

Exploring data storage innovations: From DNA to holography

Zhiheng Wang¹, Liheng Xu², Runkun Yang^{3,4}

¹Shanghai united international school (jiaoke) campus, Shanghai, 200080, China

²Guanghua Cambridge International School, Shanghai, 201319, China

³Pennon Education Group, Qingdao, 266041, China

⁴20204251@stu.hebmu.edu.cn

Abstract. This paper delves into the evolution of data storage technologies, focusing on the solid-state drive (SSD) and hard disk drive (HDD) and their working principles. It explores the advancements in DNA as a storage medium, highlighting its exceptional data density and potential as a long-term storage solution. The paper discusses the challenges and limitations of DNA data storage, emphasizing the need for further research to overcome these hurdles. Additionally, the paper introduces holographic data storage as a novel optical storage technology and discusses its potential applications. It also touches on the decline of the optical disc industry due to the emergence of alternative storage technologies. The use of metal nanoparticles (MNPs) in optical disks is explored, emphasizing their role in high-contrast imaging and data storage. The paper concludes by highlighting the diverse methods of data storage, each with its unique advantages and drawbacks, and anticipates future improvements in electronic information processing technology, leading to faster data transmission and increased data storage capacity within the same volume of storage media.

Keywords: SSD, HDD, Data Storage Technology.

1. Introduction

For over 30 years since the invention of the solid-state drive (SSD), there has been a revolution in the realm of hard disks. With the advancement of technology, both the cost and efficiency of SSDs have improved significantly. Therefore, it is worthwhile to explore the differences between solid-state drives and hard disk drive (HDD). Additionally, other intriguing methods of storage are DNA data storage and optical data storage which have different advantages and disadvantages, and the current methods have been presented alongside a glimpse into the future.

2. Working principle of the HDD and SSD

2.1. Working principle of the HDD

The HDD has been in existence for more than 60 years and is widely used in computers today. Inside the HDD, there is a disk that contains a cobalt-chromium-tantalum alloy, which can be magnetized to become a permanent magnet. Beneath the disk, a DC motor enables it to rotate at about 7200 revolutions per minute. Each side of the disk can be divided into more than 50,000 tracks, and each track can be further divided into several sectors. These sectors typically include the preamble zone, address, data, error correction code, and gap.

Another crucial component is the head stack, with one stack located above and one below the disk. Each arm at the end of the stack holds a slider and a read/write head. At the other end of the stack, there are two strong neodymium magnets, and the Ampere's force acting on the stack can move it across the disk as rapidly as 20 times per second.

Each domain on the disk represents a bit. In the write head of the stack, there is a coil of wire capable of generating a strong magnetic field, which can induce the alloy in the disk to become a permanent magnet. Furthermore, the direction of the magnet in the domain can be determined by the direction of the magnetic field in the coil. If the nearby domain's magnetic field is the same, it represents 0; otherwise, it represents 1.

To read from the disk, a Giant Magnetoresistance (GMR) component, made of ferromagnetic and non-magnetic materials, can detect changes in magnetic field strength flux within it by altering its resistivity. This allows us to distinguish between 0 and 1 by monitoring the resistance of the GMR.

2.2. Working principle of the SSD

After discussing the hard disk drive, let's delve into the working principle of the SSD. SSD is a technology that has been around for over 30 years. Although it may resemble an HDD, both the components and the working principles of an SSD are entirely distinct from those of an HDD.

An HDD accesses rotating magnetic storage media using a head mounted on a mechanical actuator. In contrast, an SSD stores data on non-volatile memory, typically NAND flash. The term 'solid state' refers to the absence of moving parts in the device and its use of silicon as the storage medium. Due to this lack of moving parts, SSDs are less fragile than HDDs, making them widely used in mobile devices such as digital cameras and cell phone memory cards like Secure Digital and CompactFlash, which offer simpler and more compact alternatives to SSDs.

The primary consideration when deciding between using an SSD and a simpler storage device is the performance requirements of the application. For example, a digital camera requires much less storage than a server or a laptop with multiple cores. These differences significantly impact the design of SSDs.

HDDs, which serve as the primary component for non-volatile storage in modern computer systems, fall under the block devices class of storage, just like SSDs. These devices use logical addressing to access data, employing brief, defined, continuous blocks of bytes as a means to retrieve data and abstract the physical media. Each block device consists of storage media, a controller, and various minor components [1]. A typical SSD is made up of an array of NAND flash memory, an SSD controller, and host interface logic (SATA, USB, and PCI Express) [2].

These are the main components of SSDs and some differences in their working principles compared to HDDs.

3. Comparison between the HDD and the SSD

3.1. Writing performance

Firstly, an SSD can read and write data at significantly higher speeds than a hard drive because it lacks moving parts. In contrast, a hard drive relies on rotating disks and read/write heads. Consequently, SSDs require much less time to access data since they don't need to account for seek times or spin delays.

Secondly, SSDs deliver faster data access thanks to their use of flash memory chips, which can read and write data directly. In comparison to mechanical hard drives, SSDs offer notably quicker random read and write speeds, along with reduced access times. To illustrate, think of SSDs as books you can open to any page instantly, while HDDs are akin to records that need to physically turn to access information [3].

3.2. Power

SSDs are entirely mechanically inert, requiring no motor to rotate platters or move magnetic arms. Consequently, they consume minimal power when idle. Additionally, SSDs read and write data much faster than mechanical hard disks, which reduces the power required for data operations. While the

primary control chip and cache chip in SSDs do consume some power, their power consumption is decreasing as manufacturing processes improve. Nevertheless, due to the inherent structure of HDDs [4], there are currently limited additional methods available to significantly reduce power consumption.

3.3. Price

The price of HDDs is notably lower than that of SSDs due to several factors. Mechanical hard disks lack the dense mechanical components found in SSDs, and their storage capacity is often greater. Consequently, as shown in Figure 1, the cost per gigabyte (GB) of mechanical hard disks is significantly lower compared to SSDs [5].

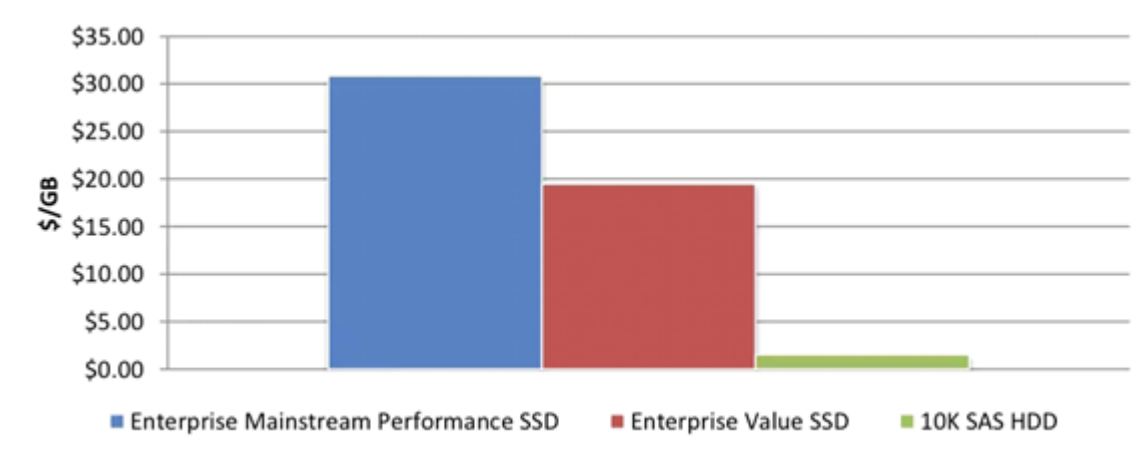


Figure 1. Compare between SSDs and HDDs.

3.4. Resistance to falling

SSDs are not susceptible to collisions or vibrations, and they excel in withstanding large shocks, making them highly reliable in terms of data preservation. Consequently, the risk of data loss is minimized with SSDs.

Considering this aspect, SSDs may come at a higher cost, but they offer a superior user experience when it comes to storage. However, if speed and other advanced features are not essential for your needs, low-cost mechanical hard drives remain a viable and cost-effective choice. HDDs can indeed be an excellent option in such cases.

4. DNA data storage

The worldwide demand for data storage is expected to reach 1.75×10^{14} GB by 2025, driven by the exponential growth in global data generation. Moreover, recent years have witnessed significant advancements in using DNA as a medium for data storage. Scientists follow a six-step procedure, as illustrated in Figure 2, to store and retrieve information in DNA:

- 1) Translate data into a specific code.
- 2) Write the code using the DNA base pairs (synthesize new oligonucleotides).
- 3) Store the DNA in a suitable environment to preserve its integrity.
- 4) Enable random access to the stored DNA.
- 5) Sequence the DNA to retrieve the code.
- 6) Decode the DNA to obtain the original data [3].

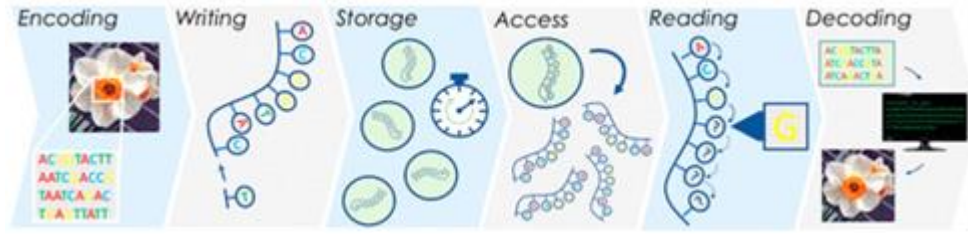


Figure 2. Six steps of processing data in DNA [6].

There are numerous advantages to using DNA as a storage medium. DNA molecules can remain viable for extended periods, with the potential to stay intact for thousands of years under favorable conditions. For instance, scientists have successfully sequenced DNA from specimens dating back 300,000 years, such as that preserved in the bones of a bear. Additionally, maintaining suitable conditions for DNA storage requires relatively low power consumption, resulting in cost-effective storage.

The Shannon index can be employed to calculate a medium's information storage capacity. The largest amount of information that a single base can store is

$$H = -\sum_i^{A,T,C,G} P(i) \log P(i) \leq \log \sum_i^{A,T,C,G} P(i) \frac{1}{P(i)} = \log 4 = 2 \text{ bit} \quad (1)$$

$P(i)$ is the probability that the base is set on a place on the chain of DNA, A T C G are four bases of deoxyribonucleotide monomers and \log is the base-2 logarithm because it is often used to calculate the unit bit. when if the probability of the occurrence of four bases is equal, $P_i = 1/4$, and the maximum information that a base can provide is 2 bits, then the density can be calculated as [7]:

$$\rho = \frac{2 \text{ bit}}{1 \text{ base} \times 325 \frac{\text{Dalton}}{\text{base}} \times 1.67 \times 10^{-24} \frac{\text{g}}{\text{Dalton}}} = 3.69 \times 10^{-21} \frac{\text{bit}}{\text{g}} = 4.61 \times 10^{20} \frac{\text{byte}}{\text{g}} \quad (2)$$

DNA offers a crucial advantage over current information processing systems beyond its durability and cost-effectiveness: data density. DNA boasts a theoretical data density of 4.5×10^7 GB/g, which means that even a small amount of DNA can hold a vast amount of data if the technology is fully developed. Internet companies, for example, could replace large warehouses full of disk drives with a tiny amount of DNA.

However, various technical limitations hinder DNA data storage from competing with disk drives at present. Long DNA sequences must be broken down into smaller segments (up to 200 bases) because the de novo synthesis of larger sequences remains challenging, necessitating the creation of numerous distinct DNA sequences. While the read principle is similar to that of hard disk drives, in data readout, researchers must sequence individual oligonucleotides to access the stored data, which can only be detected through fluorescence. This process requires a substantial workforce, costly fluorophores, and expensive optical equipment.

Furthermore, despite its significant impact on the design and coding of data storage systems, DNA degradation is still defined imprecisely and too narrowly. Although many studies have quantified it using mass spectrometry or next-generation sequencing, a DNA storage system can deteriorate in various ways. Factors like base composition, length, and surroundings can lead to DNA strand loss, changes in DNA sequences, or even chain breakages. Occasionally, sequencing alone may not recover the original information due to chemical changes in the bases. System design should carefully consider how to protect DNA from environmental factors [8].

One approach to address this problem is to construct three-dimensional structures using DNA. Leveraging the specific base-pairing abilities, researchers can create different three-dimensional structures or platform shapes for storing data [6].

5. Optical data storage technology

5.1. Holographic data storage

Storing information using specific holographic photos is a new optical storage technology called holographic data storage. In the writing process, data is transformed into a two-dimensional image, and this image data forms a highly concentrated beam of light. This concentrated beam of light then interferes with a reference laser beam, and the resulting interference pattern is stored in holographic media.

To read the data, researchers only need to illuminate the holographic media with the reference laser beam, which reconstructs the two-dimensional image. This process allows them to obtain the original data, as shown in Figure 3.

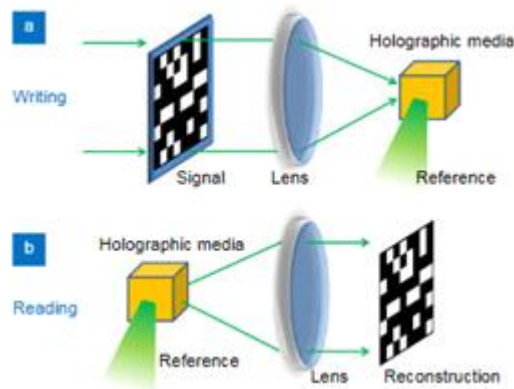


Figure 3. Writing and reading process of holographic data storage [9].

The optical disc industry has produced various storage devices like CDs, which currently boast a maximum recording capacity of up to 200 GB. However, despite their significant success in the past, the optical industry is gradually becoming outdated. According to reports from the storage market, other storage technologies have now taken its place.

5.2. Metal nanoparticles

Over time, reliable and high-yield preparation techniques have been developed, making it easy to synthesize metal nanoparticles (MNPs) with various geometries and sizes ranging from 1 to 100 nm. Due to their excellent chemical stability, gold and silver are the preferred choices for creating MNPs. There have been several instances of MNP utilization in optical disks.

One approach involves utilizing MNPs for high-contrast imaging, which is suitable for the process of retrieving stored data. Additionally, silver MNPs, forming different structures, play a crucial role in physically storing data on optical disks. Researchers employ electron beam lithography to position these MNPs on a transparent substrate, which can be observed using optical microscopy.

Data readout is achieved through conventional optical far-field techniques, capturing light at various frequencies or wavelengths. The increase in data density is constrained by the spectrum width of the nanoparticles' resonance lines. Nonetheless, this density surpasses that of traditional optical storage techniques [10].

6. Conclusion

In conclusion, this paper has provided an insightful exploration of various data storage technologies, shedding light on their mechanisms, advantages, and limitations. We began by examining the evolution of traditional hard disk drives (HDDs) and the modern solid-state drives (SSDs), emphasizing the significant improvements in speed, power efficiency, and durability offered by SSDs. Additionally, we

discussed the affordability of HDDs and their relevance in scenarios where high-performance requirements are not paramount.

DNA data storage emerged as an exciting prospect for the future, offering unparalleled data density and longevity. We outlined the six-step process involved in encoding and decoding data using DNA, highlighting its potential to address the ever-growing demand for data storage.

The paper also introduced holographic data storage as a cutting-edge optical technology with promising applications, despite the declining fortunes of the optical disc industry. The utilization of metal nanoparticles in optical disks was examined, showcasing their role in enhancing data storage density.

In anticipation of the future, we envisage the continuous evolution of electronic information processing technology, resulting in faster data transmission and greater data storage capacities within compact storage mediums. The landscape of data storage is dynamic and multifaceted, offering various options, each with its advantages and drawbacks. As technology advances, striking a balance between cost, performance, and data integrity will remain a key consideration for individuals and organizations alike.

Authors Contribution

All the authors contributed equally and their names were listed in alphabetical order.

References

- [1] Min, C., Kim, K., Cho, H., Lee, S. W., & Eom, Y. I. (2012, February). SFS: random write considered harmful in solid state drives. In FAST (Vol. 12, pp. 1-16). Dorman L I 1975 Variations of Galactic Cosmic Rays (Moscow: Moscow State University Press) p 103
- [2] Cornwell, Michael, and Pure Storage. "Anatomy of a Solid-State Drive | Communications of the ACM." Anatomy of a Solid-State Drive, dl.acm.org/doi/10.1145/2380656.2380672. Accessed 9 Sept. 2023. Szytula A and Leciejewicz J 1989 Handbook on the Physics and Chemistry of Rare Earths vol 12, ed K A Gschneidner Jr and L Erwin (Amsterdam: Elsevier) p 133
- [3] Martin, Matthew, and Matthew Martin. "SSD VS HDD: What Is the Difference between SSD and HDD Matthew Martin." Guru99, 25 July 2023, www.guru99.com/ssd-vs-hdd.html.
- [4] Seo, E., Park, S. Y., & Uргаonkar, B. (n.d.). Empirical analysis on Energy Efficiency of flash-based SSDs - Usenix. Empirical Analysis on Energy Efficiency of Flash-based SSDs. https://www.usenix.org/legacy/events/hotpower08/tech/full_papers/seo/seo.pdf
- [5] Solid state drive vs. hard disk drive price and Performance Study - Dell. (n.d.). https://i.dell.com/sites/csdocuments/Shared-Content_data-Sheets_Documents/en/ssd_vs_hdd_price_and_performance_study.pdf
- [6] ACS Nano 2022, 16, 11, 17552–17571 Publication Date: October 18, 2022 <https://doi.org/10.1021/acsnano.2c06748>
- [7] National Science Review, Volume 7, Issue 6, June 2020, Pages 1092–1107, <https://doi.org/10.1093/nsr/nwaa007>
- [8] Matange, K., Tuck, J.M. & Keung, A.J. DNA stability: a central design consideration for DNA data storage systems. Nat Commun 12, 1358 (2021). <https://doi.org/10.1038/s41467-021-21587-5>
- [9] Lin X, Liu J P, Hao J Y, Wang K, Zhang Y Y et al. Collinear holographic data storage technologies. Opto - Electron Adv 3, 190004 (2020). DOI: 10.29026/oea.2020.190004
- [10] Gu, M., Zhang, Q. & Lamon, S. Nanomaterials for optical data storage. Nat Rev Mater 1, 16070 (2016). <https://doi.org/10.1038/natrevmats.2016.70>