

# Network Backup

## *The Basics of Configuring a Network Backup*

**Winston Hait**

**Senior Product Manager**

**Syncsort, Inc.**

**50 Tice Boulevard**

**Woodcliff Lake, NJ 07675**

**Phone: (201) 573-5179, Fax: (201) 930-8290**

**E-Mail: [whait@syncsort.com](mailto:whait@syncsort.com)**

# Overview - The 8 Basic Steps

- ◆ Inventory
- ◆ Examine Network/Hardware
- ◆ Document Your Requirements
- ◆ Identify Priorities
- ◆ Choose Backup Software
- ◆ Implement Hardware/Software
- ◆ Implement Strategy
- ◆ Put into Production/Monitor/Refine

# Step 1: Inventory (Enterprise)

◆ How many sites?

◆ Currently?

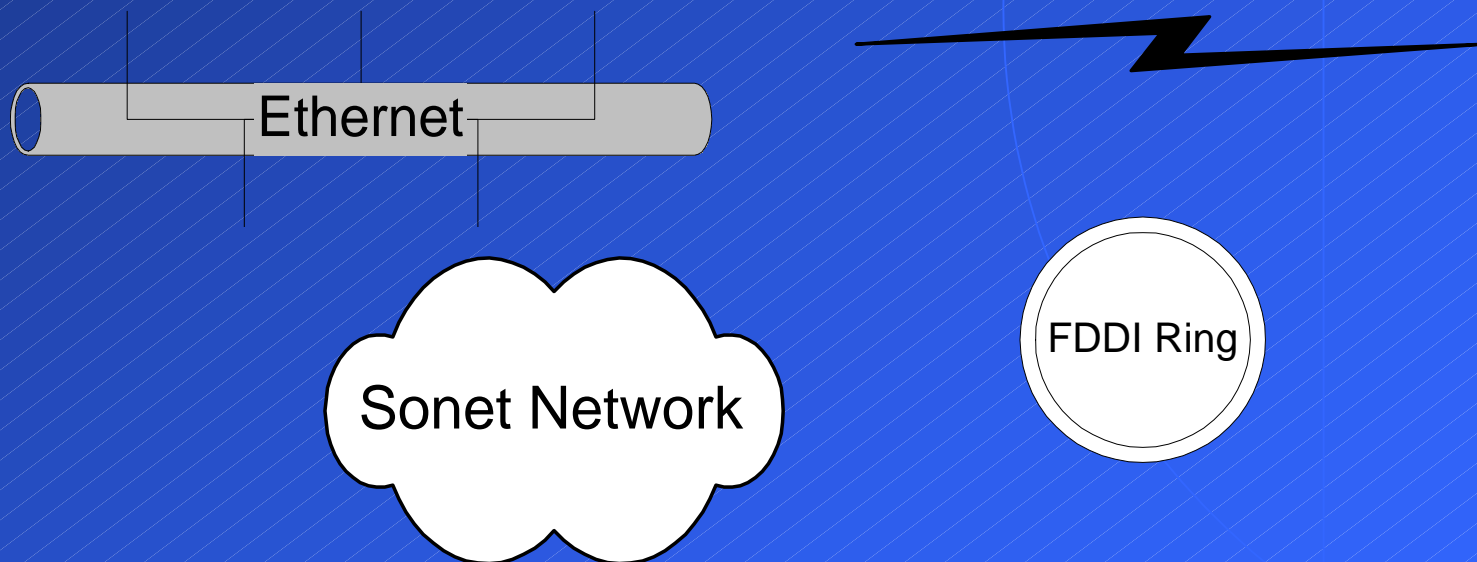


◆ In the Future?



# Step 1: Inventory (Network)

- ◆ What is the Network backbone?



# Step 1: Inventory (Machines)

## ◆ Machines: Operating systems?



NT/ Win98 workstation

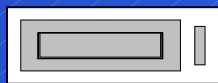


UNIX Machine

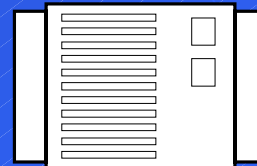


WinNT Server

## ◆ Machines: Any tape devices?

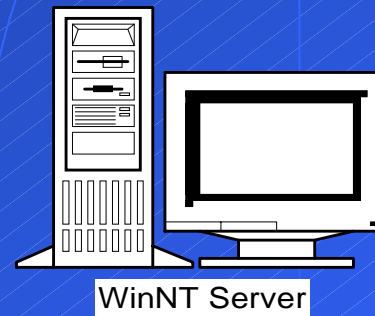
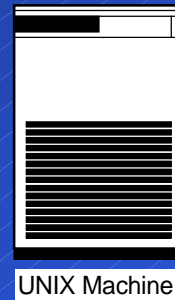


tape drive



Tape Library  
(Jukebox)

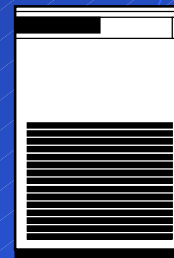
# Step 1: Inventory (Applications)



- ◆ Applications running on each machine
- ◆ How Critical?
- ◆ Databases?
- ◆ E-Mail?

# Step 1: Examine the inventory (Logistics)

◆ Location



UNIX Machine

◆ Location



NT/ Win98 workstation



WinNT Server

◆ Location

Step 1:

Inventory: Where is the most data?

- ◆ Why is location so important?

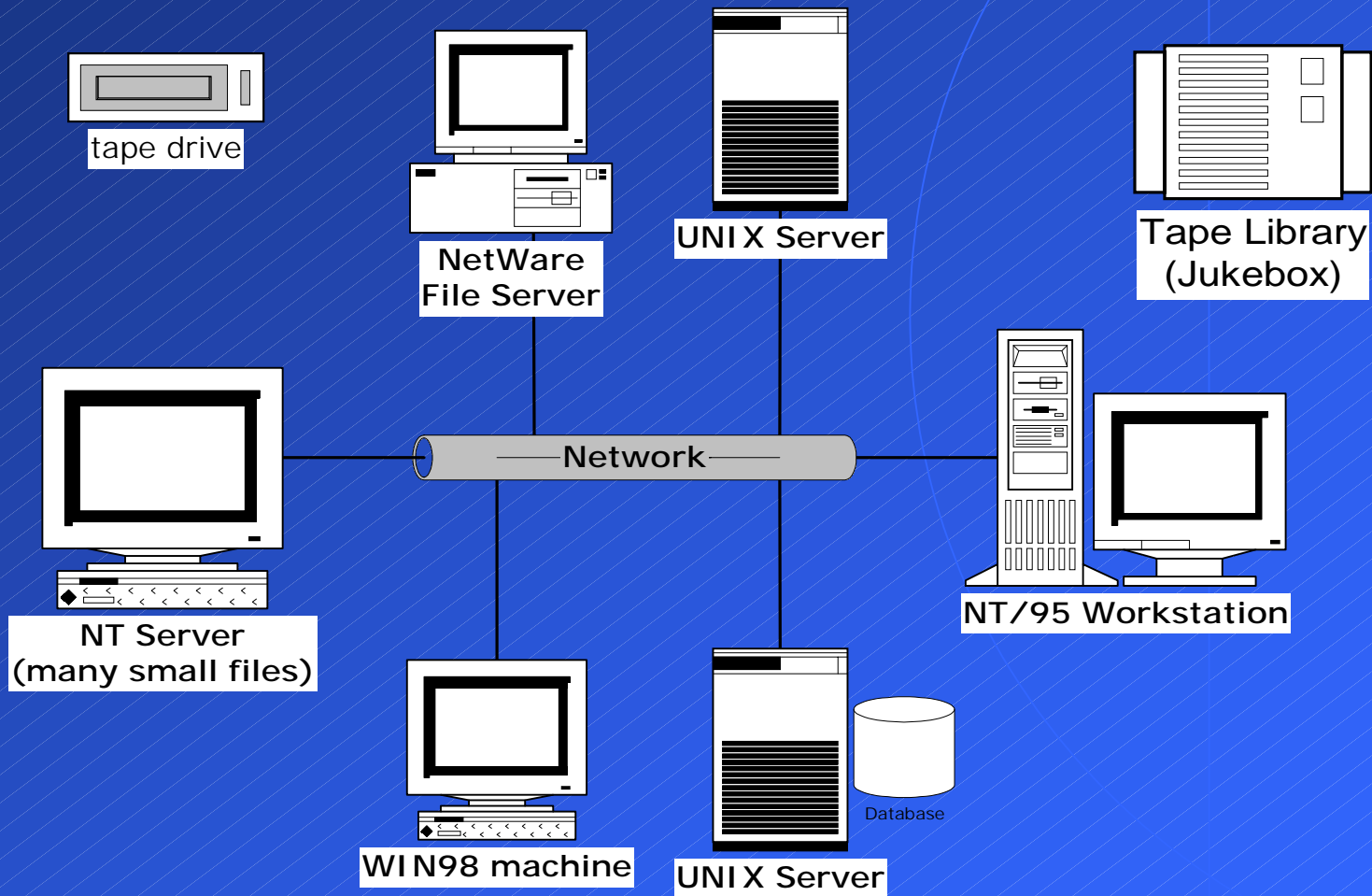
- ◆ It's the location of the ...

◆ D A T A

7



# Step 1: Evaluate the Inventory



## Step 2 - Review

Let's review  
networking capacities  
and  
backup media  
(hardware)

## Step 2 - Review: Examine network speeds

*In a perfect world...*

- ◆ 10Base-T: 3.6GB/hr (=60MB/min=1MB/sec)
- ◆ 100Base-T: 36 GB/hour
- ◆ 1000Base-T: 360 GB/hour
- ◆ FDDI: 36 GB/hour
- ◆ ATM: 270 GB/hour
- ◆ Fibre Channel: 360 GB/hour

Step 2 - Review:  
Examine network speeds

*Let's be practical...*

Divide by 2

Fibre - Think 90%

## Step 2 - Review: Hardware Specifications: 1/2"

3480 comp*	1.5 MB/sec 3 MB/sec	.2GB cap .4GB	5.4GB/hour 10.8GB/hour
3490 comp*	3 MB/sec 6 MB/sec	.4GB cap .8GB	10.8GB/hour 21.6GB/hour
3490E comp*	3 MB/sec 6 MB/sec	.8GB cap 1.6GB	10.8GB/hour 21.6GB/hour
3590 Magstar comp*	9 MB/sec 18 MB/sec	10GB cap 20GB	32.4GB/hour 64.8GB/hour

\* based upon 2:1 compression

# Step 2 - Review: Hardware Specifications: 4mm

DDS2 comp*	336KB/sec 772KB/sec	4GB cap 8GB	1.2GB/hour 2.4GB/hour
DDS3 comp*	1.2MB/sec 2.4MB/sec	12GB cap 24GB	4.32 GB/hour 8.64 GB/hour

\* based upon 2:1 compression

## Step 2 - Review: Hardware Specifications: 8mm

8mm Mammoth	3MB/sec	20GB	10.8GB/hour
comp*	6MB/sec	40GB	21.6GB/hour

8mm Mammoth-2	6MB/sec	40GB	21.6GB/hour
comp*	12MB/sec	80GB	43.2GB/hour

8mm AIT	3MB/sec	25GB cap	10.8GB/hour
comp*	6MB/sec	50GB	21.6GB/hour

8mm AIT-2	6MB/sec	50GB cap	21.6GB/hour
comp*	12MB/sec	100GB	43.2GB/hour

\* based upon 2:1 compression

# Step 2 - Review: Hardware Specifications: DLT

DLT 4000 comp*	1.5 MB/sec 3.0 MB/sec	20GB cap 40GB	5.4GB/hour 10.8GB/hour
DLT 7000 comp*	5 MB/sec 10 MB/sec	35GB cap 70GB	18GB/hour 36GB/hour
DLT 8000 comp*	6 MB/sec 12 MB/sec	40GB cap 80GB	21.6GB/hour 43.2GB/hour

\* based upon 2:1 compression



# Step 2 - Review: Hardware Specifications: Latest

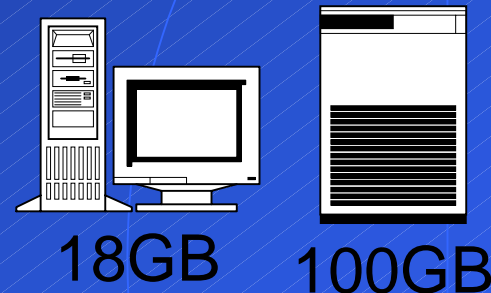
Ecrix (8mm) comp*	3 MB/sec	33GB	10.8GB/hour
	6 MB/sec	66GB	21.6GB/hour
9840 (STK) comp**	10 MB/sec	20GB	36GB/hour
	40 MB/sec	80GB	144GB/hour
SuperDLT comp*	15 MB/sec	100-500GB	54GB/hour
	30 MB/sec	200GB-1TB	108GB/hour
LTO comp*	15 MB/sec	100-500GB	54GB/hour
	30 MB/sec	200GB-1TB	108GB/hour

\* based upon 2:1 compression

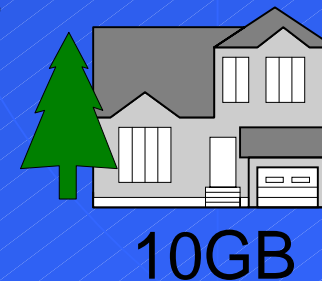
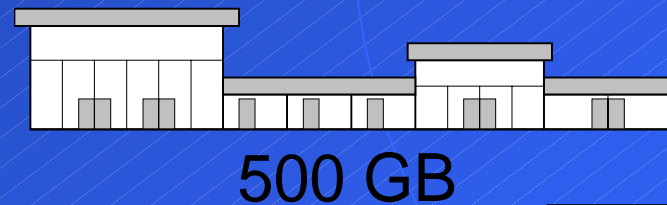
\*\* based upon 4:1 compression

# Step 3: Document Backup Requirements

- ◆ Total data by machine



- ◆ Total data by location



## Create a spreadsheet!

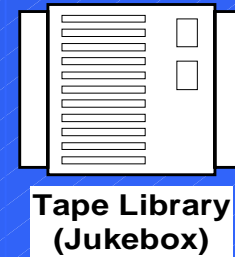
17

Node	Partition	Tape #	base on	When	total data	Best case time	day of week	time	how long?
A	C,D,E	1	Saturday	3AM	Base 75Gig	13 hours	<i>Differential</i>	8PM	1 hour
A	G,H,I	2	Saturday	3:15AM	Base 75Gig	13 hours	<i>M-TH</i>	8:15	
A	J	1	Friday	7PM	Base 50 Gig	10 hours		8PM	
A	K	2	Friday	7:15PM	Base 50 Gig	10 hours		8:15	
B	D,E,F	3	Saturday	4AM	Base 75Gig	13 hours	<i>Differential</i>	8PM	1 hour
B	G,H,I	4	Saturday	4:15AM	Base 75Gig	13 hours	<i>M-TH</i>	8:15	
B	J	3	Friday	8PM	Base 50 Gig	10 hours		8PM	
B	K	4	Friday	8:15PM	Base 50 Gig	10 hours		8:15	
C		1,2,3,4	Base on	7PM	<b>BASE 200Gig</b>	14 hours	<i>Differential</i>	8PM	1 hour
ATL 4/52			Friday				<i>M-TH</i>		
D			Base on	7PM	BASE 100 Gig	4 hours	<i>Differential</i>	8PM	1 hour
ATL 4/52			Saturday				M-F		
E			Base on	2AM	BASE 100 Gig	4 hours	<i>Differential</i>	8PM	1 hour
ATL 4/52			Sunday				M-F		
F			Base on	8AM	BASE 25 Gig	4 hours	<i>Differential</i>	8PM	1 hour
G			Sunday		BASE 75 Gig		M-F		
ATL 4/52									
H			Base on	2PM	BASE 85Gig	5 hours	<i>Differential</i>	8PM	1 hour
I			Sunday		BASE 25 Gig		M-F		
J					BASE 10 Gig				
ATL 4/52									
K			Base on	8PM	BASE 25 Gig	1 Hour	Catalog Backup		
			Sunday				Every Day		

# Sample spreadsheet!

## Step 3: Document Backup Requirements

- ◆ What is the backup window? Every day? Weekends?
- ◆ How much data changes daily? By percentage of the machine
- ◆ Data retention requirements? Any legal issues?
- ◆ Physical Storage considerations? (Automated Tape Library vs. Offsite)



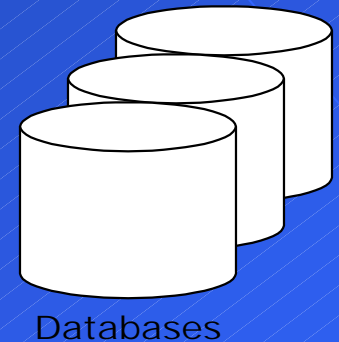
19

## Step 3: Document Backup Requirements

- ◆ What about Databases?
- ◆ On-line database requirements

## Step 3: Database Backup Possibilities

- ◆ Cold backup - Physical snapshot that requires database down.
- ◆ Hot backup - The application must run 24/7, therefore the database is active: Perform logical backup.
  - ◆ Mirroring - (e.g. EMC TimeFinder)

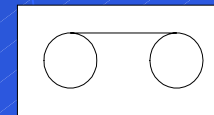


# Step 3: When to Perform the Backups...

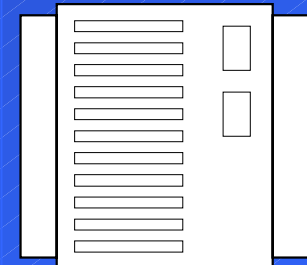
- ◆ Unattended Backup Required/Desirable?
- ◆ Now...
- ◆ Future Growth
- ◆ Introduction of backups



tape drive



Tape Drive



Automated Tape Library

Let's  
cook up  
a plan!

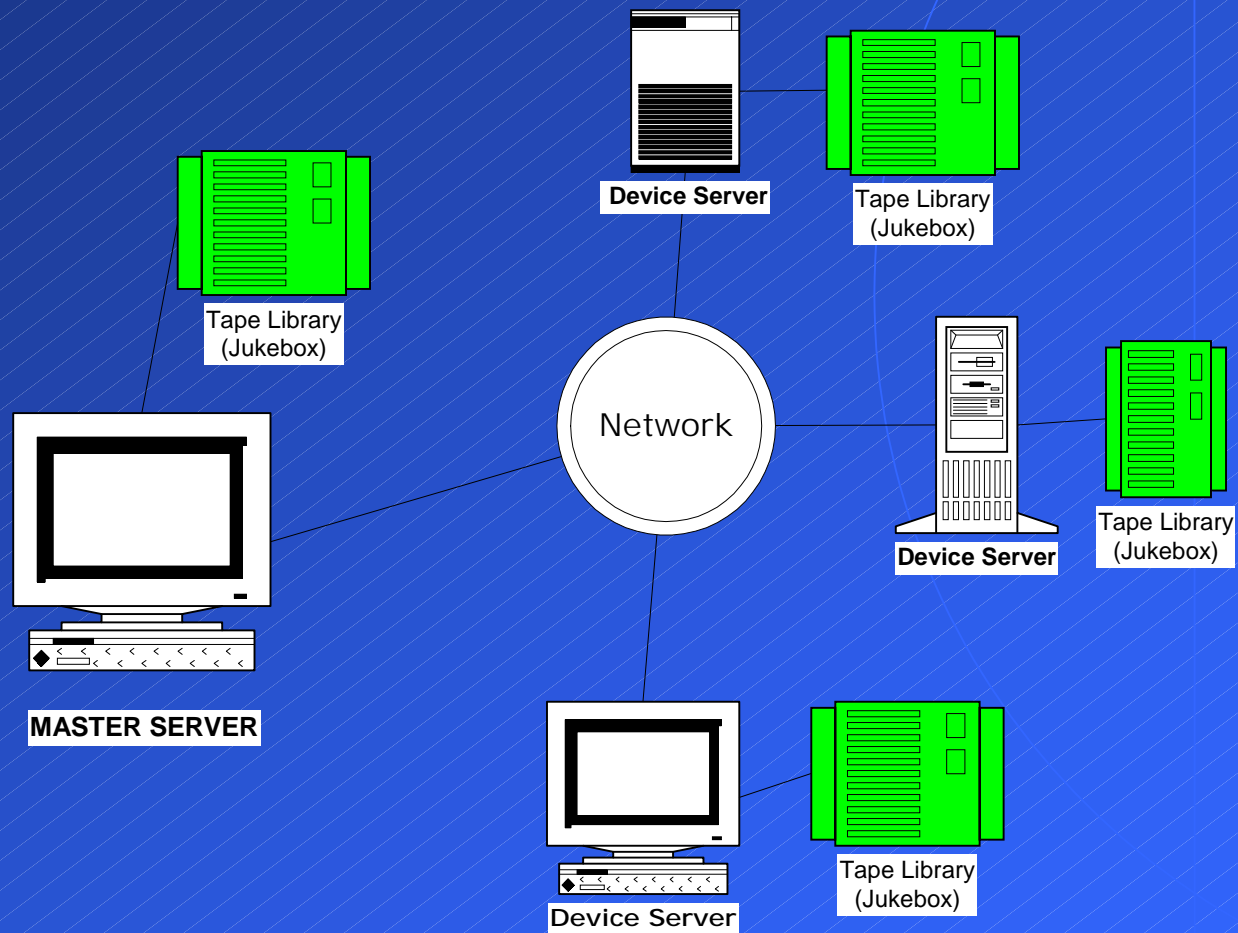




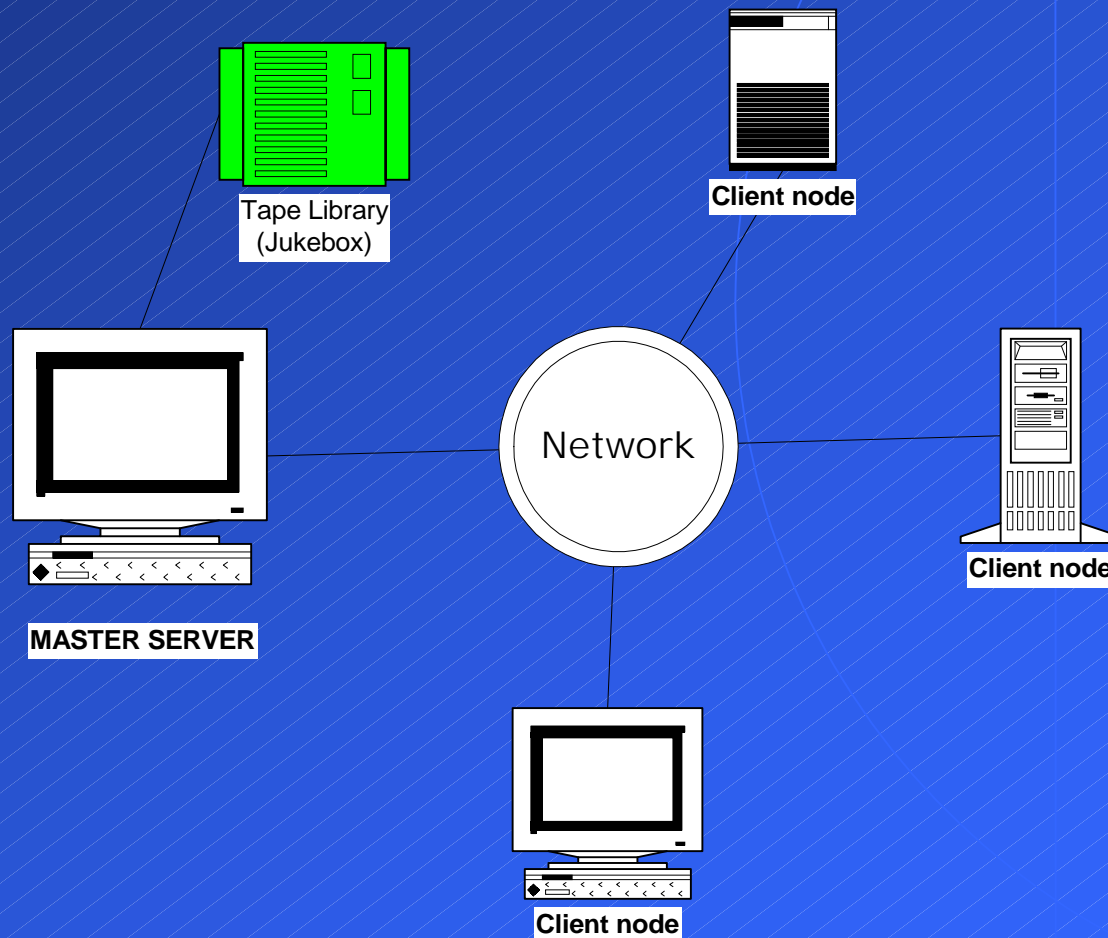
## Step 4: Identify Priorities

- ◆ Use all the information you gathered
- ◆ Specify minimum requirements and features
- ◆ Local Backups (Distributed) vs. Over the Network (Centralized)
- ◆ Do you need a SAN? At what cost?
  - ◆ Multi-hosting
  - ◆ Arbitrated Loop
  - ◆ Mesh (Switched) Fabric

# Step 4: Identify Priorities (Fully Distributed)



# Step 4: Identify Priorities (Centralized)



## Step 4: Identify Priorities (Multi-Hosting)

- ◆ Multi-hosting – direct SCSI-connect server to tape library
  - ◆ Cheap – cost of SCSI cables
  - ◆ Fast – each connection can move 40-80MB/sec
    - ◆ Most tape drives write between 3 and 15MB/sec
  - ◆ Off-loads traffic and maximizes tape investment
    - ◆ Speeds are fast enough to keep tape spinning all the time
  - ◆ Only downside – distance - 25 meter limit of SCSI

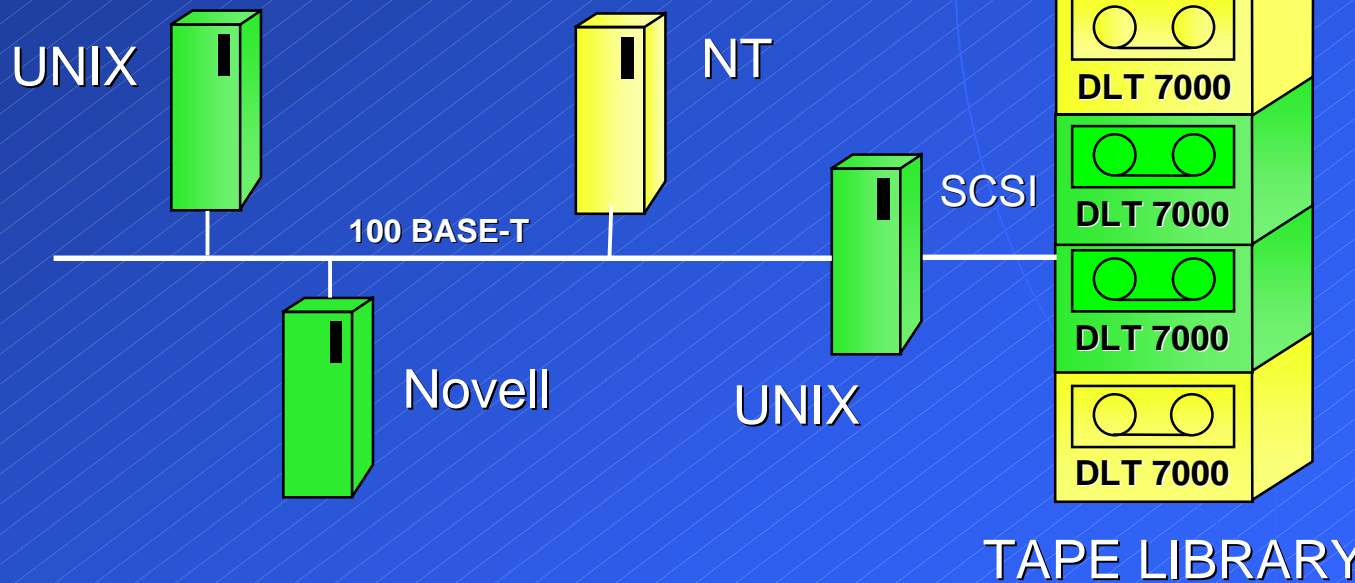
27

# Step 4: Identify Priorities (Multi-Hosting)

Network Speed = 10 MB/sec

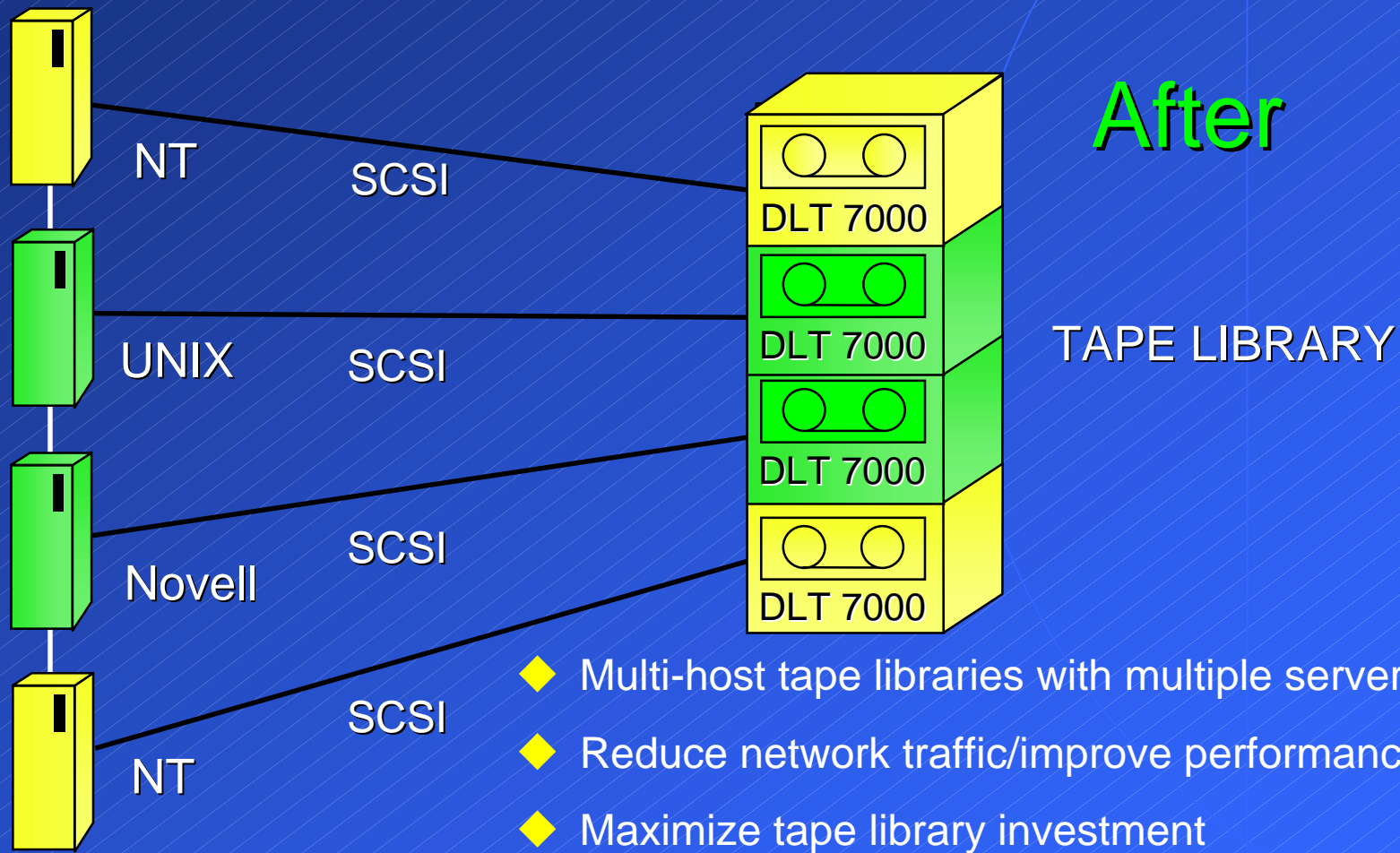
Library Speed = 20 MB/sec

-10 MB/sec  
Lost Throughput



28

# Step 4: Identify Priorities (Multi-Hosting)



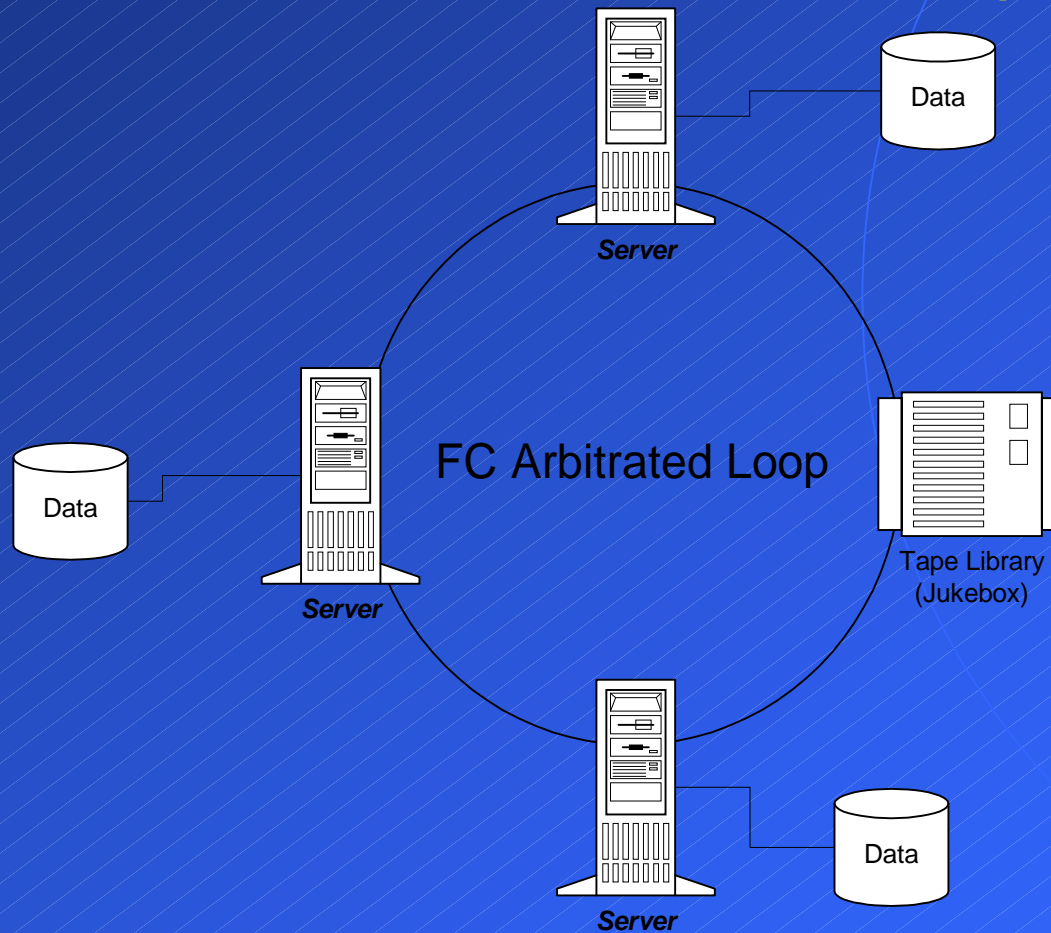
- ◆ Multi-host tape libraries with multiple servers
- ◆ Reduce network traffic/improve performance
- ◆ Maximize tape library investment

## Step 4: Identify Priorities (Fibre Channel-Arbitrated Loop)

- ◆ 100 MB/sec speed
- ◆ Arbitrated - only one conversation at a time without redundant loops and hubs
- ◆ Not a bandwidth consideration, but latency of arbitrating connections
- ◆ 126 devices on one loop vs. 16 for SCSI
  - ◆ Current suggested/recommended is 3-6 servers
  - ◆ Must bring down entire loop to add/remove device
    - ◆ Or use a hub
    - ◆ Or use node-bypass circuitry (increases price)

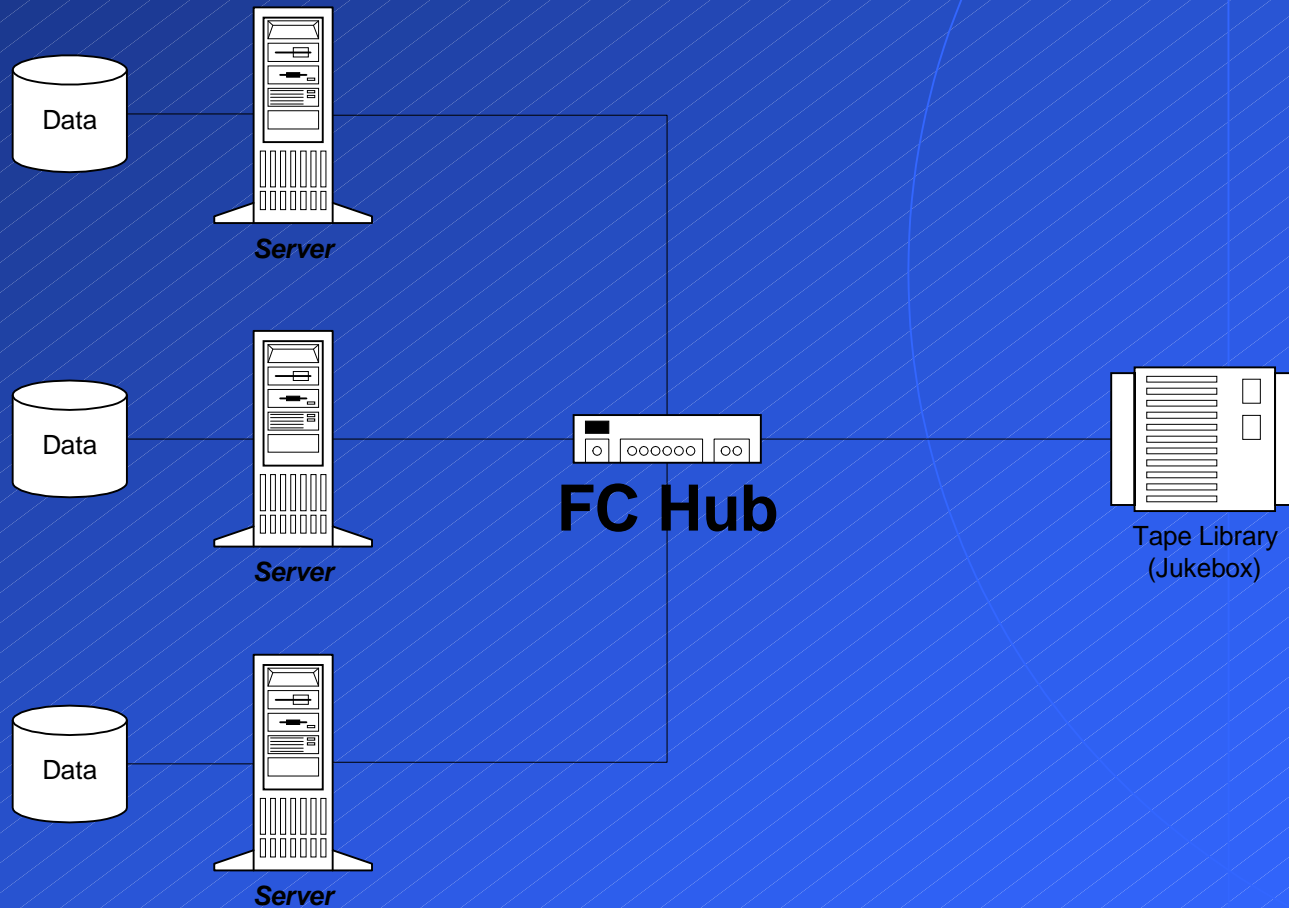
30

# Step 4: Identify Priorities (Fibre Channel-Arbitrated Loop)





# Step 4: Identify Priorities (Fibre Channel-Arbitrated Loop)

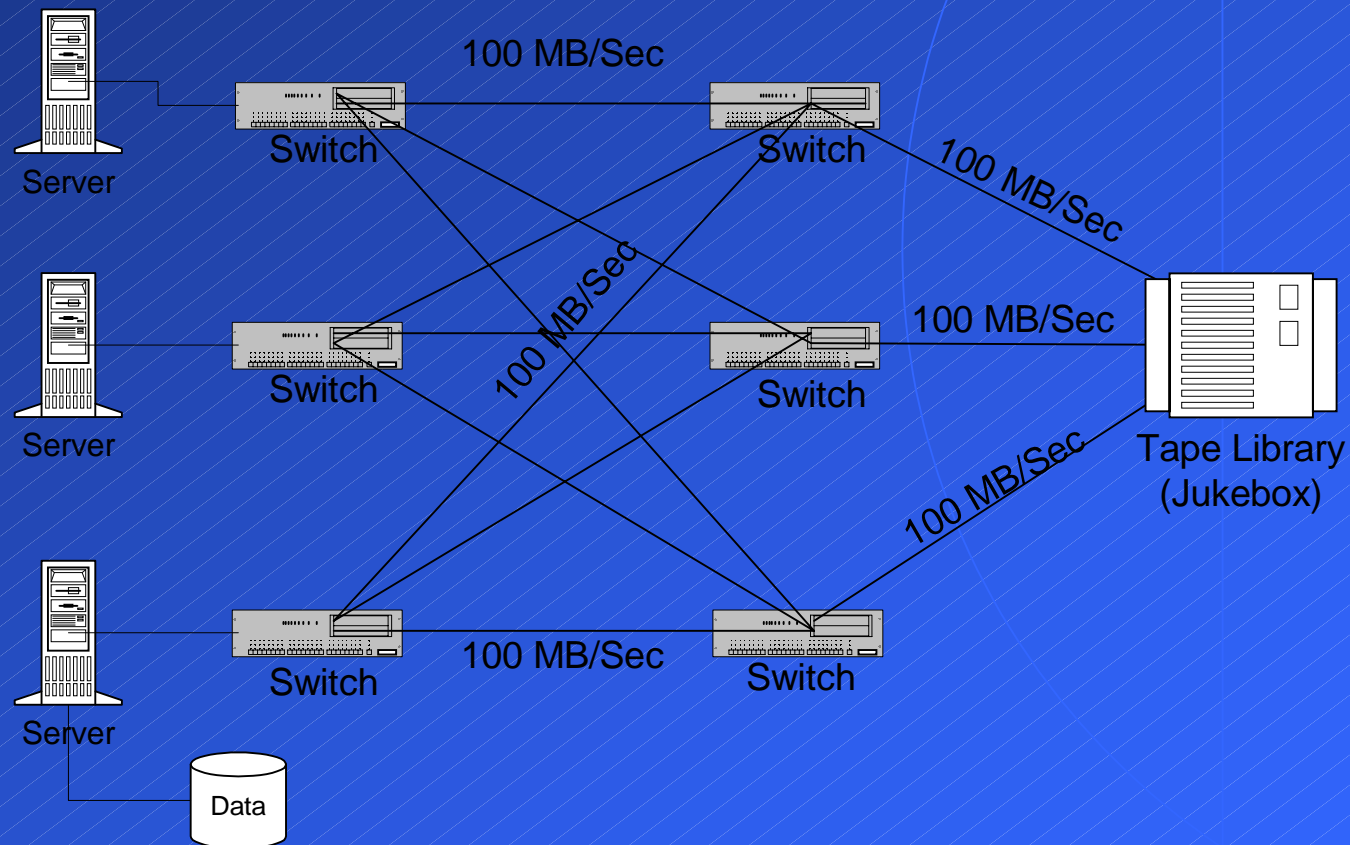


## Step 4: Identify Priorities (Fibre Channel - Switched Fabric)

- ◆ More expensive
  - ◆ Multiple switches, hubs and routers
  - ◆ Can help to have SCSI to FC converters at the end points, but the fabric in the middle is where the costs climb quickly
- ◆ Fast speeds
  - ◆ Multiple paths allow multiple channels of gigabit speed in system
    - ◆ 8 simultaneous 100MB/sec pipes through a 16 port switch
- ◆ Redundancy/Rollover
- ◆ Zoning

33

# Step 4: Identify Priorities (Fibre Channel-Switched Fabric)



34

# Step 4: Identify Priorities

## Comparison: SCSI vs. Fibre Channel

- ◆ SCSI
  - ◆ Risk is low
    - ◆ Time tested
    - ◆ Widely installed
    - ◆ Simple to use
    - ◆ Inexpensive - mostly just cost of cables
  - ◆ Limited to maximum distance of 25 meters from server
  - ◆ Slower than FC
    - ◆ SCSI III - 80 MB/sec vs. FC - 100MB/sec

35

# Step 4: Identify Priorities

## Comparison: SCSI vs. Fibre Channel

- ◆ Fibre Channel - Risk is higher, but is dropping
  - ◆ No standard...yet: FibreAlliance & SNIA
  - ◆ Greater distance – up to 10km on a single run
  - ◆ Heterogeneous systems are still 2-3 years away
  - ◆ Cost is higher, but dropping
    - ◆ Managed hub - \$500-700/port, Switches - \$1,500/port
    - ◆ Fully redundant switches (failover) - \$6,000/port
  - ◆ Zoning -
    - ◆ Security
    - ◆ Fail-over capabilities
    - ◆ Management / Load balancing

# What is the best for you?

- ◆ First question - how important is availability
  - ◆ - i.e. how much is it worth?
- ◆ ERP systems with verbose, highly redundant subsystems
  - ◆ Easily worth it
    - ◆ Easy to justify
    - ◆ Link aggregation can help scale up as need grows
    - ◆ Helps handle data transfer bursts
- ◆ Large amount of data, few number of servers
  - ◆ Multi-hosting may be best

# What is the best for you?

## Server Farms

- ◆ Greatest advantage is distance, or lack there of
  - ◆ SCSI
    - ◆ Distance isn't a problem
    - ◆ Simple – SCSI is an accepted standard
    - ◆ Inexpensive – just the cost of cables
    - ◆ Multiple servers can be connected to one tape library
    - ◆ Configure a server with more than one SCSI adapter
    - ◆ Multi-hosting - Multiple channels from one source
  - ◆ Fibre Channel
    - ◆ More expensive
    - ◆ Faster speeds - not that important if use multiple SCSI connects

38

# What is the best for you? Campus or MAN

- ◆ FC or SCSI
  - ◆ Distance is solved with FC connects or SCSI routers
    - ◆ Simplifies management
- ◆ Arbitrated Loop (AL)
  - ◆ May work if only have a few servers - dedicated loops
- ◆ Switched Fabric (SF)
  - ◆ If you have great value to the data on a 24/7 basis
  - ◆ Expensive, but reliable, flexible and dependable

39



# What is the best for you?

## Enterprise

- ◆ Combination based on distance and needed availability
- ◆ Can look to multiple “sites” within the organization
  - ◆ One main location, but each building it’s own SAN with switched, aggregated links to main data center
  - ◆ Will have to use the same vendor for everything to ensure interoperability for now
- ◆ Depending on amount of data to move, a combination of SCSI, FC-AL and FC-Switched may be the best

## Step 5: Choose Backup Software

- ◆ Evaluate (features)
  - ◆ Test
  - ◆ Implement

# Step 6: Implement hardware/software

- ◆ Identify Equipment/Needs
  - ◆ Main server (catalog)
    - ◆ Location
    - ◆ CPU, Memory, hard disk
    - ◆ Other applications to run on server?
  - ◆ Device servers
  - ◆ Client nodes

## Step 6: Implement hardware/software

- ◆ Implement strategy within backup window
- ◆ Local Backups are fastest
- ◆ Determine and test schedules
- ◆ Minimize network traffic

## Step 7: Implement Backup strategy

- ◆ Work with System Administrator to determine authorizations
- ◆ Define appropriate retention schemes. Double check.
- ◆ Refine backup schedules.

44

## Step 8: Put it into Production

- ◆ Monitor effectiveness of the Backups
- ◆ Survey users regarding restores
- ◆ Monitor growth - compare with predictions

45

# Summary - The 8 Basic Steps

- ◆ Inventory
- ◆ Examine Network/Hardware
- ◆ Document Your Requirements
- ◆ Identify Priorities
- ◆ Choose Backup Software
- ◆ Implement Hardware/Software
- ◆ Implement Strategy
- ◆ Put into Production/Monitor/Refine

46

# Questions?

