What is the Next Relevant Storage Technology for your Archive Tomorrow?

Author: David Thomson
Senior VP Sales and Marketing

Version: 1.0
Date: **September**, 2022
Status: Released
Distribution: General Distribution

www.qstar.com

Introduction

We have already looked at tape, object storage and cloud storage as the three main archive options for today. These will continue into the future with object storage and cloud storage built to use the latest higher capacity disk or flash drives and automatically migrate from old to new, while LTO tape has recently added 5 new generations to their program going from today's LTO9 with a roadmap to LTO14.

Besides these existing technologies, several new (or improved) technologies will be vying for market share in this massive and massively growing market sector, over the next decade. These include improved optical technologies, holographic storage, glass storage and DNA-based archive technologies.

Why Optical Storage?

Optical Storage, for many, is a past technology, not a future technology, with its hay day being the late 1980s and into the early 2000s. In 1995, the original mass-market optical archive solution was CD-R (compact-disk recordable) manufactured by Phillips which was a WORM (write once read many) format, followed by CD-RW (or CD+RW) which was a rewritable media. Limited to 650MB, all the original development was on record (write) speeds. This was followed by DVD at 4.7GB and Blu-ray (BD) at 25GB initially with higher capacities later. The most important feature for optical use, was backward compatibility, in that the current generation drive could always read all formats back to CD.

In addition, "enterprise-class" optical, which was always in a cartridge, revolved around 5.25-inch phase-change (PC) or magneto-optical (MO) technology providing five generations of its technology from 650MB (1987) to 9.1GB (2001) before Ultra Density Optical (UDO) replaced it with 30GB and later 60GB media (2007). The most recent enterprise optical solution (starting in 2013) has been Sony's Optical Data Archive (ODA) with three generations – 1.5TB, 3.3TB and 5.5TB, based on Blu-ray disk technology. All enterprise-class optical offers both rewritable or WORM media options.

Optical storage creates "spots" on media using lasers at a high-power setting and then reads those spots using a lower power setting. As the laser heads fly over the media, they can be re-positioned very quickly allowing for random access (unlike tape storage which is always sequential). Optical, in the past, struggled mainly with performance as the time taken to write data was slow in comparison to tape, the maximum BD-R write performance is 72MB/sec. With ODA it is a more respectable 150MB/sec. The other issue is cost of media – ODA is around three times the cost per TB of LTO.

Next Generation optical is being championed by Folio Photonics and promises its Folio Disk media (called DFD) at $3 per TB (LTO is currently at $8 per TB), so if that is true one of the significant disadvantages of previous optical solutions is removed. Capacity is seen as less important, if they can keep cost of drives and media very low – so initially capacities will start at 750GB and cost of media at $10 per disc (or $13 per TB). Total performance can be improved through the use of larger numbers of drives or optical heads writing in parallel. DFD will only be available in a WORM (write once read many) format and its solutions are scheduled to be available 2024 / 2025. Crucially, it will need library partners to integrate drives and media and create multi-PB solutions to stand any chance of competing in the archive market space.

Why Holographic Storage?

Holographic storage was first shown in 2005 by InPhase Technologies, they went out of business in 2011 after failing to make a commercial solution. In 2009, they were aiming at a solution providing 300GB of capacity and 20MB/sec (very modest compared to today's technology). Apple now owns the assets of InPhase, although there is no indication they are using this to create a storage product.

GE Global Research in 2011 was also researching holographic storage and discussed at that time a 500GB of capacity on a 120mm disk with a roadmap into multiple TB media.

Microsoft Research are currently investigating the possibility of using holographic storage (HSD) in the future. Holographic storage stores data in 3D and therefore could be very dense and could possibly allow for rewritable media as well as WORM. Potentially, instead of flat media, square or oblong manufactured "crystals" could be used.

Holographic storage does not store data on the surface of the media, but uses the total volume of the media (crystal) provided. A laser is used to write a tiny hologram inside the crystal, where each "block" or "page" can hold 100KB – 1MB and these are read by adjusting the angle of the laser to "focus" on a particular page. As the media is not moving, only a lens set up in front of the laser moves, read performance could be very high.

Research is continuing in universities around the world and so commercially available holographic storage is still a possibility.

Why Glass Storage?

In 2019, Microsoft presented their "Project Silica" concept for archiving data on glass. In the same way that holographic storage uses a multi-layer process to pack large amounts of content in a small space, glass storage also uses a layered approach with a laser encoding data and polarized light reading solution onto 75 x 75 x 2 millimeter media.

They described the principle of "voxels", a 3D version of a pixel where more than 100 layers of voxels can be laid down and read from a glass media. Naturally, this is a WORM (write once read many) storage device only, data cannot be overwritten or deleted and its intentions are for very long-term archive purposes. Their first real-world example was to preserve the 1978 Superman Movie.

Glass is a very stable media and is not affected by humidity, reasonable temperature changes or magnetic fields. It can also be cleaned easily without impacting data stored. This leads to a permanent "store it once" approach where data migration, required for all current archive storage technologies, is eliminated. Data life has been suggested to be over 500 years (one report suggests 10,000 years!!), significantly greater than any current storage technology.

The 3 inch by 3 inch square of glass can store up to 100GB of data, but performance is currently limited to hundreds of kilobytes per second – way below any other form of archive storage today.

Why DNA-based Storage?

DNA storage seems to be the technology most likely to be used for long-term archive, based on the number of organizations studying it, and that SNIA has supported the working group (DNA Data Storage Alliance) to create interoperable solutions. Illumina Inc, Microsoft, Twist BioScience Corp, and Western Digital formed the group and have been joined by a further 34 organizations to date.

DNA storage has the ability to store the same amount of data a million times denser than SSD, or 200PB per gram of DNA. It is a permanent form of storage with an estimated "shelf-life" in the hundreds or possibly thousands of years, depending on storage conditions. Again, this is limited to WORM (write once read many) storage only as it would be impossible to change the data without rewriting everything again with the new information substituted for the old.

DNA consists of four "base-pairs" that combine two by two (adenine with thymine and guanine with cytosine) creating a double-helix of information. These pairs can be substituted into the digital storage world of 0s and 1s, a concept that underlies all storage products of today. Researchers have already encoded content into a DNA sequence and read that data with 99.99% accuracy.

Speed of write, and subsequently reads, will certainly be the biggest issue for this technology along with the size of any device and (as the Alliance has understood from the beginning) the interoperability. There is little point developing competing solutions without the ability for one system to read content laid down by another, if long-term data preservation is the destination. Cost today is also very high, but the roadmap anticipates significant cost reductions as the technology moves into its commercial phase.

DNA data can also be easily replicated – a critical requirement – as the helix is destroyed when reading. One possible method being investigated is to imbed the DNA sting into a bacterium – which can self-replicate.

How QStar can Help

QStar has been supplying storage gateway solutions for ALL archive technologies for 35 years. Our current solutions support tape, optical, object storage, cloud and disk, presenting one or more of these technologies through SMB, NFS and/ or S3 interfaces.

QStar solutions are designed with the concept of data outliving the storage it is originally placed on. Migrating data from old to new is in-built to our solutions and can be handled as a background task. QStar will support any of these technologies that come to market as a commercial offering and will continue to provide data migration from current technology to new technology in the archive space.

For more information see www.qstar.com or email sales@qstar.com.

Summary

All four technologies covered above have strong potential for archive with many of them only supporting a WORM format. DNA storage certainly seems to be the technology where most money and research time is being directed. It also looks like it will be at least 10 years before any commercial solution will be available. Holographic storage looks the least likely to become available. It appears Microsoft is the only organization still seriously investigating this technology after previous efforts failed to make it to the commercial stage. Likewise, Glass storage appears chiefly to be a Microsoft research project. Glass is definitely a more stable, longer-term medium than disk or tape and could be another option for long-term data storage. Having a proven demonstration with Warner Brothers gives the technology a boost but without other organizations supporting it, this looks like a research project only. Optical storage is perhaps the easiest solution to bring to market, with a proven pedigree to look back on. It can only succeed if the costs predicted can be achieved, and the only way that can happen is mass acceptance. Scalability is the biggest concern as media capacities are still small.

For Holographic, Glass and DNA storage, Microsoft is involved in all three. The reason is undoubtedly for long term preservation of data in the Cloud. As Cloud storage is relatively new (less than 20 years), the requirement for long term preservation of data is only now becoming pressing. The same issues will exist at Google and Amazon. The rapid growth of the "deep archive" class with its lowest costs of storage creates a long-term problem for Cloud providers.

Today most of this data is on tape, and although tape has good data preservation characteristics, it lacks any significant backward compatibility. LTO7 data cannot be read on current LTO9 drives. The quandary is to either keep old drives available to read that data, or to write the same content again to new tape media (either through backup processes or media migration). Both options come with cost and data consistency issues.

Far preferrable would be to store content once to a medium that lasts for hundreds of years, and has guaranteed compatibility for reading content on future devices. Many analysts are predicting a huge data gap, where humanity will create data at a faster rate than we can manufacture data storage devices. DNA storage could be the long-term answer to this gap in production.

Optical Storage for your Archive Tomorrow	
Pros	**Cons**
Proven technology	Capacity and performance have limited enterprise optical
Random access to data on same media	Probably WORM only
Commercial availability, low-cost of entry	Requires many drives to make up for low capacity and performance
Existing Operating System and application support	

Holographic Storage for your Archive Tomorrow	
Pros	**Cons**
Longer term storage than optical	Technology tried and failed already
Possible commercial availability	Low Capacity and performance
	Lack of developers investigating this technology

Glass Storage for your Archive Tomorrow	
Pros	**Cons**
Very long-term archive solution	Low capacity and performance
	WORM only
	Lack of developers
	Possibly no commercially available product – Cloud providers only

DNA Storage for your Archive Tomorrow	
Pros	**Cons**
Very dense storage solution	Performance issues for writing and reading
Long term archive when stored correctly	Destroys content by reading
Possibly simple to replicate	WORM only
Significant number of organizations researching	No commercially available product due to cost – Cloud providers only