

Deploying SAN Transports

2GBps FC vs. iSCSI vs. Infiniband

Wayne Rickard

wayne@gadzoox.com

Sr.VP/CTO, Gadzoox Networks

16241 Laguna Canyon Road, Suite 100
Irvine, CA 92618
(949) 453-1270 FAX: (949) 453-1271





Agenda

◆ Industry Trends

- Protocol convergence
- Technologies

◆ SAN Drivers

- Short term (LAN Free, Server Free, Storage Consolidation)
- Long term (Virtualization)

◆ Stacking up the Transports

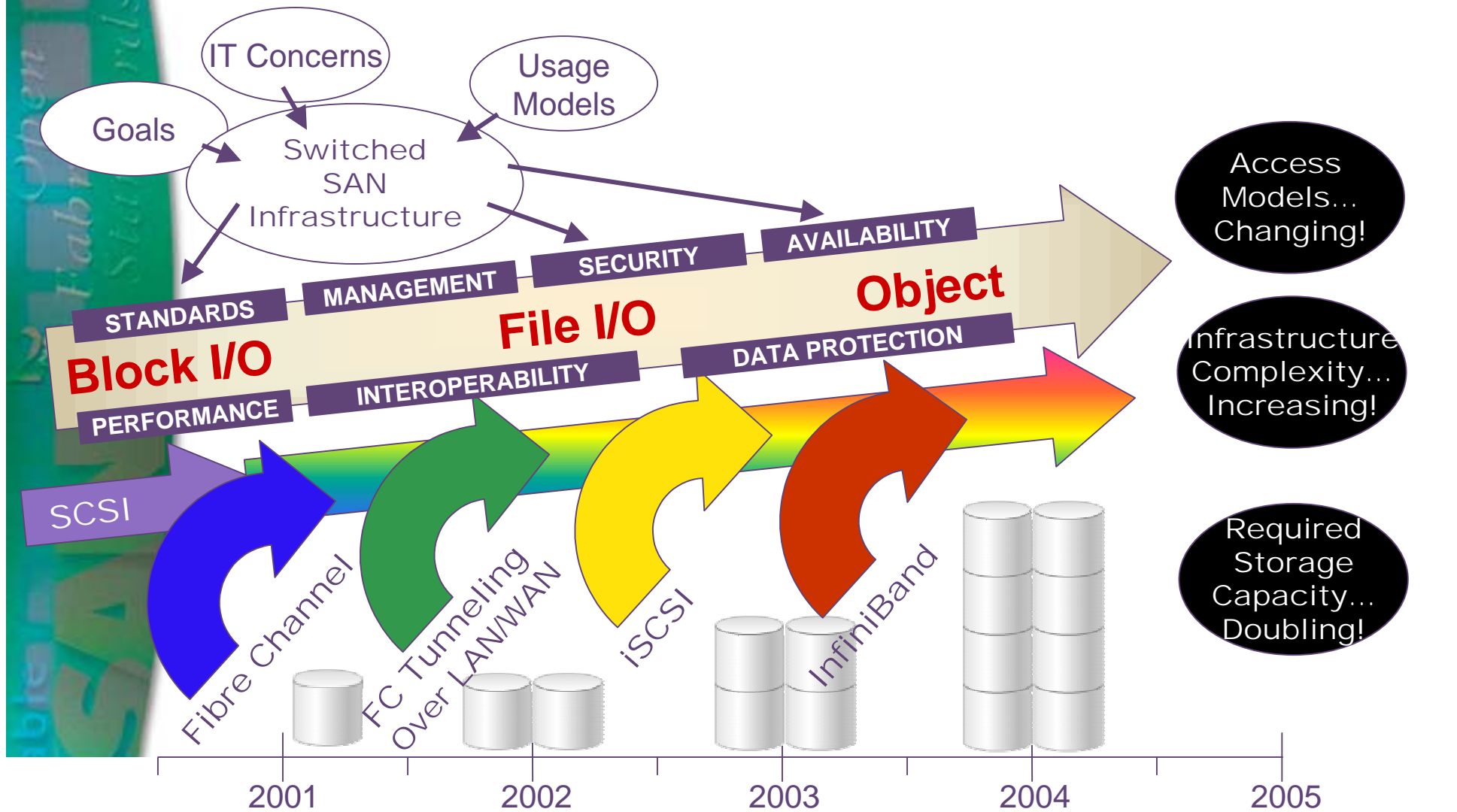
- Fibre Channel
- Ethernet
- InfiniBand

Industry Trends

What about Storage over IP?
What about Clustering?
What about InfiniBand?



Storage Market & Technology Trends



SAN Applications

Who needs a SAN?





The Five IT Goals

- ◆ Better IT Staff Efficiencies
 - Improved Service Levels
 - Reduce IT Costs
- ◆ Deferred Investment in Capital Assets
 - Make Storage investment decisions independent of Servers and Applications
 - Extend the life of IT investments
- ◆ Quick Solution Deployment
 - Standardization of Infrastructure
 - Increase scalability of capacity, performance
- ◆ Eliminate Downtime - 24 x 7 operation
- ◆ Never get paged on the weekend



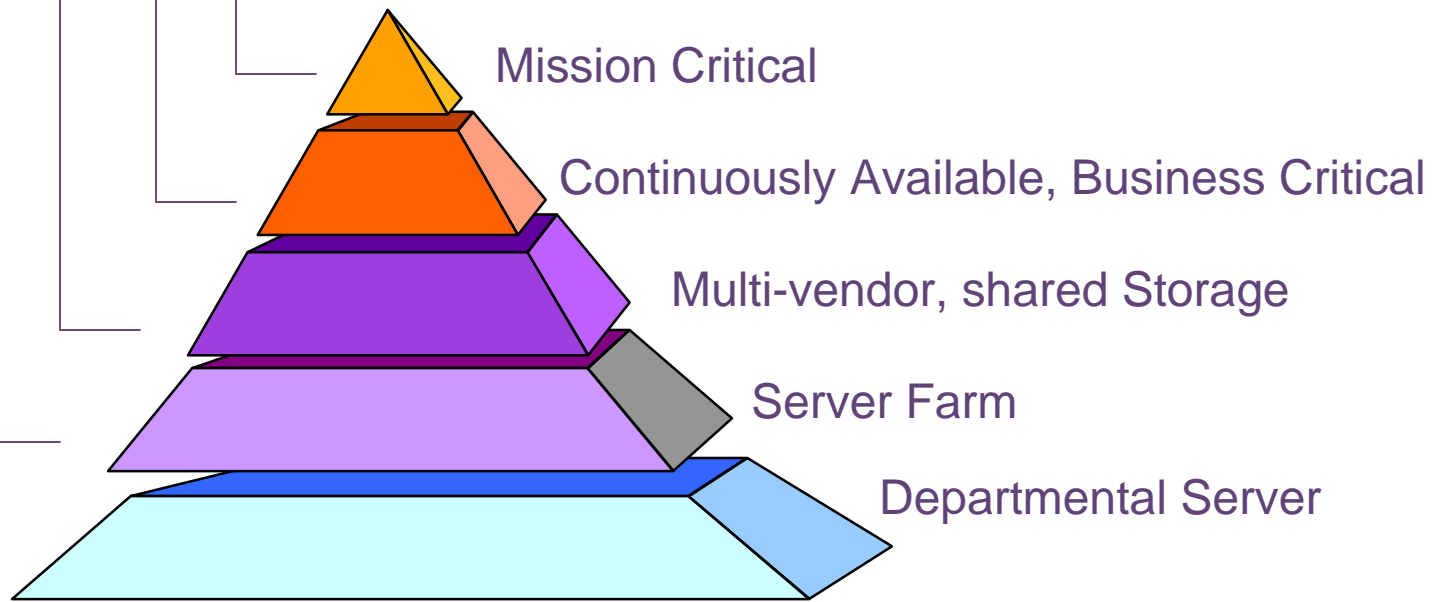
The Five Driving Applications

- ◆ Backup
- ◆ High Availability
 - Cluster Failover
- ◆ Centrally Managed Storage Resources
 - Drag & Drop Storage Allocation
 - HSM
 - Automated hotspot management
- ◆ Server Consolidation
- ◆ Disaster Recovery

Storage Freedom
 Open Enterprise Standards
 Gigabit
 Scalable
CAN

Users - IT Community Needs

✓	✓	✓	✓	✓	Simplify Backup
	✓	✓	✓	✓	Consolidate Enterprise Storage Management
		✓	✓		Server Consolidation/Share Storage Pools
			✓	✓	High Availability/Data Clusters
				✓	Disaster Recovery



Server/Storage Usage Models



Models with high SAN Needs

What: Traditionally a market for midrange and mainframe. Now also targeted by largest Wintel vendors as a growth market, enabled by **clusters**.

Requirement: SANS to support efficient, reliable, networked access **by applications** to data.

Focus: **Availability & Reliability**

Where: Business Unit

Continuously Available, Business Critical

Multi-vendor, shared Storage

What: An emerging market, enabled by IT Managers who need to be able to de-couple **Storage purchases** from Server purchases.

Requirement: Storage connections to SANs are open, interoperable, and OS-agnostic.

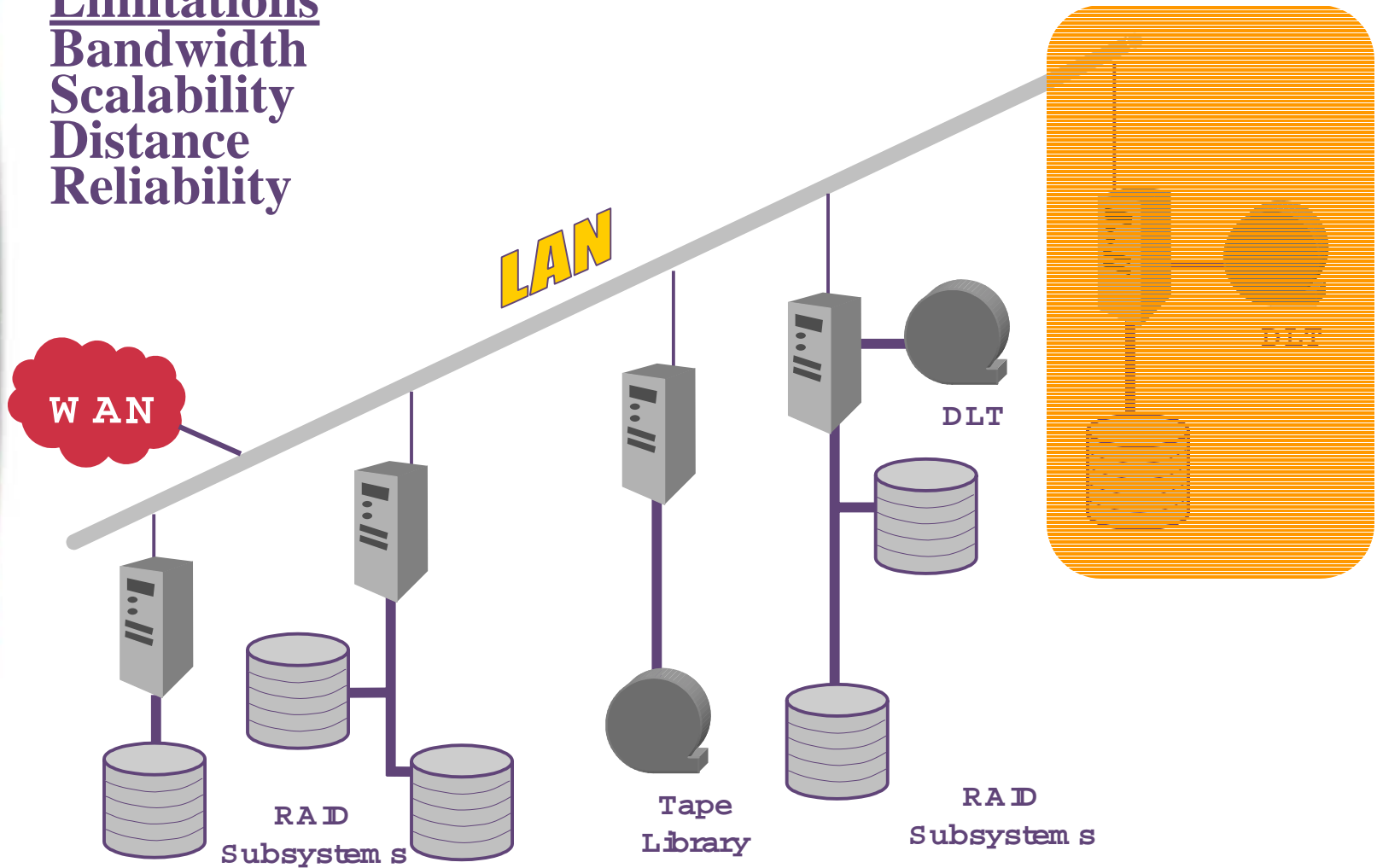
Focus: **Manageability**

Where: Firmwide Infrastructure

Old Model

Limitations
Bandwidth
Scalability
Distance
Reliability

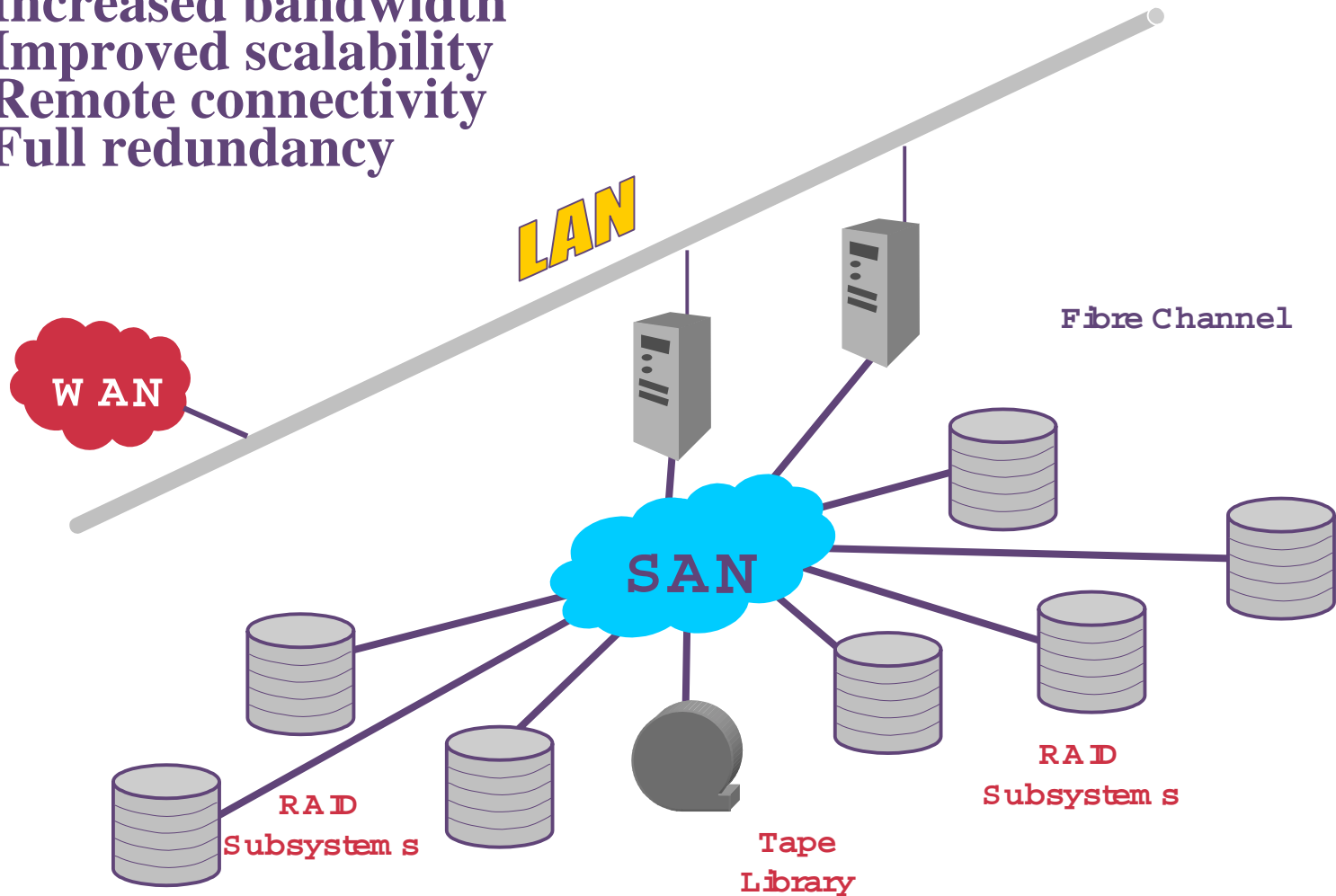
Unit of Scalability



Storage Networking Model

Benefits

Increased bandwidth
Improved scalability
Remote connectivity
Full redundancy



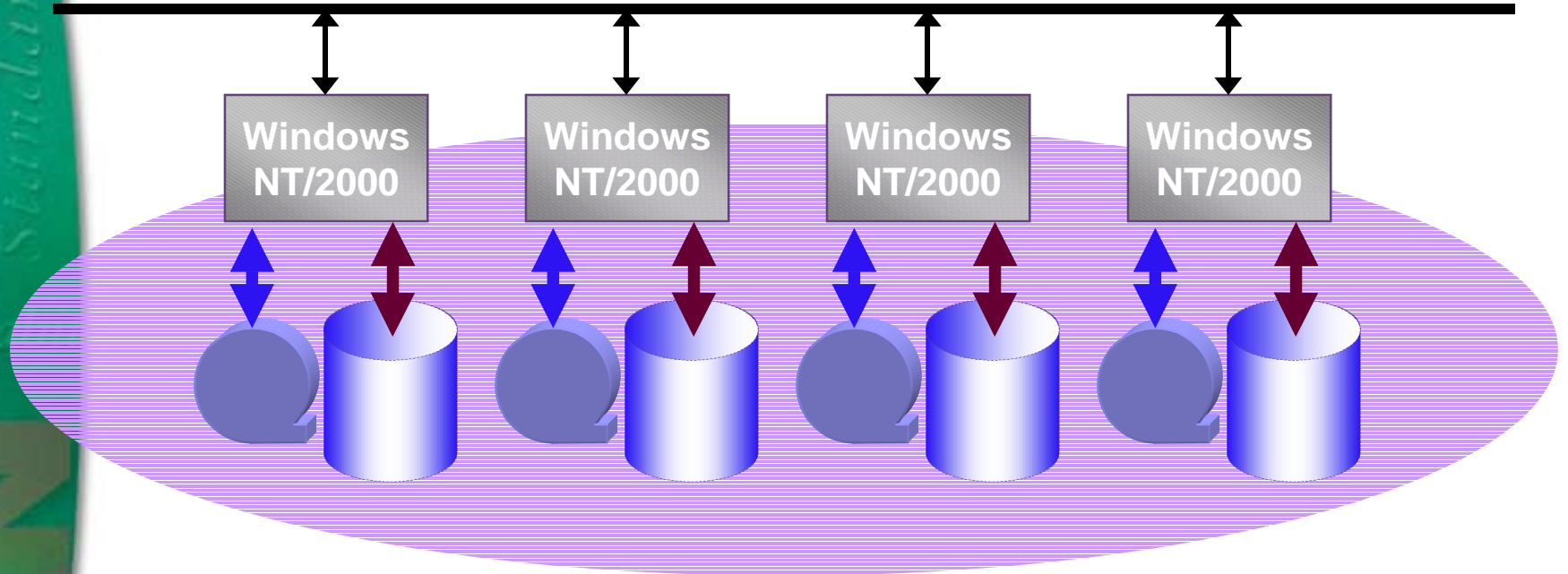
How will SANs Be Adopted?

Stages of Deployment
A Model for Shared
Storage



What We're Starting From...

LAN



- Unit of replication duplicates storage, server, and backup resources
- Multiple single points of failure
- Multiple points of Management
- Leveraging Backup investments congests the LAN
- No shared storage infrastructure**



The SNIA Shared Storage Model

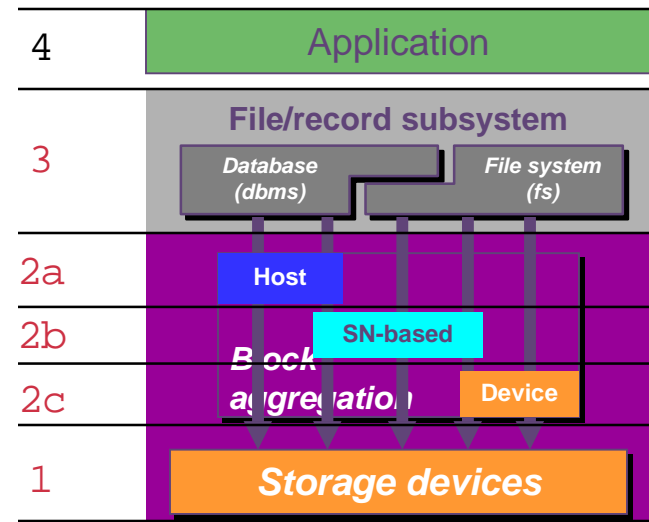
- ◆ A simple model for shared storage architectures
- ◆ Use it to describe common implementations graphically
- ◆ Expose, for each one:
 - What services are provided
 - Where *interoperability* is required
 - Pros and cons of the architecture



The SNIA storage model

A layered view

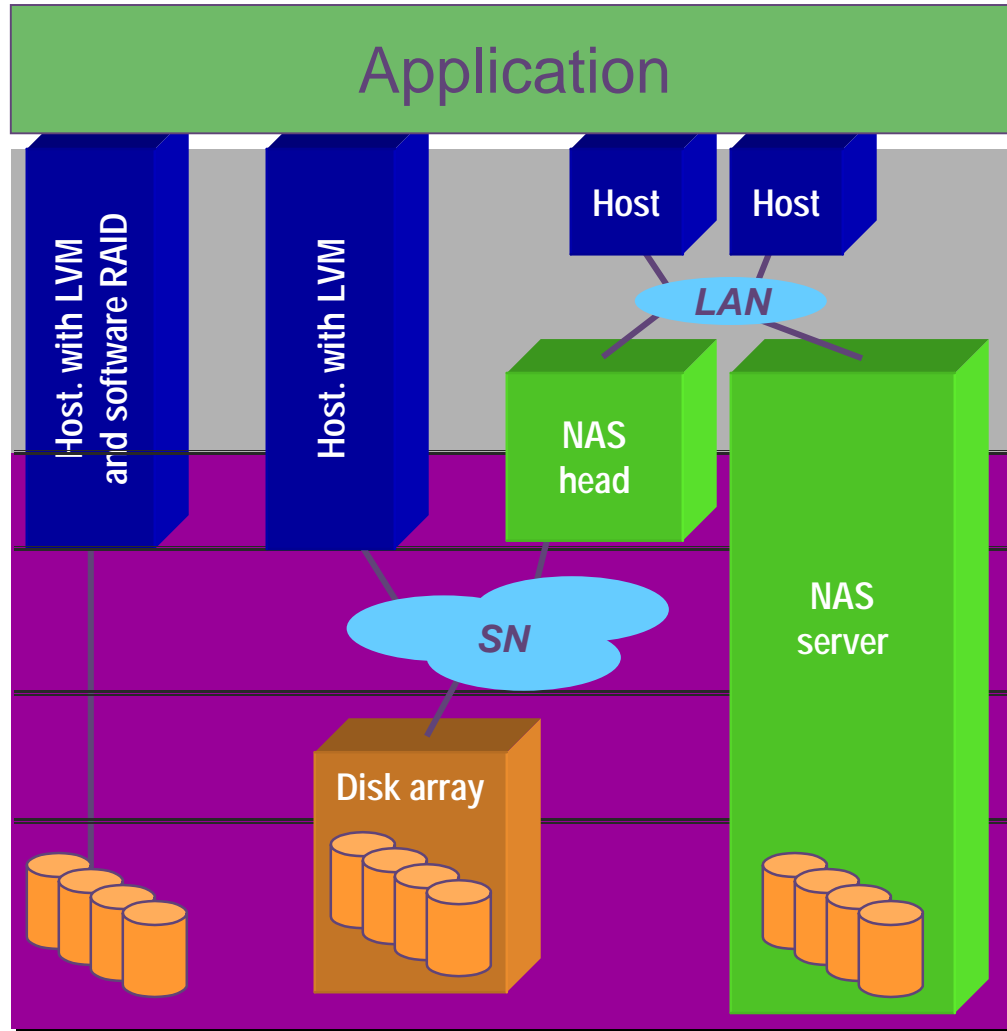
- 4. Application
- 3. File/record subsystem
 - 3a. Database
 - 3b. File system
- 2. Block aggregation
 - 2a. Host block aggregation
 - 2b. SN block aggregation
 - 2c. Device block aggregation
- 1. Storage devices



The SNIA storage model

Sample architectures

File/record subsystem
 Block subsystem



1. Direct attach
2. SN attach
3. NAS head
4. NAS server

← Host block-aggregation

← SN block-aggregation

← Device block-aggregation

The SNIA Storage Networking Model “Infrastructure for Data”



File/Record Subsystem

Storage Interconnect

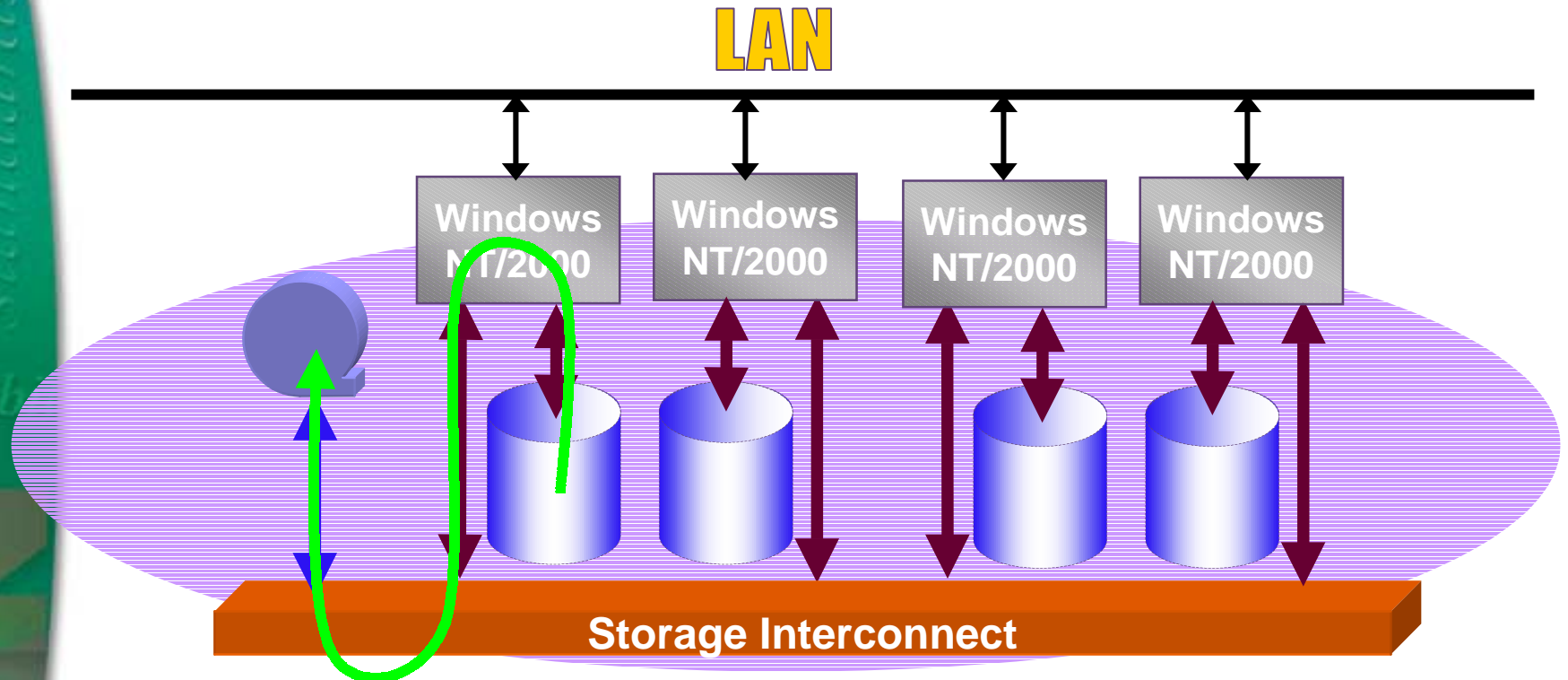
Block Aggregation

Device Interconnect

Storage Device

- ◆ **Storage Clients:**
“Consumers” such as file systems, DBMS, applications
- ◆ **Virtualization:**
Aggregators (such as LUN managers, RAID, etc.) and attribute-extenders
- ◆ **Storage Devices:**
Disk Subsystems, Tape

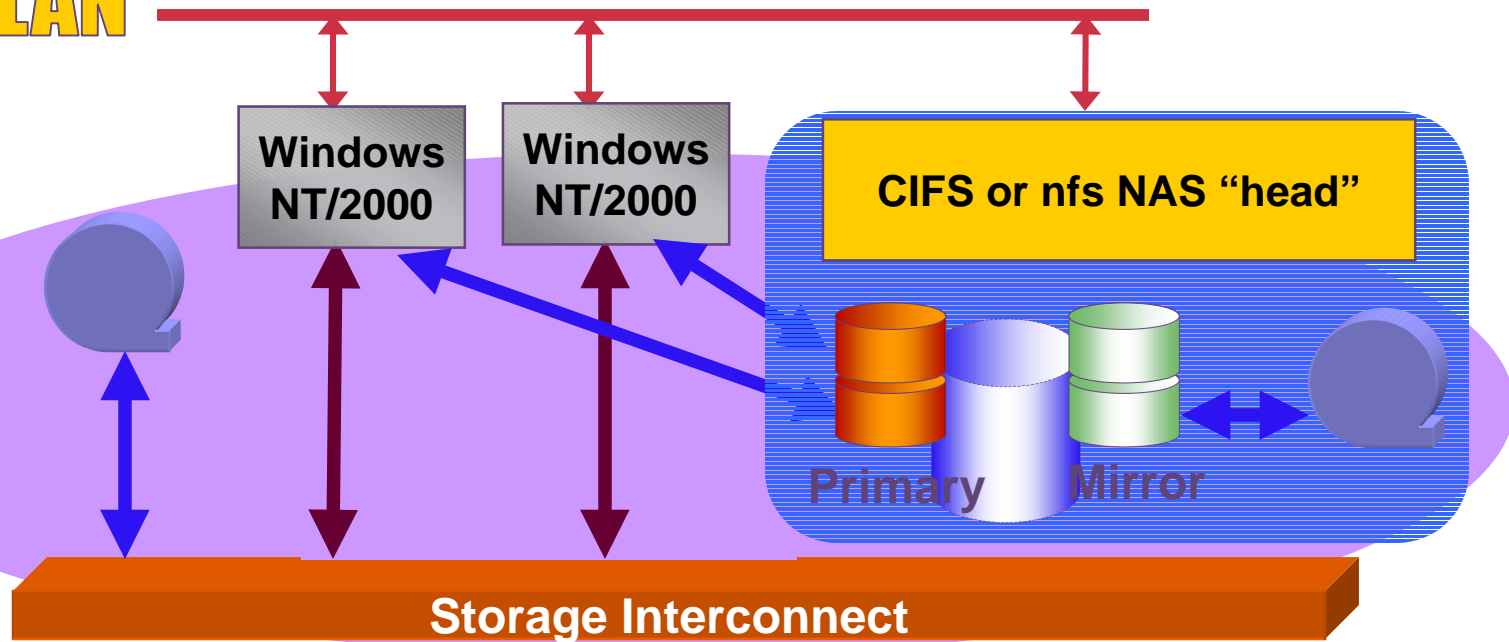
LAN-Free Backup



- Simplest SAN Deployment - Only Tape is moved to SAN
- Sharing Tape Resources requires Backup SW Application Support
- Very low risk for highest immediate return

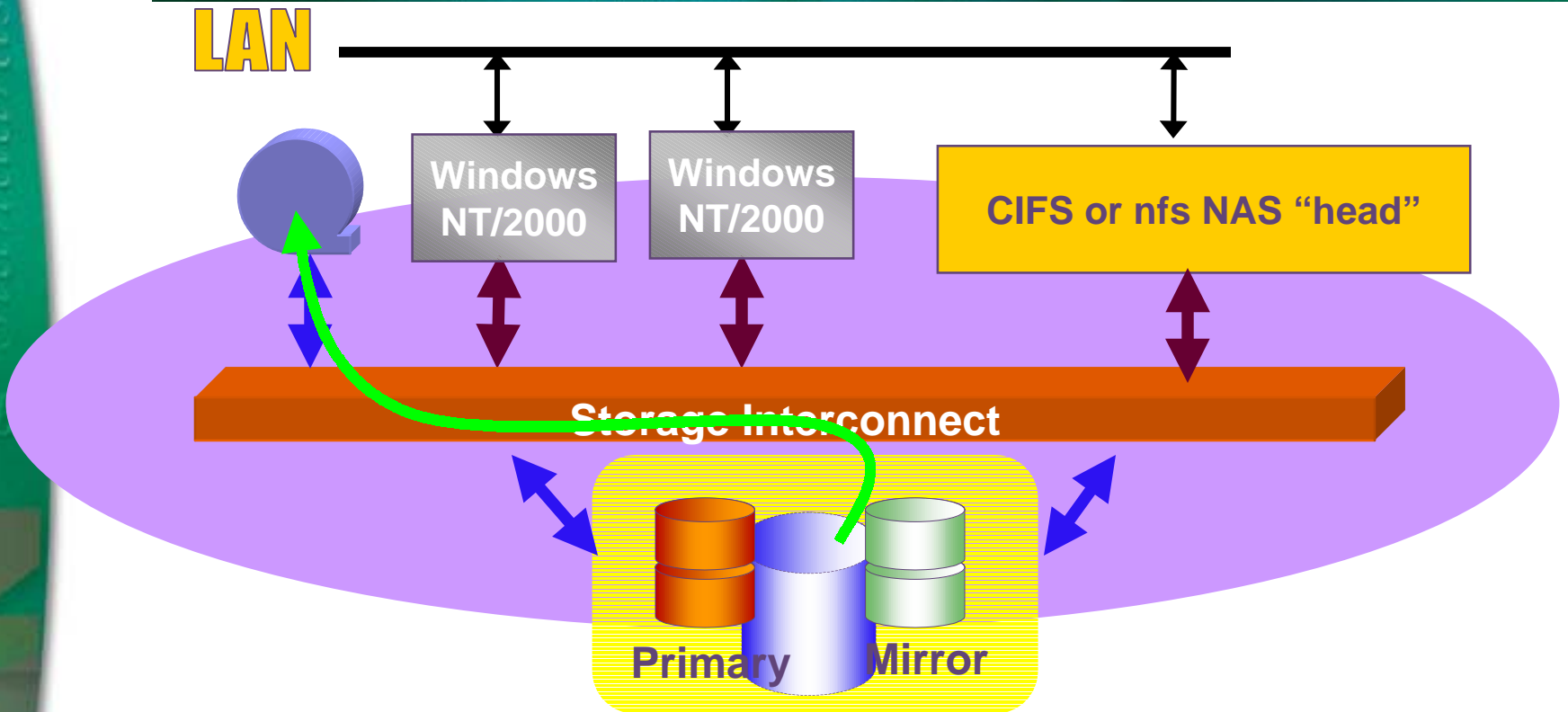
Increasing Storage Availability

LAN



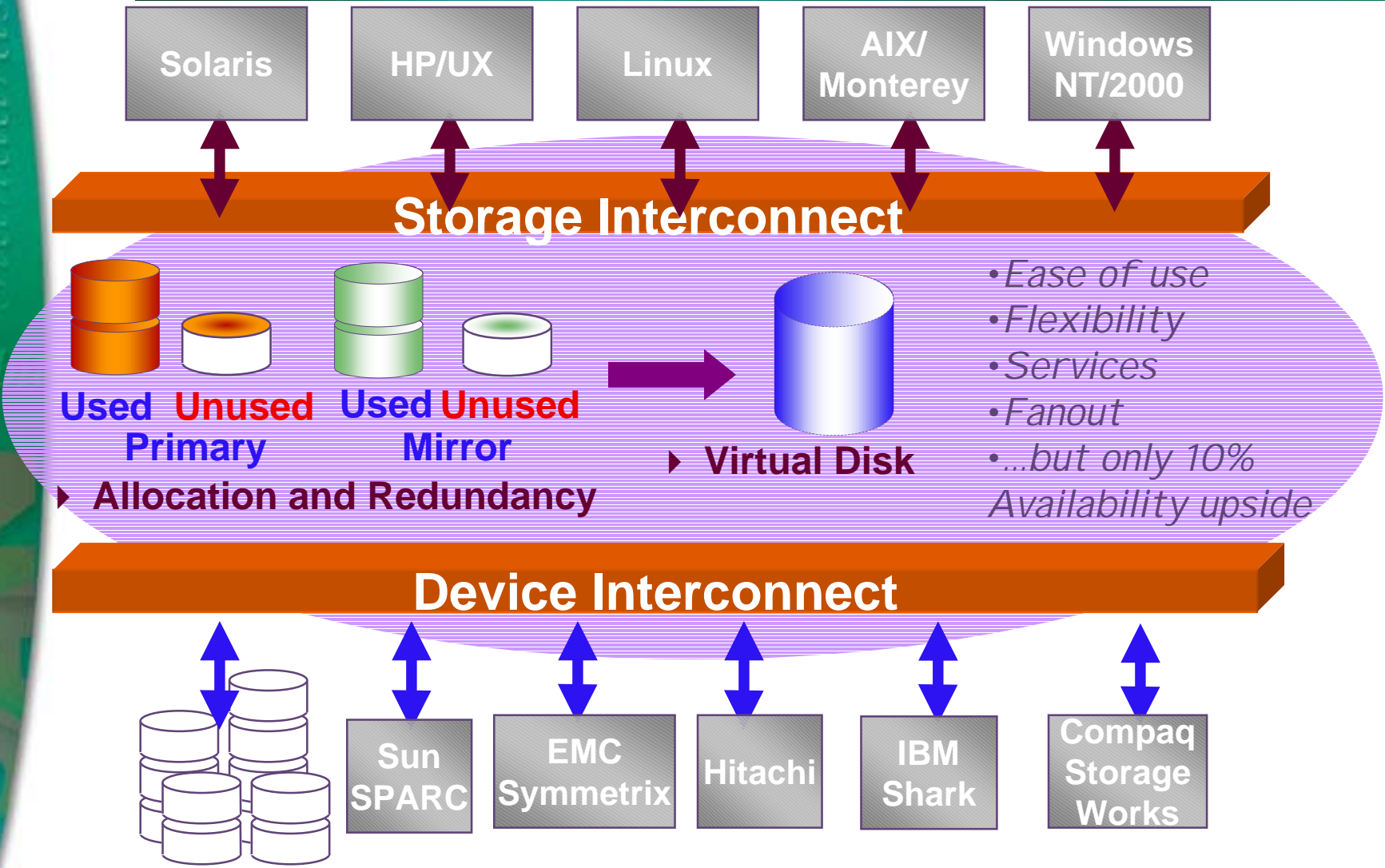
- ☺ Centralized Storage Resource Pool
- ☺ Centralized Storage LUN Management and live Backup
- ☺ Design Storage Subsystem for desired SLA - uptime, I OPS
- ☹ Reduced Vendor & Configuration choices
- ☹ "Infrastructure in a Box" limits scalability
- ☹ Expensive

Increasing Storage Scalability

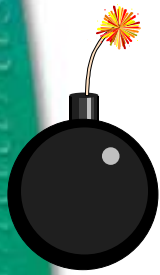


- ☺ Centralized Storage Resource Pool
- ☺ Centralized Storage LUN Management and live Backup
- ☺ Enables Server-Free Backup

SAN "Grail": Multi-Vendor Shared Storage



10%? What Really Fails?



5% - environment

43%

Ethernet
Hubs & Switches



15% - users

Tape
Library

Disk Arrays

Disk Drives

RAID
Controller

Host Bus
Adapter

9% - Network

Client
SW

18%

Storage

Servers

LAN

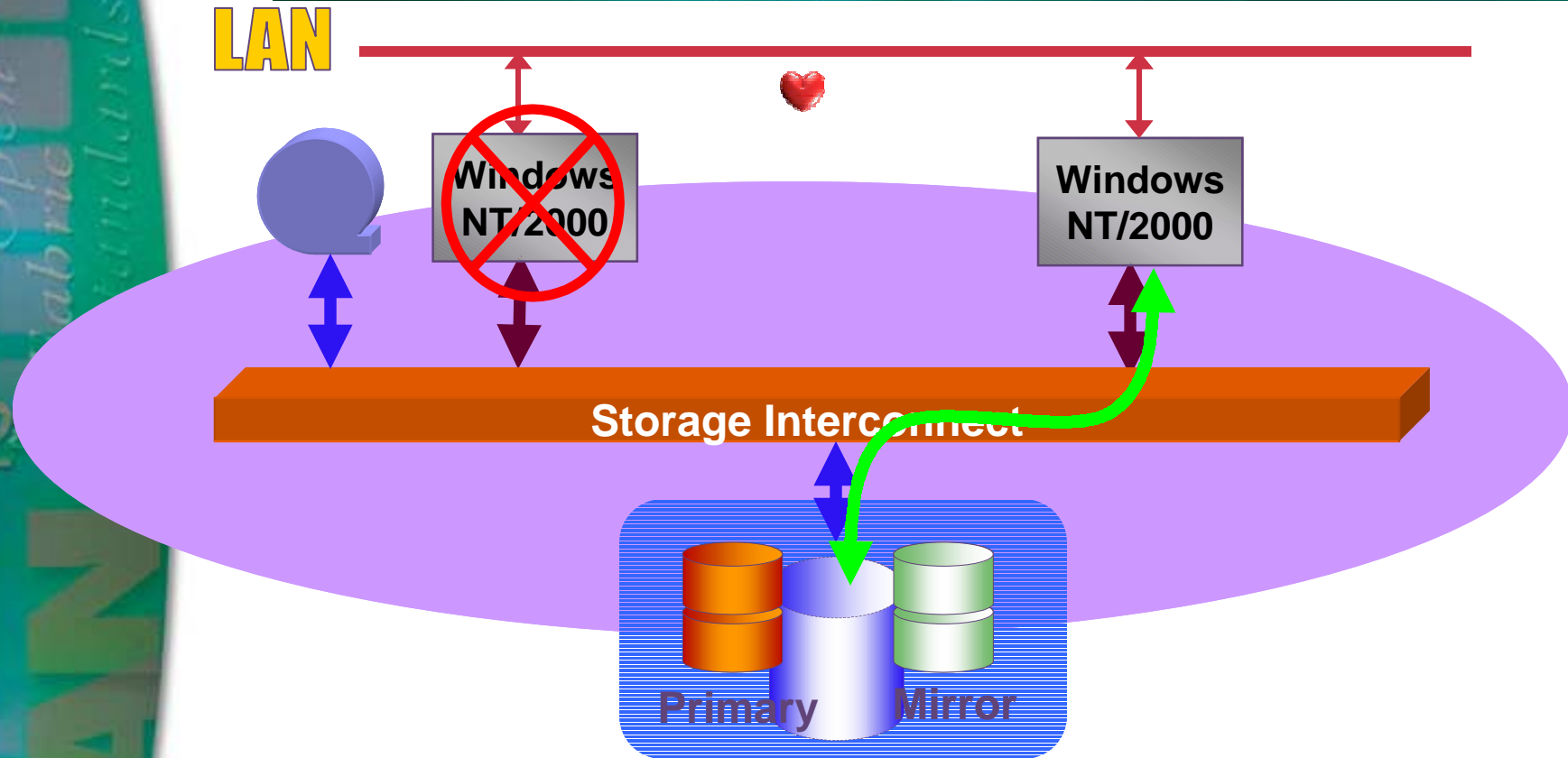
Clients

10% - all HW

Source: IEEE Computer, 1996

Increasing Application Availability

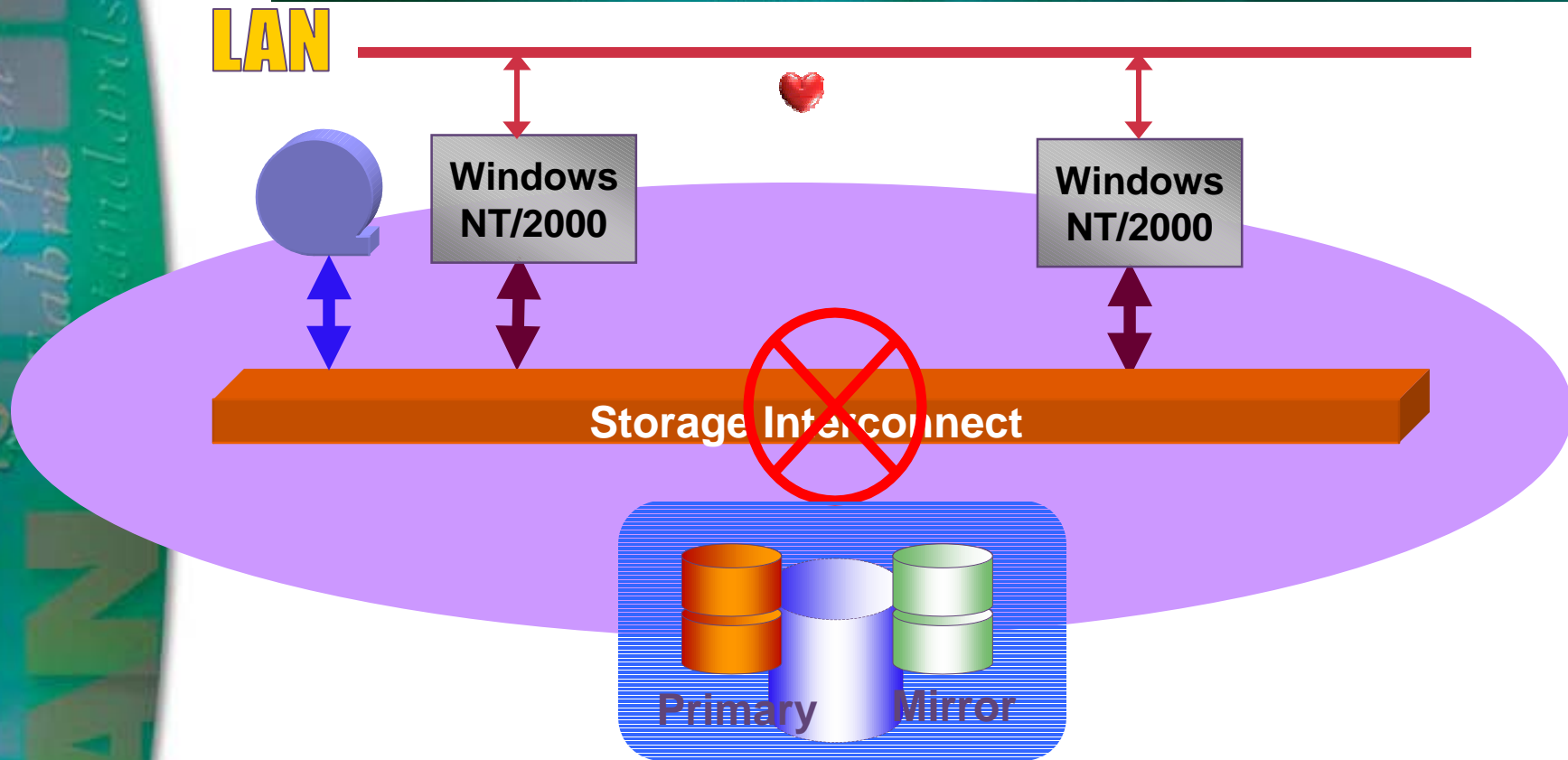
LAN



- ☺ Business Operations protected from Application failure
- ☺ Infrastructure allows cluster to easily scale (2x, 4x, 8x, etc.)

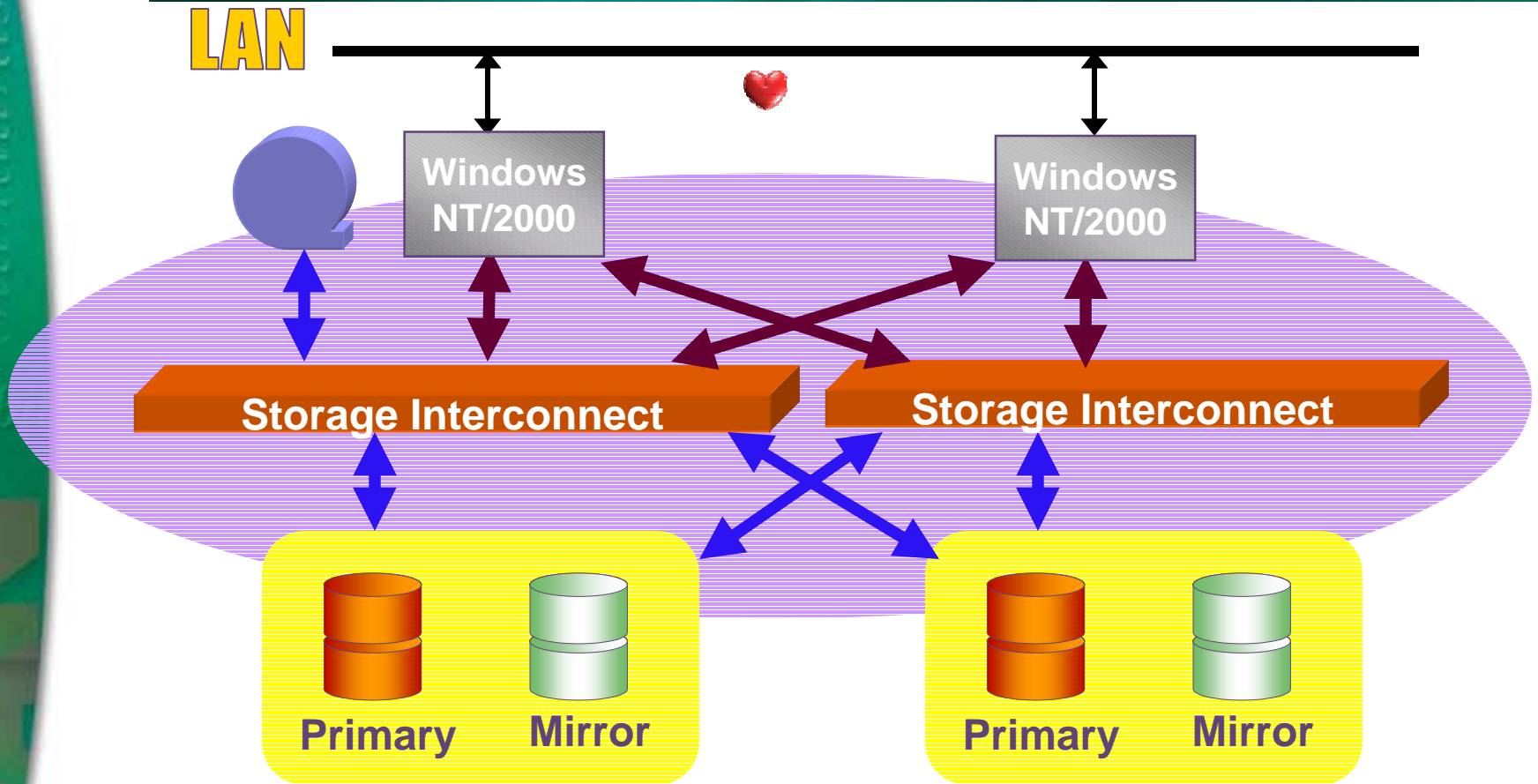
Increasing Application Availability

LAN



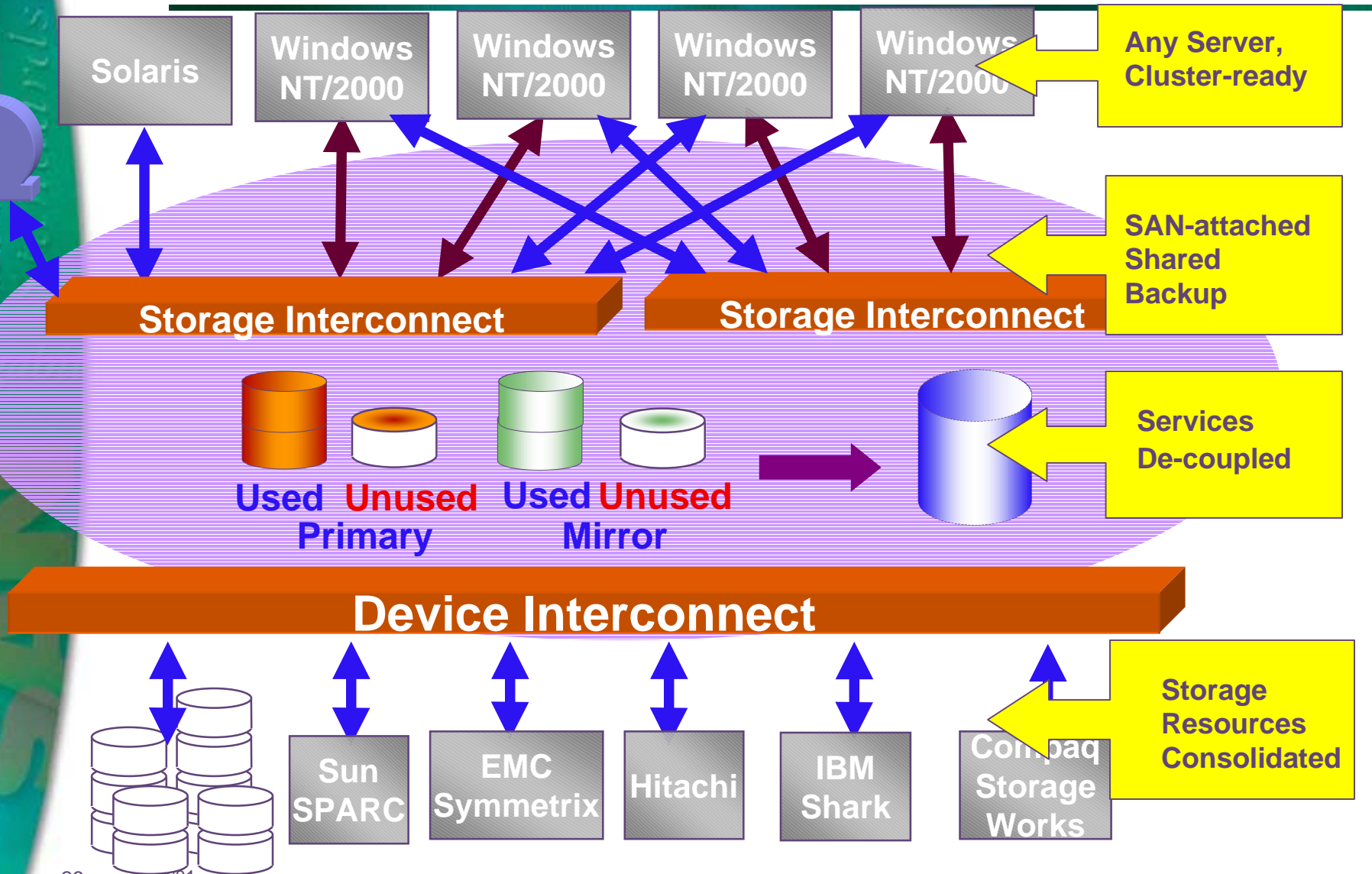
- ☺ Business Operations protected from Application failure
- ☺ Infrastructure allows cluster to easily scale (2x, 4x, 8x, etc.)
- ☹ Infrastructure itself is now single point of failure

Single-Vendor HA Data Cluster

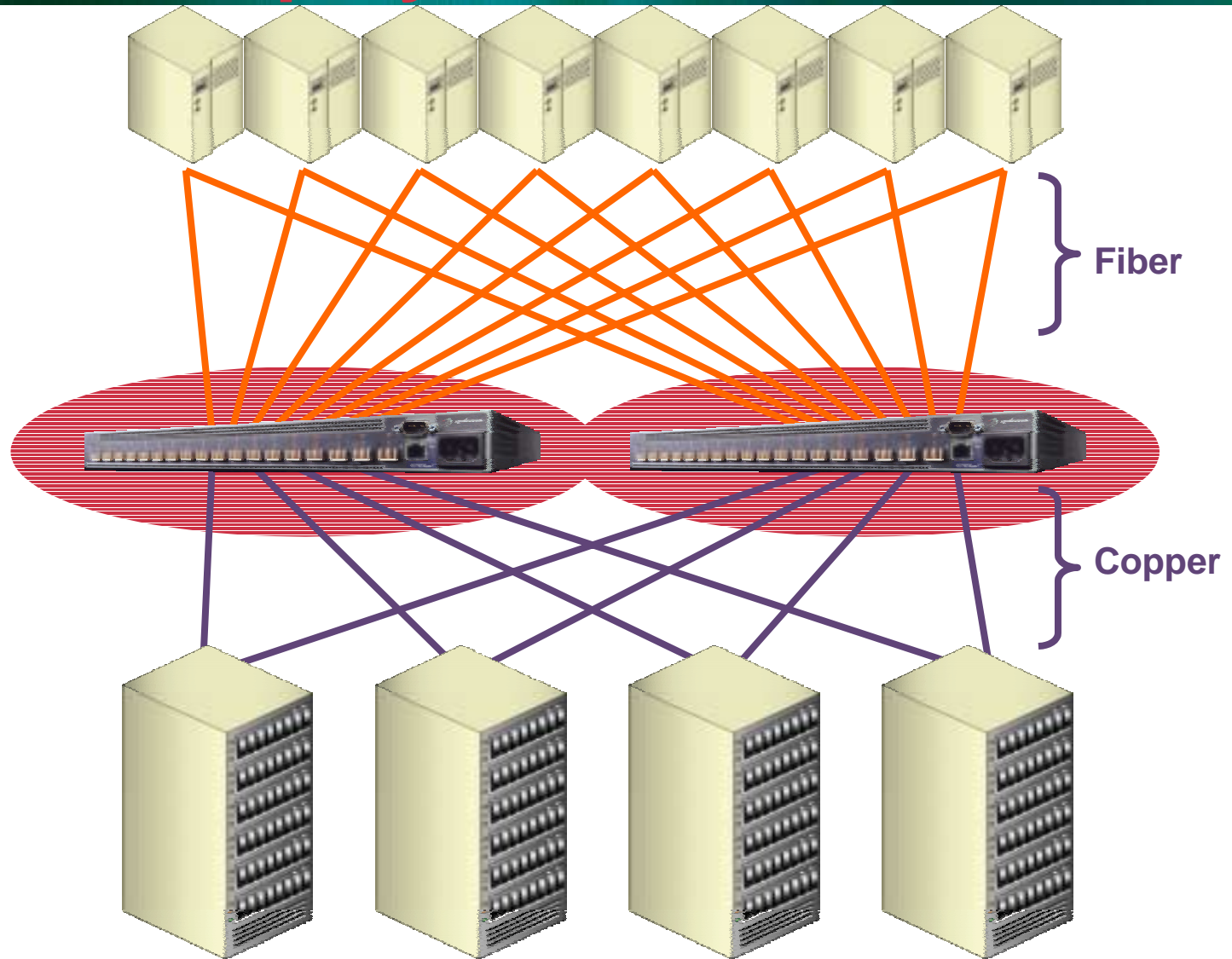


- Has close to 50% Availability upside
- Leverages switched SAN infrastructure

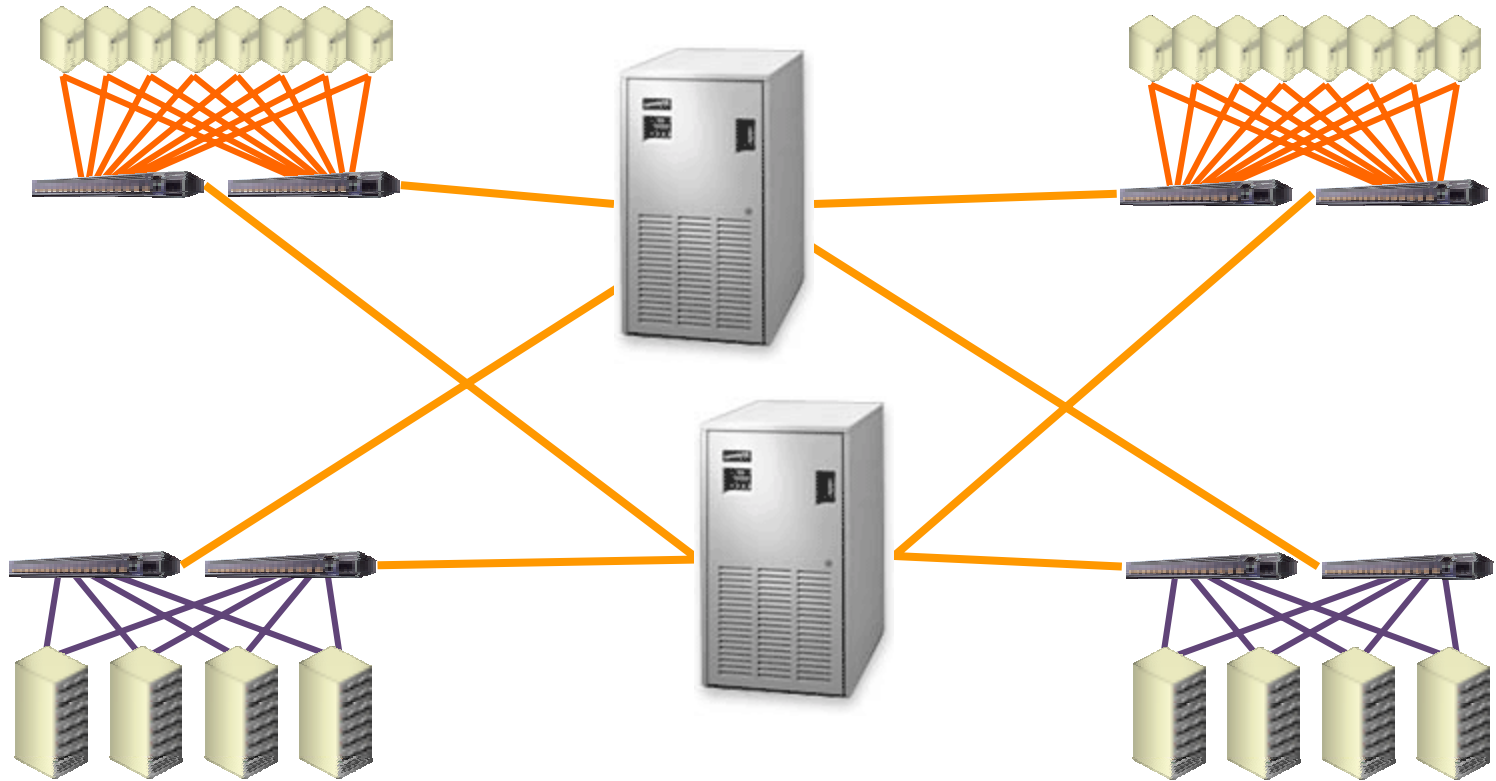
Optimal Server/Storage Interconnect



Typical Business Unit SAN Deployment



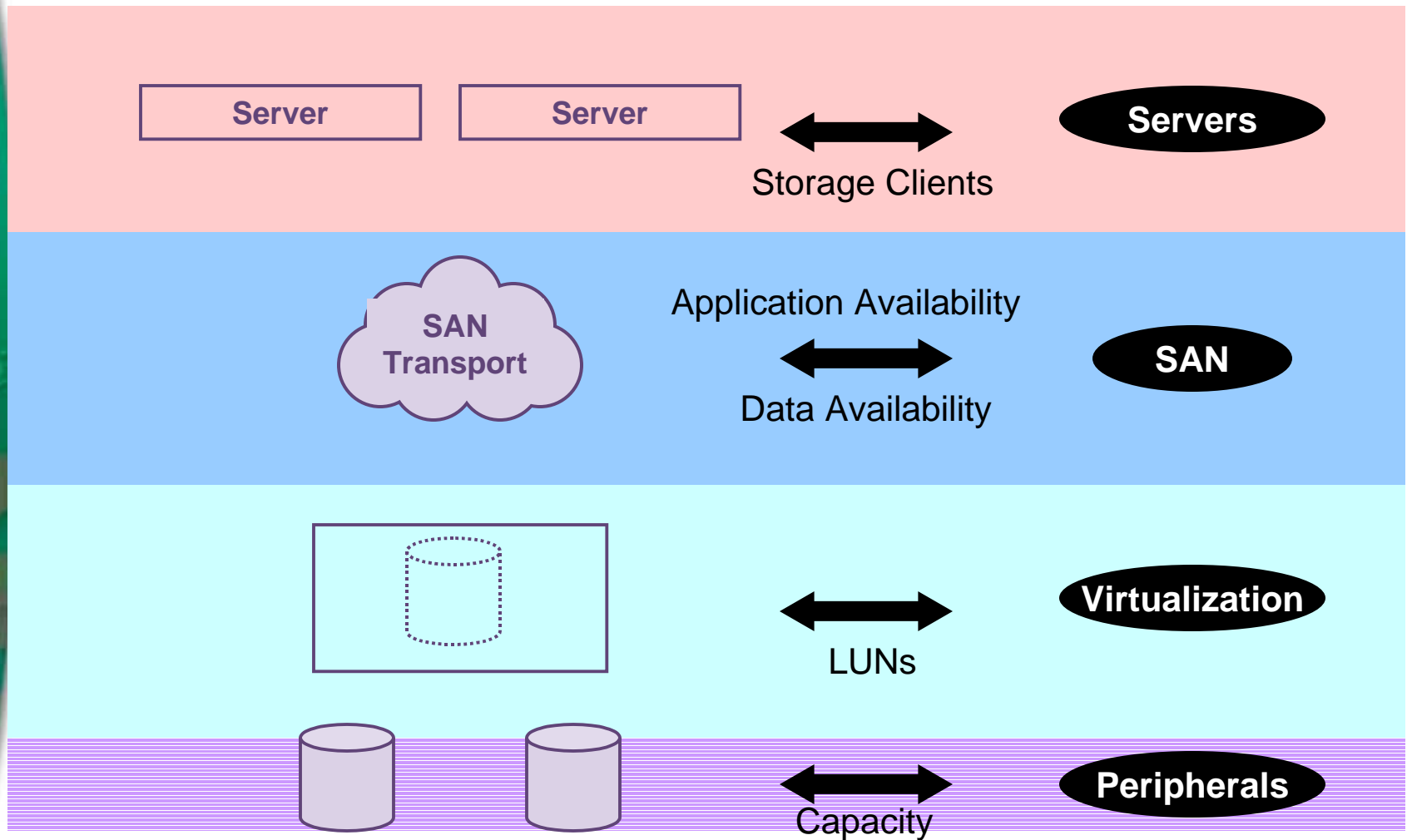
HA Enterprise SAN Architecture



Large Enterprise SAN:

- ➔ Redundant Fabric Edge
- ➔ No single-point-of-failure Director Core
- ➔ Centralize HA Storage Resources
- ➔ Centralize large tape libraries

SAN Scaling



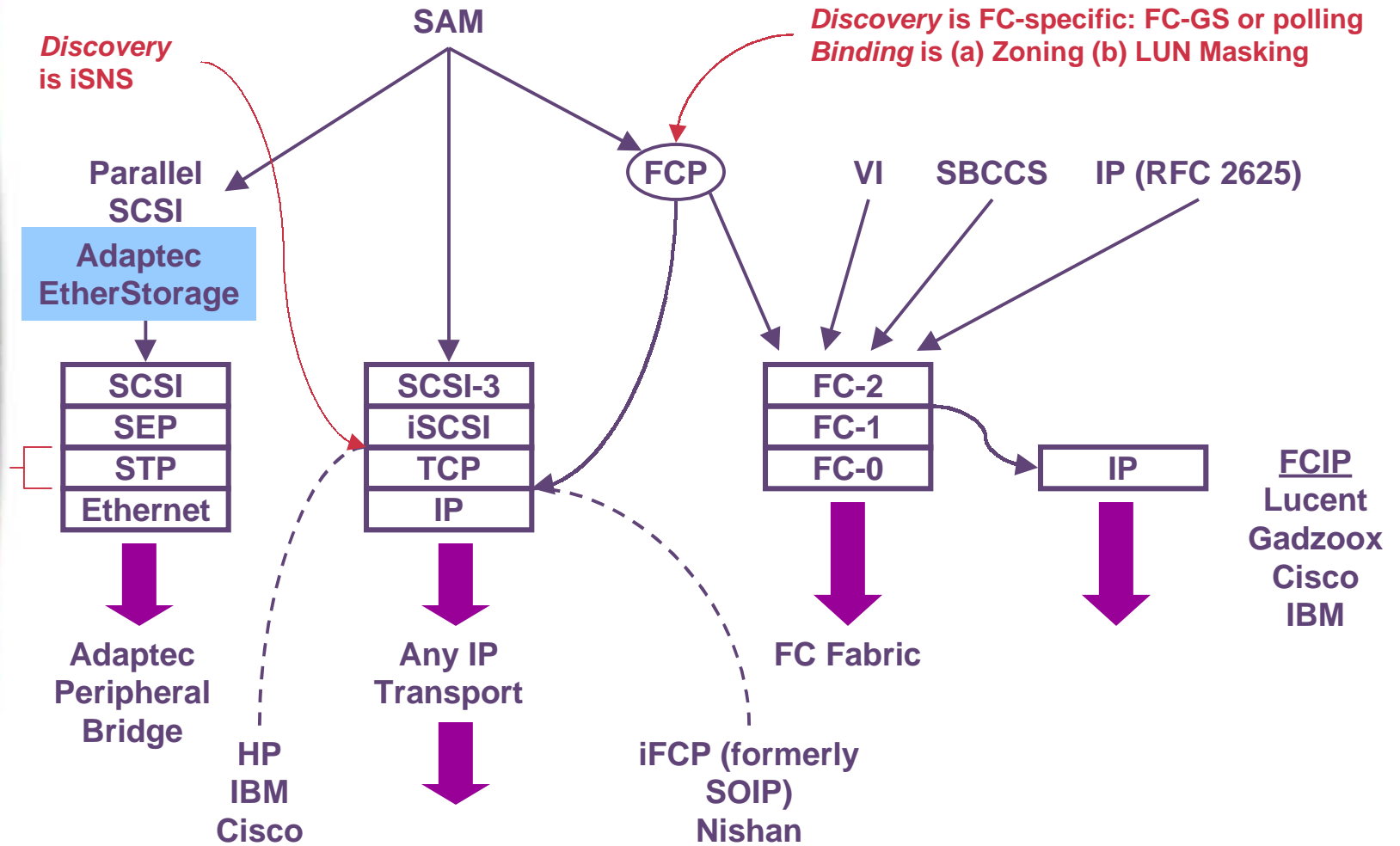
Storage on IP Networks

What is Storage over IP?

FCIP iFCP
iSNS iSCSI



Alternate Storage I/O Stacks



Proprietary

What is FCIP (FC-over-IP)?

- ◆ FCIP is a tunneling protocol for Fibre Channel FC-2 frames
- ◆ Leverages any IP-based network to connect SAN islands over LANs, MANs, or WANs
- ◆ Independent of link-level protocols such as Gigabit Ethernet, Sonet, ATM, or DWDM
- ◆ Standardization effort is under the IPS working group in IETF
- ◆ RFC submitted by Lucent and Gadzoox

What is FCIP (FC-over-IP)?

- ◆ Much new interest in FCIP work. New members of author list include Cisco, Compaq, Rhapsody, Aarohi, CNT
- ◆ FCIP group is working on:
 - Port modeling issues: how the IP port fits into the FC fabric
 - Common encapsulation method for FCIP and iFCP
 - Timeouts and error recovery

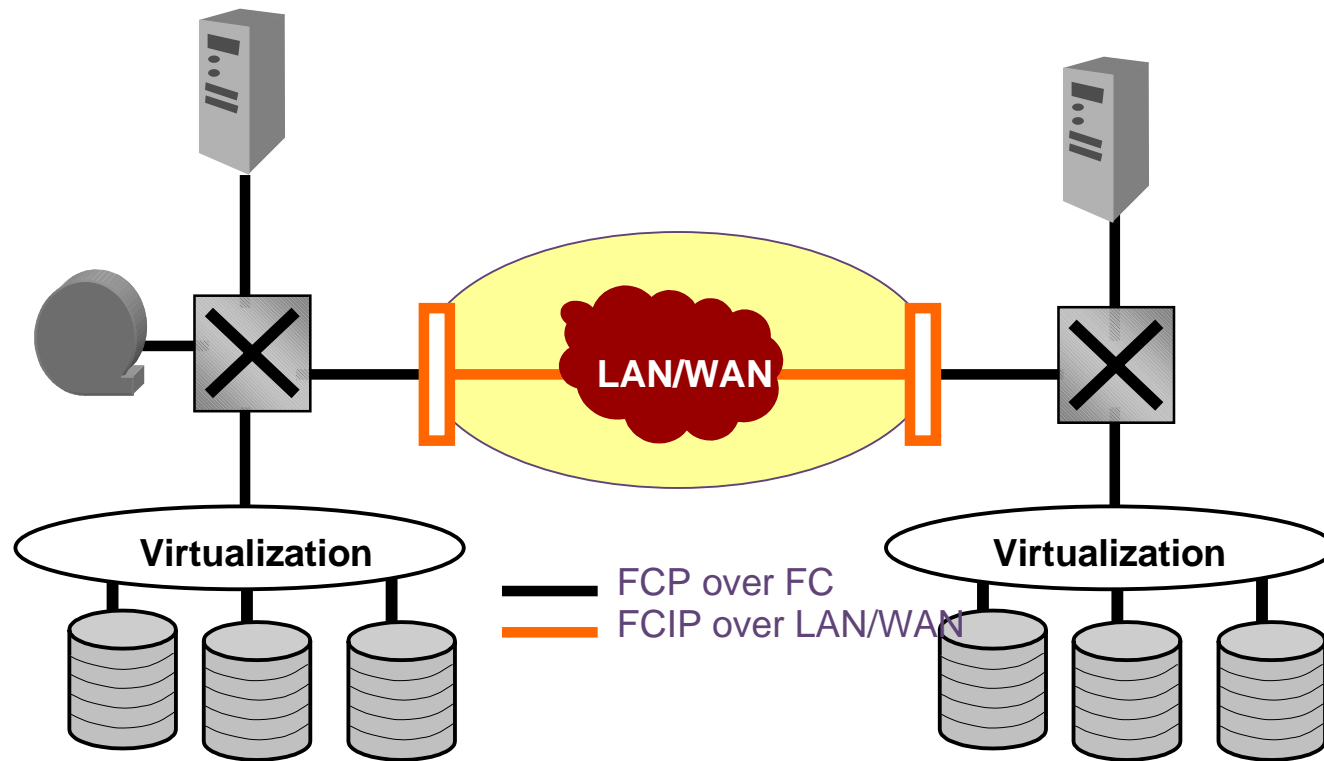
What is iFCP?

- ◆ iFCP uses Fibre Channel only at the edge (servers and storage nodes) then uses FC to Ethernet gateways to access a TCP/IP Switched Backbone
- ◆ Approach assumes LAN Switches are preferable to FC Switches
 - Distance
 - Addressing
 - Familiarity and Ease of Management
 - Trunking and QoS
- ◆ Requires a new Name Service (iSNS)
- ◆ Supported by Nishan

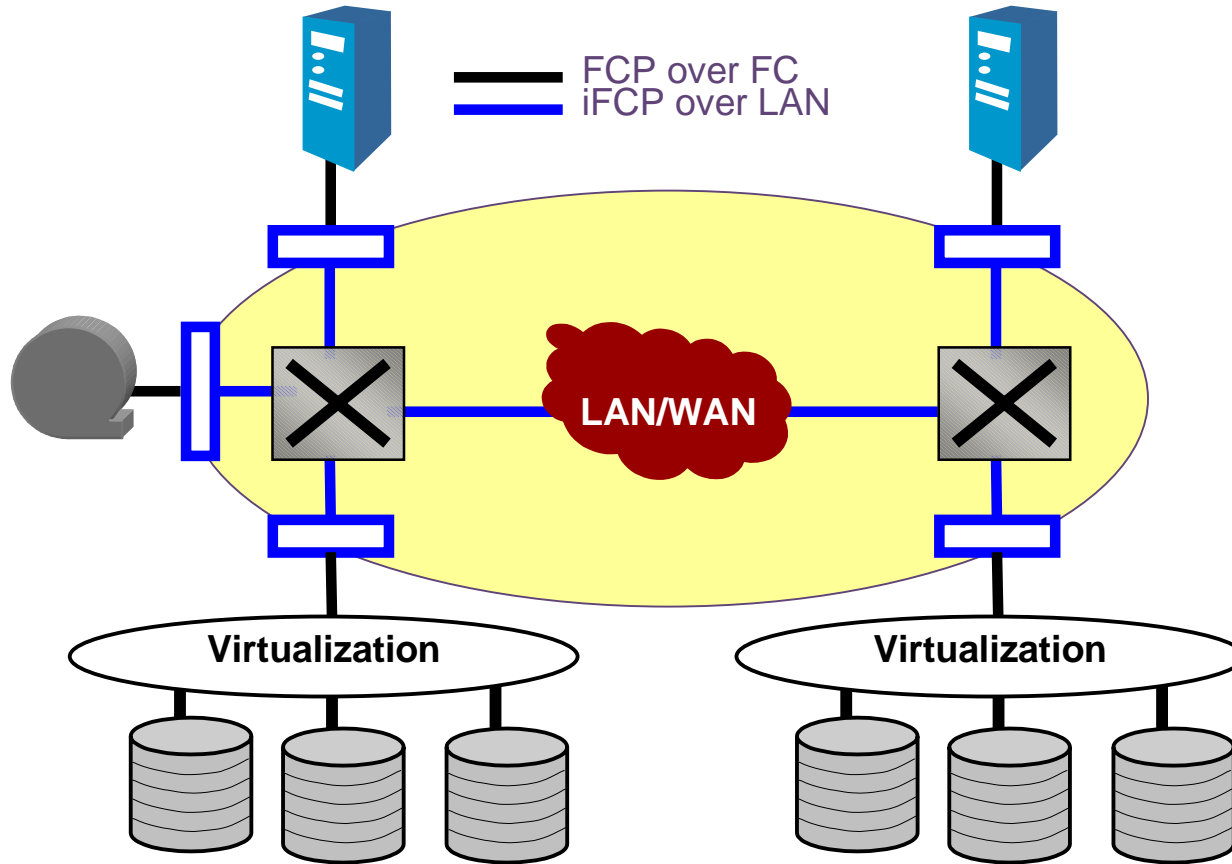
What is iSCSI?

- ◆ iSCSI maps the SCSI-3 protocol directly to TCP/IP
- ◆ Anticipates the arrival of Ethernet NICs and drivers that can support server-side block I/O
- ◆ Local-area Storage networks will be built using Ethernet switches
- ◆ Standardization effort is major task of IPS working group under IETF
- ◆ Supported by HP, IBM, Cisco, all major Networking companies
- ◆ Strength: Heavyweight supporters and Ethernet branding
- ◆ Weakness: Poor performance without changes to Ethernet NIC and TCP/IP stack. Needs new storage drivers for each OS.

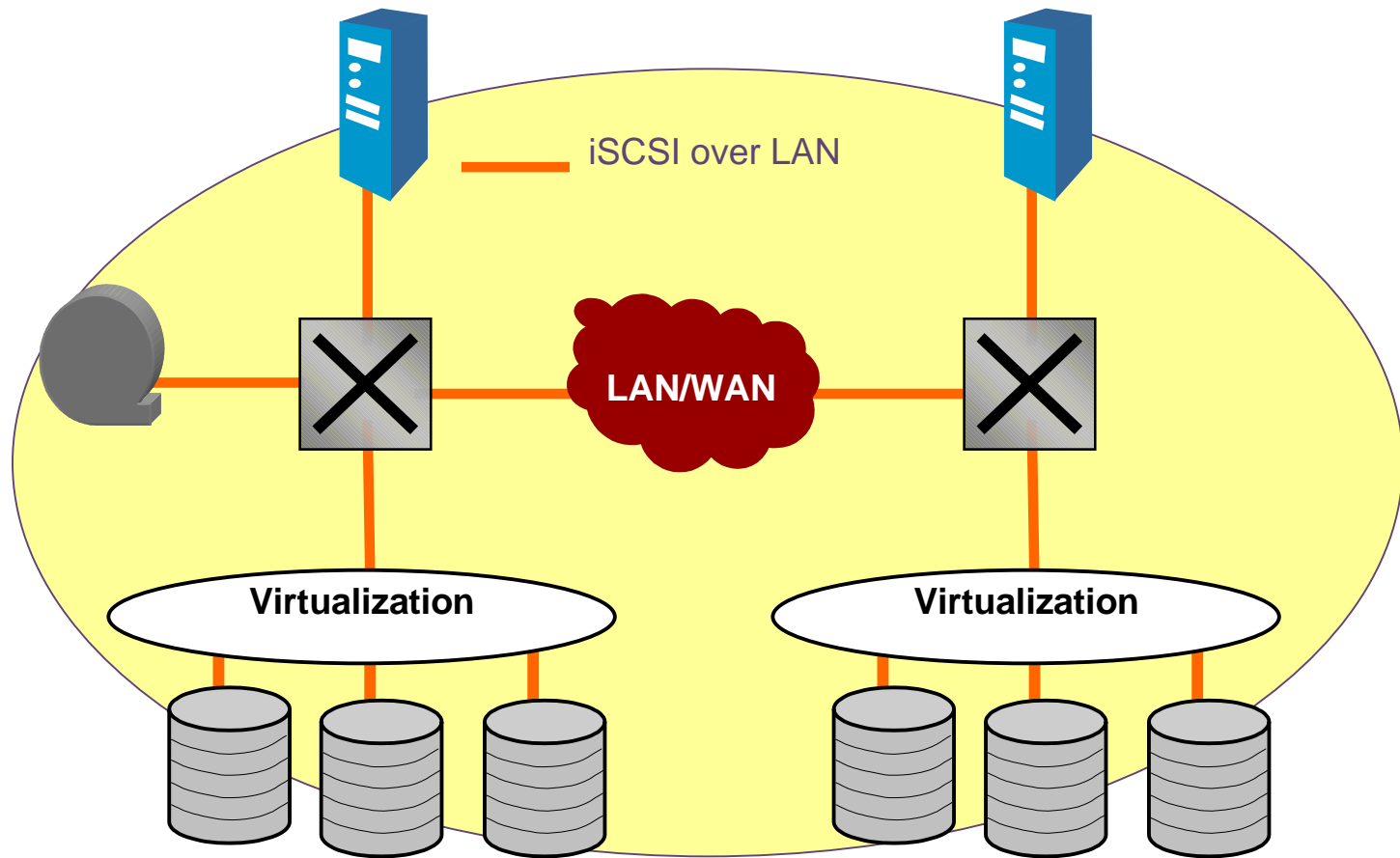
FCIP (FC over IP)



iFCP

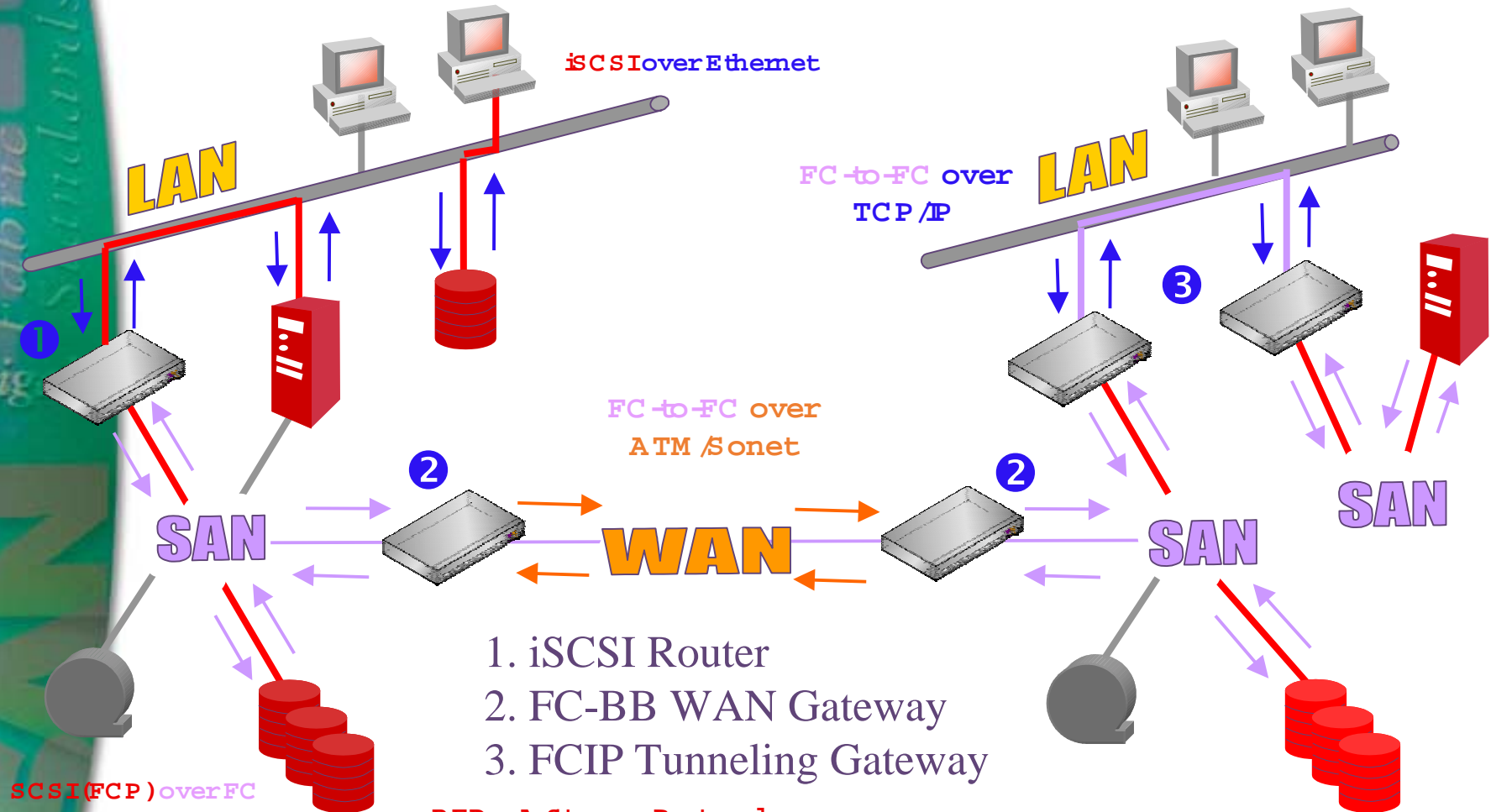


iSCSI



A Transport-Independent SAN

Storage Freedom
 Gigabit Ethernet
 Scalable SAN

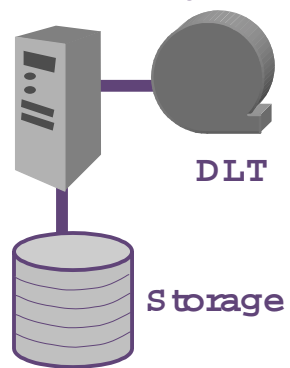


1. iSCSI Router
2. FC-BB WAN Gateway
3. FCIP Tunneling Gateway

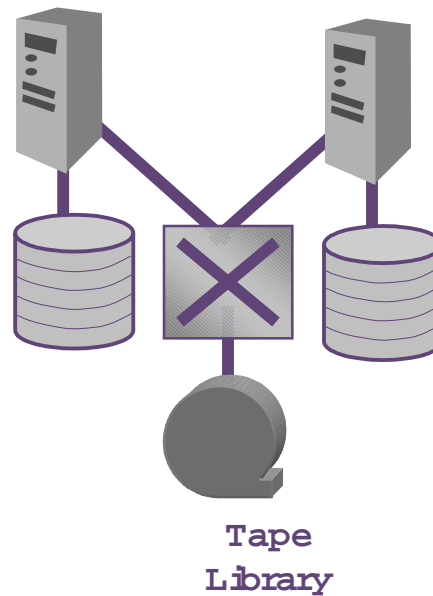
RED - A Storage Protocol
 BLUE - TCP/IP Protocol, Ethernet Transport
 VIOLET - FC Frames, FC Transport
 ORANGE - WAN Frames, WAN Transport

Summary

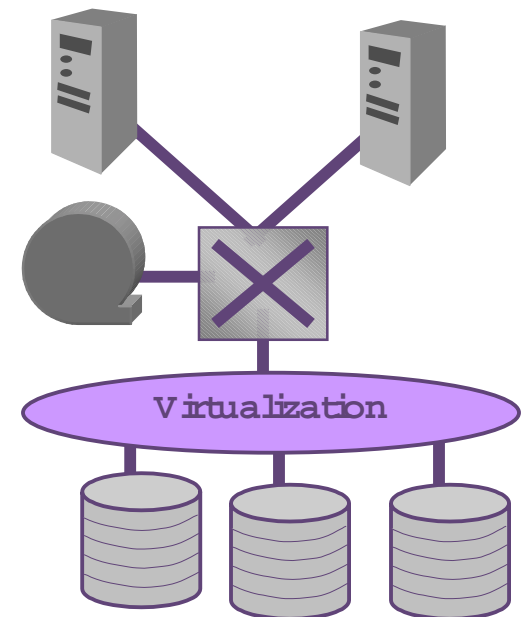
Today



Backup

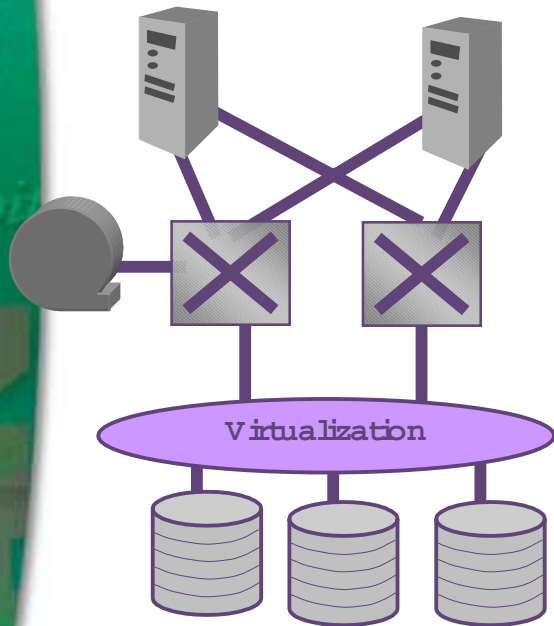


Consolidation

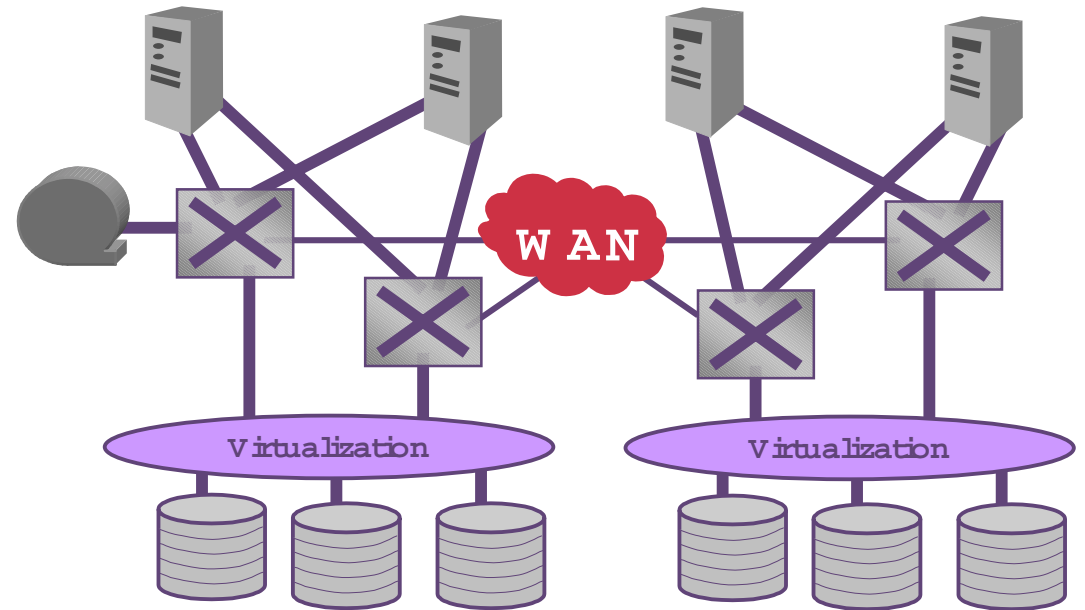


Summary

Cluster



Remote Mirror



What is “Clustering”?

Depending on who you ask...

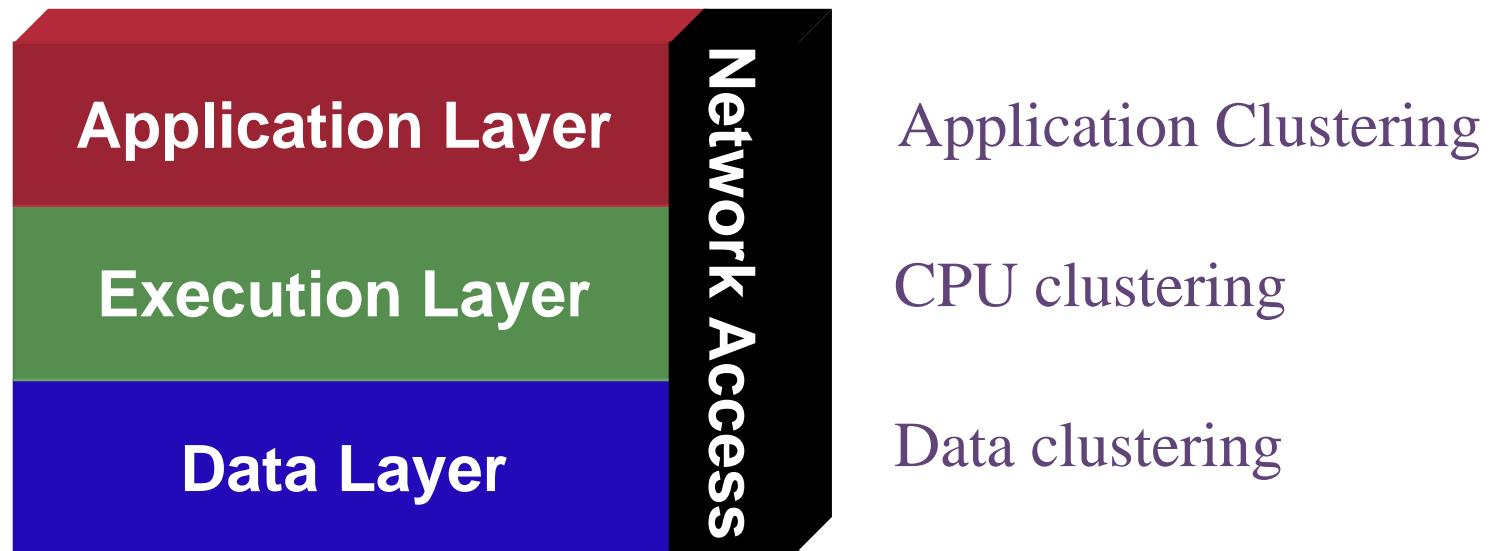
Compute
Clusters

Data
Clusters

Application
Clusters



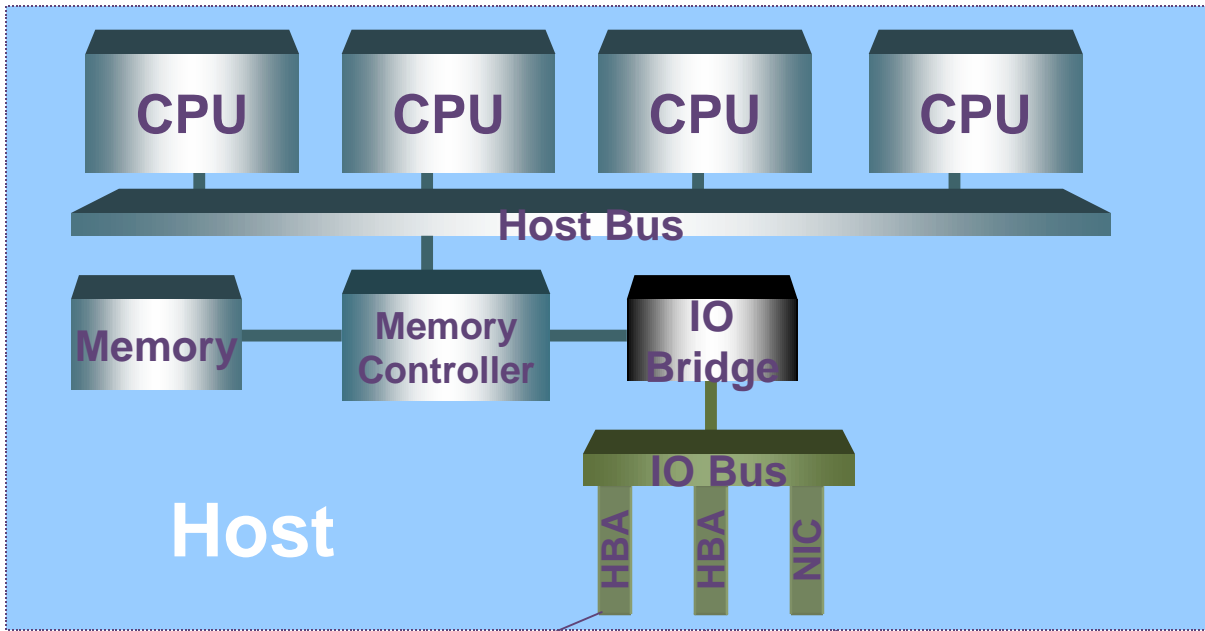
What is Clustering?



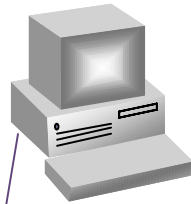
These Three Things are Distinctly Different!

Current Interconnect Configuration

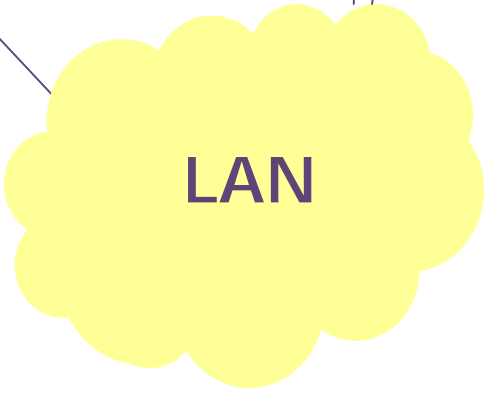
3
Networks



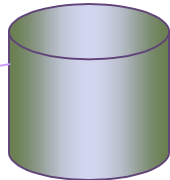
...and application clustering lives here.



Compute clustering lives here...



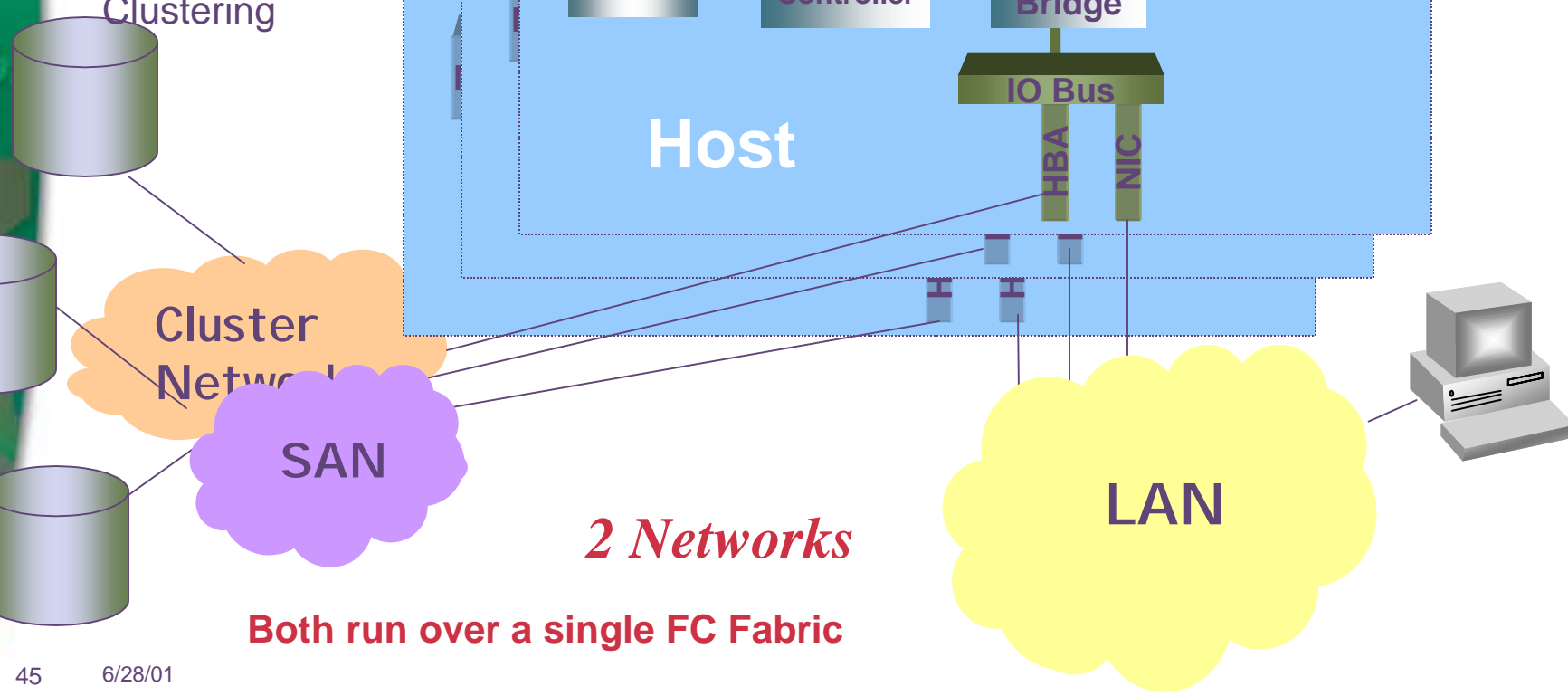
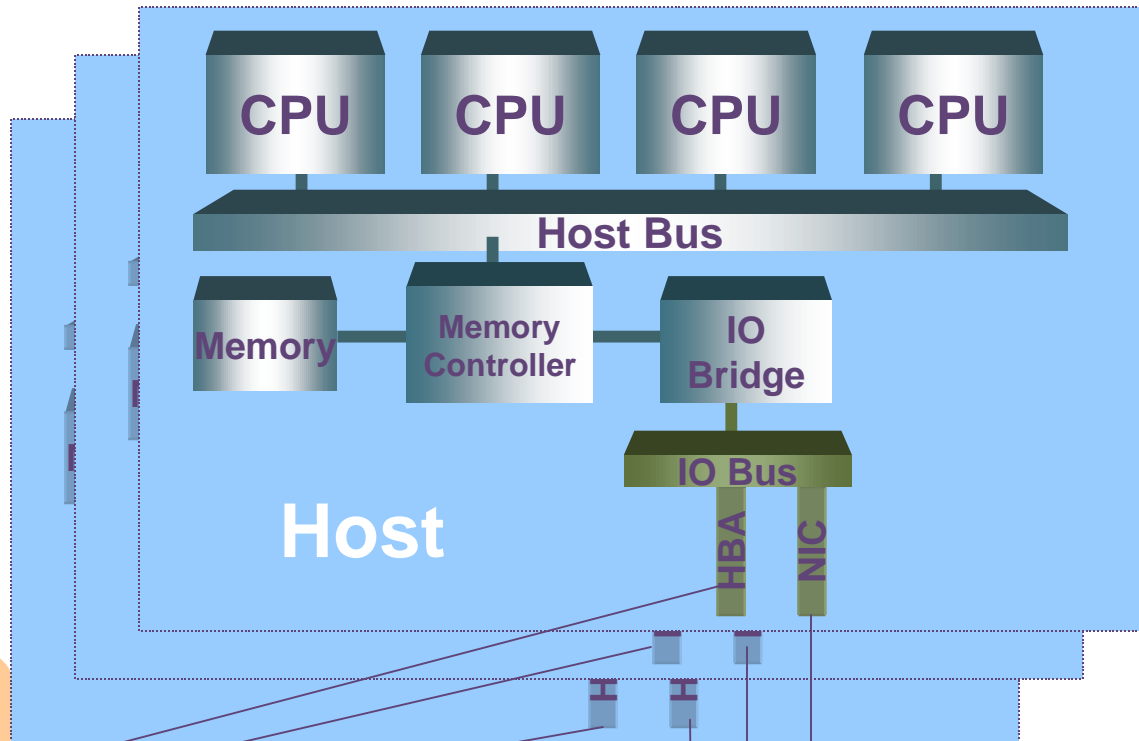
Data Clustering lives here...



FC Clustering Configuration

Use **FC-VI** for
Compute-
Clustering

Use **FCP** for
Data
Clustering

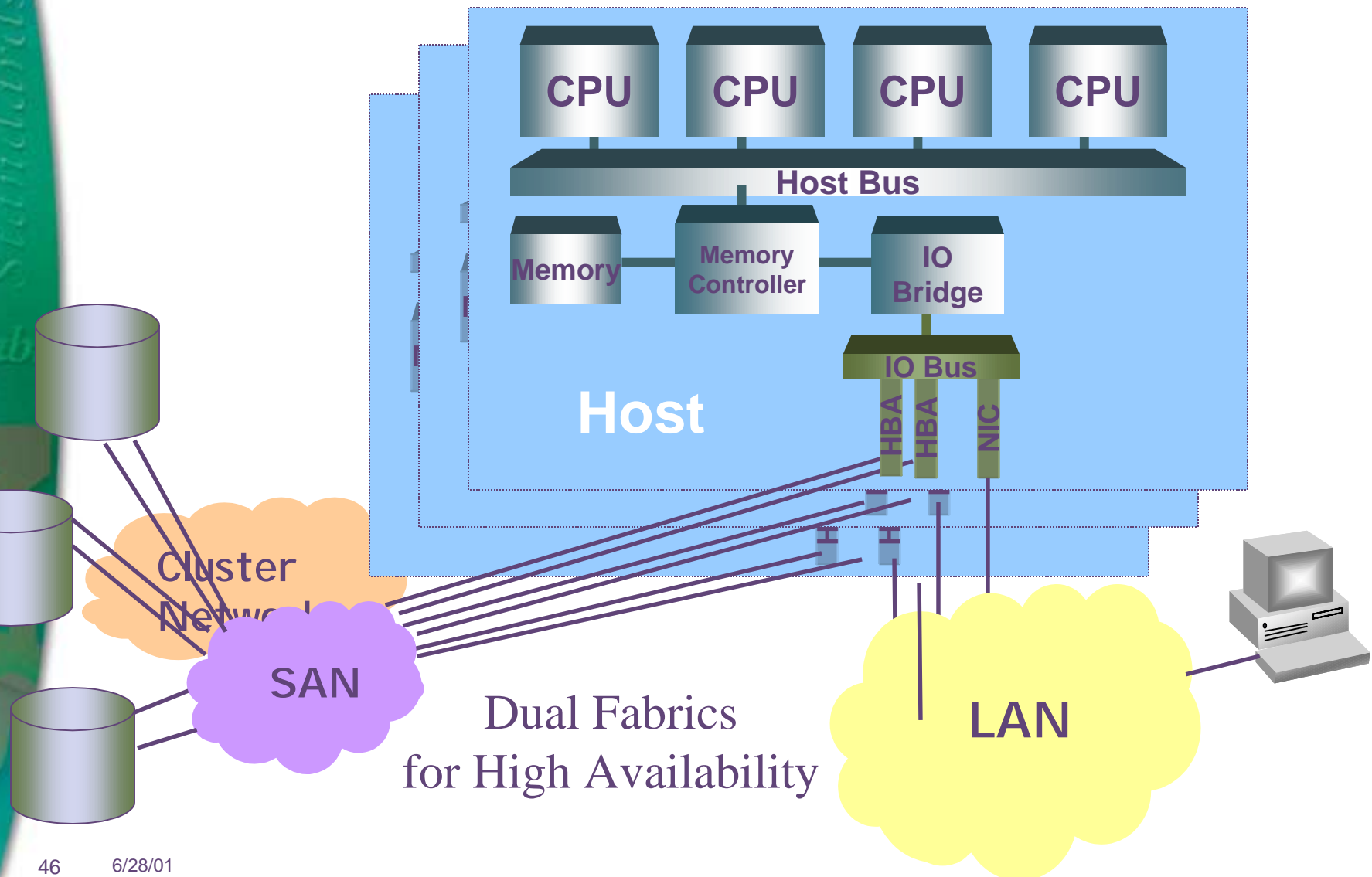


2 Networks

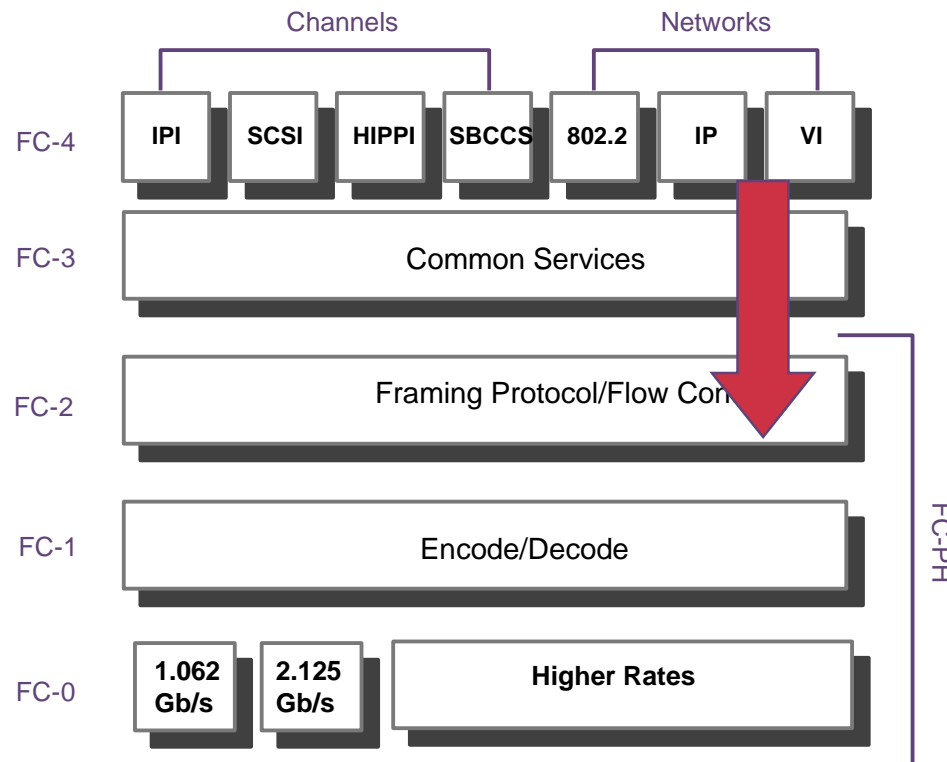
Both run over a single FC Fabric

FC Failover Configuration for DATA

Storage Freedom
Fibre Channel Standards
Gigabit Ethernet
Scalable



VIA over Fibre Channel



◆ VI compliant HBA's

- Pre-Built Headers
- Single Frame Sequences
- HW queues

◆ No FC switch impact

- 2-level priority provides QoS for FC-VI traffic

Clustering with Fibre Channel

◆ Strengths

- Usability
- Good Fault Tolerance Model
- Multi-protocol Support
- It already Exists
- Big effort in VI support
- Commodity Pieces
- Basis for Storage Area Networks (SAN)
- Riding the performance curve up

◆ Weaknesses

- Fibre Channel is feature rich
- No low level interrupts
- No sync Mem<->Mem
- Not Path/Attribute Aware
- Needs priority in mixed protocol environment
- Proprietary solutions exist

**InfiniBand vs.
Fibre Channel:
Only one winner?**

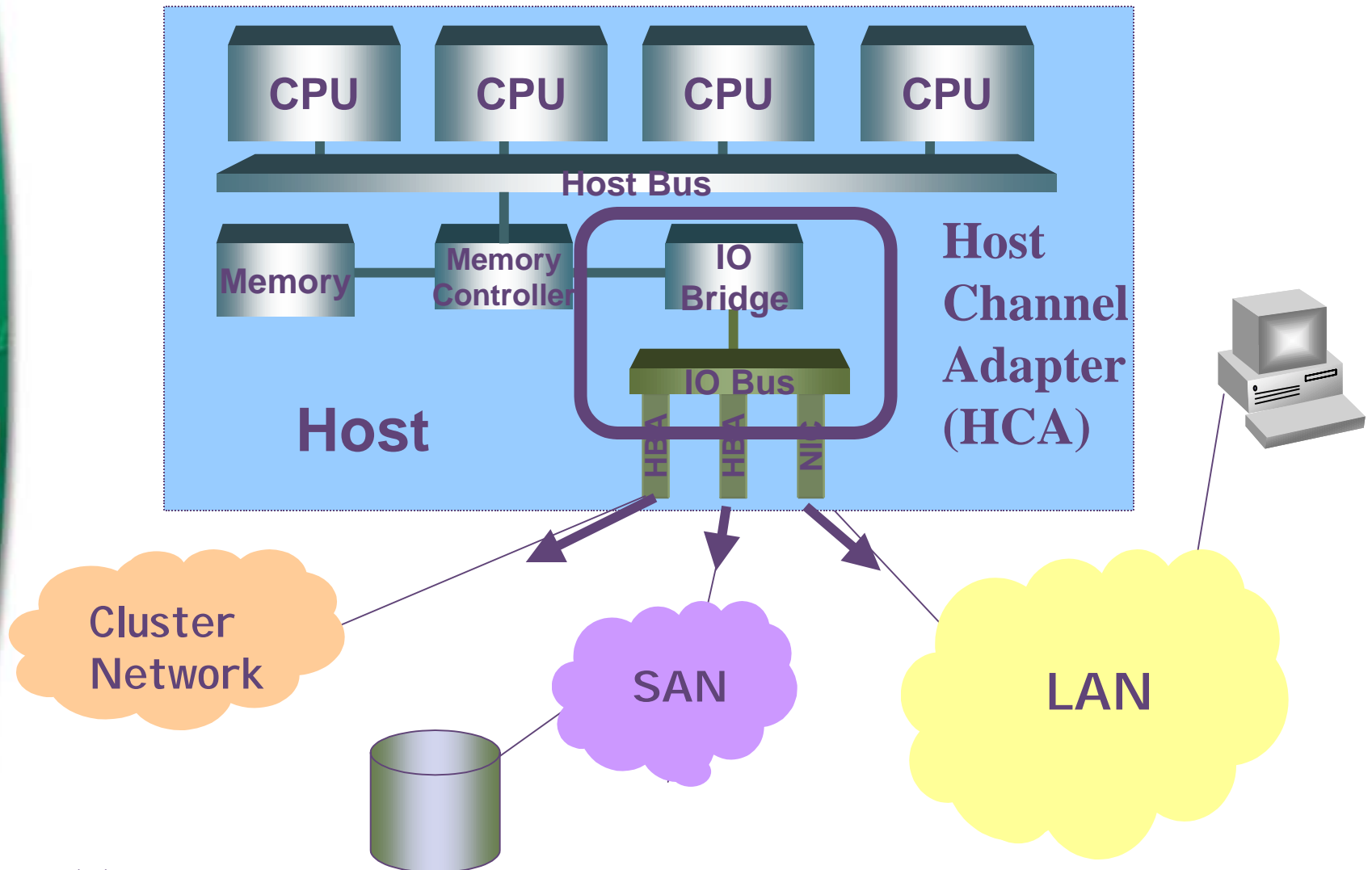




Important Questions

- ◆ Is InfiniBand likely to replace Fibre Channel in:
 - Compute clusters?
 - Data clusters?
 - Primary Server connections to storage?
- ◆ When will the technology be deployed?

Driven by Server Design Constraints



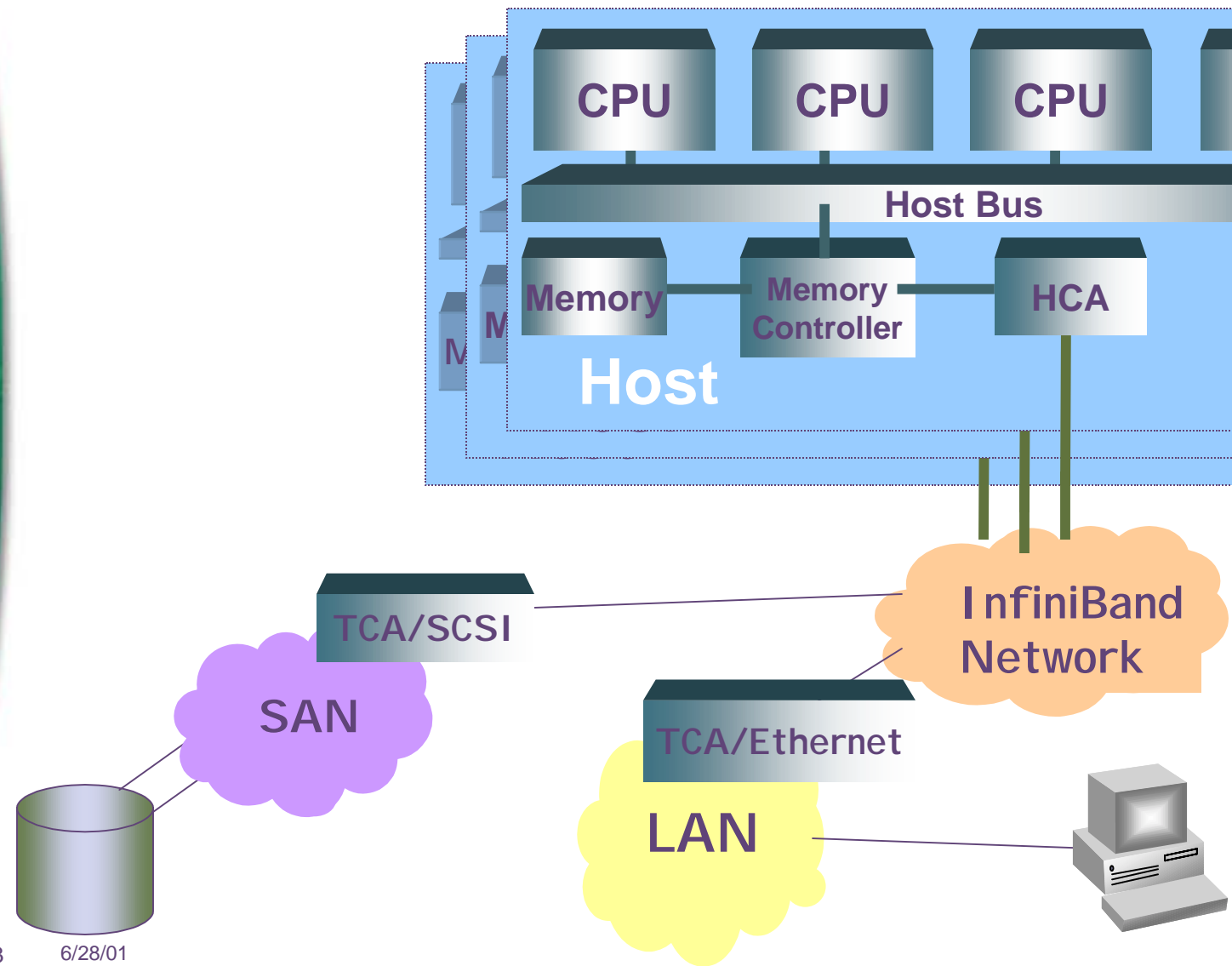


A New Clustering Network

- ◆ Parallel (PCI) Bus replaced. A New Network is inserted in Series
- ◆ Splits existing HBA control layers between Server and Peripheral
- ◆ I/O Bridging operations pushed into the existing Networks
- ◆ Separates I/O bridging operations from Server

Storage Freedom
Open Fabric Standards
Gigabit
Scalable SAN

An InfiniBand Network



InfiniBand

◆ Strengths

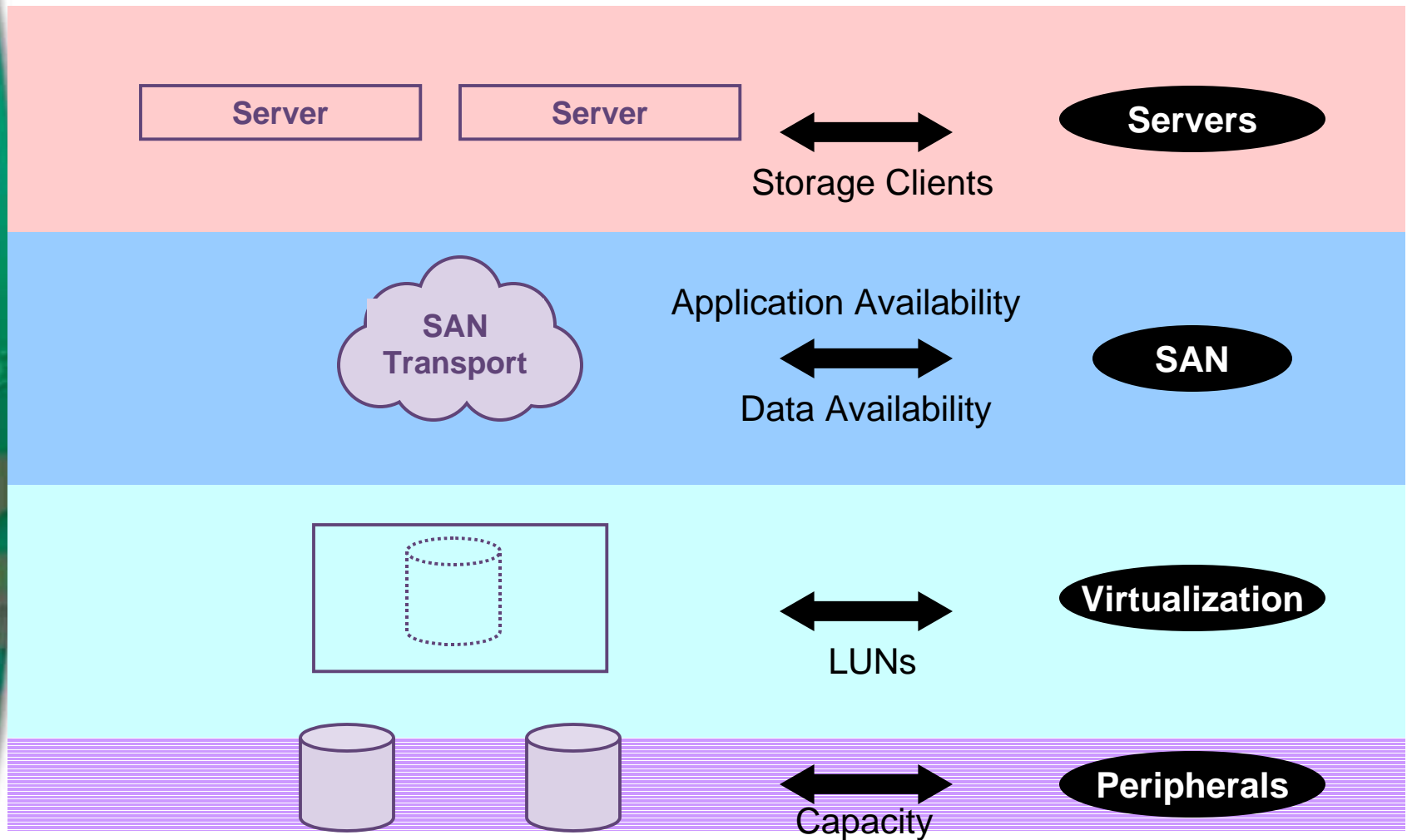
- Good Clustering Interconnect
- Path/Attribute aware
- Good Discovery Model
- Multi-tier Credit Model
- Eliminates Bus
- Broad industry support
- **A new market!**

◆ Weaknesses

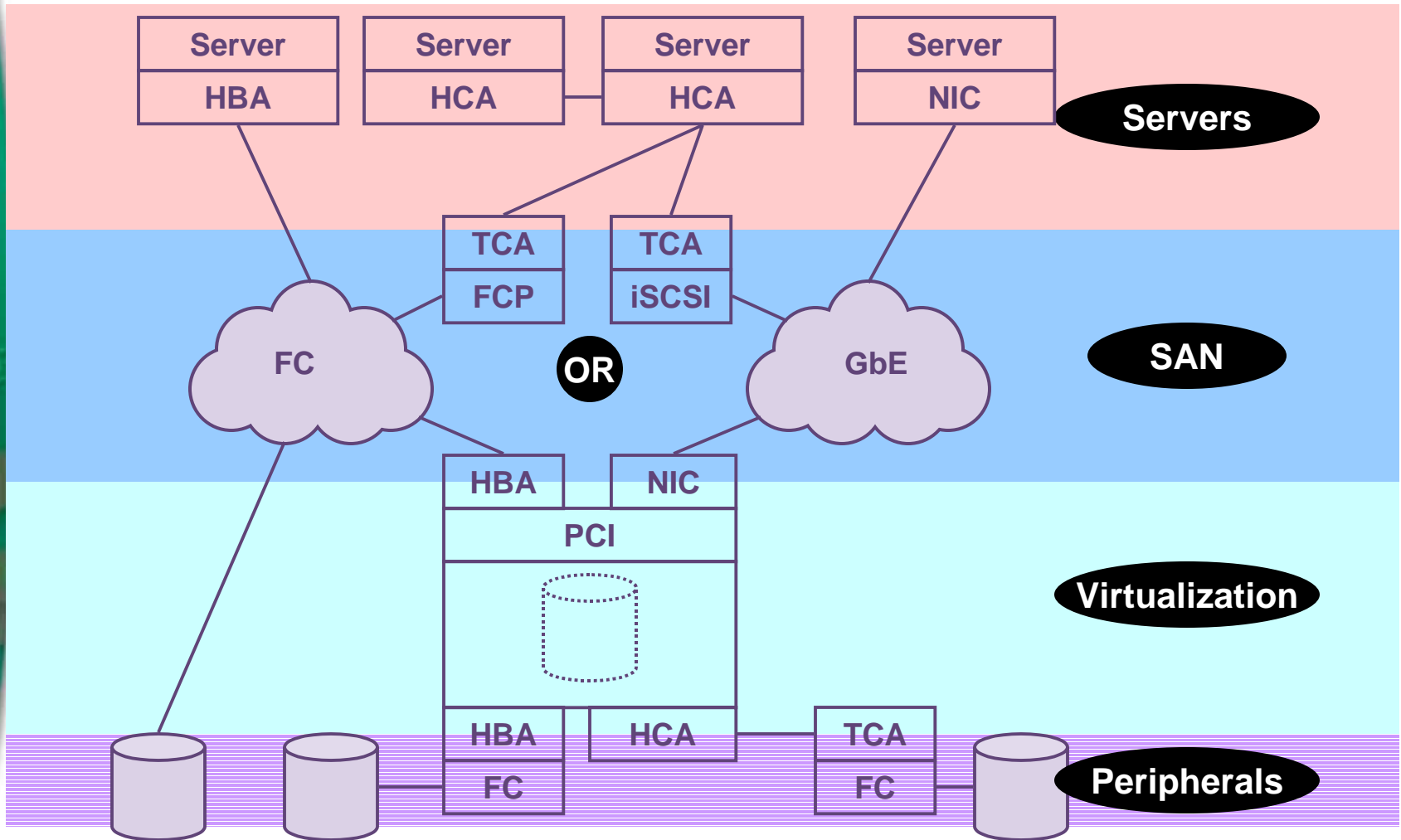
- No storage support
- Not open yet - will have competing implementations
- Market interest is at 4x link rate

In pact on FC: Diminishing importance of FC-VI for *Compute*-Cluster Interconnect. FC remains important for Dual-hosted *Data Clusters*

SAN Scaling



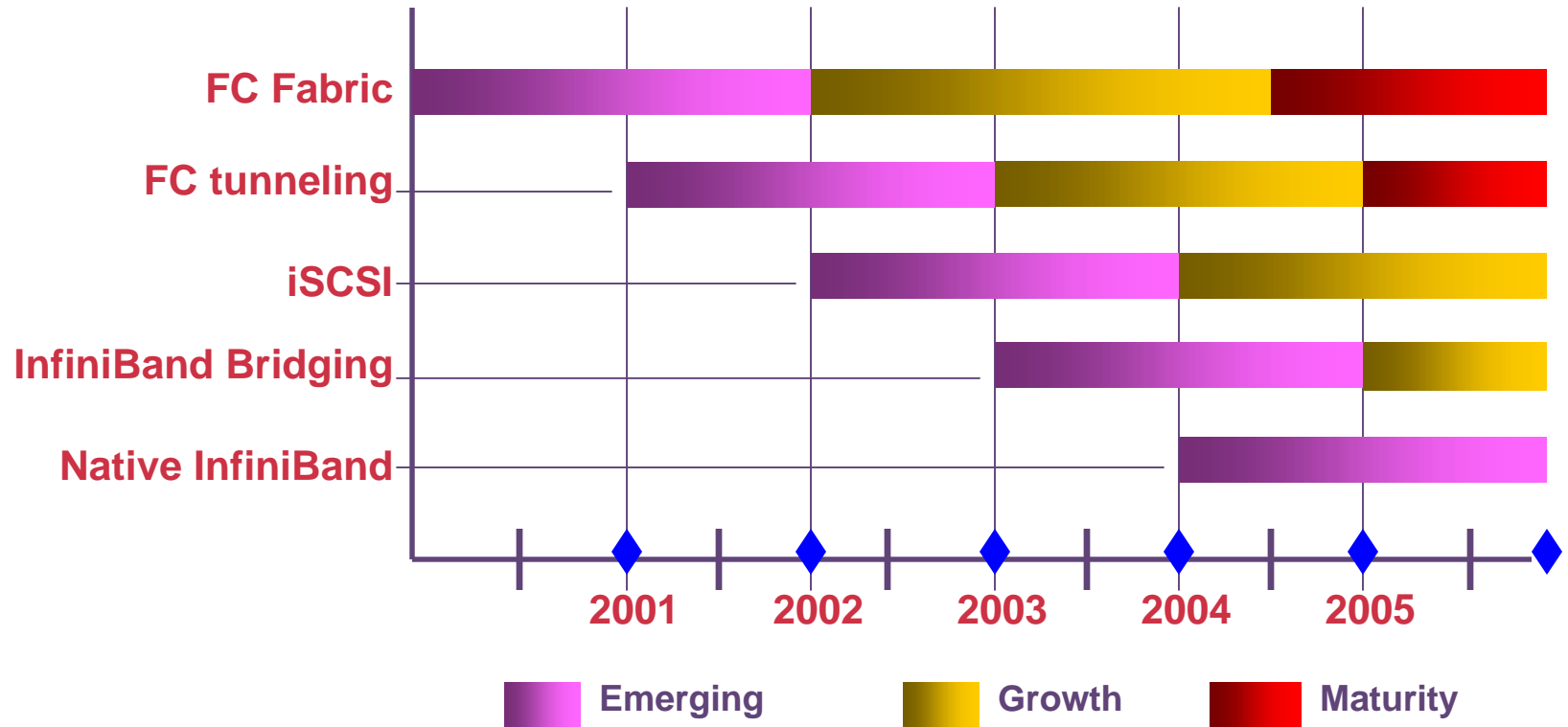
Topology Independent SAN



Storage Freedom
Open Fabric Standards
Gigabit
Scalable SAN

Gadzoox View on Protocol Adoption

Fibre Channel Dominates in Near-Term



2gigabit

Storage
Freedom



gadzoox[®]
NETWORKS

Open
Fabric
Standards

Thank you

Scalable