A Technical Crash Course on SANs

HP World, Chicago, 2001

Instructor: Allison Banister

Cambridge Computer Services, Inc.

Introduction

Course Overview

About Cambridge Computer Services

Class Agenda

Goals of This Class

About Cambridge Computer Services

- Over 10 years in the field of storage systems and storage management technologies.
  - Sales
  - Integration
  - Consulting
  - Training
- Large percentage of our business is subcontracted training and integration for industry giants (EMC, Compaq, Legato, etc.)
- Headquartered in Boston, MA
  - Clients all over the world

Other SAN-Related Activities

- Hired to write O'Reilly book on storage area networks and network attached storage.
  - Watch for it! Fall, 2001.
- Participating in Storage Network Industry Association SAN certification program.
  - Classes in Boston and on site all over the world.
- Lectures at major conferences
  - HP World, Usenix LISA 2001, PC Expo SAN Summit, Disaster Recovery 2001, Contingency and Planning Management 2001.
- Private consulting and integration services.

Web Sites

- www.cambridgecomputer.com
  - Main corporate site
  - White papers and other interesting tidbits
- www.san101.com
  - Launches this winter. Free educational materials on the latest developments in SAN technology.
- www.sanconnection.com
  - Launches this winter as a SAN industry portal. In the meantime, it contains a listing of discounted SAN equipment.

Class Agenda

Chapter 1	Introduction to Storage Area Networks
Chapter 2	SAN Infrastructure: SCSI & Fibre Channel
Chapter 3	Basics of SAN Partitioning
Chapter 4	SAN Backup Technologies
Chapter 5	Overview of SAN Management
Appendix	Alternatives to Fibre Channel

Goals of This Class

- Understand basic storage problems today and how SANs address them
- Understand how common technologies (SCSI, Ethernet, RAID) relate to SANs
- Learn about fibre channel
- Understand the basic components of a SAN
- Understand the architectures of SAN applications such as disk storage and tape backup
- Gain perspective on the future of SANs

Chapter 1

An Introduction to Storage Area Networks

Defining the Storage Problem

How SANs Help

Introduction to Building a SAN

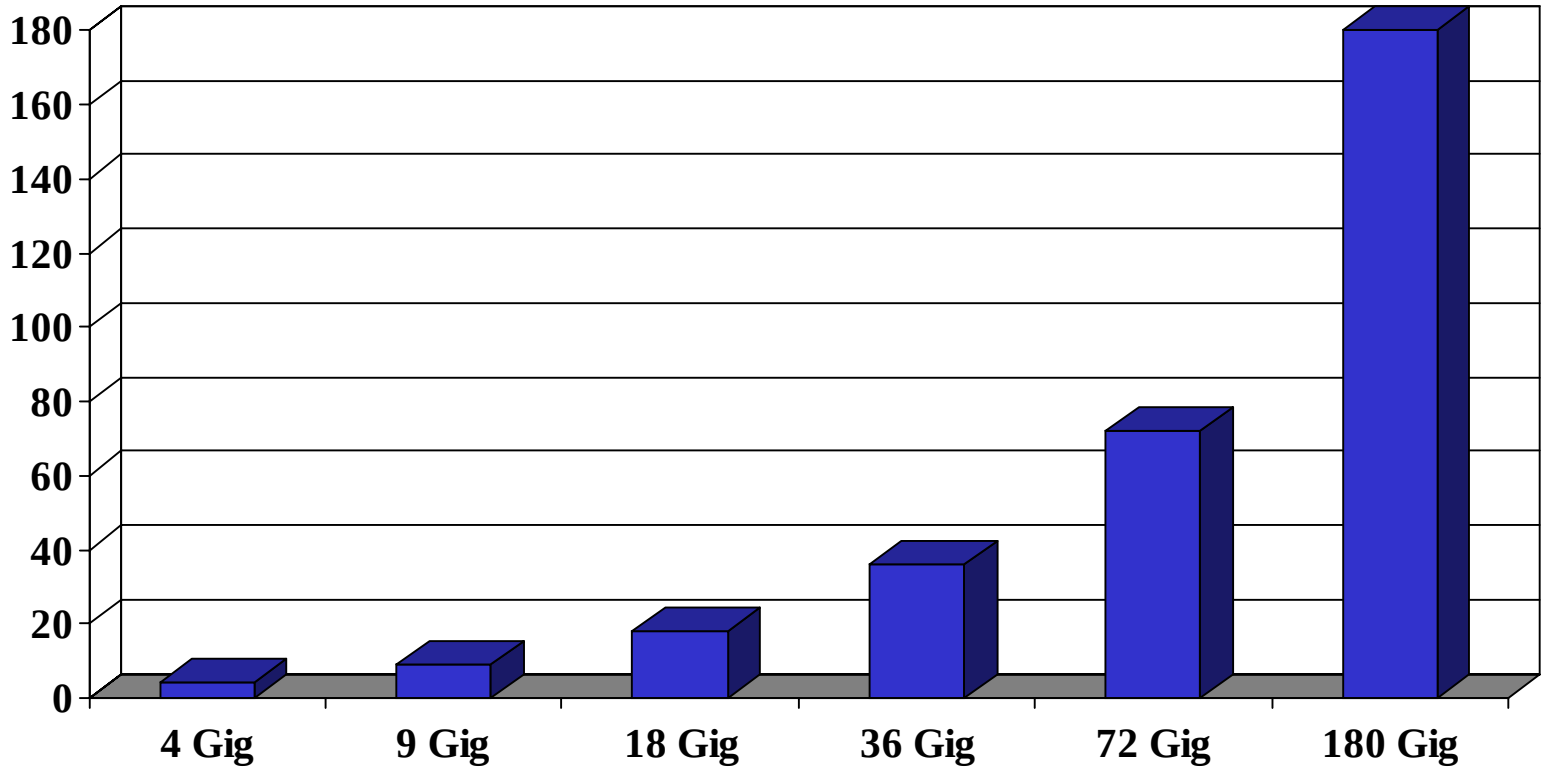
The Storage Problem

Disk storage is growing faster than the infrastructure for storage management.

Expanding Disk Storage

- Capacities double on average every 18 months.
- Drives keep getting bigger and cheaper.
- Drives are usually aggregated in a disk array.
- 75 GB drives cost only about +/- $1000.00.
- One Terabyte JBOD costs as little as $20K.
- Nature hates a vacuum.

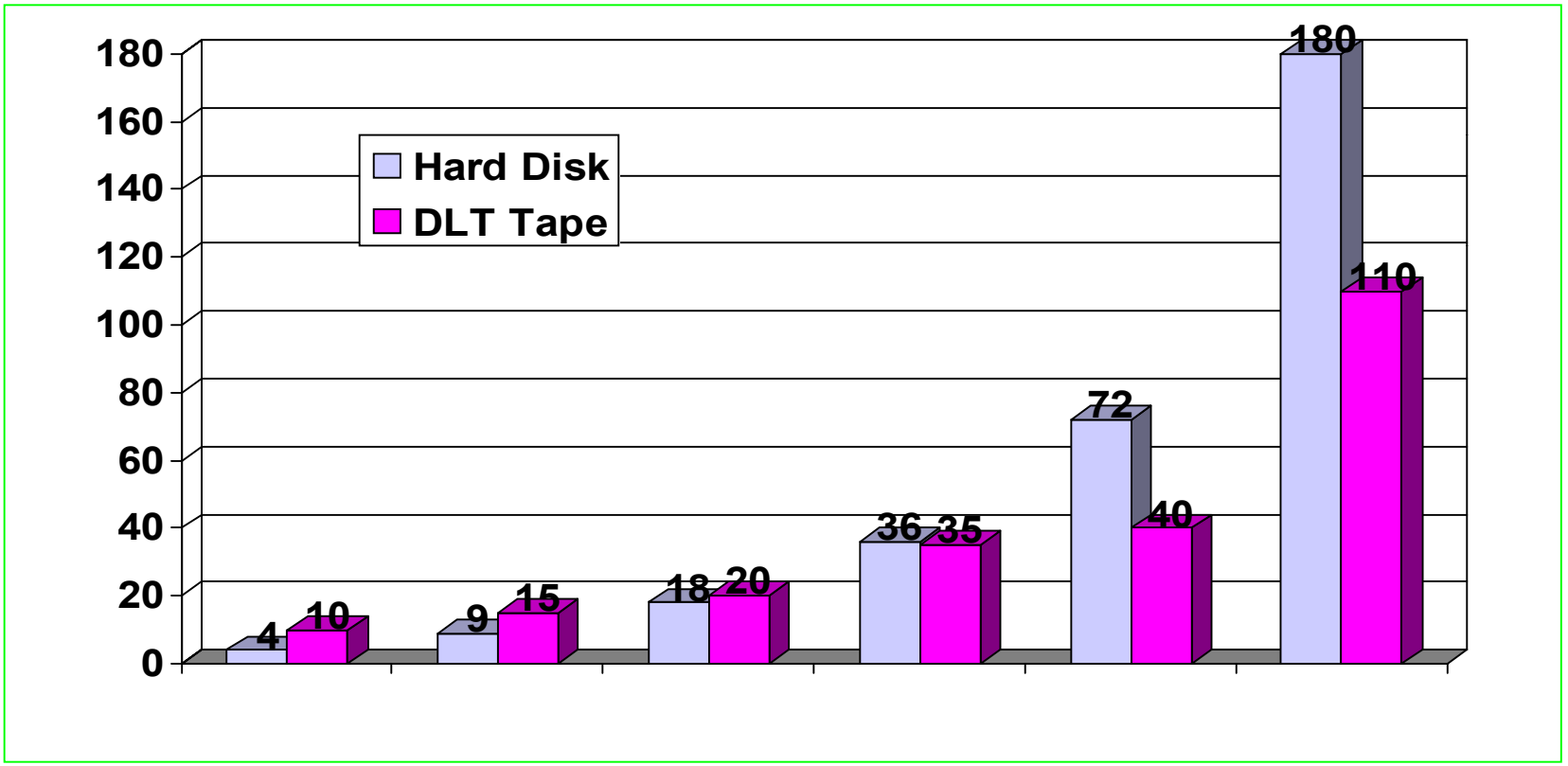
Recent Hard Drive Growth



What do we mean by infrastructure?

- Network Throughput
- I/O processing abilities of computers
  - Operating systems capabilities
  - I/O Bus Capabilities
- Backup software capabilities
- Tape Drive Capabilities

Disk vs. DLT Tape Drive Growth



Tape drives used to be bigger than hard drives. This changed with the 36 gig drive.

Network Throughput Issues

- If Disk Doubles, total throughput must double too
- Network bandwidth is also growing at a much slower rate
  - 100 Base-T is NOT 10x faster than 10 Base-T on point-to-point transfers
  - Point-to-point Gigabit Ethernet is not all it is advertised to be:
    - Nominally faster than 100mb on NT and 2000
    - A lot faster – but not blazing – on UNIX
  - Theoretical bandwidth limited by IOPS

Gigabit Ethernet NT Throughput

Windows NT lacks network throughput:

NT to NT via TCP/IP on 100Mbps:

25 Mbps, or only 25% utilization = 3 MB/s

NT to NT via TCP/IP on 1,000Mbps:

29 Mbps, or only 3% utilization - not significantly better

Source: Network World, 7/5/99 - West Virginia University

Storage Driving Factors

- The day is still only 24 hours long.
- But other forces have virtually reduced it:
  - 24x7 e-commerce
  - Global economies
  - Longer business days
- Corporate dependence on data has increased, and downtime costs more.
- Tolerance for downtime has decreased.
- Companies have a narrower “backup window.”

Some Numbers to Memorize

3 MB/SEC	~ 10 GB/hr (10.8GB)
6 MB/SEC	20GB/hr. Maximum native speed of DLT-8000 and AIT-2.
6 MB/SEC	A very generous amount of data to push over the network to a single NT server. Assumes special hardware.

- At 6MB/SEC:
  - 1 TB will take 50 hours to backup/restore.
  - 100 GB will take 5 hours to backup/restore.

Storage Management Challenge

- Storage on demand: Allocate disk where and when you need it.
- Better way for backups and restores: Full restores are too time-consuming to be practical. Must never lose data.
  - Live mirrors in separate storage system, ideally in separate building.
  - Disk separate from servers
  - Server clustering
- Outsource storage.

SAN-Enabled Solutions

Disk Sharing	• Dynamic Disk Storage Allocation
Fault Tolerance	• Server Clustering and Application Fail-Over • Data Vaulting • Storage Service Provisioning
Tape Sharing	• Shared Tape a/k/a “LAN Free Backup” • Serverless Backup

Understanding SANs

SNIA Definition

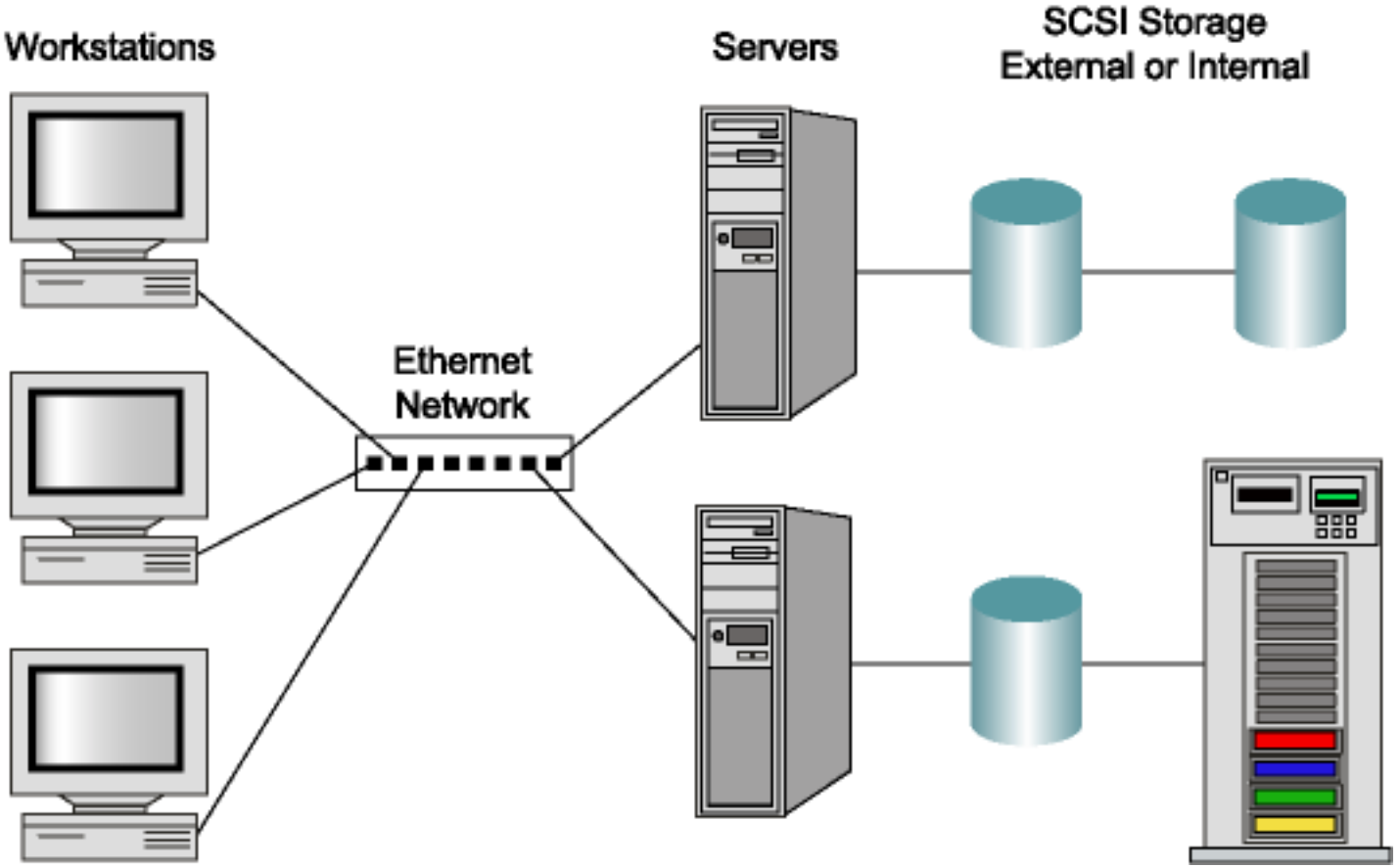
A network whose primary purpose is the transfer of data between computer systems and storage elements and among storage elements. Abbreviated SAN. A SAN consists of a communication infrastructure, which provides physical connections, and a management layer, which organizes the connections, storage elements, and computer systems so that data transfer is secure and robust."

Source: Storage Networking Industry Association
http://www.snia.org

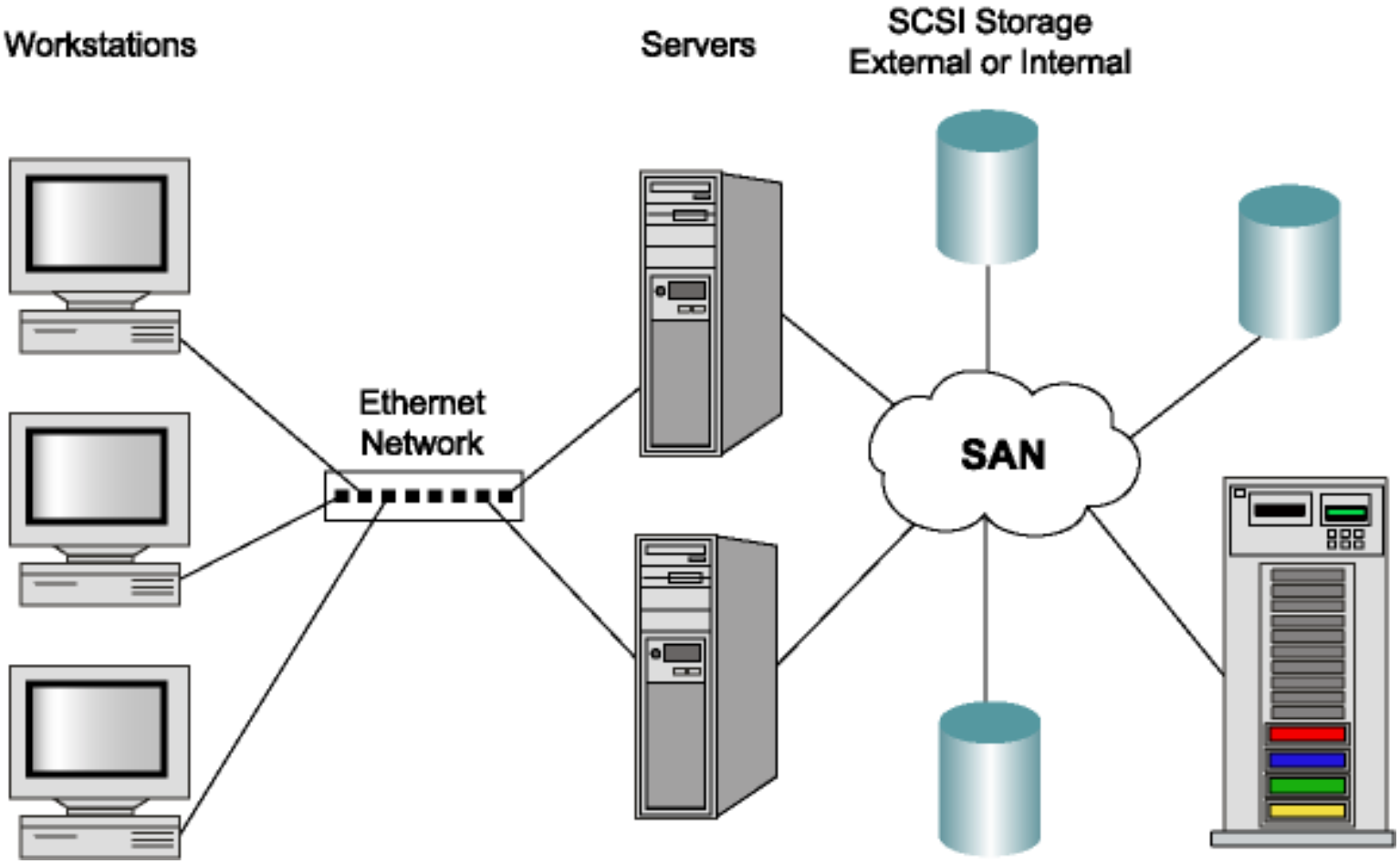
A Storage Area Network:

- Consists of a network of storage devices
- Provides unique benefits, including speed, scalability and fault tolerance
- Generally uses fibre channel today (although other protocols are supported)
- Adds flexibility by:
  - Separating infrastructure (cabling, etc.) from storage
  - Separating storage from host computers

Storage Infrastructure Today

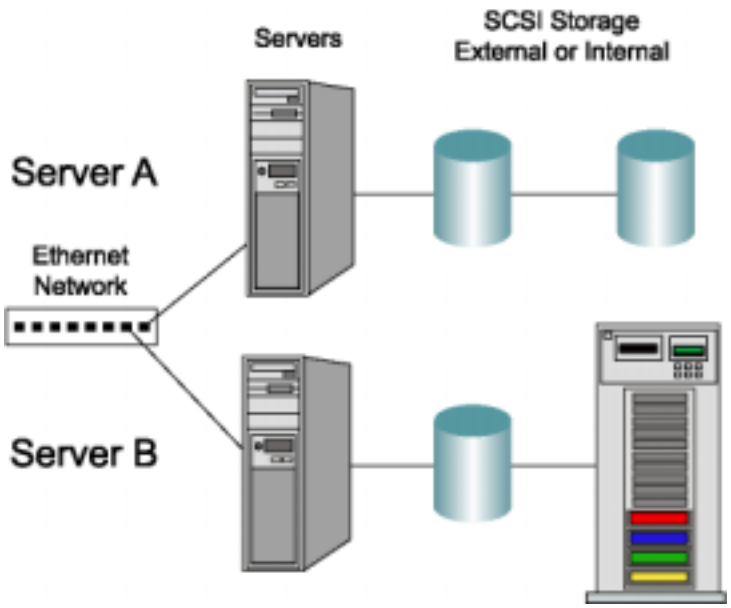


The Same Components as a SAN

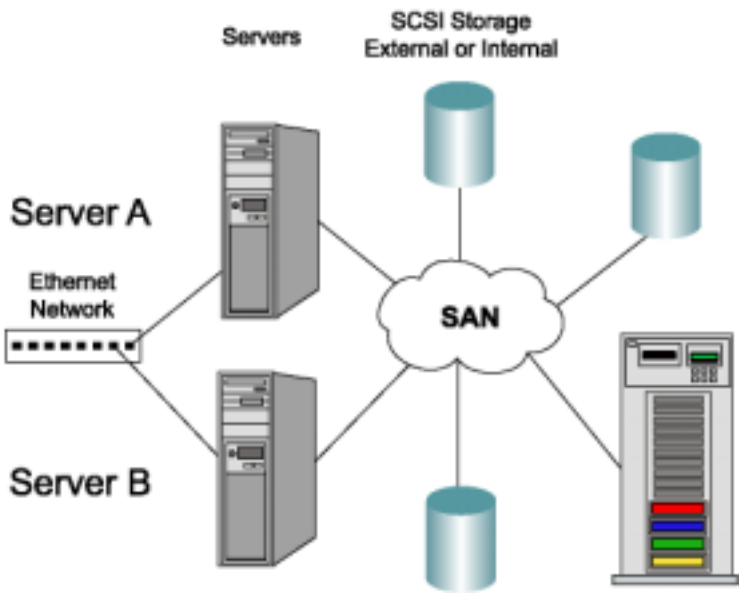


Traditional LAN vs. SAN

Before

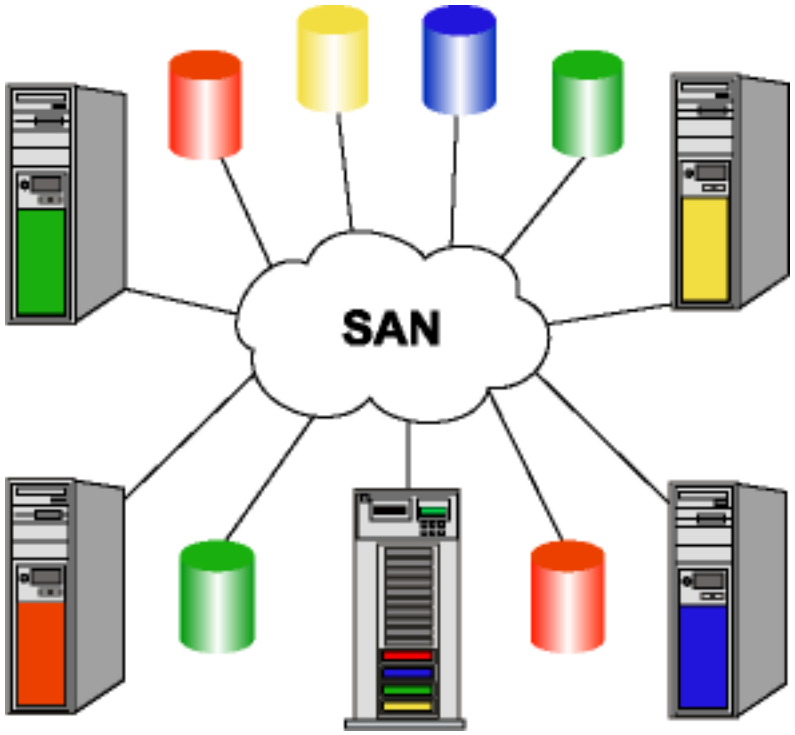


After



A SAN simply connects multiple hosts to a common set of storage devices.

SAN: A Practical Definition

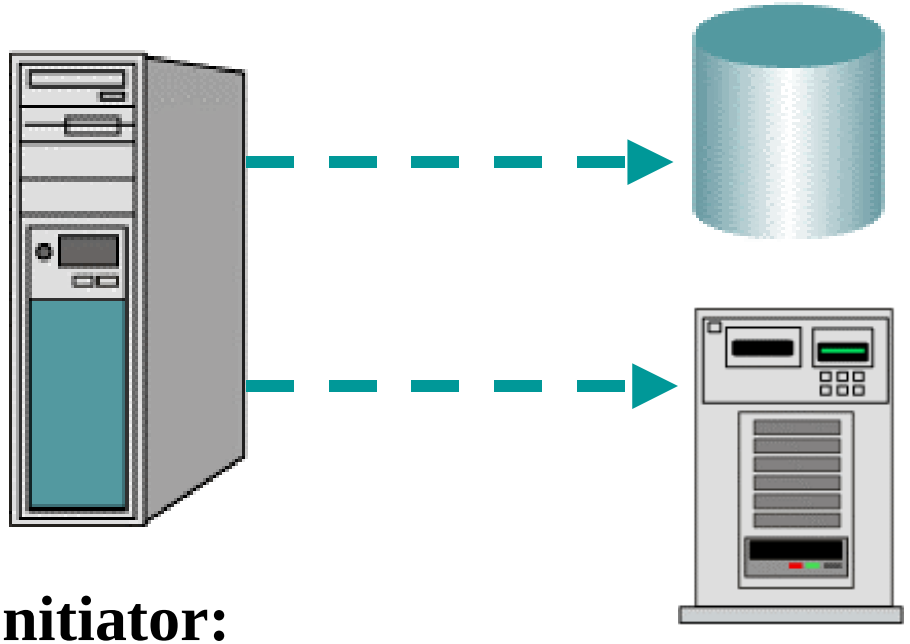


"A bunch of SCSI hosts and storage devices all plugged in together."

Our Definition of a SAN

A storage channel consisting of at least one storage device (target) with at least two host computers (initiators).

Hosts and Initiators



Initiator:
A device that makes requests.

Target:
A device that responds to requests.

SCSI? What about Fibre Channel?

- Most current SANs use Fibre Channel
- In most current implementations, ***fibre channel is simply a form of SCSI***.

- NOTE: Fibre channel is not the only data transport suitable for SANs. We will talk about Gigabit Ethernet and other alternatives toward the end of the lecture.

Chapter 2

SAN Infrastructure

SCSI & Fibre Channel

Fibre Channel Advantages

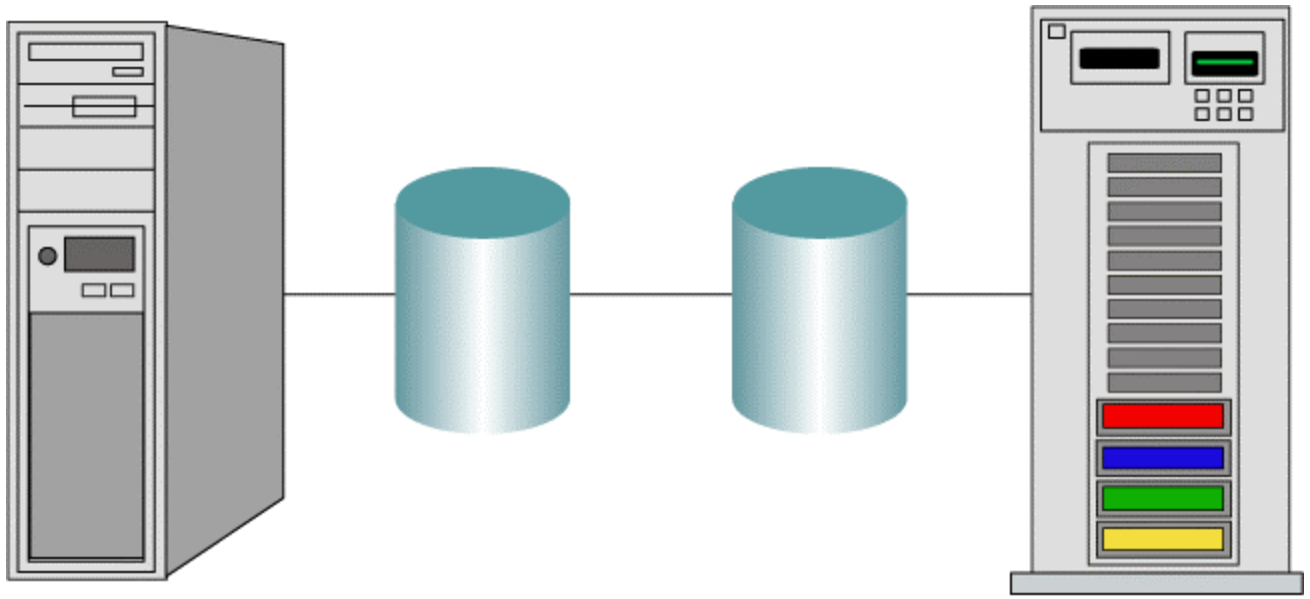
Fibre Channel Topologies

Ingredients of a SAN

SCSI

SCSI Review

- Most popular method of connecting storage devices
- SCSI devices are connected on a daisy-chained bus



SCSI Review

Signal Type	Cable Length
Single-Ended	6 meters
Differential (HVD)	25 meters
Low Voltage Differential (LVD)	12 meters

Do not mix differential with single-ended or LVD
If you mix single-ended and LVD the whole bus drops to single-ended.

SCSI Addressing

- SCSI uses IDs to distinguish devices on the same bus.
- LUNs (Logical Unit Numbers) are a subset of SCSI IDs
  - SCSI ID = Street Address
  - LUN = Apartment Number

SCSI Type	# IDs	ID Range	LUNs per ID
Narrow SCSI (8 Bit)	8 ID's	0-7	8 (64 total)
Wide SCSI (16 Bit)	16 ID's	0-15	Unlimited*

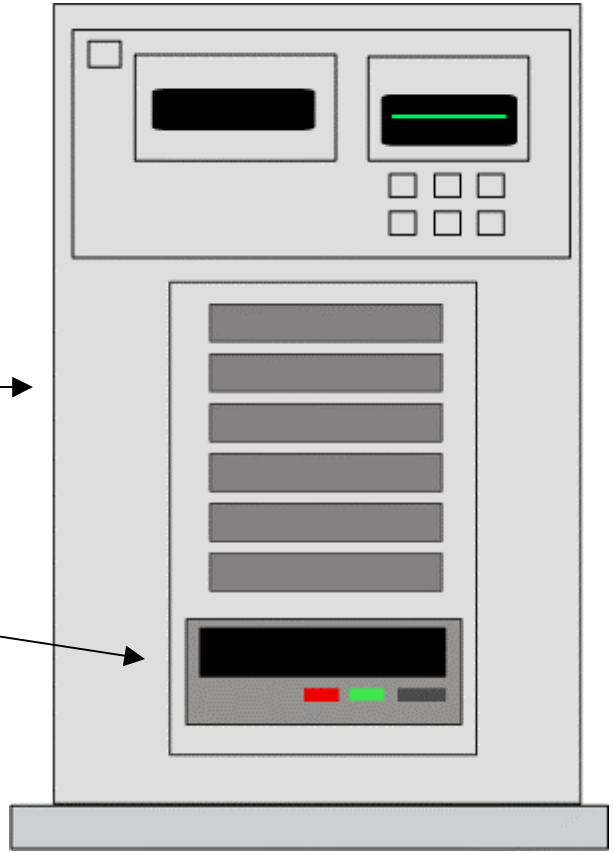
**The actual number of LUNs depends on operating system and driver support.*

Example: SCSI LUNs in a Tape Loader

Tape Library
SCSI ID = 3

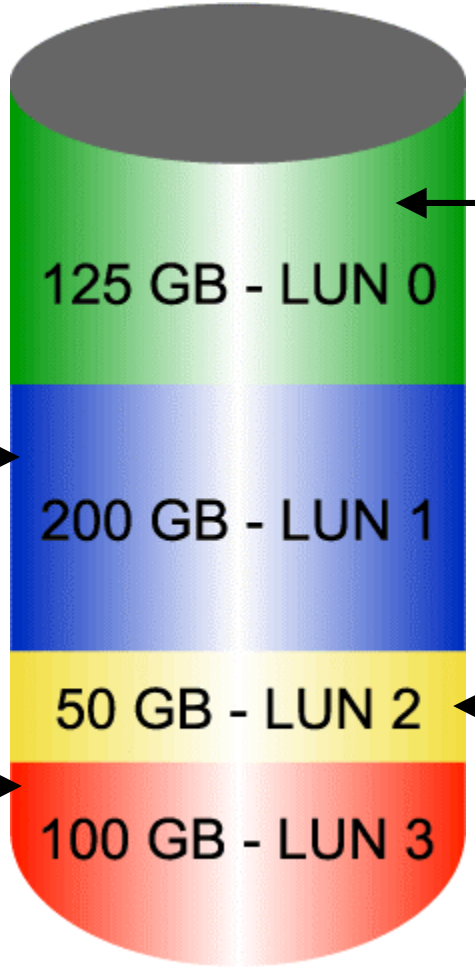
Robotics = ID3, LUN 0

Tape Drive = ID 3, LUN 1



Example: SCSI LUNs in a RAID Array

RAID System
SCSI ID = 1



125 GB Partition: ID 1, LUN 0

200 GB Partition: ID 1, LUN 1

50 GB Partition: ID 1, LUN 2

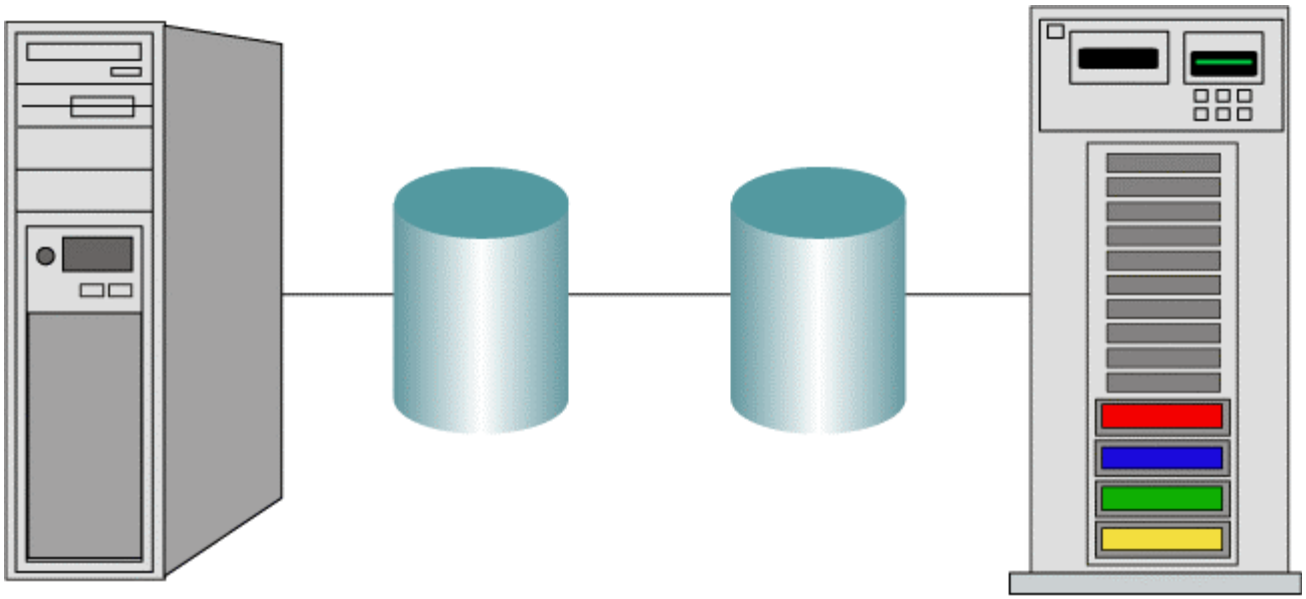
100 GB Partition: ID 1, LUN 3

Review of SCSI Addressing

- The operating system uses three things to identify SCSI devices:
  - SCSI Channel
  - SCSI ID
  - SCSI LUN
- Most of the time the ID is the differentiator
  - Many SCSI systems only have one channel
  - Most of the time the LUN is set to zero
- But: all three factors make up the complete address

Problems with Traditional SCSI

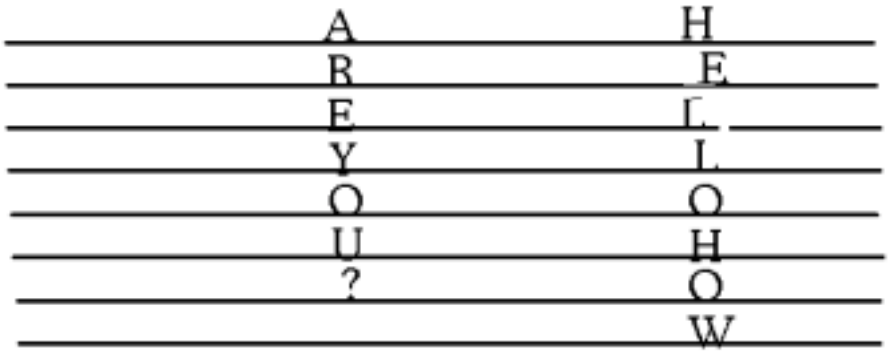
- Standard SCSI is connected with a connection technology described as a “Parallel bus.”

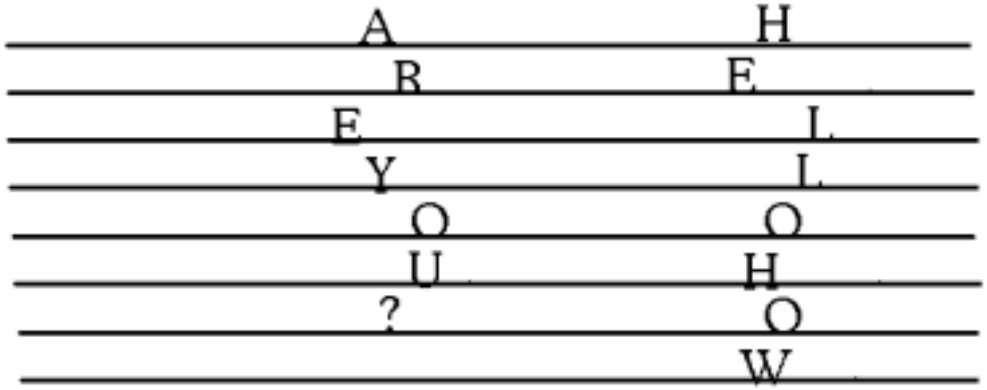


Problems with Traditional SCSI

- Strict cabling requirements
  - Parallel bus technology requires stringent cabling specifications.
  - Violation of specs can lead to system unreliability.
- Relationship between speed and cabling
  - As parallel SCSI gets faster and faster, the cable rules become more and more restrictive.
- Example: LVD Cabling
  - Maximum cable length of 12 meters--but not with all 15 devices.

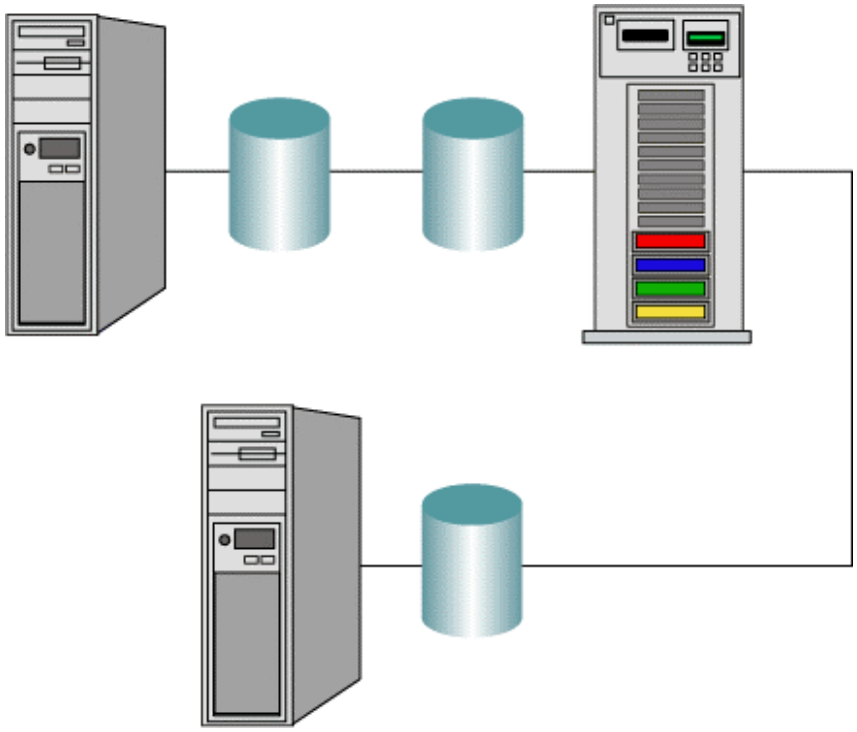
Parallel SCSI Skew





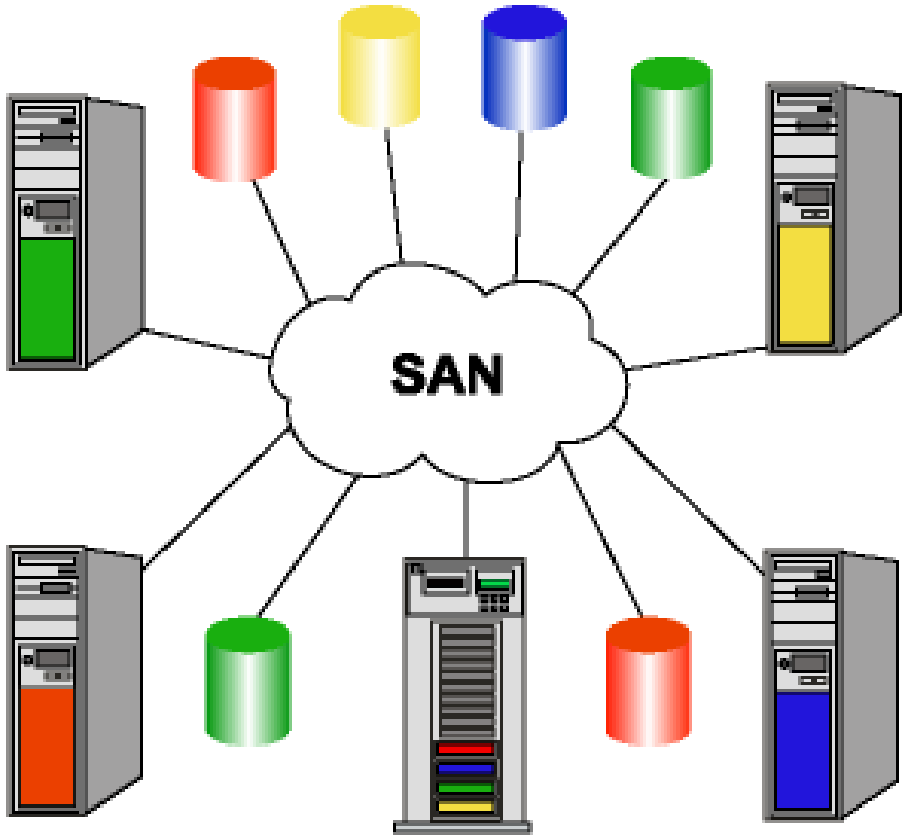
Parallel SCSI & SANs

- You can build a simple SAN using parallel SCSI.
- But parallel SCSI is not well-suited for SANs.



Complex SANs

Complex SANs require more sophisticated technology.

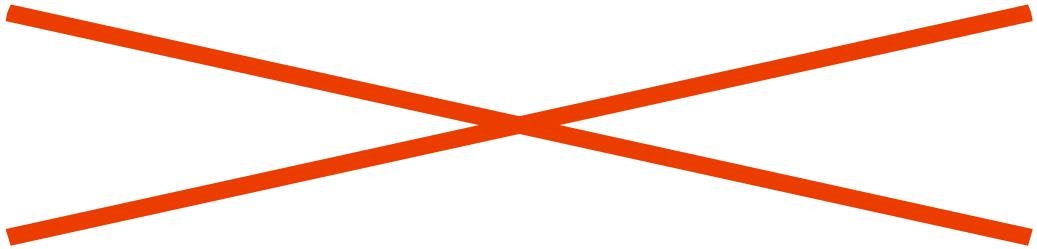


The Good News

- SCSI can be separated into
  - Data transport
  - SCSI command protocol

- The SCSI protocol can be run on alternative data transports.
  - Similar to the way that TCP/IP and other network protocols can run over Ethernet, Token Ring, FDDI, etc.

Fibre Channel

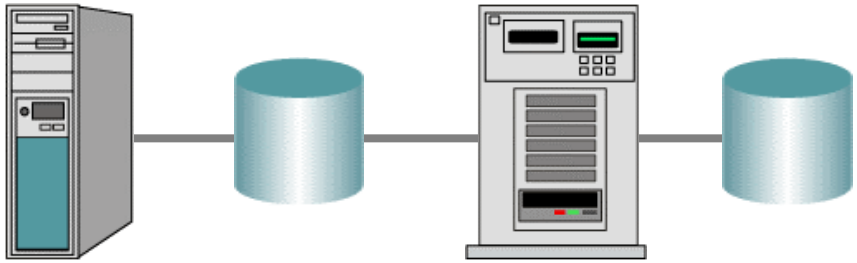


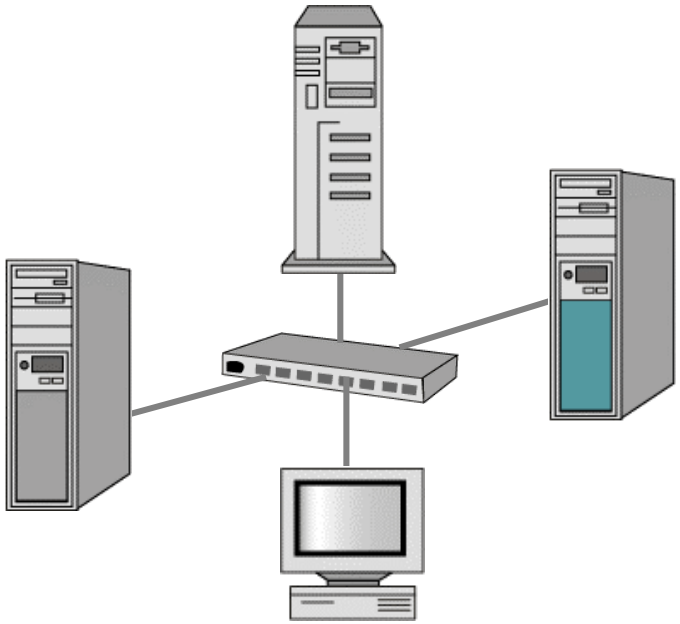
Fibre Channel Facts

- Fibre channel:
  - Has been a standard since 1994 .
  - Represents the evolution of SCSI.
  - Combines elements of both networks and channels into a system highly favorable for storage.
  - Supports multiple familiar protocols.
  - Is not new technology--certainly not “bleeding edge.”
  - Is becoming more affordable and commonplace.
- Software applications that enable powerful fibre channel storage systems are now becoming available

Protocol Types

Channel





Network

Channels

- Closed, structured, predictable environment
- Intended for transferring large blocks of data
- All devices are discovered in advance
- Changes require reconfiguration in software/ configuration table
- Master-slave environment
- Error-free delivery paramount; time secondary
- Low overhead: hardware handles most processing
- Example: SCSI

Networks

- Open, unstructured, unpredictable environment
- Intended for transferring many small blocks of data
- Devices not discovered in advance
- Any device can communicate with any other device
- Changes in configuration are handled “on the fly”
- Peer-to-peer environment
- Time is paramount; error-free delivery secondary
- Higher overhead: more software involvement
- Example: TCP/IP

Serial SCSI

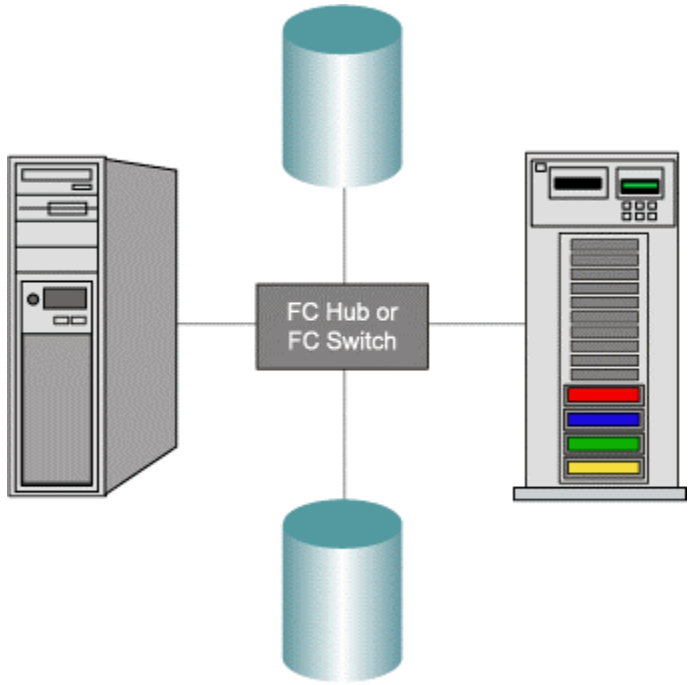
- Industry anticipated problems of parallel SCSI bus.
- ANSI released SCSI-3 standard in 1994
  - Standard for serial SCSI in a star topology
  - Also known as “SCSI-3 Serial”
- Fibre channel uses Fibre Channel Protocol (FCP), a version of the SCSI-3 serial standard

Fibre Channel: Serial SCSI

- 100% SCSI compatible.
- Is a communication protocol, not a physical standard.
  - That means that it is a very broad and flexible set of specifications.
- Uses industry-standard copper and optical cabling.
  - Fibre channel cabling is identical to Gigabit Ethernet.
  - Although the cabling is different than SCSI, it carries the same SCSI protocol.

Fibre Channel = SCSI in a Star

- Ethernet evolved from a bus architecture (10Base-2) to a star architecture (10Base-T & 100Base-T)
- With fibre channel, SCSI has followed suit

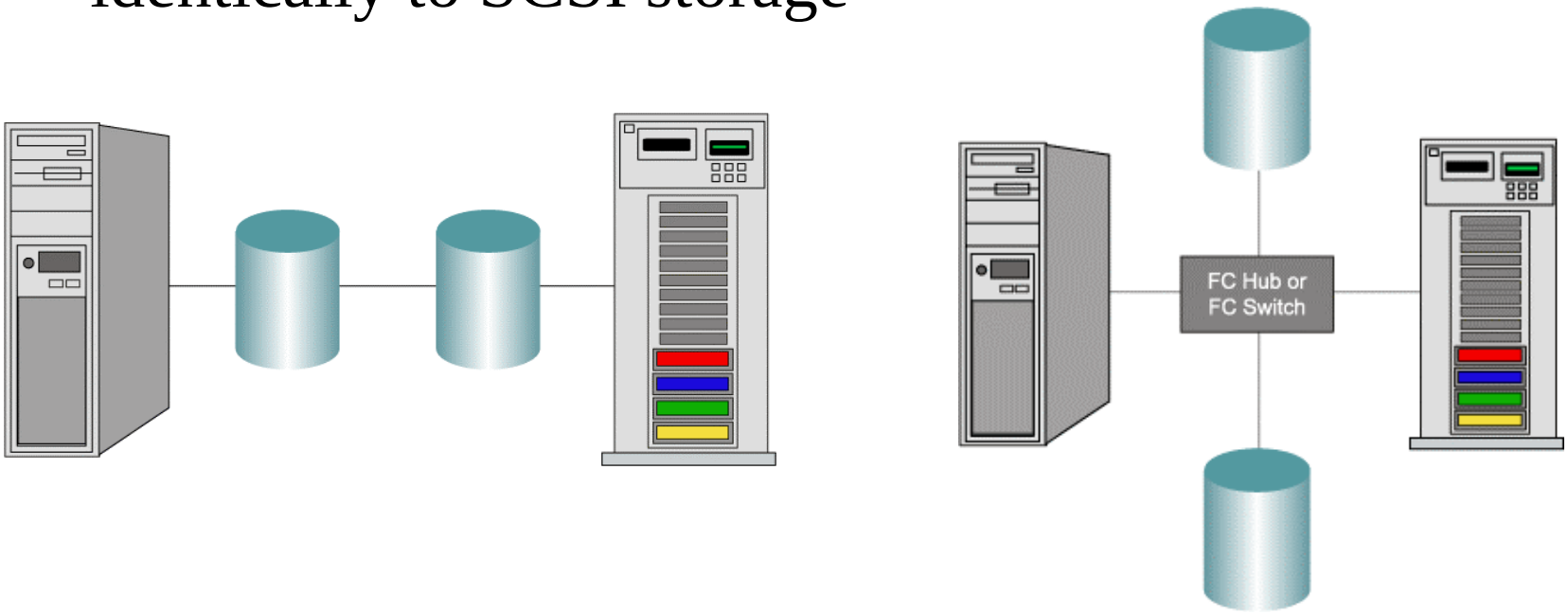


Benefits of Star Topology

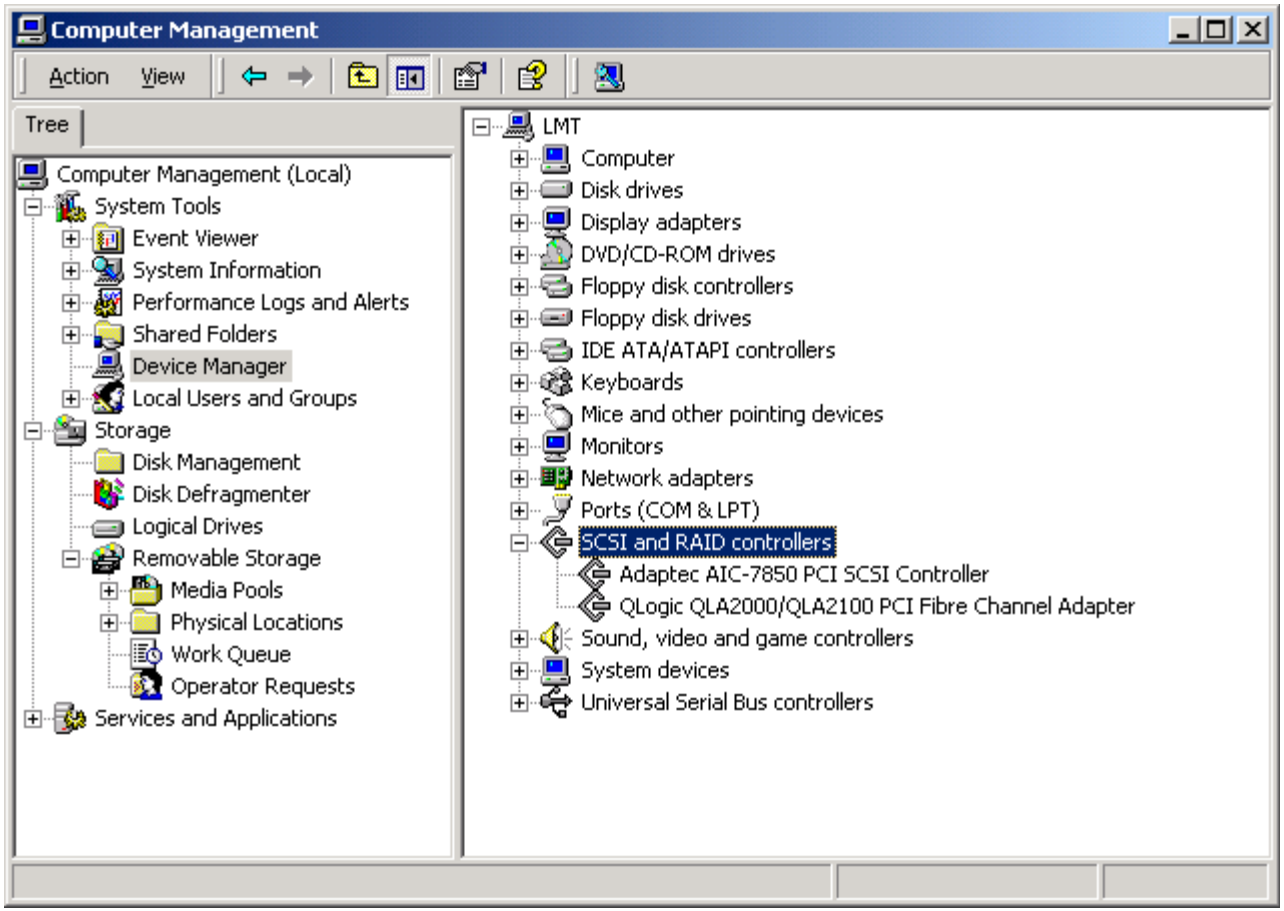
- Cable fault isolation
  - Bad cable or connection does not bring down entire system.
  - Easy to troubleshoot a failed connection.
- Cable organization
  - Cables less bulky and easier to organize.
  - Each machine simply wires into a hub or switch.
- Hubs and switches enable network management
  - Can tie in to SNMP systems like OpenView, TNG, etc.

Fibre Channel: Host's Perspective

- Host computers can not tell the difference between Fibre channel and SCSI.
- Fibre channel storage is installed and configured identically to SCSI storage



Fibre Channel HBA Installation



Fibre Channel Supported Protocols

- Network protocols
  - TCP/IP
- Channel protocols
  - SCSI
  - HIPPI
  - VI

Fibre Channel v. Parallel SCSI

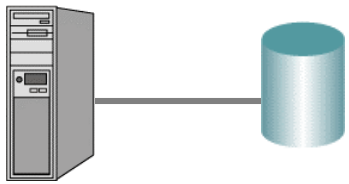
- More fault tolerant
  - Cable fault isolation
  - Simpler cables and no terminators
- Cables are easier to run, more reliable and capable of going longer distances.
- Can be managed centrally
- Has a roadmap for greater bandwidth
- Uses standards for switching

Parallel SCSI Advantages vs. FC

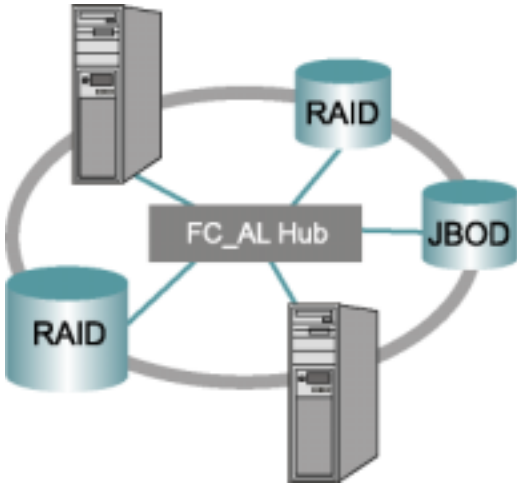
- Cost
  - Hard drives cost about the same.
  - Fibre Channel cabinets are far more expensive than SCSI.
- SCSI will always be used to connect devices inside a storage cabinet.
  - Hard drives in a RAID disk system might be SCSI.
  - Tape drives in a library are almost definitely SCSI.
- NOTE: It is rarely advantageous to have fibre channel disk drives inside a storage array.

Fibre Channel Topologies

Fibre Channel Topologies

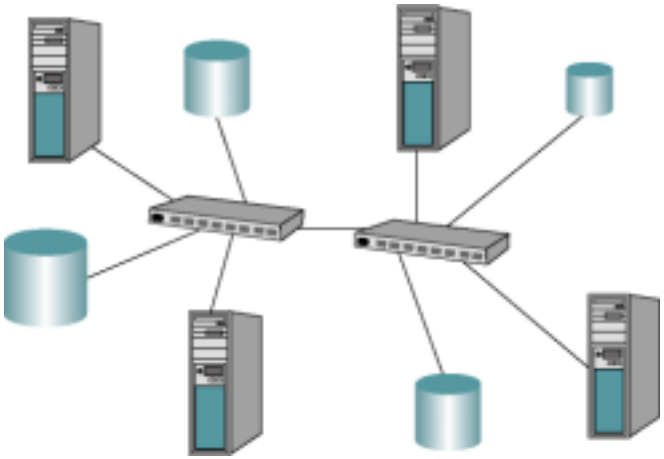


Point-to-Point



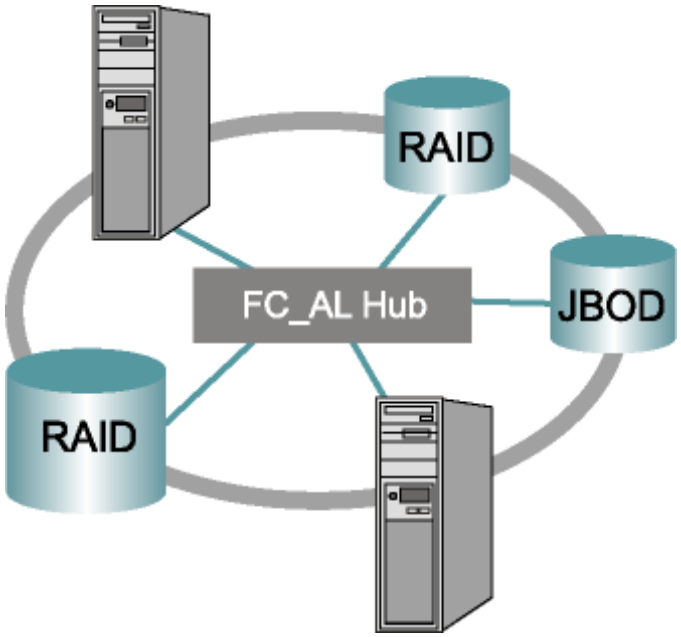
Fibre Channel Arbitrated Loop (FC_AL)

Switched Fabric



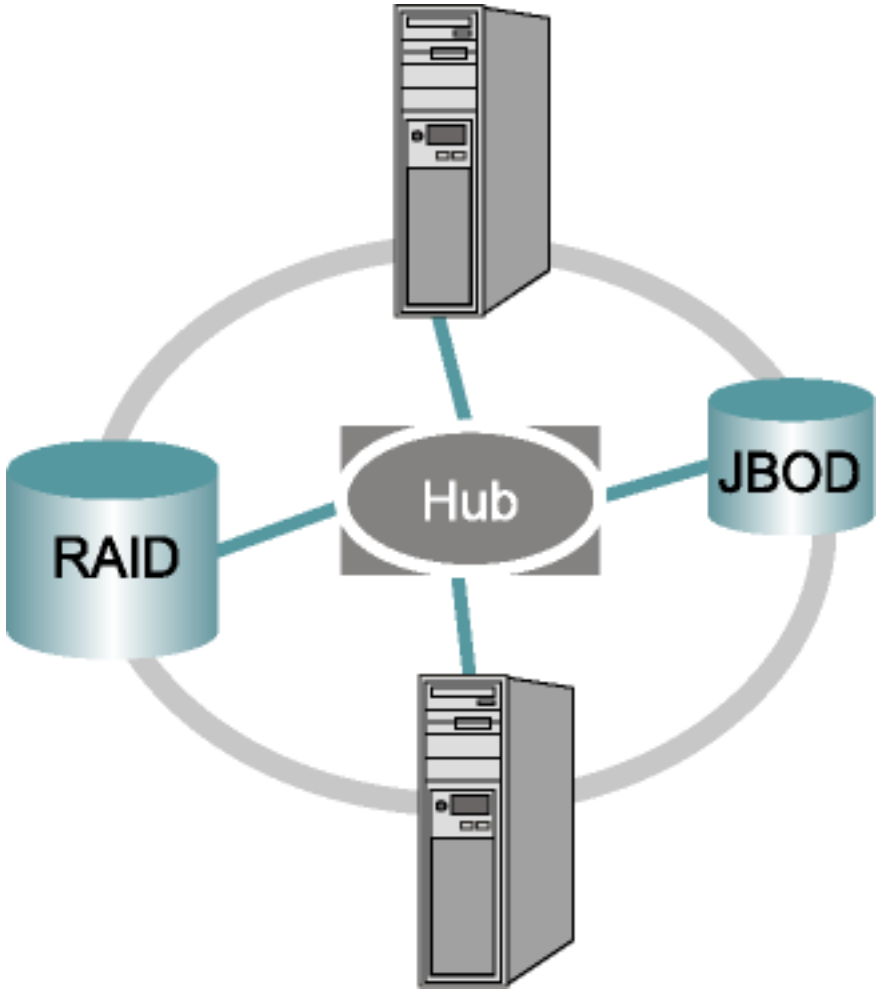
Fibre Channel Arbitrated Loop (FC_AL)

- Supports up to 127 devices.
- Loop dynamically assigns IDs via "loop initialization."
- Only two devices on a loop can communicate at once.
- Bandwidth is usually shared.
- Loops can be connected to fabrics using an uplink port.
- Used by fibre channel hard drives.



Arbitrated Loop Hubs

- Create a “virtual loop” inside the hub
- Share bandwidth between connections
- Provide a “loop bypass circuit” if a connection goes down.
- Are not “intelligent” members of the loop.



FC_AL Advantages

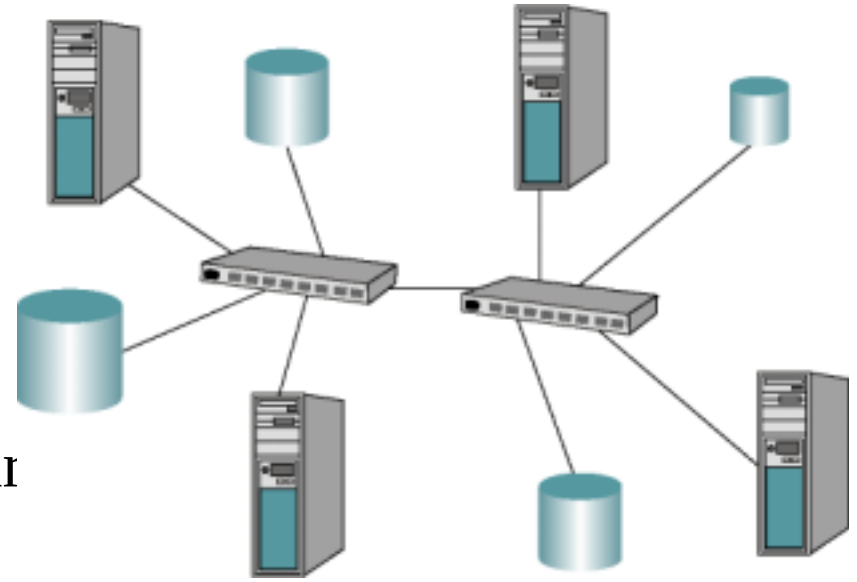
- Simple to implement
- Relatively fast and reliable
  - Ideal for small SANs
  - Well suited for private storage connections (e.g. intracabinet connections; connecting disk cabinets to a single server)
- Relatively inexpensive
  - FC_AL hub: $250 per port
  - FC_AL HBA: $1400 per port

FC_AL Disadvantages

- Supports a comparatively small number of devices
- Device additions and subtractions require re-initialization of entire loop
- Bandwidth limitations
  - Only 1 “conversation” per loop
  - Problem is alleviated but not solved by “loop switches” and “zoned hubs”
- Distance degrades performance

Fibre Channel Switched Fabric

- Supports over 16 million devices.
- Requires a fabric switch for "intelligence."
- Supplies full duplex aggregated bandwidth.
- Multiple "conversations" can occur at once with full, dedicated bandwidth.



Uses a log-in/log-out system rather than initialization/arbitration.
Offers intelligent, centralized management of fabric.

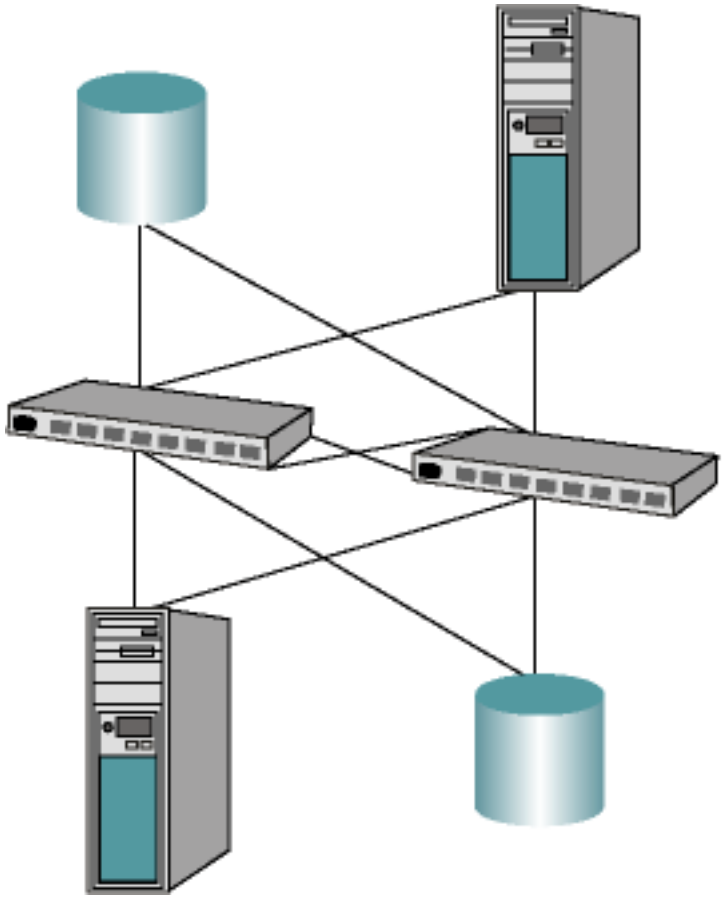
Switched Fabric Advantages

- Supports many more devices
- Performance does not degrade due to:
  - Long cable segments
  - Increased numbers of devices
- Scalability
- Lower overhead
- Designed to facilitate redundancy
- Facilitates centralized management

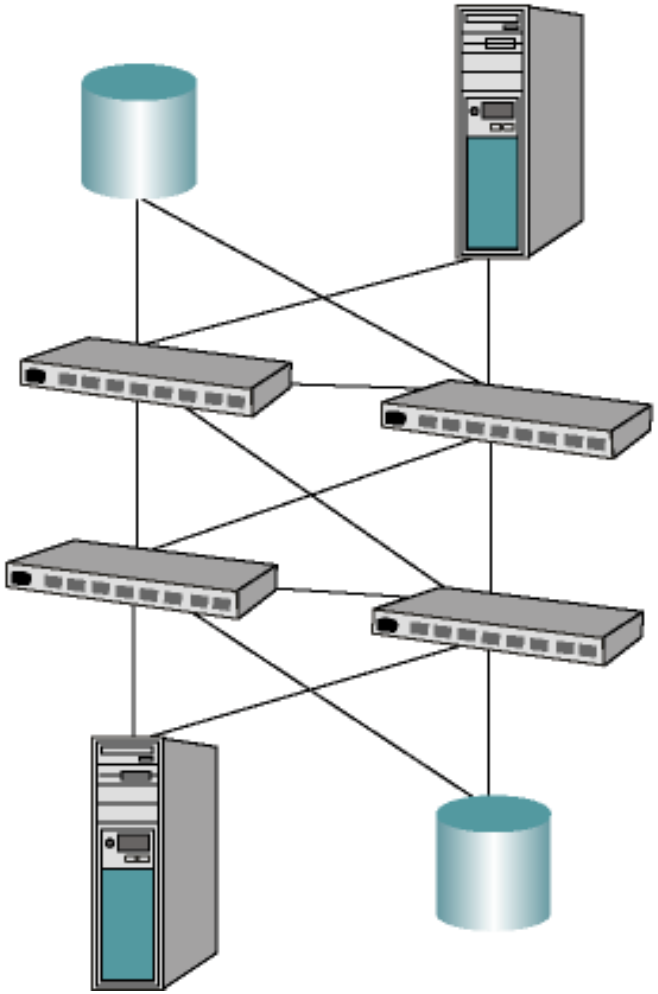
Switched Fabric Disadvantages

- Comparatively expensive
- Lack of standardization until recently
  - Interoperability concerns
- Most devices still FC_AL only
  - Many fabrics simply emulate FC_AL

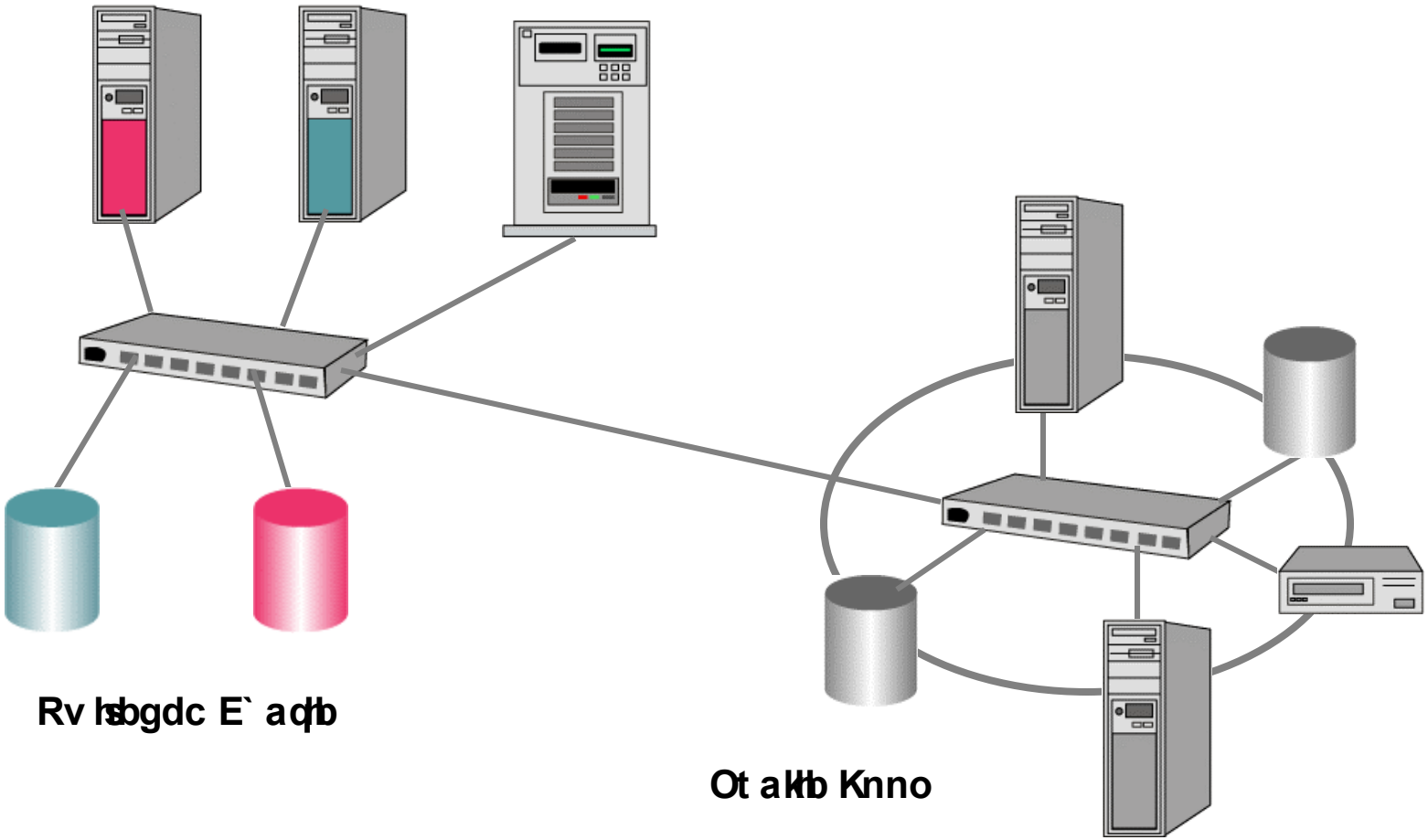
Simple Redundant Fabric



Full Fabric “Mesh”



Mixed Fabric and Loop Topologies



Fabric Versus Loop

FC_AL	Switched Fabric
127 Devices	16 million+ devices
Shares bandwidth (except when using loop switches)	Aggregates bandwidth
Limited to one conversation at a time per loop	Supports multiple conversations with dedicated bandwidth
Optimized for small departmental SANs	Optimized for large enterprise SANs
Performance may suffer when cascading switches	Designed for switch cascading
Industry standard since 1994	Interoperability standards now emerging
Supported by almost every FC device	Supports many devices by emulating FC_AL

Fibre Channel Addressing

World Wide Name (WWN)

- 64-bit unique name
- Similar to Ethernet MAC address
- Tied to hardware (assigned to ports and nodes)
- Usually assigned by the IEEE (each manufacturer is assigned a range)
- Globally unique
- Most reliable way to address a specific device

You Do Not Need to Know...

- Technical details of fibre channel addressing
- Ins and outs of loop initialization and arbitration
- Specifics of fabric addressing and communication
- Flow control
- Classes of service
- Fibre Channel layers
- 8b/10b encoding
- Port names/types

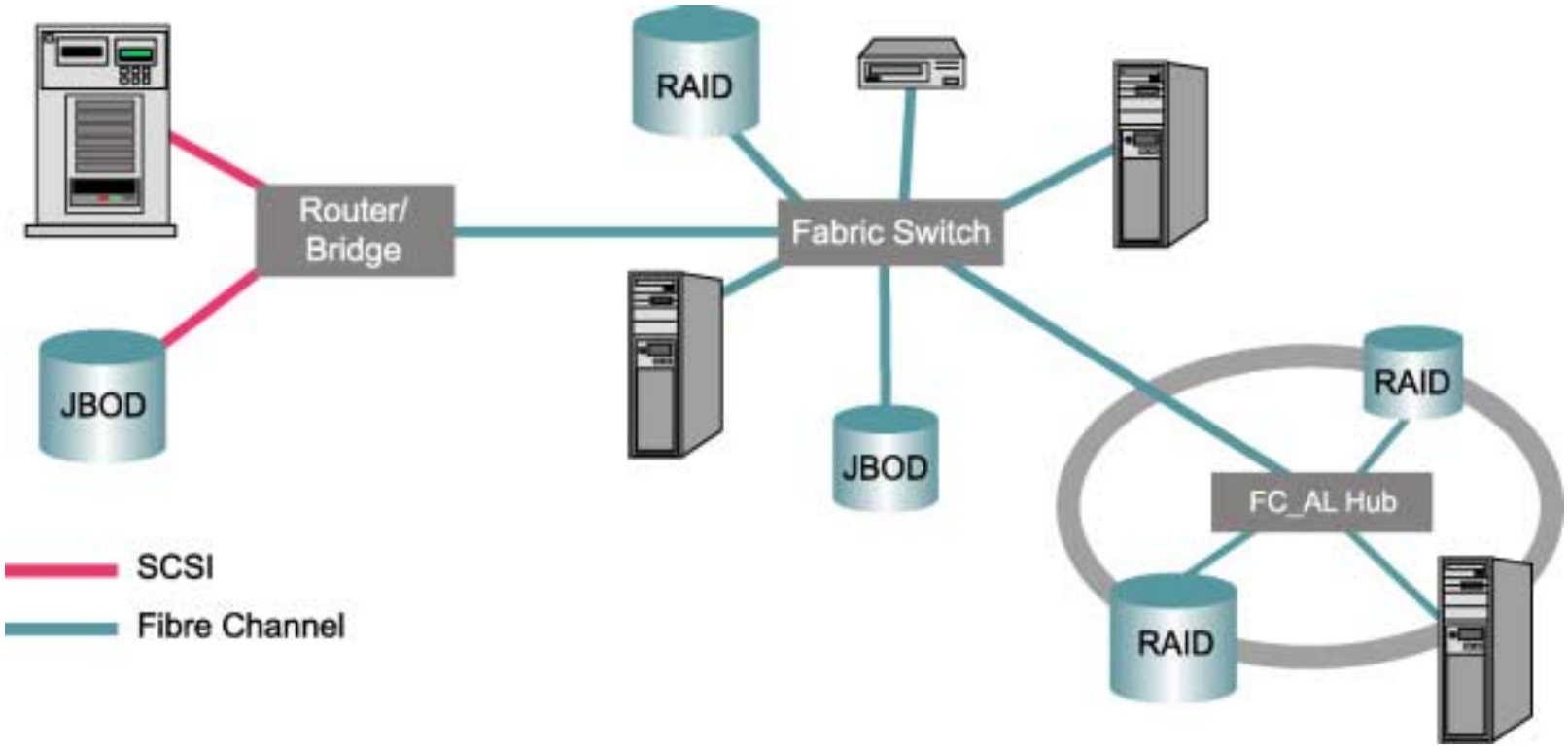
SAN Ingredients

(a.k.a. “What’s inside the cloud?”)

Components of Today's SAN

- Host Computers
- Fibre Channel Host Bus Adapters (HBAs)
- Tape Devices
- Cables and Connectors
- Fibre Channel Hubs
- Fibre Channel Switches
- Fibre Channel-SCSI Bridges and Routers
- JBOD & RAID subsystems
- Data Movers / Extended Copy Agents
- "SAN Appliances"
- SAN-Enabled Application Software
- SAN Management Software

SAN Components

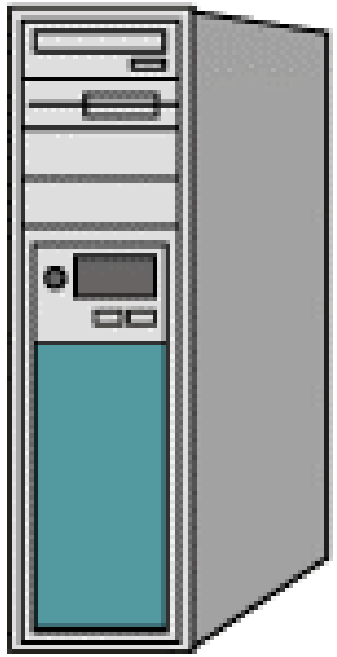


General Considerations

- Loop vs. fabric compatibility
- 1Gbps vs. 2 Gbps
  - Can be combined, but drop to lowest common denominator
  - 1Gbps is plenty and will be around for a long time
- Interoperability
  - Consult the manufacturer or other available literature

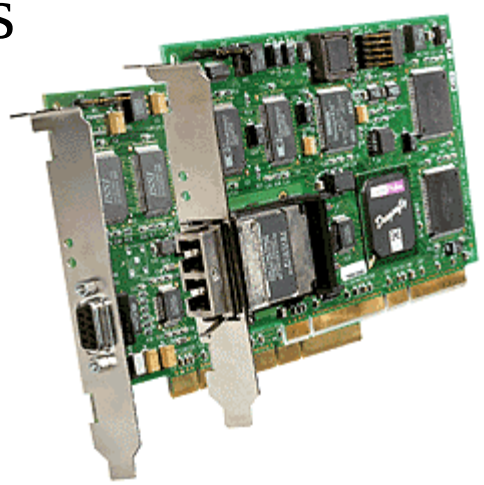
Host Computer

- Computer (generally a server) with the ability to communicate with other devices on the SAN
- Must have fibre channel HBA installed or be connected via SCSI to a device that can connect SCSI to the SAN.
- Generally acts as a SCSI initiator



Fibre Channel Host Bus Adapters (HBAs)

- I/O cards installed in a computer's PCI slot
- Looks like a SCSI card to the OS
- Looks like NIC if running IP or other network protocol.
- Connects hosts to fibre channel devices
- Both modular and fixed connections available
- Suppliers: Emulex, Qlogic, JNI, Major OEMs (Compaq, HP, Dell, Sun, IBM)



HBA Selection Criteria

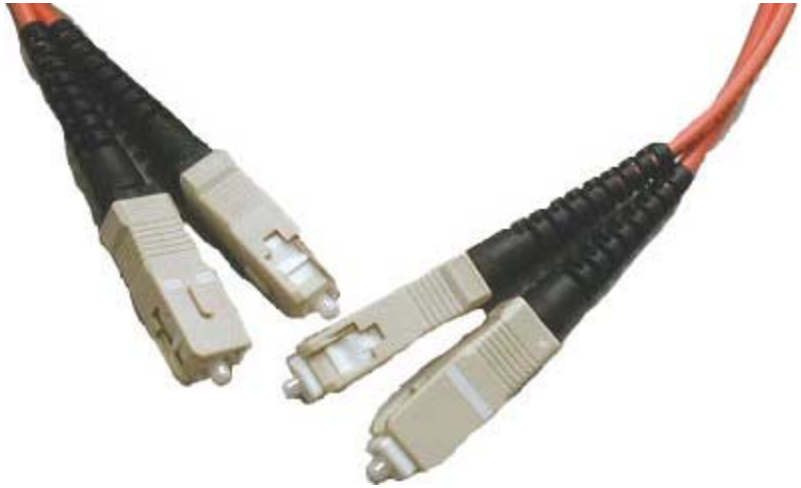
- Bandwidth (1Gbps/2Gbps)
- Single/multiple channel
- For PCI 33mhz or 66mzh
- Supported topologies (Full fabric/FC_AL)
- Supported protocols (e.g. IP, SCSI-FCP)
- Operating systems
- LUN masking support
- Co-processor included
- Teaming support

Cabling

- Twinaxial Copper (electrical)
  - Length: Max 30 meters
  - Least expensive
  - Generally used “inside the enclosure”
- Multimode optical (short wave fiber optic)
  - Length: Max 500 meters
  - More expensive than copper, less than single mode optical
- Single mode optical (long wave fiber optic)
  - Length: 10km
  - Longest range, but also most expensive

Cable Connectors

- Copper
  - DB-9
  - HSSDC



- Optical
  - SC: Most common optical connector; most are full duplex "SC-SC" connectors
  - ST
  - LC

Cable Selection Criteria

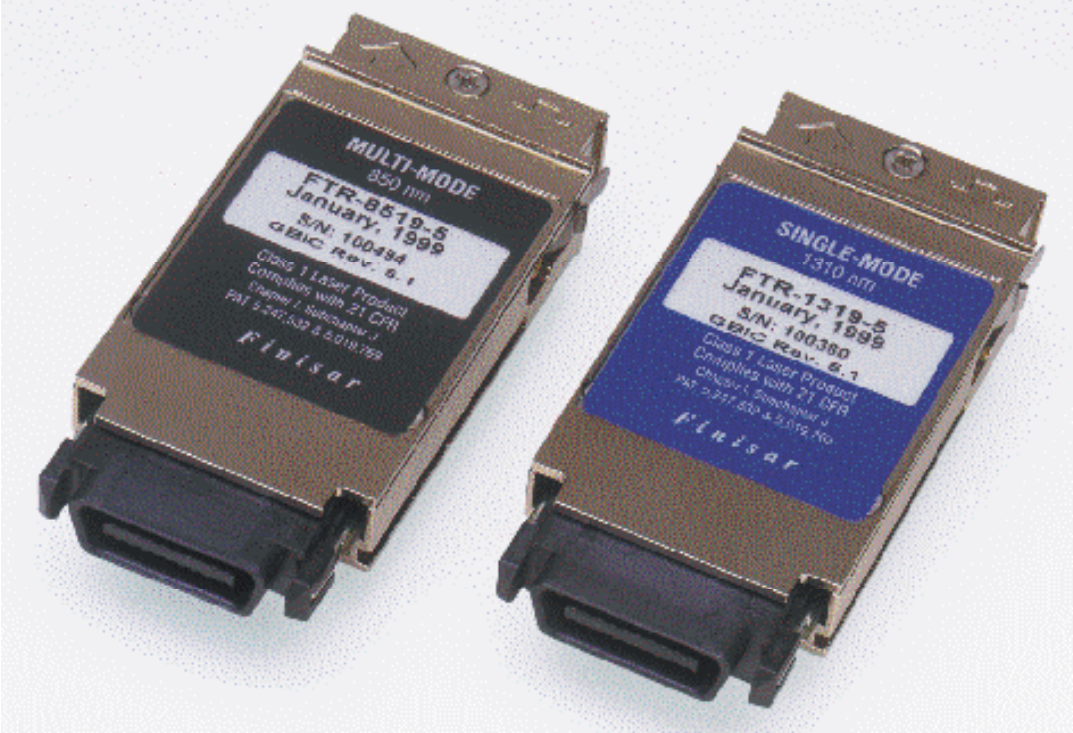
- Length needed
- Application
- Cost
- Connector type
- Intra/inter-cabinet

- Recommendation: Optical cables are ultimately cheaper and easier to run.

GBICs

- GBIC: Gigabit Interface Connector
- Used to connect cabling to fibre channel devices
- Most devices have empty slots; the GBIC inserted determines the type of cable connection
- Available in plastic, metal, fin types
- Hot swappable
- Available for all types of cabling
- Suppliers: Vixel, IBM, Finisar, AMP

GBIC: Gigabit Interface Connector



SFF/SFP Modular Transceivers

- SFF: "Small Form Factor"
- SFP: "Small Form Factor Pluggable"
- Replaces GBIC in 2GB and future technologies.
- Requires optical cables with LC-style connectors
- Smaller than GBIC, allowing more connections on a hub, switch or HBA
- SFP preferred because is hot-pluggable
- Suppliers: Finisar, Stratos Lightwave
  - More to come as 2Gbps speeds grow in popularity

Other Connectors

- GLM: Gigabit Link Module
  - 1Gbps
  - Mounts onto HBA like daughter card
- MIA: Media Interface Adapter
  - Converts copper to optical (and vice versa)
  - Suppliers: AMP, Stratos Lightwave

Fibre Channel Arbitrated Loop Hubs

- Very similar in concept to Ethernet Hubs
- Each device plugs into the hub
- Hubs can be “cascaded” with other hubs or switches
- Devices share available bandwidth
- Only one “conversation” can occur at once
- FC_AL only
- Up to 127 devices
- Port bypass circuitry protects continuity of loop
- Suppliers: Vixel, Gadzoox, Emulex, Major OEM’s (Compaq, Dell, HP, etc.)

Hub Selection Criteria

- Number of ports
- Port types
  - Fixed port vs. modular
  - GBIC vs. GLM-based vs. SFF transceiver
- 1Gbps vs. 2Gbps Fibre Channel
- Hot pluggable and/or redundant components
- Managed/unmanaged
- Support for zoning

Arbitrated Loop Switches

- Support up to 127 devices
- Bandwidth aggregated, not shared
  - Improved scalability
- Support zoning
- Recommended alternative to a loop hub: better value for the cost
- Suppliers: Vixel, Gadzoox
- Fabric switches might be licensed only for FC-AL support to compete with loop switches.

Fabric Switches

- Provides central connection point for devices
- Bandwidth aggregated
- Allows switch “cascading” with other fabric switches or either type of hub
- Support up to 16 million devices

- Support zoning
- Support “hot” device insertion/removal
- Multiple “conversations” can occur simultaneously
- Suppliers: Brocade, Gadzoox, McData, Vixel, Major OEM’s (Compaq, Dell, etc.)

Loop vs. Fabric Switches

Loop Switch	Fabric Switch
• Supports 127 devices • Easier to configure and manage • More affordable • Faster in single-switch environments • Intended for applications in which scalability is not the primary concern	• Supports 16 million devices • Requires more configuration • Faster in multi-switch environments • Required for large, enterprise SANs and connections to high end storage

Suitable Fabric Switch Environments

- Applications in which several devices are generating I/O at the same time
- Applications where performance is a concern
- Applications involving longer cable lengths
- Multi-drive backup applications
- SANs involving both disk and tape sharing
- SANs that must be able to scale significantly over time

Switch Selection Criteria

- Number of ports
- Port types
  - Modular (GBIC or SFF) v. fixed
- Bandwidth (soon 2gb switches will be available)
- Hot pluggable and/or redundant components
- Management features

Switch/Hub Selection Criteria

- Cost
- Scalability requirements (e.g. expected growth)
- Storage environment
- Number of devices
- Speed/bandwidth requirements
- Redundancy requirements

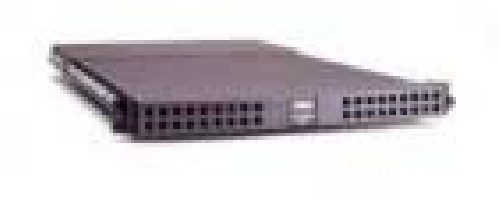
Fibre Channel-SCSI Bridges & Routers

- Used to connect Fibre Channel and SCSI devices together on the same SAN
- Most often used to plug tape backup devices and/or SCSI RAID and JBOD devices into SAN
- Some can be used to connect SCSI hosts to SAN

- “Bridge” implies limited or no awareness of SCSI protocol.
- “Router” implies awareness of SCSI LUNs.
- “Gateway” implies more scalable Router with advanced features.
- Suppliers: Crossroads, Chaparral, Pathlight, major OEMs

FC-SCSI Routers: Tape Applications

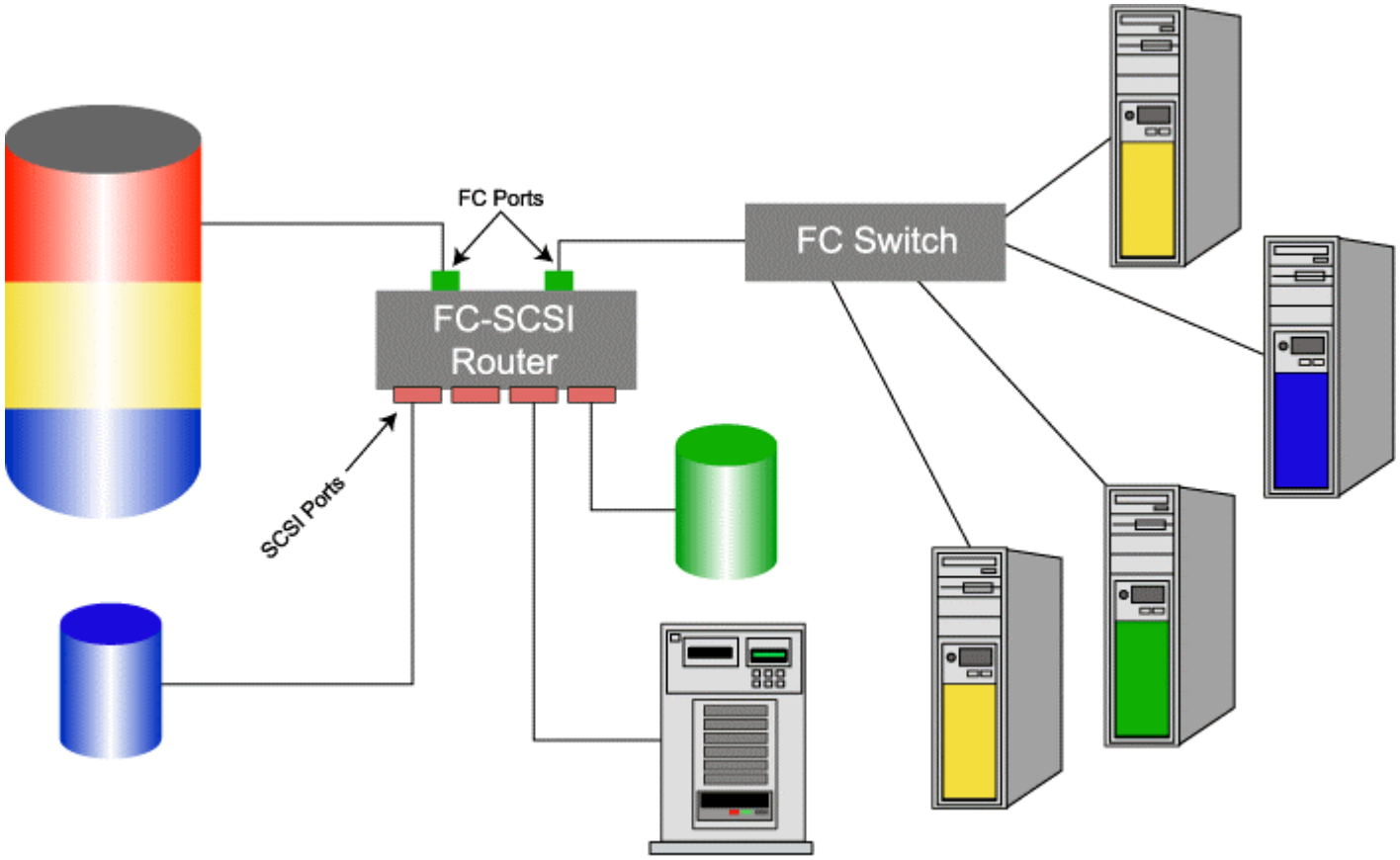
- Most “fibre channel” tape drives and libraries simply put a router or bridge inside the enclosure and come pre-connected.
- Some “native” fibre channel devices now available



FC-SCSI Router Selection Criteria

- Number of SCSI and Fibre Channel ports
- Maximum Bandwidth and I/O Handling
- SCSI type: HVD, LVD, single-ended
- Port types (Modular v. Fixed)
- Management features
- Hot plug and/or redundant components
- LUN masking support
- Data Mover Support
- Other advanced software

SAN with FC-SCSI Router



Tape Devices

- Include drives, libraries
- Used for backup applications
- All mainstream tape drives are parallel interface SCSI
  - FC tape devices include some FC-SCSI “converter”
  - FC-SCSI “converter” can just as easily be external.
    - Best of breed v. manufacturer’s assurance that it works together
- A few tape drives are native fibre channel, but the specification for fibre channel tape is not widely implemented.

Data Movers / Extended Copy Agents

- Workstation or special device that copies data from one storage device to another without first reading the data into a host computer
- Used for “serverless” backup applications
- Used for SAN-based disk mirroring applications
- Move data via SCSI Extended Copy command

Data Mover Location

- Data movers are packaged in three different ways:
  - A software option on a FC-SCSI router
  - A standalone device that plugs into the SAN
  - A software process running on a computer connected to the SAN

SAN Appliances

- Marketing term that could mean anything.
- Usually a PC running some OS with some software on it, packaged as an appliance to make it sound turnkey.

- Recommendation: If the vendor calls it an appliance, figure out what the catch is.

SAN Management Software

- Management software on a SAN plays a similar role to management software on a LAN
  - Certain events trigger alerts
  - Alerts are sent via SNMP (over Ethernet) to a management console
- Advantage of fibre channel: can be monitored with SNMP management tools
  - Hub or switch must support monitoring

SAN-Enabled Application Software

- Cluster and Failover Software
  - OEMs (Compaq, HP, SUN, SGI, etc.)
  - Operating Systems (Microsoft, Linux)
  - ISVs (Legato, VERITAS, etc.)
- Backup Software
  - CA ARCServe 2000
  - Legato NetWorker, SmartMedia, Celestra
  - Veritas BackupExec, NetBackup
- Disk Sharing & Virtualization
- SAN file systems

Some Leading SAN Software Vendors

Backup Vendors	SAN Management
• Backup Vendors • Legato Systems • VERITAS Software • Computer Associates • Tivoli • Syncsort • BakBone	• SAN Management • DataCore • StorageApp • Tivoli • Veritas • StoreAge • FalconStor • Prisa Networks • Major OEMs (e.g. Dell, Compaq, HP)

Chapter 3

Basics of Partitioning the SAN

Zoning

LUN Masking

Two Main Technologies for Partitioning

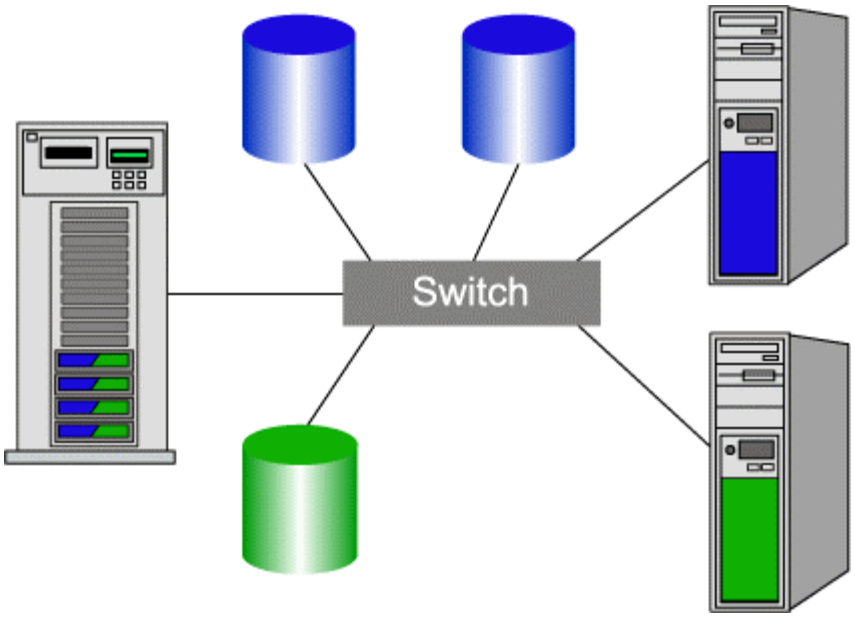
- **Switch Zoning**
  - Layer 2 filtering
  - Protocol independent
- **LUN Masking**
  - SCSI Protocol filtering at the specific device level
  - Same as LUN Assignment

Switch Zoning

- Makes mini virtual SANs out of all of the devices on the SAN
- Can be configured by port number or WWN
- Often sold as an add-on option
- Any given device can be a member of multiple zones

Switch Zoning

A single device can be in multiple zones:

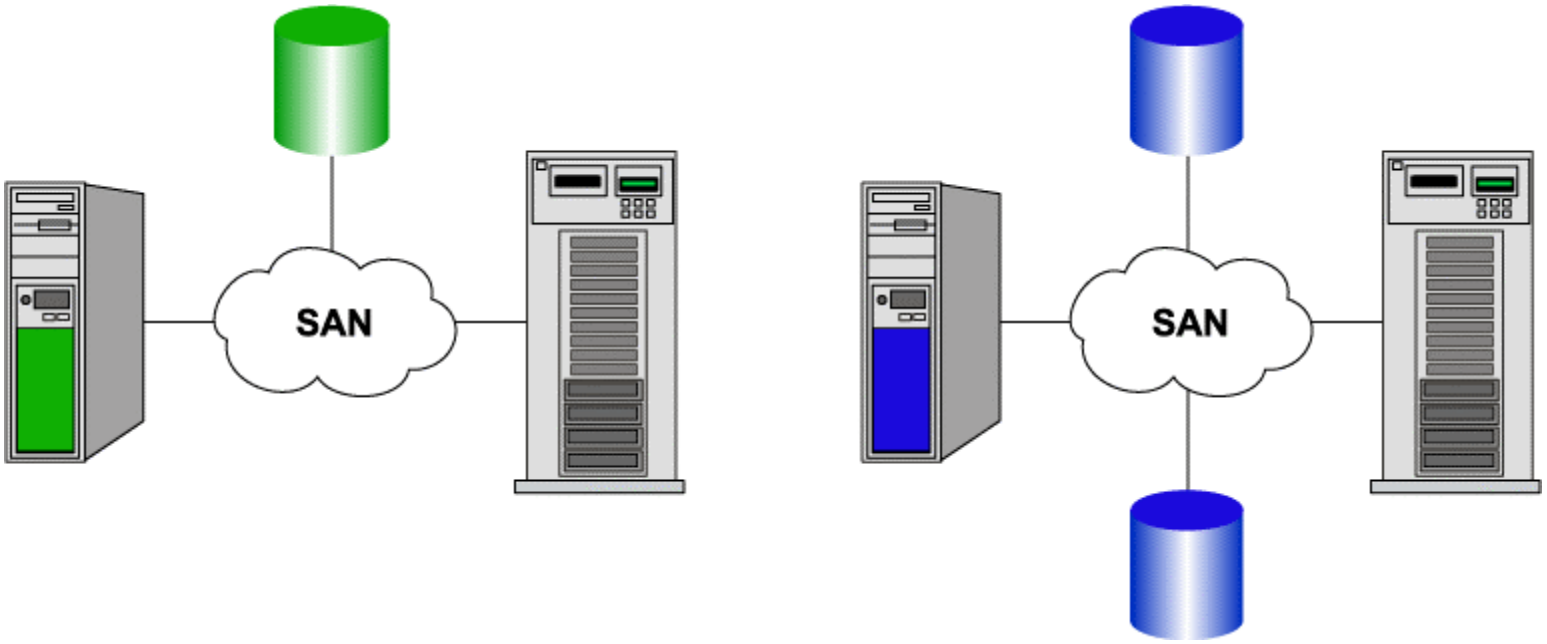


- **Blue devices:** Zone 1
  - Blue hard drives cannot be used by green server
  - Green hard drive cannot be used by blue server.
- **Green devices:**

  Zone 2
  - Green server can only use the green disk
- **Tape library robotics:**

  Zone 1 AND Zone 2
  - Both servers can share the tape library

Switch Zoning



The switch creates two smaller “virtual” SANs

Zoning Limitations

- Switch zoning is OSI Layer 2
- Not SCSI protocol aware
- Only recognizes device IDs (not LUN-aware)
- Cannot sub-divide a single device
  - Cannot be used for sharing a central disk array or tape library
- Zoning alone is not usually enough
- Often sold separately

LUN Masking

- LUN: Logical Unit Number
- Subset of the SCSI or FC_AL or Fabric ID
  - If ID = street address, LUN = apartment number
- Sub-Partitions in a RAID system are usually presented as LUNs
- Individual tape drives in a library are presented as LUNs

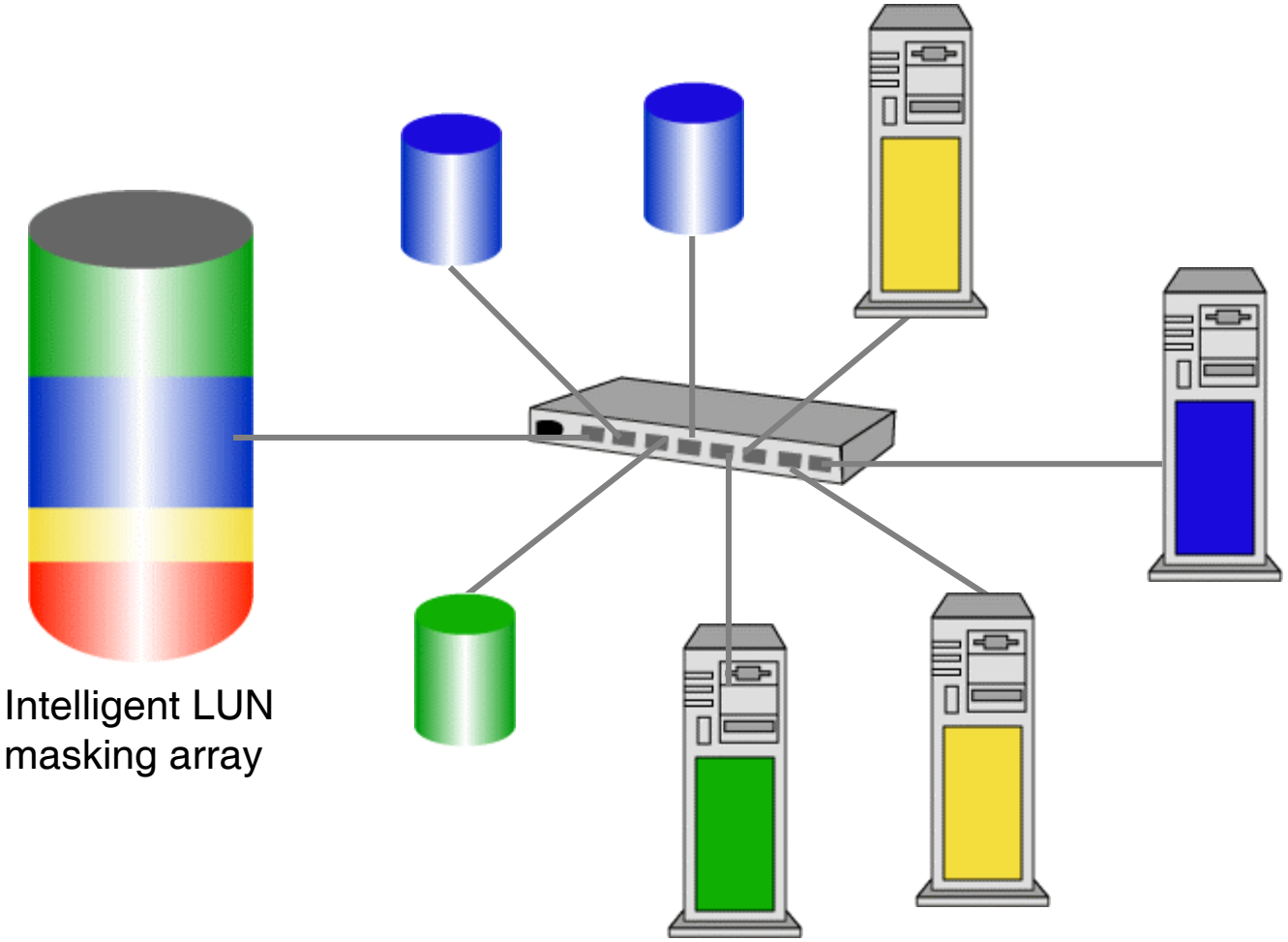
Possible LUN Masking Locations

- Host bus adapter
- Disk system
- Somewhere in the middle
  - FC-SCSI Router
- Not likely to be on the switch because switch is layer 2 device that is not SCSI protocol aware.

Devices Capable of LUN Masking

- Host Bus Adapter (driver or firmware)
- Fibre Channel - SCSI Routers
  - Usually used for connecting tape drives to SAN
  - Can be used for connecting hosts and disks
- Intelligent Disk Controllers
  - Like EMC Symmetrix and Modern Arrays
- Disk Virtualizers (a.k.a. “SAN Appliances”)

LUN Masking on Disk Controller



Summary of Partitioning

- Switch Zoning
  - Macro view division of SAN into logical mini SANs.
  - Happens at the switch, if the switch can do it.
- LUN Masking
  - Detailed sub-division of resources
  - Happens at the target, initiator, or somewhere in between. Many devices can do it.

Chapter 5

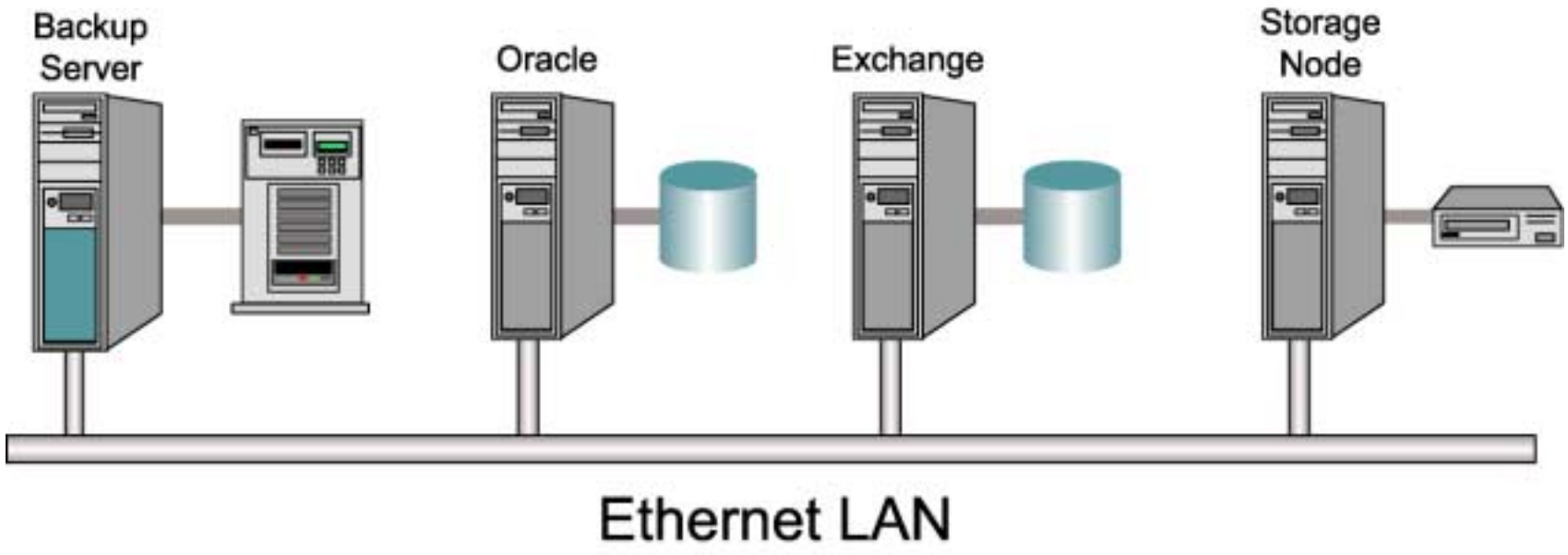
SAN Backup Technologies

LANless Backup (a/k/a Tape Library Sharing)

Serverless Backup

Clientless Backup

Traditional Network Backup System

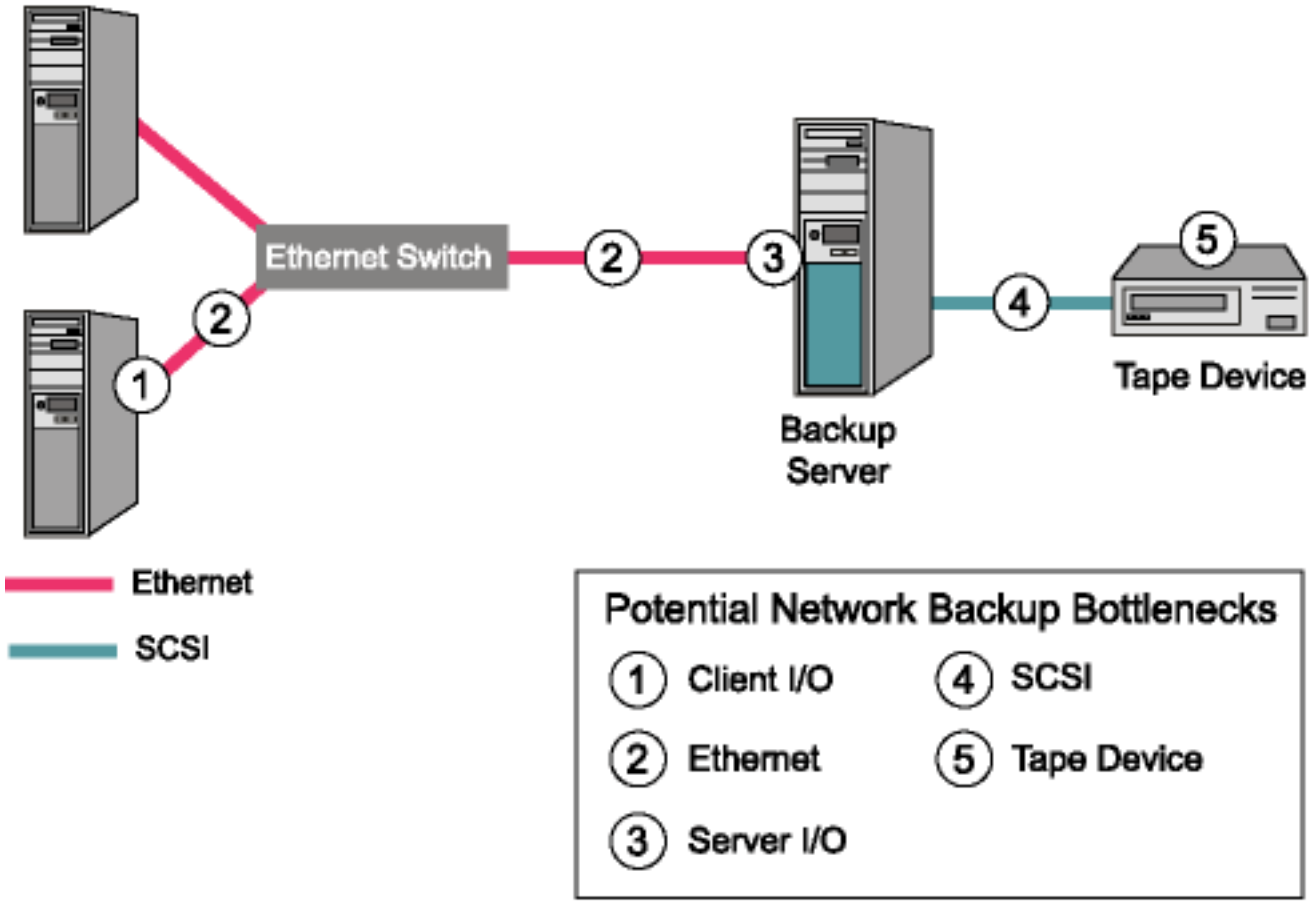


"Brainless" Backup

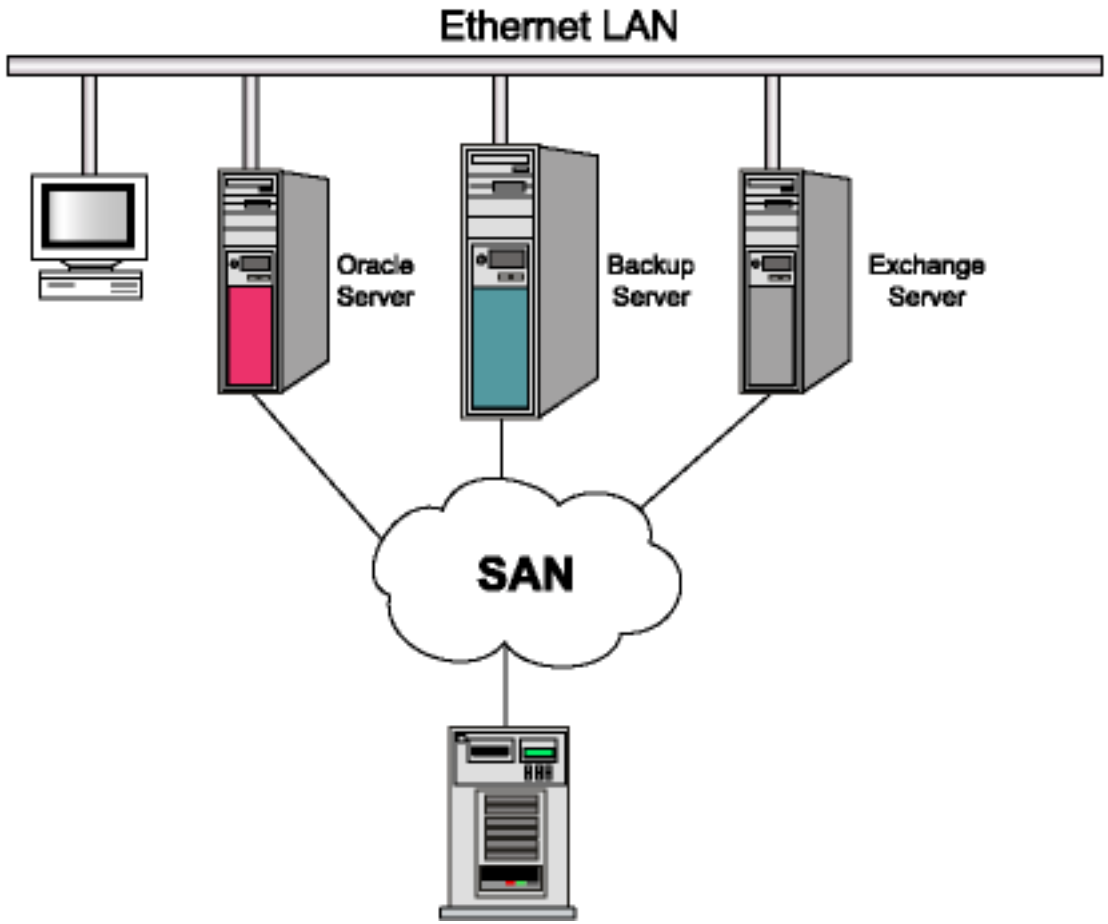
- Addition of Gigabit NIC backbone to improve server backup performance

- This is not a good idea
  - I/O Processing of the central backup server is still a big bottleneck
  - Fibre channel loop switching is actually cheaper than Gigabit Ethernet
  - A bit more money buys you technology for running SAN over Gigabit Ethernet. (More on this later.)

Tape Backup: Data Path Bottlenecks



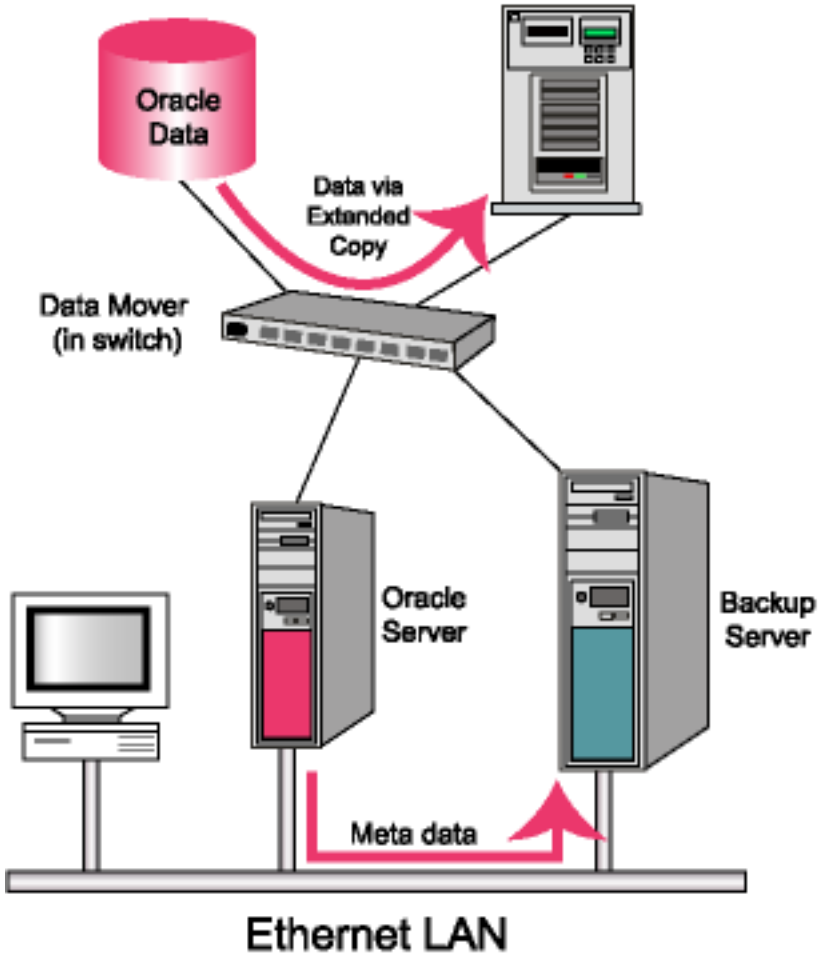
LAN-Less or LAN-Free Backup



Serverless Backup

- Data is sent from disk to tape without passing through the application server.
- Application server only has to send file system meta data; raw blocks of data are sent automatically
- Uses data-mover technology to send data straight from disk to tape with SCSI-3 extended copy command
- Disk storage MUST be on the same SAN as the tape backup device.

Serverless Backup

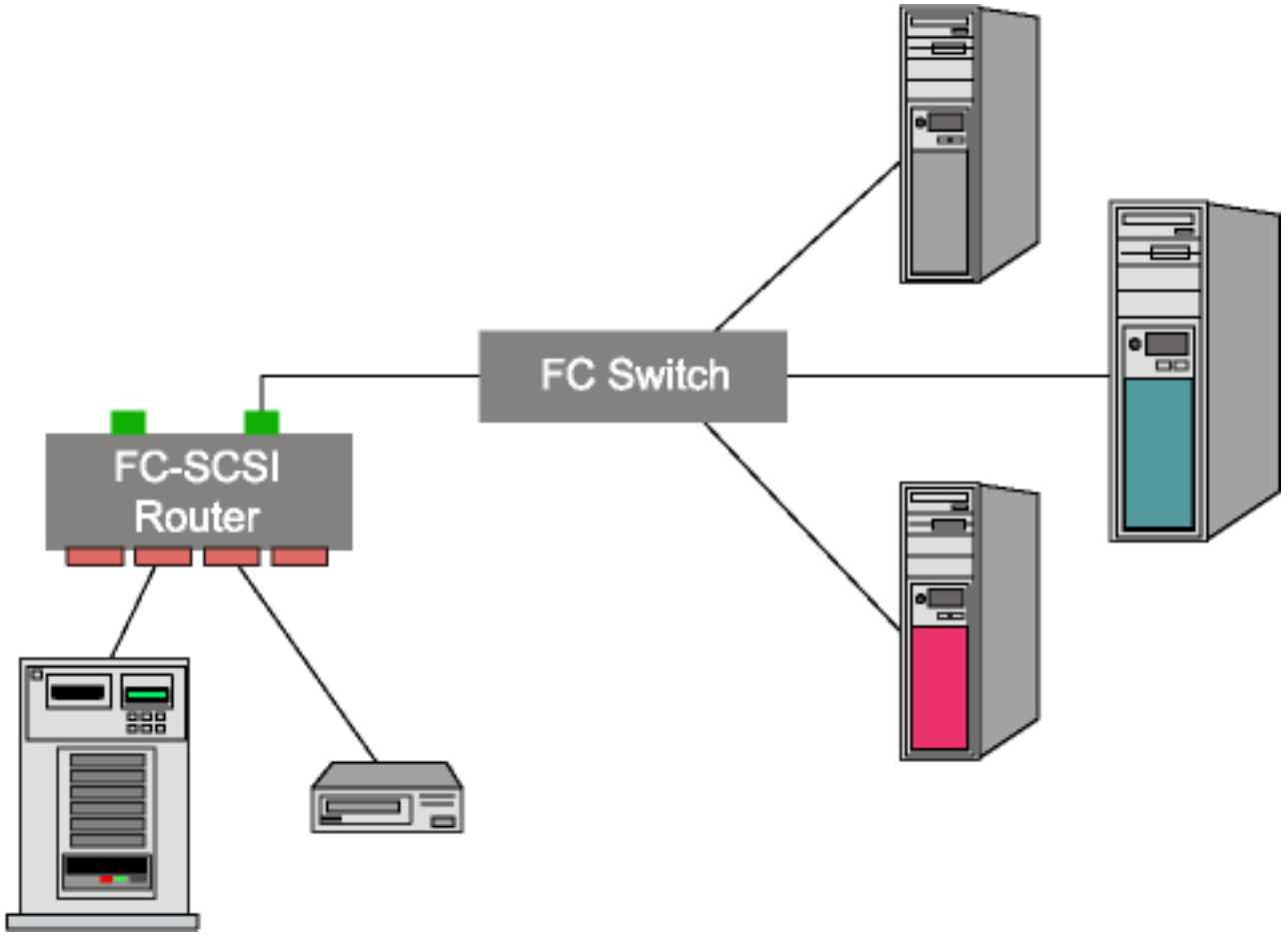


Client-less Backup

- Data of one application server is backed up via another server.
- Mirrors are broken or data is duplicated to another disk volume.

- This can be difficult to automate for some applications.

SAN-Based Backup Architecture



Tape Backup: SAN conclusions

- Fibre Channel and SANs enable smarter, better tape backup
- Because storage is a separate resource on a SAN, SAN-based backups become a storage issue rather than a network issue
- Soon storage will be divorced even from its owner-server, to further minimize problems.

Chapter 5

Overview of SAN Management

SAN Management

- Not to be confused with allocation of disk resources (a.k.a. “resource allocation” and “disk virtualization”.)

- Includes:
  - Element (Device) Configuration and Monitoring
  - Performance monitoring
  - Alerts and events via SNMP
  - Asset tracking management
  - Network topology diagramming

SAN Management

- Enables system administrators to set up, maintain & control the SAN
- SAN management tools perform the:
  - Configuration of SAN components
  - Monitoring SAN components
  - Document the SAN infrastructure

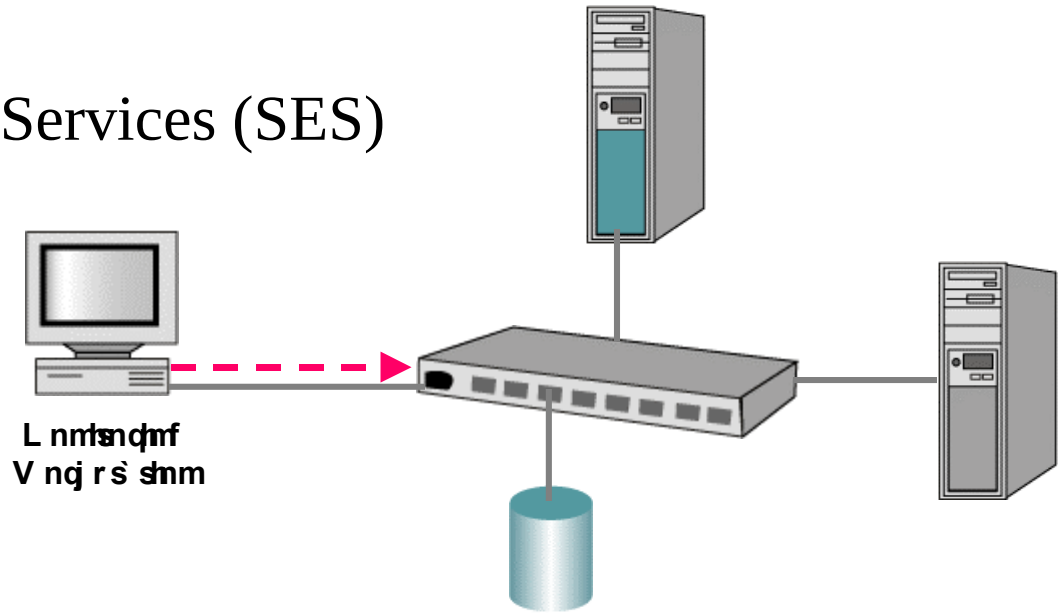
Management Tools: Types

- In-band
  - IP client using FC over IP
  - FC tester software
- Out-of-band
  - Using the console port or IP client

Do not confuse with disk virtualization terms.

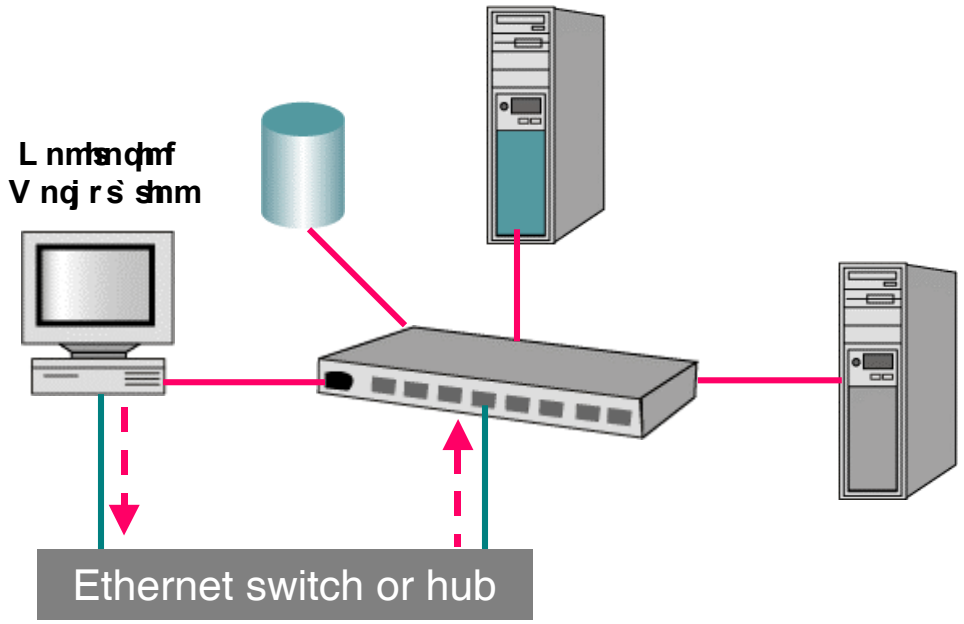
Benefits of In-band

- Provides a detailed topological map of the SAN
- No additional connections to the SAN
- In-band tools:
  - SNMP/IP
  - SCSI Enclosure Services (SES)



Benefits of Out-of-band

- Connection is not disabled when the Fibre Channel network is down
- Out-of-band tools:
  - Ethernet
  - Telnet
  - Serial ports



SAN Management Tools

- Vendor supplied tools (proprietary)
- Via console port (proprietary)
- Via IP
  - Telnet and/or Web client (proprietary)
- Via SNMP
  - Unicenter, TNG, OpenView, Patrol, etc.
- Via SES (storage enclosure services)
- Via 3rd party application
  - Connex, Prisa, Qlogic, Vixel, etc.
- API Tools
  - Veritas V2

Appendix A

Alternatives to Fibre Channel

Alternatives to Fibre Channel

- Ethernet
  - iSCSI
  - Storage over IP
  - IP Tunnelling
- ESCON
- Infiniband

Fibre Channel Advantages v. Ethernet

- FC is faster for large chunks of data, like those used in mirroring, copying, backup, and restore.
  - Larger block size means fewer I/Os required to move same amount of data.
- Channel properties of fibre channel better support SCSI protocol.
  - SCSI is dominant protocol for moving data.
  - Efforts underway to standardize SCSI over Ethernet and SCSI over IP. Additional Ethernet overhead require to reliably carry SCSI.
- Short term roadmap has FC going to 2Gb this year.

Ethernet SAN Challenges

- IP is a network protocol. SCSI is a channel protocol.
  - Packets must arrive in the order in which they were sent.
  - SCSI devices cannot compensate for error rates common in IP/Ethernet transmissions.
  - Processing overhead needed to do iSCSI reliably is extraordinary.
- Solvable Problem
  - UDP instead of TCP
  - Co-processed adapter cards
  - Smarter software

Ethernet Advantages over FC

- Lower cost
- Ubiquitous
- Long term roadmap of Ethernet (10 Gigabit) exceeds long term roadmap of fibre channel (4 Gigabit).
- Ethernet Industry is powerful
  - Technical problems are easy to solve
  - Storage is viewed as lucrative!
- As processors get faster and I/O bus problems get resolved, Ethernet could compete more effectively.

SCSI over IP

- Industry standardized effort is called iSCSI
- Uses IP protocol to carry SCSI commands
- Takes advantage of existing IP infrastructure
- Rudimentary drivers available to run iSCSI with Gigabit Ethernet cards.
- iSCSI host adapters with I/O processing in ASICs and dedicated processors due early next year.
- Plans in the works for iSCSI peripherals and gateways.

Conclusions and

Questions & Answers