Bull

Storage Area Network (SAN)
Installation and Configuration Guide

AIX

ORDER REFERENCE
86 A2 57EF 01

Bull

Storage Area Network (SAN) Installation and Configuration Guide

AIX

Software

February 2002

BULL CEDOC
357 AVENUE PATTON
B.P.20845
49008 ANGERS CEDEX 01
FRANCE

ORDER REFERENCE
86 A2 57EF 01

The following copyright notice protects this book under the Copyright laws of the United States of America and other countries which prohibit such actions as, but not limited to, copying, distributing, modifying, and making derivative works.

Copyright © Bull S.A. 1992, 2002

Printed in France

Suggestions and criticisms concerning the form, content, and presentation of this book are invited. A form is provided at the end of this book for this purpose.

To order additional copies of this book or other Bull Technical Publications, you are invited to use the Ordering Form also provided at the end of this book.

Trademarks and Acknowledgements

We acknowledge the right of proprietors of trademarks mentioned in this book.

AIX® is a registered trademark of International Business Machines Corporation, and is being used under licence.

UNIX is a registered trademark in the United States of America and other countries licensed exclusively through the Open Group.

The information in this document is subject to change without notice. Groupe Bull will not be liable for errors contained herein, or for incidental or consequential damages in connection with the use of this material.

Preface

Who should read this book

This book is intended for readers who are not familiar with Storage Area Network (SAN) but who need to install a SAN on their IT system. The first three chapters provide an introduction to SAN technology, infrastructure, and components. The next two chapters describe how to design a SAN and how to install the Bull SAN. The last chapter contains technical details of internal SAN software operation.

The book is organized as follows:

- Chapter 1. Introduction
- Chapter 2. SAN Infrastructure
- Chapter 3. SAN Components
- Chapter 4. SAN Methodology
- Chapter 5 SAN Installation
- Chapter 6. How the OS handles SAN Objects
- Glossary
- Index

Related Publications

Cabling Guide for Multiple Bus system,	86A170JX
Site Monitoring for Disaster Recovery,	86A283JX
Recovery from Cluster Site Disaster User's Guide,	86A286JX

DAS

DAS Subsystems – DAE Rackmount Models – Installation and Service,	86A145KX
DAS Subsystems – DAE Deskside Models – Installation and Service,	86A146KX
DAS4500 Deskside Models – Installation and Service,	86A101EF
DAS4500 Rackmount Models – Installation and Service,	86A102EF
DAS4700 Hardware Reference,	86A170EF
DAS4700–2 Rackmount Model Hardware Reference,	86A185EF
DAS4700–2 Configuration Planning Guide,	86A186EF
DAS5300 Deskside Models – Installation and Service,	86A101EF
DAS5300 Rackmount Models – Installation and Service,	86A102EF
DAS5700 Series Standby Power Supply (SPS) – Installation and Service,	86A148KX
DAS Subsystems DC SPS (Direct Current Standby Power Supply) – Installation and Service,	86A120KX
Planning a DAS Installation – Fibre Channel Environments,	86A194JX
DAS – Configuring and Managing a DAS – AIX Server Setup,	86A220PN
DAS – Navisphere for AIX – Setup and Operation,	86A247KX

Navisphere Manager Installation and Operation	86A204EF
Navisphere 4.X Supervisor Installation and Operation	86A205EF
EMC Navisphere 4X Agent Installation	069000880
EMC Navisphere ATF	069000980

Fibre Channel adapter

PCI Fibre Channel Adapters Installation and Configuration Guide	86A195HX

S@N.IT! V2 (formerly known as ASM)

S@N.IT! User's Guide	86A259EF
SRB	86A286EE

Brocade

Silkworm 2000 Entry Family	53–0000010
Silkworm 2800 Hardware Reference Manual	53–0001534
Fabric OS TM Version 2.2	53–0001560
Fabric OS TM Version 2.0	53–0001535

Symmetrix

Symmetrix Storage Reference Guide	86A185KX
PowerPath V2.0 for Unix Installation Guide	300–999–266
Volume Logix V 2.3 Product Guide	300–999–024
EMC Fibre Channel Interface V2.0 for AIX Product Guide	200–999–642
SymmWin V5.0 Product Guide	300–999–074
Conn Enterprise Storage Network System, Topology Guide	300–600–008
Enterprise Fibre Director Model 1032 User's Guide	300–600–004
EMC ControlCenter Integration Package for Windows and Unix Version 4.1 Product Guide	300–999–087

Microsoft Windows NT

Server Setup	014002899

Table of Contents

Chapter 1. Introduction to Storage Area Networking

This chapter describes in general terms the benefits, the challenges, and risks of using a Storage Area Network (SAN).

Overview

The term SAN designates a new type of storage architecture in which the storage systems are attached to a high speed network dedicated exclusively to storage. It involves a whole new network totally distinct from existing communication networks, as is illustrated in Figure1. The application servers (usually UNIX or Windows NT based) access the storage resource through the SAN. Most of the local storage resources are off–loaded from the applications servers, are managed separately from them, and are consolidated at the data centre, site, or enterprise level.

The SAN architecture is represented below:

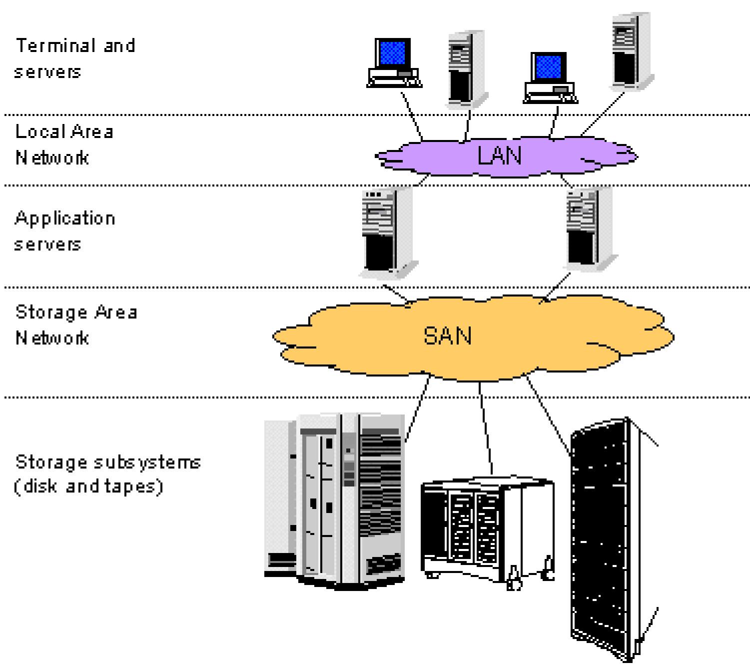


Figure 1. SAN architecture

In this document, the term SAN refers both to the high speed network infrastructure and the whole storage architecture, including servers, storage subsystems and management software.

Fibre Channel is currently the preferred technology for implementing SAN architecture. The technology is open, widely accepted, and is defined by ANSI standards. To simplify the migration and integration of legacy equipment, SAN infrastructure based on Fibre Channel technology supports multiple protocols. For example, the infrastructure can convey SCSI protocols, widely used in UNIX and Intel based servers, ESCON for IBM mainframes, and IP to offer networking capability. But the purpose of SAN is not to replace LANs. All these protocols can simultaneously use the same cables.

This new storage architecture is very different from traditional architectures, where each storage system is connected to a single or sometimes to a limited group of servers. That is why they are sometimes called private storage architectures. SAN creates a revolution similar to the one created by LAN deployment in the last decade. SAN is a powerful solution that enables data storage management of an unprecedented efficiency and flexibility.

The benefits of a SAN are:

Cost reduction: The cost of storage systems depends on the cost of the disks, the cost of the enclosure, and the cost of management software. In a private storage architecture, it is usually necessary to buy many small or medium size enclosures.

With SAN architecture, it is possible to reduce the number of enclosures by storage sharing through the network. Also, disk space in the enclosures can be shared.

Another factor in cost reduction is the performance of the high–speed network, which allows the use of larger storage systems that are usually less expensive (per Gb) than several smaller systems.

Easy administration: The separation of servers and storage resources enables the storage systems and their management applications to be centralized, consolidated, and homogenized.

Storage administrators do not need to be trained for server operating systems and management software. Management is simplified (less variety and number of equipment, less tools, less consoles), and consolidated (aggregation of information from several storage systems).

Failure detection and prevention, and trend analysis is much easier, because of the lower number of systems to be monitored.

High security and availability: SAN improves the availability and security of the storage. SAN offers a meshed infrastructure with multiple paths. Redundant interfaces for servers and storage systems provide highly available data paths between servers and storage. Because of networked access, the storage remains available even in the case of server failure. The cost reduction achieved by storage concentration justifies the deployment of RAID solutions, which can be used by any server.

The systems themselves can be more secure, with internal high availability features (hot plug, redundancy). The reduced number of storage systems to manage facilitates backup operations.

Disaster Recovery solutions also increase the availability and security by monitoring master and distant backup sites.

Improved performance: SAN improves the performance in various ways:

- any application server can be connected directly to the storage systems at (multi–)gigabit speed. Compared to today's network storage solutions, based on protocols operating on LANs (NFS, DFS, and SMB), data access is greatly improved
- data can be shared between hosts, eliminating multiple copies
- performance is boosted by the use of high–end storage systems with a high internal bandwidth, larger read and write caches, and efficient processing engines.

Quick evolution capability:	The storage concentration enables very quick adaptation to the evolution of information processing. The storage allocation is extremely flexible and scaleable. You can migrate applications from server to server, add or remove servers without any impact on the data: no need to move data from one storage system to another, no need to add disks. The data remains at its location, only routing and access rights need to be modified. Administrators can add storage without stopping application servers.

Challenges and Risks

Designing a SAN is like designing a LAN: it is a network, and the design must integrate analysis of traffic requirements, the number and location of devices to be connected, topology constraints, and other network related factors. The design must also take into account the availability constraints, which are much higher for a SAN than for legacy LANs. Without redundant SAN architecture, failure of a Fibre Channel (FC) component may stop tens of applications servers, resulting in no access to their data.

It is necessary to understand the behaviour of infrastructure equipment such as hubs, switches, bridges, and all the proprietary features implemented by hardware vendors to simplify interoperability, improve reliability, and manage the infrastructure. This knowledge must be combined with a good understanding of all the proprietary fail–over solutions supported by each server, disk array, tape library, SAN infrastructure device and management software.

FC technology is not yet mature enough to guarantee interoperability between any device as does LAN technology. Thus, the design must take into account the qualified links, or certified interoperability, published by all SAN vendors, including Bull.

The last risk is that much equipment and software (like operating systems for servers) have not yet integrated the SAN dimension. They still behave like private SCSI storage, trying to discover all the storage resources, or writing to disks when they recognise an OS compatible format. Sharing storage resources in a heterogeneous server environment must be carefully analyzed and managed.

SAN management

SAN management covers aspects not covered by previous storage management tools:

- SAN infrastructure (hubs, switches, bridges)
- resource sharing.

Managing SAN infrastructure is similar to managing a LAN. It entails managing the hubs, switches, bridges, and any other devices in the network infrastructure, analysing the topology, the status of links, managing virtual networks (using zoning), and monitoring data flow, performance metrics, QoS, and availability. The product offering is still not complete, due to the immaturity of the technology, to the lack of standards, and to the small size of hardware vendors of SAN infrastructure. No single vendor provides a completely integrated and dedicated solution that includes all the necessary devices plus applications covering those devices.

Managing resource sharing is also critical to safe operation in SAN environments. SAN enables each server to access each tape drive or each disk logical volume attached to the SAN. It is very flexible, but very complex for the administrators of the servers, because they may have to manage hundreds of resources, duplicated on all the applications servers. That is why it is critical, even in a small SAN (less than 8 servers), to deploy management tools to control resource sharing. Again, there is not yet a dedicated solution. Tape drive allocation and sharing is usually managed by the backup software, which must integrate specific SAN features. For the disk resources, LUN access control technology can be used at either server, or disk array level. Zoning can also be used to create virtual networks that simplify management.

Chapter 2. SAN Infrastructure

This chapter discusses the concepts and technologies applied in SAN infrastructure.

FC is the only technology available on the market that allows SANs to be deployed with a high level of performance. It is a highly reliable, gigabit interconnection technology that allows concurrent communications between servers, data storage systems, and other peripherals using SCSI and IP protocols. One of the advantages of FC technology is the high flexibility of the network design: the use of switches and hubs enables you to create large infrastructures with a high level of performance and interconnection of various equipment.

Media speed

FC operates at various frequencies. The current de–facto standard is 100 MB/s, full–duplex, per physical connection. Equipments operating at 200 MB/s (2 Gbit/s) are available now. 100 MB/s and 200 MG/s equipments can be mixed on the same SAN, but only 200 MB/s capable equipments directly connected one to each other will operate at 200 MB/s. But it is not the ultimate evolution of this technology: products operating higher speed are emerging. The 400 MB/s speed is standardized, but no product is available today.

Cabling

The FC standards define several media. The most frequently used solutions and their characteristics are described in the next table.

Media	Maximum cable length	Typical connector	Cable structure
62.5 µm multi–mode optical fiber	175 m	SC	2 fibers
9 µm single mode optical fiber	10 km	ST, SC	2 fibers
copper	30 m (active interface) 3–15 m (passive interface)	DB9, HSSDC	4 wires
50 µm multi–mode optical fiber	500 m	SC	2 fibers
50 µm multi–mode optical fiber	300 m (2 Gbit/s)	LC	2 fibers

- 50 µm multi–mode optical fibers is the preferred cable for FC. When used at 2GBit/s, the lenght is limited to 300 m.
- 62.5 µm has been introduced to re–use LAN cabling, where this type of fiber is widely used.
- Single mode fibers are reserved for specific applications, such as long distance interconnection for site interconnection and/or disaster recovery solutions. Using more powerful lasers and/or more sensitive receivers, this type of cabling enables connections over greater distances than the 10 km defined by the standard.
- Copper cabling has been widely used due to its lower cost. It is now less and less used, because the price of optical equipment has rapidly decreased, and because of the sensitiveness of the copper technology to EMI, ESD and ground problems. The copper cables are specific to FC technology, and are not compatible with UTP or STP cables used in LANs, nor with the copper cables used in telco.

FC Networks

FC networks are implemented in conformance with the ANSI standards, and use some specific network equipment.

Point–to–point connection

The point–to–point connection between two devices is the simplest cabling solution. The cable types and distance limitations are described in Cabling, on page 2-1.

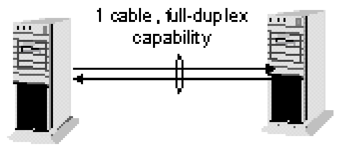


Figure 2. Point–to point cabling

Loops and hubs

The loop has been very popular since the beginning of FC networking. It enables interconnection of up to 126 devices at very low cost. The principle of this topology is to link the transmitter of each port to the receiver of the next one, thus creating a loop. The loop protocol is defined by the FC–AL standard.

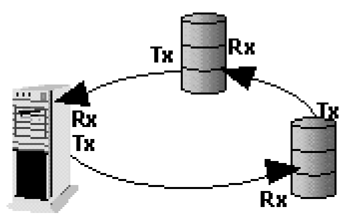


Figure 3. Loop cabling

The loop has several main drawbacks:

- The bandwidth is shared by the connected devices and only two devices can communicate simultaneously.
- Each device in the loop is a potential point of failure: if the device fails, the loop is broken and none of the connected devices can communicate.
- Inserting or removing devices causes loop disruption.
- No hub equipments at 200 MB/s are available yet.
- Connectivity is limited to 126 devices.

The bandwidth and connectivity limitations are inherent in the protocol used. The risk of loop disruption is addressed by using FC hubs, which implement a loop within the device in conjunction with bypass circuits for each hub's port. The bypass circuits maintain the loop integrity in the case of unused ports or faulty external devices. They also simplify the cabling, converting loop cabling to a star cabling, as illustrated below.

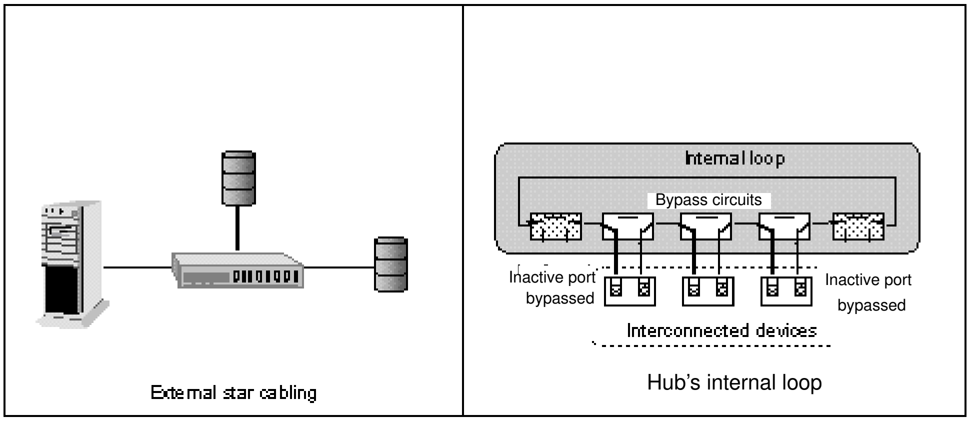


Figure 4. Star cabling

In practice, a hub may contain one or more logical loops. The distribution of ports in several loops is performed either with hardware switches that control loop continuity or loopback, or through software configuration.

Fabrics and switches

A fabric is a network where, contrary to loops, several pairs of devices can communicate simultaneously, and can use the whole bandwidth. The FC fabric can be compared to other switched networks, like Ethernet, ATM or frame relay networks. Each switch's port analyses the frame headers, selects the destination port, and puts the frame in the appropriate transmission queue. Non–blocking designs and advanced forwarding algorithms allow a very low latency (a few micro–seconds), and a full bandwidth per port, regardless of effective traffic conditions. The fabric and switch protocols are defined by the FCP, FC–AL, FC–FLA, FC–PLDA standards.

The advantage of switches versus hubs are:

- switches enable the deployment of large networks,
- switches offer higher performances and scalability (no bandwidth sharing, selection of the best path),
- switches provide better fault isolation,
- switches and fabrics are more stable than loops (less reconfigurations and traffic interruptions),
- they have a built–in management function.

The disadvantage is usually a higher price per port.

The smallest FC fabric consists of a single switch. Larger fabrics can be interconnected by cascading switches. The protocol permits the connection of millions of devices.

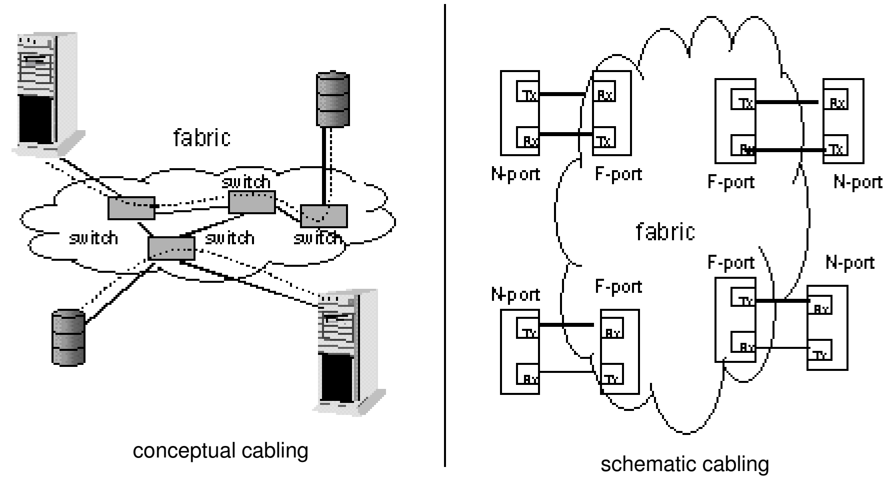


Figure 5. Fabrics and switches cabling

Integrated routing algorithms automatically route circuits and frames across the fabric, through the best path. Resilient multi–switch fabrics can be deployed, with automatic data routing and re–routing, self–healing and almost unlimited scalability.

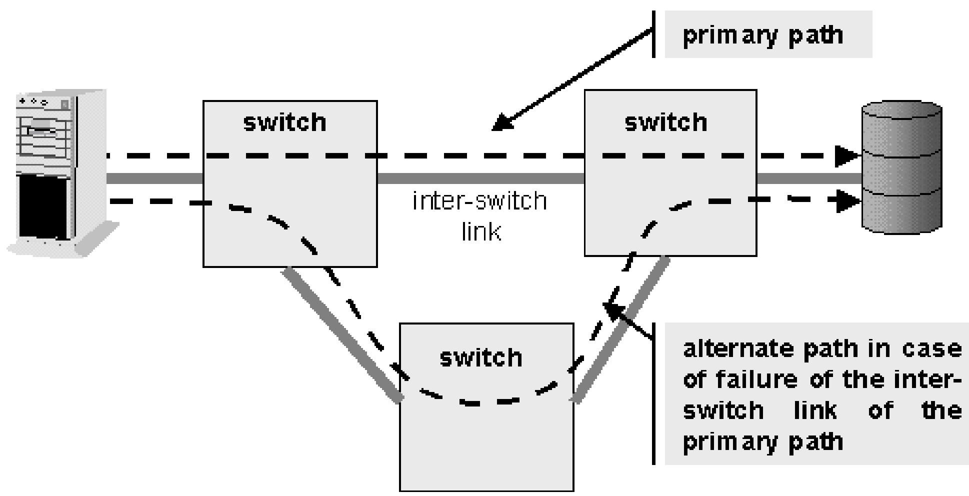


Figure 6. Switches: automatic data routing

Multiple inter–switch links can be aggregated, to avoid the risk of a bandwidth bottleneck which may arise when using a single link operating at the same speed as the switch ports.

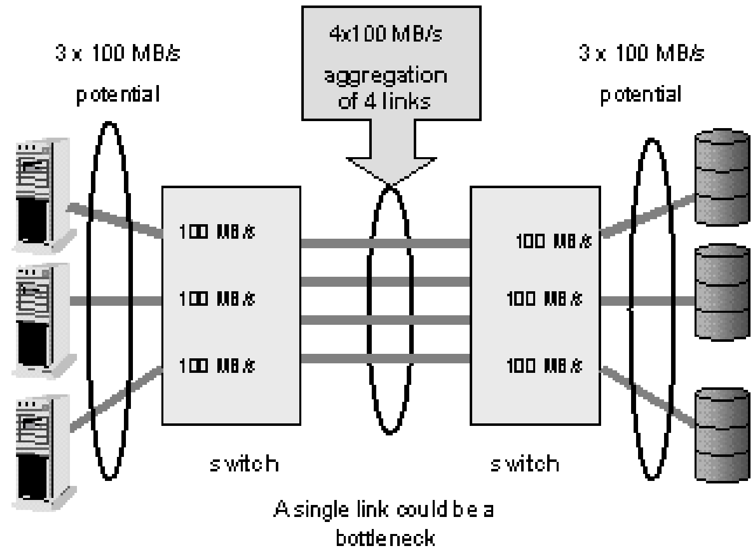


Figure 7. Switches: bandwidth

The speed of the links could be either 100 MB/s or 200 MB/s depending on each equipment speed support. The speed is auto–negociated.

In the future, higher speed links (400 MB/s or even more) will allow the building of hierarchical network infrastructures, with multi–gigabyte core backbones.

Interconnection of fabrics and loops

Fabric and switch ports are quite expensive. Thus, the FC standards allow mixing of switches and hubs. A switch port can be shared between multiple devices through a hub connected to the switch port. They share the 100 MB/s or 200 MB/s offered by the switch port. Such devices are called public loop devices. They can communicate with any other devices of the fabric or with devices of other public loops. The public loops are always connected at the edge of a fabric. Loops and hubs cannot be used to interconnect switches.

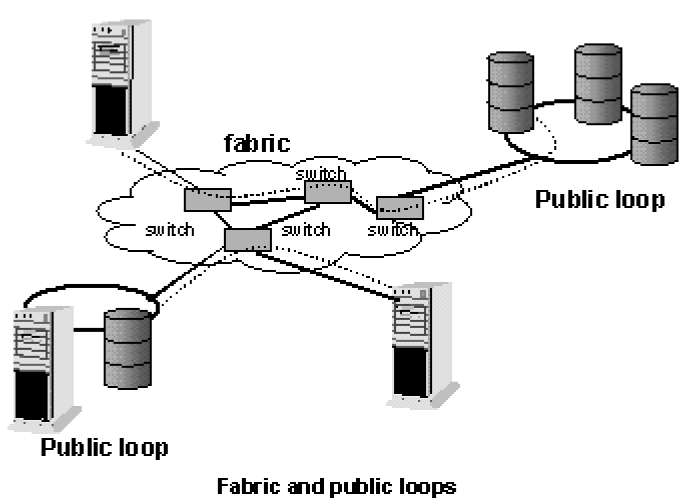


Figure 8. Fabrics and loops interconnection

Zoning, switched loops, mixed private loops and fabrics

Using these basic topologies and protocols to operate in private or public loop and fabric environments, various features have been implemented by SAN infrastructure vendors. These features are proprietary, usually compatible with FC standards, and bring new interconnection and operation capabilities.

Fabric zoning

Fabric zoning is a feature that enables the creation of several logical areas on a unique physical FC network. Devices such as servers or disk subsystems can communicate only with other devices in the same zone. With some implementations, zones can overlap, allowing a device to belong to several zones.

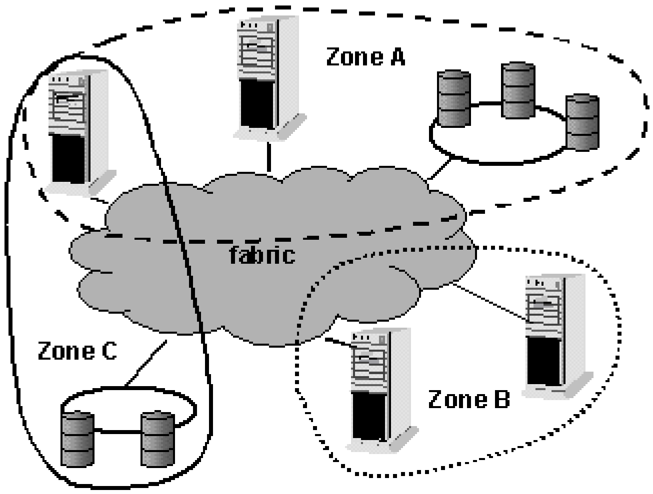


Figure 9. Fabric zoning

Zoning can be used as a security service, preventing some resources from being used by non–authorised users (for example, a disk subsystem with confidential data can be isolated in a zone accessible only to certain users).

Zoning can also be used to solve some interoperability problems: a zone can be defined for all the servers and storage resources used by AIX servers, another zone being dedicated to Windows NT servers and storage resources.

Loop zoning

Loop zoning allows the creation of several independent loops through the same hub. The zoned loops are usually defined by a list of hub ports. Contrary to fabric zoning, it is usually not possible to define overlapping zones, nor zones across multiple interconnected hubs. The zoned loops can be enforced at the hardware level (by splitting the hub’s internal loop), or emulated by software.

Switched loop

A switched loop is the synthesis of switching technologies and FC private loop protocol. It enables private loop devices, which do not support fabric connection, to benefit from improved performances compared to a simple loop. This technology may be used temporarily to enable optimization of the performances of current hardware.

With this technology, several small private loops, called looplets, are interconnected by a switching device called a switched hub or storage switch. The device creates a single logical private loop compliant with the FC–AL standard (with up to 126 devices), but enables several simultaneous communications between pairs of ports, as does a fabric switch. The device intelligently controls loop protocols, like initialisation and arbitration protocols, and enables several devices to acquire the loop simultaneously as long as they are not in the same looplet.

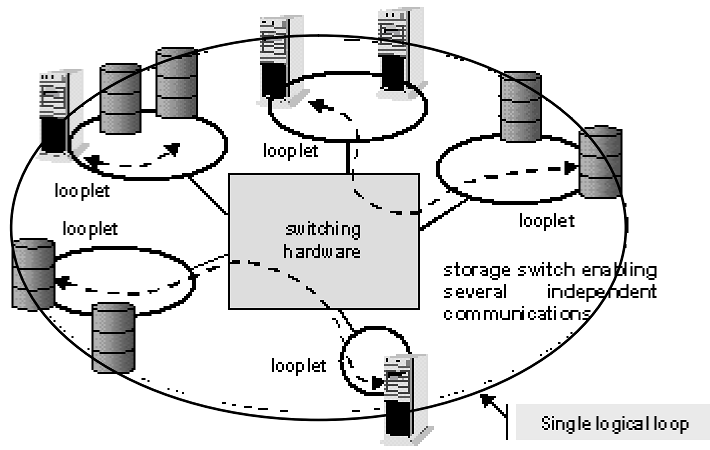


Figure 10. Switched loops

This feature is available in managed hubs relying on an internal switching architecture, or can be emulated on fabric switches. Depending on the implementation, the logical loop can be limited to a single hub/switch, or can spawn across multiple hubs/switches.

A storage switch increases the performance compared to hubs without using fabric protocols, but it is more limited than a fabric in term of addressing, interconnection, and routing.

Mixing private loops and fabrics

The connection to a fabric, either in point–to–point or through hubs requires specific protocols which are not yet available on all devices. Thus, fabric switch vendors have implemented various proprietary solutions to support, on the same hardware, full fabric operations, private loop emulation, and bridging capabilities between fabric capable devices and private loop devices.

The characteristic of each hardware device must be carefully checked before mixing these features in a SAN implementation.

Interconnection with Legacy Technologies

FC/SCSI bridges

FC/SCSI bridges provide an FC interface for SCSI devices. A bridge can connect servers or storage subsystems that do not have FC interfaces.

There are two modes of operation:

- FC initiator (e.g. server with FC adapter) to SCSI target (e.g. library).

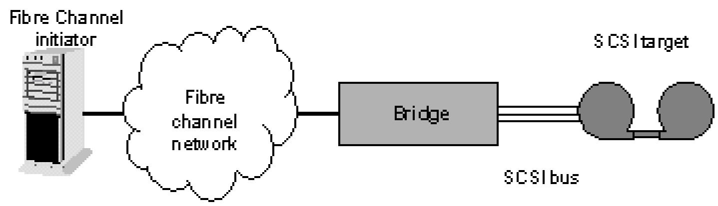


Figure 11. FC/SCSI Bridges: FC initiator

- SCSI initiator (e.g. server with SCSI adapter) to FC target (e.g. FC disk subsystem). This mode is not advised due to its low performance.

The bridges have a lot of implementation dependant features, notably:

- support of SCSI or FC initiators, or both,
- number of SCSI busses and FC interfaces,
- supported FC topologies (point–to–point, fabric, private or public loop),
- mode of operation SCSI/FC address translation,
- device discovery,
- performance level.

FC/IP bridges

The FC/IP Bridge devices interconnect legacy IP networks (Ethernet, FDDI, ATM) to an FC network supporting IP. The market for such devices is very small, because FC is now positioned as a storage technology (SAN architecture), and no longer simply as a potentially high speed LAN.

FC/WAN bridges

FC supports long distance links, but it is often difficult to deploy a private cabling in the public domain.

Thus proprietary devices have been designed by hardware vendors to interconnect SAN islands using services from telco operators like ATM or E3/T3. Some of these devices implement real fabric services or loop emulation, and are fully transparent to the FC protocols. Others support only SCSI over FC, and behave like FC/SCSI bridges.

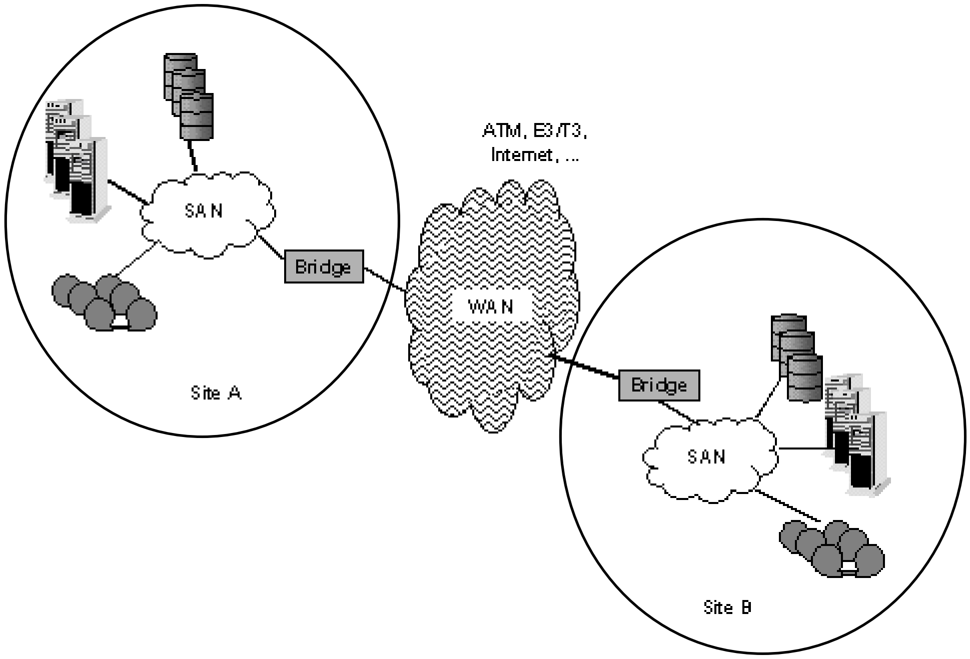


Figure 12. FC/WAN Bridge example

In–depth analysis is mandatory before integrating such devices in SAN solutions.

Chapter 3. SAN Components

This chapter describes the basic hardware and software components of the Bull Storage Area Network (SAN) offer and lists the specific requirements for these components.

SAN hardware components

The basic hardware components of the Bull SAN offer are:

Host servers

Escala PCI servers : E series, T series, EPC and PL series, running AIX 4.3 and later.

Express 5800 servers running at least Windows NT4/SP6A.

Host Bus Adapters (HBA)

FC Host Bus Adapters connect each server to the SAN network. They must be provided by Bull.

Number of adapters per server depends on the type of server: a minimum of 1 adapter is supported per server, a maximum of 16 on the high range.

The following types of adapters are supported:

– Emulex LP7000E with copper interface,

– Emulex LP8000 with copper or optical fiber SC2 interface,

– Emulex LP850 with copper or optical fiber SC2 interface.

– Emulex LP9002L with optical fiber LC2 interface,

Mid–range disk array sub–systems (DAS products)

The DAS family supports point–to–point, FC–AL and fabric FC topologies.

– The DAS 4500, 4700, 5300, 5700 and DAE 5000 Series storage subsystems constitute a powerful, highly scalable, family of FC storage solutions. They provide JBOD or RAID capabilities, terabytes of disk storage capacity, high transfer rates, flexible configurations, and high availability.

– DAS 4500 and DAS 4700 are the high–end models of the DAS product line. They can run 2 core software applications: Direct Attach or Access Logix which provide LUN masking capability. Storage Processors (SPs) are equipped with optical GBIC interfaces.

– DAS 4700–2 is the new 2 Gbit/s DAS model. It operates at 1 Gbit/s in a 1 Gbit/s network and at 2 Gbit/s in a 2 Gbit/s network.

– DAS 5700, 5300, 3500 are the medium and low end models of the DAS product line. SPs are equipped with copper interfaces.

– All these sub–systems are fully redundant, having dual SPs with mirroring cache, dual power supply and dual standby power supply. They support from 10 to 60 or 70 fibre disk drives.

– Disk drives are FC 10krpm drives with a capacity of 9GB, 18GB, 36GB or 73GB.

High–range disk array sub–systems (Symmetrix products)

The Symmetrix family is the high end subsystem of the Bull FC offer. It offers the best features in term of scalability, connectivity, and information protection and sharing.

The Symmetrix family includes:

- Symmetrix 3000 with models Symmetrix 3930 (256 disks), 3830 (96 disks) and 3630 (32 disks).
- Symmetrix 8000 with models Symmetrix 8830 (384 disks) and 8530 (96 disks).

All models support ultra–SCSI and multi–mode fibre directors. Specific Remote Link Directors offer Symmetrix Remote Data Facility (SRDF).

All models support FC point–to–point, loop, and fabric topologies at 1 Gbit/s and 2 Gbit/s.

FC Switches

FC switches enable the building of FC fabrics by cascading the switches.

The following models of switches are supported in the Bull SAN offer:

- Silkworm 2040:

  The Silkworm 2040 is an entry level FC switch designed to address the SAN requirements of small workgroups, server farms and clusters. It does not support cascading.

  The Silkworm 2040 is an 8 port switch, including 7 ports with fixed Short Wave Length optical ports (for multimode optical fibers), and 1 configurable GBIC based port.
- Silkworm 2800:

  The SilkWorm 2800 switch is a 16–port FC (GBICs) switch that provides connectivity for up to 16 FC compliant device ports.
- Silkworm 3800:

  The SilkWorm 3800 switch is a 16–port FC (SFP) switch that provides connectivity for up to 16 FC compliant device ports at 1 or 2 Gbit/s.
- Connectrix: 32–port GBIC switch, full high availability architecture.

FC hubs

FC hubs enable the building of a FC Arbitrated Loop (FC–AL) configuration. The following hubs are supported:

- Gadzoox hub:

  The Gadzoox Bitstrip™ TW FC Arbitrated Loop (FC–AL) hub is specifically designed to be a cost effective entry point into high availability, gigabit storage networking. It is a 9–port hub equipped with fixed copper DB9 ports.

  This hub is being replaced by the Vixel 1000 hub.
- Vixel 1000:

  The Vixel 1000 is also an unmanaged hub equipped with a 7–port optical or copper GBICs.
- Vixel 2100:

  The Vixel 2100 is an 8–port managed hub with optical and copper GBIC.

Gigabit Interface Converters (GBICs)

GBICs are interface modules that connect the different devices (switches, hubs, SPs) to the FC network.

There are three types of GBIC:

- Shortwave Optical GBIC: allows connection up to 500m via a 50µ multimode cable (SC2).
- Longwave Optical GBIC: allows connection up to 10 kms via a 9µ monomode cable (SC2).
- Copper Active DB9 GBIC: allows connection to a copper cable.

GBICs do not exist for all devices. Refer to the respective specifications for specific information.

The GBICs used on each device must be the one shipped with it – they are not interchangeable.

SFP (Small Form Factor Pluggeable Media)

SFPs are interface modules that connect 2 Gbit/s capable devices (switches) to the FC network. They offer a dual LC (LC2) connector.

2 types of SFP are available:

- shortwave optical SFP: up to 300 m via a 50µm multi–mode optical fiber,
- longwave optical SFP: up to 300 m via a 9 µm single mode optical fiber.

MIA (Media Interface Adapter)

A MIA is an externally mounted converter for connecting an optical Dual–SC (SCù) fiber cable to a native copper–interface device.

A MIA is not allowed on “DB9–GBIC”, or on LP8000 copper adapters.

SAN software components

The basic software components are:

FC driver

Allows communication between the operating system and the HBA.

The driver is delivered on the Bull Enhancement CD–ROM or on the Bull Addon CD–ROM for AIX, and with the server attachment kit for Windows.

DAS software products

Navisphere Manager for AIX

Allows the configuration and management of Fibre DAS5300 and DAS57x0 disk array sub–systems. It runs on an AIX server or PowerConsole and communicates with the Navisphere Agents located on the servers.

It offers distributed management to manage all the DAS5300 and 57x0 arrays of the site connected to AIX servers from a central console.

Navisphere Manager for AIX is delivered on the Bull Enhancement CD–ROM for AIX.

Navisphere Manager for Windows NT

Allows the configuration and management of all types of FC disk array sub–systems DAS5300, 57x0, 4500, 4700 Direct Attach and Access Logix (mandatory for DAS4500 Access Logix and DAS4700). It communicates with the Navisphere Agents located on the servers.

Offers distributed management of all the arrays of the SAN from a central console.

Navisphere Manager for NT is delivered either on the Bull Enhancement CD–ROM or on a specific CD–ROM.

Navisphere Supervisor for Windows NT

Allows the configuration and management of FC disk array sub–systems DAS5300, 57x0, 4500 Direct Attach. Allows the management of only one DAS connected to one server at a time. Does not allow the management of DAS4500 Access Logix and DAS 4700.

Navisphere Supervisor for Windows NT is delivered on the Bull Enhancement CD–ROM for NT.

Navisphere Agent

Ensures communication, through the internet link, between the Navisphere Manager (Manager for AIX or Supervisor for Windows NT) and the Licensed Internal Code (LIC) on the SPs.

Navisphere Agent for AIX is delivered on the Bull Enhancement CD–ROM for AIX.

Navisphere Agent for Windows NT is delivered on a specific CD–ROM.

Navisphere Event Monitor for Windows NT

Notifies the administrator when selected storage system events occur. It must be installed on the same machine as Navisphere Supervisor.

Event Monitor is delivered on a separate CD–ROM.

Navisphere Analyzer

Collects graphs and archives performance data for FC arrays.

ATF (Application Transparent Failover)

Failover software installed on a server connected to the DAS and which guarantees access to data on the storage system when one SP fails. ATF is delivered on a separate CD–ROM.

Access Logix

Provides data protection, shared storage access, and security. Allows the management and control of heterogeneous data access in distributed, extended storage networks.

SnapView (DAS 4700)

Creates snapshots of production data, enables fast restoration of that data, and enables nondisruptive backups.

MirrorView (DAS 4700)

Provides synchronous mirroring of critical data and, by enabling failover to a secondary site, allows fast disaster recovery.

Symmetrix software products

Also see page iv for details of Symmetrix documentation.

Symmetrix software provides a variety of information protection/business continuance options including mirroring, the optimum RAID level for both performance and availability.

The following software offerings supplement the Symmetrix Enterprise Storage philosophy:

Symmetrix Remote Data Facility

Symmetrix Remote Data Facility (SRDF) provides fast enterprise–wide data recovery in the event of a data center outage.

TimeFinder

TimeFinder supports the online creation of multiple independently addressable Business Continuance Volumes (BCVs) of information allowing other processes such as backup, batch, application development and testing, and database extractions and loads to be performed simultaneously with OLTP and other business operations.

Symmetrix provides centralized, sharable information storage that supports changing environments and mission–critical applications. This leading–edge technology begins with physical devices shared between heterogeneous operating environments and extends to specialized software that enhances information sharing between disparate platforms.

Symmetrix Enterprise Storage Platform

Symmetrix Enterprise Storage Platform (ESP) software provides simultaneous mainframe and open systems support for Symmetrix systems.

InfoMover

InfoMover extends information sharing by facilitating high–speed bulk file transfer between heterogeneous mainframe, UNIX, and Windows NT host platforms without the need for network resources.

Symmetrix systems improve information management by allowing users to consolidate storage capacity for multiple hosts and servers. Powerful tools are offered and it dramatically simplifies and enhances configuration, performance, and status information gathering and management.

ControlCenter

ControlCenter™ is a family of applications providing extensive management tools. ControlCenter’s monitoring, configuration, control, tuning, and planning capabilities simplify management of the Symmetrix supporting open systems and mainframe environments, while keeping track of changing requirements.

PowerPath

PowerPath™ optionally offers a combination of simultaneous multiple path access, workload balancing, and path failover capabilities between Symmetrix systems and supported server hosts.

Volume Logix

Volume Logix allows multiple UNIX and Windows NT server hosts to share the same Symmetrix Fibre Adapter port while providing multiple volume access controls and data protection. The Volume Logix software also provides volume access management for non–supported Volume Logix GUI and Host Utilities platforms that are connected via FC.

Symmetrix Optimizer

Symmetrix Optimizer provides automatic performance tuning for Symmetrix by balancing performance and eliminating potential bottlenecks. It reduces the effort and expertise required to achieve optimal Symmetrix performance by automatically balancing I/O activity across physical drives.

Brocade Silkworm Software

Brocade WEB TOOLS

- Brocade WEB TOOLS provide a graphical and intuitive management interface to monitor and control the Silkworm switches.

Brocade Telnet interface

- Brocade Telnet interface enables access to a command line interface allowing access via SNMP to the switch MIB.

Brocade SES

- Brocade SES (**S**CSI **E**nclosure **S**ervices) is also offered to enable easy integration in management applications.

Brocade Zoning

- Brocade Zoning allows the restriction of access to devices by creating logical subsets of devices for additional management of, and security in, the SAN.

S@N.IT! (formerly known as ASM)

S@N.IT! enables storage administrators to efficiently manage SAN for AIX and for Windows NT servers. S@N.IT! is a multi–tiered administration tool, which consolidates the management of multiple AIX and Windows NT servers, SAN, and RAID disk subsystems.

The S@N.IT! greatly simplifies management tasks by providing:

- server–centric view of SAN infrastructures and disk subsystems
- graphical display of SAN topology, with automatic refresh of the display whenever configuration changes occur (new components added, component status changed)
- consolidation of information between other management tools: AIX, Open Symmetrix Manager, Navisphere for DAS
- centralization of event monitoring (faults detected on disk subsystems and switches are reported in S@N.IT!)
- host–based LUN Access Control to hide unauthorized resources from AIX and Windows NT servers.

Amongst other advantages, S@N.IT!:

- is ideal for storage consolidation (no need to see resources used by other servers), in heterogeneous AIX and Windows NT server environments
- makes EMC/Typefounder usable
- enforces data confidentiality policies
- reduces the risk of administrator errors.

The S@N.IT! application comprises three functional parts:

- SAN agents that discover information about SAN components and perform actions on the hosts and subsystems: there are SAN agents for both AIX and Windows NT systems
- client managers that are responsible for the end user interface and are used in the WebSM server (AIX platforms only)
- a central agent that gathers information from SAN agents and answers requests from client managers (AIX platforms only).

Hardware and software minimum requirements

Component	Minimum requirement
AIX Server	AIX 4.3.3 FC adapters DCCG141–A000 : firmware 3.02x1 minimum DCCG147–0000 : firmware 3.20x4 DCCG148–0000 : firmware 3.20x4 DCCG154–0000 : firmware 3.82x1 FC driver : 1.3.3.31
Windows NT server	Windows NT4.0, SP6 FC adapters Emulex LP8000 : firmware 3.20x4 Emulex LP850 : firmware Fibre Channel driver : port driver 1.27a3
DAS sub–systems	DAS 5300 Direct Attach: flare/PROM 5.24.05/2.09 DAS 5700: flare/PROM 5.11.14/3.23 DAS 5720 (capped): flare/PROM 5.11.64/3.23 DAS 4500 Direct Attach: flare/PROM 5.32.13/2.09 DAS 4500 Access Logix: flare/PROM 6.32.13/2.09 DAS 4700 Direct Attach: flare/PROM 8.43.03 DAS 4700 Access Logix: flare/PROM 8.43.53 DAS 4700–2 Direct Attach: flare/PROM 8.43.03 DAS 4700–2 Access Logix: flare/PROM 8.43.53 DiskArray 4.13.5 FcDiskArray 1.0.1.4 ATF for AIX 2.1.14 for AIX V4, 2.2.3 for AIX V5 Navisphere Agent for AIX 1.2.3 forr AIX V4, 1.3.0 for AIX V5 Navisphere Manager for AIX 1.1.6.6 ATF for Windows NT 2.0.4 Navisphere Agent for Windows NT 5.2.5 Navisphere Manager for Windows NT 5.2.5
Symmetrix sub–systems	Symmetrix 3000 : microcode 5566 Symmetrix 8000 : microcode 5566 PowerPath for AIX TimeFinder for AIX ECC (EMC ControlCenter) Symmetrix Optimizer for AIX PowerPath for Windows NT TimeFinder for Windows NT Symmetrix Optimizer for Windows NT
S@N.IT!	Windows: 3.0.0.3. AIX: san_it.all 1.0.5.11 & san_it.devices 1.0.5.3
Switches	Silkworm 2040 and Silkworm 2800 Firmware 2.1.6 or 2.2
Hubs	Gadzoox Bitstrip Vixel 1000 Vixel 2100

Note: Refer to the System Release Bulletin (SRB) delivered with the Bull Enhancement CD–ROM, or to the Release Notes delivered with the products to get the current versions.

Chapter 4. Methodology to Build SAN Configurations

This chapter describes a method of building a SAN in a logical and systematic manner, which is more likely to lead to a correctly working network at your first attempt.

Overview

The methodology consists of the following sequential steps:

1. collect the data (see page 4-2)
2. size the servers and storage systems (see page 4-3)
3. design the SAN topology (see page 4-4)
4. checking operational constraints (see page 4-6)
5. determine a management policy (see page 4-7)
6. write a consolidated report (see page 4-7).

Step 1 Collect the Data

1.1 Identify all the servers, applications, and data–sets

Server <name>	
Name	**Purpose**
Model	
OS	
New or existing	
Purpose	Production test development backup failover
List of applications	
User data (Gb) in internal storage	
User data (Gb) in external storage	
Trend for 12 months	Flat, X Gb more
Identified data–sets	
Global data backup policy	File, raw volume, snapshot, on–line using an application service
Multipathing for failover (see Note 2. below)	Y/N
Application takeover on another server	Y/N (if Y, which one) Manual/HA–CMP
Remote copy for disaster recovery	Y/N
Distance constraints	
Performance requirements (disk and backup)	
Number of FC HBAs	

Notes: 1. Check the impact of connecting existing server (for example, is an upgrade required?)

2. Multipathing for failover requires at least 2 adapters.

1.2 Get characteristics of each data–set

data–set <name>	
For application	
Purpose	
Backup size (Gb)	
Trend for 12 months	
RAID protection	RAID 1, 5, 10, S, 0
Remote copy	Y/N
Data backup policy (see Notes below)	File, raw volume, on–line (using application service), AIX on–line backup, array–based snapshot
Multipathing for failover	Y/N
Access required from another server	Y/N (if yes, which one) Manual/HA–CMP
Performance requirement, type of traffic	

Notes: 1. When remote copy and/or array based snapshot are required, create additional data–sets.

2. For AIX online backup, double the backup size, but manage a single data–set.

Step 2 Size the Servers and Storage Systems

2.1 Create connectivity matrix and data–set maps

	Array 1	Array 2
Disk array model		
Comments		
Configurability		
number of HDDs (used and maximum)		
number of host interfaces (used and maximum)		
maximum performance per host interface and globally for the array		
Data–set mapping		
data–set 1		
data–set 2		
data–set n		
Host attachment		
Server 1	Number of expected links, redundancy requirement	
Server n		

Notes: 1. The number of HDDs used requires a detailed array configuration (step 2.2).

2. Data–sets created for remote copy or array–based snapshot cannot be connected to the same server as the original data.

3. Host to array connectivity matrix can be described in a separate table.

2.2 Configure each disk array

Disk array: <name>				
Disk groups				
Number of HDDs	RAID protection, BCV	Data–set	Number of LUNs	Port visibility
		Catalog	X	
		Ordering	Y	SP–A (default) SP–B (failover)
Number of HDDs	RAID protection, BCV	Data–set	Number of LUNs	
Total number of HDDs			Total number of LUNs	

Notes:1. Check capacity and addressing limits (number of LUNs, etc.)

2. If necessary, modify data–set placement and iterate the process

3. Array and data–set performance may be estimated (based on number of HDDs)

2.3 Identify hot–spots

This is necessary to be aware of risks and, when possible, to refine performance analysis.

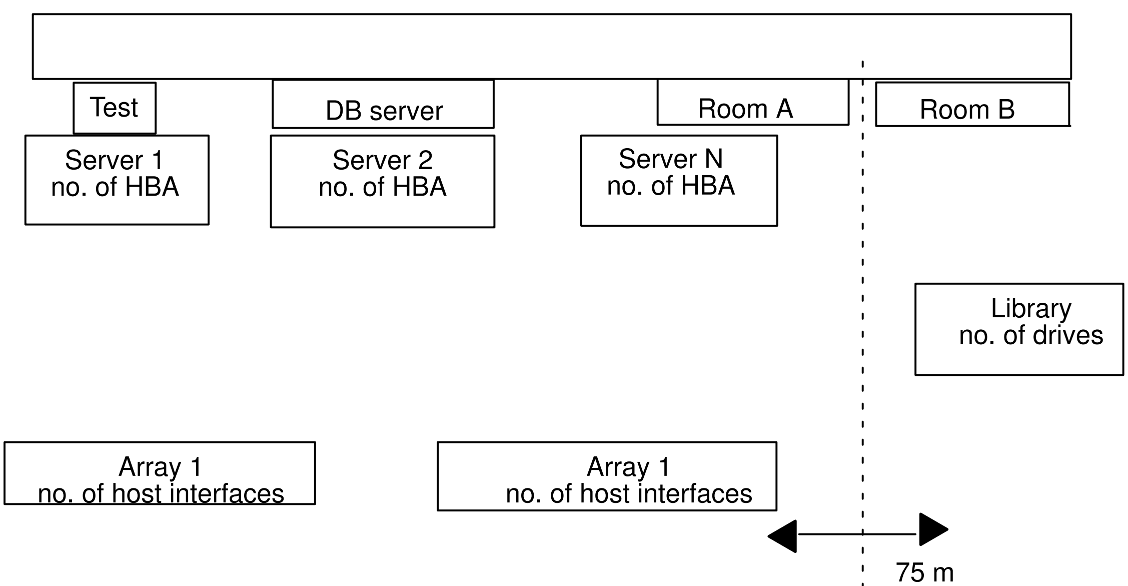
- List all the servers/HBA which must access a large number of target ports (on disk arrays).
- List all the disk array ports used by a large number of servers.
- Try to do some performance analysis for these hot spots.

2.4 Configure the libraries

Library definition – connectivity matrix		
	Library 1	Library 2
Library model		
Configurability		
type of drives		
number of drives (used and maximum)		
number of cartridges (used and maximum)		
connectivity (number of FC and/or SCSI devices)		
Robotic control		
SCSI or through SAN		
name the controlling server		
Host attachment		
Server 1	Client (IP/S) SAN, direct access (SCSI or SAN) used drives	

Step 3 Design the SAN Topology

3.1 Start with an empty map



The following sequence is recommended in conjunction with the above template:

1. Group servers by location and function.
2. Group arrays and libraries by location and function.
3. Identify site boundaries, distance information, rooms, etc.
4. Draw existing SAN infrastructure (in the case of an upgrade).
5. Draw interconnection of arrays (e.g. Symmetrix and SRDF).
6. Identify servers with multipathing software.

3.2 Analyze the structure of the SAN to be designed

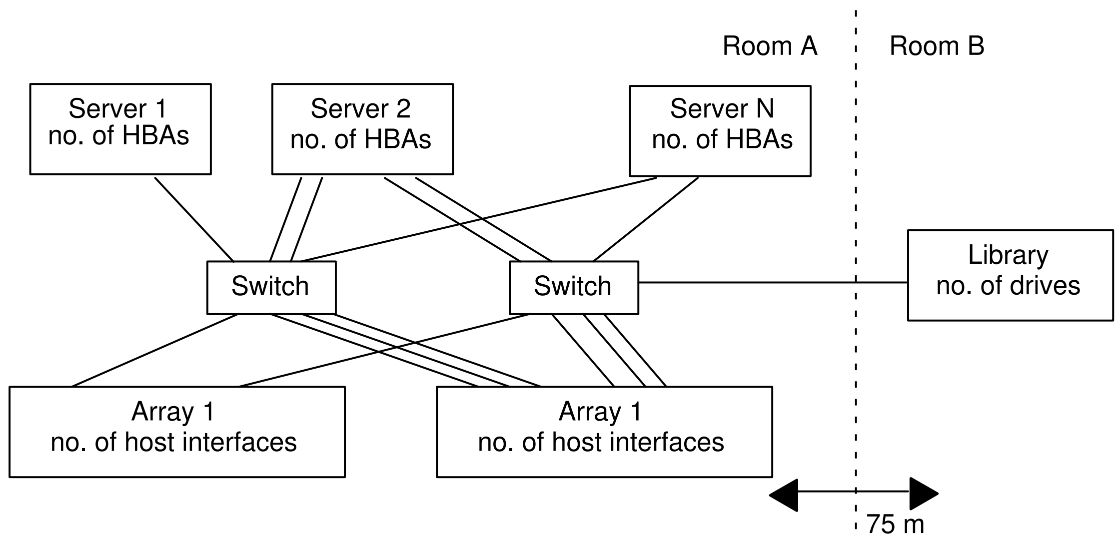
Use the connectivity matrix to determine whether you have independent storage islands.

In the case of independent islands, determine whether you will propose:

- a single SAN, which has the advantages of:
  - having more flexibility for future evolution
  - may require less hardware than multiple islands
- multiple islands, which have the advantages of:
  - they may simplify cabling (e.g. direct attachment vs switches and hubs)
  - they may avoid large multi–switch fabrics.

In the case of multiple rooms or sites, check distance constraints, existing cabling (to be re–used). Determine whether you need switches in each room.

3.3 Interconnect hardware according to your connectivity matrix



When direct attachment of server to array ports is not applicable, design the interconnection using:

- pairs of switches when redundancy is required
- standalone switches
- hubs are acceptable for small configurations.

Add cables to implement the connectivity matrix defined in the data–set mapping table

Iterate (adding switches if necessary) as often as required.

At this step, the SAN infrastruture speed (1 Gbit/s or 2 Gbit/s) as well as media choice (copper, optical SC2 or LC2 connector for fibre channel) must be taken into account.

3.4 Check performance constraints

Check for potential performance bottlenecks of the proposed infrastructure.

If necessary, iterate on SAN topology design.

Step 4 Check Operational Constraints

4.1 Identify zones to be defined in the SAN

Determine whether there is a need for zoning:

- to simulate independent islands
- to isolate areas to meet customer's requirements
- to share the infrastructure with servers and storage not approved by Bull.

4.2 Specify the LUN management tools

Document which LUN access control product will be used:

1. S@N.IT!
2. Volume Logix (Symmetrix)
3. Access Logix (DAS 4500)
4. A combination of all three
5. None (acceptable for 2–node clusters – otherwise risky).

4.3 Check multipathing risks

The interconnection of devices creates more connections than required in the connectivity matrix. Check that these extra connections do not lead to operational problems.

If there are problems:

- use zoning or masking to resolve them

and/or

- modify LUN assignment to array ports

and/or

- modify cabling.

Check list:

- is LUN visible through multiple ports?
- are there multiple copies of the LUN (in the same array, or in other arrays)?
- are there servers with multiple HBA (with or without multipathing software)?

4.4 Consolidate all the information

Create a topology map with all the interconnections and zones.

Draw up a list of the required hardware to prepare ordering with NOEMIE (HBA, hubs, switches, bridges, cables, etc.).

Draw up a consolidated definition of LUN visibility (bindings) to array ports.

Define zones and LUN masking for servers to ensure correct operation of servers.

Step 5 Determine a Management Policy

5.1 Draw the management infrastructure

List all the software to be installed and identify all the components to prepare NOEMIE ordering. The list should include:

- array management
- library management
- SAN infrastructure device management
- backup management
- resource access control
- SAN management
- HA and clustering management
- an updated topology map.

Step 6 Write a Consolidated Report

Prepare a proposal for your customer that includes:

- an explanation of the proposed solution
- a list of hardware and software (derived from NOEMIE)
- the data–set distribution on the arrays
- a topology map
- all the sizing and performance data.

Make sure the following information is available for the installation phase:

- disk array configuration
- zoning and masking requirements
- software installation procedure.

Chapter 5. Installing a SAN

Before Installing a SAN

Before installing a SAN, check the configuration to be setup, verify that all hardware and software components are available, and that the component–related documentation is to hand.

- All available hardware and software components are described in Chapter 3. SAN Components.
- Related documentation is described in About this book: Related publications on page iii.

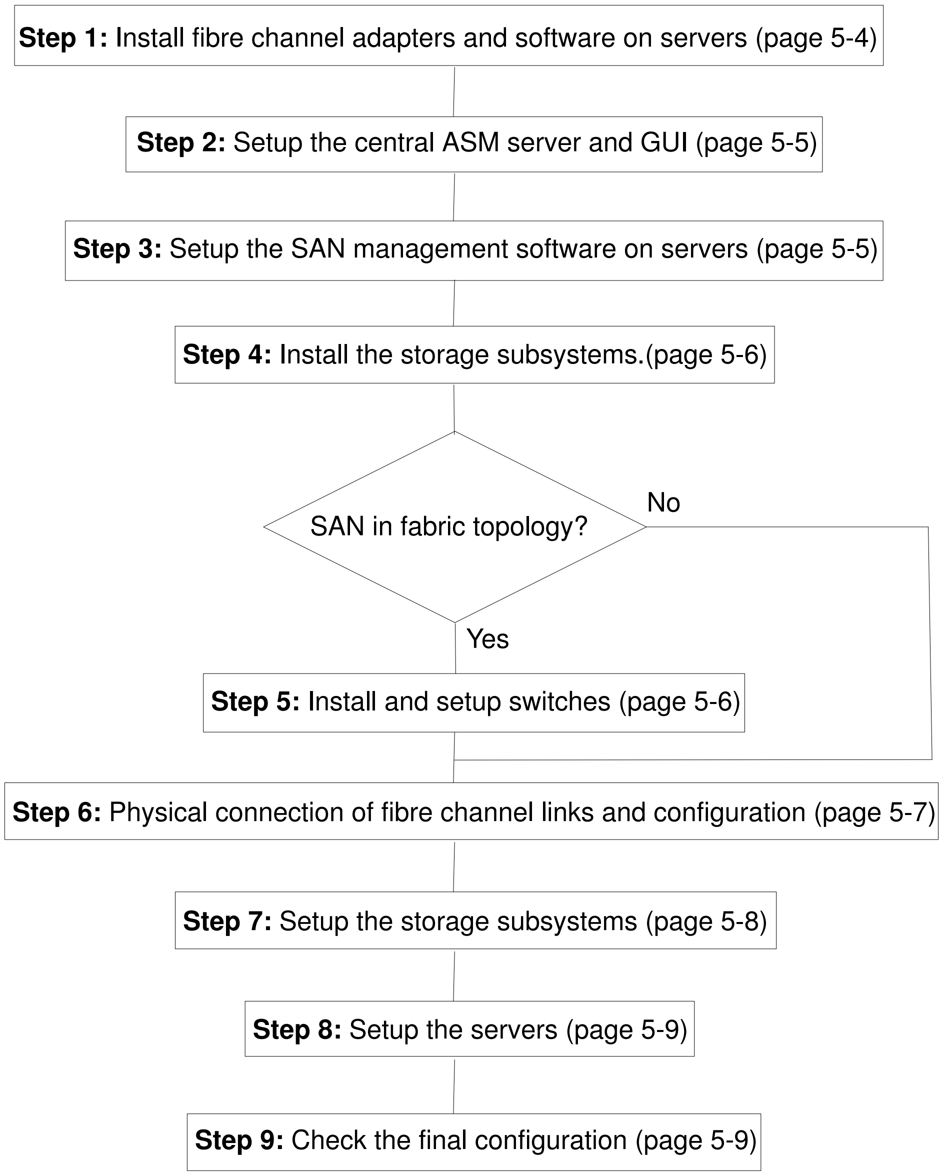
A global cabling scheme must be available, as well as labels on both ends of the fibre–channel cables.

To better manage them through the SAN management software, a naming convention for SAN switches, disk subsystems and servers must be defined.

Note: The procedures in this section presume that all hosts that will be installed in the SAN have the operating system software and TCP/IP services installed and running.

SAN Installation Overview

This section shows the main steps to follow when installing a new SAN. Each step is described in detail on the indicated page:



Note: The physical connection of systems and disk systems in the SAN is not one of the first step of the SAN installation – a large part of the hardware and software installation and configuration must be done before cabling (done in step 6) the SAN.

The following illustration shows an example configuration which the subsequent procedure describes. The AIX server with the S@N.IT! central server role can be a separate AIX station outside the SAN scope (for example a powerconsole). In the example, one of the AIX servers on the SAN is chosen as the S@N.IT! Central server for easier understanding.

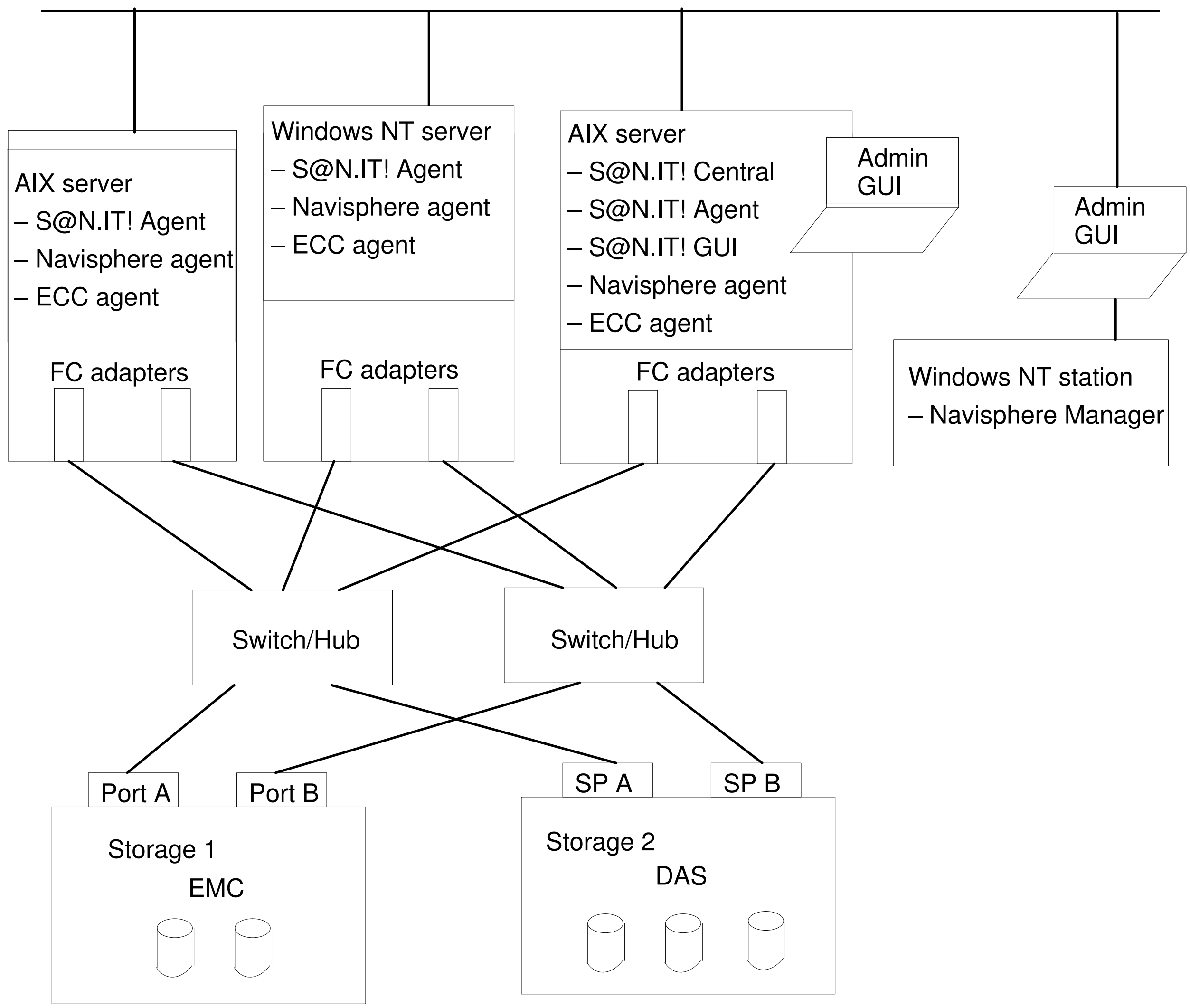


Figure 13. Typical SAN configuration.

Step 1: Install software and FC adapters on servers

See Page 3-4 for a description of the available software.

Installation on AIX Servers

1. Install the driver for the FC adapters.

▶ See *PCI Fibre Channel Adapters Installation and Configuration Guide* and *Bull Enhancement System Release Bulletin (SRB)*.

2. Install other SAN–related softwares (DiskArray, ATF, CDA, Navisphere, Powerpath, HACMP, etc.) as required.

   The software is delivered on Bull Enhancement CD–ROM, or Optional Medias. See the associated documentation and System Release Bulletin for details.

3. Install S@N.IT! packages:

   Java filesets are prerequisites on all servers and WebSM filesets are prerequisites for the S@N.IT! GUI.

   The file `/etc/san/SANManager.cfg` must be configured according to the host role in SAN (see details in Step 2 for central server and Step 3 for SAN agents).

4. Activate the LUN access control through the S@N.IT! GUI.
5. Shutdown AIX and power down the server.
6. Install the PCI Fibre Channel HBAs in the server:

   Verify the positions of the jumpers.

   Attach MIAs if needed.

7. Power up the server:

   Check the availability of the adapters with `lsdev` command (see page 6-7 for listing).

   Verify, and set if required, the topology value in `fchan` driver according to whether the configuration is loop or switch topology (see page 6-7 for an example).

   The S@N.IT! LUN access control is initially created empty and is inactive.

   If the server is attached to Symmetrix boxes, the installation of the basic ECC is needed to monitor box errors.

Installation on Windows NT servers

1. Shutdown Windows NT and power down the server.
2. Install HBAs in the server:

▶ See *Microsoft Windows NT Server Setup* documentation.

3. Power up the server.
4. Install the FC HBA port driver:

   For DAS, the driver is delivered within Fibre Connection Kits. Choose Arbitrated Loop or Fabric driver according to your SAN configuration.

   For Symmetrix, the driver should be retrieved from the Emulex website.

▶ See *Microsoft Windows NT Server Setup* documentation.

5. Install SAN management software:

   execute the `SSMSetup.exe` file. Enter the IP name of AIX central server.

The `Manager.cfg` configuration file is updated and the scheduler is started. The LUN Access Control is initially created empty and is activate.

▶ See *S@N.IT! User's Guide and Software Release Bulletin S@N.IT ! for Windows NT* documentation.

6. Install and configure other SAN–related software:

   if ATF is to be installed, install it before the Navisphere agent software.

   Reboot the server.

7. Install Navisphere Agent and optional CLI software.

▶ See *EMC Navisphere ATF, EMC Navisphere Agent and CLI* documentation.

   If the server is attached to Symmetrix boxes, then installation of the basic ECC is needed to monitor box errors.

Step 2: Setup the central S@N.IT! server and GUI

1. The S@N.IT! central role must be installed on only one AIX host server.
2. The server must have a graphical screen connected to it when it is to be used to launch the S@N.IT! GUI.

Verify that Netscape software is installed and that the browser is available and configured. Managing the switches through Netscape (WebTools) requires Java2 software.

Edit the `/etc/san/SANManager.cfg` file to configure the SAN management software (set CentralRole, MyHostname and CentralHost parameters). Also set ClientRole to be able to launch the GUI.

If this AIX server is also a full member of the SAN and has access to disk subsystems (as in the configuration shown in figure 13), then also set the LocalRole parameters.

Start the SAN manager daemon using rc.sanmgt start command and launch the graphical application (wsm command, then click on SAN Manager icon). At this stage, only the central server should be visible in the main window.

▶ See *SAN Manager User's Guide* documentation.

Step 3: Install the SAN management software on servers

On each SAN server edit the configuration file (`/etc/san/ SANManager.cfg` file on AIX and `Manager.cfg` file on Windows NT) to set:

- the server name,
- the central server host name,
- the LocalRole.

Start the S@N.IT! daemons on each server (`rc.sanmgt` start on AIX, S@N.IT! icon on Windows NT) and verify that the list of available SAN servers is correct in the S@N.IT! GUI window.

- Activate LUN access control in the S@N.IT! main window GUI. Verify that the red dot on the host icon becomes a yellow question mark.
- Reboot the server to make the LUN Access Control active.

Step 4: Install the storage subsystems

- Physical installation and internal cabling.
- Set SP Address ID switches of the DAS to unique numbers in the SAN.
- Power up the subsystems:

  Check the statuts lights on the front of the SPSs, DPEs and DAEs according to the DAS documentation.
- Set–up the Navisphere Management station for DAS subsystems on either AIX or Windows NT (recommended):

Navisphere Supervisor or Manager for Windows NT (recommended)

The Windows NT station that supports Navisphere must be connected to the LAN.

The Navisphere Supervisor supports all DAS models, but only a single DAS on a single system at the time. It does not support Acces Logix flare on DAS 4500 and DAS 4700.

▶ See *EMC Navisphere 4X Supervisor Installation* or *EMC Navisphere Manager Installation* documentation.

Installation of Navisphere for Windows NT is mandatory to manage DAS 4500 and DAS 4700, and to benefit from RAID Group support.

Navisphere Manager on AIX (for backward compatibility)

The Navisphere Manager on AIX does not support DAS 4500, DAS 4700, nor Storage Group/Raid Group features.

Install Navisphere Manager on the AIX servers which the Navisphere GUI will be used on (the servers connected to the LAN).

Set up licences for multiple servers.

▶ See *Navisphere for AIX Setup and Operation* documentation.

Installing Symmetrix management software

▶ Refer to *EMC ControlCenter Version 4.1 Installation Guide* to install the EMC administration components.

▶ Refer to *PowerPath V2.0 for Unix: Installation Guide* to install PowerPath on the servers.

▶ Refer to *Volume Logix V 2.3 Product Guide* to install and manage Volume Logix.

All the LUN addressing on the boxes is assumed to be performed either by Bull Technical Support or by EMC technical support (except when using the LUN dynamic re–allocation option).

Step 5: Install and set–up switches (SAN in fabric topology)

- Connect the switches to the LAN.
- Power–up the switches.
- Set basic switch IP parameters (address, Subnet Mask, Domain), either by front–panel keys (2800 model) or through a serial line (2040 model) – see the related documentation for details.
- For Connectrix switches, indicate the IP address of the switch and not that of the administration PC.
- Check the Ethernet network connection (for example, by trying a telnet session or trying a ping).

Step 6: Physical connection and configuration of FC links

The following procedure is a step–by–step, connect–and–verify cabling sequence that is easier and simpler than a cable–everything–and–verify–afterwards approach which could be extremely difficulty in a complex SAN architecture.

Connections can be confirmed by observing LEDs on the interface of each device. The LEDs are lit when an incoming electrical or optical signal is present, so verify LEDs on both ends of the cables to be confident of a good physical connection. Also check that cables are properly secured on their connectors (DB9 cables must be fully screwed on and you should hear a click when SC or LC optical connectors are correctly connected). If MIAs are needed to perform electrical to optical conversions, make sure that they are also correctly connected. Cable labelling should also be performed at this stage for easier maintenance.

The S@N.IT! GUI enables the connections and cabling to be checked.

Basic steps are described below. Storage subsystems and hosts specific setups are detailed in the steps 7 and 8.

- Interconnect the switches or hubs to build the SAN backbone.
- Connect the AIX central server to the LAN (cable from the adapter to the switch).
- To use the switch management tools, the switch must be recognized by the S@N.IT! central server. Perform for each adapter:

  ```
  cfgmgr -vl fchanX
  ```

  The main window appears with the switch object in it and set to `monitored`.
- Brocade switch configuration

  Through the S@N.IT! GUI, launch a management session (telnet or web tool) and:
  – check the switch status
  – verify `Licences` and `Firmware` version
  – set the logical name
  – configure zoning if required.
- Connectrix switch configuration

  Using the PC in the rack connected to the switches:
  – check the status the switches
  – program zoning.

  Note: The S@N.IT! GUI is not able to launch the Connectrix administrative tool. This can only be done on the PC.
- Connect each DAS disk subsystem to the SAN.

  If a switch is used, check the status light on the switch port, and check the port status with the switch management tool.

  If HUBs are used, check the status light on the HUB port.
- Reboot the central host.

  Check that the new entries (fcp, sp, atf) with `lsdev` and `lsattr` commands (see chapter SAN internals) really reflect the SAN objects.

  At this time no `hdisk` objects should be present because the initial state of LUN Access Control mechanism is that all the available disks on the SAN are hidden.

  To configure the Navisphere management agent on the central server, identify the sp devices of DAS subsystems with `lscfg -Ccarray` command

- Update `agent.config` configuration file of the Navisphere agent and start the agent with `rc.naviagent start`.

▶ See *Navisphere Agent Installation for Windows NT* or *Navisphere for AIX Setup and Operation* documentation.

At this time, it is possible to launch, through the S@N.IT!, a previous version of Navisphere AIX management session for FC Direct Attach DAS.

- Connect the remaining servers to the SAN (if necessary, see Step 8 for details)
- Setup each DAS subsystem using Navisphere tool (Raid groups, LUN binding): see the following section.

Step 7: Setup the storage subsystems

DAS subsystem setup

Note: Using Access–Logix: this feature provides a way to LUN Access Control, directed by the disk subsystems rather than by the servers, and may be useful when the operating systems is other than AIX and/or WindowsNT on the SAN.

A mandatory requirement of the S@N.IT! is that S@N.IT! agent hosts must see the same LUNs on each subsystem port. To be able to use the S@N.IT! LUN access control facility, it may be necessary to re–configure the SAN (maybe using zoning) to meet that requirement.

Using Windows NT Navisphere:

- select a host to be the managing host
- select the storage system to manage
- verify flare level and download it if necessary
- set storage systems general properties: memory cache
- set fair access, configuration access, data access if using Acces Logix flare
- create RAID Group and LUN
- create Storage Group if any
- select all LUNs
- logically connect hosts.

▶ See *EMC Naviphere Manager or Supervisor Installation & Operation* documentation.

Symmetrix Subsystem Setup

Note: Using Volume Logix: this feature provides a way to LUN Access Control, directed by the disk subsystems rather than by the servers, and may be useful when the operating systems is other than AIX and/or WindowsNT on the SAN.

A mandatory requirement of the S@N.IT! is that S@N.IT! agent hosts must see the same LUNs on each subsystem port. To be able to use the S@N.IT! LUN access control facility, it may be necessary to re–configure the SAN (maybe using zoning) to meet that requirement.

On each Symmetrix, the following configuration must be done:

- configure Front End bit according to the system (Loop, Fabric, or Volume Logix)
- create RAID Volumes and LUNs
- create SRDF volumes and associated remote links ports (FC or ESCON)
- create Gatekeepers (for OSM)
- assign LUNs to hosts ports.

▶ See *SymmWin V5.0 Product Guide* documentation and the Symmetrix installation team for details of this operation.

Step 8: Setup the Servers

AIX Server Setup

Once the AIX server cables are connected to the SAN, run the cfgmgr command. The green LED on the FC adapter must be ON (not flashing).

- Verify the LED and name server entry on the switch with the `mshow` command, or the LED on the hub.

Verify the driver and array devices with `lsdev` and `lsattr` commands.

At this time, no hdisk should be present (other than the internal ones).

Windows NT Server Setup

Once the server is connected to the SAN, the green LED on the FC adapter must be ON.

- Verify the LED and name server entry on the switch, or the LED on the hub.
- Reboot to make the adapter configuration operational.

Setup SAN LUN access control

At this step, all disk subsystems, switches, and hosts should be visible in the S@N.IT! GUI main window.

Domains can be created to simplify SAN management, logically isolating sets of hosts and subsystems depending on the usage criteria.

Using the S@N.IT! GUI, allocate to each host the necessary LUNs for each subsystem. All the LUNs need not be allocated in this first step: some of them may be reserved for future increases of disk space on a server.

Once a LUN has been added to the LUN Access Control, the corresponding `hdisk` device object is automatically created on the AIX server.

On Windows NT server a message asks whether the Disk Administrator should be run to complete the operation.

▶ See *AIX* S@N.IT! *User's Guide* documentation.

Step 9: Check the final configuration

In the S@N.IT! GUI main window, no host icon should have a RED dot or a yellow question–mark.

There should be no RED icons. If there are, refer to the equipment maintenance guide and fix the problem.

- Using the graphical topology display of the S@N.IT! GUI, verify that the SAN description is exactly the one you expected.

The SAN is now configured and secure.

- The disk space on each host can now be used as required : to make volume groups, file systems, etc on AIX server, the disk administrator on Windows NT servers can be started and partitions formatted.
- If Navisphere agent is installed on servers other than the central host, it must be customized and the agent started:

  on AIX Server:

  – identify the `sp` devices

  – update Navisphere file `agent.config` and start the Navisphere agent.

on Windows NT Server under Navisphere Agent Utilities:

– clear the device list.
– `Auto Detect` action must display each DAS `sp` entries.
– `Scan Bus` action displays mapped SCSI devices declared in the LUN Access Control.
– Saving the new declaration will update the Navisphere Windows NT configuration file and restart the agent.

- For each DAS using ATF, it is recommended that ATF device status is checked to verify that no device has trespassed during configuration:

on AIX server:

use `smit atf` command for a global view of each subsystem, or `atfi` command for a view of the individual LUN properties in Navisphere Manager. AIX `errlog` file also logs LUN trespasses

after checking the configuration, perform restore operations if needed, using `smit atf` or `atf_restore` commands

in Navisphere Manager, perform `an update host information` operation on the LUNs to get the correlation between the LUN and the hdisk object on AIX

▶ See *Configuring and Managing a DAS and Using ATF* documentation.

on Windows NT server:

LUN trespass can be checked using ATF console. Individual LUN properties in Navisphere Manager show the current and default position. The Event viewer also logs LUN trespasses

after checking the configuration, perform the restore operations if necessary, using `atf_restore` utility in the ATF console.

- In Navisphere Manager, perform `update host information` operation on the LUNs to get the correlation between the LUN and the Windows NT partition.

▶ See *EMC Navisphere ATF* documentation.

Chapter 6. How the OS Handles SAN Objects

This chapter describes how AIX or Windows NT sees SAN objects and how the OS presents them on the GUI. The chapter is intended for readers who are qualified to system administrator or support level.

SAN Objects on AIX

The following objects are described:

FC adapters

There is one `fchan<x>` object for each FC PCI adapter:

```
# lsdev -Cc adapter | grep fchan
fchan0  Available 10-38    PCI Fibre Channel Adapter
fchan1  Available 20-50    PCI Fibre Channel Adapter
fchan2  Available 20-58    PCI Fibre Channel Adapter
fchan3  Available 20-60    PCI Fibre Channel Adapter
```

The location code is the physical PCI slot in the system (specific values depending on the type of platform).

The adapter topology can be displayed with the `lsattr` command:

```
# lsattr -El fchan0 -a topology

topology        loop       Fibre Channel topology
True
```

CAUTION:
Point–to–point topology must be used when the adapter is connected to a switch.
Set the topology with the `chdev` command while this adapter is in the Defined state:

```
# chdev -l fchan0 -a topology=loop
```

or

```
# chdev -l fchan0 -a topology=pt2pt
```

FC interface

There is one `fcp<x>` object per remote fibre–channel port:

```
# lsdev -Cs fcp
fcp0 Available 04-06-3M Fibre Channel FCP Interface
fcp3 Available 04-07-3M Fibre Channel FCP Interface
fcp1 Available 04-06-3N Fibre Channel FCP Interface
fcp2 Available 04-06-3O Fibre Channel FCP Interface
fcp4 Available 04-07-3N Fibre Channel FCP Interface
fcp5 Available 04-07-3O Fibre Channel FCP Interface
```

The location code is in the format: AA–BB–XY

AA–BB is the location code of the parent fchan adapter

XY is a specific encoding of the index in an internal table called node table:

X is in the range [0–E] (0 to 14)

Y is in the range [0–Z] (0 to 35)

The index in the node table is: X*36 + Y (0< X*36+Y <512).

It is linked to the minor number of the /dev/fcp<x> special file

(X*36 + Y = minor_number modulo 512,

and minor_number DIV 1024 = fchan_number)

Example:

```
# ls -l /dev/fcp0

crw-rw-rw-   1 root     system    21,2178 Sep 15 11:32 /dev/fcp0
```

2178 div 1024 = 2   –> adapter=fchan2

2178 mod 512 = 130

```
# lsdev -Cl fcp0

fcp0 Available 20-58-3M Fibre Channel FCP Interface
```

3M –> 3*36 + 22 = 130

Disks

There is one `hdisk` object per subsystem logical disk (LUN):

on AIX 4.3.3

```
# lsdev -Cl hdisk13
hdisk13 Available 04-07-3N-1,11 Bull FC DiskArray RAID 5 Disk Group
```

on AIX 5.1

```
# lsdev -Cl hdisk15
hdisk15 Available 20-60-3N     Bull FC DiskArray RAID 5 Disk Group
```

```
# lsattr -El hdisk15
FLU           27             Flare Logical Unit Number (FLU)  False
reserve_lock yes            Issue SCSI reserve at startup     True
unit_id      a10005300162-1BArray Serial Number-FLU' or WWN  False
max_transfer 0x40000        Maximum transfer size             True
max_coalesce 0x20000        Maximum coalesce size             True
pvid         none           Physical volume identifier       False
clr_q        yes            Clear adapter queue on error      True
q_err        no             Clear device queue on error       True
q_type       simple         Tag for command queuing           True
queue_depth  16             Maximum device queue depth        True
target_id    72192          SCSI/LOOP/ALPA ID of SP          False
scsi_id      0x11A00        SCSI/LOOP/ALPA ID of SP          False
lun_id       0x1B000000000000 Logical Unit Number (or VLU)   False
parent       atf2           Real SP or ATF Parent Device     False
```

location code: AA–BB–CC–X,Y

where AA–BB–CC is the location code of the fcp<x> object to which the disk is attached

and X,Y the LUN–id (X and Y are in decimal form)

X*16 + Y = LUN_ID

for example: 1,11 –> LUN–id is 1*16+11=27

On AIX 5.1 the fields X and Y no longer exist, however they are accessible with the `lsattr` command.

Note: All the examples shown are those from AIX 4.3 system.

```
# lsdev -Cc disk -t CLAR*
hdisk1  Available 04-06-3N-0,0   Bull FC DiskArray RAID 1 Disk Group
hdisk2  Available 04-06-3N-0,3   Bull FC DiskArray RAID 0 Disk Group
hdisk3  Available 04-07-3O-0,1   Bull FC DiskArray RAID 5 Disk Group
hdisk4  Available 04-07-3O-0,2   Bull FC DiskArray Individual Disk
hdisk5  Available 04-06-3O-0,1   Bull FC DiskArray RAID 5 Disk Group
hdisk6  Available 04-06-3O-0,4   Bull FC DiskArray RAID 1 Disk Group
hdisk7  Available 04-07-3N-0,2   Bull FC DiskArray RAID 5 Disk Group
hdisk8  Available 04-06-3M-0,1   Bull FC DiskArray Individual Disk
hdisk9  Available 04-06-3M-0,10  Bull FC DiskArray RAID 5 Disk Group
hdisk10 Available 04-07-3M-0,0   Bull FC DiskArray Individual Disk
hdisk11 Available 04-07-3M-0,9   Bull FC DiskArray RAID 5 Disk Group
hdisk12 Available 04-07-3N-13,13 Bull FC DiskArray RAID 5 Disk Group
hdisk13 Available 04-07-3N-1,11  Bull FC DiskArray RAID 5 Disk Group
```

`Bull FC DiskArray xx Disk` corresponds to a FC disk.

`Bull Diskarray xx Disk` corresponds to a SCSI disk, if applicable.

SP

One `sp<x>` object per DAS subsystem SP:

`sp<x>` is of array type:

```
# lsdev -Cc array
sp0 Available  Bull DiskArray Storage Processor
sp1 Available  Bull DiskArray Storage Processor
sp2 Available  Bull DiskArray Storage Processor
sp3 Available  Bull DiskArray Storage Processor
```

No location code is available for this object.

Parent object (i.e. the FC port to which the SP is connected) is an attribute of the object itself:

```
lsattr -El sp0
sysser_spnum a10005100062,1 System serial number and SP number False*
autoHotSwap  true           Auto hot swap on SP replacement    True
default_HBA  fcp0           Default Host Bus Adapter          False
max_lun_id   223            Maximum LUN number allowed         True
queue_depth  255            SP queue depth                     True
sp_reboot_delay 15          SP Reboot Delay in seconds         True
```

(*) SP number = 0 for SPA or 1 for SPB.

ATF Objects

A specific `atf<x>` object is created for each DAS subsystem, covering the two access points (`sp` and `fcp` objects). Only one `hdisk` object is created for each logical disk, whatever the link (`fcp`) they are connected to. The `hdisk` location code is linked to the one of the `fchan` and `fcp` objects they are connected to at configuration time.

```
# lsdev -Cl atf1

atf1    Available       Bull DiskArray ATF Pseudo-Device
```

No location code is available for this object.

Links with other objects are available with the `atfi` command:

```
# /usr/sbin/atf/atfi -d atf0 -p
atf0:
        ATF version        = 4.0.3.7
        Array serial No.  = a10005100062
        Device number      = 61.0
        I/O count          = 0
        Flags              = ACTIVE OPEN
        Pending ioctls     = 0
sp0:
        Slot in array    = B
        SP devno         = 60.0
        Target ID        = 0
        Queue depth      = 255
        SP ID            = 1
        Issued cmds      = 0
        Pending cmds     = 0
fcp0:
        Adapter devno    = 58.130
sp3:
        Slot in array    = A
        SP devno         = 60.3
        Target ID        = 0
        Queue depth      = 255
        SP ID            = 0
        Issued cmds      = 0
        Pending cmds     = 0
fcp3:
        Adapter devno    = 58.1154
Unit Information:
----------------
                        Pending Issued  Current
Name            State   Cmds    Cmds    Path
---------- ---------- -------- -------- ----------
hdisk10         GOOD    0       0       DEFAULT
hdisk8          GOOD    0       0       DEFAULT
hdisk11         GOOD    0       0       DEFAULT
hdisk9          GOOD    0       0       DEFAULT
Path Information:
----------------
Unit      HBA   SP    ID    LUN   FLAGS
-------   ----  ----  ----  ---   -----
hdisk10   fcp0  sp0   130   0
hdisk10   fcp3  sp3   130   0     d c

hdisk8    fcp0  sp0   130   1     d c
hdisk8    fcp3  sp3   130   1

hdisk11   fcp0  sp0   130   9
```

```
hdisk11   fcp3  sp3   130   9     d c
hdisk9    fcp0  sp0   130   10    d c
hdisk9    fcp3  sp3   130   10
```

(*) d and c indicate respectively the default and current SP.

Powerpath Objects

A specific `powerpath0` driver object is created for the whole AIX system at Powerpath installation time.

```
#lsdev -Cl powerpath0

powerpath0 Available PowerPath Control Device
```

A specific `hdiskpower<x>` disk object is created for each pair of paths (`fcp`) (or more if there are more than 2 physical paths to the subsystem).

Use the `powermt` command to link `hdisk<x>` with `hdiskpower<x>` and `fcp<x>` objects:

```
#lsdev -Cc disk

hdisk0      Available 04-A0-00-0,0 Bull 4.2 GB 16 Bit SCSI Disk
Drive.
hdisk1      Available 04-02-3O-0,0 SYMMETRIX RAID-1(local)
hdisk2      Available 04-02-3O-0,2 SYMMETRIX RAID-1(local)
hdisk3      Available 04-02-3O-1,1 SYMMETRIX RAID-S(local)
hdisk4      Available 04-02-3O-2,1 SYMMETRIX Disk Drive
hdisk5      Available 04-02-3O-2,3 SYMMETRIX Disk Drive
hdisk6      Available 04-06-3L-3,0 SYMMETRIX RAID-1(local)
hdisk7      Available 04-01-0D-3,0 SYMMETRIX RAID-1(local)
hdisk8      Available 04-02-3O-3,0 SYMMETRIX RAID-1(local)
hdiskpower0 Available 04-02-3O-0,0 PowerPath Device
hdiskpower1 Available 04-02-3O-0,2 PowerPath Device
hdiskpower2 Available 04-02-3O-1,1 PowerPath Device
hdiskpower3 Available 04-02-3O-2,1 PowerPath Device
...
#powermt display dev=0
Pseudo name=hdiskpower0
Symmetrix frame ID=000182600940; volume ID=009
state=alive; policy=SymmOpt; priority=0; queued-IOs=0
=================================================================
------------- Host Devices ------------  - Symm -  --- Path ----
 -- Stats ---
### HW-path      device     director  mode    state  q-IOs errors
================================================================
  2 fcp7         disk19    FA 14a    active  open       0      0
  0 fcp3         hdisk1    FA 14b    active  open       0      0
  1 fcp4         hdisk24   FA 16a    active  open       0      0
```

LUN Identification in AIX, Navisphere NT, or S@N.IT! GUI

The LUN of a DAS storage system is identified by the Flare Logical Unit (FLU) at the level of the storage system and by the Virtual Logical Unit (VLU) at the level of the operating system.

The FLU and VLU values are identical except for the DAS with Access Logix flare. The mapping between the FLU and VLU values is given by the `lsattr` command or by S@N.IT!.

Example LUN of a DAS4500 with Access Logix flare:

```
# lsattr -El hdisk4
FLU         171        Flare Logical Unit Number (FLU)            False
unit_id 600601FA8890000086BA75F6D166D511 'Array Serial Number-FLU' or
WWN     False
pvid    000000767de3eb8b0000000000000000 Physical volume identifier  Fal
se
q_type       simple     Tag for command queuing                      True
queue_depth  16         Maximum device queue depth                   True
reserve_lock yes        Issue SCSI reserve at startup                True
target_id    0          SCSI/LOOP/ALPA ID of SP                     False
lun_id       2          Logical Unit Number (or VLU)                False
parent       atf1       Real SP or ATF Parent Device                False
```

The `Z0` field in `lscfg` command is the corresponding hexadecimal value of FLU :

```
# lscfg -vl hdisk4
  DEVICE            LOCATION          DESCRIPTION

  hdisk4            04-07-3O-0,2      Bull FC DiskArray
Individual Disk

        Manufacturer................DGC
        Machine Type and Model......DISK+
        ROS Level and ID............0176
        Serial Number...............F60005001208DISK
        Device Specific.(Z0)........AB
        Device Specific.(Z1)........00513547
        Device Specific.(Z2)........CL
```

You see this mapping when using the NT Navisphere Manager or the S@N.IT! LUNs panels shown below:

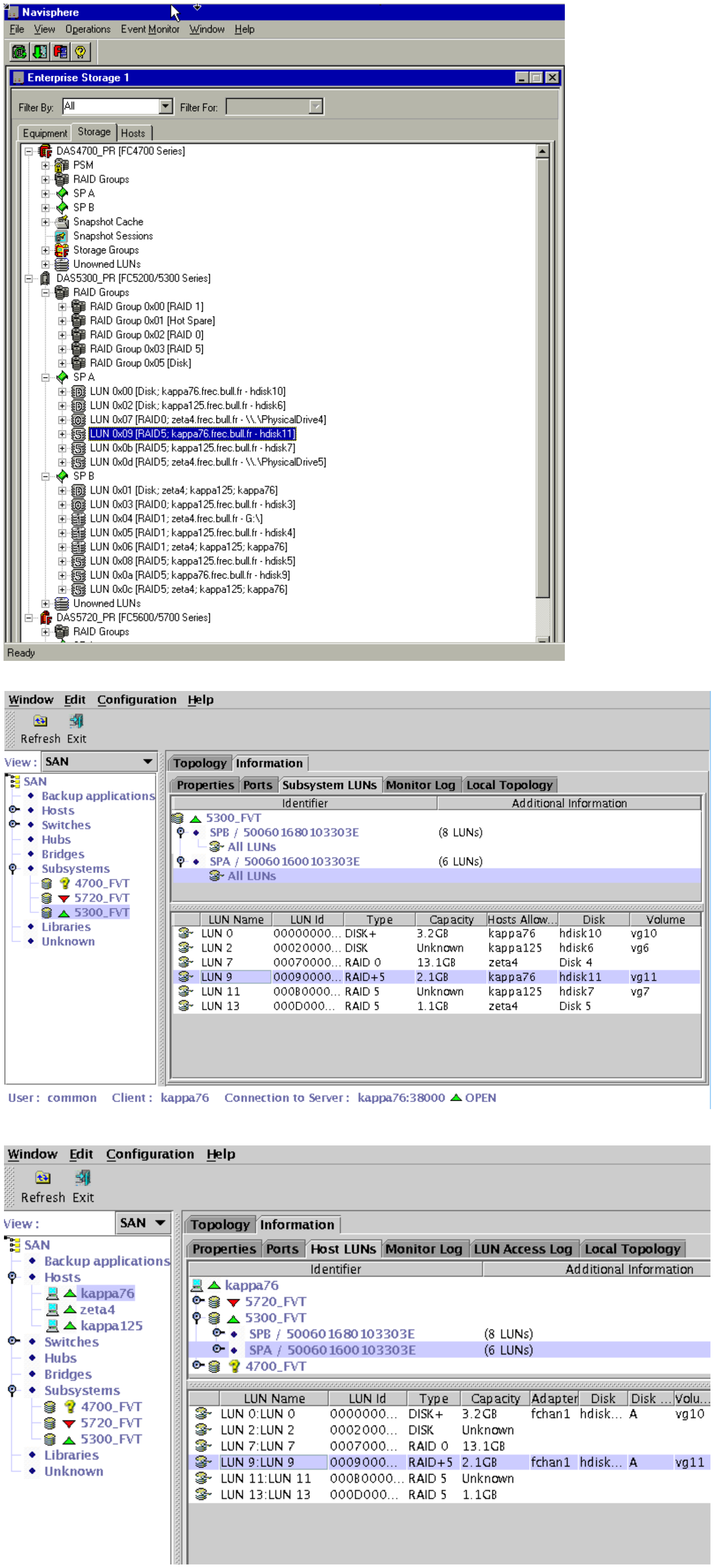


Figure 14. LUNs mapping of a DAS4500

DAS Subsystem Identification

The unit_id field of the lsattr command in AIX identifies the DAS subsystem.

The unit_id corresponds to a WWN for DAS4500 with Access Logix, and to a DAS serial number–FLU in other cases:

```
# lsattr -El hdisk4
FLU       171      Flare Logical Unit Number (FLU)                False

unit_id            600601FA8890000086BA75F6D166D511      'Array Serial
Number-FLU' or WWN  False

pvid  000000767de3eb8b0000000000000000      Physical volume identifier
False

q_type      simple  Tag for command queuing                         True
queue_depth 16      Maximum device queue depth                      True
reserve_lock        yes                     Issue SCSI reserve at startup
True
target_id   0       SCSI/LOOP/ALPA ID of SP                        False
lun_id      2       Logical Unit Number (or VLU)                   False
parent      atf1    Real SP or ATF Parent Device                   False


# lsattr -El hdisk13
FLU         27            Flare Logical Unit Number (FLU)          False
unit_id     a10005300162-27            'Array Serial Number-FLU' or WWN
False
pvid        none          Physical volume identifier               False
q_type      simple        Tag for command queuing                   True
queue_depth 16            Maximum device queue depth                True
reserve_lock              yes              Issue SCSI reserve at startup
True
target_id   0             SCSI/LOOP/ALPA ID of SP                  False
lun_id      27            Logical Unit Number (or VLU)             False
parent      atf2          Real SP or ATF Parent Device             False
```

The unit_id field can be used to link AIX object to subsystems identification in AIX S@N.IT! for all DASs other than DAS with Access Logix.

For the DAS with Access Logix, go to the attributes of the `atf` object to get the serial number of the array (also available with `smit atf` command):

```
# lsattr -El atf1
system_serial   F60005001208    Array Serial Number          False
```

Note: The DAS serial number can be seen by the NT Navisphere Manager in the storage system properties dialog box, or on the S@N.IT! GUI screen.

See below for examples of how NT Navisphere Manager and the S@N.IT! display the `unit_id` of a DAS.

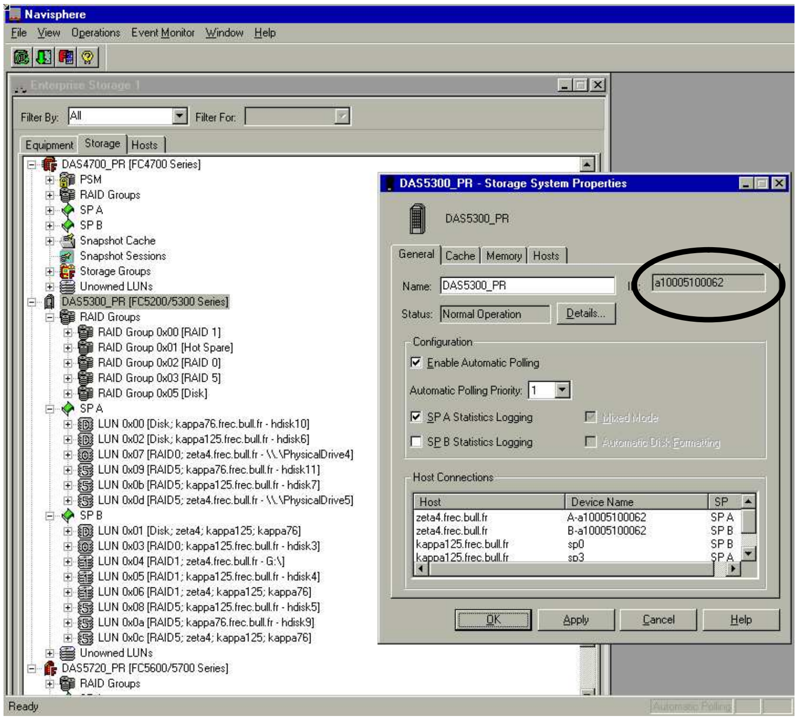


Figure 15. DAS subsystem identification with Navisphere manager

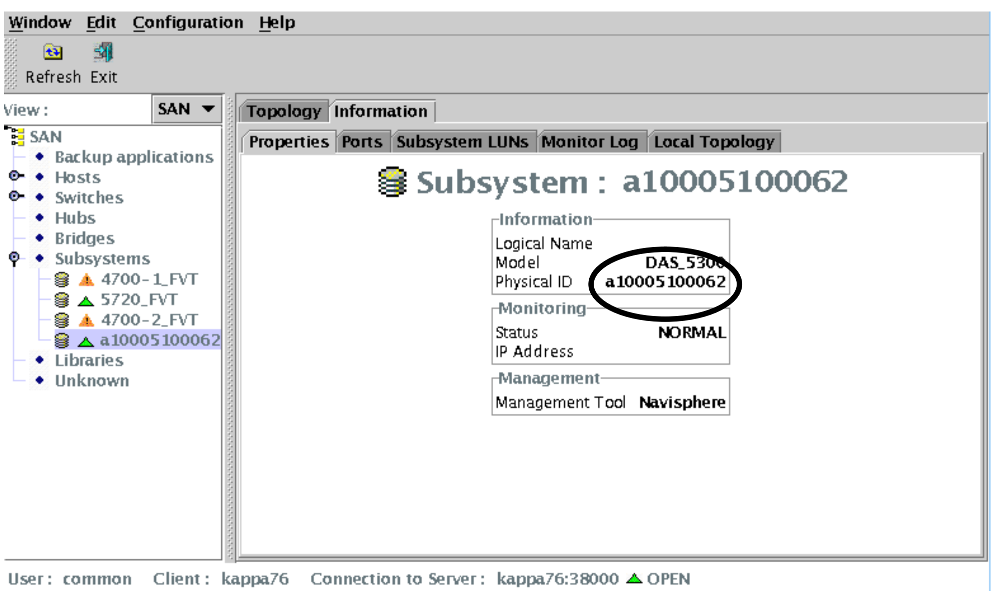


Figure 16. DAS subsystem identification with S@N.IT!

Storage Processor (SP) FC–AL Identification

Each DAS storage processor must have an unique FC address ID in a configuration.

The FC–AL protocol translates the FC–AL address ID to an 8–bit address ALPA.

This ID can be set either by a switch on the back of the SP for a DAS5700 or a DAS4500, on the SP board for a DAS5300, or via the manager on a DAS4700.

```
# lsattr -El hdisk5

FLU           2                                 Flare Logical Unit Number (FLU)  False
unit_id       60060160FD102A57A5B520C05A63048C  'Array Serial Number-FLU' or WWN False
pvid          none                              Physical volume identifier       False
q_type        simple                            Tag for command queuing          True
queue_depth   16                                Maximum device queue depth       True
reserve_lock  yes                               Issue SCSI reserve at startup    True
target_id     51                                SCSI/LOOP/ALPA ID of SP          False
lun_id        2                                 Logical Unit Number (or VLU)     False
parent        atf1                              Real SP or ATF Parent Device     False
```

This SP ID can be found with the NT Navisphere Manager, as shown below:

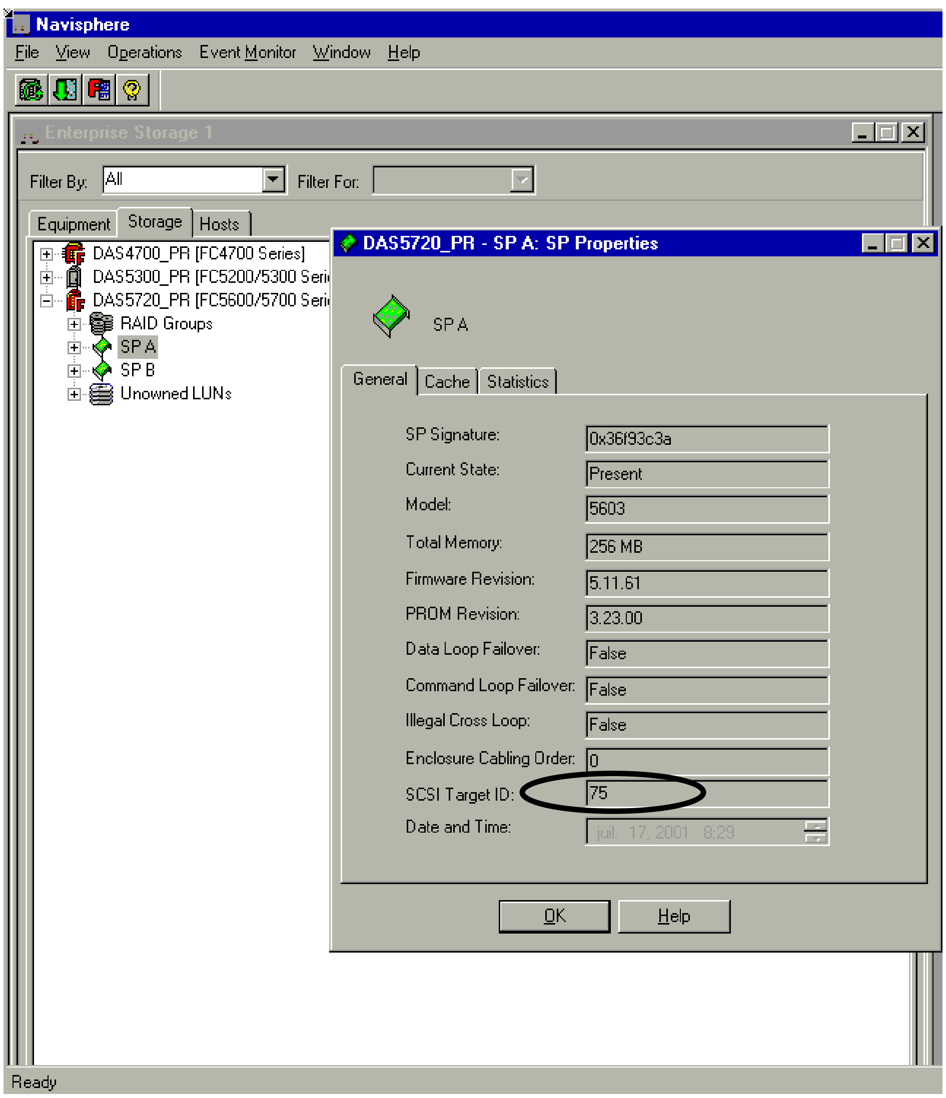


Figure 17. SP ID with Navisphere on a DAS5720

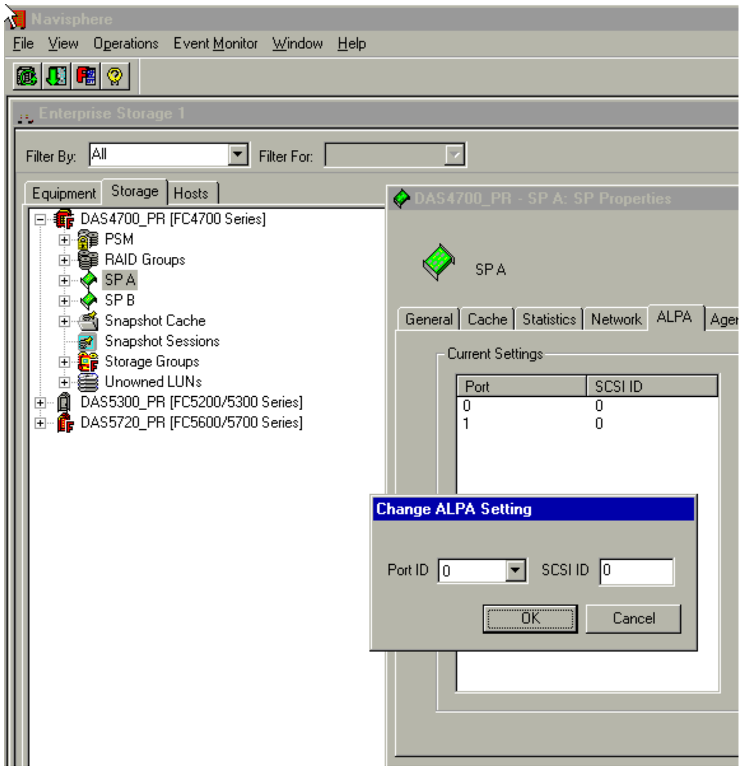


Figure 18. SP ALPA settings

Name–server table display:

```
switch1:admin> nsshow
The Local Name Server has the following 14 entries:
Type  Pid      COS   PortName                  NodeName                  TTL(sec)
N     011100;  2,3;  10:00:00:00:c9:20:c0:ba;  20:00:00:00:c9:20:c0:ba;  na
      IP address: 0.0.0.0.0.0.0.0.255.255.0.0.78.132.240.150
      NodeSymb: [14] fchan0 kappa76
      Fabric Port Name: 20:01:00:60:69:10:15:0f
N     011200;  2,3;  10:00:00:00:c9:21:6d:8f;  20:00:00:00:c9:21:6d:8f;  na
      FC4s: FCIP FCP
      Fabric Port Name: 20:02:00:60:69:10:15:0f
N     011300;  2,3;  10:00:00:00:c9:21:2b:25;  20:00:00:00:c9:21:2b:25;  na
      NodeSymb: [15] fchan2 kappa125
      Fabric Port Name: 20:03:00:60:69:10:15:0f
N     011400;  2,3;10:00:00:00:c9:21:18:15;    20:00:00:00:c9:21:18:15;  na
      NodeSymb: [15] fchan3 kappa125
      Fabric Port Name: 20:04:00:60:69:10:15:0f
N     011500;  2,3;  10:00:00:00:c9:20:eb:fe;  20:00:00:00:c9:21:6d:8f;  na
      FC4s: FCIP FCP
      Fabric Port Name: 20:05:00:60:69:10:15:0f
N     011600;  2,3;  10:00:00:00:c9:20:d1:63;  20:00:00:00:c9:20:d1:63;  na
      IP address: 0.0.0.0.0.0.0.0.255.255.0.0.61.1.0.145
      NodeSymb: [14] fchan1 kappa76
      Fabric Port Name: 20:06:00:60:69:10:15:0f
N     011700;  3;    50:06:01:68:01:03:30:3e;  50:06:01:60:02:03:30:3e;  na
      FC4s: FCP
      Fabric Port Name: 20:07:00:60:69:10:15:0f
N     011800;  3;    50:06:01:60:01:03:30:3e;  50:06:01:60:02:03:30:3e;  na
      FC4s: FCP
      Fabric Port Name: 20:08:00:60:69:10:15:0f
NL    0119ef;  3;    50:06:01:60:11:03:24:b8;  50:06:01:60:88:03:24:b8;  na
      FC4s: FCP [DGC 0176]
      Fabric Port Name: 20:09:00:60:69:10:15:0f
N     011a00;  3;    20:00:00:60:16:32:d1:ad;  20:00:00:60:16:32:d1:ad;  na
      FC4s: FCP
      Fabric Port Name: 20:0a:00:60:69:10:15:0f
N     011b00;  3;    20:00:00:60:16:31:e5:1b;  20:00:00:60:16:31:e5:1b;  na
      FC4s: FCP [DGC 0511]
      Fabric Port Name: 20:0b:00:60:69:10:15:0f
NL    011ce8;  3;    50:06:01:68:11:03:24:b8;  50:06:01:60:88:03:24:b8;  na
      FC4s: FCP [DGC 0176]
      Fabric Port Name: 20:0c:00:60:69:10:15:0f
NL    011d01;  2,3;  10:00:00:00:c9:20:d4:17;  20:00:00:00:c9:20:d4:17;  na
      NodeSymb: [14] fchan1 rho17–1
      Fabric Port Name: 20:0d:00:60:69:10:15:0f
NL    011e01;  2,3;  10:00:00:00:c9:20:de:a2;  20:00:00:00:c9:20:de:a2;  na
      NodeSymb: [14] fchan0 rho17–1
      Fabric Port Name: 20:0e:00:60:69:10:15:0f
```

```
switch1:admin> switchshow
switchName:     switch1
switchType:     2.3
switchState:    Online
switchRole:     Principal
switchDomain:   1
switchId:       fffc01
```

```
switchWwn:      10:00:00:60:69:10:15:0f
switchBeacon:   OFF
port  0: sw     No_Light
port  1: id     Online  F-Port  10:00:00:00:c9:20:c0:ba
port  2: sw     Online  F-Port  10:00:00:00:c9:21:6d:8f
port  3: sw     Online  F-Port  10:00:00:00:c9:21:2b:25
port  4: sw     Online  F-Port  10:00:00:00:c9:21:18:15
port  5: sw     Online  F-Port  10:00:00:00:c9:20:eb:fe
port  6: sw     Online  F-Port  10:00:00:00:c9:20:d1:63
port  7: sw     Online  F-Port  50:06:01:68:01:03:30:3e
port  8: sw     Online  F-Port  50:06:01:60:01:03:30:3e
port  9: sw     Online  L-Port  1 private, 4 phantom
port 10: sw     Online  F-Port  20:00:00:60:16:32:d1:ad
port 11: sw     Online  F-Port  20:00:00:60:16:31:e5:1b
port 12: sw     Online  L-Port  1 private, 4 phantom
port 13: sw     Online  L-Port  1 public
port 14: sw     Online  L-Port  1 public
port 15: --     No_Module
```

DAS objects are identifed by FCP DGC for all DASs except DAS4500 with Access Logix which is not identified (see `Pid 011b00` in the table above and `port 9` in the table below).

`Port_id, port_name` and `node_name` can be used to identify the corresponding `fcp<x>` object in AIX.

Example: `port 9` in the following name server table:

```
switch1:admin>
switchName:     switch1
switchType:     2.3
switchState:    Online
switchRole:     Principal
switchDomain:   1
switchId:       fffc01
switchWwn:      10:00:00:60:69:10:15:0f
switchBeacon:   OFF
port 0: sw      No_Light
port 1: sw      Online  F-Port  10:00:00:00:c9:21:6d:8f
port 2: sw      No_Light
port 3: sw      No_Light
port 4: sw      No_Light
port 5: sw      Online  F-Port  10:00:00:00:c9:21:2b:25
port 6: sw      No_Light
port 7: sw      Online  F-Port  10:00:00:00:c9:20:d1:63
port 8: --      No_Module
port 9: sw      Online  F-Port  50:06:01:68:fd:10:2a:57
port 10: --     No_Module
port 11: sw     Online  F-Port  20:00:00:60:16:fd:00:1e
port 12: --     No_Module
port 13: --     No_Module
port 14: --     No_Module
port 15: --     No_Module
```

```
# lsattr -El fcp0
loop_id   N/A               Loop id                      False
alpa      N/A               Arbitrated Loop Physical Address False
port_id   0x011700          Port id                      False
port_name 0x500601680103303E                        Port name
False
node_name 0x500601600203303E                        Node name
False
```

SP objects are linked to fcp objects:
lsattr –El sp0
sysser_spnum a10005100062,1 System serial number and SP number False

autoHotSwap	true	Auto hot swap on SP replacement	True
default_HBA	fcp0	Default Host Bus Adapter	False
max_lun_id	223	Maximum LUN number allowed	True
queue_depth	255	SP queue depth	True
sp_reboot_delay	15	SP Reboot Delay in seconds	True

Each AIX server entry in the name–server table of the switch is identified by the name of the server and the `fchan<x>` (see Pid 7 in the name–server table on page 6-14).

Port_id, port_name and node_name can be used to identify the corresponding fchan<x> object in AIX.

Example: Pid 5 in the name server table on page 6-14.

lsattr –El fchan1

bus_intr_lvl	15	PCI Bus interrupt level	False
intr_priority	3	Interrupt priority	False
bus_mem_addr	0xc0002000	PCI Bus memory address	False
bus_mem_addr_rg	0xc0001100	PCI Bus memory address	False
xmt_que_size	256	Transmit queue size	True
topology	pt2pt	Fibre Channel topology	True
class	3	Fibre Channel class	True
network_on	yes	Network operations allowed	True
fcp_on	yes	FCP operations allowed	True
frame_512	no	Frame size limited to 512 bytes	True
num_iocbs	5000	Number of IOCBs to allocate	True
num_bufs	600	Number of ELS/IP buffers to allocate	True
fabric_timeout	0	Extra FCP timeout for fabrics	True
fcpclass	3	Fibre Channel FCP class of service	True
loop_id	N/A	Loop id	False
alpa	N/A	Arbitrated Loop Physical Address	False
port_id	0x011600	Port id	False
port_name	0x10000000C920D163	Port name	False
node_name	0x20000000C920D163	Node name	False

The World Wide Name (WWN) of the switch in `switchshow` command is used by the S@N.IT! to uniquely identify the switch:

```
switchshow
switchName:     switch1
switchType:     2.3
switchState:    Online
switchRole:     Principal
switchDomain:   1
switchId:       fffc01
switchWwn:      10:00:00:60:69:10:15:0f
switchBeacon:   OFF
```

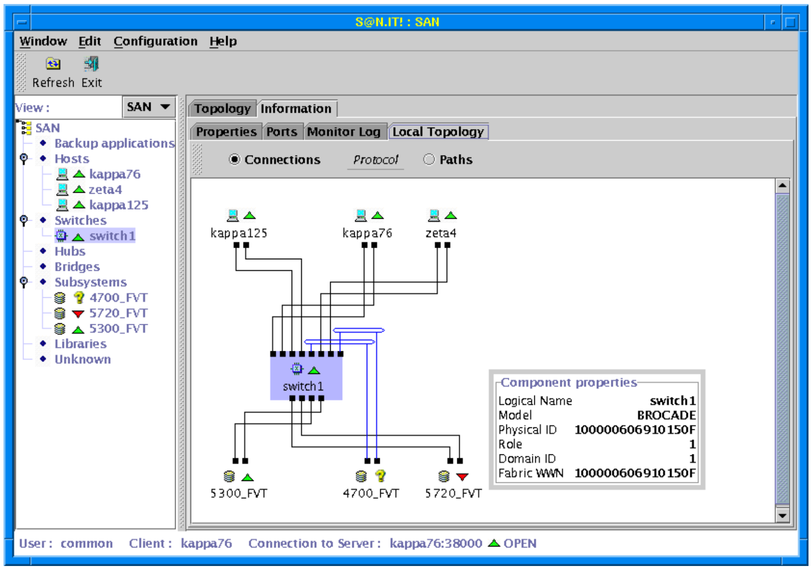


Figure 19. View of the switch through the S@N.IT!

Zoning with cascaded switches

When you wish to zone a fabric with 2 or more linked switches, create the zones on one of the switches (including ports of the other ones), then the configuration will be broadcasted to the other switches on the `cfgsave` command. It is not necessary to include the E–ports in the zones.

See the Brocade document *Fabric OS*and use the on–line command `zonehelp` for more information.

Windows NT Objects

Subsystems visibility

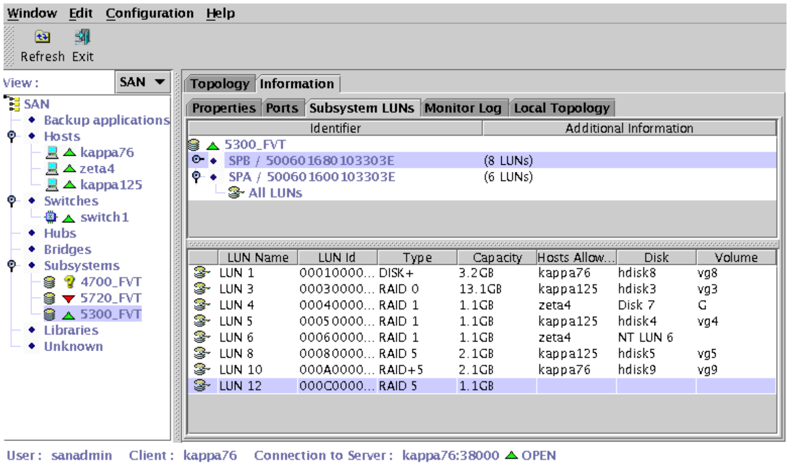


Figure 20. View of subsystems on Windows NT with S@N.IT!

Disk X (A) corresponds to the available path configured at boot time on the Windows NT host.

N*T LUN X (D)* corresponds to the path defined at boot time of the same object.

Emulex adapter and driver identification

The Emulex adapters can be identified using either the `elxcfg` (Emulex Configuration) tool or the SCSI Adapters box in the Control Panel.

Warning: Do not change the LUN mapping checkbox – doing so may cause unpredictable results in S@N.IT! Access Control.

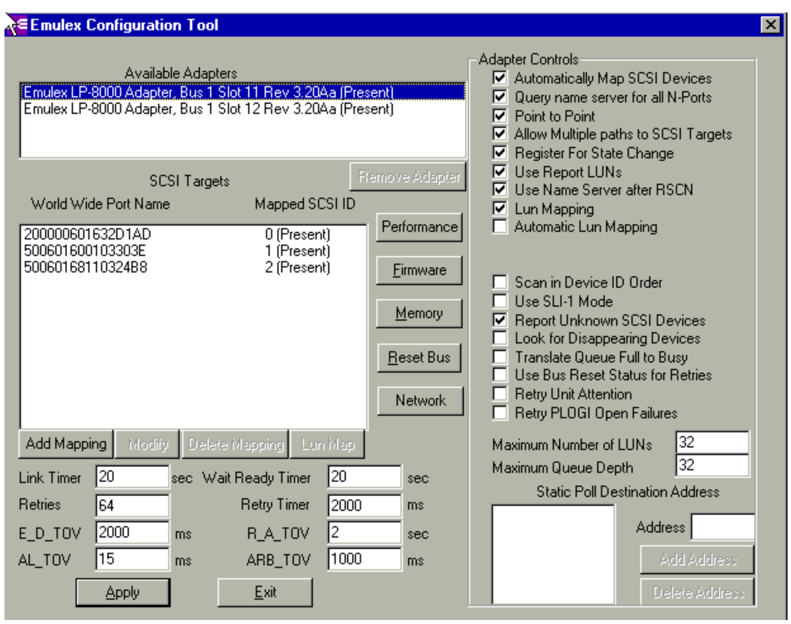


Figure 21. Emulex adapter identification using `elxcfg`.

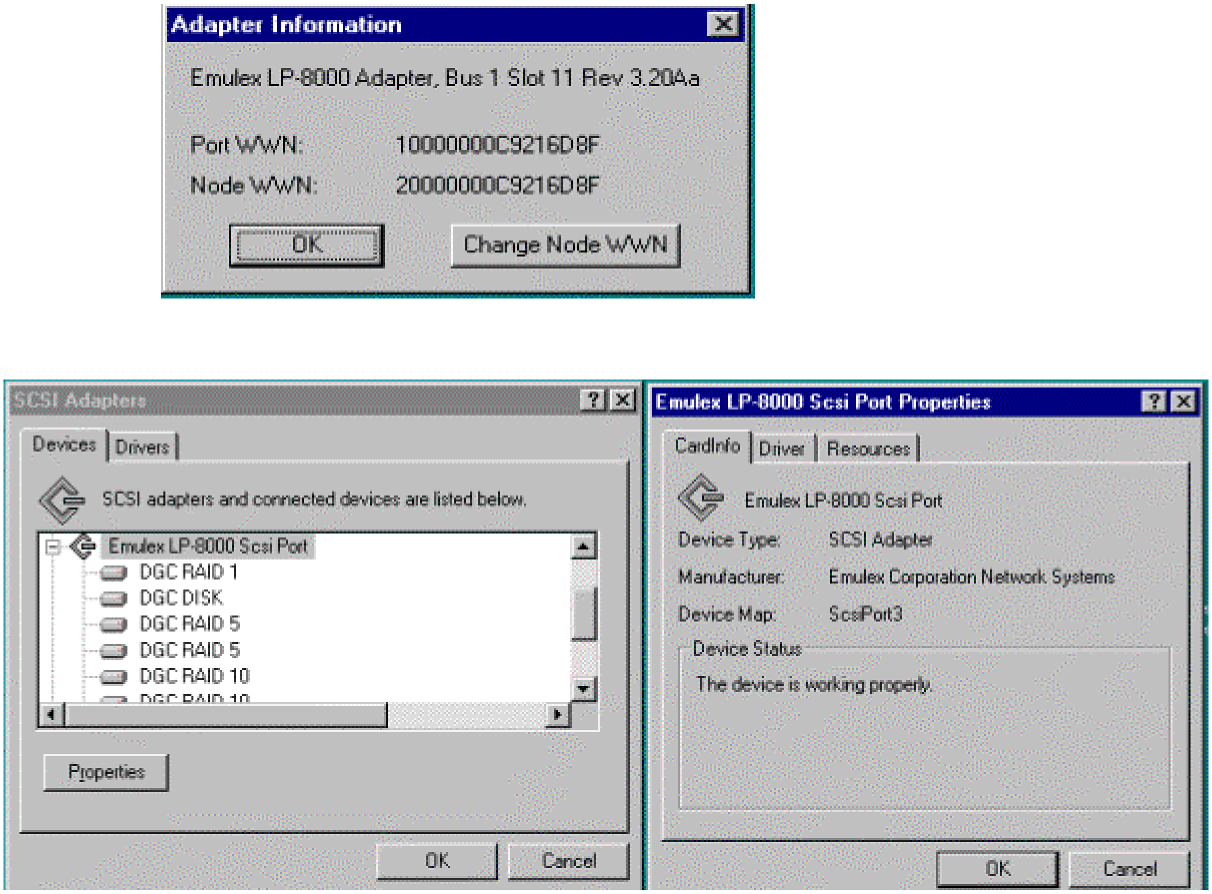


Figure 22. Emulex adapter identification using the SCSI Adapters box

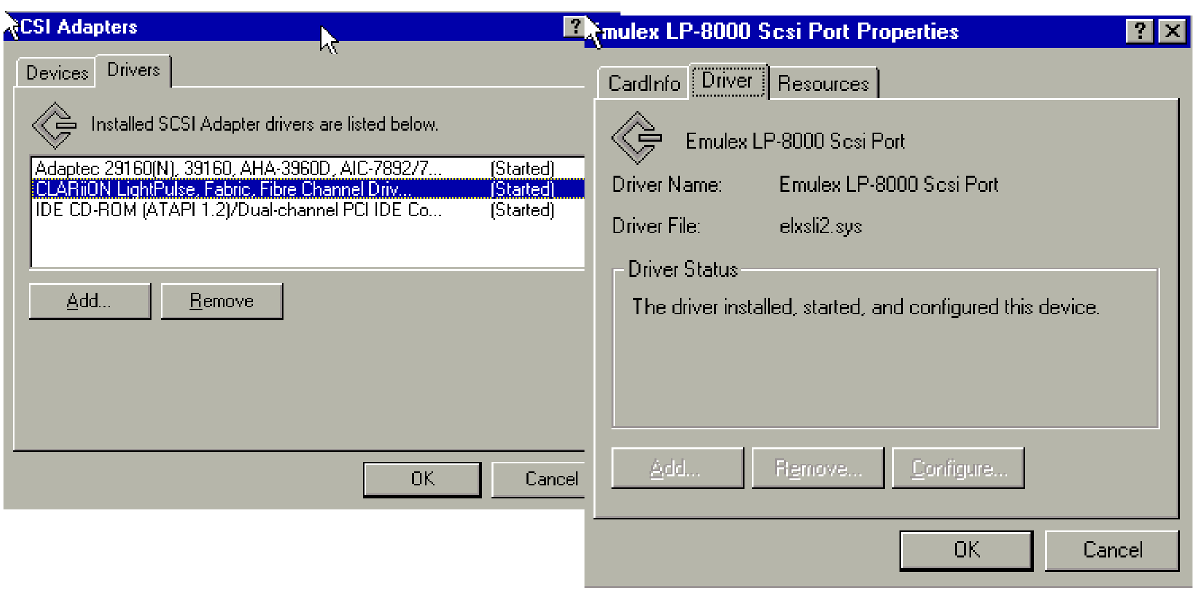


Figure 23. Driver type and installation/configuration state

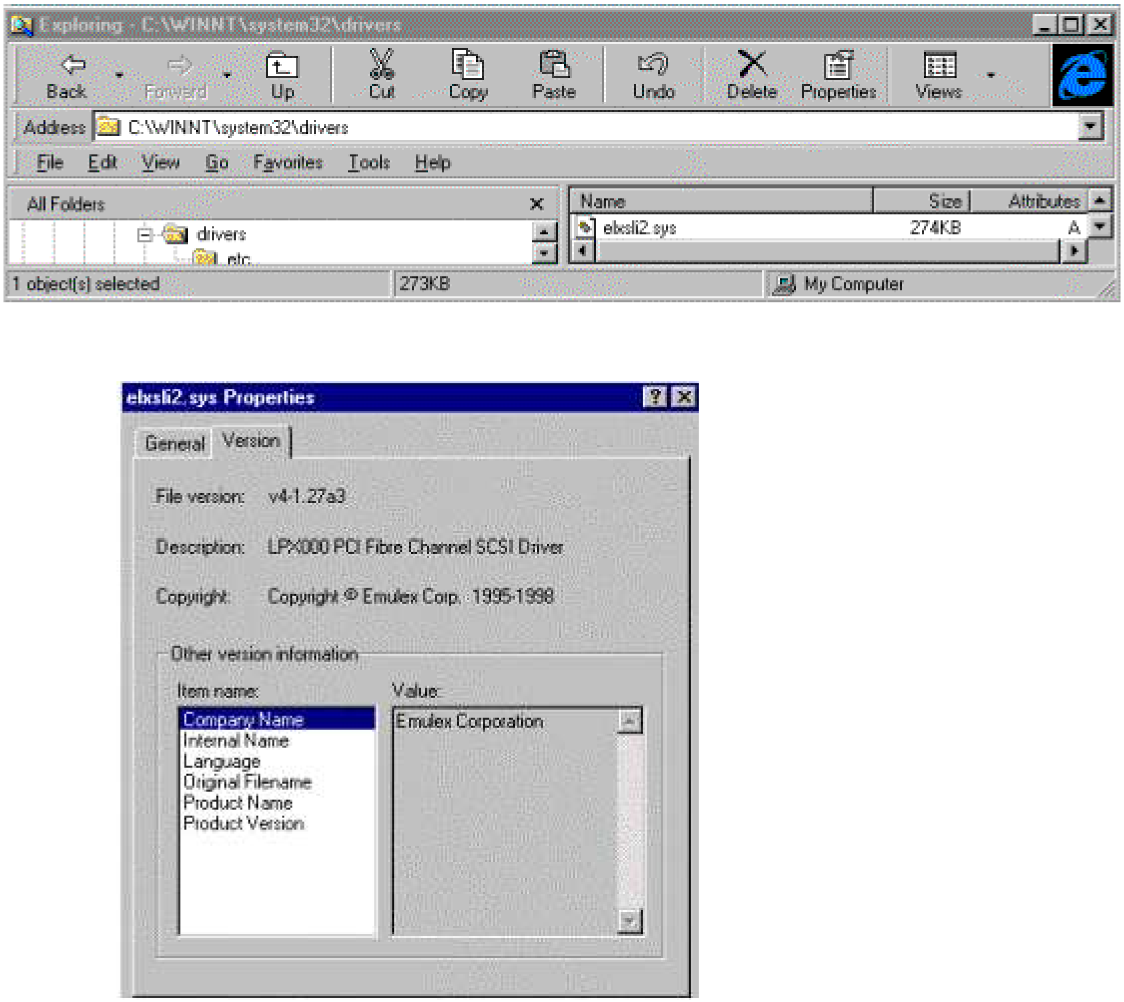


Figure 24. Driver version with `elxsli.sys` file properties on Windows NT

Errlogger

The error log file can be read using the event viewer, which is accessible from the NT Navisphere window by clicking:

```
Start
Programs
Administrative tools
Event Viewer
```

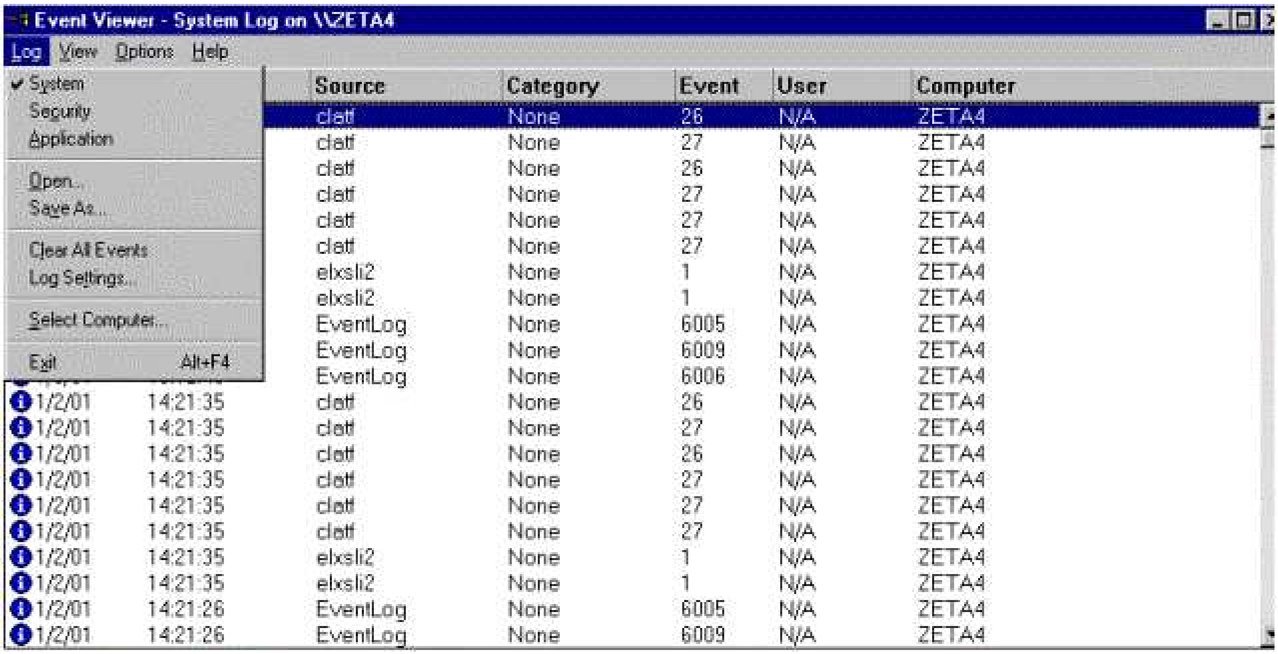


Figure 25. Event viewer

ATF events are posted in the **System** Log.

The Navisphere Agent or Monitor events and the S@N.IT! events are posted in the **Application** Log.

ATF events, and ATF Console

With a double click on a selected line, you get details of an event log:

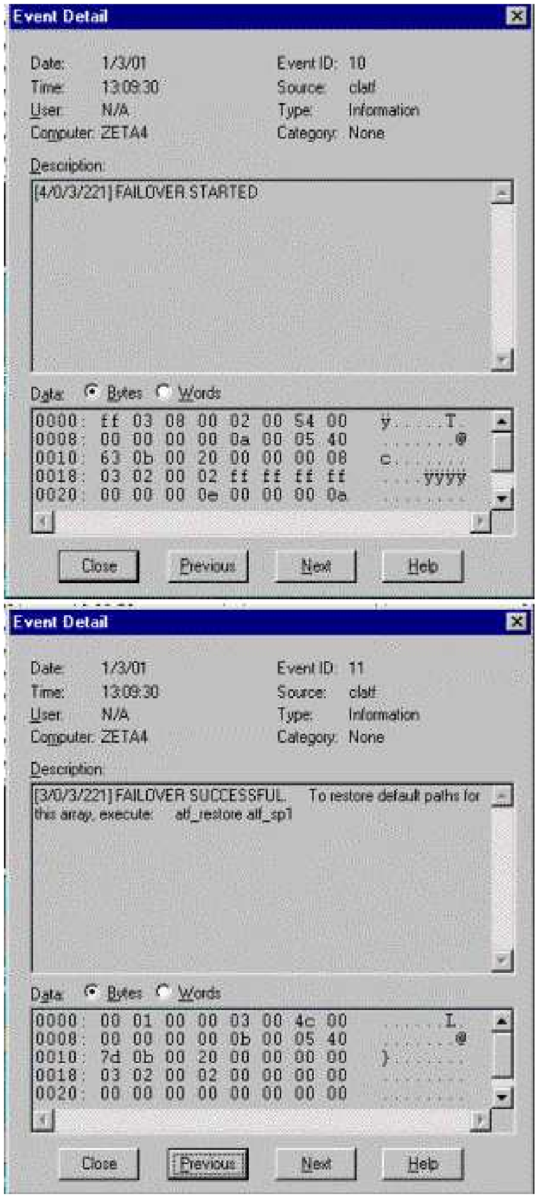


Figure 26. ATF Event Detail with Event Viewer tool

The LUN information in ATF Event Detail is the same as in the Navisphere agent configuration and in the ATF console information:

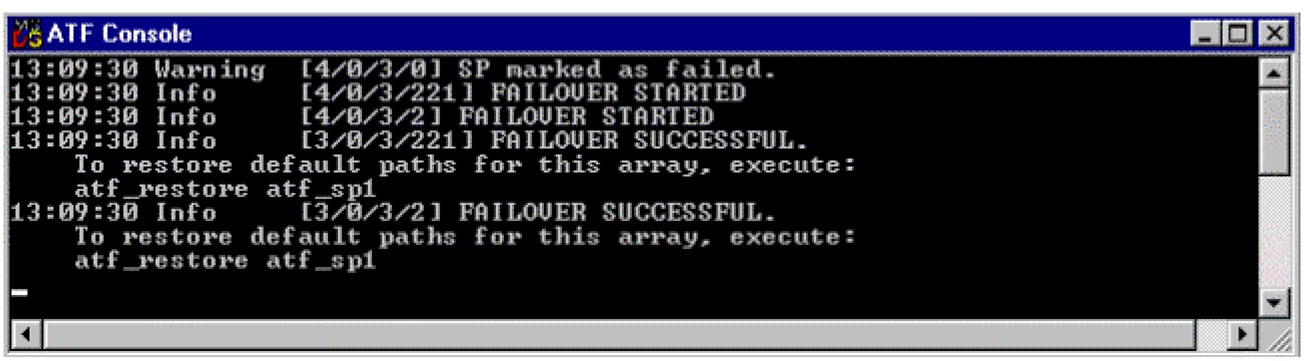


Figure 27. ATF Event Detail

Trespass/Restore

Use the ATF Console to follow the report of the LUN trespass.

Trespass occurs on failover or on `trespass` and `restore` commands.

Check the LUN SP owner after trespass, using the Windows NT Navisphere Manager or the Subsystem LUNs panel in the AIX S@N.IT!.

The following example shows a single LUN trespass using the `trespass` command:

- determine what subsystem (`atf_spX`) and which storage processor are concerned by the trespass using the Navisphere Agent and Manager
- launch the `atf_trespass` command and check the result on:
  - the ATF console
  - the event viewer
  - the Navisphere manager and the SAN.

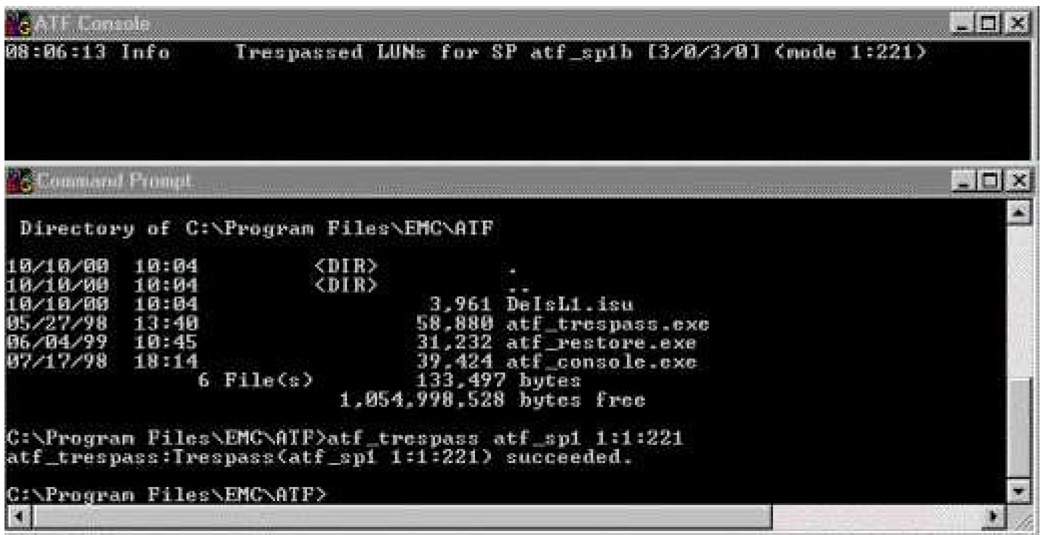


Figure 28. ATF trespass

Here, in the command `atf_trespass atf_sp1 1:1:221`

- `atf_sp1` corresponds to the subsystem DAS5300
- the first digit `(1)` corresponds to the SP B to re–assign the LUN to.
- the second digit corresponds to the mode, here `(1)` for a single LUN trespass
- 221 is the LUN ID in decimal mode as seen in Navisphere or S@N.IT! GUI.

Associated event log in the Event Viewer:

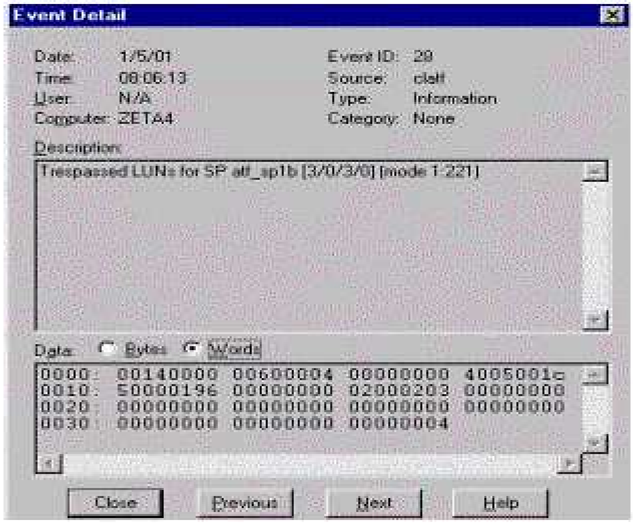


Figure 29. Event log

Note: A bitmap indicating all LUNs affected by the `restore` command is given from address 0020. Each digit indicates a set of LUNs affected by the operation, using hexadecimal values by digit, here:

```
00000000 00000000 00000000 00000000
00000000 00000000 00000004
```

7 x 8 = 56 digits –> 56 x 4 = 224 bits, i.e. 224 possible LUN values.

00 … 04 in hexadecimal corresponds to 0000 0000 …. 0000 0100 in binary => LUN 221

The following example uses the `restore` command to restore the LUN to its default owner SP:

Figure 30. ATF restore

Associated event log in the Event Viewer:

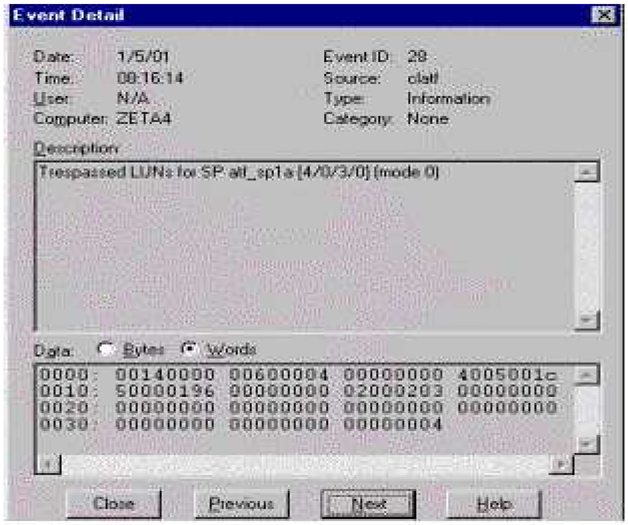


Figure 31. Event log

LUN Management with the S@N.IT!

To create or add a new LUN on a Windows NT host:

1. Bind the LUN using the Navisphere Manager. The new LUN appears on the subsystem as not accessible.

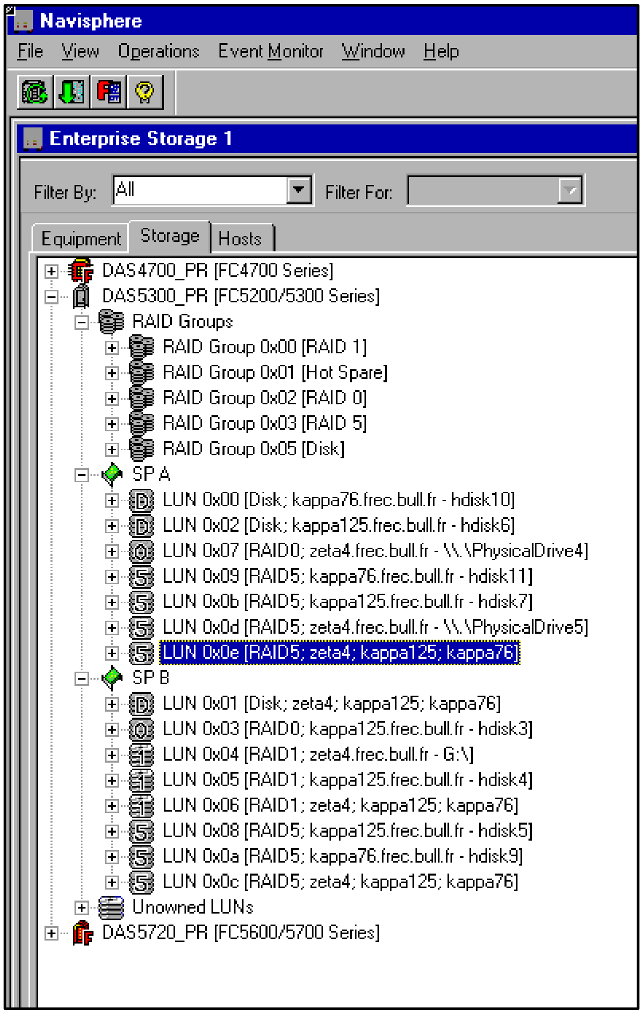


Figure 32. LUN management with the S@N.IT!

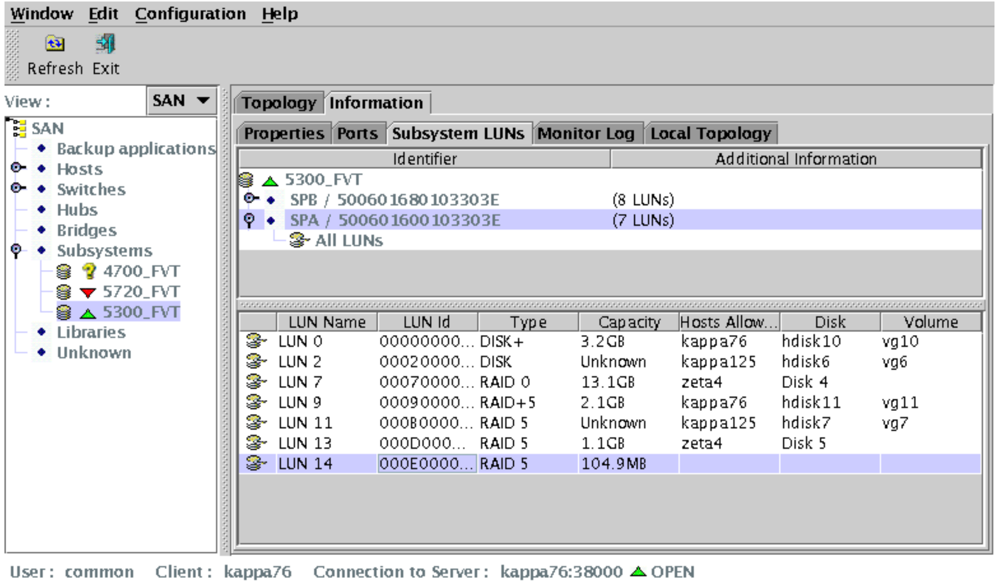


Figure 33. Identification of LUNs

2. Allow access to this LUN with the S@N.IT!. The LUN is seen by NT as LUN X.

See *S@N.IT! User's Guide* documentation.

For DASs with Access Logix flare:

- create a Storage Group (SG) if necessary
- assign LUNs to the SG
- connect a host to the SG
- in S@N.IT!, use the `Host LUNs` option to allow/deny access.

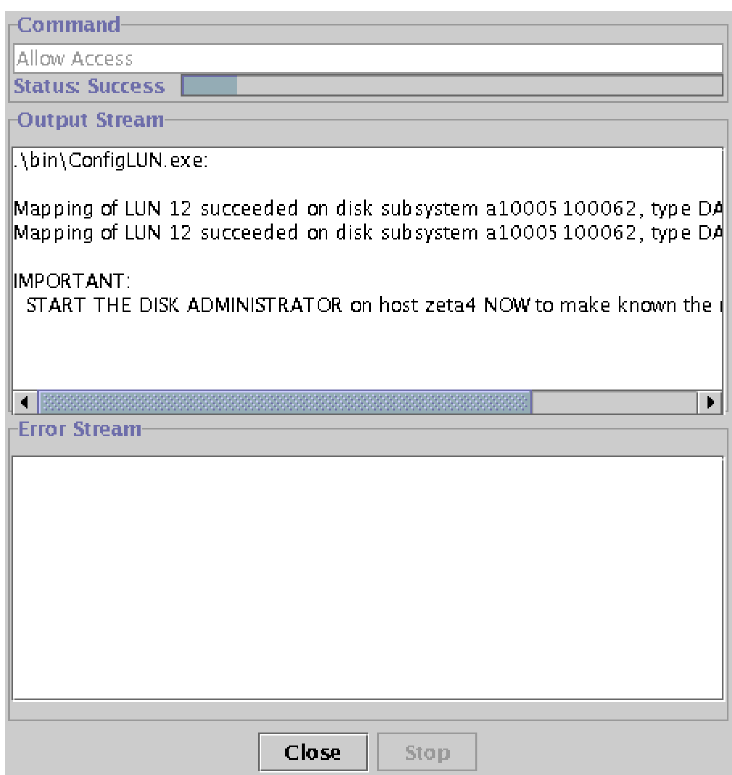


Figure 34. Configure LUNs screen

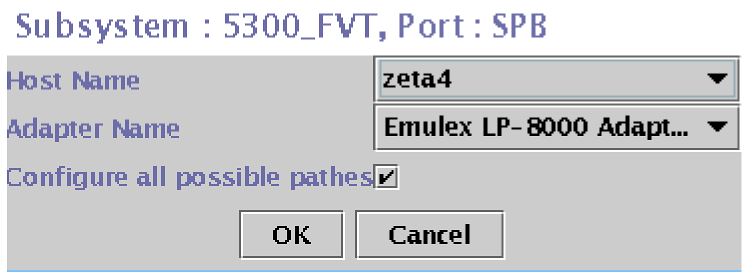
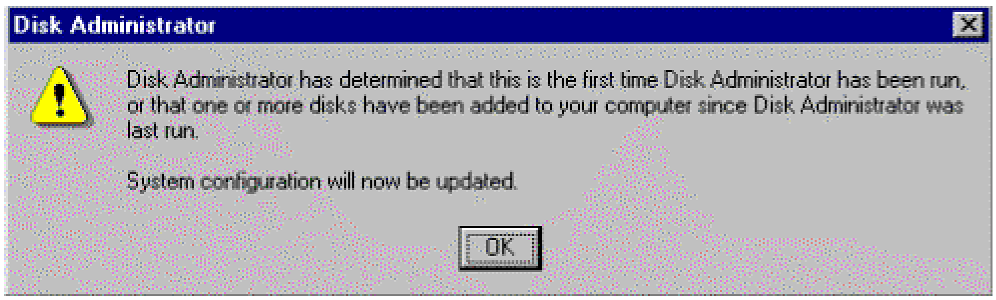


Figure 35. Allow LUN access with the S@N.IT!

3. Run the Disk Administrator on the Windows NT host .

If this is the first time the LUN is allowed, a Windows NT signature must be written on it.



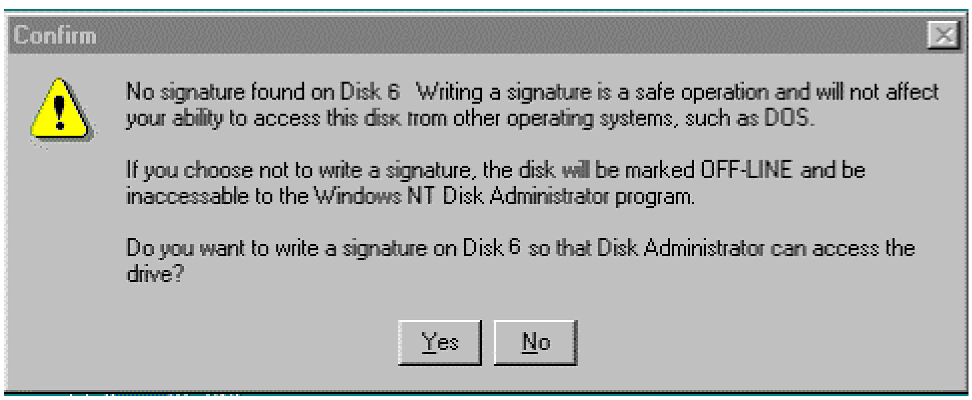


Figure 36. NT signature

When the access to one or several LUNs on a DAS subsystem is changed to allow or deny the access from a Windows NT host with ATF, a reboot is needed to allow ATF to take into account the new configuration (refer to the S@N.IT! User's Guide).

CAUTION:
Make sure the Windows NT LUN Access Control is activated before running the Disk Administrator tool. Otherwise, all the disks of the subsystems are seen and, if a signature is written on a disk assigned to an AIX server, the AIX data will be lost.

Glossary

This glossary contains abbreviations, key–words and phrases found in this documentation.

AIX (Advanced Interactive eXecutive).
IBM UNIX™ operating system derived from AT&T UNIX™ System V.

S@N.IT!
A software that enables storage administrators to efficiently manage SAN for AIX and Windows NT Servers.

ATF (Automatic Transparent Failover).

ATM (Asynchronous Transfer Mode).

Bandwidth
This is the range of signal frequencies that can be carried on a communications channel. The capacity of a channel is measured in bits per second. Bandwidth varies according to the sort and method of transmission.

BCV (Business Continuance Volume)
See EMC documentation for more information.

CD (Channel Director).

DAS (Disk Array Sub–system)
A fibre channel storage device that provides JBOD or RAID capabilities, terabytes of disk storage capacity, high transfer rates, flexible configurations, and high availability.

DAE (Disk Array Enclosure)
A storage device that includes an enclosure, up to 10 or 30 disk modules (depending on model), one or two FC LCCs, and one or two power supplies.

DPE (Disk array Processor Enclosure)
A storage device that includes an enclosure, up to 10 disk modules, one or two SPs, one or two FC LCCs, and one or two power supplies.
A DPE can support up to 11 DAEs (each with up to 10 disk modules) in addition to its own 10 disk modules, for a total of 120 disk modules.
You can attach a DPE to one or more servers or external hubs in many different configurations.

ECC (EMC **C**ontrol**C**enter**)**
A proprietry software that provides extensive network management facilities.

EMI (Electromagnetic Interference)

E–port
A port used to connect two switches together to build a fabric.

ESD (Electrostatic Discharge)

ESP (Enterprise Storage Platform)
A software that provides simultaneous mainframe and open systems support for Symmetrix systems.

Fabric
The term fabric is used to refer to a set of interconnected switches, even if the set is limited to a single switch.

FC–AL (Fibre Channel Arbitrated Loop)
An arrangement of Fibre Channel stations such that messages pass from one to the next in a ring.

FCP (Fibre Channel Protocol)
Protocol that allows FC devices to communicate with SCSI devices.

FDDI (Fiber Digital/Distributed Data Interface).

FC (Fibre Channel)
A high–speed, high–bandwidth serial protocol for channels and networks that interconnect over twisted–pair wires, coaxial cable or fiber optic cable.

GBIC (Gigabit Interface Converter).
An interface module that connects different devices (switches, hubs, SPs) to the FC network.

F–port
A port used in a fabric.

GUI (Graphical User Interface).

HBA (Host Bus Adapter)
A host adapter is a SCSI card in a host computer that allows the host to communicate with SCSI devices.

HDD (Hard Disk Drive)

HSSDC (High Speed Serial Data Connection).

Hub
A device to which several others are attached, providing a common point of connection to all other devices in a network.

ID
The unique address of a SCSI device. 8–bit SCSI can have up to eight IDs; 16–bit up to sixteen IDs; 32–bit up to 32 IDs. There must be a minimum of one target and one initiator on the bus. SCSI IDs range from #00 to #07 for 8–bit; #00 to #15 for 16–bit; and #00 to #31 for 32 bit systems.

JBOD (just a bunch of disks)
Another name for DAE.

LAN (Local Area Network)
A local area network is a short–distance network used to link a group of computers together within a building.
10BaseT Ethernet is the most commonly used form of LAN. A hub serves as the common wiring point, enabling data to be sent from one machine to another over the network. LANs are typically limited to distances of less than 500 meters and provide low–cost, high–bandwidth networking within a small geographical area.

LCC (**Link Control Card**).

Location code
Address where a communications adapter is located in the machine. The format is AA–BB–(CC).

L–port
A port that is used in a loop.

LUN (Logical Unit Number)
One or more disk modules (each having a head assembly and spindle) bound into a group — usually a RAID group.
The OS sees the LUN, which includes one or more disk modules, as contiguous disk space.

MIA (Media Interface Adapter).

MP (Multi-processor).

MSCS (Microsoft Cluster Server).
High availability in windows NT: distributed architecture, DB server and CI must be on a separate node.

N–port
A port that is used on a node.

OLTP (Online Transaction Processing)
An integral component of information processing systems, OLTP systems gives multiple users simultaneous access to large databases.

PCI (Peripheral Component Interconnect)

QoS (Quality of Service).
A way to allocate resources in the network (e.g., bandwidth, priority) so that data gets to its destination quickly, consistently, and reliably.

RAID (Redundant Array of Independant (or Inexpensive) Disks).
A RAID provides convenient, low–cost, and highly reliable storage by saving data on more than one disk simultaneously.
At its simplest, a RAID–1 array consists of two drives that store identical information. If one drive goes down, the other continues to work, resulting in no downtime for users.
RAID–1 is not a very efficient way to store data, however. To save disk space, RAID–3, –4, and –5 stripe data and parity information across multiple drives (RAID–3 and –4 store all parity data on a single drive). If a single disk fails, the parity information can be used to rebuild the lost data. Unfortunately, there is a performance trade–off: depending on the type of RAID used, reading and writing data will be slower than a single drive.

SAN (Storage Area Network)
A high speed network that establishes a direct connection between storage elements and servers or clients.

SCSI (Small Computer Systems Interface)
An intelligent peripheral I/O interface with a standard, device independent protocol that allows many different peripheral devices to be attached to the host's SCSI port.
Allows up to 8, 16 or 32 addresses on the bus depending on the width of the bus. Devices can include multiple initiators targets but must include a minimum of one of each.

SFP (Small Form Factor Pluggeable Media)
SFP is a specification for a new generation of optical modular transceivers. The devices are designed for use with small form factor (SFF) connectors, and offer high speed and physical compactness. They are hot–swappable.

SNMP (Simple Network Management Protocol).
A means of communication between network elements allowing management of gateways, routers and hosts.

SP (Storage Processor)
A board with memory modules and control logic that manages the storage system I/O between the server FC adapter and the disk modules.
The SP in a DPE storage system sends the multiplexed fibre channel loop traffic through a link control card (LCC) to the disk units.
For higher availability and greater flexibility, a DPE can use a second SP.

SPOF (Single Point Of Failure)

SRDF (Symmetrix Remote Data Facility)
A software that provides fast data recovery capability in the event of a data center outage.

STP (Shielded Twisted Pair)

Switch
A network device that switches incoming protocol data units to outgoing network interfaces at very fast rates, and very low latency, using nonblocking, internal switching technology.

UTP (Unshielded Twisted Pair)

WebSM (Web–based System Manager).

WWN (World Wide Name)
A Name_Identifier which is worldwide unique, and represented by a 64–bit unsigned binary value.

Index

L

M

N

P

R

S

T

W

Z

Vos remarques sur ce document / Technical publication remark form

Titre / Title : Bull Storage Area Network (SAN) Installation and Configuration Guide

Nº Reférence / Reference Nº : 86 A2 57EF 01

Daté / Dated : February 2002

ERREURS DETECTEES / ERRORS IN PUBLICATION

AMELIORATIONS SUGGEREES / SUGGESTIONS FOR IMPROVEMENT TO PUBLICATION

Vos remarques et suggestions seront examinées attentivement.
Si vous désirez une réponse écrite, veuillez indiquer ci-après votre adresse postale complète.

Your comments will be promptly investigated by qualified technical personnel and action will be taken as required.
If you require a written reply, please furnish your complete mailing address below.

NOM / NAME : ______________________ Date : ________

SOCIETE / COMPANY : ______________________

ADRESSE / ADDRESS : ______________________

____

Remettez cet imprimé à un responsable BULL ou envoyez-le directement à :

Please give this technical publication remark form to your BULL representative or mail to:

BULL CEDOC
357 AVENUE PATTON
B.P.20845
49008 ANGERS CEDEX 01
FRANCE

Technical Publications Ordering Form

Bon de Commande de Documents Techniques

To order additional publications, please fill up a copy of this form and send it via mail to:
Pour commander des documents techniques, remplissez une copie de ce formulaire et envoyez-la à :

BULL CEDOC
ATTN / Mr. L. CHERUBIN
357 AVENUE PATTON
B.P.20845
49008 ANGERS CEDEX 01
FRANCE

Phone / Téléphone : +33 (0) 2 41 73 63 96
FAX / Télécopie +33 (0) 2 41 73 60 19
E–Mail / Courrier Electronique : srv.Cedoc@franp.bull.fr

Or visit our web sites at: / Ou visitez nos sites web à:

http://www.logistics.bull.net/cedoc

http://www-frec.bull.com   http://www.bull.com

CEDOC Reference # Nº Référence CEDOC	**Qty** Qté	**CEDOC Reference #** Nº Référence CEDOC	**Qty** Qté	**CEDOC Reference #** Nº Référence CEDOC	**Qty** Qté
__ __ ____ _ [__]		__ __ ____ _ [__]		__ __ ____ _ [__]	
__ __ ____ _ [__]		__ __ ____ _ [__]		__ __ ____ _ [__]	
__ __ ____ _ [__]		__ __ ____ _ [__]		__ __ ____ _ [__]	
__ __ ____ _ [__]		__ __ ____ _ [__]		__ __ ____ _ [__]	
__ __ ____ _ [__]		__ __ ____ _ [__]		__ __ ____ _ [__]	
__ __ ____ _ [__]		__ __ ____ _ [__]		__ __ ____ _ [__]	
__ __ ____ _ [__]		__ __ ____ _ [__]		__ __ ____ _ [__]	
[_ _] : **no revision number means latest revision** / pas de numéro de révision signifie révision la plus récente					

NOM / NAME : ____________________ Date : __________

SOCIETE / COMPANY : ____________________

ADRESSE / ADDRESS : ____________________

____

PHONE / TELEPHONE : ____________________ FAX : ____________________

E–MAIL : ____________________

For Bull Subsidiaries / Pour les Filiales Bull :

Identification: ____________________

For Bull Affiliated Customers / Pour les Clients Affiliés Bull :

Customer Code / Code Client : ____________________

For Bull Internal Customers / Pour les Clients Internes Bull :

Budgetary Section / Section Budgétaire : ____________________

For Others / Pour les Autres :

Please ask your Bull representative. / Merci de demander à votre contact Bull.

PLACE BAR CODE IN LOWER
LEFT CORNER

BULL CEDOC
357 AVENUE PATTON
B.P.20845
49008 ANGERS CEDEX 01
FRANCE

ORDER REFERENCE
86 A2 57EF 01

Utiliser les marques de découpe pour obtenir les étiquettes.
Use the cut marks to get the labels.

AIX

Storage Area
Network (SAN)
Installation and
Configuration
Guide

86 A2 57EF 01

AIX

Storage Area
Network (SAN)
Installation and
Configuration
Guide

86 A2 57EF 01

AIX

Storage Area
Network (SAN)
Installation and
Configuration
Guide

86 A2 57EF 01