
# Long-Term Data Preservation: Comparing Tape and Cloud Storage

White Paper | Spectra Logic

# Table of Contents

## Abstract

As data retention requirements continue to grow, organizations must determine the best long-term storage strategy. When comparing cloud and tape storage solutions, factors such as reliability, security, cost, and accessibility must be considered to make an informed decision. This paper explores the evolving landscape of storage technologies and how a hybrid approach that leverages both cloud and tape can offer an optimal solution for many organizations.

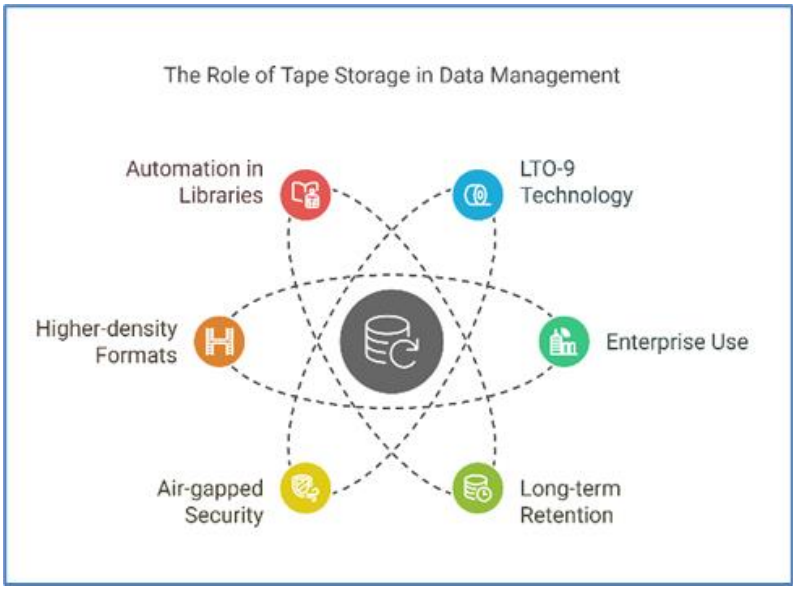
## Introduction

A glance at today's IT landscape suggests that cloud technology dominates the market. With major providers like Amazon AWS, Google Cloud, and Microsoft Azure continually innovating, cloud storage has become integral to enterprise data strategies. However, tape storage remains a crucial component in archival and backup solutions. Despite being declared "obsolete" multiple times, tape technology has quietly evolved, offering increased capacities, improved efficiencies, and cost advantages that challenge the all-cloud narrative. This paper evaluates cloud and tape storage to understand their roles in modern data preservation strategies.

## Perceptions of Leading-Edge Technology

The cloud has gained a reputation as the go-to storage solution for its agility, scalability, and accessibility. Enterprise cloud adoption continues to rise, with public cloud spending expected to exceed $1 trillion by 2026[1]. Cloud providers offer seamless integration with analytics, AI, and other services, making cloud storage a strategic asset beyond data archiving.

Meanwhile, despite significant advancements, tape storage is often seen as a legacy technology. LTO-9, for example, offers up to 45TB of compressed storage per cartridge, demonstrating tape's continued relevance. Many Fortune 100 companies still use tape for their data protection strategy, particularly for long-term retention and air-gapped security. The introduction of higher-density tape formats and automation in tape libraries has positioned tape as an attractive long-term storage medium for enterprises managing petabytes of data.




## Reliability

Reliability, stability and life expectancy are always top considerations for any organization when choosing a storage platform. Since the time data needs to be stored continues to expand, the need to store and retrieve data at any given point has become crucial to all businesses.

[1] Forrester Consulting, Cloud Costs Are Out of Control: Integration and Modernization Can Help Rein Them In, sponsored by Boomi, April 2, 2024.

Tape boasts an impressive lifespan of up to 30 years, making it one of the most durable storage options. Regular maintenance ensures that tape libraries can function indefinitely, with data migration enabling long-term accessibility. Meanwhile, the Amazon S3 Standard storage class is designed for 99.99% availability[2], but true reliability depends on redundancy across multiple cloud regions—a feature that increases costs.

While cloud storage ensures high availability, outages do occur. In 2022 alone, major cloud providers experienced downtime incidents that impacted enterprises worldwide. Data stored in cloud environments is only as reliable as the provider's infrastructure, requiring organizations to establish redundant storage strategies for critical information.

## Security

Security remains a primary concern for cloud storage. Despite advancements, cloud data breaches are still common, with high-profile incidents underscoring vulnerabilities in misconfigured services and unauthorized access. Ransomware continues to be a significant threat, and while cloud providers offer data recovery options, these often come with additional costs.

Tape storage, by contrast, provides inherent security through air-gapping, making it immune to network-based cyberattacks. Once a tape is removed from the library, it is physically inaccessible to hackers. Encrypted tape storage further enhances security, offering a robust last line of defense against data compromise. Many organizations now adopt a "3-2-1-1" backup strategy (3 copies of data on two different media types, 1 copy offsite and 1 immutable copy, ensuring that at least one copy of data remains offline and impervious to cyber threats.

[2] AWS, General S3 FAQs, https://aws.amazon.com/s3/faqs/, February 3, 2025

## Speed

Speed can be a significant factor in choosing storage. Backing up and archiving massive amounts of data can take hours for any project or set of files to be transferred to the storage medium.

Storage speed considerations differ based on workflow needs. Cloud storage offers convenience, but large-scale data transfers depend on internet bandwidth, creating bottlenecks. While cloud retrieval times vary based on the service tier, some deep archive options may require hours or days to access stored data.

Modern tape solutions have significantly improved write speeds, with LTO-9 achieving up to 900 MB/s (compressed). High-performance tape systems can efficiently handle large-scale backups faster than the cloud for sequential data writes.

## Storing Data

Cloud storage is perceived as fast and agile, but to varying degrees, there is limited control and flexibility for how data can be uploaded. Uploading an initial dataset to the cloud depends on how much bandwidth an organization has and how much bandwidth is available to the cloud provider. Cloud providers, or backup software providers that write to the cloud, can quickly update stored data. Depending on the need, cloud bandwidth can be scalable to accommodate larger file transfers into the cloud. Still, the speed of the archive job is typically based on the organization's internet bandwidth. On average, if a business were transferring the same 60TB to the cloud instead of tape, with a 1Gbps bandwidth connection, this transfer would take over six days to archive. That is assuming the 1Gbps connection was dedicated entirely to this transfer and not being used for anything else, and the 1Gbps connection was completely stable for six days.

Most of the time, a business can only use a quarter of its bandwidth for uploads to the cloud. With a dedicated 250Mbps connection, 60TB would take 25 days—a whole month—to place that project into the cloud! A 1PB upload into the cloud on the same 250Mbps connection could take 418 days, or one year and two months, to complete!

Previously, tape had a reputation for being slow. This may have been true 40 years ago, but tape has advanced exponentially like all technologies. Today, tape can transfer data extremely fast, and like any on-premises device, its speed is scalable. Organizations have complete control over upgrades to the tape library and can add more tape drives and expansion frames to enhance performance. In fact, for complete write workflows, tape even outpaces disk with speeds ranging from 400 MB/s uncompressed up to 900 MB/s per drive with compression. A modest 24-drive installation can write over 75TB of data in an hour, 1.8PB in less than a day, and 13PB in a week with compressible data at a 2.5:1 compression ratio.


Cloud storage offers flexibility, but costs escalate with increased data volume and access requirements. While the cloud provides seamless scaling, managing expenses requires careful planning to avoid unexpected retrieval fees. Conversely, tape storage offers a predictable cost structure. Large-scale data storage on tape remains cost-effective, particularly for cold storage applications where frequent access is unnecessary.

The cost of storing data in cloud environments can be misleading. While initial storage costs appear low, access fees, retrieval times, and data movement costs must be considered. Organizations with extensive archival needs often find tape storage to be the most economical option over time.

## Retrieving Data

While storing data is straightforward, retrieving data can be a bit more complicated. Typically, the speed of data recovery depends on how much data needs to be recovered. In the event of disaster recovery, tape is the fastest option. Retrieving data requires loading the tape into a tape drive, reading the data, and transferring it over the network to the designated source. Tape could take longer if a set of specific files are being recalled. The needed files could be written on multiple tapes, and since tape is read

linearly, the required file could be at the front or the end of the tape. If the tape is stored offsite, this will add a few hours to the retrieval time. There are modern applications, like Spectra BlackPearl® Object-based Tape, which creates objects and provides a type of on-premises cloud that helps the library identify where a specific file is located, which speeds restore times drastically. Retrieving data from tape requires a structured process, with files restored sequentially. Newer tape management systems improve retrieval efficiency, but tape remains better suited for bulk data recovery rather than ad-hoc file access.

The retrieval of archived cloud data has historically been more complicated than portrayed. Cloud storage retrieval varies by tier. Services like Amazon S3 Glacier Deep Archive offer low-cost storage but can take up to 12 hours for retrieval. Higher-tier cloud storage allows for faster access but at significantly increased costs. Many organizations have begun implementing a hybrid model, where frequently accessed data is stored in the cloud while long-term archives remain on tape.

The table below outlines Amazon's lowest-cost storage tiers and the SLA for accessing data, both in time and cost. Once a restore request has been made, downloading and accessing the data will take the time listed.

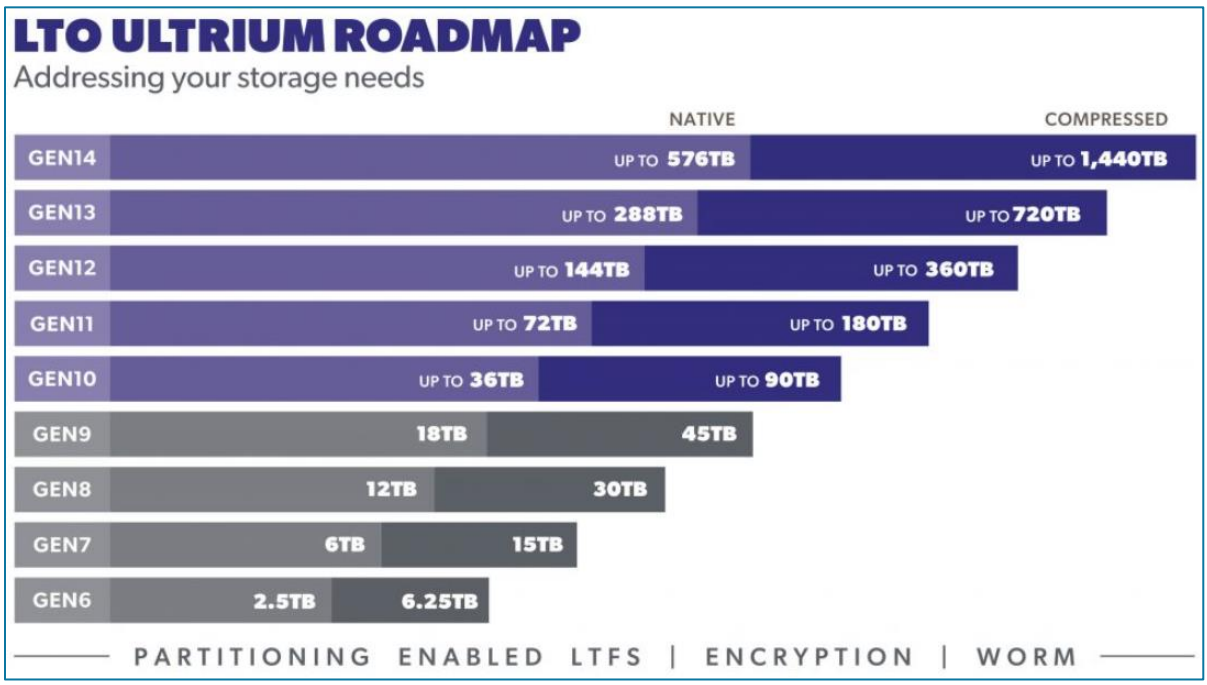
| **AWS Storage and Retrieval Costs** | **S3 Glacier Flexible Retrieval** | **SE Glacier Deep Archive** |
|---|---|---|
| Time to Access | 1 to 12 hours | within 12 hours |
| Availability | 99.99% | 99.99% |
| | | |
| **Storage** | | |
| Cost per GB | $0.0045 | $0.0020 |
| Cost per TB | $4.50 | $2.00 |
| | | |
| **Retrieval** | | |
| Cost per 1,000 objects | $0.033 | $0.06 |
| Cost of Restore per TB | $11.00 | $22.00 |

Source: AWS S3 Pricing, Region - US West Region (N. California), https://aws.amazon.com/s3/pricing/, Feb 3, 2025.

## Data Growth

Data storage technology is improving year over year, and tape technology is continuously increasing its density every few years. The improvements are exciting but can lead organizations down a confusing path of navigating upgrades and cycling outdated storage technology and media.

In the next two years, many organizations are expected to quadruple the amount of data that needs to be stored. This anticipated increase caused tape technology experts to double down on efforts to maximize the density of the media. In late 2021, LTO-9 tape media was released with an astonishing 18TB per tape cartridge, a fifty percent capacity increase over LTO-8 media, which was 12TB, native. LTO-10 will again substantially increase the per-cartridge capacity. During these advancements, there is confusion about when to buy new media and drives and when to transfer data from old media onto new media. Another potential complication is when and how to upgrade the physical hardware. Upgrades, from tape drives to the whole library, can feel costly and complicated.




Source: LTO Ultrium website, February 3, 2025

While it may seem counter-intuitive, waiting to upgrade tape media is recommended. Tapes have a shelf life of 30 years, and modern libraries are designed to have near-perfect conditions for tape media. Tape drives can also read and write back at least one generation; e.g., LTO-9 tape drives can read

LTO-7 and LTO-8 media. While this also means that organizations would need to keep older tape drives in the library, most tape libraries support multiple generations of tape drives. Many companies wait for the cost of a new generation of tape media to come down to a reasonable price per gigabyte and then purchase new drives and media. By utilizing this method, a slow upgrade can occur over time instead of shutting down production on the tape library to write data to new media every three to five years.

Many organizations think that by utilizing cloud storage, they won't have to think about the ever-changing state of storage media because the cloud service provider will take care of everything for the business. This is only partially true. By utilizing the cloud, there is no need to worry about an on-premises storage device. Organizations don't have to build a plan to transition to new generations as they come out. The drawback is that all cloud companies utilize physical storage devices like disk and/or tape to store data.

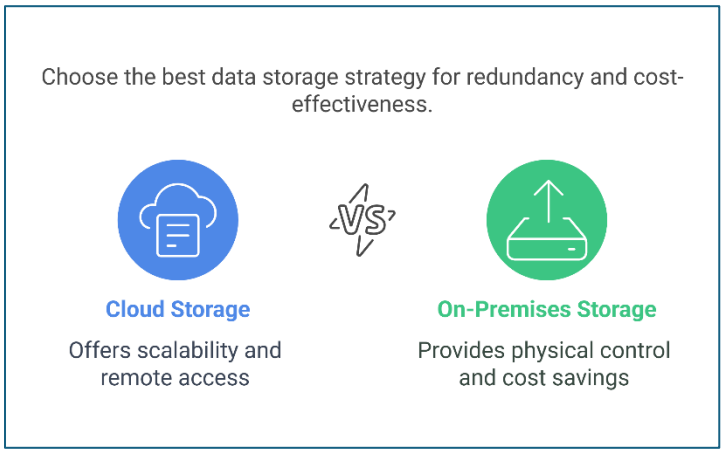
This affects businesses because, after a specific tier of storage or percentage of physical storage space is hit, most cloud providers have a new pricing structure that is more expensive than the previous one.

Moreover, most cloud storage costs are based on the total capacity of files, the total number of files being stored, retrieval and access costs, transfer and transfer acceleration costs, and replication costs. All these various categories also have costs that vary by region, so cloud costs for the eastern U.S. could differ from those for the western U.S. For example, Amazon S3 has six different pricing tiers for "standard" storage in the United States. Then, based on the storage method chosen, there are another five billing categories for data access, including whether an account moves from infrequent storage access into standard access. An account can transfer from infrequent access to standard access for two reasons: (1) the organization opts to transfer to standard access for faster access speeds, or (2) the business requests too many files in a month and is automatically transferred into standard access. The second scenario perfectly shows how a hybrid cloud and tape environment can save organizations money. Data can be stored in the cloud at the deepest level, which is what Amazon Glacier does, and the same data can also be stored on tape. If a sizable portion of data needs to be retrieved, it can be retrieved from the tape copy at no additional cost without the risk of being moved to a more expensive tier of cloud archive.

## Redundancy

Redundancy must always be factored into data growth calculations. Multiple copies are essential for disaster recovery and critical for any data storage failure.

Storage failure can happen on any media and can affect a single file up to multiple terabytes of data. With tape storage, redundancy is easy. A set of files is written on two separate tapes. Ideally, one of these tapes is kept on-premises, and one is stored offsite in a secure location. If corruption occurs on a set of files, the secondary tape can be returned onsite and used to restore the data.




Cloud storage can be as easy as putting two copies in the cloud and hoping they are stored in separate regions or locations with both copies remaining uncorrupted. However, the uncertainty has led many users to store a single copy in the cloud and store a copy on another storage media such as tape or disk. Putting multiple copies in the cloud can also impact pricing tiers, with twice the amount of data in the cloud doubling the price. Many experts suggest that redundancy is necessary and always recommend that an on-premises storage option is best paired with cloud storage.

Copies can be retrieved from cloud storage if a disaster impacts the data center. Once the data center is up and running, the data can be duplicated from the cloud and put back onto tape. Without a cloud mandate, organizations can still achieve best practices in data storage while saving money by leveraging tape: (a) making two copies to tape, with one ejected and taken offsite, or (b) making two copies to tape at two different customer-managed locations.

Modern object storage systems now allow seamless integration between cloud and tape, enabling businesses to automate redundancy policies effectively. Organizations seeking optimal data protection invest in solutions incorporating both technologies, leveraging their strengths.

## Expected Future Costs

Total Cost of Ownership (TCO) is a key differentiator between tape and cloud. Tape remains the most cost-effective solution for large-scale, long-term storage. Studies show that tape storage costs can be as low as 1/5th of cloud storage[3] over ten years.

While cloud pricing has been stable in base storage rates, retrieval and egress fees can significantly impact budgets. When calculating cloud expenses, organizations must account for hidden costs such as API requests, replication fees, and long-term retention charges.

With Tape Archive Platform-as-a-Service (TAPAS) from Spectra Logic and Geyser Data, organizations have even greater choice with the combined benefits of traditional tape technology in a cloud-based, as-a-service model. TAPAS provides seamless integration with Amazon S3 and S3 Glacier, offering superior performance without requiring specialized skills or upfront investments. This cost-effective storage solution ensures secure, long-term retention with the freedom of immediate access—without the burden of hefty egress or restore fees.

[3] LTO Ultrium TCO Calculator for Data Storage Solution, comparison of LTO-9 versus Amazon s3 Glacier Flexible Retrieval for 1PB, 10-year retention, 10% growth, and 1% monthly data retrieval.

## Parting Thoughts

While cloud storage is often the preferred choice for accessibility and scalability, tape offers unmatched cost efficiency and security. A hybrid approach that leverages both strengths can provide organizations with a flexible, resilient, and cost-effective data preservation strategy. Whether for compliance, archival, or disaster recovery, the right combination of tape and cloud storage ensures that data remains protected, accessible, and affordable for years to come.

When trying to find the right storage strategy, the answer comes down to the ideal storage solution that optimizes each medium's strengths while reducing storage failure risks. Tape and cloud together behind an object storage device is the most potent combination. With the object storage system handling the workflow, tape as a primary backup, and the cloud as a disaster recovery option, data can be protected elegantly, seamlessly and cost-effectively from unfortunate circumstances. If this hybrid combination cannot be deployed, today's tape storage is the recommended platform for an economical long-term data storage and archival approach.

## About Spectra Logic

Spectra Logic modernizes IT infrastructures to preserve, protect and defend data from days to decades, whether on-premises, in a single cloud, across multiple clouds or in all locations at once. Our cost-effective solutions help organizations efficiently manage, migrate, and store long-term data, from terabytes to exabytes, with features that make ransomware resilient.

To learn more, visit www.SpectraLogic.com


Toll Free: 800-833-1132 • 303-449-6400 • 6285 Lookout Road • Boulder, CO 80301 USA
V3-030325