

# Digital Collections

## Storing Stuff for Long Term

Alan Poston

Xyratex

# Assumptions

- This is no such thing as “long term” storage
- 3 to 5 years is tops
- Power and cooling are key considerations
- Fun with Math
- Technology selections
- Technology predictions (or lack thereof)

# Fun with Math

- 5,20, 50 PB are the design points.
- 5 PB with 10% retrieval rate is about 16 MB/sec ( $.5\text{PB} / (365 \times 24 \times 3600)$ )
- BUT, you have to store 5PB... so you need 160 + 16.
- To get to the next generation storage you have to copy this forward in 3 months or less ( $160 + 16 + 640$  or 816 MB/sec)
- X4 for 20 PB... 3.3 GB/sec and x10 for 50 (8.16 GB/sec)
- 176 MB/sec to a replicate system on another fault zone

# Flash, Disk, Tape

- Flash is still too expensive and it's not needed
- Tape still suffers from needing expensive robotics and latency is a big issue, it's more stable than disk (because it's not mounted all the time) and it consumes less power
- Disk is the current price/performance winner, especially if you can do predictive caching and leave most disks powered down most of the time

# Future Storage Technology

- I've quit trying to predict the death of disk
- I've almost quit trying to predict the death of tape
- I funded Holographic Storage 20 years ago and it's still not here
- Other solid state memory systems (biological, etc) may show up in 5 years.
- Replicated, power controlled rotating disk is the way to go for the next 5 years perhaps assisted with flash.
- WD recently announced Helium disk drives at probably 6GB capacity.