STORServer ®

Business Continuity Appliances

Disk-To-Disk Backup:

Making a Bigger D

Presented by Kelly J. Lipp

CTO STORServer, Inc.

Overview

- ◆ **Why a Bigger D?**
- ◆ **When to use Bigger D…**
- ◆ **When might disk be better than tape?**
- ◆ **Storage Configuration options**
- ◆ **TSM Parameter Selection**
- ◆ **Storage Pool Hierarchies**
- ◆ **Additional considerations:**
  - ▪ **Compression**

Why More Disk?

- ◆ **Remember: TSM was the first D2D backup solution: we've always done this!**

- ◆ **Market perception that disk is faster and more reliable than tape.**

- ◆ **Market perception that disk is the only way to solve backup problems.**

- ◆ **The price of disk has decreased making the use of disk more attractive.**

- ◆ **Bottom Line: are you having a problem with your current TSM solution that will be solved by the addition of disk?**

When to use more disk…

- When more disk is available.

- When more disk is cheaper than more tape.

- When it is true that more disk will make the backups faster or more reliable.

- **Business reasons, not technology should drive the decision.**

Is Disk really better than Tape?

- ◆ **Intrinsic technology is very similar.**
- ◆ **Top LTO3 speeds similar to disk spiral transfer rates.**
- ◆ **When backing up large objects (Exchange, SQL, Oracle, etc.) tape will be as fast, or perhaps faster than tape.**
- ◆ **Tape technology uses hardware compression.**
- ◆ **400 GB SATA Disk: 180 € – 200 €**
- ◆ **LTO3 Tape: 80 € (plus amortized cost of drive/library)**
- ◆ **OK, disk is random, I'll give you that!**
- ◆ **Fundamentally, disk is not better than tape!**

When is disk better?

- ◆ **If the bulk of your restores are single files from yesterday, disk wins!**

- ◆ **Perhaps for backing up large file servers where tape mounts might be problematic.**

- ◆ **Anyone with real life restore experience?**

Disk Storage Configurations

- Since the vast majority of our data will reside in primary storage pools on disk, hardware data protection is essential.

- Overall design must offer best space vs. performance option.

- Continue to use copy storage pools for large and small disasters.

- TSM features designed for dynamic disk environments. This talk assumes dedicated environments.

SATA vs. SCSI

- **SCSI MTBF much longer than SATA**
- **SCSI intrinsically more reliable**
- **Standard SATA drives not rated for 100% duty cycle environments.**
- **RAID Edition SATA drives available from some manufacturers.**
- **SATA: inexpensive, SCSI: not.**
- **Technologies appear to be converging, however.**
- **Choice: Generally SATA for cost.**

RAID Configurations

- **Must RAID, but the old RAID rules still apply: most RAID options are not optimized for Write…**
- **Must consider duty cycle of drives.**
- **RAID 1 Still the best, but most expensive.**
- **Clever RAID 5 solutions with Write Back Cache offer most cost effective solution.**
- **Choice: Large Write Back Cache and RAID 5 with spare drives.**

Storage Pool Options

- ◆ **File vs. Disk Device Class**
- ◆ **Should we front a file device class pool with a disk device class pool?**
  - ▪ **Reduces mount limit requirements.**
  - ▪ **Reduces the number of file volumes (in the short run).**
- ◆ **Should we have a tape pool behind our disk pool?**

Answer: File Device Class Pool

- ◆ **Since we get reclamation using a file device class, and we don't with disk device class we choose file device class storage pools for large disk environments.**

- ◆ **Still might front large file device class pool with a disk device class pool.**

File Device Class Configuration

- **Define devclass onlinefile devtype=file**
  - Maxcapacity: smaller might mean more reclamation, more efficient use of space. Too small, too much reclamation. Strike a balance between reclamation and space efficiency.
  - Directory: with 5.3, multiple directories are allowed making it easy to expand.
  - Mountlimit: if the associated pool is the target of a copy group, this value must reflect the number of simultaneous backups sessions expected.

Storage Pool Configuration

- ◆ **Define stgpool onlinefile filedev1**
  - ▪ **Migdelay: set to file retention days.**
  - ▪ **Migcontinue: Probably set to NO.**
  - ▪ **High/LoMig: 99/98.**
  - ▪ **Maxscratch: divide maxcapacity into total storage available. Or use a space trigger.**
  - ▪ **Other parameters as you would normally.**

Space Triggers

- **Using scratch volumes is less efficient than pre-defining volumes or using a space trigger.**

- **Space Triggers Available for automated, controlled Storage Pool expansion by adding volumes.**

- **Set Maxscratch on the pool lower and use a space trigger on the pool to expand.**

- **Define spacetrigger stg spaceexpansion=25 stgpool=onlinefile**
  creates additional 5GB volumes when existing volumes are filled to 80%

Defining Volumes

- **Define volume onlinefile d:\of.dsm numberofvolumes=50 formatsize=5000M to predefine volumes to the pool.**

- **This command creates 50 volumes of form d:\df001.dsm where df is used as a prefix.**

- **Using space trigger or manual volume definition dramatically reduces disk fragmentation and will improve performance.**
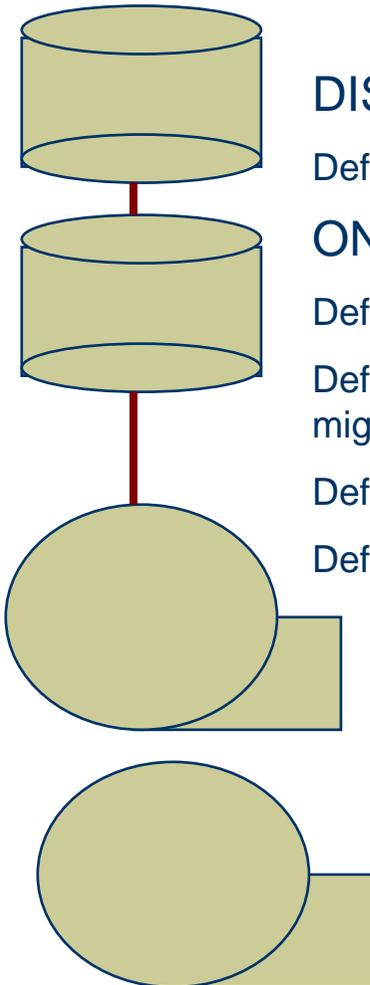
Hierarchy Considerations

- ◆ **Copy Group Target: Disk or File?**
- ◆ **Disk Advantages: Fewer mountlimit issues.**
- ◆ **File Advantages: No migration issues.**
- ◆ **Answer: Personal preference.**

Sample Hierarchy

DISKPOOL: Disk Device Class

Define stgpool diskpool disk nextstg=onlinefile

ONLINEFILE: File Device Class

Define devclass filedev1 devtype=file directory=e:,f: maxcap=5000M

Define stgpool onlinefile filedev1 highmig=99 lowmig=98 migdelay=15 migcontinue=no maxscratch=100 nextstg=onlinetape

Define spacetrigger stg spaceexpansion=25 stgpool=onlinefile or

Define volume onlinefile e:\of.dsm formatsize=5000M numberof=50

ONLINETAPE: Tape Device Class

Define stgpool onlinefile lto maxscratch=100

COPYPOOL: Tape Device Class

Define stgpool copypool lto pooltype=copy maxscratch=100

Considerations

- ◆ **Compression: Tape drives compress data, disk drives do not.**

- ◆ **OS Compression: CPU overhead on the TSM Server.**

- ◆ **Client Compression: CPU overhead on each client.**

- ◆ **When sizing don't forget that you are used to tape compression!**

Considerations

- **Performance: limited testing comparing file vs. disk device class indicates that file device class pools are faster than disk device class pools.**

- **Tape and disk are about the same speed on large objects. Disk wins when objects are small.**

Summary

- ◆ **TSM is still the best D2D solution in the marketplace today, especially with the changes in 5.3.**

- ◆ **Because TSM uses a pool based storage model, D2D solutions are not as critical to our success, but are still very useful to us.**

- ◆ **Design D2D2T solutions based on sound business justification.**