Designing a Backup Architecture That Actually Works

W. Curtis Preston
President/CEO
The Storage Group
curtis@thestoragegroup.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

What are the design options?

SAN: __________, __________, and ________________ backup
NAS: NDMP filer to self  filer to filer  filer to server, & ________________

LAN-based backups
   • Standard method
   • Central backup server with ________________
     clients backing up across the LAN
   • Simplest, least expensive design
LAN-free backups

How does this work?

- Reserve/Release
- Third-party
- Levels of drive sharing
- Restores

Client-free backups

Backup transaction logs to disk

Establish backup mirror

Split backup mirror and back it up
Client-free restores

Server-free backups
Server directs client to take a copy-on-write

- Client and server record block and file associations
- Server sends XCOPY request to SAN

Server-less Restores

- Changing block locations
- Image level

- File level restores
Backing up a filer: NDMP

- Filer to ______
- Filer to ______
- Filer to ______
- ________ to filer
- Similar to ___________ backups

Using NDMP

- Level of functionality depends on the ______________________ vendors
- Robotic Support
- Filer to Library Support
- Filer to Server Support
- Direct ____________ restore support
- Image level backup

Using disk

- _____-based disk arrays as low as $3/GB
- Virtual _______ _______
Using disk
- Use as a target for all _______________ backups. (Full, too, if you can afford it)
- For off-site storage, ___________ all disk-based backups to tape
- ______________ disk-based backups _____ ____________

Sizing the backup system
Give it enough power
- Not enough ___________ __________
- Tape drives that aren’t fast enough
- Not enough ___________ in the tape library
- Not enough bandwidth to the _______________

Don’t give it too much power
- Streaming tape drives must be _______________
- If you don’t, you will ____________ __________ your tape drives and ______________ aggregate performance
- Must match the speed of the ________________ to the speed of the tape
- You can actually increase your throughput by using _____________ tape drives

Server Size/Power
- I/O performance more important than ____________ power
- CPU, memory, I/O expandability paramount
- Avoid ________________ by ________________ prospective server under load
Catalog/database Size

- Determine number of ________________ (n)
- Determine number of ___________ 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 ________________, per backup
- Use a 1.5 multiplier for growth and error
- Index Size = (n + (i*d)) * c * 250 * 1.5

Number of Tape Drives – All Tape

- LAN-based Backup
- ________________ as many backup drives as your network will support
- Use only as many drives as ______ ____________ _______ _____________
- Use the other half of the drives for ______________________

Number of Drives – Disk/Tape Combo

- D2D2t
  - Buy disk system large enough to satisfy entire ________________
    ________________ without deletion
  - Library should be large enough to hold _______________ days of backups.
    (Only needs to hold duplicated tapes until ______ _______ ________)

Number of Drives – LAN-Free backup

- Most large servers have enough ________________ to back themselves up
  within a reasonable time
- Usually a simple matter of ________________
  - 8 hr window, 8 TBs = 1 TB/hr = 277 MB/s
  - 30 10 Mb/s drives, 15 20 MB/s drives
    - Must have sufficient ________________ to tape drives
- Filesystem vs. raw recoveries
- Allow drives and time for duplicating
Library Size - slots (all tape environment)
• Should hold __________ onsite tapes
• On-site tapes automatically expire and get reused
• Only offsite tapes require ____________________.
• Should monitor library via a script to ensure that each pool has enough free tapes before you go home
• Watch for those ______________________ messages

Library Size - slots (disk/tape environment)
• Do all backups to __________ wherever possible
• Library only needs to hold the latest set of copies (three or four days worth).
• Disk-based backups automatically __________________ and space gets reused
• Only off-site tapes require phys. mgmt.
• Should monitor library and disk via a script to ensure that each pool has enough free space before you go home
• Watch for those downed drive messages

Configuring your server
• Backup all drives
• Make sure you are __________________ your drives
• Create an automated monitoring system
• Establish standards wherever possible, and use them!

Resources
• Directories of products to help you build a better backup system
  http://www.storagemountain.com

Questions:
curtis@thestoragegroup.com