Staying afloat in today’s Storage Pools

Bob Trotter
IGeLU 2009 Conference
September 7 - 9, 2009
Helsinki, Finland
Topics to Discuss

• Background
• Storage Today
• What we do
• Storage Issues
• Trends
Background

• We don’t have DigiTool
• Digital Library of Georgia (DLG) Archive
• W. J. Brown Media Archives
• Special Collections Building
• Technical Stuff
Background: DLG Archive

• **Georgia Government Publications**
  – State Depository
  – Digitizing since 1994
  – Around 40,000 documents

• **Georgia HomePLACE**
  – Image Collections from around Georgia
  – Funded by IMLS

• **Digitization Guidelines**

Image from Sanborn Maps Collection
Welcome to the Digital Library of Georgia

The Digital Library of Georgia is a gateway to Georgia's history and culture found in digitized books, manuscripts, photographs, government documents, newspapers, maps, audio, video, and other resources.

The Digital Library of Georgia connects users to 500,000 digital objects in 105 collections from 60 institutions and 100 government agencies. Though this represents only a fraction of Georgia's cultural treasures, the Digital Library of Georgia continues to grow through its partnerships with libraries, archives, museums, government agencies, and allied organizations across the state.

Based at the University of Georgia Libraries, the Digital Library of Georgia is an initiative of GALILEO, the state's virtual library.

Partners & Sponsors

- [GALILEO](https://www.galileo.info)
- Georgia Public Library Service (Board of Regents of the University System of Georgia)
- Georgia Humanities Council
- Institute for Museum and Library Services
- National Endowment for the Humanities
- The University of Georgia Libraries
- More Partners »
DLG Archive

• Newspapers
  – **Red and Black** student newspaper
  – **Georgia Newspaper Project**
    • Since the early 1950s
    • Microfilm scanners for digitization
DLG Archive

• Video

– Civil Rights Digital Library (CRDL)

• Digitized newsfilm, radio, and images
Background: Media Archives

• **Peabody Award Collection**
  - Over 40,000 Award winning Titles
    - Radio programs dating from 1940
    - Television programs starting in 1948

• **Newsfilm Collections**
  - WSB (Atlanta) – over 5 million feet of film
  - WALB (Albany) – over 1,600 cans of film
Digitized Video

- SAMMA Lite system for video digitization
  - Creates streaming derivative, and archival copy
  - Archival copy straight to LTO tape
  - Grant funded project working on 2,200 hours
  - Total of Around 60,000 hours of video
- Will require 600-800 TB archival storage
New Special Collections Building

• Break ground this summer
• Huge archival vault for collections and archival copies
• Data center
  – Processing/storing archival data
  – I have already been asked to provide specs for what hardware will be installed
New Special Collections Building
Background: Technical Stuff

• Talk about this a little later
Storage Today: What Is It

• Storage Array: The basic unit of storage
  – Controller tray
    • Contains the brains; the storage OS that handles the I/O, caching, RAID, etc.
    • Also has disks
  – JBOD: Just a Bunch Of Disks
    • Has to be attached to a Controller tray
    • Depending on Array type: 2 – 7 / controller
  – Available now with 1 Terabyte drives
    • Multiple Terabyte coming
Storage Today

• Many different layouts
  – DAS: Direct Attached Storage
    • Usually over Fibre Channel
  – SAN: Storage Area Network
    • Switch based collection of arrays
  – NAS: Network Attached Storage
    • Large arrays with built in computer front end
    • Accessable via NFS, Samba, etc.
Storage Today

• Tiered Storage / Hierarchical Storage
  – Data storage environment consisting of two or more kinds of storage delineated by differences in at least one of these four attributes: Price, Performance, Capacity and Function.
Tiers

• Performance / Price
  – Architecture of array
    • FC-AL: Fibre Channel – Arbitrated Loop
      – Fastest most expensive
    • SAS: Serial Attached SCSI
      – Not quite as fast or expensive
    • SATA: Serial ATA
      – Least Expensive; slightly slower than SAS
Tiers

• Example of 3 Tiered Storage
  – Tier 1: FC-AL array(s): DAS or SAN
    • Sun StorageTek 6140s or 6540
  – Tier 2: SATA array(s): DAS or NAS
    • Sun X4500(s)
  – Tier 3: Tape Robotic device
    • Sun SL3000 Modular Library System
Storage Today

• 1 TB != 1 TB
  – Vendor TB based on 1,000 Byte K
  – Actual TB (what computer reports) based 1,024 byte K
DLG Storage

- Sun Storage arrays: 3310, 3511, 6140
- Exceeded backup window
- When expanded with 3511s (~2005)
  - Had just recently retrofitted a room in MLC to expand our data center
  - Purchased two 3511s and another server
  - Set up an offsite mirror in data center at MLC
  - Use modified rsync for nightly syncing of offsite mirror
DLG Storage

– Also set up automated zipping of archive files
– Provide staff with scripts to unzip when needed
– Weren’t sure how much this would help, but on scanned text documents we get about 40% compression
DLG Storage

Current setup (DAS)

- Each side of mirror is:
  - One - 6140 controller tray with 16 x 500GB SATA drives
  - One 16 TB JBOD (16 x 1 TB SATA drives)

Just purchased two 12 TB Sun 2540s (12 x 1 TB SATA drives) for expansion
DLG Storage

- 24 TB total raw space for each side of mirror
  - 21.84 TB as the computer sees it
- Each mirror comprised of 3 RAID5 (single parity) volumes
- With hot spares and parity drives comes to 16.1 TB usable storage
- Average of 30% of space taken by recoverability overhead
Media Archival Storage

• Grant funded project working on 2,200 hours of the Peabody Awards collection
  – Originally designed to write archival version to LTO tape
  – Many problems with both creating tar files and writing the LTO tapes from windows box
  – Total archival output of project estimated at around 24 TB
  – Had money left in grant so started looking for dense, inexpensive arrays
Media Archival Storage

- Settled on 48 TB Sun X4500 Storage Server (Thumper) and purchased two
  - It is an array with built in server front end: 2 x dual core AMD; 16 GB Memory; 8 onboard disk controllers; customized version of Solaris x86; ZFS
Configuring the X4500

• Now we had to configure it
  – Two drives used for mirrored root file system for server.
  – Found lots of information on web and settled on a configuration recommended by several experts.
  – One large zpool with four 11 disk raidz2 groups, and two hot spares.
  – This gives you a total of just over 32 TB of available space (reported by the system).
Configuring the X4500

– That’s right, it went from 48 to 32 TB.

  • First 2 went to the root filesystem
  • Another 2 for hot spares
  • With raidz2 you have 2 parity drives for each of the four groups = 8 more
  • Total of 12, but 48 – 12 = 36.
  • But remember 48 != 48; 48 ~= 43.8; and 12 ~= 10.9.
  • So, 43.8 – 10.9 = 32.9
Capacity/Recoverability Dilemma

• With that much capacity loss, I wanted to try other configurations.
  – Only gained back ~3.6 TB
  – Decided it was not worth the recoverability hit
Our Future

• Consolidate all archives to 3 Tiered Storage configuration.
• With Hierarchical Storage Management software controlling it.
• Like the 3 Tiered configuration I mentioned earlier.
• Sun’s SAM-FS (Storage Archive Management-File System)
Tiered Storage

• Archiving and Staging
Trends

• Proactive data protection
• Self Healing
• Copy on Write
• Scrubbing

Self-Healing Data in ZFS

1. Application issues a read. ZFS mirror tries the first disk. Checksum reveals that the block is corrupt on disk.

2. ZFS tries the second disk. Checksum indicates that the block is good.

3. ZFS returns good data to the application and repairs the damaged block.
Trends

• Integration of Solid State Devices (SSD)
• Much Faster
• Also Much more expensive ($10,000 for 100 GB Read Flash Accelerator)
More Topics

• Storage Virtualization
  – the process of abstracting logical storage from physical storage

• iSCSI, IP SAN

• Deduping
Eat up with Acronyms

- RAID – Redundant Array of Inexpensive Disks
- SCSI – Small Computer System Interface
- JBOD – Just a Bunch Of Disks
- SATA – Serial Advanced Technology Attachment
- SAS – Serial Attached SCSI
- iSCSI – SCSI over internet connection
- HBA – Host Bus Adapter
Other Companies

- Isilon Systems
  - OneFS OS
  - On each node
  - Vertical Striping
Other Companies

• Xiotech
  – Intelligent Storage Element
  – Sealed DataPac
  – Factory remanufacturing
  – Comes with 5 Year Warranty
Questions?

rwt@mail.libs.uga.edu