Managing the Data Explosion

The Strategic Advantage of Tape

Brad Johns
Brad Johns Consulting LLC
October 2024

Managing the Data Explosion - Executive Overview

In today's digital age, the demand for data storage is growing at an unprecedented rate. Enormous amounts of data are being created and stored, driven by factors such as high-definition video, surveillance footage, social media, 5G networks, virtual reality, electronic medical records, the Internet of Things (IoT), scientific research, and the rise of artificial intelligence (AI).

Several other factors contribute to the growth of the enterprise storage install base[1]. Many organizations are reluctant to delete data because there is a risk that they may require the deleted information for litigation, compliance, or business reasons. Also, deleting data requires a deep understanding of the information and establishing processes for its disposal, which requires time and money. Finally, with the advent of GenAI, organizations are discovering that there may be significant value in their historical data, sparking a new wave of interest in retaining data for potential use in developing their own AI applications.

In collaboration with Furthur Market Research, our research indicates staggering growth in the enterprise storage install base. We estimate that over 5.7 zettabytes of compressed enterprise storage were installed in 2023[2]. Further, we are projecting an annual install base growth of 24.5% between 2023 and 2030, resulting in an install base of 26.6 zettabytes in 2030, growing by a factor of 4.5 times over seven years. This tremendous growth will challenge many organizations' ability to manage vast amounts of data cost-effectively, underscoring the issue's urgency while meeting sustainability objectives and power constraints.

Even though at least 70% of all enterprise data is inactive, with access frequency ranging from minutes to decades, it is stored on SSDs or HDDs. Energy is consumed, and carbon emissions are generated unnecessarily to keep this data online. Surprisingly, while tape storage's financial and environmental benefits are well-known, we estimate that tape comprised only 16% of the shipped 2023 enterprise industry storage capacity. Expanding the use of tape storage represents a tremendous opportunity for organizations to significantly reduce overall costs and energy consumption while helping organizations achieve their sustainability goals.

Given all the potential benefits of utilizing tape storage for inactive data, why aren't more organizations using tape storage? Historically, to use tape storage, unique software and skills were required. However, innovative solutions that seamlessly integrate into the IT infrastructure using the Amazon S3 object interfaces have reached the market. Some recent examples include Point Software and Systems S3-to-Tape software PoINT Archival Gateway[3], Grau Data XtreemStore[4], and the IBM Storage Deep Archive[5].

Object storage is ideally suited for the storage and management of inactive data. By incorporating the S3 interface, these solutions enable customers to utilize standard object storage interfaces for their applications and users while integrating tape storage to minimize the need for unique infrastructure and skills. Utilizing these innovative solutions and leveraging existing tape storage infrastructure can improve sustainability and security while reducing costs.

Mitigating Energy Consumption in Data Centers: The Power-Saving Potential of Tape Storage

Data centers are a significant and growing portion of energy consumption. The Electric Power Research Institute (EPRI)[6] estimates that data centers consume 4% of all power in the United States. They suggest that due to GenAI, data center power consumption could grow to 9% of all power consumed by 2030. Data storage accounts for a sizable portion of data center energy requirements. We estimated in 2020 that 17% of all power in data centers was used by storage, a figure that could rise to 29% by 2030 due to the increasing use of SSDs for GenAI applications.

Power availability has become critical for organizations looking to expand their data centers. It is increasingly difficult to obtain the necessary power for large data centers, which can quickly require hundreds of megawatts. For example, northern Virginia is a major data center hub but faces significant power supply challenges. The demand for data center capacity has led to multiyear delays in power availability, and relief is not expected until 2026, when new transmission projects are completed.

Tape storage helps address this challenge. It requires dramatically less energy than disk storage. For example,

---

[1] We define the enterprise install base as the total capacities installed of all enterprise-class SSDs, HDDs, and tape. This definition specifically excludes shipments of consumer-grade SSDs, HDDs, and flash modules.

[2] Furthur Market Research - The impacts of Generative AI on Enterprise Data, https://furthurdata.com/

[3] Point Software and Systems - https://www.point.de/en/products/point-archival-gateway/archiving-directly-to-tape-via-s3/

[4] Grau Data - https://www.graudata.com/produkte/tape-object-archive/xtreemstore/

[5] IBM - https://www.ibm.com/products/deep-archive

[6] EPRI - Powering Intelligence: Analyzing Artificial Intelligence and Data Center Energy Consumption, https://www.epri.com/research/products/000000003002028905

using a publicly available TCO tool[7], if 100 PB of inactive data is stored on tape for ten years, 237 megawatt hours of power will be required. However, if the same 100 PB is stored on disk, 12,568 megawatt hours will be required. The tape solution consumes 98% less power over ten years.

Figure 1 illustrates the power savings.

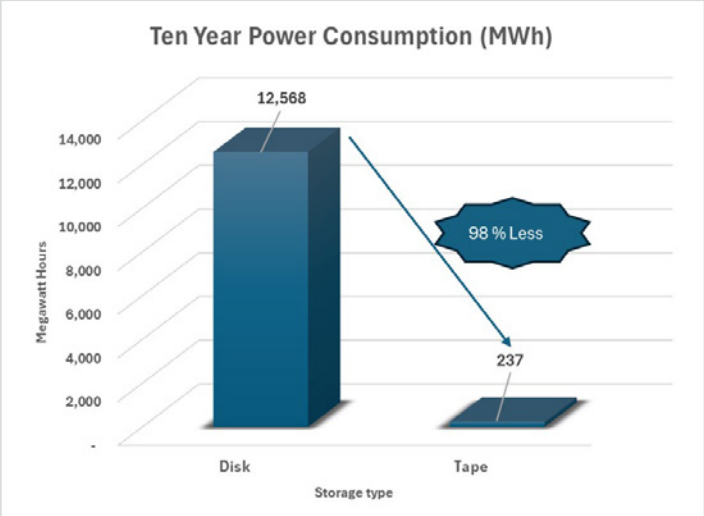

FIGURE 1 - POWER CONSUMPTION

Reducing Carbon Footprint and eWaste with Tape Storage

Power consumption is just one aspect of sustainability; CO2e emissions over the product lifecycle, including manufacturing, distribution, usage, and disposal, are also critical considerations in purchasing decisions. Tape storage can significantly reduce overall environmental impacts in the product lifecycle.

For example, compare the estimated CO2e emissions of storing 100 PB on HDDs to storing it on LTO 9 tape media. The estimate includes the CO2e generated due to energy consumption during the media's use and the emissions associated with acquiring raw materials, manufacturing, and the final disposal of the storage media. Storing 100% of the data on hard disk drives generates 2,663 tons of CO2e over ten years.[8] Fujifilm provided estimates for the LTO 9 tape media lifecycle CO2e emissions. The tape solution generated only 79 tons of CO2e, a 97% reduction.

In addition to lowering carbon emissions, moving data to tape substantially reduces the amount of electronic waste (eWaste). A ten-year retention requirement will require at least one HDD refresh, requiring the purchase of replacement HDDs in year 6. Based on historical capacity growth, we estimate the refresh will utilize 38 TB HDDs. Storing all 100 PB on hard disk drives and refreshing them after five years generates 8.2 tons of eWaste. Tape storage has a longer life and does not require a refresh. As a result, storing the data on tape produces only 1.2 tons of eWaste, an 85% reduction.

Figure 2 shows the eWaste reduction.

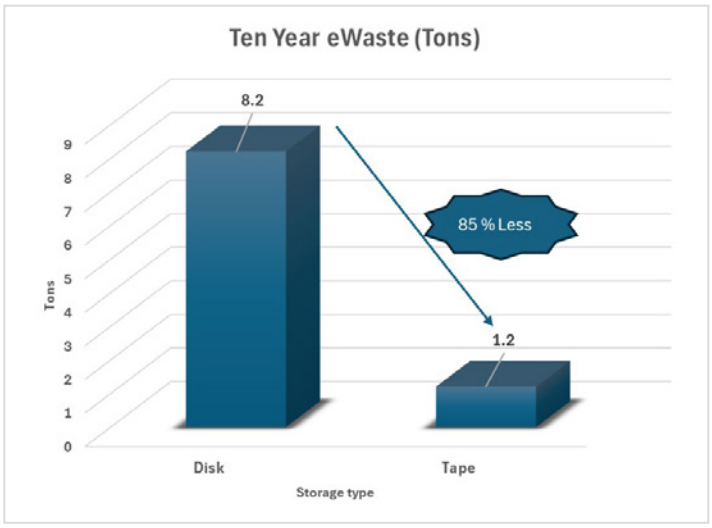

FIGURE 2 - eWaste

Strengthening Data Protection: The Role of Tape Storage in Cybersecurity

With the increasing number of cyber-attacks, protecting stored data is crucial to improving IT security. Tape storage offers significant protection from data loss or corruption. One of the key security advantages of tape storage is its "air-gapped" nature. Air-gapped means there is no direct network connection to the tape media. Access is only via a tape library or a manual process. Because tape storage is air-gapped, it mitigates the impact of a security breach, providing a robust layer of protection against cyber threats.

Best data management practices typically suggest keeping three copies of the data on two different media types: one disk resident copy locally, one disk resident copy remotely, and one tape copy. However, with increasing concerns for data protection, leading companies are now adding one additional tape copy completely offline and offsite. This strategy, in combination with encryption and WORM, provides robust data protection.

---

[7] Fujifilm TCO tool - https://www.fujifilm.com/us/en/business/data-storage/resources/tco-tool

[8] Fortunately, the data storage industry provides excellent data on the environmental impact of its products' manufacture, use, and disposal. The basis for the hard disk drive CO2e is the Seagate Exos X18 enterprise hard disk drive Seagate - https://www.seagate.com/global-citizenship/product-sustainability/exos-x18-sustainability-report/

Reducing Long-Term Storage Costs: Tape vs. Disk a Ten-Year TCO

With data growth exceeding 25% in the coming years, storage costs will continue to rise despite the declining $/TB storage cost. The total cost of storage includes not only the initial capital outlays but also annual maintenance and support charges, the energy cost for power and cooling, network bandwidth costs, and systems management personnel costs. In addition, the cost of technology refreshes must be included in a ten-year Total Cost of Ownership analysis.

Tape storage is significantly less expensive than disk storage over ten years. The analysis assumes the disk storage system will be replaced at the end of five years. However, the tape libraries remain on the data center floor for at least ten years. The original tape drives and media generations are upgraded after five years, offering denser storage without total replacement. Developing a ten-year TCO for storing 100 PB, the estimated tape storage TCO is $4.9 million, and the disk storage cost is $17.6 million. Tape storage is 72% less expensive than disk storage.

Figure 3 illustrates the Ten Year TCO.

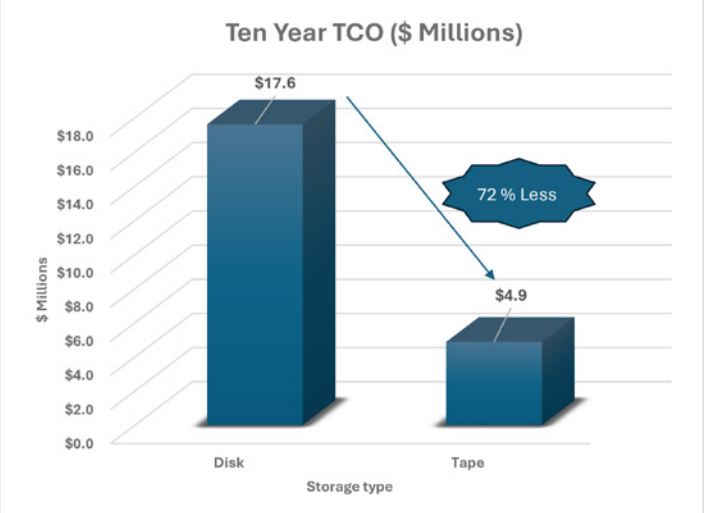

FIGURE 3 - TEN YEAR TCO

Executive Summary

In the current digital landscape, data creation is surging at an unprecedented pace, driven by technologies like AI, IoT, 5G networks, and more. This explosive growth presents a significant challenge for enterprises as they struggle to manage vast amounts of data while controlling costs and meeting sustainability objectives. Even though a substantial portion of enterprise data is inactive, much remains stored on power-hungry SSDs and HDDs, leading to unnecessary energy consumption and higher operational costs. Tape storage, though underutilized, offers a compelling solution to these challenges.

Data centers consume substantial amounts of energy, with storage systems accounting for a substantial share. Projections indicate that as AI and other technologies continue to evolve, the power demands of data centers could nearly double by 2030. Tape storage, however, requires significantly less energy than disk storage. For example, storing 100 PB of data on tape over ten years uses 98% less power than storing the same amount on disk. This dramatic reduction in energy consumption positions tape storage as a vital component for organizations looking to expand their data centers while managing power constraints.

Beyond power consumption, sustainability is an increasingly important consideration for enterprises. The lifecycle CO2e emissions of storage media, from manufacturing to disposal, are crucial. Tape storage reduces energy usage and significantly lowers carbon emissions and electronic waste. A ten-year analysis shows that using tape instead of HDDs to store 100 PB of data results in a 97% reduction in CO2e emissions and an 85% reduction in eWaste. These environmental benefits make tape storage an attractive option for organizations committed to reducing their carbon footprint.

In addition to its cost and environmental benefits, tape storage offers robust data security. The “air-gapped” nature of tape storage, which is offline and inaccessible via network connections, provides a strong defense against cyber threats. This security advantage, combined with best practices like maintaining multiple copies of data on different media types, makes tape storage a critical tool in safeguarding sensitive information from cyber-attacks.

Despite these advantages, the adoption of tape storage has been limited by the need for specialized software and skills. However, new solutions that integrate tape storage seamlessly into existing IT infrastructures using S3 object storage interfaces are emerging. These innovations lower the barriers to tape storage adoption, enabling organizations to take full advantage of its cost, energy, and security benefits without needing unique infrastructure or expertise. As data growth continues, leveraging these solutions will be crucial for organizations aiming to manage their data efficiently and sustainably.

Regarding the Information Used in This Report

Brad Johns Consulting LLC believes that the information in this report was accurate as of the publication date. Information is provided "AS IS" without warranty of any kind.

Trademarks and Special Notices

Linear Tape-Open, LTO, the LTO Logo, Ultrium, and the Ultrium logo are trademarks of HPE, IBM Corp., and Quantum in the US and other countries.

Other company, product, or service names may be trademarks or service marks of others.

About the Author

Brad Johns is the owner and President of Brad Johns Consulting LLC. He has over forty years of experience in the Information Technology industry. He specializes in storage industry economic analysis and consulting. He started his information technology career with the Data Processing Division of IBM in 1978 and held several sales management, consulting, and marketing positions. He joined the IBM Storage Systems Division in 1997 and was responsible for product management and marketing until his retirement in 2010. He holds an MBA and a BA in Economics from the University of Arizona.

Email: brad.johns@bradjohnsconsulting.com
Website: bradjohnsconsulting.com

Copyright © 2024 Brad Johns Consulting L.L.C