Archival Storage - Tape versus UDO

Introduction

The need for trustworthy, long-term archival storage of important business information has never been greater. As the volume of digital records grows and data retention regulations proliferate, organizations are scrambling to develop long-term data archive strategies that allow them to meet their corporate and legal obligations, to protect themselves against crippling litigation and to make the most of their valuable corporate assets.

Two of the most traditional archival storage technologies are magnetic tape and 5.25 inch UDO (Ultra Density Optical). Both of these can be used for archival storage, but the performance attributes of tape and UDO are very different so should be compared in the context of a professional archive environment. In addition, we look specifically at medical image archives to provide a real world analysis of performance in that environment.

For the purpose of this report, the technologies of focus will be UDO generation 2 (UDO2) and Ultrium LTO generation 4 (LTO-4).

Archival Storage Criteria

The priorities for a long-term data archive are not the same as those for on-line storage or for backup and disaster recovery applications. When comparing tape and UDO in high capacity automated libraries, it is necessary to consider those requirements that are most important for archival storage.

- On-line Storage
  - Uninterrupted availability
  - High performance read / write access

- Backup and Disaster Recovery
  - High media capacity
  - High performance read / write streaming
  - Low storage cost per GB

- Archival Storage
  - Data trustworthiness / authentication
  - Extended data longevity
  - Rapid read / write random access
  - Low Total Cost of Ownership over decades

Tape versus UDO

Tape storage has been designed primarily for backup and disaster recovery applications. Using tape for data archives is quite common since it provides professional quality hardware, high performance data streaming and the flexibility of removable media. However, tape does not offer the same level of data authentication, media longevity and high performance random access as UDO. In addition, using tape to store data for decades requires hardware and media maintenance considerations that can dramatically increase the Total Cost of Ownership for a tape archive.

Authentication

As with magnetic disk, tape is a rewritable media so does not provide the same level of audit trail authentication as Write Once optical. UDO uses true Write Once media that does not allow recorded data to be altered or rewritten. In addition, UDO media is inherently unalterable thereby guaranteeing that data does not change due to environment effects, etc. While some tape products do offer WORM emulation using a hardware lock to prevent the rewritable tape from being overwritten, this technique is simply not as secure as a true Write Once media for environments where data authenticity is paramount.

Media Longevity

Magnetic tape is a very good choice for high capacity, short-term storage such as backup and disaster recovery, but when used in long-term archive additional maintenance considerations come into play. Tape is a relatively delicate contact media, which degrades with use, can become physically damaged and is adversely affected by swings in environmental conditions. Data stored on tape can also be lost from exposure to magnetic fields. In order to ensure tape data remains available over extended periods of time, it is strongly recommended that tapes be frequently re-tensioned and data be periodically refreshed (read and rewritten). Managing re-tension and refresh cycles for hundreds of tapes written over many years is a complex and costly task with serious consequences if not managed properly. By contrast, UDO media is non-magnetic, highly stable, does not degrade with use and does not demand the extra burden of re-tensioning and refreshing. UDO media has a greater than 50 year media life, can be stored in a much broader range of temperatures/humidity and requires virtually zero media maintenance over decades of active use.

Environmental Specifications	UDO2	LTO-4
Operating Temperature	5 - 45°C	10 - 40°C
Operating Relative Humidity	5 - 90%	20 - 80%
Archive Temperature	-10 - 55°C	16 - 32°C
Archive Relative Humidity	5 – 95 %	20 - 80%

Access Performance

Many archive storage environments have a large number of concurrent users that require access to data within seconds. Tape struggles to meet high data access service levels because of very slow load and seek times. Tape takes minutes to access a single random file, whereas UDO can perform the same task in seconds. Reduced tape performance adds up to slower overall access cycles, resulting in many fewer possible access cycles per hour. UDO drive in an automated library can service 5 to 15 times more requests per hour than a similarly configured tape drive. Environments with many concurrent users will require more tape drives than optical drives to service the same level of access pattern frequency. The higher drive to media ratio required in a tape archive has a very direct and negative impact on the overall system cost.

Drive / Library Specs	UDO2	LTO-4
Load Time	5 sec	12 sec
Unload Time	3 sec	15 sec
Average Seek Time	25 msec	57 sec
Average Rewind Time	0 sec	54 sec
I/O Transfer Rate	12 MB/sec	120 MB/sec
Media Exchange Time[1]	6 sec	6 sec
Average File Random Access Time (media loaded)	25 msec	57 sec
Average Access Cycle (old media unload + new media load)	14 sec	148 sec
Access Cycles per hour	257 cycles	24 cycles

[1]6 seconds is expected exchange time on a Plasmon library. Many tape libraries are significantly slower, as they don't have dual pickers like the Plasmon library which reduces the media exchange time.

- **A single UDO2 drive provides 11x more access cycles than a single LTO-4 tape drive.**

Total Cost of Ownership

At first glance, the cost per GB for tape storage appears to be extremely low, but in order to get an accurate picture, it's necessary to consider the Total Cost of Ownership for additional hardware, media and administration essential to a long-term tape archive. A tape archive will require more drives than a UDO archive to achieve the same level of level of access cycle performance. Additional tape media is required each time data is refreshed, and for redundant media copies made for insurance against physical tape damage. Significant additional administration and drive resource is also needed for managing a re-tension and refresh programme in order to ensure that data written to tape remains accessible. Adding up the additional hardware, media and administration costs results in a much higher TCO for tape archives than first imagined. Not only does UDO offer a very competitive TCO when compared to tape in a long-term archive, it does so with greater authentication, faster access performance and much simpler media management.

Healthcare Environment – At a Glance

It is very interesting to compare the use of an LTO-4 based archive to a UDO2 based archive. Although it is important to consider both technologies from a write ingestion and random (recall) read performance standpoint, it is important to note that both technologies are fronted by Disk technology for caching in all *Best Of Class* archive solutions. Hence, write ingestion performance is typically masked by the Disk cache on the front end, and the user application will see Disk performance when storing images. However, for the situation where data needs to be recalled from the actual LTO-4 or UDO2 media, this performance can not be masked by the Disk cache (unless the disk capacity is matched to the archive capacity). So, below we look at several scenarios for comparing recall performance of both technologies in the context of Healthcare Imaging.

For the purposes of this report, we have chosen an average study size of 125 MB (consists of multiple files typically). Clearly, many of the modalities will create much smaller files or studies, but this size is more representative of the larger images created by CT, Cardiac Angiography and Breast MRI.

Also, for simplicity of comparison purposes, it is assumed that all files of a single study are stored sequentially on the media. It should be noted that this is not an issue for UDO2, since it can random access seek in 25msec, but for tape this is an important consideration. Also, it should be recognized that sequential access of the files in a single study can degrade if the tape drive requires a repositioning cycle before reading a file (i.e. unable to stream through each file). This can occur when the host computer is unable to speed match with the drive due to various host based overheads (both software and hardware induced). This study will assume that these degradations are not encountered in this study.

Time to Read a Study	UDO2	LTO-4
Media not loaded in Drive	25 sec	149 sec
Media loaded in drive	11 sec	58 sec
Number of studies/hr recalled (on one media)	327	62
Number of studies/hr recalled (random distributed in library)	144	24

Summary

To properly compare technologies, they need to be put in context. Tape has many strong qualities that make it an ideal backup and disaster recovery technology, but to use it in a long-term archival storage environment demands a number of performance and economic trade-offs. By contrast, UDO was designed specifically for archival storage so meets the technical and economic requirements of a professional archive strategy.