IEEE
Advancing Technology for Humanity

INTERNATIONAL ROADMAP FOR DEVICES AND SYSTEMS™

INTERNATIONAL ROADMAP FOR DEVICES AND SYSTEMS

2023 UPDATE

MASS DATA STORAGE

THE IRDS IS DEVISED AND INTENDED FOR TECHNOLOGY ASSESSMENT ONLY AND IS WITHOUT REGARD TO ANY COMMERCIAL CONSIDERATIONS PERTAINING TO INDIVIDUAL PRODUCTS OR EQUIPMENT.

Table of Contents

Table of Figures

Table of Tables

Contributors

Tom Coughlin, Chair
Roger F. Hoyt, co-chair

Name	Affiliation
Ed Childers	IBM
Tom Coughlin	Consultant
Ron Dennison	Consultant
Simeon Furrer	IBM
Roger Hoyt	Consultant
John Hoffman	Ernst&Young/DNA Data Storage Alliance
Dave Landsman	Western Digital/DNA Data Storage Alliance
Mark Lantz	IBM
Kevin Lu	Folio Photonics
Niranjan Natekar	Western Digital
Ken Singer	Folio Photonics
Doug Wong	Kioxia

Summary

SOLID STATE

Over the past 30 years (1993 to 2023), the NAND flash market has grown from zero to become a nearly $60B dollar market not only by NAND displacing existing storage media, but also by NAND-based products enabling new markets. The initial market for NAND flash was audio tape replacement in digital telephone answering machines, but the market that jump started NAND flash adoption was its use in digital cameras. A proliferation of small form flash memory cards followed the advent of digital photography: PCMCIA (PC Cards), CompactFlash, Smart Media, Multi Media card, and Secure Digital (SD) cards. The demand for flash grew along with the digital camera market. As the cost of NAND flash fell, the market grew as floppy disks and writeable CD's began to be replaced by USB flash drives. The transition from audio tape and CD players to digital audio MP3 players was also enabled by the falling cost of NAND flash. At what point alternative memory technologies like MRAM, FeRAM, ReRAM, or others, might displace DRAM or NAND is unknown. Today, all of these alternative technologies are still more costly per bit than DRAM or NAND flash, and this prevents them from being selected as replacements except in those rare circumstances where the cost is less important than certain important attributes they provide.

HDD

Of current mass data storage technologies, in terms of storage capacity shipped, hard disk drives (HDD) are by far the largest single component of the mass data storage industry. Today the HDD market continues its decline in unit volume primarily due to displacement by solid state drives. However, the demand for data center nearline storage continues to grow and technology advances such as helium filling, more heads and disks, dual actuators and heat assisted magnetic recording (HAMR) have fueled continuing capacity increases. Seagate is now shipping 32 TB drives and anticipates 50 TB drives in 2026. Future capacity growth will depend on the further development of HAMR as well as new technologies such as next generation TDMR and heated dot magnetic recording.

TAPE

The continued exponential growth in the creation of digital data combined with the recent slowdown in areal density and capacity scaling of hard disk drives is driving demand for cost effective data storage solutions. Magnetic tape technology is particularly well suited to help meet this demand due to its very low total cost of ownership and its efficient data center footprint that is approaching 5PB/ft^2. Part of the TCO advantage of tape arises from its very low power consumption, which also contributes to its small CO2 footprint. The natural physical airgap of tape solutions also provides an additional level of protection against accidental data deletion and cyber security threats. Moreover, recent tape areal density demonstrations provide confidence that tape has the potential to continue scaling areal density for multiple future generations with cartridge capacities expected to reach hundreds of TB per cartridge within the next decade. All these benefits have resulted in an increased adoption of tape, particularly among hyperscale cloud companies, and are driving growth of the tape market.

OPTICAL

Optical storage media is undergoing a shift from its traditional role in consumer media distribution to a focus on enterprise and institutional archival storage. To enhance capacity while minimizing costs, emerging optical technologies are exploring storage solutions in the third dimension and beyond. Robust library systems tailored for optical media are emerging to meet the stringent requirements of enterprise-level storage demands.

The low maintenance and operating energy costs of optical media, coupled with infrequent remastering needs, position it as a naturally advantageous solution in the sustainability-conscious landscape of data archiving and preservation. Optical Write-Once, Read-Many (WORM) technologies, with their air-gap feature, offer cybersecurity advantages. In terms of energy consumption, optical technologies exhibit the lowest levels both intrinsically and in the context of data center environmental control, promising significant reductions in greenhouse gas emissions.

Despite these benefits, critical challenges persist in the optical data storage domain, including the need to lower initial costs, expand capacity, increase speed, and improve error management. Exploring possibilities such as multidimensional media, remote-write libraries, femtosecond lasers, and high-speed display and imaging technologies opens new opportunities for enterprise optical data storage.

DNA STORAGE

The world is attempting to digitize unprecedented amounts of information. This information can be valuable if mined, stitched together, or otherwise searched and analyzed; however, the cost of saving the massive amount of data associated with this information is beginning to overwhelm the ability to pay for it using conventional storage technologies. This trend is leading system designers to look for new storage technologies which can sustain the densities, access flexibility, and TCO needed for this wave of digitization. One of the candidate technologies being considered is synthetic DNA. DNA is a potentially compelling storage medium due to its ~1bit/nm^3 bit size, the fact that it is incredibly stable at room temperature if kept dry, and that it can be read in the future even if the original writing/reading technology no longer exists. This combination of factors enables the potential of the proverbial "datacenter in a shoebox" as compared to incumbent storage technologies: small footprint, low power, no fixity checks or technology refresh. In other words, compelling TCO.

While DNA data storage is not yet ready for productization today, with the huge advances in medical/scientific DNA technology and applications over the past several decades, plus academic research and commercial biotech both targeted at DNA data storage over the past decade, the basic foundations for synthetic DNA as a data storage medium have been demonstrated[17]. It is thus reasonable to expect that a path to a synthetic DNA data storage ecosystem will come into focus over the next 5-10 years.

Solid State Storage

NAND Flash Storage

Situation Analysis

Over the past 30 years (1993 to 2023), the NAND flash market has grown from zero to become a nearly $60B dollar market not only by NAND displacing existing storage media, but also by NAND-based products enabling new markets. See also the IRDS NVM Technology Roadmap at this site: https://irds.ieee.org/images/files/pdf/2022/2022IRDS_MM_Tables.xlsx.

The initial market for NAND flash was audio tape replacement in digital telephone answering machines, but the market that jump started NAND flash adoption was its use in digital cameras. A proliferation of small form flash memory cards followed the advent of digital photography: PCMCIA (PC Cards), CompactFlash, SmartMedia, MultiMedia card, and Secure Digital (SD) cards. The demand for flash grew along with the digital camera market. As the cost of NAND flash fell, the market grew as floppy disks and writeable CDs began to be replaced by USB flash drives. The transition from audio tape and CD players to digital audio MP3 players was also enabled by the falling cost of NAND flash. Figure 1 shows NAND flash shipment revenue from 2013 through 2022 and estimated for 2023.

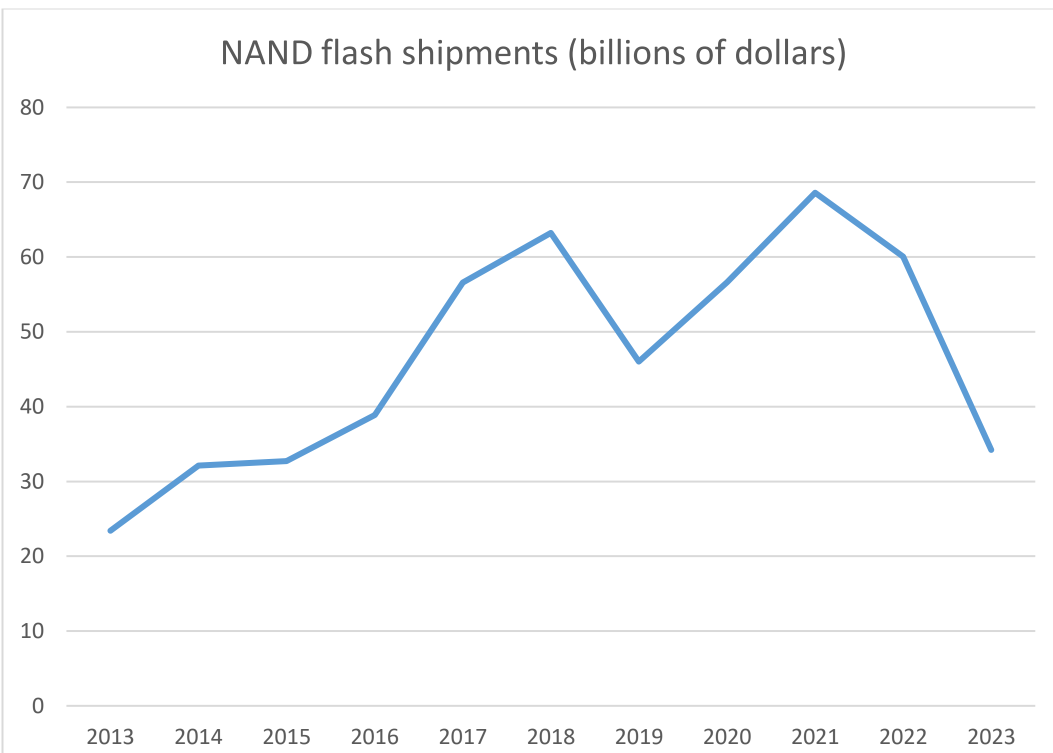


Figure 1. NAND Flash Shipments

Source: Forward Insights

Other products enabled by the high density and low bit cost of NAND include portable GPS devices and Personal Digital Assistants (PDAs), but the next big application was the creation of the smart phone, and that application continues to drive a large segment of the NAND flash market. It is only within the last decade that NAND flash has become inexpensive enough to start displacing traditional rotating media (i.e., hard disk drives) and today (2023), solid state drives (SSD) are the largest market for NAND flash memory. **Figure 2** shows NAND flash byte shipments from 2013 through 2022 and estimated for 2023.

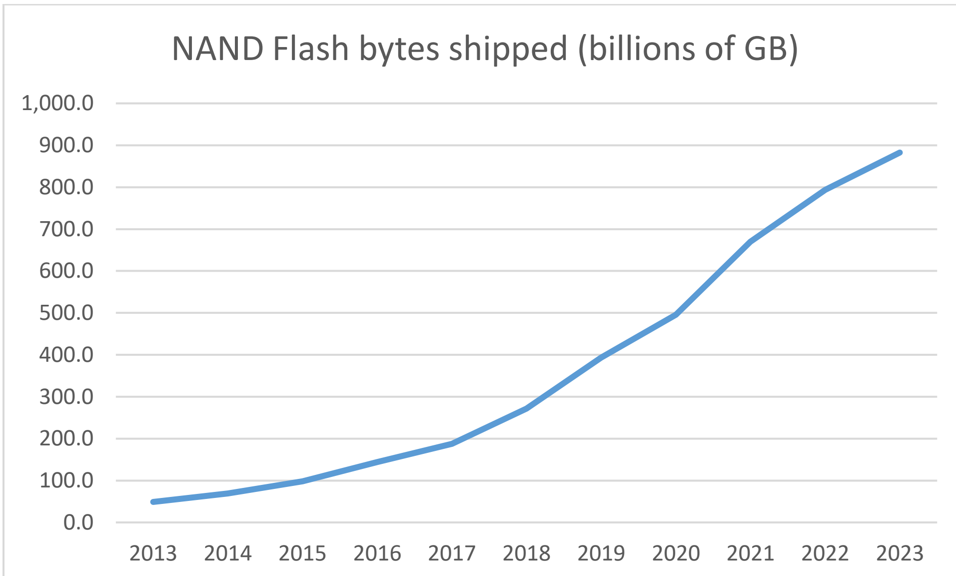


Figure 2. NAND Flash Shipments

Source: Gartner

As flash-based SSD costs have fallen, HDDs have been displaced: first in consumer PCs, and increasingly in data centers, as SSDs have come to occupy the tier of frequently, and randomly, accessed data. SSDs and HDDs continue to coexist because HDDs (and tape) will continue to offer the lowest cost per bit for the foreseeable future, but the designers of storage systems and servers recognize the benefit of SSDs for improving data access time and reducing power consumption, and the development of SSD form factors specifically for this market indicates this.

Figure 3 shows a view of the memory and storage hierarchy comparing cost and performance on a log-log scale. DRAM and SRAM (L1, L2, L3) are volatile memories while NAND, HDD, and tape are non-volatile. The key point is that memory performance is correlated with cost per bit.

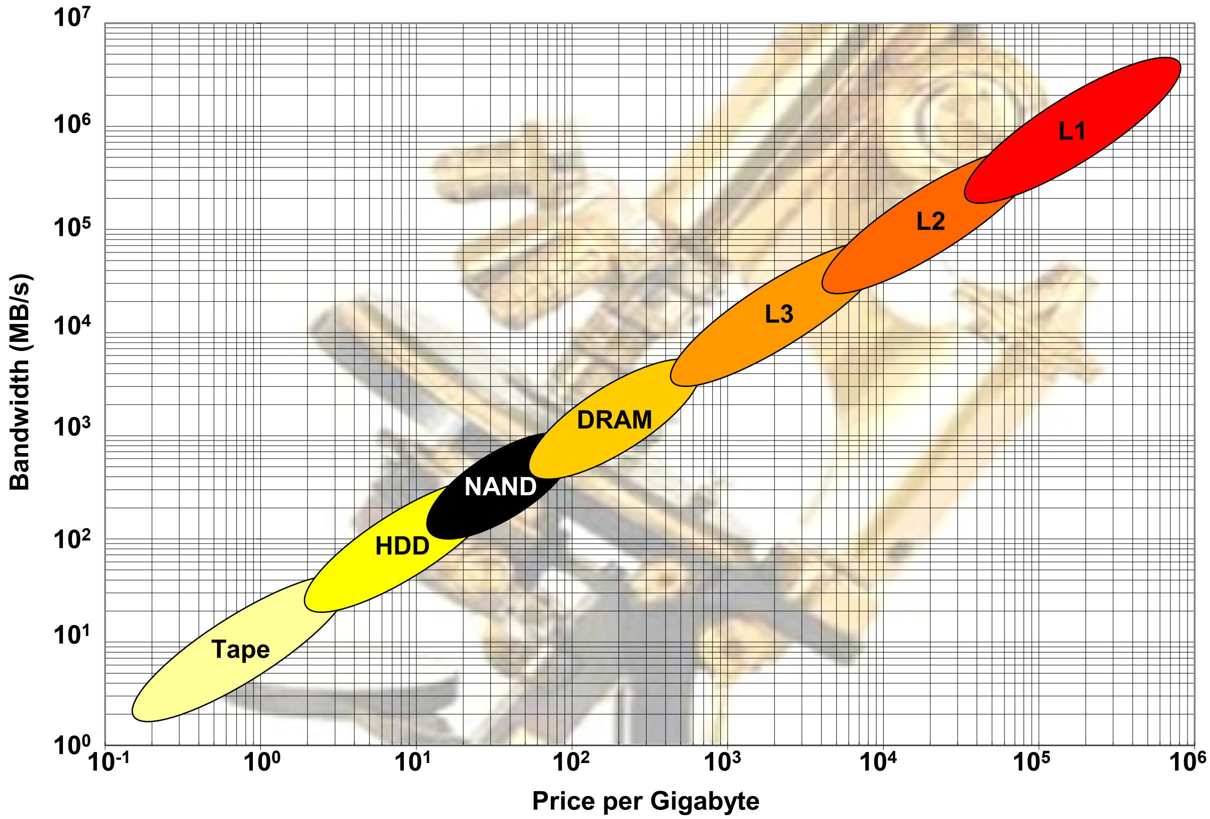


Figure 3. The Memory/Storage Hierarchy in Cost vs. Performance

Source: Objective Analysis

While NAND flash is sometimes directly connected to an SoC (system on chip) or microcontroller, it is most often used with a controller chip specifically designed to support a specific interface. An SSD controller might support PATA, SATA, SAS or NVMe. A controller might be packaged in the same package as the NAND flash itself, as is the case for eMMC (embedded MultiMedia Card) or UFS (Universal Flash Storage).

Due to cost and capacity, SSDs exist as a faster, but more expensive, storage tier. HDDs are used for bulk mass storage and SSDs are used for speed. Frequently requested and typically randomly accessed data resides in the SSDs, while less-frequently used and typically sequentially accessed data is kept in high-capacity HDDs.

When measured in cost per gigabyte (GB), an SSD is more expensive than an HDD. The cost/GB gap between SSD and HDD has been shrinking over the past decade but is unlikely to cross due to the fact that although NAND bit density per die continues to increase, so does the areal bit density of HDDs. **Figure 4** shows the average selling price of a GB of NAND flash memory from 2013 through 2022 and estimated for 2023. **Figure 5** shows the average selling price of a GB of memory in an SSD from 2013 through 2022 and estimated for 2023. SSD prices are somewhat higher than the price of raw NAND flash memory. Because of this, SSDs first replaced those HDDs that were being used in a way that increases I/O speed at the expense of capacity.

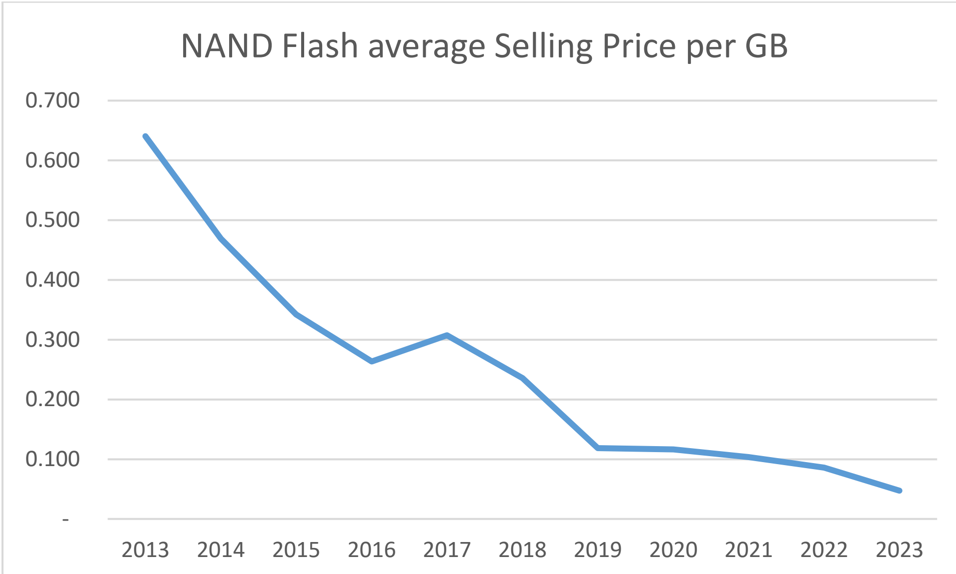


Figure 4. NAND Flash ASP/GB

Source: Gartner

In the past, storage systems used a few HDDs in a faster storage tier that used "short stroking" or "destroking" to provide faster data access. In a short-stroked HDD, the system used only the tracks on the outside edge of the platters, limiting actuator travel distance and accessing the physical media where data passes most rapidly under the heads, thus increasing overall IO performance. Today's redesigned storage systems have abandoned short-stroked HDDs, replacing a number of short-stroked HDDs with a single SSD. This increased performance while lowering costs.

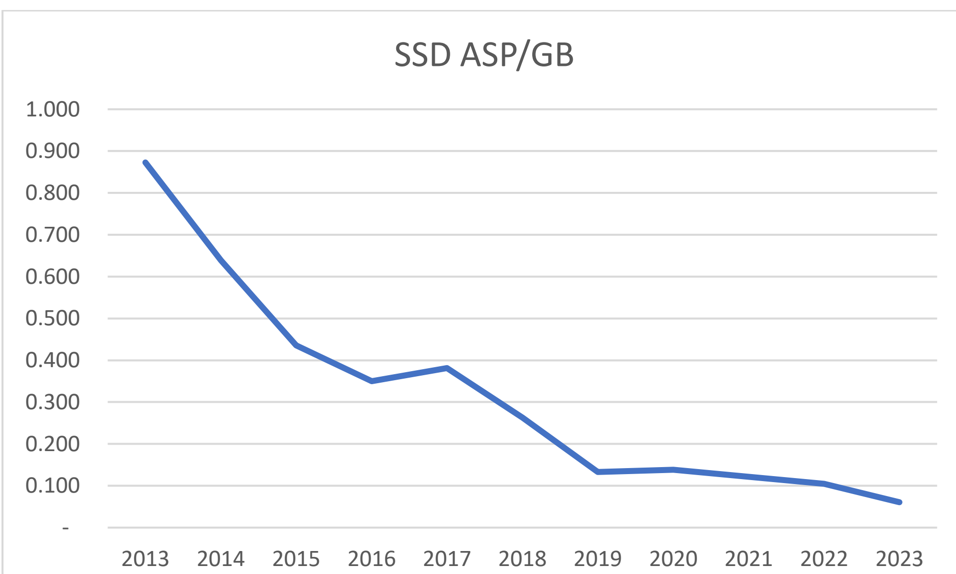


Figure 5. SSD ASP per Gigabyte History

Source: Gartner

Five years ago, SSD penetration percentages in consumer or client laptops was still relatively low. The reason was higher cost and no appreciable performance improvement using the SATA protocol. Today, the SSD penetration rate exceeds 95% due to the falling cost of SSDs as well as the performance improvement gains and reduced boot times of using NVMe (Non-Volatile Memory Express). NVMe was designed to take advantage of the shorter latency associated with solid state non-volatile memory (not just NAND flash, but also any future non-volatile memory technology).

BUSINESS/TECHNICAL ISSUES

There are many parts of the flash memory value chain, some of which are more profitable than others. Flash card makers and many SSD manufacturers compete by perfecting manufacturing efficiency and inventories, and through good responsiveness to the market, but this market is undifferentiated so margins are low. Other SSD makers use proprietary controller chips, or other differentiators to justify a higher price for their product. In both flash cards and SSDs the flash memory tends to account for 80% or more of the bill of materials, so changing the controller or other non-flash portions of the system may result in a high payback. Makers of higher-value devices including smart phones, and tablet PCs can reap larger profits (from storage) thanks to the greater differentiation of these products.

In undifferentiated markets the bulk of the profits and the greatest challenge lay in the production of the NAND flash chips used to make those products rather than in other technologies and manufacturing capabilities.

Intellectual property is an important part of the flash memory business. Flash manufacturers invest heavily in their intellectual property portfolios and patent protection has been very useful in assuring that this research effort is properly rewarded. Prior to its acquisition by Western Digital, SanDisk signed royalty agreements with nearly all flash chip, card, and controller manufacturers to generate a royalty stream that offset a significant portion of the company's R&D cost.

Over its lifetime NAND's price per gigabyte has declined rapidly, allowing the technology to displace entrenched storage solutions including photographic film, floppy disks, magnetic tape, and rotating optical media. Today SSDs and video applications consume the largest per-device quantity of NAND flash storage capacity.

NAND flash chips are manufactured by Samsung and SK Hynix in Korea, WD and Kioxia in a joint venture in Japan, Micron in the US and Singapore, YMTC in China, and Macronix and Winbond in Taiwan.

MANUFACTURING EQUIPMENT

NAND flash chips are manufactured using standard semiconductor processing equipment. Cost is a key focus, so NAND manufacturers rapidly migrate process technologies and moved from 200mm wafers to 300mm wafers starting in 2007. In fact, the process migration of NAND flash has become so important that flash became the process driver for those companies who manufacture both flash and another technology (i.e. DRAM). A "process driver" is the technology used to perfect the manufacturing process of a company's most advanced and lowest-cost technology. However, the move from planar 2D NAND to vertical 3D NAND has changed the need to create the smallest sized features and instead, the availability of advanced deposition and etch technologies are driving the development of the latest 3D NAND designs.

Flash is manufactured using a production process similar to that of standard CMOS logic, a process used to manufacture most semiconductors. Historically, the main difference between flash and CMOS logic was that flash added a thin gate oxide layer and a polysilicon floating gate to the process to enable the construction of a floating gate. This floating gate, the key to any electrically programmable nonvolatile memory, can be programmed either by hot electron injection or by tunneling, both of which force electrons through the thin gate oxide insulator. Even though flash adds only a small number of additional layers to a standard CMOS process, the development of a flash process is quite complex, and provides a barrier to entry into this market. **Figure 6** shows cross-sections of the floating-gate memory cell in reading, programming and erasing. This cell is programmed using tunneling, rather than hot electron injection.

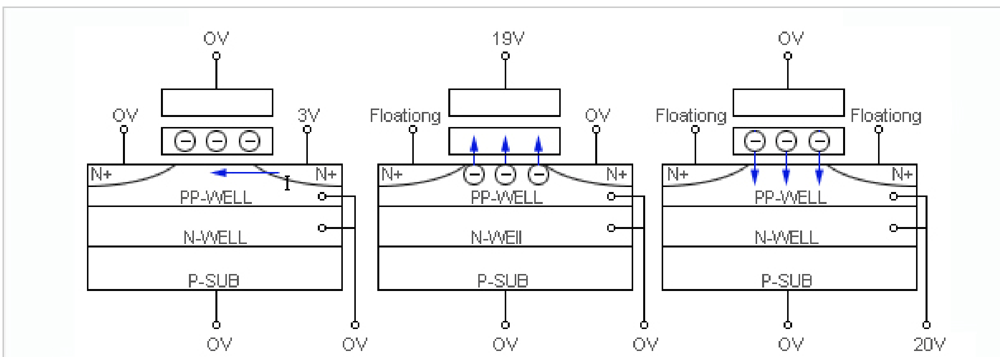


Figure 6. Cross-Section of planar floating-gate flash memory cell

Left: Reading, Middle: Programming, Right: Erasing
(Source: Samsung Semiconductor Company)

While floating gate NAND flash memory continues to be manufactured, more than 90% of the NAND flash memory bits produced today are 3D NAND flash memory. The industry has transitioned from the 2D planar process, in which the memory array lies on the surface of the silicon wafer, to a vertical 3D process in which memory cells are formed in stacked layers. No more development is being done using the 2D planar process.

The reason for the transition is to increase the number of bits per die, which is necessary in order to reduce bit cost. Planar 2D flash has undergone continuous cost decreases by following Moore's Law by shrinking the size of the transistors every 18 months or so. But this constant shrinking eventually brought flash to the limits of scaling. As the memory cell transistor got smaller, the number of electrons it stored also got smaller, which reduced write/erase cycle endurance and data retention time. When the lithography node hit the mid-teens (15nm), it was no longer feasible to shrink further. At this point, the die density was approximately 128 Gbits for an MLC (2 bit per cell) device.

At this point, the NAND flash market needed to transition to a new technology. 3D NAND flash. 3D (three-dimensional) NAND (**Figure 7**) stacks bit cells rather than shrinking them. The number of bits per chip increases in proportion to the number of layers. Although this process does scale up in total chip density each generation, it is less cost effective than the density increase achieved by a lithography shrink in the past. This 2007 Toshiba invention describes a vertical NAND structure using conventional semiconductor materials but a relatively complex process using several steps that have never previously been used in chipmaking.

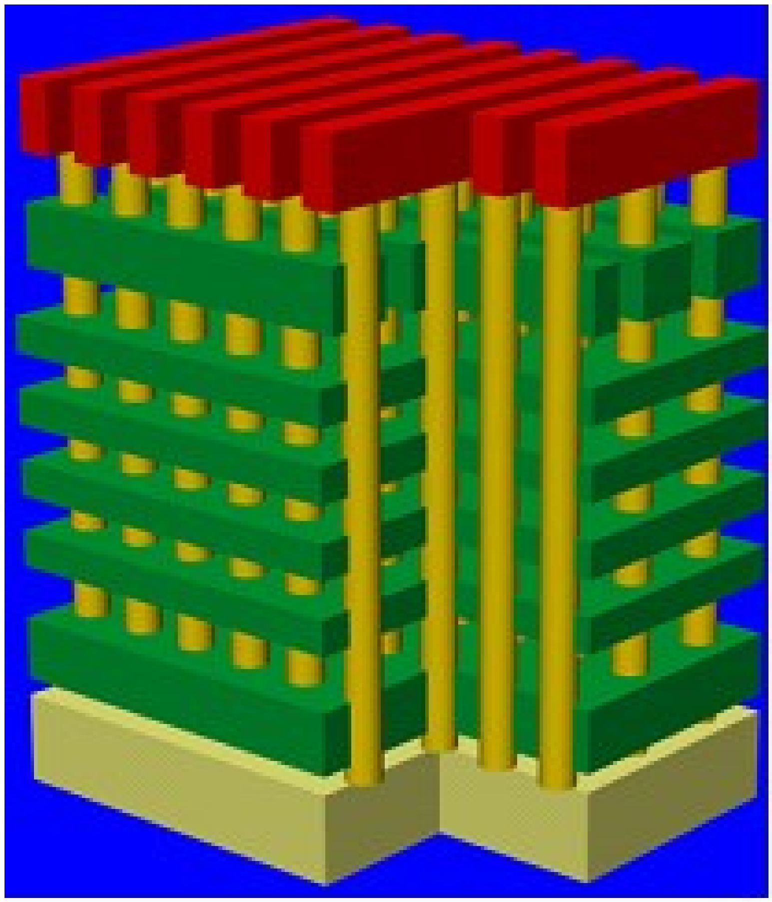


Figure 7. Toshiba's BiCS Vertical NAND Structure (Courtesy of Toshiba Corp (Kioxia))

Each manufacturer now uses a different variant of this process. Samsung has been shipping 3D NAND-based SSDs since July 2014. Other companies delayed their launch of the technology until they saw a clear path to profitability, but as of 2023, NAND vendors have now been in mass production of 3D NAND for several generations and 3D NAND comprises >90% of the total NAND flash market in bits.

The conversion from yesterday's planar to today's 3D process was a significant transition since the manufacturing equipment mix is different between the two. Planar NAND uses a mix that is higher in the need for lithography tools, while 3D NAND production uses more deposition and etch tools. Historically, flash manufacturers have been able to migrate from one process to the next through the addition of incremental numbers of new tools that are absolutely necessary to the process migration, using already-installed tools for all of the other steps. The equipment mix needed for 3D NAND was sufficiently different that migration could only be performed through the addition of significantly more tooling than in previous process migrations.

In the last iNEMI report, it was speculated that 3D NAND would eventually take over the market from the existing 15nm planar NAND flash. This has now occurred and 3D NAND bits represents most of the NAND market.

Now that 3D NAND is the mainstream process, the challenge becomes how to continue the path toward more bits per die and lower bit cost. The accepted way to continue cost reductions with 3D NAND is to add layers to the chip. Manufacturers started the first generation of 3D NAND at 24-36 layers. For a number of years flash makers thought that the number of layers would be limited to about 100 layers due to the difficulties of etching a deep enough and narrow enough hole to build the columns in Figure 7. In 2015 SK Hynix revealed a solution to this problem called “String Stacking” which builds a chip of perhaps 100 layers, then builds another 100-layer chip on top of that, and so forth. There is no clear limit on how high such a stack can be practically made, but it is clear that the 3D NAND architecture has many years of life ahead of it.

There will be a point in the future in which current 3D NAND technology can no longer scale cost efficiently to higher density chips. While areal density per die increases with increasing layer count, there is also an incremental additional cost per layer. In the era of 2D floating gate NAND, lithography shrinks directly enabled smaller memory arrays and resulted in lower bit cost. However, bit cost reduction is more challenging in the 3D era because feature sizes are not shrinking at the same rate. At some point, there is the expectation that alternative technologies will eventually replace flash, but it is likely for any future memory to still be composed of a 3D physical interconnection of transistors and have layer count limitations for the manufacturing process.

Manufacturing Processes

The NAND flash market is very cost sensitive and capital intensive. NAND manufacturers can effectively compete only by using a manufacturing process on a par with that of their competition. Process advancements can only be achieved through the use of the most expensive tooling, and this limits the market to a handful of key players who have access to billions of dollars of capital. As a rule of thumb, a manufacturer will need to upgrade its wafer fabrication plants from one process node to the next every 2 years.

Process line widths are measured in nanometers (nm), or 10^{-9} meters. NAND manufacturers have migrated the vast majority of their production from a planar mid-teens nm process to the 3D NAND process. While first generation 3D NAND products started at 24-36 layers, today, the highest layer count 3D NAND currently shipping is 232 layers.

Materials

The materials used to manufacture NAND flash memories are common to nearly all other semiconductor processes. Semiconductor manufacturing requires extraordinarily pure materials, since the key to making a semiconductor is to manipulate small impurities within extremely pure silicon crystals. Typical materials include raw silicon wafers, photographic emulsion and developers, high-purity gases, including hydrogen, oxygen, nitrogen, and silane, aluminum and copper sputtering targets, boron, arsenic, tantalum, and other dopants, de-ionized water, and an abundant dependable supply of electricity. Some processes use wet steps that depend upon a supply of acids and other reactive liquids.

From a materials standpoint the 3D NAND process is driving relatively minor changes since most manufacturers need to convert from a conductive polysilicon floating gate to a charge trap, most commonly made using silicon nitride (which is already used as a passivation layer over the top of all silicon semiconductors), and in some cases this will be joined by the replacement of a polysilicon top gate with tantalum and the gate dielectric with alumina. Silicon nitride, alumina, and tantalum are regularly used in semiconductor processing and are well understood and abundantly available.

Memory makers anticipate that one or more of the new "emerging" nonvolatile technologies that are currently being researched will become attractive at some point in the future. These technologies include magnetic materials, ferroelectrics, chalcogenide glasses, and other materials that will be added to the silicon process. Some of these technologies are already in production. Although these new technologies may not constitute an important part of today's market, they could come into widespread use within the timeframe of this roadmap's outlook.

One issue with these new technologies is that they introduce a new material into the process, and new materials always create new problems in the fabrication plant since they are not as well understood as those technologies that have been in volume production for a long time. For this reason there will be false starts as these newer technologies vie to replace established silicon-only semiconductor processes.

Quality/Reliability

NAND flash as a raw storage medium has many idiosyncratic behaviors and failure modes, so a controller is always used with NAND-based devices to perform error correction and manage wear-out. The quality and reliability of storage devices that use NAND flash (i.e., SSDs, flash cards, USB flash drives) is measured through three main indices:

- Data integrity as provided by the controller,
- Memory cell wear or endurance
- Lifetime of the data in the device (data retention)

Each of these will be addressed in order:

Data integrity: The cost per gigabyte of NAND flash is lower than that of any other semiconductor memory. NAND's low cost is achieved by trading off price against data integrity: the lower the data integrity, the lower the price, so NAND has been designed in a way so that bit errors are anticipated and allowed to occur. Advanced forward error correction is required to use NAND flash. Bit errors are corrected external to the flash chip in a controller chip that uses the same sort of error correction code (ECC) approaches used to recover bit errors in hard disk drives. This is a very well understood technology. In today's 3D NAND, LDPC is now the most commonly used ECC with a correction strength of approximately 120-160 bits/1kByte.

All NAND chips include an area for the storage of parity bits (ECC) that coexist with the data array in order to enable error correction. Controller designers differentiate their products by using these parity and coding bits in different ways. Although some companies pride themselves on having algorithms that provide higher data integrity than their competition, all modern algorithms provide data that users accept as error-free. Users are comfortable that SSDs, flash cards, and USB flash drives will accurately store either code or data, and no single supplier has had to overcome a reputation of supplying media that does not accurately replicate a file.

Another concern with NAND flash is the wear mechanism. All EEPROM (electrically erasable programmable read only memory) technologies (i.e. NOR flash and NAND flash) exhibit wearout mechanisms that put a limit on the number of write/erase cycles sustainable. This characteristic, called "endurance", varies with the number of bits stored in a memory cell. NAND flash chips are categorized in terms of how many bits can be stored in a cell:

- Single Level Cell (SLC), which stores one bit per transistor
- Multi-Level Cell (MLC), which stores 2 bits per transistor
- Triple Level Cell (TLC), which stores 3 bits per transistor
- Quadruple Level Cell (QLC), which stores 4 bits per transistor
- Penta Level Cell (PLC), which stores 5 bits per transistor

TLC (as well as QLC and PLC) terms a little inaccurate because there are actually more levels than implied, but the names have stuck. In SLC, there is a single threshold voltage that differentiates between a 0 and a 1 state. In MLC, there are 3 threshold voltages to distinguish between the 4 states necessary to store 2 bits of information. But for TLC, there are more than 3 "levels" needed to differentiate the 8 charge states necessary to store 3 bits.

The more bits per cell that a NAND chip can store, the fewer write/erase cycles it will be able to sustain. A typical SLC NAND flash chip made in a 24 nm planar floating gate process may be able to withstand 50k-100k write/erase cycles while a 15nm MLC chip is rated to only 3k cycles. The transition to 3D has allowed the individual cells to get larger again, so a TLC 3D NAND cell can also achieve 3k cycle endurance. QLC and PLC cells of the same size achieve fewer cycles.

The controller that manages the flash in an SSD, card, or USB flash drives detects endurance failures and either corrects the failed bits or maps the blocks that contain excess failed bits out of the device, disallowing their subsequent use. This is possible as long as there are spare blocks available. Flash controllers attempt to spread the wear out across all blocks (wear leveling), so that all blocks have a similar amount of wear. Sophisticated SSD controllers also perform write coalescing and write gathering to reduce the number of times a block is written to. Regardless of how well all of these techniques work, however, at some point, the flash in a device will be worn out.

Data retention: The charge in a NAND flash determines the bit values. This charge can be expected to leak over time and temperature, creating bit errors that eventually become numerous enough that uncorrectable read errors occur.

NAND flash chips are typically specified to retain data for 10 years when new and typically 1 year at the end of life (i.e. the maximum rated write/erase cycles), when powered down. This compares against 5 years for magnetic media, and up to 100 years for recordable optical media. JEDEC (Joint Electron Device Engineering Council) published a number of specifications enabling the standard measurement and characterization of endurance and data retention in NAND-based chips and devices. As a chip experiences more cumulative write/erase cycles, the data retention time will decrease. Such failures are the result of lattice disruptions in the insulating oxide which create leakage paths. As the device experiences more cumulative write/erase cycles, it develops more lattice disruptions, increasing this leakage. This is the end of life wearout mechanism for all types of flash and results in insufficient data retention time.

Unlike magnetic media, NAND is unaffected by extraneous magnetic fields. NAND flash is less sensitive to heat than optical and magnetic media, and is often specified to operate at up to 85ºC and to withstand storage temperatures of up to 150ºC.

Environmental Technology

Every semiconductor wafer fabrication plant generates hazardous waste, and the handling of that waste is counted among the many criteria that impact the design and operation of a plant. Gas emissions are typically cleaned by scrubbers on a facility's roof. Wastewater is often purified and recycled within the facility, to avoid concerns of contaminating groundwater, streams, or other bodies of water.

All semiconductor manufacturers perform on-site decontamination of factory effluents and have regular programs of hazardous waste removal. The contaminants removed from effluents are sealed in drums and conveyed by bonded and insured carriers to hazardous waste dumps.

Semiconductor production generates a low volume of such hazardous wastes, and despite the high costs of removing it, the cost factor of such handling does not place a material impact on overall costs. Should a fab in one country not be held to the same environmental standards as those that apply in a competing country, it would not give that fab a meaningful cost advantage over its competition.

Test, Inspection and Measurement (TIM)

One important challenge for flash manufacturers is the implementation of sufficient test procedures. Historically, NAND chip densities (capacities) have doubled each generation. and the time required for testing increase in proportion to the density of the chip. To prevent this from impacting manufacturing costs, NAND flash manufacturers have started to apply statistical methods to avoid having to test every bit in the device.

Internal self-test mechanisms are also used, and will increase in sophistication over time. These internal test mechanisms will allow less test equipment to be used to test more chips. Although this approach will not help reduce the WIP costs of testing, it will cut the capital expenditures required in this area.

Roadmap of Quantified Key Attribute Needs

The data in Table 1 is based on the IEEE IRDS Semiconductor Roadmaps (International Roadmap for Devices and Systems) to estimate where NAND flash is headed over the next several years. This roadmap can be found at https://irds.ieee.org/

Table 1 shows a prediction of how NAND flash processes develop over time and the transition as the industry has moved from a 2D planar process to a 3D NAND structure.

Table 1. NAND Flash Chip Roadmap

	2013	2015	2017	2019	2021	2023	2025	2027	2029
Density	64Gb (MLC)	128Gb (MLC)	256Gb (TLC)	256Gb (TLC)	512Gb (TLC)	1Tb (TLC)	1Tb/2Tb (TLC+)	2Tb/4Tb (TLC+)	4Tb/8Tb (TLC+)
Planar Process	19nm	15nm	N/A	N/A	N/A	N/A	N/A	N/A	N/A
3D Layers			32-48	64-96	112-176	Low 200s	High 200s	300+	500+

While the development of new planar 2D NAND processes has ended, the 3D era of NAND flash is still maturing. Just a few years ago, it was thought that 100 layers might be a practical limit, but in 2023, 3D NAND with layer counts exceeding 200 layers is now in production. TLC 3D NAND is mainstream today, but the percentage of QLC continues to grow.

Critical Issues

Limitations of the Flash Process

Flash makers have for many years understood that standard planar flash designs would not be able to shrink below some certain process geometry. NAND flash manufacturers generally expected that the last process generation at which standard NAND could be manufactured was around 15nm and this prediction turned out to be accurate. The main problems were cross-coupling between adjacent cells, and decreasing data retention time due to the decreasing number of electrons that could be stored in each memory cell transistor.

For the 3D stacked cell, the physical memory cell transistor is significantly larger than in the 2D process. The memory transistor is now a cylindrically shaped cell that can store more charge than the 2D cell and made possible the transition from 2 bits per cell to 3 bits per cell. QLC is also now available and demonstrations of 5 bit per cell (PLC) have also been done. It now seems possible to stack hundreds of layers in tiers to achieve high areal density per die. For each additional layer in the 3D NAND, there will be an incremental increase in die density, but also an incremental increase in wafer cycle time and in cost per wafer.

The path to continuing to increase the die density and reduce the cost per bit will involve: minimizing the growth in layer count, increasing the density of memory cells per layer (increasing areal density per layer), decreasing the size of the memory holes, increasing the number of bits stored per transistor (TLC to QLC to PLC), and maintaining a uniform high aspect ratio etch for each memory hole. These are design and manufacturing issues and must be solved to maintain the good yields necessary for profitable manufacturing.

Ultimately, there will be a limit to the number of layers that 3D NAND can practically achieve, but it will be some time before that limit is reached. As of 2023, projections of up to 1000 layers has been discussed. Any alternative memory technology that could potentially replace 3D NAND in the future will be faced with similar manufacturing issues. On a 3D NAND with 128 layers, the depth of the memory hole (120nm in diameter) was reported to be 6-8 microns or about 55nm per layer. On a recent 200-layer 3D NAND, the stack was reported to be 5.5 microns or about 20-30nm per layer.

A practical problem with the increasing density of 3D NAND flash die each generation is the reduction in the number of dies necessary to create an SSD of a given capacity and the performance limit per die. Using a 1Tb TLC die (128GB per die), only 2 die are needed to create a 256GB SSD. The performance of this SSD will be bottlenecked by the lack of parallel NAND channels to the SSD controller as well as the speed of each NAND channel. Work is now taking place in JEDEC to improve the interface speed and performance of future NAND flash devices.

COMPETING TECHNOLOGIES IN THE SHORT TERM

The single most important enabler for success in the major semiconductor memory markets is low cost per bit. In the semiconductor memory market for discrete external memories, DRAM and NAND flash are dominant and are expected to remain so for the next decade due to economies of scale.

NAND flash memory continues to make solid inroads into the established mass storage markets. SSDs and all flash arrays are popular as a fast storage layer between HDDs and DRAMs, and SSDs have replaced high spindle-speed HDDs (10k & 15k RPM). However, the ultra-high capacity HDD market will remain unavailable to flash for the foreseeable future due to its lower cost per bit.

TECHNOLOGY NEEDS: RESEARCH, DEVELOPMENT, IMPLEMENTATION

The IRDS is carefully evaluating the technologies that will become necessary to allow NAND flash to continue to be viable through the next decade. This technology development will be collectively funded by all participants in the semiconductor memory business.

Silicon is very well understood and this gives the material a great cost advantage over competing technologies. Silicon always gains the upper hand in manufacturing costs.

GAPS AND SHOWSTOPPERS

Over the past couple of decades, researchers expressed a belief on several occasions that insurmountable obstacles to scaling flash memory were at hand. Once that obstacle was reached, other researchers have succeeded in devising some ingenious method of sidestepping it. This has happened for at least four planar NAND process generations, and when planar NAND flash finally did hit the wall, it was replaced by 3D NAND flash which will enable NAND flash technology to continue to dominate the non-volatile memory market for at least the next decade.

The cost of a gigabyte of NAND flash memory has continually decreased at an average rate of 30% annually since its invention. In 2002 NAND flash transitioned from SLC to MLC, doubling the bits per cell. Volume production of 2D TLC commenced in 2010 with TLC being the dominant type of 3D NAND today. Some four-bit QLC NAND has shipped in the past, but this technology is expected to become a significant portion of the market in the future thanks to the robustness of today's 3D NAND flash technology.

RECOMMENDATIONS ON PRIORITIES AND ALTERNATIVE TECHNOLOGIES

A very high level of ongoing capital spending is necessary in the flash business. NAND suppliers must be willing to continue to invest even when the market suffers a price collapse. This is the recipe for success in any undifferentiated semiconductor market.

High-capacity data storage is experiencing incredible demand growth due to cloud providers and hyperscale data centers. The key metric has always been total cost of ownership, but performance and energy consumption per rack. While HDD continues to offer a lower cost per GB, the gap continues to shrink.

3D NAND flash also the storage requirements of mobile devices include low power drain, small size and low cost. Compact, inexpensive, efficient, high-density nonvolatile storage is critical for smart phones, tablets, laptops and other portable consumer devices.

Today's most promising alternative technologies are MRAM (Magnetic Random Access Memory), which is viewed as a potential successor to DRAM, SRAM and NOR Flash, ReRAM (Resistive Random Access Memory), which might replace NOR flash and SRAM in the future, FRAM (Ferroelectric RAM), and PCM (Phase-Change Memory). All these technologies are non-volatile and in mass production but have yet to encroach on the established DRAM and NAND flash markets.

Again, the problem is the lack of economies of scale, so today, these promising alternative memory technologies are mostly confined to the embedded memory market such as RFID, mass-transit fare cards, power meter readers, gaming systems as well as various other consumer, automotive and industrial markets.

OTHER EMERGING SOLID STATE MEMORY TECHNOLOGIES

SITUATION ANALYSIS

Since the 1960s, semiconductors have been undergoing a constant pace of ~30% annual cost reductions based on regular reductions in the size of the transistors on the chip. This process, known as "scaling" involves reducing the length and width of transistors mainly through lithographic techniques. The fact that this trend follows a constant slope was noticed in 1965 by Intel founder Gordon Moore and has been given the name "Moore's Law". Gordon Moore's 1965 paper noted that the number of transistors on a chip doubled every year or two, and, based on this trend, Moore was able to project the transistor count of chips a decade into the future.

Memory chips have been following Moore's Law since that time, doubling their density (the number of bits on a chip) approximately every two years by continually shrinking the size of the transistors on the chip. This doubling of transistor density each generation was achieved by shrinking the lithography by approximately 30% each generation. Since area is linear dimension squared, 0.70 * 0.70 = 0.49, or about half the size required for a given number of transistors for each generation.

If density scaling continues to follow Moore's Law, both DRAM and 2D NAND flash were expected to reach a point where they could no longer scale due to the fact that both of these memories use charge storage to represent bits, and the charge that the memory cells could store would eventually be too small to detect. 3D NAND flash has sidestepped this issue by increasing layer count rather than reducing memory cell size, but DRAM scaling has slowed significantly.

At what point alternative memory technologies like MRAM, FeRAM, ReRAM, or others, might be able to displace DRAM or NAND is unknown. Today, all of these alternative technologies are still more costly per bit than DRAM or NAND flash, and this prevents them from being selected as replacements except in those rare circumstances where the cost is less important than certain important attributes they provide.

The alternative memory technologies are all "persistent" or "nonvolatile", that is, they retain their data even without power, and this allows them to be used as storage as well as working memory. This is a key differentiator between NAND flash and DRAM – DRAM is fast (though not as fast as SRAM) but it cannot store data when powered down (it is a volatile memory). NAND flash is about one thousand times slower than DRAM, but retains its data without power. This allows NAND flash to be used as storage while DRAM can only be used when powered.

The IEEE IRDS Road Map indicates that 3D NAND technology could extend the life of NAND flash to 2030 and beyond, enabling chip densities to continue to increase even using process lithography that would no longer be the most aggressive.

However, embedded NOR flash, used for code storage in embedded devices has reached a scaling limit at 28nm and this has opened up opportunities to replace embedded NOR with persistent memories that can scale to smaller size. MRAM and ReRAM are beginning to replace NOR flash in some embedded products. Likewise, embedded SRAM may also face scaling limits at around 14nm and SRAM cells are very large because of the many transistors they use. For this reason, there are some embedded devices that are starting to use MRAM or ReRAM to replace slower SRAM caches in some embedded products[i].

FERROELECTRIC RANDOM ACCESS MEMORY (FRAM OR FERAM)

Ferroelectric memory, commonly known as FRAM or FeRAM, is a technology that has been in the marketplace for over 30 years and one could argue that it has shipped in more products than any other emerging memory technology because of its widespread use in embedded RFID applications (such as fare cards). Discrete devices are available Fujitsu, Infineon, and Rohm, while some Texas Instrument microcontrollers have embedded FeRAM along with ASICs from Fujitsu. Densities range from 4kb to 16Mb.

Most FRAMs are used in mass-transit fare cards, RAID drives, gaming systems, and power meters, where data must be rapidly written into nonvolatile storage while consuming very little power. In RFID cards this power is generated from the interrogating radio waves. In other applications power from a small charge on a capacitor is used to move data from a RAM into an FRAM in the

[i] Emerging Memories Branch Out Report, Coughlin Associates and Objective Analysis, 2023

event of a power failure. All of these applications can justify FRAM's higher cost structure because of the extraordinarily low energy required to write data into the chip.

The "Ferroelectric" name is a misnomer given to Perovskite crystals that exhibit a state change similar to the hysteresis curve of magnetic media. An atom located in the center of these crystals can be pushed from one side of the crystal to the other by an electric current. The atom will stay in that position until a current in the opposite direction moves it back. Current flow through the device follows the familiar "S" curve of hysteretic devices, and can be used in its two stable states to represent a logic "1" or a logic "0". This storage mechanism is insensitive to radiation, unlike the charge-storage mechanism used by DRAM and NAND flash, so there is some interest in its use in avionics and military applications, although PCM and MRAM currently hold more appeal than does FRAM in these markets.

A ferroelectric memory is manufactured by layering the Perovskite material onto standard silicon CMOS logic as the dielectric of a storage capacitor. This adds cost to the wafer, making it uncompetitive with established technologies. Ferroelectric materials disagree with silicon, so a barrier layer must be used to protect the silicon from the ferroelectric material. This barrier layer commonly consists of platinum, whose high cost does not increase materials costs appreciably since insignificant quantities are required, but since platinum is very difficult to etch the cost of processing a wafer is higher than that of standard silicon. As a result, and also because of the small manufacturing volume for FRAM chips, FRAM wafers are more costly to produce than flash wafers.

A further difficulty is that FRAMs have a larger two-transistor memory cell, and therefore a larger die size, than single-transistor technologies like DRAM and flash memory, so their cost would be higher even if the wafer processing costs were the same as those of established technologies.

In theory FRAMs' inherent die size and processing cost disadvantages can be solved, but this solution has not been put into practice. Today's FRAMs cost significantly more to manufacture than NAND flash memories of the same capacity. This indicates that it will be a long time before FRAMs find use in mass-storage applications.

Fujitsu offers embedded FeRAM LSI products (RFID, authentication, and custom). Texas Instruments offers an MSP430 microcontroller with FRAM (MSP430FR5xxx). Work by larger manufacturing companies has brought much-needed R&D spending to the technology, without which FRAM could never reach cost competitiveness with pure-silicon memories.

Applications that use FRAM find this technology attractive since it offers the high speed and low power dissipation of SRAM with a cost structure that is theoretically similar to that of DRAM, and is nonvolatile like flash. The blend of these three features means that the device could displace all of today's existing memory technologies should it reach the point where it can be manufactured cost-effectively.

Magnetoresistive RAM (MRAM)

Magnetoresistive random access memory has also been in development since the mid-80s. Data storage in these devices is in the form of storing magnetic spins vs electron charge storage utilized in flash memory. Fundamentally, it is composed of a magnetic free layer separated from a magnetic pinned layer by an insulating tunneling layer. The resistance of this magnetic tunnel junction (MTJ) is determined by the orientation of the magnetic free layer with respect to the pinned layer. If the two layers have the same magnetic alignment, the resistance through the magnetic tunnel junction is minimized. If the two layers have opposite magnetic alignment, then the resistance is maximized. The two types of MRAM currently in production are toggle MRAM and spin torque transfer (STT-MRAM).

Other MRAM technologies are being developed with active research ongoing. Two types of MRAM memory, Spin-orbit torque (SOT) and voltage controlled magnetic anisotropy (also known as magnetoelectric RAM or MeRAM) are being considered as faster and lower energy MRAM storage for future applications.

As a technology, it has been promoted as a potentially universal memory, replacing the functionality of both DRAM and flash in one memory. Discrete devices are available from Everspin, Honeywell, and Avalanche Technology. The main applications for stand-alone MRAM to date have been avionics and military applications due to the insensitivity to radiation, fast write speeds, and unlimited endurance. Currently, densities up to 1 Gb (STT-MRAM) are available.

MRAM is showing up in many embedded devices as a replacement for NOR flash for code storage because of embedded NOR flash scaling limits and as a replacement for larger slower and higher-level SRAM cache for applications such as AI inference devices. As the scaling used in embedded devices decreases it is expected that the use of MRAM in embedded electronics will increase, driving up production volume and lowering the costs of manufacturing.

RESISTIVE RAM (RRAM OR RERAM)

The resistive RAM or ReRAM is a large umbrella category that covers a number of subcategories, namely PCM, PMCm, Oxygen Depletion Memory.

PHASE-CHANGE MEMORIES (PCM, PCRAM, OR PRAM): OVONIC UNIFIED MEMORY (OUM)

A number of companies are investigating phase-change memories, a type of ReRAM that uses the crystalline/amorphous phases of a material to determine whether a bit is a "1" or a "0". Most of these companies were originally under license from Ovonyx, the company that invented and owned the technology that it called the OUM or Ovonic Unified Memory, named for the material's inventor Stanford Ovshinsky. Intel, Micron, IBM, SK hynix, Samsung, and BAE are the most visible, and many other licensees may be conducting undisclosed research.

In 2012 Ovonyx was liquidated, and many of the OUM patents were acquired by Micron Technology.

PCM uses current to heat a small volume of a phase change material virtually identical to that used for the storage layer in CD-RW disks. The heat changes the state of the material back and forth between a crystalline and a polycrystalline structure, thus changing its resistance to a read current. This change is reversible depending on the cooling profile. Today's materials have been well characterized but they are not fully understood. Research at the University of Pennsylvania points to an alternate method of creating an amorphous area that could lead to improvements in both programming methods and materials development.

Today's PCM chips use chalcogenide glasses. Some chalcogenides (alloys containing one or more group VI elements) exhibit reversible change between the disordered (amorphous) and ordered (crystalline) atomic structure, the most common being $Ge_2Sb_2Te_5$. When the crystal is amorphous it exhibits a high electrical resistance, which lowers when the material crystallizes. The application of highly localized heat drives rapid, reversible changes between the amorphous and crystalline structure. A resistive heating element generates Joule heating, the temperature of which determines whether the material comes to rest at an amorphous or a polycrystalline phase.

PCM has good technical specifications:

- Fast switching (<30ns),
- Good endurance (>10^{16} write/erase cycles),
- Imperviousness to cosmic radiation,
- Theoretical ease of integration into CMOS processes
- Scalability to smaller dimensions

PCM is an expensive form of memory today because the die size is larger than that of flash implying higher costs, and wafer costs are high. The material promises to be able to migrate beyond flash's scaling limit, and this scalability is a compelling reason for researchers to consider PCM as a replacement for NAND flash. PCM technology appears scalable to at least 5nm. Both programming current and operating voltage decrease with process geometry, giving PCM an advantage over flash. PCM should be relatively simple to add to existing CMOS or bipolar structures. Stacked PCM structures promise to produce effective cell sizes smaller than those of FRAM.

PCM commercial products have been introduced by Micron-Numonyx and BAE, and this technology is one reason Micron gave for its 2010 acquisition of Numonyx. Samsung also produced PCM for a time and used these chips in a low-end Samsung cell phone, but it appears that no parts shipped outside of Samsung. Samsung is believed to have cancelled its PCM development. The only company currently producing SoCs with embedded PCM is STMicroelectronics for automotive applications.

In 2015, Intel and Micron announced a technology they call "3D XPoint Memory" (with the word "XPoint" pronounced as "Crosspoint"). This was a form of PCM, although the companies just said that it was a resistive memory. We will talk more about 3D XPoint Memory (which Intel sold as Optane memory) later..

Programmable Metallization Cell Memory (PMCm)

A good bit of research has been devoted to Programmable Metallization Cell Memory, or PMCm. This ReRAM technology performs similarly to the chalcogenide glasses used in PCM but counts upon metal threads migrating through the dielectric. These metal threads take the place of the crystalline/amorphous areas in the PCM to provide either a high-resistance or a low resistance path.

Adesto Technologies is shipping a product it calls "CBRAM" (Conductive Bridge RAM) which is based on this sort of cell manufactured using chalcogenide glass with copper threads.

Another firm, Crossbar Technologies, promised to have a product by 2015, but has not yet shipped or sampled a product. It has more recently been promoting use of this technology to create a physically unclonable function (PUF) chip for product verification.

The benefits of these devices are similar to those of PCM, but where PCM is temperature sensitive, requiring special solder flow processes to attach a pre-programmed PCM chip to a PC board, PMCm is not temperature sensitive and can withstand the rigors of a standard soldering process.

At this early stage in the product's life, it is too early to determine whether PMCm will achieve a competitive cost advantage over other alternative technologies.

OXYGEN DEPLETION MEMORY

Recent strides have been made in the development of Oxygen Depletion Memory, another form of ReRAM that employs a poorly-understood technique of moving oxygen atoms into and out of a glass layer to increase or decrease that layer's electrical resistance. The two companies that have disclosed the most information on this technology are Unity Semiconductor, which was acquired by Rambus, but later shuttered when the technology failed to perform as anticipated, and Hewlett Packard, whose R&D labs made promises of the technology's development in the early 2010's, but have missed every promised milestone to date.

The HP (now HPE) version of the technology has been dubbed "Memristor" and it was positioned as a fourth missing basic electronic component after the resistor, capacitor, and inductor. It was to have been the sole memory type used in HPE's ambitious new computer architecture dubbed: "The Machine". Although HPE and partner SK hynix produced prototypes, insufficient progress was made to continue to rely on the memristor, and the first rendition of The Machine was converted away from memristor memory to DRAM use in the summer of 2015.

Oxygen depletion memory technology has had a number of near misses, but has never quite made it to the sampling stage, so today it is farther behind than PCM, PMCm, MRAM, or FRAM.

COMPARING THE TECHNOLOGIES

This section will attempt to point out the comparative strengths and weaknesses of each technology. As has been mentioned earlier, the most important factor is cost, and the memories that account for the dominant share of the market are those with a lower cost structure than their counterparts: DRAM and NAND flash.

NAND flash continues to be the cost leader in the cost-driven memory market by a very wide margin, making it difficult to determine which of the emerging technologies is more likely to supersede today's memory technologies.
The fact that NAND is likely to maintain a decided lead for a number of years also complicates matters. Since it will be a very long time before NAND flash memory is displaced by an emerging

technology, any one of these emerging technologies might achieve a breakthrough over the next several years that might put it well ahead of the other emerging technologies.

Table 2 lists a number of salient features of today's three leading memory types: DRAM, SRAM, and NOR and NAND flash, and contrasts these against the attributes promised by MRAM, FRAM, PCM and ReRAM: an umbrella category that groups together PMCm, and Oxygen Depletion Memories.

Table 2. Solid State Memory Technology Comparison[i]

	Established Memory Types				Emerging Memory Types				
	SRAM	**DRAM**	**NOR Flash**	**NAND Flash**	**FRAM**	**ReRAM**	**Toggle MRAM**	**STT**	**PCM**
Nonvolatile?	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cell Size (f^2)	40-500	6-10	10	0.03-5	10-32	4-50	16-32	40-160	4-50
Read Time (ns)	1-100	30	50	10,000	20-50	10-20	3-20	3-15	5-20
Write Time (ns)	<1	1-10	10^{5-7}	10^{4-6}	50	10^{1-5}	1-10	1-100	10^{6-9}
Endurance	∞	∞	10^5	10^{3-5}	10^{15}	10^{3-9}	∞	10^{6-15}	10^{6-9}
Write Energy	Low	Low	High	Med	Low	Low	Rather High	Low	Low
Write Voltage	None	2	6-8	12	2-3	1.2	3	1.5	1.5-3

Source: Objective Analysis & Coughlin Associates

Any of the four emerging technologies (FRAM, MRAM, PCM or ReRAM) will provide clear technical advantages over the established memory types if it can be brought into price competition with any of them. The price issue is paramount. Should any of the established technologies reach a scaling limit one of the newer technologies is likely to take over its market.

Figure 8 shows another way of comparing many of the same attributes of these technologies. Note that Ferro refers to FeRAM. This radar chart is set up so that each vector is either well satisfied (100%) by a certain technology, or poorly satisfied (0%). The values represent a subjective rather than an absolute valuation. The intent is to show that the new memory technologies (represented by the dashed lines) have stronger technical attributes than today's leading technologies, (solid lines.) This advantage leads to the dashed lines' running along the outside of the chart, rather than dipping towards the center.

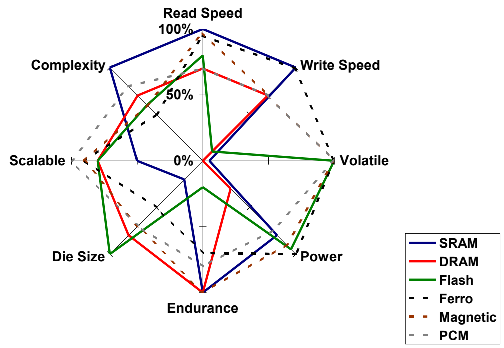


Figure 8. Radar Chart of Memory Technical Attributes

Source: Objective Analysis

The main advantage that any of these technologies has over flash is an ability to continue to scale past the flash scaling limit. The line "Scalability" shows how each technology is expected to follow standard semiconductor lithographic scaling. This scaling is the means that has been used for the past half century to reduce the cost of semiconductors – the transistors are made smaller every generation. Notice that scaling is limited for DRAM and NAND flash. If a technology can be scaled but it does not have good cell density or array efficiency then it will still have trouble displacing flash.

STORAGE CLASS MEMORY (SCM)

This section will focus on a category of non-volatile memory called Storage Class Memory (SCM). This category is important since it set the stage for the 3D XPoint memory that Micron and Intel introduced and then suspended development on. The experience with 3D XPoint (Intel's Optane memory) points out the difficulties of displacing or augmenting existing standalone memories.

SITUATION ANALYSIS

The sections above explained the need for emerging technologies to replace today's entrenched technologies once today's technologies reach their scaling limit. We also noted that the emerging technologies are all "nonvolatile", that is, they retain their bits even without power, and this allows them to be used as storage.

The ability to store data is a key differentiator between NAND flash and DRAM – DRAM is very fast but loses its data when powered down, while NAND flash is one thousand times slower than DRAM, but retains its data without power. This means that NAND flash can be used as storage while DRAM can only be used as memory.

Emerging technologies fall somewhere between DRAM and NAND flash. They are almost as fast as DRAM, and they retain their contents without power. This implies that the system's memory, which has been viewed as volatile storage since semiconductor memory replaced magnetic core memory in the mid-1970s, could once again become nonvolatile or "persistent" as it was back in the era of magnetic core memories. IBM gave these new memory technologies the blanket name of "Storage Class Memory" or "SCM" since the company believes that new software architectures will be developed to take advantage of the fact that memory will be persistent and can be used as storage.

With this conversion from DRAM and NAND flash to SCM will come significant changes in computer architecture. Today's software and CPU architecture has evolved since the mid-1970s under the assumption that all memory is volatile and that all storage is many orders of magnitude slower than memory. IBM's SCM initiative was designed to anticipate upcoming changes in which storage and memory would become the same, with the expectation that the underlying assumptions of slow storage and fast, volatile memory would need to be reconsidered.

There have been numerous advancements designed to enable this change:

- The Storage Networking Industry Association (SNIA) established a nonvolatile memory (NVM) programming model to streamline software stacks that communicate with SCM
- New memory modules (DIMMs) that pair DRAM with NAND flash to emulate nonvolatile memory modules became a distinct market
- The Joint Electron Device Engineering Council (JEDEC, a standards-setting body for semiconductors) defined naming conventions for three kinds of nonvolatile DIMMs: NVDIMM-N, NVDIMM-F, and NVDIMM-P
- Intel and Micron announced a new SCM memory type they call 3D XPoint memory in the summer of 2015 that was to be used as an SCM in future Intel processing platforms (to be discussed in detail below)
- Intel introduced SSDs using its Optane memory (3D XPoint) as well as DIMM devices for their processors
- Intel added new instructions to its standard Intel Architecture (IA) processors to support SCM functionality.

THE INTEL/MICRON 3D XPOINT MEMORY

In the summer of 2015 Micron Technology and Intel Corporation announced a new memory type called 3D XPoint (three D cross point). This was marketed as Optane (Intel) from 2017 to 2022. Bit storage is based on a change in the bulk resistance rather than by charge storage as in the case of flash. Some industry participants believe that this technology is a re-branding of the PCM (phase change material) technology that these companies had been developing for years, but Intel and Micron has said that the technology is different than PCM and uses chalcogenide materials that are faster and more stable than traditional PCM materials.

3D XPoint could be considered to be a type of ReRAM (but so are traditional PCM devices). In early 2021, Micron announced it would cease production of 3D XPoint and the fab in Lehi was sold to Texas Instruments. In late July of 2022, Intel announced winding down the Optane division.

As a technology, 3D XPoint fit in between NAND and DRAM. It was sold at approximately half the price of DRAM, but about 5X the cost of NAND. Slower than DRAM, but faster than NAND, there was a potential market niche since it was a fast non-volatile memory. It was positioned as a new layer in the memory hierarchy – storage class memory.

In that case, 3D XPoint could be considered to be a type of ReRAM. However, in 2021, Micron announced it would cease production of 3D XPoint and the fab in Lehi, Utah was sold. In 2022, Intel announced the winding down of the Optane division.

As a technology, 3D XPoint fit in between NAND and DRAM. It was approximately half the price of DRAM, but about 5X the cost of NAND. Slower than DRAM, but faster than NAND, there was a potential market niche since it was a fast non-volatile memory. It was positioned as a new layer in the memory hierarchy – storage class memory.

BUSINESS/TECHNICAL ISSUES

The most important business problem standing in the way was that 3D XPoint memory had to be sold as prices lower than DRAM. Achieving economies of scale is difficult when DRAM and NAND are the two highest volume memories in production today.

As a result, it was likely that 3D XPoint memory needed to be sold below cost for an extended period with the hope that demand will drive higher production volume and thus lower costs. If this happened and the costs of production dropped below the prices the memory were sold for, losses would turn into profits.

Intel had good reason to be willing to do this. By selling 3D XPoint memory at a loss the company would be able to assure that the system's speed keeps up with that of the processor, allowing Intel to continue upgrading its customers to newer, faster, more expensive processors. If Intel can sell a processor at a price that is $100 higher by losing $20 on the system's memory then there would be an overall profit. For other companies it is less clear why they would be willing to sell the technology below cost to incubate the market. Indeed, Micron never manufactured any significant 3D Xpoint products for sale although it did announce some 3D XPoint products.

MANUFACTURING EQUIPMENT

The equipment used to manufacture this new technology was be identical to that used to produce standard DRAM and NAND flash chips as well as other types of semiconductors. Micron produced the first 3D XPoint samples in its Lehi, Utah USA NAND flash manufacturing plant using a 20nm double-patterning process that is based on the company's 20nm NAND flash process.

MANUFACTURING PROCESSES

The structure of the 3D XPoint memory appears in **Figure 9** below.

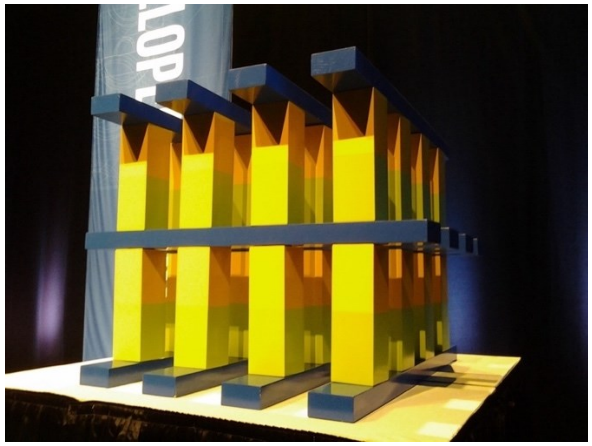


Figure 9. Physical Mock-Up of a 2-Layer 3D XPoint Memory Structure

It appears that a lot of the processing that is used for the 3D XPoint memory is based on the basic layered memory architecture that was pioneered by Matrix Semiconductor in the early 2000s prior to SanDisk's acquisition of that company. A microphotograph of the Matrix chip appears in **Figure 10**.

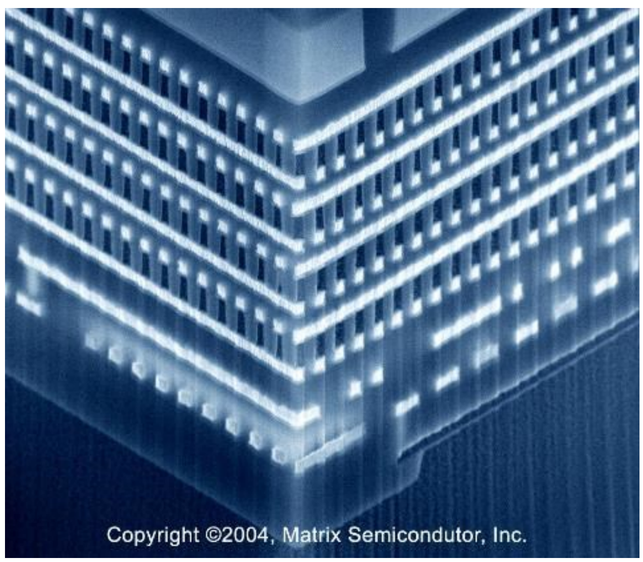


Figure 10. Stacked Crosspoint Memory Array

Although these processes were new and are more difficult than standard semiconductor processes, they were sufficiently well understood that they would not pose any undue difficulties. Today's PCM materials are based on germanium, antimony, tellurium, and aluminum.

Gaps and Showstoppers

The success of this effort depends upon Intel's support of a new memory layer in its future platform designs. Although Intel had a strong commitment to this effort and the size of Optane memory cells was $1/10^{th}$ that of DRAM several factors prevented this technology from every reaching breakeven.

Optane SSDs never achieved the performance boost for users because of limitations in the SSD interfaces for these products and so this market didn't develop the way Intel had hoped. The DIMM-based Optane products could only be used on Intel processors. Intel was a major driver of the compute express link (CXL) memory technology, which could be used to allow pooling of heterogenous types of memory with different performance characteristics, such as Optane memory. An earlier introduction of CXL might have increased the demand for Optane memory. Unfortunately, CXL did not come on the market until after Intel decided to discontinue development of Optane.

The bottom line is that, although 3D XPoint has a smaller cell size than DRAM the wafer production volume never reached the levels required for the production costs to drop below the price that Optane had to sell for to be competitive to DRAM (about half the price of DRAM). This inability to scale production for various reasons was the major reason why Optane memory did not become the SCM it was hoped it would become.

It is estimated that Intel likely subsidized Optane memory production by over $7B between 2016 and 2022. After selling the company's NAND memory business to SK hynix in 2020 and with Micron's withdrawal from making 3D XPoint memory in early 2021, Intel announced in late July 2022 that it would not continue to develop new Optane memory, although it would continue to ship and support existing Optane memory products.

Recommendations on Priorities And Alternative Technologies

Although Optane memory was not able to achieve the volume it needed to become an established player in the storage and memory hierarchy, its development spurred many other technologies that benefit mass storage and non-volatile memory. In particular, it spurred development of PCIe-based NVMe and CXL protocols and stimulated the development of software and firmware that could optimally use fast non-volatile memory.

Other companies are trying to fill the gap in the storage and memory hierarchy left by the withdrawal of Optane. These include NAND-based CXL devices from Samsung and SK hynix (and likely other NAND flash manufacturers) that can either expand attached DRAM memory at a lower price and introducing non-volatility or as part of a shared heterogeneous memory system using CXL. In order to deal with the lower endurance of the higher-level NAND-flash based products they are often using SLC flash to get closer to the high endurance (and performance) that Optane would have offered.

Introducing standalone memories using new non-volatile technologies has big barriers due to the need to get production volume up in order to bring prices down where the technology can be competitive. However, introducing these memories in embedded products is much easier as long at the new memory production processes can be easily combined with current CMOS processes. For most of the emerging memories this is being done at the back end of the line, after CMOS production is done.

For this reason, various emerging non-volatile memories are finding applications as memory in embedded products serving many consumer, automotive and industrial applications. These include replacing NOR flash for code storage and slower SRAM cache memory. Both of these memories face scaling issues in embedded applications and SRAM cells are much larger than the emerging memory cells and so more memory can be put in a given die area. Various companies such as Fujitsu, STMicroelectronics and Texas Instruments are utilizing FeRAM and PCM in their embedded devices.

All the major foundry companies, including TSMC, Global Foundries, Samsung and UMC are offering MRAM or ReRAM embedded memories and several SoC devices have come on the market, made in these foundries, for consumer, automotive and industrial applications. Growing volume of embedded memory may help drive overall production volumes and thus decrease the costs of using these non-volatile memories. This could create a virtuous cycle driving these memories to higher volumes and making them attractive for new applications.

For this reason, academic and industrial laboratories continue to develop and refine new non-volatile memories and industrial, automotive, consumer and aerospace applications will drive adoption of these technologies in embedded products and for standalone applications as well. Although 3D XPoint will not drive these changes, efforts on software and interfaces that can support high performance and non-volatile memory that 3D XPoint inspired, will enable more non-volatile memory applications.

Industry trade groups such as JEDEC, SNIA as well as professional technical standards organizations, such as the IEEE, can play an important role in setting standards for these products that will enable wide spread adoption and interoperability. Joint development and manufacturing efforts such as the 3D XPoint efforts by Micron and Intel, could help drive the adoption of future non-volatile memory technologies. Foundry support for these technologies will be critical for the creation of niche markets for new non-volatile memories.

At this point, it is not clear what kind of memory could become that new Storage Class Memory, but the system infrastructure is now being developed that will enable its usage – Compute Express Link.

COMPUTE EXPRESS LINK (CXL)

CXL is a high speed serial link based on PCIe that enables high speed connectivity between CPUs and devices (such as accelerators) and memory and is designed for high performance data centers. It uses three protocols: PCIe-based block input/output protocol (CXL.io), a new cache-coherent protocol for accessing system memory (CXL.cache), and device memory (CXL.mem). CXL was initially developed by Intel, and then by the CXL Consortium which was established in March 2019 by founding members: Alibaba Group, Cisco Systems, Dell EMC, Meta, Google, Hewlett Packard Enterprise (HPE), Huawei, Intel Corporation and Microsoft.

There were several competing consortia developing memory cache coherent standards prior to CXL: OpenCAPI, Gen-Z, and CCIX. In February of 2022, Gen-Z specifications and assets were transferred to CXL. On August 1, 2022, the OpenCAPI specifications and assets were transferred to the Compute Express Link (CXL) Consortium. On August 3, 2023, the CCIX consortium signed a letter of intent to transfer CCIX assets to CXL. So as of late 2023, the CXL consortium has emerged as the industry focal point for coherent I/O. (www.ComputeExpressLink.org)

CXL 1.0 was released in March of 2019. In September of 2019, the CXL Consortium was incorporated at CXL 1.1 was released. The CXL 2.0 spec was released in November of 2020, and CXL 3.0 was released in August of 2022. CXL 1.0/1.1 allows a host CPU to access shared memory on accelerator devices with a cache coherent protocol. CXL 2.0 adds support for CXL switching, to allow connecting multiple CXL 1.x and 2.0 devices to a CXL 2.0 host processor, and/or pooling each device to multiple host processors.

CXL 1.0/1/1/2.0 are based on PCIe 5.0 and supports x16, x8, and x4 widths natively. At 32GT/s, a bandwidth of up to 64 GB/s can be achieved. CXL 3.0 supports PCIe 6.0 at 64 GT/s and up to 256 GB/s using a x16 link. It also expands CXL switches to multi-level, allows peer to peer direct memory access, enhanced coherency and memory sharing. The Global Fabric Attached Memory device is also a new feature in CXL 3.0 which enables disaggregated memory separate from hosts (CPU, GPU, and other processing devices).

The three types of CXL devices are:
Type 1 (using CXL.io and CXL.cache) – specialized caching devices/accelerators (such as smart NIC) with limited or no local memory. Such devices require coherent access to host CPU memory.
Type 2 (using CXL.io, CXL.cache and CXL.mem) – general-purpose accelerators (GPU, ASIC or FPGA) with DDR, GDDR, or HBM locally attached memory.
Type 3 (using CXL.io and CXL.mem) – DRAM memory expansion boards and persistent memory. Type 3 devices provide host CPUs with access to large amounts of DRAM or non-volatile memory. Memory expansion beyond the directly attached DRAM via a CXL device will be possible with a Type 3 device.

CXL is an open industry standard and an important enabling technology. The existence of three prior consortia demonstrates the industry need. The standard offers a low latency and high bandwidth connectivity standard between processors and devices such as memory, accelerators, and smart I/O devices. It is expected to support the growing computational requirements for machine learning, artificial intelligence, and high-performance compute.

If one were to summarize the promise of CXL in two words, it would be "endless memory". Memory on a CXL board could be shared between processors, or banks of memory could be allocated to specific processors from the memory pool. While the initial CXL memory products are for DRAM expansion, it is not clear that this is the best use case. DRAM on CXL increases the latency and reduces the performance without directly reducing the cost. It offers to processors and accelerators a big pool of slower DRAM. This might be useful if the DRAM capacity is the main constraint in processing a workload. But a reasonably fast non-volatile storage class memory, that is lower in cost than DRAM, would be a better fit here.

Hard Disk Drive (HDD) Technology

SITUATION ANALYSIS

Of current mass data storage technologies, in terms of storage capacity shipped, hard disk drives (HDD) are by far the largest single component of the mass data storage industry. From its initial introduction by IBM in 1956, this technology grew to play a vital role in computing data systems. At this time the HDD market is undergoing a significant transition in markets with shrinking unit volumes yet growing overall bit capacity due to a number of trends. These include the declining desktop and notebook computer markets (signaling the demise of the 2.5-inch form factor in these segments) as well as the nearly complete displacement of high-performance enterprise class HDDs by SSDs (10k and 15K rpm drives have disappeared from the market leaving only 7200 RPM and slower). These declines are somewhat offset by strong growth of HDDs, in unit volume and even more in bit capacity, as warm/cool tier high capacity storage in data centers (nearline storage). It is anticipated that the nearline market will stabilize the overall unit decline in HDD volume by the last half of this decade as shown in **Figure 11**.

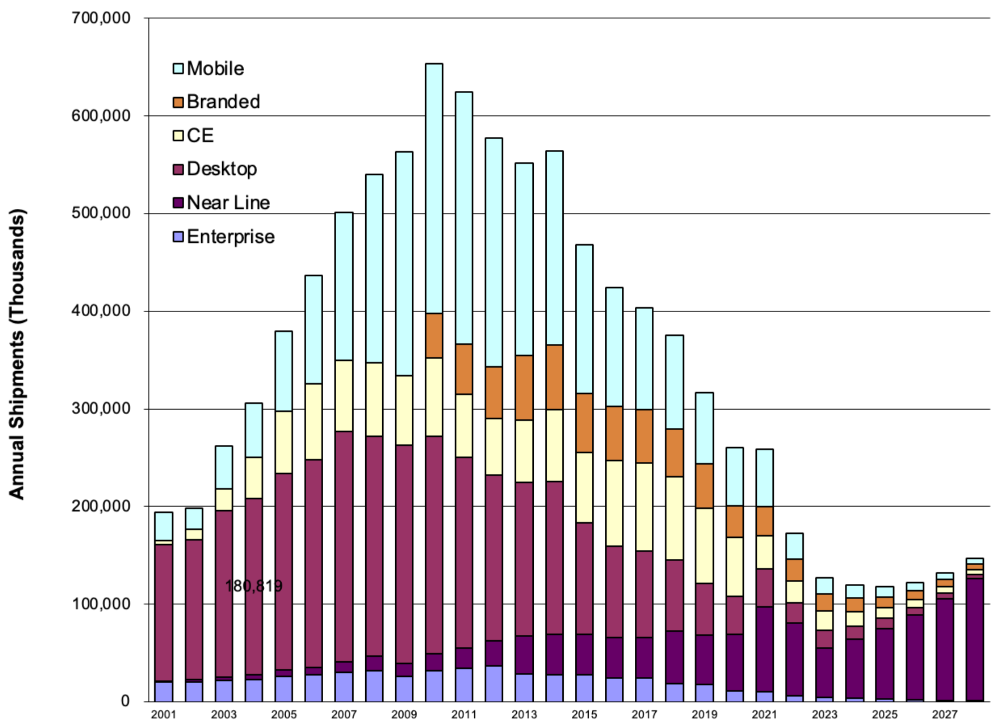


Figure 11. Shipped Disk Drive Volumes vs. Time, by Application

Source: Coughlin Associates

HDD's are also undergoing a significant transition in recording technology. Perpendicular Magnetic Recording technology has nearly reached its limit in areal density near 1.2Tbpsi. Higher HDD capacities are currently being achieved mainly through putting more platters in the drive enabled by the use of helium in place of air. Future areal density growth will depend on new technologies such as heat assisted magnetic recording, three-dimensional magnetic recording, and heated dot magnetic recording.

BUSINESS/TECHNICAL ISSUES

While HDD unit volumes have declined, with possible renewed growth later in the decade, prices per GB have continued to decrease (although somewhat slower than in the past due to slower areal density growth, Figure 12**),** driven by a number of key scientific and technical advances in magnetic head, disk, interface, mechanical, signal processing, and manufacturing efficiencies. While this has made HDD solutions suitable for a wider range of applications, particularly consumer and entertainment applications, it has also created unrelenting pressure on the companies that develop and manufacture HDD products to maintain profitability and business viability. More stable HDD prices due to industry consolidation and better control of manufacturing inventories have resulted in continued profitability by HDD companies, which could support the costs of introducing the next wave of technology innovation.

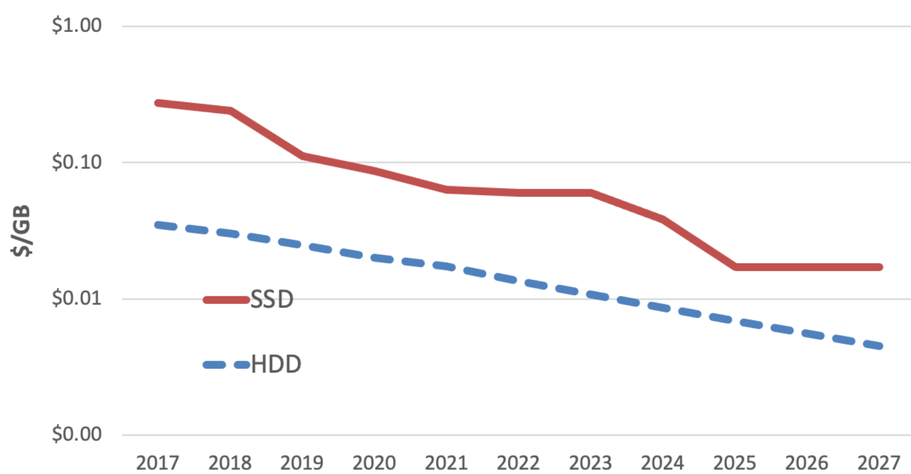


Figure 12. Raw Storage Average Retail Price vs. Time

Source: Coughlin Associates

MANUFACTURING EQUIPMENT

The magnetic mass storage industry depends upon a complex set of manufacturing tools. The critical lithography, pattern transfer and deposition equipment for head and disk production is similar to that of the semiconductor industry, with similar critical dimensions, a unique materials set and some different process requirements, such as extremely thin deposition layer thicknesses (< 1nm), magnetic orientation, and 3-dimensional nanoscale topography. A unique aspect of HDD manufacturing involves air bearing slider fabrication processes. HDD magnetic recording heads require precise sub-micrometer lapping and polishing techniques for sub-nanometer control of head-media spacing, magnetic head element and laser mount (for HAMR) geometry control. Another unique feature of magnetic disk (as well as tape recording) is that a continuous surface produced in a single manufacturing process is used for the recording media (unlike solid state and optical storage devices).

Mechanical assembly tools and test equipment are provided by a number of vendors, both North American and international based. Manufacturers use a number of different providers of software for manufacturing planning, process flow, inventory control, etc. The capital equipment required for a single HDD manufacturing line can be in excess of $50-100 million. Thus, the decision on commitment of additional production capacity must be made carefully. Key considerations here include costs of manufacturing and manufacturing test equipment (such as drive burn-in and test), and the need to reduce costs to meet customer requirements.

Storage industry associations such as the International Disk Drive Equipment and Materials Association (IDEMA), Advanced Storage Research Consortium (ASRC), the Storage Networking Industry Association (SNIA) and, in the past, the Information Storage Industry Consortium (INSIC) give visibility to the equipment manufacturers, organize technical meetings on critical issues, and form working committees to analyze and adopt common standards for component specifications, manufacturing tests, storage management and product performance specifications.

MANUFACTURING PROCESSES

Development of key manufacturing processes can be carried out internally, providing proprietary competitive advantage for drive vendors. Trade secrets for head and disk manufacturing and test can be easily buried in manufacturing processes that are difficult to copy. However, these secrets are sometimes not well enough understood, making process reproduction challenging. **Figure 13** shows the general steps involved in manufacturing a magnetic recording head for a disk drive from the wafer to the completed head gimbal assembly (HGA).

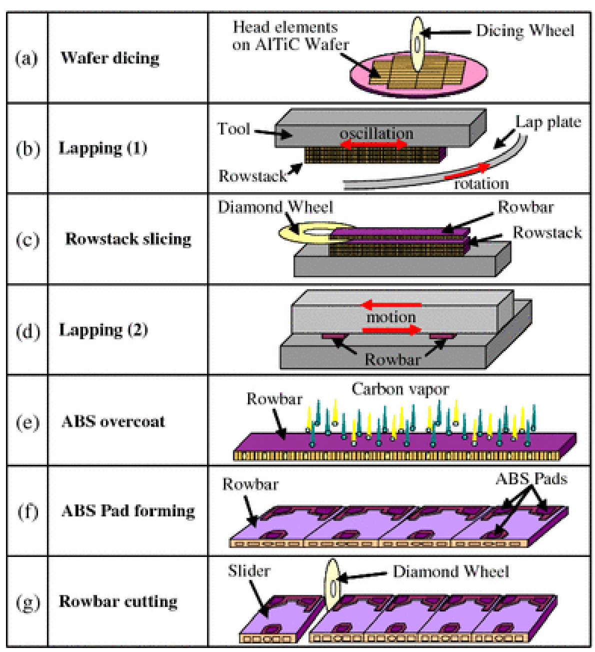


Figure 13. Hard Disk Drive Head Manufacturing Process

*Source: Naniwa, I., Sato, K., Nakamura et al. Microsyst Technol **15**, 1619–1627 (2009)*

Measurements are critical in manufacturing processes. They are needed initially in process development, and are vital in the transfer of information and components between vendors. This has generated a need for improved process characterization and control, along with an understanding at the engineering design stage of how a manufacturing process will operate in volume production. Examples of process characterization and control in the storage industry are the adoption of statistical tolerancing for mechanical components and assemblies, and the critical role of supply chain and industry control software tools. These software tools lead to greater efficiencies in supply chain management by automating the control of inventory and order anticipation using the Internet.

The storage industry has seen some HDD companies adopt the use of independent production contractors, which introduces an additional factor into the relationship between product design and manufacturing. This includes the key role of manufacturing vendors in printed circuit board (PCB) assembly and test, drive subassembly (head stack and actuator assembly), and hard disk drive integration testing. In the past there were some drive companies that had contract manufacturers build all or a substantial amount of their shipping drive volume. These drive companies could be called fab-less drive companies. The industry has moved to almost all vertical integration, with key head and disk technology development and manufacturing being brought increasingly in-house.

For example, Western Digital Corporation acquired the assets of both Read-Rite and Komag corporations for head and disk manufacture. HGST (Hitachi Global Storage Technology—acquired by Western Digital) and Seagate Technology established internal head and disk production capability. This integration allows close communication between the component manufacturers and the HDD plant, to assure critical process problems are addressed early in the manufacturing of a new product. The third remaining HDD company, Toshiba, buys all their heads from outside suppliers, TDK and Showa Denko (now part of Resonac).

Up until a few years ago, the rapid advances in areal density of magnetic recording meant that production lifetimes were shorter than in many industries. This required fast ramp-up to full production, and made manufacturing process development a key constraint in new product introduction. It also dictated that major changes from one product to the next must be more evolutionary, to reduce risk associated with the new technologies. High manufacturing yields are vital to assure a company’s ability to compete in this fast-paced industry. In the past, quite a few technically strong companies have been forced to leave the industry because of their inability to execute their yield ramps and cost controls on new product introduction.

MATERIALS

The leading companies are carrying out advanced materials development for their components and the final products. **Figure 14** and **Figure 15** identify typical components in a magnetic hard disk drive while **Figure 16** shows the relationship of the head and disk indicating the head flying above the disk and important characteristics of the head "slider."

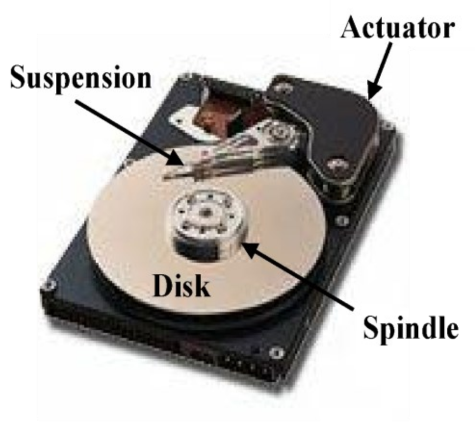


Figure 14. Typical Hard Disk Drive with Key Components Identified (Photo of a Samsung Hard Disk Drive)

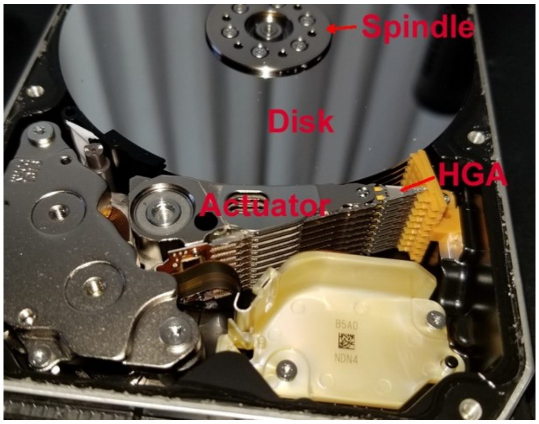


Figure 15. Close-up of Actuator Positioning a Head Suspension (HGA) on a Disk (Photo of a WD Drive)

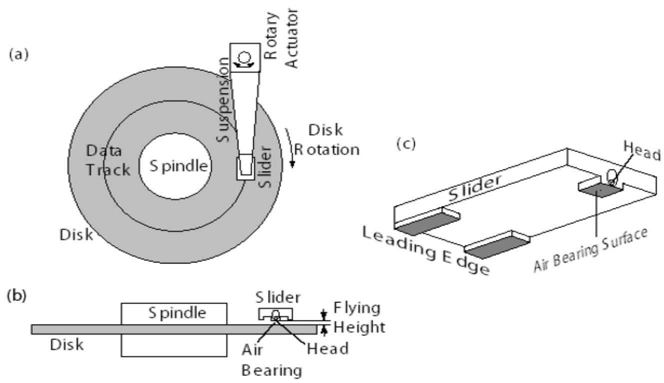


Figure 16. Relationship of Head and Disk Including "Flying Height" of the "Slider"

Source: S. Wang and A. Taratorin, Magnetic Information Storage Technology, 1999

Improvements in materials and processes have enabled several Japanese companies to become suppliers of critical components and subassemblies to the magnetic recording storage industry. Examples include disk substrates (high purity aluminum, as well as glass) and spindle motors (one company in Japan (Nidec), supplies all disk drive companies with motors). The only independent head vendor is a Japanese/Chinese partnership (TDK/SAE). TDK Headway/SAE has become well positioned to offer volume production of state-of-the-art magneto-resistive heads. The only remaining independent magnetic disk company is Showa Denko (now part of Resonac).

For the mass storage industry, critical advances are required in magnetic and non-magnetic materials for functional head read and write performance at high areal densities and high frequencies. These advances include: disk media and overcoats as well as heads capable of supporting densities in excess of 160 Gb/cm^2 (1,000 Gb/in^2 or 1 Tb/in^2); changes in media substrate technology; and dielectric films less than 1 nm thick for advanced GMR (Giant Magneto-resistive) and TMR (Tunneling Magneto-resistive) heads. In addition, technologies such as SMR (Shingled Magnetic Recording), (TDMR) Two-Dimensional Magnetic Recording, HAMR (Heat Assisted Magnetic Recording) an example of Energy Assisted Magnetic Recording and Heated Dot Magnetic Recording (HDMR), will require HDD designs with new processes and materials.

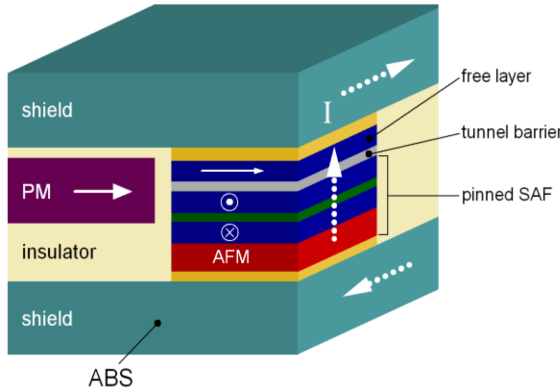
HAMR technology requires an investment in optical materials and processing as well as new media materials such as those based on FePt alloys. HAMR heads also require access to appropriate lasers for the media heating.

Ongoing improvements in permanent magnet materials will be a critical enabler for more efficient spindle motor and actuator design, and introduction of piezo-electric or MEMS (Micro-Electro-Mechanical-Systems)-based actuators required new materials development. The first commercial high-performance disk drives featuring dual stage actuators appeared on the market in 2005 in high performance enterprise disk drives and this technology is now used in all hard disk drives. **Figure 17** shows the basic structure of a modern TMR head transducer both as a schematic (16a) and a TEM (Transmission Electron Microscope) photo showing the deposited layers (16b). **Figure 18** shows an example of a piezo-electric-based micro-actuator for increasing the drive servo bandwidth and thus the available track density of recording.

Starting in 2005 several disk drive companies began to ship perpendicular recording HDDs for mobile computer applications**. Figure 19(a)** is an illustration of the older longitudinal recording technology while **Figure 19(b)** shows perpendicular recording. The soft magnetic underlayer (SUL) below the magnetic recording layer in the perpendicular media increases the effective magnetic write field during the recording process allowing writing on higher coercivity media.

Perpendicular magnetic recording allows making thermally stable magnetic recording at higher areal densities. For this reason, all drives in current production utilize perpendicular recording. This transition was very rapid and has allowed areal densities to be pushed to near 1.2 Tbpsi.

Competitive advantages can accrue to drive and component manufacturers if they play a leading role in the development of key material technologies since time to market for disk drive products is so critical to a company's success.)



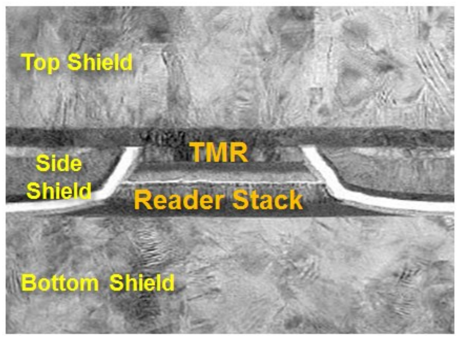


Figure 17. (a) Schematic of TMR head layers, including the Permanent Magnet (PM), Synthetic Anti-Ferromagnetic (SAF), and Anti-Ferromagnet (AFM) Structures and (b) Air Bearing Surface (ABS) view of a TMR head

Source: Seagate Technology

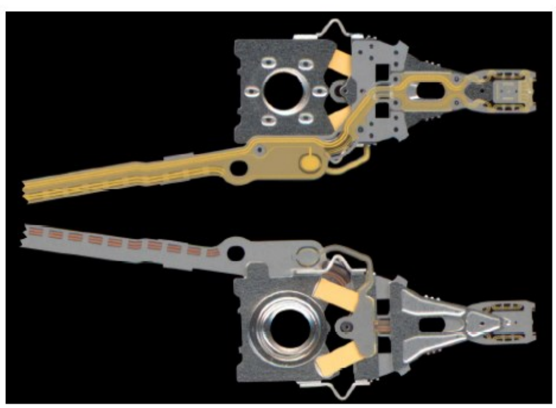
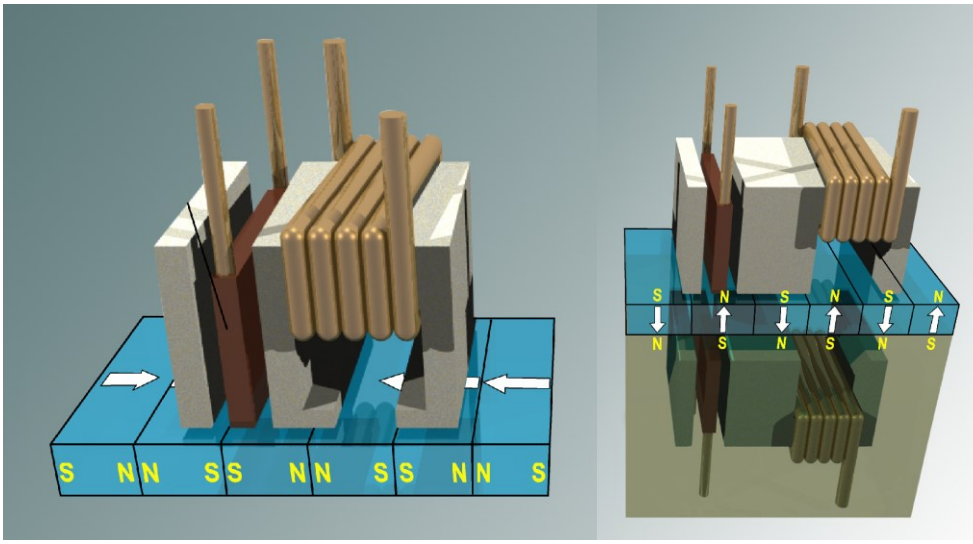


Figure 18. Example of a Piezo-Based Head Microactuator

Source: Hutchinson Technology Corp, now TDK



a) Longitudinal b) Perpendicular

Figure 19. Schematic Illustration of Magnetic Recording

Source: Seagate

ROADMAP OF KEY ATTRIBUTE NEEDS

Table 3 displays the current roadmap outlook of key attribute needs for Hard Disk Drive (HDD) storage technology, including capacity and performance metrics, as well as key head, disk, interface, and data channel characteristics.

Table 3. Magnetic Mass Data Storage Technology Roadmap – HDD

	Unit	2022	2025	2028	2031	2034	2037
Industry Metrics							
Form Factor (dominant form factor is bold)	inches	3.5, **2.5**	3.5, **2.5**	3.5	3.5	3.5	**3.5**
Capacity	TB	1-22	2-40	6-60	7-75	8-90	10-100
Market Size	units (M)	166	173	208	249	299	359
Cost/TB (avg.)	$/TB	13.6	6.91	3.46	2.60	2.00	<2.00
Design/Performance							
Areal Density	Tb/in^2	>1.0	>2.0	>4.0	>6.0	>8.0	>10.0
Rotational Latency	ms	2-12	2-12	2-12	2-12	3-12	3-12
Seek Time*	ms	3-5	3-5	3-5	2-5	1.5-5	1-4
RPM		4.2-10K	4.2-10K	4.2-7.2K	4.5-7.2K	4.5-7.2K	4.5-7.2K

Data rate	MB/sec	140-600	140-600	180-1,200	200-1,200	210-1,800	220-1,800
Power *Key Component Requirements*	watts	1.9-14	1.9-14	3-14	3-14	3-14	3-14
Read Head	type	TMR	TMR	TMR	TMR	TMR / CPP-GMR	CPP-GMR / LSV
Slider	type & size (% of micro, 3.86 mm^3)	5%	5%	5%	5%	<5%	<5%
Disk	type	AlMg, Glass	AlMg, Glass,	AlMg, Glass,	Glass, New Substrate,	Glass, New Substrate,	Glass, New Substrate,
Disk Static Coercivity	Oe	5,000-6,000 HAMR ≥30,000	5,000-6,000 HAMR ≥30,000	5,000-6,500 HAMR ≥30,000	5,000-20,000 HAMR ≥30,000	6,000-40,000	20,000-50,000
Magnetic Recording Technology		Perpendicular TDMR, ePMR, HAMR, MAMR, SMR	Perpendicular TDMR, , HAMR, SMR	Perpendicular TDMR, HAMR, SMR	Perpendicular TDMR, HAMR, SMR	Perpendicular TDMR, HAMR, SMR	Perpendicular TDMR, HAMR, SMR, Heated Dot Recording
Electronics/Channel	type	LDPC Iterative GPR (Turbo), Pattern Dependent Noise Predictive GPR	LDPC Iterative GPR (Turbo)	LDPC Iterative GPR (Turbo), TDMR	LDPC Iterative GPR (Turbo), TDMR	Soft ECC, TDMR	Soft ECC, TDMR
Channel Bandwidth	MHz	500-2,000	500-2,000	500-2,000	500-2,500	>2,500	<3,000
SNR	dB	<20	<20	<20	<20	<18	<18
Actuator	type	Conventional/Micro, +DSA	Conventional/Micro, +DSA	Micro, + DSA	Micro, + DSA	Micro, + DSA	Micro, + DSA
Spindle	type	Fluid	Fluid	Fluid	Fluid	Fluid	Fluid

**Seek time is one third full stroke seek time and does not include micro-actuator local track or rotational latency.*

CRITICAL ISSUES

BUSINESS VITALITY

Ongoing vitality and success in the HDD industry is critically dependent upon companies' ability to sustain satisfactory profitability in a competitive environment of continuous global price reduction and mounting pressure from alternative technologies. The ability to deal with eventual price reduction will require constant improvements in development and operational efficiency, while innovations in technology and manufacturing will allow leadership against competing technologies to be maintained.

AREAL DENSITY GROWTH RATE

As mentioned in this report there are several promising technologies such as HAMR, HDMR (Heated-Dot Magnetic Recording), TDMR, dual stage actuators, high sensitivity readers etc. that look likely to enable higher areal densities. See **Figure 20**. However, these technologies are expensive to develop and bring to market and may carry with them higher manufacturing costs and production yield risks. There are now fewer players in the disk arena and less US government funding for advanced technologies in this field. In a low margin industry such as HDD industry has been, the lack of funding could seriously inhibit commercialization. It is thus likely that new HDD technologies will continue to increase the annual areal density growth rate but it is unlikely that the 50% or higher average area density growth rate seen in the past will return. ASRC's areal density roadmap suggests 20% CAGR from 2022 to 2035.

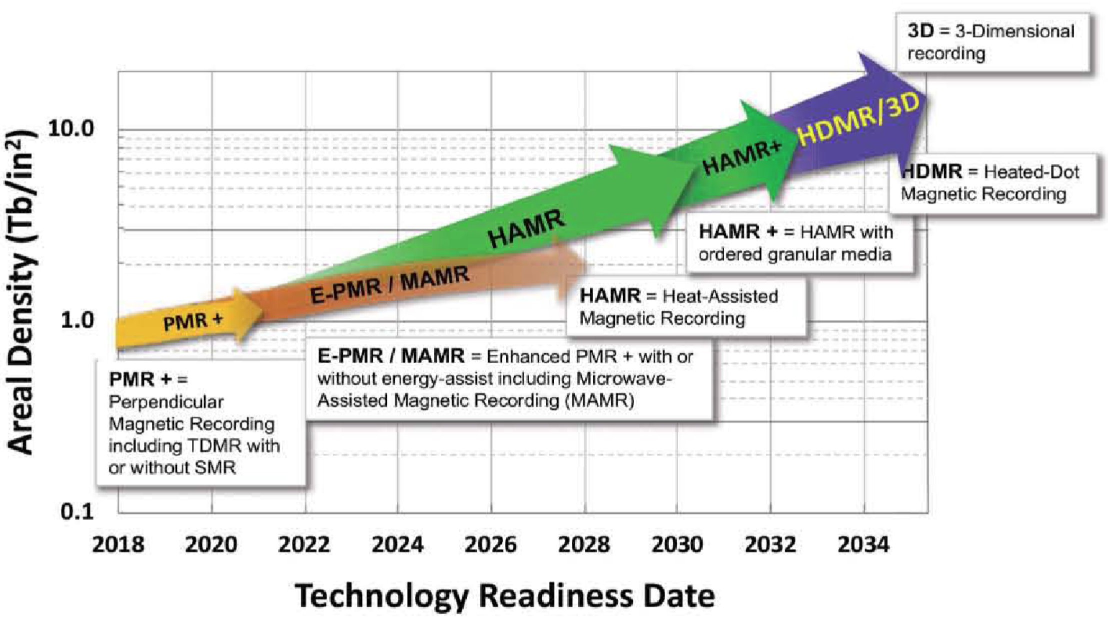


Figure 20. 2022 Areal Density Roadmap

Source: ASRC

COMPETITION FROM SOLID STATE DRIVES (SSDS)

SSDs are pressing HDDs in the area of cost at the low-end e.g. portable computer products, where capacities are relatively modest and entry cost, form factors and power consumption favor SSDs. At the high end, the superior performance (access time) of SSDs is gradually eliminating the HDD high performance enterprise (server) business. The remaining market where there is growth potential for HDDs is as high capacity nearline storage in data centers.

POWER CONSUMPTION

The power consumption of HDDs has become an increasing important issue as energy costs have risen and customers have become more environmentally conscious. Some of the power consumption of hard disk drives can be reduced by making the drives themselves more efficient (better power supplies, more efficient electronics, use of helium rather than air, etc.) but a large part of the problem is the excessive amount of time many drives spend in the idle mode i.e. spinning but not transferring data. SNIA has worked with the EPA (Environmental Protective Agency) on reducing the idle time in arrays and on developing an Energy Star rating for disk drives. The sealed Helium drives, including many data center and enterprise HDDs enable over a twenty percent reduction in idle power consumption.

SECURITY

In-disk drive encryption using FDE (Full Disk Encryption) is an excellent solution to the protection of data on HDDs. It is being introduced in more types of HDDs with time and is gradually being accepted by the marketplace. The HDD companies offer in-drive encryption in many of their hard disk drives.

TECHNOLOGY NEEDS: RESEARCH, DEVELOPMENT, IMPLEMENTATION

For a given set of storage components (heads, media, spindles, etc.), writing a higher amount of data to a given medium translates to an overall lower cost per gigabyte. Head and media development work continues at corporate and university laboratories. Much of the advanced work done on media and heads is basic science (where direct utilization of the information is less clear), and support has to be found for this work. Because of business demand, companies manufacturing magnetic storage products may not support long-range work needed to ensure future solutions in all areas.

Universities are a natural place for the work to be conducted, and there have been seven centers at major North American universities (Carnegie Mellon University, University of California San Diego, University of California Berkeley, University of Minnesota, and University of Alabama, as well as other groups world-wide such as those in Europe (Queen's University of Belfast and the University of York, in the UK) and Asia (Japan, Korea and Taiwan Universities as well as the National Univ. of Singapore, and A-Star in Singapore) have contributed to advanced magnetic recording development.

In addition, the guidance of industry-wide programs under the former IDEMA Advanced Storage Technology Consortia (ASTC), and currently with the ASRC, have provided guidance for long term "main line" storage programs for many years, and have fostered collaboration throughout the industry. This collaboration was also enabled by metrology technologies developed by the National Institute of Standards and Technology (NIST).

Research areas needing development to support continued growth of magnetic recording include heated dot recording media, new reader designs, heat assisted magnetic recording (HAMR), microwave assisted recording (MAMR), two-dimensional magnetic recording (TDMR), recording channel and preamplifier design, multiple actuator servo systems, skew mitigation, and advanced motor and multiple actuator design.

Disk drive track pitches are now < 100nm in size. Locating data and staying centered on such narrow track widths during reading and writing, while the medium moves past the transducer at almost 100 miles (160 km) per hour, is a huge engineering challenge. New actuator designs, shingled write, heat assisted magnetic recording and heated dot magnetic recording (formerly patterned bit media) technologies may offer solutions for further decreasing track pitch.

Taking advantage of advances in materials and processing, the disk drive industry is employing aspects of micro-electromechanical systems (MEMS) in order to increase servo system bandwidth and overcome motor bearing non-repeatable run-out and other eccentricities. This enables narrower track widths and faster data transfer rates. Dual stage actuators in the head suspensions using piezoelectric actuators built into the flexure are now used in all HDDs and many include a third level of actuation in the suspension.

GAPS AND SHOWSTOPPERS

The rapidly increasing areal recording density in magnetic recording is creating a problem. As the density of recording increases, for a given form factor recording device, the data that can be stored on it increases. As the storage capacity on a surface increases, the time it takes to access a given piece of data increases and the time to read or write all the data on the disk increases as well. Although the average performance of disk drives to access data is increasing by about 10% per year as shown in **Figure 21**, the disk access density (Access density = I/Os per second per gigabyte) of disk drives is continually decreasing as disk drive capacity increases faster than disk drive performance.

There are two main categories of challenge for HDDs

ACCESS DENSITY

While the areal density of HDDs has increased over the years, leading to devices with ever higher total capacity, the inherent seek and rotational latencies associated with HDDs has created a related problem that the I/O operations per second (IOPS) per byte (i.e., Access Density), is decreasing. Given that the growth market segment for HDDs is migrating to high capacity nearline storage, the Access Density problem is becoming increasingly acute, as shown in Figure 21. Lower Access Density, in turn, translates into a $/TB problem.

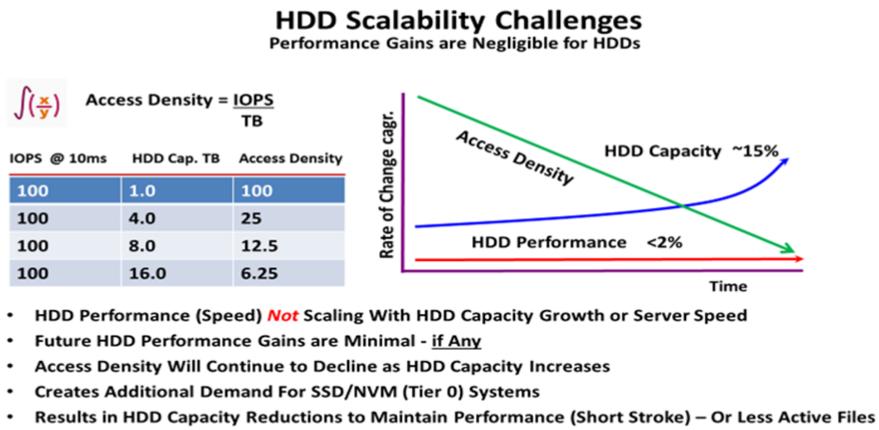


Figure 21. Disk Drive Access Density Is Decreasing

Source: Horison Information Strategies

A disk I/O is an input or output operation consisting of the seek time, latency time, and data transfer time for a disk drive. As a consequence, it takes longer to get to a given piece of data and especially, to move significant amounts of data. As the capacity of hard disk drives increases, the potential for a larger number of concurrent users or applications goes up, resulting in longer queue depths and slower response times. Late in 2017, Seagate announced multiple actuators to increase drive IOPS. In its first generation, Seagate's Multi Actuator technology uses two actuators on same pivot which enables a potential doubling in IOPS while maintaining the same capacity. Today both Seagate and Western Digital are shipping dual actuator disk drives and the number of actuators on a given pivot could increase in the future, helping to keep the Access Density at reasonable levels.

Another strategy to mitigate Access Density decrease is caching and multi-path storage architectures. Hybrid HDDs (drives featuring flash memory for caching and buffering were introduced in the marketplace in 2007. Also, some Solid State Drives (SSD) are marketed for caching and application acceleration in combination with HDDs to provide a significant performance boost while keeping the total $/GB costs at an acceptable level.

In May 2022 Western Digital introduced its OptiNAND SSDs. The 64GB of NAND in OptiNAND host metadata, provides a non-volatile write cache (Armor-cache), enables high resolution track interface management and enables enhanced servo capabilities. OptiNAND is used in Western Digital's 28TB UltraSMR HDD that offers 17% higher capacity than the non-SMR version at 24TB.

SNR

The biggest technical hurdle to overcome in extending the areal density of magnetic recording for HDDs is increasing the signal to noise ratio (SNR). It takes approximately 5-6 dB signal to noise increase over the previous generation of HDD to double the density. This additional SNR can come from the channel, the media or the heads. Better channels, improved reader and writer SNR and improved media have all contributed to increasing areal density. There are still channel improvements possible, but they are at the cost of increasing complexity (and thus more processing) and we are approaching the theoretical limits of data decoding.

TMR (Tunneling Magneto-Resistive) heads are in all modern HDDs. Write head improvements in conjunction with improved media (smaller grains, higher coercivity, H_c) have been responsible for significant gains in SNR. However, today's write heads have reached the limits of known materials in achieving write field strength with magnetic saturation fields of 2.4 Tesla.

The higher write field is necessary due to the increased magnetic anisotropy field, H_k, of the media grains. In addition to increasing SNR, the grain size in the media must be decreased without leading to thermal decay (thus requiring higher media coercivity). The introduction of HAMR and other energy assisted recording technologies, as well as heated dot magnetic recording (HDMR, formerly bit patterned media) offer solutions to extend density in the face of these increasingly difficult constraints. With the introduction of HAMR, track pitch will be defined by the thermal spot size and with HDMR, the patterned media bit spacing; therefore HAMR NFT (lithography, etch and deposition) and patterned media process (assisted self-assembly) equipment will become important as well.

In late 2023 Seagate began volume shipments of 32 TB HDDs with 10 disks using heat assisted magnetic recording (HAMR). Seagate projects higher areal density drives using HAMR in the next few years. Seagate plans for 50TB+ HDDs with areal density over 3.0 TBpsi by 2025, following more than a 20% annual increase in areal density.

Disk drive track density is a key technology to increasing areal density, particularly since linear density may be limited by thermal decay sensitivity. The track pitch is defined by the width of the write (and read) heads. These dimensions are created using lithographic techniques borrowed from the semiconductor industry.

RECOMMENDATIONS ON PRIORITIES AND ALTERNATIVE TECHNOLOGIES

Key to the continued health of the magnetic mass storage industry is innovation in a wide variety of technologies. These include: head magnetic properties for better writing performance and signal-to-noise ratios on read-back, magnetic media to support higher areal density, improved materials, advanced interface and air bearing designs, precision mechanical components, contamination control, high speed electronics, and signal processing technologies. Drive sizes less than 1.8" have disappeared-likely forever. 2.5-inch drives are still in production, but their number will continue to fall as sales of HDDs into PCs slows with growing adoption of SSDs. 3.5-inch HDD, particularly high-capacity drives for nearline applications will be the remaining HDD form factor in a few years.

SHINGLED MAGNETIC RECORDING (SMR)

One approach to improving areal density in hard disk drives is shingled writing. With few changes in existing heads and disks this technology has shown a practical increase in areal density of 10-18 percent (Western Digital's UltraSMR with OptiNAND has 18% higher capacity that the conventional magnetic recording or CMR version). Shingled recording (**Figure 22**) uses a recording head with a stronger, but asymmetric write field. This head is utilized with sequential track writing and a track pitch that is significantly smaller than the head effective magnetic write field width. The overlapping write process leaves behind a written track, which while much narrower than the write head width, is easily read by the even narrower TMR read head.

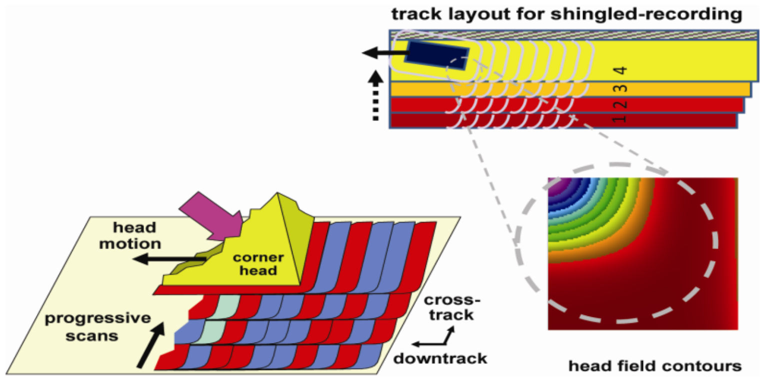


Figure 22. Shingled Recording and Head Magnetics

Source: Hitachi GST

Note that shingled writing has been used for years in magnetic tape recording. Shingled recording was introduced in LTO-2 tape products in 2003 (nearly a decade before the introduction of shingled recording in HDDs in 2013.

While this technology allows random reads, it does not readily accommodate random writes. Due to the nature of the write process, a number of tracks adjacent to that being written are overwritten or erased, in whole or in part, in the direction of the shingling progress, creating so-called "zones" on the media, which behave somewhat analogously to erase blocks in NAND flash. This implies that some special areas on the media must be maintained for each recording zone, or group of zones to allow random write operation, or random writes must be cached in a separate non-volatile memory.

Further, the writing of sequential data must be maximized to minimize the writing of random data. To accomplish this, special algorithms, caching, and metadata (data which describes data) are needed in a shingled HDD architecture. In 2022 Seagate said that nearly 25% of their nearline drives had been SMR and SMR has also been widely used in external backup HDDs. WDC has made similar statements.

SMR operation is most suited for sequential write workloads and has the potential to negatively impact performance for random write-intensive applications. To better manage these impacts in enterprise systems, particularly increased write latency, standards were developed which allow the host system to manage the SMR device's write characteristics.

Sealed Helium Drive

The density of helium is one-seventh that of air, enabling advantages such as less drag force acting on the spinning disk stack, thus substantially reducing the required spindle motor power and the fluid flow forces buffeting the disks and the arms are substantially reduced allowing for disks to be placed closer together and to place data tracks closer together (see **Figure 23**). The lower shear forces and more efficient thermal conduction of helium also mean the drive runs cooler and emit less acoustic noise. Finally, the use of helium also allows the use of larger diameter disks, thereby increasing total drive capacities.

Figure 23. Western Digital 22TB and Seagate - 20TB Sealed Helium Drives

Shipment of production models of these drives commenced in 2013 but market growth was slow at first until production volumes improved. Seagate Technology and Western Digital began shipping similar products in 2015. Both vendors are now shipping 20+TB models and these products are popular for data center and enterprise applications due their higher energy efficiency and enabling more disks and heads per drive.

Heat Assisted Magnetic Recording (HAMR)

Perpendicular recording is used in all commercial disk drive products. Areal Densities have been demonstrated using perpendicular recording at 2 Tb/in^2 . In spite of this progress, serious challenges face continued use of pure perpendicular recording without some enhancement. Conventional perpendicular recording or so-called conventional magnetic recording (CMR) is not extendable much beyond 1 Tb/in.^{2}. CMR density is constrained by (a) the need to reduce media grain size for higher areal density (AD), (b) the requirement to increase magnetic anisotropy to maintain thermal media stability and (c) the physical limits of write fields to record on the media. These constraints are commonly referred to as the Trilemma.

Heat Assisted Magnetic Recording (HAMR) is an alternative to increase magnetic recording areal densities by heating during the recording process. HAMR uses the fact that the coercivity (H_C) drops continuously, to practically zero, when a magnetic material is heated near its magnetic ordering temperature or Curie temperature (T_C). The HAMR writing process uses this effect by heating the media to an elevated temperature where H_C is below the writing field of the head. The heated medium region is then rapidly cooled down to ambient temperature after the head field is applied to write on the medium. This permits writing of media with much higher room temperature H_C and with smaller grains than conventional perpendicular recording, resulting in much higher areal density. A sketch illustrating the HAMR writing process is shown in Figure 24.

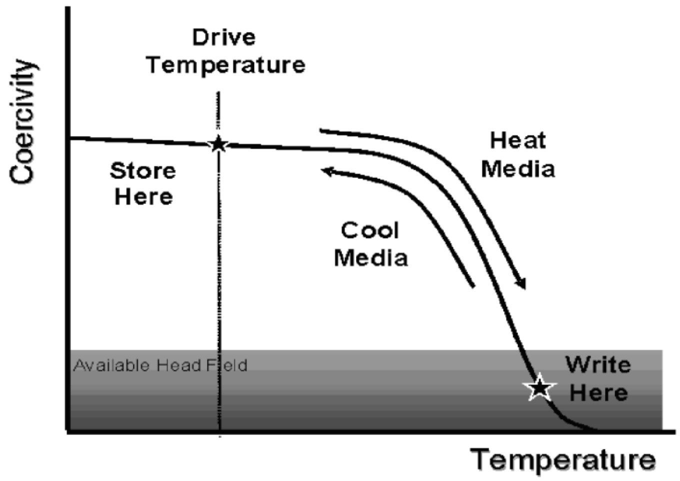


Figure 24. Curie Point Writing Using Head Assisted Magnetic Recording

Source: Seagate Corporation

A HAMR recording system needs the additional capability to heat extremely small spots on the medium along with all the features of CMR. Seagate's shipping HAMR system uses laser light focused on a Near Field Transducer (NFT) as heat source[i]. This system couples laser light into a specially shaped waveguide for light delivery to a NFT to form a deep submicron thermal spot on the media. A schematic overview of a proposed HAMR recording system is shown in **Figure 25**, to introduce its most important components. **Figure 26** shows an approach where the laser is built into the head itself.

Both the HAMR head and media in these two approaches can be manufactured on existing head and media manufacturing lines with some modification to existing processes and with additional components (such as lasers) and assembly tooling. Seagate has been shipping HAMR drives for data center and enterprise evaluation over the last couple of years and plans to ship these drives in volume starting in 2023.

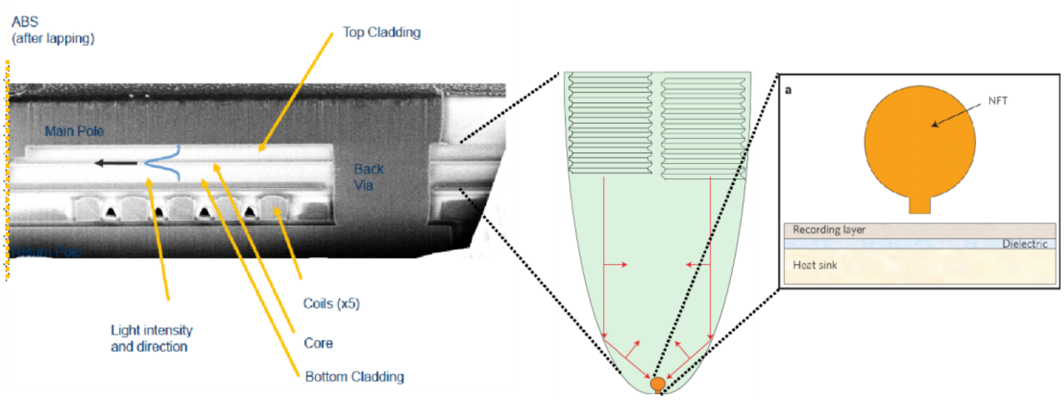


Figure 25. Sketch of basic proposed HAMR recording showing head design w/light impinging on the grating

Source: Seagate Technologies

---

[i] Rottmayer et al, "Heat Assisted Magnetic Recording", IEEE Trans. Mag., Vol. 42, No.10, p 2417, 2006

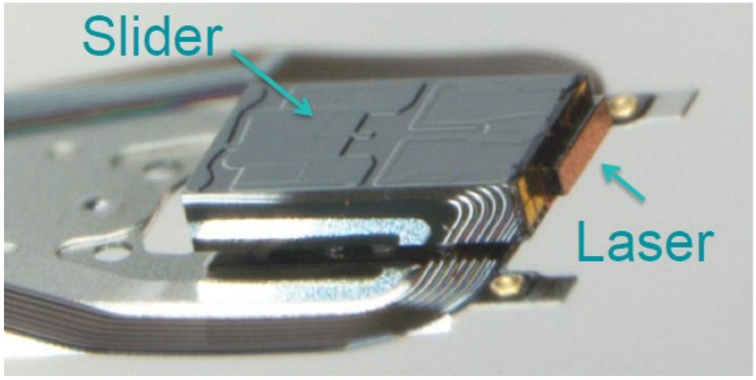


Figure 26. HAMR head with the laser source built into the head

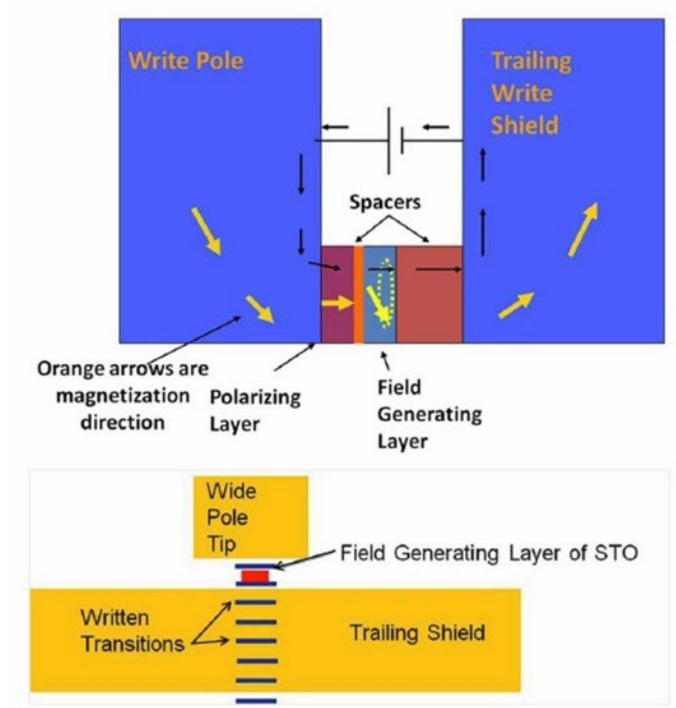
Source: Seagate Technologies

ENERGY ASSISTED PERPENDICULAR MAGNETIC RECORDING (EPMR)

ePMR is a technology currently being used by Western Digital to increase the capacity of their HDDs. ePMR increases BPI by applying an electrical current to the main pole of the write head throughout the write operation. This bias current enables more consistent, and faster switching of the write head, thus reducing timing jitter. Higher BPI is achieved when individual bits of data can be written closer together, which leads to higher areal density. Western Digital discovered this effect and is currently the prime user of this technology.

MICROWAVE ASSISTED MAGNETIC RECORDING (MAMR)

Microwave Assisted Magnetic Recording (MAMR) is an alternate Energy Assisted Magnetic Recording concept to extend AD. MAMR seeks to improve writability by supplementing the perpendicular write field with a localized RF field which is typically supplied by a small Spin Torque Oscillator (**Figure 27**).



Side view (top) of a spin torque oscillator in the write gap with electrical current (black arrows) flowing between the poles. ABS view (bottom) shows that the pole width is much greater than the STO width which determines the track width.

Figure 27. MAMR concept

Source: Mallory et al, IEEE Trans Magn, 50, 3001008, 2014

Modeling has suggested MAMR may extend AD to near 4 Tbpsi if STO requirements can be achieved and suitable media engineered. (Ref. J Zhu, IEEE Trans Mag 40, 3200809, 2014). To date very limited experimental AD gains have been publicly demonstrated. A version of MAMR (Flux Control-MAMR) is being used by Toshiba in some current high-capacity HDDs and the company has shown a path to using Microwave Assisted Switching MAMR (MAS-MAMR) for future HDD capacity gains. (**Figure 28**)

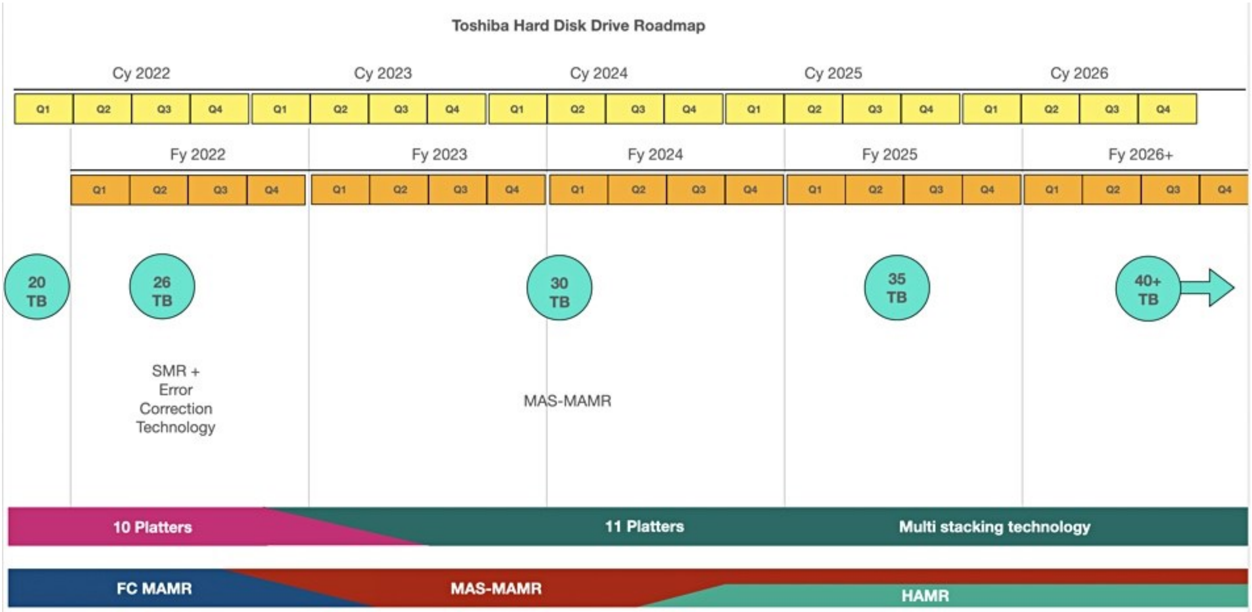


Figure 28. Toshiba MAMR/HAMR Roadmap derived by Chris Mellor in The Register Jan. 13, 2022

TWO DIMENSIONAL MAGNETIC RECORDING (TDMR)

Today's granular media is composed of grains with no directional information on the media. See Figure 29. Thus, the surface of the magnetic medium is a two-dimensional environment. Conventional magnetic recording defines which direction is along-track and which direction is cross-track.

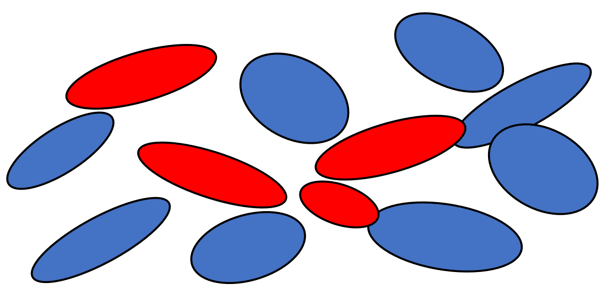


Figure 29. Media surface is a two dimensional environment

Source: Seagate Technology

The TDMR concept utilizes the two-dimensional magnetic surface for higher Areal Density (AD) and potentially better reading performance. Multiple generations of TDMR are envisioned. (Ref: TDMR Roadmap, F1, TMRC 2015) with progressively growing read head complexity to facilitate more sophisticated data encoding and density. Specifically, it is based on:

- Reading multiple adjacent tracks either by multiple read elements (See **Figure 30**) or multiple spins, and
- Processing those tracks jointly.

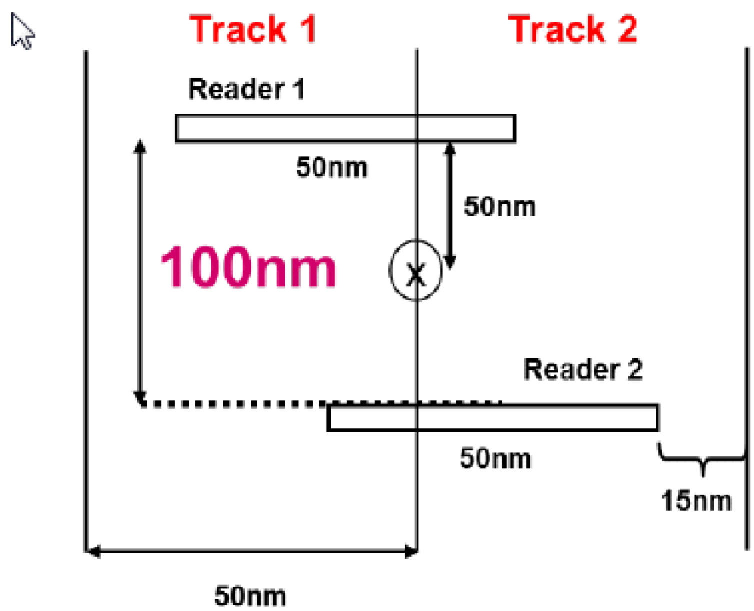


Figure 30. Dual reader TDMR concept.

Source: IDEMA ASRC

TDMR eliminates the direct relationship between the track pitch and read-head cross-track profile; it allows Inter-Track-Interference (ITI) between tracks which can be resolved by processing adjacent tracks jointly. This enables different Linear Density (LD) and Track Density (TD) points for the overall system to yield higher AD and possibly better read performance. Initial TDMR products were introduced by Seagate in 2017. The technology is currently used on many high-capacity HDD models.

HEATED DOT MEDIA

The combination of HAMR and patterned media (Heated Dot Magnetic Recording – HDMR) will be required to bring areal densities of magnetic recording to 10 Tb/in^2. Patterned media has discrete elements of the magnetic material distributed in an orderly fashion across the disk surface. These patterned dots become the magnetic bits when the head writes on them. Patterned media will be combined with HAMR recording to create 10 Tb/in^2 magnetic recording areal density within the next 15 years. **Figure 31** compares conventional magnetic recording media to patterned media.

In patterned media, the servo information is defined by lithography. While this would potentially save equipment costs for servo writers, the positioning accuracy of lithographically defined servo information needs improvement to be useful. These risks and in particular the perceived capital expenses for moving to patterned media have convinced the HDD industry to deal with patterned media as advanced research with introduction into volume manufacturing more than 10 years out. **Figure 32** shows some examples of patterned media including servo and address patterns.

In patterned media the nano-scale track and bit dimensions would be defined only by lithography. There are several technical candidates for this enhanced lithography.

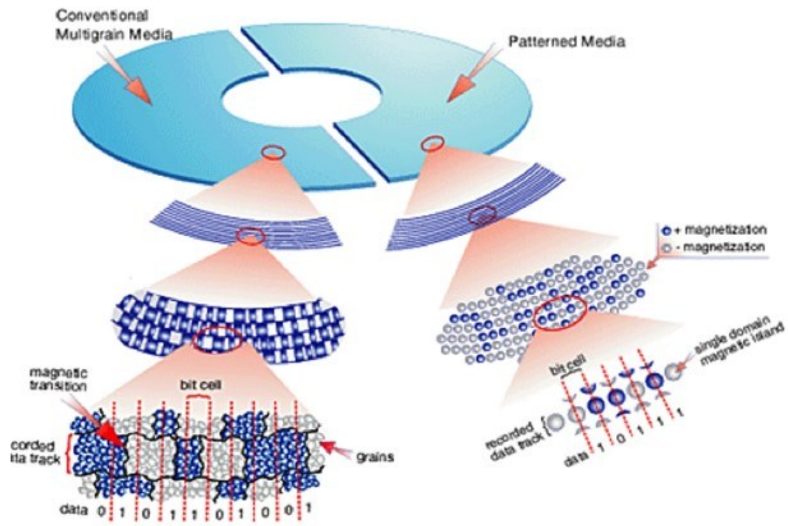


Figure 31. Comparison of conventional magnetic media to patterned media

Source: Hitachi Global Storage Technologies

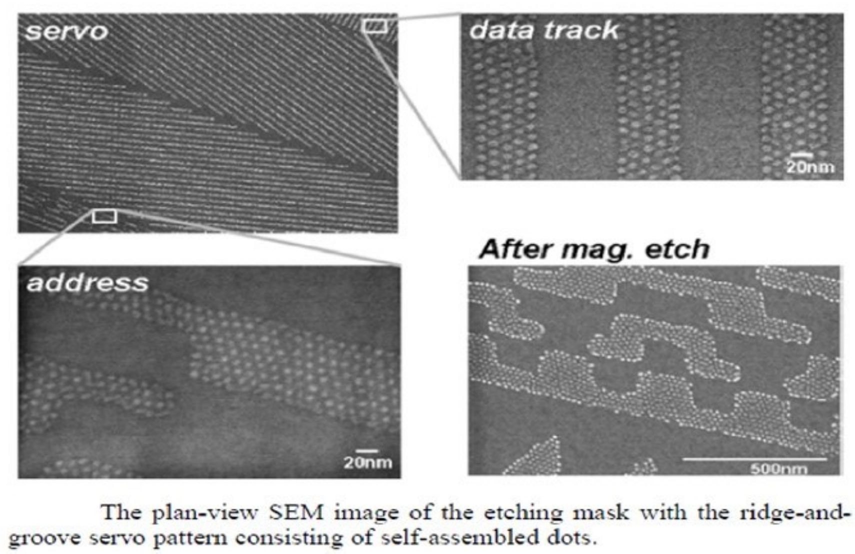


Figure 32. SEM images of self-assembled patterned magnetic recording media showing servo and address patterns[ii]

The likely approach for creating the physical media patterns is the use of block co-polymer self-assembly and pattern multiplication (i.e., SADP, SAQP, SAOP) to reduce the pitch such as used in semiconductor. The use of Di-Block Co-polymer (DBCP) self-assembly can be used to define a patterned media. This pattern must then be transferred to form the patterned media. This method has produced bit densities of 5Tbpsi with a pitch near 15nm. DBCP technology is the result of certain physical and chemical properties of unique polymeric substances.

[ii] Yoshiyuki Kamata, Akira Kikitsu, Naoko Kihara, Seiji Morita, Kaori Kimura, and Haruhiko Izumi, "Fabrication of Ridge-and-Groove Servo Pattern Consisting of Self-Assembled Dots for 2.5 Tb/in^2 Bit Patterned Media," *IEEE Trans. Mag*., **47**, 51-54 (2011)

These have the properties of self-assembly with highly consistent pitch which can be combined with a long-range guide pattern such as a simple, low-density e-beam generated pre-pattern. This minimizes e-beam lithographic write times since the patterns produced function as alignment marks to assist the ordering of the self-assembly process. In addition, the technique could be used to write pre-patterned imprint masters and consequently could reduce the costs of making e-beam masters.

Figure 33 indicates how block copolymers can be used to create these patterns. To increase SNR it is further desirable for the media bits to be rectangular rather than round. WDC and Seagate have overlaid printing of concentric rings with radial lines to produce rectangular bits. Recording demonstration to 1.6 Tbpsi have been demonstrated by Seagate.

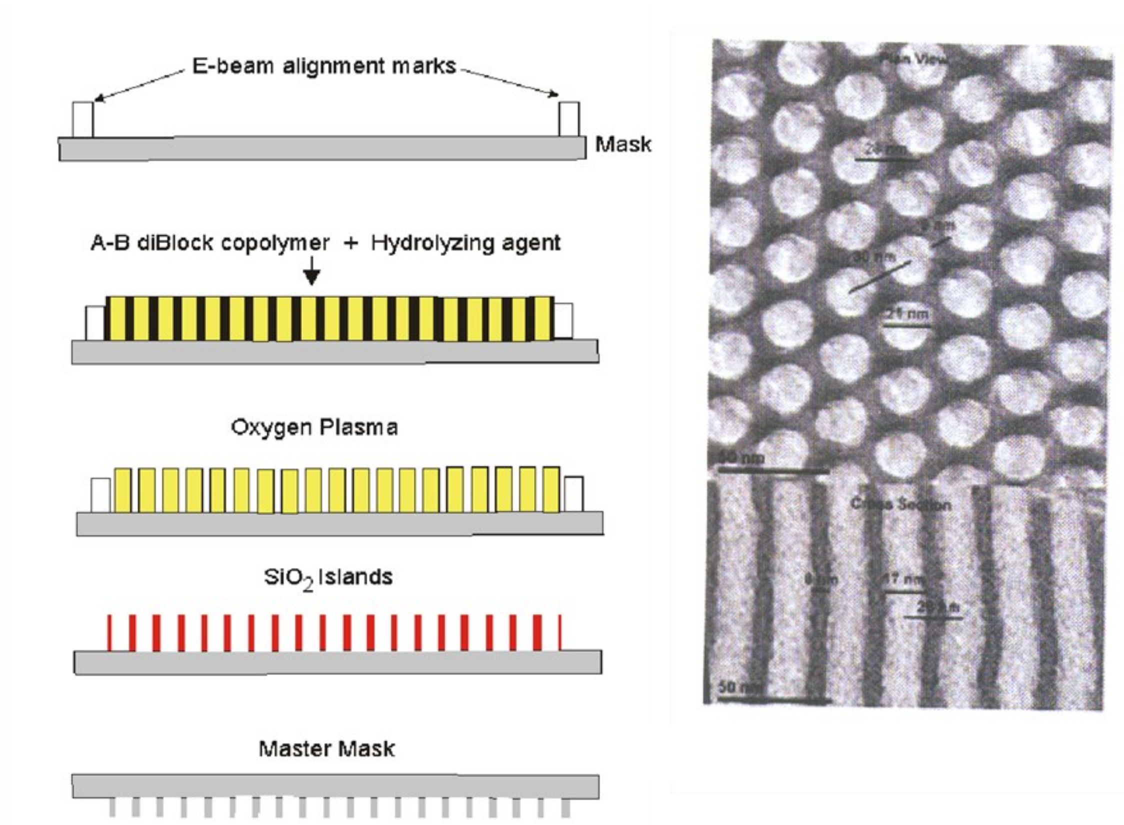


Figure 33. Block co-polymer self-assembly

Source: Photo Courtesy of IBM Almaden Research Center

The use of e-beam exposure, today's process for fabrication precision optical masks, can be considered the ultimate in lithographic processing, although as previously indicated today's e-beam tooling and resist formulations result in very low throughputs. It is likely that e-beam lithography will be used to create master patterns that can be used to make working imprint devices for creating imprinted patterns on magnetic surfaces. The process of creating working imprint surfaces from e-beam defined masters could involve either two or three generations of working stampers. One e-beam lithography generated master can be used to create several generations of imprinting surfaces, each of which may be capable of imprinting 100,000 disks.

This procedure is feasible since the e-beam technology required for nanoimprint master exposure could require many hours of production time, which can be amortized across a great many disks.

Patterned media requires magnetic bits to be located by lithography, whereas today's magnetic media accomplishes bit locations along a track through prewritten servo patterns and the actual write process from the magnetic head. Defects originating from lithographic errors in either exposure (from masks defects) or during the etching process include shorting of adjacent bits (bridging), missing bits, and diameter or positional variations. The latter would result in timing errors and reduced magnetic fields. These issues must be resolved to make this technology ready for production use.

Patterned media can be made by a variety of processes from "self-organizing" structures that could include lithographic and nanoimprinting techniques. **Figure 34** shows how patterned media may be manufactured.

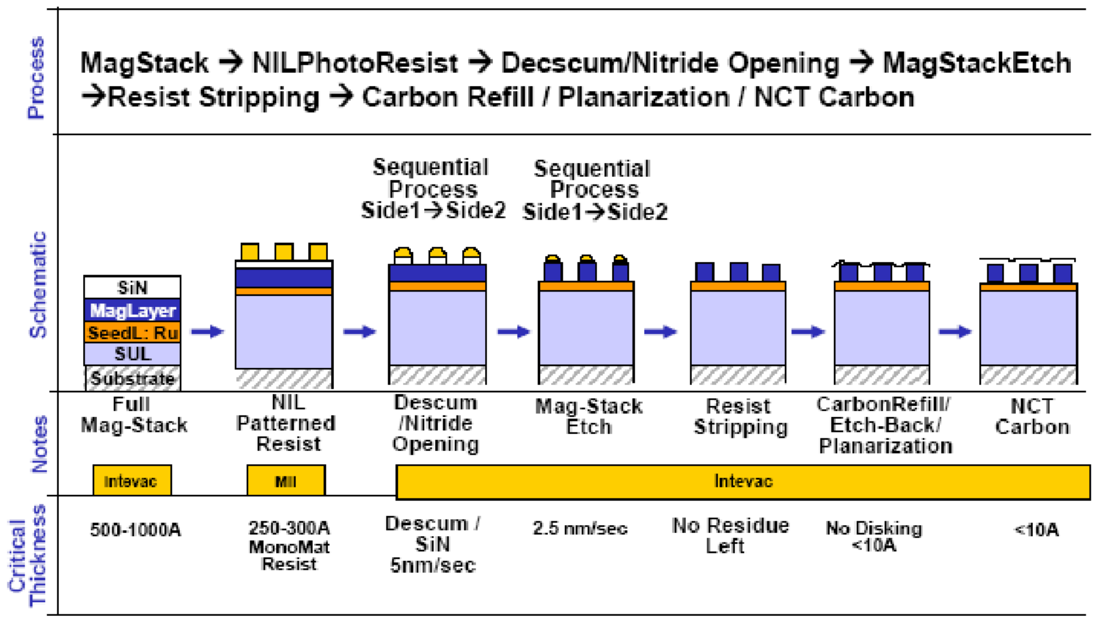


Figure 34. Possible patterned media production process

Source: Intevac

FULL DISK ENCRYPTING DRIVES

Full disk encrypting (FDE) drives encrypt everything that is written to the drive and decrypt everything that is read from the drive. So, encrypted data exists only while "at rest", while stored on the drive. The purpose of FDE is to provide the following:

- Protection from loss or theft of the drive (or the computer containing the drive): Since the data is encrypted, access by unauthorized users is not possible. All 50 states, the District of Columbia, Guam, Puerto Rico and the Virgin Islands have laws requiring private businesses, and in most states, governmental entities as well, to notify individuals of security breaches of information involving personally identifiable information, UNLESS the data is encrypted.

- Re-purposing or end-of-life: Under privileged (administrator) control, the cryptographic key can be erased, rendering the encrypted data inaccessible, thus "sanitizing" the drive, which can then be used as a re-purposed FDE drive (new key) or disposed.
- "Rapid erase": key erasure is nearly instantaneous, so that 'sanitization' does not take hours that traditional methods require, such as data overwriting.

The cryptographic function is implemented in dedicated hardware (circuitry) directly on the drive, so that system performance is not impacted. The cryptographic key never leaves the drive, eliminating one layer of key management required for software-based solutions. Instead, authentication keys are used to unlock the drive and invoke the FDE function.

Initially, FDE drives were available for portable devices, like laptops, due to the ever-increasing mobility of the workforce and the high frequency of laptop loss and theft. However, the data center is also subject to hard drive "mobility", not only misappropriation of the drives, but leaving the data center for maintenance and end-of-life. For this reason, FDE is now available on all new design enterprise and data center class drives. In addition, having the cryptographic function directly on the drive (instead of somewhere else "upstream" in the storage system), simplifies data center planning and configuration management, supports expansion with no performance impact, and eliminates the need for data classification.

FDE in hardware, directly on an HDD, has been standardized by the Storage Workgroup of the Trusted Computing Group (TCG). A subgroup of TCG has spelled out the details of (authentication) key management for enterprise drives and multiple drive configurations within an enterprise. The TCG Storage Specification defines an architecture for a variety of security functions built directly into storage devices, including FDE. (download available at: https://www.trustedcomputinggroup.org)

The Institute of Electrical and Electronic Engineers promulgated a comprehensive data sanitization standard covering data security in 2022, IEEE 2883-2022, that recognizes FDE as cryptographic erase, a medium security methodology.

Magnetic Tape Storage

SITUATION ANALYSIS

Linear tape technology uses the same basic magnetic recording principles as hard disk drives (HDDs) and leverages many of the technologies developed by the higher volume and more advanced HDD industry. For example, linear tape drives first adopted AMR, then GMR and most recently TMR reader technologies that were first developed for use in hard drives up to 10 years before being used in tape drives. The latest enterprise class tape drive, the IBM[i] TS1170, which was released in Aug. 2023, operates with a native cartridge capacity of 50TB. Although state of the art tape drive bit length is only about 3.5x larger than that of the highest capacity disk drives, tape drive track width is about 12x larger than disk and hence tape areal density is about 42x lower than HDD.

However, the latest generation of tape cartridges contain more than 1000m of 12.65mm wide tape, e.g., the TS1170 user tape length is more than 1200m, resulting in a useable recording surface area that is about 100 times larger than that available in a state-of-the-art high-capacity HDD with ten platters that are > 95mm in diameter. Hence the capacities of the latest generation enterprise tape format is roughly double the highest capacity HDDs on the market at the time of its release (i.e. 22TB CMR and 26TB SMR). The fact that tape systems operate at a roughly 42x lower areal density than state of the art HDD implies that tape systems have the potential to continue scaling areal density and capacity for multiple future generations before having to face the challenges resulting from the super paramagnetic effect (i.e., the magnetic recording trilemma) that the HDD industry is currently struggling with.

Tape drive developers have adopted a variety of technologies originally developed for disk drives. These include head technologies and data channel algorithms such as NPML (Noise Predictive Maximum Likelihood detection). The challenge for tape drive developers is to adapt these technologies for use in tape drives and to develop additional technologies to deal with the unique challenges that arise from parallel recording on a flexible tape media that runs in contact with the head. Rather than a single active recording channel as in a disk drive, state of the art tape drives record and simultaneously read verify 32 tracks in parallel.

In recent generations of IBM Enterprise and LTO[ii] drives, this is achieved with a three-module head design composed of a read module with 32 data readers and two servo readers sandwiched between two data writer modules which each contain 32 data writers and two servo readers, as illustrated in **Figure 35**. In this architecture, the left writer module is used for writing data in the forward tape direction and the right writer module is used for writing in the reverse direction and the center reader module is used for data read back and read-while-write verification in both directions. In addition to the reliability provided by read-while-write verification, tape drives also implement powerful error correction codes (ECC) to achieve bit error rates that are four orders of magnitude better than disk products.

---

[i] IBM is a trademark of International Business Machines Corporation registered in many jurisdictions worldwide.
[ii] LTO is the trademark of HP, IBM, and Quantum in the Unites States and other countries.

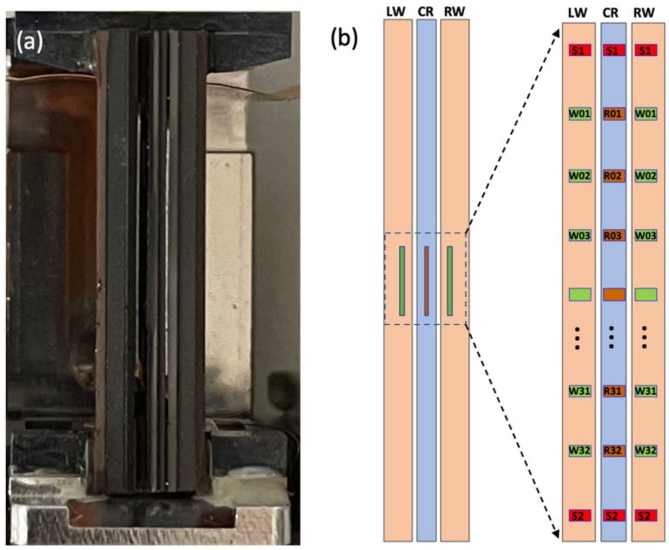


Figure 35. (a) Photograph of a 3 module, 32 channel tape head, (b) Illustration of a 3 module 32 channel tape head. The center reader module (CR) contains 32 data readers (R01-R32) and 2 servo readers (S1 and S2), and the left (LW) and right writer (RW) modules each contain 32 writers (W01-W32) and two servo readers (S1 and S2). The read transducers in the reader module are aligned with the write transducers in the writer modules to enable read while write verification.

The multi-channel nature of modern tape drives necessitates the design of a set of custom ASICS for driving all the parallel channels, i.e., driving the writers, pre-amplification and filtering for the readers and to implement the digital data and servo channels as well as the ECC algorithms. Compared to HDD, the single channel data rate of a tape drive is much lower, which simplifies some aspects of the ASIC designs, however support for many parallel channels creates additional complexity.

Moreover, the removeable nature of tape media and requirements for interchangeability and backwards compatibility as well as the use of contact recording in which debris can temporarily be deposited on the head, all necessitate that the data and servo channels of tape drives are designed to be much more adaptable than those implemented in HDD. In addition, the support of a wide range of tape speeds required for the variable data rates provided by tape drives in order to match the host’s transfer rate imposes additional ASIC design challenges.

The continued scaling of ASIC technology to smaller transistor sizes has enabled tape drive designers to develop new ASICs with the capability to process the increasing number of parallel write and read transducers in the head without significant increases in the ASIC die size. Compared to HDD, the number of tape drives shipped per year is orders of magnitude lower and hence the number of ASICS required is also much lower. This challenge of low ASIC volumes is compounded by the rising costs of designing custom ASICs as the technology nodes scale to smaller transistor sizes and more expensive lithography and mask technology. The tape industry therefore amortizes the cost of designing and manufacturing these ASICs by reusing ASICs in multiple generations and across different product families.

The most recent enterprise and LTO tape formats, IBM TS1170 and LTO-9, both support the built-in file system called the Linear Tape File System (LTFS) that was first introduced with the LTO-5 tape format. LTFS expands the capabilities of tape drives and tape library systems to allow Network Attached Storage (NAS) like single tape behavior and in library systems allows faster access time to data. RESTful APIs using LTFS tapes have also enabled object-based storage built around the use of magnetic tape.

BUSINESS/TECHNICAL ISSUES

Historically, tape capacity has been scaled primarily by areal density scaling with additional capacity gains achieved through improvements in format efficiency and reducing the tape thickness to increase tape length. As with HDD, tape media advances are key to continued tape scaling. As areal density is increased, improvements in media are required to compensate for the loss in signal and signal to noise ratio (SNR) resulting from the smaller bit size and reductions in the thickness of the magnetic recording layer that help to enable increases in linear density.

Tape media recording performance improvements are typically achieved through a combination of technologies that include reducing the size of the magnetic particles in the recording layer to reduce media noise, developing new particles with improved magnetic properties, improving the dispersion of the magnetic particles, orienting the particles during the coating process, reducing variations in the thickness of the recording layer and reducing the media roughness in order to improve magnetic spacing.

Unfortunately, making the media smoother tends to increase tape-head friction which can degrade recording performance and reduce the ‘runability’ of the tape. These effects can be minimized through careful engineering of the tape surface roughness, improvements in lubrication technology and optimization of the geometry and topography of the tape head to minimize contract area. Additional tape path components and features will also likely be adopted to deal with increases in startup friction/stiction and tribology issues during tape transport that arise from smoother media.

In state-of-the-art linear tape drives, 32 equally spaced data tracks are written and read back in parallel in one of four data bands that are about 2.7 mm wide as illustrated in **Figure 36**. The written track width is much smaller than the spacing between adjacent writers and hence many passes back and forth along the length of tape are required to fill a data band. The roughly 2.6 mm span of the array of transducers in combination with tape dimensional stability (TDS), i.e., changes in width of the thin flexible tape, makes it challenging to keep all the writers aligned with the desired track locations during write operations and all the readers centered over the written tracks during readback.

The dimensional stability of the tape depends on thermal expansion, hygroscopic expansion, operating tension and stress induced creep during storage in the cartridge reel. A user can record a tape at one extreme of environmental conditions and will expect to be able to recover the data at a different environmental extreme, perhaps years later. Historically, TDS was dealt with by including a component in the track width budget to account for the changes in tape width due to TDS and therefore ensure that the readers don’t get too close to the track edges during readback.

The other major component of the tracking budget accounts for track following errors, i.e., errors in positioning the head relative to the moving tape during read and write operations.

Reductions in track width with each generation of tape drives have therefore necessitated improvements in TDS and improvements in track following accuracy to ensure reliable operation. Recently, tape drives have implemented active TDS compensation to ensure that all 32 read and write transducers are placed at the correct location on tape in the presence of changes in tape width due to TDS effects. This innovation relaxes somewhat the need for continual improvements in tape dimensional stability which were becoming increasingly difficult to achieve. Hence, continued track density scaling will require continued incremental improvements in active TDS compensation and improvements in track following fidelity.

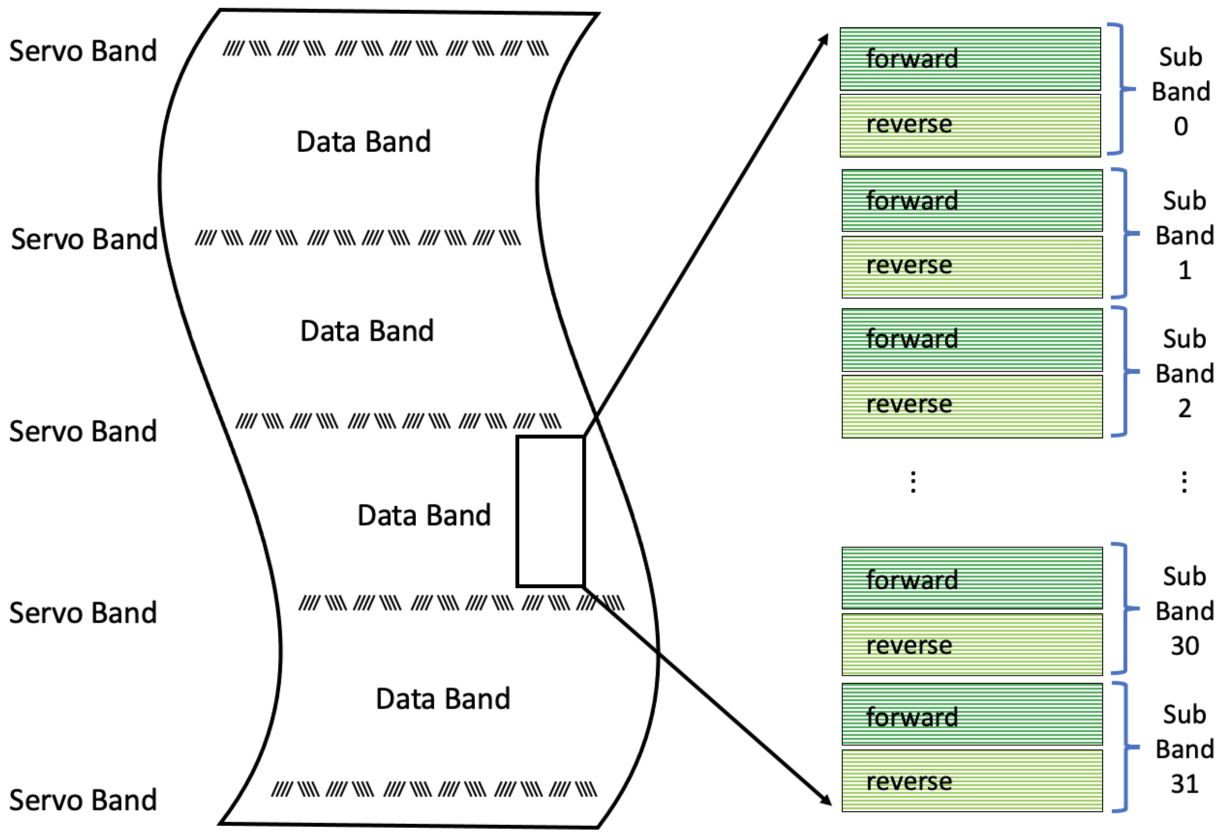


Figure 36. Illustration of the four data band / five servo band tape layout used in the LTO format. LTO generations 7 to 9 use a 32-channel format illustrated on the right side of the figure in which each data band is subdivided into 32 sub-data bands. 32 tracks are written in parallel with forward wraps (tracks) written in the upper part of each sub-data band and reverse wraps written in the lower part. Multiple passes back and forth along the length of tape are required to fill each data band.

The current tape market is dominated by the Linear Tape Open (LTO) format with a smaller share held by the IBM TS11xx enterprise format. Current LTO media manufacturers include Sony as well as Fujifilm who also manufactures media for IBM's TS11xx tape drives. HPE, IBM and

Quantum all provide TPC (Technology Provider Companies) certified LTO drives. Multiple companies including Dell, Fujitsu, HPE, IBM, Quantum, Spectra Logic, Tandberg, etc., offer tape library solutions.

Tape systems provide a very low total cost of ownership for storing large volumes of data. For example, a recent 10-year TCO (total cost of ownership) study from ESG found an 86% cost reduction for an LTO8 tape solution over an HDD based solution [1]. In addition, tape systems can achieve very high streaming data rates but with the penalty of access latencies in the range of tens of seconds. Hence, tape systems are particularly well suited for archival data storage applications, i.e., the long-term preservation of infrequently accessed data. The built-in physical air gap provided by the removeable nature of tape cartridges provides an additional layer of protection against cyber-crime or accidental data deletion.

The removeable nature of tape also results in a very low power consumption per PB of capacity that contributes to both the low TCO and the very low CO2e footprint of tape systems. A recent IBM study compared the 10-year CO2e footprint of a 27 PB tape archive to an open compute project (OCP) Bryce Canyon HDD solution and found a 96% lower CO2e for the tape solution and more than 90% less power consumption. The combination of these benefits has been driving growth in the tape market with particularly strong growth amongst hyperscale cloud companies. In 2021 the overall tape market grew by about 14%, with approximately 2/3 of that growth driven by the Hyperscale market segment [2]. The tape market is projected to continue growing with a projected growth of 64% by 2025.

Figure 37 shows an LTO-9 (Linear Tape Open Generation 9) full-height tape drive and cartridge. Tape drives and cartridges are often used in automated library systems. **Figure 38 (a)** and **(b)** shows the IBM and the Spectra Logic Tape Library offerings, respectively. Not all of the library types shown in Figure 38 are designed for use with enterprise drives and at the time of writing, only the TS4500 provides support for the very recently introduced TS1170 drive. Therefore, to facilitate comparisons between different library types, the maximum library capacities shown in the figure are calculated assuming LTO9 technology.

Assuming the use of TS1170 technology instead would result in much higher maximum native capacity, e.g. for the TS4500 library the maximum capacity would be 877 PB instead of 417 PB with LTO9 technology. (Note that the max number of LTO slots in the TS4500 library is larger than the max number of enterprise slots due to the slightly smaller size of LTO cartridges.)


Figure 37. LTO-9 FH Tape Drive and Cartridge

Although somewhat less form factor driven than disk, tape drive offerings have now settled into the Full Height and Half Height form factors. The very low-end of the tape market that was served by the 4mm and 8mm tape formats shrunk drastically due to competition from small form-factor removable disks and the advent of flash memory alternatives, in the late '90's and early 2000's. No new low end tape drives have been introduced on the market for many years.

The entry level of the mid-range tape market is served the LTO HH drive which can be installed in a server, connected to a server via a 'stand-alone box' or installed in an 'auto-changer' such as the TS2900 shown in **Figure 38 (a).** In the middle markets and enterprise sectors, automated tape libraries have become commonplace and range in capacity from 10s to 1,000s of tapes with maximum native capacity of a single library now exceeding an exabyte (1 EB = 10^{18} bytes).

The wide-spread adoption of tape by hyperscale cloud companies has motivated the tape industry to design solutions to specifically address the unique requirements of this market segment such as modular designs for ease and speed of deployment as well as optimal reliability in erasure coded environments. The Diamondback library recently introduced by IBM is one example of this [3]. The Diamondback library can be shipped with media and drives pre-installed and can be installed in the datacenter in less than 30 minutes. The library is designed for easy self-service, with most major components replaceable in two minutes or less. It fits in the same floor space as a standard open compute project (OCP) rack and provides up to 69.5PB compressed capacity using LTO9 technology in a single 8 ft^2 (0.7 m^2) footprint.

Another recent trend in tape storage is the enablement of object storage to tape. For example, Spectra Logic offers a tape-based object store with an S3 interface that uses their BlackPearl gateway. Another example is ActiveScale Cold Storage from quantum that provides an S3 interface to both HDD and Tape based storage with a common name space. Point Archival Gateway is an example of a software solution that enables a tape-based object store with a standardized S3 interface and supports tape libraries from multiple vendors as a backend.

This solution also enables erasure coding of objects across multiple tape drives (RAIT) or across multiple tape libraries (RAIL), an architecture adopted by many hyperscale cloud tape users [4]. Fujifilm also recently announced a software product called Fujifilm Object Archive that provides an S3 compatible interface to enable object storage with a tape library back end [5]. Objects are written to tape using an open format developed by Fujifilm called OTFormat [6]. Even more recently, IBM announced Diamondback S3, an S3 based object storage solution that uses the IBM Diamondback library. There are also open-source initiatives to enable object storage using a tape back-end such as Open Stack Swift HLM (high latency media), that supports both tape and optical disc backends [7].

	TS2900	TS4300	Diamondback	TS4500
Max Capacity* (native)	162 TB	5.04 PB	27.8 PB	417 PB 877 PB **
Max Capacity (compressed)	405 TB	12.6 PB	67.5 PB	1.04 EB
Max # of drives (H/FH half/full height)	1	21 HH 7 FH	14 FH	128 FH
Max # of slots	9	280	1584	23170
# frames at max capacity	n/a	1	1	18
Capacity/floor space PB/ft² (PB/m²)	n/a	1.2 (12.9)	3.51 (37.82)	2.27 (24.44)

*assumes LTO-9; usable capacity may vary depending on number of slots reserved for system operations
** assumes TS1170; usable capacity may vary depending on number of slots reserved for system operations

(a) IBM Tape Libraries

	Stack	T380	T680	T950	Tfinity ExaScale
Max Capacity* (native)	10 PB	6.8 PB	12 PB	~180 PB	1.01 EB
Max Capacity (compressed)	25 PB	17 PB	30 PB	450 PB	2.53 EB
Max # of drives (H/FH half/full height)	42 HH 21 FH	12 FH	12 FH	120 FH	144 FH
Max # of slots	560	380	670	10020	56400
# frames for Max Capacity	1	1	1	8	45
Capacity/floor space PB/ft² (PB/m²)	2.19 (23.6)	1.37 (14.72)	1.5 (16.1)	2.58 (27.8)	2.58 (27.65)

* assumes LTO-9

(b) Spectra Logic Tape Libraries

Figure 38. (a) IBM Tape Libraries and (b) Spectra Logic Tape Libraries

ROADMAP OF QUANTIFIED KEY ATTRIBUTE NEEDS

In 2021, the LTO tape consortium announced the latest generation of LTO magnetic tape, LTO-9. This tape format supports 18 TB of native storage capacity and advertises a 45 TB cartridge capacity with 2.5:1 compression. The maximum native drive data rate is 400 MB/s and up to 1000 MB/s compressed, assuming 2.5:1 compression. Like LTO-8, the LTO-9 cartridge format uses Barium Ferrite (BaFe) particle-based media and tunneling magneto-resistive read sensors. LTO-9 drives can read and write LTO-8 cartridges and support widely used encryption standards.

LTO-9 tape drives operate at an areal density of about 12 Gb/in^2 (giga-bits per square inch). On Aug 23, 2023, IBM announced the TS1170 Enterprise tape drive with 50 TB native capacity, 400 MB/s native transfer rate and areal density of about 26 Gb/in^2 using Strontium Ferrite (SrFe) based media. In contrast to LTO9 and previous enterprise drives, the TS1170 does not provide any backwards compatibility.

The tape industry began transitioning from the previously used metal particle (MP) technology to hexagonal platelet shaped BaFe particles around the time frame of LTO-6, with LTO-6 supporting both MP and BaFe media technologies and LTO-7 media based exclusively on BaFe. The scaling potential of MP technology was limited by the need for a thin glass coating to prevent oxidation of the particles. This limited the minimum particle size of MP technology to about 3000 nm^3, beyond which it is difficult to maintain particle coercivity.

In contrast to MP, both BaFe and SrFe are already oxides and hence do not require a coating and can therefore be scaled to smaller particle sizes. Both LTO-9 and the IBM TS1160 media use BaFe particles with a partial perpendicular orientation in the magnetic recording layer of the tape. The TS1170 media uses SrFe particles and also has a partial perpendicular orientation in the magnetic recording layer. SrFe is from the same family of hexagonal ferrous oxides as BaFe, but has a higher saturation magnetization and higher coercivity and hence can be scaled to smaller particle sizes. LTO tape uses polyester-based substrates whereas the both the TS1160 and TS1170 media use a more stable but also more expensive aramid-based substrate. The coercivity of both BaFe and SrFe particles can be tuned over a wide range using substitution elements, similar to doping a semiconductor.

Hence there is potential to continue reducing the size of BaFe and SrFe particles to enable higher areal densities while increasing the particle coercivity to maintain thermal stability of the recorded data. Increasing the coercivity of future tape media will necessitate the use of tape write heads that produce stronger magnetic fields. This is an area in which tape developers can take advantage of materials and technologies developed for HDD that currently use media with much smaller magnetic grains and higher coercivity than tape media.

Table 4 summarizes a set of major tape industry metrics and their expected evolution over the next 15 years. **Figure 39** presents the latest LTO Consortium roadmap which describes the five currently planned future LTO generations [8]. **Figure 40** shows the areal density scaling projections of the 2019 INSIC roadmap for magnetic tape along with historical scaling data for HDD and tape as well as areal density demonstrations [9]. The LTO roadmap does not provide a timeline, however, recent LTO generations have been released roughly every 2.5 years. **Table 4** projects an areal density of 28.3% CAGR (compound annual growth rate) out to 2031 that will enable a 32% CAGR in capacity scaling. Note that a 32% CAGR corresponds to a doubling in capacity every 2.5 years, corresponding to the recent trend in time between LTO generations. Beyond 2031, tape areal density scaling is projected to continue scaling with a somewhat smaller ~27% CAGR and capacity with a 30% CAGR.

Table 4. Magnetic Mass Data Storage Technology Roadmap – Tape

	Unit	2021	2023	2025	2027	2029	2031	2033	2035	2037
Native Capacity	TB	20	35	61	106	184	321	543	917	1550
Max Data Rate (native)	MB/s	400	529	700	925	1224	1618	2140	2830	3743
Areal Density	Gb/in^2	12.0	19.7	32.3	53.3	88.0	145.4	233.3	374.8	602.6
Tape Speed (for data)	m/s	5.7	6.9	4.1	4.9	5.9	7.1	4.3	5.1	6.1
Volumetric Density	TB/in^3	1.5	2.7	4.6	8.1	14.1	24.5	41.4	69.9	118.2
Form Factor*	n.a.	HH, FH	HH, FH	HH*, FH	HH*, FH	HH*, FH	HH*, FH	HH*, FH	HH*, FH	HH*, FH
Key Requirements										
Read Head	type	TMR	TMR	TMR	TMR	TMR	TMR	TMR cpp-GMR	TMR cpp-GMR	TMR cpp-GMR
Write Head	type	High B_s	High B_s	High B_s	High B_s Shielded	High B_s Shielded	High B_s Shielded Monopole	High B_s Shielded Monopole	High B_s Shielded Monopole	High B_s Shielded Monopole
Number of channels	n.a.	32	32	64	64	64	64	128	128	128
Magnetic Film	n.a.	BaFe	BaFe SrFe	BaFe SrFe ε-FeO	BaFe SrFe ε-FeO	SrFe ε-FeO Sputtered	SrFe ε-FeO Sputtered	SrFe ε-FeO Sputtered	SrFe ε-FeO Sputtered	SrFe ε-FeO Sputtered
Recording Tech.		Perp.	Perp.	Enhanced Perp.	Enhanced Perp.	Enhanced Perp.	P+SUL	P+SUL	P+SUL EAMR	P+SUL EAMR
Tape Thickness	μm	5.0	4.8	4.6	4.4	4.25	4.1	3.9	3.8	3.6
Substrate Material	type	PEN Aramid Adv. Sub.	PEN Aramid Adv. Sub.	PEN Aramid Adv. Sub.	PEN Aramid Adv. Sub.	PEN Aramid Adv. Sub.	Aramid Adv. Sub.	Aramid Adv. Sub.	Aramid Adv. Sub.	Aramid Adv. Sub.

Notes: Volumetric Density: Capacity divided by volume of an LTO cartridge (4 x 4.1 x 0.8 = 13.12 in^3)

Form Factor: HH = half height = 5.25" half-height internal form factor
FH = half height = 5.25" full-height internal form factor
*Note, due to power density and space constraints HH drives will likely remain in a 32-channel format with lower data rate when FH drives transition to a 64-channel format.

Read Head: TMR = tunneling magneto-resistive
cpp-GMR = current perpendicular to the plane giant magneto-resistive

Magnetic Film: BaFe = barium ferrite
SrFe = strontium ferrite
ε-FeO = Epsilon iron oxide

Recording Technology: Perp. = perpendicularly oriented media,
Enhanced Perp. = perpendicular media with improved orientation
P+SUL = perpendicular media with a soft underlayer
EAMR = energy assisted magnetic recording

Substrate Material: PEN = polyethylene napthalate,
Aramid = Aromatic Polyamid
Adv. Sub. = Advanced Substrate

It is interesting to note that the areal density projected in the last entry of **Table 4** for the year 2037 is still about 50% lower than the areal density of current HDD. This provides a level of confidence that the roadmap projections are attainable. In addition, several research papers demonstrating the future scaling potential of tape technology have been published in the last few years that provide further evidence that the projections for the next decade are attainable. For example, in 2017, IBM in collaboration with Sony, reported a single channel tape recording demonstration of 201 Gb/in^2 using a sputtered tape based on a CoPtCr-SiO2 perpendicularly oriented recording layer [10].

This areal density corresponds to a potential cartridge capacity of 330 TB assuming similar formatting overheads as an IBM TS1155 drive and considering the increased tape length enabled by the reduced thickness of the demo media. In 2020, IBM in collaboration with Fujifilm, reported a single channel tape demonstration of 317 Gb/in^2 using a perpendicularly oriented SrFe particle-based recording layer [11].

This areal density corresponds to a potential cartridge capacity of 580TB for the 4.3 µm thickness of the demo media and assuming similar formatting overheads as a TS1160 drive. Even more recently, in 2022, Western Digital studied the recording performance of a sputtered magnetic tape developed by Sony Media Solutions Corporation that used a sputtered CoPtCr-SiO2 granular recording layer with a thin CoPtCrB capping layer and reported an areal density of 400 Gb/in^2 [12].

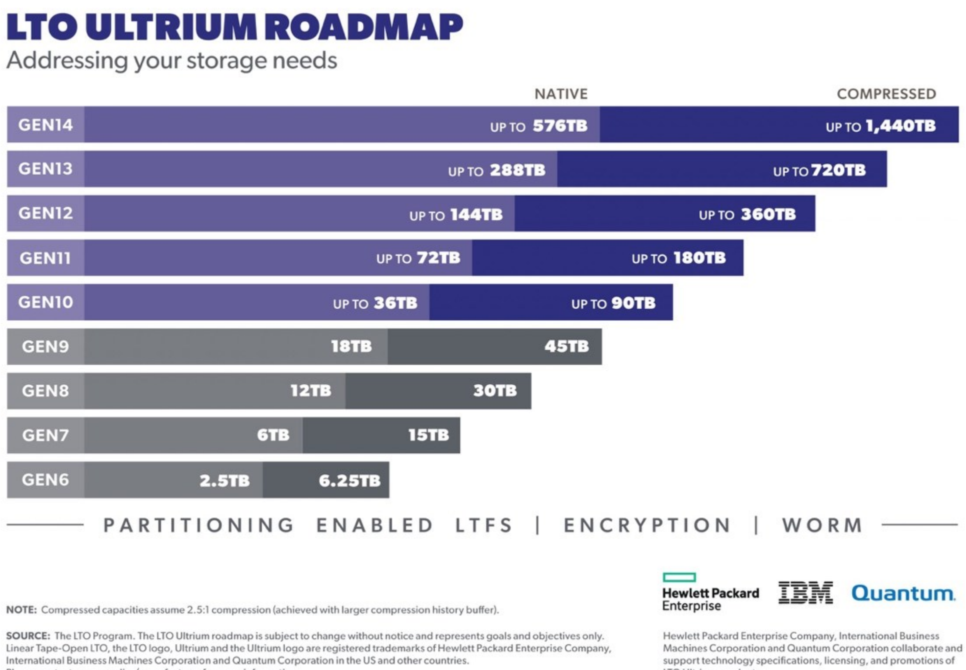


Figure 39. LTO Consortium Roadmap [8]

Source: The LTO Program. Hewlett Packard Enterprise, International Business Machines Corp., and Quantum Corporation

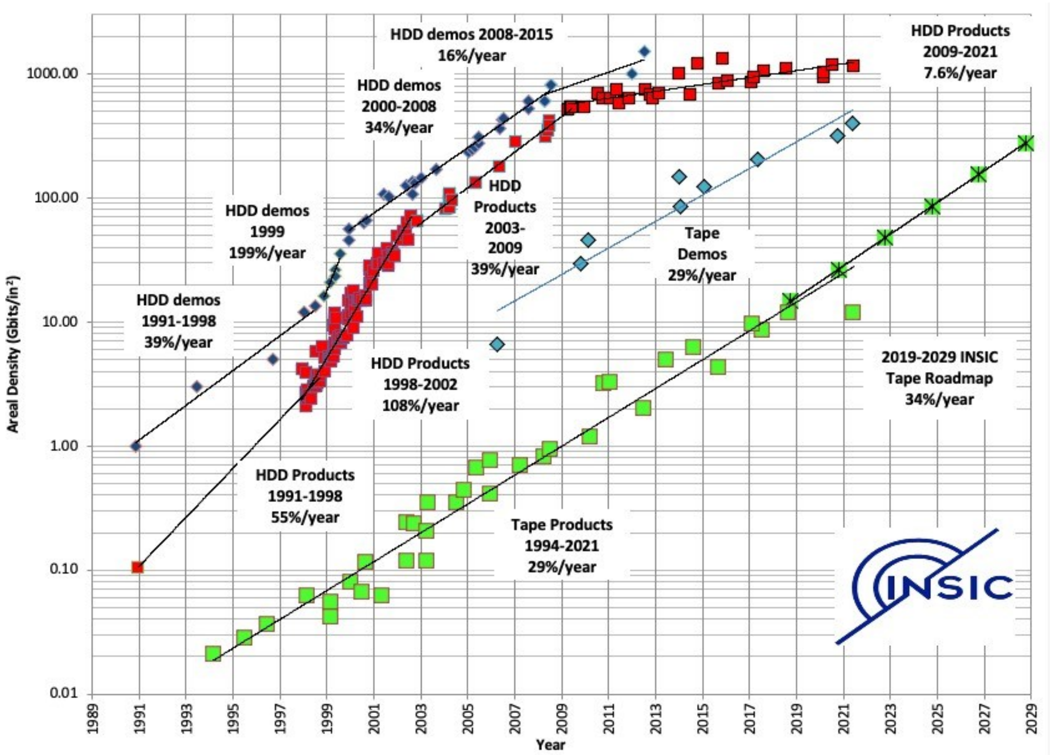


Figure 40. INSIC Tape Areal Density Roadmap and Areal Density of Tape and HDD Demonstrations and Products [9].

Source: INSIC

CRITICAL ISSUES

Magnetic tape storage will continue to benefit from adapting the head and media technologies developed by the HDD industry and will have to continue developing advanced servo and signal processing solutions for the unique challenges that arise in tape systems. In addition, systems level considerations (performance, error rate, archiving, reliability) will continue to have a major influence on the design and implementation of tape storage solutions. The rapid growth in adoption of tape technology by hyperscale cloud companies who buy large quantities of tape but have a unique set of requirements compared to traditional users of tape will likely influence future product roadmaps and design tradeoffs between rates of capacity and data rate scaling.

TECHNOLOGY NEEDS: RESEARCH, DEVELOPMENT, IMPLEMENTATION

A critical parameter in continuing to increase areal density in tape drives is the media signal to noise ratio (SNR). In linear tape recording, the media is the dominant source of noise. The noise follows particle-counting statistics and is therefore sensitive to the reader width. Modulation noise and defects also contribute to the noise environment and are significant for sampled amplitude detection channels such as PRML.

The most effective way to decrease particle noise is to reduce the particle size. However, as the particles get smaller the capability to maintain coercivity and remanence and the capability to disperse the particles uniformly becomes more difficult. This will continue to be crucial to advances in tape recording and will consume significant development resources.

Another critical area related to the media is the need to continue to reduce the head-media spacing to support higher linear densities. Tape is a contact recording technology in which the flexible tape is run in contact with the tape head and the spacing between the read/write transducers and the magnetic particles is determined primarily by the surface roughness of the media. Making the media smoother reduces spacing but at the same time tends to increase tape-head friction which can degrade recording and servo performance and increase tape and head wear.

To control friction, non-magnetic spacer particles are introduced to reduce the contact area between the tape and head and hence reduce friction. These particles are also somewhat abrasive and help to clean debris that might otherwise build up on the tape bearing surface of the head. To protect the sensitive TMR read transducers, the read elements are coated with a thin wear resistant coating that also contributes to the spacing.

To achieve the areal densities projected in the later phases of the roadmap will require the development of ultra-thin wear coatings and very smooth media such that head-tape spacing can approach that used in current HDD products. In addition, the development of low friction head technologies that optimize the geometry and topography of the tape bearing surface will be required to enable the use of such very smooth media. An alternative strategy to address the spacing challenge is to move away from contract recording and adapt the air bearing technology developed for HDD to tape recording.

Increasing the track density in linear tape recording will require improvements to track following servo systems. This includes the process that originally records the servo information on the tape and the servo system’s ability to follow the tracks on a flexible media. The potential for the required improvements in tape track following performance have already been shown in the previously mentioned IBM tape areal density demonstrations that included a track following component.

In the most recent demonstration, IBM in collaboration with Fujifilm demonstrated a track following accuracy characterized by the standard deviation of the track following error of 3.2 nm or less over a tape speed range of 1-4m/s. Here the main challenge will be to implement and adapt the lab technologies used to achieve this performance in commercial tape drives and media manufactured at scale.

The requirement to increase cartridge capacity will lead to thinner tape, which increases the difficulty of handling and guiding the tape. The introduction of flangeless tape paths significantly reduced the potential for tape edge damage; however, thinner media will necessitate the need for improvements in tape path mechanics and tension control during tape transport.

The dimensional stability of the tape is a large part of the off-track budget for a multi-channel tape system. The recent introduction of active tape dimensional stability compensation has relaxed the need for continual improvement in the dimensional stability of tape substrate materials, however continual improvements in the accuracy of the compensation mechanism will be required to enable continued track density scaling. While not trivial, active TDS compensation shares many of the same challenges, uses the same position measurement signals and hence benefits from advances in track following technology. The nanometer scale accuracy that has been demonstrated for tape track following control provides some confidence that similar levels of accuracy are attainable for TDS compensation.

Beyond automating backup and archiving applications, tape libraries have enabled the use of tape in hierarchical storage management (HSM) applications for improving performance with the migration and recall of data sets in a variety of storage environments. Data storage management hardware and software, including virtual tape, has enabled increased utilization of the tape cartridge while improving the performance and retrieval times for the tape environment by combining faster access storage such as HDDs or flash memory to the tape storage system.

Archival tape storage has evolved in a similar direction, with the development of automated high-performance tape storage systems having native capacities larger than one exabyte (EB) and achieving data center floor space densities of up to 4.77 PB/ft^2 (TS4500 library with TS1170 tape technology). Magnetic tape continues to offer the highest probability of success in achieving the maximum total volumetric density with acceptable data rate combined with the lowest price per terabyte of any storage technology. The improvement in tape usage, drive and media reliability has made tape acceptable in a wide variety of high-capacity, moderate to low activity applications. Fixed content, archive, compliance and disaster recovery applications are the primary uses of magnetic tape today as the traditional backup and restore market has largely moved to magnetic disk-based architectures.

Linear tape cartridges have sustained about 30-40% annual growth rate in storage capacity over the last decade through the combination of breakthroughs in four areas. These are the incorporation of magneto-resistive, giant magneto-resistive and tunneling magneto-resistive heads, advances in track following servo technology, advanced in ECC and data channels as well as advances in media technology. Several years ago, several tape cartridges were needed to back up a single disk drive; today a single tape cartridge can often backup multiple disk drives.

Today tape remains the greenest, i.e. the most energy efficient and lowest CO2e, of all mass storage technologies. These attributes are expected to increase the economic appeal of tape systems in the coming years.

GAPS AND SHOWSTOPPERS

Continuing to achieve the required head tracking precision and the continued advancement of media are two key challenges for tape's continued progress. Progress in the ease of use and integrations of tape into hierarchical storage systems combined with continued areal density growth during the recent slow-down in areal density scaling of HDD has created an opportunity for tape to increase the capacity and cost advantages of tape storage over HDD.

New tape architectures are addressing indexing, tags and naming conventions and enabling a move toward an "object-oriented" approach to keep the vast storage reservoirs tape provides usable. These concepts benefit from more intelligent drive and cartridge systems that can rapidly locate specific objects or data, enabled by technologies such as the LTFS file system. All the tape drive vendors now have file systems access to tape, effectively creating Network Attached Storage (NAS) tape. In addition, many vendors are moving to object storage support on tape. Object storage with magnetic tape in cloud-based storage environments is a future growth opportunity for tape archiving.

Recommendations on Priorities and Alternative Technologies

Tape storage will continue to play a key role as the foundation of the storage hierarchy "pyramid", providing an archival non-volatile solution for vital data records of businesses and governments, as well as (indirectly) for consumer data stored in "the cloud". As such, the underlying magnetic recording, as well as drive and library systems technologies, needs to be continuously tracked to understand and follow improvements in performance and reliability. Of particular interest will be how other storage technologies such as hard disk drives and solid-state drives may be used to augment and enhance tape systems performance and efficiency. Moreover, adding a file system and object storage capability to tape storage systems has had a dramatic impact on the design of tape library systems and their use. The large-scale adoption of tape technology by hyperscale cloud companies is also likely to influence future development directions for tape drives and libraries.

References

1 A. Acilla, "Quantifying the Economic Benefits of LTO-8 Technology" Accessed: Dec. 3, 2022. [Online]. Available: https://www.lto.org/wp-content/uploads/2018/08/ESG-Economic-Validation-Summary.pdf

2 IDC Worldwide Branded Tape Share Report Four Quarter Roll-up through 4Q 2021

3 IBM Diamondback Tape Library, Accessed: Dec. 23, 2022. [Online]. Available: https://www.ibm.com/products/diamondback-tape-library

4 "Point Archival Gateway", Accessed: Dec. 3, 2022. [Online]. Available: https://www.point.de/en/products/point-archival-gateway/

5 "Object Archive", Accessed: Dec. 3, 2022. [Online]. Available: https://www.fujifilm.com/us/en/business/data-storage/data-management/data-archive/object"

6 "OTFormat Specification", Accessed: Dec. 3, 2022. [Online]. Available: https://activearchive.com/wp-content/uploads/2020/09/OTFormat_Specification_ver1.0.0-1.pdf

7 Github repository for SwiftHLM (Swift Hight-Latency Media) middleware, Accessed: Dec 3, 2022. [Online]. Available: https://github.com/ibm-research/swifthlm

8 Linear Tape Open Roadmap, Accessed: Dec. 3, 2022. [Online]. Available: https://www.lto.org/roadmap/

9 Information Storage Industry Consortium (INSIC) 2019 International Magnetic Tape Storage Roadmap, Accessed: Dec. 3, 2022. [Online]. Available: http://www.insic.org/wp-content/uploads/2019/07/INSIC-Technology-Roadmap-2019.pdf

10 S. Furrer et al., "201 Gb/in2 recording areal density on sputtered magnetic tape," IEEE Trans. Magn., vol. 54, no. 2, Feb. 2018, Art. no. 3100308. DOI: 10.1109/TMAG.2017.2727822

11 S. Furrer et al., "317 Gb/in2 Recording Areal Density on Strontium Ferrite Tape" IEEE Transactions on Magnetics, JULY 2021, 57, 7, pp. 1-11, DOI: 10.1109/TMAG.2021.3076868

12 P.-O. Jubert, Y. Obukhov, C. Papusoi, and P. Dorsey, "Evaluation of Sputtered Tape Media with Hard Disk Drive Components", IEEE Transactions on Magnetics, Vol. 58, No. 4, April 2022 , DOI: 10.1109/TMRC53175.2021.9605133

Optical Archival Data Storage

SITUATION ANALYSIS

Optical discs are a nearly ubiquitous form of digital data storage as over 256 billion[i] units [1] have been sold over the last 40+ years (compared to 10 billion HDD and 352 million LTO units). CDs, DVDs and BR discs comprise the media collections of consumers all over the world including CDs that are still playing pristine, high-fidelity music after four decades. Perfect fidelity DVDs have also demonstrated decades of longevity.

CDs and DVDs have been shown to survive the most extreme environments of up to 80°C when stored in automobiles. This outstanding longevity and robustness to environmental conditions make optical storage an important approach for the burgeoning need for a carbon-friendly and energy efficient enterprise active archival data storage medium.

The previous applications of optical discs have included replicated discs for media distribution and write-once read many (WORM) and rewritable technologies for data storage, mostly for consumer applications. This report will focus on the recent developments in WORM optical media and their forward-looking roadmaps toward applications in enterprise archival storage. Following a brief summary of the historical consumer-oriented technology and applications, the value proposition of optical storage for enterprise archival data storage will be described.

The incumbent technology is characterized by optically written data marks on thin layers in single and multiple photosensitive layers deposited on plastic substrates. Discs are spun at high speed with focus and tracking servos securing rapid access to data locations in three dimensions.

Multilayer disc media will be reviewed, while new structures in the market and in development will be reviewed in this report, including a description of the media, drives and libraries comprising the current and future optical media technologies.

CONSUMER MARKET

Research into optical data storage and its practical applications has been ongoing for many decades, with analog image microfiche widely considered as the first optical storage medium, introduced as early as the 1800s [2]. The first true widely-adopted digital data storage system was the replicated Compact Disc, introduced in 1982, adapted from audio (CD-DA) to data storage (the CD-ROM format) with the 1985 Yellow Book, and re-adapted as the first mass-market optical storage medium with CD-R and CD-RW in 1988. Subsequently, DVD and Blu-Ray discs have been developed in the 1990s and 2000s respectively to accommodate full-length movie distribution in ever-increasing display resolution [3].

CD/DVD/BLU-RAY

The original concepts of CD/DVD drives continue to be the basis for the incumbent multilayer enterprise data storage media, with innovations built on the basics: dynamic focus and tracking servos and PRML and related data channels, among others. Indeed, the disc form factor is common

---

[i]Estimate based on Global Consumer CD unit sales and Global Consumer DVD and BD $ sales

to all the multilayer technologies described below. Digital data is encoded with marks of various lengths with accepted error correction codes. The evolution of the layered disc toward higher capacity has relied on shortening the laser wavelength and increasing the numerical aperture of the objective lens to enable shrinking the size of the diffraction limited spot and shortening the track pitch, as depicted in **Figure 41**. Other innovations described below also increased the data density/disc capacity. Capacity increases also involved the addition of layers.

Since the focus of CDs and DVDs was mass media distribution, few improvements were made throughout their lifecycle, and the development effort has been mostly focused on standardization of manufacturing and cost reduction. The transition from CD to DVD was mainly driven by the storage of feature films versus music collections. The transition to Blu-ray was driven by increases in video display resolution.

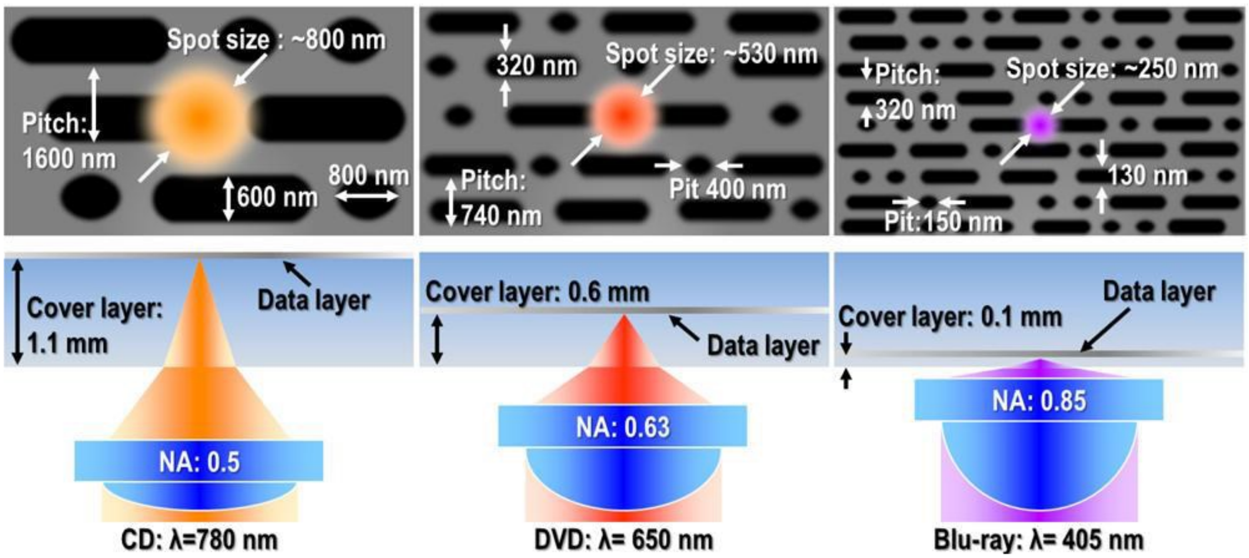


Figure 41. Decrease in laser wavelength and spot size with increasing numerical aperture (NA) lenses. [4].

Reproduced with permission: American Chemical Society

Blue-Laser Disc is defined by Blu-ray Disc (BD) format and all other types of optical data storage technology (for example, UDO) that use blue lasers for writing/reading. Thus, Blu-ray became the natural successor to DVDs and HD DVDs with its superior capacities, (25 GB per layer vs 4.7 GB per layer) which were essential to distribute 4K and ultra-high-definition videos.

The introduction of Blu-Ray coincided with the popularization of streaming media over the internet. With competition from both the lower cost of DVDs and the improved convenience of streaming and digital downloads from vendors, Blu-ray never experienced the same level of success as DVDs and CDs. Blu-ray sales only barely surpassed DVDs sales recently, which was a shadow of its former volume two decades ago. The trend away from consumer applications of optical media is depicted in **Figure** 42[5].

Personal computer applications of optical storage have also been a traditional market for optical media. In this case, the on-board drives have been used for software distribution on read-only discs as well as local data storage on writable and re-writable discs.

Traditionally, read-only CDs and DVDs (CD-DA, CD-ROM, VCD, DVD-ROM, BD-ROM, etc.), are used for large software distributions. However, with the advent of downloadable software and increased broadband availability, optical discs became a less popular or desirable means of distributing software [6], [7]. More recently, major gaming consoles are releasing devices without optical drives despite the piracy concerns associated with the industry [8].

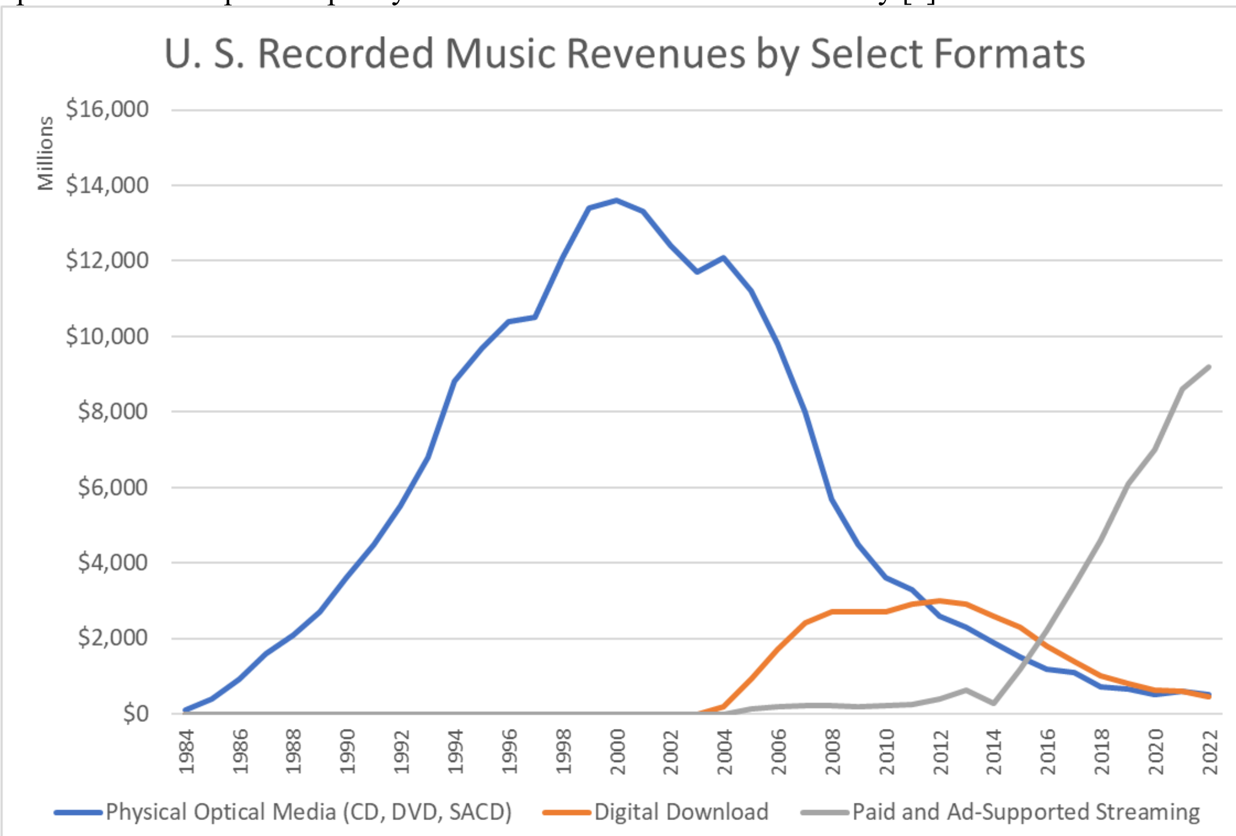


Figure 42. Music delivery. DVD sales exhibit a similar trend to CD sales with a peak in the year 2004 [9], [10].

For a time, write-once and rewritable discs (CD-R, CD-RW, DVD+/-R, DVD+/-RW, DVD-RAM, BD-R, BD-RE, etc.) were commonly used for general file operations using simulated file systems such as UDF [11]. However, this market declined quickly with the introduction and widespread use of NAND-based SSDs. The performance and capacity of SSDs make them the obvious choice in these applications.

Discontinued/Marginal Consumer Media

Very few major developments in the optical storage space pertaining to consumer applications took place over the past 5 years. Please refer to the previous edition of the iNEMI roadmap [12] for more details on technologies such as Magneto-Optical, VMD, Millennium Discs, and others.

Enterprise Use Case Value Proposition

Just as the demand for archival data storage is rapidly expanding due to e.g. data science demands and the internet of everything, both business and technical issues are straining the incumbent magnetic media. These stressors include the breakdown of Kryder's law [13], the necessity for a sustainable approach to data storage, and the oligopoly of both storage producers and consumers.

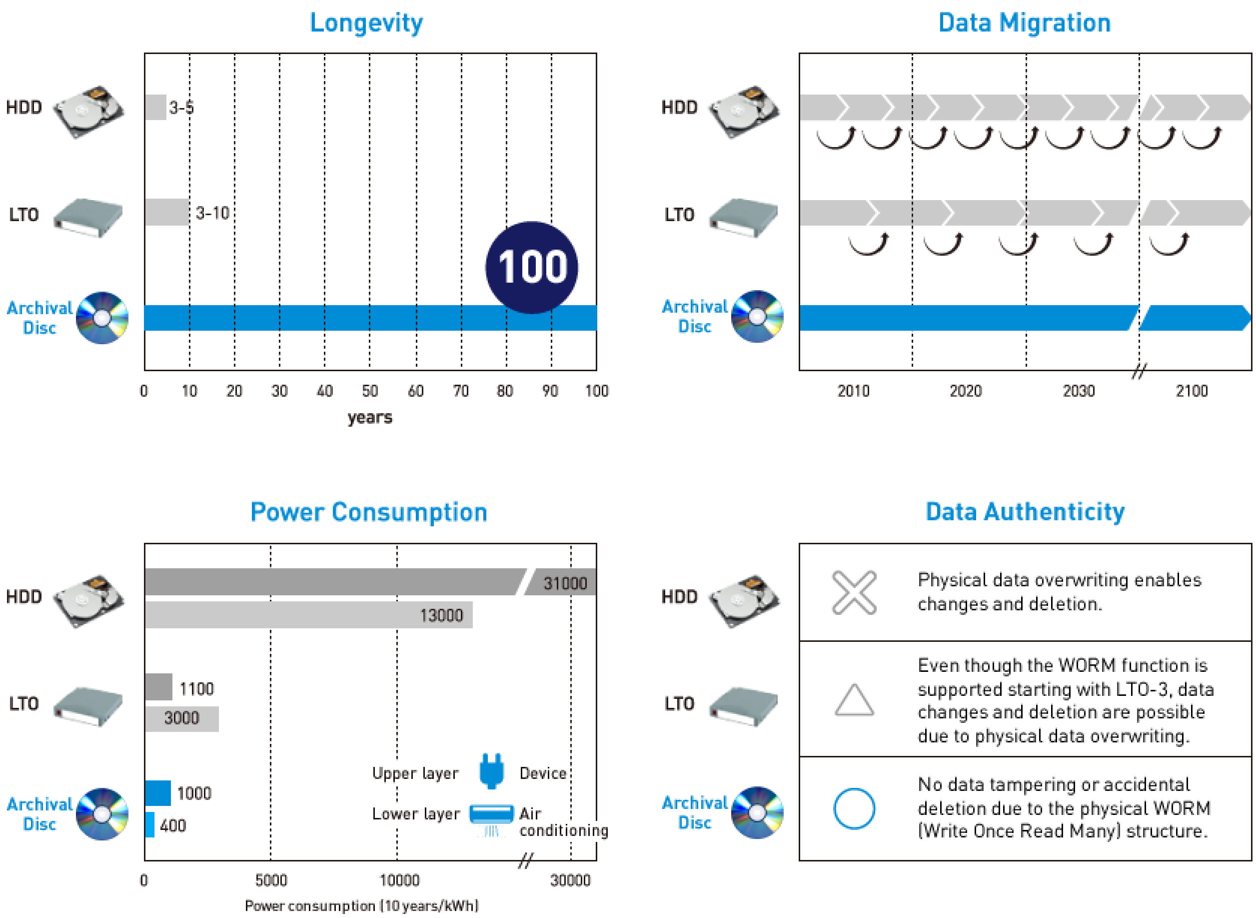


Figure 43. Storage media comparison. [14]

Reproduced with Permission: Panasonic Holdings Corporation

Optical data storage has the potential to meet these challenges. Several elements of the value proposition of optical data storage are shown in **Figure 43**[15]. These advantages arise from the robust nature of marks produced by WORM photothermal processes in the photosensitive materials comprising the active layers. As the photothermal conversion occurs at hundreds of degrees, the resulting WORM marks are extremely robust, leading to a long lifetime as the disc materials also have century-scale lives. Figure 44 depicts the lifetime/temperature trade-off for AD technology indicating that even at the unlikely occurrence of continuous storage at 35 C, decades of life are expected [16], [17]. The environmental robustness and longevity have been verified over the decades since the introduction of optical discs as mentioned in the introductory paragraph.

These generic attributes of optical media also minimize power consumption and carbon footprint, which is increasingly important for enterprise archives facing increasingly larger cold data storage sizes. Unlike magnetic media, optical media can generally be stored in relatively uncontrolled ambient environments, leading to significant energy/cost savings as indicated in **Figure 44** and **Table 5**. As a result of this and their longevity, optical media generally require zero energy to store, an important factor in their energy consumption and carbon footprint.

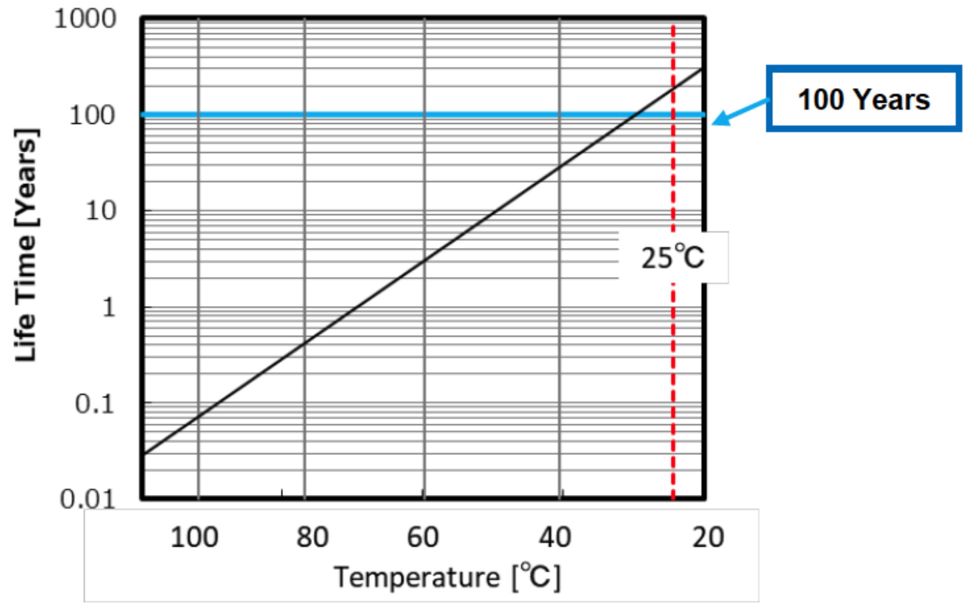


Figure 44. Accelerated lifetime test results for Sony/Panasonic Archival Disc. [18], [19]

Reproduced with permission: Sony Group Corporation and Panasonic Holdings Corporation

Table 5. Common Enterprise Archival Storage Media

	Unit Capacity	Media (Cost/TB) [20]	Transfer Speed (MB/sec)	BER with Error Corr.	Archiving Power Consumption (/time/Capacity)[ii]	TCO (Total Cost of Ownership ($MM/100GB/20 years)[iii]
HDD[iv]	22 TB 2.2 TB/disk	>$10	291	1E-15	~440	7.6
Tape[v]	18 TB (Native)	>$5	400[vi]	1E-19	~41	2.9
BDXL[vii]	100 GB to 200 GB	<$40	18-36	1.00E-18	~14	0.7
Archival Disc	0.5 TB	$30-45	30-45[viii]	1E-23	~14	<0.7

The long lifetime is especially important for data archiving for decades as the media does not need to be remastered. Typical remastering intervals for HDD and magnetic tape are 3-10 years, so that archiving data for 50 years can result in an order of magnitude of cost savings accounting for both media and remastering infrastructure.

The robustness of optical media has important cybersecurity implications as the WORM data is impervious to tampering. Because the discs are stored separately from the drive, the system possesses an "air gap" to provide passive media protection from cybersecurity breaches. In addition, optical media are impervious to electromagnetic pulse (EMP) and immersion in water.

Current commercial optical data storage technologies are based on concepts developed almost 50 years ago for consumer applications. The pivot of optical storage to data center archiving presents opportunities for new technical approaches at the system level, allowing novel concepts at the device/media level as well. In addition, advances in optical components driven by the imaging and display industries introduce new components for such systems. These advances will be described below.

TECHNICAL REQUIREMENTS (BUSINESS/TECHNICAL ISSUES)

Despite its advantages, optical archival storage has not found widespread use as neither its capacity nor its cost is competitive with HDD and tape. At 1,500 PB per year, the optical archival market is less than half of 1% of the total storage market.

[ii] Optical Technology | Optical Data Archiver freeze-ray series | Panasonic Global
[iii] Low Power and Cost Efficient Data – HIE Electronics
[iv] WD Gold Enterprise Class SATA Hard Drive Up To 22TB | Western Digital
[v] Fuji LTO Ultrium 9 Tape (16659047). Fujifilm LTO9 Ultrium Tape Data Cartridges with Barium Ferrite (BaFe) (tapeandmedia.com)
[vi] Transfer speed for a tape cartridge, which leverages multi-track parallel transferring. LTO-9 models can write on 32 data tracks at a time.
[vii] Disk Prices (US)
[viii] Transfer speed per individual optical discs. Higher transfer speed can be achieved through multi-unit drives and various library systems. See *Archival Disc Libraries* Section for detail.

Optical storage technologies have evolved from CD to DVD to Blu-ray as capacity requirements have increased for music, video, gaming, and personal storage. The basic disc and drive technology have not undergone fundamental change. On the disc side, the basic disc format has remained constant with track density reflecting the increasing capacity permitted using higher numerical aperture lenses and shorter wavelength lasers.

The development of wide bandgap semiconductor lasers such as those using GaN, the subject of the 2014 Nobel Prize in physics, was a major innovation leading to 405 nm wavelength lasers for Blu-ray and, now, Archival Disc (AD) technology. Sony and Panasonic introduced a two-sided disc along with improvements in material and drive technologies to achieve the significant capacity advances in Archival Disc technology. Drive improvements include progress in PRML, other data channel implementations, and error correction schemes.

Generations of optical discs have shared an additive manufacturing process consisting of an injection molded polycarbonate substrate with the tracking pattern embossed on the upper surface. For CD and DVD technology, the relatively low numerical aperture of the objective lens provided a large working distance so that a polycarbonate slab can be used as the cover of the written layers. The transition to the blue laser and 0.85 numerical aperture required a shorter working distance, well below the 1mm scale.

This required the development of both a deposited thin cover layer and a hard protective coat. Various materials have been used as active media over the years, but current blue laser re-writable technologies are based on photothermal inorganic phase change composite materials that change from reflective crystalline to non-reflective amorphous states upon writing. Typically, phase change composites contain Te, Ge, Sb along with other elements.

On the other hand, for WORM technology like recordable Blu-ray discs, various inorganic materials such as phase-change composite, metal-oxide and metal-alloy have been used. These inorganic materials are UV and temperature stable and have long lifetimes, making them suitable for archival applications. Alternatively, photosensitive organic dyes have been used for recordable Blu-ray discs using Verbatim's spin-coating process instead of the traditional sputtering process for inorganic materials.

The capacity of Blu-ray discs can be potentially increased by adding more recording layers [21], but an additive, sequential layer deposition process even at 98% yield for a single layer would result in a prohibitively low manufacturing yield for more than 3 or 4 layers. **Table 6** is a summary of some current and novel optical products and technologies.

Table 6. Summary of select current and novel optical products/technologies

Product Name	Development State	Recording Medium	Service Segments	Max Capacity GB/disc	Media Cost per TB	Read/Write Speed (MB/sec)[ix]	BER with Error Corr.
CD/DVD	Commercial	Photosensitive organic dyes	Consumer	0.7-30	$15-$30	7.2-66/7.2-66	1.00E-9 1.00E-15[x]
Blu-ray	Commercial	Phase change inorganic composites; photosensitive organic-dyes	Consumer Enterprise	50	$15+	36-72/36-72	~1.00E-18, Varies by drive[xi]
BDXL	Commercial	Phase change inorganic composites	Consumer Enterprise	200	<$40	18-36/9-36	~1.00E-18, Varies by drive
Archival Disc[xii]	Commercial	Inorganic oxides	Enterprise	300-500	<$33	90	~1.00E-23
Folio Disc	Prototyping	Photosensitive dyes dispersed in polymer matrix – reflective or fluorescent media	Enterprise	500+ (projected)	<$5	TBD	TBD
Piql[xiii]	Commercial	polyester photographic film	Enterprise	N/A	N/A	40	TBD
DOTS	Prototyping	metallic alloy sputtered on polyester photographic film	Enterprise	1200+ (projected)	TBD	1,000	TBD

The photosensitive material possesses a strong optical absorption feature at the 405 nm wavelength so that the absorbed optical energy of the writing pulse is quickly converted to heat which increases the temperature beyond a phase transition or decomposition event. Both processes, then, are characterized by an absorbed-heat threshold converting the active material to produce the mark. The threshold nature provides that the much-lower reading power level of the incident laser does not significantly affect the written marks, thus ensuring millions of reading cycles. This is also an important factor in the photostability of the written data to ambient light.

The production of data marks is a complex process, involving not only the photothermal event but also the transport of heat over the relevant time and length scales. The mark is written with a particular writing strategy, which is a sequence of laser pulses near the nanosecond time scales carefully crafted to minimize thermal transport and create the smallest marks, at best at the ~10-100 nm length scale in blue laser media, well below the diffraction limit of light at 405 nm wavelength. The writing strategy depends not only on the materials, but also on the dimensions of the layers as they affect heat confinement and transport.

INCUMBENT MEDIA: MULTILAYER DISC

BDXL

The Blu-ray Disc Association released specifications for two multi-layer variants of Blu-ray discs in 2010, dubbed Blu-ray XL (BDXL) [22]. With this release, BDXL increased Blu-ray capacity to 100 GB (rewritable) or 128 GB (WORM). The fourfold increase in capacity is achieved by using three or four recordable or rewritable BD layers without increasing their areal density. This opened

[ix] Pioneer Just Made A New Optical Disc Drive For PCs And Yes It's 2022 | HotHardware
[x] DVD Benchmark (hometheaterhifi.com)
[xi] Error-correction codes for optical disk storage [5643-58] (psu.edu)
[xii] Folio Photonics Wants To Kill HDDs With Film (forbes.com)
[xiii] https://www.ejournals.eu/pliki/art/20806/

the door for Blu-ray disc enterprise archival applications, which have gained traction in some sectors.

In 2014, a double-sided version of Blu-ray standard was released (BD-DSD), which allows 200GB of storage between 6 layers (3 on each side) [23]. However, BD-DSD requires specialized optical drives to process and is not widely available. Since there is currently no roadmap for BDXL development, this solution will likely be capped at the current capability.

In March 2022, Pioneer released a new optical disc drive that delivers 8x/6x speed recording on triple/quadruple-layer BDXL discs, a substantial improvement over the 4x/2x speed drive released in 2017 [24]. It is unclear whether there will be an ongoing roadmap moving forward. **Figure 45** shows the layer structure for 3 and 4 layer BDXL discs.

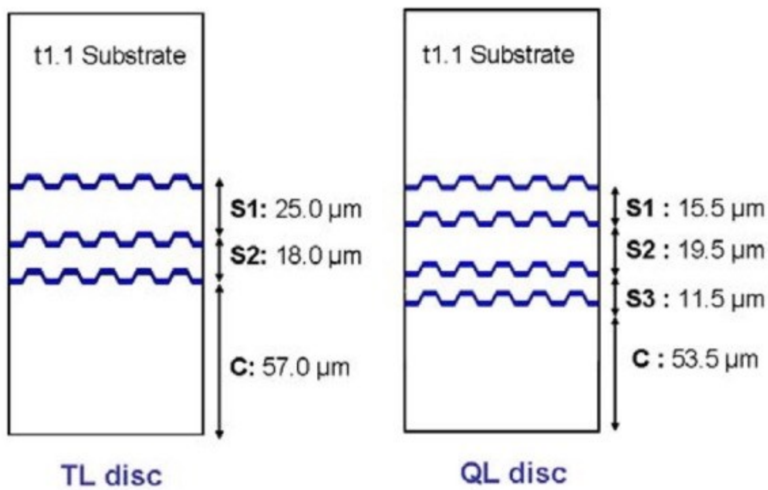


Figure 45. BDXL layer structure. [25]

Reproduced with permission: Blu-ray Disc Association

SONY + PANASONIC ARCHIVAL DISCS

Sony and Panasonic have successfully developed a next-generation optical disc for enterprise storage with an initial capacity of 300GB (for 2016 shipment). The Archival Disc (AD) has the same dimensions as current Blu-ray discs and will also be readable for at least 50 years. The disc has three layers per side.

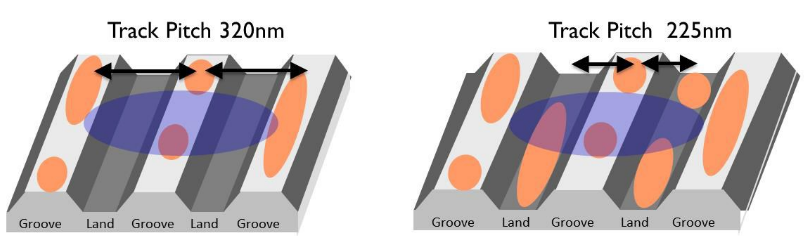


Figure 46. Land-and-groove recording. Blu-ray Disc on the left vs. AD disc on the right. [26]

Reproduced with permission: Sony Group Corporation and Panasonic Holdings Corporation

A major breakthrough of the Archival Discs was the utilization of land-and-groove recording technology, nearly doubling the areal density per layer compared to conventional Blu-ray discs (**Figure 46**). This also enables the Archival Discs to use existing optical units with the same 405nm laser at 0.85 NA. Sony and Panasonic indicate that the crosstalk noise generated between adjacent tracks is cancelled out by their newly developed crosstalk-cancelling technology. AD is backward compatible with existing Blu-ray standards. Panasonic AD drives also support BD discs.

Sony and Panasonic have been steadily making improvements on the AD product. In 2020, Sony and Panasonic released the second generation of discs with 500GB capacity, which was achieved primarily by shortening the data bit-length, thus improving linear recording density [27]. In addition, the use of new oxide-based materials improves recording rate, disc capacity and durability. This improvement was made possible by AD's advanced inter-symbol interference elimination technology, which will be fitted to 2nd generation drives to rectify reduced playback spot resolution caused by higher recording density. This generation of AD also benefited from a new data format, which improved linear recording efficiency by 7%, and new channel modulation/advanced error correction code to reduce data error rates.

Their literature describes the next generation to have an estimated capacity of 1TB and indicates that this higher capacity will be achieved through signal-processing technologies including multi-level recording technology as depicted in **Figure 47**.

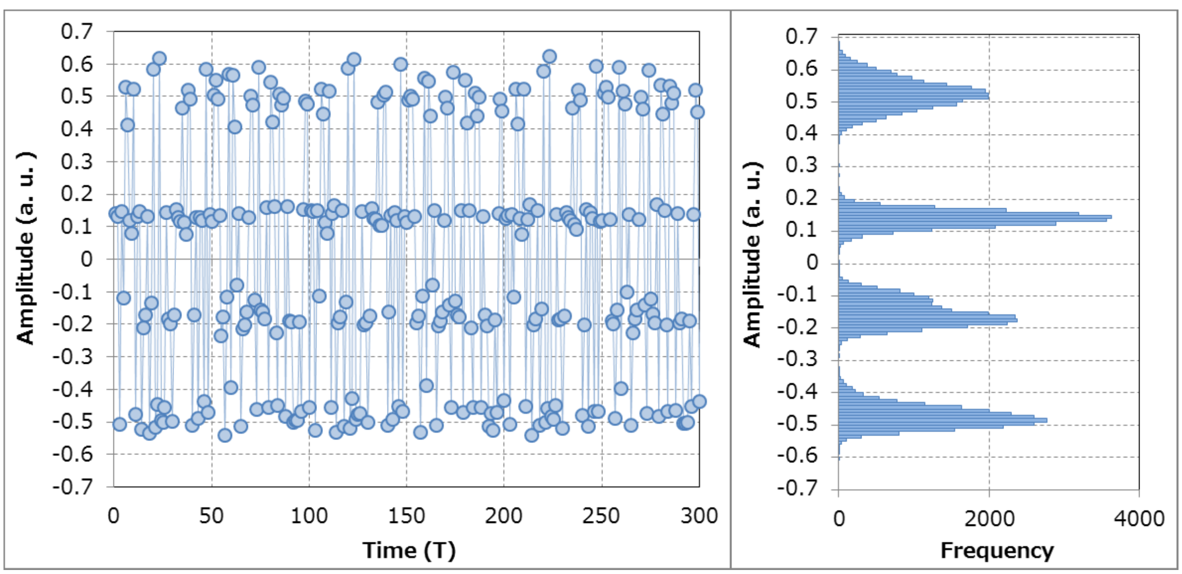


Figure 47. 3rd generation AD multiple-level recording/playback performance [28]

Reproduced with permission: Sony Group Corporation and Panasonic Holdings Corporation

Sony and Panasonic's 3rd generation AD would double the capacity of 2nd generation by encoding twice as much data with the same format and physical disc structure by effectively encoding 2 bits of information per each data spot (**Figure 48**). Over the course of each generation, disc capacity is raised without changing base optical parameters or media structure, minimizing manufacturing cost and ensuring the AD system's backward compatibility.

Sony and Panasonic note that Archival Discs would last 50 to 100 years archival life depending on various environmental factors. The discs have been demonstrated to be readable after being immersed in seawater for a period of 5 weeks. Sony also states that since optical discs are not magnetic, they are unaffected by geomagnetic events and other sources of electromagnetic pulses (EMP).

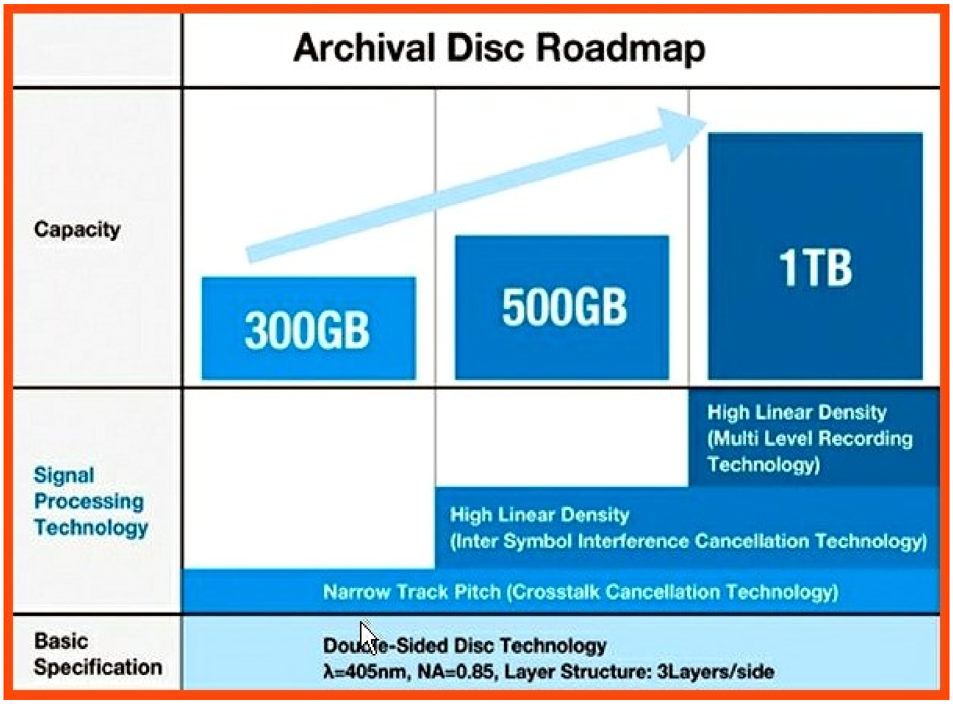


Figure 48. Archival disc capacity roadmap. [29]

Reproduced with permission: Sony Group Corporation and Panasonic Holdings Corporation

Sony and Panasonic depict the roadmap for higher-capacity ADs. As the graphic in **Figure 48** shows, the third generation will have a 1 TB capacity. Sony and Panasonic have not disclosed plans for higher capacity beyond 1TB. Consider the data in **Table 7**. Here the basic disc specs are given.

Table 7. Main parameters of AD discs. Reproduced with permission: Sony Group Corporation and Panasonic Holdings Corporation [30]

【Main parameters】	【AD 300GB Specifications】	【AD 500GB Specifications】
Disc diameter	120mm	
Total nominal thickness	1.2mm	
Double sided disc	Triple Layer (TL)/Side	
Cover Layer thickness	57.0um	
Recording polarity	High to Low	
Recording method	Land & Groove	
Data Zone inner radius/ outer radius	24mm/58mm	
Track pitch	0.225um	
Addressing method	Wobbled Grooves with addresses	
Maximum user data transfer rate	359.65Mbps	
Channel modulation	17PP	110PCWA
Error correction code	64KBLDC+BIS	256KB Extended-RSPC
Total efficiency	81.738%	87.850%
Nominal Channel bit length	53.0099nm	34.216nm
Nominal Data bit length	79.5149nm	51.324nm
User data capacity 120mm	300.00572GB	500.12357GB

DATA PRESERVATION

New products on the market and in development are focusing on immutable storage at the millennium scale. These will address preservation markets and may also be applicable in markets for more active data storage. Piql AS has products in the market and more are described later as under development.

PIQL AND PIQLFILM[xiv]

Backed by the European Union and the Norwegian government, the Norwegian company **Piql** has developed a photosensitive film, **piqlFilm**, that both protect and preserves data for 1000 years [31]. What Piql has done is in principle to convert an established and well proven information carrier, the 35 mm black and white photographic film, traditionally used for analogue data storage, into a digital storage and preservation medium that can be used to store any digital data [32]. The technology preserves data in the form of high-resolution QR codes that is decodable using open-source software. The piqlFilm has been developed in collaboration with three film manufacturers, Kodak (US), Harman Technology (UK) and Filmotec (DE). **Figure 49** shows the piqIFilm package

Figure 49. piqlFilm package

[xiv] This section authored by Piql AS Managing Director Rune Bjerkestrand

The piqlFilm is a black and white, negative silver-halide film on polyester base, see **Figure 50**. It has extremely fine grains (20 nm to 40 nm) and high resolving power (>1000 line-pairs per mm). Further, piqlFilm has truly unique security and longevity properties, ideally suited for offline/"cold storage" of valuable and/or irreplaceable data and information. Data stored on piqlFilm cannot be hacked, modified, or deleted, neither can data be destroyed by electromagnetic weapons or nuclear radiation. The piqlFilm is packed in the piqlBox to protect it physically and over time. The piqlBox is manufactured using a specially designed and manufactured polymer material that has no negative impact on the piqlFilm and data over time and that has the same longevity as the piqlFilm, i.e. 1000 years. The same goes for the label used on the piqlBox.

The piqlFilm is made self-contained and self-explanatory due to the fact that it contains human readable instructions (i.e., in addition to the digital data, see **Figure 51**) on how to understand the storage medium and how to deal with it in the future. Further it contains all file format descriptions and relevant source code for programs needed to render or view information in the future. Piql has even developed a virtual machine that makes future data retrieval independent of specific hardware or operating systems available at that point in time. This makes the solution resilient against the accelerating developments and obsolescence of specific software and hardware.

Figure 50. Digital and analog data on piqlFilm

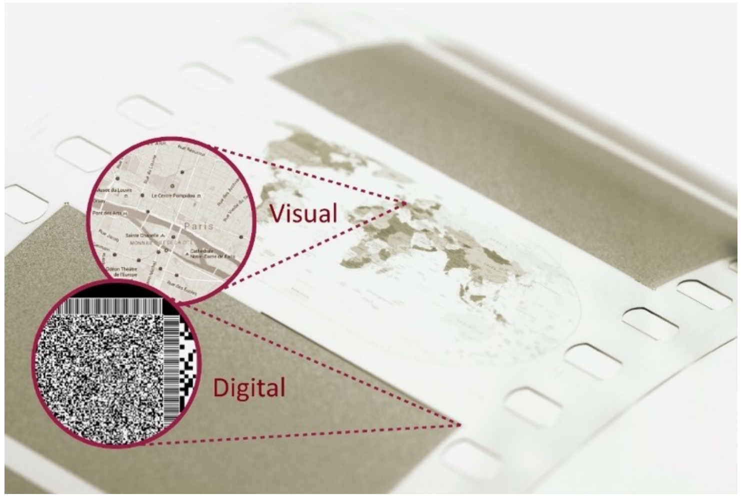


Figure 51. piqlFilm. Reproduced with permission: Piql AS

Since the piqlFilm is migration-free, passive and requires no energy to keep data alive, it is a truly sustainable storage technology with a close to zero carbon footprint over time.

Data is ingested through Piql's software platform, **piqlConnect**, where data upon being organized by the client, is transferred to a specific machine, the **piqlWriter**, that encodes the received bit stream into QR-codes, such as shown in **Figure 52**.

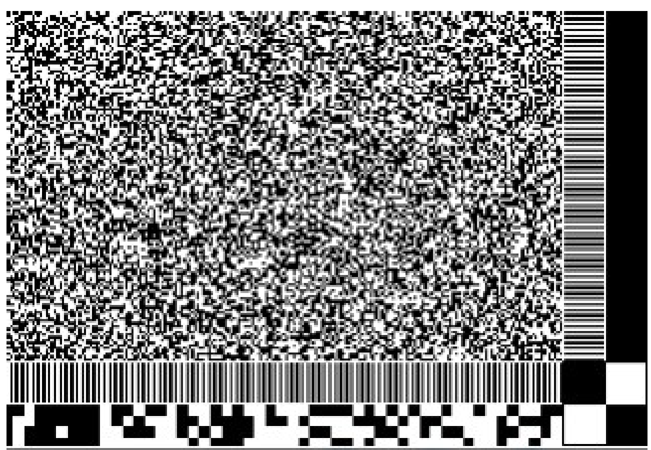


Figure 52. Sample binary data on piqlFilm

The piqlWriter receives the encoded binary data as a set of files representing the writable data-frames (images) and writes them on the piqlFilm at 20 data-frames per seconds, or roughly 40MB/s. Imaging is based on a Texas Instruments Digital Micromirror Device (DMD) imaging sensor, consisting of more than 8,800,000 micro-mirrors capable of writing datapoints of 6µm size. The piqlWriter uses a monochromatic green LED light which is modulated by the DLP. Each datapoint can be written in black, white or shades of grey.

The photographic nature of the film enables piqlFilm to store both analog and digital data, which is essential for ultra-long-term data preservation purposes, where future data retrieval without current technology is an important consideration (**Figure 50**) [33]. The decoding method is written in human-readable text that can be read with a magnifying glass, and the digital data stored is searchable as with any other digital media.

One fundamental principle for Piql when they designed their solution was to ensure that data could be read back tomorrow, next year or a thousand years into the future, without the need for specific hardware or software. What in principle is needed to retrieve the data from piqlFilm is:

- a light source (to illuminate the data-frames on the piqlFilm),
- a magnifying glass (to hold over the piqlFilm where illuminated),
- a camera (that can capture the image (i.e. the QR code) as seen through the film and the magnifying glass),
- and a computer (that can interpret and convert the QR code into the original file format).

All the instructions, the file format descriptions and the software that is needed (i.e. the source code), is all included on the piqlFilm and can be read by the human eye and understood by a non-technical person. Even the instructions how to build an automated reading device is included on the piqlFilm.

To automate and industrialize the readback of the piqlFilm, Piql has developed the **piqlReader (Figure 53)**.

The piqlReader is a high speed, high-resolution digital data reader. It captures data-frames from the film and restores digital and visual data. The piqlReader is used for two purposes; The first use is data verification after the film has been written and subsequently processed. This ensures that data written on film is verified to be both readable (with low error rates) as well as authentic (i.e. checksum of the files are verified). The second use is data retrieval upon request. When a client requires data from the film, the piqlReader is used for accessing the data.

Figure 53. A piqlReader

The piqlReader features precision mechanical film handling components, customized optics, a diffused 405nm LED light source, a built-in line-scan 12K camera and a high-performance computer. It runs open-source software (available on GitHub) for capturing images and decoding them to digital data in real-time. Retrieving digital data from film requires two processes running in parallel. The first process is capturing the image from the film, and the second is decoding that captured image file. The piqlReader reads in continuous mode and can quickly access the needed file, whether at the beginning or at the end of a piqlFilm. Data retrieval speed is 24 MB/s or 12 data-frames per second.

Piql has established a global network of distributors; Piql Official Resellers, Piql Partner and B2B partners (i.e. larger system integrators/solution providers) that deliver Piql's services for **data protection, archival and long-term digital preservation**. Clients across the world can reach these services through the SaaS platform, piqlConnect, that is the connection to/from the offline, off-grid piqlFilm as well as online storage. The Piql Partners (located across the world on all continents) manage a local piqlVault where the piqlFilm is securely stored.

Piql's services have been delivered to prestigious clients like the Vatican Library, European Space Agency, GitHub, various National Archives, National Museums, National Libraries, banks, large corporates, SME's, research institutions, various public agencies (legal, infrastructure, utilities, nuclear, health, defense, public administration and more).

Piql is also the initiator of the **Arctic World Archive** (AWA), a repository for World Memory located on the Svalbard archipelago in the Arctic Ocean, between the top of Norway and the North Pole.

AWA was established in 2017 [34]. based on a vision to ensure that our digital memory is not lost or manipulated but shall be available for future generations in its authentic form, - in a world where few places are safe from natural and man-made disasters, hacking, cyberattacks, wars, and terror. Data is stored on the piqlFilm that is kept in a secure vault in the dept of the permafrost in an Arctic Mountain. The lifetime of the data is expected to be beyond 2000 years when stored in a cold and dry climate like in the Arctic World Archive.

OPTICAL STORAGE SYSTEMS

BDXL LIBRARIES

Amethystum, KDS, and NETZON

Some Chinese archival storage management system providers have created library systems that integrate optical, magnetic, and NVM storage media in integrated systems to achieve cost-performance optimization in enterprise storage. This approach allows the volume density to be comparable to other storage solutions, while streamlining optical-magnetic-digital system integration. Major library makers include Amethystum, KDS, and NETZON (**Figure 54**). These solutions are popular in the Chinese market.

Figure 54. NETZON HDL 10368 holds 10368 discs enclosed in 864 cartridges and 36 parallel drives. [35]

Reproduced with permission: Suzhou NETZON Information Storage Technology Co., Ltd.

In the realm of pure optical media library development, the library makers take a similar approach [36]. BDXL library makers pack as many discs as possible, with up to ~12000 Blu-ray XL discs packed in a cabinet, realizing a 2.5 PB/cabinet deployment (1.25 PB for the standard 19 "42U cabinet) [37]. Multiple drives allow for high parallel data access speed. **Table 8** shows specifications for the Amethystum ZL BDXL Optical Storage System.

Table 8. Amethystum ZL series BDXL Optical Storage System Specifications [38]

Model	ZL600	ZL1800	ZL2520	ZL6120	ZL12240
Storage	60TB	180TB	504TB	1224TB	2448TB
Disc number	600	1800	2520	6120	12240
Maximum Drive Number	6	6	12	24	48
Network Interface	1Gb/10Gb Ethernet				
Avg. grab time	14s	14s	60s	60s	60s
Max Transmission rate	162MB/s	162MB/s	324MB/s	548MB/s	1296MB/s
Voltage	100-240V AC / 47-63Hz				
Size	19-inch 24U cabinet	19-inch 37U cabinet	19-inch 25U rack	19-inch 42U cabinet	1000x800x20 90mm
Weight (Fully Loaded)	150kg	240kg	180kg	454kg	1028kg
Operating Environment	10°C to 35°C, 20% to 80% Humidity				

Consumer Library Designs for Enterprise Use

Some consumer-grade optical archival systems (Optical Jukeboxes) manufacturers such as Zerras and Kintronics have developed enterprise solutions by upscaling their consumer products [39], [40]. These lower capacity libraries have short (<10 seconds) access time and the minimal number of robotic mechanisms may prove desirable for near-line storage applications that rely on the flexibility and modularity it provides.

Archival Disc Libraries

The technology for a modern data archiving system based on optical discs requires two primary elements: a robust, high-capacity optical disc and system architecture suitable and effective for enterprise data centers.

The collaboration between Sony and Panasonic on the Archival Disc has established system-level products, but the high cost of the media has hobbled market penetration. The systems described below have recently been discontinued according the Sony and Panasonic websites.

Sony Everspan

Everspan aimed for an ultra-high-capacity storage solution, but has been displaced by Sony PetaSite [41]. It consisted of three types of units: The Base Unit, the Robotic Unit, and up to 14 Expansion Units, based on the 'triplet' rack developed by the Open Compute Project (OCP) and containing up to 43,520 Archival Disks each. When using 300GB ADs, the total capacity of a single Everspan system is 181 Petabytes and up to four systems can be linked. Everspan has a relatively high I/O rate – the robotic read-write array features not one, but eight lasers for a total read speed of up to 18 Gigabytes per second.

Sony PetaSite

In 2020, Sony released the 3rd generation of Optical Disc Archive Cartridge alongside the corresponding drives [42]. The third-generation disc cartridge holds 11 AD 500GB discs with a maximum capacity of 5.5TB, compared to the previous generation's 3.3TB. The third generation ODS-D380U/F features eight lasers. With two assemblies positioned at the top and two at the bottom, the drive can read and write both sides of the disc simultaneously. The drives were able to improve the read/write speed to 375MB/187.5MB (3Gbps/1.5Gbps) while maintaining the backward compatibility of the 1st and 2nd generation cartridges. **Figure 55** and **Table 9** show the Sony PetaSite library system and give the library system specifications.

Figure 55. Sony PetaSite library system. [43]

Reproduced with permission: Sony Group Corporation

Table 9. Sony PetaSite library system specifications. [44]

Reproduced with permission: Sony Group Corporation

	ODS-L30M	ODS-L60E	ODS-L100E
Maximum number of drives	2	4	0
Maximum number of cartridges	30	61	101
Data capacity *1	Up to 165TB	Up to 335.5TB	Up to 555.5TB
Host interface	Fibre Channel 8Gbps		
Maintenance interface	Gigabit ethernet		
Power requirements	100V AC to 240V AC, 50Hz/60Hz		
Power consumption (max.) *2	409 W	472 W	–
Heat load (max.)	1,395 BTU/Hr	1,610 BTU/Hr	–
Operating temperature	5°C to 35°C (41°F to 95°F)		
Operating humidity	20% to 80% (relative humidity)		
Weight *3	31 Kg	25 Kg	23 Kg
Dimensions (W, H, D) (excluding protrusions)	445 x 308 x 940 mm (17 5/8 x 12 1/4 x 37 1/8 inches)		

*1 Recording capacity depends on the usage environment. Actual recordable capacity may be less than indicated on the cartridge.
*2 The wattage values for the ODS-L30M/L60E are measured with the maximum number of ODSD380F drive units. ODS-L30M without drives is 179 W. The ODSL60E without drives is 12 W. Each drive max power is 115 W.
*3 Includes the media, drives and rack rails. Excludes the rack.

Sony's latest Optical Disc Archive, PetaSite Scalable Library, return to the traditional 42U rack form factor [45]. Sony achieves this scalability by deploying a "Master Unit" and up to 5 "Extension Units". The ODS-L30M forms the initial building block of the PetaSite modular library solution. It provides robotics for an entire rack with support for two drives and 30 cartridges. To increase archive performance and capacity, the ODS-L60E modular extension unit supports up to 4 additional drives and 61 cartridge slots. If additional archive capacity is all that is required, then the ODS-L100E can be added to provide an additional 101 slots of cartridge capacity. Note the

PetaSite Library uses the ODA optical drives, which read/write AD discs without removing them from their individual cartridges, essentially mitigating the dust issue presented in the Everspan system. Note this solution is also superior in terms of environmental tolerance, featuring a wide range of operating conditions that minimize energy cost from air conditioning. With 4 drives installed in an ODS-L60E library, a total of 750MB/s per rack transfer speed can be achieved.

Panasonic Freeze-ray

Panasonic has a system configuration similar to that of Sony's, with the key difference being that discs are handled and stored in magazines, and discs are written/read individually after being extracted from the magazines mechanically [46]. Concurrent read/write is possible with multiple drives per unit. This system was introduced in 2016 in collaboration with Facebook to store rarely accessed data cheaply for extended periods in data centers. The system was designed for BDXL initially, then later configured to accommodate the AD discs.

Panasonic's latest Data Archiver Writer contain optical units with typical transfer speeds of 54MB/s. However, the multi-unit writer is capable of simultaneously read from or write to 3 double-sided discs, or 6 disc-sides at a total of 324 MB/s.[xv] With LB-DH6 Data Archiver's 2-writer configuration, Freeze-ray data archivers can achieve a maximum transfer speed of 648 MB/s per rack. The Panasonic Freeze-ray is shown and specified in **Figure 56** and **Table 10**.

Figure 56. Panasonic's Freeze-ray LB-DH6 (left) and LB-DH7 (right) data archiver systems. [47]

Reproduced with Permission: Panasonic Holdings Corporation

[xv] Optical Disc Data Archiving: A New Age of Cold Data — The Storage Revolution Begins Now Whitepaper (panasonic.net)

Table 10. Freeze-ray specifications [48]

	LB-DH6 Data Archiver	LB-DH7 Data Archiver
Number of mountable magazines	532	532
Compatible Magazine Types	6TB (Gen 2 AD), 3TB (Gen 1 AD), 1.2TB (BD)	3TB (Gen 1 AD), 1.2TB (BD)
Capacity	3.19 PB	1.9 PB
Number of writer units (per rack)	2 unites	2 units
Total data transfer rate	648MB/s Read, 432MB/s Write	432 MB/s Read/Write
Host interface	SAS/iSCSI/FC (by server for Data Archiver control software)	
Command protocol	SCSI (MMC, SMC)	
Max Height when mounted in a 19-inch rack with EIA panels	46U	46U
Input power	**DC +24V, +12V**	

Technical Issues

Current optical media solutions face an array of technical issues that are being addressed by the industry. The commercially available products don't offer any long-term development roadmap that can keep up with the exponential growth of archival data needs.

Capacity

The latest Archival Disc (AD) features a per disc capacity of 500GB. A 6TB, 12-disc cartridge is comparable in form factor with LTO tape cartridges and LFF HDDs. This capacity is considerably less than HDDs and LTO tape cartridges. With the forecasted future capacity of 1TB/disc in the next few years without further development plans published, current commercial optical storage solutions will likely lag behind its magnetic counterparts.

Cost

Table 6 lists the approximate costs of various data archiving media. It is apparent that current optical technologies are not cost competitive from an up-front cost perspective. While lenient environmental requirements and longer remastering cycle partially offset the media price disadvantage, it still represents a large short-term commitment for enterprise clients to use optical storage as the archival medium of choice.

Speed

While read-write speed represents a challenge for nearly all digital storage technologies, a particular challenge facing the optical enterprise storage solution is time-to-first-bit (TTFB). This issue can be tackled by more efficient library systems that are able to retrieve and load discs into multiple drives concurrently and using metadata to deliver initial data quickly. Even with current TTFB, optical libraries deliver initial data much faster than tape.

Potential Solutions (Roadmap of Quantified Key Attributes Needs)

The prevailing approach building on the multilayer concept, centers around storage in 3rd dimension (and beyond) to maximize storage density both by increasing the areal density and the number of layers. Concepts for wavelength multiplexing and multilevel data schemes are being researched. In addition, concepts such as multiphoton writing and holographic storage are being investigated.

Opportunities for all these approaches mentioned involve a pivot from the old, consumer-based systems approach toward new concepts at the system level appropriate for data center deployment. A particularly important example is the possibility of a system implementation that separates the write hardware from the read hardware, a remote-write library. This approach, having read-only drives, takes best advantage of the WORM nature, assuring an “air gap” for maximum cyber security. In addition, it allows for using femtosecond laser technology for new multiphoton writing mechanisms. In such remote write libraries, the high cost of femtosecond lasers can be shared and amortized among, perhaps, millions of TB-scale media.

The high intensity of femtosecond lasers introduces new levels of multiphoton processes, such as multiphoton ionization. This highly nonlinear response opens the door for many more potential materials beyond the two-photon materials investigated years ago, with examples below. We note that these highly nonlinear processes present the opportunity for data marks far below the diffraction limit as only the central portion of the focal volume exceeds the writing threshold in the nonlinear regime.

Additionally, as with any nonlinear writing process, absorption occurs only in the focal volume so that materials transparent to the laser wavelength can be employed. The penetration and number of written layers then become limited by the ability of optical system.

Historical and current optical discs are based on confocal imaging optics for both write and read. The focused spot on the disc forms the data mark. For reading, the spot is imaged in a similar confocal arrangement onto a segmented detector for implementing both focus and tracking servos as well as the high-speed read channel. The time-dependent linear read channel is processed in a similar manner to HDD read channels. Panasonic and Sony’s great improvements in the read channel in developing the Archival Disc (as well as materials innovations) have resulted in bit error rates as good as any media.

Advances in spatial light modulators (SLM) have provided the opportunity to focus an array of writing spots all at once as described in the Piql section above. As described below, these devices provide the opportunity for using a single objective lens for a high intensity write laser in a broader array of materials. This has the potential to greatly increase the write speed as SLMs can have, at least, thousands to millions of pixels. The writing threshold of the material and laser power will limit how many marks can be written with a single laser pulse. The ability to store data at areal densities using SLMs that approach or exceed the current state of the art for traditional high NA writing, to our knowledge, has not yet been demonstrated.

On the read channel side, wide-field imaging can create a two- or three-dimensional images where modern high-speed imaging and image processing techniques could interpret the data. For example, machine learning has already been shown to provide high fidelity linear read channels in HDDs and might be adapted to higher dimensions with multidimensional neural networks [49]. Read channels based on arrays of bits at once opens the door for bitwise writing schemes to enter the page-wise read domain currently occupied by holographic storage.

While these new write and read channel opportunities seem quite exciting, no performance data for high-speed multi-dimensional write and read channels have been published, most notably, bit error rates. In current technologies, exquisite servo controls are necessary to provide the high fidelity write and read demanded by data storage and recovery and it is unclear what will be needed in these new approaches.

FUTURE MULTILAYER TECHNOLOGY

The limitations of the manufacturing process used to make current blue laser discs has stimulated intense development of higher areal density as in the Archival Disc. Folio Photonics Inc. (www.foliophotonics.com) is adapting a widespread multilayer polymer manufacturing process to both increase the number of layers and reduce costs [50]. The process is an adaptation of the co-extrusion manufacturing process to produce active layers with ~100 nm thickness and separated by ~10 micron thick buffer layers [51]. The process can produce dozens of layers "all at once" in a roll-to-roll process (**Figure 57**). As the process is easier to scale up than down, many layers are produced at much lower cost. Thus, the process promises to address both the cost and capacity problems of existing technologies.

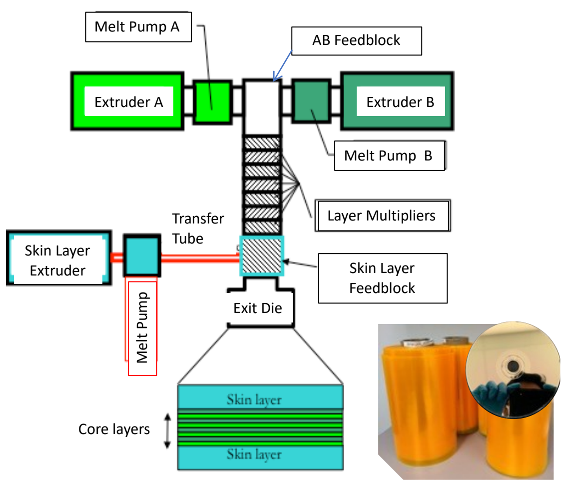


Figure 57. Folio Photonics manufacturing process and product. [52]

Reproduced with permission: Folio Photonics, Inc.

Both reflective and fluorescent data strategies have been demonstrated by a photothermal threshold scheme similar to existing technology. The number of layers is not limited by the manufacturing method, but by the laser power budget and spherical aberration correction in the objective lens of the optical pickup unit.

Unlike current multilayer Blu-ray technology, the Folio Photonics disc does not have the spiral tracking feature embossed on every layer. Such discs have been investigated by several companies in the past and are known as super multilayer discs (**Figure 58**) [53].

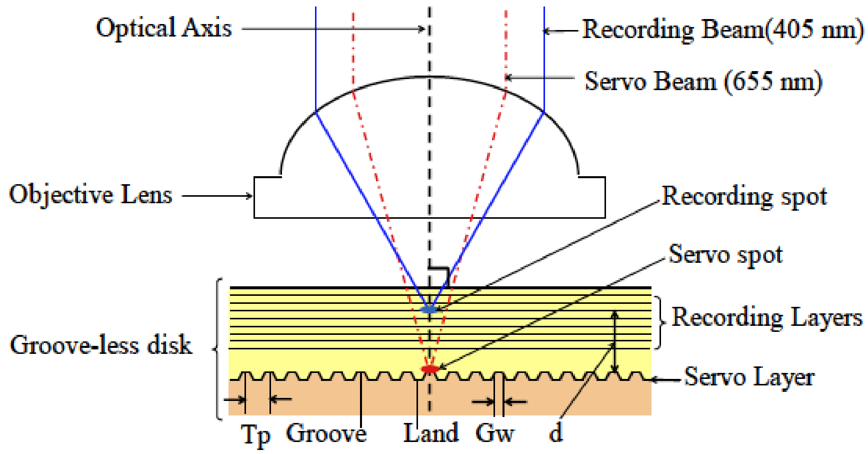


Figure 58. Guide layer tracking scheme. [54]

Copyright (2015) The Japan Society of Applied Physics, Reproduced with permission.

In these cases, tracking is carried out by guide layer tracking where an additional laser is focused on the disc substrate where the tracking pattern is embossed [55]. Folio's scheme for creating a focus error signal is shown in **Figure 59** [56].

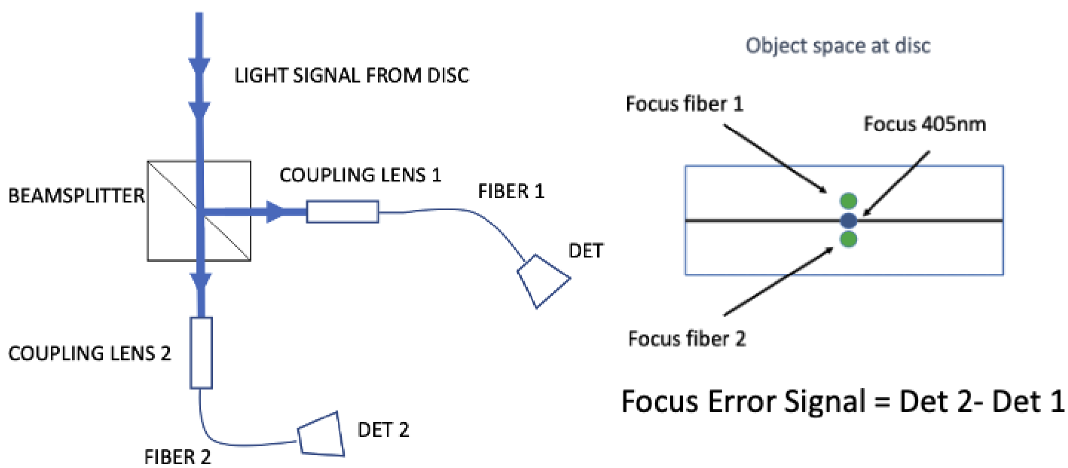


Figure 59. Scheme of Folio Photonic's confocal fiber focus error signal scheme for fluorescent signals. [57]

Reproduced with permission: Folio Photonics, Inc.

Folio Photonics has proved the feasibility of its optical pickup unit for writing and reading eight layers at speed. The company believes that its unique roll-to-roll co-extrusion fabrication of multilayer films, a process commonly used for other low-cost industry applications, enables the company to produce discs for sale at $5/TB, making the solution to be cost-competitive with LTO tapes and HDDs. With its first commercial product expected to launch in the general market in 2026 following testing and qualification, the company is in the process of creating a platform ecosystem with library companies and other partners.

Folio's current plans center on traditional confocal laser diode writing and linear read channels to take advantage of existing supply chains and know-how. The future areal density roadmap includes multilevel writing and wavelength multiplexing. Future implementations could take advantage of the system level opportunities and new optical and image processing technologies described above.

Folio's first product will be a two-sided 8-layer disc expected to hold from 800GB to 1TB of data, with a roadmap to add additional layers over time as shown in **Table 11**.

OPTICAL MASS DATA STORAGE TECHNOLOGY ROADMAP

Table 11. Optical Mass Data Storage Technology Roadmap – BDXL, Archival Disc, and Folio Photonic Disc

Technology	Metrics	Unit	2021	2023	2025	2027	2029	2031	2033	2035	2037
BDXL	Capacity	TB	0.2	0.2	No Further Roadmap Announced						
	Media Cost	$/TB	40	40							
	Data Transfer Rate (Max)	MB/sec	27	36							
	Layer Count (Double Side)	Layers	6	6[i]							
	Areal Density (per Layer)	GB/Layer	33	33							
Archival Disc	Capacity	TB	0.5	0.5	1[ii]			No Further Roadmap Announced			
	Media Cost	$/TB	33	30	30[iii]						
	Data Transfer Rate (Max)	MB/sec	375	375	750[iv]						
	Layer Count (Double Side)	Layers	6	6	6						
	Areal Density (per Layer)	GB/Layer	83	83	167						
Folio Photonic Disc	Capacity	TB			0.5-1.0	1.0-2.0	4	8	16	32	64
	Media Cost	$/TB			5	3	2	1	0.6	0.35	0.2
	Data Transfer Rate (Max)[v]	MB/sec			40	80	320	500	600	800	1000
	Layer Count (Double Side)	Layers			16	24	32	40	40	54	64
	Areal Density (per Layer)	GB/Layer			33	85	125	200	400	600	1000

[i]BD-Double-Sided Disc (BD-DSD) has no plan to expand to quad-layer per side
[ii]There is currently no set release date for the 3rd generation AD
[iii]Assumes constant rate of cost/TB reduction relative to reduction from 300GB AD to 500GB AD
[iv]Assumes no change in drive speed
[v]Assumes multiple optical pickup units per drive
Note: Folio Photonic Disc will utilize its 2nd gen technology starting in 2031

EMERGING CONCEPTS

In this section, new concepts in optical data storage are summarized, namely, multiphoton writing and holographic. These involve new twists on methods investigated in the last two decades. New concepts include the use of ultrafast lasers in remote-write libraries, where WORM media are written with specialized write-only drives that allow the high cost of ultrafast lasers to be amortized among many thousands (or more) media. Group 47 and Cerabyte use 1- and 2-d spatial light modulators to write many data marks at once and 2-d image analysis for reading. These concepts have been commercialized by Piql for data preservation which have less rigorous requirements for storage density and speed. Time will tell if these new concepts achieve the performance to replace the traditional bit-wise disc technology in data center applications, where bit-wise methods also develop apace.

MULTI-PHOTON

Project Silica (Glass)

A recent result at the University of Southampton has made it possible to store data in fused silica (i.e., quartz glass) [58]. By focusing a femtosecond laser inside a block of fused silica, a 3D nanostructure can be formed, which is a permanent change to the physical structure of the material via multiphoton ionization.

Microsoft Research is collaborating with the University of Southampton to develop a cloud-scale archiving system based on this technology [59]. The effort focuses on a design aimed at cloud-scale performance by addressing such issues as the entanglement of the write and read throughputs, the refresh cycle described above and constrained workloads by broadening the storage tier.

The use of silica has the potential benefit of true 3D optical storage, since the opacity of the quartz glass is significantly less than the relatively opaque layers of a classical optical disc, the effect of scattering and noise will be reduced, enabling multi-layer read/write. Due to the small nonlinear optical response of silica, an amplified femtosecond laser is required for writing single bits. The writing speed is thus given by the laser pulse repetition rate. Elongated marks exhibiting form birefringence can be polarization multiplexed within a single voxel. In addition, multilevel storage within a voxel has also been demonstrated. Together these are dubbed 5D optical storage. Seven bits per voxel have been demonstrated. This multibit voxel storage leads to high storage densities. This scheme is depicted in Figure 60.

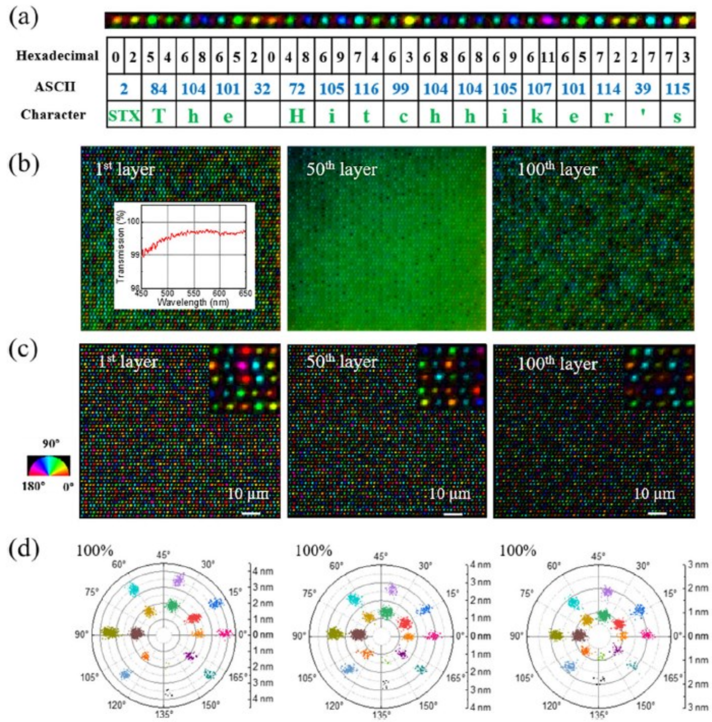


Figure 60. 5D storage of "The Hitchhiker's Guide to the Galaxy."
a) Illustration of data encoding and decoding. b) The birefringent images of data voxels of different layers. Inset is the transmission of 100-layer data in the visible range. c) The birefringent images after removing the background of (b). Insets are enlargements of small region (10 μm x 10 μm). d) Polar diagram of the measured retardance and azimuth of all voxels in (c). [60]

The permanent nature of quartz glass also allows the realization of ultra-long-term storage, since fused silica can retain its properties for hundreds of years while withstanding extreme conditions. The nanograting in fused silica has high thermal stability and high optical damage threshold. Project Silica indicates that its glass media can withstand temperatures over 1000 C, referring glass storage as the modern day "stone etching".

A recent video posted by Microsoft has revealed an interesting remote-write library system and described their technology in some depth [61]. The system shelves the raw silica media on passive storage panels. Silica media can be retrieved using shuttles, free-roaming robotics that can be added or removed from the system based on changing archival access demand. The library leverages an inherent air-gap design that eliminates the shuttles' ability to move written Silica media into writers. The decentralized and modular design of the library system mitigates the risk of large-scale incidents and safeguards archival data integrity.

An early related publication as well as a recent one indicate a slow writing speed below 100kB/sec [62], [63]. However, the recent video suggests advances at the drive and system levels have achieved "aggregate system level throughputs comparable to system-level tape archive deployments." We speculate that high-bandwidth spatial-light-modulators and high-rep rate lasers are factors in this speed increase. The video also indicates volumetric storage densities much higher than tape via 200-layer writes within the volume. For reading, machine learning is used to decode a 2-D array of data using a convolutional neural network. We look forward to learning more about the performance of the system.

Group 47 (DOTS)

A California-based start-up Group 47 (www.group47.com) has developed the Digital Optical Tape System (DOTS) that preserves data for more than 200 years. DOTS leverages a legacy Kodak-developed phase change media composed of a metallic alloy sputtered on a polyester-based film, which is immune to electromagnetic fields (including EMP), chemically inert, and will not oxidize and contains no chemical binders that can degrade hygroscopically. Data may be stored as binary dots and blanks, as well human-eye readable images. DOTS has experimentally demonstrated a 200 to 2000 year archival longevity and a -9°C to 66°C temperature tolerance [64].

DOTS's reflective encoding technique is based on the metallic alloy's phase change properties. Writing is a threshold photothermal process with a threshold of 0.3nJ/mm^2. Both the reading and writing processes are accelerated by novel paralleling and image encoding schemes [65]. The film can be encoded and read continuously using the optical guide tracks and various guide spots, which enables the medium to maintain data integrity and robustness against the effects of variable media velocity and even physical deformation (**Figure 61**). DOTS' write head uses a diffractive spatial light modulator (SLM) as part of the laser multiplier that enables simultaneous writing of over 10,000 spots, allowing a single pass to fill up the tape. Writing is carried out using a 532 nm wavelength laser. Lines of data are read by imaging the data using an oversampled linear detector array.

While DOTS is yet to be commercialized, several successful prototype projects, including one awarded by the Central Intelligence Agency (CIA), have demonstrated the solution successful proof-of-concept. The library will use the remote write library concept. Group 47 also has plans to release open source read-only units and license them to be manufactured royalty-free to accelerate adoption. Group 47 indicates that commercial shipments in the next two years with a 1.2TB Native capacity per tape cartridge and an impressive transfer speed of 1GB/s.

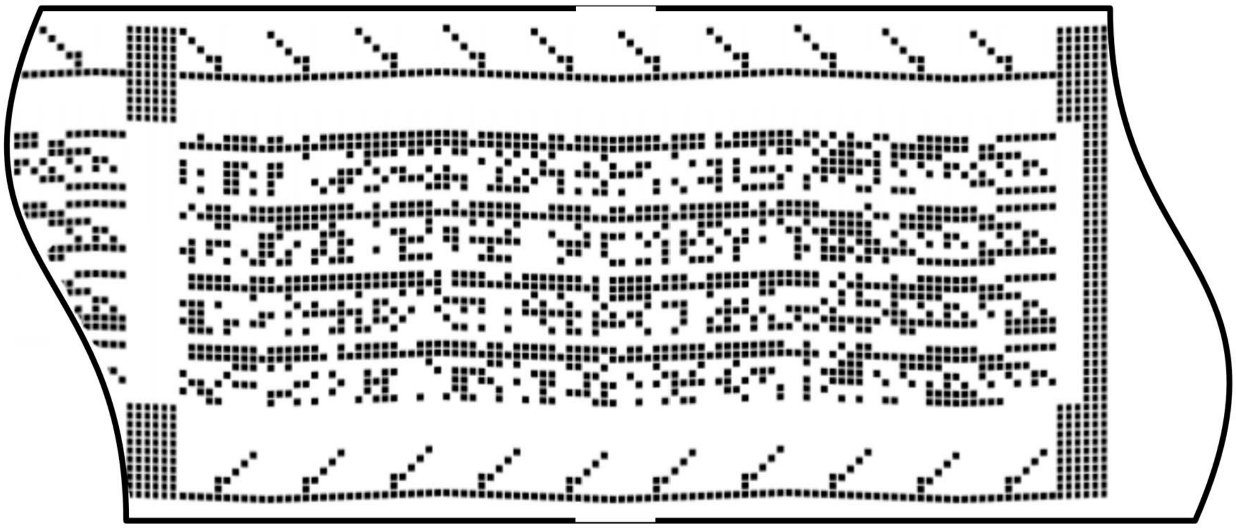


Figure 61. Sample DOTS data encoding mechanism. [66]

Reproduced with permission: Group 47, Inc.

Cerabyte[xvi]

Cerabyte (www.cerabyte.com), founded in 2020, is developing a patented [67] data storage medium featuring a sputtered deposited extremely durable ceramic nano layer (10 nm thick) with a broad absorption spectrum, a "grey ceramic", which allows ultra-fast write with threshold as low as 0.1 nJ per bit. The ceramic is deposited on both sides of a flexible ultra-thin planar substrate with only 100 µm thick foldable glass [68] or 10 µm thick ribbon glass [69] for potential tape development. These substrates and the coating of a ceramic nano layers leverage existing display glass production capacities of today 350 million m2 per year [70]. Thus, the company expects to achieve media cost below $ 1 per TB by 2030.

Accelerated aging tests at temperatures of -273 °C to 500 °C indicate potential storage lifetime in the millennia range. Furthermore, the data is not corrupted even when exposed to electromagnetic pulses, UV and gamma rays.

Data is written encoded in an array of data matrices by using a 2-D digital micro mirror device (DMD) with up to 2 million elements simultaneously written by femtosecond laser pulses in the UV spectrum with a rep rate of several kHz (**Figure 62**). This implies a writing speed in excess of 1 GB/s with less than 1 W average power.

[xvi] This section authored by Cerabyte (Ceramic Data Solutions Holding GmbH) CEO, Christian Pflaum

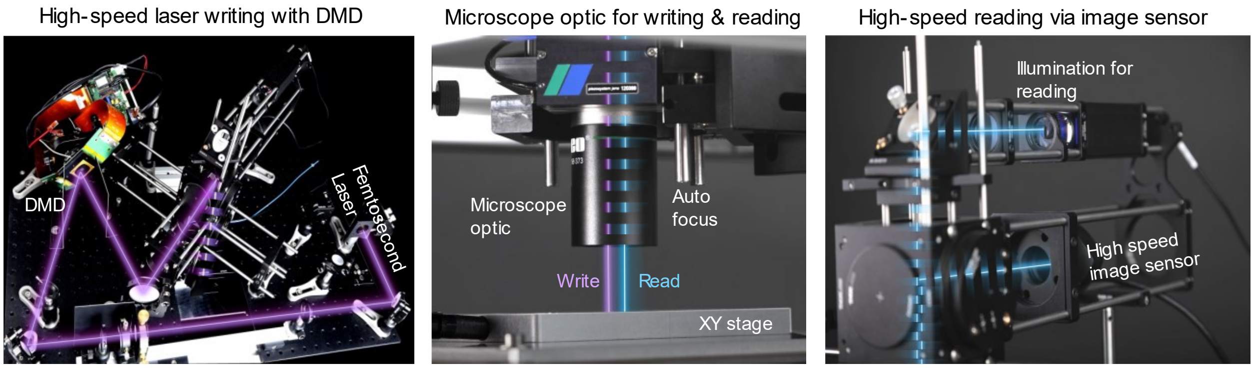


Figure 62. Cerabyte writing scheme

Reading is performed at GB/sec rate using high speed image sensors, UV illumination and parallel, high speed image processing for decoding. The high-resolution images are captured at more than 500 fps followed by parallel processed by a FPGA, which produces a data stream using the 2-D image while applying conventional error correction methods in a second processing step. Both reading and writing is carried out across the substrate by scanning the microscope optics using high-speed XY stages kept in focus using a piezo driven auto focus system. This setup enables random access.

Plans call for hundreds of 9 by 9 cm media sheets to be stacked in individual cartridges to reduce storage volume as shown in **Figure 63**. Cerabyte uses the form factor of mainstream magnetic tape cartridges. The media access scheme enables random access enabling a faster time to first bit than tape.

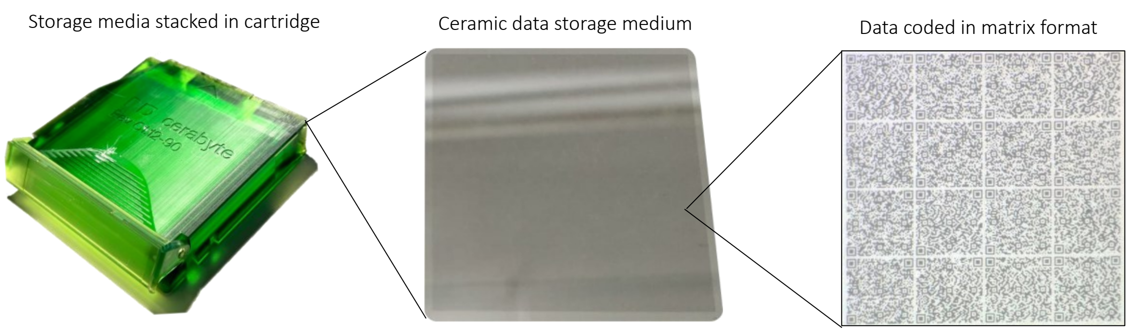


Figure 63. Cerabyte's storage concept

Cerabyte employs a commercially available library unit utilizing a remote write architecture. The library will locate and retrieve the cartridge, then unload and unstack the substrates for positioning the addressed substrate in the optical unit.

Cerabyte indicates that a demo system with a single write & read head unit achieving 100 MB/s write/read speeds and a storage capacity of 1PB per 19" rack will be developed in 2023. The first product for corporate archiving systems is scheduled to be launched in 2024 with 500 MB/s write/read speeds and a capacity of 5 PB/rack up to ten 19" racks. In 2025 a 10-30PB rack system for cloud data centers will be launched with 1 GB/s+ write/read speeds, which is projected to increase over time (**Figure 64**) and will achieve capacities and bandwidth attractive for hyperscaler use cases by the end of the decade.

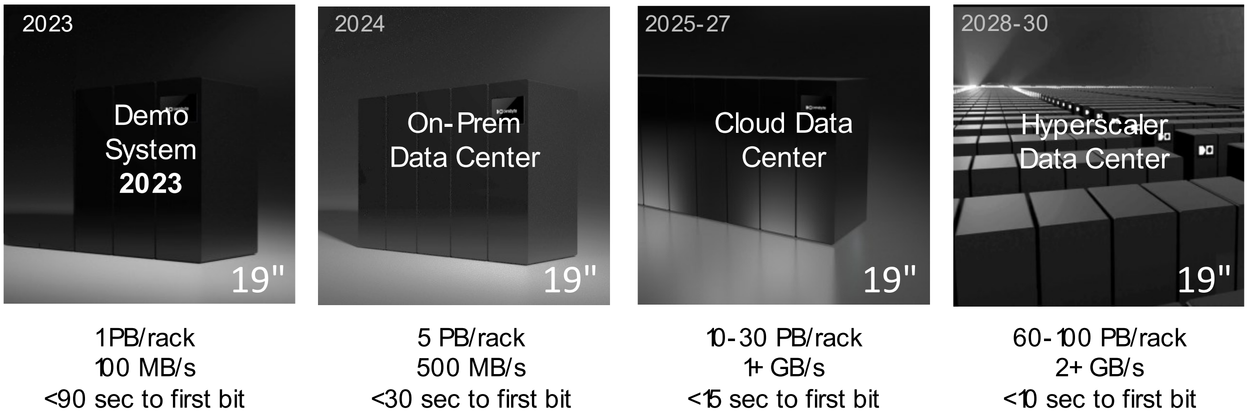


Figure 64. Cerabyte's product roadmap of enterprise archiving systems

HOLOGRAPHIC

While holographic data storage was proposed as early as the 1960s. By the 2000s, leading research institutions including Aprilis and Bell Labs, have failed to deliver commercial success despite significant research progress. Brief commercial attempts such as one by InPhase Technologies, which released a $180, 300GB holographic storage disc (HSD) in 2007 [71], all failed to capture the market due to their high costs [72].

The technology provides a series of benefits, including high data density, massively parallel write and read, fast access time of <50 us, fast transfer rate of ~ 1 Gbit/s, air gap and WORM or rewritable formats [73]. Despite these technological advantages, HSD development remains in the R&D stage today, decades after the technology's introduction. While many successful prototypes have been built, none have gained significant commercial traction.

Holographic data storage (HDS) allows multiple holograms to be written in the same fixed volume of rather thick media (200+ um), with theoretical recording density limit on the order of tens of Tb/cm^3. Data recorded holographically are read back in "pages" instead of bit streams. HDS is a volumetric technique, making its density proportional to $1/l^3$, while for optical discs the density is proportional to $1/l^2$. The storage density is a function of the number of holograms multiplexed into a volume of the recording media and principally determined by its refractive index contrast and media thickness.

A wide range of materials was investigated as holographic media, including photorefractive crystals, photorefractive organic materials, photo-addressable polymers, photochrome systems, photopolymers, systems with electrocyclic ring closure, etc. Multiple materials properties such as dimensional stability (resistance of the material to shrinkage or expansion during the recording process and thermal changes), optical quality, scattering level, linearity and volatility affect the fidelity of the recording and read-out process. The recording and read-out transfer rates depend on the photosensitivity and the diffraction efficiency.

For the WORM media format, photopolymers are preferred materials that can be easily manufactured in production scale. However, their shrinkage during writing is a critical issue that should be solved before industrial applications is possible.

A lot of research efforts were made for developing writing and reading holographic systems. Several multiplexing technologies were proposed the "page-wise" approach, including two-beam angle multiplexing [74], coaxial multiplexing [75], and coaxial shift multiplexing [76]. Using these methods several prototypes were built to demonstrate drive system capabilities to handle dimensional changes of the materials due to thermal expansion.

PROJECT HSD (HOLOGRAPHIC)

Announced in 2020, as one of the two cloud-focused storage projects led by Microsoft Research, Project HSD explores the potential of Holographic storage, designing mechanical-movement-free, high-endurance cloud storage that is both performant and cost-effective [77], [78].

Holographic data formats can be similar to that of 2-dimensional QR codes, coded inside a $LiNbO_3$ crystal. By leveraging two beams: the data beam and the reference beam, the 2-D codes can be encoded as interference patterns. By changing the angle of the reference beam, one can record multiple sets of interference patterns within the same block of crystal. Data can then be read using the reference beam that recreates the interference pattern, before capturing the 2-D code using a camera. The crystal can be then erased using UV light and be reused indefinitely to store more data.

On the physical level, the Iron ions doped in the crystal add an additional donor level and a deep trap state to the energy levels of the crystal. When a region is exposed to the bright part of the interference pattern, the extra electrons are excited, moving to the conduction band, before decaying preferentially to the deep iron trap level [79].

Their focus on scalable optical systems, storage systems and machine learning are important new directions that can move holographic enterprise optical archiving forward. We speculate that Project HSD's approach addresses previous challenges.

CHALLENGES (CRITICAL ISSUES)

Currently, most of the enterprise archival market is dominated by LTO magnetic tape and HDD. Magnetic tapes arguably offer the lowest cost/GB and a relatively long lifetime, positioning them to be an effective solution for "deep archival" use cases. However, its long TTFB and the lack of random access prevents it from delivering value in more active archival applications. LTO also suffers from a relatively short lifetime and an expensive and not always backward-read-compatible drive. HDD offers performance that exceeds the need of archival applications but suffers from a high operating cost and short lifespan, which results in frequent remastering, further increasing TCO. The challenges for optical storage media to compete with magnetic tape and HDD in the near to medium term can be categorized into three aspects – capacity, cost, and read/write speed.

CAPACITY

While capacity technically can be expressed as a function of cost, the value of a higher GB/volume is self-explanatory. High data density can be an insurmountable requirement for large data center deployments as space and other operating constraints are often pre-determined factors regardless of budget.

Layered optical discs traditionally drive capacity growth by increasing both the number of layers per disc and areal density. However, spherical aberrations and light-attenuation would eventually limit the number of layers that can be efficiently recorded and retrieved. However, multiphoton writing/reading demonstrations can increase volumetric and areal density. In addition, 3D approaches such as multiphoton in non-layered media and holographic storage promise high volumetric density.

COST

The capacity challenge can be addressed as an aspect of the broad cost challenges. LTO-8 tapes are priced at around $5/TB, which is an order of magnitude cheaper than the optical competitors, including the BDXL and Sony-Panasonic Archival discs. Optical disc is not competitive compared to the much more expensive HDD either, despite the slowing of HDD capacity improvement and cost reduction. With a shrinking market, optical disc manufacturers are benefitting less from the economy of scale. To overcome this challenge, optical media must move to reduce the cost in media manufacturing that utilizes more efficient processes, lower-cost materials, and/or other cost-cutting innovations. System-level costs for enterprise data storage including drives and libraries should be competitive with the competition with the caveat that optical data storage approaches have an intrinsically lower total cost of ownership owing to their long life, high energy efficiency, and low carbon footprint.

SPEED

The known approaches to optical data storage involves libraries of platters and spooled media. Platter libraries have a much lower time to first bit of any spooled media, likely in the seconds to tens of seconds. While the read/write speed has been a challenge for most storage media due to the rapid growth of data size, drive and system-level innovations have largely overcome this hurdle on the enterprise scale by implementing multi-channel approaches such as storing across multiple drives within a disc library, or as Sony has, by creating a drive with multiple pickup units. Machine learning using neural networks for reading has the potential to substantially increase reading speed by reading in multiple dimensions. In addition, recent results using machine learning in the data channel of hard drives show considerable promise improved performance [80], [81]. Page-wise reading in holographic storage would also result in high throughput.

SUMMARY & KEY POINTS

- Optical storage media is pivoting from the historical consumer media distribution focus to the enterprise/institutional archival storage use case.
- To expand capacity while reducing cost, new optical technologies will continue to exploit the 3rd dimension and beyond.
- Robust library systems designed for optical media are being engineered to meet the robust requirements of enterprise-level storage needs.
- Optical media's low maintenance/operating energy cost and remastering frequency will realize a natural advantage in the sustainability-conscious data archive landscape.
- Optical WORM technologies the air-gap offer advantages in cybersecurity.

- Optical technologies present the lowest energy consumption, both intrinsically and in data center environment control.
- Optical technologies promise significant greenhouse gas reduction.
- Critical challenges facing optical data storage are lowering initial cost, expanding capacity, increasing speed, and error management.
- Non-disc, novel technologies will likely deliver breakthroughs in capacity, cost, and speed to the data storage industry at-large, though likely not in the near-term.
- The possibility of remote-write libraries, femtosecond lasers and high-speed display and imaging technologies provide new opportunities for enterprise optical data storage.

CONCLUSIONS & RECOMMENDATIONS

The demand for archival data storage is growing rapidly due to the growth of permanent access driven by demands of data science, internet of things, and requirement for permanent storage. At the same time, the pace of capacity growth and cost reduction in incumbent magnetic technologies are slowing due to the approach of physical limits implied by Kryder's law. In addition, the oligopoly of both consumers and producers are stressing the data storage market. Optical archiving technologies are now emerging based on multilayer disc libraries, but still suffer from high cost and low capacity. New approaches to multilayer manufacturing aimed at lowering cost and increasing capacity are in development, while other approaches based on holographic storage and multiphoton writing are being researched.

For too many years, R&D in the development of optical data storage has focused too much on media capacity while downplaying the importance of speed and cost. In the future, the challenges of developing drives and libraries need to be considered from day one. Partnerships among media, drive, and library system players within and between organizations will be key to widespread commercialization of optical archiving technologies and the strong value proposition they provide.

REFERENCES

1 "IFPI GLOBAL MUSIC REPORT 2023". IFPI. https://globalmusicreport.ifpi.org/ (accessed September 1st. 2023)

2 "IFPI GLOBAL MUSIC REPORT 2023". IFPI. https://globalmusicreport.ifpi.org/ (accessed September 1st. 2023) web.archive.org/web/20150518082128/http:/5eshs.hpdst.gr/abstracts/139""The history of ideas "the optical disc as a "unique" carrier of information in the systems management". European Society of the History of Science. Archived from the original on 2015-05-18.

3 "The history of ideas "the optical disc as a "unique" carrier of information in the systems management". European Society of the History of Science. Archived from the original on 2015-05-18.

4 E. Hwu & A. Boisen, "Hacking CD/DVD/Blu-ray for Biosensing," *ACS Sens.* 2018, 3, 7, 1222–1232, July 2018, doi: 10.1021/acssensors.8b00340.

5 "U.S. Sales Database." RIAA. https://www.riaa.com/u-s-sales-database/ (accessed December 1st, 2022)

6 S. Whitten, "The death of the DVD: Why sales dropped more than 86% in 13 years." CNBC. https://www.cnbc.com/2019/11/08/the-death-of-the-dvd-why-sales-dropped-more-than-86percent-in-13-years.html (accessed December 1st, 2022)

7 M. Siegler, "Yep, Apple Killed The CD Today." TechCrunch. https://techcrunch.com/2010/10/20/a-compact-death/ (accessed December 1st, 2022)

8 N. Statt, "Sony announces PlayStation 5 Digital Edition with no disc drive." The Verge. https://www.theverge.com/2020/6/11/21288493/ps5-playstation-5-digital-edition-no-disc-drive-hardware-specs-price-sony (accessed December 1st, 2022)

9 "U.S. Sales Database." RIAA. https://www.riaa.com/u-s-sales-database/ (accessed December 1st, 2022)

10 S. Whitten, "The death of the DVD: Why sales dropped more than 86% in 13 years." CNBC. https://www.cnbc.com/2019/11/08/the-death-of-the-dvd-why-sales-dropped-more-than-86percent-in-13-years.html (accessed December 1st, 2022)

11 "ECMA-167." Ecma International. https://www.ecma-international.org/publications-and-standards/standards/ecma-167/ (accessed September 1st, 2023)

12 2019 iNEMI Roadmap - Mass Data Storage. iNEMI. (accessed December 1st, 2022)

13 C. Walter, "Kryder's Law: The doubling of processor speed every 18 months is a snail's pace compared with rising hard-disk capacity, and Mark Kryder plans to squeeze in even more bits." https://www.scientificamerican.com/article/kryders-law/ (accessed September 1st, 2023)

14 "Optical Technology: Optical Data Archiver freeze-ray series." Panasonic Global. https://panasonic.net/cns/archiver/optical_technology/index.html (accessed December 1st, 2022)

15 "Optical Technology: Optical Data Archiver freeze-ray series." Panasonic Global. https://panasonic.net/cns/archiver/optical_technology/index.html (accessed December 1st, 2022)

16 "White Paper: Archival Disc Technology." 1st Edition. Panasonic.net. July 2015 https://panasonic.net/cns/archiver/pdf/E_WhitePaper_ArchivalDisc_Ver100.pdf (accessed December 1st, 2022)

17 "White Paper: Archival Disc Technology." 2nd Edition. Panasonic.net. January 2020 https://panasonic.net/cns/archiver/pdf/E_WhitePaper_ArchivalDisc_2nd_Edition.pdf (accessed December 1st, 2022)

18 "White Paper: Archival Disc Technology." 1st Edition. Panasonic.net. July 2015 https://panasonic.net/cns/archiver/pdf/E_WhitePaper_ArchivalDisc_Ver100.pdf (accessed December 1st, 2022)

19 "White Paper: Archival Disc Technology." 2nd Edition. Panasonic.net. January 2020 https://panasonic.net/cns/archiver/pdf/E_WhitePaper_ArchivalDisc_2nd_Edition.pdf (accessed December 1st, 2022)

20 J. Moore, "The Escalating Challenge of Preserving Enterprise Data." Further Market Research. August 2022 https://asset.fujifilm.com/master/americas/files/2022-08/5082fc03fc44d80b691c87a1a96febd5/Furthur_Market_Research_WP_080322_FINAL.pdf (accessed December 1st, 2022)

21 "White Paper: Blu-ray Disc Format - 1. B Physical Format Specifications for BD-R." 5th Edition. October 2010 https://blog.ligos.net/images/The-Reliability-Of-Optical-Disks/BD-R_physical_specifications-18326.pdf (accessed December 1st, 2022)

22 "Format Specification - R3 Format Specification (BDXL™)." blu-raydisc.info. https://blu-raydisc.info/format-spec/r3-spec.php (accessed December 1st, 2022)

23 "Blu-ray Disc Association Announces New Double-Sided Disc Specification for 'Big Data' Storage." blu-raydisc.com. 20140818-Blu-ray-Disc-Association-Announces-New-Double-Sided-Disc-Specification-for-Big-Data-Storage.pdf (blu-raydisc.com) (accessed December 1st, 2022)

24 P. Lilly, "Pioneer Just Made A New Optical Disc Drive For PCs And Yes It's 2022." HotHardware. https://hothardware.com/news/pioneer-optical-disc-drive-pc (accessed December 1st, 2022)

25 "White Paper: Blu-ray Disc Format - 1. B Physical Format Specifications for BD-R." 5th Edition. October 2010 https://blog.ligos.net/images/The-Reliability-Of-Optical-Disks/BD-R_physical_specifications-18326.pdf (accessed December 1st, 2022)

[26] “White Paper: Archival Disc Technology.” 1st Edition. Panasonic.net. July 2015 https://panasonic.net/cns/archiver/pdf/E_WhitePaper_ArchivalDisc_Ver100.pdf (accessed December 1st, 2022)

[27] “White Paper: Archival Disc Technology.” 2nd Edition. Panasonic.net. January 2020 https://panasonic.net/cns/archiver/pdf/E_WhitePaper_ArchivalDisc_2nd_Edition.pdf (accessed December 1st, 2022)

[28] “White Paper: Archival Disc Technology.” 2nd Edition. Panasonic.net. January 2020 https://panasonic.net/cns/archiver/pdf/E_WhitePaper_ArchivalDisc_2nd_Edition.pdf (accessed December 1st, 2022)

[29] “White Paper: Archival Disc Technology.” 1st Edition. Panasonic.net. July 2015 https://panasonic.net/cns/archiver/pdf/E_WhitePaper_ArchivalDisc_Ver100.pdf (accessed December 1st, 2022)

[30] “White Paper: Archival Disc Technology.” 2nd Edition. Panasonic.net. January 2020 https://panasonic.net/cns/archiver/pdf/E_WhitePaper_ArchivalDisc_2nd_Edition.pdf (accessed December 1st, 2022)

[31] “Long-term Data Storage.” piql. https://piql.com/services/long-term-data-storage/ (accessed September 1st, 2023

[32] “Technology.” piql. https://piql.com/about/technology/ (accessed September 1st, 2023)

[33] M. Smith, “How To Back Up Your Digital Photos... To Film.” Fstoppers. https://fstoppers.com/originals/how-back-your-digital-photos-film-537562 (accessed December 1st, 2022)

[34] “Arctic World Archive.” Wikipedia. https://en.wikipedia.org/wiki/Arctic_World_Archive (accessed September 1st, 2023)

[35] “Suzhou NETZON Information Storage Technology.” hit-netzon.com.cn. https://www.hit-netzon.com.cn/HDL10368-en.html (accessed December 1st, 2022)

[36] “Suzhou NETZON Information Storage Technology.” hit-netzon.com.cn. https://www.hit-netzon.com.cn/HDL10368-en.html (accessed December 1st, 2022)

[37] “Commercial Products.” Amethystum.com. https://www.amethystum.com/article/10171.html (accessed December 1st, 2022)

[38] “Commercial Products.” Amethystum.com. https://www.amethystum.com/article/10171.html (accessed December 1st, 2022)

[39] “Zerras.” ICEBOX. https://www.zerras.com/icebox/ (accessed December 1st, 2022)

[40] “Disc Library Systems.” Kintronics. https://kintronics.com/disc-library-systems/ (accessed December 1st, 2022)

[41] M. Smolaks, “Optical disc as the future of data archiving.” DCD. https://www.datacenterdynamics.com/en/analysis/optical-disc-as-the-future-of-data-archiving/ (accessed December 1st, 2022)

[42] “Optical Disc Archive Generation 3.” Sony Pro. https://pro.sony/s3/2020/05/28164824/Sony_ODA_Generation3_Brochure-PDF.pdf (accessed September 1st, 2023)

[43] “Optical Disc Archive Generation 3.” Sony Pro. https://pro.sony/s3/2020/05/28164824/Sony_ODA_Generation3_Brochure-PDF.pdf (accessed September 1st, 2023)

[44] “Optical Disc Archive Generation 3.” Sony Pro. https://pro.sony/s3/2020/05/28164824/Sony_ODA_Generation3_Brochure-PDF.pdf (accessed September 1st, 2023)

[45] “Cold Data Storage Technology.” Sony Pro. https://pro.sony/ue_US/technology/optical-disc-archive (accessed December 1st, 2022)

[46] “光ディスクデータアーカイブシステム - パナソニック コネクト” panasonic.com. (accessed September 1st, 2023); “Products | Optical Data Archiver freeze-ray series.” Panasonic Global. https://panasonic.net/cns/archiver/product/#Outline (accessed December 1st, 2022)

[47] “光ディスクデータアーカイブシステム - パナソニック コネクト” panasonic.com. (accessed September 1st, 2023); “Products | Optical Data Archiver freeze-ray series.” Panasonic Global. https://panasonic.net/cns/archiver/product/#Outline (accessed December 1st, 2022)

[48] “光ディスクデータアーカイブシステム - パナソニック コネクト” panasonic.com. (accessed September 1st, 2023); “Products | Optical Data Archiver freeze-ray series.” Panasonic Global. https://panasonic.net/cns/archiver/product/#Outline (accessed December 1st, 2022)

[49] Y. Qin and J. -G. Zhu, "Automatically Resolving Intertrack Interference With Convolution Neural Network Detection Channel in TDMR," *IEEE Transactions on Magnetics*, vol. 57, no. 2, pp. 1-6, Feb. 2021, Art no. 3100306, doi: 10.1109/TMAG.2020.3017373.

[50] Folio Photonics, Inc.

[51] C. Mellor, “Folio Photonics archive optical disk technology is real.” Blocks and Files. https://blocksandfiles.com/2022/08/31/folio-photonics-archive-optical-disk-technology-is-real/ (accessed December 1st, 2022)

[52] Folio Photonics, Inc.

[53] Y. Tanaka, T. Ogata, & S. Imagawa, “Decoupled direct tracking control system based on use of a virtual track for multilayer disk with a separate guide layer,” *Jpn. J. Appl. Phys.* 54 09MB03, August 2015, doi: 10.7567/JJAP.54.09MB03.

[54] Y. Tanaka, T. Ogata, & S. Imagawa, “Decoupled direct tracking control system based on use of a virtual track for multilayer disk with a separate guide layer,” *Jpn. J. Appl. Phys.* 54 09MB03, August 2015, doi: 10.7567/JJAP.54.09MB03.

[55] M. Ogasawara, K. Takahashi, M. Nakano, M. Inoue, A. Kosuda, and T. Kikukawa, “Sixteen-Layer Write Once Disc with a Separated Guide Layer,” *Jpn. J. Appl. Phys.* 50 09MF01, September 2011, doi: 10.1143/JJAP.50.09MF01.

[56] Patent: US 11,456,010 B2: SYSTEMS AND METHODS FOR INCREASING DATA RATE AND STORAGE DENSITY IN MULTILAYER OPTICAL DISCS

[57] Patent: US 11,456,010 B2: SYSTEMS AND METHODS FOR INCREASING DATA RATE AND STORAGE DENSITY IN MULTILAYER OPTICAL DISCS

[58] “Project Silica.” Microsoft. https://www.microsoft.com/en-us/research/project/project-silica/ (accessed September 1st, 2023)

[59] P. Anderson et al., “Glass: A New Media for a New Era?” *10th USENIX Workshop on Hot Topics in Storage and File Systems* (HotStorage 18) July 2018.

[60] H. Wang et al., “100-Layer Error-Free 5D Optical Data Storage by Ultrafast Laser Nanostructuring in Glass,” *Laser & Photonics Reviews*, 16(4), pp. 2100563, January 2022, doi: 10.1002/lpor.202100563.

[61] “Research talk: Storing data for millennia.” Microsoft Research. Youtube. https://youtu.be/V7L_wdEuQXs (accessed September 1st, 2023)

[62] J. Zhang et al., “Seemingly Unlimited Lifetime Data Storage in Nanostructured Glass,” *Phys. Rev. Lett.* 112, 033901, January 2014, doi: 10.1103/PhysRevLett.112.033901.

[63] A. Jain et al., “Optimization of Multi-Layer Data Recording and Reading in an Optical Disc,” *Photonics* 9(10), 690. September 2022, doi: 10.3390/photonics9100690.

[64] Group 47

[65] US Patents: US 9,208,813 B2, US 9,508,376 B2, US 9,640,214 B2, US 10,033,961 B2, US 10,067,697 B2

[66] Group 47

67 Patents: WO2021028035 Ceramic data storage media; WO2022033800 Data recording on ceramic data storage media: WO2022194354 Super resolution reading of ceramic data storage media

68 "Foldable Glass | Ultra-thin Bendable Glass Technology." Corning. https://www.corning.com/worldwide/en/innovation/materials-science/glass/Foldable-Glass.html (accessed September 1st, 2023)

69 "Glass-ribbon." Nippon Electric Glass Co., Ltd. https://www.neg.co.jp/en/product/ep/glass-ribbon (accessed September 1st, 2023)

70 "Display production capacity by type 2016-2025." Statista. https://www.statista.com/statistics/1057441/display-panel-production-capacity-type/ (accessed September 1st, 2023)

71 P. Miller, "InPhase 300GB holographic storage solution out the door." Engadget. https://www.engadget.com/2007-02-13-inphase-300gb-holographic-storage-solution-out-the-door.html (accessed September 1st, 2023)

72 P. Shread, "Holographic Storage Appears." Enterprise Storage Forum. https://www.enterprisestorageforum.com/hardware/holographic-storage-appears/ (accessed September 1st, 2023)

73 S. S. Orlov, E. Bjornson, W. Phillips, Y. Takashima, X. Li and L. Hesselink, "High transfer rate (1 Gbit/sec) high-capacity holographic disk digital data storage system," *Conference on Lasers and Electro-Optics (CLEO 2000). Technical Digest. Postconference Edition. TOPS Vol.39 (IEEE Cat. No.00CH37088),* San Francisco, CA, USA, 2000, pp. 190-191, doi: 10.1109/CLEO.2000.906896.

74 "A matter of format," Nature Photon 2, 401, July 2008, doi: 10.1038/nphoton.2008.118.

75 P.Koppa et al., Multiphoton and Light Driven Multielectron Processes in Organics: New Phenomena, Materials and Applications (Eds.: F. Kajzar, M.V. Agranovich), Kluwer, Dordrecht, 2000

76 K. Tanaka et al., "High Density Recording of 270 Gbit/in.^{2} in a Coaxial Holographic Recording System," *Jpn. J. Appl. Phys.* 47 5891, July 2008, doi: 10.1143/JJAP.47.5891.

77 "How does holographic storage work?" Microsoft Research. YouTube. https://www.youtube.com/watch?v=4EADwGV5Gv8&t=20s&ab_channel=MicrosoftResearch (accessed December 1st, 2022)

78 "Project HSD: Holographic Storage Device for the Cloud." Microsoft Research. https://www.microsoft.com/en-us/research/project/hsd/ (accessed December 1st, 2022)

79 "The physics of hologram formation in iron doped lithium niobate." Microsoft Research. YouTube. https://www.youtube.com/watch?v=CfME3e7aSNk&t=2s (accessed December 1st, 2022)

80 A. Aboutaleb et al., "Deep Neural Network-Based Detection and Partial Response Equalization for Multilayer Magnetic Recording," *IEEE Transactions on Magnetics*, vol. 57, no. 3, pp. 1-12, March 2021, Art no. 3101012, doi: 10.1109/TMAG.2020.3038435.

81 Y. Qin and J. -G. Zhu, "Automatically Resolving Intertrack Interference With Convolution Neural Network Detection Channel in TDMR," I*EEE Transactions on Magnetics*, vol. 57, no. 2, pp. 1-6, Feb. 2021, Art no. 3100306, doi: 10.1109/TMAG.2020.3017373.

DNA Data Storage

SITUATION ANALYSIS

The storage technologies outlined in the other chapters of this roadmap will continue to evolve and scale, but the reality is that the world is attempting to digitize unprecedented amounts of information. This information can be valuable if mined, stitched together, or otherwise searched and analyzed; however, the cost of storing the massive amount of data associated with this information is beginning to overwhelm the ability to pay for it using conventional storage technologies, and a cost-effective scaling path using traditional storage technologies remains unclear.

Further, the operational costs of refreshing data, or creating copies, using existing storage technologies is becoming prohibitive, with the refresh of some large archives needing to start, or nearly so, by the time the previous refresh finishes. These factors are, in turn, causing potentially valuable data, and even the potential for newly discovered knowledge, to be thrown away. This lost opportunity is shown as the "Zone of Potential Insufficiency" (**Figure 65**).

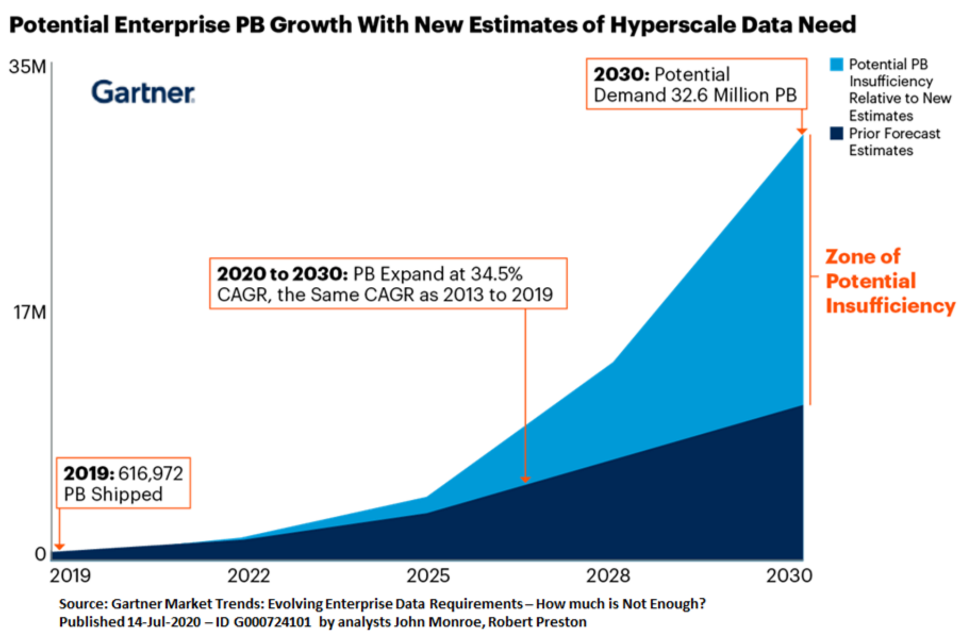


Figure 65. The Zone of Potential Insufficiency[1]

All of these factors are leading system designers to look for new storage technologies which can sustain the capacity, access flexibility, and TCO needed for the massive wave of digitization, and one of the leading technologies being considered is synthetic DNA.[i]

---

[i] For background on DNA data storage, see two presentations from the "DNA Memories" tutorial at the 2023 IEEE 15th International Memory workshop (IMW 2023): 1) "Integrating biomolecules and semiconductors to build a data storage system", Andres Fernandez, Twist Bioscience; and 2) "DNA Sequencing for Data Storage", Boyan Boyanov, Illumina. Also see "Preserving our Digital Legacy: An Introduction to DNA Storage, DNA Data Storage Alliance."

DNA (**Figure 66**) is a potentially compelling storage medium due to its ~1bit/nm^3 bit size (**Figure 67**), the fact that it is very stable at room temperature if kept dry, thus potentially avoiding fixity checks and media technology migration, and that the medium is decoupled from the read/write device, meaning it can always be read. This combination of properties enables the potential of the proverbial "datacenter in a shoebox"; in other words, compelling TCO.

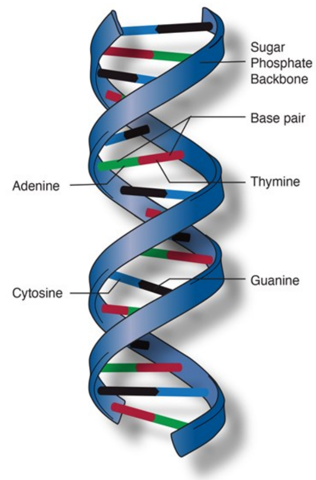


Figure 66. DNA Double Helix: National Human Genome Research Institute (NHGRI)

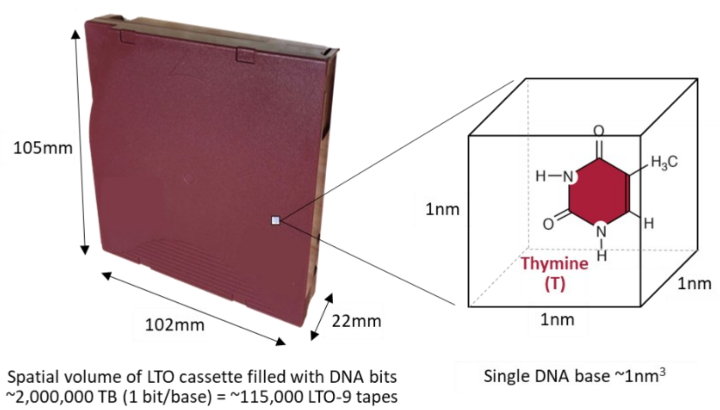


Figure 67. DNA Density; Preserving our Digital Legacy: An Introduction to DNA Data Storage; DNA Data Storage Alliance[4]

While an area of active research and development, DNA data storage (DDS) is not generally commercially viable today and thus does not lend itself to a clear roadmap like the other technologies documented in this report. For example, it is unclear what price/performance points will drive broad commercialization of DDS, since the potential use cases are highly dependent on techniques for writing (synthesis) and reading (sequencing) digital data in synthetic DNA which, while well-grounded in decades of medical/scientific applications, are nascent in the context of data storage.

This means that, in this nascent phase of the DDS ecosystem, not only will synthesis and sequencing capabilities continue to be tailored for specific use cases, the business value of any solution (i.e., the price users are willing to pay) will be significantly affected by the capabilities of the technology as it is deployed in production systems. Notwithstanding, this chapter outlines the main challenges and issues related to the state of the DDS ecosystem today.

CHALLENGES AND ISSUES

TCO

Storage TCO encompasses all capital expenditures (Capex) and operational expenditures (Opex) over the lifetime of the storage solution. Capex includes costs of equipment/hardware, software and infrastructure. Opex includes consumables, labor, energy and natural resources (e.g., water), and maintenance and support. Another important contribution to both Capex and Opex is data migration, either when media reaches end of life, or when the capabilities of new media justify transferring data.

As seen in recent industry analyses,[2,3] tape offers significant TCO advantages over other magnetic media which, when combined with performance, make it the dominant technology in the archival tier.

Comparing projected costs of DDS to existing technologies like tape is not that useful because the use case for a medium like DNA, especially initially, will not involve replacing tape, or any other existing storage technologies, but complementing them in the storage hierarchy (**Figure 68**).

Moreover, the relationship of the value of a DDS solution to a particular use case will evolve dynamically as DDS technology emerges commercially. The first use cases of DDS will likely be for deep/cold archives (100+ years duration), with Write Once/Read Never-to-Seldom access patterns, and assuming no additions or changes to the data during storage. For customers with such use cases, the value of the data may outweigh a high (relative to incumbent media) cost of writing and reading, as well as (again relative to incumbent media) low throughput and high latency (time to first byte).

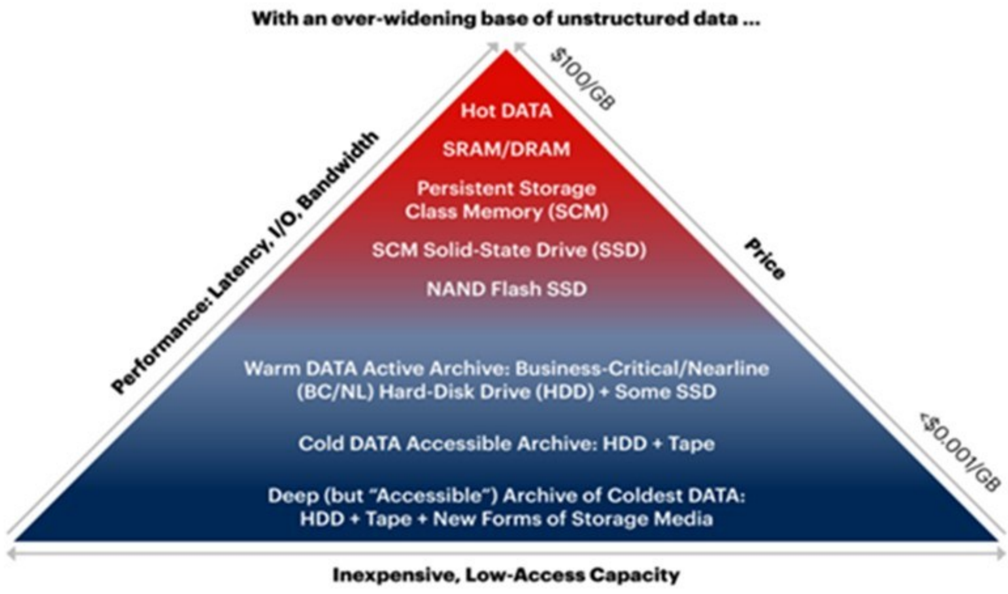


Figure 68. The Evolving Storage Pyramid[4]

Source: Gartner Product Manager Insight: Tape to the Future? October 23, 2020[ii]

As DNA write and read technologies improve in cost and performance, and are realized in more working systems, the solutions will enable more flexible use cases, with more diverse access patterns, resulting in different tradeoffs between read/write speed and cost versus the value of the data (i.e., what is an acceptable TCO). Aside from increases in throughput and reduced latency, the evolution of these TCO tradeoffs will involve many other aspects of using DNA as a storage medium in a system, including:

a) Costs of write and read substrates and storage containers, their reusability, the costs of reagents, resource costs (e.g., labor, water, electricity).
b) Cost of rack or system-level components necessary to automate storage and handling of substrates, containers and/or reagents.
c) Transition from manual to automated instruments.
d) Optimized error rates (i.e., reduced number of molecules required per TB); and
e) Standardized interfaces between each step of the DDS workflow.

Lastly, all of the above will affect, and be affected by, how DNA data storage is delivered. For example, writing and reading may be services, or on premises at a data center, or both.

By far the most definitive and important thing we can say about DDS TCO today is that when DNA based media is stored in controlled environments (temperature and humidity), it is extremely stable. Figure 69 shows DNA preservation methods studied in Organick[5], Grass[6], Coudy[7], and Bonnet[8] that demonstrate that properly controlled preservation techniques can ensure molecular stability (i.e., data stability), even at room temperature, over very long periods. Grass, Organick,

[ii] Gartner Product Manager Insight: "Tape to the Future?" J. Monroe and R. Preston. 23 October 2020, ID G000724101.

and Coudy additionally showed that the observed levels of molecular stability were sufficient to enable the encoded data to be successfully recovered.

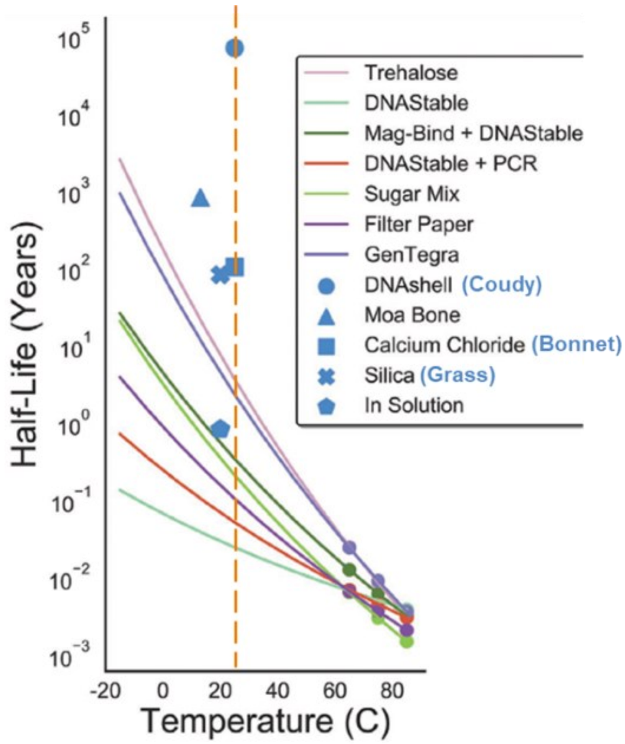


Figure 69. Half-life of verious DNA preservation methods. [5]

Source: From Figure 2b, Organick et al

The demonstrated stability of DNA can effectively eliminate fixity checks and technology migration, both increasingly large factors in archival storage TCO. In particular, regarding technology migration, since DNA based data is stored in a pool, and not written into a device with a fixed substrate as media (i.e., HDD, SSD, Tape), and the DNA molecular structure is universal, it will always be physically possible to read DNA media back from an archive, even if the devices used to write or read the media when it was created are no longer available. This means that DNA based data need not undergo the increasingly costly technology migration which characterizes today's archival storage technologies.

Figure 70 illustrates the implications of DNA data storage cost over time. The commercial price to store a petabyte of data using list prices for Tape (Fujifilm Calculator) and Cloud (AWS public pricing) was used as a baseline. The price points for DNA-based storage were picked arbitrarily, for illustration. Beyond 10 years, the advantageous properties of DNA-based storage relative to incumbent media begin to dominate; that is, the DNA data storage TCO becomes increasingly attractive.

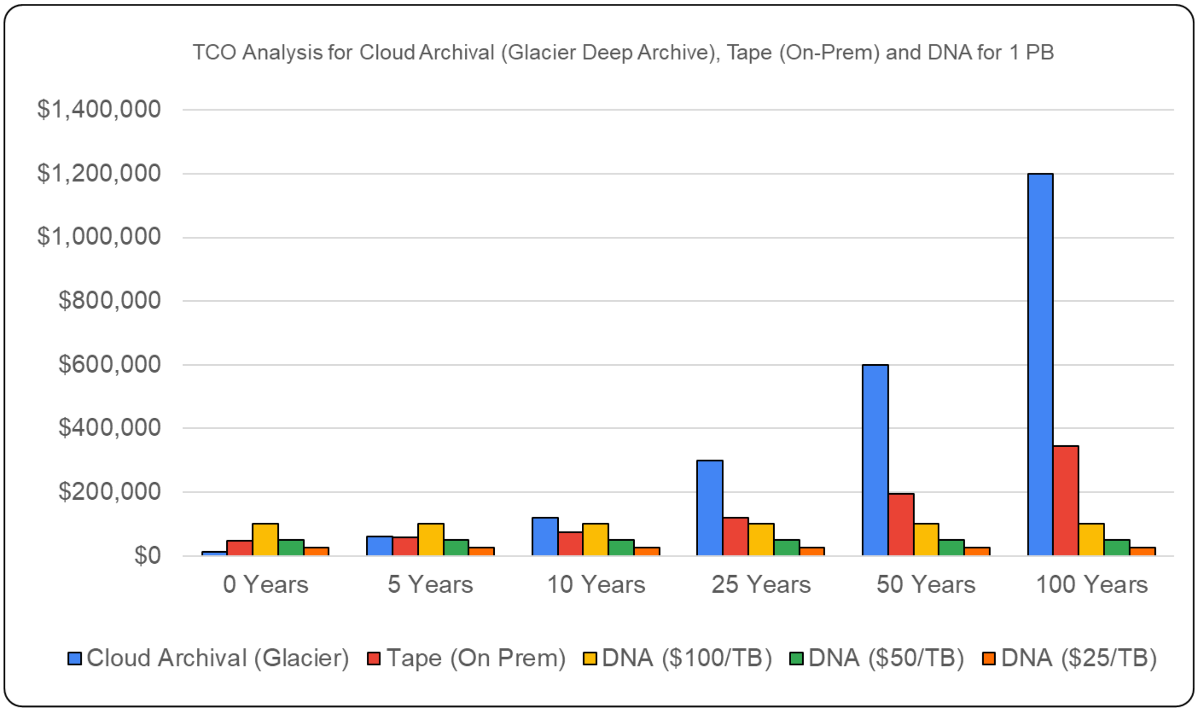


Figure 70. Estimated Cost of Writing and Storing – Legacy vs. DNA[4]

The key challenges for DNA data storage TCO are not related to storage, but to read/write performance and cost. We explore this in the following two sections.

WRITING DNA (SYNTHESIS)

The main challenge for DDS in terms of writing is that today's DNA synthesis performance is several orders of magnitude less capable than that for incumbent media (**Figure 71**). Per the TCO discussion above, it is probably unnecessary that DNA data write speed be directly competitive with incumbent media. For example, deep/long archival users may be content with write throughput in the range of a few 100 megabytes per day (few kilobytes/s). That said, regardless of the entry point, solutions will need some minimum performance characteristics at a cost that facilitates a competitive TCO given the use case.

Media	Write Latency (time to 1st byte)	Write Throughput
DNA Data Storage	Minutes to hours	~100 MB/day = 0.001 MB/s
Tape	Seconds to minutes	~400 MB/s (uncompressed)
HDD	Up to tens of seconds	~300 MB/s
SSD	Single digit seconds or less	~500 MB/s (non-NVMe)
Flash	Single digit seconds or less	~1000 MB/s

Figure 71. Write latency and throughput for various storage solutions: DNA Data Storage Alliance

Parallelism is foundational to advancing DNA synthesis throughput. Since the speed of chemical reactions is inherently bounded, a common approach to increasing throughput is to take advantage of the small size of DNA molecules to execute millions to billions of synthesis operations in parallel.

To spur DNA synthesis innovation for DNA data storage, the IARPA Molecular Information Storage program (MIST) set synthesis goals (**Figure 72**) and challenged industry and academia to develop solutions.

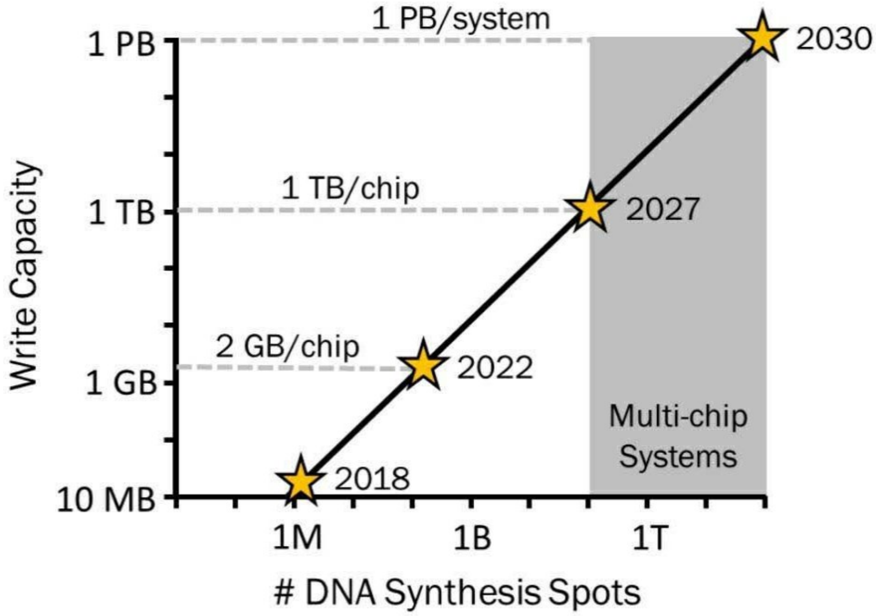


Figure 72. 2022 IARPA Roadmap for DNA synthesis, courtesy David M. Markowitz. Assumes single stranded DNA, 150 nucleotides in length (20 nt flanking primers), encoded at 1 bit/nucleotide.

BASE-BY-BASE SYNTHESIS

Electrochemical array-based synthesis, in which the fluidics necessary to do synthesis are combined with CMOS semiconductor technology to control synthesis chemistry on discrete and very small areas on a chip[9,10], is one of the main modalities being pursued. In 2021, Nguyen et al[11] showed synthesis of synthetic DNA in an electrochemical array, successfully synthesizing 150 nucleotide long sequences in 650nm wells, at a feature pitch of 2um, or a density of 32 million synthesis spots per cm^2.

Key to the success was demonstrating the ability to localize acid diffusion at each synthesis site, which is critical to keeping the synthesis chemistry localized to each well and thus ensuring that the synthesis process is precisely controlled with acceptable error rates. This effort advanced the state of the art in array-based synthesis density by nearly 3 orders of magnitude. While this chip was not scaled to full production, if it were, at the densities demonstrated, such a device could yield performance on the order of a few kilobytes/s/cm^2 (assuming each unique DNA sequence encodes 10 bytes of data and is written over 24 hours) or a few hundreds of megabytes per day. While far lower than incumbent media, this might be adequate for deep/cold archival data storage applications. In evidence of further progress, Twist Bioscience stated in 2022 that they had developed a chip which has a capacity of 1GB per run with about 100M synthesis spots, approaching the trajectory of the MIST goals. Other base-by-base synthesis techniques are also being pursued, such as ink-jet printing[12,13] and light directed synthesis[14,15].

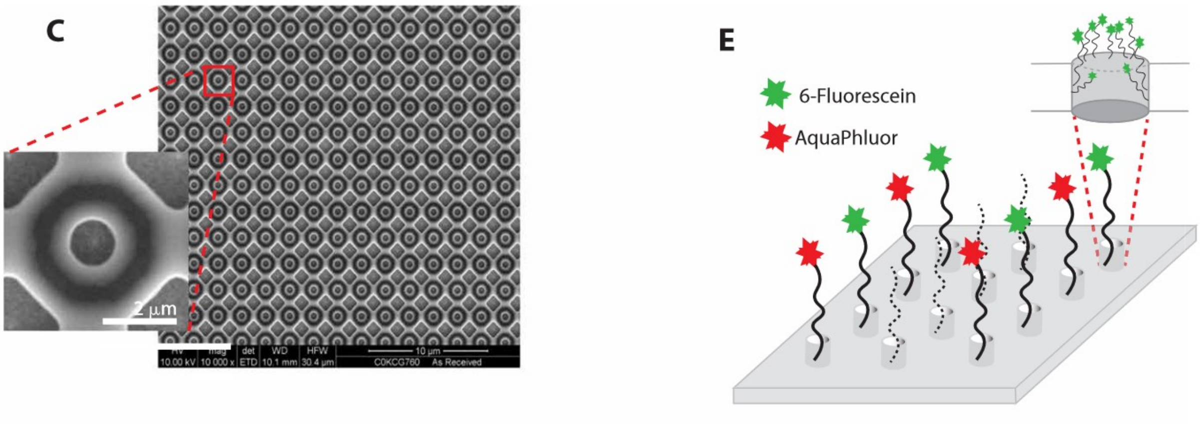


Figure 73. Electrochemical DNA synthesis on a nanoscale array. (c) An overview of the nanoscale DNA synthesis array with scanning electron microscopy images of the 650-nm electrode array and enlarged view of one electrode. (e) Illustration of the wells patterned with ssDNA oligos with multiple copies of each oligo per synthesis location.[11]

DNA Assembly Based Synthesis

In addition to the above techniques, which write molecules base-by-base, typically in strands containing 150-200 nucleotides, there are synthesis techniques being pursued by companies such as Catalog and Biomemory, where a library of predefined sequences, somewhat akin to movable type in a mechanical printing press (i.e., symbols) is used to encode the sequences. The motivation is that these symbols (themselves built with base-by-base synthesis, but needing to be built only once), in combination with certain encoding strategies, can then be assembled into longer strands of DNA by various DNA assembly techniques (e.g. ligation) that can enable different synthesis efficiencies than base-by-base methods.[16] Catalog currently claims a peak throughput of about a terabit per day in their first generation platform (i.e., in the hundreds of megabytes per day range that was estimated for a scaled version of the electrochemical array-based technology above).

Non-Sequence Based Techniques

Lastly, there is research into synthesis techniques which encode data not in the sequence of bases in the synthetic DNA molecules, but by using (and guiding) the self-assembly properties of DNA molecules to assemble molecular structures in which the structure itself encodes information; so-called "Structure-Based DNA Data Storage"[17]. These techniques are not as far along in scaling for data storage as are sequence-based techniques, but they represent active areas of work providing yet further potential alternatives in throughput and information density.

In summary, while significant scaling advances are still required, the fundamentals of synthesizing DNA encoded with digital data on scalable technology platforms have been shown to work at speeds which could be practical for early deep/cold archive applications, and substantial investment in R&D continues within the ecosystem.

READING DNA (SEQUENCING)

At present, the two primary modes of DNA sequencing most promising for DDS are Sequencing by Synthesis (SBS), **Figure 74**, and Nanopore sequencing, **Figure 75**.

SBS detects DNA bases indirectly. It gets its name because it starts with a single-stranded template strand of DNA and then synthesizes a complementary DNA strand from that template. As each base is added in the complementary strand it is identified (typically optically), and the base in the original template strand can then be identified.

Nanopore sequencing detects bases directly. A strand of DNA is passed through a pore in a membrane surrounded by an electrolyte solution. With an electrical bias applied across the membrane, the DNA strand moves through the nanopore, the disruption of ionic current or tunneling current is registered, enabling the direct detection of the bases in the strand. Nanopores today are almost entirely biological; however, research is underway to create solid state nanopore sequencing[18].

In general, today, on a raw, per-base basis, SBS is more accurate/slower and nanopore is less accurate/faster. In terms of scaling in production, there is a battle for throughput/cost. We'll examine both.

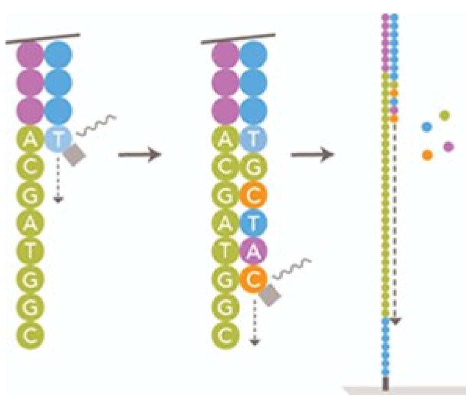


Figure 74. Sequencing by Synthesis: Illumina

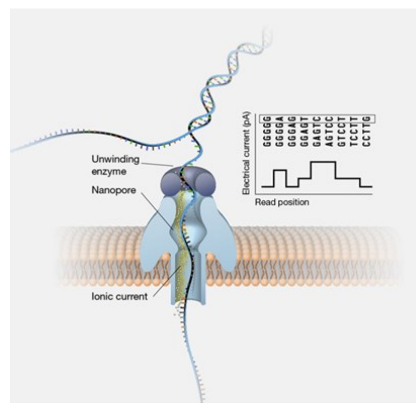


Figure 75. Nanopore Sequencing: National Human Genome Research Institute (NHGRI)

As tracked by NHGRI (**Figure 76**) DNA sequencing underwent a rapid cost reduction (and performance increase) curve starting in 2008, due to the advent of SBS; however, the rate of change has slowed and leveled out at just under $10/GBase. Illumina stated in 2022 that that their highest end products can achieve roughly $6/Gbase, that there is a "direct line of site" to $1/Gbase, and that there are "no conceptual hurdles" to $0.1/Gbase. Assuming we encode 1 bit/base, the $0.1/Gbase milestone is probably on the horizon over the next 3-5 years and would bring us to a price of roughly $800/terabyte. This is still quite high for commercial deployment, and more cost scaling is needed, but note that estimates of DNA sequencing costs today are all based on medical/scientific use cases and business models, and various factors could further mitigate cost:

- Data storage applications can accept higher error rates than medical/scientific applications. For example, the cost of sequencing the human genome is frequently used as a benchmark for sequencing cost. In general, the coverage factor (average # of times each nucleotide in the sequence is sampled to get the final read done) is 30X. In data storage, the DNA codec can both avoid nucleotide sequences that cause problems during sequencing and can also recover from errors during decode. Such factors might, for example, enable coverage factors nearer to 10X. This, in turn, would result in significantly higher effective throughput during sequencing, with attendant lower costs.

- Related to the previous point, the 1bit/base encoding efficiency used in the estimates here may be overly conservative; densities between 1 and 2 bits/base would raise effective throughput and thus lower cost.

- The highest performance form of sequencing today is SBS. While further enhancements in SBS are inevitable, nanopore sequencing is a promising new technology which, especially with the advent of solid state nanopore development, could achieve breakthroughs that initiate another major downward cost (and upward performance) trend.

- Ecosystem considerations and business models could help drive lower sequencing prices as/if data storage begins driving demand for solutions, providing new demand to the ongoing and increasing medical/scientific demand.

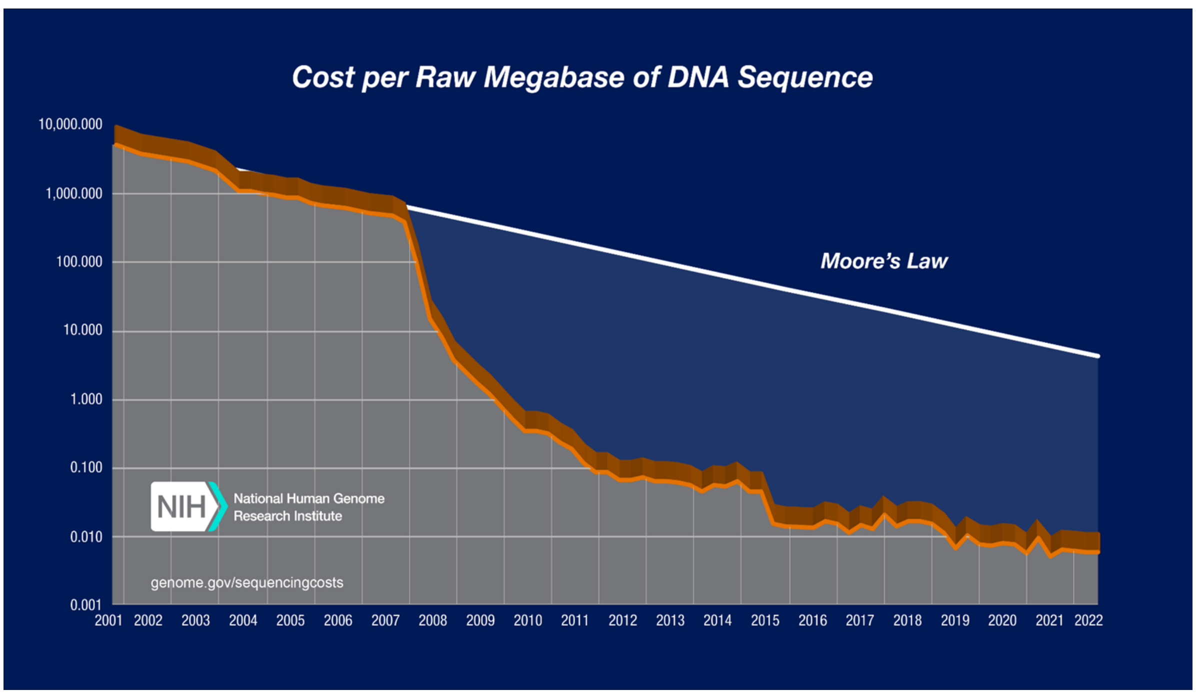


Figure 76. Cost per Raw Megabase of DNA Sequencing, DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP) Available at: www.genome.gov/sequencingcostsdata

Let's look at raw performance to get a sense of how much data, and how quickly, we can read from DNA. The fastest sequencing solutions today are approaching 7 terabases/day which, again using 1bit/base, is approaching 1 terabyte/day. Is this enough for deep archival recovery? Perhaps.

Is it enough for more, so-called, “Active Archive” use cases, where reads are more frequent? Probably not; these use cases will likely require read speeds on order of a few hundred terabytes/day or higher.

As with DNA synthesis, the costs and performance of DNA sequencing need to continue scaling to become cost competitive in solutions; however, we are likely approaching the requirements of some use cases, and, also as with synthesis, the fundamental technologies have been demonstrated, and scaling is now the task.

SUSTAINABILITY

Another important attribute of DDS is sustainability. At first order, DDS has inherent advantages for sustainability due to its extremely high data density and small volumetric footprint. In comparison to sprawling data centers, there is simply not much of a comparison.

More systematically, the most important metrics that contribute to sustainability are (1) greenhouse gas emissions (kgCO_2equivalents/TB), (2) energy consumption (MJ/TB), and (3) water consumption (L/TB). Recent life cycle assessment (LCA) simulations[19,20] have indicated that the sustainability profile of DDS approaches compare well to tape and other incumbent storage media. This will have to be demonstrated at scale.

The major sustainability issue for DDS today is the use of oil-based chemicals needed for Phosphoramidite synthesis, which are toxic and flammable; however, the projection is that enzymatic synthesis, which uses aqueous based processes, will mitigate this issue.

BIOSECURITY

It is important to note that no organisms or living cells are required for DNA data storage; synthetic DNA for data storage is constructed and manipulated through well-controlled chemical processes, and the sequences generated for it are controlled by software codecs. It is already the case that organizations, such as the International Gene Synthesis Consortium[21], have been formed to both conform with governmental regulations, and further enhance sequence screening protocols to reduce the risk of generating sequences associated with pathogenic organisms. Moreover, due to the fact the sequences encoded for data storage are not constrained by biological/medical requirements, DNA codecs can and will embed biosecurity requirements as they exist today, and as required with the evolution of the DNA data storage ecosystem.

SUMMARY

Synthetic DNA as a storage medium is compelling due to its ~1bit/nm^3 bit size, its molecular stability, and the property that it is decoupled from the read/write device. These aspects, if delivered in commercial systems, can enable compelling TCO for high capacity and long time scale archival use cases, as well as potentially for other use cases as the capabilities of the technology evolve.

While DNA data storage is not yet ready for productization today, with the huge advances in medical/scientific DNA technology and applications over the past several decades, plus academic research and commercial biotech both targeted at DNA data storage over the past decade, the basic foundations for synthetic DNA as a data storage medium have been demonstrated[17]. It is thus reasonable to expect that a path to a commercial DNA data storage ecosystem will come into focus over the next 5-10 years.

REFERENCES

[1] Gartner Market Trends: Evolving Enterprise Data Requirements —How Much Is Not Enough? (ID G00727554)

[2] Arcilla, A. Quantifying the Economic Benefits of LTO-8 Technology.

[3] TCO Calculator for Data Storage | Fujifilm [United States]. https://www.fujifilm.com/us/en/business/data-storage/resources/tco-tool.

[4] Preserving our digital legacy: An introduction to DNA data storage", DNA Data Storage Alliance; https://dnastoragealliance.org/dev/wp-content/uploads/2021/06/DNA-Data-Storage-Alliance-An-Introduction-to-DNA-Data-Storage.pdf

[5] Organick *et al.* An Empirical Comparison of Preservation Methods for Synthetic DNA Data Storage. *Small Methods* **5**, 2001094 (2021).

[6] Grass RN, Heckel R, Puddu M, Paunescu D, Stark WJ. Robust Chemical Preservation of Digital Information on DNA in Silica with Error-Correcting Codes. Angew Chem Int Ed Engl. 2015; 54(8): 2552-2555

[7] Coudy, D., Colotte, M., Luis, A., Tuffet, S. & Bonnet, J. Long term conservation of DNA at ambient temperature. Implications for DNA data storage. PLOS ONE 16, e0259868 (2021).

[8] Bonnet, J. et al. Chain and conformation stability of solid-state DNA: implications for room temperature storage. Nucleic Acids Res. 38, 1531–1546 (2010).

[9] Kosuri, Sriram, and George M. Church. "Large-scale de novo DNA synthesis: technologies and applications." Nature methods 11.5 (2014): 499-507

[10] Jensen M.A., Davis R.W, Template-Independent Enzymatic Oligonucleotide Synthesis (TiEOS): Its History, Prospects, and Challenges; Biochemistry 2018, 57, 12, 1821–1832, March 13, 2018; https://doi.org/10.1021/acs.biochem.7b00937

[11] Nguyen, B. H. et al. Scaling DNA data storage with nanoscale electrode wells. Sci. Adv. 7, eabi6714.

[12] Catalog, https://www.catalogdna.com

[13] Oligonucleotide Library Synthesis | Agilent. https://www.agilent.com/en/product/sureprint-oligonucleotide-library-synthesis/oligonucleotide-library-synthesis

[14] Efcavitch, J. W. & Holden, M. T. Homopolymer encoded nucleic acid memory (2021); US Patent #: US-11174512-B2

[15] Antkowiak, P.L., Lietard, J., Darestani, M.Z. et al. Low cost DNA data storage using photolithographic synthesis and advanced information reconstruction and error correction. Nat Commun 11, 5345 (2020). https://doi.org/10.1038/s41467-020-19148-3

[16] Roquet et al. DNA-based data storage via combinatorial assembly. https://doi.org/10.1101/2021.04.20.440194

[17] Doricchi et al. Emerging Approaches to DNA Data Storage: Challenges and Prospects. ACS Nano 2022 16 (11), 17552-17571 DOI: 10.1021/acsnano.2c06748

[18] Xue, L., Yamazaki, H., Ren, R. et al. Solid-state nanopore sensors. Nat Rev Mater 5, 931–951 (2020). https://doi.org/10.1038/s41578-020-0229-6

[19] Nguyen, B. H. et al. Architecting Datacenters for Sustainability: Greener Data Storage using Synthetic DNA.

[20] Mytton, D. Data centre water consumption. Npj Clean Water 4, 1–6 (2021).

[21] https://genesynthesisconsortium.org