Concepts In Storage Area Network Design

- Jim Gursha, President
- High Performance System Solutions, Inc.
  - 1735 York Avenue, 32H
  - New York, NY 10128

- HP Interex West Symposium
  Session # 3148 March 2004

HPSS
1735 York Avenue, 32H
New York, NY 10128
212-831-0291/917-359-2087
jimgursha@sandisks.com

Topics

- Storage Area Network Rationale On A Departmental and Enterprise Basis.
- Consolidation Efficiencies of A Storage Area Network.
- Server Effectiveness
- Storage Savings
- Fiber Channel SAN's
- SAN Fabric
- Fabric Benefits
- Basic Fabric Design
- Switch Interconnectivity
- SAN Based Backup/Restore.
- Network Considerations For SAN's.
- High Availability and Disaster Tolerant SAN Considerations.
- HP SAN's Virtualization Concepts, Differences and Considerations.

SAN Rationale
Departmental Vs Enterprise

Consolidation Efficiencies

What is Consolidation?

- Server and storage consolidation is a process gaining greater efficiencies and simplicity in your IT infrastructure. This process involves evaluating your IT environment and implementing a strategy of combining servers and storage to address your business requirements for:
 - Reducing costs
  - Increasing manageability
  - Increasing availability
  - Increasing performance
 - Increasing flexibility

Consolidation Candidates

- Candidates for server and storage consolidation are companies which are experiencing:
    - Recent mergers or acquisitions and need to integrate their IT infrastructure
    - High management costs for large numbers of servers and storage systems
    - Loss of or reduced support staff and high IT staff training costs
    - Older servers and storage systems which need to be replaced or come off lease
    - Need to replace outdated and non-standard platforms with a corporate standards

Storage Connect Architectures

Application

File System

Storage

Direct Attach Storage

(DAS)

Application

Local Area Network

File System

Storage

Network Attach Storage

(NAS)

Application

File System

Storage Area Network

Storage

Storage Area Network

(SAN)
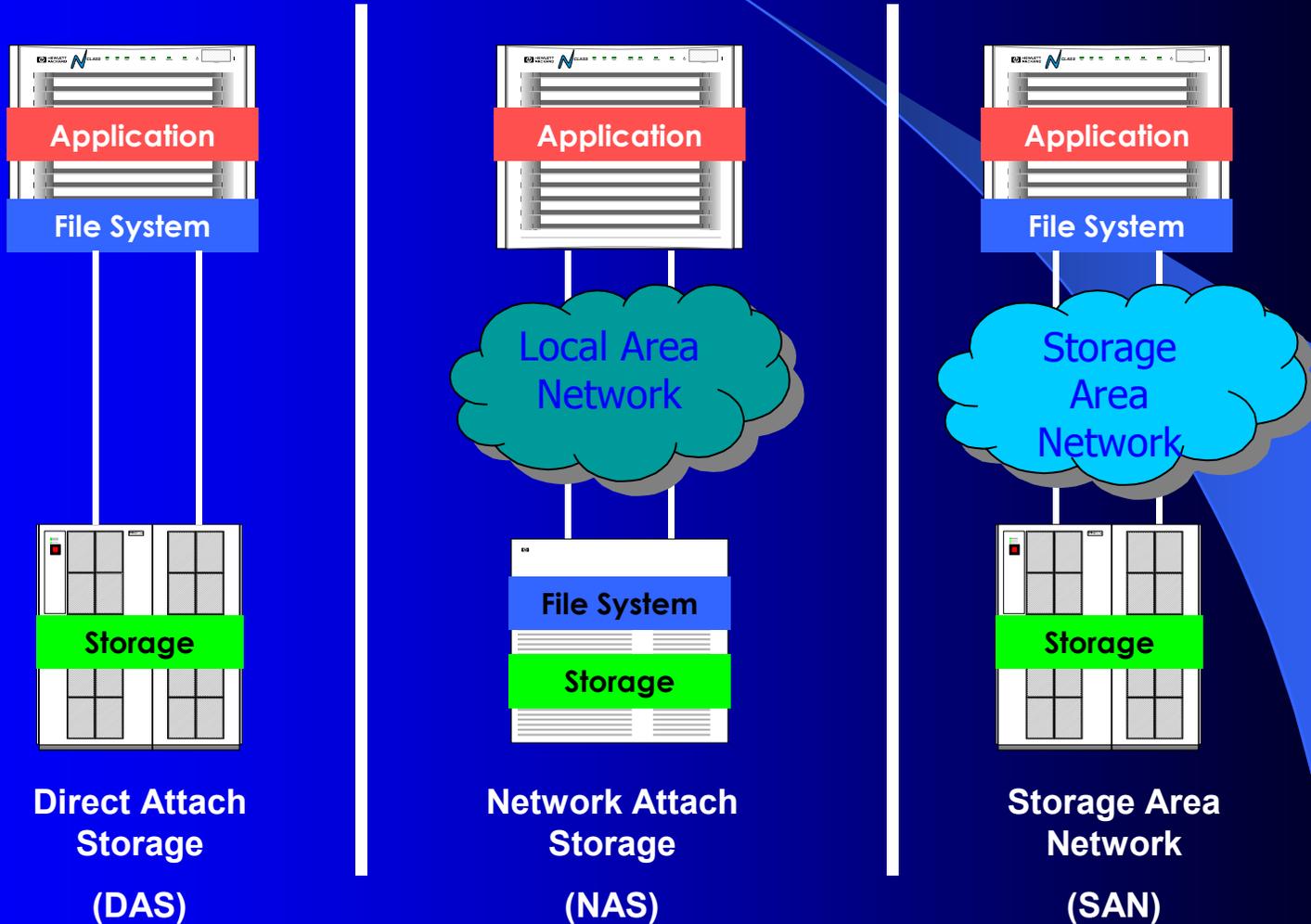
Example DAS Server Storage (No Consolidation)

LAN USERS

Server   Server   Server   Server

Disk
capacity
utilization

Internal
storage
FULL!

Add
a new
Server
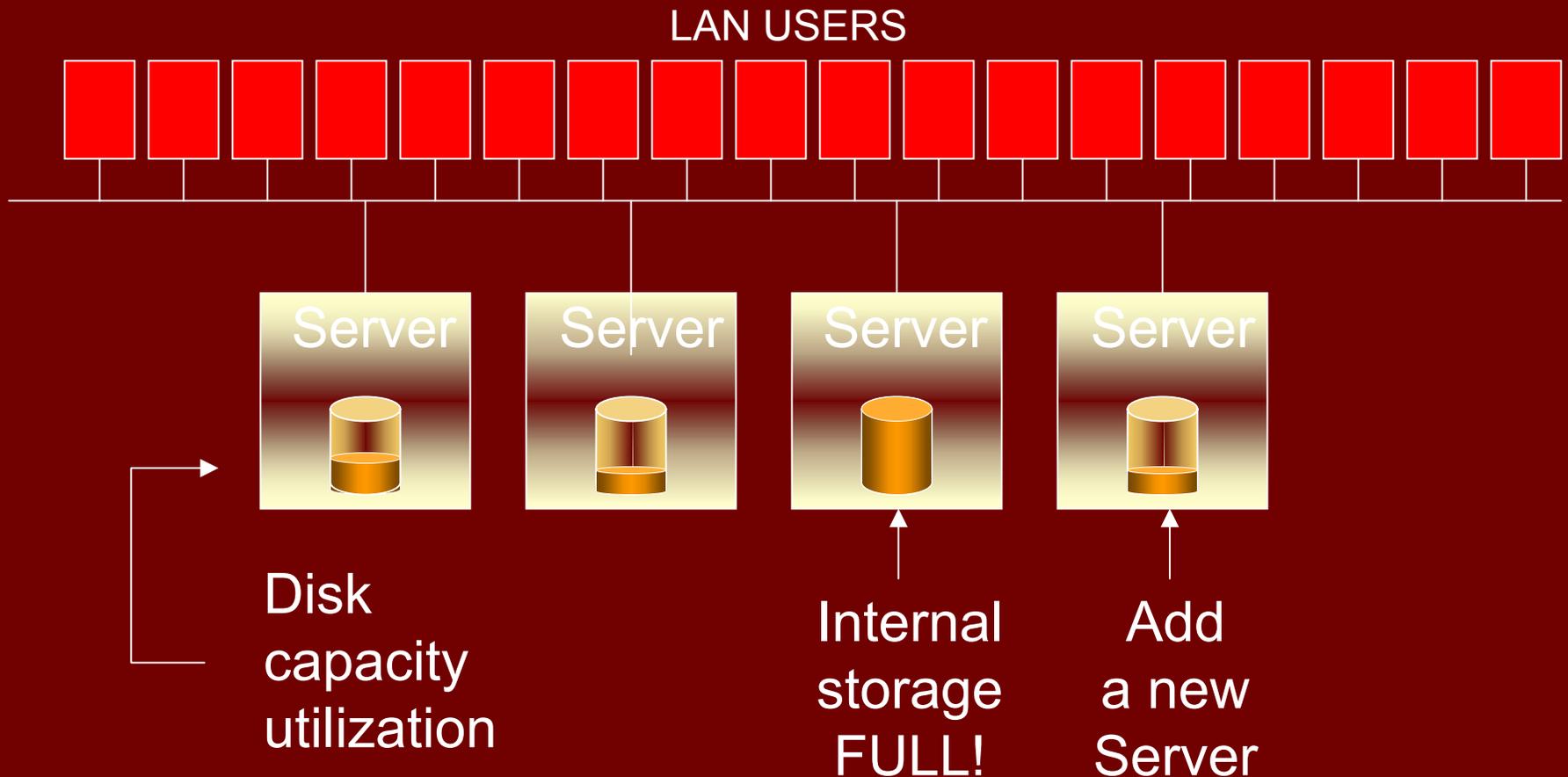
SAN Consolidation Strategy

LAN USERS

Server   Server   Server   Server   Server

SAN Serverless
Back Up

SAN Storage
Consolidation!

Dynamic Storage
Growth
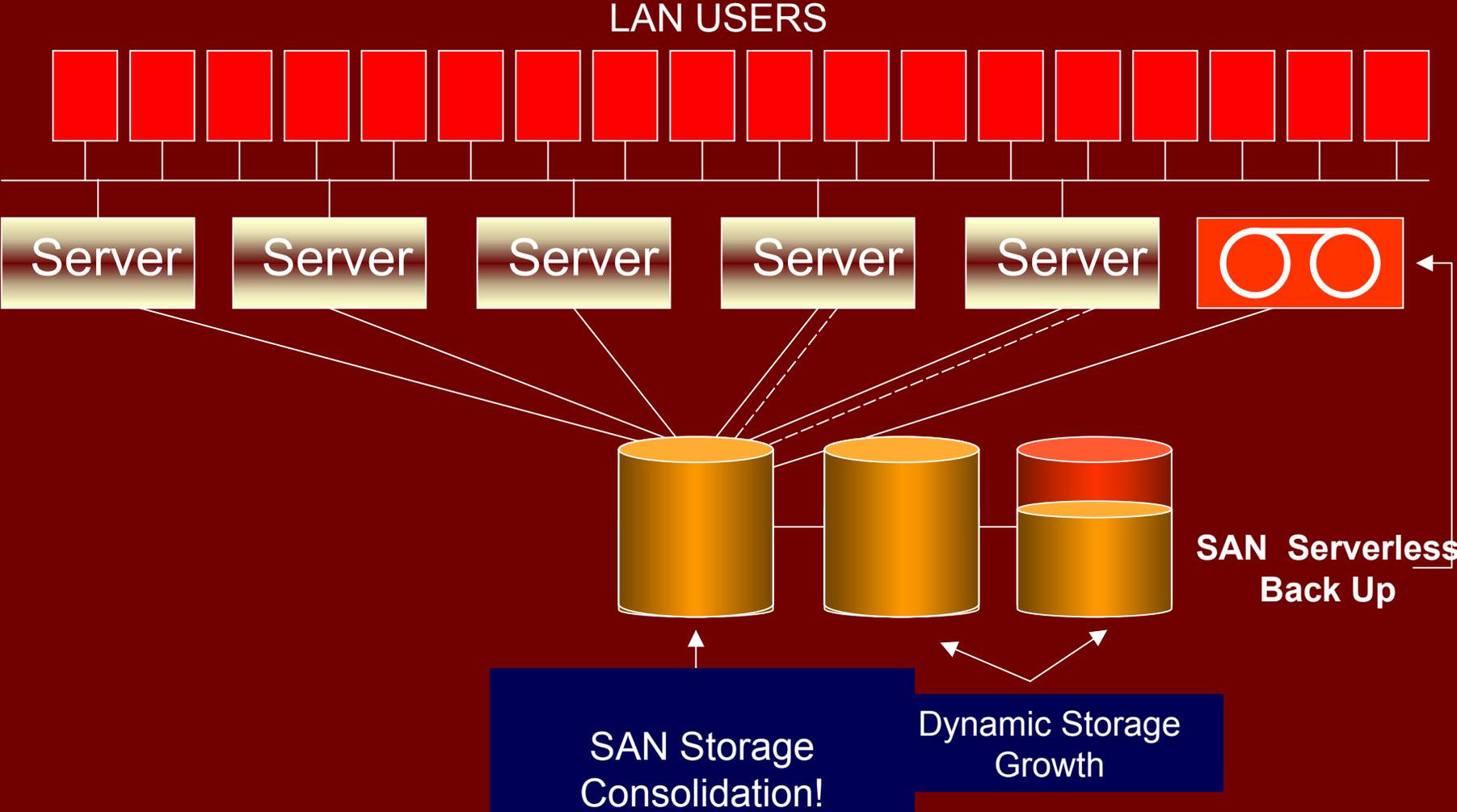
Ultimate Consolidation Including NAS
(Where You Want To Be)

Benefits

- A single storage pool to manage
- Fewest servers to manage
- SAN investment optimized
- Unlimited NAS scalability
- Maximum storage flexibility and performance

SAN

Exchange Server

Oracle Server

NAS Executor E7000

Windows Client

Novell Client

UNIX Client

LINUX Client
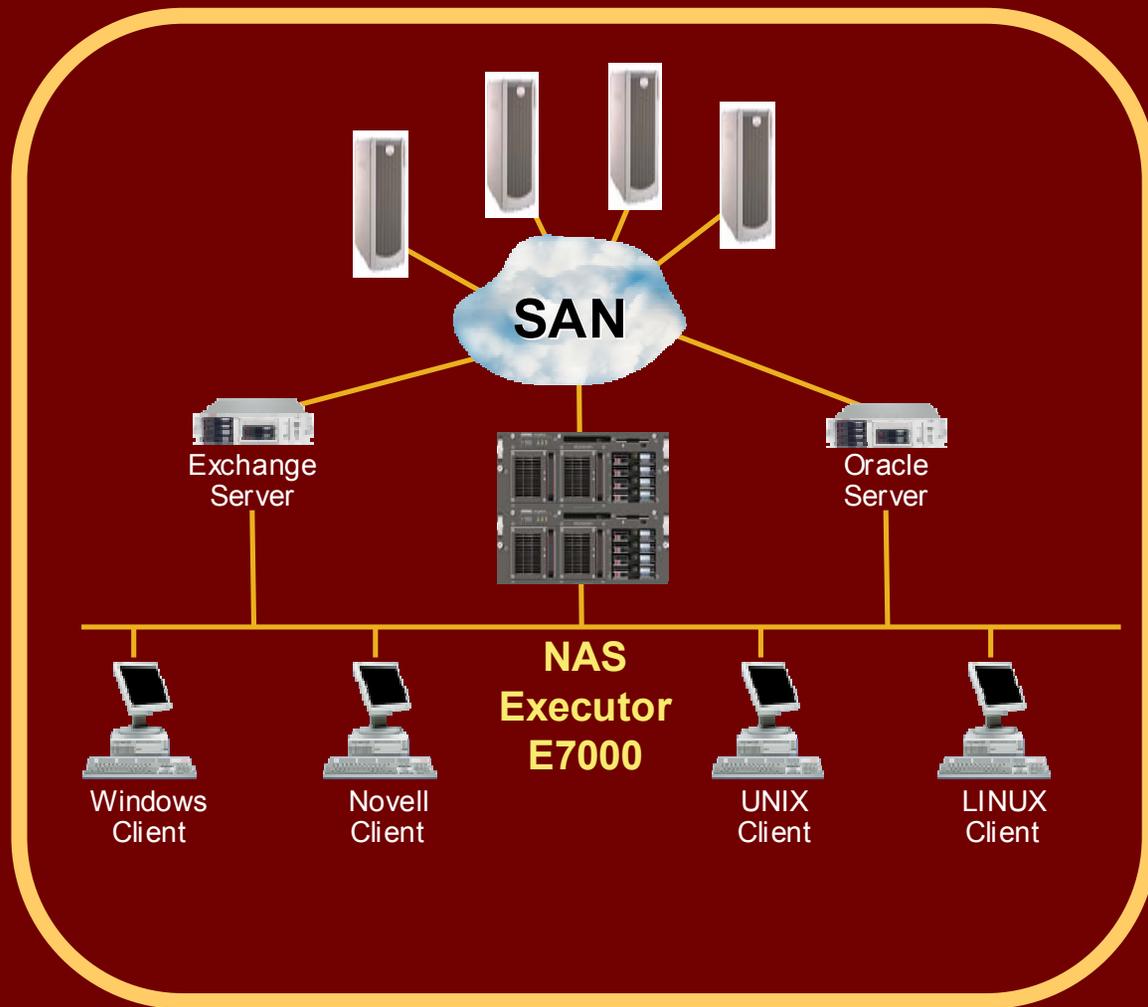
HP XP Series

Consolidation requires
- Broad connectivity
- Very high throughput
- Large capacity
- Capable management tools
  - Security, performance, allocation, availability
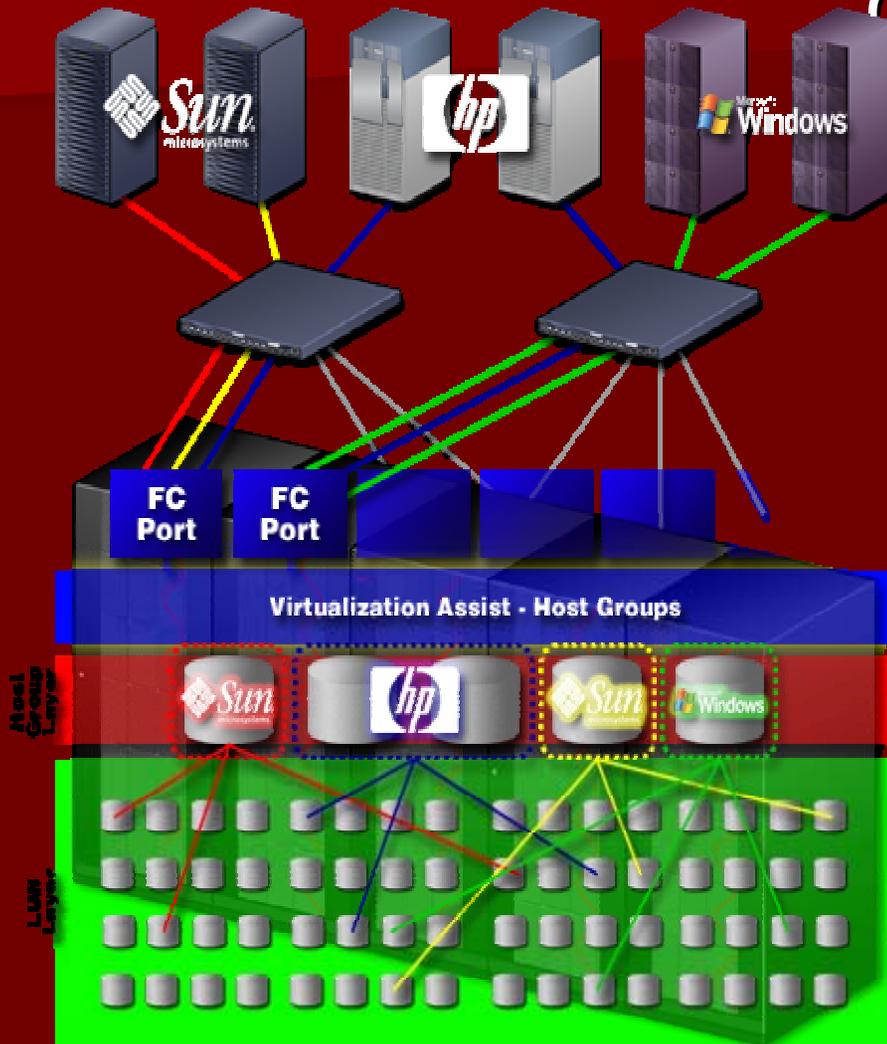
Storage Consolidation with SAN
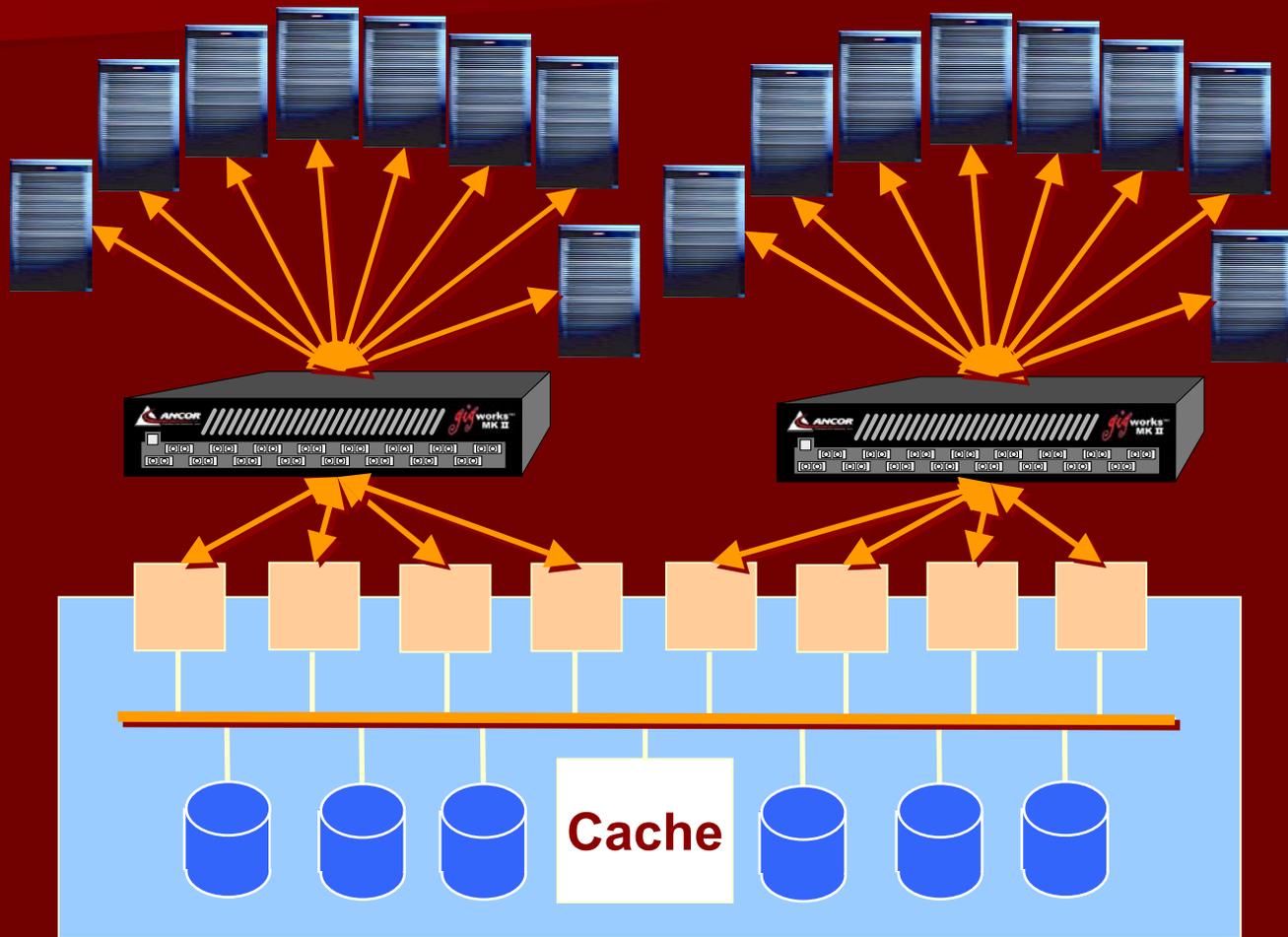
Server Effectiveness

HP ProLiant server from 4P to 8P
Enhanced enterprise performance

ProLiant BL40p, DL560, DL580, ML570 - up to 4 processors

ProLiant DL740, DL760 - up to 8 processors

NEW:

A complete family of Itanium-based servers: HP's Integrity Servers

High end
- Launch: 6/30/03 — **Superdome 16,32,64-way**
- Fall 2003 — HPTC Cluster

Midrange
- Launch: Fall 2003 — 8-way, 16-way

Entry
- Launch: 6/30/03 — **rx2600 (2-way)**, **rx5670 (4-way)**
- Fall 2003 — Carrier grade (2-Way)

2004
Itanium
NonStop

Windows Server 2003

Storage Savings

Driving Trends

Storage Growth Is Growing *Very Fast!*

Chart legend:
- □ Staffing Expenditures to Manage Data=5%/YR
- ■ Corporate Data Storage Growth=100%/YR

Years: 2000, 2001, 2002, 2003, 2004, 2005
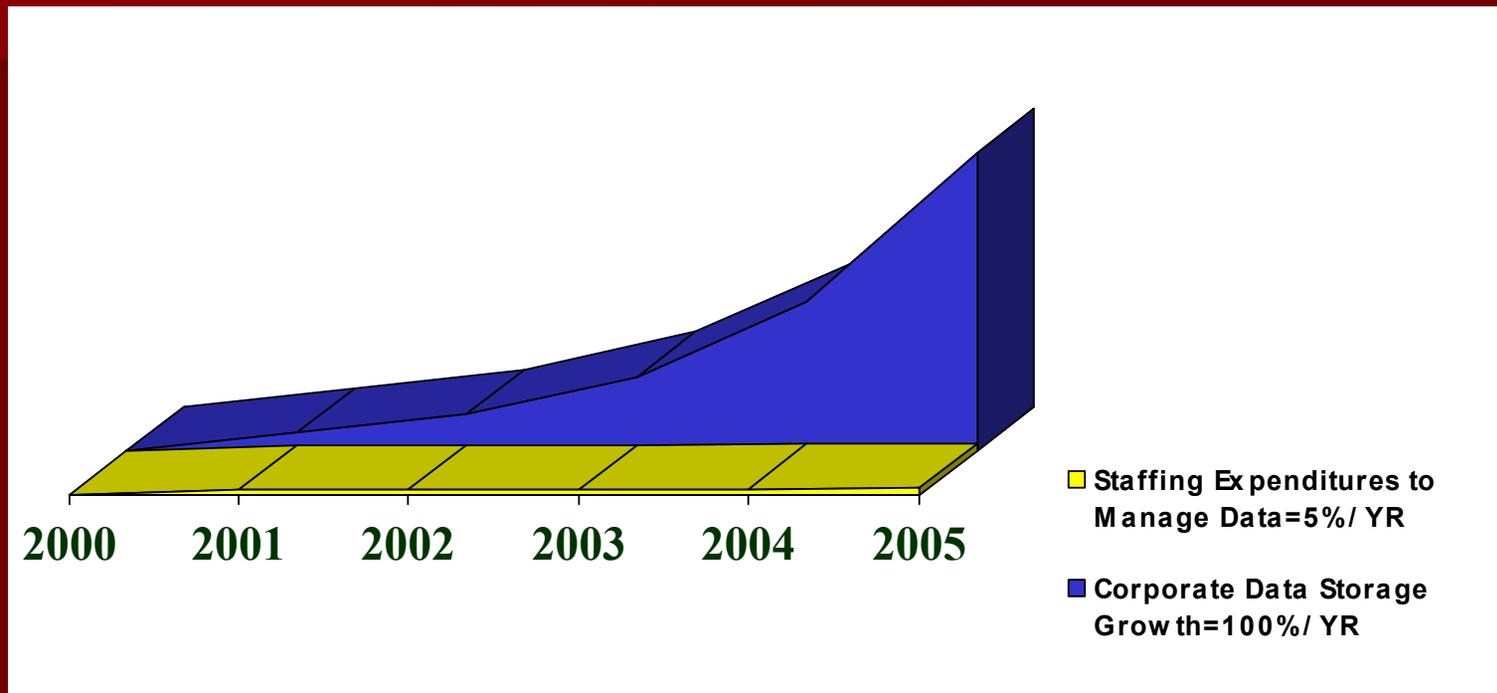
IT staffing budgets increasing 5% a year

Corporate data growth increasing 100% year

Storage Administrator Implications

SAN Customer Benefits
(Storage Savings)

➢Better disk capacity utilization(average unused storage in a DAS environment typically >50%).

➢Manage with less people(typically one storage administrator can manage 4 or more times the storage).

➢significantly shorter backup windows.

➢Drives and tape on the same storage network.

➢Reduced LAN traffic.

➢Flexible environment for future growth.

➢Usually should be able to add storage, switches, tape to the SAN while applications run.

Fiber Channel SAN's

Host Bus Adapters

- Individual Component Utilization Is A Necessary Part Of SAN Architecture and Implementation.

- Selecting the Right HBA Will Lower Overall SAN Costs.

Host Bus Adapters

- Connect the Server to the SAN.
- Alleviate the Server From Some I/O Processing.
- Typically, Assist in the Execution of Parts of Communications Protocol.
- Compatibility Across HBA's.

Fiber Channel Switches

- ➢ Hubs
  - • 7 port and 12 port models
  - • Much lower cost than switches
  - • Used in smaller SANs
- ➢ Switches
  - • 8, 16, 64 and 128 port models
  - • Better isolation than hubs
  - • Switches are more $ than hubs.
  - • Scalable to large SANS
  - • Cascading, Zoning, Quickloop
  - • Higher MTBF, easier repair
- ➢ GBICs (GigaBaud Interface Cards)
  - • Short and long wave variants
  - • LW to 100Km
  - • SW to 500m

FC Switch Products

SAN Core 2/64 **SAN Director 2/64**

2GB Products Are Replacing 1Gb!

SAN Edge 2/32
SAN Edge 2/16

SAN Switch 2/16

SAN Switch 2/16 EL
SAN Switch 2/8 EL

SAN Director 64

MSA Fabric Switch 6

SAN Switch Integrated 32/64

SAN Switch 16-EL
SAN Switch 16

MSA1000 only

SAN Switch 8
SAN Switch 8-EL

Current 1Gb Products
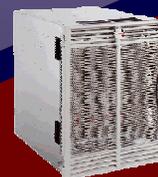
Fiber Channel Implementation

Switches implement FC Fabric

- Switches act like network routers

- *FC Switches provide non-blocking Dynamic Paths Through The Fabric*

- Multiple communication paths

- Bandwidth is aggregate of paths

- Switches provide logical isolation

Departmental SAN's

Modular SAN Array 1000
Technical Overview

- **Key Elements**
  - Scaleable to 3TB
  - Fully redundant components
  - 2Gb FC host connections
  - Ultra3 SCSI -- 1" universal drives
  - High performance controller
  - Boot from SAN capability
  - ProLiant & x86 servers

- **Competitive Advantages**
  - *Built on Compaq's exclusive DtS Architecture*
  - Embedded 2Gb fabric switch option
  - Advanced Data Guarding

Enterprise SAN's

Scalable Storage

Enterprise Virtual
Array Family

Enterprise Virtual
Assisted Array
XP Series

Modular Array
Family

Modular SAN
Array 1000

Enterprise Class High End

Enterprise Class
Midrange

Business Class
(Departmental SAN)
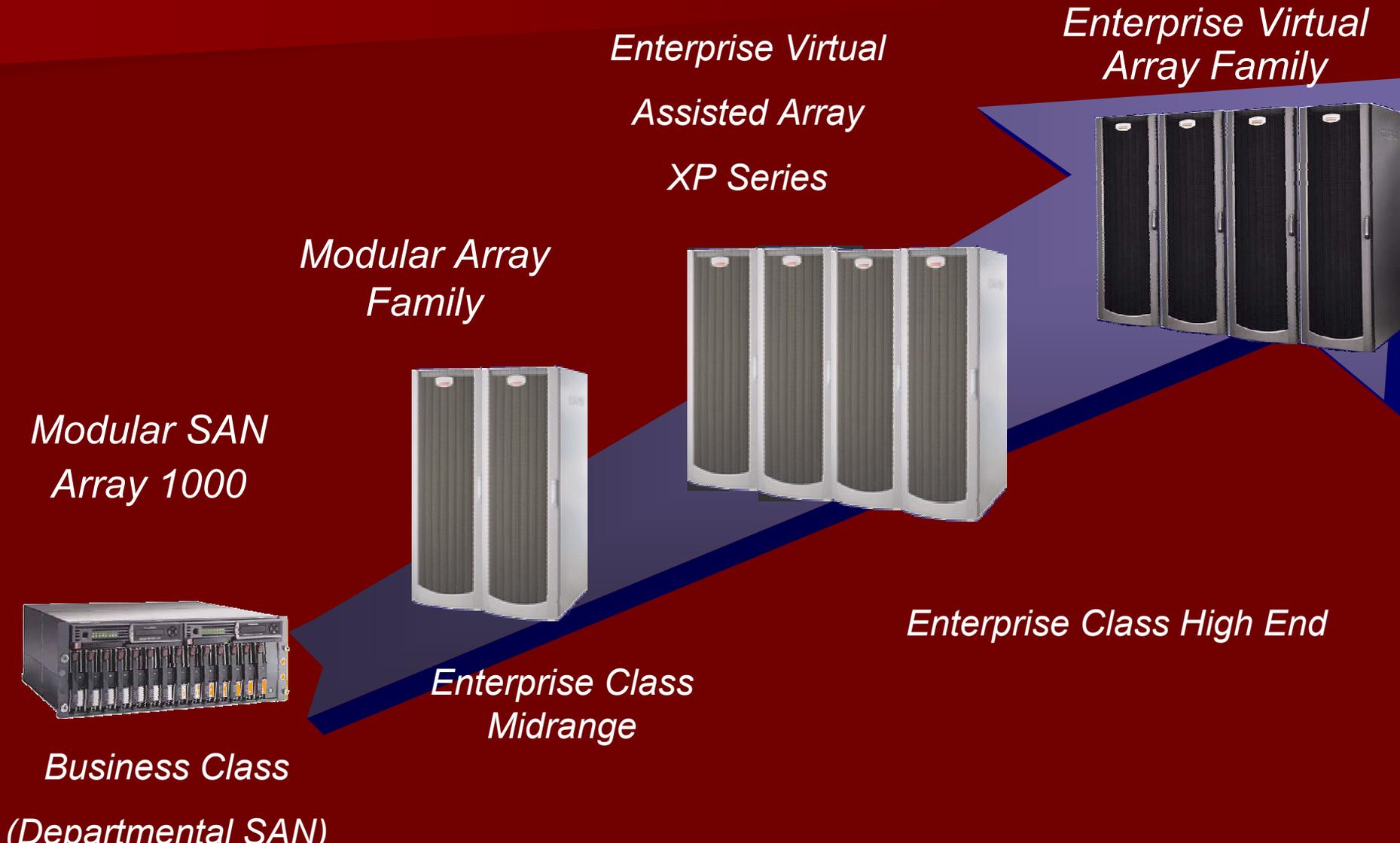
StorageWorks Enterprise Virtual Array

- Modular, scalable and highly available design
  - Redundant Power
  - Redundant Cooling
  - Distributed hot spare disk drives
  - Mirrored Cache with battery backup
- Multi-level virtual RAID architecture
  - **RAID V0**
  - **RAID V1**
  - **RAID V5**
- Full 2Gb fibre channel front to back
  - 1Gb fibre channel SAN compatible
- Co-exist with HSG80 class Modular Array's
- Compaq SANworks SecurePath™ compatible
- High performance HSV110 Array Controllers
  - **Support up to 240 disks**
  - **Capacity support up to 12.3TB in single cabinet**
- Centralized, unobtrusive manageability

Significantly Higher "Utilization" of Purchased Capacity

- **Up to <u>twice</u> the typical 40-50% Open Systems utilization . . . Based upon dynamic pool/LUN expansion, etc.**

- **Importance to <u>Business/IT</u>:**
  - **Customer minimizes purchase of <u>un</u>usable capacity**
  - **Just-In-Time capacity increments for application growth**
    - **Even to the point of adding one disk-at-a-time**
    - **Dynamic Pool/LUN Expansion (w/Server support)**
  - **No "droop" effect in performance**
  - **No intensive storage administration "gyrations"**

- **<u>BOTTOM LINE</u>:**
  Much lower "effective" price/MB . . . Easily "justifiable"

Simplify with XP Series

- **Storage that is easy:**
  - Who would think a single device could do so much?
    - Heterogeneous connectivity
    - Convergence of data types
    - Multi-Protocol flexibility
  - Multi-dimension scalability
    - Capacity, performance, connectivity, workload mix
  - Eliminates hardware and software complexity
    - Less stress, reduced expense
  - Self-managing
    - Automated quality of service
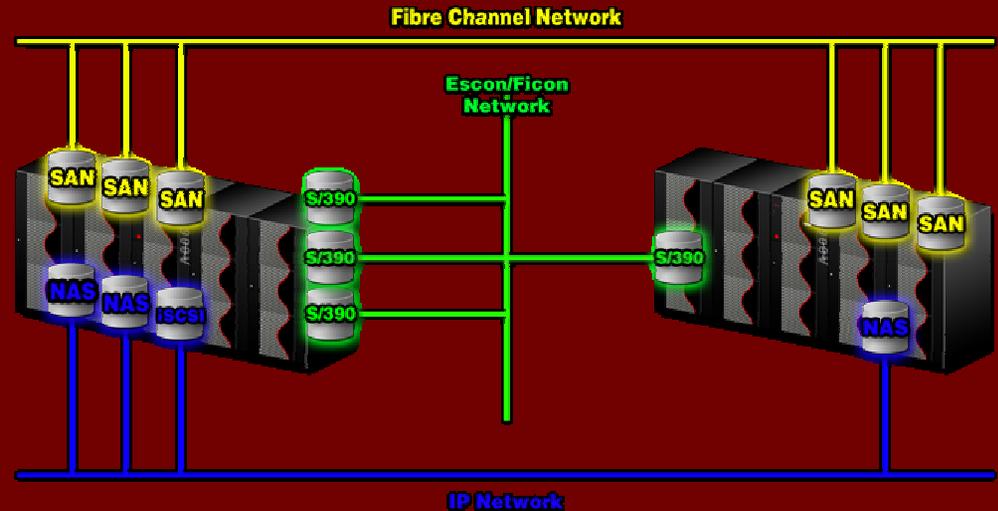
XP Consolidation

- **Host Storage Domains**
  - **Eliminates SAN Islands (Host types)**
  - **Less physical complexity & network cost**
  - **Enables much larger TB/unit**
- **FICON, PAV and Mod 27**
  - **Eliminates ESCON Bottlenecks**
  - **Larger S/390 systems**
- **Performance Maximizer (QoS)**
  - **Allows consolidation without compromise**
  - **Reduces management**
- **Multi-Protocol support for ubiquitous access**
  - **Fibre Channel, FICON, ESCON, others in the future**
  - **Provides consolidation point for all storage**

Fibre Channel High Performance Disk Arrays

Cache

Switch

Switch

Cache

**Dual 100/200mb/s Fiber Loops
RAID Striping**
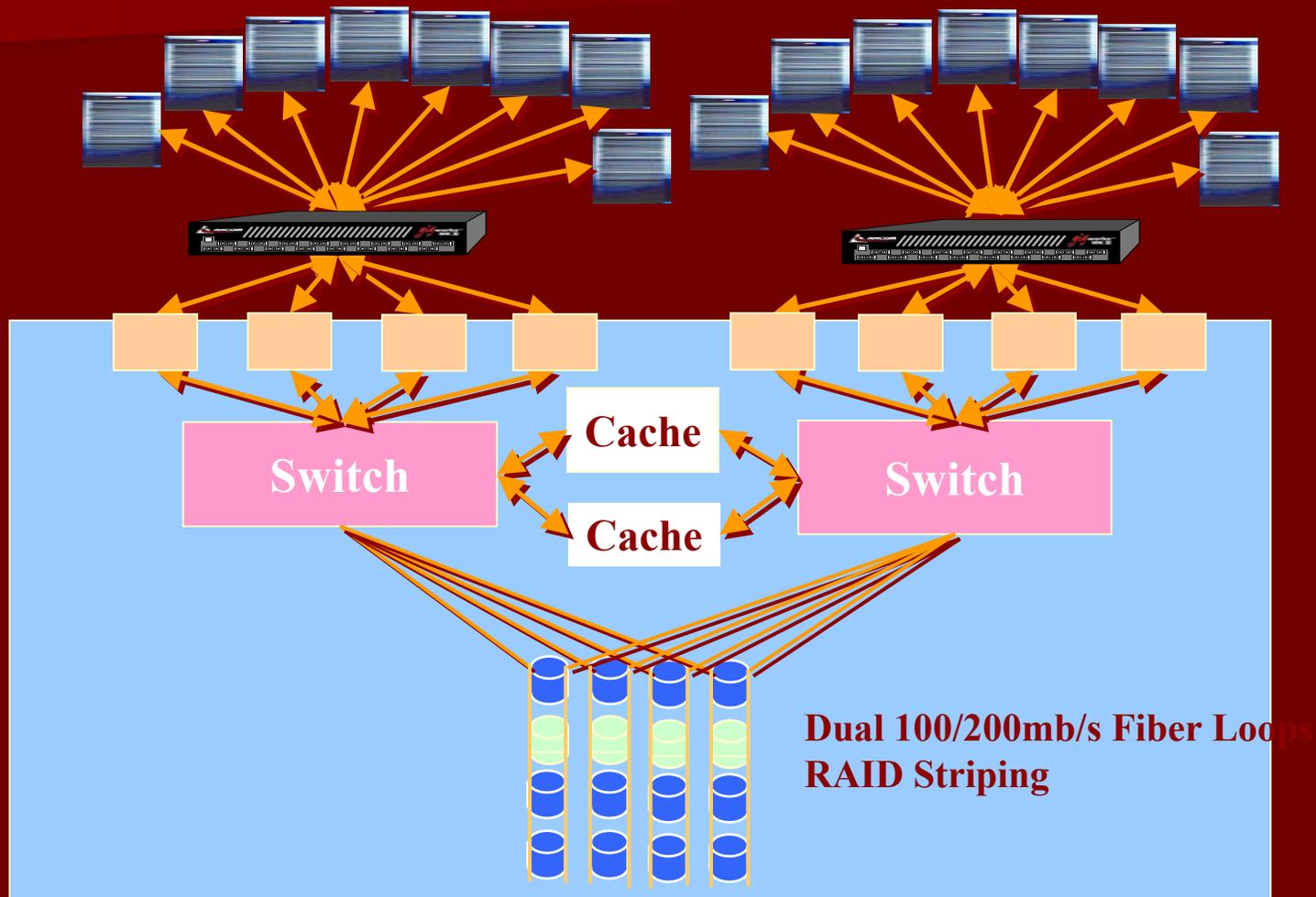
The Future - Open Virtual Storage

SAN Wide Virtual View

COMPAQ
digital
HEWLETT PACKARD
Microsoft Windows NT Server
Sun microsystems
Novell
IBM

"D1" "C1" "N1" "S1" "SG1" "I1" "I2"

Shared Pool of Storage
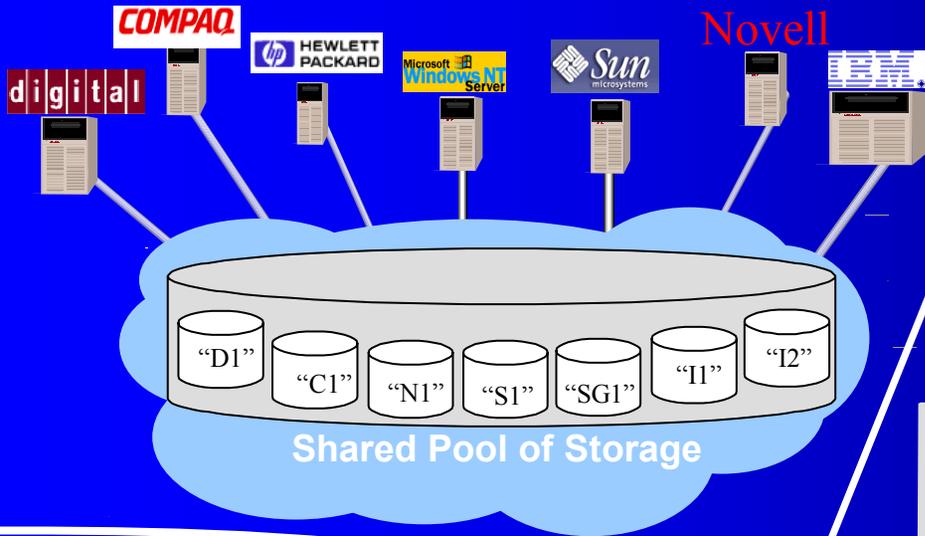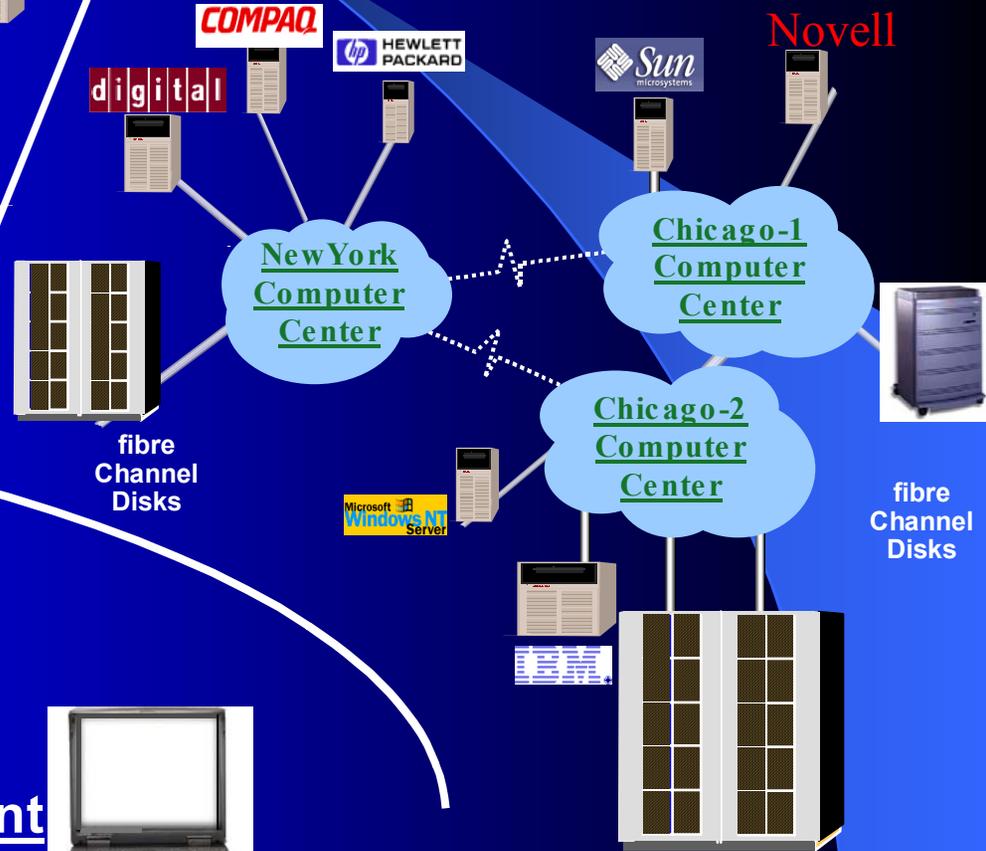
Individual System View

"D1"
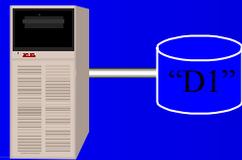
Global Virtual Management

Actual Geography

COMPAQ
digital
HEWLETT PACKARD
Sun microsystems
Novell

New York Computer Center

Chicago-1 Computer Center

Chicago-2 Computer Center

Microsoft Windows NT Server

IBM

fibre Channel Disks

fibre Channel Disks

RAID Implementations

RAID types

data
data
mirror
mirror

RAID 1 2D+2D

data
data
data
parity

RAID 5 3D+1P

data
data
data
data
mirror
mirror
mirror
mirror

RAID 1 4D+4D

data
data
data
data
data
data
data
parity
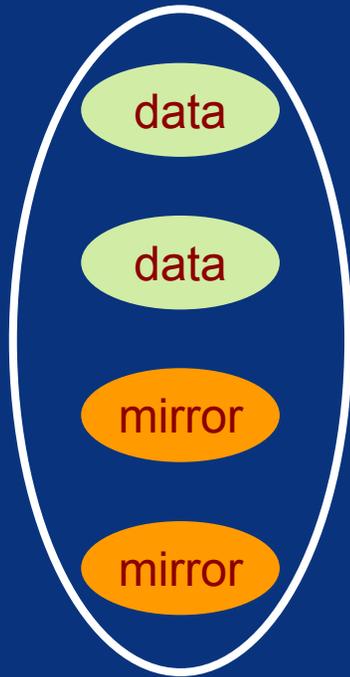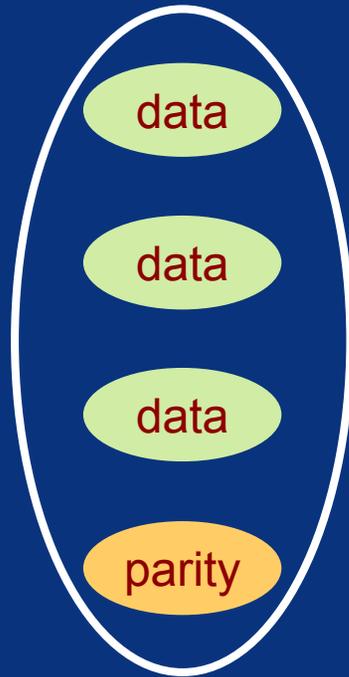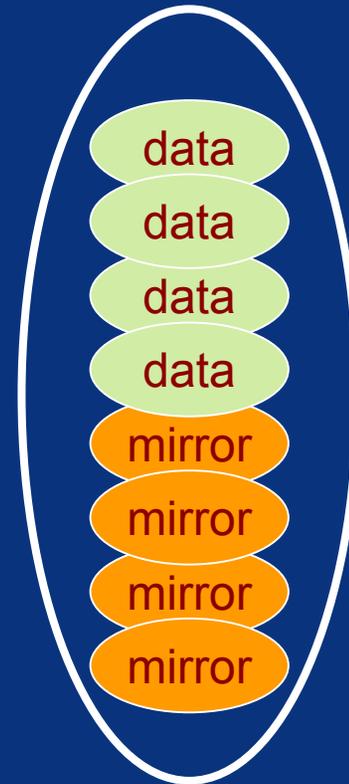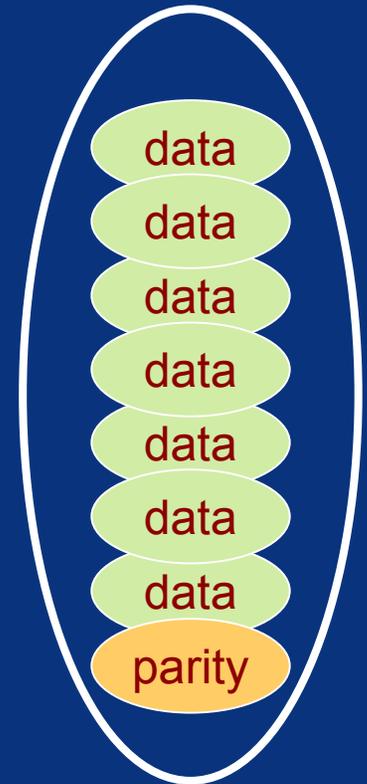
RAID 5 7D+1P

| 50% | 75% | 50% | 87.5% |

storage efficiency, usable/raw

RAID 0+1

- Striping First plus Mirroring
  - Dual Pair RAID group configuration
  - Higher performance in very random environments
  - 32 Slots pre-fetched for sequential access

advanced data guarding

hp
exclusive
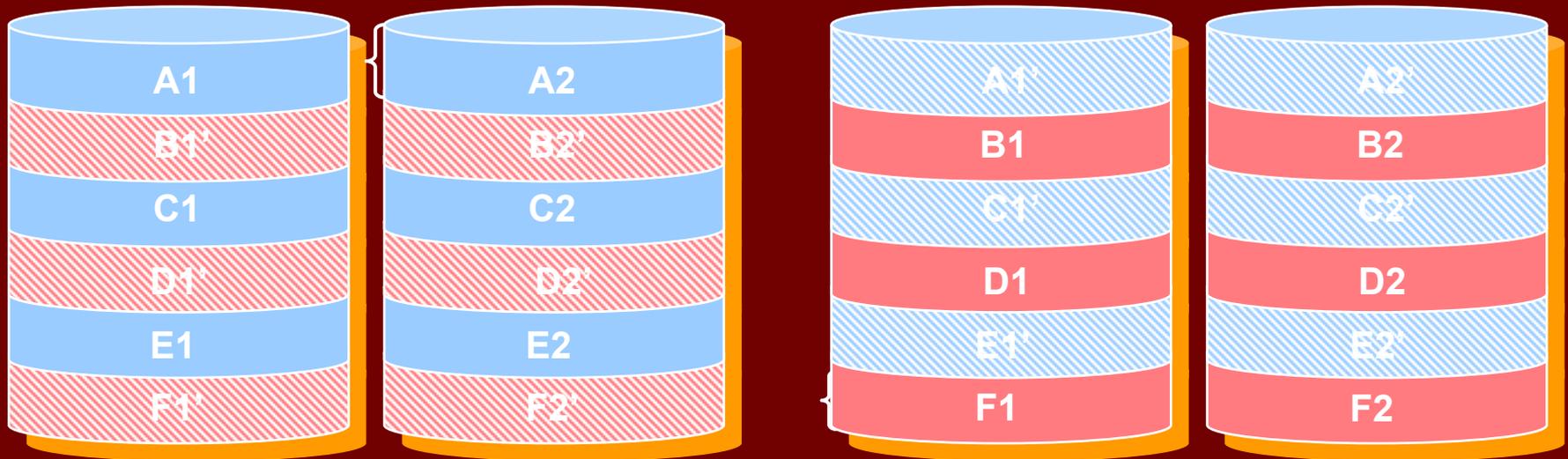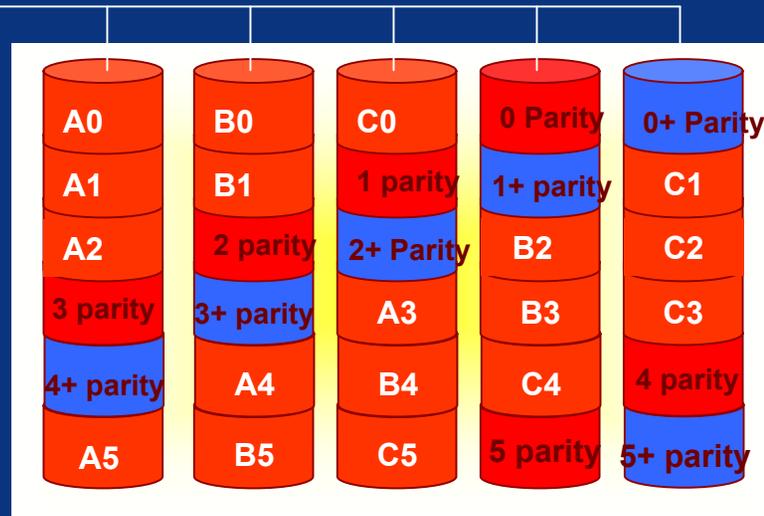
- similar to RAID 5, data and parity is distributed across all drives
- the capacity equal to two drives is reserved for two sets of parity data (RAID 5 distributes one set of parity data)
- RAID ADG can withstand two simultaneous drive failures without downtime or data loss
- supports online RAID level migration from RAID 1 or RAID 5

SAN RAID Design Before Analysis

AIX Development Database Server Oracle

AIX Production Database Server Oracle

AIX Production MUMPS Database Server

HP Alpha Cluster Tru64 EMAIL Server 5000 Users

HP Alpha Cluster Tru64 Email Server 5000 Users

HP Alpha Tru64 EMAIL Server 1800 Users
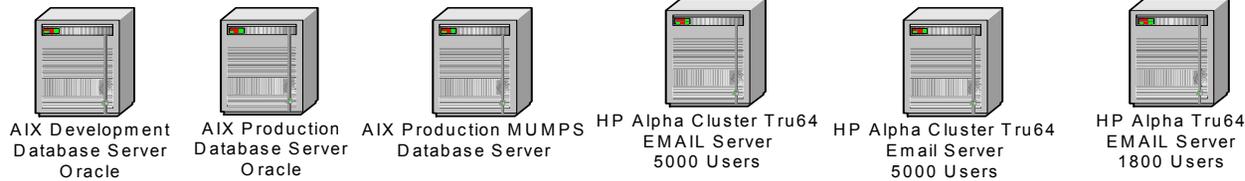
Most Servers Dual Homed HBA's (Connections Not SHown)

ILinuxI-Band Virtualization 1 HBA to Clients, 1 HBA to Disks, 1 HBA Spare
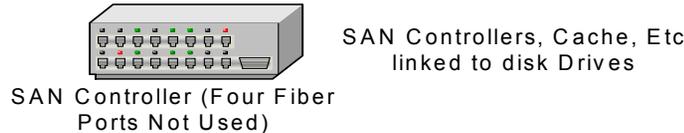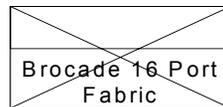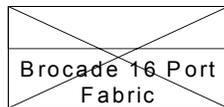
Win2000 Database

Linux In-Band Virtaizton 1 HBA to Clients, 1 HBA to Disks, 1 HBA Spare

Brocade 16 Port Fabric

Brocade 16 Port Fabric

Shelf L0 -- Dual Fiber Loop With RAID-5 Sets Using 72GB Disks (Note - All RAID Set Disks Physically Adjacent)

SAN Controllers, Cache, Etc linked to disk Drives

SAN Controller (Four Fiber Ports Not Used)

LUN L2   LUN L2   LUN L2   LUN L2   LUN L4   LUN L4   LUN L4   LUN L4   LUN L6   LUN L6   LUN L6   LUN L6
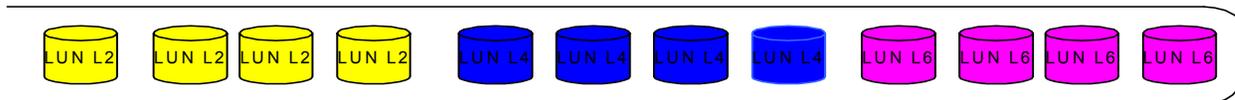
LUN L2 Three Oracle Databases w Redo Logs on 4 Disk RAID-5 Set

LUN L4 MUMPS IDX Database (CISR50 128GB) 4 Disk RAID-5 Set

LUN L6 Backup EMAIL Database (86GB) 4 Disk RAID-5 Set

Shelf L1 -- Dual Fiber Loop With RAID-5 Sets Using 72GB Disks (Note - All RAID Set Disks Physically Adjacent)

LUN L3   LUN L3   LUN L3   LUN L3   LUN L5   LUN L5   LUN L5   LUN L5   LUN L7   LUN L7   LUN L7   LUN L7
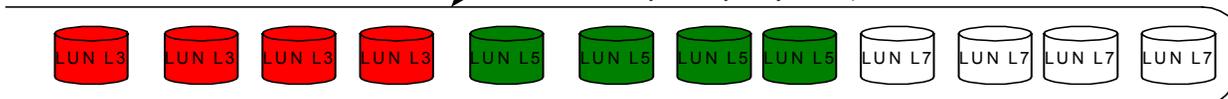
LUN L3 MUMPS IDX Database (CISR50   128GB) 4 Disk RAID-5 Set

LUN L5 One Oracle Database (64 GB) w Redo Logs & Prod EMAIL Database (120GB) on 4 Disk RAID-5 Set

LUN L7 Not Used

SAN RAID Design After Analysis

Most Servers Dual Homed HBA Connetions. (Servers and connections not shown)

Brocade 16 Port Fabric

Brocade 16 Port Fabric

Linux In-Band Viirt
1 HBA to Clients, 1 HBA to Disks, 1 HBA Spare

Brocade 16 Port Fabric

Brocade 16 Port Fabric

Linux In-Band Virtt
HBA to Clients, 1 HBA to Disks, 1 HBA Spare

SAN Controllers, Cache, Etc linked to disk Drives

SAN Controller (Four Fiber Ports Not Used)

Shelf L0 -- Dual Fiber Loop With RAID-1+0 Sets Using 72GB Disks

| LUN L2 | LUN L3 | LUN L5 | LUN L4 | LUN L6 | MIRL2 | MIRL3 | MIR L5 | MIR L4 | MIR L6 | Redo Logs | MSTR | Spare |

LUN L2 Three Oracle Databases w Separate Redo Logs on RAID-1+0 Sets

LUN L4 MUMPS IDX Database (CISR50 128GB) RAID-1+0 Sets

LUN L6 Backup EMAIL Database (86GB) RAID-1+0 Sets

Shelf L1 -- Dual Fiber Loop With RAID-1+0 Sets Using 72GB Disks

Data

| LUN L2 | LUN L3 | LUN L5 | LUN L4 | LUN L6 | MIRL2 | MIRL3 | MIR L5 | MIR L4 | MIR L6 | Redo Logs | Spare |

LUN L3 MUMPS IDX Database (CISR50 128GB) RAID-1+0 Sets

LUN L5 One Oracle Database (64 GB) w Separate Redo Logs & Prod EMAIL Database (120GB) on RAID-1+0 Sets

Shelf L2 -- Dual Fiber Loop With RAID-1+0 Sets Using 72GB Disks

| LUN L2 | LUN L3 | LUN L5 | LUN L4 | LUN L6 | MIRL2 | MIRL3 | MIR L5 | MIR L4 | MIR L6 | Redo Logs | Spare |

| LUN L2 | LUN L3 | LUN L5 | LUN L4 | LUN L6 | MIRL2 | MIRL3 | MIR L5 | MIR L4 | MIR L6 | MSTR MIR | Spare |

Shelf L3-- Dual Fiber Loop With RAID-1+0 Sets Using 72GB Disks

When *NOT* To Use RAID-5

➢In general, RAID-5 disk configuration is not recommended for use with most relational databases (Do Not USE with Oracle databases)!

➢Write performance suffers very significantly compared to a RAID 1+0 (mirroring and striping) configuration (checksum Calculation & Update Degrades I/O Performance!)

➢Time window availability not withstanding significant increases in volume of data will probably cause total elapsed time to exceed acceptable time limits.

RAID 1+0 Or RAID 0+1

In

RAID 1+0 Or RAID 0+1?

➢ RAID 0+1 configuration disks striped together into sets, followed by the mirroring of the sets.

➢ RAID 1+0 configuration drives are mirrored followed by the striping together of the resultant mirror sets.

➢ RAID 1+0 better I/O (read and write)

➢ Performance increase because of simultaneous I/O across various spindles.

➢ RAID 1+0 *MORE* expensive (you get what you pay for !)

SAN Fabric

Fabrics

- **Scalability means more than adding ports**
  - Supporting data growth while sustaining infrastructure performance, availability, and management capability

- **Enable scalability**
  - Design fabric that supports key environment attributes
    - Examples: Data Store Size and location, Backup, Distance, Applications

- **Fabric design drives product selection**
  - Switched Infrastructure
    - Building block size
    - Availability
    - Link performance
 - Management
    - Device to Fabric-wide

Large Fabric Design

Switch Fabric
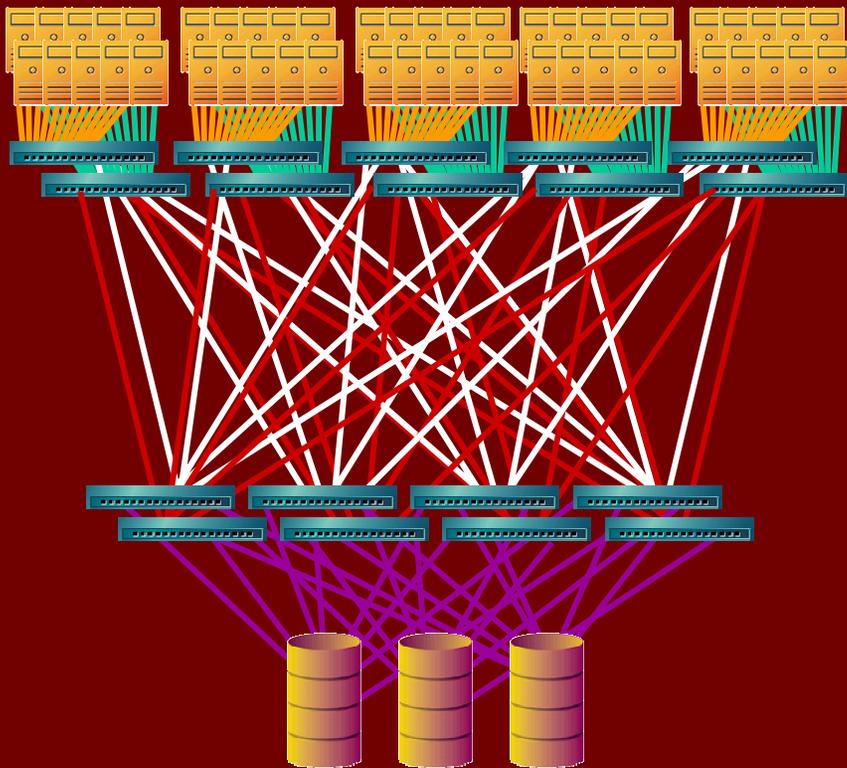
Director Fabric

Enterprise Storage
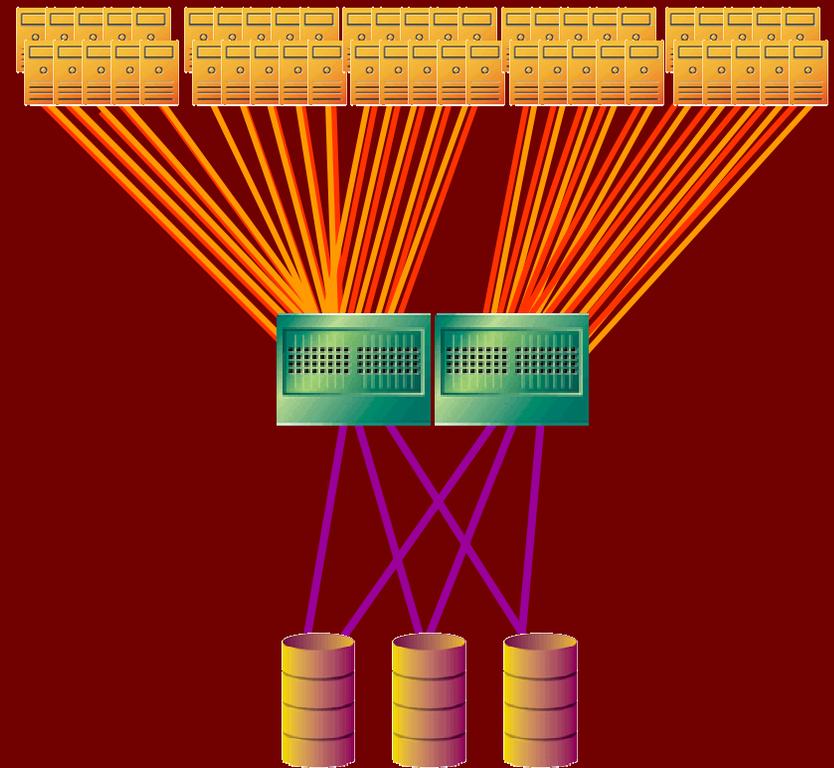
Enterprise Storage

Fabric Architecture Design
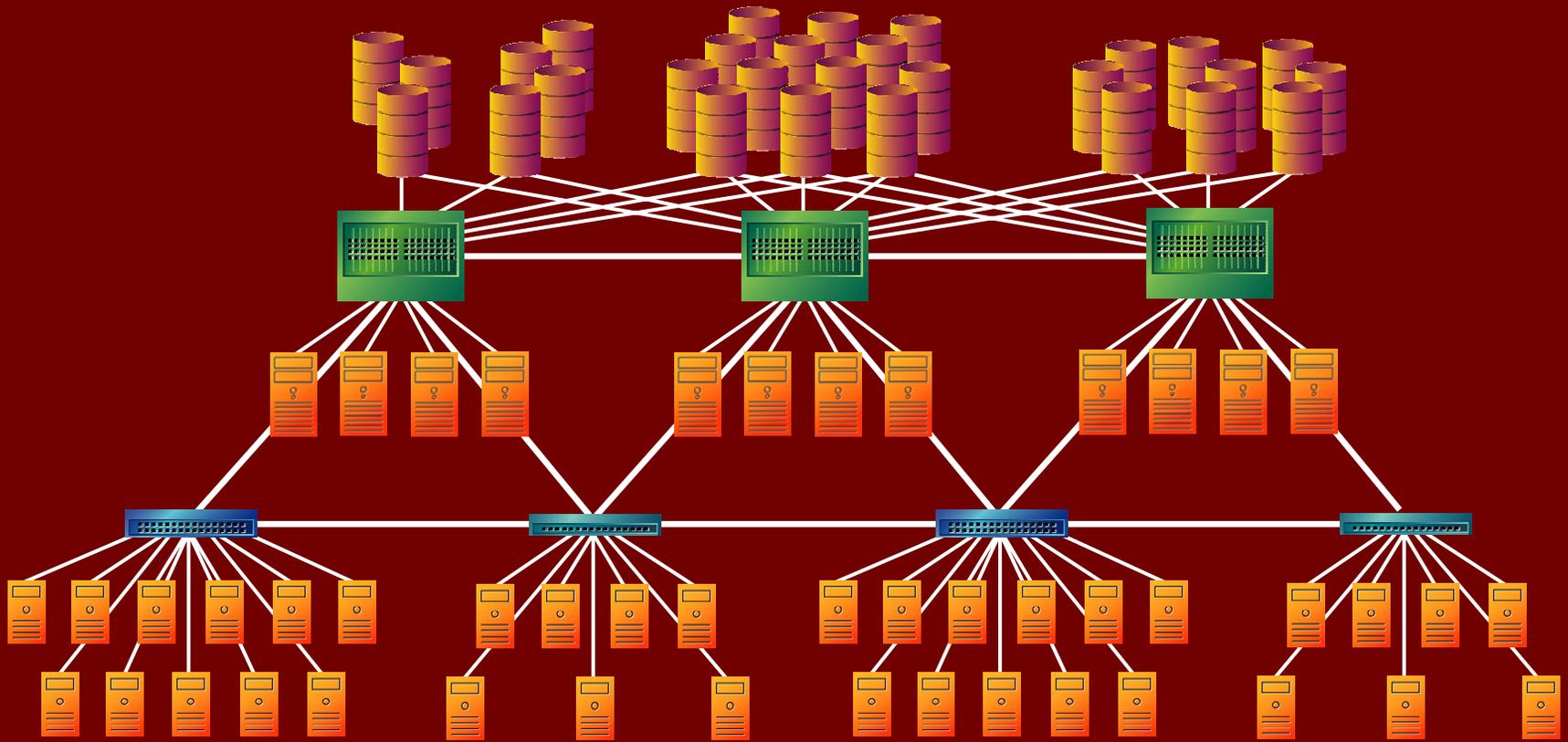
Directors and Switches

- **Availability *Depends On Vendor***

  – Director: ~99.999% downtime < 5 minutes / year

  – Fabric Switch: ~99.9%
    - downtime 8.8 hours / year
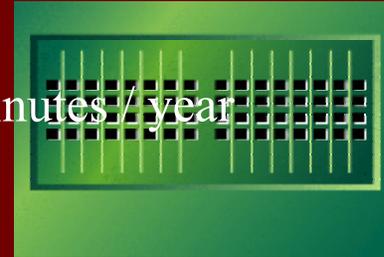
- **Any-to-Any Connectivity**

  – Director: large port count

  – Fabric Switch: small to medium port count
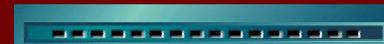
- **Serviceability**

  – Director: non-disruptive

  – Fabric Switch: potentially disruptive

- **Scalability**

  – Director: better fabric scaling

  – Fabric Switch: Use Multiple Switches

Improve Storage Utilization

DAS

NT   UNIX   NT

NT 50%   **UNIX** 50%   **NT** 50%

NT
UNIX
NT — 80%

SAN

NT   UNIX   NT
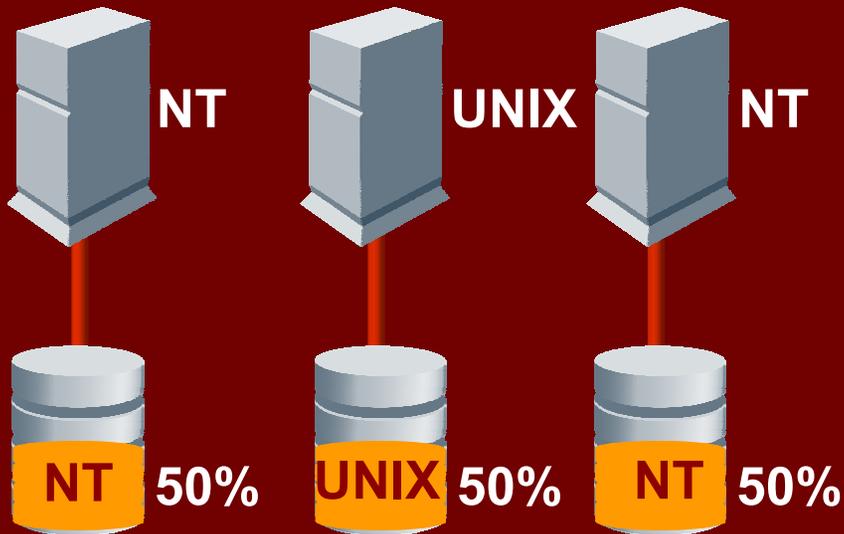
- Limited scalability and bandwidth

- >50% of storage maybe unused

- Admin costs 5-6 times acquisition cost
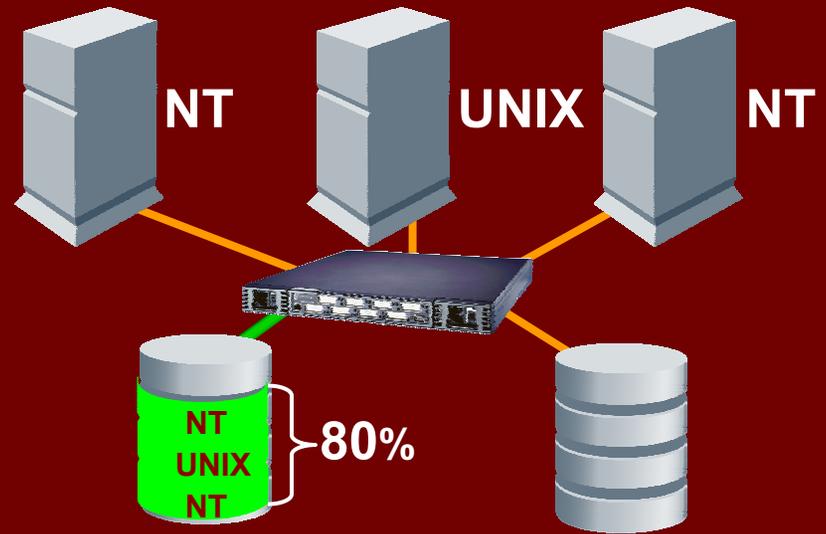
- Reduce storage headroom

- Enable just-in-time provisioning

- Combine UNIX and NT storage growth needs