


Storage Decisions

Hosted by STORAGE SearchStorage.com

Designing a Backup Architecture That Actually Works

W. Curtis Preston
President/CEO
The Storage Group



Storage Decisions

Hosted by STORAGE SearchStorage.com

What will we cover?

- **What are the design options?**
 - LAN-based, LAN-free, Client-free, Server-free
 - NDMP
 - Using disk in your backup system
- **What should I do with them?**
 - Sizing your server

Storage Decisions

Hosted by STORAGE SearchStorage.com

What Backup Software Are You Using?

- A. Backup Exec
- B. Backup Express
- C. BrightStor Enterprise Backup
- D. BrightStor ArcServe
- E. Commvault/Galaxy
- F. OmniBack/Data Protector
- G. NetBackup
- H. NetWorker
- I. NetVault
- J. TSM
- K. Other

Storage Decisions

Hosted by STORAGE SearchStorage.com


How Satisfied Are You With Your Backup Software?

- A. It's rocks. I'd buy stock if I could.
- B. I'd change some things if I could.
- C. These guys better start getting a clue.
- D. What were those other products again?

Storage Decisions

Hosted by STORAGE SearchStorage.com

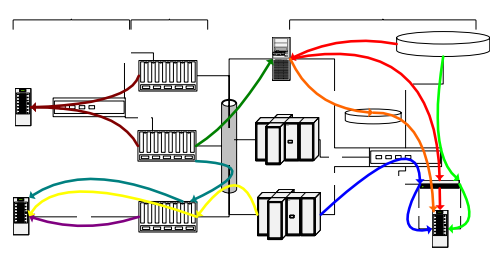
What Are the Design Options?



Storage Decisions

Hosted by STORAGE SearchStorage.com

SAN: LAN-free, Client-free, and Server-free backup
NAS: NDMP filer to self, filer to filer, filer to server, & server to filer



Storage Decisions Hosted by STORAGE SearchStorage.com

LAN-based Backups

- Standard method
- Central backup server with network clients backing up across the LAN
- Simplest, least expensive design

The diagram illustrates a central backup server at the top, connected via blue arrows to three client servers below. Each client server is also connected to a disk icon, representing its local data source.

Storage Decisions Hosted by STORAGE SearchStorage.com

LAN-free Backups

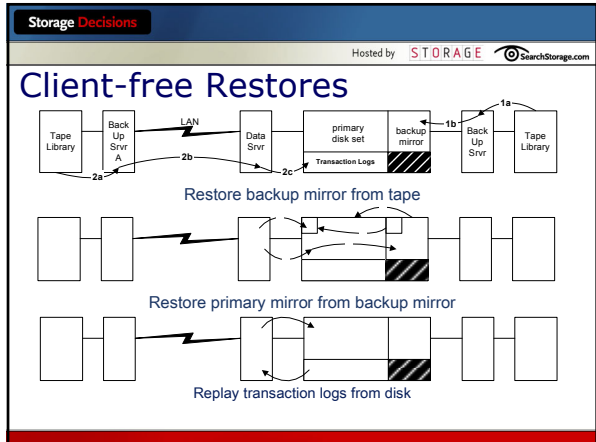
- How does this work?
 - SCSI Reserve/Release
 - Levels of drive sharing
 - Third-party queuing system
 - Restores

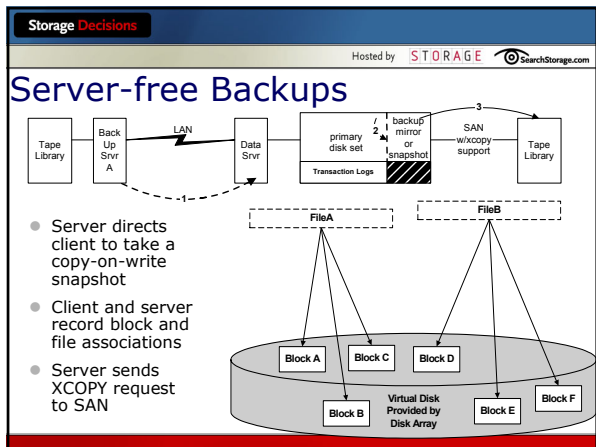
The diagram shows three disk servers on the left, each connected to an FC switch. The FC switch is connected to a SCSI/FC Router, which in turn connects to a Tape Library via multiple SCSI channels.

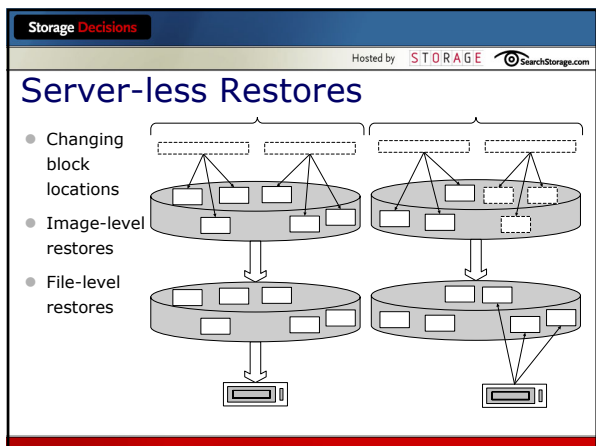
Storage Decisions Hosted by STORAGE SearchStorage.com

Client-free Backups

The first diagram shows a Data Svr connected to a primary disk set and a backup mirror. Transaction logs are sent to a Back Up Svr A, which then backs up to a Tape Library. The second diagram shows the establishment of a backup mirror between the primary disk set and a Back Up Svr. The third diagram shows the process of splitting the backup mirror and backing it up to a Tape Library.







Storage Decisions

Hosted by STORAGE SearchStorage.com

Backing up a Filer: NDMP

- Filer to self
- Filer to filer
- Filer to server
- Server to filer
- Similar to server-free backups

Storage Decisions

Hosted by STORAGE SearchStorage.com

Using NDMP

- **Level of functionality depends on the DMA and filer vendors**
 - Robotic support
 - Filer-to-Library support
 - Filer-to-Server support
 - Direct access restore support
 - Image-level backup

Storage Decisions

Hosted by STORAGE SearchStorage.com

If You Have Filers, How Are They Backed up?


- NDMP w/local tape drives on most filers
- NDMP filer-to-filer or filer-to-server
- Backup client running on the filer (usually a Windows or Linux filer)
- NFS/CIFS mount to a backup client/server
- Other

Storage Decisions

Hosted by STORAGE SearchStorage.com

Disk to Disk to Tape (D2D2T)


- ATA Arrays as low as \$3/GB.
- Use ATA as target for all incremental backups. (Full, too, if you can afford it.)
- For offsite storage, duplicate all disk-based backups to tape.
- More information & ideas available at my other talk: *Integrating Disk Into Your Backup System*



Storage Decisions

Hosted by STORAGE SearchStorage.com

Sizing the Backup System




Storage Decisions

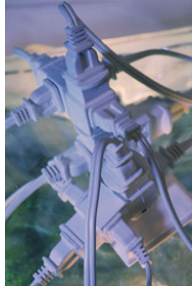
Hosted by STORAGE SearchStorage.com

Give it Enough Power

- Not enough tape drives
- Tape drives that aren't fast enough
- Not enough slots in the tape library
- Not enough bandwidth to the server



Don't Give it Too Much Power



- Streaming tape drives must be streamed.
- If you don't, you will wear out your tape drives and decrease aggregate performance.
- Must match the speed of the pipe to the speed of the tape
- You can actually increase your throughput by using fewer tape drives.

Server Size/Power

- I/O performance more important than CPU power
- CPU, memory, I/O expandability paramount
- Avoid overbuying by testing prospective server under load.

Catalog/database Size

- Determine number of files (n)
- Determine number of days in cycle (d)
- (A cycle is a full backup and its associated incremental backups.)
- Determine daily incremental size (i = n * .02)
- Determine number of cycles on-line (c)
- 150-250 bytes per file, per backup
- Use a 1.5 multiplier for growth and error
- Index Size = (n + (i*d)) * c * 250 * 1.5

Storage Decisions

Hosted by STORAGE SearchStorage.com

What is the predominant drive for open systems backup in your data center today?

- A. AIT
- B. SDLT
- C. LTO
- D. SAIT
- E. STK 9x40
- F. IBM 3xx0
- G. DDS
- H. VXA
- I. Other

Storage Decisions

Hosted by STORAGE SearchStorage.com

If you could start from scratch, what would you buy tomorrow?

- A. AIT
- B. SDLT
- C. LTO
- D. SAIT
- E. STK 9x40
- F. IBM 3xx0
- G. DDS
- H. VXA
- I. Other

Storage Decisions

Hosted by STORAGE SearchStorage.com

Number of Tape Drives
– All Tape

- **LAN-based Backup**
 - *Buy* twice as many backup drives as your network will support.
 - *Use* only as many drives as the network will support. (You will get more with less.)
 - Use the other half of the drives for duplicating.

Number of Drives - Disk/Tape Combo

● LAN-based Backup

- Buy disk system large enough to satisfy entire onsite retention period without deletion
- Buy enough tape drives to duplicate each night's backups. Duplicate each night's backups to tape, then take them out and send them offsite.
- Library should be large enough to hold three to four days of backups. (Only needs to hold duplicated tapes until they're sent offsite.)


Number of Drives - LAN-Free backup

- Most large servers have enough I/O bandwidth to back themselves up within a reasonable time.
- Usually a simple matter of mathematics:
 - 8 hr window, 8 TBs = 1 TB/hr = 277 MB/s
 - 30 10 Mb/s drives, 15 20 MB/s drives
- Must have sufficient bandwidth to tape drives
- Filesystem vs. raw recoveries
- Allow drives and time for duplicating.

Library Size - slots (all-tape environment)

- Should hold *all* onsite tapes
- Onsite tapes automatically expire and get reused.
- Only offsite tapes require physical management.
- Should automatically monitor library to ensure that each pool has enough free tapes before you go home.
- Watch for those downed drive messages.


Storage Decisions

Hosted by **STORAGE** 

Library Size -slots (disk/tape environment)

D2D2t	D2d2T
<ul style="list-style-type: none"> • Do all backups to disk • Disk big enough to hold all on-site copies • Disk-based backups automatically expire and space gets reused • Library needs to hold the latest set of copies (three or four days worth). • Only off-site tapes require phys. mgmt. • Should automatically monitor library <i>and disk</i> to ensure that each pool has enough free space before you go home • Watch for those downed tape and disk drive messages 	<ul style="list-style-type: none"> • Do some backups to disk • Library big enough to hold all on-site copies • Disk-based backups automatically expire after copied to tape • Disk needs to hold last night's backups before they go to tape • Only off-site tapes require phys. mgmt. • Should automatically monitor library <i>and disk</i> to ensure that each pool has enough free space before you go home • Watch for those downed tape and disk drive messages


Storage Decisions

Hosted by **STORAGE** 

D2D2t

- Do all backups to disk
- Disk big enough to hold all on-site copies
- Disk-based backups automatically expire and space gets reused
- Library needs to hold the latest set of copies (three or four days worth).
- Only off-site tapes require phys. mgmt.
- Should automatically monitor library *and disk* to ensure that each pool has enough free space before you go home
- Watch for those downed tape and disk drive messages

Storage Decisions

Hosted by **STORAGE** 

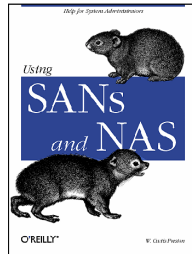
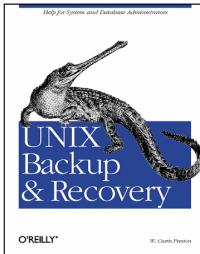
D2d2T

- Do some backups to disk
- Library big enough to hold all on-site copies
- Disk-based backups automatically expire after copied to tape
- Disk needs to hold last night's backups before they go to tape
- Only off-site tapes require phys. mgmt.
- Should automatically monitor library *and disk* to ensure that each pool has enough free space before you go home
- Watch for those downed tape and disk drive messages

Configuring Your Server

- **Back up all drives.**
- **Make sure you are streaming your drives.**
- **Create an automated monitoring system.**
- **Establish standards wherever possible, and use them!**

Resources



Resources

- **Directories of products to help you build a better backup system**
<http://www.storage-mountain.com>
- **Send questions to:**
curtis@thestoragegroup.com

More Questions?

Mr. Preston will be available for one-on-one questions at the "Ask the Experts" area of the Exhibit Hall this evening from 5:00-6:00 pm.

Sponsored by CA
