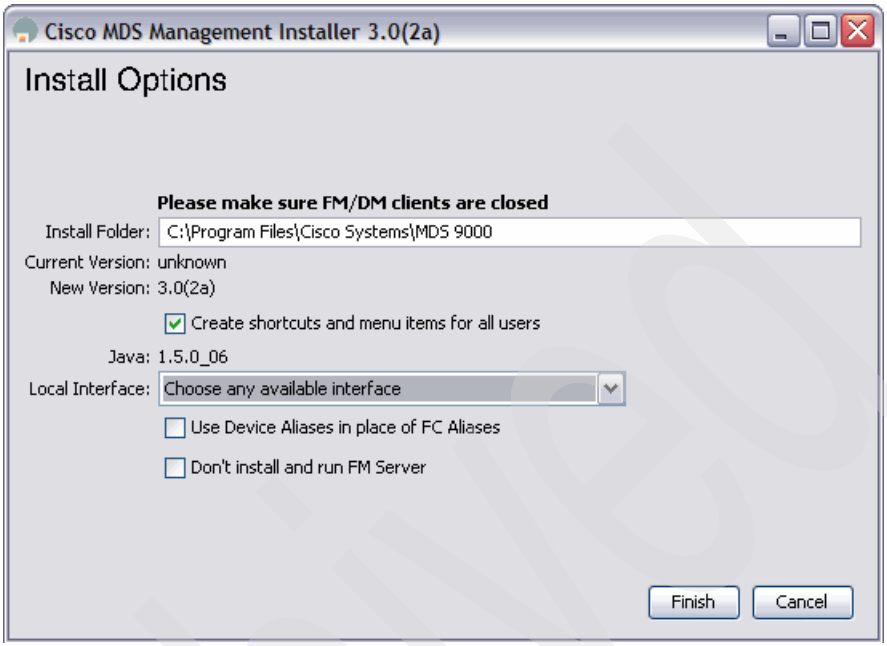


Figure 11-4 Initial FM installation panel

3. When the installation completes, you should see the Fabric Manager login window, as shown in Figure 11-5.

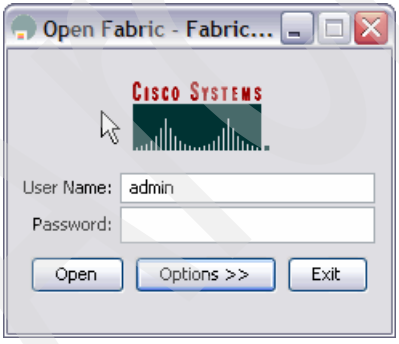


Figure 11-5 Fabric Manager initial logon window

The Fabric Manager is now ready for use.

Click **Options** to input the IP of the switch you want to manage, as shown in Figure 11-6.

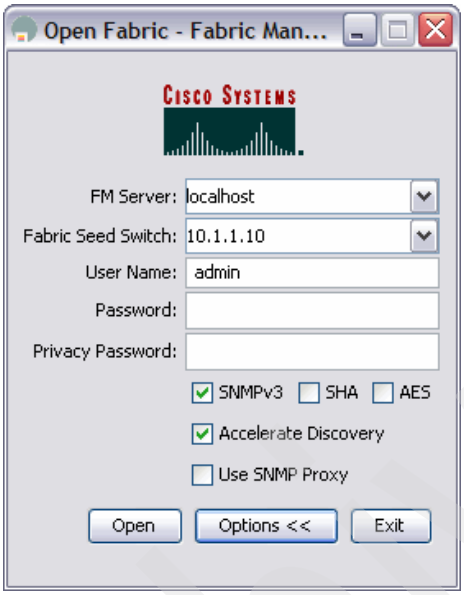


Figure 11-6 FM Options

Enter your logon credentials, a window similar to the one in Figure 11-7 opens.

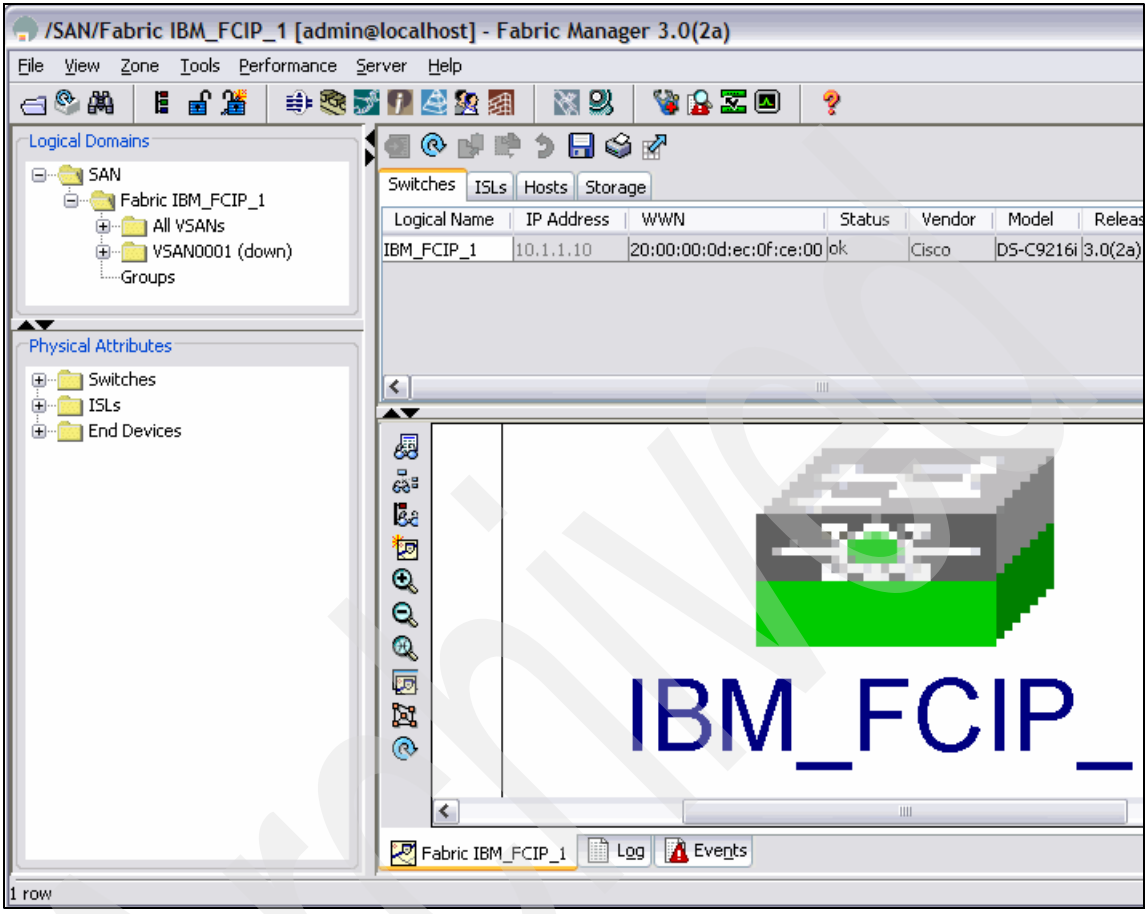


Figure 11-7 Fabric Manager

11.2.6 Installing Device Manager

To install Device Manager, complete the following steps:

1. Start the installation of Device Manager from the Web browser by clicking the **Cisco Device Manager** link, as shown in Figure 11-2 on page 311.

2. When the installation completed, the Device Manager login window opens, as shown in Figure 11-8.

Figure 11-8 Device Manager initial login window

The Device Manager is now ready for use.

Entering your credentials again produces something similar to Figure 11-9. If you are not working with a Cisco MDS 92116i, your display will be slightly different.

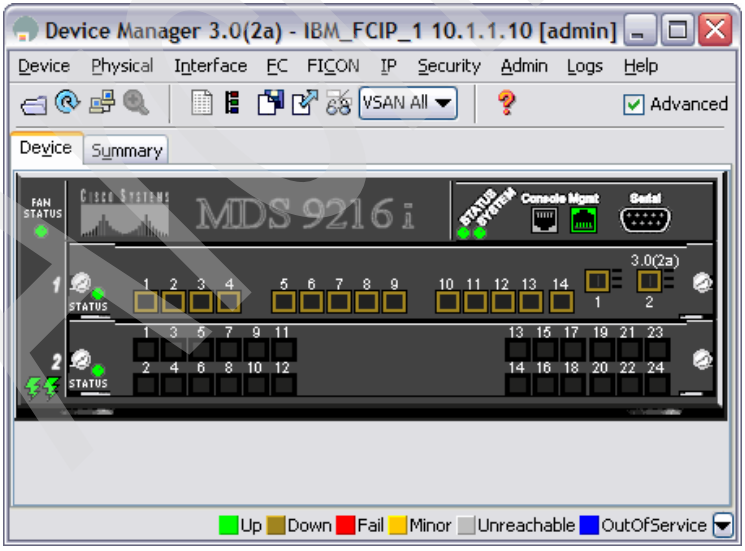


Figure 11-9 Device Manager

Note: The user ID and password that is required to access both Fabric Manager and the Device Manager is the one you created during the initial setup process.

11.3 Managing the Cisco SAN with Fabric Manager

Fabric Manager is a centralized tool used to manage the Cisco SAN fabric and the devices connected to it.

11.3.1 Getting started

You can start the Fabric Manager from the icon on your desktop or Windows Start menu. Enter the IP address or host name of your switch, the user name, and password and click **Open**.

11.3.2 User interface

When you start the Fabric Manager, you will see the logical view of your switch fabric, as shown in Figure 11-10.

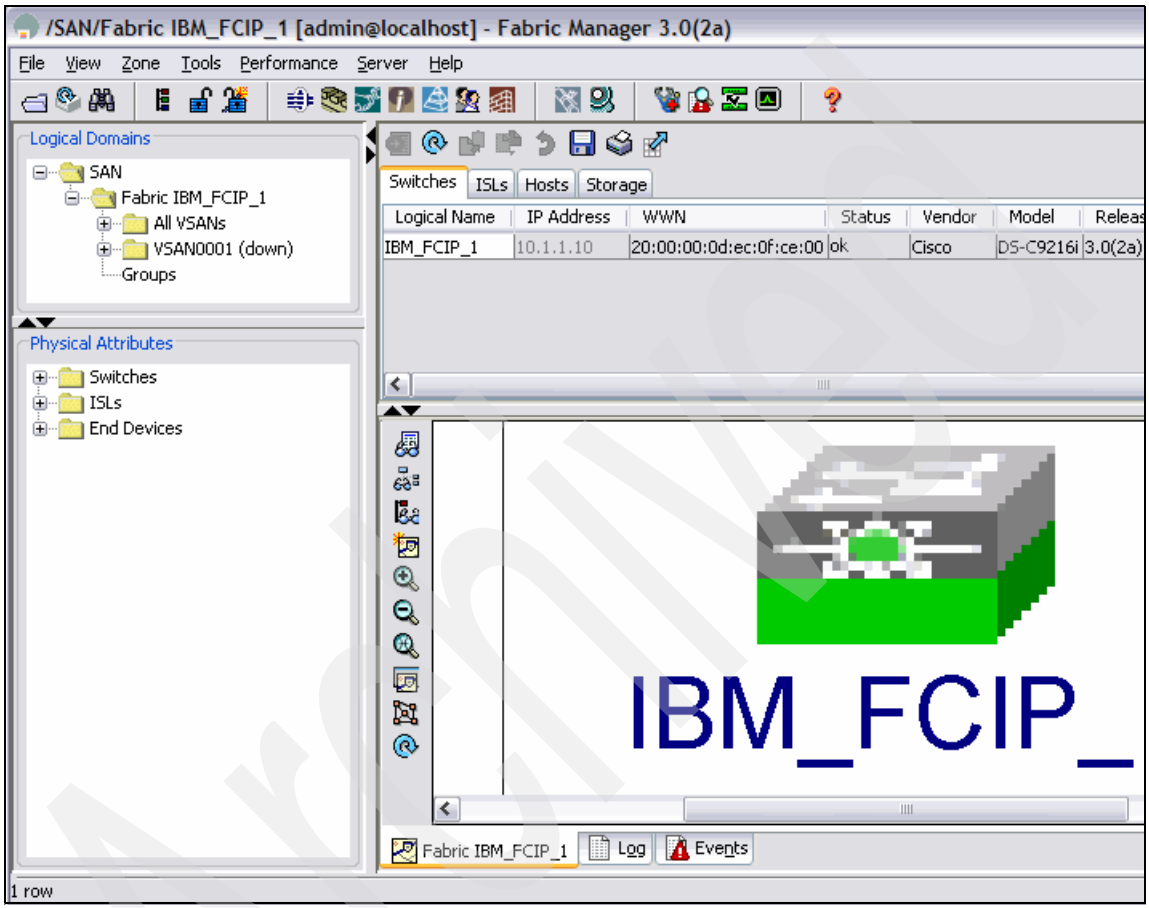


Figure 11-10 Fabric Manager logical view

The window contains the graphical representation of your switch fabric on the bottom right, an information area on the top right, and a navigation menu on the left. Any of the areas can be hidden to give more space to the other windows. The content of the information area changes automatically to represent the selection chosen in the navigation menu, and the current selection is shown on top of the information area.

There are two navigation panes available, and the panes are displayed on the left side of the window. The Logical Domains pane is a representation of the VSANs

defined in the network and the zone sets, zones, and zone members within each VSAN. The Physical Attributes pane is a representation of all the physical assets in the SAN and can also be used to configure most operating parameters of all of the switches in the SAN.

SNMP timeouts

The Fabric Manager uses the SNMP protocol to communicate with the switch. SNMP is a stateless protocol, and when you apply changes to the switch, the Fabric Manager sends a request packet with the changes to the switch and waits for a response packet.

Depending on your network, either the request packet or the response packet might end up being dropped. This results in a SNMP timeout message.

If you get an SNMP timeout message, you do not know which of the packets was dropped. This means that you do not know if your changes were applied to the switch or not. We recommend that you click the **Refresh Values** button, as shown in Figure 11-11, to ensure that the information in the Fabric Manager is up to date before making any further changes.

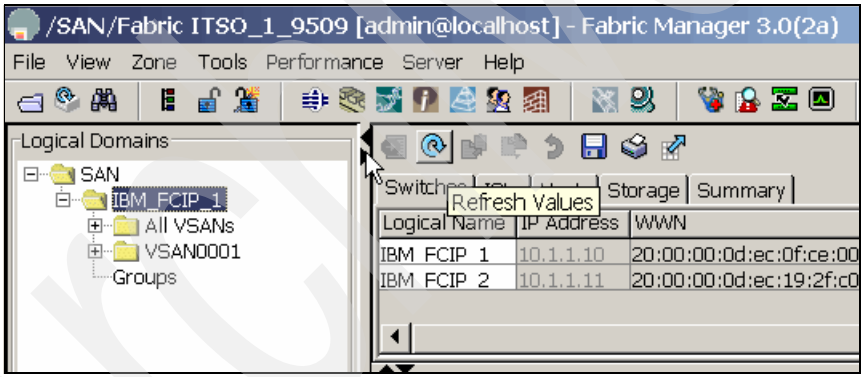


Figure 11-11 Refresh Values being displayed

Stopping Fabric Manager

If you made changes to the Cisco running configuration that have not yet been copied to the startup configuration, you will get a message similar to that shown in Figure 11-12 when you exit from or leave a Fabric Manager session.

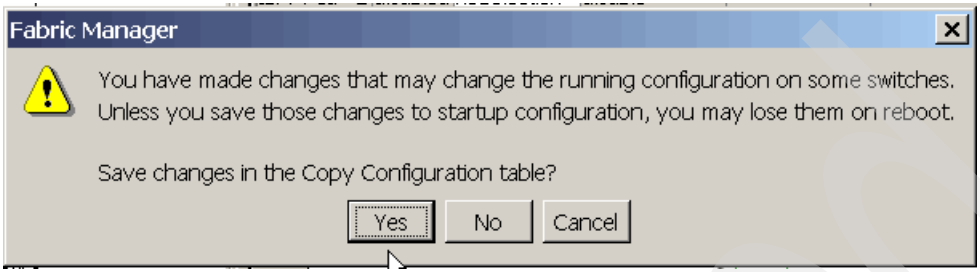


Figure 11-12 Unsaved running configuration warning

Click **Yes** to go to the Copy Configuration window, and then click **Apply Changes** to do the actual copy, and wait for the copy processes to finish. After all of the copy processes are finished, close the Fabric Manager.

You can also display the Copy Configuration panel from the Physical Attributes - Switches pane, as shown in Figure 11-13. Here, you can place check marks in the appropriate row and click **Apply Changes**.

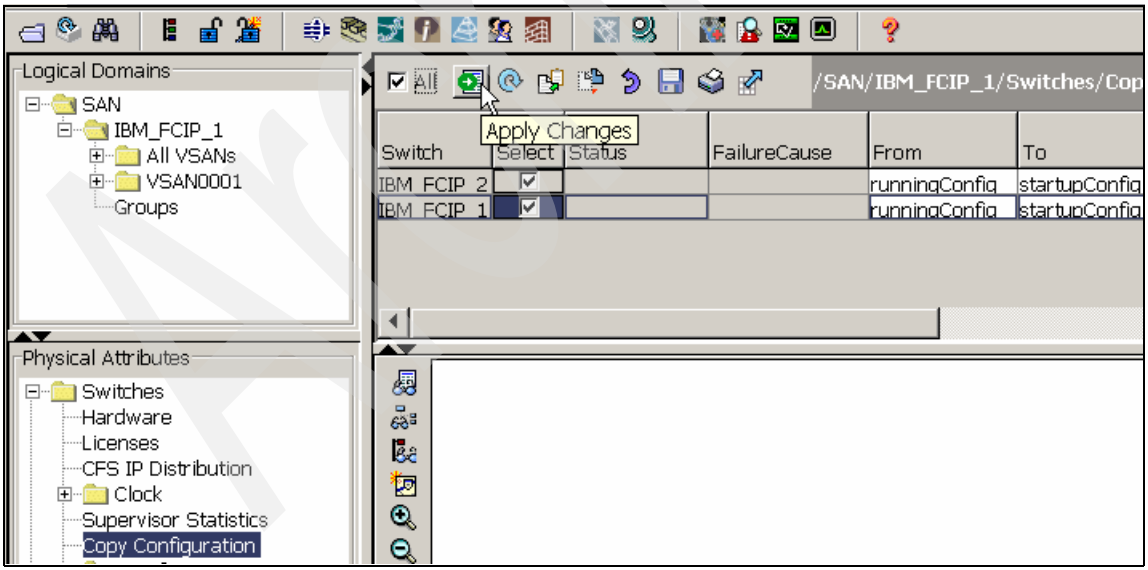


Figure 11-13 Saving running configuration changes

11.4 Managing zones and zone sets

On the switches in the Cisco MDS 9000 Family, each VSAN has its own zones and zone sets. Only one zone set can be active in a VSAN at any given time. The zone set can contain multiple zones, and a zone can belong to multiple zone sets.

We build the open systems topology shown in Figure 11-14. Our topology is very simple. It consists of two Cisco MDS 9216i switches linked through two FCIP connections configured into a single PortChannel. This type of configuration is typical for a remote DR type scenario, where one MDS 9216i is at the production site, and the other is at the DR site. We do have two DS4300s in our fabric, but these are not the focus of this document.

Tip: For further information about VSANs, zoning, zone sets, and so on, see the IBM Redbook *IBM TotalStorage: Implementing an Open IBM SAN*, SG24-61166.

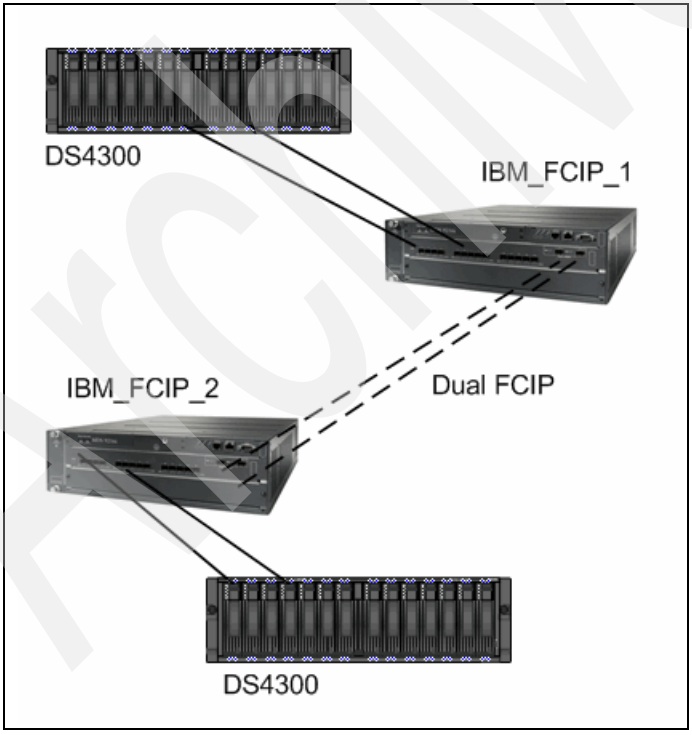


Figure 11-14 Working topology

11.4.1 Updating firmware

It is best to go to the IBM Web site and check to see if your switch has the latest available firmware.

The following IBM Web site provides the links to all the IBM SAN switches:

`http://www.ibm.com/servers/storage/support/san/index.html`

To update the firmware:

1. We select the **Cisco MDS 9216i Fabric Switch** from the list and go to the **Download** tab to get to the following URL:

   `http://www.ibm.com/servers/storage/support/san/index.html`

2. Clicking the hotlink on the displayed page produces a warning about leaving the IBM site and linking directly to the Cisco site. Click **Continue**.

3. Enter your CCO ID and password. If you do not have one, follow this link to sign up:

   `http://tools.cisco.com/RPF/register/register.do`

4. After logging in, you need to select your switch platform. The SAN-OS releases open. These releases consist of a *kickstart* image and a *system* image for each level of code, so you need to download both files. At the time of writing, the latest SAN-OS was v3.0(2a). Figure 11-15 gives a sample of what you will find on the Cisco code Web site. You can see both the kickstart and system images for 3.0(2a), as well as earlier levels.

Select a File to Download

Sort by : Filename Go

Filename	Release	Date	Size (Bytes)
m9200-ek9-kickstart-mz.3.0.2a.bin Kickstart image for Cisco MDS 9200 Series Fabric Switches	3.0(2a)	18-JUL-2006	14655488
m9200-ek9-mz.3.0.2a.bin System image for Cisco MDS 9200 Series Fabric Switches	3.0(2a)	18-JUL-2006	66678928
m9200-ek9-kickstart-mz.3.0.2.bin Kickstart image for Cisco MDS 9200 Series Fabric Switches	3.0(2)	08-MAY-2006	14725120
m9200-ek9-mz.3.0.2.bin System image for Cisco MDS 9200 Series Fabric Switches	3.0(2)	08-MAY-2006	66556293

Figure 11-15 Web code sample

   Download both of the images for the level you want to install.

5. Copy the code to your favorite type of server. The Cisco SAN-OS supports all of the following methods: TFTP, SFTP, SCP, and FTP.

Now that we have the needed code and it is being served by our server, we will open Fabric Manager and install it.

Important: We recommend that you read the *Release Notes* that are provided with your firmware download before proceeding.

6. Click the **FM** icon that was installed on your desktop.
7. With FM open, select the **Software Install** icon, as shown in Figure 11-16.

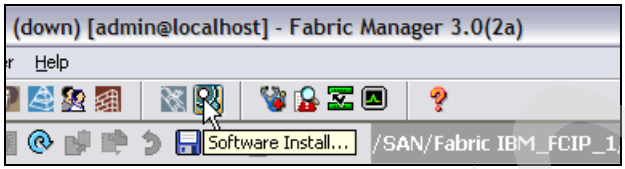


Figure 11-16 FW install selection

8. The first panel of the wizard asks you to select the switches you want to upgrade, as shown in Figure 11-17. Select your switch or switches and click **Next**.

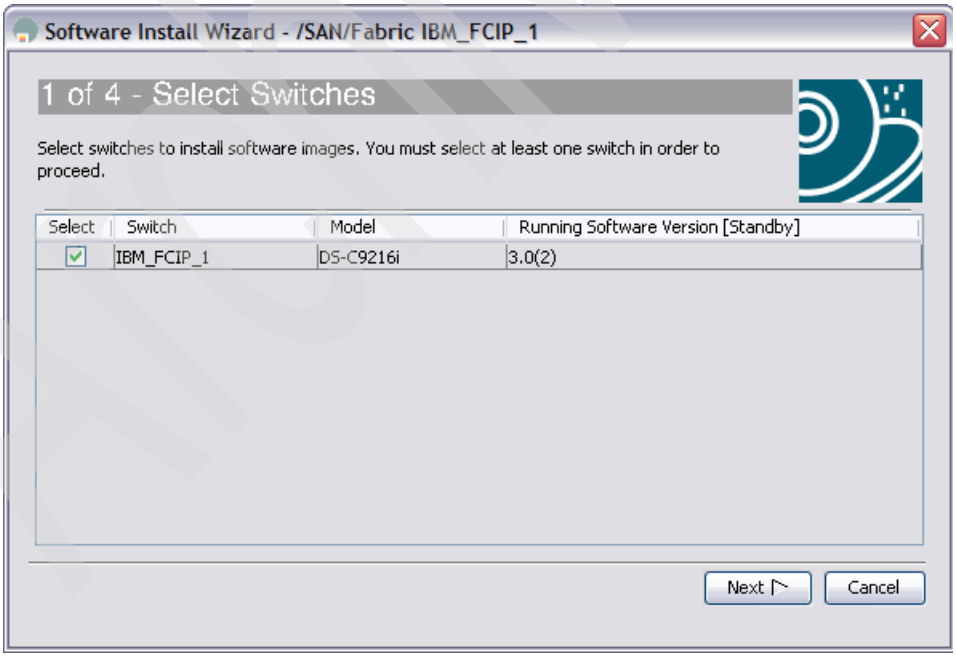


Figure 11-17 FM software update 1 of 4

9. Select the type of server to use (Figure 11-18). We use an FTP server. Enter the IP address of this server. Click **Verify Remote Server and Path** to ensure that you can reach your server.

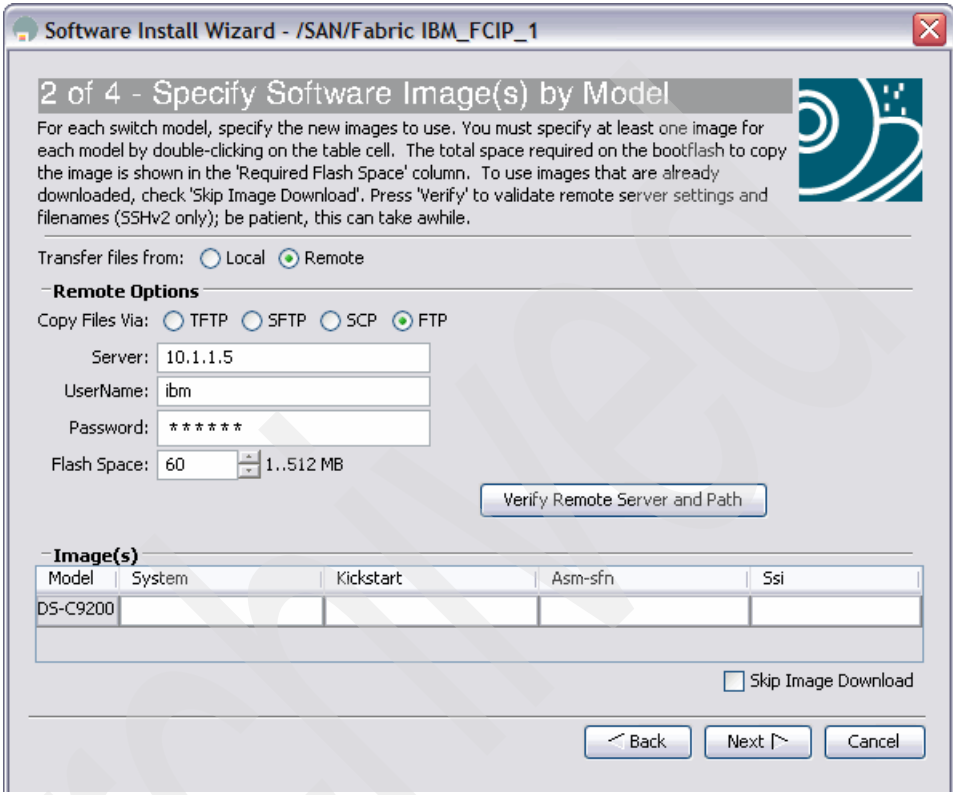


Figure 11-18 FM software update 2 of 4 with server credentials

Enter the full name of the code, as shown in Figure 11-19, and then click **Next**.

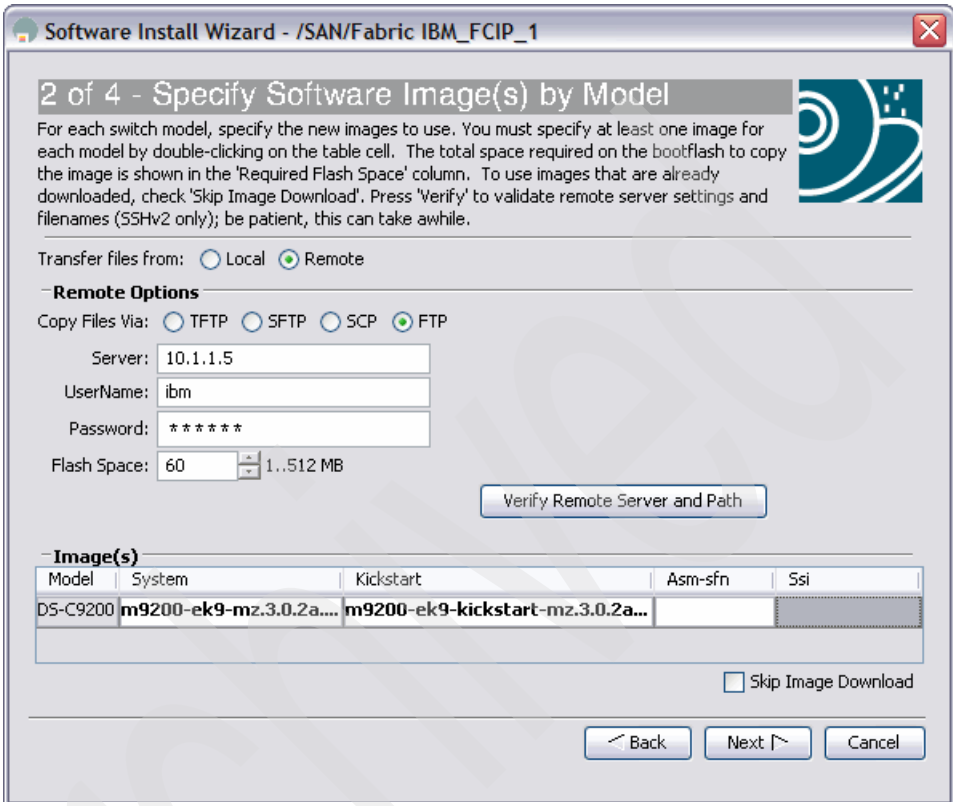


Figure 11-19 FM software update 2 of 4 with file names

10. The software has verified that we have enough space in our flash for our new images (Figure 11-20). Click **Next**.

    If your panel does not say Ok, clean up by selecting the "**...**" in the left corner as indicated by the cursor position.

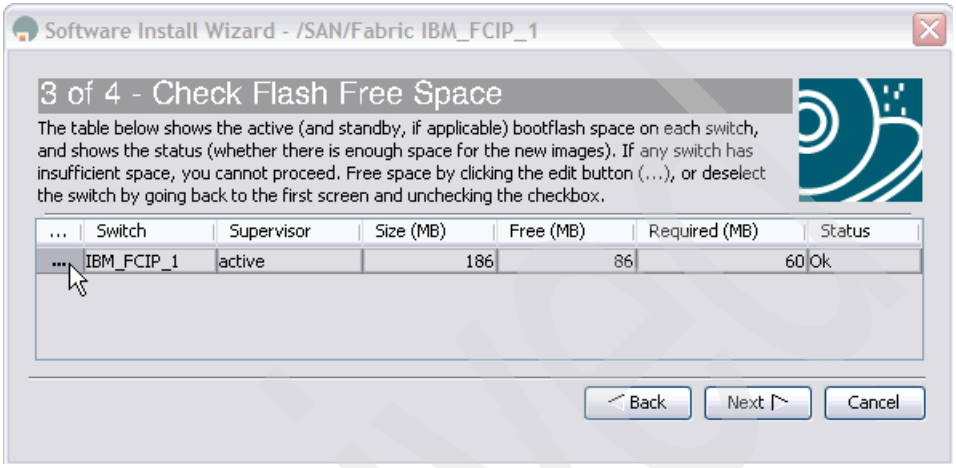


Figure 11-20 FM software update 3 of 4 flash check

Figure 11-21 lists the contents of our flash memory. Simply select the images you do not need and click **Delete**. Click **Close** to see if your flash is OK yet. If not, continue deleting until it is.

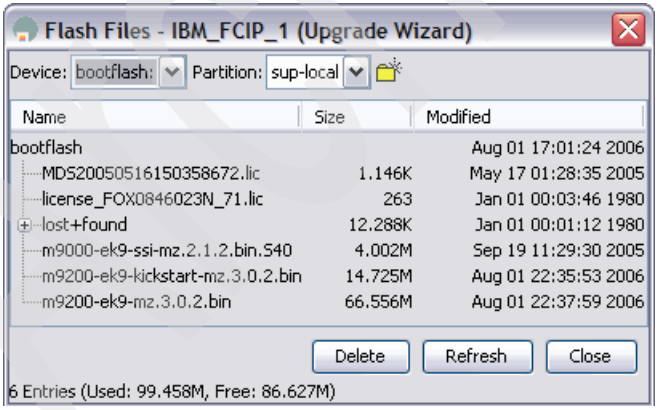


Figure 11-21 Flash memory content listing

11. In panel 4 of 4 (Figure 11-22), verify the contents, and click **Finish**.

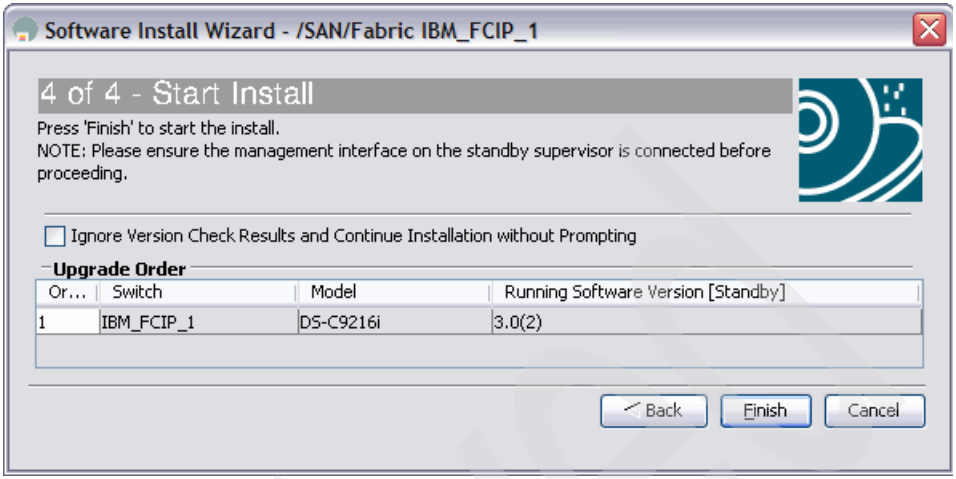


Figure 11-22 FM software update 4 of 4

Figure 11-23 shows the update progress.

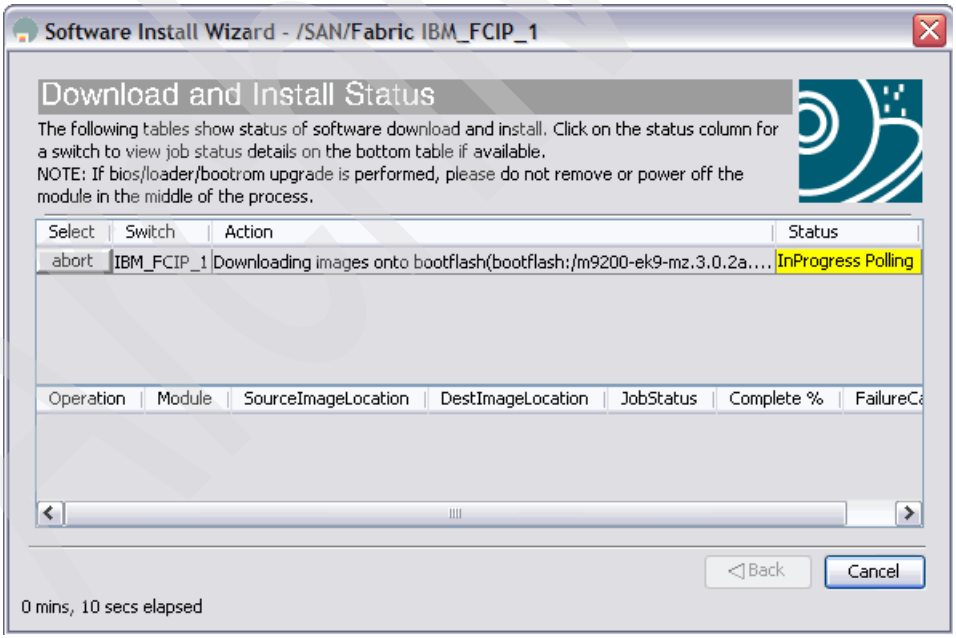


Figure 11-23 FM software update progress

The displayed information changes as the update progresses, as shown in Figure 11-24. In this case, it is performing a version compatibility check.

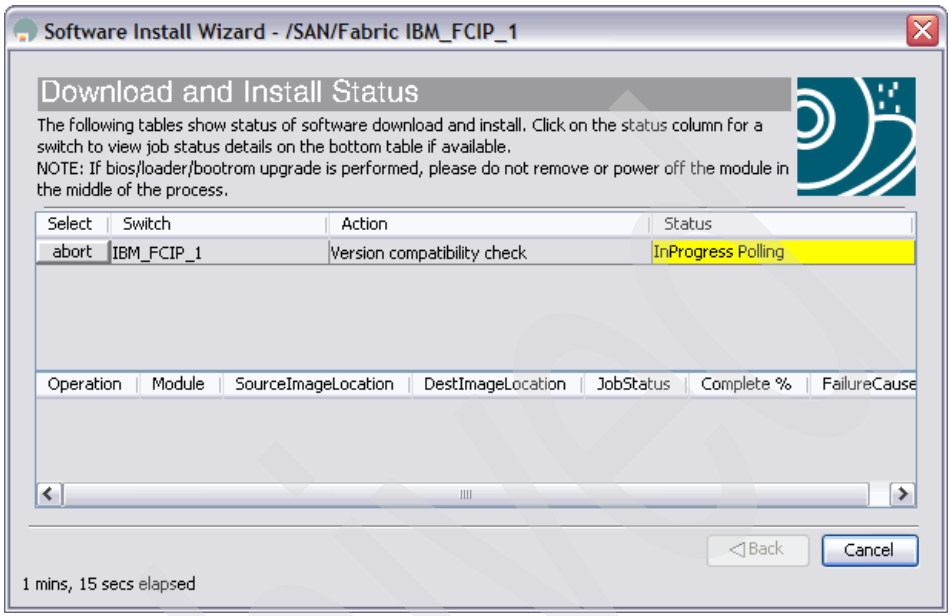


Figure 11-24 FM software update progress Version compatibility check

12. The window shown in Figure 11-25 opens, waiting for your input. Because we are upgrading a single supervisor switch (MDS 9216i) it states that the update will be disruptive. If you are updating a Cisco MDS 95xx switch with dual supervisors, it will show that the update is concurrent. Click **Yes** to allow the update to continue.

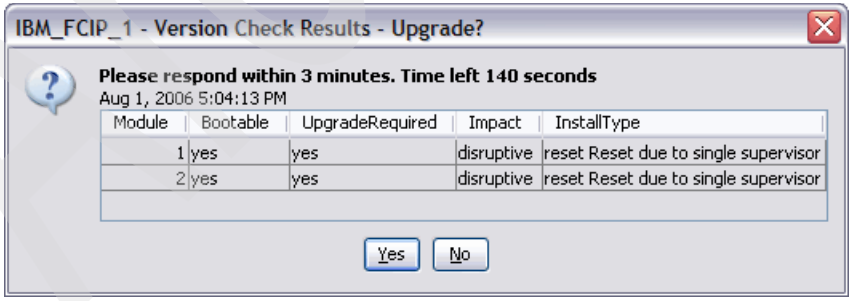


Figure 11-25 FM software update verify window

Figure 11-26 shows the installation progress.

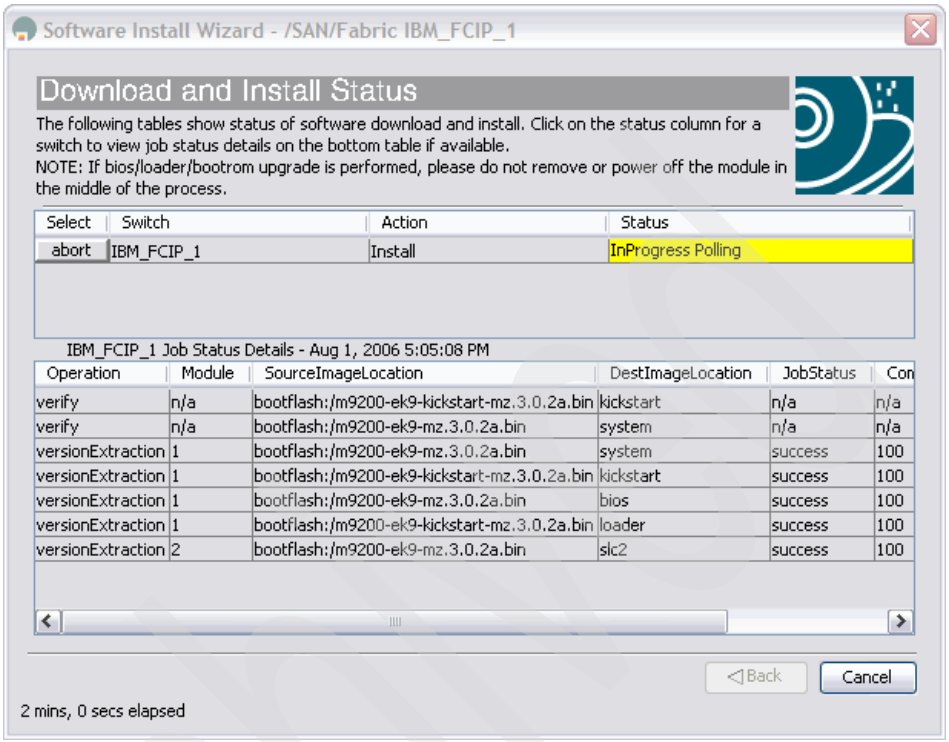


Figure 11-26 FM software update progress

After a period of time (usually measured in minutes), an update completed successfully window opens (Figure 11-27).

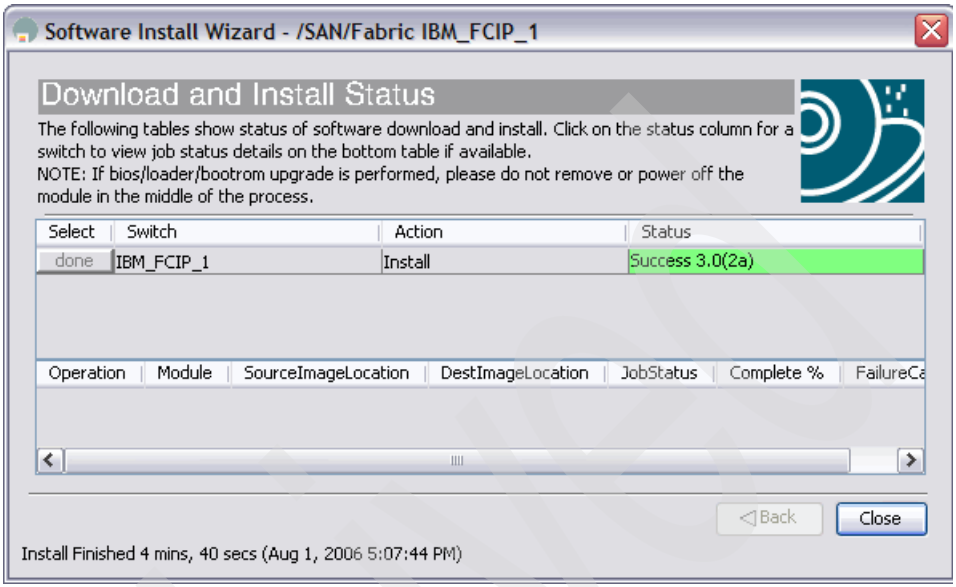


Figure 11-27 FM software update success

13. To verify the updates, select **Switches** → **Hardware** from the Physical Attributes pane and look in the pane above the SAN map, as shown in Figure 11-28. Look in the S/W Rev column as indicated.

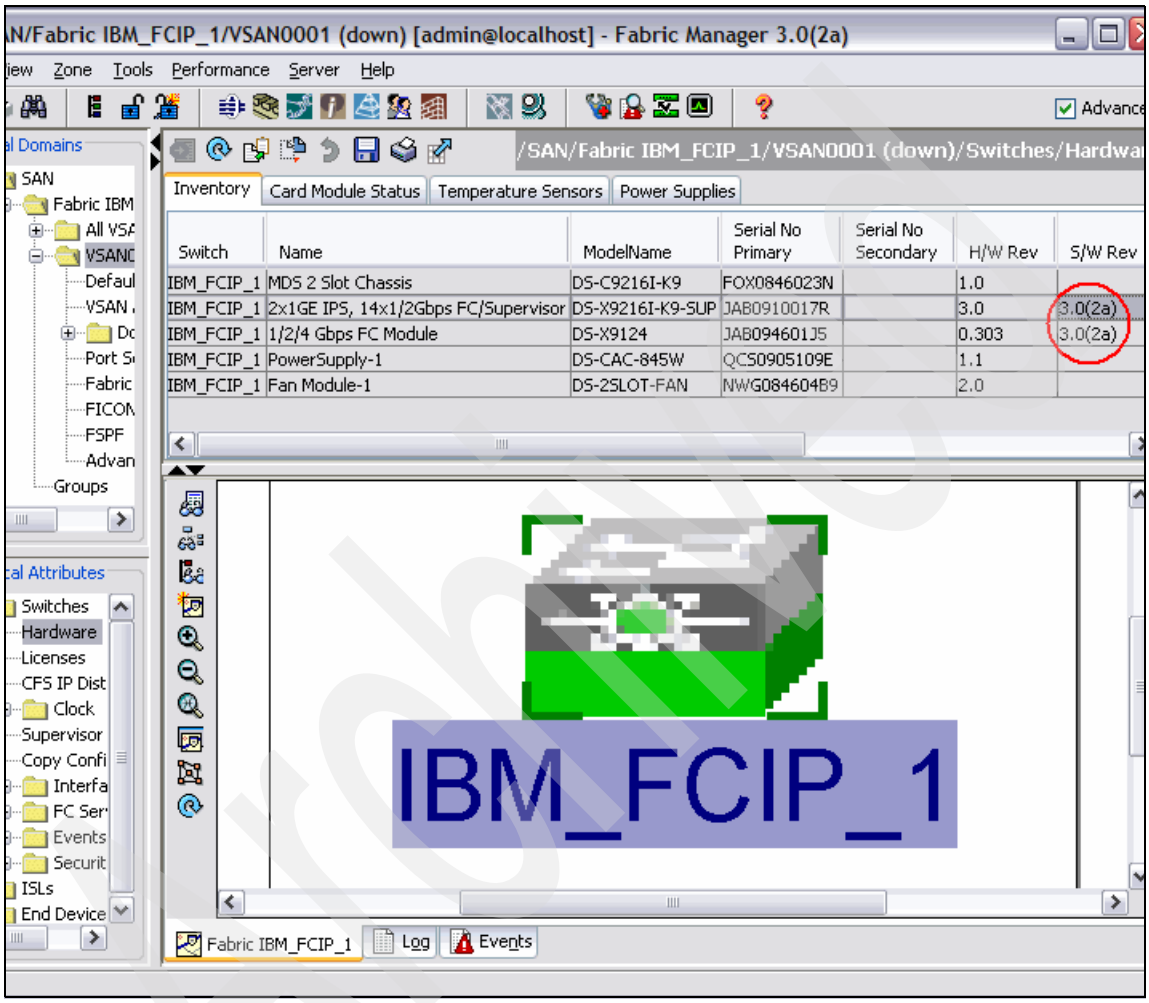


Figure 11-28 Verifying the software update

The switch is now at the latest firmware level. We recommend that you update the firmware yearly at an absolute minimum. Also, whenever you plan to add a new blade or start to use a new feature, it is a good practice to check your current firmware's compatibility and upgrade as needed.

12

Cisco FCIP implementation

FCIP provides the capability to extend a SAN over existing IP networks provided sufficient bandwidth is available. For short distances, SANs can be extended using traditional FC ISLs and multimode fiber. For longer distances, extend SANs using single-mode fiber with Coarse Wave Division Multiplexing (CWDM) or Dense Wave Division Multiplexing (DWDM) equipment. FCIP provides a third alternative for extending SANs over IP networks where IP is the most viable transport option either due to cost or distance.

Typical scenarios for FCIP SAN extension include, but are not limited to, data replication for disaster recovery using either a synchronous or asynchronous transmission. Synchronous transmission is very sensitive to latency, and because of this, FC ISLs, CWDM, or DWDM might be a better choice. Remote access to storage targets or control units from local initiators or channels is another application that can benefit from FCIP SAN extension.

In this chapter, we configure two FCIP links from the 9216i (IBM-FCIP-1) to the 9216i (IBM-FCIP-2), as shown in Figure 12-1. We then bundle these FCIP links into a PortChannel.

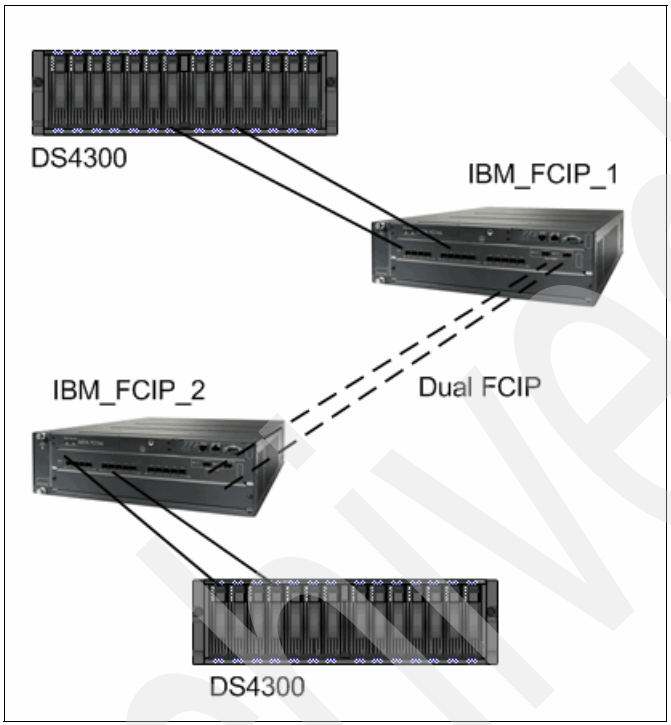


Figure 12-1 Test scenario

12.1 FCIP concepts

The following FCIP concepts were taken in their entirety from the “IP Services” part of the *Cisco MDS 9000 Family Fabric Manager Configuration Guide*, 78-16493-01, at the following URL:

http://www.cisco.com/en/US/products/ps5989/products_configuration_guide_book09186a0080667aa0.html

To configure the IPS module for FCIP, you need a basic understanding of the following concepts:

- FCIP and VE_Ports
- FCIP links
- FCIP profiles
- FCIP interfaces

FCIP and VE_Ports describe the internal model of FCIP with respect to Fibre Channel inter-switch links (ISLs) and Cisco's enhanced ISLs (EISLs).

FCIP defines virtual E (VE) ports, which behave exactly like standard Fibre Channel E_Ports, except that the transport in this case is FCIP instead of Fibre Channel. The only requirement is for the other end of the VE_Port to be another VE_Port. A virtual ISL is established over an FCIP link and transports Fibre Channel traffic. Each associated virtual ISL looks like a Fibre Channel ISL with either an E_Port or a TE_Port at each end.

FCIP links consist of one or more TCP connections between two FCIP link end points. Each link carries encapsulated Fibre Channel frames. When the FCIP link comes up, the VE_Ports at both ends of the FCIP link create a virtual Fibre Channel (E)ISL and initiate the E_Port protocol to bring up the (E)ISL. By default, the FCIP feature on any Cisco MDS 9000 Family switch creates two TCP connections for each FCIP link:

- One connection is used for data frames.
- The second connection is used only for Fibre Channel control frames, that is, switch-to-switch protocol (all class F) frames. This arrangement is used to provide low latency for all control frames.

To enable FCIP on the IPS module, an FCIP profile and FCIP interface (interface FCIP) must be configured. The FCIP link is established between two peers; the VE_Port initialization behavior is identical to a normal E_Port. This behavior is independent of the link being FCIP or pure Fibre Channel, and is based on the E_Port discovery process (ELP, ESC). After the FCIP link is established, the VE_Port behavior is identical to E_Port behavior for all inter-switch

communication (including domain management, zones, and VSANs). At the Fibre Channel layer, all VE and E_Port operations are identical.

Note: Table 11-1 on page 301 contains a listing and description of all the port modes.

The FCIP profile contains information about local IP address and TCP parameters. The profile defines the following information:

- The local connection points (IP address and TCP port number)
- The behavior of the underlying TCP connections for all FCIP links that use this profile

The FCIP profile's local IP address determines the Gigabit Ethernet port where the FCIP links terminate.

The FCIP interface is the local endpoint of the FCIP link and a VE_Port interface. All the FCIP and E_Port parameters are configured in context to the FCIP interface. The FCIP parameters consist of:

- The FCIP profile, which determines which Gigabit Ethernet port initiates the FCIP links and defines the TCP connection behavior
- Peer information
- Number of TCP connections for the FCIP link
- E_Port parameters, the trunking mode and trunk allowed VSAN list

12.2 FCIP licensing

The SAN Extension over IP (SAN_EXTN_OVER_IP) license for the IPS-8 module includes:

- FCIP protocol
- FCIP compression
- FCIP write acceleration

Note: For SAN-OS 2.0 and later, the SAN Extension licence also includes:

- IVR
- Tape acceleration

The 9216i is shipped with the required licences for FCIP. The included licence is "SAN Extension over IP package for integrated IP ports," and includes all of the protocols that are in the SAN_EXTN_OVER_IP licence.

The licenses are needed on a per module basis. For example, if you had two IPS-8 modules running FCIP in a switch, you need two SAN_EXTN_OVER_IP licenses. If you added an IPS-8 module to your MDS9216i, you still need the SAN_EXTN_OVER_IP license for that module.

You can verify license and feature information in Fabric Manager by opening the **Switches** folder in the Physical Attributes pane and selecting **Licenses**, as shown in Figure 12-2.

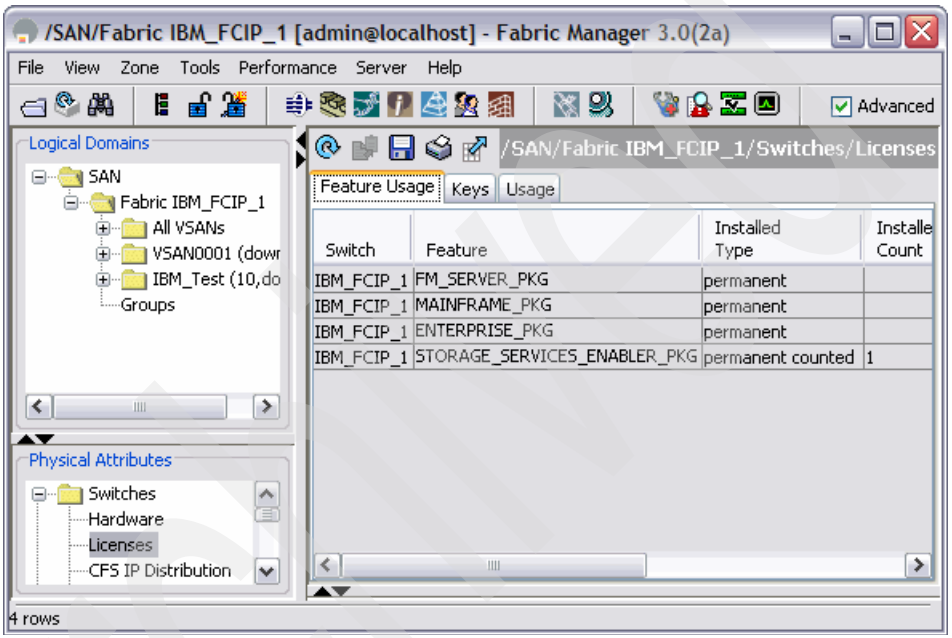


Figure 12-2 Verifying features in FM

Select the **Keys** tab to display the licence keys, as shown in Figure 12-3. Your display might look different, but will still display the switch licensed information.

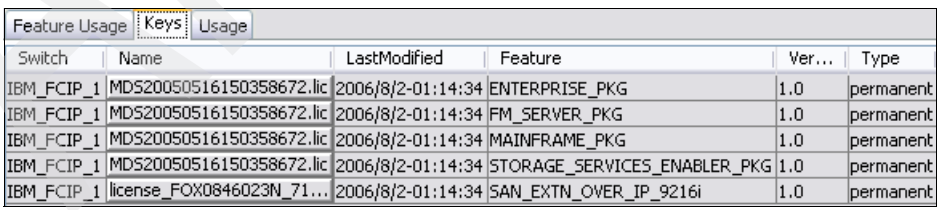


Figure 12-3 Licence Keys in FM

12.3 Configuring FCIP

In this section, we discuss FCIP configuration using FM, DM, and the CLI. We predominantly use the FM FCIP wizard to complete the configuration of FCIP. We then use the FM PortChannel wizard to bundle the two FCIP links into a single PortChannel. We then discuss configuring a single FCIP link with the CLI.

To configure FCIP:

1. The first step in configuring FCIP is to enable the GigE interfaces we want to use and setting the MTU. The path to do this in FM is in the Physical Attributes pane. Select **Switches** → **Interfaces** → **Gigabit Ethernet**, as shown in Figure 12-4.

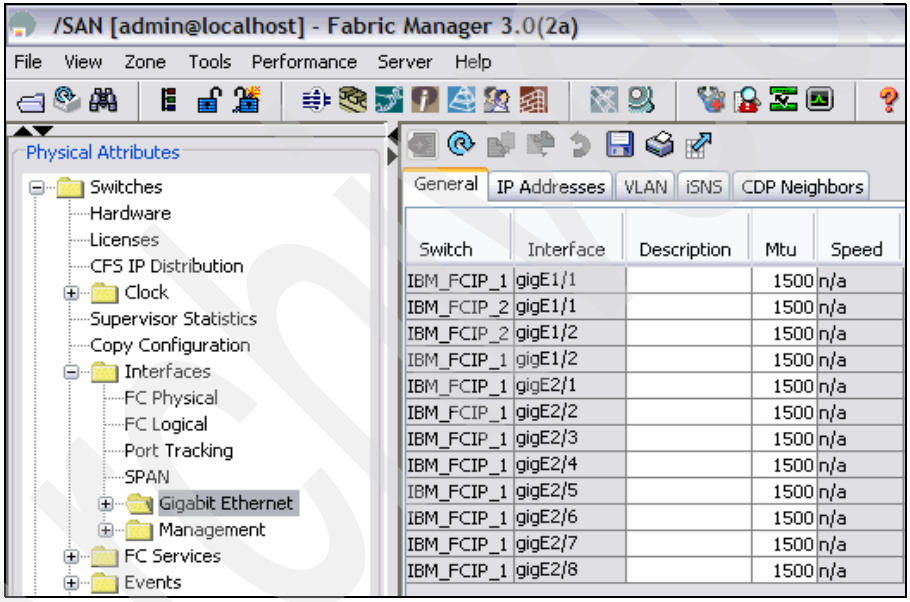
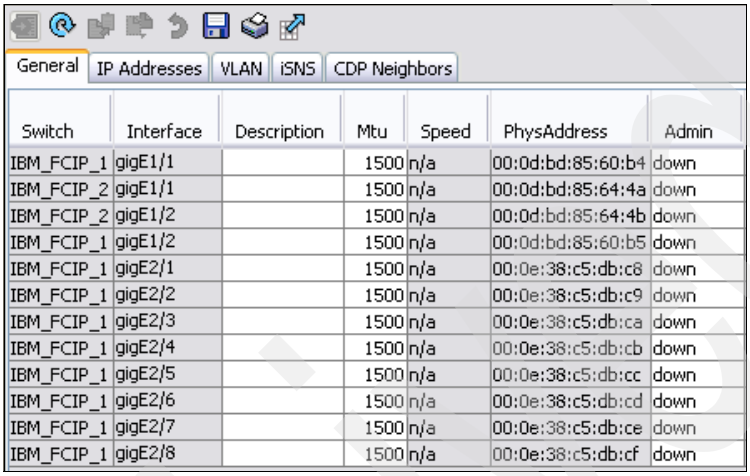


Figure 12-4 FM path to view and modify Gigabit Ethernet interfaces

This displays the information in the pane above the FM map, as shown in Figure 12-5.

Notice that we have an IPS-8 module in slot 2 of our chassis. We use a port from this line card for one of our connections. We could have used the two GigE ports that come standard with the MDS 9216i.



Switch	Interface	Description	Mtu	Speed	PhysAddress	Admin
IBM_FCIP_1	gigE1/1		1500	n/a	00:0d:bd:85:60:b4	down
IBM_FCIP_2	gigE1/1		1500	n/a	00:0d:bd:85:64:4a	down
IBM_FCIP_2	gigE1/2		1500	n/a	00:0d:bd:85:64:4b	down
IBM_FCIP_1	gigE1/2		1500	n/a	00:0d:bd:85:60:b5	down
IBM_FCIP_1	gigE2/1		1500	n/a	00:0e:38:c5:db:c8	down
IBM_FCIP_1	gigE2/2		1500	n/a	00:0e:38:c5:db:c9	down
IBM_FCIP_1	gigE2/3		1500	n/a	00:0e:38:c5:db:ca	down
IBM_FCIP_1	gigE2/4		1500	n/a	00:0e:38:c5:db:cb	down
IBM_FCIP_1	gigE2/5		1500	n/a	00:0e:38:c5:db:cc	down
IBM_FCIP_1	gigE2/6		1500	n/a	00:0e:38:c5:db:cd	down
IBM_FCIP_1	gigE2/7		1500	n/a	00:0e:38:c5:db:ce	down
IBM_FCIP_1	gigE2/8		1500	n/a	00:0e:38:c5:db:cf	down

Figure 12-5 FM Gigabit Ethernet display

2. We change the Mtu to 2300, set the Admin status to Up, and applied the changes to the ports in our configuration (Figure 12-6 on page 340). We chose 2300 as the Mtu value because the maximum Fibre Channel frame is 2148 bytes, as shown in Table 12-1.

Table 12-1 FC frame

Start of frame	Fibre Channel header	Payload	CRC	End of frame	Total bytes
4 bytes	24 bytes	2112 bytes	4 bytes	4 bytes	2148

If you then add the FCIP encapsulation usage, the new frame size is 2246 bytes, as shown in Table 12-2.

Table 12-2 FC frame with FCIP encapsulation

Max FC frame	FCIP header	TCP options	TCP header	IP header	Ethernet header and CRC	Total bytes
2148 bytes	28 bytes	12 bytes	20 bytes	20 bytes	18 bytes	2246

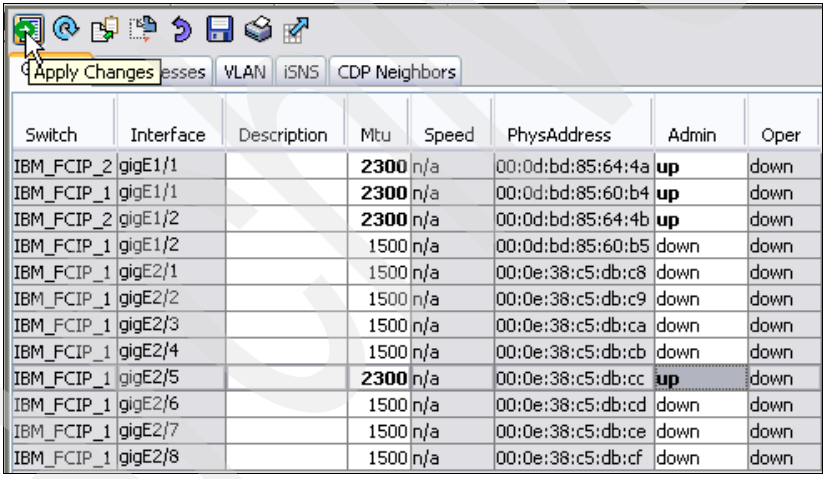
Finally, if you are using VSANs or 802.1q, or both, you have a total frame size of 2246-2258 byes, as shown in Table 12-3.

Table 12-3 FC frame using VSANs, 802.1q, or both

FCIP frame	Optional VSAN header for TE_Ports	Option 802.1q header for GigE	Total bytes
2246 bytes	8 bytes	4 bytes	2246-2258

Restriction: Jumbo frames are a technique for maximizing the throughput of Ethernet networks by increasing the frame size from the default 1518 bytes up to 9,000 bytes. Remember that if you use jumbo frames, they must be supported on each piece of networking equipment in the IP path.

Here we change our Mtu and Admin states. Simply, double-click the appropriate cells and either type in the changes or select the value from the drop-down list. Click the **Apply Changes** icon in the top-left corner to apply the changes.



Switch	Interface	Description	Mtu	Speed	PhysAddress	Admin	Oper
IBM_FCIP_2	gigE1/1		**2300**	n/a	00:0d:bd:85:64:4a	**up**	down
IBM_FCIP_1	gigE1/1		**2300**	n/a	00:0d:bd:85:60:b4	**up**	down
IBM_FCIP_2	gigE1/2		**2300**	n/a	00:0d:bd:85:64:4b	**up**	down
IBM_FCIP_1	gigE1/2		1500	n/a	00:0d:bd:85:60:b5	down	down
IBM_FCIP_1	gigE2/1		1500	n/a	00:0e:38:c5:db:c8	down	down
IBM_FCIP_1	gigE2/2		1500	n/a	00:0e:38:c5:db:c9	down	down
IBM_FCIP_1	gigE2/3		1500	n/a	00:0e:38:c5:db:ca	down	down
IBM_FCIP_1	gigE2/4		1500	n/a	00:0e:38:c5:db:cb	down	down
IBM_FCIP_1	gigE2/5		**2300**	n/a	00:0e:38:c5:db:cc	**up**	down
IBM_FCIP_1	gigE2/6		1500	n/a	00:0e:38:c5:db:cd	down	down
IBM_FCIP_1	gigE2/7		1500	n/a	00:0e:38:c5:db:ce	down	down
IBM_FCIP_1	gigE2/8		1500	n/a	00:0e:38:c5:db:cf	down	down

Figure 12-6 FM Gigabit Ethernet General tab

3. Next, we enable the FCIP feature. This must be done before we can configure FCIP. To enable FCIP using FM:

   a. From the Physical Attributes pane, open the **Switches** → **ISLs** folder, and select **FCIP**, as shown in Figure 12-7.

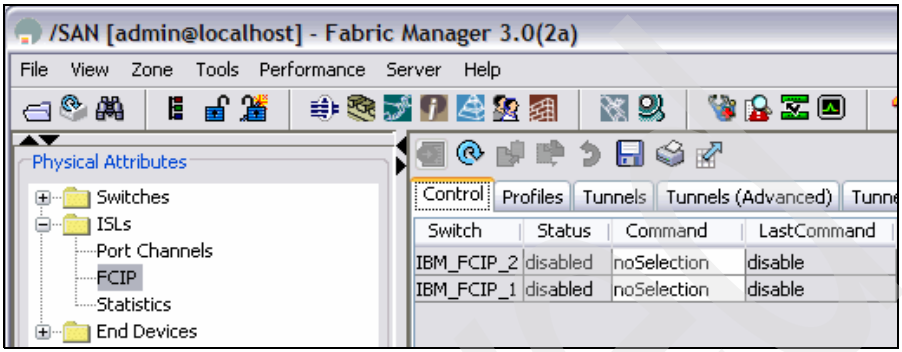


Figure 12-7 FM path to view or modify FCIP configuration

   b. The FCIP Control information tab should now be displayed. Use the pull-down list in the Command column, and select **enable**, as shown in Figure 12-8.

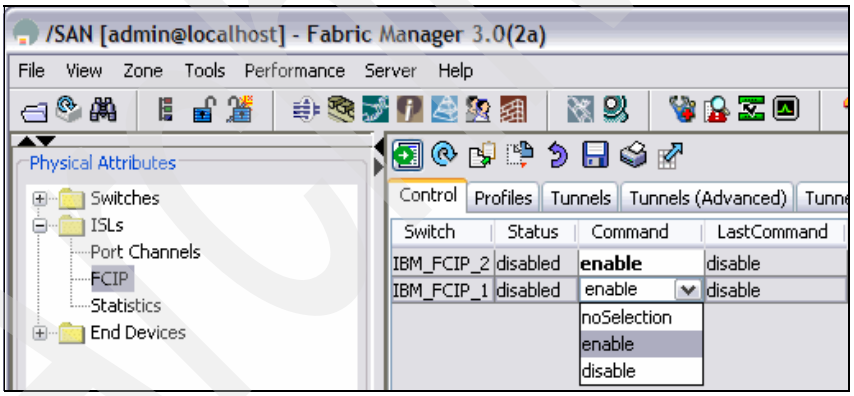


Figure 12-8 FM FCIP Control panel enable

c. Apply the changes by clicking the **Apply** icon. Figure 12-9 shows the FCIP protocol enabled for both our switches.

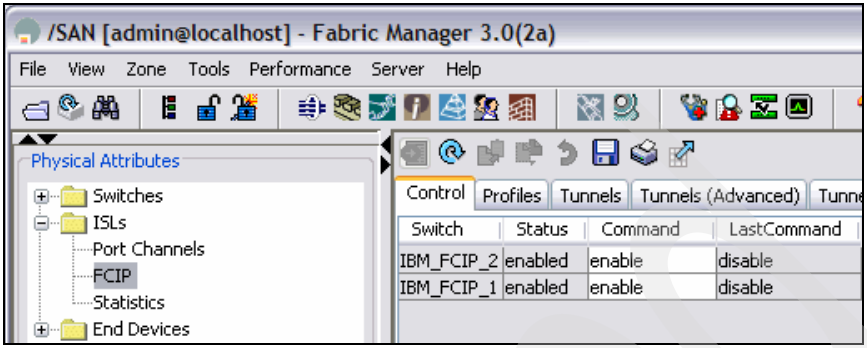


Figure 12-9 FM FCIP enable

4. Next, we select the FCIP tunnel wizard in FM. This automates the configuration process. We can run the FCIP tunnel wizard again if we want to make modifications to the original configuration.

> **Tip:** Before starting the FCIP tunnel wizard, ensure that you select **SAN** in the Logical Domains pane on the left side of FM, as shown in Figure 12-10 on page 342. This enables you to select your switches from the drop-down menus in the wizard. Also, ensure that you have the physical GigE interfaces cabled to your LAN.

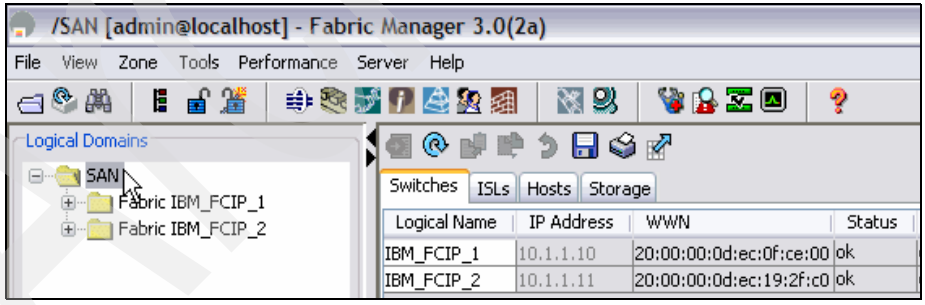


Figure 12-10 San selection

a. Select the FCIP wizard, as shown in Figure 12-11.

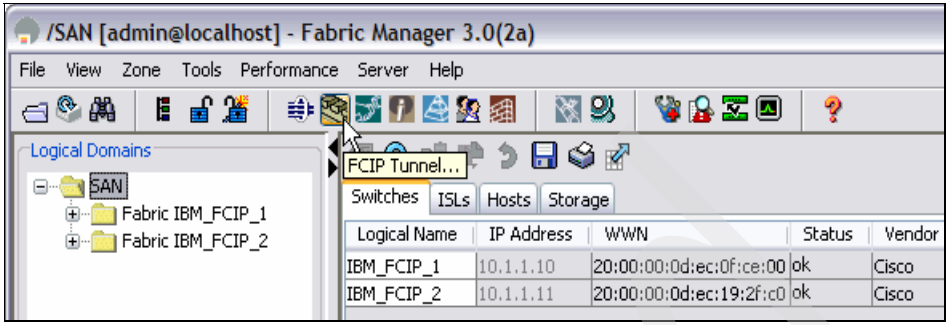


Figure 12-11 FM FCIP wizard

b. In the first panel of the FCIP tunnel wizard (Figure 12-12), we select the switches (or you can enter the IP addresses) between which to configure an FCIP link. Click **Next** after making your choice.

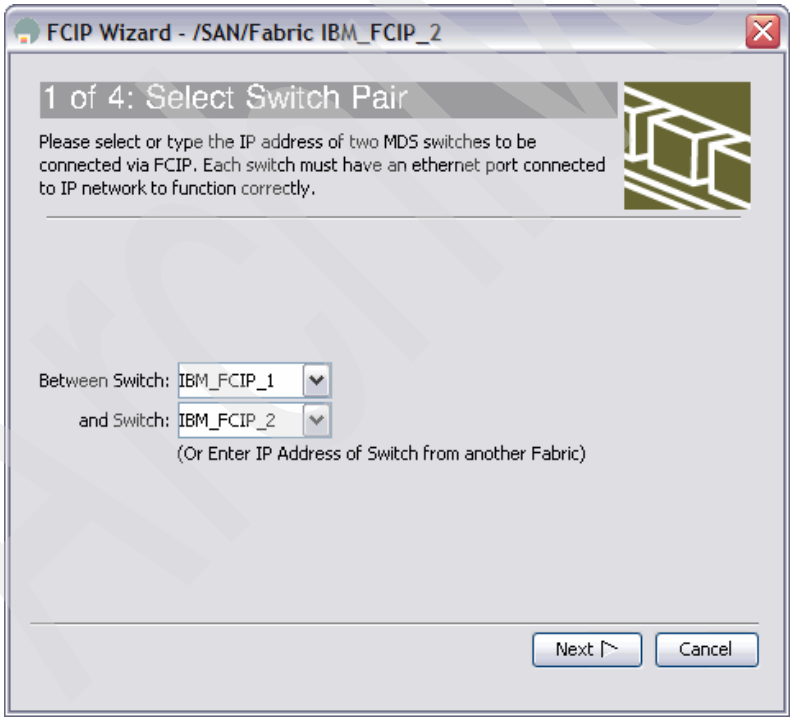


Figure 12-12 FM FCIP Wizard panel 1 of 4

c. In the next panel, shown in Figure 12-14 on page 345, we configure which Gigabit Ethernet interfaces will make up the link. Highlight the correct interface from each switch and click **Next**. In this example, we connect interface gigE1/1 on IBM_FCIP_2 to the interface gigE1/1 on IBM_FCIP_1.

Note: The interfaces will show failed, as shown in Figure 12-13 on page 344, if you do not have the GigE interfaces connected to your LAN. Cable your physical ports to your LAN, and restart the wizard to correct this.

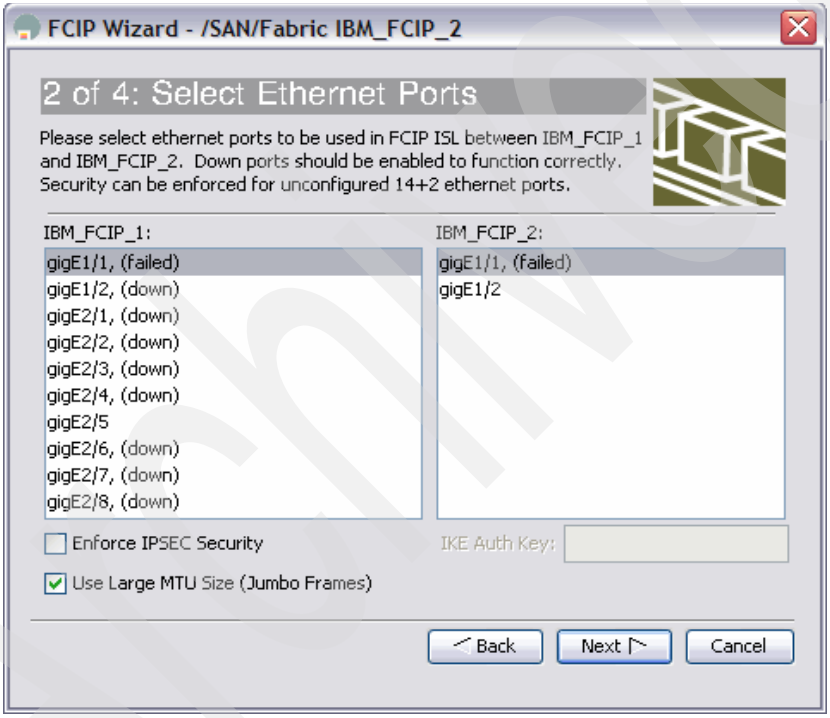


Figure 12-13 GigE ports showing failed status

d. Select the ports to connect your FCIP Tunnel, as shown in Figure 12-14, and click **Next**.

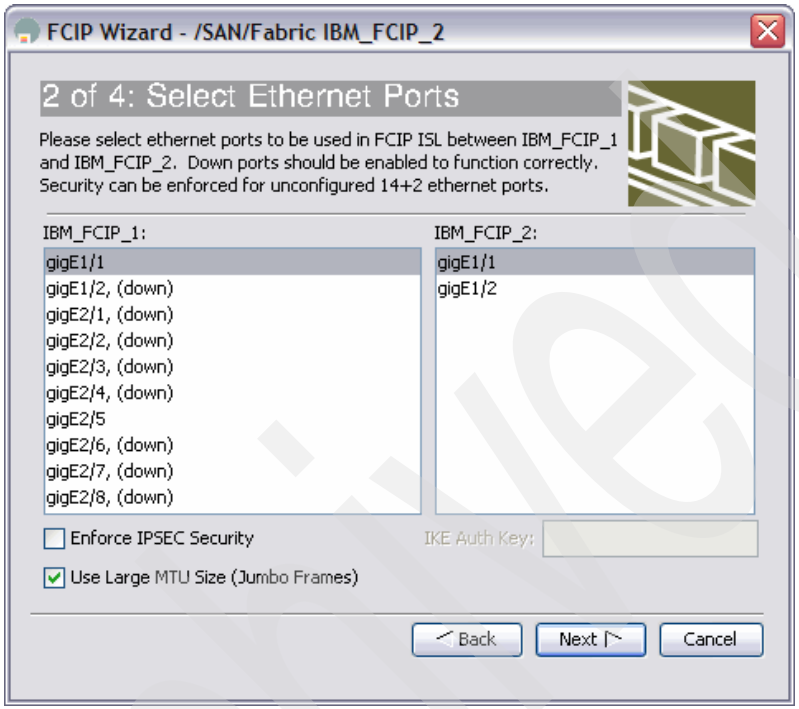


Figure 12-14 FM FCIP Wizard panel 2 of 4

e. In the next panel (Figure 12-15 on page 346), we configure the FCIP tunnel properties. This is where we perform most of the tuning changes. We discuss this in greater detail in 12.5.1, “Advanced FCIP profile configuration” on page 362, 12.5.2, “Advanced FCIP interface configuration” on page 364, 12.5.3, “Configuring FCIP write acceleration” on page 366, and 12.5.4, “Enabling FCIP compression” on page 367.

For our example, we use the default settings.

Important: Although we take the default settings, do not leave the min, max, and RTT at their defaults. You must configure real values, and ensure that the min is greater than 1/20 of the max.

In 12.5.6, “Calculating round-trip time (RTT)” on page 370, we document a tip for providing a proper estimate of the RTT value.

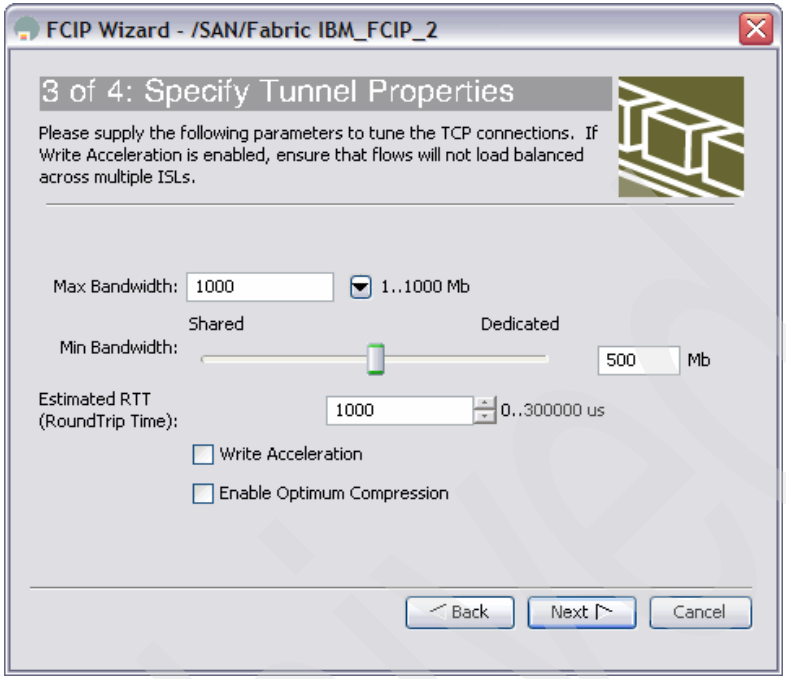


Figure 12-15 FM FCIP Wizard panel 3 of 4

f. In the next panel (Figure 12-16 on page 347), we code the IP addresses of the GigE interfaces. We use a single subnet, but this is not a requirement. Addresses from different subnets can be used, combined with static routes defined on each switch indicating the correct path. We also select the default VSAN for this interface and configure the trunking behavior. In this case, we select **auto** for the trunk behavior. Table 12-4 explains trunk behavior based on individual switch trunk configurations.

Table 12-4 Trunk behavior

Trunk config switch 1	Trunk config switch 2	Trunk behavior
On	On	On
Off	On	Off
Off	Off	Off
Auto	Auto	Off
Auto	On	On
Auto	Off	Off

When we create the PortChannel (discussed later in this chapter), it will be configured to **trunk**. After clicking **Finish** (Figure 12-16), the first FCIP link is created.

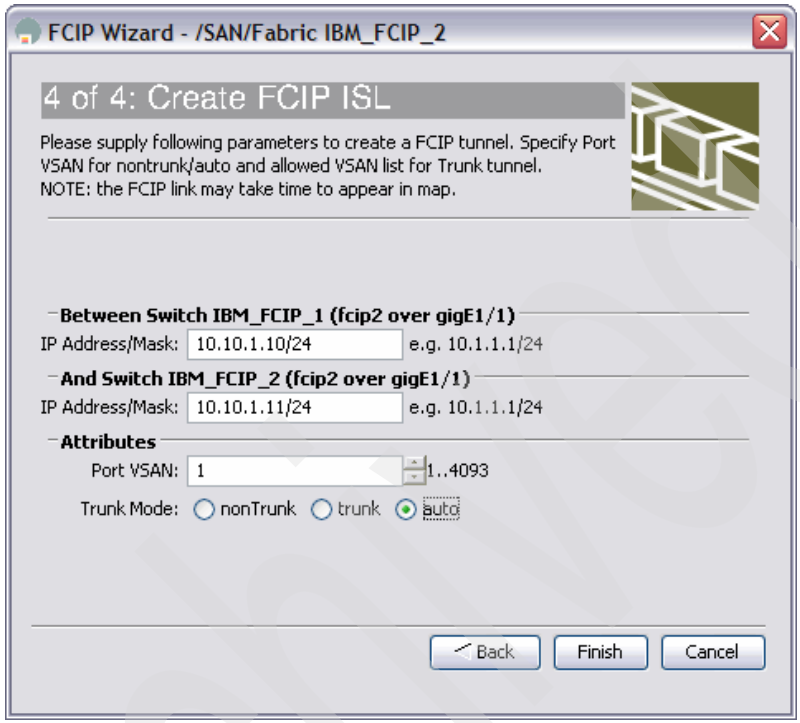


Figure 12-16 FM FCIP Wizard panel 4 of 4

5. We again select the FCIP tunnel wizard, shown in Figure 12-11 on page 343, and are ready to select which interface will make up the second FCIP link, as shown Figure 12-17. The second link will use theGigE2/5 port on IBM_FCIP_1 and the GigE1/2 on the other MDS 9216i.

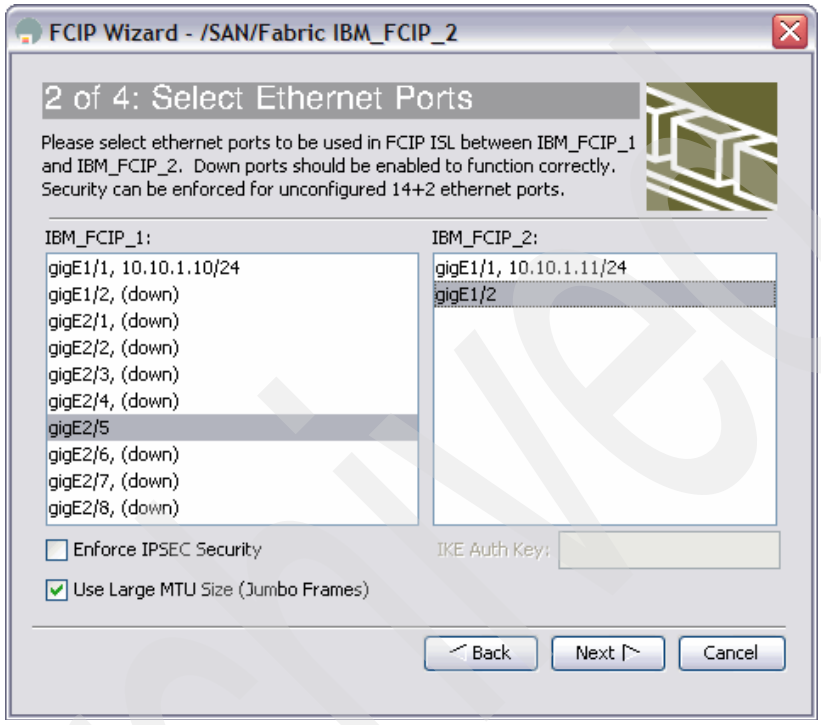


Figure 12-17 FCIP port selection

a. Configure the tunnel properties for this second link (Figure 12-18).

> **Tip:** It is best to consult the individual OEM vendor documentation when determining if proprietary data replication applications support write acceleration or IP compression.

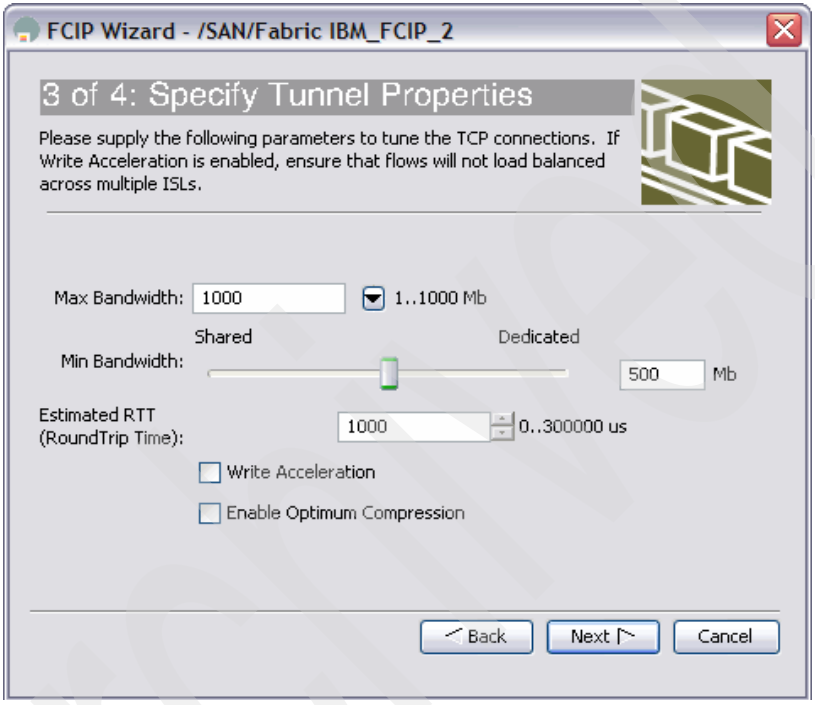


Figure 12-18 FM FCIP Wizard

b. We configure the IP addresses and trunk behavior for this second link and click **Finish** (Figure 12-19).

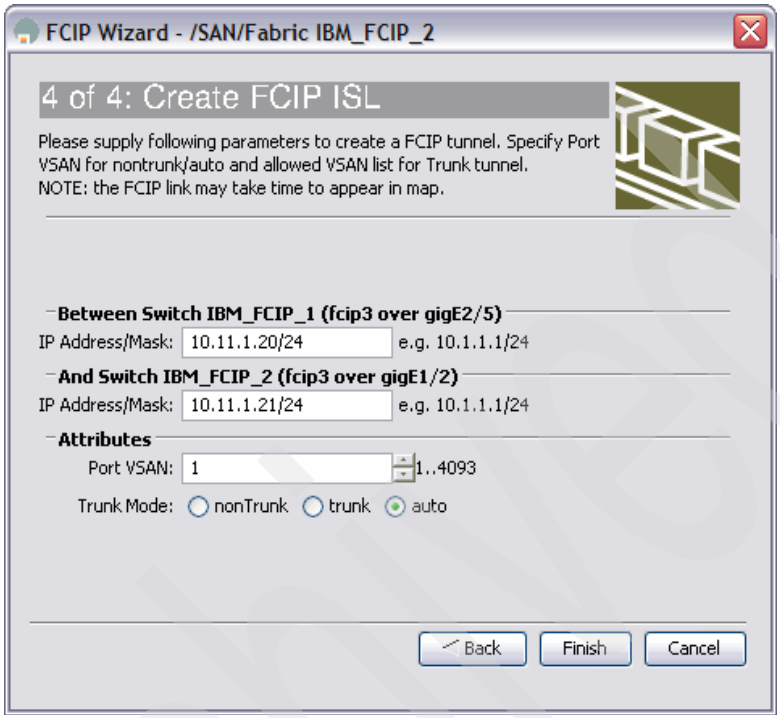


Figure 12-19 FM FCIP Wizard

In Figure 12-20, we can see the two individual FCIP links in the FM map. We place the cursor over one of the links, and the message window indicates this is fcip2 (highlighted). Notice that the FCIP links show up as dotted lines to distinguish them from FC ISLs.

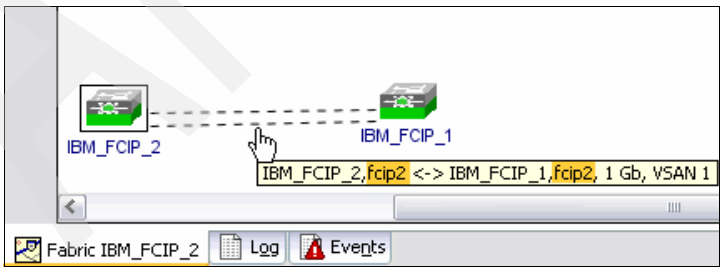


Figure 12-20 FM map display of FCIP interface fcip2

In Figure 12-21, we place the cursor over the other FCIP link and the message window indicates that it is fcip3.

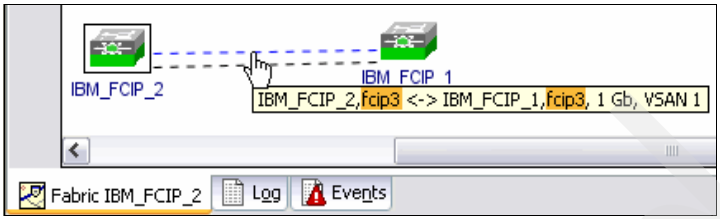


Figure 12-21 FM map display of FCIP interface fcip3

Note: If your window displays a thicker and darker line for the link, as shown in Figure 12-22, click the icon to Expand/Collapse the links. You will then see the links represented by two individual lines, as shown in Figure 12-21.

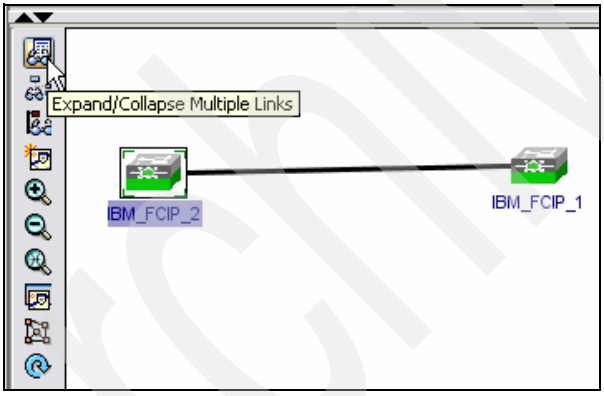


Figure 12-22 Multiple Links view

6. The next step is to bundle the two FCIP links into a PortChannel for high availability and bandwidth reasons using the PortChannel wizard, shown in Figure 12-23.

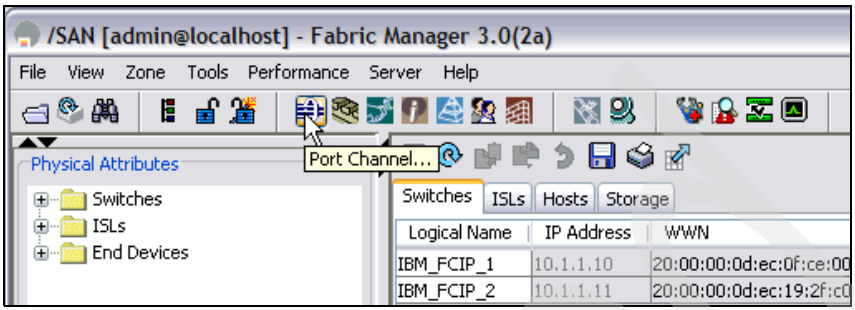


Figure 12-23 FM Port Channel wizard

a. In the first panel of the wizard (Figure 12-24), we identify the switch pair that will be linked by the PortChannel. In our case, **Create New** is automatically selected (because there no existing links to edit).

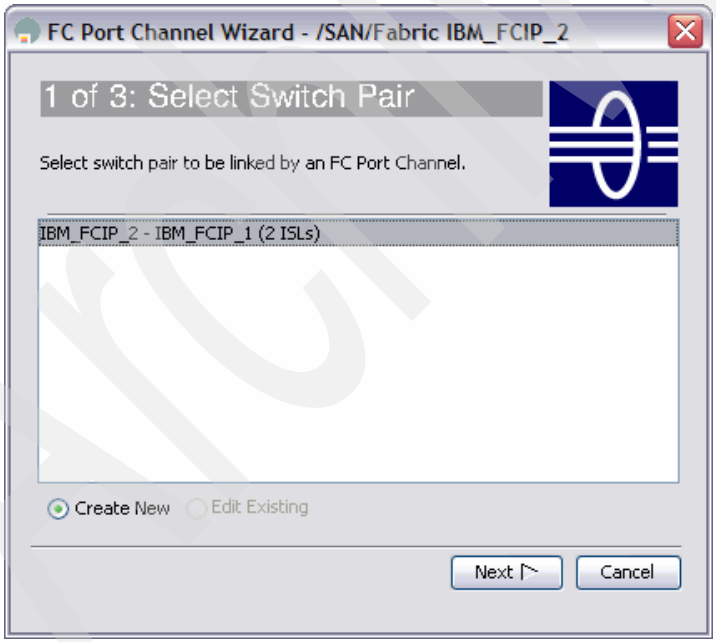


Figure 12-24 FM Port Channel Wizard panel 1 of 3

b. Highlight which links will make up the PortChannel, and using the arrow keys, move them from the Available to the Selected column (Figure 12-25). Click **Next** to proceed.

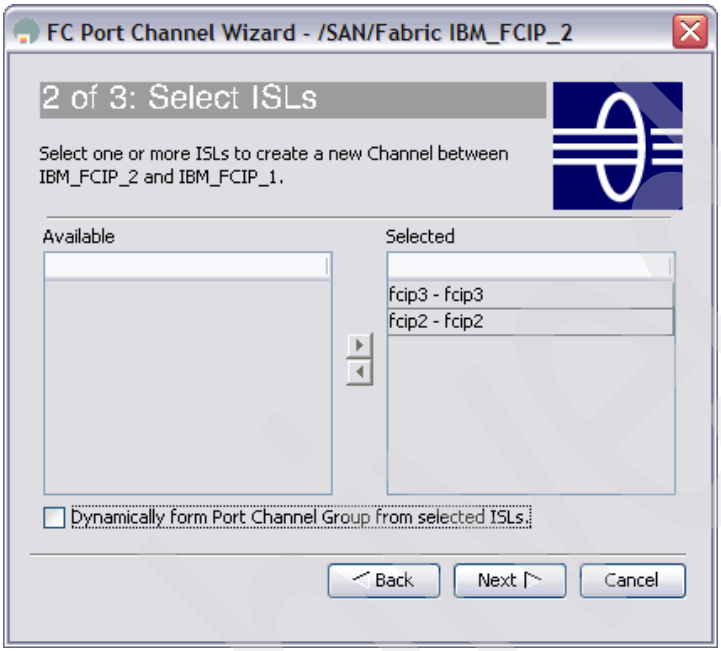


Figure 12-25 FM Port Channel Wizard panel 2 of 3

> **Tip:** Optionally, select the **Dynamically form Port Channel Group from selected ISLs** check box if you want to dynamically create the PortChannel and make the ISL properties identical for the Admin, Trunk, Speed, and VSAN attributes.

c. In the last panel of the Port Channel Wizard, shown in Figure 12-26 on page 354, we configure the Channel Id, Description, and the Trunk Mode behavior of the PortChannel. We then click **Finish** to create this PortChannel.

   This panel contains the following options:

   - VSAN List: This lists the VSANs that the PortChannel will allow to traverse the link. We allowed *all* of them to use our link.

- Trunk Mode: You can enable trunking on the links in the PortChannel. Select trunking if your link is between TE_Ports. Select nontrunking if your link is between E_Ports (for example, if your link is between an MDS switch and another vendor's switch). Select auto if you are not sure.
- Force Admin, Trunk, Speed, and VSAN attributes to be Identical: This option ensures that the same parameter settings are used in all physical ports in the channel. If these settings are not identical, the ports cannot become part of the PortChannel.
- Speed: The port speed values are auto, 1Gb, 2Gb, 4Gb, and autoMax2G.

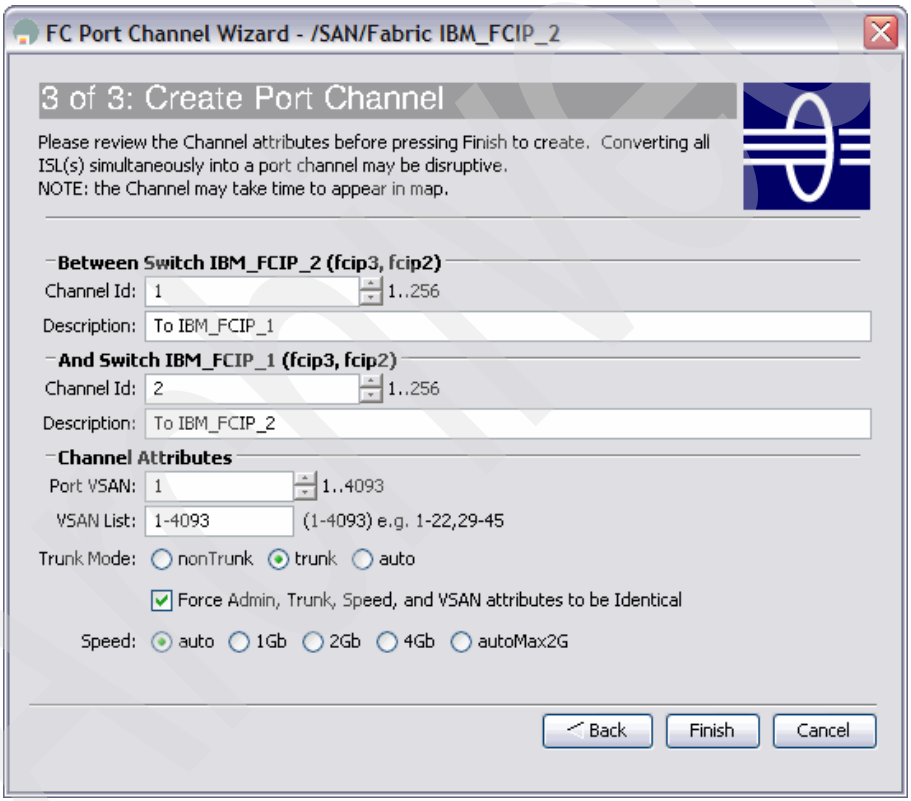


Figure 12-26 FM PortChannel wizard panel 3 of 3

Looking at Figure 12-27, we see that there are now three dotted lines for the FCIP ISL, and the message window indicates it consists of both fcip2 and fcip3. We also see it is a TE_Port (trunking), consisting of channel 1 and channel 2.

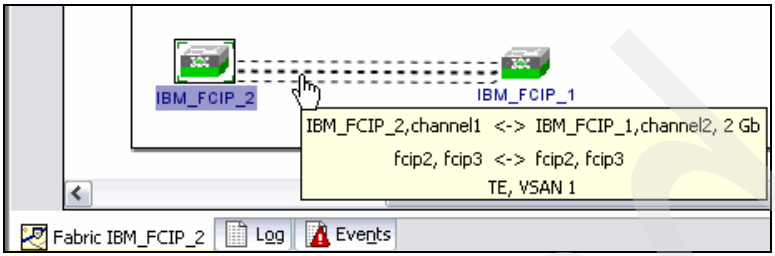


Figure 12-27 FM map display of PortChannel 2

In our example, we have allowed *all* VSANs to traverse this PortChannel. In a production environment, ensure that only the necessary VSANs are configured for its use. The reason for this is the limited bandwidth associated with the WAN connections.

To limit the VSAN access to this PortChannel, in the Physical Attributes pane in FM, open the **Switches** → **Interfaces** folder and select **FC Logical**, as shown in Figure 12-28.

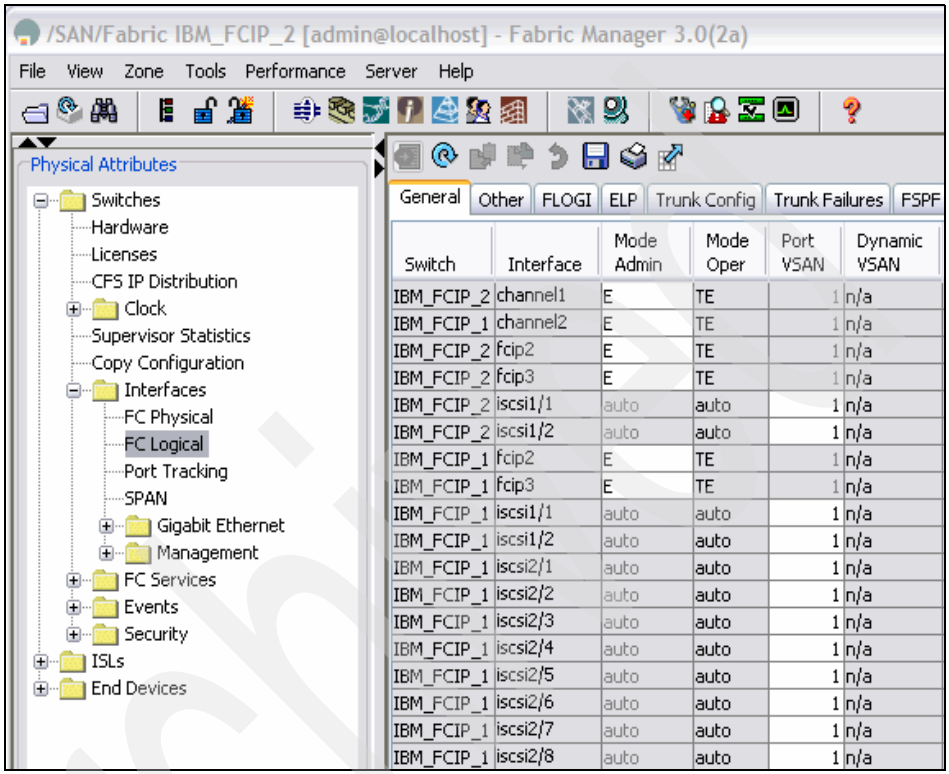


Figure 12-28 Logical Interfaces

Select the **Trunk Config** tab, and edit the Allowed VSANs column for the PortChannels you are working with, as shown in Figure 12-29. Ensure that you configure both channels with the same values. Remember to apply the changes.

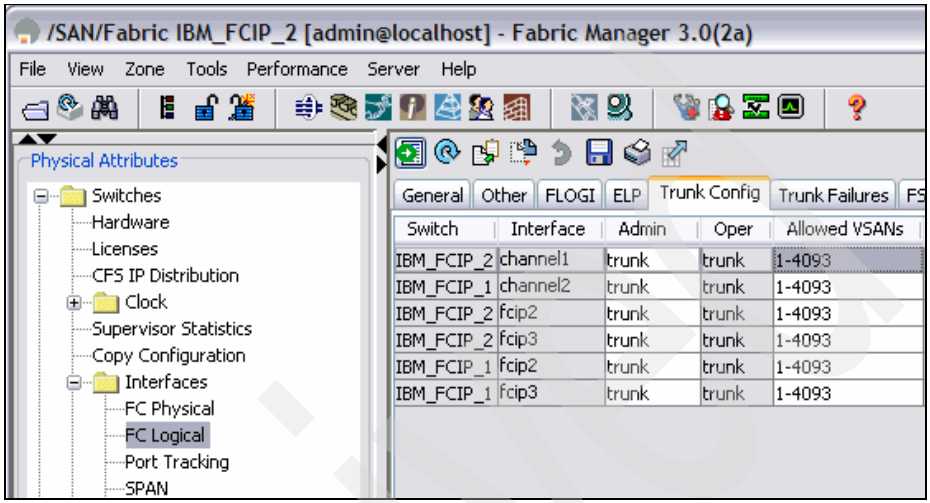


Figure 12-29 Allowed VSAN trunk configuration

12.4 Verification

This section decibels how to verify that the FCIP-bundled PortChannel is operating normally.

To verify that the FCIP-bundled PortChannel is operating normally, perform the following steps:

1. The first step is to use the FM map. Place the cursor over the dotted line representing the FCIP PortChannel, and verify that it consists of the desired FCIP links and that the correct VSANs are trunking, as shown in Figure 12-30 on page 358. We are currently trunking the default VSAN 1.

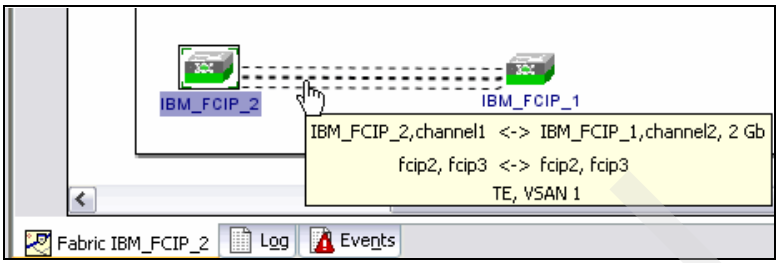


Figure 12-30 FM map display of FCIP PortChannel 2

2. To view the next set of FCIP tunnel displays from FM (Figure 12-7 on page 341), in the Physical Attributes pane, open the **ISLs** folder and select **FCIP**.

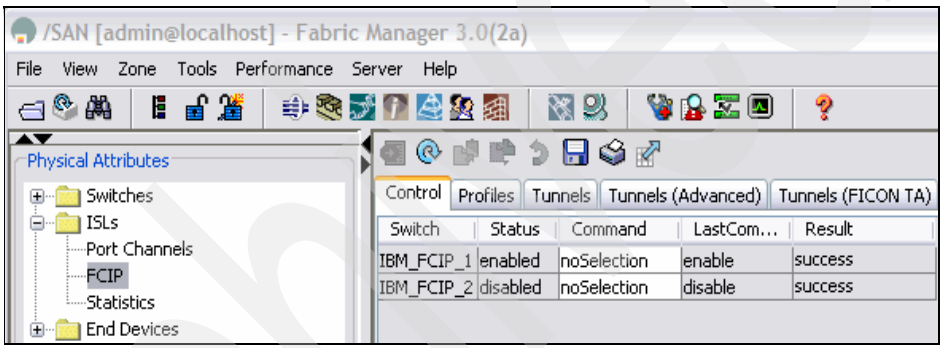
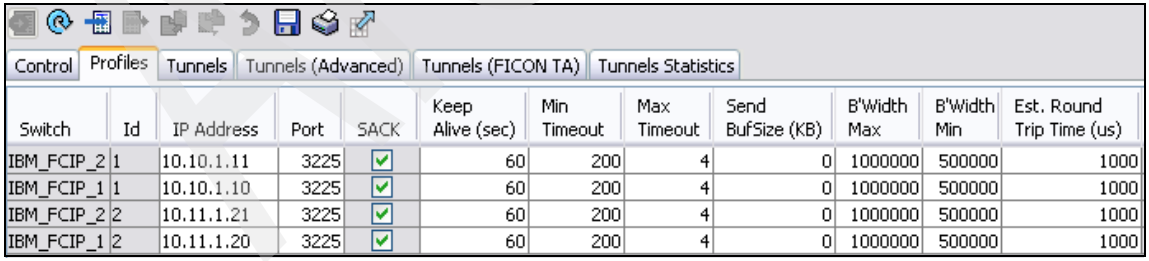


Figure 12-31 FCIP tunnel

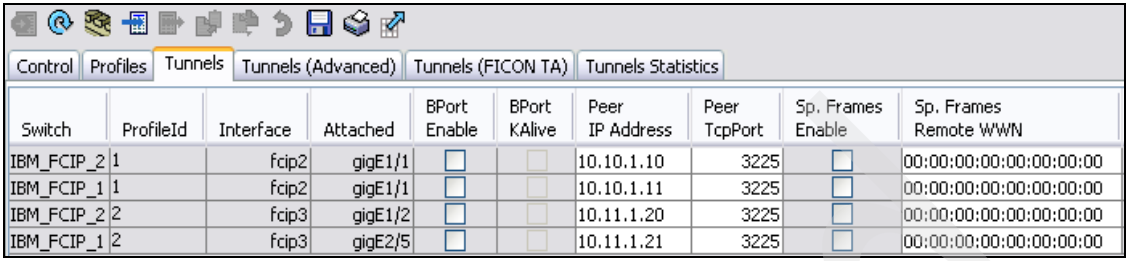
3. First, we look at the Profiles tab, as shown in Figure 12-32. We should see four sets of entries: the two FCIP links for *each* switch. We can verify the configuration and modify it as needed.



Switch	Id	IP Address	Port	SACK	Keep Alive (sec)	Min Timeout	Max Timeout	Send BufSize (KB)	B'Width Max	B'Width Min	Est. Round Trip Time (us)
IBM_FCIP_2	1	10.10.1.11	3225	☑	60	200	4	0	1000000	500000	1000
IBM_FCIP_1	1	10.10.1.10	3225	☑	60	200	4	0	1000000	500000	1000
IBM_FCIP_2	2	10.11.1.21	3225	☑	60	200	4	0	1000000	500000	1000
IBM_FCIP_1	2	10.11.1.20	3225	☑	60	200	4	0	1000000	500000	1000

Figure 12-32 FM FCIP Profiles display and configuration panel

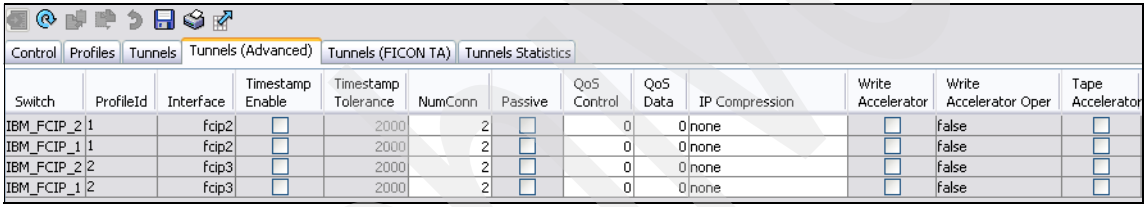
4. Use the Tunnels tab (Figure 12-33) to view and modify the FCIP interface configuration.



Switch	ProfileId	Interface	Attached	BPort Enable	BPort KAlive	Peer IP Address	Peer TcpPort	Sp. Frames Enable	Sp. Frames Remote WWN
IBM_FCIP_2	1	fcip2	gigE1/1	☐		10.10.1.10	3225	☐	00:00:00:00:00:00:00:00
IBM_FCIP_1	1	fcip2	gigE1/1	☐		10.10.1.11	3225	☐	00:00:00:00:00:00:00:00
IBM_FCIP_2	2	fcip3	gigE1/2	☐		10.11.1.20	3225	☐	00:00:00:00:00:00:00:00
IBM_FCIP_1	2	fcip3	gigE2/5	☐		10.11.1.21	3225	☐	00:00:00:00:00:00:00:00

Figure 12-33 FM FCIP Tunnels display and configuration panel

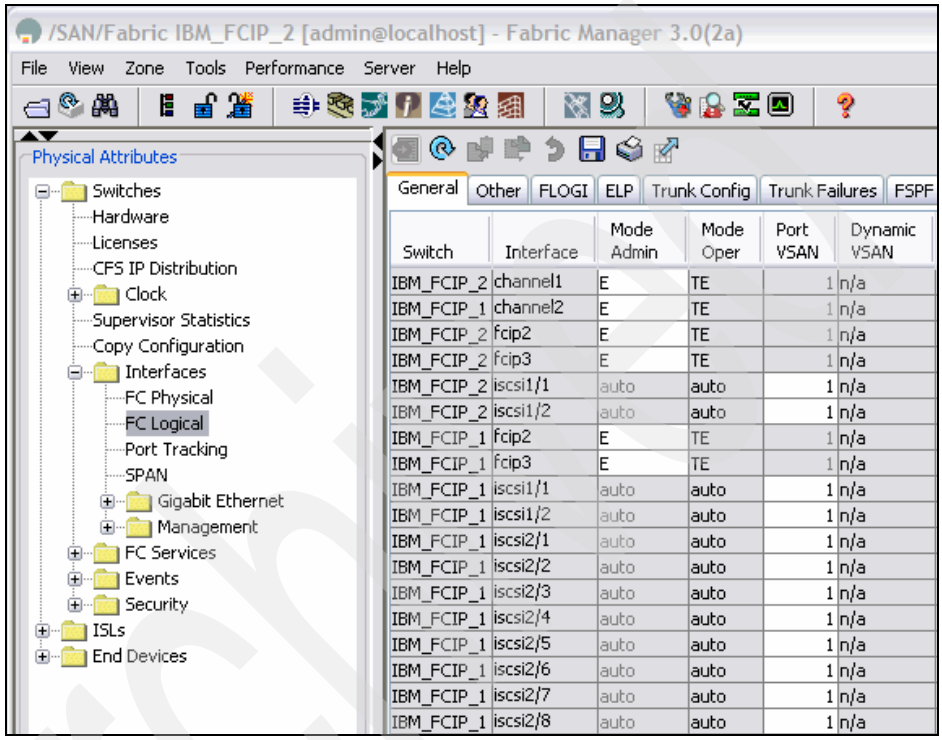
5. View and modify the advanced Tunnel options such as Timestamp, QoS, IP Compression, and Write Accelerator in the Tunnels (Advanced) tab (Figure 12-34). We discuss these advanced features in 12.5, “Advanced configuration concepts” on page 362.



Switch	ProfileId	Interface	Timestamp Enable	Timestamp Tolerance	NumConn	Passive	QoS Control	QoS Data	IP Compression	Write Accelerator	Write Accelerator Oper	Tape Accelerator
IBM_FCIP_2	1	fcip2	☐	2000	2	☐	0	0	none	☐	false	☐
IBM_FCIP_1	1	fcip2	☐	2000	2	☐	0	0	none	☐	false	☐
IBM_FCIP_2	2	fcip3	☐	2000	2	☐	0	0	none	☐	false	☐
IBM_FCIP_1	2	fcip3	☐	2000	2	☐	0	0	none	☐	false	☐

Figure 12-34 FM FCIP Tunnels (Advanced) panel

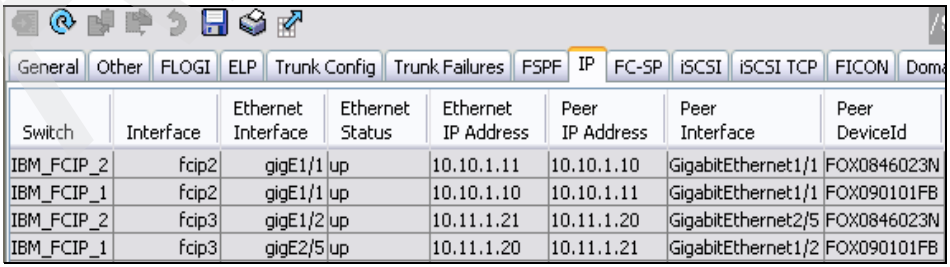
6. Next, we look at the FC logical information in FM to verify the PortChannel operation. In FM, from the Physical Attributes panel, open the **Switches** → **Interfaces** folder, and select **FC Logical**, as shown in Figure 12-35. The pane above the FM map changes, and the General tab shows us our channels. In our scenario, the two channels, one and two, make up our ISL.



Switch	Interface	Mode Admin	Mode Oper	Port VSAN	Dynamic VSAN
IBM_FCIP_2	channel1	E	TE	1	n/a
IBM_FCIP_1	channel2	E	TE	1	n/a
IBM_FCIP_2	fcip2	E	TE	1	n/a
IBM_FCIP_2	fcip3	E	TE	1	n/a
IBM_FCIP_2	iscsi1/1	auto	auto	1	n/a
IBM_FCIP_2	iscsi1/2	auto	auto	1	n/a
IBM_FCIP_1	fcip2	E	TE	1	n/a
IBM_FCIP_1	fcip3	E	TE	1	n/a
IBM_FCIP_1	iscsi1/1	auto	auto	1	n/a
IBM_FCIP_1	iscsi1/2	auto	auto	1	n/a
IBM_FCIP_1	iscsi2/1	auto	auto	1	n/a
IBM_FCIP_1	iscsi2/2	auto	auto	1	n/a
IBM_FCIP_1	iscsi2/3	auto	auto	1	n/a
IBM_FCIP_1	iscsi2/4	auto	auto	1	n/a
IBM_FCIP_1	iscsi2/5	auto	auto	1	n/a
IBM_FCIP_1	iscsi2/6	auto	auto	1	n/a
IBM_FCIP_1	iscsi2/7	auto	auto	1	n/a
IBM_FCIP_1	iscsi2/8	auto	auto	1	n/a

Figure 12-35 FM path for displaying FC Logical information

7. Select the **IP** tab (Figure 12-36). We see both of our FCIP links and verify that they are using the GigE interfaces that we configured. For each link on both switches, we see a Neighbor Device indicating that the links are up.



Switch	Interface	Ethernet Interface	Ethernet Status	Ethernet IP Address	Peer IP Address	Peer Interface	Peer DeviceId
IBM_FCIP_2	fcip2	gigE1/1	up	10.10.1.11	10.10.1.10	GigabitEthernet1/1	FOX0846023N
IBM_FCIP_1	fcip2	gigE1/1	up	10.10.1.10	10.10.1.11	GigabitEthernet1/1	FOX090101FB
IBM_FCIP_2	fcip3	gigE1/2	up	10.11.1.21	10.11.1.20	GigabitEthernet2/5	FOX0846023N
IBM_FCIP_1	fcip3	gigE2/5	up	10.11.1.20	10.11.1.21	GigabitEthernet1/2	FOX090101FB

Figure 12-36 FM IP panel for FC Logical interfaces

8. To view the operational status for our configured ISLs (Figure 12-37), select **ISLs** from the Physical Attributes pane.

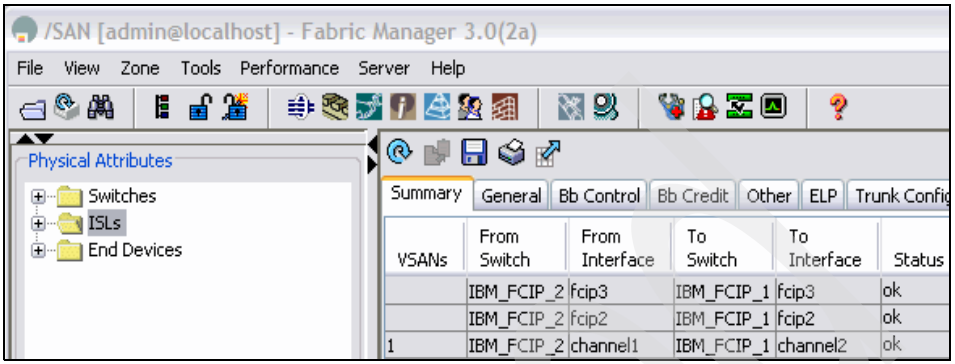


Figure 12-37 FM path to ISL status display

9. Select the **General** tab, as shown in Figure 12-38.

> **Note:** This screen capture shows just a sampling of the information available here.

Note that you can access many of the previously referenced tabs through one of the tabs located across the top of the pane.

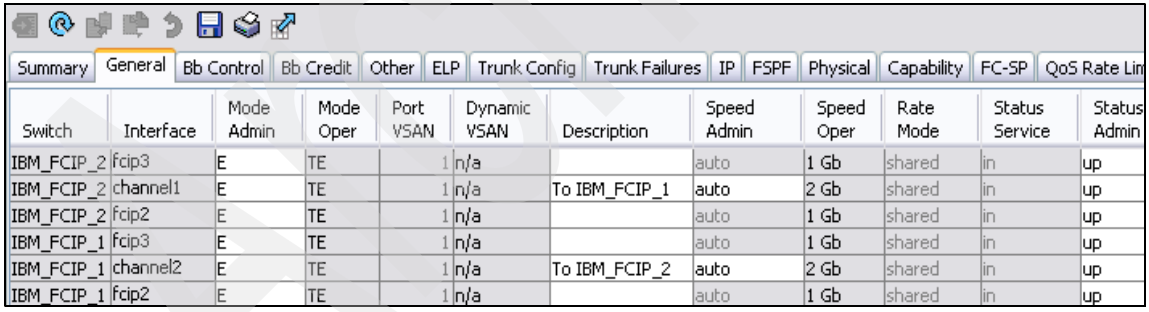


Figure 12-38 FM PortChannel summary

12.5 Advanced configuration concepts

The following advanced concepts were was taken in their entirety from the “Configuring IP Storage” chapter of the *Cisco MDS 9000 Family Fabric Manager Configuration Guide*, 78-16493-01, at the following URL:

http://www.cisco.com/en/US/products/ps5989/products_configuration_guide_book09186a0080667aa0.html

12.5.1 Advanced FCIP profile configuration

A basic FCIP configuration uses the local IP address to configure the FCIP profile. In addition to the local IP address and the local port, you can specify other TCP parameters as part of the FCIP profile configuration:

- Configuring TCP Listener Ports: The default TCP port for FCIP is 3225. Change this port using the **port** command.
- Minimum Retransmit Timeout: The TCP minimum-retransmit-time option controls the minimum amount of time TCP waits before retransmitting. By default, this value is 200 milliseconds.
- Keepalive Timeout: The TCP keepalive timeout option enables you to configure the interval between which the TCP connection verifies if the FCIP link is functioning. This ensures that a FCIP link failure is detected quickly even when there is no traffic. If the TCP connection is idle for more than the specified time, keepalive timeout packets are sent to ensure that the connection is active. This command can be used to tune the time taken to detect FCIP link failures. The first interval during which the connection is idle is 60 seconds (default). When the connection is idle for 60 seconds, eight keepalive probes are sent at 1-second intervals. If no response is received for these eight probes and the connection remains idle throughout, that FCIP link is automatically closed. Only the first interval (during which the connection is idle) can be changed from the default of 60 seconds. This interval is identified using the keepalive timeout option. The valid range is from 1 to 7200 seconds.
- Maximum Retransmissions: The Tcp max-retransmissions option specifies the maximum number of times a packet is retransmitted before TCP decides to close the connection.

- Path MTU: Path MTU (PMTU) is the minimum MTU on the IP network between the two end points of the FCIP link. PMTU discovery is a mechanism by which TCP learns of the PMTU dynamically and adjusts the maximum TCP segment accordingly (RFC 1191). By default, PMTU discovery is enabled on all switches with a default timeout of 3600 seconds. If TCP reduces the size of the max segment because of PMTU change, the reset timeout specifies the time after which TCP tries the original MTU.
- SACK: TCP might experience poor performance when multiple packets are lost within one window. With the limited information available from cumulative acknowledgments, a TCP sender can only learn about a single lost packet per round trip time. A selective acknowledgment (SACK) mechanism helps overcome the limitations of multiple lost packets during a TCP transmission. The receiving TCP sends back SACK advertisements to the sender. The sender can then retransmit only the missing data segments. By default, SACK is enabled on Cisco MDS 9000 Family switches.
- Window Management: The optimal TCP window size is computed using the max-bandwidth option, the min-available-bandwidth option, and the dynamically-measured round-trip-time (RTT). The interaction and the resulting TCP behavior is outlined here:
  - The configured round-trip-time option determines the window scaling factor of the TCP connection. This option is only an approximation. The measured RTT value overrides the round-trip-time option for window management. If the configured round-trip-time is too small compared to the measured RTT, the link might not be fully utilized due to the window scaling factor being too small.
  - If the average rate of the FC traffic over the preceding RTT is less than the min-available-bandwidth * RTT, every FC burst is sent immediately at the min-available-bandwidth rate, provided no TCP drops occur.
  - If the average rate of the FC traffic is greater than min-available-bandwidth * RTT, but less than max-bandwidth * RTT, and then if the FC traffic is transmitted in burst sizes smaller than the configured CWM value, all the bursts are sent immediately by FCIP at the max-bandwidth rate.
  - If the average rate of the FC traffic is larger than the min-available-bandwidth * RTT and the burst size is greater than the CWM value, some traffic will not be sent immediately.
  - The maximum-bandwidth option and the measured round-trip-time together determine the maximum window size.

  - The min-available-bandwidth option and the measured round-trip-time together determine the threshold below which TCP aggressively maintains a window size sufficient to transmit at min-available-bandwidth. The software uses standard TCP rules to increase the window beyond the one required to maintain the min-available-bandwidth in order to reach the max-bandwidth. The defaults are max-bandwidth = 1 G, min-available-bandwidth = 15 Mbps, and round-trip-time = 1 ms.

- Buffer Size: The send-buffer-size option defines the required additional buffering, beyond the normal send window size, that TCP allows before flow controlling the switches egress path for the FCIP interface. The default buffer size is 0 KB.
- Quality of Service: The quality of service (QoS) feature specifies the differentiated services code point (DSCP) value to mark all IP packets (type of service, TOS, field in the IP header).
  - The control DSCP value applies to all FCIP frames in the control TCP connection.
  - The data DSCP value applies to all FCIP frames in the data connection.
  - If the FCIP link has only one TCP connection, that data DSCP value is applied to all packets in that connection.
- Monitoring Window Congestion: The congestion window monitoring (CWM) option determines the maximum burst size allowed after an idle period. If the FC traffic burst is smaller than the configured CWM value, every packet is sent immediately, provided that no TCP drops were detected in the previous RTT. If the FC traffic burst is larger than the configured CWM value, the excess packets will be sent during succeeding RTTs. By default, the tcpcwm option is enabled and the default burst size is 10 KB.

Tip: We recommend that this feature remain enabled to realize optimal performance. Increasing the CWM burst size can result in more packet drops in the IP network, impacting TCP performance. Only if the IP network has sufficient buffering, try increasing the CWM burst size beyond the default to achieve lower transmit latency.

12.5.2 Advanced FCIP interface configuration

You can establish a connection to a peer by configuring one or more of the following options for the FCIP interface:

- Peer IP Address: The basic FCIP configuration uses the peer's IP address to configure the peer information. You can also specify the peer's port number to configure the peer information. If you do not specify a port, the default 3225 port number is used to establish connection.

- Special Frames: You can alternatively establish an FCIP link with a peer using an optional protocol called special frames. You can enable or disable the special-frame option. On the peer side, the special-frame option must be enabled in order to establish the FCIP link. When the special-frame option is enabled, the peer IP address (and optionally the port or the profile ID) only needs to be configured on one end of the link. After the connection is established, a special frame is exchanged to discover and authenticate the link. By default, the special frame feature is disabled.

  Tip: Special frame negotiation provides an additional authentication security mechanism because the link validates the WWN of the peer switch.

- Configuring Active Connection: Use the passive-mode option to configure the required mode for initiating an IP connection. By default, active mode is enabled to actively attempt an IP connection. If you enable the passive mode, the switch does not initiate a TCP connection and merely waits for the peer to connect to it. Ensure that both ends of the FCIP link are not configured as passive mode. If both ends are configured as passive, the connection will not be initiated.
- Configuring the Number of TCP Connections: Use the tcp-connection option to specify the number of TCP connections from a FCIP link. By default, the switch tries two (2) TCP connections for each FCIP link. You can configure 1 or 2 TCP connections. For example, the Cisco PA-FC-1G Fibre Channel port adapter which has only 1 (one) TCP connection interoperates with any switch in the Cisco MDS 9000 Family. One TCP connection is within the specified limit and you can change the configuration on the switch using the **`tcp-connection 1`** command. If the peer initiates one TCP connection, and your MDS switch is configured for two TCP connections, the software handles it gracefully and moves on with just one connection.
- Enabling Time Stamps: Use the time-stamp option to enable or disable FCIP time stamps on a packet. The time-stamp option instructs the switch to discard packets that are outside the specified time. By default, the time-stamp option is disabled. The acceptable-diff option specifies the time range within which packets can be accepted. If the packet arrived within the range specified by this option, the packet is accepted. Otherwise, it is dropped. By default, if a packet arrives within a 1000 millisecond interval (+ or - 1000 milliseconds), that packet is accepted. If the time-stamp option is enabled, be sure to configure NTP on both switches.

- B Port Interoperability Mode: Although E_Ports typically interconnect Fibre Channel switches, some SAN extender devices, such as Cisco's PA-FC-1G Fibre Channel port adapter and the SN 5428-2 storage router, implement a bridge port model to connect geographically dispersed fabrics. This model uses B port as described in the T11 Standard FC-BB-2.
  - B ports bridge Fibre Channel traffic from one E_Port to a remote E_Port without participating in fabric-related activities such as principal switch election, domain ID assignment, and Fibre Channel routing (FSPF). For example, Class F traffic entering a SAN extender does not interact with the B port. The traffic is transparently propagated (bridged) over a WAN interface before exiting the remote B port. This bridge results in both E_Ports exchanging Class F information, which ultimately leads to normal ISL behavior such as fabric merging and routing.
  - FCIP links between B port SAN extenders do not exchange the same information as FCIP links between E_Ports and are, therefore, incompatible. This is reflected by the terminology used in FC-BB-2: While VE_Ports establish a virtual ISL over a FCIP link, B ports use a B access ISL.
  - The IPS module supports FCIP links that originate from a B port SAN extender device by implementing the B access ISL protocol on a Gigabit Ethernet interface. Internally, the corresponding virtual B port connects to a virtual E_Port, which completes the end-to-end E_Port connectivity requirement.
  - The B port feature in the IPS module allows remote B port SAN extenders to communicate directly with a Cisco MDS 9000 Family switch, therefore eliminating the need for local bridge devices.
  - When an FCIP peer is a SAN extender device that only supports Fibre Channel B ports, you need to enable the B port mode for the FCIP link. When a B port is enabled, the E_Port functionality is also enabled and they coexist. If the B port is disabled, the E_Port functionality remains enabled.

12.5.3 Configuring FCIP write acceleration

The FCIP write acceleration feature in SAN-OS 1.3(3) or later enables you to significantly improve application performance when storage traffic is routed over wide area networks using FCIP.

Important: Do not enable write acceleration on FCIP links carrying FICON VSANs. At the time of writing, it is not supported with DS8000™ copy services.

When FCIP write acceleration is enabled, WAN throughput is maximized by minimizing the impact of WAN latency for the command to transfer ready acknowledgement.

The write acceleration feature is disabled by default and must be enabled on both sides of the FCIP link. If it is only enabled on one side of the FCIP tunnel, the tunnel will not initialize.

In some cases, data sent by the host is queued on the target before the target issues a Transfer Ready. This way, the actual write operation might be done in a less time than the write operation without the write acceleration feature being enabled.

Tip: FCIP write acceleration will not work if the FCIP port is part of a PortChannel or if there are multiple paths with equal weight between the initiator and the target port. Such a configuration might cause either SCSI discovery failure or broken write or read operations. With SAN-OS 2.0 and later, FCIP write acceleration will work with PortChannels (but not equal cost FSPF paths).

12.5.4 Enabling FCIP compression

The FCIP compression feature allows IP packets to be compressed on the FCIP link if the feature is enabled on that link. By default, FCIP compression is disabled. When enabled, the software defaults to using the auto mode (if a mode is not specified).

There are four compression modes that you can select:

- mode1 is a fast compression mode for high bandwidth links (> 25 Mbps).
- mode2 is a moderate compression mode for moderately low bandwidth links (between 10 and 25 Mbps).
- mode3 is a high compression mode for low bandwidth links (< 10 Mbps).
- auto (default) mode picks the appropriate compression scheme based on the bandwidth of the link (the bandwidth of the link configured in the FCIP profile's TCP parameters).

The IP compression feature behavior differs between the IPS module and the MPS-14/2 module—while mode2 and mode3 perform software compression in both modules, mode1 performs hardware-based compression in MPS-14/2 modules and software compression in IPS modules. The MDS 9216i has the same hardware components available as the MPS-14/2 module, therefore behaving the same.

Note: The compression modes in Cisco SAN-OS Release 2.0(1b) and later are incompatible with the compression modes in Cisco SAN-OS Release 1.3(1) and earlier.

While upgrading from Cisco SAN-OS Release 1.x to Cisco SAN-OS Release 2.0(1b) or later, we recommend that you disable compression before the upgrade procedure, and then enable the required mode after the upgrade procedure.

12.5.5 FCIP high availability

The following high availability solutions are available for FCIP configurations:

- Fibre Channel PortChannels

  Figure 12-39 shows an example of a PortChannel-based load balancing configuration. To perform this configuration, you need two IP addresses on each SAN island. This solution addresses link failures.

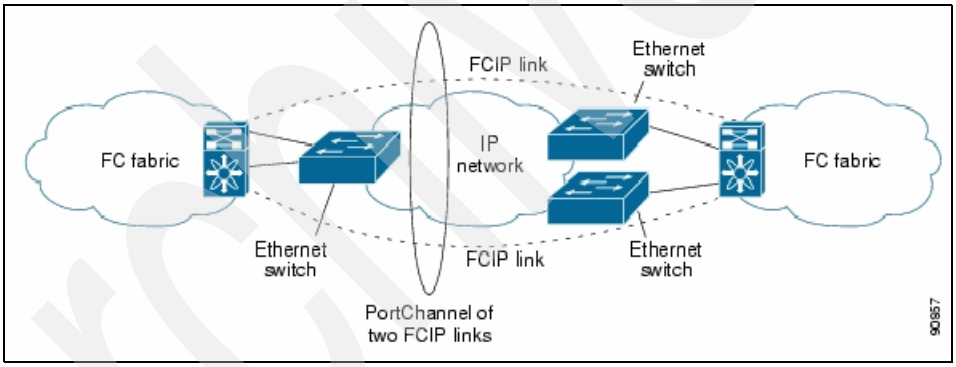


Figure 12-39 PortChannel-based load balancing

The following characteristics set Fibre Channel PortChannel solutions apart from other solutions:

- The entire bundle is one logical (E)ISL link.
- All FCIP links in the PortChannel should be across the same two switches.
- The Fibre Channel traffic is load balanced across the FCIP links in the PortChannel.

- Fabric Shortest Path First (FSPF)

  Figure 12-40 displays an FSPF-based load balancing configuration example. This configuration requires two IP addresses on each SAN island, and addresses IP and FCIP link failures.

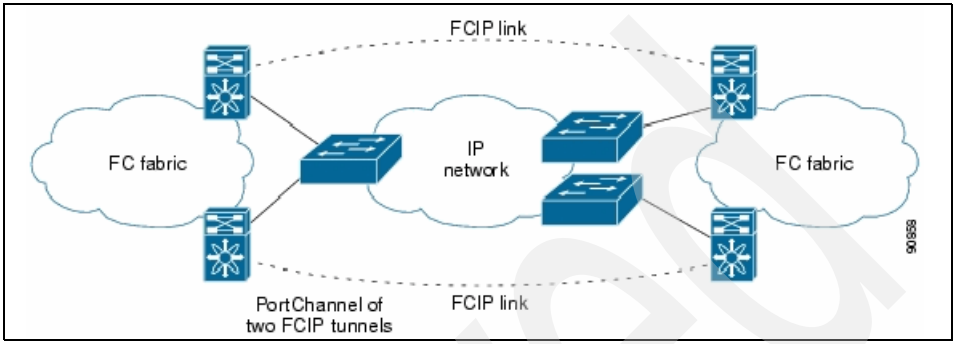


Figure 12-40 FSPF-based load balancing

  The following characteristics set FSPF solutions apart from other solutions:

  – Each FCIP link is a separate (E)ISL.
  – The FCIP links can connect to different switches across two SAN islands.
  – The Fibre Channel traffic is load balanced across the FCIP link.

- Virtual Router Redundancy Protocol (VRRP)

  Figure 12-41 displays a VRRP-based high availability FCIP configuration example. This configuration requires at least two physical Gigabit Ethernet ports connected to the Ethernet switch on the island where you need to implement high availability using VRRP.

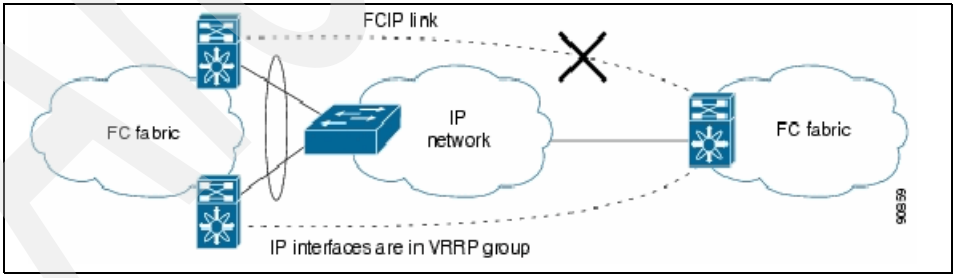


Figure 12-41 VRRP-based high availability

  The following characteristics set VRRP solutions apart from other solutions:

  – If the active VRRP port fails, the standby VRRP port takes over the VRRP IP address.

- When the VRRP switchover happens, the FCIP link automatically disconnects and reconnects.
- This configuration has only one FCIP (E)ISL link.

- Ethernet PortChannels

  Figure 12-42 displays an Ethernet PortChannel-based high availability FCIP example. This solution addresses the problem caused by individual Gigabit Ethernet link failures.

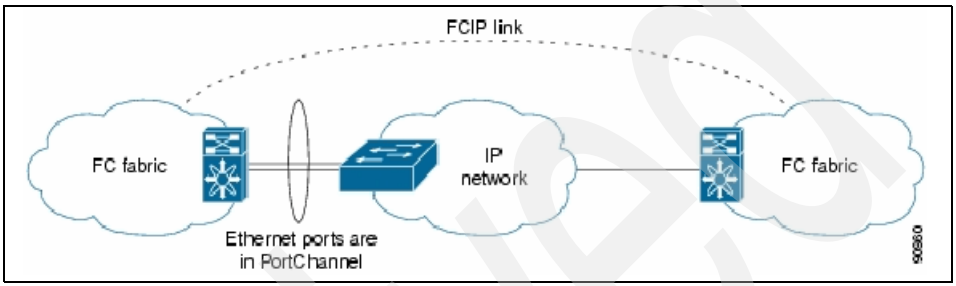


Figure 12-42 Ethernet PortChannel-based high availability

The following characteristics set Ethernet PortChannel solutions apart from other solutions:

- The Gigabit Ethernet link level redundancy ensures a transparent failover if one of the Gigabit Ethernet links fails.
- Two Gigabit Ethernet ports in one Ethernet PortChannel appear like one logical Gigabit Ethernet link.
- The FCIP link stays up during the failover.

12.5.6 Calculating round-trip time (RTT)

To determine the round-trip time (RTT) or latency, use the **ping** command in the Cisco MDS 9000 Family. Set the target **ping** IP address to the IP address of the Gigabit Ethernet port of the peer IPS Module, set the repeat count to 10, set the datagram size to 2112, and set the timeout (in seconds) to 1 s. We recommend that you use the average latency expressed in seconds for this calculation.

The following output shows a sample **ping** procedure and its results.

Example 12-1 Ping command and output

```
w16-excal-1# ping
Target IP address: 5.5.5.2
Repeat count: 10
Datagram size: 2112
```

```
Timeout in seconds: 1
Extended commands (y/n): n
PING 5.5.5.2 (5.5.5.2): 2112 data bytes
2120 bytes from 5.5.5.2: icmp_seq=0 ttl=255 time=3.5 ms
2120 bytes from 5.5.5.2: icmp_seq=1 ttl=255 time=3.4 ms
2120 bytes from 5.5.5.2: icmp_seq=2 ttl=255 time=3.4 ms
2120 bytes from 5.5.5.2: icmp_seq=3 ttl=255 time=3.4 ms
2120 bytes from 5.5.5.2: icmp_seq=4 ttl=255 time=3.4 ms
2120 bytes from 5.5.5.2: icmp_seq=5 ttl=255 time=3.4 ms
2120 bytes from 5.5.5.2: icmp_seq=6 ttl=255 time=3.4 ms
2120 bytes from 5.5.5.2: icmp_seq=7 ttl=255 time=3.4 ms
2120 bytes from 5.5.5.2: icmp_seq=8 ttl=255 time=3.4 ms
2120 bytes from 5.5.5.2: icmp_seq=9 ttl=255 time=3.4 ms
--- 5.5.5.2 ping statistics ---
10 packets transmitted, 10 packets received, 0% packet loss
round-trip min/avg/max = 3.4/3.4/3.5 ms
```

The delay value that would be used from Example 12-1 on page 370 is 3.4 ms.

12.5.7 Configuring FCIP with the CLI

In this section, we demonstrate the steps required to configure an FCIP tunnel with the CLI.

You need to use your favorite Telnet utility and connect to both switches that will be participating in the FCIP connection. We use Putty, a freeware program that you can download from the Web.

To configure FCIP using the CLI, perform the following steps:

1. Log in to the switches, and check the Gigabit Ethernet interface configuration. If it has not been configured, you will need to do this. In our case, it has been configured, as shown in Example 12-2.

> **Note:** The Ethernet interfaces do not have to be configured for the same subnet. You must configure a static route (or routes) to allow the two interfaces to communicate.

Example 12-2 Checking the Gigabit Ethernet interface configuration

```
IBM_FCIP_1# show interface gigabitethernet 1/1
GigabitEthernet1/1 is up
    Hardware is GigabitEthernet, address is 000d.bd85.60b4
    Internet address is 10.10.10.20/24
    MTU 1500  bytes
```

```
    Port mode is IPS
    Speed is 1 Gbps
    Beacon is turned off
    Auto-Negotiation is turned on
    5 minutes input rate 1424 bits/sec, 178 bytes/sec, 1 frames/sec
    5 minutes output rate 1488 bits/sec, 186 bytes/sec, 1 frames/sec
    1393 packets input, 145042 bytes
      0 multicast frames, 0 compressed
      0 input errors, 0 frame, 0 overrun 0 fifo
    1428 packets output, 159572 bytes, 0 underruns
      0 output errors, 0 collisions, 0 fifo
      0 carrier errors
```

2. We do the same for the switch to which we are connecting, which we refer to as switch 2.

Example 12-3 Checking the Gigabit Ethernet interface configuration

```
IBM_FCIP_2# show interface gigabitethernet 1/1
GigabitEthernet1/1 is up
    Hardware is GigabitEthernet, address is 000d.bd85.644a
    Internet address is 10.10.10.30/24
    MTU 1500  bytes
    Port mode is IPS
    Speed is 1 Gbps
    Beacon is turned off
    Auto-Negotiation is turned on
    5 minutes input rate 0 bits/sec, 0 bytes/sec, 0 frames/sec
    5 minutes output rate 0 bits/sec, 0 bytes/sec, 0 frames/sec
    1426 packets input, 162062 bytes
      0 multicast frames, 0 compressed
      0 input errors, 0 frame, 0 overrun 0 fifo
    1394 packets output, 143956 bytes, 0 underruns
      0 output errors, 0 collisions, 0 fifo
      0 carrier errors
```

3. Now, we enable FCIP and configure a profile back on switch 1.

Example 12-4 Enabling FCIP on switch 1

```
IBM_FCIP_1# config t
Enter configuration commands, one per line.  End with CNTL/Z.
IBM_FCIP_1(config)# fcip enable
IBM_FCIP_1(config)# fcip profile 1
IBM_FCIP_1(config-profile)# ip address 10.10.10.20 (this ip is the ip
of the interface we want to associate with our FCIP tunnel)
```

```
IBM_FCIP_1(config-profile)# exit
IBM_FCIP_1(config)# exit
IBM_FCIP_1# show fcip profile
-------------------------------------------------------------------------
ProfileId       Ipaddr                                          TcpPort
-------------------------------------------------------------------------
1               10.10.10.20                                     3225
IBM_FCIP_1# show fcip profile 1
FCIP Profile 1
    Internet Address is 10.10.10.20  (interface GigabitEthernet1/1)
    Listen Port is 3225
    TCP parameters
        SACK is enabled
        PMTU discovery is enabled, reset timeout is 3600 sec
        Keep alive is 60 sec
        Minimum retransmission timeout is 200 ms
        Maximum number of re-transmissions is 4
        Send buffer size is 0 KB
        Maximum allowed bandwidth is 1000000 kbps
        Minimum available bandwidth is 500000 kbps
        Estimated round trip time is 1000 usec
        Congestion window monitoring is enabled, burst size is 50 KB
        Auto jitter detection is enabled
```

4. We now move to switch 2.

Example 12-5 Enabling FCIP on switch 2

```
IBM_FCIP_2# config t
Enter configuration commands, one per line.  End with CNTL/Z.
IBM_FCIP_2(config)# fcip enable
IBM_FCIP_2(config)# fcip profile 2
IBM_FCIP_2(config-profile)# ip address 10.10.10.30
IBM_FCIP_2(config-profile)# exit
IBM_FCIP_2(config)# exit
IBM_FCIP_2# show fcip profile
-------------------------------------------------------------------------
ProfileId       Ipaddr                                          TcpPort
-------------------------------------------------------------------------
2               10.10.10.30                                     3225
IBM_FCIP_2# show fcip profile 2
FCIP Profile 2
    Internet Address is 10.10.10.30  (interface GigabitEthernet1/1)
    Listen Port is 3225
    TCP parameters
```

```
        SACK is enabled
        PMTU discovery is enabled, reset timeout is 3600 sec
        Keep alive is 60 sec
        Minimum retransmission timeout is 200 ms
        Maximum number of re-transmissions is 4
        Send buffer size is 0 KB
        Maximum allowed bandwidth is 1000000 kbps
        Minimum available bandwidth is 500000 kbps
        Estimated round trip time is 1000 usec
        Congestion window monitoring is enabled, burst size is 50 KB
        Auto jitter detection is enabled
```

5. Now, we configure the tunnel. We move back to switch 1.

Example 12-6 Configuring the tunnel on switch 1

```
IBM_FCIP_1# config t
Enter configuration commands, one per line.  End with CNTL/Z.
IBM_FCIP_1(config)# interface fcip 1
IBM_FCIP_1(config-if)# use-profile 1
IBM_FCIP_1(config-if)# peer-info ipaddr 10.10.10.30
IBM_FCIP_1(config-if)# no shutdown
IBM_FCIP_1(config-if)# exit
IBM_FCIP_1(config)# exit
IBM_FCIP_1# show fcip summary
--------------------------------------------------------------------------
Tun prof Eth-if peer-ip Status T W T Enc Comp Bandwidth max/min (us)
                              E A A
--------------------------------------------------------------------------
1   1    GE1/1 10.10.10.30   DOWN  N N N  N   N    1000M/500M  1000
```

Notice that the link status is down. This is expected, because we have not finished configuring the other end of the link.

6. We move onto switch 2.

Example 12-7 Configuring the tunnel on switch 2

```
IBM_FCIP_2# config t
Enter configuration commands, one per line.  End with CNTL/Z.
IBM_FCIP_2(config)# interface fcip 2
IBM_FCIP_2(config-if)# use-profile 2
IBM_FCIP_2(config-if)# peer-info ipaddr 10.10.10.20
IBM_FCIP_2(config-if)# no shutdown
IBM_FCIP_2(config-if)# exit
IBM_FCIP_2(config)# exit
IBM_FCIP_2# show fcip summary
```

```
-------------------------------------------------------------------------------
Tun prof Eth-if peer-ip Status T W T Enc Comp Bandwidth max/min (us)
                                E A A
-------------------------------------------------------------------------------
2   2    GE1/1 10.10.10.20 TRNK  Y N N  N   N    1000M/500M  1000
```

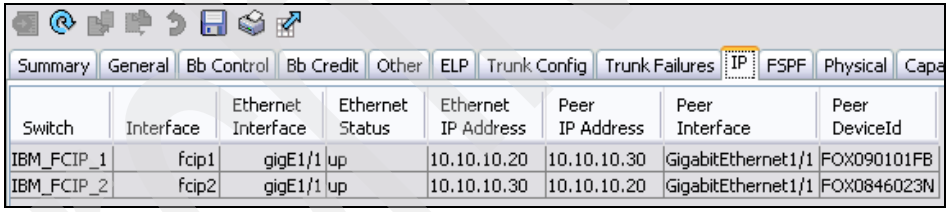
Notice that the tunnel is up and trunking (shown in bold).

7. We double-check this on switch 1.

Example 12-8 Checking tunnel and trunking on switch 1

```
IBM_FCIP_1# show fcip summary
-------------------------------------------------------------------------------
Tun prof Eth-if peer-ip Status T W T Enc Comp Bandwidth max/min (us)
                                E A A
-------------------------------------------------------------------------------
1   1    GE1/1 10.10.10.30 TRNK  Y N N  N   N    1000M/500M  1000
Both sides show the tunnel up and active.
```

8. We log in to FM and check the FCIP status there. Click **ISLs** in the Physical Attributes pane and select the **IP** tab from the window that displays above the SAN map. Figure 12-43 shows the results.



Switch	Interface	Ethernet Interface	Ethernet Status	Ethernet IP Address	Peer IP Address	Peer Interface	Peer DeviceId
IBM_FCIP_1	fcip1	gigE1/1	up	10.10.10.20	10.10.10.30	GigabitEthernet1/1	FOX090101FB
IBM_FCIP_2	fcip2	gigE1/1	up	10.10.10.30	10.10.10.20	GigabitEthernet1/1	FOX0846023N

Figure 12-43 FCIP CLI configuration summary

Getting more CLI information

You can download the *Cisco MDS 9000 Family Command Reference, Release 3.x* from the Cisco Web site at:

http://www.cisco.com/en/US/products/ps5989/products_command_reference_book09186a008066602d.html

Part 3

IBM m-type family

In this part, we discuss OEM products from McDATA.

13

IBM TotalStorage m-type family routing products

This chapter describes the IBM TotalStorage m-type routers. IBM offers the following m-type routers:

- IBM TotalStorage SAN04M-R (2027-R04): This is an entry-level SAN router in a 1U rack space that is designed for open systems IBM TotalStorage SAN solutions. It contains the following ports:
  - Two Fibre Channel (FC) ports at 1 Gbps speed
  - Two 1 Gigabit Ethernet ports
- IBM TotalStorage SAN16M-R (2027-R16), which provides:
  - The 16-port base SAN router in a 1U rack space
  - The standard edition of SANvergence Management software
  - Rack-mount kit
  - Fully populated 2 Gbps shortwave small form-factor pluggable (SFPs) on all ports

13.1 Product description

The IBM TotalStorage SAN m-type routers support applications that require interconnection of SAN islands to provide any-to-any connectivity to fabric-connected devices.

13.1.1 IBM TotalStorage SAN04M-R

IBM TotalStorage SAN04M-R (2027-R04) is the McDATA Eclipse 1620 SAN Router. It provides four ports in a 1U rack space. Two ports are Fibre Channel 1 Gbps ports; the other two ports are intelligent ports for IP connectivity. Each of the two IP ports is provided with two connectors: a standard RJ45 and an SFP. Either one of those can be used, but *not* both at the same time. The RJ45 connects Fast Ethernet, and the SFP is for Gigabit Ethernet connections.

The IP ports support Internet Fibre Channel Protocol (iFCP) or Internet Small Computer Systems Interface (iSCSI) connectivity. The base functionality supports Fibre Channel, iFCP, Ethernet, and iSCSI with a maximum number of 50 iSCSI server connections. The optional firmware version (iFCP Enterprise) adds compression and fast write functionality. For enhanced management of the SAN router, clients can order Enterprise SANvergence Management Software.

Note: At the time of writing, the SANvergence Manager is still available. However, it will soon be retired and the SANvergence Manager functions will be integrated in Enterprise Fabric Connectivity Manager (EFCM) 9.0.

SAN router features

The SAN router supports iSCSI, iFCP, and R_Port for trunking to both IP backbones and existing Fibre Channel fabrics. It connects to a wide range of end systems, including Fibre Channel, and Fibre Channel initiators and targets. The SAN routers support TCP/IP routing over extended distances at wire speed. They offer the following benefits:

- SAN internetworking for scalable and fault-tolerant SANs
- Support for full fabric, private, and public loop Fibre Channel devices
- Patent-pending fast write technology for maximizing throughput across long distances
- Compression for increased bandwidth

SAN router physical description

All the ports are located on the front of the SAN04M-R router. Only the Fibre Channel SFPs are field-replaceable units (FRUs). The cooling fans are on the rear and they are not hot-swappable.

When viewed from the front, standard power connections are on each side of the device. The unit contains two independent power supplies for redundancy and higher availability. Power supplies are not hot-swappable. The power supplies are input rated between 100 and 240 volts alternating current (V ac), at between 47 Hz and 63 Hz.

Fibre Channel and IP ports

There are two user-configurable Fibre Channel ports located on the front of the SAN router. The SFPs provide Fibre Channel connectivity at 1 Gbps. These ports can be configured as:

- FC_Auto (default)
- FL_Port
- F_Port
- L_Port
- R_Port

An LED to the left of each Fibre Channel port indicates the configuration and status of the associated port.

There are two intelligent ports for IP connectivity. Each IP port has two connectors, one standard RJ45 and one SFP. Either connector can be used, but not both at the same time. RJ45 connects Fast Ethernet, and SFP is for Gigabit Ethernet connections. The IP ports support iFCP or iSCSI connectivity.

Management ports

Two management ports are located on the front of the SAN router. An RS-232 serial port can be connected to a VT100 terminal emulator for access to the command-line interface (CLI). An RJ45 port can be connected to the local area network (LAN) for out-of-band management. The RJ45 management port can be accessed by any PC on the LAN with a Web browser or by the Enterprise Fabric Connectivity Manager (EFCM) server.

Operational features

Table 13-1 on page 382 describes the SAN router features. Some features are optional and might not be present in all SAN router software versions.

Note that both SAN16M-R and SAN04M-R do not support Metro Fibre Channel Protocol (mFCP) links and mSAN routing since firmware release 4.7 and later. We discuss mFCP and mSAN routing in “mFCP and iFCP” on page 389.

Table 13-1 Features of the SAN router

Feature	Description
Intelligent ports	Two intelligent ports can be configured for iSCSI or iFCP.
iFCP standards track protocols	The SAN router supports the Internet Engineering Task Force (IETF) draft standard for iFCP, which provides connectivity and networking for existing Fibre Channel devices over a TCP/IP network.
iSCSI	The SAN router is capable of providing iSCSI connectivity.
R_Port	Support for FC-SW2 standard E_Port and Brocade interoperability mode enables you to fully integrate the SAN router into an existing Fibre Channel SAN that includes one or more Fibre Channel switches. In the McDATA Open mode, R_Port supports Cisco and Qlogic E_Ports as well.
Fast write	The fast write software feature available on intelligent ports improves the performance of write operations between Fibre Channel initiators and targets in a wide area network (WAN). The improved speed depends on the WAN round-trip time (RTT), available buffer space on the target, number of concurrent I/Os supported by the application, and application I/O size.
Zoning	Using SANvergence Manager, network management software, or the CLI, you can create zones across networks. Use zone sets for periodic reallocation of network resources. For example, you can have one set of zones for daytime data transactions and another set of zones for nighttime backups.
Real-time and historical system logs	The Element Manager and LogViewer can be used to look at current system log messages from the connected SAN router.
Compression	Compression technology that is available on intelligent ports identifies repetitive patterns in a data stream and represents the same information in a more compact and efficient manner. By compressing the data stream, more data can be sent across the network even if slower link speeds are used.
Jumbo frames	Because the maximum Fibre Channel payload size is 2112 bytes, two regular Ethernet frames are required. The jumbo frame option extends the Ethernet payload to 2112 bytes. With the support of jumbo frames, a Fibre Channel frame can be mapped to one Ethernet frame, providing more efficient transport. For iSCSI traffic, up to 4 K size frames are supported.

Element Manager overview

The SAN Router Element Manager, a Web-based Java applet, is used to configure, monitor, and troubleshoot the router. Table 13-2 summarizes the configuration and monitoring functions of the Element Manager software.

Table 13-2 SAN Router Element Manager

Feature	Description
SAN router configuration	SAN router inband IP address Date and time System properties Default zoning behavior Password management SNMP traps
Port configuration	Fibre Channel and TCP ports (supporting iSCSI and iFCP) Management port Static routing
iFCP gateway configuration	iFCP setup Remote connection configuration Port redundancy configuration
iSCSI configuration	Device configuration RADIUS server configuration
SAN router operations	System log Upgrade firmware Reset the system Configuration backup and restore
Monitoring	Device view LEDs and icons, system information icons Message log setting Polling interval
Reports and statistics	Address Resolution Protocol (ARP) table Gigabit Ethernet port statistics Fibre Channel port statistics Fibre Channel device properties MAC forwarding table Storage Name Server (mSNS) Internet Protocol forwarding table remote gateway statistics Graphics port traffic statistics Ping iFCP Compression rate statistics VLAN configuration statistics

13.1.2 IBM TotalStorage SAN16M-R

IBM TotalStorage SAN16M-R (2027-R16) is the McDATA Eclipse 2640 SAN Router. It provides:

- The 16-port base SAN router in a 1U rack space
- The standard edition of SANvergence Management software
- Rack-mount kit
- Fully populated 2-Gbps shortwave SFPs on all ports

Twelve ports are user configurable as a 1 Gbps or 2 Gbps Fibre Channel or as a Gigabit Ethernet. The remaining four ports are intelligent Gigabit Ethernet ports, which support optional extended distance iFCP connectivity when activated. Base functionality of the 12 user-configurable ports provides SAN routing on up to two Fibre Channel ports, Fibre Channel fabric support, and iSCSI support on Gigabit Ethernet ports.

Clients can order three optional firmware versions: iFCP with fast write and compression on the four intelligent Gigabit Ethernet ports, SAN routing on any of the 12 user configurable ports, and comprehensive bundle with full iFCP and SAN routing capability. For enhanced management of the SAN router, clients can order the Enterprise SANvergence Management Software.

SAN router features

The SAN router supports iSCSI, iFCP, and R_Port for trunking to both IP backbones and existing Fibre Channel fabrics. The SAN router connects to a wide range of end systems including Fibre Channel, and Fibre Channel initiators and targets. SAN routers support TCP/IP routing over extended distances at wire speed. The SAN router offers:

- SAN internetworking for scalable and fault-tolerant SANs
- Support for full fabric, private, and public loop Fibre Channel devices
- Patent-pending fast write technology for maximizing throughput across long distances
- Compression for increased bandwidth

SAN router physical description

All ports and connectors are located on the front of the SAN router, except for the power connectors. The rear of the SAN router contains the power connectors and cooling fans. The FRUs are the optical transceivers and power supplies, which include internal fans.

Each of the two power connections supplies ac power to a different power supply for power redundancy and backup. Either power supply can support the SAN

router operation, but we recommend that you connect both, each to a different power source. If one power supply fails, the SAN router continues to operate, but you must replace the failed power supply immediately.

Fibre Channel ports

The 12 user-configurable Fibre Channel ports (labeled 1 through 12) are on the front of the SAN router. These port connections are SFP connectors that provide 1 Gbps or 2 Gbps Fibre Channel or 1 Gbps Gigabit Ethernet connectivity. These ports can be configured as:

- FC_Auto (default)
- FL_Port
- F_Port
- L_Port
- R_Port

To the left of each Fibre Channel port is an LED that indicates the configuration and status of the associated port.

Ethernet ports for IP connectivity

The SAN router provides four intelligent ports for Gigabit Ethernet connectivity, labeled 13 through 16. The red labeled ports are intelligent ports that can be configured for iFCP. The white labeled ports can be configured for mFCP.

Any intelligent port (red labeled) can be configured for iSCSI. Each port has 256 Mb in buffers: 96 transmit, 96 receive, and 64 for processing usage.

Management ports

Two management ports are on the front of the SAN router. An RS-232 serial port can be connected to a VT100 terminal emulator for access to the CLI. An RJ45 port can be connected to the LAN for out-of-band management using the SAN Router Element Manager or the SANvergence Manager. The RJ45 management port can be accessed by any PC on the LAN with a Web browser.

Operational features

Table 13-3 on page 386 summarizes the SAN router features. Some features are optional and might not be present in some SAN router software versions.

Table 13-3 Features of the SAN router

Feature	Description
Intelligent ports	Two intelligent ports can be configured for iSCSI or iFCP.
iFCP standards track protocols	The SAN router supports the IETF draft standard for iFCP, which provides connectivity and networking for existing Fibre Channel devices over a TCP/IP network.
iSCSI	The SAN router is capable of providing iSCSI connectivity.
R_Port	Support for FC-SW2 standard E_Port and Brocade interoperability mode enables you to fully integrate the SAN router into an existing Fibre Channel SAN that includes one or more Fibre Channel switches. In the McDATA Open mode, R_Port supports Cisco and Qlogic E_Ports as well.
Fast write	The fast write software feature available on intelligent ports improves the performance of write operations between Fibre Channel initiators and targets in a WAN. The improved speed depends on the WAN RTT, available buffer space on the target, number of concurrent I/Os supported by the application, and application I/O size.
Zoning	Using SANvergence Manager, network management software, or the CLI, you can create zones across networks. You can use zone sets for periodic reallocation of network resources. For example, you can have one set of zones for daytime data transactions and another set of zones for nighttime backups.
Real-time and historical system logs	The Element Manager and LogViewer can be used to look at current system log messages from the connected SAN router.
Compression	Compression technology available on intelligent ports identifies repetitive patterns in a data stream and represents the same information in a more compact and efficient manner. By compressing the data stream, more data can be sent across the network even if slower link speeds are used.
Jumbo frames	Because the maximum Fibre Channel payload size is 2112 bytes, two regular Ethernet frames are required. The jumbo frame option extends the Ethernet payload to 2112 bytes. With the support of jumbo frames, a Fibre Channel frame can be mapped to one Ethernet frame, providing more efficient transport. For iSCSI traffic, up to 4 K size frames are supported.

Element Manager overview

The SAN Router Element Manager, a Web-based Java applet, is used to configure, monitor, and troubleshoot the router. Table 13-4 lists the Element Manager software configuration and monitoring functions.

Table 13-4 SAN Router Element Manager

Feature	Description
SAN router configuration	SAN router inband IP address Date and time System properties Default zoning behavior Password management SNMP traps
Port configuration	Fibre Channel and TCP ports (supporting iSCSI and iFCP) Management port Static routing
iFCP gateway configuration	iFCP setup Remote connection configuration Port redundancy configuration
iSCSI configuration	Device configuration RADIUS server configuration
SAN router operations	System log Upgrade firmware Reset the system Configuration backup and restore
Monitoring	Device view LEDs and icons, system information icons Message log setting Polling interval
Reports and statistics	Address Resolution Protocol (ARP) table Gigabit Ethernet port statistics Fibre Channel port statistics Fibre Channel device properties MAC forwarding table Storage Name Server (mSNS) Internet Protocol forwarding table remote gateway statistics Graphics port traffic statistics Ping iFCP Compression rate statistics VLAN configuration statistics

13.2 SAN router architecture

This section describes the basic functions, features, and internal architecture of the McDATA routers. It also explains basic terminology used in McDATA SAN routing.

13.2.1 SAN routing terminology

The introduction of SAN routing technology brought with it new jargon and terminology. Routing has caused us to define new terms so that we can describe routed SANs and their properties effectively. In addition to the standard SAN routing terms, each vendor uses different terms along with their products. Table 13-5 introduces some terms.

Table 13-5 McDATA SAN routing terms

Term	Definition
R_Port	The SAN routing port. It is used on McDATA SAN router side to identify a connection to a Fibre Channel switch. The opposite end to an R_Port on the Fibre Channel switch is the E_Port.
mSAN	Metro area SAN. This is the actual fabric, which is interconnected through the SAN router to one or more other fabrics.
iSAN	Internetworked SAN. iSAN is a logical name that represents one or more mSANs (fabrics) interconnected through SAN routers across larger distances (outside the metro area).
IRL	Inter-router link. This is an IP-based connection between two SAN routers. It uses the iFCP.
iFCP	Internet Fibre Channel Protocol. This protocol connects two or more mSANs together. Usually, it uses the external, high-latency, lower-bandwidth networks outside the metro area.

Figure 13-1 on page 389 shows an example of interconnected mSANs in an iSAN. From firmware release 4.7 and later, an mSAN can consist only of one SAN04M-R or SAN16M-R. These can be interconnected through one or more IRLs, using the iFCP. Inter-router iFCP connections provide path failover capability. We recommend that you always use at least two connections for availability. All of the mSAN routers are interconnected together through an external WAN and come together to build an iSAN using the iFCP.

Figure 13-1 on page 389 also shows an example of a typical routed SAN. It is a set of individual fabrics interconnected by SAN routers. A routed SAN functions as a single, large SAN and provides any-to-any connectivity, while keeping the

particular fabrics autonomous. Should any event occur within an individual fabric, such as component failure, fabric reconfiguration, or a registered state change notification (RSCN) broadcast, such an event will not be propagated into the other fabrics.

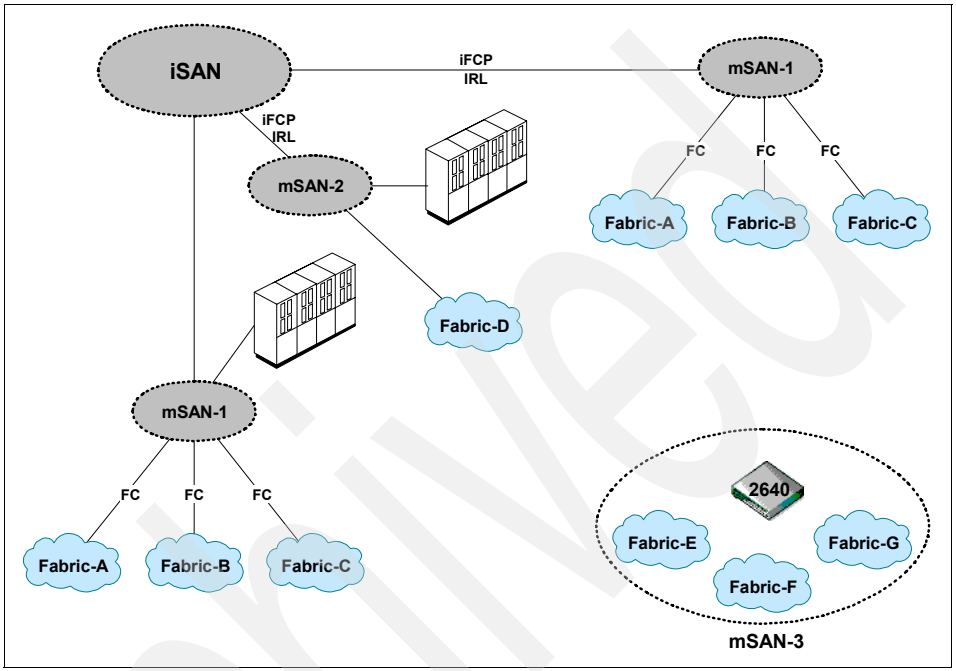


Figure 13-1 mSAN and iSAN interconnections

13.2.2 SAN routing features

This section describes the various features used in McDATA SAN routing.

mFCP and iFCP

mFCP and iFCP interconnect McDATA routers. Even though these protocols are similar to each other from a high-level perspective, they are used in different environments and have different features.

The TCP stack of both iFCP and mFCP is governed by a Saturn processor. Each iFCP port has its own dedicated Saturn processor. Apart from the TCP stack, the processors accommodate fast write and compression as well.

The mFCP interconnects another SAN16M-R router in a metro area, by using a high-speed, low-latency, usually dedicated LAN or VLAN. There is no mFCP

support on the SAN04M-R router. Also, since firmware version 4.7, there is no support for mFCP connections *between* SAN16M-R routers.

The Gigabit Ethernet connection for mFCP must be full duplex, with symmetric flow 802.3x control. However, it does not use fast write, compression, and rate limiting. It is a McDATA proprietary User Datagram Protocol (UDP) and currently requires a direct connection. mFCP uses Gigabit Ethernet ports only.

Both routers that use mFCP must be on the same subnet. Two SAN16M-R routers cannot be interconnected together with a Fibre Channel-based IRL. UDP by design does not support retransmission, packet reordering, or duplicate packets detection. Therefore, a fast and reliable connection is required. To prevent buffer overflow on each router's side, which would lead to the drop of packets, the SAN16M-R uses 802.3x Symmetric Flow Control.

In addition, up to four mFCP links can be aggregated into a single pipe by using 802.3ad Link Aggregation. Each packet uses Differentiated Service Code Point (DSCP) to ensure quality of services (QoS). DSCP is defined in RFC 2598 and defines Expedited Forwarding. In the McDATA mFCP implementation, Expedited Forwarding has set the "do not fragment" flag. The result is that packets are transmitted at the maximum possible maximum transmission unit (MTU) and do not fragment if a device with a lesser MTU set is in the path. Rather, the packet is dropped. Make sure that you use the appropriate MTU size when planning to implement SAN16M-R routers with an mFCP link or links into your environment.

Tip: For RFC 2598 and related RFCs on the Web, see:

`http://www.rfc-editor.org/`

iFCP is used in completely different environments, having lower bandwidth and high latency links over greater distances, usually across a WAN. The iFCP is a TCP-based protocol. Therefore, a dedicated CPU runs the TCP stack within the SAN16M-R. Only the red labeled ports on the SAN16M-R have the dedicated CPU, and you cannot run iFCP on any other ports but these.

TCP itself provides a variety of services, such as packet retransmission, packet reordering, and duplicate packet detection. However, to transmit storage traffic effectively (to use the available bandwidth at a maximum possible sustainable rate), further optimization is implemented. These are the fast write algorithm, LZO compression, and rate limiting.

The 802.3ad Link Aggregation is not implemented in iFCP. One iFCP link can serve up to 64 TCP connections per single R_Port with the SAN04M-R or 256 TCP connections per single R_Port with the SAN16M-R. A TCP session is

established as a result of PLOGI, PDISC, or ADISC in one of the iSAN-connected fabrics.

iFCP maps Fibre Channel frames to an IP datagram. One Fibre Channel frame is mapped per single IP datagram. Figure 13-2 shows the Fibre Channel to IP encapsulating mechanism.

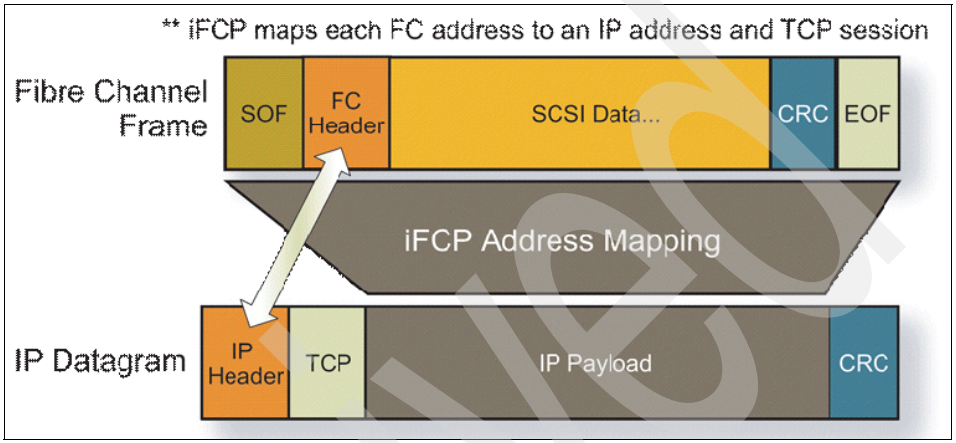


Figure 13-2 Fibre Channel frame to IP datagram encapsulation

Note: To compare different gateway protocols, such as iFCP and FCIP, refer to 1.2, “Gateway protocols” on page 4.

Table 13-6 summarizes the fundamental differences between mFCP and iFCP.

Table 13-6 mFCP and iFCP comparison

Technology	mFCP	iFCP
OSI Layer 4 Protocol	UDP	TCP
Intelligent ports with CPU only	No	Yes
Fast write	No	Yes
Compression	No	Yes
Rate limiting	No	Yes
802.3ad Link Aggregation	Yes	No
802.3.x Flow Control	Yes	Yes
Must be on the same inband IP subnet	Yes	No

The inband IP address is the router's internal SAN address, as shown in Figure 13-7 on page 398.

Fast write

When a SCSI write operation is performed in a native Fibre Channel environment, multiple handshakes occur before a block of data is sent from a SCSI initiator to a SCSI target. First, the SCSI initiator sends a message with the total amount of data it wants to send. The SCSI target responds with a transfer ready message (FC_XFER_RDY) and indicates how much data it is prepared to receive. When the amount of data to send is settled between the SCSI initiator and target, the initiator sends the data. After a successful write, the SCSI target sends another FC_XFER_RDY and the handshake starts over again.

When SCSI write operations are performed over greater distances, each handshake message directly impacts the time to complete the write operation. This is not that big a problem in a pure Fibre Channel environment, which is by design a low-latency, high-throughput environment. However, transferring SCSI handshakes and data over distance using the IP environment has a significant impact on performance, as shown by the example in Figure 13-3 on page 393.

McDATA Fast Write helps mitigate the impact of the higher latency of IP networks over distance. When a write operation is started by an initiator, the local router forwards it to the remote router. That is the normal SCSI operation, which ensures that the commands are delivered to the target in the same order as initiated on the initiator.

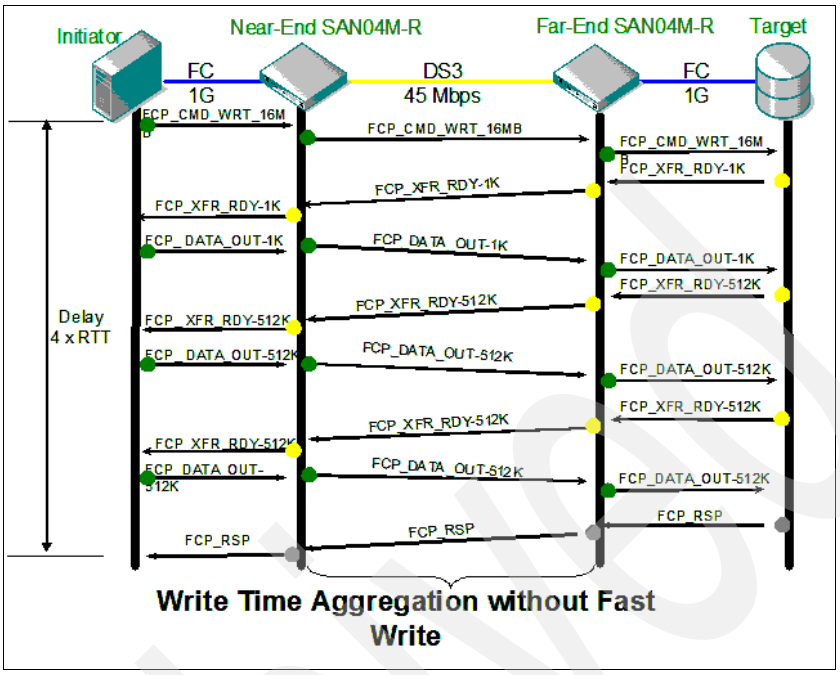


Figure 13-3 SCSI write over high-latency environment without fast write

The remote router, however, acts like a virtual target to the local initiator and sends back the FC_XFER_RDY message, prompting it to send the whole data segment for the write operation. The initiator sends the data and does not require any other handshake messages. From the remote target point of view, the remote router acts as an virtual initiator. FC_XFER_RDY messages are exchanged between this virtual initiator and remote target, so the handshake round-trip messages do not have to travel across the high-latency link. Figure 13-4 on page 394 shows the flow of the SCSI write operation between remote sites using fast write.

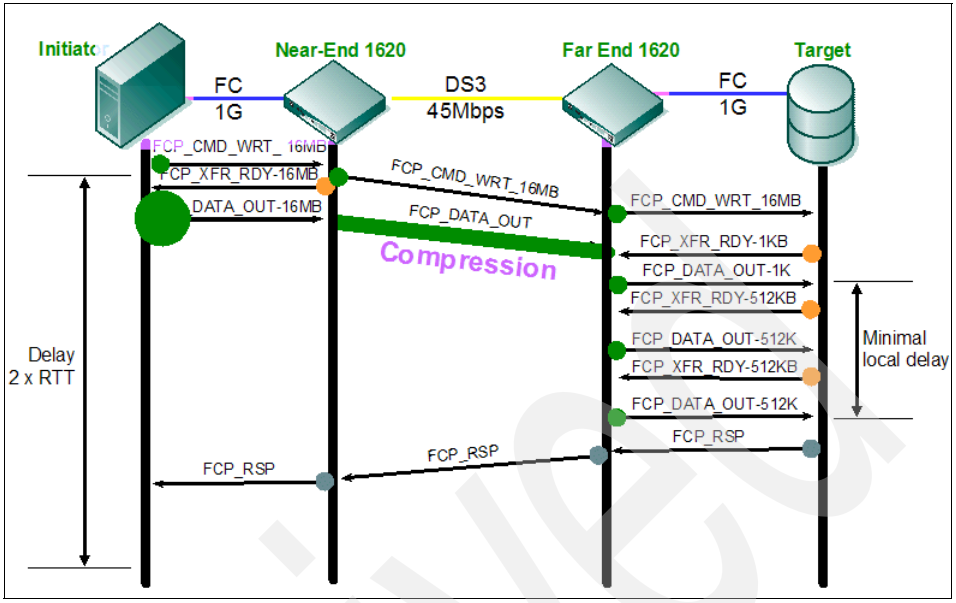


Figure 13-4 SCSI write over high-latency environment with fast write

Fast write tracks the status of all of its open Fibre Channel sessions at both the local and remote site. Any error condition on the target is detected by the initiator during the final SCSI completion message and leads to error recovery. Also, the Fibre Channel protocol has incorporated a checksum mechanism in its design to ensure data integrity. The fast write does not interfere by the final SCSI completion message; this is sent at the end of the session to the real initiator. All of these aspects ensure data integrity, so there is no real danger of corrupt data when using fast write.

iSCSI gateway

The m-type routers enable communication between iSCSI initiators and Fibre Channel targets. Fibre Channel storage can be either directly attached to the router or can use standard R_Port connections from connected fabrics. iSCSI initiators can be directly connected to the intelligent ports on m-type routers, or they can use an intermediate network.

The iSCSI protocol is designed to map the SCSI protocol over TCP/IP. Figure 13-5 shows SCSI to IP mapping.

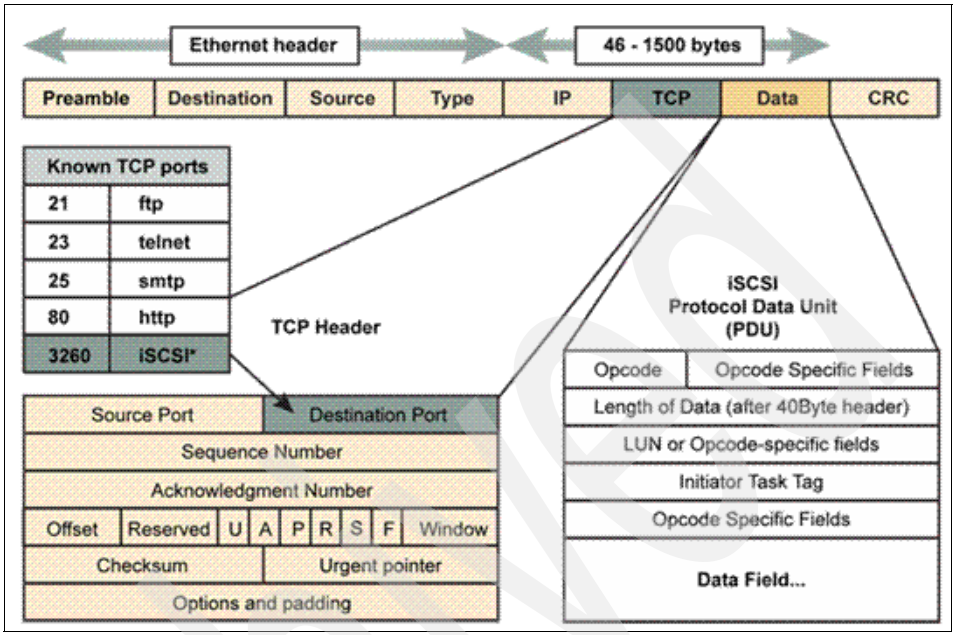


Figure 13-5 iSCSI protocol

In an iSCSI session, initiators establish iSCSI sessions with targets. Session IDs are generated to uniquely identify individual conversations between specific iSCSI nodes within the corresponding network entities. An initiator logging on to a target includes its iSCSI name and an initiator session ID (ISID). A target, responding to the login request, generates a unique target session ID (TSID), in combination with its iSCSI name. A single ISID/TSID session pair can have multiple TCP connections between them, as per the results of login negotiation. However, if multiple TCP connections for that session have been established, individual command and response pairs must flow over the same TCP connection. This is known as *connection allegiance*.

Figure 13-6 shows how you can use an m-type router to connect iSCSI hosts to your existing Fibre Channel environment.

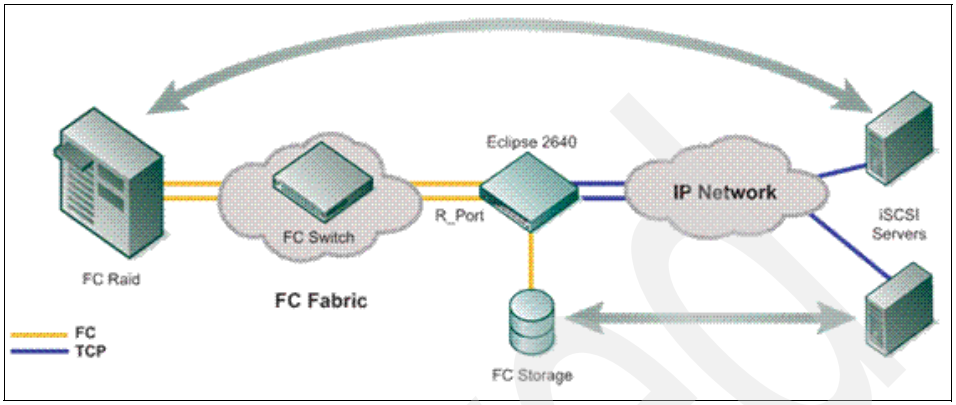


Figure 13-6 Connecting iSCSI servers to an existing fabric using m-type router

Selective Acknowledgement

TCP Selective Acknowledgment (SACK) is defined by RFC 2018. On a “lossy” network, when using SACK, the receiving host informs the sender that the data has been received. That means the receiving host can acknowledge packets out of order. The sender then retransmits only those packets that have been lost. For example, if the receiving host acknowledges packets 1, 2, 3, 4, 6, and 8, then only packets 5 and 7 are retransmitted.

Without SACK, the sender has to retransmit all packets *after* the first missing packet, or in this example, packets 5, 6, 7, and 8.

Compression

Another optimization implemented in the McDATA SAN routers (both SAN04M-R and SAN16M-R) is compression. The LZO algorithm is implemented in the McDATA routers and runs on a dedicated Saturn processor on each iFCP port. It is not done on the shared CPU of the Eclipse SAN router itself. You can choose from four different compression modes:

- LZO

  This is a frame-based algorithm. If you have many active initiator-target sessions opened on the iFCP link in your environment, this method works best.

- Fast LZO with history

  This mode provides a 2-byte compression at a time with keeping the 8-byte history. This uses more memory, but offers a higher compression ratio. We recommend that you use this mode in an environment with a faster link (at least T3 and faster) and with a fewer active initiator-target sessions.

- LZO with history

  This mode provides one byte of compression at a time along with keeping the 8-byte history. It offers a higher compression ratio at the expense of speed. Therefore, we recommend that you use this mode in an environment with fewer initiator-target sessions and a slower link (such as T3 and E3).

- Deflate

  This mode is suited to provide the best compression ratio compared to other modes, but at the cost of speed. Therefore, we recommend that you use this mode on slower lines, such as 10 Mbps Ethernet and slower.

Rate limiting

Rate limiting prevents ingress traffic from entering faster than egress traffic, which results in buffer overflow and dropping packets. Dropped packets cause TCP to resort to flow-control, which leads to a throughput decrease.

Interoperability mode

McDATA supports OEM interoperability through the use of Open Fabric mode. All m-type family SAN routers have set the Open Fabric mode as the default mode. With Open Fabric mode, worldwide name (WWN) zoning is available, and port zoning is not. Features that are implemented differently by each vendor might be unavailable, too. McDATA and Brocade interoperability is not supported by IBM except by an RPQ.

13.2.3 SAN routing architecture

We can differentiate between routing at an mSAN level, within a particular router, and routing at an iSAN level, among routers interconnected through an iFCP link. Routers do not need to be on the same subnet.

To provide routing services, m-type family routers use the following internal network architecture:

- Router internal IP address
- SAN internal IP address for each port
- External IP address for each non-Fibre Channel port

Table 13-7 summarizes the default IP address settings for the SAN04M-R and SAN16M-R routers.

Table 13-7 Default IP address settings for SAN04M-R and SAN16M-R routers

Description	SAN04M-R default	SAN16M-R default
Router internal IP address	192.168.111.100	0.0.0.0
SAN internal IP addresses Subnet mask	192.168.111.103 192.168.111.104 255.255.255.0	0.0.0.0 0.0.0.0
External IP addresses Subnet mask	0.0.0.0 0.0.0.0	0.0.0.0 0.0.0.0
Default gateway	0.0.0.0	0.0.0.0
Management IP address subnet mask	192.168.100.100 255.255.255.0	192.168.100.100 255.255.255.0

Figure 13-7 shows a diagram of the router internal network architecture.

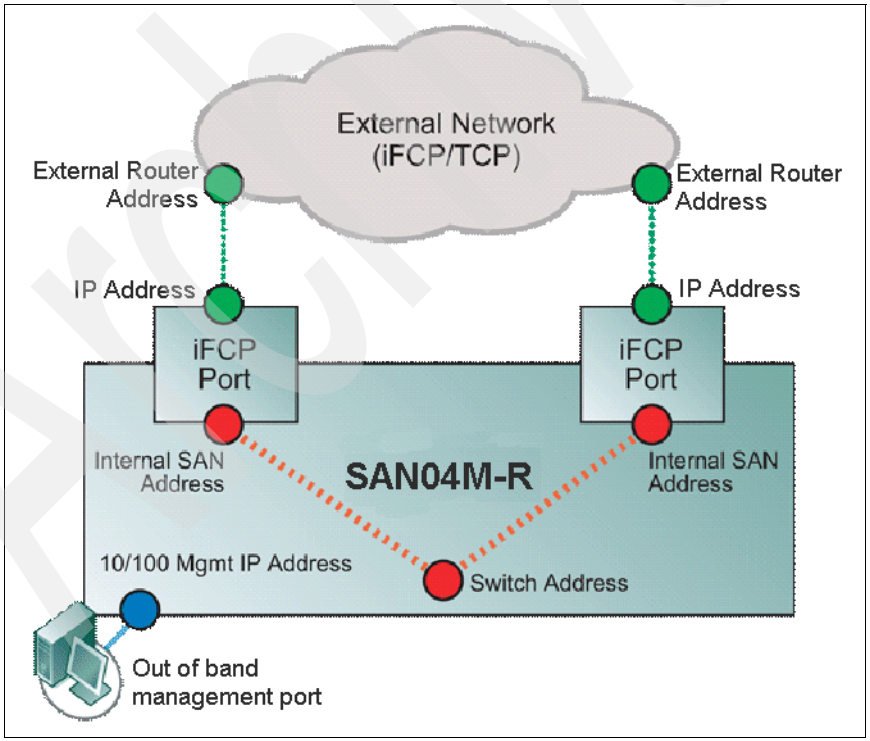


Figure 13-7 SAN router internal network architecture

Routing at the fabric level

Even though routing at the fabric level is not done by SAN routers, we discuss it here briefly for the sake of completeness. At the fabric level, switches are interconnected through E_Ports. Path selection (routing) is governed by Fabric Shortest Path First (FSPF) based on calculating the cost of the particular path. The Simple Name Server (SNS) service is used to register Fibre Channel nodes. All switches keep the SNS updates and are, therefore, all aware of the devices in the SNS.

Routing at the mSAN level

At the mSAN level, fabrics are interconnected together, but not merged. The interconnection is done using one or more SAN routers. There can be no more than two SAN16M-R routers in a single fabric and only one SAN04M-R in a single fabric. SAN16M-R routers can be connected together with up to four IRL mSAN links.

Fabric manager

Routing at the mSAN level uses each router's R_Ports. One of the R_Ports connected to a particular fabric is in charge of controlling the routing to and from that particular fabric. This port is called the *fabric manager*. The fabric manager is always the port with the lowest node worldwide name (nWWN). It is selected during the fabric build when all nWWNs are sent across the fabric. Only the R_Port nWWNs are eligible to become the fabric manager. To distinguish the R_Ports from other ports in the fabric, especially E_Ports, the following method is used.

Each fabric component vendor has assigned to it a particular address range, the Organizational Unique Identifier (OUI). Within this pool of addresses, McDATA reserved a subset of addresses for routers only. When nWWNs are sent across the fabric build, the m-type family routers recognize the R_Ports and assign fabric manager R_Port to the port with the lowest nWWN. After the fabric manager is selected, it cannot be changed. It remains the fabric manager until another fabric build occurs. The SAN router cannot start a fabric build; it can only be initiated by a switch within the particular fabric.

Name server in mSANs

In an mSAN, two name servers are used: the primary Metro Simple Name Server (primary mSNS) and the secondary Metro Simple Name Server (secondary mSNS). The primary mSNS keeps a database of all the WWN nodes from that particular mSAN and other mSANs. It propagates this database to the entire iSAN (other mSANs' primary SNSs). The database of the primary mSNS is fed with data from each fabric's SNS. The primary mSNS uses the fabric manager for communication. The secondary mSNS serves as a client to the primary mSNS.

Its database contains only those WWN nodes entries that are locally connected to the router.

Both mSNSs must be on the same inband IP address subnet. Otherwise, they will not be able to communicate with each other. The inband IP address is the internal IP address of the router, as shown in Figure 13-7 on page 398.

The primary mSNS is selected automatically by the fabric build, or it can be set manually. Primary and secondary mSNSs are synchronized at the time of primary mSNS selection only, which occurs during the fabric build. After that, updates are sent through RSCNs. RSCNs are sent by using subnet broadcasts, for example to all devices on 10.0.0.255 and 255.255.255.0. Universal broadcasts, such as 255.255.255.255, are not used. If there is a change in the database in the secondary mSNS, a unicast packet is sent to the primary mSNS. The primary mSNS then sends a broadcast to all listening m-type routers.

The mSNS database stores the following information:

- All storage entities in a local IP network, such as storage devices, Fibre Channel hosts, iSCSI initiators, and m-type routers
- WWN addresses for each Fibre Channel node
- Type of protocol and its properties for each registered entity
- iSCSI initiator name

The following services are provided by an mSNS:

- mSNS Registration service
- mSNS State Change Notification service
- mSNS Keyed Query Service

The mSNS is responsible for partitioning the mSAN into zones.

Zoning

Zoning at the mSAN level is governed by mSNS. It provides a similar functionality to Fibre Channel zoning. Only devices that are members of a particular zone are allowed to communicate with each other. To share devices across an mSAN and between two SAN16M-R routers, perform the following steps:

1. Create mSAN zones with unique mSAN zone IDs (range from 1 through 512) on both routers. Zone names can be different unless they have the same mSAN zone ID.
2. If necessary, set a maximum bandwidth limit for your zones.
3. Add local devices into each zone on both routers.

Now zone members are mutually shared across the mSAN.

Routing domains and Fibre Channel network translation

The fabrics connected together by a router are kept separate. They do not merge together into one large fabric, and the Class F traffic is not allowed to pass through the router. Considering this, it is necessary to have a mechanism to enable traffic among those fabrics, the Fibre Channel network translation mechanism.

To allow cross-fabric communication, a device from one fabric must be presented with a unique fabric ID (domain ID) in the remote fabric. The problem is that usually the address space (fabric domain IDs) is reused in fabrics and, therefore, is not unique in mSANs. Two domain IDs, known as *routing domains*, are reserved for the purpose of unique addressing among fabrics. These reserved routing domains are:

- 0x7E for routing among fabrics within the mSAN boundaries
- 0x7F for routing among mSANs and for devices directly connected to one of the Fibre Channel ports on the router

If a remote nWWN node wants to communicate with another nWWN node within a local fabric, its fabric ID will always start with 0x7E.

Next, to locate the fabric that particular node comes from, four area IDs are available per single fabric and mapped to the routing domain. Four area IDs provide for addressing up to 1024 devices (4 x 256).

Finally, each fabric's R_Port has assigned to it the fabric ID of that particular fabric. This must be configured manually from the router's configuration interface.

Table 13-8 shows the mapping of the area IDs to fabric IDs for domain 0x7E.

Table 13-8 Area ID to fabric ID mapping for routing domain

Routing domain ID	Area ID	Fabric ID
0x7E	1 - 4	1
	5 - 8	2
	9 - 12	3
	13 - 16	4
	17 - 20	5
	21 - 24	6

Each egress traffic from the particular fabric undergoes network address translation (NAT) into 0x7E. Its fabric ID is mapped to a corresponding area ID range.

For example, if an initiator wants to communicate with a target in another fabric, the following actions occur:

1. The initiator queries the primary mSNS to see if the target is within the same zone and what its fabric ID is.
2. Because the target is in another fabric, the initiator sends a Fibre Channel frame to the router's routing domain 0x7E.
3. The router then performs the NAT back on the original domain of the target fabric and forwards the frame (received from the initiator) to the appropriate R_Port connected to the fabric where the target resides.

Note: The format as to how the routing domain IDs are propagated to each fabric depends on the mode in which the fabric operates. In the McDATA Open Fabric mode, routing domains are 0x7E and 0x7F. However, in both the McDATA and Brocade native mode, the routing domains are presented as 30 and 31, respectively.

Path selection

Two aspects influence the selection of the R_Port when mSAN routing occurs:

- Path cost is considered using the standard FSPF mechanism.
- If more than one destination with the same FSPF cost exists, one path is selected using a round-robin algorithm.

However, traffic from one initiator is always sent through the same R_Port, unless the following actions occur:

- The target is disconnected from the fabric.
- A R_Port reset occurs.
- The fabric is rebuilt.

Only the traffic from multiple devices is subjected to round-robin among equal cost R_Ports (if there are any), but not from one particular device. However, SAN routers do not have control over the inter-switch link (ISL) selection between the switch and director and a SAN router.

Routing at the iSAN level

The iSAN consists of two or more mSANs interconnected with an existing WAN link, enabling the sharing of devices among mSANs over greater distances. The WAN infrastructure consists of existing IP switches and routers, as shown in Figure 13-8.

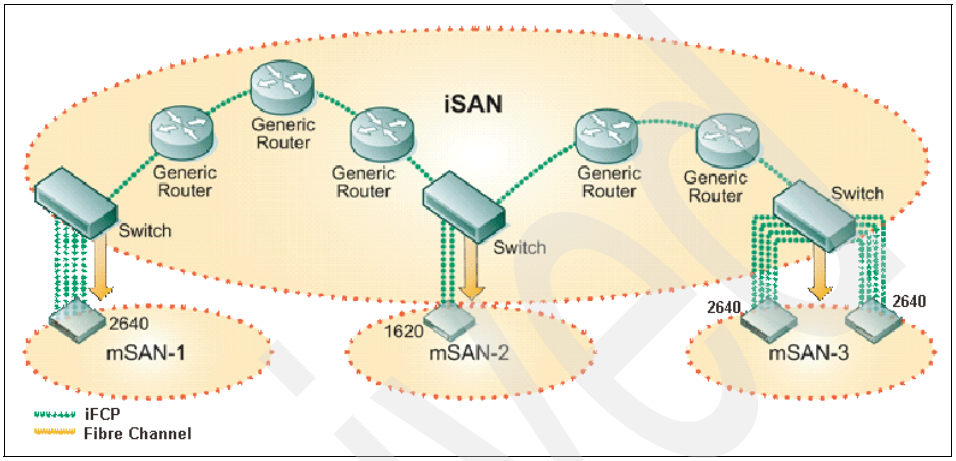


Figure 13-8 An example of iSAN architecture

From the iSAN perspective, we can observe that the following actions take place:

- Routing at the fabric level, fabric-level zoning, and the primary mSNS in control
- Routing at the mSAN level, mSAN-level zoning, and both primary and secondary mSNS in control
- Routing at the iSAN level, iSAN zoning, using the same mechanism as mSAN zoning, primary mSNS in control

The path is selected using FSPF. Note that the cost behind a router (everything presented as routing domain 0x7E) is not advertised at the mSAN or iSAN level. This means that equal cost might be seen from the FSPF perspective, not taking into consideration any other hops behind a router. As shown in Figure 13-9, servers from fabric 1 see both iSAN and mSAN storage on the right side of the picture as directly attached to the router (presented as 0x7E), without considering FSPF within the fabric.

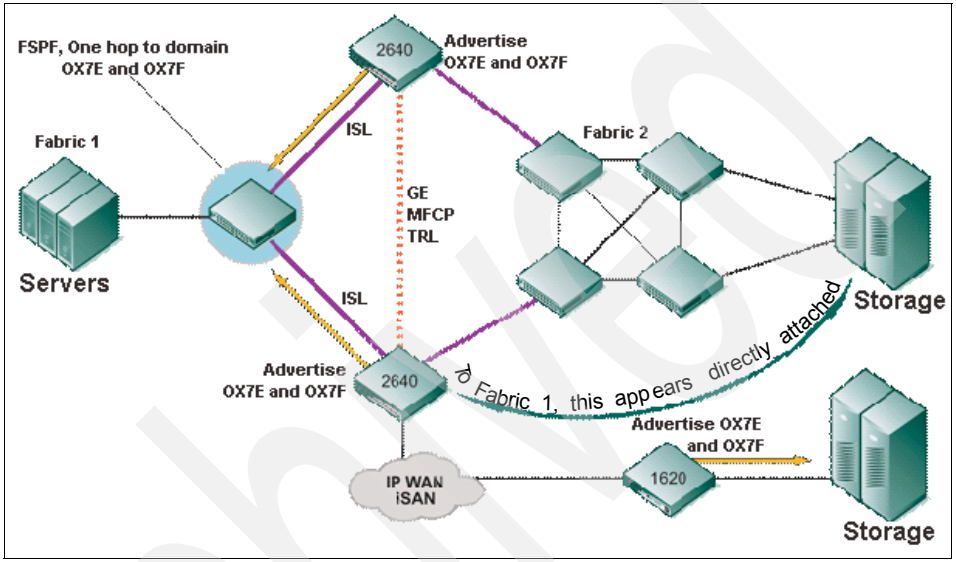


Figure 13-9 Additional FSPF costs behind the 0x7E domain

14

IBM TotalStorage m-type family solutions

Routers provide access to data that is located in a different fabric. The principal uses for storage routing are:

- Storage area network (SAN) extension over IP networks
- Lowering connection costs using Small Computer System Interface over IP (iSCSI)
- Achieving isolation and interoperability between different business units
- Managing scalability as your SAN environment grows
- Migrating from an older storage environment to a newer one

There are two IBM m-type OEM products to accommodate these features: the SAN04M-R and the SAN16M-R. The former one is especially suitable for iSCSI consolidation, simple SAN extension, and migration. The latter one fits well for SAN extension and scalability and fabric isolation solutions.

14.1 SAN fabric local FC-FC routing

Figure 14-1 shows the local Fibre Channel-Fibre Channel (FC-FC) routing solution.

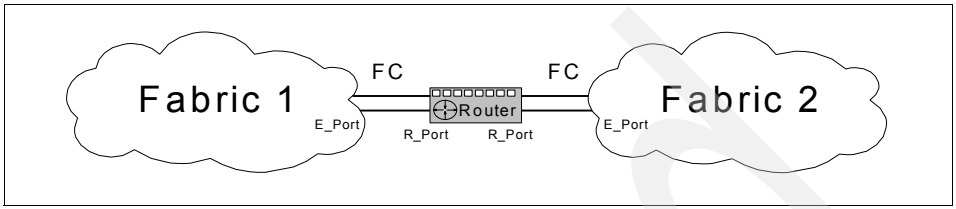


Figure 14-1 Local FC-FC routing between two SAN fabrics

In this case, a SAN16M-R is connected to different SAN fabrics using redundant Fibre Channel connections. From each fabric's switch point of view, the connection to the router appears as a standard E_Port, while on the router, the port is the R_Port. We can extend this configuration to span up to six SAN fabrics, each connected with two Fibre Channel connections.

If you have multiple switches in your fabric, we recommend that you distribute the connections to routers across them for maximum availability. If you are using a core edge fabric, we recommend that you connect routers to the core switches.

Some of the solutions that can be provided by local FC-FC routing include:

- Scalability

  Fibre Channel addressing can theoretically support up to 16 million nodes in a single fabric. However, the practical limit for the number of nodes is much lower. This is similar to TCP/IP networks, where the network address space is divided into smaller subnets of limited number of IP addresses, and traffic is routed between them. Usually, the practical limit of a fabric from both technical and management standpoint is somewhere between 250 and 1000 nodes.

  FC-FC routing enables you to divide the environment into several fabrics, while providing access to shared resources between fabrics.

- Multiple SAN administrators

  In many cases, enterprises have several small SAN islands that are managed by different SAN administrators, often as a result of mergers or acquisitions. Using FC-FC routing between the fabrics allows each fabric to be managed separately from other fabrics. It also prevents the propagation of any management errors to other fabrics.

The mSAN configuration is the only part of the fabrics that needs to be coordinated among the SAN administrators. Because the mSAN zones need to be defined on all fabrics before they can route traffic, the devices in each fabric are protected against unplanned access from the other fabrics.

- Interoperability between storage vendors

  With FC-FC routing, you can separate fabrics built on different vendor's equipment to their own SAN fabrics, as shown in Figure 14-1 on page 406. This helps to avoid problems with interoperability. This solution also allows each edge fabric to be supported and managed by the corresponding storage vendor, while enabling storage access between different SAN fabrics.

- Interoperability between old and new fabrics

  In many cases when implementing a new SAN fabric, you already have an existing fabric. The existing fabric might have some parameter settings that you want or need to have set up differently in the new fabric. One good example is the Core PID setting.

  By using FC-FC routing to connect the fabrics, you do not need to change the settings in the old fabric and are free to choose the settings you need for the new fabric. You can also use only a single Fabric OS level in any fabric, independent on the Fabric OS levels supported by the old hardware.

- Migration between old and new fabrics

  The storage hardware is usually replaced with new hardware every three to five years. When refreshing the disk hardware, it might make sense to refresh the SAN hardware as well, especially if the new disk vendor is different from the former vendor.

  FC-FC routing enables you to implement the new disk subsystems and SAN fabric in the final configuration. It also enables you to connect the complete new environment to the current SAN fabrics. This way, you have simultaneous access from the servers to both old and new disk subsystems and can use server-based tools, such as Logical Volume Manager (LVM), to migrate the data from the old disks to the new disks.

  After you migrate any host to the new disks, you can move the Fibre Channel ports of the server to the new SAN fabric. Because you can do this one server at a time, the outage needed is minimized.

- Storage consolidation

  Many enterprises implement a separate SAN fabric for tape backups. Without FC-FC routing, this requires a separate Fibre Channel adapter in each server that needs to be connected to the backup devices, as well as the additional fiber cabling to support these adapters. If you set up FC-FC routing between the normal SAN fabrics and the backup fabric, you can share the tape devices across any adapters in any fabric, as required.

Another example of storage consolidation is to implement a single IBM TotalStorage SAN Volume Controller (SVC) cluster across multiple SAN fabrics.

14.2 SAN extension with iFCP

Fibre Channel distances have been traditionally limited to either local fiber runs, using 9 micron long wave Fibre Channel, or high-quality wide area networks (WANs) such as SONET and SDH in combination with coarse wavelength division multiplexing (CWDM) or dense wavelength division multiplexing (DWDM) multiplexers.

The advent of Internet Fibre Channel Protocol (iFCP) has meant that applications that can tolerate the high latencies of IP networks can now make Fibre Channel connections across standard corporate IP WANs. The advantages of this are that it uses a widely available and well-understood infrastructure, which translates into lower cost.

We are still in a phase where people want iFCP over standard networks to be a panacea for all SAN extension applications. The inherent latencies involved are around 5 microseconds per km traveled in each direction with added latencies at every step (for example, up to 100 microseconds per router or firewall). This generally prevents iFCP from being used effectively for applications such as synchronous replication or online transaction processing (OLTP). For example, a high-quality network of 1000 km might have a latency of around 20 milliseconds. Given that a disk I/O might only take 10 milliseconds itself, the problem with a 20 millisecond latency becomes obvious.

Because some corporate WANs provide uncertain quality of service (QoS), storage router vendors tend to be cautious about quoting distances for iFCP. They generally recommend that high quality WANs are necessary to provide services over anything more than 200 km or 300 km.

In practice, the most common uses for iFCP are asynchronous replication and non-critical access over campus or metro distances. A client might choose to implement Fibre Channel mapped into IP on a campus scale simply because the IP links are already in place. On a short IP network, the main problem becomes QoS because the latencies are not so large. The principles are the same whether running over 500 meters or 5000 km. All that varies is the link latency, the service reliability, and consistency.

Compression

iFCP compression in the m-type family routers increases the effective WAN bandwidth. Although Gigabit Ethernet ports for IP Storage Services can theoretically achieve up to a thirty to one (30:1) compression ratio, typical ratios in the field are less than two to one (2:1).

The compression feature is implemented in both SAN04M-R and SAN16M-R routers and can be used with iFCP links. For more information regarding compression in m-type family routers, see “Compression” on page 396.

Using jumbo packets can also improve throughput. However, keep in mind that jumbo packets need to be turned on through the entire data path.

SAN extension over 700 km distance example

Figure 14-2 shows an example of an asynchronous replication running over IP at a 700 km distance.

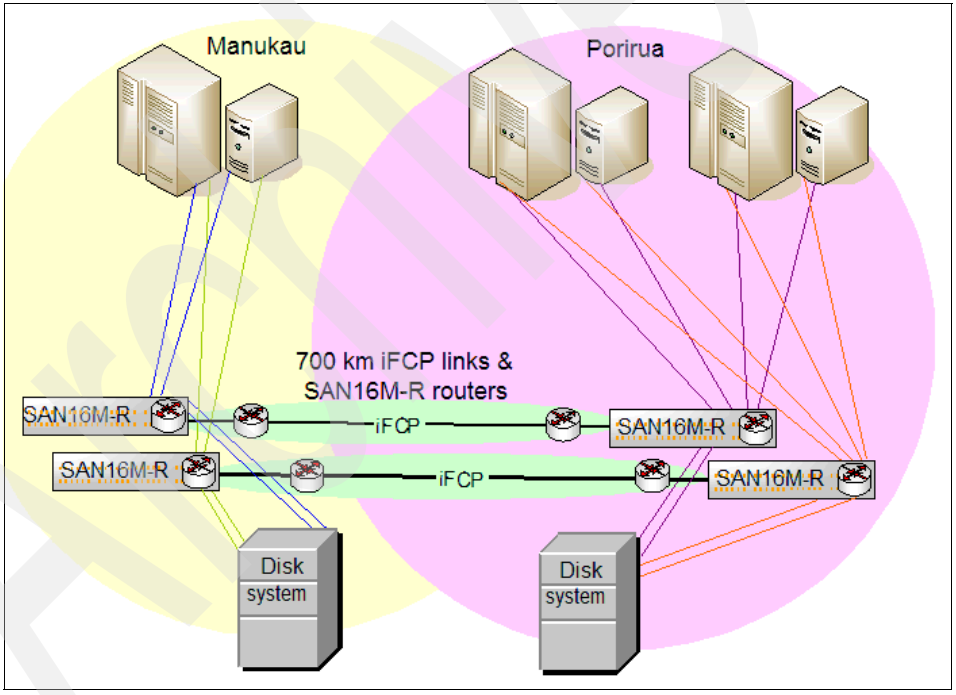


Figure 14-2 SAN extension over IP using iFCP over 700 km distance

Using iFCP fast write

The fast write mechanism is an attempt to mitigate the transport latency associated with long-distance SCSI I/O operations. Fast write performance depends heavily on the traffic profile of the SCSI operations being performed. The following I/O characteristics are well-suited to benefit from fast write:

- Long-distance, high-latency network
- Write intensive I/O
- High number of small SCSI writes (rather than low number of large writes)
- Disk system with low write latency

These characteristics suggest that the best use for fast write is in disk system replication. However, some replication solutions, such as IBM Metro Mirror and IBM Global Mirror, already use a mechanism similar to fast write.

Across a 100 km link, replication using fast write can be expected to deliver around 10% improvement in throughput and a similar percentage reduction in latency on a given FCIP network.

Learn more details in “Fast write” on page 392.

14.3 Low-cost connection with iSCSI

There are three common ways to create low-cost connections to disk storage:

- Fibre Channel Arbitrated Loop (FC-AL)

  Using FC-AL does not require a switch port for each server, because up to 126 devices can share a single port. However, one Fibre Channel host bus adapter (HBA) is still required for each server.

- Network-attached storage (NAS) gateway

  Using an NAS gateway, you need only provision Fibre Channel ports for the gateway device, rather than for each server. Also, no Fibre Channel HBAs are required for the servers. Therefore, the primary costs are in the cost of the gateway itself, the cost of upgrading your Ethernet network to handle the increased traffic, and establishing a virtual LAN (VLAN) for this new traffic.

Note: Some block I/O applications cannot be accessed effectively through an NAS gateway.

- iSCSI

  iSCSI can be thought of as an IP SAN. Using iSCSI, you do not need to provision Fibre Channel ports for each server. Also, no Fibre Channel HBAs are required, but iSCSI imposes a processing usage on each server. In some cases, a high performance Ethernet card with a TCP/IP offload engine (TOE) function might be advisable. Again, you need to look at the costs associated with upgrading your Ethernet network, such as setting up a VLAN. Because iSCSI delivers block I/O, applications compatibility is not an issue.

The m-type family routers have the capability for iSCSI. The SAN04M-R with standard firmware version can accommodate up to 12 iSCSI initiators. If this is not enough, clients can order the optional firmware version, which enables 50 iSCSI initiators. The SAN16M-R can accommodate up to 50 iSCSI initiators per port and up to 200 per single router with the advanced firmware version.

Figure 14-3 shows how you can use iSCSI to provision disk storage to non-critical servers. iSCSI can also be used for critical servers. However, in general, you can expect lower performance and lower reliability on an Ethernet network than on a Fibre Channel network. Use iSCSI multipathing when using iSCSI for critical servers.

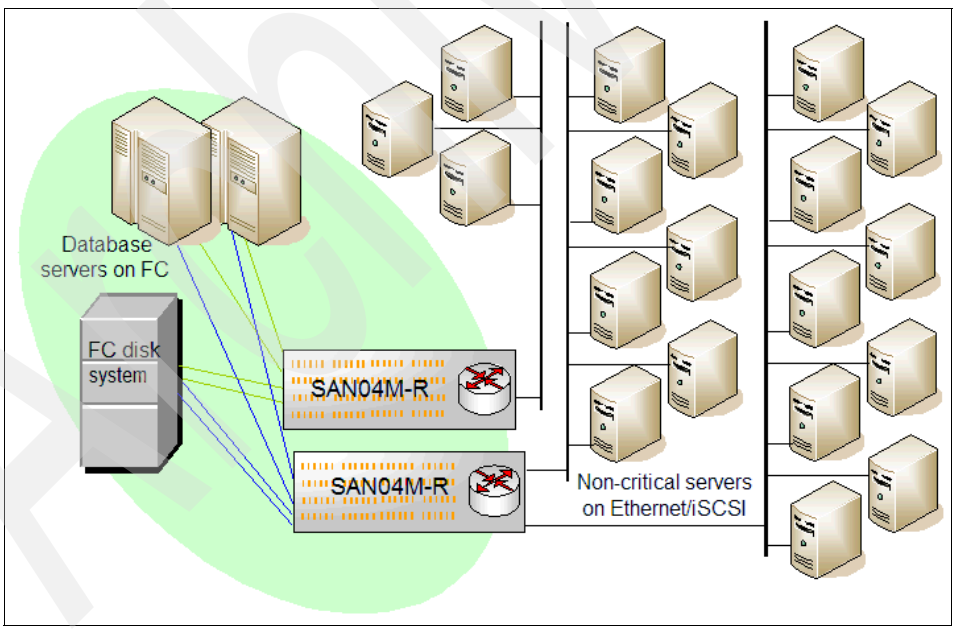


Figure 14-3 Using iSCSI routing to provision disk storage to non-critical servers

14.4 Isolation and interoperability using SAN routing

This section shows some examples of ways in which you might use SAN routing.

14.4.1 Separating production from development

In addition to your main production environment, you might have a development or test environment that is subject to frequent reconfiguration and rebooting. Or, you might have an environment that is subject to a higher risk of failure due to less rigorous change controls. You need to isolate this from your production systems, but test systems also need occasional access to data that is stored on the production disk systems.

Figure 14-4 shows how to achieve this fabric isolation using two SAN routers.

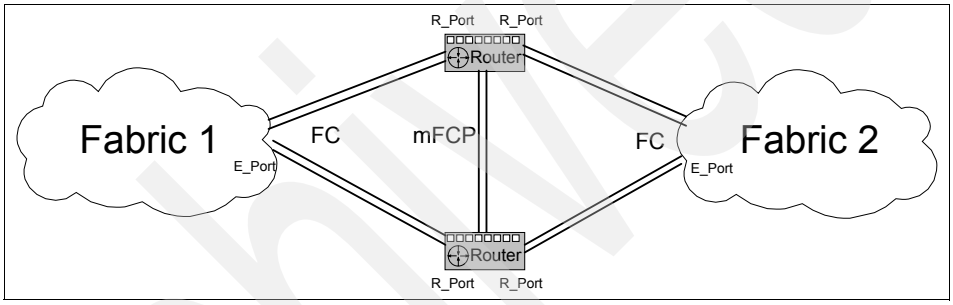


Figure 14-4 Separating fabrics using two routers

14.4.2 Separating corporate subsidiaries

A corporation can also choose to isolate subsidiary companies from each other while providing some shared services such as centralized backup. A shared-services provider can also use this approach to host multiple clients on the same physical infrastructure.

Figure 14-5 shows an example where separate subsidiaries share a physical infrastructure and allow shared access to the backup infrastructure.

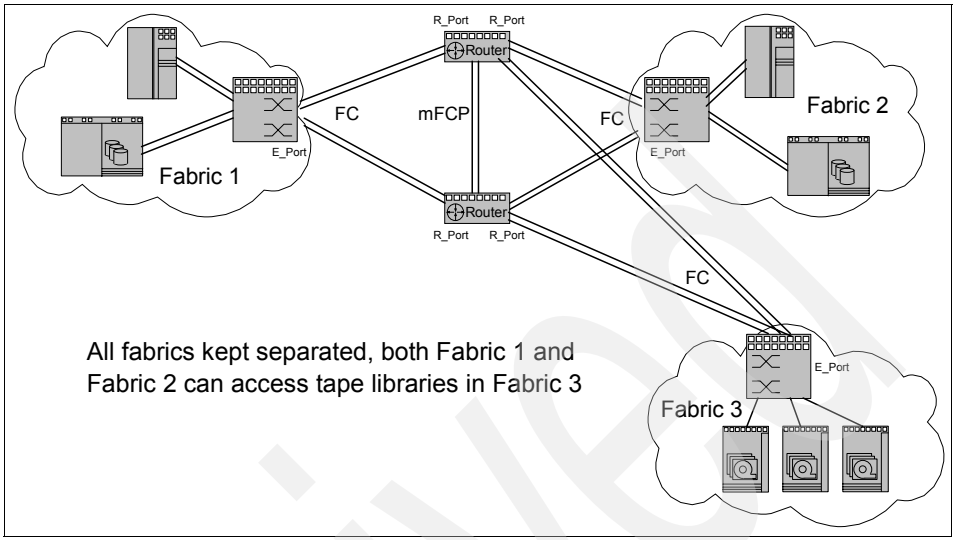


Figure 14-5 Using m-type SAN routers to isolate subsidiaries

14.4.3 Isolation of multivendor switches and modes

You can have Fibre Channel switches from multiple vendors that each require different mode settings and behave slightly differently in the network. You might want to incorporate them into your network, but keep them isolated either for departmental reasons or to keep the different modes of operation separate from each other. SAN routers give the architect the confidence to combine switches from other vendors into the network, knowing that each SAN island has its own separate fabric services.

The Brocade fabric in this case includes initiator devices attached to the Brocade switch, the Brocade switch itself, and an inter-switch link (ISL) to the SAN router. The router provides and manages the routing between the Brocade and McDATA fabrics.

Figure 14-6 shows how switches from Brocade and McDATA can be incorporated into the network and yet be isolated when using SAN routers to interconnect them.

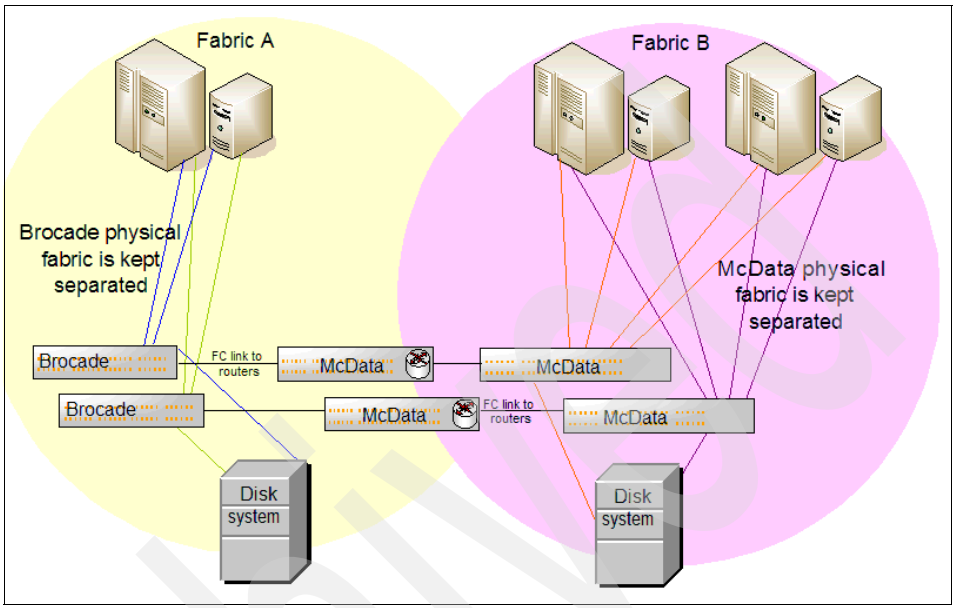


Figure 14-6 Using SAN routers to provide multivendor isolation and integration

14.5 Migrating existing storage to a new environment

You can use SAN routers to migrate data from your existing storage into a new environment. For example, assume that you have an HP XP512 storage system shared among AIX 5L, HP-UX, and Windows servers. For historical reasons, each server platform has its own SAN fabrics and connections to the XP512. Each SAN fabric consists of a single 16-port, 1 Gbps switch.

Figure 14-7 shows the initial environment. For the sake of clarity, we show only some of the servers.

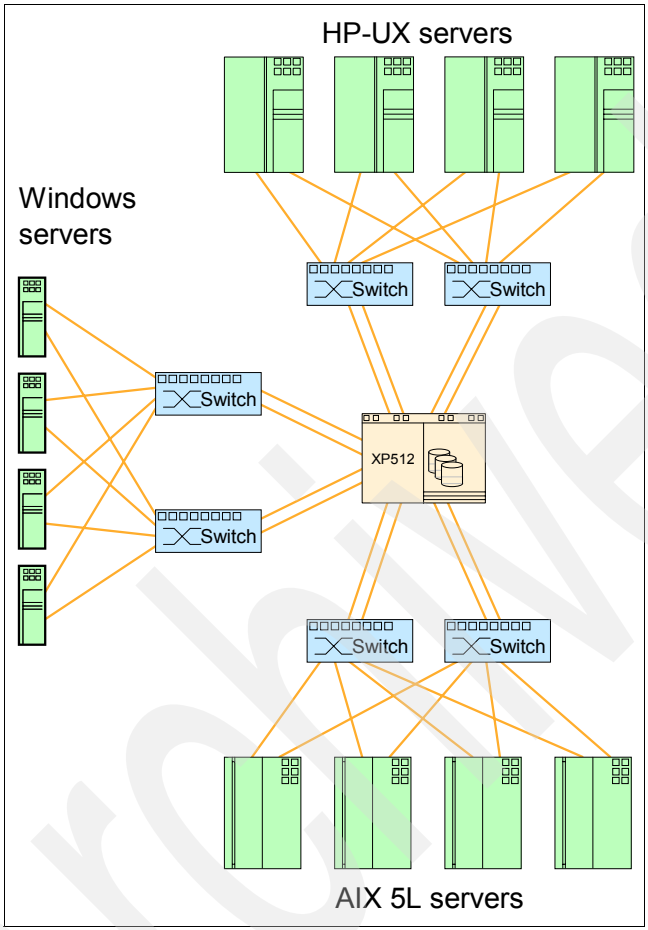


Figure 14-7 Initial storage environment

The newly implemented solution consists of the following components:

- IBM TotalStorage DS8100 disk subsystem
- Two IBM TotalStorage SAN256M Directors with 64 ports each
- Two IBM TotalStorage SAN16M-R routers

Figure 14-8 shows the interim solution.

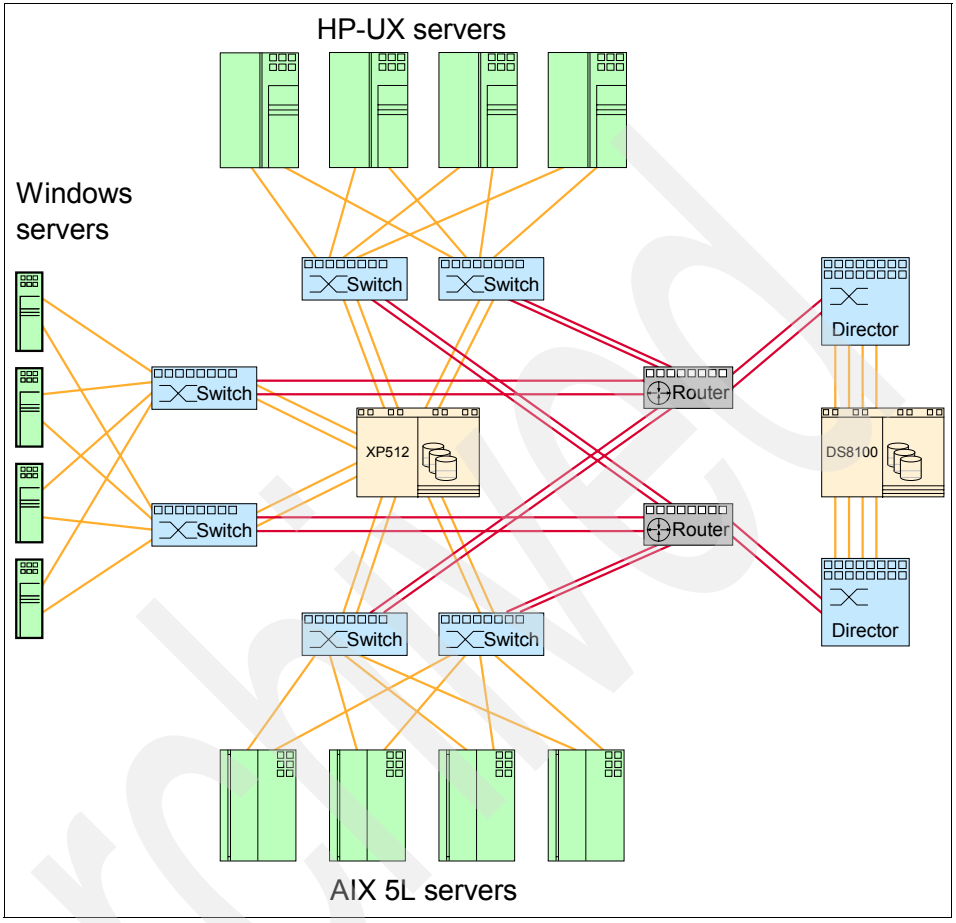


Figure 14-8 Migration interim storage environment

When the new environment is installed, we are ready to perform the migration. The migration typically includes the following steps:

1. Create mSANs for the server to access the DS8100.
2. Install the IBM Subsystem Device Driver (SDD) package and any other DS8100-specific software on the server.
3. Allocate new storage in the DS8100 to the servers.

4. Migrate all server data from the old storage to the new storage using the operating system based tools:
   - Native LVM for AIX 5L
   - PVLinks for HP-UX
   - Veritas Volume Manager for Windows
5. Create zones to allow the server to access the storage from the new SAN fabrics.
6. Disconnect the server from the old switches and move it to the new directors.
7. Delete the mSAN zones created in step 1.
8. Disconnect the routers and old equipment.

After completing the migration, we do not have any servers connected to the old switches and the XP512 is idle. At this time, we can remove the old storage hardware from the environment. The IBM TotalStorage SAN16M-R routers are also freed and can be used for other purposes, such as a SAN extension over iFCP.

Figure 14-9 shows the final solution after the migration.

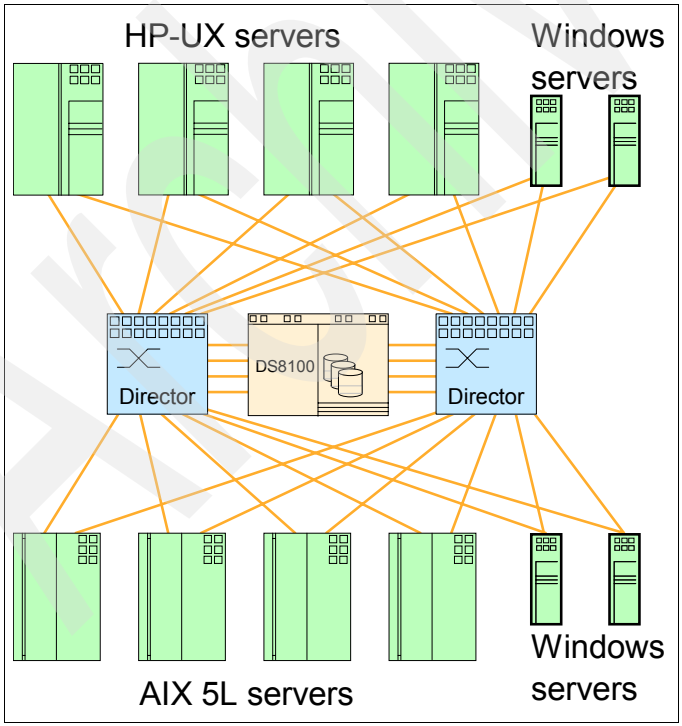


Figure 14-9 How the environment looks after the migration

15

IBM TotalStorage m-type family best practices

In this chapter, we discuss various items that you need to plan before you introduce a storage area network (SAN) router into your SAN environment. This chapter does not present a comprehensive list. However, we think it offers a good starting point. Because each environment is different, there is no one list that provides answers to every question.

15.1 The planning checklist

When planning to introduce a SAN router to your environment, consider the following areas:

- Determine the amount of data you want to transfer among your interconnected fabrics. How much data of the total amount will change over a specific period of time? You need this information to size your link properly.
- Determine the wide area network (WAN) link type, its quality, and the number of independent paths of that link. Is it a shared or dedicated link? If it is shared, what other applications use this link? The best practice is to use dedicated links where possible; a shared link with a quality of service (QoS) mechanism is considered to be second best practice. We do not recommend best-effort shared links d for storage traffic.

  What is the maximum transmission unit (MTU) for your WAN? Do you use IP Security (IPSec)? Does your network support MTU auto-discovery? IPSec usually changes the MTU to a lower value, which can result in packet loss when your router's MTU setting is higher than that.

  What is the average number of dropped packets? If this value is too large, the performance of the TCP-based communication is usually poor. The guideline is that the network is considered to be well-performing if there is an average packet drop of one packet per million.
- If multiple WAN links are available, what are the differences among them? These include the bandwidth, latency, reliability, MTU size, and number of hops.
- Determine the type of security being used on your links. You need to open ports on your firewall, for example. We list all the ports an m-type router uses in 15.5.1, "Ports used by m-type SAN routers" on page 429.
- Do your network components support 802.3x flow control? The best practice is to use flow control to avoid packet drop due to buffer overflow.
- How many remote locations do you have? How many of them are within the mSAN area and how many are internetworked SANs (iSANs)? Do you have enough ports in your fabrics and IP infrastructure to accommodate routers and guarantee the aggregate bandwidth?
- Do you need to implement bandwidth management at the zone level? You can set up a zone with a minimum and maximum bandwidth that zone will use. This might be essential, especially on shared lines.
- What level of high availability is required? Will you need a backup WAN link? What will be the impact on the production systems if your primary link fails for a certain amount of time?

Decide how zones will be implemented. You can have your router append to your fabric's switches, or the zones can reside on routers only. If the latter is implemented, then every time you change a zone on the router level, you also need to change your zones at the switch level.

The best practice is to have your router append zones to fabric switches. All router zones have the SoIP_ prefix, so none of your zones in particular fabrics will be modified.

- Create and maintain a matrix of all firmware and driver levels, operating system versions, and maintenance levels. Consult with the vendors if there is a need for any upgrades prior to the installation.

15.1.1 The installation checklist

Consider the following items prior to a SAN router installation:

- Documentation

  Have your list of IP addresses, port numbers, virtual LANs (VLANs), link types, fiber ports, protocols, equipment type and model, firmware, and driver version list ready and up to date.

- Cabling

  For Gigabit Ethernet ports, you need cables with an LC connector on one end, the SAN router side, and an SC connector on the other, the Ethernet switch side.

- Router management

  Prepare a management workstation and connect it to the network that can reach the management ports of SAN routers. Reserve ports and IP addresses for connecting management ports. Set the router's management ports to Auto Negotiation (best practice). Make sure that your firewalls (if there are any between SAN routers and the management application) allow Simple Network Management Protocol (SNMP) and Hypertext Transfer Protocol (HTTP) communication.

- Prepare a roll-back plan to return back to your starting point.

15.1.2 Running a pilot solution

Keep in mind that the primary task of your SAN is to serve your production servers. Therefore, it is not a good idea to introduce new components directly into the live environment without prior testing.

Always follow the safe path and run a pilot solution. Your pilot solution should consist of all types of components that you have installed in your production

environment. Develop the test cases and run them in your pilot environment. Try any possible “what if” scenarios. The costs of downtime of a small test environment are incomparable to the costs of a downtime of your production SAN environment.

15.2 Fabric considerations

Before introducing any new component into your fabrics, it is a good practice to check the firmware levels of all the components and check with the product vendors to learn if there are any compatibility issues at given firmware levels. The m-type SAN routers can be interconnected together through mSAN or iSAN links only if they are on the same Enterprise Operating System (E/OS) level. Running different versions of E/OS levels is not supported.

Before upgrading the software on one of the interconnected routers, it is good practice to disable the mSAN, iSAN, or both ports. This will help to avoid undesired communication among the routers while they are not running the same level of code.

Do not use domain IDs FE and FF because the McDATA routers use these addresses for internal routing. Attaching a router to a Fibre Channel switch with these addresses will cause a fabric conflict.

Always make a backup of your router configuration before any router’s software upgrade. By doing so, you can avoid having to re-create your lost zoning configuration and setting vital parameters, such as interoperability mode and domain IDs.

A good practice is not to fix anything that is not broken. You do not need to upgrade your router’s Fabric OS just because there is a new release available. Plan your upgrades carefully and perform them only when you have a good reason to do so.

15.3 Bandwidth and capacity planning

One of the main impacts on the final bandwidth requirements and capacity of the link is the type of application you will run across your interconnected fabrics.

Generally speaking, the highest and the most constant bandwidth requirements usually are ones that use synchronous data replication. Even higher bandwidth requirements can include a remote fabric tape library access. However in this case, the need for bandwidth usually occurs during the backup windows and tape

management and vaulting tasks of your backup system. In case of asynchronous replication, the bandwidth requirements might be substantially lower.

15.3.1 Aspects that influence communication performance

This section discusses various aspects that have a significant influence on the link performance.

Link bandwidth

The link bandwidth is the most obvious factor that affects performance. It is also one of the key metrics used when provisioning the link. In storage environments, the link bandwidth used should always be the guaranteed bandwidth from the service provider.

If the guaranteed bandwidth is anything less than a full Gigabit Ethernet, it needs to be configured into the routers at both ends of the link to avoid overrunning the link. We recommend that you set the maximum allowed speed of the inter-router link (IRL) ports between 77% and 96% of the guaranteed bandwidth of the link at both ends of the link, depending on your link quality.

Latency and round-trip time

Link latency is a metric of the round-trip time (RTT) it takes for a packet to cross the link. RTT is the time it takes for a datagram to be received and returned to the sender over a network. The key factors contributing to the link latency include:

- Distance
- Router and firewall latencies
- Time of frame in transit
- Network congestion

Distance

The speed of light in optical fiber is approximately 208,000 km/s. Therefore, the delay caused by a FC connection is approximately 4.8 μs/km. To calculate the round trip latency, we have to count this delay both ways.

For example, for a 100 km link, the round trip latency is approximately:

100 km x 4.8 μs/km x 2 = 960 μs

Similarly, for a 1000 km link, the round trip latency is 9600 μs, or 9.6 ms.

Router and firewall latencies

Any delay caused by routers and firewalls along the network connection needs to be added to the total latency. The latency varies a lot depending on the routers or

firewalls and the traffic load. It can range from a few microseconds to several milliseconds.

You also need to remember that the traffic generally passes through the same routers both ways. Therefore, for round-trip latency, you need to count the one-way latency twice.

If you are purchasing the routers or firewalls yourself, we recommend that you include the latency of a particular product in the criteria you use to choose the products. If you are provisioning the link from a service provider, we recommend that you include at least the maximum total round-trip latency of the link in the service level agreement (SLA).

Time of frame in transit

The time of frame in transit is the actual time that it takes for a given frame to pass through the slowest point of the link. It therefore depends on both frame size and link speed.

The maximum size of the payload in a Fibre Channel frame is 2112 bytes. The Fibre Channel headers add 36 bytes to this, giving a total Fibre Channel frame size of 2148 bytes. When transferring data, Fibre Channel frames at or near the full size are used.

Jumbo frames: Jumbo frames are a technique for maximizing the throughput of Ethernet networks by increasing the frame size from the default 1518 bytes up to 9000 bytes. To gain the maximum benefit, all devices in the network have to support the frame size. Non-jumbo routers break the frames down to 1518-byte frames.

Note: The m-type router supports a maximum IP packet size of 4096 bytes.

If we assume that we are using jumbo frames in the Ethernet, the complete Fibre Channel frame can be sent within one Ethernet packet. The TCP and IP headers and the Ethernet medium access control (MAC) add a minimum of 54 bytes to the size of the frame, giving us a total Ethernet packet size of 2202 bytes, or 17616 bits. The SAN router currently uses half-sized jumbo frames.

For smaller frames, such as the Fibre Channel acknowledgement frames, the time in transit is much shorter. The minimum possible Fibre Channel frame size is one with no payload. With Internet Fibre Channel Protocol (iFCP), the minimum size of a packet with only the headers is 96 bytes or 768 bits.

Table 15-1 on page 425 provides details about the transmission times of an iFCP packet over some common WAN link speeds.

Table 15-1 FCIP packet transmission times over different WAN links

Link type	Link speed	Large packet	Small packet
Gigabit Ethernet	1250 Mbps	14 µs	0.6 µs
OC-12	622.08 Mbps	28 µs	1.2 µs
OC-3	155.52 Mbps	113 µs	4.7 µs
T3	44.736 Mbps	394 µs	16.5 µs
E1	2.048 Mbps	8600 µs	359 µs
T1	1.544 Mbps	11,400 µs	477 µs

If we cannot use jumbo frames, each large Fibre Channel frame needs to be divided into two Ethernet packets. This doubles the amount of TCP, IP, and Ethernet MAC usage for the data transfer.

Normally, each Fibre Channel operation transfers data in only one direction. The frames going in the other direction are close to the minimum size.

Congestion and dropped packets

An example of how congestion can occur is when more ingress ports are communicating to a single egress port. The buffers fill faster than they can drain out. The result is packet drop, which leads to retransmission.

15.3.2 Throughput and efficiency

Table 15-2 shows link speeds, approximate usage, efficiency, and throughputs (with no usage) for most common network link types. We use these parameters in our sizing examples in 15.3.3, “The amount of data and link sizing” on page 426.

Table 15-2 Example of throughput and efficiency of different network links

Link type	Link speed	Usage	Efficiency	Throughput
100baseT Ethernet	125 Mbps	6.30%	93.70%	11.71 MBps
Gigabit Ethernet	1250 Mbps	6.30%	93.70%	117.13 MBps
OC-12 SONET	622.08 Mbps	8.92%	91.08%	70.82 MBps
OC-3 SONET	155.52 Mbps	8.92%	91.08%	14.164 MBps
T3	44.736 Mbps	5.05%	94.95%	5.31 MBps
T1	1.544 Mbps	4.42%	95.58%	0.184 MBps

15.3.3 The amount of data and link sizing

You can use different sizing methods and estimations depending on your environment and the application.

Link sizing for synchronous data replication

For synchronous data replication, you need to find the peak in write operations over a period of time. The theoretical maximum for a peak in write operations is influenced by many factors, such as the central processing unit (CPU), application efficiency, networks, and storage devices.

For example, let us assume that the set of logical unit numbers (LUNs) that you need to replicate has a peak of 110 MBps during the system's busy hours. What speed of the WAN link is required for data replication?

The link speed of Gigabit Ethernet is 1000000000 bits per second. The link efficiency, under ideal conditions, is 93.7%. We can achieve a maximum throughput of 937 Mbps of storage data only, which is 117.13 MBps (937/8=117.125). We need only 110 MBps during peak periods for the data replication, so Gigabit Ethernet is the answer in our example.

However, you need to monitor the utilization of your link and any possible increase in the write peaks as the environment grows over time.

Our calculation is simplified and does not take into consideration further link optimizations, such as compression or fast write. Compression usually leads to less usage, so greater efficiency is achieved. However, using compression for synchronous replication cannot be considered as a best practice, because it introduces additional data processing, which introduces more latency. Latency then causes longer response times.

Here are other assumptions.

- There is a dedicated network (no shared bandwidth).
- We do not use fast write.
- Latency does not lead to congestion and dropped packets.
- We omit retransmits.

Link sizing for asynchronous data replication

When planning the link size for asynchronous data replication, we do not need bandwidth to accommodate data replication during peak hours. Instead, we need to find the total amount of data to be replicated and the amount of data to be changed in a period of time. While the former quantity is usually easy to determine, the latter is usually estimated as a best guess based on experience.

A good start is to assume that 20% of the total amount of data changes during the day. Your application's specific tools or operating system utilities can help you estimate this quantity.

For example, assume that we have 8 TB of data, with 20% changed over a 24-hour period. That is an average 69 GB change in one hour (20% from 8 TB = 1639; 1639 / 24 = 68.3 GB per hour). We can calculate it with a conservative two to one (2:1) compression ratio, which is near 35 GB per hour (68.3 / 2 = 34.15).

Let us calculate the required throughput in MBps. Because 35 GB equals 35840 MB, 35 GB per hour translates into 35840 MB per hour, and this is equal to 9.96 MBps. Therefore, the throughput equal to that of 100BaseT Ethernet suffices in this case.

Link sizing for tape data backup

If planning for a link for remote tape communication, consider the total amount of data and the time period in which the data should be transferred from your hosts or backup servers to tapes. Another sizing factor might be the number of drives in the library. However, many of the new generation tape drives are capable of operating at data transfer rates able to consume the bandwidth of 1 Gbps Ethernet. For example, the IBM 3592 is capable of achieving a theoretical throughput of 120 MBps with compression, 40 MBps without compression. The conclusion is that you cannot feed, for example, two IBM 3592 drives with a 1 Gbps Ethernet link because this link gives you only a 117 MBps rate maximum. You need 240 MBps.

Let us consider an example where we have 10 TB of data to be transferred every weekend from primary data pools of our backup server to the remote tape library. The time for the data to be transferred cannot exceed 30 hours. We can calculate that:

10485760 MB / 108000 seconds (30 hours) = 97.1 MBps

We need a Gigabit Ethernet link to accommodate this amount of data within the given period of time.

15.3.4 Fast write and IBM products

Do not use fast write with IBM data replication products, such as Metro Mirror, Global Mirror, and Global Copy. These already include write optimization in the product itself, so they would not benefit from McDATA's Fast Write on the router level.

IBM data replication products do not send round-trip messages (write command and transfer ready) to start sending data. Instead, they start to send the data immediately to the remote subsystem. This is similar to what the McDATA Fast Write does. Therefore, the best practice is not to use fast write with IBM Metro Mirror, Global Mirror, and Global Copy.

15.4 Planning for availability

This section describes some basic considerations to achieve a higher level of availability in your interconnected SANs using m-type routers. When planning your interconnected SAN for availability, always determine the level of availability that you really need. Calculate how much risk you can tolerate and what would be the impact on your production systems if you lost the fabric interconnectivity for any given period of time.

15.4.1 Hardware limitations

Routers, in general, are designed as a switch-class product, not a director-class product. Therefore, a situation might occur when the router must be rebooted or replaced.

The m-type routers are not chassis-based, and the main electronic components are not redundant, except the power supplies and fans. In case of a power supply or fan failure, the router will remain operational, but a service window must be scheduled to replace it. In the case of any other component failure, the entire router must be replaced.

15.4.2 Multiple paths and path failover on a router level

To increase availability, we recommend that you use multiple paths to your fabrics and Metro Fibre Channel Protocol (mFCP) paths to a single router. One router can have multiple iFCP paths. However, only one of them can be active at a time.

Note: mFCP is only supported in firmware release 4.6.2 and the upcoming 5.0 release.

The backup iFCP path can be activated automatically by the router itself if the primary path fails. This is done by the heartbeat between two routers connected by an iFCP link. If the heartbeat is lost for a particular period of time (10 seconds), the backup path is activated. You can bring the primary path back online either manually or let the router do this automatically.

With the SAN16M-R router, you can design your mSAN for path failover. That means you can configure two routers in your mSAN, interconnect them with an mFCP link, and connect each fabric's switch with two alternate paths, each to a different SAN16M-R router.

Figure 15-1 shows the path failover scenario at a router level. If one of the routers fails, all the traffic is transferred through the remaining router.

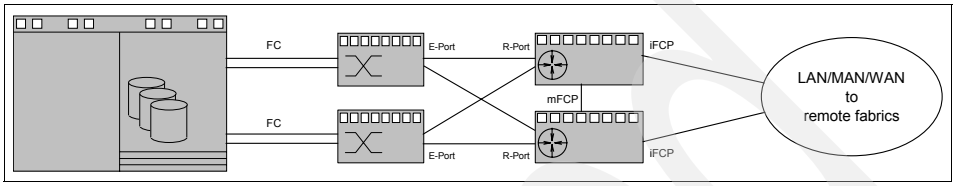


Figure 15-1 Path failover on router level

It is a good practice to keep local storage traffic and traffic for replication separated, or to dedicate one pair of storage ports for data replication and the other pair for local traffic.

15.4.3 Fault isolation

If the WAN link in Figure 15-1 fails, fabric reconfiguration does not occur, because fabrics are not merged. Each fabric has its own Storage Name Server (SNS) database, domain space, and principal switch. No F-Class traffic is sent beyond an R_Port, and only port format registered state change notifications (RSCNs) are propagated to remote fabrics.

15.5 Planning for security

With the introduction of SAN routers, the storage traffic leaves the known boundaries of Fibre Channel and traverses through IP networks to remote locations. This brings new security challenges and items to plan and consider before you can interconnect SANs.

15.5.1 Ports used by m-type SAN routers

For routers to communicate, many TCP and User Datagram Protocol (UDP) ports must be enabled as appropriate. In addition to this, some other ports must be enabled to manage routers. Table 15-3 on page 430 lists all the ports that are being used by m-type router products.

To use virtual private network (VPN) with IPSec, we highly recommend that you secure communication among your m-type routers.

Table 15-3 Ports used by m-type routers

Description	Type	Port	Protocol
Inter-switch control	TCP	37121	iFCP
Redundancy control	TCP	37122	iFCP
Data	TCP	3420	iFCP
Data	TCP	3260	iSCSI
MTU discovery	UDP	7	UDP Echo
SNMP traps	UDP	162	SNMPv1
Telnet	TCP	23	Telnet
FTP	TCP	20	FTP control
FTP	TCP	21	FTP data
TFTP	UDP	-	TFTP
Java applet	TCP	80	HTTP
Applet switch communication	UDP	161	SNMPv1
Switch message log communication	UDP	37009	
iFCP/mFCP ping results communication	UDP	37010	

15.5.2 Zoning

The Fibre Channel router ports need to be secured using zoning. You can only zone at port (pWWN) level. However, pWWN zoning is not supported between mSANs that are interconnected by an iFCP connection.

One of the oldest best practices from the early SAN days is to have one initiator and one target in a single zone and not to have more than two ports in a single zone. This was a necessity in the past, but it is too conservative today. The host bus adapters (HBAs), their firmware, and OS drivers are much more mature than they were couple of years ago.

There is no single, best recommendation on how to use zoning in your environment. In general, we can organize the zones as follows:

- One initiator and one target per zone

  This is not suitable for midrange and large environments. Imagine having to administer your zones in an environment of 50 servers, each with four HBAs and two storage subsystems, each with six HBAs and two tape libraries with eight drives each. It would not be manageable to keep with this zoning approach.

- One initiator and multiple targets per zone

  If you still want to stay conservative, but do not want to run into zone administration usage, consider this rule for your zone configuration.

- Multiple initiators and multiple targets per zone

  If your level of conservatism is not too high, and if you can take a calculated risk of possible “chatter” among your initiator HBAs within the zone, consider this approach.

Zoning at the mSAN or iSAN level should only be done for devices that you really need to have access to in remote fabrics. If there are only certain periods of time when you need to access a particular device in another mSAN, for example, once a month, we recommend that you export (zone) this device only for that particular period of time that you need to have it exported.

15.6 Scalability and limitations

This section describes some best practices and limitations that you need to consider when planning for interconnecting fabrics using SAN routers. Many of the numerical figures that we mention here are usually the technical limits of particular devices or practical recommendations. However, most of the SAN environments in the real world, at the time of writing, are within these limits.

Note: The design options related to mFCP are only available in firmware version 4.6.2, or the upcoming 5.0 release.

There can be a:

- Maximum number of two SAN16M-R routers in an mSAN

  There can be only one SAN04M-R in an mSAN, because it does not support an mFCP link. You cannot go beyond this limit.

- Minimum of two mFCP links

  The mFCP links are configured in pairs. If you only activate one, you will receive an error message stating that mFCP will not be usable until its pair is also active. The pairs are 1-2, 3-4, 5-6, 7-8, 9-10, and 11-12.

- Maximum number of four mFCP connections between the SAN16M-R in an mSAN

  We recommend that you use at least two mFCP connections (as per the previous point) for availability, regardless of performance requirements. You cannot go beyond this limit.

- Maximum number of four inter-switch links (ISLs) to each fabric from all mSAN routers

  Because you can have a maximum of two SAN16M-R routers, we recommend that you have two ISLs from each router to the fabric. In a case where you only have one SAN16M-R router, it can have up to four ISLs to the fabric. A good practice is to always have at least two ISLs from one or both routers in an mSAN to each fabric.

- Maximum number of six fabrics connected in an mSAN

  This applies whether you have one or two routers in the mSAN. The number of connected routers in an mSAN does not affect this limit.

- Maximum number of 48 switches in an mSAN

  The number of connected routers in an mSAN does not affect this limit. Because you can have up to six fabrics, the number of switches in each fabric can vary, but it cannot exceed the total number of 48.

- Recommended average number of 12 switches per single fabric

  You can have more than 12 switches in one fabric, but fewer than 12 switches in another fabric to keep the average around 12. For example, you can have 16 switches in one fabric and 8 in another fabric (16 + 8 = 24, 24 / 2 = 12). You also need to keep in mind that there is a maximum number of 48 switches in an mSAN. You cannot exceed this limit.

- Maximum number of 1024 connected devices in a single fabric

  Usually, this is the limit of the SNS database size in the fabric itself.

- Maximum number of 504 devices imported from one single fabric

  The total number of imported devices from all fabrics per single SAN16M-R is 512. This is the limit of the mSNS database.

In addition to these limits, Table 15-4 on page 433 shows the tested limits at the time of writing.

Table 15-4 Tested scalability limits at the time of writing

Metric	SAN04M-R	SAN16M-R
Maximum number of Fibre Channel fabrics	2	6
Maximum number of switches (domains) per fabric	12	16
Total number of Fibre Channel switches in all interconnected fabrics in an mSAN	24	24
Combined maximum imported Fibre Channel devices from all fabrics with an mSAN	64	256
Maximum Fibre Channel devices in a connected fabric	1024	1024
Maximum number of R_Ports connected to a single fabric	2	4
Maximum number of SAN16M-R zones (recommended and tested, possible 512 and 1024, respectively)	128	256
Maximum number of loop devices off a single router FL_Port	8	32
Maximum number of loop devices attached to a single router	16	384
Maximum iFCP plus iSCSI sessions on a single Gigabit Ethernet-TCP port (initiator-target pairs)	64	64
Maximum iFCP plus iSCSI sessions (initiator-target pairs) per router	64	256
Maximum number of iSCSI initiators per router port	50	50
Maximum number of iSCSI initiators per Eclipse SAN router	50	200
Maximum number of iSCSI initiators in one mSAN (two routers, mFCP)	N/A	200
Maximum number of iSCSI sessions in one mSAN (two routers, mFCP)	N/A	256
Maximum iFCP point-to-multipoint connections per router port (one “site” to many “sites”)	8	8
Maximum iFCP point-to-multipoint connections (one “site” to many “sites”) per SAN router	16	32
Maximum number of Eclipse SAN routers in an mSAN	1	2
Maximum mFCP connections between two Eclipse SAN routers	N/A	4

16

IBM TotalStorage m-type family real-life routing solutions

This chapter describes the details of some real-life solutions that were implemented with the IBM TotalStorage m-type family routing products. We discuss the following solutions:

- Backup consolidation
- Migration to new storage environment
- Long-distance disaster recovery over IP

Important: The solutions and sizing estimates that we discuss or make in this chapter are unique. Make no assumptions that they will be supported in, or apply to, each environment. We recommend that you engage IBM to discuss any proposal.

16.1 Backup consolidation

In this scenario, we present a solution to consolidate the local area network (LAN)-free tape backups from two storage area network (SAN) fabrics.

16.1.1 Client environment and requirements

The client has two existing SAN fabrics and is currently using ArcServe software to back up the Microsoft Windows servers in the SAN fabrics to a tape. The client also has several application servers that do not have SAN attachment. Figure 16-1 shows the client's environment.

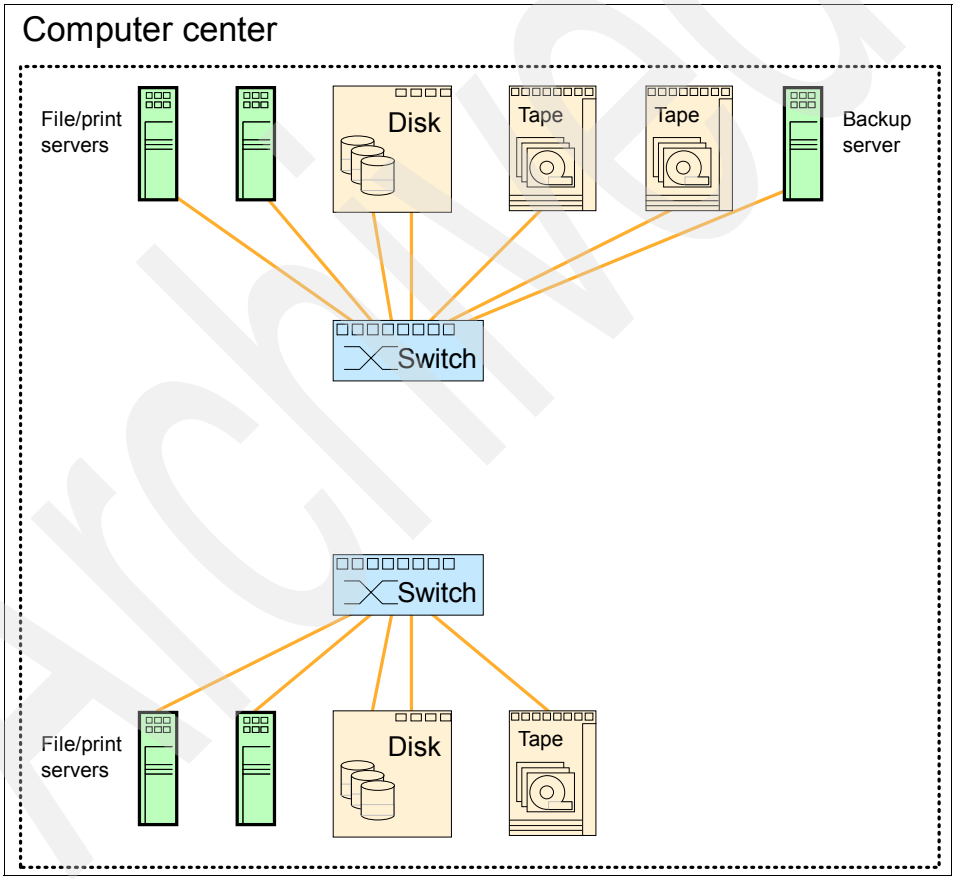


Figure 16-1 Current backup environment

The client has the following requirements for the new solution:

- Consolidate the tape backups to a single Tivoli Storage Manager environment.
- Provide for LAN-free backups from both current SAN fabrics.
- Implement the new backup system to a location separate from the computer center.
- Leverage the existing investment to SAN hardware.

In the first SAN fabric, the client currently has 160 GB of disk space, which is projected to grow to 630 GB in the near future. In the second SAN fabric, the client has 100 GB of disk space.

16.1.2 The solution

Our solution has the following new components:

- IBM System p server for Tivoli Storage Manager
- IBM 3583-L72 tape library with four Fibre Channel drives
- IBM TotalStorage SAN32M-2 switch for the backup environment
- IBM TotalStorage SAN16M-R router

Figure 16-2 on page 438 shows the new backup environment.

We locate the router in the computer center to minimize the need of fiber connections between the computer center and the backup site. All other components are located in a single rack at the backup site.

We connect each of the current fabrics and the new backup switch to the router with two inter-fabric links (IFLs) for redundancy. The client provides the two long wave fiber connections required between the computer center and the backup site.

The Tivoli Storage Manager server will use its internal disks for both Tivoli Storage Manager databases and disk storage pools. Therefore, it does not need any access to the existing client SAN fabrics. The tape drives are divided evenly to the two Fibre Channel adapters in the Tivoli Storage Manager server.

We create a separate mSAN zone for each server in any SAN fabric that needs to have access to the tape drives. The mSAN zone will contain the worldwide name (WWN) of the host bus adapter (HBA) of the server and the WWNs of all the tape drives.

Because our new environment is only used for daily backups, it does not have as high availability requirements as SAN fabrics used for disk access. Therefore, it is adequate to have a single backup switch and a single router in the solution.

The application servers that are not connected to any SAN fabric are backed up to the Tivoli Storage Manager server over a LAN connection.

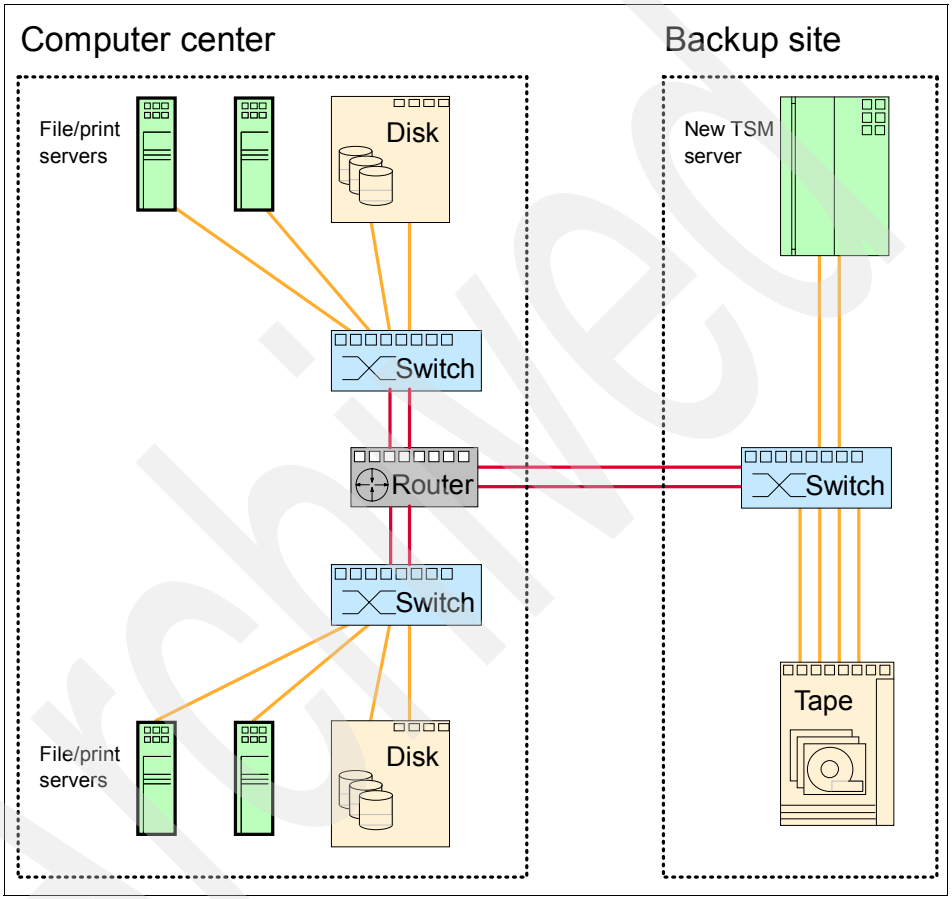


Figure 16-2 New backup environment

16.1.3 Failure scenarios

This section describes how the failure of different components affects the operation of our solution:

- Power failure

  The Tivoli Storage Manager server, the tape library, and all of the SAN fabric components in the environment have dual redundant power supplies connected to different power circuits. Therefore, a power failure in one circuit does not have any effect on the operation.

- IFL failure

  If an IFL fails, the system remains operational, but the maximum bandwidth available is reduced by 50%.

- Router failure

  If the SAN router fails, it is impossible to run LAN-free backups. In this situation, the Tivoli Storage Manager client automatically uses a LAN-based method for any backups and restores. The Tivoli Storage Manager server and the servers not using LAN-free backups are not affected.

- Backup switch or Tivoli Storage Manager server failure

  The failure of either the backup switch or the Tivoli Storage Manager server prevents any backup and restore activity.

16.2 Migrating to a new storage environment

The following scenario presents a solution to migrate the client's current storage environment to a new environment.

16.2.1 Client environment and requirements

The client has a Hewlett-Packard (HP) XP512 storage system that is shared between AIX 5L, HP-UX, and Windows servers. Due to historical reasons, each server platform has its own SAN fabrics and connections to the XP512.

Each SAN fabric consists of a single 16-port, 1 Gbps switch. Because the lease period of the environment expires within a few months, the client needs a new solution to replace the current environment.

Figure 16-3 shows the initial environment. For clarity, you see only some of the servers.

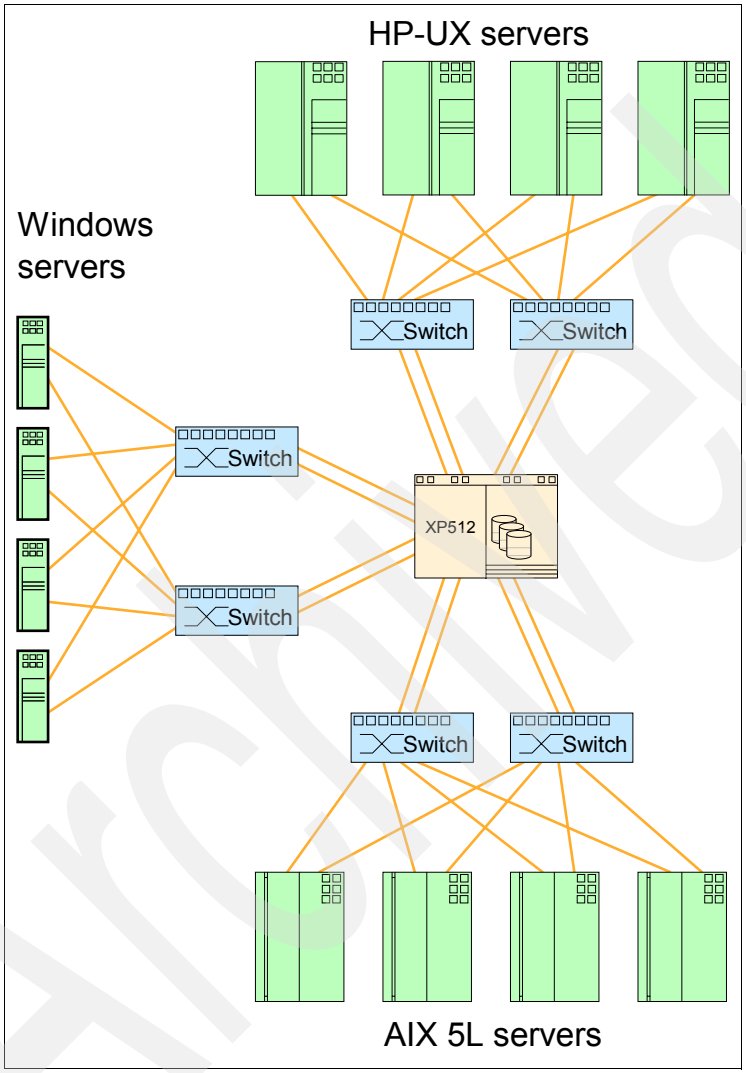


Figure 16-3 Initial storage environment

The client has the following requirements for the new solution:

- New hardware to replace the current disk system and SAN fabric
- Flexibility in allocating ports between different platforms
- Scalability to support future applications
- Minimized amount of downtime of servers due to migration

16.2.2 The solution

Our solution has the following new components:

- IBM TotalStorage DS8100 disk subsystem
- Two IBM TotalStorage SAN140M Directors with 64 ports each
- Two IBM TotalStorage SAN16M-R routers

We install the components of the new storage environment and connect the environment to the old environment with IFLs, as shown in Figure 16-4.

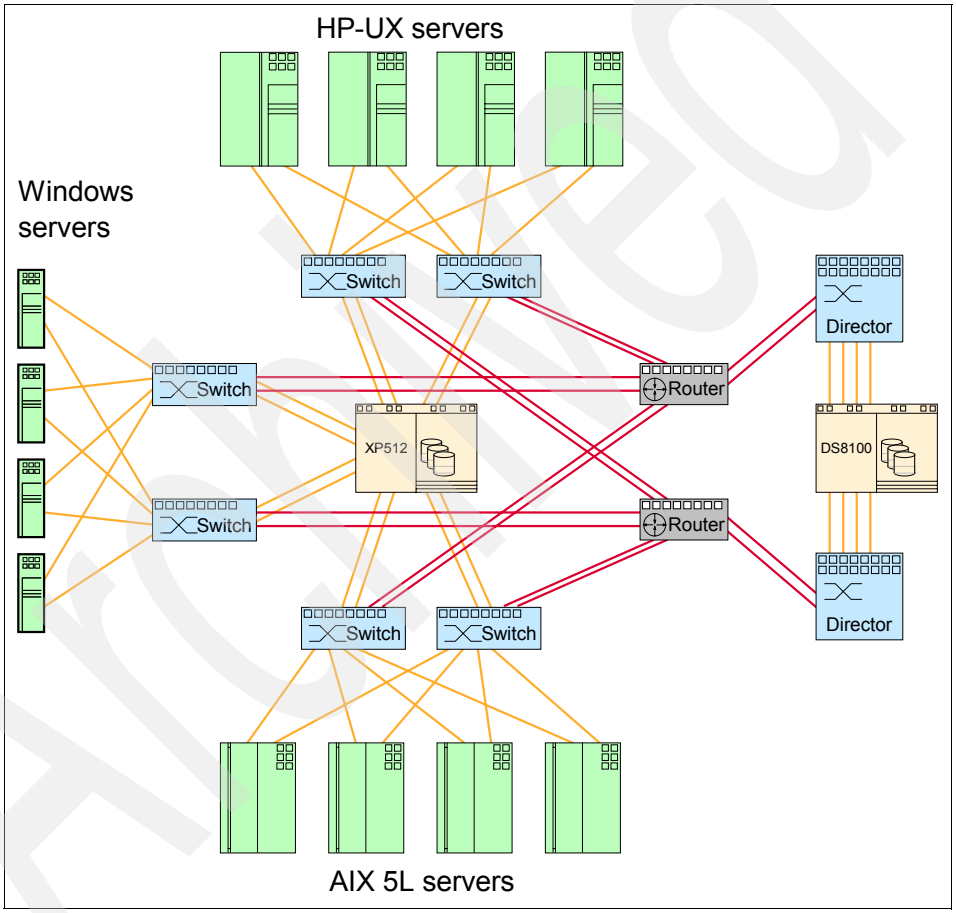


Figure 16-4 Interim environment for migration

In the new environment, all the DS8100 ports are shared among all servers. Because we only migrate a few servers at the same time, using a limited number of IFLs does not cause any performance degradation to the servers.

When the new storage environment is completely installed, we start migrating the servers, one server or a group of servers at a time, using the following procedure:

1. Create mSAN zones to allow the server to access the DS8100.
2. Install the IBM Subsystem Device Driver (SDD) package and any other DS8100-specific software on the server.
3. Allocate new storage in the DS8100 to the servers.
4. Migrate all server data from old storage to new storage using the operating system-based tools:
   - Native LVM for AIX 5L
   - PVLinks for HP-UX
   - Veritas Volume Manager for Windows
5. Create fabric zones to allow the server to access the storage from the new SAN fabrics.
6. Disconnect the server from the old switches and move it to the new directors.
7. Delete the mSAN zones created in step 1.

The only step that requires server downtime in the procedure is step 6. If the new cabling is prepared before, this step should take little time.

After migrating all of the servers, we should have no servers connected to the old switches and the XP512 should be idle. At this time, we can remove the old storage hardware from the environment. The IBM TotalStorage SAN16M-R routers are also freed and can be used for other purposes, such as SAN extension over Internet Fibre Channel Protocol (iFCP).

The following figure shows the final storage environment (Figure 16-5).

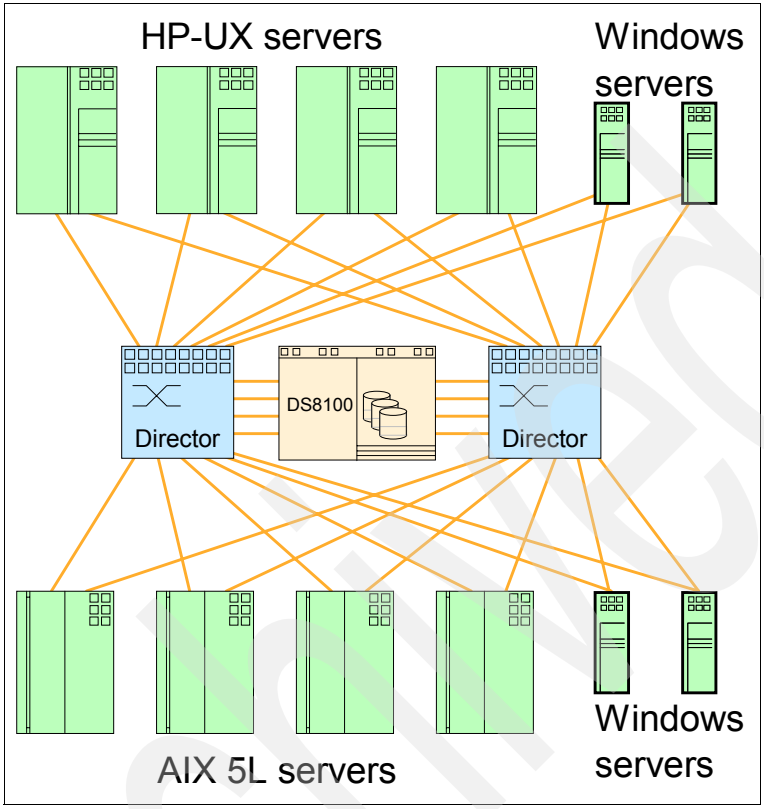


Figure 16-5 Final storage environment

16.3 Long-distance disaster recovery over IP

In this scenario, we present a solution that allows for long-distance disaster recovery (DR) over an IP connection.

16.3.1 Client environment and requirements

The client has three different SAN islands that need to be connected:

- Development SAN at the primary site
- Production SAN at the primary site
- DR SAN at the DR site

The distance between the primary site and the disaster recovery site is 600 km. The amount of data in the production environments is expected to grow to 5 TB within two years, and we expect 3% of the data to change during the peak hour.

The client has the following requirements for the solution:

- Provide asynchronous replication for production data from the primary site to the DR site, with a five minute recovery point objective (RPO) and a five minute recovery time objective (RTO).
- Keep the dual fabrics of each SAN both physically and logically separate.
- Provide access to a point-in-time copy of productive data from the test environment at the development SAN.
- Provide for LAN-free backup from the development network to the tape library in the production network.

The detailed list of the current environment is:

- Production environment at the primary site
  - Dual SAN fabrics, based on IBM TotalStorage SAN140M Directors
  - IBM TotalStorage DS8100 disk subsystem with eight Fibre Channel ports
  - IBM TotalStorage 3584 tape library with six IBM 3592 tape drives
  - Eight IBM System p servers, with dual Fibre Channel adapters
  - Sixteen IBM System x servers, with dual Fibre Channel adapters
- Development environment at the primary site
  - Dual SAN fabrics, based on IBM TotalStorage SAN 32M-2 switches
  - IBM TotalStorage DS6800 disk subsystem with four Fibre Channel ports
  - Eight System p servers, with dual Fibre Channel adapters
  - Sixteen System x servers, with dual Fibre Channel adapters
- Disaster recovery environment at the disaster recovery site
  - Dual SAN fabrics, based on IBM TotalStorage SAN140M Directors
  - IBM TotalStorage DS8100 disk subsystem with eight Fibre Channel ports
  - IBM TotalStorage 3584 tape library with six IBM 3592 tape drives
  - Eight System p servers, with dual Fibre Channel adapters
  - Sixteen System x servers, with dual Fibre Channel adapters

Figure 16-6 shows the environment. For clarity, you see only some of the servers and connections.

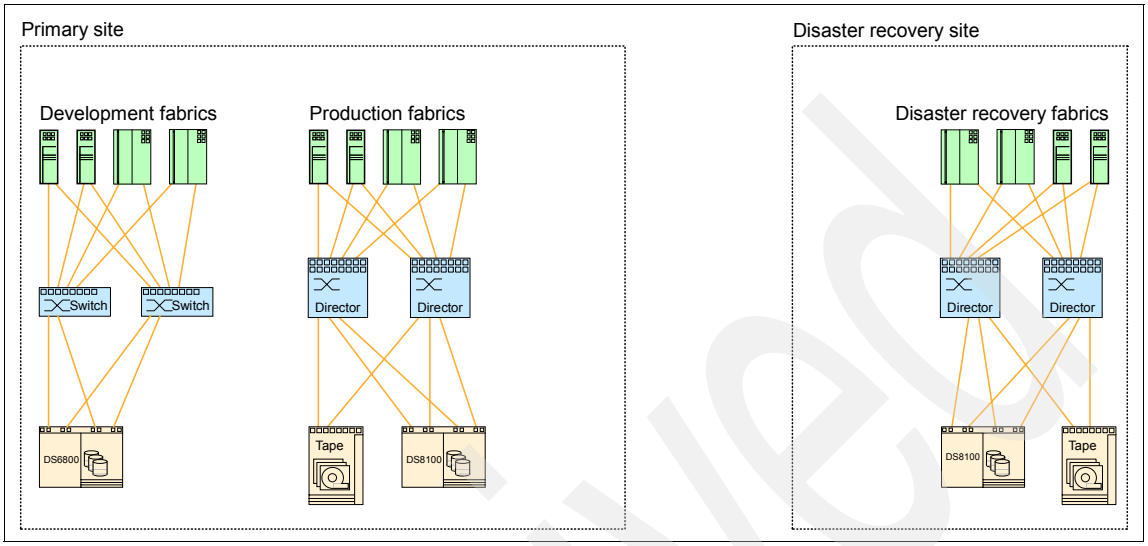


Figure 16-6 Client environment

16.3.2 The solution

Our solution has the following components:

- DS8100 Global Mirroring feature for asynchronous replication
- Four IBM TotalStorage SAN16M-R routers (2027-R16)
- Four IP links between the 2027-R16 routers from the primary site to the disaster recovery site
- IBM eRCMF software to provide automatic failover of both the System p and System x servers

Figure 16-7 shows the complete solution.

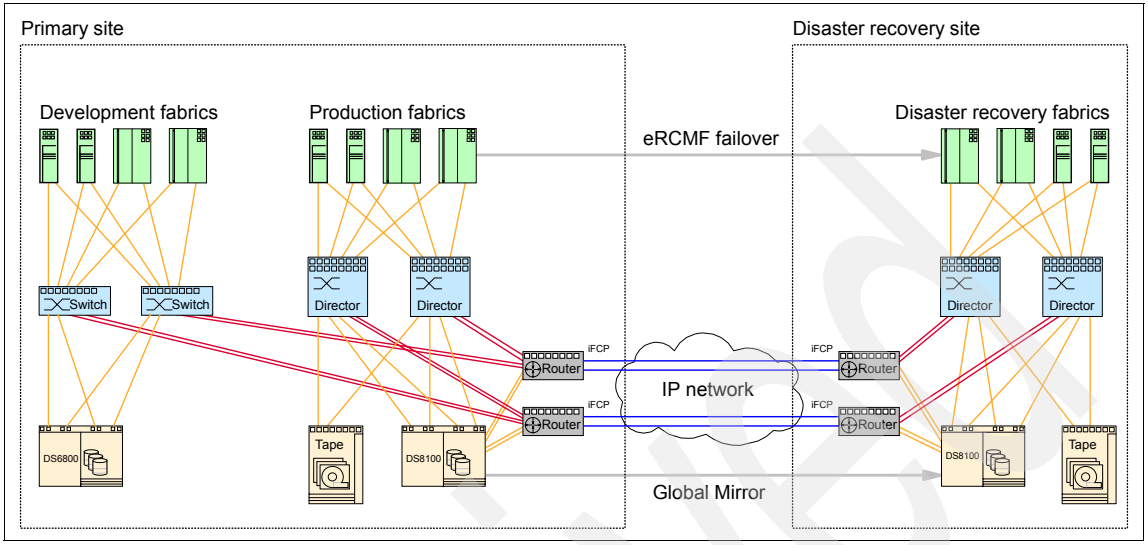


Figure 16-7 Disaster recovery solution

iFCP link sizing

Because we are using the iFCP links for Global Mirror between the DS8100 systems, we need to take into account any changes to the data when sizing the links. Based on the client's requirements, the amount of data changing during the peak hour is 3% of 5 TB, or approximately 150 GB. If we assume that the changes are evenly divided over the hour, the changes are 2.5 GB per minute, or approximately 42 MBps or 336 Mbps. We use this number as a basis for our link sizing.

If we divide the amount evenly across four links, we get traffic of 84 Mbps over each link. However, to allow for the loss of one link or any peaks in the traffic, we divide the traffic only across three links, giving us 33% extra bandwidth and 112 Mbps traffic over each link. We also plan to have a maximum of 90% utilization on the link, so the minimum link speed we need is 125 Mbps.

Each link can be implemented over an OC-3 line that has the capacity of 155 MBps. An alternative is to use a Multiprotocol Label Switching (MPLS)-based shared connection, but due to possible router latency issues, we prefer the private OC-3 -based connection.

The most significant part of the OC-3 link latency is the propagation time of the light within the fiber. For a 600 km connection, with 1200 km round trip, it is:

1200 x 4.8 μs = 5.8 ms

We round this up to 6 ms to account for the packet transmission time over the 155 Mbps OC-3 link.

16.3.3 Normal operation

In normal operation, the production servers use only the DS8100 disks in the primary site. The DR servers are connected to the DS8100 in the DR site, but do not have the disks mounted or any applications running. The development servers use the DS6800 disks in the primary site and some capacity from the DS8100 in the primary site.

For the DS8100 disk subsystems, four of the eight ports are used for host attachments; the remaining four are used for the Global Mirroring. The ports used for Global Mirroring are directly connected to the routers.

In addition to the normal zoning, we define the following mSAN zones to our environment:

- Separate mSAN zones for the HBAs of any server in the development fabric that needs access to the DS8100, containing:
  - The HBA of the server
  - Both Fibre Channel ports of the DS8100 used for host attachment in the fabric
- A separate mSAN zone for the HBAs of any server in the development fabric that needs access to LAN-free backup, containing:
  - The HBA of the server
  - All Fibre Channel ports of the tape drives in the primary site connected to the fabric

In addition, we define zones for each Global Mirror connection in the backbone fabric.

16.3.4 Failure scenarios

This section describes how the failure of different components affects the operation of our solution:

- Power failure

  All of the SAN fabric components in the environment have dual redundant power supplies connected to different power circuits. Therefore, a power failure in one circuit does not have any effect on the operation.

- iFCP link failure

  The failure of a single iFCP link reduces the available bandwidth between the sites by 25%. However, because we assumed three available links in our sizing, the performance of the system will still remain adequate.

- Development fabric switch failure

  The failure of a switch in the development fabric reduces the Fibre Channel bandwidth available for the development and test servers by 50%. The traffic is automatically routed through the remaining paths by the SDD. The production environment is not affected.

- Primary site router failure

  If the router at the primary site fails, the capacity of the Global Mirror connection will be reduced by 50%. However, because we rounded up our link speed, we still have about 300 Mbps or about 90% of the peak hour capacity available.

  In addition, it reduces the Fibre Channel bandwidth available between the development and test servers, and the storage in the production fabrics, by 50%.

- Primary site director failure

  The director failure at the primary site reduces the Fibre Channel bandwidth available for production servers by 50%.

  In addition it reduces the Fibre Channel bandwidth available between the development and test servers, and the storage in the production fabrics, by 50%.

- DR site router failure

  If the router at the DR site fails, the capacity of the Global Mirror connection will be reduced by 50%. However, because we rounded up our link speed, we still have about 300 Mbps or about 90% of the peak hour capacity available.

- DR site director failure

  The director failure at DR site reduces the Fibre Channel bandwidth available for DR servers by 50%. However, in normal situations, those servers are idle, so this reduction affects the system only in a case where the production workload is already running at the DR site.

- Primary site DS8100 port failure

  If a port used for host access in the DS8100 at the primary site fails, the Fibre Channel bandwidth available for host access is reduced by 25%.

  If a port used for Global Mirror in the DS8100 at the primary site fails, the remaining Fibre Channel ports can sustain the full Global Mirror performance.

- Primary site DS8100 failure

  If the DS8100 at the primary site fails, all hosts lose access to it. This event can be promoted to site failure, and production can resume at the DR site.

- Primary site DS8100 port failure

  If a port used for host access in the DS8100 at the DR site fails, the Fibre Channel bandwidth available for host access is reduced by 25%. However, in normal operation, those servers are idle, so this reduction affects the system only in the case where the production workload is already running at the DR site.

  If a port used for Global Mirror in the DS8100 at the primary site fails, the remaining Fibre Channel ports can sustain full Global Mirror performance.

- DR site DS8100 failure

  If the DS8100 at the DR site fails, the Global Mirror connections change to Suspended state. The DS8100 at the primary site will accumulate changes to the data and copy the changed data over to the DR site when the DS8100 becomes available.

- Primary site failure

  If the complete primary site fails, the IBM eRCMF software starts the production at the DR site automatically. Although manual failover is also possible, it is difficult to manually reach the RTO target.

17

IBM TotalStorage m-type router implementation

In this chapter, we show the steps necessary to implement a remote connection between a host and a disk subsystem across an iFCP connection using two IBM TotalStorage SAN04M-R multiprotocol SAN routers.

This includes:

- Installing the SAN04M-R
  - Connecting power
  - Initial management network configuration
  - Element Manager connectivity
  - Changing default passwords
- Installing the SANvergence Manager application
- Upgrading the router firmware
- Initial router configuration
- Connecting a fabric to the router
- Configuring an iFCP connection between two SAN04M-R routers
- Zoning devices across the iFCP connection

17.1 Installing the router

In this section, we demonstrate how to install an IBM TotalStorage SAN04M-R multiprotocol SAN router, including setting the management TCP/IP address. It is assumed that the router has already been unpacked from its box and installed in a rack (or on a table), and is ready to be connected to the power source and to the management network.

17.1.1 Connecting the power

To power up the router:

1. Connect a power cord to the router and to a live power outlet. At this point, the router powers up.
2. Connect a second power cord to the other router power socket and another live power outlet.
3. Both green power LEDs (numbered 2 and 3 in Figure 17-1) light.
4. Wait a few minutes for the router to boot before proceeding to start the configuration. After the port lights go out, it is ready.

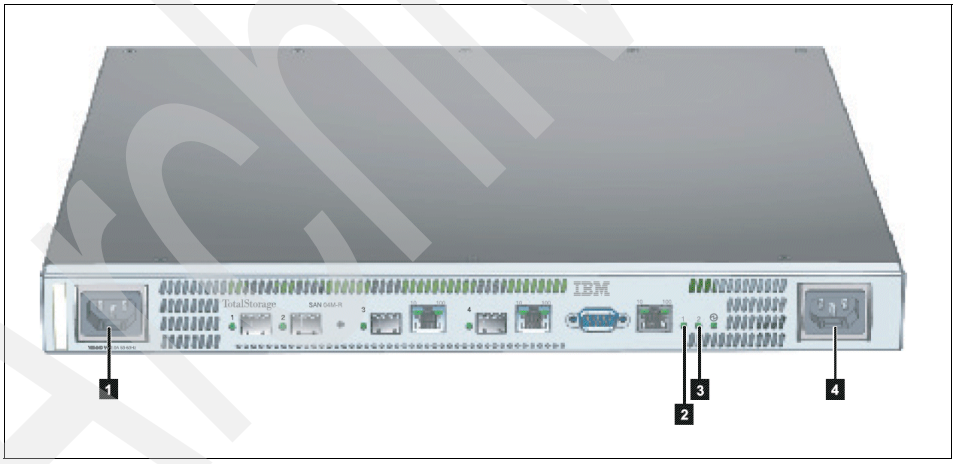


Figure 17-1 Power connections and LEDs

Note: It does not matter which power socket on the router is connected first.

Tip: We recommend that you power the unit from independent power boards for redundancy.

17.1.2 Configuring the management IP address

The next step requires that a static TCP/IP address be allocated for the management port of the router and that the appropriate subnet mask be known. In this example, we use an address of 9.43.86.120 and a 255.255.255.0 subnet mask.

If your management server is on a different subnet, you also need to know its TCP/IP address and subnet mask to be able to configure a permanent route to it on the router. In our case, the management server is in the same subnet, so we do not need to do this, but an example is shown in “Permanent route” on page 455.

Important: The management IP address *must* be on a different subnet from any that will be used to route storage traffic.

Using the supplied null modem cable, connect the router RS-232 serial port to a VT100 terminal. We use HyperTerminal on a Windows mobile computer and set the connection settings as shown in Figure 17-2.

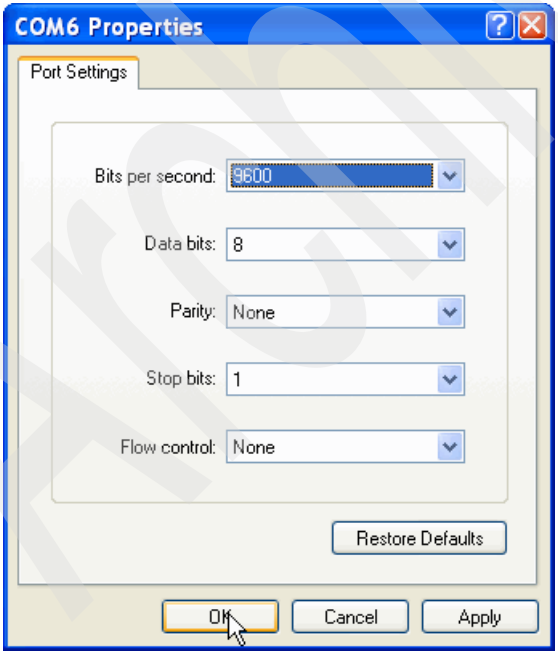


Figure 17-2 HyperTerminal connection properties

When connected, follow these steps to change the management IP address:

1. Press Enter to display the CLI prompt. You will see the following text.

Example 17-1 CLI login

```
McDATA ECLIPSE 1620 CLI Login (SW Rev 4.7.0.EF)
(Mgmt IP : 9.43.86.121, Build Date Oct 19 2005, 17:36:37)
Access Mode (read/modify):
```

2. At the Access Mode (read/modify) prompt, enter:

   ```
   modify
   ```

3. At the Password (Community string) prompt, enter:

   ```
   private
   ```

4. Enter the following command to display the current settings:

   ```
   show mgmt
   ```

   You will see the currently set management IP address.

Example 17-2 show mgmt command output

```
SAN04M-R# show mgmt
  Mgmt port address current          : 9.43.86.121
  Mgmt port mask current             : 255.255.255.0
  Mgmt port address on next reset    : 9.43.86.121
  Mgmt port mask on next reset       : 255.255.255.0
  Mgmt port MAC address              : 00:01:0F:05:36:40
  Mgmt Port Speed                    : Auto
```

5. We want to change the management address to 9.43.86.120. Enter the following command:

   ```
   set mgmt portaddr 9.43.86.120 255.255.255.0
   ```

   The following output appears in the HyperTerminal window.

Example 17-3 Setting the management IP address

```
SAN04M-R# set mgmt portaddr 9.43.86.120 255.255.255.0
  Succeeded in setting Management Port Address.
  Succeeded in setting Management Port Mask.
  Note:
    SAN Router reboot is required in order to make the changes
effective.
```

6. If the management server is on a different subnet, define a permanent route. Enter:

   ```
   set mgmt permroute IP_address subnet_mask gateway
   ```

   Where:

`IP_address`	TCP/IP address of remote management server
`subnet_mask`	Subnet mask of remote management server
`gateway`	IP address of the local network gateway

   Note: The management address and any permanent routes are not lost during a router reset to defaults.

7. Save the configuration to flash memory with:

   ```
   save
   ```

   The command produces the following output.

Example 17-4 Saving configuration to flash memory

```
SAN04M-R# save
Saving changes to flash...
```

8. Finally, reset the SAN router with the **`reset system`** command.

Example 17-5 Resetting the router

```
SAN04M-R# reset system

Relogin required after system reboots.
Resetting System...
```

Note: The save and system reset takes a few minutes to complete.

You can now close the terminal connection, unplug the null modem cable, and replace the maintenance port dust cap.

Permanent route

Because a default network gateway has not been defined for the management port, it can only communicate within its local subnet. In order to access the management port from outside of that subnet, a permanent route needs to be defined on the router.

As an example, suppose we want to have management access to the router from our site's wireless network, which has a network address of 9.1.56.0 and a subnet mask of 255.255.252.0 (giving a host address range of 9.1.56.0-9.1.59.255). We issue the following command:

```
set mgmt permroute 9.1.56.0 255.255.252.0 9.1.39.1
```

9.1.39.1 is the network gateway for the router management subnet.

In practice, any address within the host address range can be specified in the IP address field, and the router will calculate the appropriate network address based on the subnet mask.

If you want to restrict access to a specific address, enter that address and a subnet mask of 255.255.255.255 before the local gateway address.

You can only define one permanent static route.

17.1.3 Management network connection

Connect the 10/100 Ethernet management port (RJ45) to your LAN and confirm that one of the link LEDs lights (the left indicates 10 Mbps, and the right 100 Mbps). From a workstation on the same subnet, try to ping the address you entered previously to validate connectivity.

The router is now ready to be managed through the Web GUI Element Manager and by the EFCM management server.

17.1.4 Element Manager verification

The Element Manager is a Java applet used to manage individual SAN routers. Simply point your Web browser at the router management port IP address and a page similar to that shown in Figure 17-3 opens.

Note: On Windows, JRE 1.5 or later is required for Element Manager.

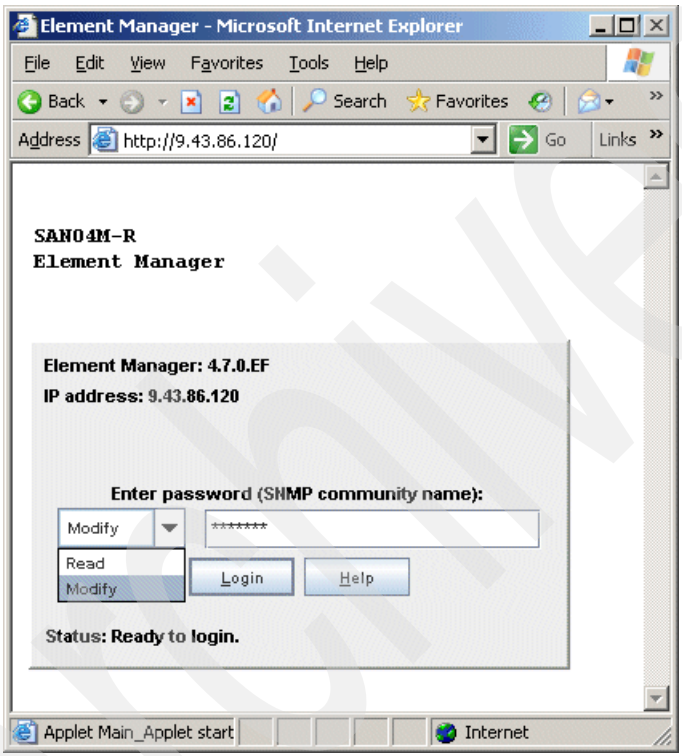


Figure 17-3 Initial Element Manager page

Select **Modify** from the drop-down list and enter `private` for the password. The Element Manager application opens, as shown in Figure 17-4.

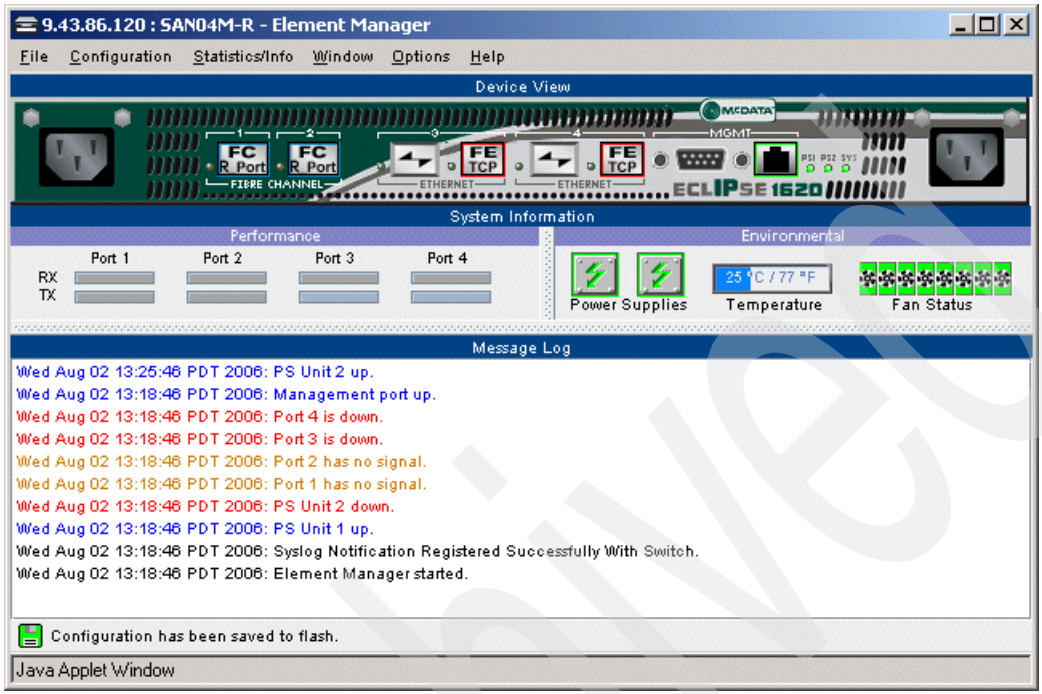


Figure 17-4 Element Manager application

17.1.5 Changing the passwords

As with any other product, change the default passwords before connecting the router to production equipment. As shown in Figure 17-5, select **Configuration → System → SNMP Communities/Hosts**.

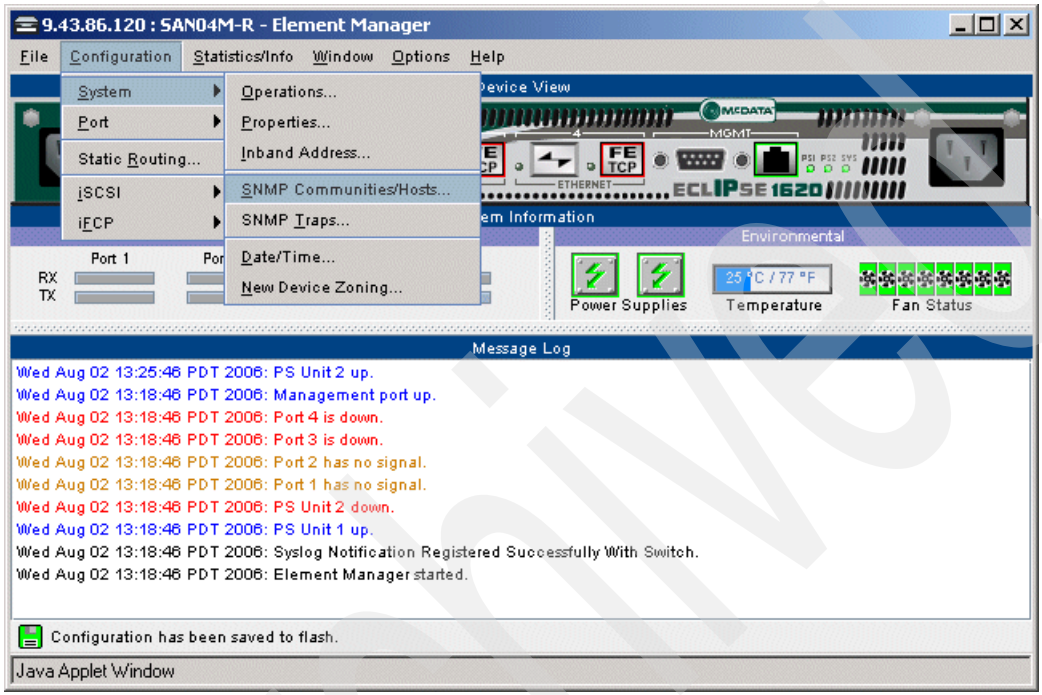


Figure 17-5 Changing SNMP community strings

Enter new Read and Write Community Strings and click **OK** (Figure 17-6). Remember to save the change to flash memory.

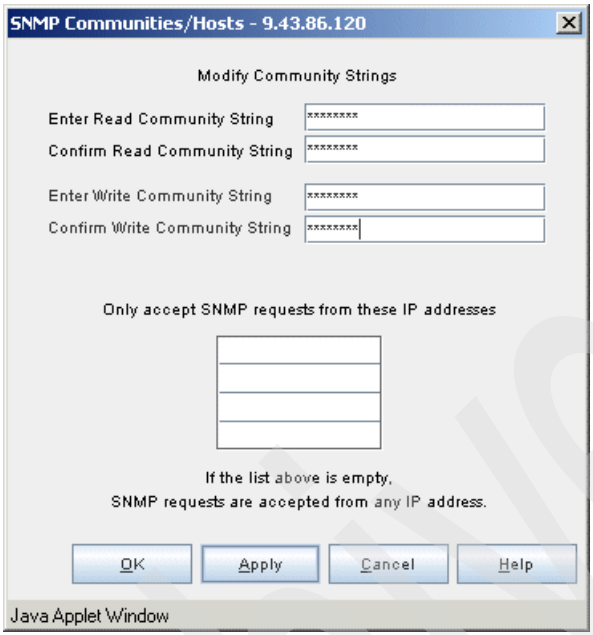


Figure 17-6 Entering new Community Strings

17.2 Installing SANvergence Manager

SANvergence Manager is a software application used to configure and manage routed SANs. It can run on a Microsoft Windows or Sun Solaris server, and in both cases, requires JRE 1.5 or later. In our case, it is installed on a Windows system. You can download the current version (4.7 at the time of writing) from the McDATA Web site.

Note: Enterprise Fabric Connectivity Manager (EFCM) 9.0 has SANvergence Manager functionality built-in. But at the time of writing, the stand-alone version of SANvergence Manager was still available for download from the McDATA Web site.

To install SANvergence Manager, perform the following steps:

1. The installation begins with extracting the files, as shown in Figure 17-7.

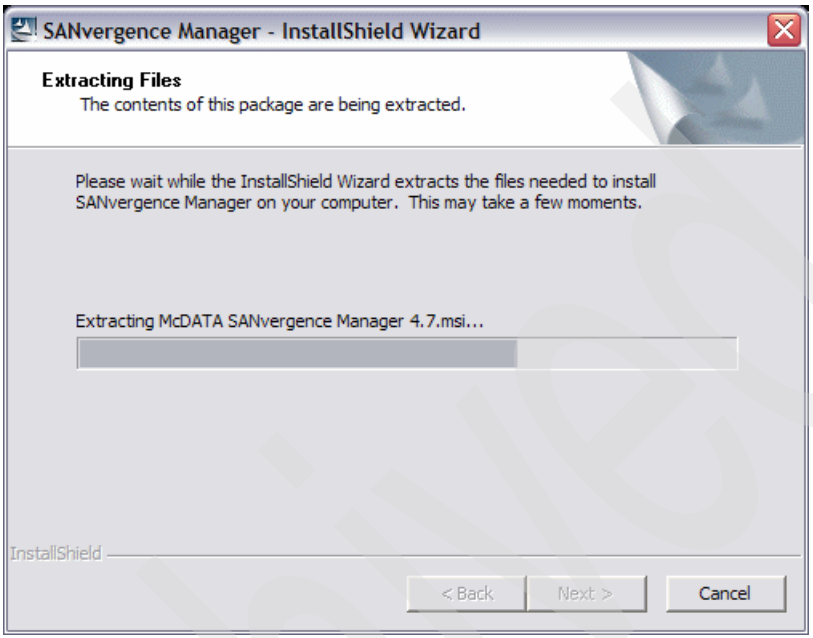


Figure 17-7 Extracting SANvergence Manager code for installation

The InstallShield Wizard launches.

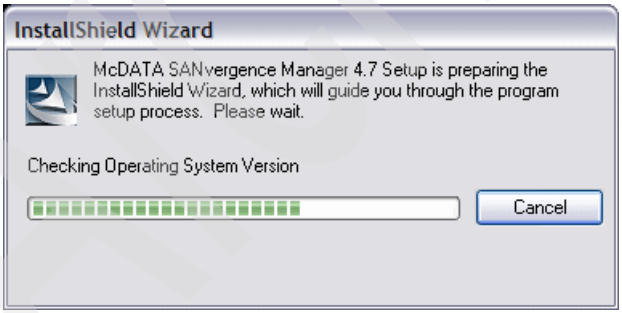


Figure 17-8 SANvergence Manager install, InstallShield starting

Then, the initial page opens, as shown in Figure 17-9.

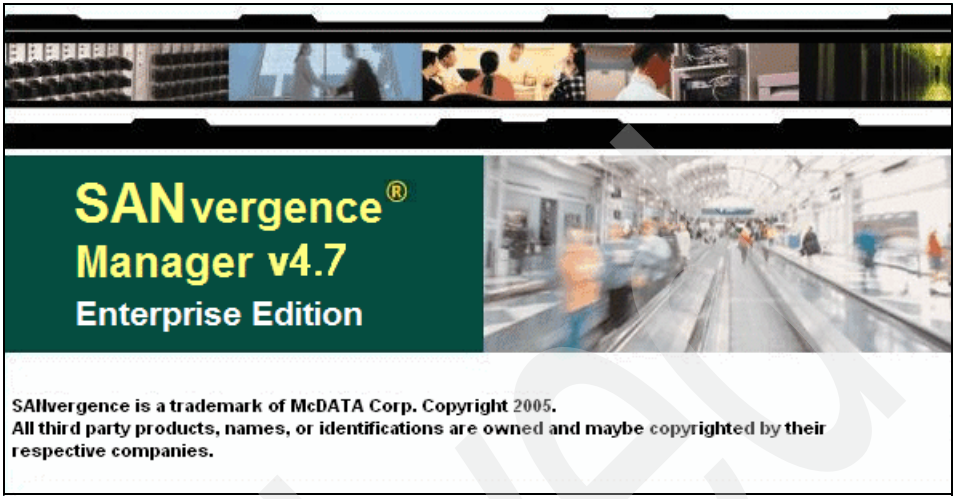


Figure 17-9 SANvergence Manager installation: Initial page

2. In the welcome window (Figure 17-10), click **Next**.

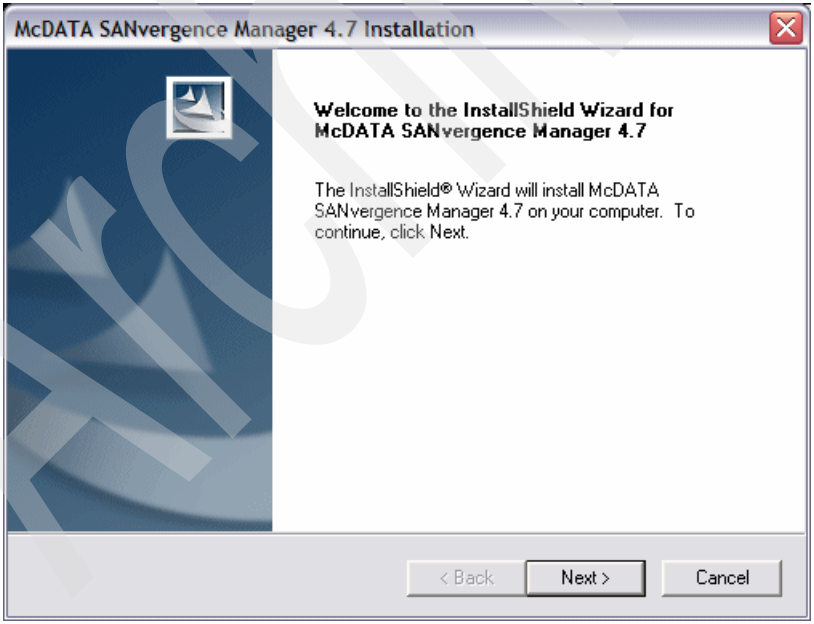


Figure 17-10 SANvergence Manager installation: Welcome window

3. Read and accept the licence agreement shown in Figure 17-11 by clicking **Yes**.

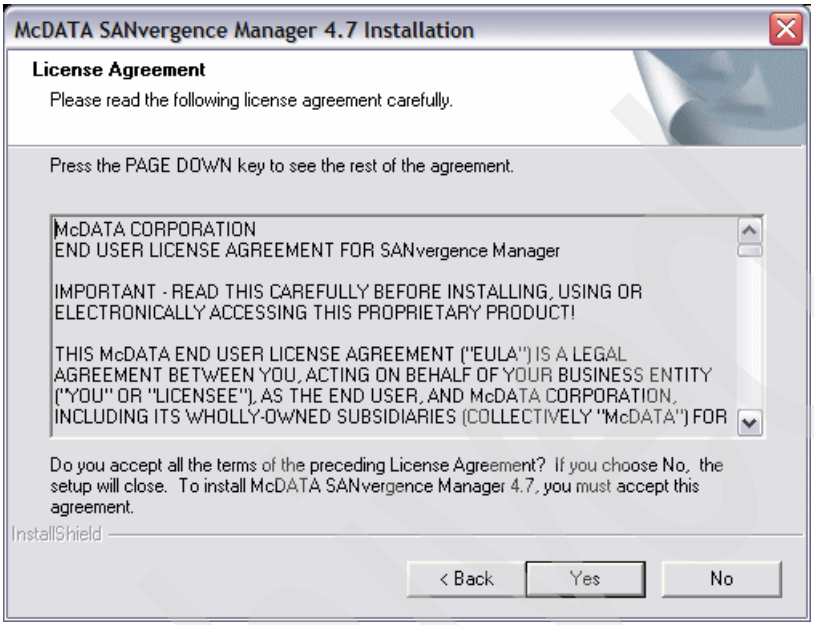


Figure 17-11 Accepting the SANvergence Manager licence agreement

4. Complete the Customer Information details and application access requirements, as shown in Figure 17-12, and click **Next**.

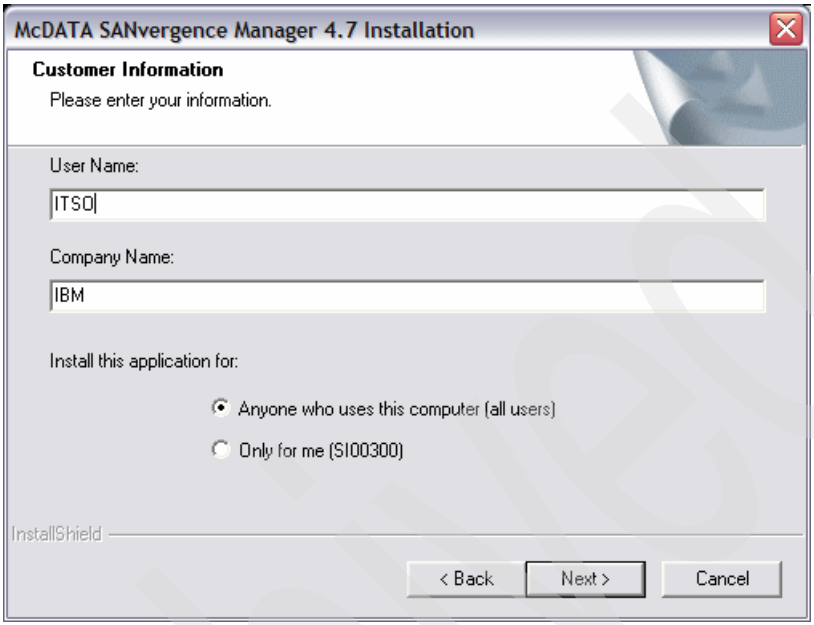


Figure 17-12 Customer Information and application access

5. Review the installation folder, shown in Figure 17-13, and click **Next**.

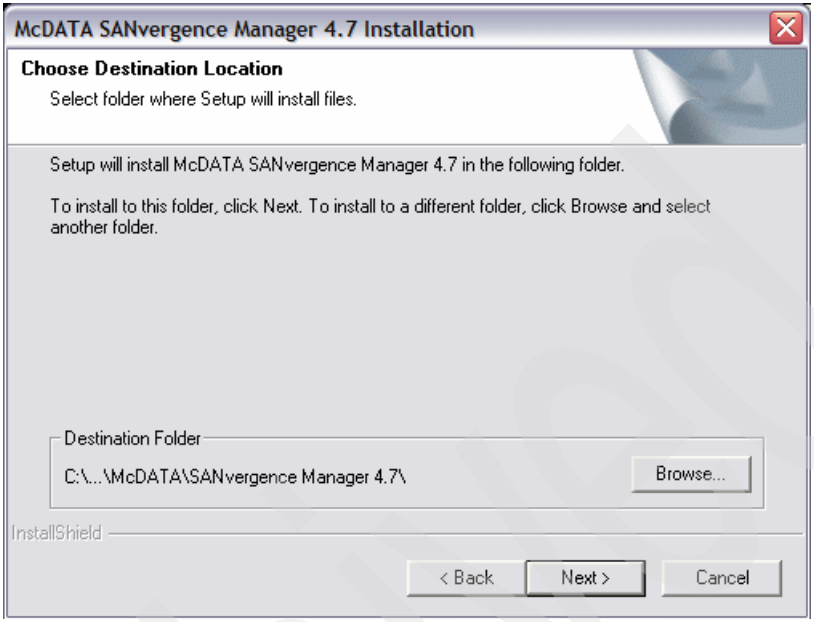


Figure 17-13 SANvergence Manager installation location

6. The installer is now ready to proceed, so click **Next** (Figure 17-14).

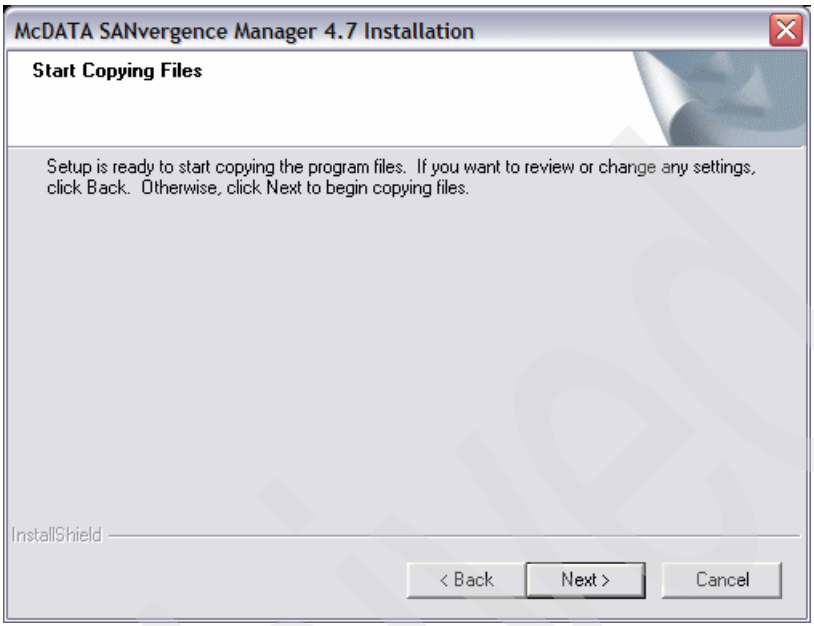


Figure 17-14 SANvergence Manager: Ready to be installed

Figure 17-15 shows the installation progress bar. This should reach 100% within a few seconds.

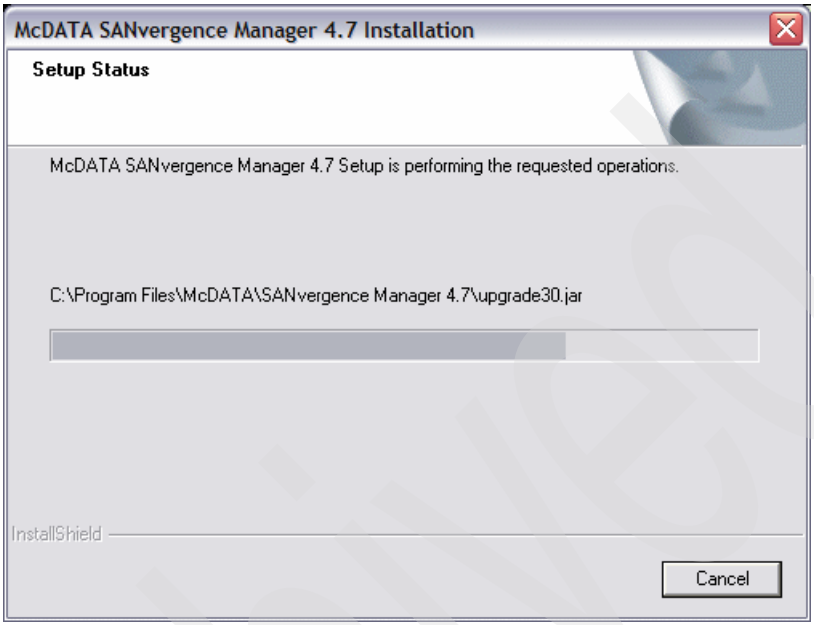


Figure 17-15 SANvergence Manager code installation progress

7. After the code installation finishes, as shown in Figure 17-16, click **Finish**.

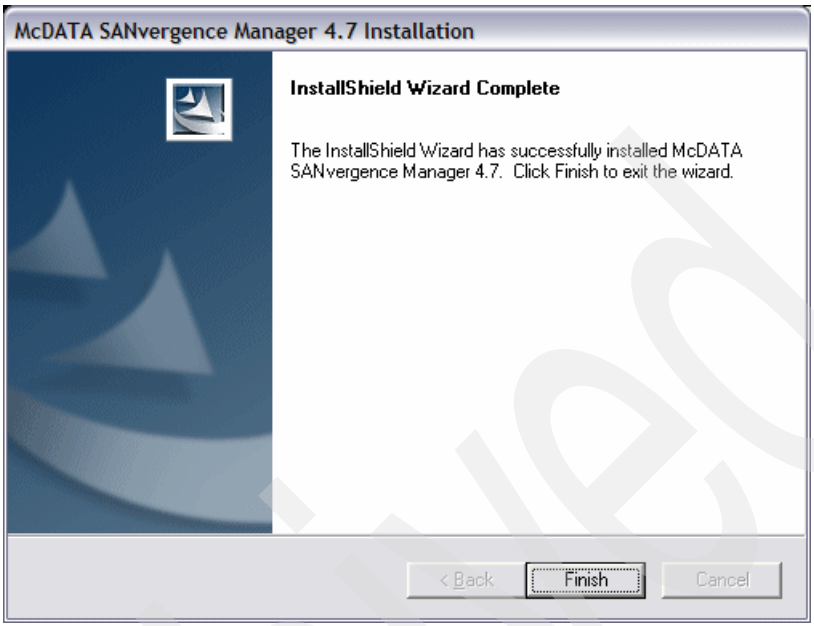


Figure 17-16 SANvergence Manager code successfully installed

17.3 Upgrading the firmware

Before configuring the router, check if the firmware requires upgrading. If it does, remember to read the *Release Notes* before beginning the upgrade. The *Release Notes* (and other guides) are available from the Technical Documents section of the McDATA Resource Library Web page:

`http://www.mcdata.com/resources/tdoc/index.html`

You must be registered to access the McDATA File Center Web site. Access this free site by clicking the **New User Registration** link at:

`http://www.mcdata.com/filecenter/template?page=index`

Figure 17-17 shows the required information needed to register.

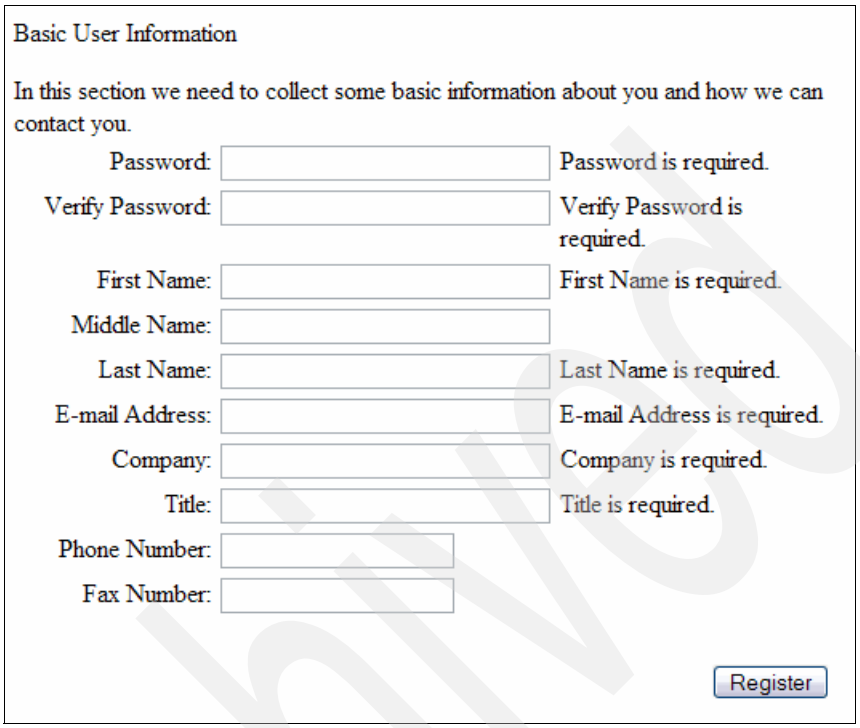


Figure 17-17 Required McDATA registration information

17.3.1 Downloading firmware from the McDATA Web site

Before proceeding, ensure that you have the serial numbers of your switches.

To download the firmware, perform the following steps:

1. Launch the McDATA logon Web page:
   `http://www.mcdata.com/filecenter/template?page=home.login`
2. Enter your credentials and click the **Login** button.

3. Using the toolbar at the top of the page, select the **Documents** tab. Select the appropriate switch family from the pane and click **search** (Figure 17-18).

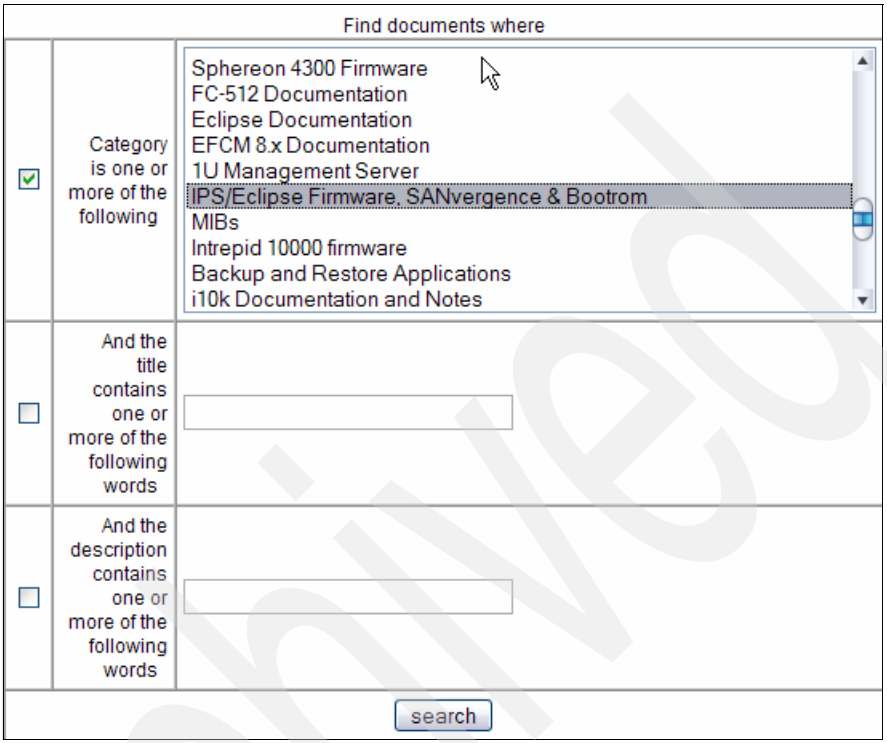


Figure 17-18 Firmware search

4. Figure 17-19 shows a sample of the search results. Find the appropriate file and select the **Add To Request** link.

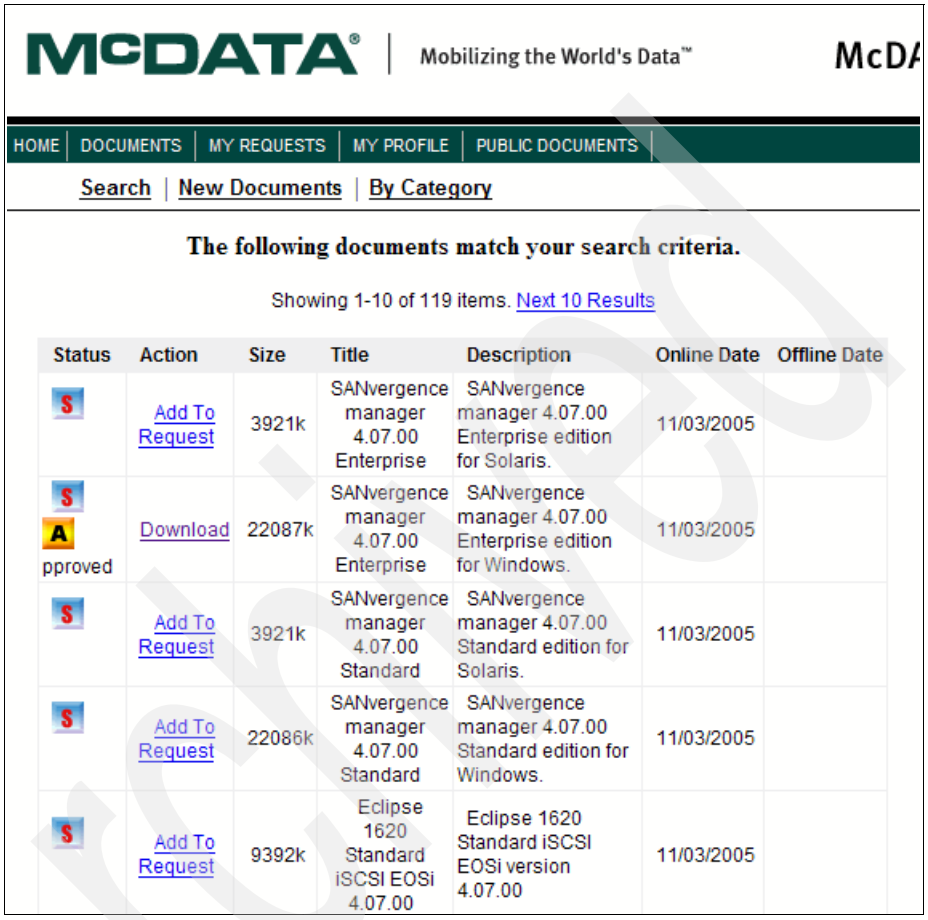


Figure 17-19 Firmware selection

5. Enter your switch's serial number on the next page and click the **Submit Request** button. Your request will display a status of Pending until it is approved, which usually takes less than an hour. You will receive an approval notification by e-mail.

> **Tip:** You can request more than one file by using the back button on your browser, making another selection, and clicking the **Submit Request** button after all your selections are complete.

17.3.2 TFTP server

The firmware upgrade process requires a TFTP server to serve the firmware files. If you do not have one, there are freeware servers available.

Generally, to server the firmware files, copy the firmware file to the TFTP server root directory, and ensure that the router TCP/IP address is authorized to access the TFTP server.

17.3.3 Upgrading the firmware with Element Manager

To upgrade the firmware with Element Manager, perform the following steps:

1. From the Element Manager, select **File** → **Firmware Upgrade**. The Firmware Upgrade window opens (Figure 17-20), showing the currently active boot image location, along with the firmware level and build date. In this case, 4.5.0 is the active level.

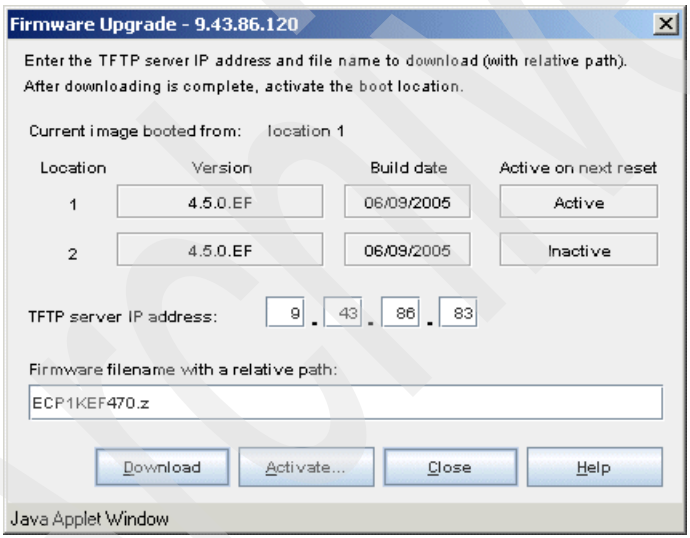


Figure 17-20 Specifying the TFTP server

2. Enter the TFTP server's IP address and the file name of the new firmware file, and click **Download**.

3. As shown in Figure 17-21, the download targets the currently inactive boot location. Click **Yes** to proceed.

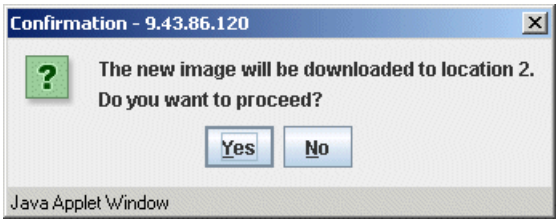


Figure 17-21 Download target confirmation

The firmware download procedure begins, and you can monitor the progress (Figure 17-22).

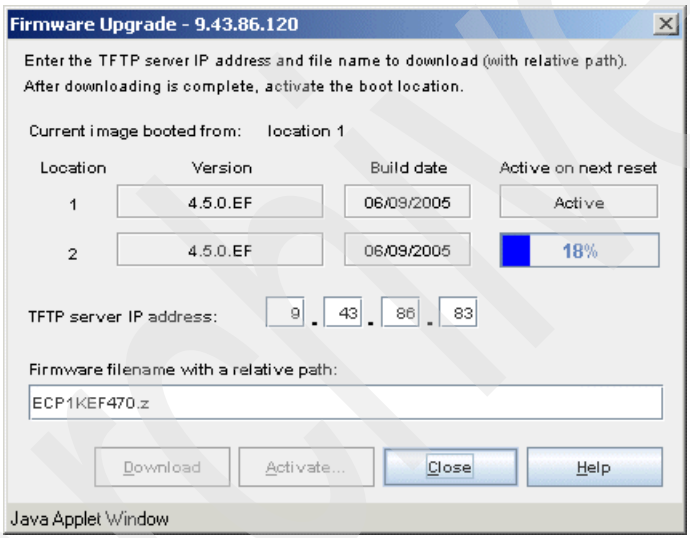


Figure 17-22 Firmware download progress bar

4. After the download completes, the message in Figure 17-23 appears. Click **OK**.

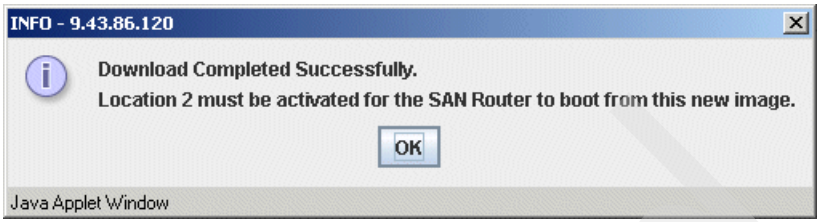


Figure 17-23 Firmware Download Completed Successfully

You are now returned to the main Firmware Upgrade window. As you can see, the new firmware has been applied, but it is not activated yet (Figure 17-24).

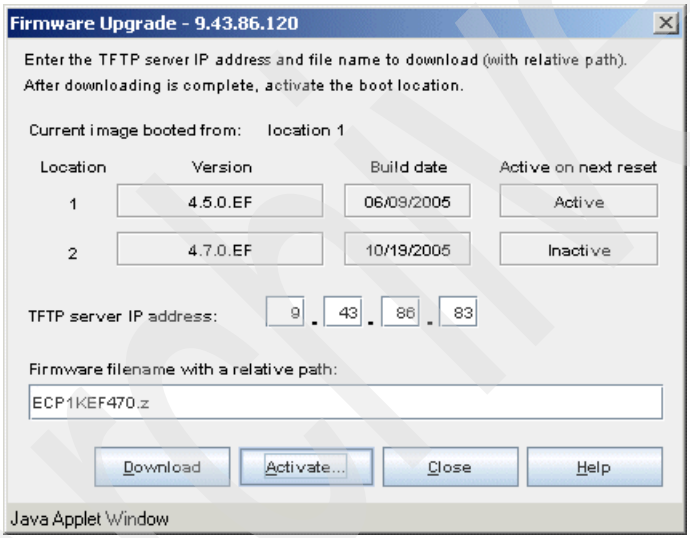


Figure 17-24 New firmware details

5. You now need to activate the boot location with the new firmware. Click **Activate** and select the new image, as shown in Figure 17-25. Click **OK**.

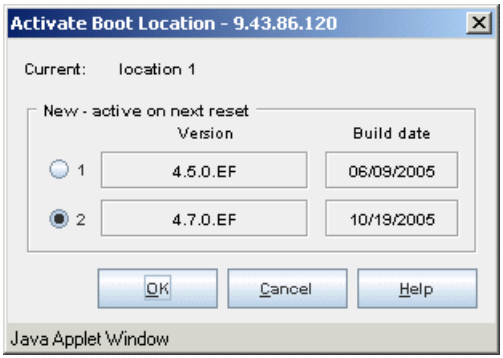


Figure 17-25 Selecting a new active boot image

6. A message opens confirming that the new image has been selected (Figure 17-26). Click **OK**.

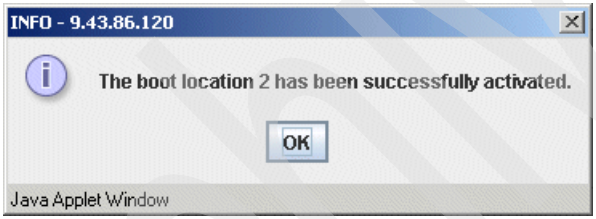


Figure 17-26 Boot image activated

7. The main Firmware Upgrade window appears again. Click **Close** to return to the Element Manager GUI.

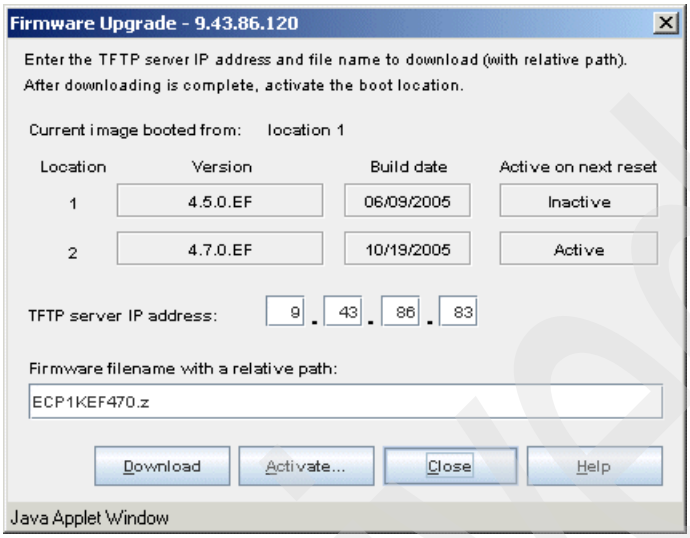


Figure 17-27 Firmware upgrade complete

Having successfully downloaded the new firmware, the router needs to be reset to make the new code active.

Note: Resetting the router will interrupt storage traffic.

To reset the router:

1. Select **File** → **Reset System** to reboot the router. Ensure that the **Reset SAN Router** radio button is selected, as shown in Figure 17-28. Click **OK**.

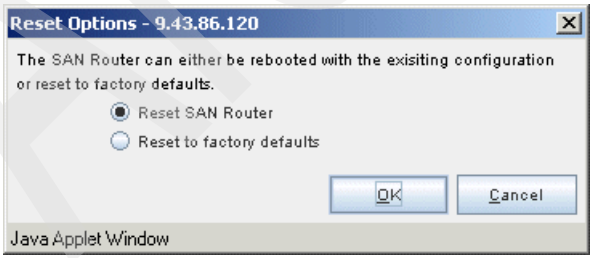


Figure 17-28 Selecting the Reset option

2. The warning shown in Figure 17-29 on page 477 appears to remind you that resetting the router will interrupt traffic. Click **OK** to continue.

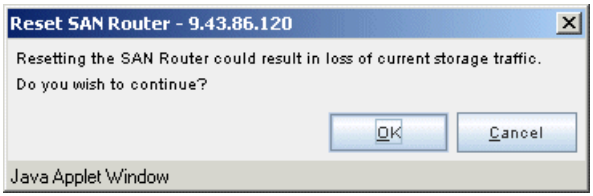


Figure 17-29 Router reset warning

The message in Figure 17-30 appears. After several minutes, the router reboots and the Element Manager window returns to normal.

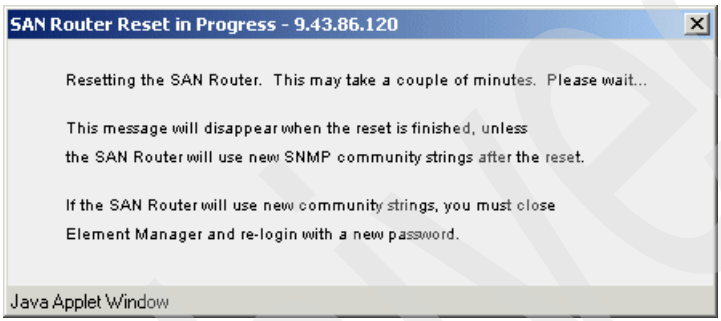


Figure 17-30 Router Reset in Progress

The router is now running level 4.7.0 code.

Client mismatch

It is possible that as a result of the firmware upgrade, the Element Manager Java application will require restarting. If so, the message in Figure 17-31 opens. Click **Yes**.

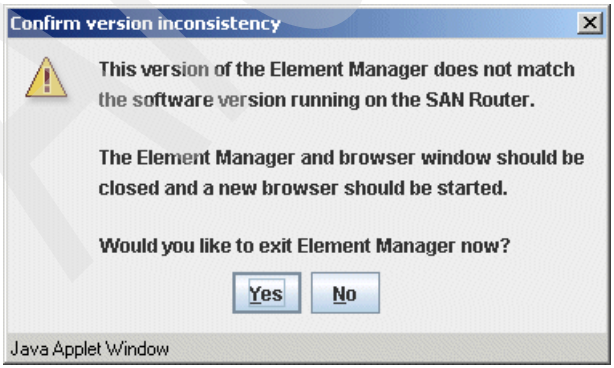


Figure 17-31 Element Manager mismatch after firmware upgrade

Note: Close your Web browser and re-open it before launching Element Manager again to ensure that the new level is used.

17.4 IP addresses

In addition to the management port address configured in 17.1.2, “Configuring the management IP address” on page 453, the router requires several more IP addresses to be set up.

17.4.1 Router inband IP address

The inband IP address is used for the internal delivery of storage traffic, and each router requires one such address. You configure it from the Element Manager by selecting **Configuration** → **System** → **Inband Address**. The SAN04M-R ships with a default inband address of 192.168.111.100 and a subnet mask of 255.255.255.0.

17.4.2 Intelligent port addresses

Each intelligent port (ports 3 and 4 on the SAN04M-R) requires two IP addresses:

- An internal address used to receive traffic from the internal network. This must be in the same subnet as the router inband IP address mentioned earlier. The router ships with default internal addresses of:
  - Port 3: 192.168.111.103 and subnet mask of 255.255.255.0
  - Port 4: 192.168.111.104 and subnet mask of 255.255.255.0
- An external address used for the iFCP or iSCSI connection over the external TCP/IP network. This must be allocated as a normal static TCP/IP address, along with a subnet mask and gateway address.

The *internal* network is the SAN fabric attached to a router FC port. The *external* network is the LAN or WAN attached to a router intelligent TCP port.

Outbound traffic arrives from the internal network, flows to the intelligent port’s internal address, is encapsulated for iFCP or iSCSI, and is transmitted from the port’s external IP address to the external network. Inbound traffic arrives from the external network at the intelligent port’s external address, is de-encapsulated, and is transmitted from the port’s internal address to the internal network using the inband address as the source address for the storage traffic.

Restriction: If the external IP address of an intelligent port is in the same network as the default internal network, the internal addresses *must* be changed to an unused network range. Otherwise, they can be left with the default values.

17.5 Example configuration

Figure 17-32 shows the configuration used for the example in this chapter.

The FC host contains an FC HBA that is attached to the router on the right. This is then remotely connected over TCP/IP using iFCP to the router on the left, which is fibre attached to the SAN16M-2 switch and thus to the FC-attached DS4400 Storage Server.

Remember that an intelligent port can only act as either an iFCP or an iSCSI port, but not both. So, if iSCSI host support is required, it needs to use a different port to attach to the IP network.

Note: The port order has been reversed on the router on the right to simplify the drawing.

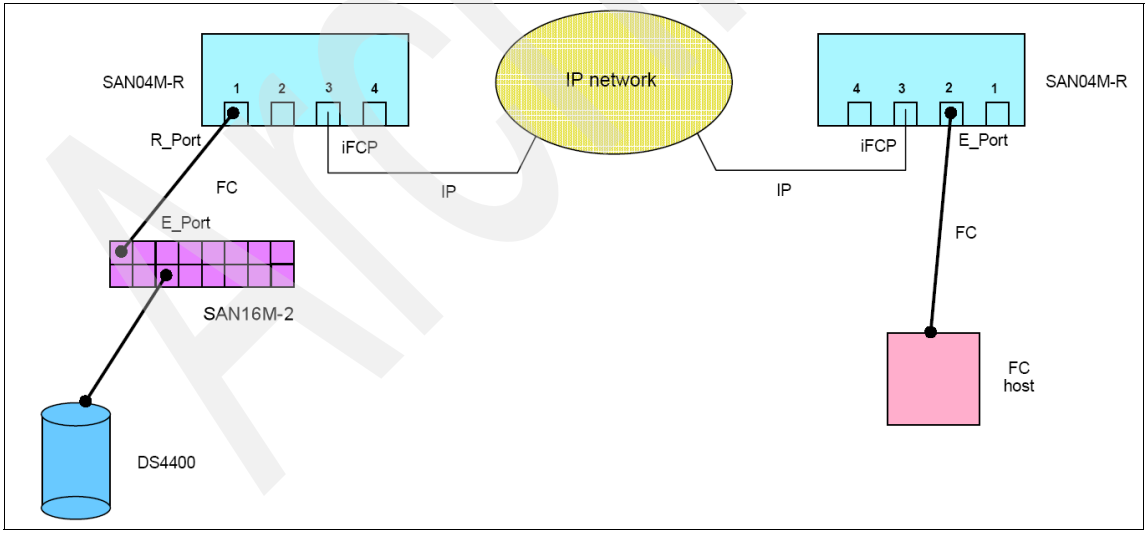


Figure 17-32 Laboratory configuration

17.5.1 Definitions

One or more SAN fabrics interconnected by a SAN router form a metro SAN, or mSAN. An mSAN is characterized by low latency, high quality, and high bandwidth ISLs. In our example, the left-side router and attached switch and disk are a very small mSAN, and the right-side router is another very small mSAN.

An internetworked SAN, or iSAN, is a collection of mSANs connected over iFCP by using two or more SAN routers, where at least one mSAN is at a distant (non-metro) location. It is characterized by high latency and low bandwidth ISLs. In our example, the whole diagram represents a small iSAN.

17.5.2 Example addresses

We use the following IP addresses:

- Left-side router
  - Internal IP address: 192.168.111.100 mask 255.255.255.0
  - Port 3 internal IP address: 192.168.111.103 mask 255.255.255.0
  - Port 4 internal IP address: 192.168.111.104 mask 255.255.255.0
  - Port 3 external IP address: 10.1.41.3 mask 255.255.0.0
  - Port 4 external IP address: 10.1.41.4 mask 255.255.0.0
- Right-side router
  - Internal IP address: 192.168.111.100 mask 255.255.255.0
  - Port 3 internal IP address: 192.168.111.103 mask 255.255.255.0
  - Port 4 internal IP address: 192.168.111.104 mask 255.255.255.0
  - Port 3 external IP address: 10.1.42.3 mask 255.255.0.0
  - Port 4 external IP address: 10.1.42.4 mask 255.255.0.0

Note: For this example, the external port addresses for both routers are in the same subnet, and thus will not require a network gateway definition.

17.6 Basic router configuration

We use the Element Manager with modify access for all of the following steps.

17.6.1 System properties

Select **Configuration** → **System** → **Properties** and enter a router name, contact, and location, as illustrated in Figure 17-33. Click **OK**.

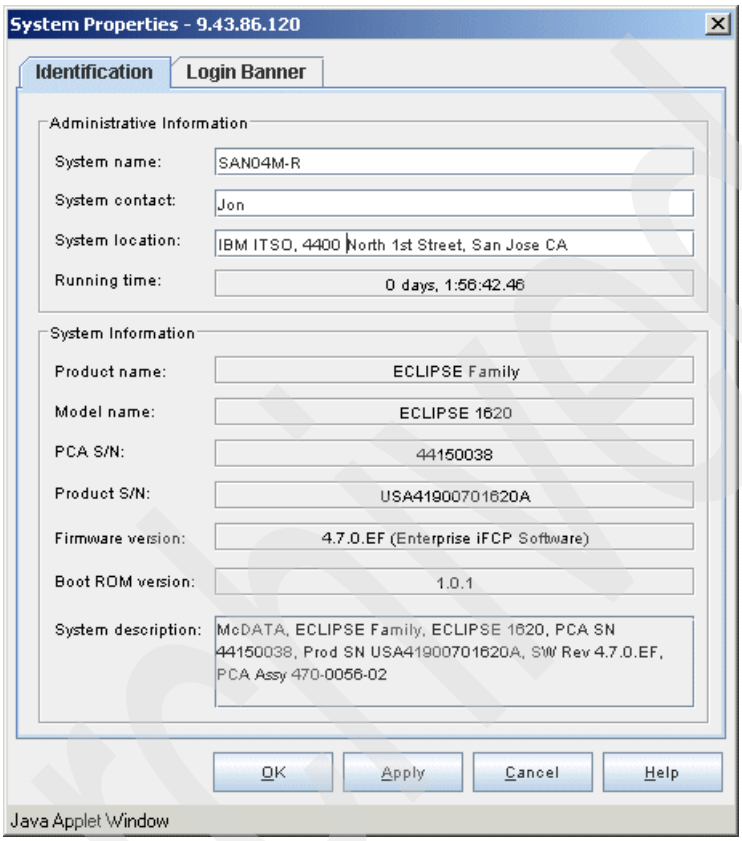


Figure 17-33 Setting system identification

Note: Figure 17-33 shows system information for the McDATA Eclipse 1620 SAN router. The IBM TotalStorage SAN04M-R SAN router displays different system information.

17.6.2 Date and time

Select **Configuration** → **System** → **Date/Time** and set the date and time, as shown in Figure 17-34. Optionally configure the router to use SNTP, either as a client, or as an SNTP server for other SAN routers.

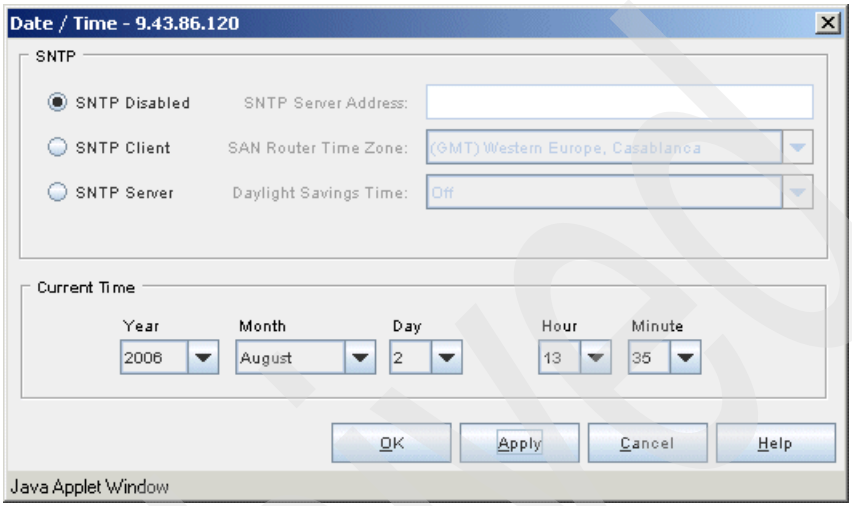


Figure 17-34 Setting the date and time

17.6.3 Cluster ID

Every router connected to an mSAN must have its own cluster ID. This ID is used by the R_Ports to register a unique virtual node WWN with connected fabrics. Select **Configuration** → **System** → **Operations** and enter a unique ID.

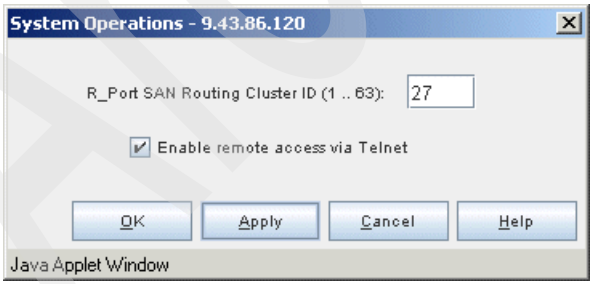


Figure 17-35 Setting the cluster ID

As reminded by the message shown in Figure 17-36, the R_Ports must be disabled and enabled for changes to this value to take effect. Click **Yes**.

Save this change to flash memory after making the other configuration changes.

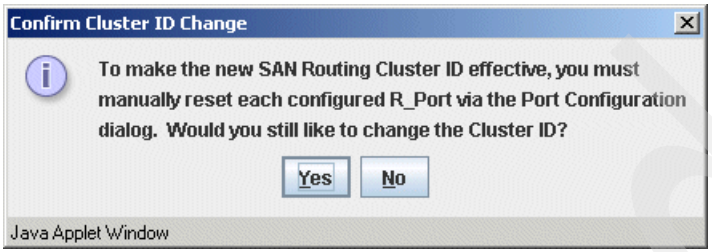


Figure 17-36 Confirming cluster ID change

17.6.4 SNMP

If you have not done so already, you should now change the SNMP community strings from their default values, by selecting **Configuration** → **System** → **SNMP Communities/Hosts** and entering new read and write values. As shown in Figure 17-37, you can also restrict the hosts allowed to send SNMP requests to the router.

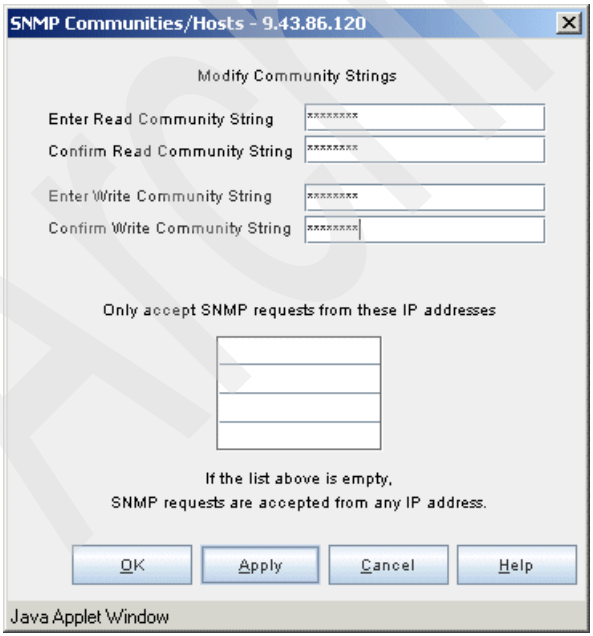


Figure 17-37 SNMP settings

To configure SNMP trap recipients, select **Configuration** → **System** → **SNMP Traps** and enter up to eight IP addresses, as shown in Figure 17-38.

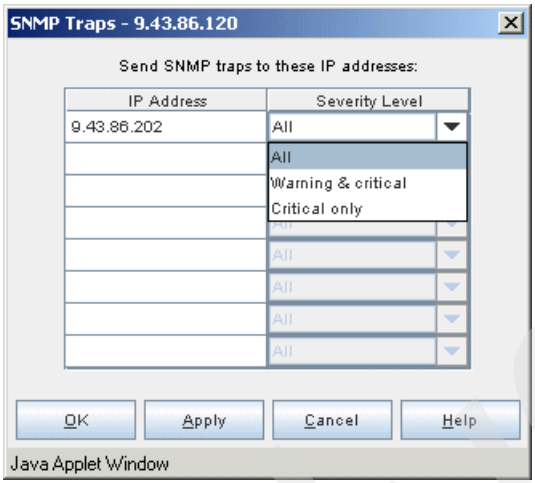


Figure 17-38 Defining SNMP Trap recipients

Save this change to flash memory after making the other configuration changes.

17.6.5 Inband IP address

As listed in 17.5, “Example configuration” on page 479, we use an inband address of 192.168.111.100 for our routers. Select **Configuration** → **System** → **Inband Address** and verify that the address and mask are correctly set. If this address is changed, a router reset will be required for it to take effect. This is highlighted in Figure 17-39 by the separate Current Configuration values column.

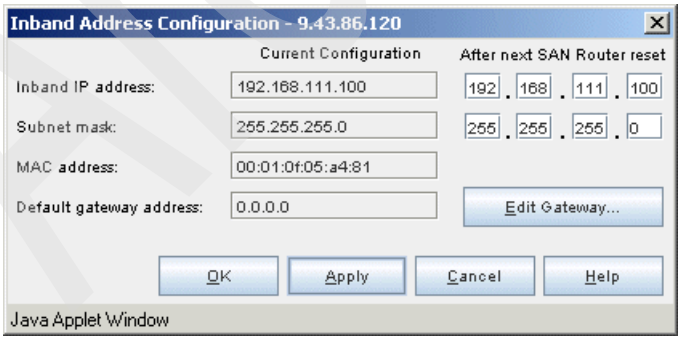


Figure 17-39 Changing the Inband IP address

Save this change to flash memory after making the other configuration changes.

17.6.6 New device zoning

You can configure the default zoning action for devices newly attached to the router. They can either be left unzoned, or they can be placed in the default router zone (zone 1). The default, and recommended, setting is to leave them unzoned so that they have to be explicitly assigned to a zone. To review this setting, select **Configuration** → **System** → **New Device Zoning**. The window in Figure 17-40 opens.

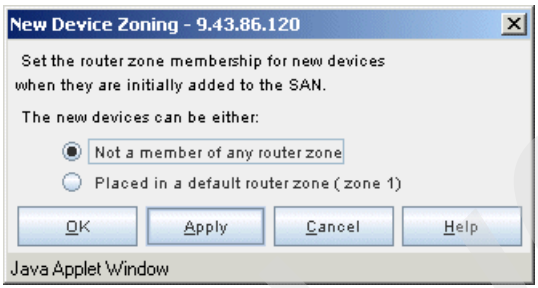


Figure 17-40 New Device Zoning default

Save this change to flash memory.

17.6.7 Saving changes to flash memory

Many configuration changes are only made permanent by updating the router's flash memory. Select **File** → **Save Configuration**. We show this in Figure 17-41 on page 486. Saving to flash memory can take about a minute to complete.

Tip: The current status of changes is shown by the colored diskette icon in the bottom left corner of the Element Manager window and its associated message. If red, the configuration has changed since it was last saved to flash.

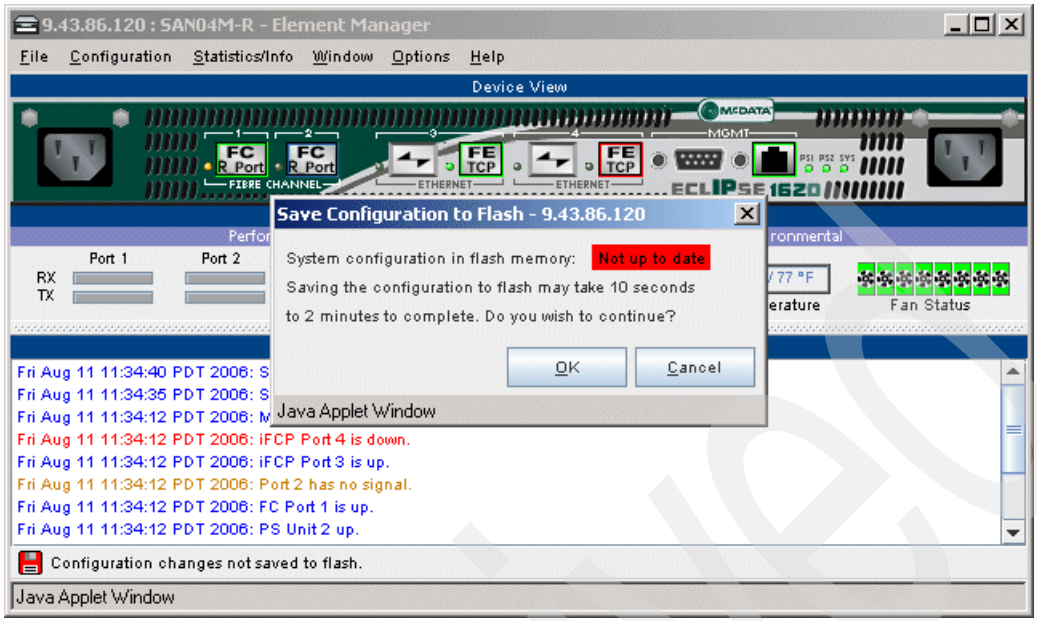


Figure 17-41 Saving the configuration change to flash memory

17.7 Connecting a fabric to the router

To connect the SAN04M-R to an FC switch, such as the SAN16M-2 in our example (see Figure 17-32 on page 479), the port must be configured as an R_Port. An R_Port is a fabric extension port like an E_Port, but intended for connecting routers to switches, rather than switches to switches.

Note: Ports 1 and 2 can also be configured as FC-Auto for direct attachment to FC devices such as servers or storage.

17.7.1 Configuring the R_Port

To configure the R_Port, complete the following steps:

1. Select **Configuration** → **Port** → **FC/Ethernet**:
   a. Select port **1** from the Port number drop-down list at the top of the window (Figure 17-42).

Figure 17-42 Configure port 1 as R_Port

   b. Enter a description for Port name and ensure that the **Enable port** check box is selected. In our case, we changed the Port name value to `FC Port 1`.

c. Select the **R_Port** radio button, and the display changes to show extra fields, as shown in Figure 17-43.

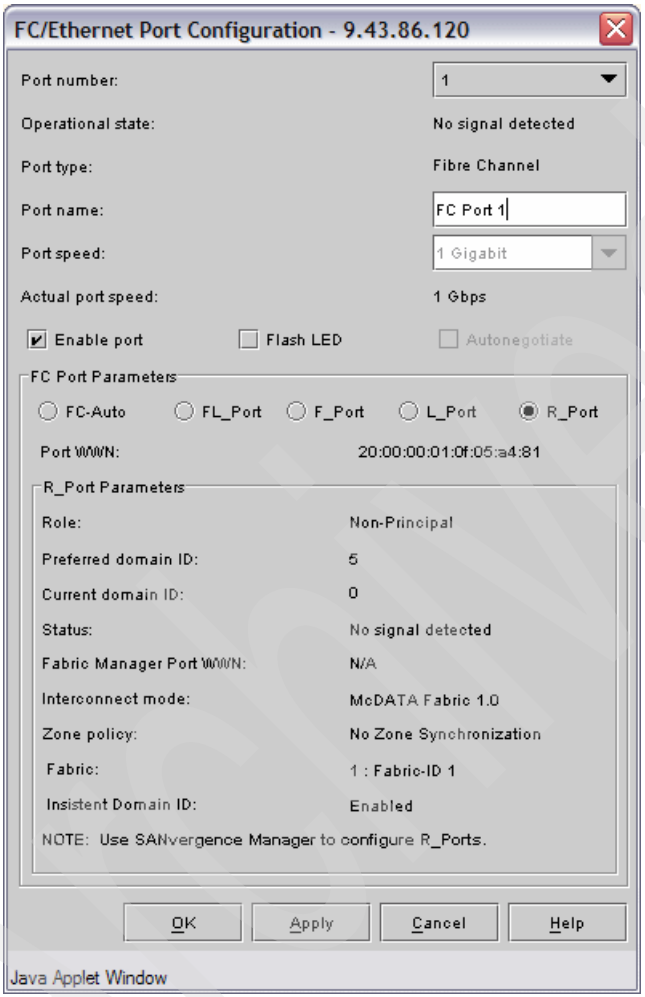


Figure 17-43 Port 1 R_Port details

d. Click **Apply** and the warning in Figure 17-44 appears.

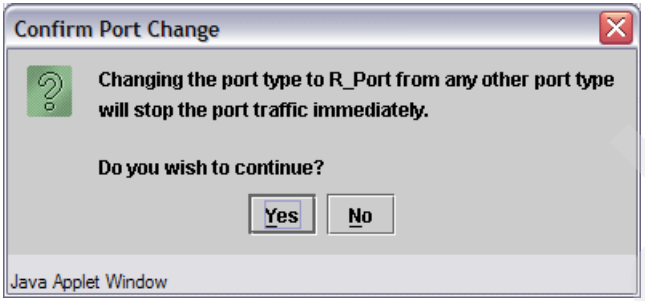


Figure 17-44 Port change warning

The R_Port parameters fields should now be similar to those shown in Figure 17-45.

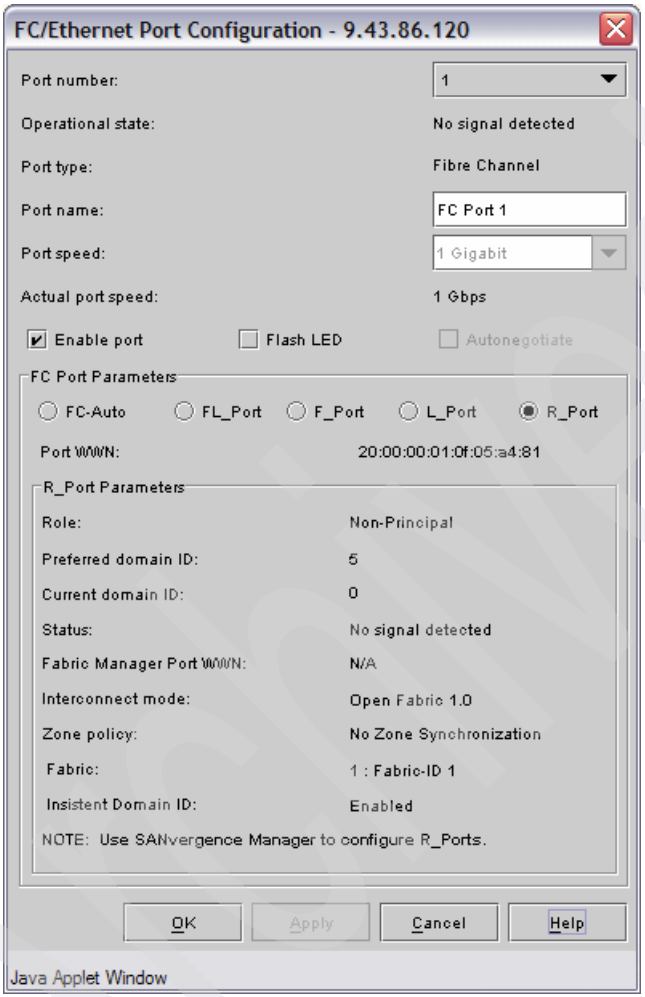


Figure 17-45 Port status after applying change

2. To continue the configuration process, log in to SANvergence Manager. Launch the application and log in as **Administrator**, as shown in Figure 17-46.

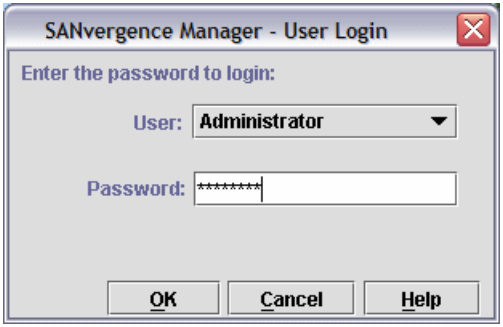


Figure 17-46 SANvergence Manager Login

3. Next, you need to log in to the mSAN. Enter the IP Address of the router, the Modify community string, provide a name for the mSAN, and select the **Out-of-band** radio button, as shown in Figure 17-47. Click **OK** to log in.

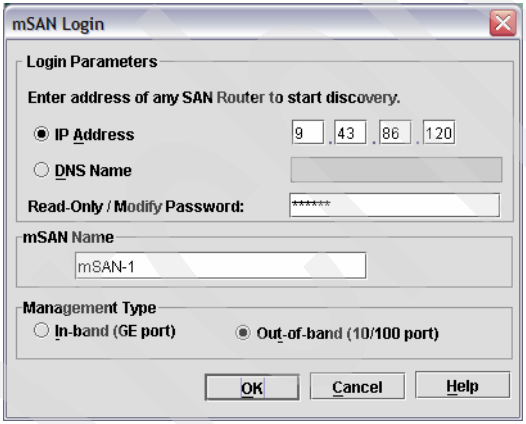


Figure 17-47 mSAN Login

The main SANvergence console shown in Figure 17-48 opens.

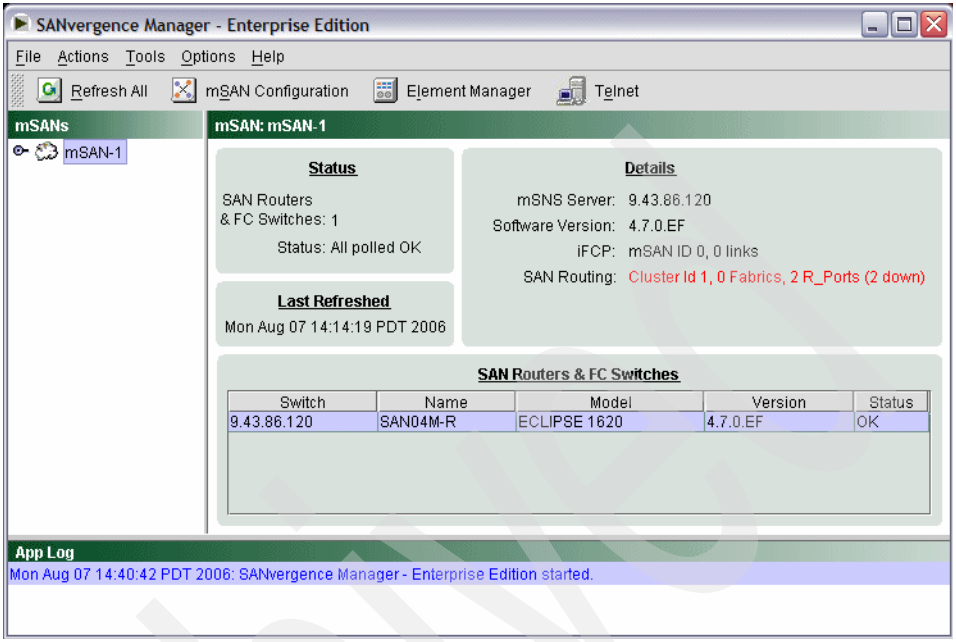


Figure 17-48 SANvergence Manager console

4. Ensure that the mSAN is highlighted in the left column and click the **mSAN Configuration** button. As soon as the mSAN information is retrieved from the router, the window shown in Figure 17-49 opens.

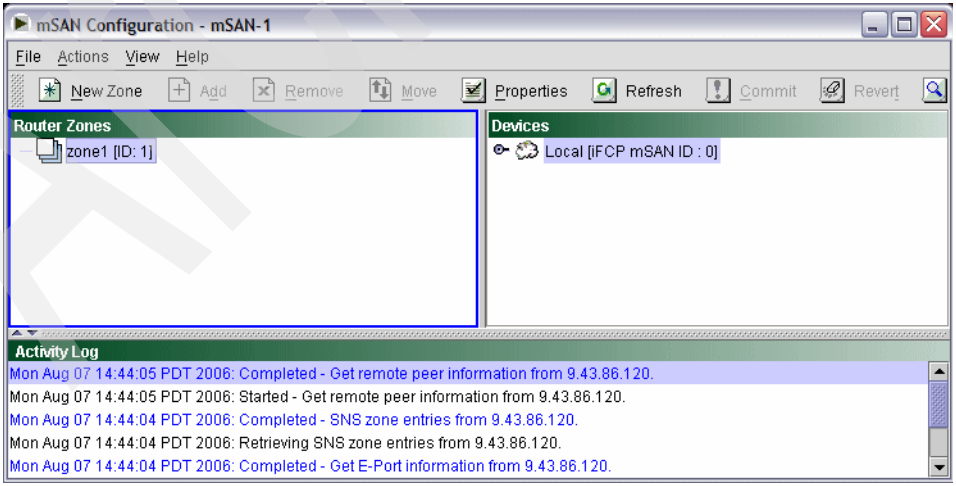


Figure 17-49 mSAN Configuration window

5. Select **Actions** → **Fabric Configuration** and the Fabric Configuration window opens.
6. On the Fabrics tab (Figure 17-50), you can set the Connection Mode and Zone Policy:
   - Even though we are setting up a connection to the SAN16M-2 switch, we decided to set Connection Mode to **Open Fabric 1.0** to allow for connectivity to open fabric-compliant switches from other manufacturers.
   - Set the Zone Policy to **No Zone Synch**.

   Click **Update Fabric** to store your changes pending commitment.

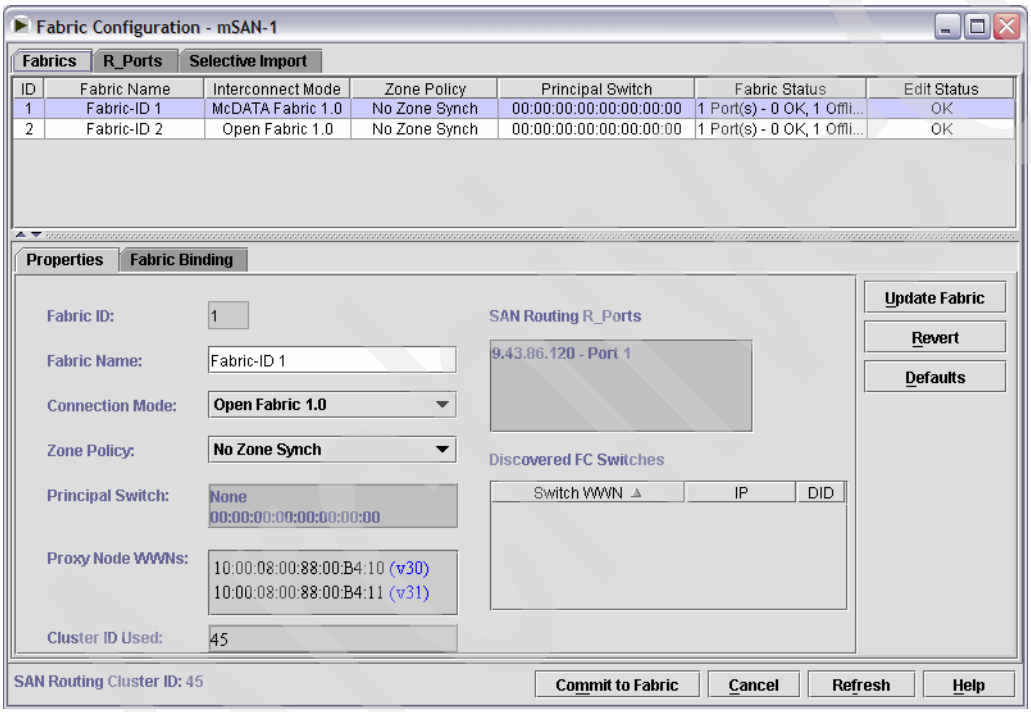


Figure 17-50 Fabric Configuration: Fabrics tab

7. Select the **R_Ports** tab, as shown in Figure 17-51 on page 494. From here, you can modify the R_Port characteristics as follows:
   - Set a Preferred Domain ID that is different from any already in the mSAN (in our case we chose 21).
   - Select the **Insistent Domain ID Enable** check box to enforce the chosen domain ID.

   Click **Update R_Port** to store your changes pending commitment.

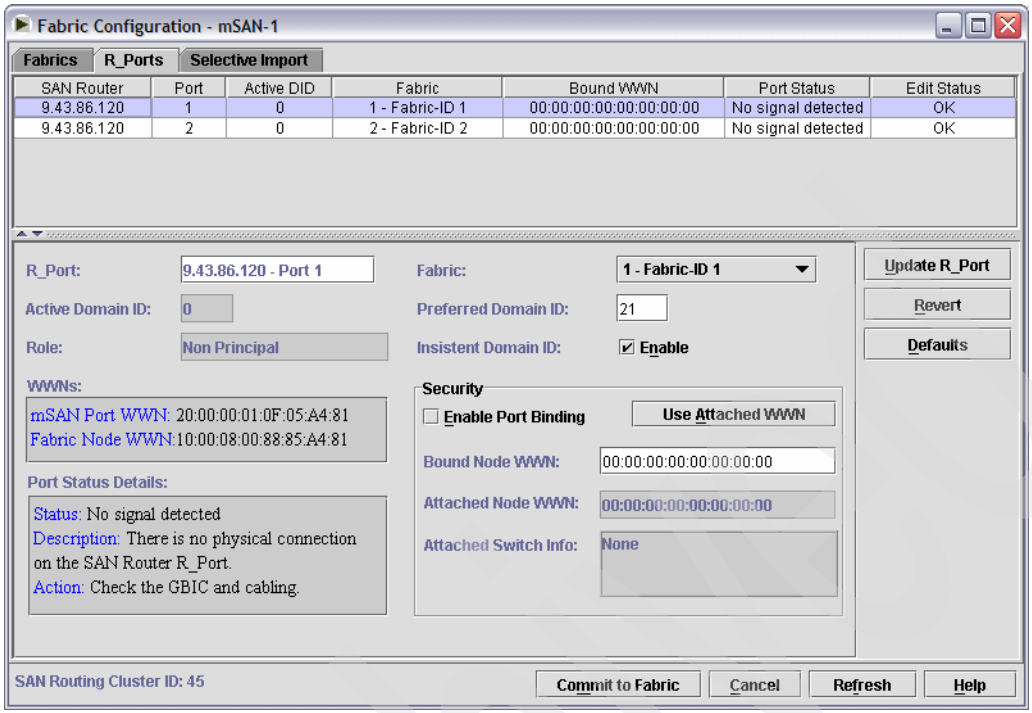


Figure 17-51 Fabric Configuration: R_ports tab

8. Click the **Commit to Fabric** button. The fabric modification warning shown in Figure 17-52 appears. Click **Commit and Save** to proceed.

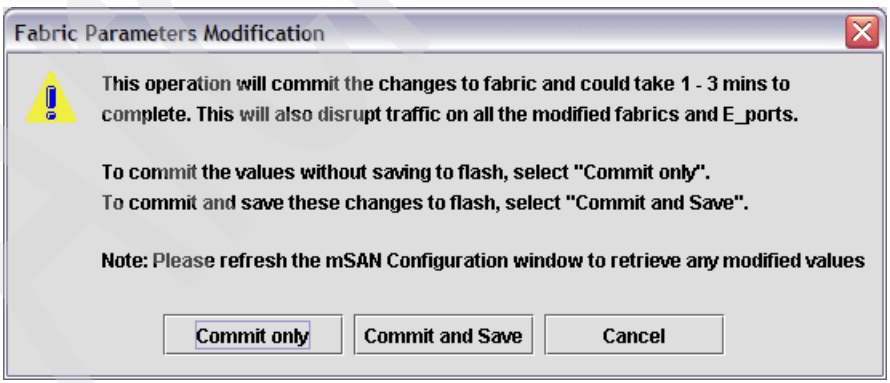


Figure 17-52 Fabric modification warning

A progress bar shows the progress of the changes.

9. Once complete, close the Fabric Configuration and the mSAN Configuration windows, and return to the router Element Manager. The router must now be reset for all the changes to take effect. Having checked that all changes have been saved to flash memory, select **File** → **Reset System**. Ensure that the **Reset SAN Router** radio button is selected and click **OK**.
10. After the reset completes, open the FC/Ethernet Port Configuration window to confirm the changes, as shown in Figure 17-53.

Figure 17-53 R_Port online

11. You can also check the status on the switch with the EFCM Element Manager, which should show the port as an E_Port, as shown in Figure 17-54.

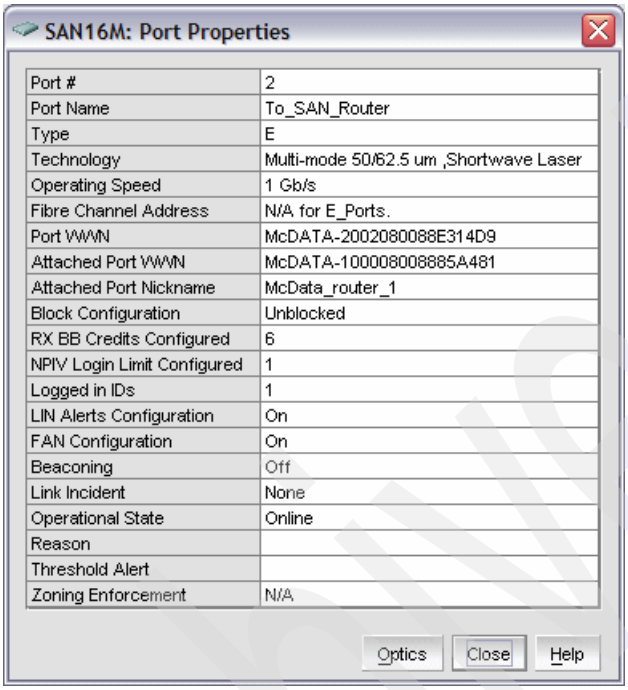


Figure 17-54 Switch E_Port online

17.7.2 Selective import

The last step is to selectively import those devices into the fabric which you want to be seen by other fabrics connected to the router:

1. From SANvergence Manager, re-open the mSAN Configuration window.
2. Select **Actions** → **Fabric Configuration**. The Fabric Configuration window opens.
3. Select the **Selective Import** tab.
4. From the Selected Fabric drop-down list, select **1 - Fabric-ID 1**.
5. Click **Retrieve** and the devices connected to the fabric will appear.

6. Highlight the device you want and click **Import** (the window should now look similar to Figure 17-55).
7. Click **Commit to Fabric** to finish.

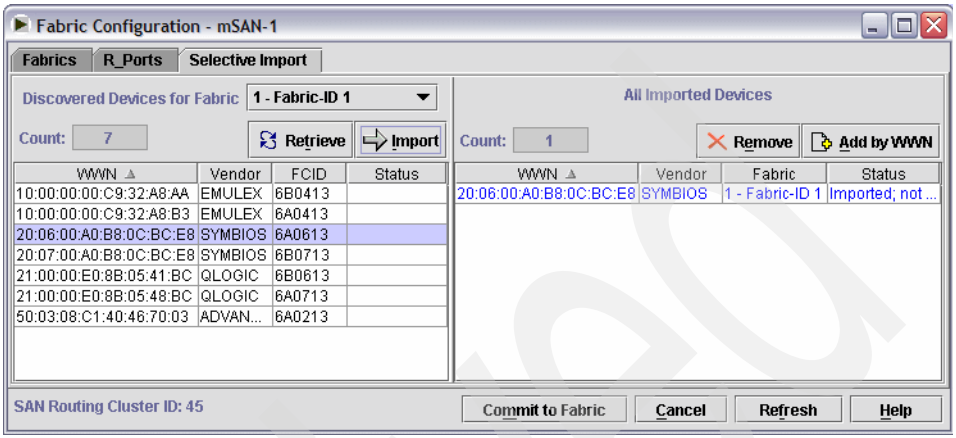


Figure 17-55 Selectively importing a device

8. Respond to the warning shown in Figure 17-56 by clicking **Commit and Save** (or **Commit only** if you will be making other changes later). This will take a short while.

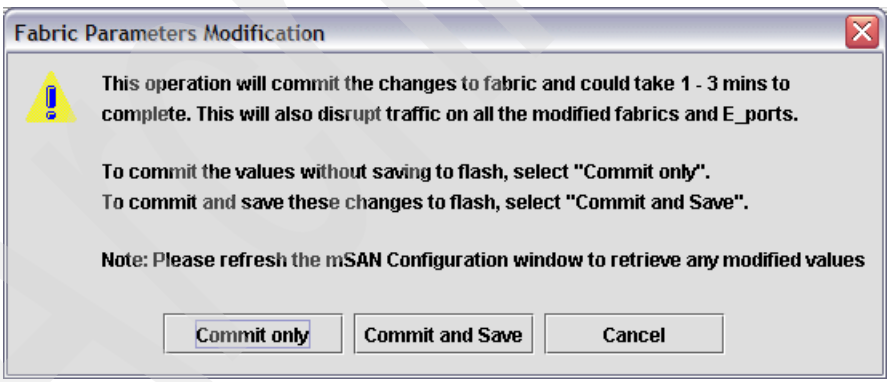


Figure 17-56 Committing the import

9. Close the Fabric Configuration window and click **Refresh** in the mSAN Configuration window. The imported device appears in the Devices panel, as shown in Figure 17-57.

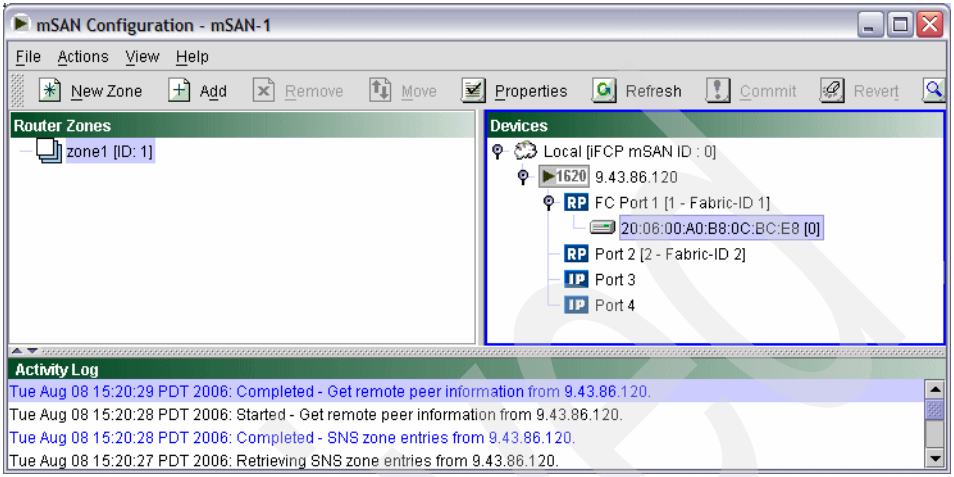


Figure 17-57 Imported storage device in mSAN Configuration window

10. Double-click the device to open the Properties window, where you can assign a port label to appear in place of the WWN highlighted in Figure 17-58.

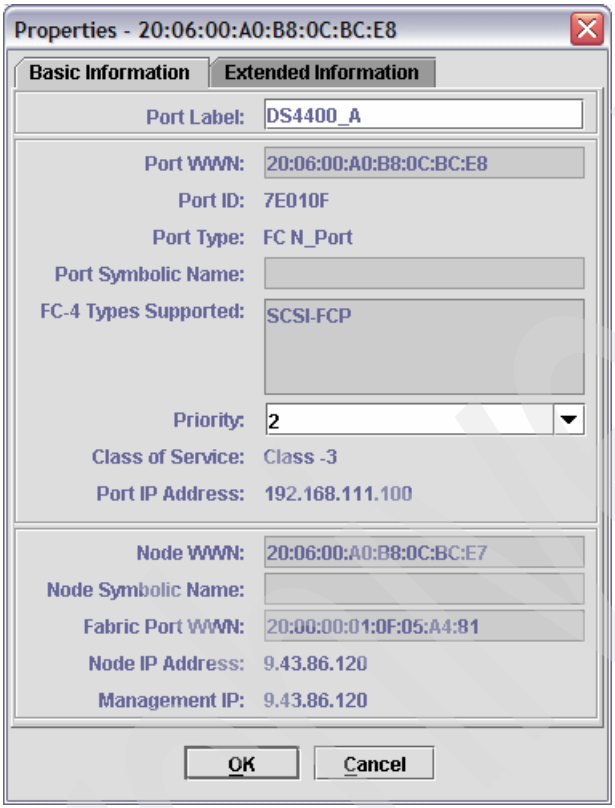


Figure 17-58 Assigning a Port Label

The new device is now available for zoning.

17.7.3 Troubleshooting

If the link does not come online, double-check the following items on both the router and the switch:

- There are no domain ID conflicts.
- At least one switch in the fabric can be the principal.
- The interop mode is the same on the router R_Port as for the fabric.
- The R_A_TOV is the same on the router R_Port as for the fabric.
- The E_D_TOV is the same on the router R_Port as for the fabric.

Note: The ED_TOV and RA_TOV can be set individually for each R_Port.

- If you are using SANtegrity Fabric Binding, the fabric membership list has been updated on all devices, including the router.
- If you are using SANtegrity Switch Binding for E_Ports, the switch membership list has been updated on both the router and the switch to which it is attaching.
- The port is unblocked on the fabric switch.
- The **Enable port** check box is selected on the router port.
- The fibre is OK by testing it between two known good unblocked ports on a switch.
- You remembered to perform a Reset System on the router after configuring the R_Port.

17.8 iFCP between two SAN04M-R routers

We now describe how to connect two SAN04M-R routers across a TCP/IP network using iFCP. This is intended for long-distance connectivity over a wide area network, but in this example, we connect over a 100 Mbps Ethernet LAN. The SAN04M-R would normally be connected using its optical Gigabit Ethernet SFP connector for optimal performance.

We assume that the RJ45 (Fast Ethernet) connector of port 3 on each of the two routers shown in 17.5, “Example configuration” on page 479 is connected to a TCP/IP network such that they are able to communicate with each other.

17.8.1 Configuring the iFCP port

To configure the iFCP port:

1. From Element Manager, select **Configuration** → **Port** → **FC/Ethernet**:
   a. Select port **3** from the Port number drop-down list at the top of the window (Figure 17-59).

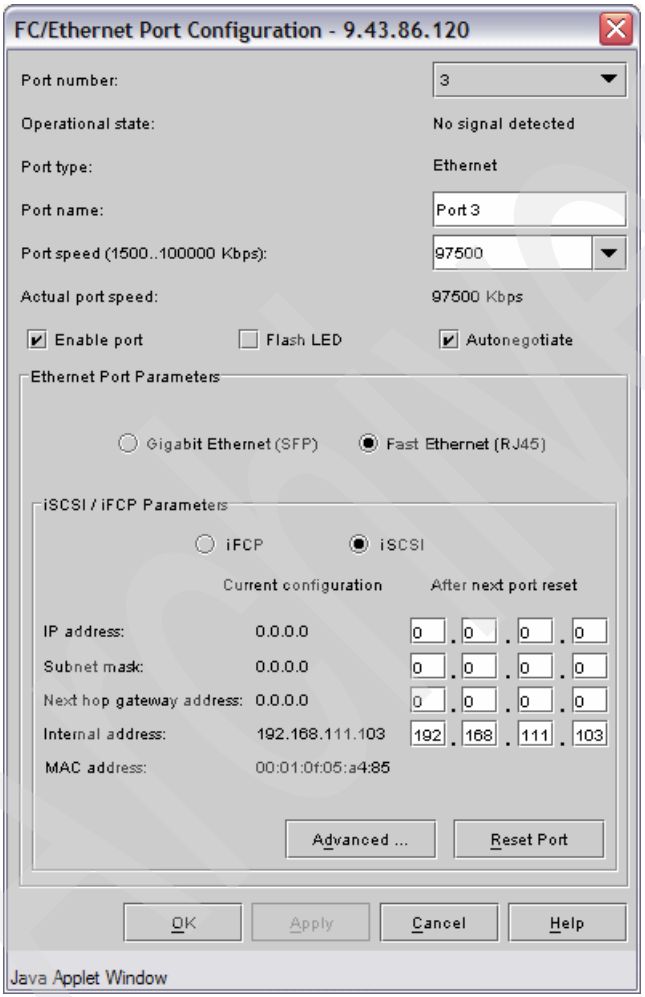


Figure 17-59 Configure port 3 for iFCP

b. Enter a description for the Port name, verify that the Port speed is appropriate for the connection, and ensure that the **Enable port** check box is selected. If appropriate for your network hardware, select **Autonegotiate**. Select **Fast Ethernet (RJ45)** and **iFCP** (Figure 17-60).

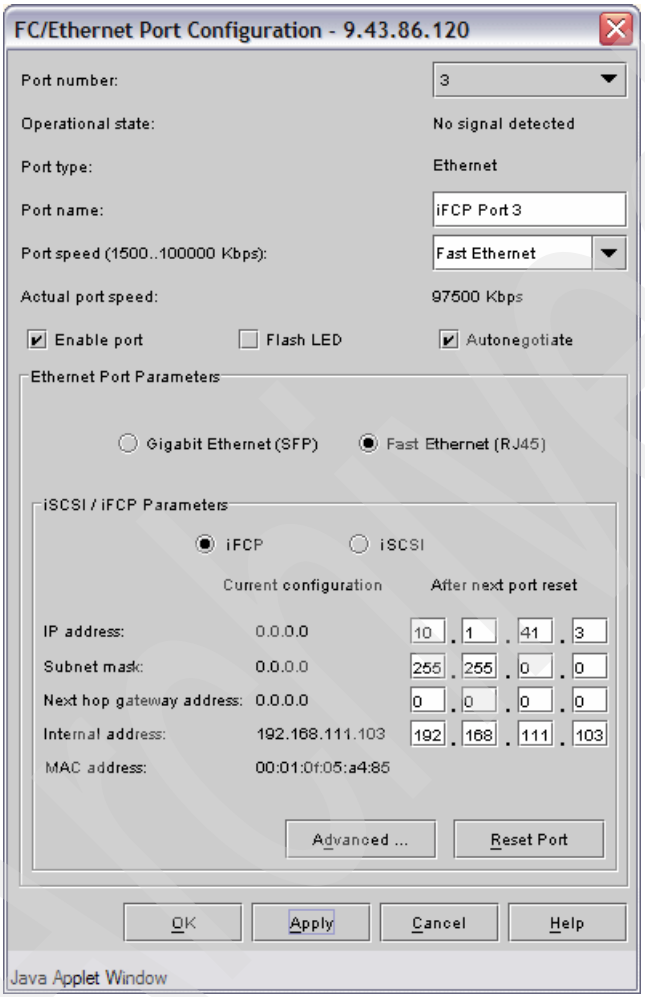


Figure 17-60 Port 3 iFCP details

c. Enter the external IP address and Subnet mask, and if the remote router is in a different subnet, enter a network gateway address. Verify that the internal port address is correct (refer to 17.5, “Example configuration” on page 479 for the values we are using).

d. Now click the **Advanced** button. The window shown in Figure 17-61 on page 503 opens.

2. On the TCP tab, ensure that the **Auto reset port on severe errors** check box is selected.

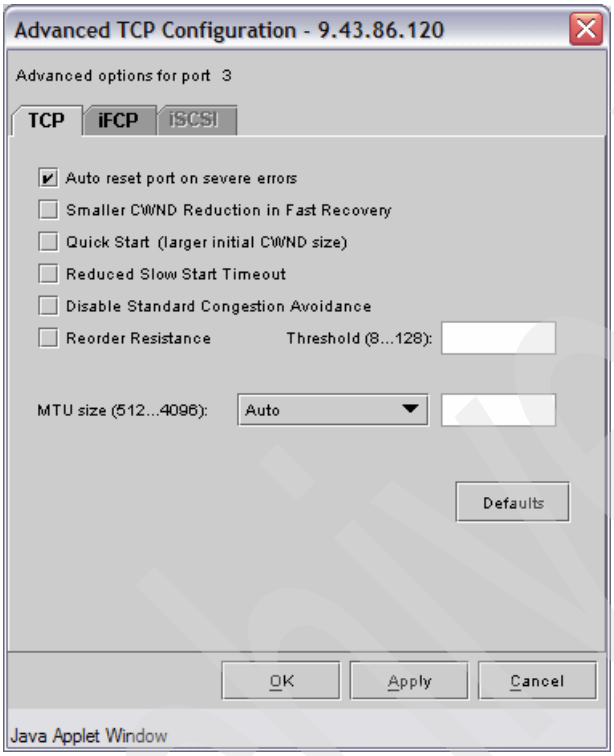


Figure 17-61 Advanced TCP configuration tab

3. Select the **iFCP** tab. The choice of Compression Level depends on the characteristics of the network being used, and whether the optional compression feature is installed. Because we are using a slow network link, compression is strongly recommended. Also, because we have support at both ends of the iFCP link for hardware compression, that is the best option to use because it is more efficient than software compression. Therefore, we select **HW** from the drop-down list shown in Figure 17-62. Refer to the *McDATA SAN Router Administration and Configuration Manual*, 620-000206, for a description of the compression options and the other advanced options.

Note: The aim is to configure the iFCP link to try and support the data rate of the router FC ports. Therefore, compression might not be required for a Gigabit network link, but most likely will be for a 100 Mbps network link.

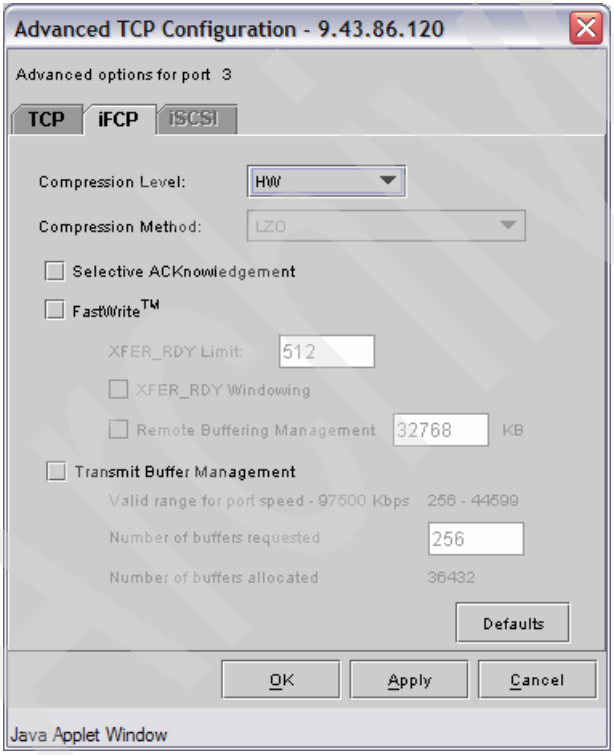


Figure 17-62 Advanced iFCP configuration tab

Click **OK** to return to the FC/Ethernet Port Configuration panel. The window should now look like the one shown in Figure 17-63.

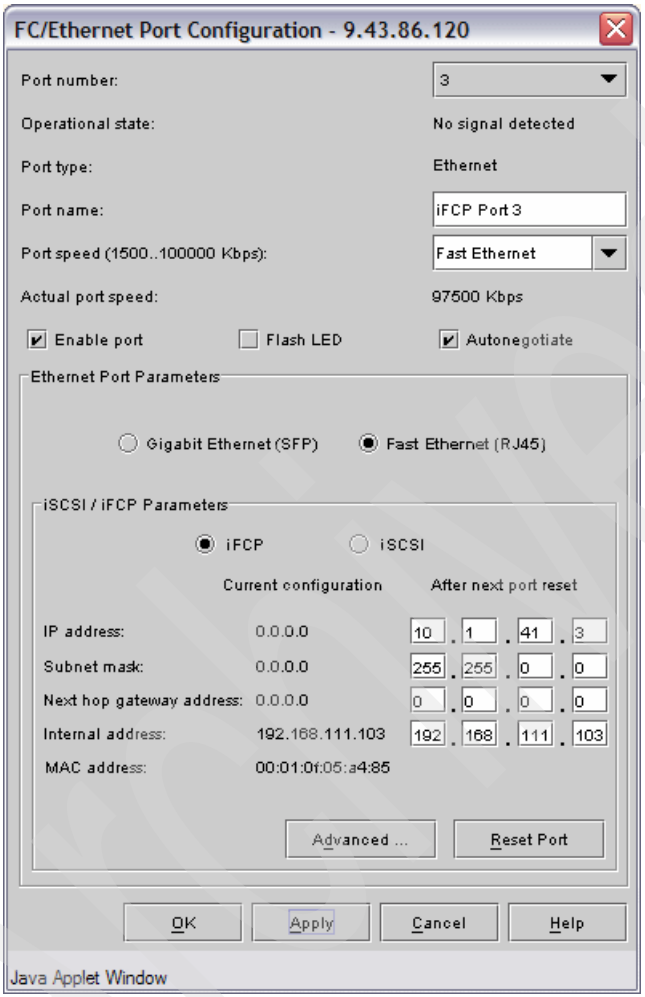


Figure 17-63 Completed iFCP Port Configuration

4. Click **OK** and the warning in Figure 17-64 appears. Click **Yes** to proceed.

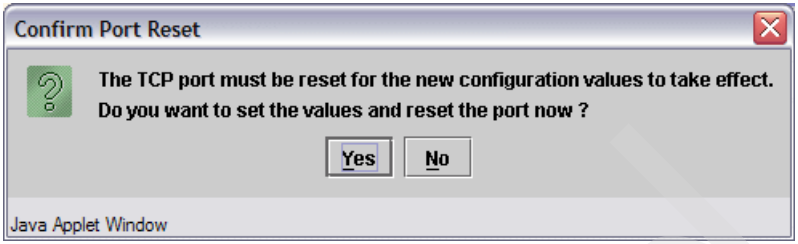


Figure 17-64 Port change warning

We save these changes to flash memory after the next step.

17.8.2 Configuring the iFCP connection

Having configured a TCP/IP port, we now need to configure the iFCP connection, including a unique mSAN ID for each router. The mSAN ID uniquely identifies the mSAN containing the router.

Setting the local mSAN ID

From Element Manager, select **Configuration** → **iFCP** → **Setup** and enter a number in the range 0 to 4,294,967,295. We use 40, as shown in Figure 17-65. Click **OK**.

Figure 17-65 Setting the mSAN ID

The warning shown in Figure 17-66 appears. Click **OK** to continue.

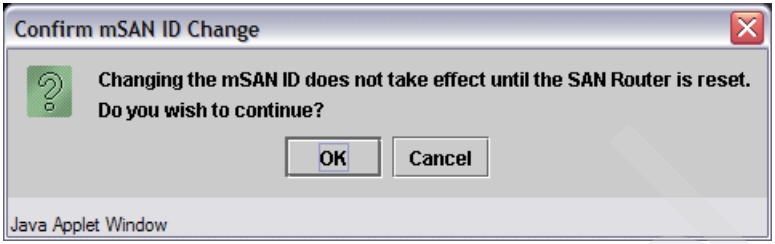


Figure 17-66 Changing mSAN ID warning

Defining the remote router

Now select **Configuration** → **iFCP** → **Remote Connections** to define the remote SAN router in the window shown in Figure 17-67.

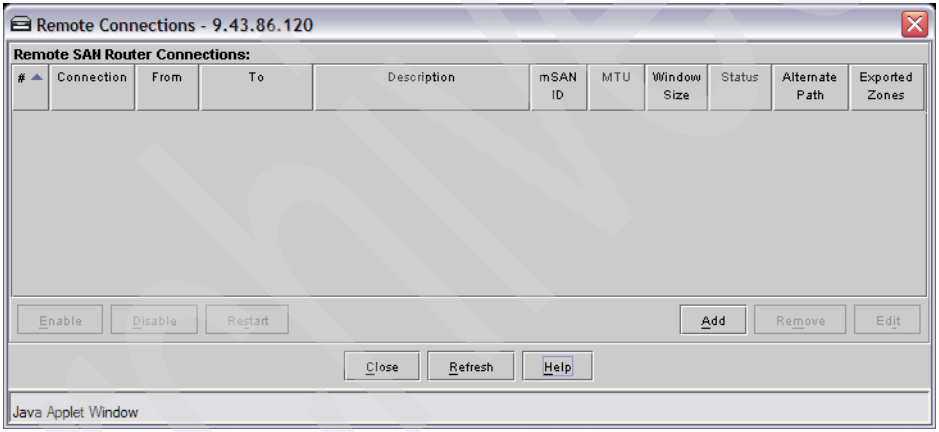


Figure 17-67 Remote Connections

Click the **Add** button to add a new connection. Enter a Connection description, select **Port 3**, and enter the TCP/IP address of the remote router port (Figure 17-68). Check that the Connection state is **Enabled**. Click **OK**.

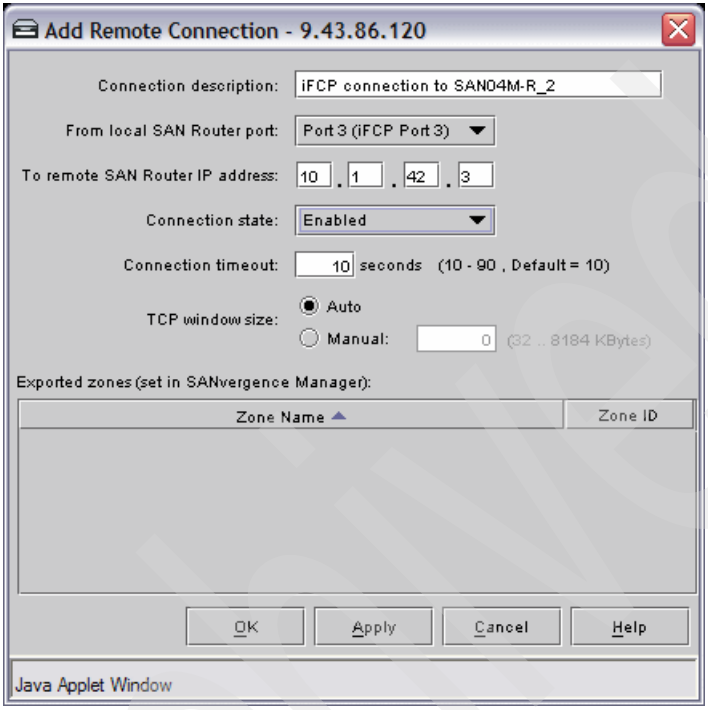


Figure 17-68 Defining remote connection for port 3

Figure 17-69 shows the new remote connection definition. Click **Close** to finish.

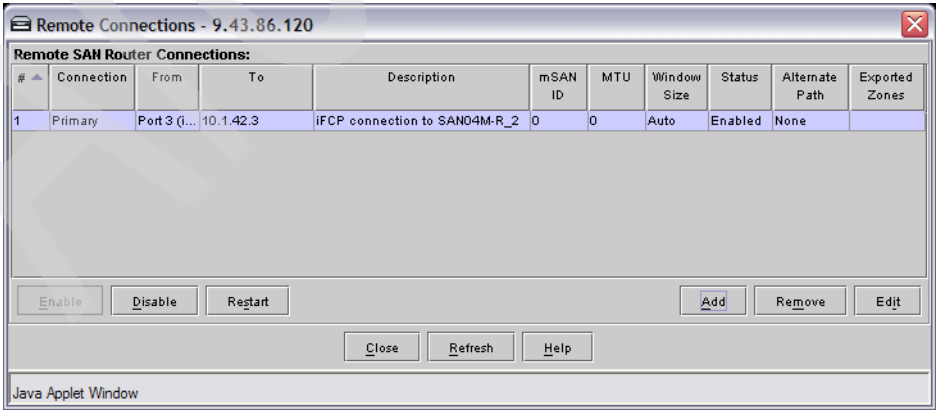


Figure 17-69 Remote connection defined on port 3

Redundant connection

It is possible to define a redundant failover connection for the iFCP route using the other IP port on the router. If you want to do this, the other port must also be configured as an iFCP port with a route to the remote router. The active workload can be balanced across these ports by exporting different zones on the two ports. Then, select **Configuration** → **iFCP** → **Port Redundancy** to symmetrically define both ports as the *backup* for each other.

17.8.3 Saving the configuration and resetting

Select **File** → **Save Configuration** to save all of the changes to flash memory. Once complete, select **File** → **Reset System** to activate the changes.

17.8.4 Configuring the remote router

Now repeat the previous configuration steps, starting at 17.8.1, “Configuring the iFCP port” on page 501, on the remote router so that it has a defined iFCP port and connection to the local router.

17.8.5 Testing the connection

The SAN router provides a ping utility to test connectivity through the IP ports. Select **Statistics/Info** → **Ping** to open the Network Utilities window shown in Figure 17-70.

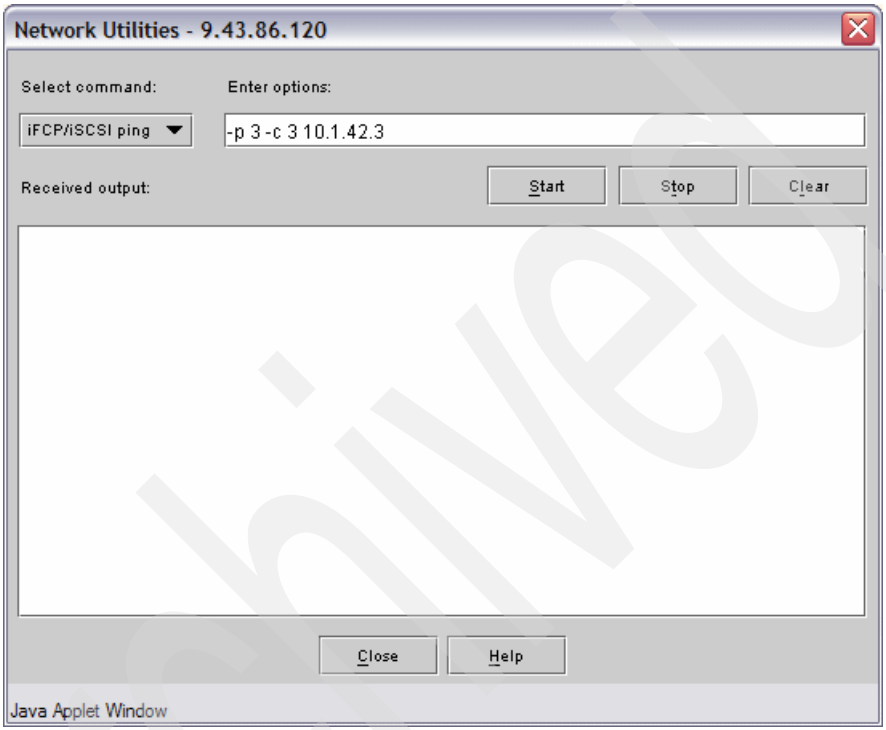


Figure 17-70 Network Utilities ping

Select **iFCP/iSCSI ping** from the Select command list, enter the following options, and click **Start**:

```
-p 3-c 3 10.1.42.3
```

Where:

- `-p 3` means send ping from port four.
- `-c 3` means send three pings.
- `10.1.42.3` is the remote router port external IP address.

The following window shows a successful ping test (Figure 17-71). Click **Close** when finished.

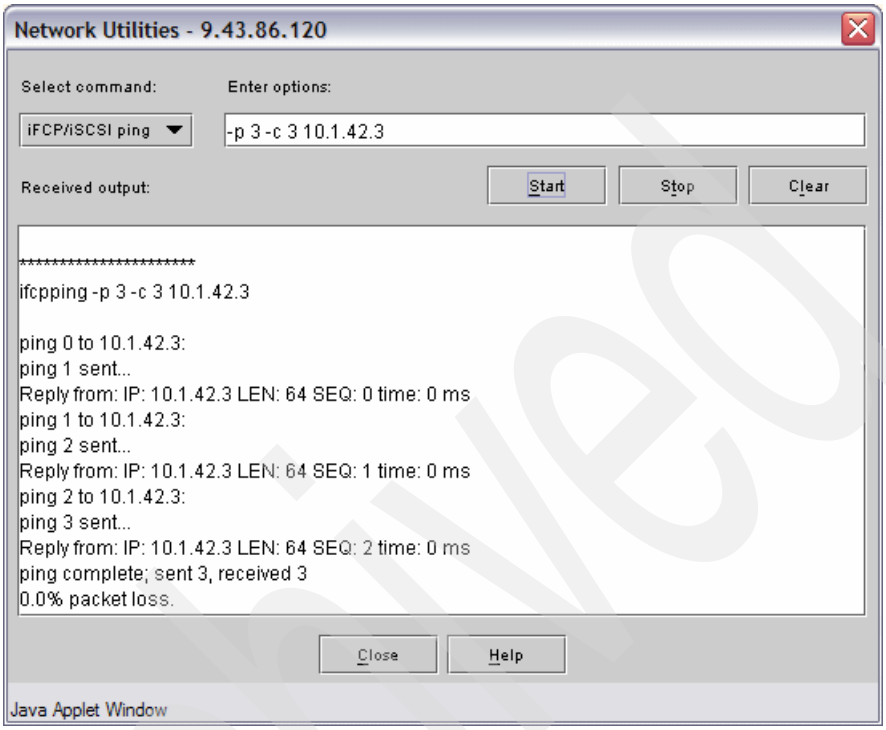


Figure 17-71 Successful iFCP ping

If you want, you can repeat the test from the remote router targeting the local router.

17.8.6 Adding remote mSAN to SANvergence Manager

If the remote router has not already been defined to the SANvergence Manager, add it now. Launch SANvergence Manager and log in. Select **Actions** → **Add mSAN** and enter the IP address of the remote router's management port and its modify password. Enter an appropriate mSAN Name and confirm that the **Out-of-band (10/100 port)** radio button is selected, as shown in Figure 17-72 on page 512. Click **OK** to finish.

Note: Notice that SANvergence Manager has remembered the local mSAN login details.

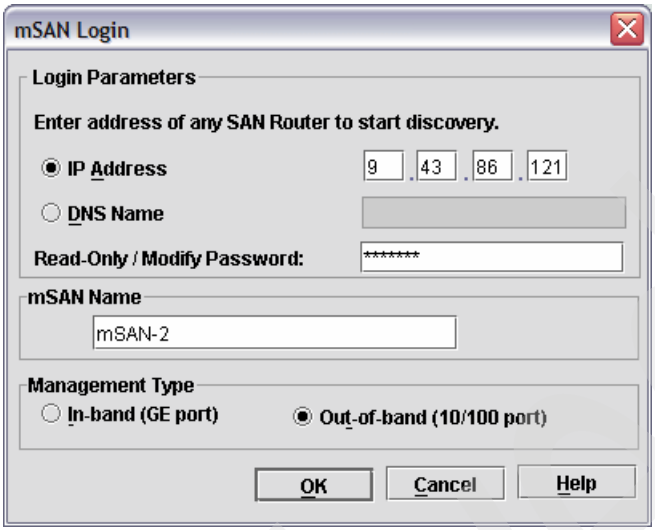


Figure 17-72 Adding an mSAN to SANvergence Manager

The SANvergence Manager window should now show two mSANs, as shown in Figure 17-73.

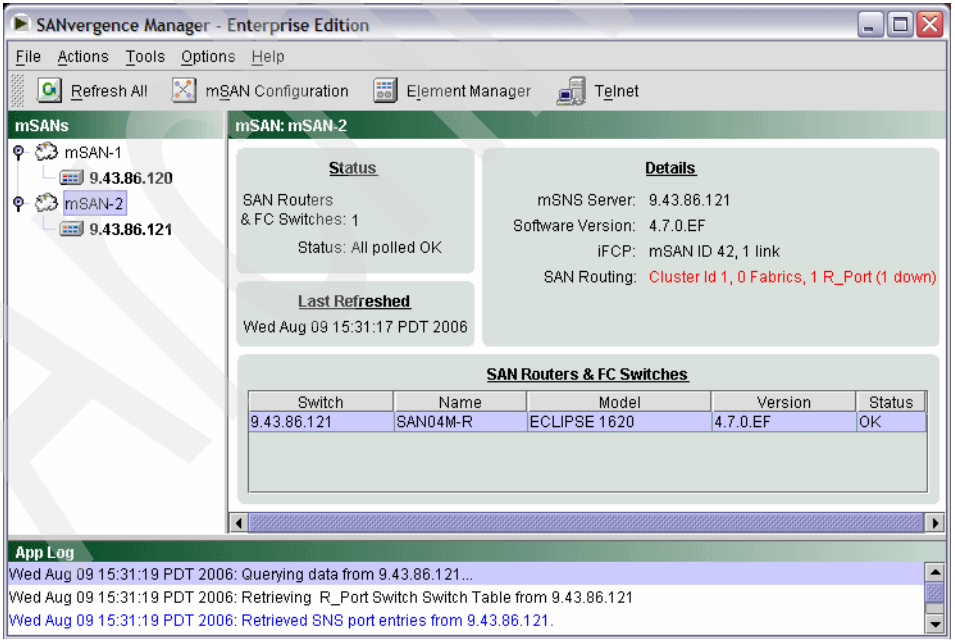


Figure 17-73 Two mSANs in SANvergence Manager

In our example, the FC-attached host is directly connected to port 1 of the router on the right, and thus does not need to be selectively imported. If you open the mSAN Configuration window for mSAN-2, the host appears in the Devices pane, as shown in Figure 17-74.

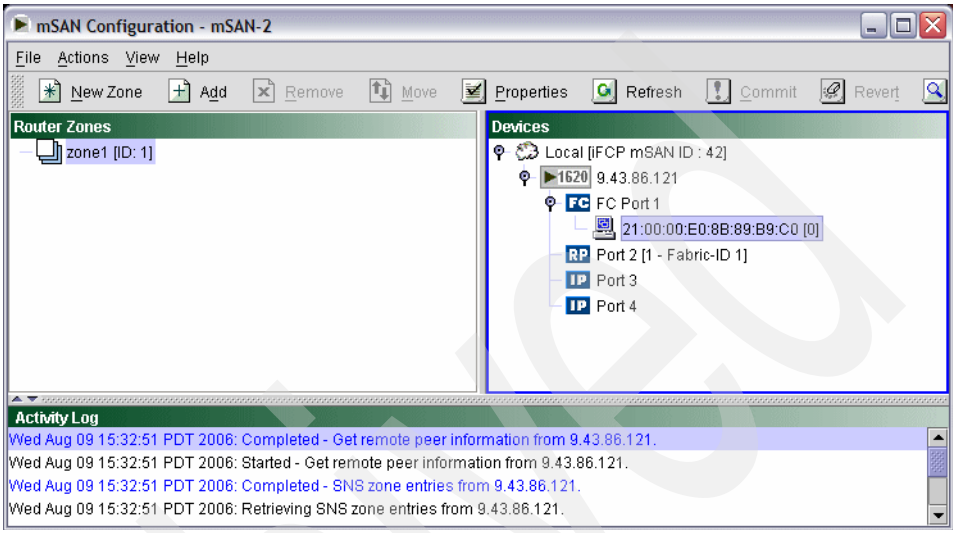


Figure 17-74 Host directly attached to port 1 in mSAN Configuration window

Double-clicking the device opens the Properties window (Figure 17-75), where you can assign a port label to appear in place of the WWN highlighted in Figure 17-74 on page 513. When finished, close the mSAN Configuration window.

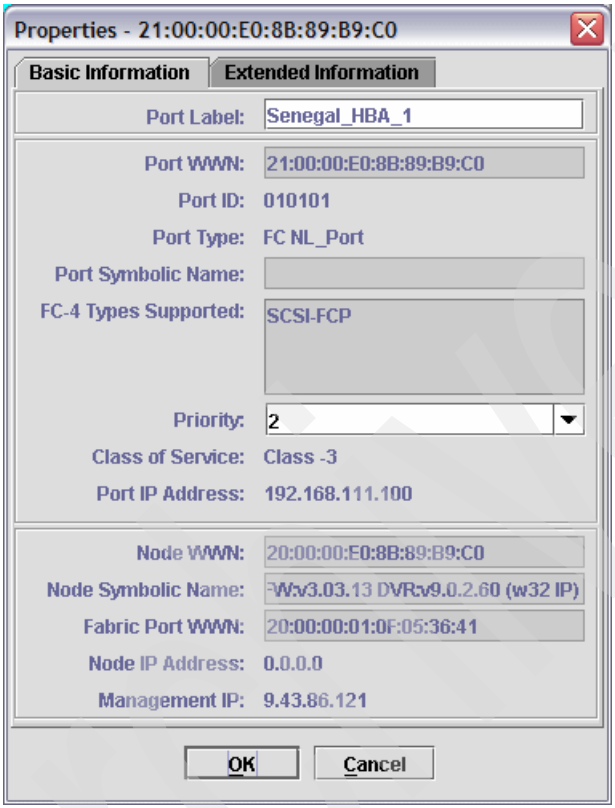


Figure 17-75 Assigning a Port Label

We are now ready to zone devices between the two mSANs.

17.8.7 EFCM view

If you are running the EFCM server to manage your m-type fabric, add the routers to the discovery list for EFCM and their details will appear. Select **Discover** → **Setup** and the window shown in Figure 17-76 opens.

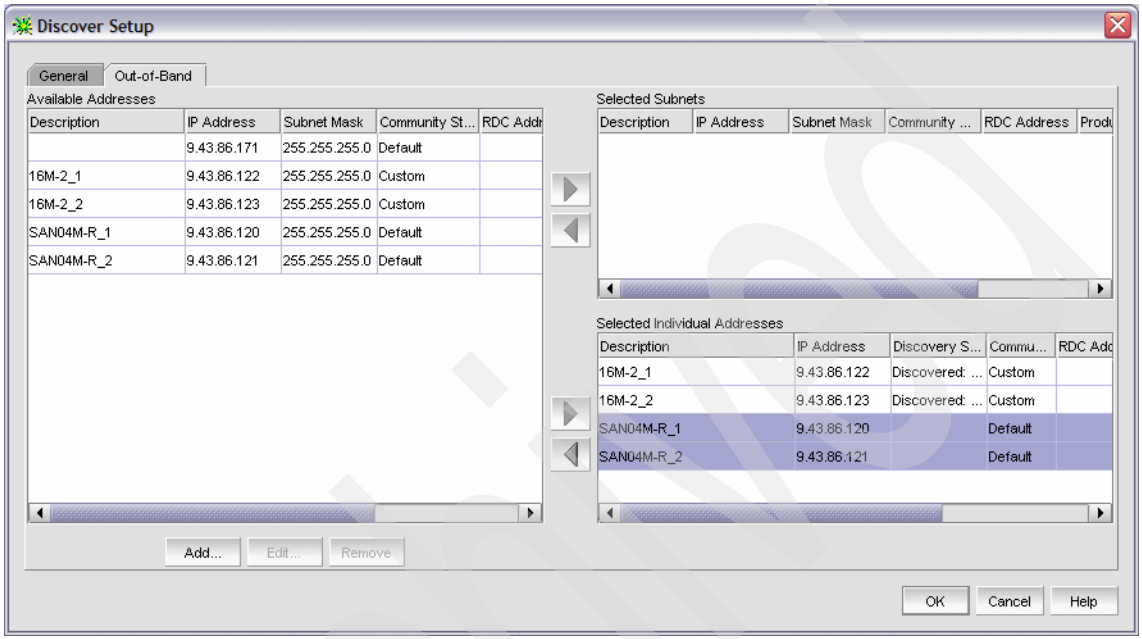


Figure 17-76 EFCM: Discover Setup

In our case, we added two SAN04M-R routers to the discovery list. Figure 17-77 shows our two routers, discovered by the EFCM.

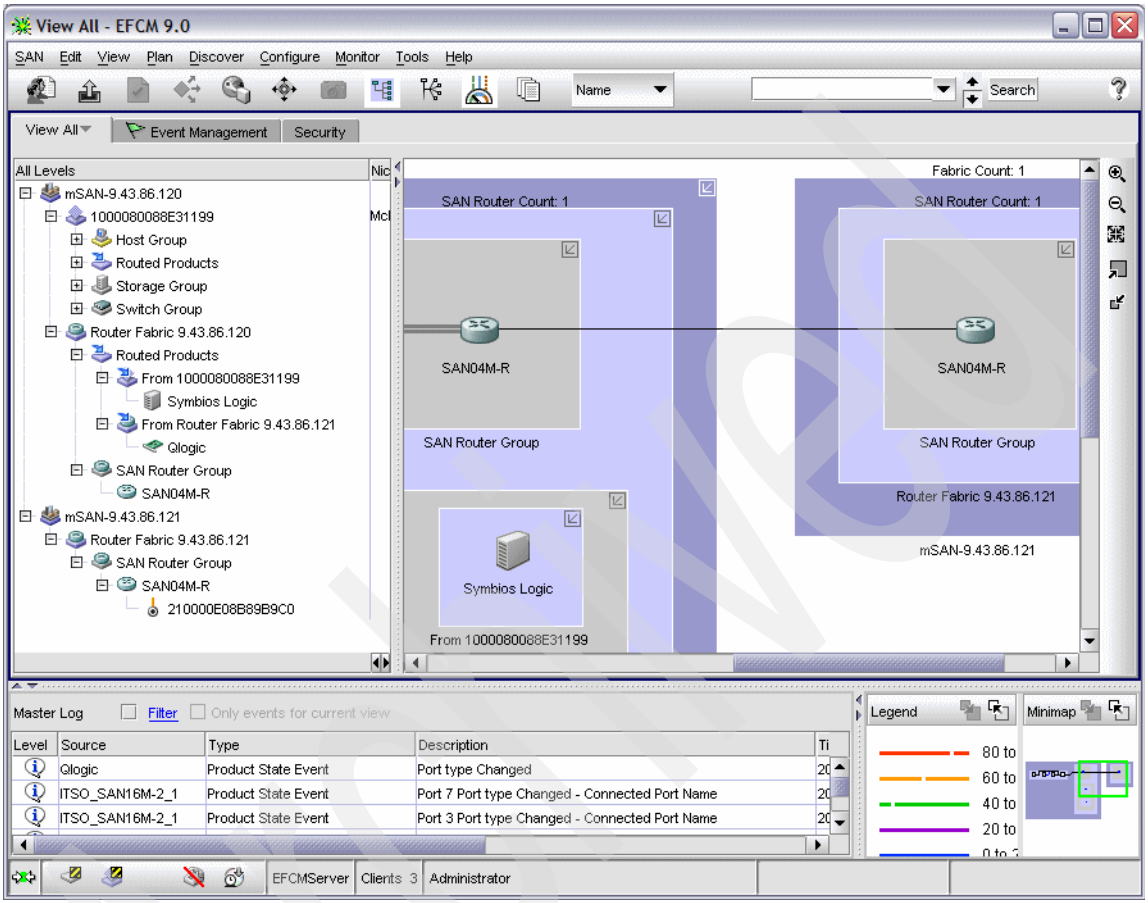


Figure 17-77 EFCM showing two SAN routers

EFCM 9.0 offers good support for SAN router configuration. By right-clicking the SAN router icon, you can access numerous actions, including:

- Element Manager
- Zoning
- Router port configuration
- SAN router configuration
- SANvergence Manager
- Telnet

We show the list of available actions in Figure 17-78.

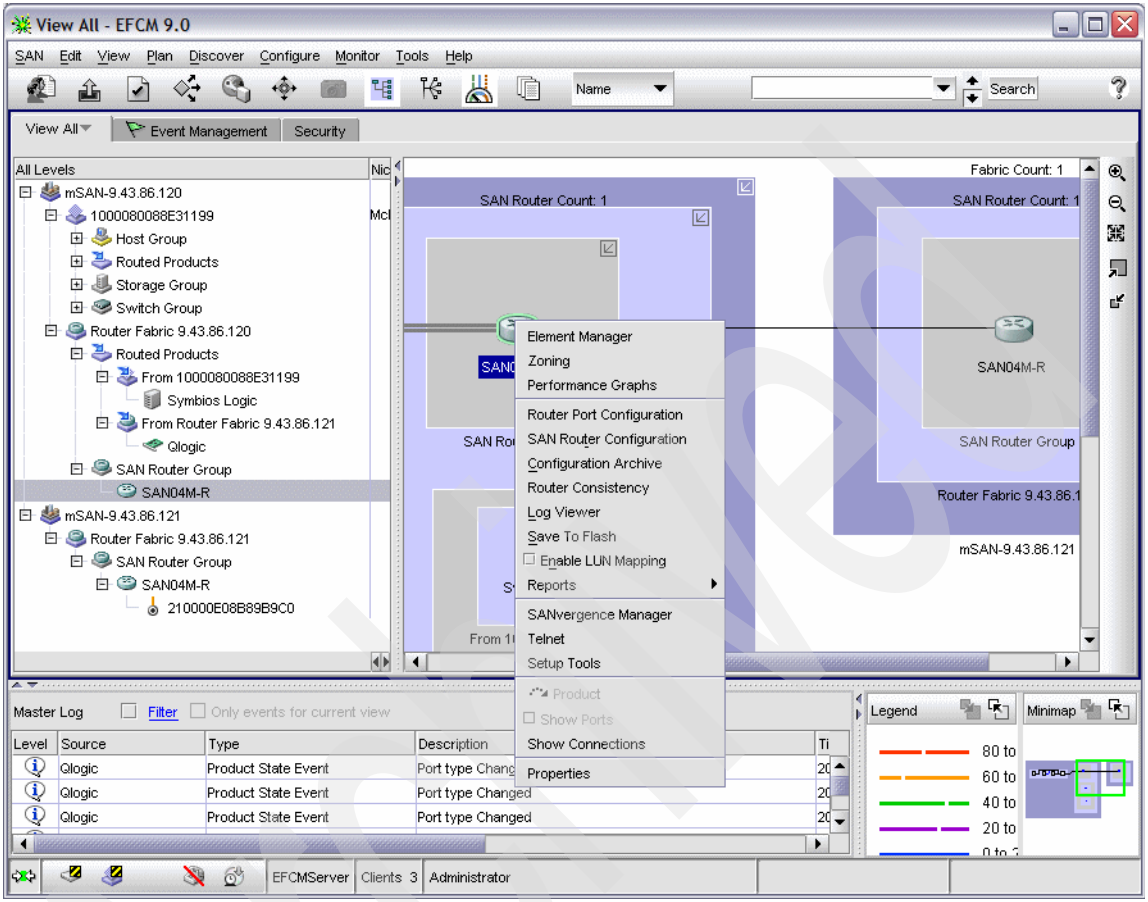


Figure 17-78 EFCM: SAN router tasks

Note: EFCM 9.0 enables you to configure all aspects of the SAN routing solution. In our example, we use stand-alone SANvergence Manager to configure the mSAN, but we can do the same from the EFCM 9.0 as well.

17.9 Zoning across iFCP

To zone devices across an iFCP link between two SAN routers, you need to create a zone on each side using a matching zone ID, and then export the zone from each router.

Using SANvergence Manager, perform the following steps:

1. Select the first mSAN, in our example, **mSAN-1**, and click **mSAN Configuration**.
   a. Create a new zone, as shown in Figure 17-79:
      i. Click **New Zone** and enter a Zone Name of `Remote_zone`.
      ii. Make a note of the Zone ID (in this case 2).
      iii. Click **OK**.

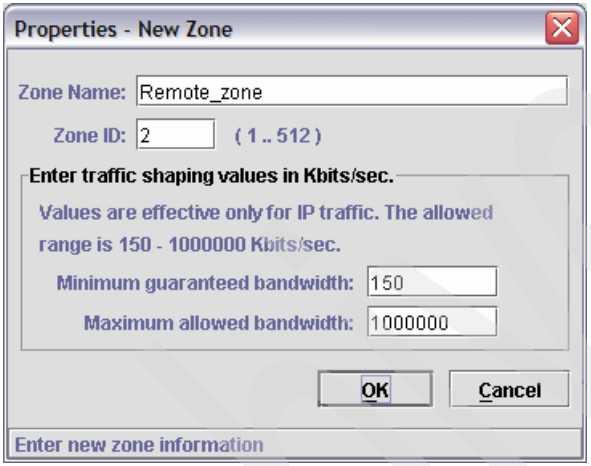


Figure 17-79 Create New Zone for mSAN-1

b. Add the storage device to the zone, as shown in Figure 17-80:

   i. Highlight **Remote_zone** in the Router Zones pane.

   ii. Expand the device tree in the Devices pane and highlight the **DS4400** port.

   iii. Right-click the storage device port and select **Add Device to Remote_zone**.

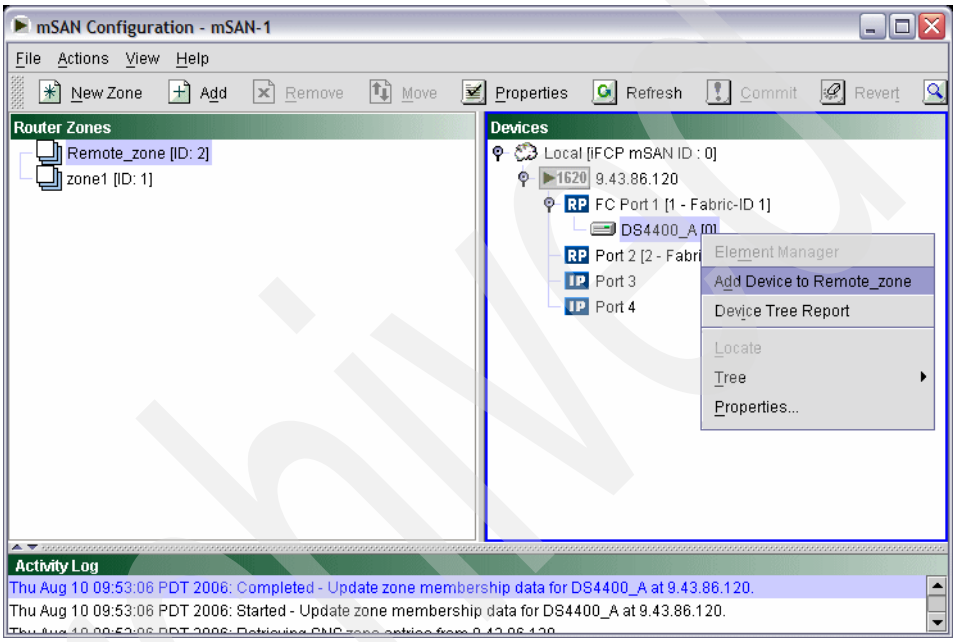


Figure 17-80 Add storage device to remote zone

c. Commit the change:

   i. Click **Commit**.

   ii. Respond **No** to the prompt to save the change to flash memory, because we will do that after exporting the new zone.

d. Export the new zone to the other router, as shown in Figure 17-81:

   i. Select **Actions** → **Export Zones**.

   ii. Select the **Remote_zone** (Zone ID 2) option.

   iii. Click **OK**.

   iv. Respond **Yes** to save the update to flash memory prompt.

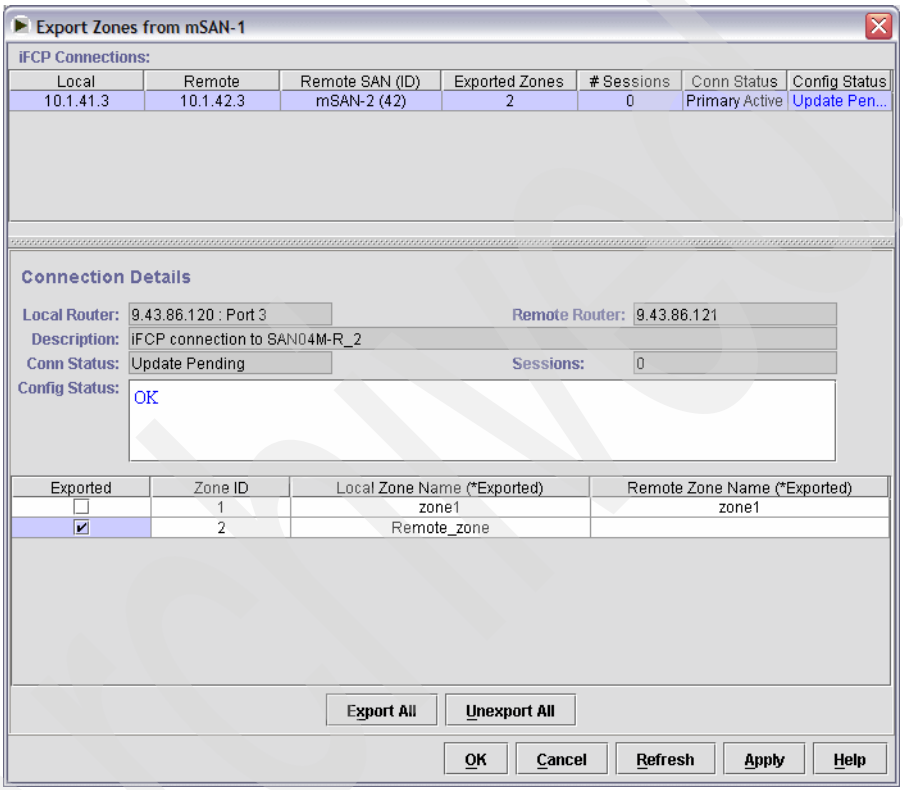


Figure 17-81 Exporting Remote_zone from mSAN-1

e. Leave the mSAN Configuration window open.

2. Now select the second mSAN, in our example, **mSAN-2**, and click **mSAN Configuration**:
   a. Create a new zone, as shown in Figure 17-82:
      i. Click **New Zone** and enter a Zone Name of `Remote_zone`.
      ii. Enter the Zone ID as noted earlier (in this case `2`).

      Important: The zone IDs *must* match, but the zone names do not have to.

      iii. Click **OK**.

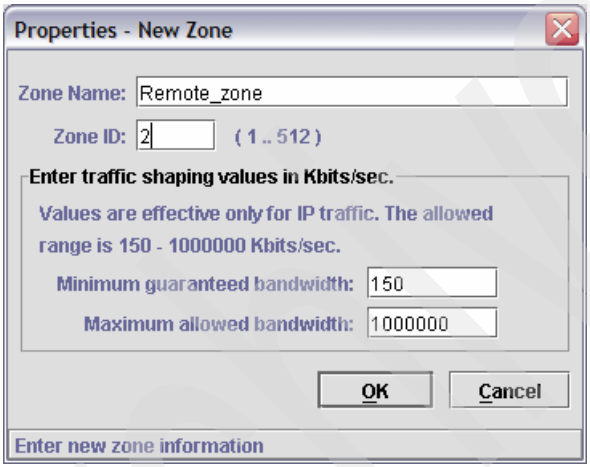


Figure 17-82 Create New Zone for mSAN-2

b. Add the host's HBA to the zone, as shown in Figure 17-83:

   i. Highlight **Remote_zone** in the Router Zones pane.

   ii. Expand the device tree in the Devices pane and highlight the host port (in our case **Senegal_HBA_1**).

   iii. Right-click the host port and select **Add Device to Remote_zone**.

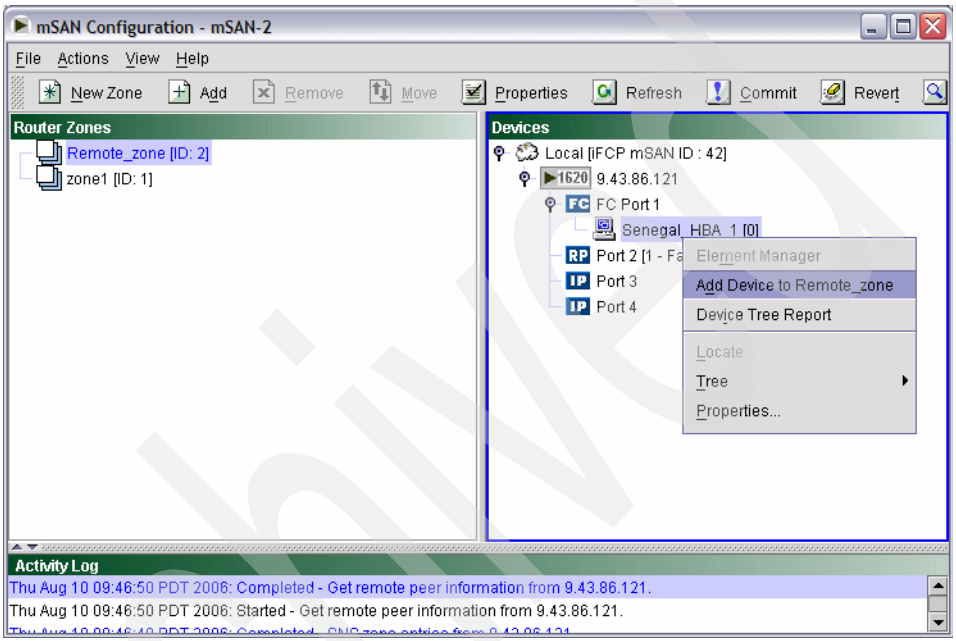


Figure 17-83 Add host's HBA to remote zone

c. Commit the change:

   i. Click **Commit**.

   ii. Respond **No** to the prompt to save the change to flash memory, because we will do that after exporting the new zone.

d. Export the new zone to the other router, as shown in Figure 17-84:

   i. Select **Actions** → **Export Zones**.

   ii. Select the **Remote_zone** (Zone ID 2) option.

   iii. Click **OK**.

   iv. Respond **Yes** to save the update to flash memory prompt.

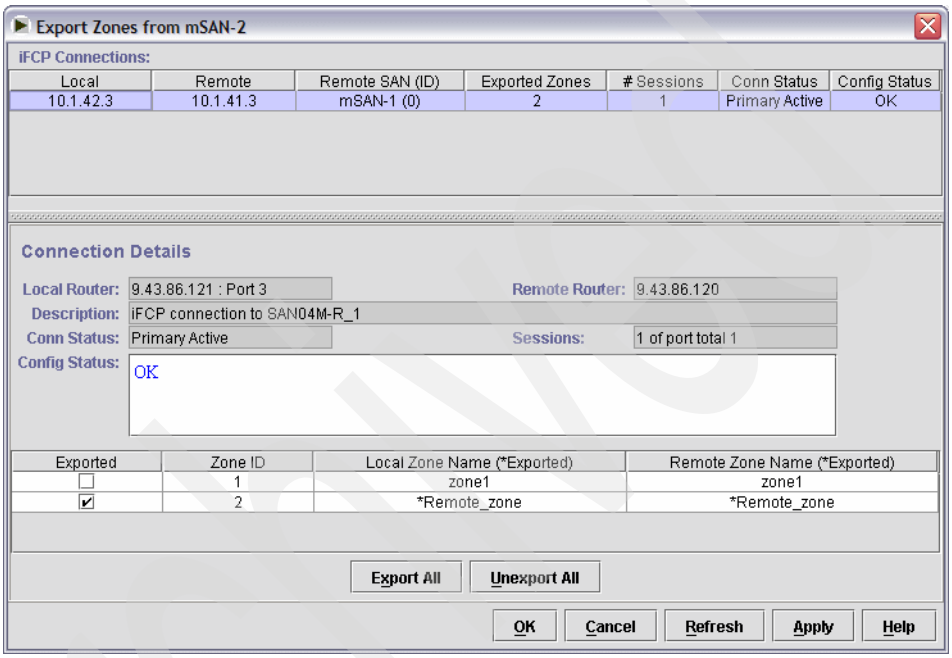


Figure 17-84 Exporting Remote_zone from mSAN-2

e. Leave the mSAN Configuration window open.

3. Click **Refresh** on the mSAN Configuration window for mSAN-1 and expand the trees in both panels. It should look like the window shown in Figure 17-85. Note the blue R next to the remote host.

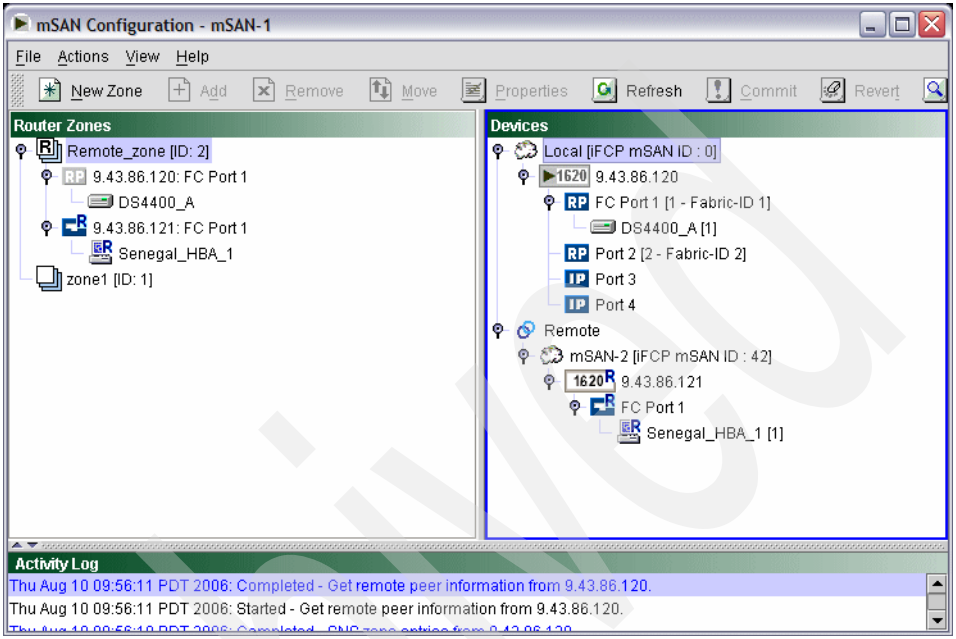


Figure 17-85 mSAN-1 showing remote host

4. Click **Refresh** on the mSAN Configuration window for the mSAN-42 and expand the trees in both panels. It should look like the window shown in Figure 17-86. Note the blue R next to the remote storage server.

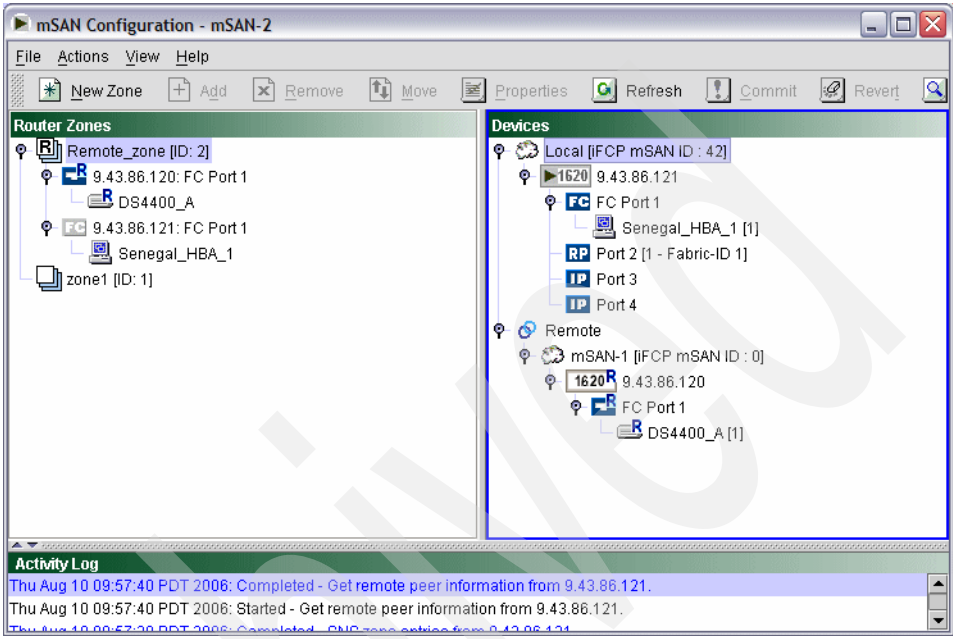


Figure 17-86 mSAN-2 showing remote storage device

You have now zoned the remote host to the local storage server.

Close the two mSAN Configuration windows.

Remote storage and Windows server

On the host, verify that the remote storage can be accessed. In our case, we used a Microsoft Windows Server 2003 host. We defined two logical drives on the DS4400:

- JA_LUN_1, 6 GB
- JA_LUN_2, 8 GB

Then, we mapped the two logical drives to our Windows host (named SENEGAL), as shown in Figure 17-87.

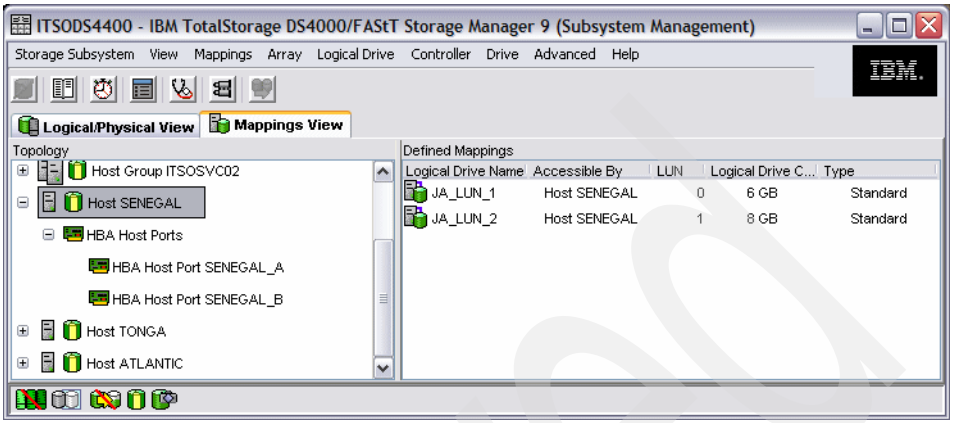


Figure 17-87 Logical drives mapped to the Windows server

Figure 17-88 shows the two logical drives as seen in the Disk Management applet on the Windows server. The partitions on the drives are being formatted.

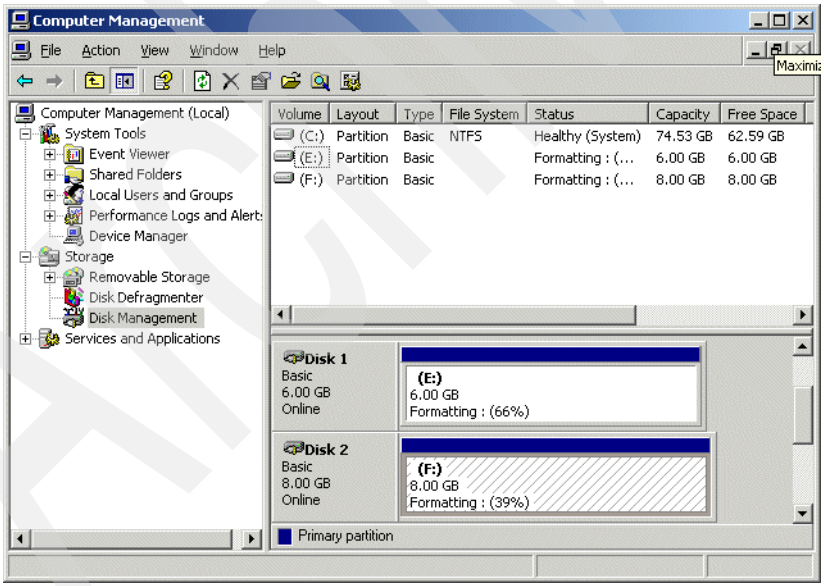


Figure 17-88 Formatting the logical drives

When the formatting completes, you can start using the two logical drives.

Glossary

8b/10b A data encoding scheme developed by IBM, translating byte-wide data to an encoded 10-bit format. The Fibre Channel (FC) FC-1 level defines this as the method to use to encode and decode data transmissions over the Fibre Channel.

active configuration In an ESCON® environment, the ESCON Director configuration determined by the status of the current set of connectivity attributes. Contrast with *saved configuration*.

adapter A hardware unit that aggregates other input/output (I/O) units, devices, or communications links to a system bus.

ADSM ADSTAR Distributed Storage Manager.

Advanced Intelligent Tape (AIT) A magnetic tape format by Sony that uses 8 mm cassettes, but is only used in specific drives.

agent In the client-server model, the part of the system that performs information preparation and exchange on behalf of a client or server application. In the Simple Network Management Protocol (SNMP), the managed system. See also *management agent*.

aggregation In the Storage Networking Industry Association Storage Model (SNIA), *virtualization* is known as *aggregation*. This aggregation can take place at the file level or at the level of individual blocks that are transferred to disk.

AIT See *Advanced Intelligent Tape*.

AL See *arbitrated loop*.

allowed In an ESCON Director, the attribute that, when set, establishes dynamic connectivity capability. Contrast with *prohibited*.

AL_PA Arbitrated Loop Physical Address.

American National Standards Institute (ANSI) The primary organization for fostering the development of technology standards in the United States. The ANSI family of Fibre Channel documents provides the standards basis for the Fibre Channel architecture and technology. See also *FC-PH*.

ANSI See *American National Standards Institute*.

APAR See *authorized program analysis report*.

arbitrated loop (AL) A Fibre Channel interconnection technology that allows up to 126 participating node ports and one participating fabric port to communicate.

arbitration The process of selecting one respondent from a collection of several candidates that request service concurrently.

Asynchronous Transfer Mode (ATM) A type of packet switching that transmits fixed-length units of data.

ATL See *Automated Tape Library*.

ATM See *Asynchronous Transfer Mode*.

authorized program analysis report (APAR) A report of a problem caused by a suspected defect in a current, unaltered release of a program.

Automated Tape Library (ATL) Large scale tape storage system, which uses multiple tape drives and mechanisms to address 50 or more cassettes.

backup A copy of computer data, or the act of copying such data, that is used to re-create data that has been lost, mislaid, corrupted, or erased.

bandwidth A measure of the information capacity of a transmission channel.

basic mode An S/390® or IBM eServer zSeries® central processing mode that does not use logical partitioning. Contrast with *logically partitioned mode*.

blocked In an ESCON and FICON Director, the attribute that, when set, removes the communication capability of a specific port. Contrast with *unblocked*.

bridge A component used to attach more than one I/O unit to a port. Also a data communications device that connects two or more networks and forwards packets between them. The bridge can use similar or dissimilar media and signaling systems. It operates at the data link level of the OSI model. Bridges read and filter data packets and frames.

bridge/router A device that can provide the functions of a bridge, router, or both concurrently. A bridge/router can route one or more protocols, such as TCP/IP, and bridge all other traffic. See also *bridge* and *router*.

broadcast To send a transmission to all N_Ports on a fabric.

byte 1) In Fibre Channel, an eight-bit entity prior to encoding or after decoding, with its least significant bit denoted as bit 0 and most significant bit as bit 7. The most significant bit is shown on the left side in FC-FS unless otherwise shown. 2) In S/390 architecture or z/Architecture™ for zSeries (and FICON), an eight-bit entity prior to encoding or after decoding, with its least significant bit denoted as bit 7 and most significant bit as bit 0. The most significant bit is shown on the left side in S/390 architecture and z/Architecture for zSeries.

cascaded switches The connecting of one Fibre Channel switch to another Fibre Channel switch, creating a cascaded switch route between two N_Nodes connected to a Fibre Channel fabric.

chained In an ESCON environment, pertaining to the physical attachment of two ESCON Directors (ESCDs) to each other.

channel 1) A processor system element that controls one channel path, whose mode of operation depends on the type of hardware to which it is attached. In a channel subsystem, each channel controls an I/O interface between the channel control element and the logically attached control units. 2) In ESA/390 or z/Architecture, the part of a channel subsystem that manages a single I/O interface between a channel subsystem and a set of controllers (control units).

channel to channel See *CTC*.

channel to converter See *CVC*.

channel-attached Devices attached directly by data channels (I/O channels) to a computer. Also refers to devices attached to a controlling unit by cables rather than by telecommunication lines.

channel I/O A form of I/O where request and response correlation is maintained through a form of source, destination, and request identification.

channel path (CHP) A single interface between a central processor and one or more control units along which signals and data can be sent to perform I/O requests.

channel path identifier (CHPID) In a channel subsystem, a value assigned to each installed channel path of the system that uniquely identifies that path to the system.

channel subsystem (CSS) Relieves the processor of direct I/O communication tasks, and performs path management functions. Uses a collection of subchannels to direct a channel to control the flow of information between I/O devices and main storage.

CHP See *channel path*.

CHPID See *channel path identifier*.

CIFS Common Internet File System.

cladding In an optical cable, the region of low refractive index surrounding the core. See also *core* and *optical fiber.*

Class of Service A Fibre Channel frame delivery scheme that exhibits a specified set of delivery characteristics and attributes.

Class-1 A class of service that provides a dedicated connection between two ports with confirmed delivery or notification of nondeliverability.

Class-2 A class of service that provides a frame switching service between two ports with confirmed delivery or notification of nondeliverability.

Class-3 A class of service that provides a frame switching datagram service between two ports or a multicast service between a multicast originator and one or more multicast recipients.

Class-4 A class of service that provides a fractional bandwidth virtual circuit between two ports with confirmed delivery or notification of nondeliverability.

Class-6 A class of service that provides a multicast connection between a multicast originator and one or more multicast recipients with confirmed delivery or notification of nondeliverability.

client A software program used to contact and obtain data from a *server* software program on another computer, often across a great distance. Each *client* program is designed to work specifically with one or more kinds of server programs, and each server requires a specific kind of client program.

client/server The relationship between machines in a communications network. The client is the requesting machine, and the server is the supplying machine. Also used to describe the information management relationship between software components in a processing system.

cluster A type of parallel or distributed system that consists of a collection of interconnected whole computers and is used as a single, unified computing resource.

CNC A mnemonic for an ESCON channel used to communicate to an ESCON-capable device.

coaxial cable A transmission media (cable) used for high-speed transmission. It is called *coaxial* because it includes one physical channel that carries the signal surrounded (after a layer of insulation) by another concentric physical channel, both of which run along the same axis. The inner channel carries the signal and the outer channel serves as a ground.

configuration matrix In an ESCON environment or FICON, an array of connectivity attributes that appear as rows and columns on a display device and can be used to determine or change active and saved ESCON or FICON Director configurations.

connected In an ESCON Director, the attribute that, when set, establishes a dedicated connection between two ESCON ports. Contrast with *disconnected.*

connection In an ESCON Director, an association established between two ports that provides a physical communication path between them.

connectivity attribute In an ESCON and FICON Director, the characteristic that determines a particular element of a port's status. See *allowed, prohibited, blocked, unblocked, as well as connected* and *disconnected*.

control unit A hardware unit that controls the reading, writing, or displaying of data at one or more I/O units.

controller A component that attaches to the system topology through a channel semantic protocol that includes some form of request/response identification.

core In an optical cable, the central region of an optical fiber through which light is transmitted and that has an index of refraction greater than the surrounding cladding material. See also *cladding* and *optical fiber*.

coupler In an ESCON environment, link hardware used to join optical fiber connectors of the same type. Contrast with *adapter*.

CRC See *Cyclic Redundancy Check*.

CSS See *channel subsystem*.

CTC Channel-to-channel. A mnemonic for an ESCON channel attached to another ESCON channel, where one of the two ESCON channels is defined as an ESCON CTC channel and the other ESCON channel is defined as a ESCON CNC channel. Also a mnemonic for a FICON channel supporting a CTC Control Unit function logically or physically connected to another FICON channel that also supports a CTC Control Unit function. FICON channels supporting the FICON CTC Control Unit function are defined as normal FICON native (FC) mode channels.

CVC A mnemonic for an ESCON channel attached to an IBM 9034 convertor. The 9034 converts ESCON CVC signals to parallel channel interface (OEMI) communication operating in block multiplex mode (Bus and Tag).

Cyclic Redundancy Check (CRC) An error-correcting code used in Fibre Channel.

DASD See *direct access storage device.*

DAT See *Digital Audio Tape*.

data sharing A SAN solution in which files on a storage device are shared between multiple hosts.

datagram Refers to the Class-3 Fibre Channel Service that allows data to be sent rapidly to multiple devices attached to the fabric, with no confirmation of delivery.

DDM See *disk drive module*.

dedicated connection In an ESCON Director, a connection between two ports that is not affected by information contained in the transmission frames. This connection, which restricts those ports from communicating with any other port, can be established or removed only as a result of actions performed by a host control program or at the ESCD console. Contrast with *dynamic connection*.

Note: The two links having a dedicated connection appear as one continuous link.

default Pertaining to an attribute, value, or option that is assumed when none is explicitly specified.

Dense Wavelength Division Multiplexing (DWDM) The concept of packing multiple signals tightly together in separate groups, and transmitting them simultaneously over a common carrier wave.

destination Any point or location, such as a node, station, or a particular terminal, to which information is to be sent. An example is a Fibre Channel fabric F_Port; when attached to a Fibre Channel N_port, communication to the N_port through the F_port is said to be to the F_Port destination identifier (D_ID).

device A mechanical, electrical, or electronic contrivance with a specific purpose.

device address 1) In ESA/390 architecture and z/Architecture for zSeries, the field of an ESCON device-level frame that selects a specific device on a control unit image. 2) In the FICON channel FC-SB-2 architecture, the device address field in an SB-2 header that is used to select a specific device on a control unit image.

device number 1) In ESA/390 and z/Architecture for zSeries, a four-hexadecimal character identifier (for example, 19A0) that you associate with a device to facilitate communication between the program and the host operator. 2) The device number that you associate with a subchannel that uniquely identifies an I/O device.

dB Decibel. A ratio measurement distinguishing the percentage of signal attenuation (loss) between the I/O power. Attenuation is expressed as dB/km.

Digital Audio Tape (DAT) A tape media technology designed for very high quality audio recording and data backup. DAT cartridges look like audio cassettes and are often used in mechanical auto-loaders. Typically, a DAT cartridge provides 2 GB of storage, but new DAT systems have much larger capacities.

Digital Linear Tape (DLT) A magnetic tape technology originally developed by Digital Equipment Corporation (DEC) and now sold by Quantum. DLT cartridges provide storage capacities from 10 GB to 35 GB.

direct access storage device (DASD) A mass storage medium on which a computer stores data. Any online storage device: a disc, drive, or CD-ROM.

disconnected In an ESCON Director, the attribute that, when set, removes a dedicated connection. Contrast with *connected*.

disk A mass storage medium on which a computer stores data.

disk drive module (DDM) A disk storage medium that you use for any host data that is stored within a disk subsystem.

disk mirroring A fault-tolerant technique that writes data simultaneously to two hard disks using the same hard disk controller.

disk pooling A SAN solution in which disk storage resources are pooled across multiple hosts rather than dedicated to a specific host.

distribution panel In an ESCON and FICON environment, a panel that provides a central location for the attachment of trunk and jumper cables and can be mounted in a rack, wiring closet, or on a wall.

DLT See *Digital Linear Tape*.

duplex Pertaining to communication in which data or control information can be sent and received at the same time, from the same node. Contrast with *half duplex.*

duplex connector In an ESCON environment, an optical fiber component that terminates both jumper cable fibers in one housing and provides physical keying for attachment to a duplex receptacle.

duplex receptacle In an ESCON environment, a fixed or stationary optical fiber component that provides a keyed attachment method for a duplex connector.

DWDM See *Dense Wavelength Division Multiplexing*.

dynamic connection In an ESCON Director, a connection between two ports, established or removed by the ESCD and that, when active, appears as one continuous link. The duration of the connection depends on the protocol defined for the frames transmitted through the ports and on the state of the ports. Contrast with *dedicated connection.*

dynamic connectivity In an ESCON Director, the capability that allows connections to be established and removed at any time.

Dynamic I/O Reconfiguration An S/390 and z/Architecture function that allows I/O configuration changes to be made nondisruptively to the current operating I/O configuration.

ECL See *Emitter Coupled Logic*.

ELS See *Extended Link Services.*

EMIF See *ESCON Multiple Image Facility.*

Emitter Coupled Logic (ECL) The type of transmitter used to drive copper media such as Twinax, Shielded Twisted Pair, or Coax.

enterprise network A geographically dispersed network under the auspices of one organization.

Enterprise Systems Architecture/390® (ESA/390) An IBM architecture for mainframe computers and peripherals. Processors that follow this architecture include the S/390 Server family of processors.

Enterprise System Connection (ESCON) 1) An ESA/390 computer peripheral interface. The I/O interface uses ESA/390 logical protocols over a serial interface that configures attached units to a communication fabric. 2) A set of IBM products and services that provide a dynamically connected environment within an enterprise.

entity In general, a real or existing object from the Latin ens, or being, which makes the distinction between an object's existence and its qualities. In programming, engineering, and probably many other contexts, the word is used to identify units, whether concrete items or abstract ideas, that have no ready name or label.

E_Port Expansion Port. A port on a switch used to link multiple switches together into a Fibre Channel switch fabric.

ESA/390 See *Enterprise Systems Architecture/390*.

ESCD Enterprise Systems Connection (ESCON) Director.

ESCD console The ESCON Director display and keyboard device used to perform operator and service tasks at the ESCD.

ESCON See *Enterprise System Connection.*

ESCON channel A channel having an Enterprise Systems Connection channel-to-control-unit I/O interface that uses optical cables as a transmission medium. Can operate in CBY, CNC, CTC or CVC mode. Contrast with *parallel channel.*

ESCON Director An I/O interface switch that provides the interconnection capability of multiple ESCON interfaces (or FICON Bridge (FCV) mode - 9032-5) in a distributed-star topology.

ESCON Multiple Image Facility (EMIF) In the ESA/390 architecture and z/Architecture for zSeries, a function that allows logical partitions (LPARs) to share an ESCON and FICON channel path (and other channel types) by providing each LPAR with its own channel-subsystem image.

exchange A group of sequences that share a unique identifier. All sequences within a given exchange use the same protocol. Frames from multiple sequences can be multiplexed to prevent a single exchange from consuming all the bandwidth. See also *sequence*.

Extended Link Services (ELS) Through a command request, solicits a destination port (N_Port or F_Port) to perform a function or service. Each ELS request consists of an Link Service (LS) command; the N_Port ELS commands are defined in the FC-FS architecture.

fabric Fibre Channel employs a fabric to connect devices. A fabric can be as simple as a single cable connecting two devices. The term is most often used to describe a more complex network using hubs, switches, and gateways.

Fabric Login (FLOGI) Used by an N_Port to determine if a fabric is present and, if so, to initiate a session with the fabric by exchanging service parameters with the fabric. Fabric Login is performed by an N_Port following link initialization and before communication with other N_Ports is attempted.

Fabric Shortest Path First (FSPF) An intelligent path selection and routing standard that is part of the Fibre Channel Protocol.

FC 1) A short form when referring to something that is part of the Fibre Channel standard. Used by the IBM I/O definition process when defining a FICON channel (using IOCP of HCD) that will be used in FICON native mode (using the FC-SB-2 communication protocol. See also *Fibre Channel*.

FC-0 Lowest level of the Fibre Channel Physical standard, covering the physical characteristics of the interface and media.

FC-1 Middle level of the Fibre Channel Physical standard, defining the 8b/10b encoding and decoding and transmission protocol.

FC-2 Highest level of the Fibre Channel Physical standard, defining the rules for signaling protocol and describing transfer of frame, sequence, and exchanges.

FC-3 The hierarchical level in the Fibre Channel standard that provides common services such as striping definition.

FC-4 The hierarchical level in the Fibre Channel standard that specifies the mapping of upper-layer protocols to levels below.

FCA See *Fibre Channel Association*.

FC-AL See *Fibre Channel Arbitrated Loop*.

FC-CT Fibre Channel Common Transport Protocol.

FC-FG See *Fibre Channel Fabric Generic*.

FC-FP See *Fibre Channel HIPPI Framing Protocol*.

FC-FS See *Fibre Channel-Framing and Signaling*.

FC-GS See *Fibre Channel Generic Services*.

FCLC See *Fibre Channel Loop Association*.

FC-LE See *Fibre Channel Link Encapsulation*.

FCP See *Fibre Channel Protocol*.

FC-PH See *Fibre Channel Physical and Signaling*.

FC-PLDA Fibre Channel Private Loop Direct Attach. See *Private Loop Direct Attach*.

FCS See *Fibre Channel standard*.

FC-SB See *Fibre Channel Single Byte Command Code Set*.

FC Storage Director SAN Storage Director.

FC-SW See *Fibre Channel Switch Fabric*.

fiber See *optical fiber*.

Fibre Channel A technology for transmitting data between computer devices at a data rate of up to 4 Gbps. It is especially suited for connecting computer servers to shared storage devices and for interconnecting storage controllers and drives.

Fibre Channel Arbitrated Loop (FC-AL) A reference to the FC-AL standard, a shared gigabit media for up to 127 nodes, one of which may be attached to a switch fabric. See also *arbitrated loop*.

Fibre Channel Association (FCA) A Fibre Channel industry association that works to promote awareness and understanding of the Fibre Channel technology and its application, and provides a means for implementers to support the standards committee activities.

Fibre Channel Fabric Generic (FC-FG) A reference to the document (ANSI X3.289-1996) that defines the concepts, behavior, and characteristics of the Fibre Channel fabric along with suggested partitioning of the 24-bit address space to facilitate the routing of frames.

Fibre Channel-Framing and Signaling (FC-FS) The term used to describe the FC-FS architecture.

Fibre Channel Generic Services (FC-GS) A reference to the document (ANSI X3.289-1996) that describes a common transport protocol used to communicate with the server functions, a full X500-based directory service, mapping of the SNMP directly to the Fibre Channel, a time server, and an alias server.

Fibre Channel HIPPI Framing Protocol (FCFP) A reference to the document (ANSI X3.254-1994) that defines how the HIPPI framing protocol is transported through the Fibre Channel.

Fibre Channel Link Encapsulation (FC-LE) A reference to the document (ANSI X3.287-1996) that defines how IEEE 802.2 Logical Link Control (LLC) information is transported through the Fibre Channel.

Fibre Channel Loop Association (FCLC) An independent working group of the FCA focused on the marketing aspects of the Fibre Channel loop technology.

Fibre Channel Physical and Signaling (FC-PH) A reference to the ANSI X3.230 standard, which contains the definition of the three lower levels (FC-0, FC-1, and FC-2) of the Fibre Channel.

Fibre Channel Protocol (FCP) The mapping of SCSI-3 operations to Fibre Channel.

Fibre Channel Service Protocol (FSP) The common FC-4 level protocol for all services, transparent to the fabric type or topology.

Fibre Channel Single Byte Command Code Set (FC-SB) A reference to the document (ANSI X.271-1996) that defines how the ESCON command set protocol is transported using the Fibre Channel.

Fibre Channel standard (FCS) An ANSI standard for a computer peripheral interface. The I/O interface defines a protocol for communication over a serial interface that configures attached units to a communication fabric. The protocol has four layers. The lower of the four layers defines the physical media and interface, the upper of the four layers defines one or more Upper Layer Protocols (ULPs), for example, FCP for SCSI command protocols and FC-SB-2 for FICON protocol supported by ESA/390 and z/Architecture. Refer to ANSI X3.230.1999x.

Fibre Channel Switch Fabric (FC-SW) A reference to the ANSI standard under development that further defines the fabric behavior described in FC-FG and defines the communications between different fabric elements required for those elements to coordinate their operations and management address assignment.

fiber optic cable See *optical cable.*

fiber optics The branch of optical technology concerned with the transmission of radiant power through fibers made of transparent materials such as glass, fused silica, and plastic.

Note: Telecommunication applications of fiber optics use optical fibers. Either a single discrete fiber or a non-spatially aligned fiber bundle can be used for each information channel. Such fibers are often called “optical fibers” to differentiate them from fibers used in non-communication applications.

FICON 1) An ESA/390 and zSeries computer peripheral interface. The I/O interface uses ESA/390 and zSeries FICON protocols (FC-FS and FC-SB-2) over a Fibre Channel serial interface that configures attached units to a FICON supported Fibre Channel communication fabric. 2) An FC4 proposed standard that defines an effective mechanism for the export of the SBCCS-2 (FC-SB-2) command protocol through Fibre Channels.

FICON channel A channel having a Fibre Channel connection (FICON) channel-to-control-unit I/O interface that uses optical cables as a transmission medium. Can operate in either FC or FCV mode.

FICON Director A Fibre Channel switch that supports the ESCON-like “control unit port” (CUP function) that is assigned a 24-bit Fibre Channel port address to allow FC-SB-2 addressing of the CUP function to perform command and data transfer. (In the Fibre Channel world, it is a means of in-band management using a FC-4 ULP.)

field-replaceable unit (FRU) An assembly that is replaced in its entirety when any one of its required components fails.

F_Node Fabric Node. A fabric-attached node.

FLOGI See *Fabric Login*.

F_Port Fabric Port. A port used to attach a Node Port (N_Port) to a switch fabric.

frame A linear set of transmitted bits that define the basic transport unit. The frame is the most basic element of a message in Fibre Channel communications, consisting of a 24-byte header and zero to 2112 bytes of data. See also *sequence*.

FRU See *field-replaceable unit*.

FSP See *Fibre Channel Service Protocol*.

FSPF See *Fabric Shortest Path First*.

full duplex A mode of communications allowing simultaneous transmission and reception of frames.

gateway A node on a network that interconnects two otherwise incompatible networks.

Gbps Gigabits per second. Also sometimes referred to as Gb/s. In computing terms, it is approximately 1000000000 bits per second. Most precisely it is 1073741824 (1024 x 1024 x 1024) bits per second.

GBps Gigabytes per second. Also sometimes referred to as GB/s. In computing terms, it is approximately 1000000000 bytes per second. Most precisely it is 1073741824 (1024 x 1024 x 1024) bytes per second.

GBIC See *Gigabit Interface Converter*.

Gigabit One billion bits or one thousand megabits.

Gigabit Interface Converter (GBIC) Industry standard transceivers for connection of Fibre Channel nodes to arbitrated loop hubs and fabric switches.

Gigabit Link Module (GLM) A generic Fibre Channel transceiver unit that integrates the key functions necessary for the installation of a Fibre Channel media interface on most systems.

GLM See *Gigabit Link Module*.

G_Port Generic Port. A generic switch port that is either an F_Port or E_Port. The function is automatically determined during login.

half duplex In data communication, pertaining to transmission in only one direction at a time. Contrast with *duplex.*

hard disk drive Storage media within a storage server used to maintain information that the storage server requires. Also a mass storage medium for computers that is typically available as a fixed disk or a removable cartridge.

hardware The mechanical, magnetic, and electronic components of a system, such as computers, telephone switches, and terminals.

HBA Host bus adapter.

HCD Hardware configuration dialog.

HDA See *head and disk assembly.*

HDD See *hard disk drive.*

head and disk assembly (HDA) The portion of an HDD associated with the medium and the read/write head.

hierarchical storage management (HSM) A software and hardware system that moves files from disk to slower, less expensive storage media based on rules and observation of file activity. Modern HSM systems move files from magnetic disk to optical disk to magnetic tape.

High Performance Parallel Interface (HPPI) An ANSI standard that defines a channel that transfers data between CPUs and from a CPU to disk arrays and other peripherals.

HIPPI See *High Performance Parallel Interface*.

HMMP HyperMedia Management Protocol.

HMMS See *HyperMedia Management Schema*.

hop A Fibre Channel frame can travel from a switch to a director, a switch to a switch, or a director to a director, which in this case is one hop.

HSM See *Hierarchical Storage Management*.

hub A Fibre Channel device that connects nodes into a logical loop by using a physical star topology. Hubs will automatically recognize an active node and insert the node into the loop. A node that fails or is powered off is automatically removed from the loop.

hub topology See *loop topology*.

Hunt Group A set of associated N_Ports attached to a single node, assigned a special identifier that allows any frames containing this identifier to be routed to any available N_Port in the set.

HyperMedia Management Schema (HMMS) The definition of an implementation-independent, extensible, common data description/schema that allows data from a variety of sources to be described and accessed in real time regardless of the source of the data. See also *WEBM* and *HMMP*.

ID See *identifier.*

identifier A unique name or address that identifies such items as programs, devices, or systems.

in-band signaling Signaling that is carried in the same channel as the information. Also referred to as in-band.

in-band virtualization An implementation in which the virtualization process takes place in the data path between servers and disk systems. The virtualization can be implemented as software running on servers or in dedicated engines.

information unit A unit of information defined by an FC-4 mapping. Information units are transferred as a Fibre Channel sequence.

initial program load (IPL) 1) The initialization procedure that causes an operating system to commence operation. 2) The process by which a configuration image is loaded into storage at the beginning of a work day or after a system malfunction. 3) The process of loading system programs and preparing a system to run jobs.

input/output (I/O) 1) Pertaining to a device whose parts can perform an input process and an output process at the same time. 2) Pertaining to a functional unit or channel involved in an input process, output process, or both, concurrently or not, and to the data involved in such a process. 3) Pertaining to input, output, or both.

input/output configuration data set (IOCDS) The data set in the S/390 and zSeries processor (in the support element) that contains an I/O configuration definition built by the I/O configuration program (IOCP).

input/output configuration program (IOCP) An S/390 program that defines to a system the channels, I/O devices, paths to the I/O devices, and the addresses of the I/O devices. The output is normally written to a S/390 or zSeries IOCDS.

interface 1) A shared boundary between two functional units, defined by functional characteristics, signal characteristics, or other characteristics as appropriate. The concept includes the specification of the connection of two devices having different functions. 2) Hardware, software, or both, that link systems, programs, or devices.

intermix A mode of service defined by Fibre Channel that reserves the full Fibre Channel bandwidth for a dedicated Class-1 connection, but allows connection-less Class-2 traffic to share the link if the bandwidth is available.

inter-switch link (ISL) A Fibre Channel connection between switches and directors.

I/O See *input/output*.

I/O configuration The collection of channel paths, control units, and I/O devices that attaches to the processor. This can also include channel switches (for example, an ESCON Director).

IOCDS See *input/output configuration data set*.

IOCP See *input/output configuration control program*.

IODF The data set that contains the S/390 or zSeries I/O configuration definition file produced during the definition of the S/390 or zSeries I/O configuration by HCD. Used as a source for IPL, IOCP, and Dynamic I/O Reconfiguration.

IP Internet Protocol.

IPI Intelligent Peripheral Interface.

IPL See *initial program load*.

ISL See *inter-switch link*.

isochronous transmission Data transmission that supports network-wide timing requirements. A typical application for isochronous transmission is a broadcast environment that needs information to be delivered at a predictable time.

JBOD Just a bunch of disks.

jukebox A device that holds multiple optical disks and one or more disk drives, and can swap disks in and out of the drive as needed.

jumper cable In an ESCON and FICON environment, an optical cable having two conductors that provide physical attachment between a channel and a distribution panel or an ESCON/FICON Director port or a control unit/device, between an ESCON/FICON Director port and a distribution panel or a control unit/device, or between a control unit/device and a distribution panel. Contrast with *trunk cable*.

LAN See *local area network*.

laser A device that produces optical radiation using a population inversion to provide *light amplification by stimulated emission of radiation* and (generally) an optical resonant cavity to provide positive feedback. Laser radiation can be highly coherent temporally, spatially, or both.

latency A measurement of the time it takes to send a frame between two locations.

LC Lucent Connector. A registered trademark of Lucent Technologies.

LCU See *logical control unit*.

LED See *light emitting diode*.

licensed internal code (LIC) Microcode that IBM does not sell as part of a machine, but instead, licenses it to the client. LIC is implemented in a part of storage that is not addressable by user programs. Some IBM products use it to implement functions as an alternate to hard-wire circuitry.

light emitting diode (LED) A semiconductor chip that gives off visible or infrared light when activated. Contrast with *laser*.

link 1) In an ESCON environment or FICON environment (Fibre Channel environment), the physical connection and transmission medium used between an optical transmitter and an optical receiver. A link consists of two conductors, one used for sending and the other for receiving, thereby providing a duplex communication path. 2) In an ESCON I/O interface, the physical connection and transmission medium used between a channel and a control unit, a channel and an ESCD, a control unit and an ESCD, or at times between two ESCDs. 3) In a FICON I/O interface, the physical connection and transmission medium used between a channel and a control unit, a channel and a FICON Director, a control unit and a Fibre Channel FICON Director, or at times, between two Fibre Channels switches.

link address 1) On an ESCON interface, the portion of a source or destination address in a frame that ESCON uses to route a frame through an ESCON Director. ESCON associates the link address with a specific switch port that is on the ESCON Director. 2) On a FICON interface, the port address (1-byte link address), or domain and port address (2-byte link address) portion of a source (S_ID) or destination address (D_ID) in a Fibre Channel frame that the Fibre Channel switch uses to route a frame through a Fibre Channel switch or Fibre Channel switch fabric. See also *port address.*

Link_Control_Facility A termination card that handles the logical and physical control of the Fibre Channel link for each mode of use.

LIP See *loop initialization primitive sequence*.

local area network (LAN) A computer network located in a user's premises within a limited geographic area, usually not larger than a floor or small building. Transmissions within a LAN are mostly digital, carrying data among stations at rates usually above one Mbps.

logical control unit (LCU) A separately addressable control unit function within a physical control unit. Usually a physical control unit that supports several LCUs. For ESCON, the maximum number of LCUs that can be in a control unit (and addressed from the same ESCON fiber link) is 16. They are addressed from x'0' to x'F'. For FICON architecture, the maximum number of LCUs that can be in a control unit (and addressed from the same FICON fibre link) is 256. They are addressed from x'00' to x'FF'. For both ESCON and FICON, the actual number supported, and the LCU address value, is both processor- and control unit implementation-dependent.

logical partition (LPAR) A set of functions that create a programming environment that is defined by the ESA/390 architecture or z/Architecture for zSeries. The ESA/390 architecture or z/Architecture for zSeries uses the term LPAR when more than one LPAR is established on a processor. An LPAR is conceptually similar to a virtual machine environment except that the LPAR is a function of the processor. Also, LPAR does not depend on an operating system to create the virtual machine environment.

logical switch number (LSN) A two-digit number used by the IOCP to identify a specific ESCON or FICON Director. This number is separate from the director's "switch device number" and, for FICON, it is separate from the director's "Fibre Channel switch address."

logically partitioned mode A central processor mode, available on the configuration frame when using the PR/SM™ facility, that allows an operator to allocate processor hardware resources among LPARs. Contrast with *basic mode.*

login server An entity within the Fibre Channel fabric that receives and responds to login requests.

loop circuit A temporary point-to-point like path that allows bidirectional communications between loop-capable ports.

loop initialization primitive (LIP) sequence A special Fibre Channel sequence that is used to start loop initialization. Allows ports to establish their port addresses.

loop topology An interconnection structure in which each point has physical links to two neighbors resulting in a closed circuit. In a loop topology, the available bandwidth is shared.

LPAR See *logical partition*.

L_Port Loop Port. A node or fabric port capable of performing arbitrated loop functions and protocols. NL_Ports and FL_Ports are loop-capable ports.

LSN See *logical switch number*.

Lucent Connector (LC) A registered trademark of Lucent Technologies

LVD Low Voltage Differential.

management agent A process that exchanges a managed node's information with a management station.

managed node A computer, a storage system, a gateway, a media device such as a switch or hub, a control instrument, a software product such as an operating system or an accounting package, or a machine on a factory floor, such as a robot.

managed object A variable of a managed node. This variable contains one piece of information about the node. Each node can have several objects.

Management Information Block (MIB) A formal description of a set of network objects that can be managed using the SNMP. The format is defined as part of SNMP and is a hierarchical structure of information relevant to a specific device, defined in object-oriented terminology as a collection of objects, relations, and operations among objects.

management station A host system that runs the management software.

MAR See *Media Access Rules*.

Mbps Megabits per second. Also sometimes referred to as Mb/s. In computing terms, it is approximately 1000000 bits per second. Most precisely it is 1048576 (1024 x 1024) bits per second.

MBps Megabytes per second. Also sometimes referred to as MB/s. In computing terms, it is approximately 1000000 bytes per second. Most precisely it is 1048576 (1024 x 1024) bytes per second.

media Plural of medium. The physical environment through which transmission signals pass. Common media include copper and fiber optic cable.

Media Access Rules (MAR) Enable systems to self-configure themselves is a SAN environment.

Media Interface Adapter (MIA) Enables optic-based adapters to interface with copper-based devices, including adapters, hubs, and switches.

metadata server In Storage Tank™, servers that maintain information (metadata) about the data files and grant permission for application servers to communicate directly with disk systems.

meter Equal to 39.37 inches, or just slightly larger than a yard (36 inches).

MIA See *Media Interface Adapter*.

MIB See *Management Information Block*.

mirroring The process of writing data to two separate physical devices simultaneously.

MM Multi-Mode. See *Multi-Mode Fiber*.

MMF See *Multi-Mode Fiber*.

multicast Sending a copy of the same transmission from a single source device to multiple destination devices on a fabric. This includes sending to all N_Ports on a fabric (broadcast) or to only a subset of the N_Ports on a fabric (multicast).

Multi-Mode Fiber (MMF) In optical fiber technology, an optical fiber that is designed to carry multiple light rays or modes concurrently, each at a slightly different reflection angle within the optical core. Multi-Mode Fiber transmission is used for relatively short distances because the modes tend to disperse over longer distances. See also *Single-Mode Fiber*.

multiplex The ability to intersperse data from multiple sources and destinations onto a single transmission medium. Refers to delivering a single transmission to multiple destination N_Ports.

name server Provides translation from a given node name to one or more associated N_Port identifiers.

NAS See *Network Attached Storage*.

ND See *node descriptor*.

NDMP Network Data Management Protocol.

NED See *node-element descriptor.*

network An aggregation of interconnected nodes, workstations, file servers, and peripherals, with its own protocol that supports interaction.

Network Attached Storage (NAS) A term used to describe a technology where an integrated storage system is attached to a messaging network that uses common communications protocols, such as TCP/IP.

Network File System (NFS) A distributed file system in UNIX developed by Sun Microsystems™. It allows a set of computers to cooperatively access each other's files in a transparent manner.

Network Management System (NMS) A system responsible for managing at least part of a network. NMSs communicate with agents to help keep track of network statistics and resources.

network topology Physical arrangement of nodes and interconnecting communications links in networks based on application requirements and geographical distribution of users.

NFS See *Network File System*.

NL_Port Node Loop Port. A node port that supports arbitrated loop devices.

NMS See *Network Management System*. A system responsible for managing at least part of a network. NMSs communicate with agents to help keep track of network statistics and resources.

node An entity with one or more N_Ports or NL_Ports.

node descriptor (ND) In an ESCON and FICON environment, a 32-byte field that describes a node, channel, ESCON Director or FICON Director port, or a control unit.

node-element descriptor (NED) In an ESCON and FICON environment, a 32-byte field that describes a node element, such as a disk (DASD) device.

non-blocking Indicates that the capabilities of a switch are such that the total number of available transmission paths is equal to the number of ports. Therefore, all ports can have simultaneous access through the switch.

Non-L_Port A Node or Fabric Port that is not capable of performing the arbitrated loop functions and protocols. N_Ports and F_Ports are not loop-capable ports.

N_Port Node Port. A Fibre Channel-defined hardware entity at the end of a link that provides the mechanisms necessary to transport information units to or from another node.

N_Port Login (PLOGI) Allows two N_Ports to establish a session and exchange identities and service parameters. It is performed following completion of the FLOGI process and prior to the FC-4 level operations with the destination port. Can be either explicit or implicit.

OEMI See *original equipment manufacturer information*.

open system A system whose characteristics comply with standards made available throughout the industry and that can be connected to other systems that comply with the same standards.

operation A term defined in FC-2 that refers to one of the Fibre Channel *building blocks* composed of one or more, possibly concurrent, exchanges.

optical cable A fiber, multiple fibers, or a fiber bundle in a structure built to meet optical, mechanical, and environmental specifications. See also *jumper cable*, *optical cable assembly*, and *trunk cable.*

optical cable assembly An optical cable that is connector-terminated. Generally, an optical cable that has been connector-terminated by a manufacturer and is ready for installation. See also *jumper cable* and *optical cable*.

optical fiber Any filament made of dialectic materials that guides light, regardless of its ability to send signals. See also *fiber optics* and *optical waveguide*.

optical fiber connector A hardware component that transfers optical power between two optical fibers or bundles and is designed to be repeatedly connected and disconnected.

optical waveguide A structure capable of guiding optical power. In optical communications, generally a fiber designed to transmit optical signals. See *optical fiber*.

ordered set A Fibre Channel term referring to four 10-bit characters (a combination of data and special characters) providing low-level link functions, such as frame demarcation and signaling between two ends of a link.

original equipment manufacturer information (OEMI) A reference to an IBM guideline for a computer peripheral interface. More specifically, it refers to IBM S/360™ and S/370™ Channel to Control Unit OEMI. The interface uses ESA/390 logical protocols over an I/O interface that configures attached units in a multidrop bus environment. This OEMI interface is also supported by the zSeries 900 processors.

originator A Fibre Channel term referring to the initiating device.

out-of-band signaling Signaling that is separated from the channel carrying the information. Also referred to as *out-of-band*.

out-of-band virtualization An alternative type of virtualization in which servers communicate directly with disk systems under control of a virtualization function that is not involved in the data transfer.

parallel channel A channel having a System/360™ and System/370™ channel-to-control-unit I/O interface that uses bus and tag cables as a transmission medium. Contrast with *ESCON channel*.

path In a channel or communication network, any route between any two nodes. For ESCON and FICON, this is the route between the channel and the control unit/device, or sometimes from the operating system control block for the device and the device itself.

path group The ESA/390 and zSeries architecture (z/Architecture) term for a set of channel paths that are defined to a controller as being associated with a single S/390 image. The channel paths are in a group state and are online to the host.

path-group identifier ESA/390 and z/Architecture term for the identifier that uniquely identifies a given LPAR. The path-group identifier is used in communication between the system image program and a device. The identifier associates the path group with one or more channel paths, defining these paths to the control unit as being associated with the same system image.

PCICC (IBM) PCI Cryptographic Coprocessor.

peripheral Any computer device that is not part of the essential computer (the processor, memory, and data paths) but is situated relatively close by. A near synonym is I/O device.

petard A device that is small and sometimes explosive.

PLDA See *Private Loop Direct Attach*.

PLOGI See *N_Port Login*.

point-to-point topology An interconnection structure in which each point has physical links to only one neighbor resulting in a closed circuit. In point-to-point topology, the available bandwidth is dedicated.

policy-based management Management of data on the basis of business policies (for example, “all production database data must be backed up every day”), rather than technological considerations (for example, “all data stored on this disk system is protected by remote copy”).

port An access point for data entry or exit. A receptacle on a device to which a cable for another device is attached. See also *duplex receptacle*.

port address In an ESCON Director, an address used to specify port connectivity parameters and to assign link addresses for attached channels and control units. In a FICON Director or Fibre Channel switch, it is the middle 8 bits of the full 24-bit Fibre Channel port address. This field is also referred to as the *area field* in the 24-bit Fibre Channel port address. See also *link address*.

port bypass circuit A circuit used in hubs and disk enclosures to automatically open or close the loop to add or remove nodes on the loop.

port card In an ESCON and FICON environment, a field-replaceable hardware component that provides the optomechanical attachment method for jumper cables and performs specific device-dependent logic functions.

port name In an ESCON or FICON Director, a user-defined symbolic name of 24 characters or less that identifies a particular port.

Private Loop Direct Attach (PLDA) A technical report that defines a subset of the relevant standards suitable for the operation of peripheral devices such as disks and tapes on a private loop.

Private NL_Port An NL_Port that does not attempt to log in with the fabric and only communicates with other NL_Ports on the same loop.

processor complex A system configuration that consists of all the machines required for operation, for example, a processor unit, a processor controller, a system display, a service support display, and a power and coolant distribution unit.

program temporary fix (PTF) A temporary solution or bypass of a problem diagnosed by IBM in a current unaltered release of a program.

prohibited In an ESCON or FICON Director, the attribute that, when set, removes dynamic connectivity capability. Contrast with *allowed.*

protocol 1) A set of semantic and syntactic rules that determine the behavior of functional units in achieving communication. 2) In Fibre Channel, the meaning of, and sequencing rules for, requests and responses used for managing the switch or switch fabric, transferring data, and synchronizing states of Fibre Channel fabric components. 3) A specification for the format and relative timing of information exchanged between communicating parties.

PTF See *program temporary fix*.

Public NL_Port An NL_Port that attempts to log in with the fabric and can observe the rules of either public or private loop behavior. A public NL_Port can communicate with both private and public NL_Ports.

QoS See *quality of service*.

quality of service (QoS) A set of communications characteristics required by an application. Each QoS defines a specific transmission priority, level of route reliability, and security level.

Quick Loop A unique Fibre Channel topology that combines arbitrated loop and fabric topologies. It is an optional licensed product that allows arbitrated loops with private devices to be attached to a fabric.

RAID See *Redundant Array of Inexpensive or Independent Disks.*

RAID 0 Level 0 RAID support. Striping, no redundancy.

RAID 1 Level 1 RAID support. Mirroring, complete redundancy.

RAID 5 Level 5 RAID support. Striping with parity.

Redundant Array of Inexpensive or Independent Disks (RAID) A method of configuring multiple disk drives in a storage subsystem for high availability and high performance.

repeater A device that receives a signal on an electromagnetic or optical transmission medium, amplifies the signal, and then retransmits it along the next leg of the medium.

responder A Fibre Channel term referring to the answering device.

route The path that an ESCON frame takes from a channel through an ESCD to a control unit/device.

router 1) A device that can decide which of several paths network traffic will follow based on some optimal metric. Routers forward packets from one network to another based on network-layer information. 2) A dedicated computer hardware or software package that manages the connection between two or more networks. See also *bridge* and *bridge/router*.

SAF-TE SCSI Accessed Fault-Tolerant Enclosures.

SAN See *storage area network*.

SAN See *system area network*.

SANSymphony In-band block-level virtualization software made by DataCore Software Corporation and resold by IBM.

saved configuration In an ESCON or FICON Director environment, a stored set of connectivity attributes whose values determine a configuration that can be used to replace all or part of the ESCD's or FICON's active configuration. Contrast with *active configuration*.

SC connector A fiber optic connector standardized by ANSI TIA/EIA-568A for use in structured wiring installations.

scalability The ability of a computer application or product (hardware or software) to continue to function because of a change in size or volume, for example, the ability to retain performance levels when adding additional processors, memory, and storage.

SCSI See *Small Computer System Interface.*

SCSI-3 SCSI-3 consists of a set of primary commands and additional specialized command sets to meet the needs of specific device types. The SCSI-3 command sets are used not only for the SCSI-3 parallel interface but for additional parallel and serial protocols, including Fibre Channel, Serial Bus Protocol (used with IEEE 1394 Firewire physical protocol), and the Serial Storage Protocol (SSP).

SCSI Enclosure Services (SES) ANSI SCSI-3 proposal that defines a command set for soliciting basic device status (temperature, fan speed, power supply status, and so on) from a storage enclosures.

SCSI-FCP The term used to refer to the ANSI Fibre Channel Protocol for SCSI document (X3.269-199x) that describes the FC-4 protocol mappings and the definition of how the SCSI protocol and command set are transported using a Fibre Channel interface.

SE See *service element*.

sequence A series of frames strung together in numbered order, which can be transmitted over a Fibre Channel connection as a single operation. See also *exchange*.

SERDES Serializer Deserializer.

Serial Storage Architecture (SSA) A high-speed serial loop-based interface developed as a high speed point-to-point connection for peripherals, particularly high-speed storage arrays, RAID, and CD-ROM storage by IBM.

server A computer that is dedicated to one task.

service element (SE) A dedicated service processing unit used to service a S/390 machine (processor).

SES See *SCSI Enclosure Services*.

Simple Network Management Protocol (SNMP) The Internet network management protocol that provides a means to monitor and set network configuration and runtime parameters.

Single-Mode Fiber (SMF) In optical fiber technology, an optical fiber that is designed for the transmission of a single ray or mode of light as a carrier. It is a single light path used for long-distance signal transmission. See also *Multi-Mode Fiber*.

Small Computer System Interface (SCSI) 1) A set of evolving ANSI standard electronic interfaces that allow personal computers to communicate with peripheral hardware, such as disk drives, tape drives, CD_ROM drives, printers, and scanners, faster and more flexibly than previous interfaces. The interface uses a SCSI logical protocol over an I/O interface that configures attached targets and initiators in a multidrop bus topology. The following table identifies the major characteristics of the different SCSI versions.

SCSI version	Signal rate (MHz)	BusWidth (bits)	Maximum DTR (MBps)	Maximum no. devices	Maximum cable length (m)
SCSI-1	5	8	5	7	6
SCSI-2	5	8	5	7	6
Wide SCSI-2	5	16	10	15	6
Fast SCSI-2	10	8	10	7	6
Fast Wide SCSI-2	10	16	20	15	6
Ultra™ SCSI	20	8	20	7	1.5
Ultra SCSI-2	20	16	40	7	12
Ultra2 LVD SCSI	40	16	80	15	12

SM Single Mode. See *Single-Mode Fiber*.

SMART Self Monitoring and Reporting Technology.

SMF See *Single-Mode Fiber*.

SNIA See *Storage Networking Industry Association*.

SN storage network. See also *SAN*.

SNMP See *Simple Network Management Protocol*.

SNMWG See *Storage Network Management Working Group*.

SSA See *Serial Storage Architecture*.

star The physical configuration used with hubs in which each user is connected by communications links radiating out of a central hub that handles all communications.

storage area network (SAN) A dedicated, centrally managed, secure information infrastructure, which enables any-to-any interconnection of servers and storage systems.

storage media The physical device onto which data is recorded. Magnetic tape, optical disks, and floppy disks are all storage media.

Storage Network Management Working Group (SNMWG) Chartered to identify, define, and support open standards needed to address the increased management requirements imposed by storage area network environments.

Storage Networking Industry Association (SNIA) A non-profit organization comprised of more than 77 companies and individuals in the storage industry.

Storage Tank An IBM file aggregation project that enables a pool of storage, and even individual files, to be shared by servers of different types. In this way, Storage Tank can greatly improve storage utilization and enables data sharing.

StorWatch Expert StorWatch applications that employ a three-tiered architecture that includes a management interface, a StorWatch manager, and agents that run on the storage resource or resources being managed. Products employ a StorWatch database that can be used for saving key management data, such as capacity or performance metrics. Products also use the agents and analysis of storage data saved in the database to perform higher value functions including the reporting of

capacity and performance over time (trends), configuration of multiple devices based on policies, monitoring of capacity and performance, automated responses to events or conditions, and storage-related data mining.

StorWatch Specialist A StorWatch interface for managing an individual Fibre Channel device or a limited number of like devices (that can be viewed as a single group). Typically provides simple, point-in-time management functions such as configuration, reporting on asset and status information, simple device and event monitoring, and some service utilities.

STP Shielded Twisted Pair.

striping A method for achieving higher bandwidth using multiple N_Ports in parallel to transmit a single information unit across multiple levels.

subchannel A logical function of a channel subsystem associated with the management of a single device.

subsystem A secondary or subordinate system, or programming support, usually capable of operating independently of or asynchronously with a controlling system.

SWCH In ESCON Manager, the mnemonic used to represent an ESCON Director.

switch A component with multiple entry and exit points (ports) that provides dynamic connection between any two of these points.

switch topology An interconnection structure in which any entry point can be dynamically connected to any exit point. The available bandwidth is scalable.

system area network (SAN) Term originally used to describe a particular symmetric multiprocessing (SMP) architecture in which a switched interconnect is used in place of a shared bus. Server area network refers to a switched interconnect between multiple SMPs.

T11 A technical committee of the National Committee for Information Technology Standards, titled T11 I/O Interfaces. Develops standards for moving data into and out of computers.

tape backup Making magnetic tape copies of hard disk and optical disc files for disaster recovery.

tape pooling A SAN solution in which tape resources are pooled and shared across multiple hosts rather than being dedicated to a specific host.

TCP See *Transmission Control Protocol*.

TCP/IP See *Transmission Control Protocol/Internet Protocol*.

time server A Fibre Channel-defined service function that allows for the management of all timers used within a Fibre Channel system.

topology An interconnection scheme that allows multiple Fibre Channel ports to communicate. For example, point-to-point, arbitrated loop, and switched fabric are all Fibre Channel topologies.

TL_Port A private to public bridging of switches or directors, referred to as Translative Loop.

T_Port An ISL port more commonly known as an E_Port, referred to as a Trunk port and used by INRANGE.

Transmission Control Protocol (TCP) A reliable, full duplex, connection-oriented end-to-end transport protocol running on top of IP.

Transmission Control Protocol/Internet Protocol (TCP/IP) A set of communications protocols that support peer-to-peer connectivity functions for both LAN and WANs.

trunk cable In an ESCON and FICON environment, a cable consisting of multiple fiber pairs that do not directly attach to an active device. This cable usually exists between distribution panels (or sometimes between a set processor channels and a distribution panel) and can be located within, or external to, a building. Contrast with *jumper cable*.

twinax A transmission media (cable) consisting of two insulated central conducting leads of coaxial cable.

twisted pair The most common type of transmission media (cable), which consists of two insulated copper wires twisted around each other to reduce the induction (interference) from one wire to another. The twists, or lays, are varied in length to reduce the potential for signal interference between pairs. Several sets of twisted pair wires can be enclosed in a single cable.

ULP Upper Level Protocols.

unblocked In an ESCON and FICON Director, the attribute that, when set, establishes communication capability for a specific port. Contrast with *blocked*.

Under-The-Covers (UTC) A term used to characterize a subsystem in which a small number of hard drives are mounted inside a higher function unit. The power and cooling are obtained from the system unit. Connection is by parallel copper ribbon cable or pluggable backplane, using IDE or SCSI protocols.

unit address The ESA/390 and zSeries term for the address associated with a device on a given controller. On ESCON and FICON interfaces, the unit address is the same as the device address. On OEMI interfaces, the unit address specifies a controller and device pair on the interface.

UTC See *Under-The-Covers*.

UTP Unshielded Twisted Pair.

virtual circuit A unidirectional path between two communicating N_Ports that permits fractional bandwidth.

virtualization An abstraction of storage where the representation of a storage unit to the operating system and applications on a server is divorced from the actual physical storage where the information is contained.

virtualization engine Dedicated hardware and software that are used to implement virtualization.

WAN See *wide area network*.

Wave Division Multiplexing (WDM) A technology that puts data from different sources together on an optical fiber, with each signal carried on its own separate light wavelength. Using WDM, up to 80 (and theoretically more) separate wavelengths or channels of data can be multiplexed into a stream of light transmitted on a single optical fiber.

WDM See *Wave Division Multiplexing*.

Web-Based Enterprise Management (WEBM) A consortium working on the development of a series of standards to enable active management and monitoring of network-based elements.

WEBM See *Web-Based Enterprise Management*.

wide area network (WAN) A network that encompasses inter-connectivity between devices over a wide geographic area. A WAN can be privately owned or rented, but the term usually indicates the inclusion of public (shared) networks.

z/Architecture An IBM architecture for mainframe computers and peripherals. Processors that follow this architecture include the zSeries family of processors.

zoning In Fibre Channel environments, the grouping together of multiple ports to form a virtual private storage network. Ports that are members of a group or zone can communicate with each other but are isolated from ports in other zones.

zSeries A family of IBM mainframe servers that support high performance, availability, connectivity, security, and integrity.

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

- *Introduction to Storage Area Networks*, SG24-5470
- *IBM Storage Solutions for Server Consolidation*, SG24-5355
- *IBM TotalStorage Enterprise Storage Server: Implementing the ESS in Your Environment*, SG24-5420
- *IBM Enterprise Storage Server*, SG24-5465
- *Introduction to Storage Area Networks*, SG24-5470
- *IBM Tape Solutions for Storage Area Networks and FICON*, SG24-5474
- *IBM TotalStorage: Implementing an Open IBM SAN,* SG24-6116
- *IP Storage Networking: IBM NAS and iSCSI Solutions*, SG24-6240
- *Implementing Linux with IBM Disk Storage*, SG24-6261
- *Implementing the IBM TotalStorage NAS 300G: High Speed Cross Platform Storage and Tivoli SANergy!*, SG24-6278
- *Using iSCSI Solutions' Planning and Implementation*, SG24-6291
- *IBM TotalStorage: SAN Product, Design, and Optimization Guide*, SG24-6384
- *Implementing the Cisco MDS 9000 in an Intermix FCP, FCIP, and FICON Environment*, SG24-6397
- *Introduction to SAN Distance Solutions*, SG24-6408
- *Introducing Hosts to the SAN Fabric*, SG24-6411
- *The IBM TotalStorage NAS Integration Guide*, SG24-6505

Other resources

These publications are also relevant as further information sources:

- *Cisco MDS 9000 Family Fabric Manager Switch Configuration Guide*, 78-16493-01
- *McDATA SAN Router Administration and Configuration Manual*, 620-000206
- Clark, T., *IP SANs: An Introduction to iSCSI, iFCP, and FCIP Protocols for Storage Area Networks*, Addison-Wesley Professional, 2001, ISBN 0201752778
- Farley, M., *Building Storage Networks*, McGraw-Hill/Osborne Media, 2000, ISBN 0072120509
- Judd, J., *Multiprotocol Routing for SANs*, Infinity Publishing, 2004, ISBN 0741423065

Referenced Web sites

These Web sites are also relevant as further information sources:

- IBM TotalStorage hardware, software, and solutions

  http://www.storage.ibm.com
- IBM TotalStorage storage area network

  http://www.storage.ibm.com/snetwork/index.html
- IBM System Storage Proven

  http://www.ibm.com/storage/proven
- Brocade

  http://www.brocade.com
- Cisco

  http://www.cisco.com
- McDATA

  http://www.inrange.com/
- QLogic

  http://www.qlogic.com
- Emulex

  http://www.emulex.com

- Finisar

  http://www.finisar.com
- Veritas (Symantec Corp.)

  http://www.symantec.com/enterprise/veritas/index.jsp
- Tivoli

  http://www.tivoli.com
- JNI

  http://www.jni.com
- IEEE

  http://www.ieee.org
- Storage Networking Industry Association

  http://www.snia.org
- SCSI Trade Association

  http://www.scsita.org
- Internet Engineering Task Force

  http://www.ietf.org
- American National Standards Institute

  http://www.ansi.org
- Technical Committee T10

  http://www.t10.org
- Technical Committee T11

  http://www.t11.org
- IBM eServer xSeries 430 and NUMA-Q Information Center

  http://publib.boulder.ibm.com/xseries/
- The Requests for Comments (RFC) document series

  http://www.rfc-editor.org/

How to get IBM Redbooks

You can search for, view, or download Redbooks, Redpapers, Hints and Tips, draft publications and Additional materials, as well as order hardcopy Redbooks or CD-ROMs, at this Web site:

ibm.com/redbooks

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

Index

C

D

E

F

G

H

I

J

K

L

M

N

O

P

Q

R

S

T

U

V

W

X

Z

SAN Multiprotocol Routing: An Introduction and Implementation

(1.0” spine)
0.875”<->1.498”
460 <-> 788 pages

SAN Multiprotocol Routing
An Introduction and Implementation

Read about the basics of the IBM approach to multiprotocol routing

Learn about the IBM products and solutions

Understand how to install routers

The rapid spread and adoption of production storage area networks (SANs) has fuelled the need for multiprotocol routers. The routers provide improved scalability, security, and manageability by enabling devices in separate SAN fabrics to communicate without merging fabrics into a single, large SAN fabric. This capability enables clients to initially deploy separate SAN solutions at the departmental and data center levels. Then, clients can consolidate these separate solutions into large enterprise SAN solutions as their experience and requirements grow and change.

Alternatively, multiprotocol routers can help to connect existing enterprise SANs for a variety of reasons. For instance, the introduction of Small Computer System Interface over IP (iSCSI) provides for the connection of low-end, low-cost hosts to enterprise SANs. The use of an Internet Protocol (IP) in the Fibre Channel (FC) environment provides for resource consolidation and disaster recovery planning over long distances. And the use of FC-FC routing services provides connectivity between two or more fabrics without having to merge them into a single SAN.

This IBM Redbook targets storage network administrators, system designers, architects, and IT professionals who sell, design, or administer SANs. It introduces you to the products, concepts, and technology in the IBM System Storage SAN Routing portfolio. This book shows the features of each product and examples of how you can deploy and use them.

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

SG24-7321-00 ISBN 0738496987