

## Media Considerations for Record Retention and Archiving

Print this Page

*OK, you've figured out which records to keep. Now, you just need to decide what you'll be keeping those records on.*

**Table One – Leading Device Characteristics for Long-Term Document Storage\***

Device	First Introduced*	Aerial Den.	Capacity Per side	Write D. R.	Read D. R.	Access Time	Media Price
5.25" DD	1988	2.2Gb/sq. in.	4.55 MB (14X)	4 MB/s	MB/s	25 MS	\$85 est.
UDO – 30	Planned 2003	7.4Gb/sq. in.	15 GB	4 MB/s	8 MB/s	25 MS	\$60 est.
UDO 120	Planned 2007	29.6 Gb/sq. in.	60 GB	9 MB/s	18 MB/s	25 MS	\$80 est.
Large FF	1984 (14")	Withdrawn	N/A	N/A	N/A	N/A	N/A
DD	1987 (12")	2.2Gb/sq. in.	15GB	3 MB/s	6 MB/s	96 MS	\$50 est.
CD	1981		650MB	.7MB/s	1.35 MB/s	190 MS	\$1
DVD	1995		4.7 MB	2.4 MB/s	4.8 MB/s	190 MS	\$4
Mag Tape	1952	Approx 0.3 Gb/ Sq. in	20GB/cartridge Uncompressed	20-40MB/s	20-40MB/s	16-20 Seconds**	\$20
Mag Disk	1956	Approx 25 Gb/ Sq. in.	60 GB ATA	150MB/s	150MB/s	1 MS	NA

\* There are many products in each class. I have shown specifications for currently shipping products. The "First Introduced" column shows the year of first introduction of the product class.  
 \*\* Automated Library time to first data. Subsequent accesses can be a few seconds depending on record location.

Records are created in anticipation of future disputes -to record activities/transactions at or near the time of their occurrence and thereby establish a people- independent memory of those events. For every business, its records constitute the corporate memory.

Historically, with the increase in both the volume of business

activities and the length of time those activities have to be available for recall, the reliance on records has become increasingly essential to achieving accuracy, reliability, and trustworthiness in the recollection of facts and the reconstruction of activities and events. This recollection process first sought to prove the positive-to retrieve from a large number of records a few needed records. For example, locating a signed shipping ticket that showed the receipt of goods and, therefore, the legal obligation to pay for them.

Now, in our increasingly regulated and litigious society, there is also a need to prove the negative. Records are vitally important to proving the negative because they can demonstrate, through document-based information, that a pattern of behavior was so consistent as to preclude the possibility of any validity to an allegation of procedural wrongdoing.

Although one of the basic tenets of America's judicial system is innocence until proof of guilt, such is not the case with regulatory compliance. Through its records, a company must be able to demonstrate that it did not do what it is alleged to have done by the regulatory agency. It is this increasing need to prove the negative that has affected many well-known companies recently including Arthur Andersen, Enron, and Microsoft.

The processes that produce e-records must be auditable to verify the records recorded are authentic and have not been altered. The storage media must be persistent<sup>1</sup> and certain other rules apply. Records-whether ink on paper, photographic, or electronic-must be admissible in a court of law. Laws and directives that govern records admissibility were written when records were composed of ink on paper. Courts seldom make special allowances for e-records and the penalties for noncompliance can be severe. Also there are legions of cases where electronic search has uncovered records a defendant had long forgotten. Our industry has presented the capability to keep too many records too long and to find them far more easily and quickly than with printed records. Retention only long enough to be legally compliant is important.

E-records must be admissible in court and that requires specific knowledge that most IT staffs do not have. To quote from an April 2002 email from a records manager friend in a mid-sized

financial services company, which I think accurately states the case, "To be fair and accurate, it is important to note that for the most part neither IT nor the business units know about or understand regulatory record keeping requirements. I have to begin with explaining that there are retention requirements and then there are storage requirements and then explain them in terms of storage media requirements, accessibility, retrievability, and how to preserve the integrity of the records."

### **Changes in Records Retention Technology**

For about 300 years only personal testimony was admissible in court. Commencing in the nineteenth century written interrogatory became admissible when a witness could not be present. Until around the middle of the last century the written interrogatory had to be ink on paper and be capable of passing an authenticity test (e.g., the paper depicted exactly what was said). Then microfilm and microfiche became admissible as long as it could be proven that they authentically reproduced the paper record pictured. Although the technology is older, the application of micrographic technology to paper documents is an early twentieth century phenomenon. The enormous advantage of film and fiche was to shrink the volume of space required to store a burgeoning number of records and to increase the contemporary state of customer service. There is still a significant film/fiche market because records on that media are universally admissible in jurisdictions worldwide and because film and fiche present the best long-term archival qualities with the greatest probability of authenticity. However film/fiche technology has inherent severe limitations, especially in our Web-world. It does not show color, animation, or audio. It is one-dimensional, does not allow very convenient linking, and is slow and labor intensive by today's norms.

Many of the integrators who specialized in the application of photographic technology to records storage migrated to optical technology when it became available, roughly between 1983-85. The virtue of the early optical technology was that the record, once burned, was immutable (the early technology actually physically changed the media), the medium was denser, and hence the aerial density<sup>2</sup> was greater, allowing more records to be stored in a finite space. It took about ten years to convince courts that records stored on optical technology would be as authentic as photographic technology and therefore admissible in court. Indeed the effort was so successful that in very many places the term "optical" was actually written into the law, which confused the situation for later. Optical technologies include compact disks (at the low end), DVDs (digital versatile disk), and 5.25-inch diameter optical disks and media (see Table One). Large form-factor optical disk recorders (12- and 14-inch diameter) and media are near end of life except that maintenance, consumables, and a few replacement units keep these products alive.

Plasmon Inc., a Cambridge, England and Colorado Springs, CO, company has corralled all of the large form-factor business. In addition, Plasmon is poised to release a new disruptive 5.25-inch diameter optical disk called UDO (Ultra High Density) based on blue-violet lasers (4.05 nm). Blue lasers record more densely than the red lasers currently in use. The development combines refinements in lasers, lenses, and media and plans to provide a rewritable and read-only version. Table One provides a few high-level specifications on UDO, which has a roadmap that takes it from 30 GB in 2003 to 60GB in 2005 to 120GB in 2007. The roadmap is based on the UDO standard announced by Sony, Hewlett-Packard, and Plasmon in 2000. More information is available at [www.plasmon.com](http://www.plasmon.com). The 30GB version will probably make its targeted availability date of August 2003.

Optical recording technology never caught on in the production IT community because relative to

magnetic technologies, aerial densities were substantially lower dictating a higher price per megabyte, the media was not rewritable (records storage needs a non-rewritable or WORM media to assure authenticity), and the transfer rates of data to/from media was slow relative to magnetic storage. An important exception is CDs that are now giving way to DVDs, both of which were created for the entertainment industry but won (and are winning) niches in the data storage industry.

## Legal Requirements and Technology

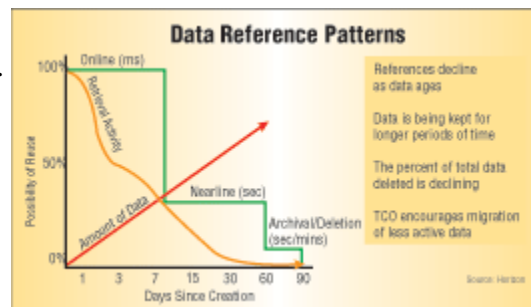
During most of the 1990s and continuing to date there have been efforts to change the laws and agency directives to decouple permissible storage technologies by name (optical) and substitute the criteria the technology must meet at the system level. For example, 17CFR275.204.2, the SEC ruling that states electronic record keeping requirements for investment advisors general statements like the investment adviser must:

- Arrange and index the records in a way that permits easy location, access, and retrieval of any record.
- Provide promptly any of the following that the SEC may request:
  1. A legible, true, and complete copy of the record in the medium and format in which it is stored.
  2. A legible, true, and complete printout of the record.
  3. A means to access, view, and print records.
- Store separately for the time required for preservation of the original record, a duplicate copy of the records on any medium allowed by the SEC for investment advisers (paper, micrographic, or electronic media).

The investment adviser must establish and maintain records to:

- Maintain and preserve records so as to reasonably safeguard them from loss, alteration, or destruction.
- Limit access to records to properly authorized personnel and the Commission (including its examiner and other representatives).
- Reasonably ensure that any reproduction of a non-electronic original record on electronic storage media is complete, true, and legible when retrieved.

17CFR275.204.2 is fairly typical of the directives that cover a large segment of the financial services business. Nowhere does it call out optical storage technology. In addition to financial services, most other industries have their own sets of laws and directives.<sup>3</sup> Before I leave this subject, it should be pointed out that not all the directives and laws agree and there are a plethora of them.



## A Few Vendors

In 1996, Storage Technology Corporation negotiated a favorable environment with the SEC that allowed storage of financial documents and checks on WORM magnetic media that was legally compliant. Images of the front and back of paper checks (about 50 KB in Greyscale) must be stored for seven years. Today, STK stores about 70% of the checks transacted in the U.S. on its magnetic tape drives equipped with the "VolSafe" (WORM) feature in automated libraries.

StorageTek, having captured the bulk of the check and financial document storage markets, has moved on to other applications in email storage, storage of digitized video, medical records, insurance applications, and various other forms of image capture, although a spokesman recently indicated it intends to refocus on document applications. The only other major magnetic tape vendor that supports a WORM tape is Sony on its new AIT-2 and AIT-3 subsystems.

On March 18, 2002 Network Appliance announced the first of the NearStore appliances that were pre-announced in December of 2001. The first appliance is the R100. The NearStore appliances are targeted at low-performance requirements where cost is a major issue. The initial usage highlighted by Network Appliance is an economical two-stage backup device where data is moved to the R100 as a low-cost, quick restore (fast recovery) device. Subsequent to that, data can be copied from the R100 to tape as needed by business requirements. Data can be moved to the R100 from any storage systems that use a file representation of data. A long list of partners has certified their software will operate with NearStore (obviously, the major backup programs are included). More information on the R100 can be found at [www.netapp.com](http://www.netapp.com) or on the Evaluator Group subscription website, [www.evaluatorgroup.com/esol](http://www.evaluatorgroup.com/esol) in the NAS products section.

On April 29, 2002 EMC announced Centera Content Addressed Storage system, an integrated hardware and software solution targeted for what EMC refers to as the "fixed content" market. Fixed content documents include items such as financial records, photographs, medical x-rays and MRIs, video, seismic data, books, and manuals. EMC classifies this as a new storage category, which it has named Content Addressed Storage (CAS). The basic architecture for Centera was acquired by EMC as part of its acquisition of FilePool, a Belgian software company, in April 2001. More information on Centera is available at [www.emc.com](http://www.emc.com) or at the Evaluator Group subscription website listed above.

**Source:** National Media Labs

**Key:** Green is acceptable, Yellow may be marginal, and Red will probably not work.

		Media Comparison																
		Magnetic Tape							Optical Disk				Paper	Film				
Length of Storage: Based on Media available in 1995	Life Expectancy of Various Information Storage Media: For Storage at 20°C (68°F) and 40% RH																	
	103	Data D-2	Data D-3	3480	3480/3490n	DLT	Data 8mm/Data VHS	DDS/4mm	QIC/QIC-100	CD-ROM	WORM	CD-R	M-O	Newspaper (high lignin)	High Quality (low lignin)	"Permanent" (buffered)	Medium-Term Film	Archival Quality (Silver)
1 week	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
2 weeks	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
3 months	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
6 months	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
1 year	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
5 years	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
10 years	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
15 years	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
20 years	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
30 years	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
50 years	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
100 years	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
200 years	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green
500 years	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green	Green

Most document storage applications employ hierarchical storage of two technologies, magnetic disk and optical disk or magnetic disk and WORM-equipped magnetic tape. Documents are usually written to the two medias simultaneously. The magnetic disk is to handle the large volume of inquiries that occur immediately following distribution of documents such as statements (typical applications allow for enough magnetic disk for ninety days storage) and optical disk or tape to store the document for the rest of the legal retention period and/or for archives. A few employ all three, magnetic disk, optical disk, and magnetic tape but that in my opinion that is overkill and reserved for unusual applications. To date, a magnetic disk-only approach like Centera has not become widely used because

the necessary integrated software was not available, regulatory approval in the U.S. has not been requested (EMC and its Centera Partners are trying to remedy that), and it is unclear how it will

compete on a life-cycle cost of ownership basis with a hierarchical approach.

All media exhibit different persistency characteristics (defined in footnote 2). Dr. John Van Bogart of National Media Labs (NML) published a series of vendor independent tables of Life Expectancy of different media available in 1995 over a range of environmental conditions. I have been unable to find these tables on the Web but many are available in Appendix Five of Authentic Electronic Records by Dr. Charles Dollar.<sup>4</sup> One of the charts is reproduced above.

## Summary

Records and especially object storage will become an evermore important part of IT professional skill set. Storage professionals will need to be cognizant of three factors when choosing media:

- Legal records retention requirements
- Admissibility of the media by the courts
- Life expectancy of data on the media (Chart Two)

*Jack Scott is the Managing Partner of Evaluator Group, a Denver based storage analysis firm. He is also Chairman of EmTAG, (Emerging Technologies Group), the technology headlights of AIIM (Association for Information and Image Management).*

## Notes

1. In this context "persistent" means that once recorded the data does not change.
2. Aerial density is one of the ultimate metrics by which all storage technologies are compared. It simply means bits per square area of measurement. Absent any unusual supporting subsystem requirements like extreme cold, it usually follows that greater aerial densities yield lower cost of storage.
3. Consult [www.cohasset.com](http://www.cohasset.com) for a more comprehensive treatment of this subject.
4. Authentic Electronic Records: Strategies for Long-Term Access by Charles Dollar, pp.217-222 available through Cohasset Associates, [www.cohasset.com](http://www.cohasset.com).

