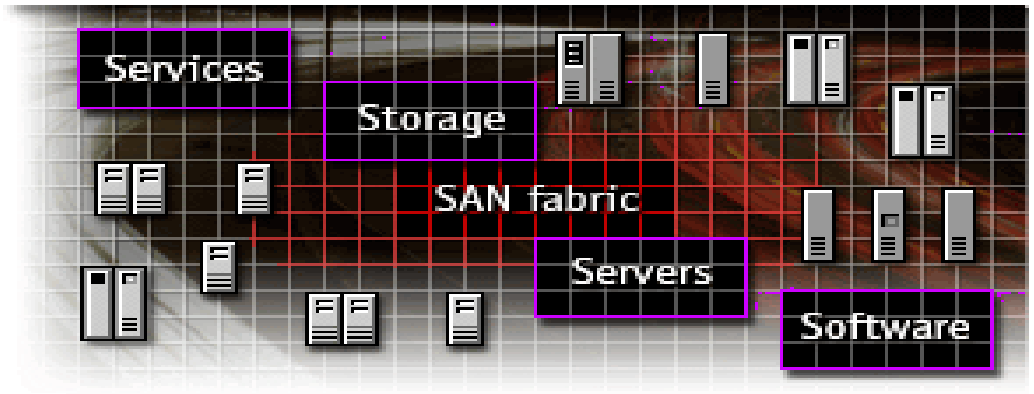
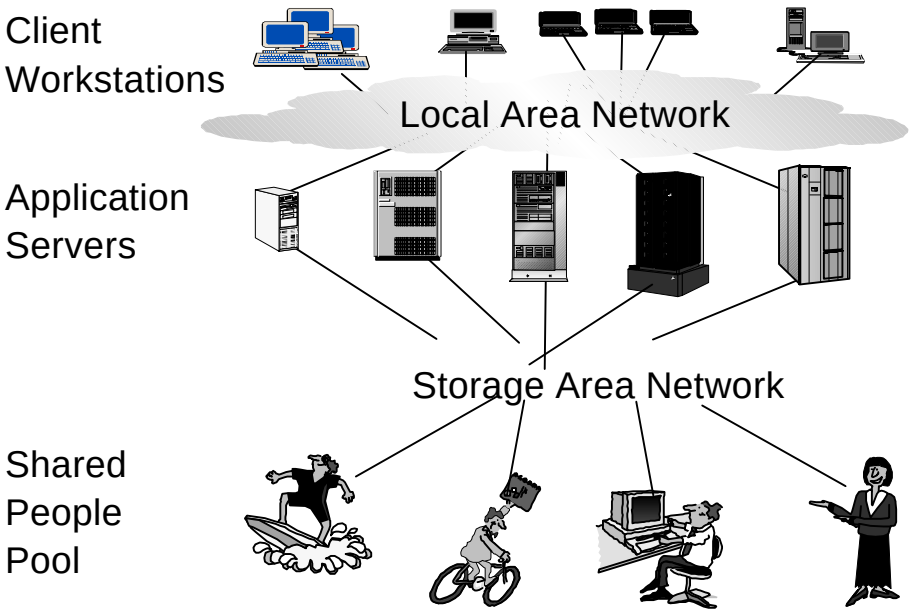
High PerformanceNetworking



Luigi Brochard
Scientific & Technical Computing Architect
IBM EMEA
luigi.brochard@fr.ibm.com
http://www.ibm.com

IBM Enterprise Server

Merging of People & Technology



With the future of 1-inch wearable disk drives: "It could fuel a trend to make every one of us a wireless node in a SAN."

Fred Moore, founder of Horison Information Strategies

Data Size!

- Gigabytes (10*9) (Billion) of disk storage - most customers
- Terabytes (10*12) (Trillion) of disk storage - large customers
- Petabytes (10*15) (Quadrillion) of disk storage - where a few customers will be by 2002 (probably UNIX/NT)
- Petabytes (10*15) (Quadrillion) of tape storage - large customers
- Exabytes (10*18) (Quintillion) of tape storage - a few accounts by 2000
- Zettabytes (10*21) (Sextillion)
- Yottabytes (10*24) (Septillion) - How Long will it take?

Speed Comparisons

Network Speeds

- ✓ "Fast" ethernet and token ring - 100 Megabits/second
- ✓ Gigabit Ethernet (1, 2, 4, 10 Gbs roadmap)
- ✓ ATM (promising to reach 2 Gbs)
- ✓ FDDI
- ✓ TCPIP over Fibre Channel - up to 60-70MBs
- ✓ Sonnet - OC3 (19 MBs) to OC256 (1,658 MBs)
- ✓ HIPPI - 800 Mbs
- ✓ GSN - 6.4 Gbs

Channel Speeds

✓ SCSI - moving from 40 to 80 MBs and up	(25 Meters)
✓ SSA - 80/160 MB/s and up	(10 KM)
✓ FCS (FC-AL, Switched) - 1, 2, 4, 10 Gbs roadmap	(500M - 10KM)
✓ ESCON - 17 MB/s	(43 KM max)
✓ FICON - 100 MB/s (native attach)	(43 KM+)

Agenda

- LAN offering and evolutions
- SAN offering and evolutions

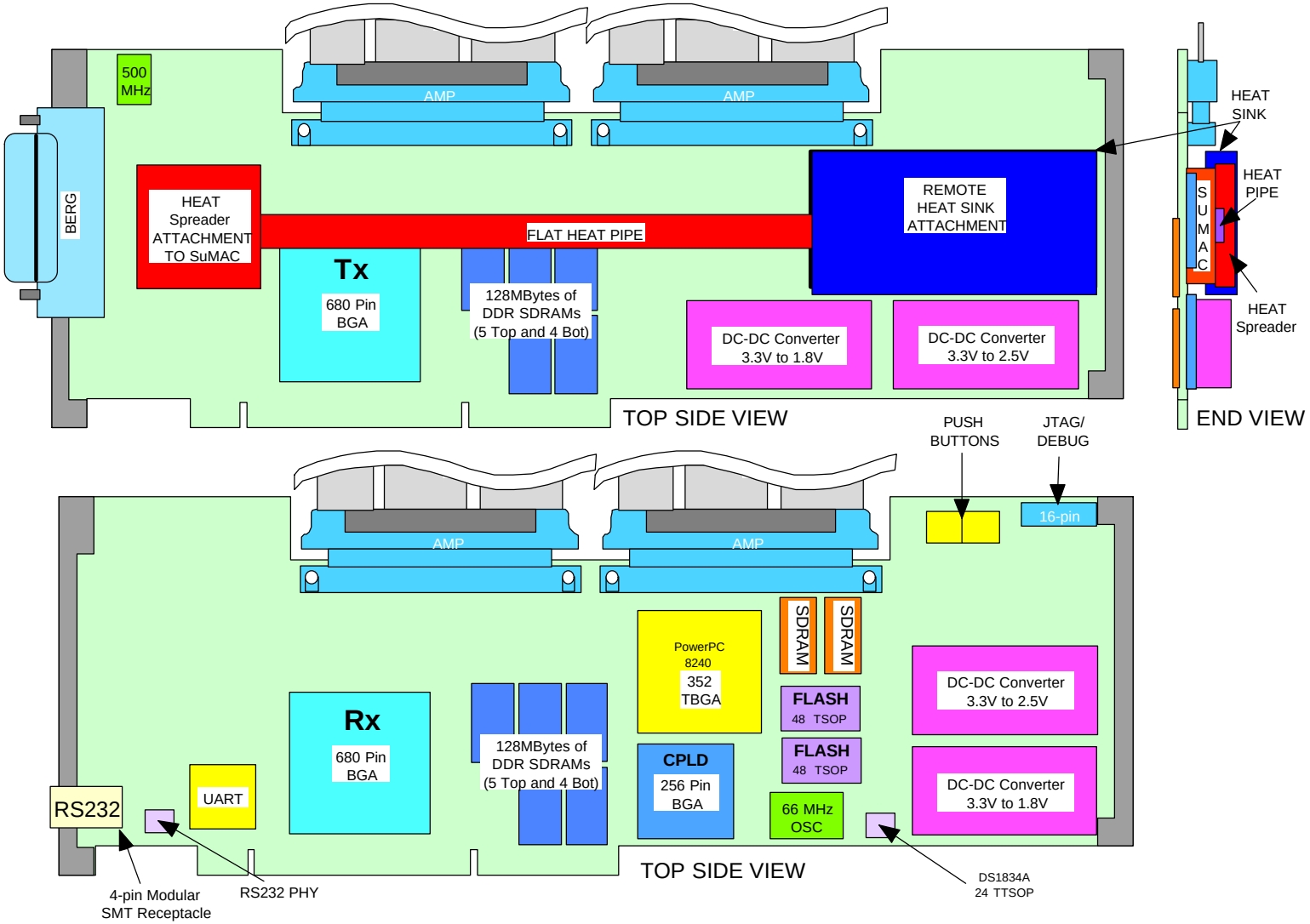
LAN offering

- RS/6000 and RS/6000 SP
  - Ethernet 100, FDDI
  - Ethernet Gigabit
  - HIPPI 800
  - ATM 155
- Netfinity
  - Ethernet 100, FDDI
  - Ethernet Gigabit
  - ATM 155

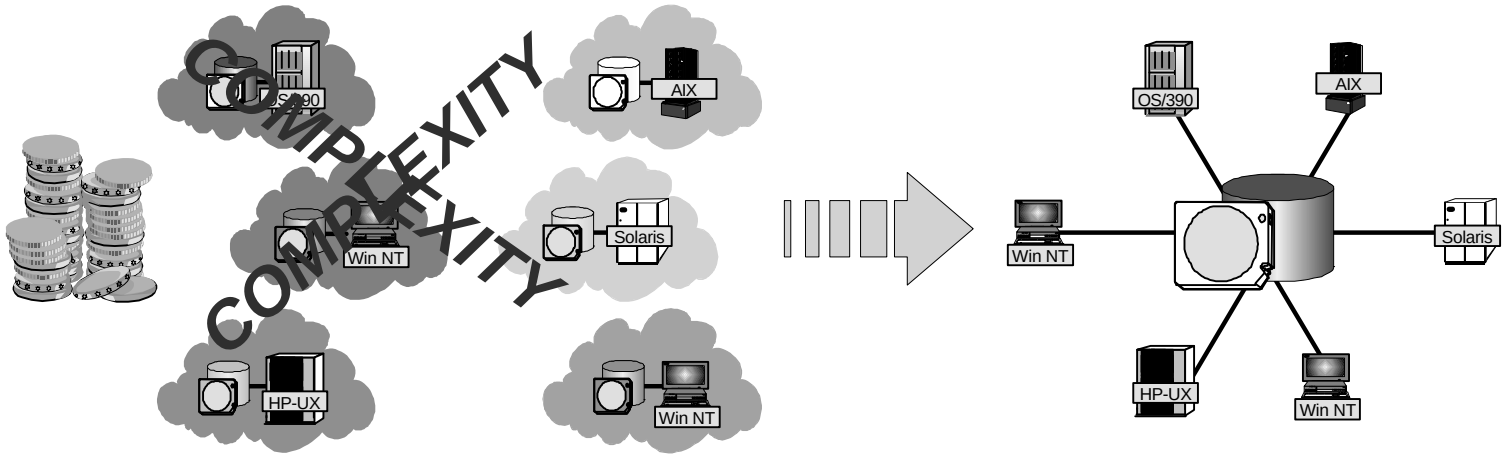
LAN directions

- GSN
  - Collaboration with Genroco:
    - 1/4 speed on PCI 64/66 RS/6000
    - test at ORNL planed on 2Q00 with AIX ST driver
  - full GSN card :
    - full speed GSN with ST HW engine
    - dual card on dual PCI-X (64/133) RS/6000 follow on
    - in development
    - demo planed 2Q01

IBM Dual PCI-X GSN Adapter



Disk and Tape Storage Consolidation



Islands of Information

- ***Distributed servers and storage***
- ***Separate storage management***
- ***Separate islands of information***

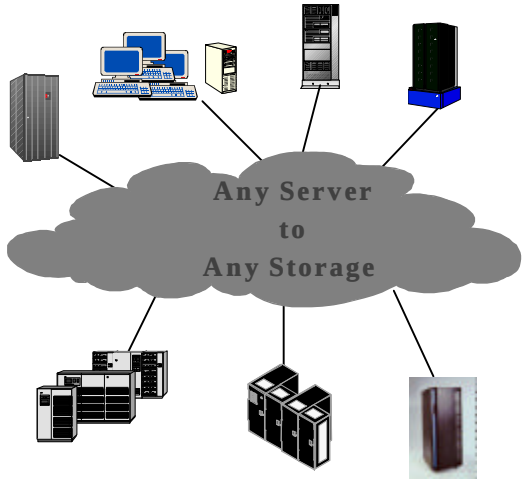
Consolidated Storage

- ***Consolidated storage systems***
- ***Consolidated storage management***
- ***Consolidated enterprise data***

Storage Area Networks can provide consolidation benefits beyond physical consolidation

Storage Area Network (SAN)

- *SAN - **Centrally managed*** high speed networks of ***multivendor*** storage subsystems, applications servers, clients and networking hardware that allow compagnies to exploit the ***value of their business informations*** *via* ***universal access and sharing of resources.***



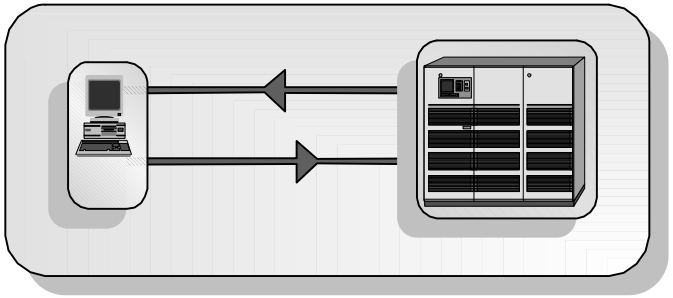
Information used to belong to the server....
NOW it belongs to the Enterprise !

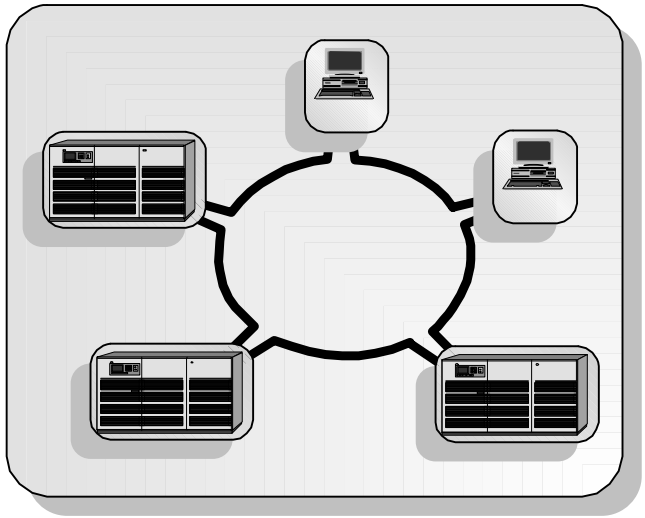
- **SAN** - Dedicated high speed network of directly connected storage elements designed to move large amounts of data between host-independant, distributed storage devices. ***(SNIA)***

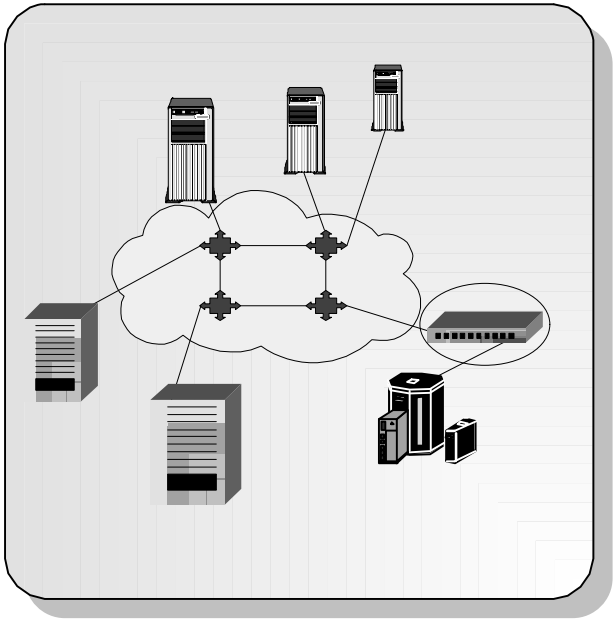
IT "Pain Levels"

"pain levels"	SAN technology enablers
storage connectivity (distance / performance / scalability / addressability)	■ 10km over single mode fiber ■ 100MB/sec bandwidth per link ■ dynamic removal/addition of hosts/devices/paths ■ cabling simplification ■ virtually unlimited addressability via switch cascading
enterprise management	■ consistent management disciplines & standardized tools /processes ■ improved consistency in quality of service ■ lower operational costs ■ enhanced management simplicity
storage consolidation	■ improved asset utilization (reduced idle capacity) ■ capacity-on-demand
backup / recovery	■ reduced backup window ■ minimal impact on production system ■ faster, more effective recovery process
high availability / clustering	■ continuous data availability with minimal performance impacts ■ automatic path selection/failover ■ enhanced load balancing
disaster tolerance (business continuance)	■ remote vaulting/mirroring over 10km (or more with repeaters)
data sharing	■ reduced data duplication ■ enhanced data currency ■ simplified cross-platform sharing

SAN Basics

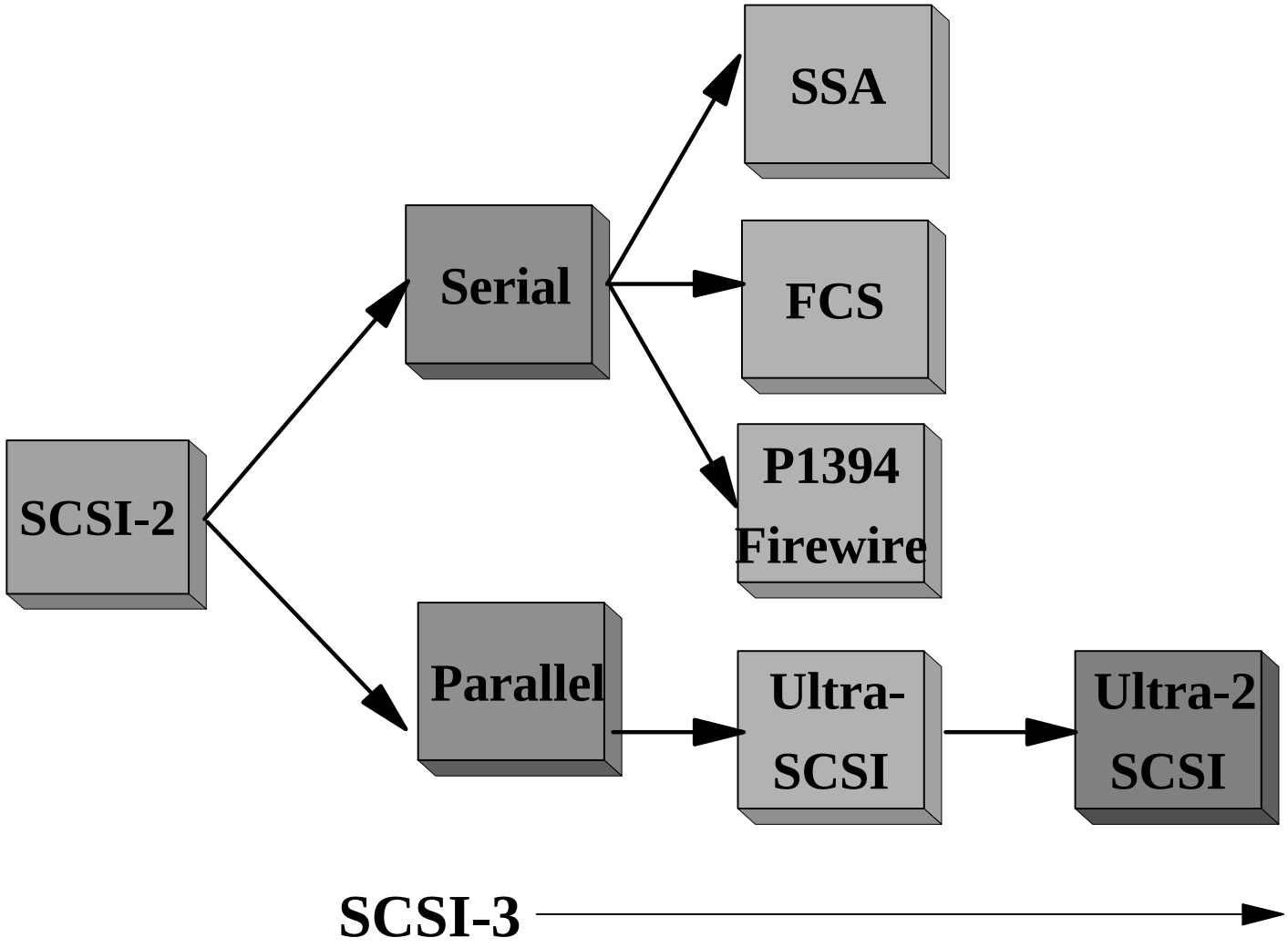






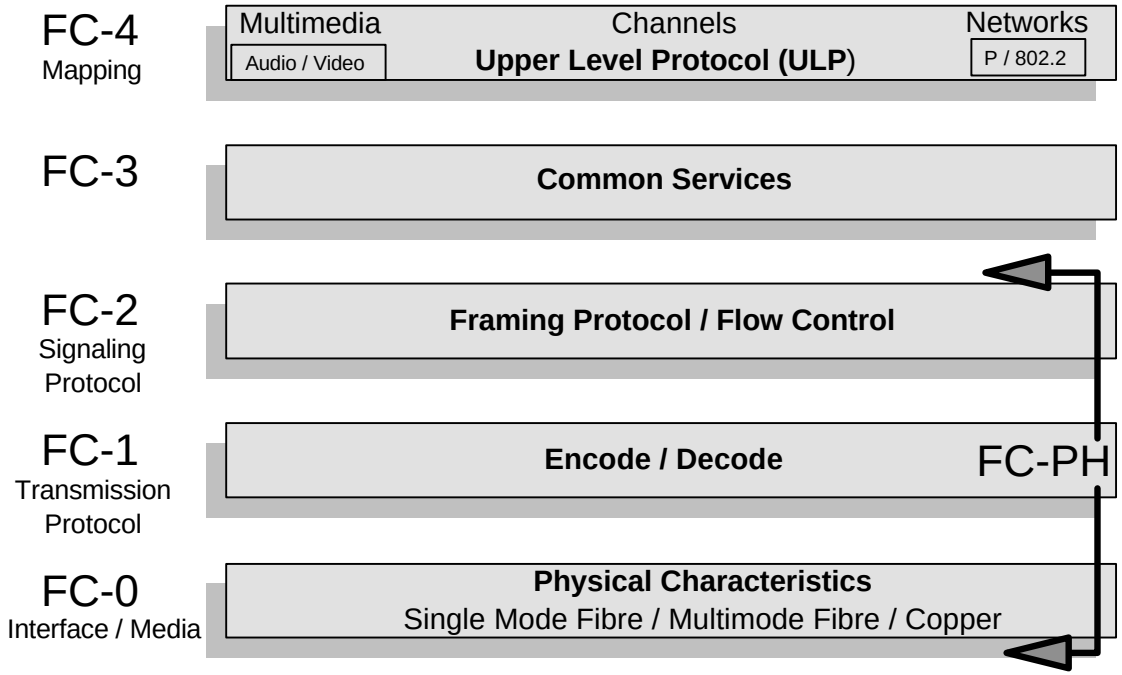
IBM Enterprise Server

SCSI-3 Interface Evolution



IBM Enterprise Server

Fibre Channel Standards



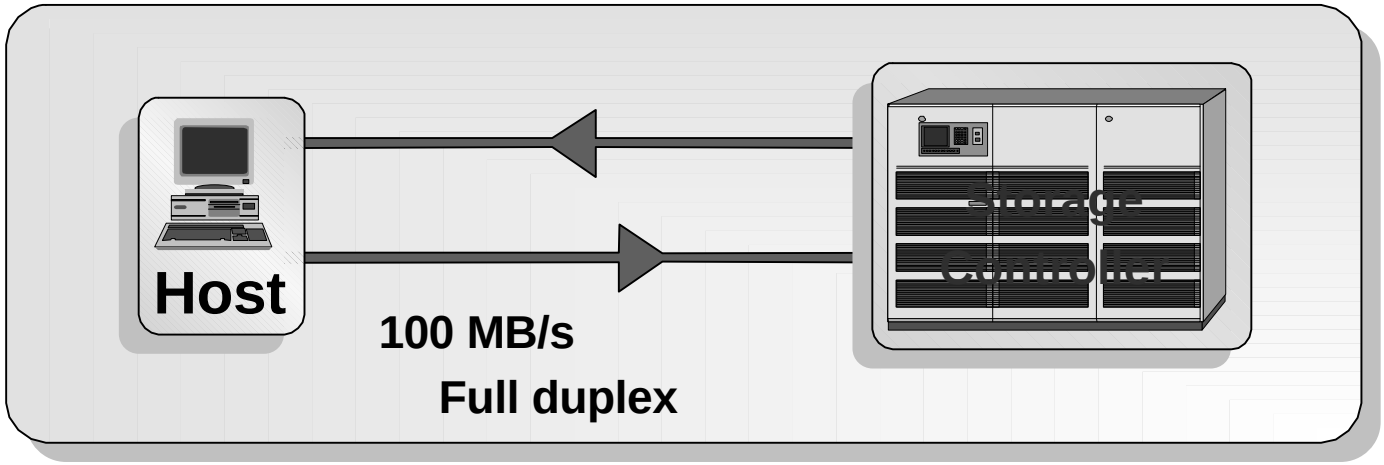
TCP/IP / SCSI / HIPPI / FICON/ Fibre Channel Protocol (FCP)

Copper - 30 Meters
Fiber Optics:
62.5 Micron - 175 Meters
50 Microns (Shortwave) - 500 Meters
9 Micron (Longwave) - 10 KM

Fibre Channel Topologies

Point to point

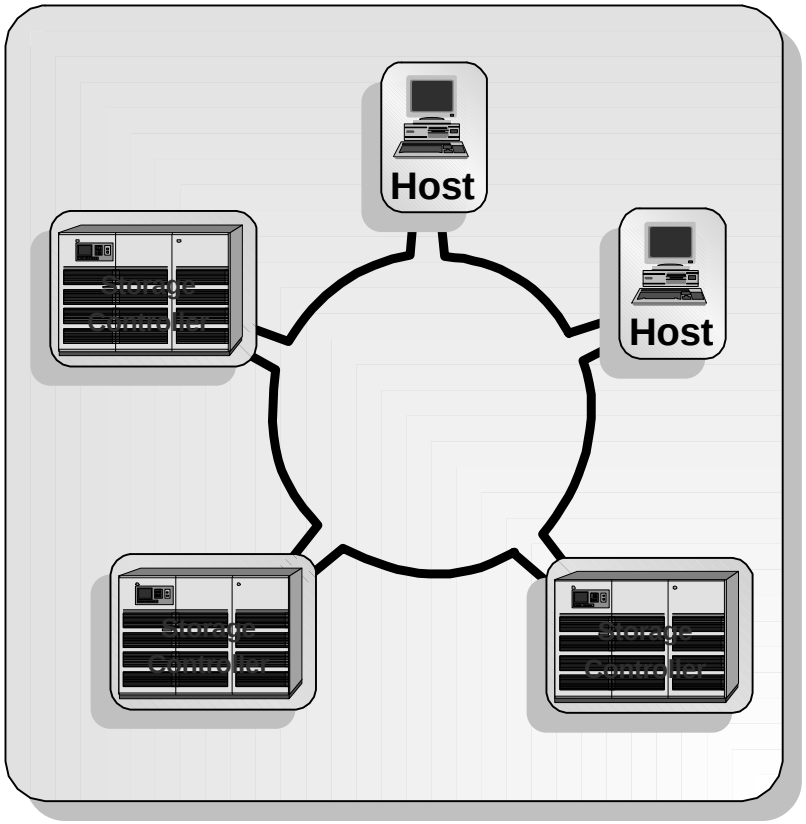
- Direct connection
- Full duplex operation
- Up to 10 KM without repeaters



Fibre Channel Topologies

FC - arbitrated loop

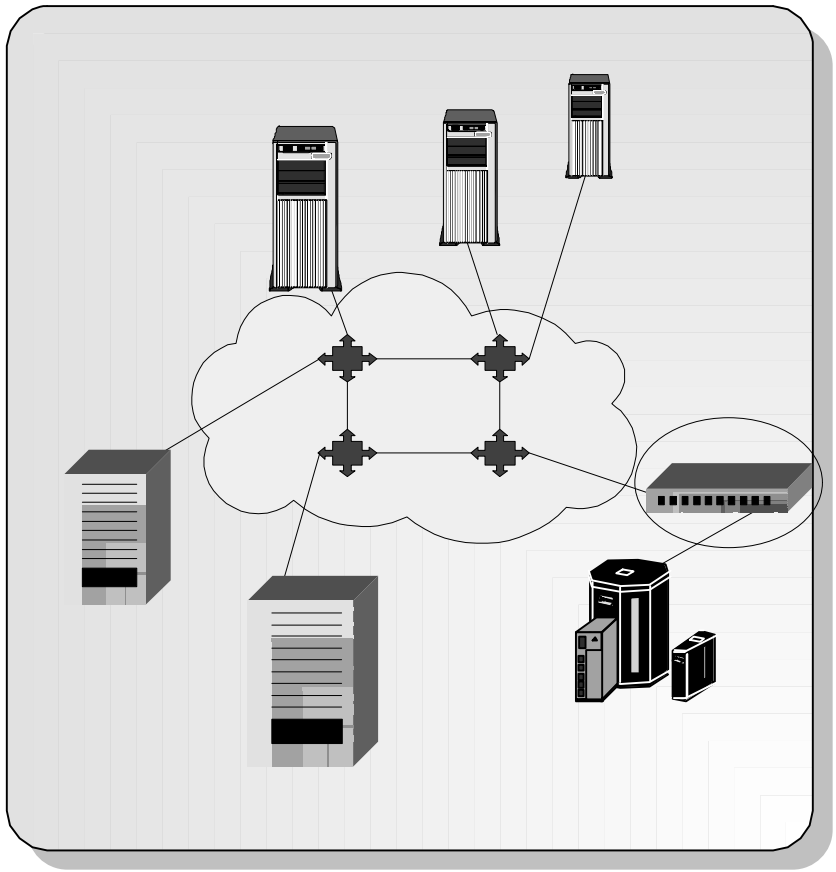
- Shared loop - up to 127 nodes
- Lower cost than point to point
- Cabling with hubs
- Performance affected by
  - Number of nodes
  - Distance
  - Workload
- Hot plugging reconfiguration
- Poor availability without use of two loops



Fibre Channel Topologies

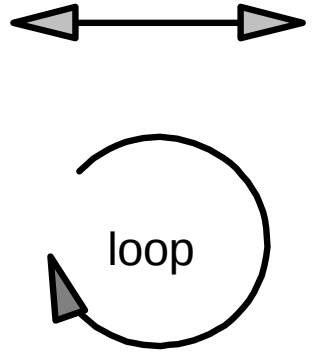
Switched fabric

- Higher bandwidth - 100 MB/s per switched node
- Very large addressability - 16M nodes in domain
- High availability with redundant paths
- High performance
- Up to 10 km distance
- Scalable, flexible re-configuration
- Fabric management required

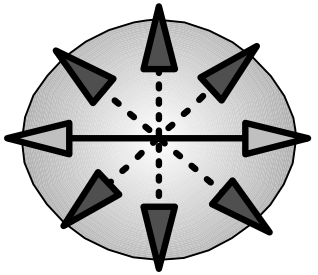


Fibre Channel Performance

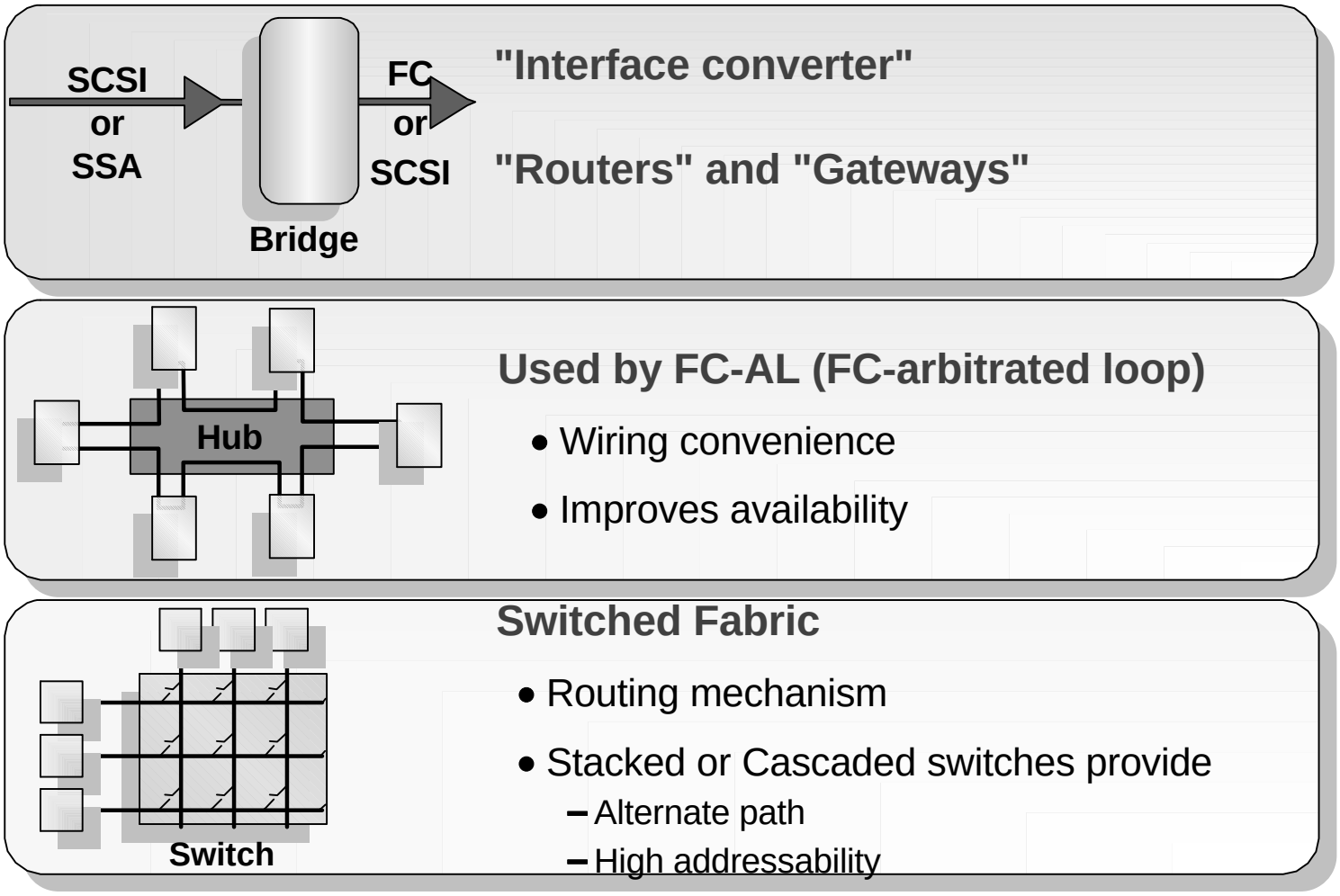
- Fibre Channel performance characteristics
  - Dedicated point-to-point for full performance
    - 100 MB/s full duplex (200 MB/s maximum)
  - FC-AL (Arbitrated Loop) for lower performance
    - 100 MB/s per loop (maximum)
    - Arbitration overhead increases as devices are added
    - Bandwidth is shared between multiple devices



- SAN FC Switch characteristics
  - Non-blocking, any-to-any design
  - 100 MB/sec point-to-point throughput
    - 8-port switch: 800 MB/s, full duplex throughput
    - 16-port switch: 1,600 MB/s, full duplex throughput



Gateways, Hubs & Switches



Fibre Channel SAN Components

- Bridges, Data Gateways and Routers
  - Interface converter between connectivity interfaces
    - Investment protection during migration
  - SAN Data Gateway (FC/SCSI)
    - SAN Data Gateway Router (FC/SCSI)
  - SAN Data Gateway for Serial Disk (SCSI/SSA)
    - Vicom SLIC Fibre Channel Router (FC/SSA)
- Fibre Channel Hubs
  - Fibre Channel Arbitrated Loop (FC-AL) loop in a box
- Fibre Channel Switches
  - Any-to-any connection for FC servers and FC storage

IBM SAN Data Gateway

- Full function gateway
- Industry standard Fibre Channel attachment to wide range of servers
- Simplify migration and accelerate Fibre Channel SAN implementation
- Extended distance between servers and storage
- StorWatch SAN DG Specialist
- *Tivoli Ready* fabric component

Fibre Channel connectivity for IBM SCSI-attached disk and tape storage

Investment protection for SCSI-attached storage

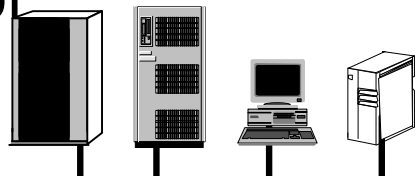
StorWatch
SAN Data Gateway Specialist

Simplifies management of SAN Data Gateways across the enterprise

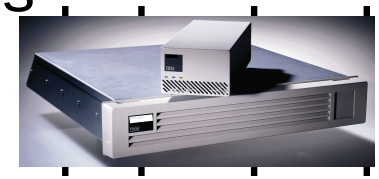
- Graphical User Interface to centrally configure, manage and service multiple SAN Data Gateways across the enterprise
- StorWatch SAN Data Gateway Specialist runs on attached host server and network attached Windows 95, 98, NT workstation

IBM 2108-S20 SAN Data Gateway for Serial Disk

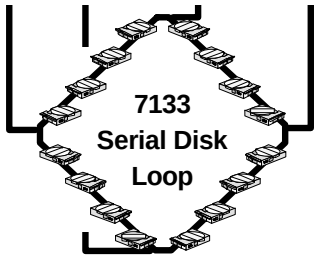
- Provides Ultra SCSI connectivity for UNIX and Intel-based servers
  - Use SCSI host adapters and drivers
- Provides host independent functions
  - Disk concatenation
  - Disk mirroring
  - Instant Copy for tape backup
- IBM StorWatch S20 Specialist
- Highly scalable serial disk loops
  - Mixture of 1 to 8 UNIX and NT hosts
- Maximize Ultra SCSI bandwidth potential



UNIX and Intel-based servers with Ultra SCSI Adapters



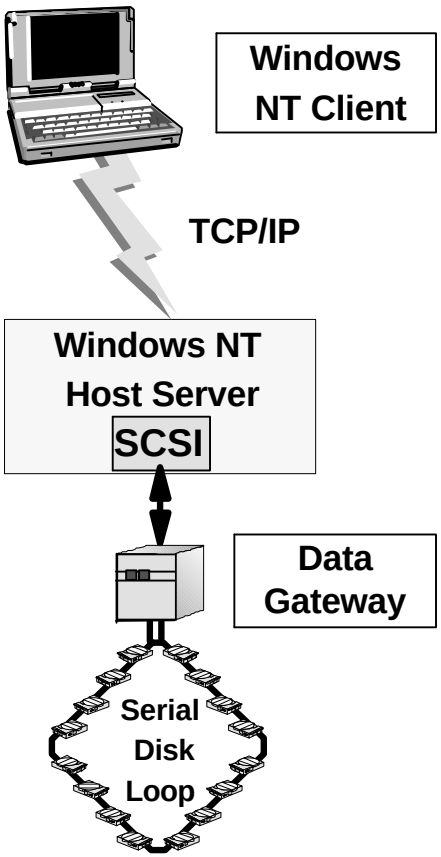
Data Gateway for Serial Disk



Separate controller protects serial disk investment

StorWatch
SAN Data Gateway S20 Specialist

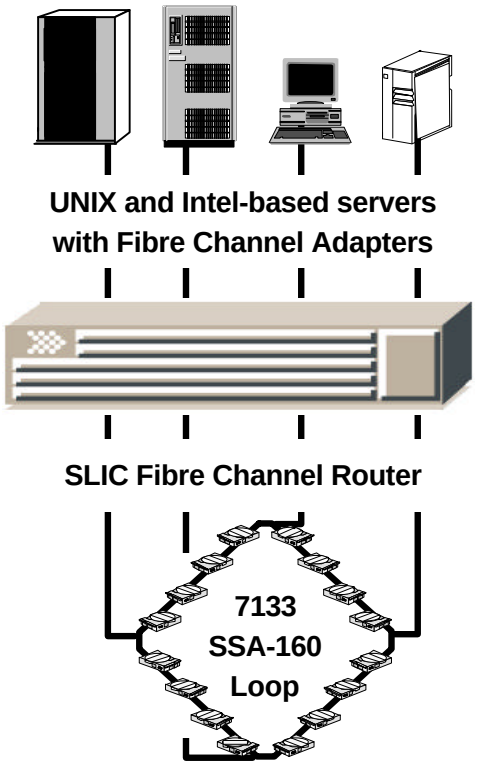
- Windows NT Management Tool
  - Locally or remotely attached system
  - Configure and manage multiple Gateways across the enterprise
- SAN Data Gateway S20 Manager
  - Create composite drives
  - Establish mirrored sets of disks
  - Instant Copy disk management*
    - Procedures and scripts in documentation
    - Customer responsible for setting up application procedures to use this function
- Service Utility
  - Monitoring, reporting and configuring Serial Disks
  - Eases management and service for all supported UNIX and Windows NT servers



* Support currently limited to Sun, HP, Compaq (Digital) and non-clustered Windows NT servers

Vicom SLIC Fibre Channel Router

- Provides Fibre Channel connectivity for UNIX and Intel-based servers
  - Use FC host adapters and drivers
- Provides host independent functions
  - Disk concatenation
  - Disk mirroring
  - Instant Copy for tape backup
- SLIC Manager
- Highly scalable serial disk loops
  - Mixture of 1 to 8 UNIX and NT hosts
- Maximize Fibre Channel bandwidth



Exploit SSA-160 performance potential

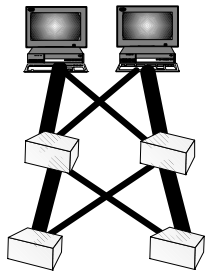
Vicom SLIC, Serial Loop IntraConnect provides SCSI, SSA and FC-AL connectivity

High-function connectivity for IBM Serial Disk storage

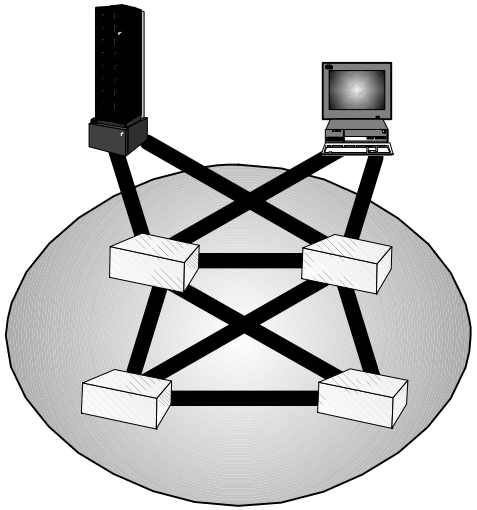
IBM Enterprise Server

Fibre Channel Connectivity

- IBM Fibre Channel Hub
  - Homogeneous server failover applications
  - Extended distance disaster recovery

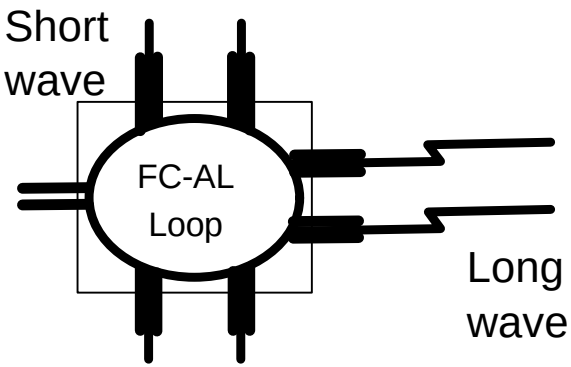


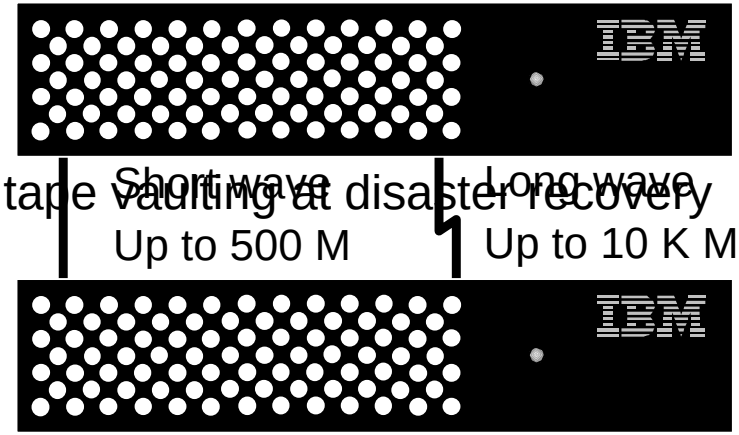
- IBM SAN Fibre Channel Fabric
  - IBM's strategic direction
  - Heterogeneous server and storage sharing applications
  - Interconnected FC switches
    - Redundant pathing
    - Scalable growth
    - Hundreds of ports



IBM Fibre Channel Storage Hub

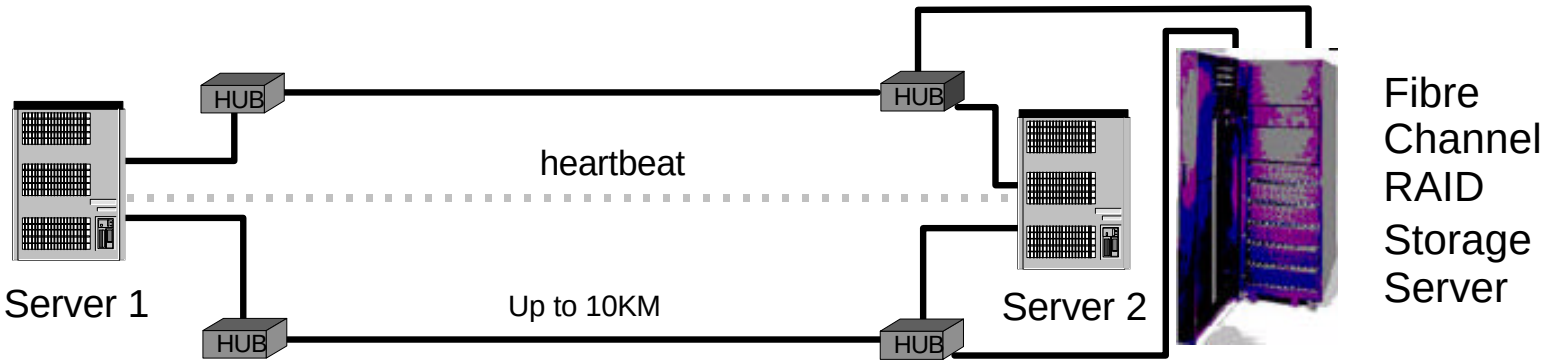
- Seven Port FC-AL Hub
  - Four shortwave optical fiber ports
  - Up to three additional ports
    - Shortwave port supports up to 500 M
    - Longwave port support up to 10 KM
  - Hot pluggable ports
    - Add nodes without down time
  - 1U (1.75 in) high rack space
- Cascade up to two hubs
  - Extended distance
  - Remote disk mirroring and remote tape vaulting at disaster recovery site distances





IBM Fibre Channel Storage Hub

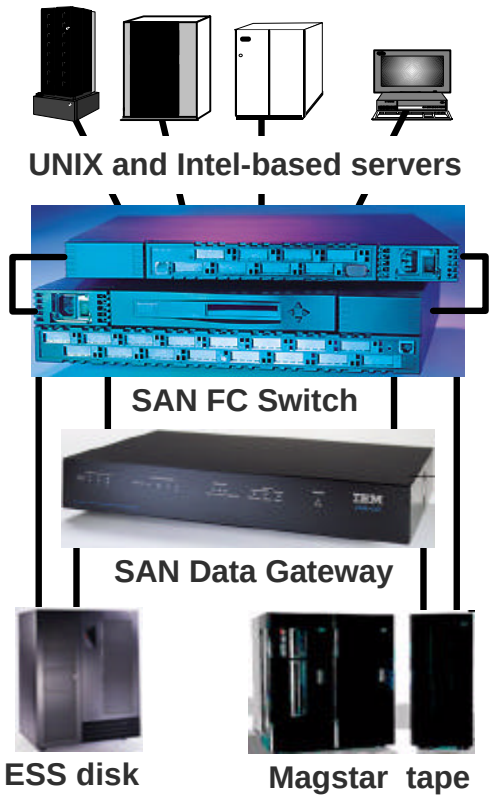
- Homogeneous Server Failover Application
  - Host cluster software manages multiple initiators
  - Normal operation is one initiator and multiple targets
  - Switch to backup initiator connections for failover
  - Normal operation is point-to-point, not arbitrated
  - Hubs convert short wave (500M) to long wave (10 KM)



Automatic Failover Configuration

IBM SAN Fibre Channel Switch

- Fibre Channel connectivity
  - Wide range of servers
  - FC switches and FC hubs
  - FC-attached disk and tape
  - IBM SAN Data Gateway
    - Ultra SCSI-attached disk and tape storage
- 8 port and 16 port models
- 100 MB/s full duplex FC ports
- Up to 10 KM distance
- StorWatch FC Switch Specialist
- *Tivoli Ready* fabric component



Enabling IBM Enterprise Storage Area Networks

IBM Enterprise Server

IBM SAN FC Switch Models

- Common features
  - Four short-wave GBICs
  - Ethernet port for Web interface
  - Rack mount or desktop
- Optional features
  - Short or long-wave GBICs
  - Redundant power supply
  - Fibre channel cables
- Eight port model
  - 1 to 4 additional GBICs
  - Serial port for telnet terminal command line interface
  - 1U form factor
- Sixteen port model
  - 1 to 12 additional GBICs
  - Two -line, 40 character LCD with four push buttons
  - 2U form factor

IBM 2108 Model S08

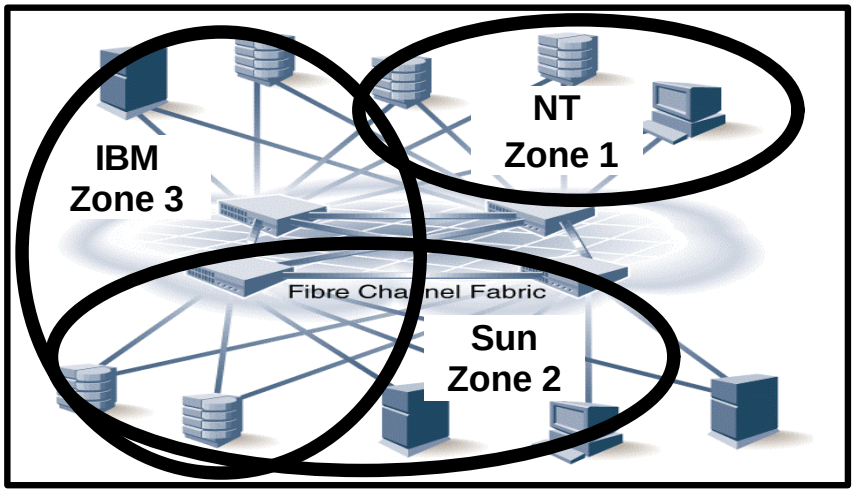
IBM 2108 Model S16

Intelligent Switch Fabric Self-management

- Automatic configuration management
  - Select type of port and initialize connection
  - Discover and register new hosts, devices, switches
  - Disable port connection if link fails
  - Re-enable port connection when link condition is resolved
- Industry standard services supported
  - Simple Name Server (SNS) registers hosts and devices
  - Registered State Change Notification (RSCN) information
  - Virtual Private SAN (VPS) limits access with zoning
  - Private loop devices registered as public devices
- Switch fabric routing
  - Dynamic path selection of most efficient routing
  - Automatic path failover

StorWatch
SAN FC Switch Specialist

- Fibre Channel SAN Zoning
  - Create logical groups within physical fabric
  - Isolate homogeneous servers and storage
  - Control data access
- Zoning management
  - Administrator sets policy
    - Dynamic configuration
    - Overlapping zones
  - Fabric enforces zoning
- Applications
  - LAN-free backup
  - Storage resource sharing

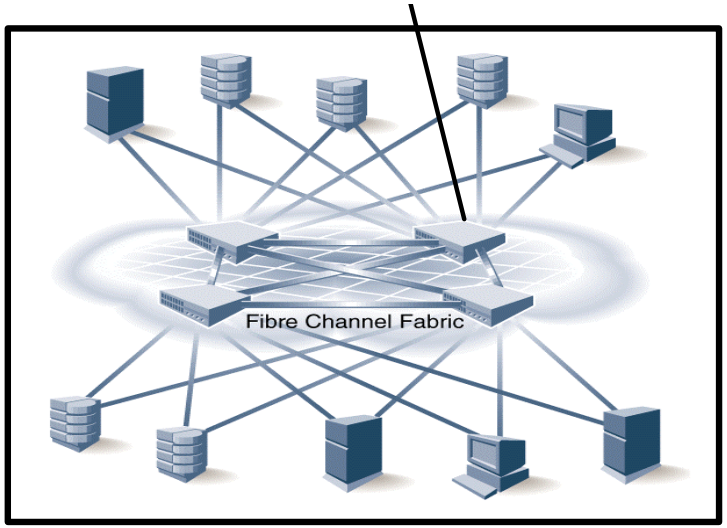


Assures data security and integrity

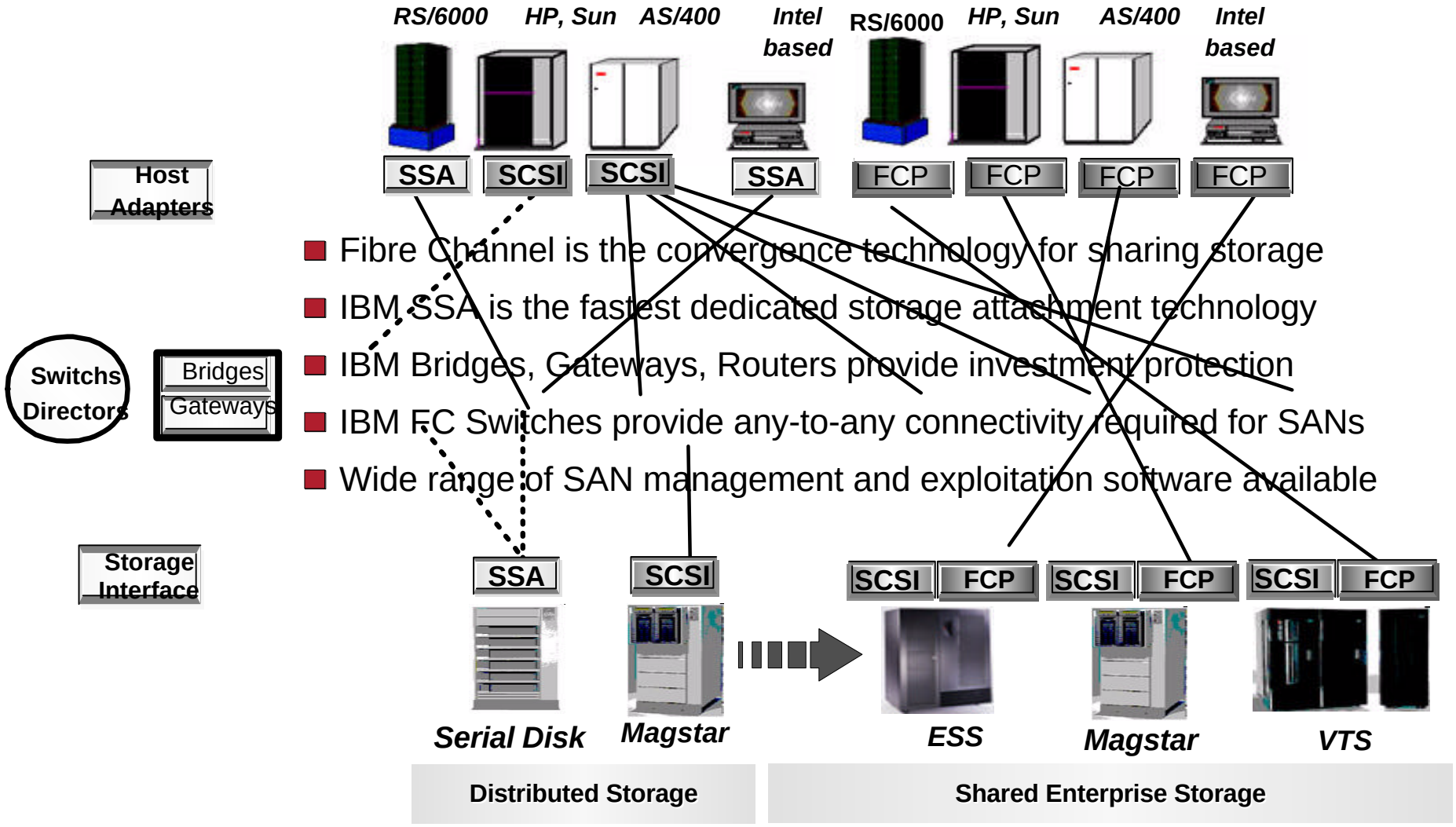
StorWatch
SAN FC Switch Specialist

- Remote switch fabric management
  - Netscape or IE browser interface
  - Windows NT or UNIX host
- Management tools
  - Fabric (SAN) View
  - Fabric Topology View
  - Name Server Table View (New)
  - Zone Administrative Interface (New)
  - General Switch View
  - Performance View
  - Port Detail View
- Administrative tools
  - Administrative Interface
  - Telnet Interface
  - Pop-up help for error conditions

Ethernet connection
to any switch

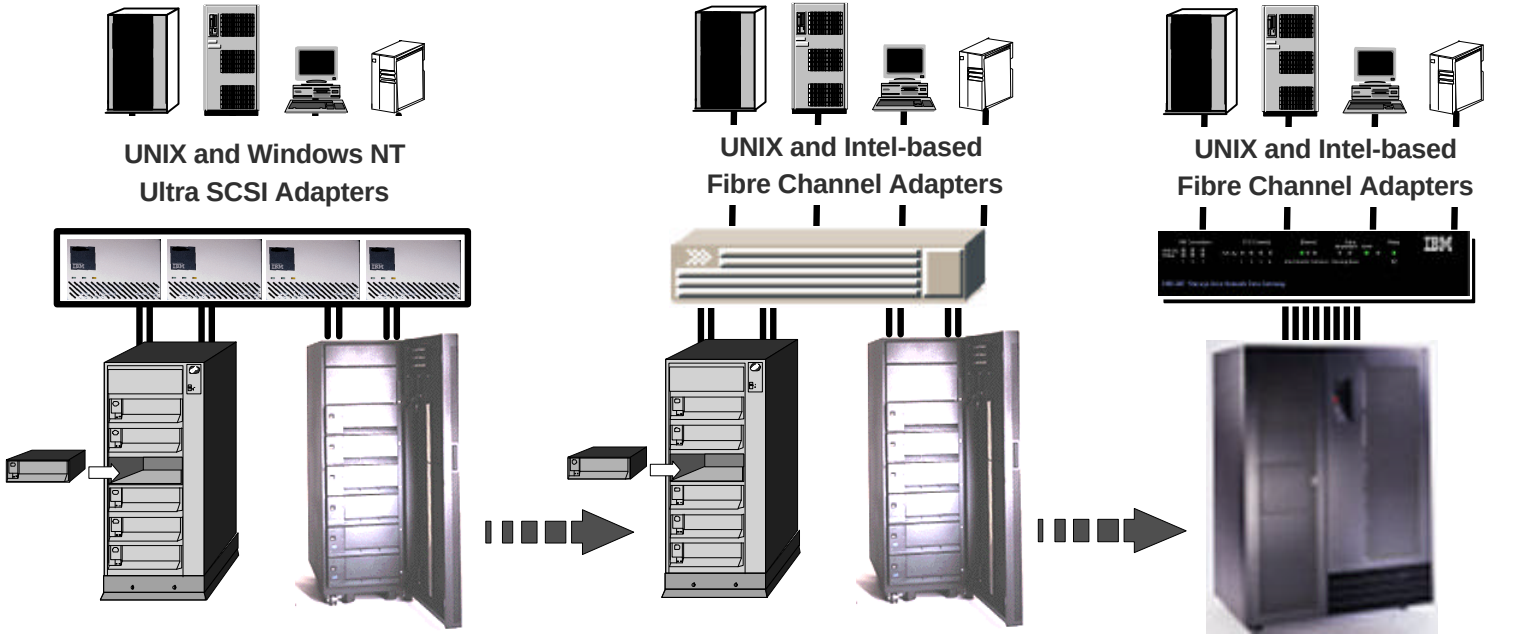


IBM Enterprise SAN



IBM Enterprise Server

Separate Disk Controller and Disk Drawers



Serial Disk storage investment protection

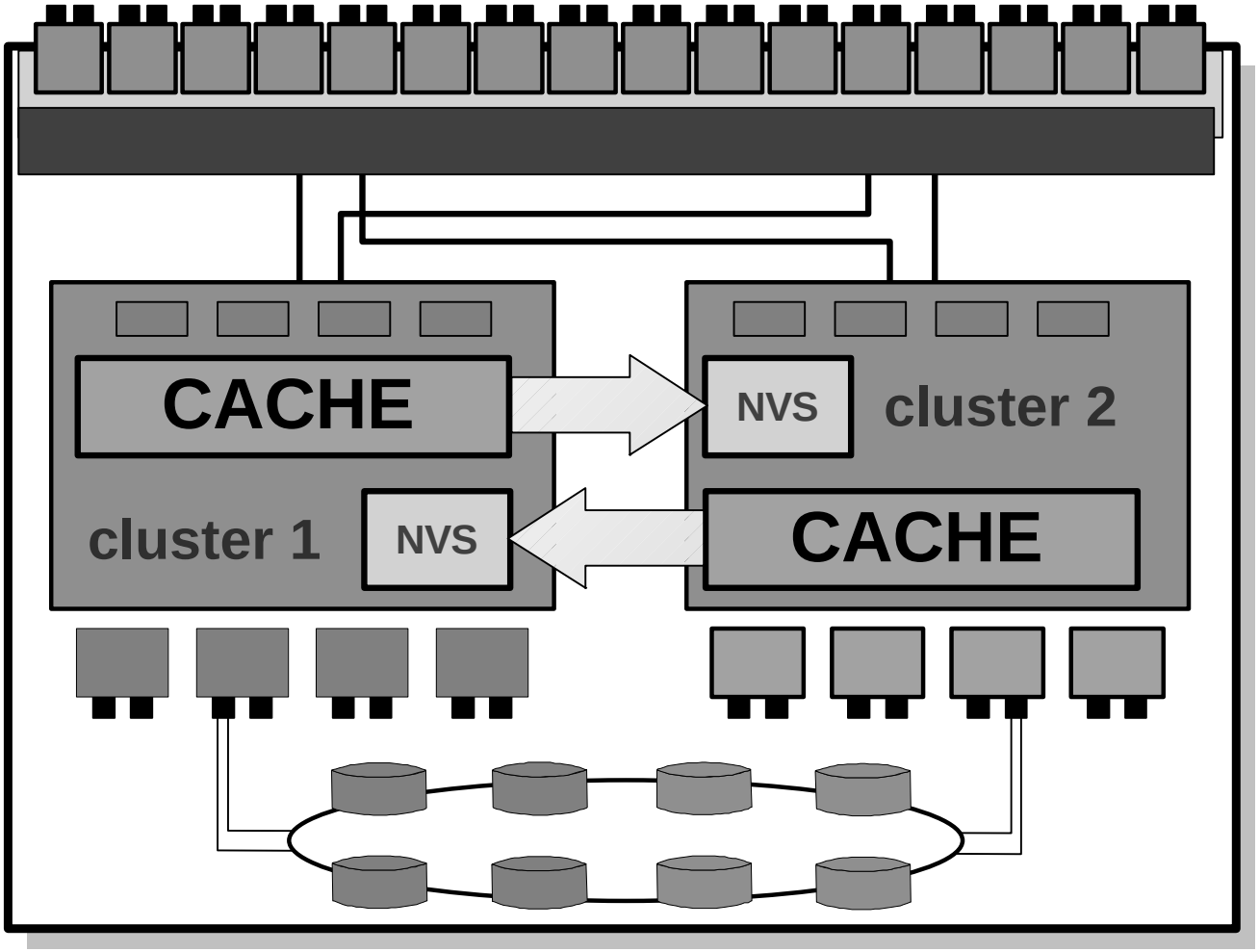
IBM Enterprise Server

ESS : The new standard

- Granularity
  - 420 GB to 11.2 TB
- High performances
  - Hardware et software optimised
- Maximum Availability
  - No *Single Point Of Failure/repair*
  - Solutions Remote Copy and FlashCopy

- Connectivity
  - ESCON, UltraSCSI, FICON, FC/AL
  - S/390, UNIX, base Intel, AS/400
- 3 years warranty

ESS - Internal Organisation



ESS : Componants and functions

Hardware componants

- Deux processeurs 4-way SMP
- 6 Go de cache / 384 Mo NVS
- Packaging et alimentations optimisés
- Supports 8-pack, 7133-020 / D40 et VSS
  - disques SSA 80 / 160 de 9 Go, 18 Go, 36 Go

Servers supportés : NT, UNIX , S/390

- Connexions serveurs :de 4 à 32 ports
  - SCSI, ESCON,
  - FICON, FC via gateway and later native
- non-RAID et RAID 5

No Single Point of Failure

- Composants redondants :
  alimentation, ventilation, adaptateurs, bus, processeurs (cluster)

No Single Point of Repair

Call home

Maintenance en service

Activation de microcode à chaud

-Alimentation 1 ou 3 phases

- Protection par batteries

Performance

- Gestion intelligente du cache
- PAV, *Multiple Allegiance*
- Plusieurs niveaux de protection du cache
- Architecture SSA 160

Flexibilité

- S/390, UNIX, base Intel, AS/400
- Gestion des Logical volumes (LUN)
- *Storage partitioning / sharing*
- 3380, 3390, flex-volumes

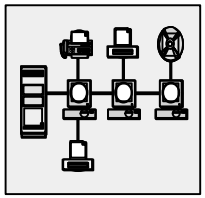
Fontions de copies

- XRC & PPRC
- FlashCopy & Concurrent Copy

Outils de management

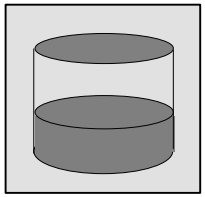
- StorWatch ESS Specialist
- StorWatch Expert
- Data Path Optimizer
- Service Alert Tools
- Data Migration Services

StorWatch ESS Expert



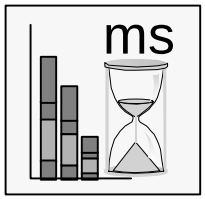
Inventory

- Synthèse des serveurs connectés par numéro de série, alias, logiciels installés, ...



Capacity

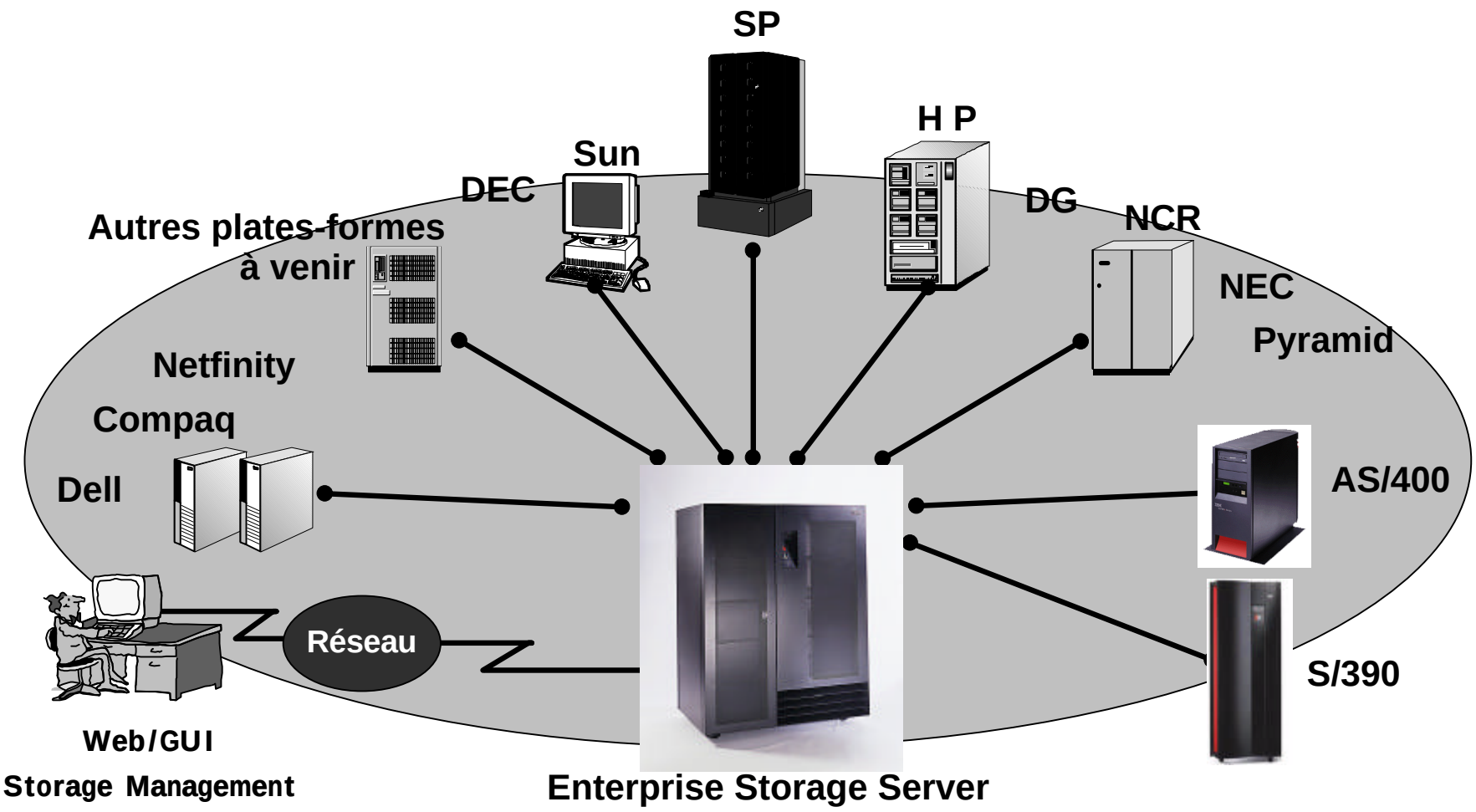
- Capacité de stockage affectée par serveur (utilisée et disponible)
- Graphique d'évolution du stockage / temps
- Synthèse des serveurs avec leur stockage
- Détails des volumes partagés entre serveurs multiples



Performances

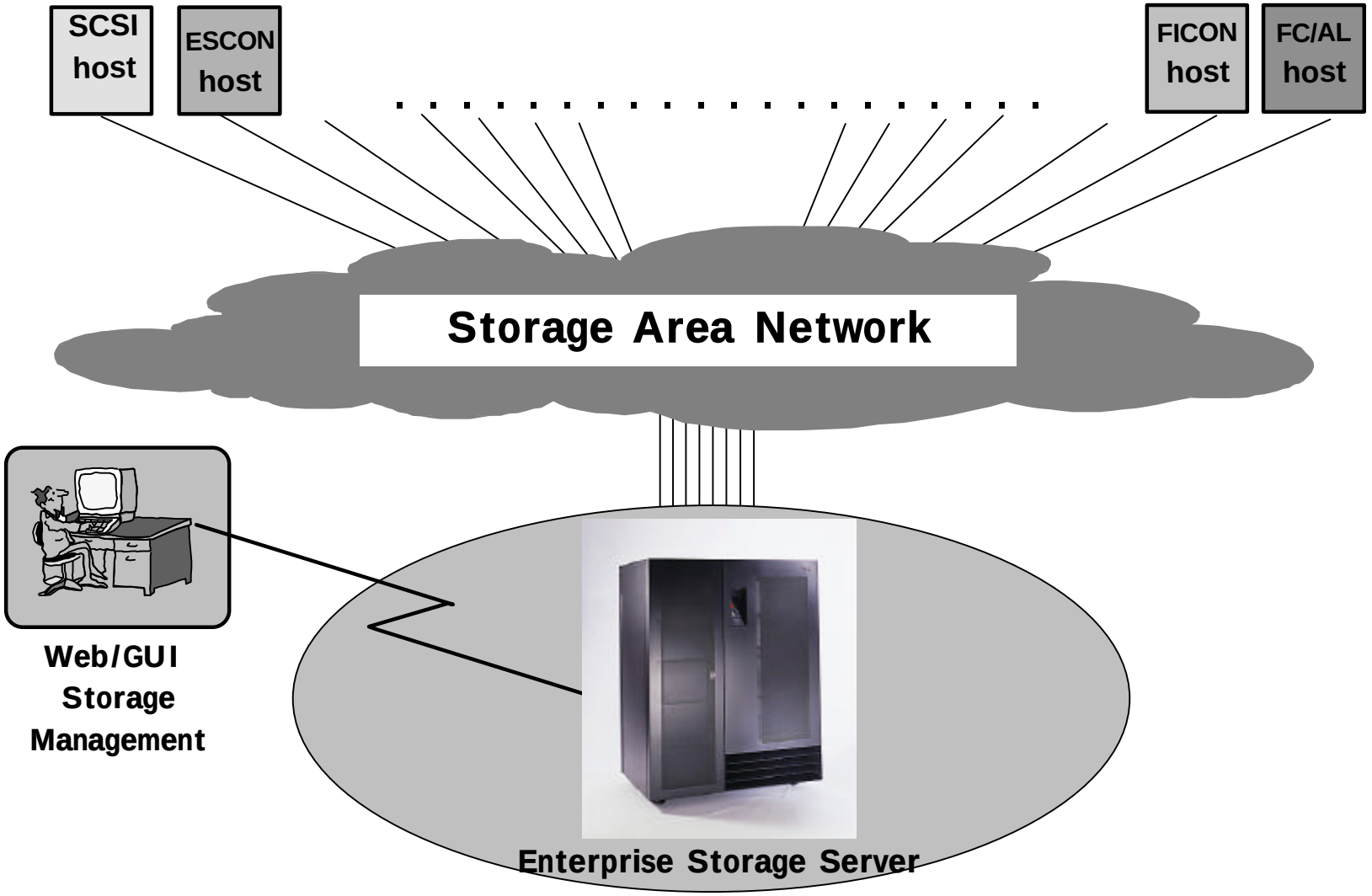
- Nombre des I/O
- Volumétrie et détail du *CPU time* de l'ESS
- Nombre d'octets transférés
- Temps de réponse en lecture / écriture
- Identification des *arrays,* adaptateurs et serveurs les plus sollicités
- Statistiques d'utilisation du cache

ESS - Wide Connectivity



IBM Enterprise Server

ESS : *SAN ready*



IBM Enterprise Server

Magstar and Magstar MP Application Support Matrix

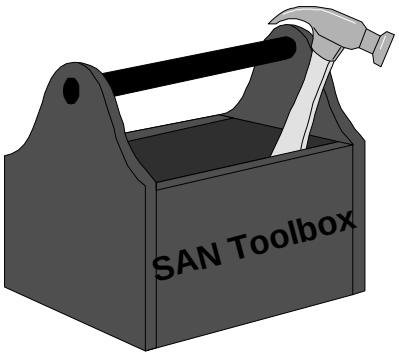
Tape Devices and Tape Libraries	Windows NT 4.0	IBM AIX 4.3	Sun Solaris
Host Base Adapters	QLogic	FC 6227	QLogic
Tivoli TSM 3.7	✓	✓	✓
Lagato Networker 5.5	✓	✓	✓
CA ARCserIT 6.6	✓	✓	
IBM SAN Data Gateway	✓	✓	✓
Magstar 3590/3494	✓	✓	✓
Magstar MP 3570/75	✓	✓	✓
IBM SAN DG Router	✓	✓	✓
3502 DLT Autochgr	✓	✓	✓

Current support: www.ibm.com/fcswitch

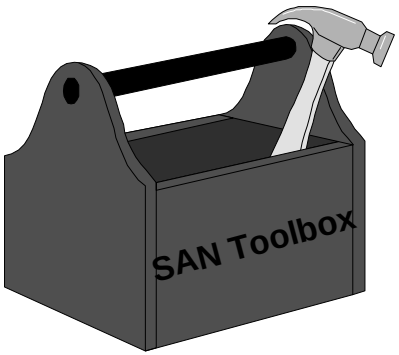
Wide range of certified options available now

SAN Solution Building Blocks

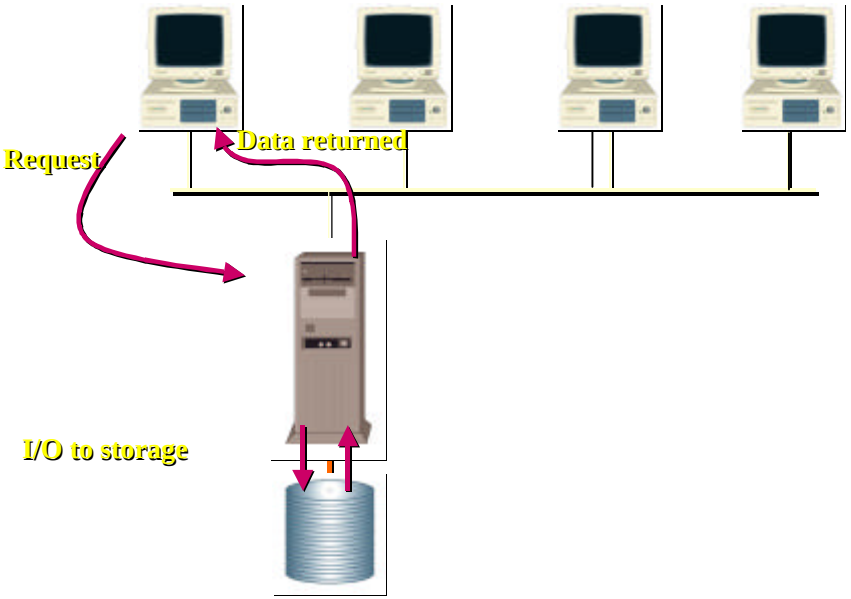
- **Servers**
  - IBM: Netfinity, RS/6000, AS/400, NUMA-Q, S/390
  - non-IBM: Sun, HP, Compaq, Dell, etc.
- **SAN Fabric Components**
  - Switches, Hubs, Gateways/Routers
- **Storage**
  - IBM: FCSS, 7133, ESS, Magstar, DLT
  - StorageWorks, STK Silo, etc.
- **Software**
  - Tivoli SAN Modules
  - non-IBM: Veritas, Legato, CA, etc.
- **Services**
  - IBM Global Services
  - Business Partners, Systems Integrators



- **Enterprise Application Management**
  - Tivoli Global Enterprise Manager, Service Desk, etc.
- **SAN Solutions - Data Management**
  - Tivoli Storage Manager (LAN-free & Server-free Backup)
- **SAN Exploitation - Resource Management**
  - Tivoli SAN Removable Media Manager (Tape Pooling)
  - Tivoli SAN File Sharing
  - Tivoli SAN Disk Manager (Disk Pooling)
  - Tivoli SAN Storage Automation for Disk Allocation
  - Tivoli SAN Data Management (Data Sharing)
- **SAN Connectivity - Network Management**
  - StorWatch SAN FC Switch Specialist
  - SAN Data Gateway Specialist
  - Tivoli NetView SAN Extensions and SAN Manager
- **SAN Hardware - Element Management**
  - StorWatch ESS Expert, ESS Specialist
  - Tivoli Management Agents (TMA)



LAN-Based Sharing Challenges



Typical LAN Environment:

- Server owns storage
- Server controls I/O
- Client waits ...
- Networks slow, high overhead
- I/O Bus fast, low overhead

LAN-based File Sharing Benefits

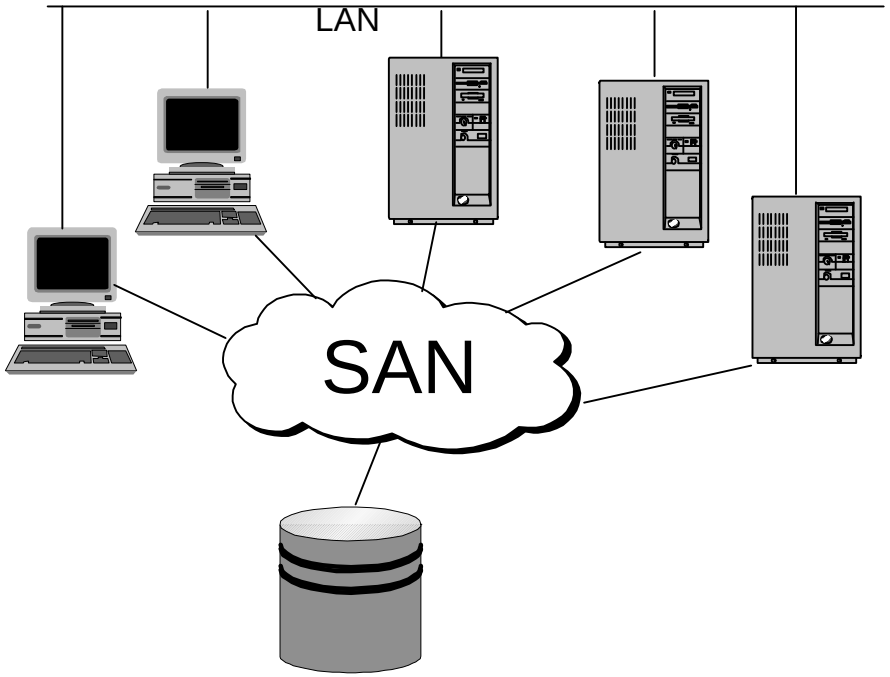
- Transparent sharing
- Heterogeneous
- Inexpensive implementation

Challenges of LAN-based File Sharing:

- **Bandwidth limitations**
- **Protocol overhead**
- **Server bottlenecks**

Tivoli SANergy File Sharing

- ✓ **Enables SAN-based file sharing between heterogeneous, open systems**
  NT, W2K, Mac, UNIX
- ✓ **Automated fail-over (new capability)**



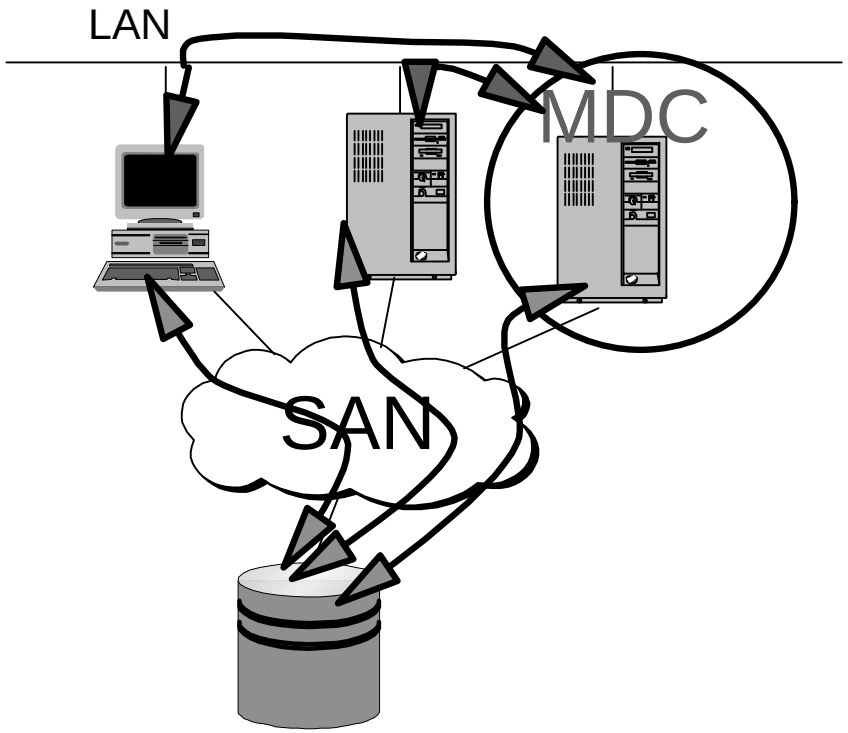
SANergy delivers:

File sharing of a LAN
- transparent file sharing
- fail-over
- file & byte-range locking
- heterogeneous support

Speed of a SAN
- high bandwidth
- direct media access
- lightweight protocol

IBM Enterprise Server

Tivoli SANergy File Sharing- how does it work?



MDC- Meta-Data Controller

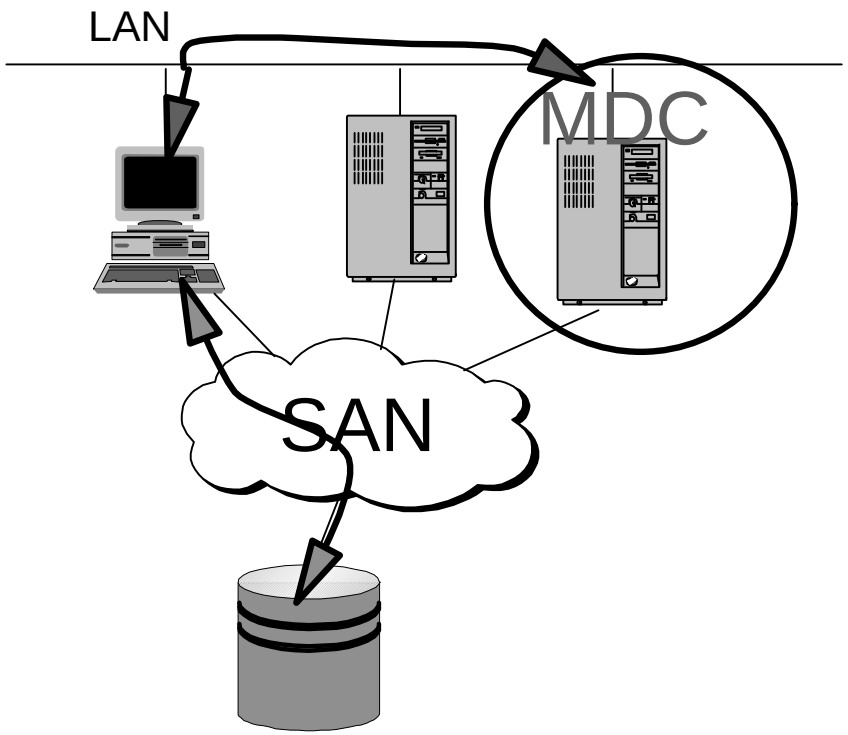
- Volumes "assigned" to a MDC
- Mounts volumes as a local drive
- Manages locks, authorization, extents
- Uses its native fs & disk format & shares the volume like a server

Client

- Mounts the volume
- Sends **meta-data** requests over LAN to MDC
- Reads **data** directly over SAN

➡ Separation of Meta-Data & Data

Tivoli SANergy File Sharing



Environment:

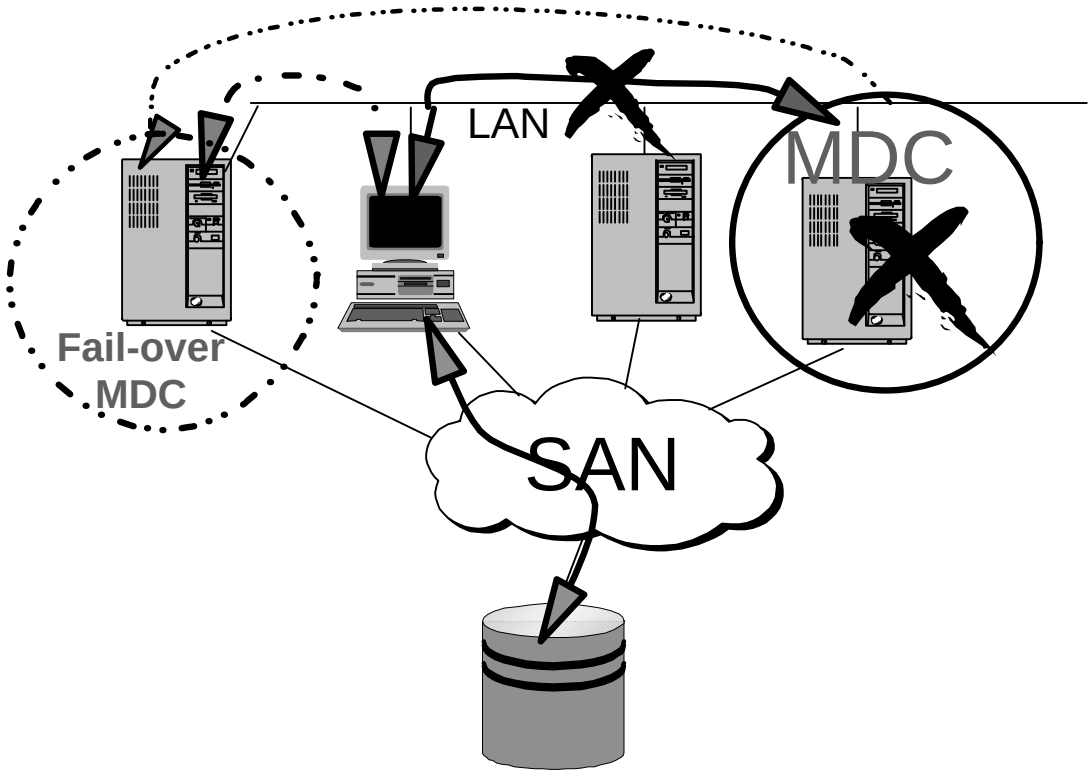
MDC- Meta-Data Controller

- NT4 NTFS
- Sun UFS
- W2K NTFS (2Q00)
- Sun SAMfs (2Q00)

Clients

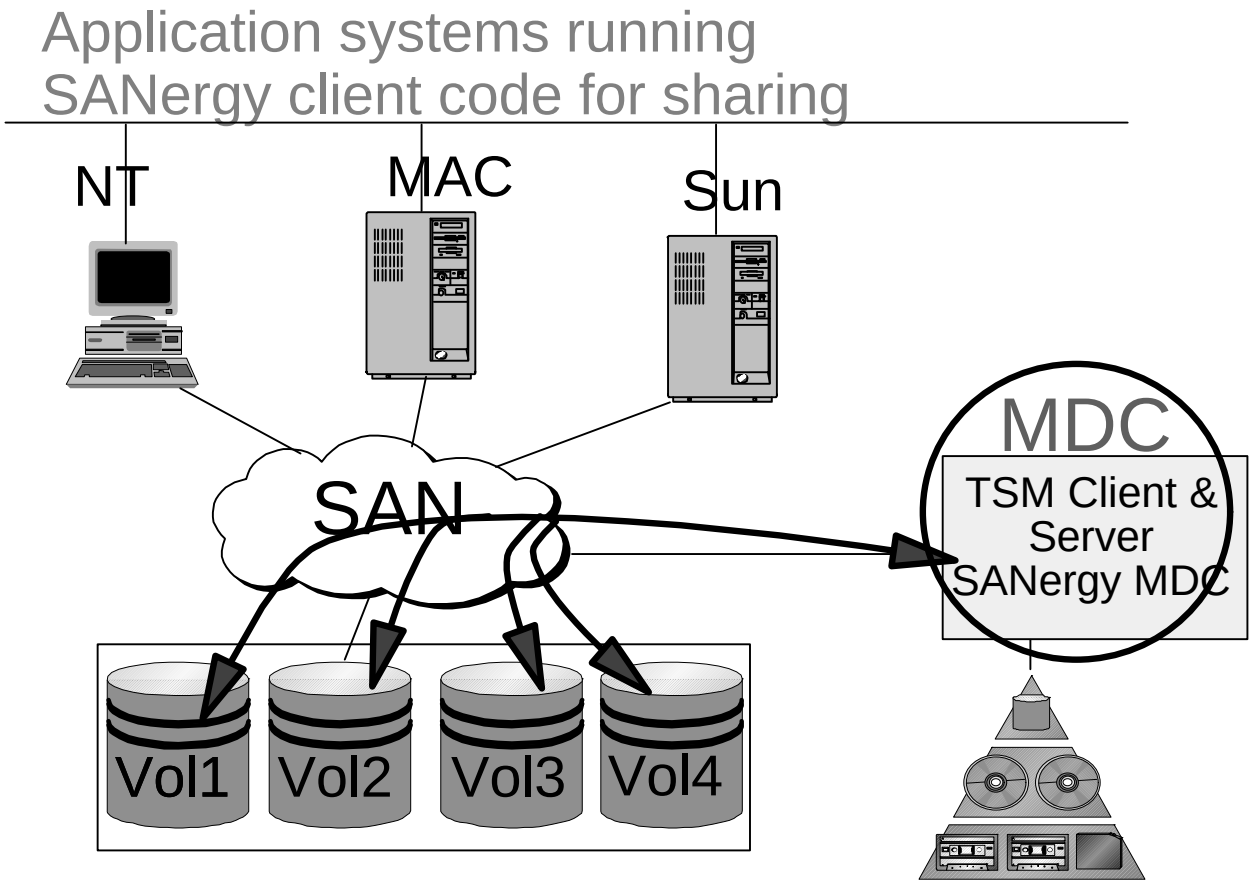
- NT4 (Intel & Alpha)
- Mac 8.1 or later
- SGI IRIX
- Solaris 2.6 or later
- AIX 4.3.2
- W2k (2Q00)

Tivoli SANergy File Sharing High Availability Option



1. Primary MDC dies
2. Stand-by MDC SW recognizes
   - Takes over volumes
   - Activates shares
3. SANergy client receives remap
   - Transparent to applications

LAN-free, Application Server-Free Backup



Tivoli Storage Manager + Tivoli SANergy

IBM Enterprise Server

Tivoli SANergy - Background

- 1998: SANergyFS V1
  - NT MDC
  - NT, Mac, UNIX clients sharing files
- 1999: SANergyFS V2.0
  - Sun MDC
  - SANergyXA, NT MDC Fail-over
- 3000 systems running SANergy

IBM Enterprise SAN Roadmap

1999

Netfinity, RS6000, S/390 fibre-enabled servers

SAN Data Gateway family

7 port FC Storage Hub

8 and 16 port SAN FC Switch

ESCON/FICON Director

Enterprise Storage Server and FC RAID Storage Server

StorWatch Specialists/Experts

Tivoli SAN Lan-Free Data Movement - Tape Resource Sharing

3494/TSM Tape Sharing

Nefinity HA 8-way Cluster

Netfinity ServerProven Solutions Program

"Tivoli Ready" Certification Competency Centers

IGS Certification Lab

IGS Fibre Transport Services

IGS Design, Planning and Implementation Capabilities

1st Half 2000

SAN Data Gateway enhancements

SAN Managed Hub

32 port SAN FC Switch

ESS native fibre (FCP) attach

SDG-attach of VTS

LTO Drive native fibre (FCP) attach

Tivoli Decision Support for Storage Management Analysis (Perf/Health)

SAN Extensions for Netview

Tivoli SANergy File Sharing

ESS Copy Functions

2nd Half 2000

S/390 FICON Director

SAN - WAN - SAN capability

ESS native fibre (FICON) attach

Magstar native fibre (FCP) attach w/ Drive and FICON w/ A60 Controller)

LTO Library native fibre (FCP) attach

Tivoli Storage Manager LAN-Free Data Movement - TDP for SQL API (NT)

Tivoli Removable Media Manager

Tivoli SAN Manager

Tivoli SAN Disk Manager

Tivoli SAN Storage Automation for Disk Allocation

Tivoli Decision Support for Storage Resource Management and SAN Management (Reporting)

2001

AS400 fibre-enabled servers

Tivoli SAN Server-less Backup

Tivoli SAN LAN-Free Client Data Transfer (NT Client and Server, Oracle API for Solaris, AIX)

Tivoli SAN Data Management (Data Sharing)

Tivoli SAN Policy Automation

2 Gb Fibre Channel

"and the beat goes on..."

IBM Enterprise Server

IBM SAN Interoperability Labs

IBM Gaithersburg - August, 1999
IBM Mainz - February, 2000
IBM Montpellier - 1H2000
IBM Tokyo - 2H2000

IBM Enterprise Server

SAN-Related Web Site URLs

IBM Enterprise SAN	www.ibm.com/san
IBM SAN Services & Interoperability Lab	www.as.ibm.com/asus/san2.html www.ibm.com/storage/ibmsan/sanlab.htm
IBM and Tivoli SAN Software Solutions	www.ibm.com/storage/ibmsan/products/sansoftware.htm www.tivoli.com/products/index/san/index.html
IBM Netfinity Server SAN Solutions	www.ibm.com/pc/ww/netfinity/san
IBM S/390 Server	www.ibm.com/s390/san
IBM RS/6000 Server	www.ibm.com/rs6000/hardware/san
IBM AS/400 Server	www.ibm.com/as400/periph/san.htm
IBM SAN Fabric	www.storage.ibm.com/ibmsan/products/sanfabric.htm
IBM Storage	www.storage.ibm.com/ibmsan/products/sanstorage.htm
IBM SAN Redbook	www.ibm.com/redbooks/abstracts/sg245470.html