

# LONG-TERM CLOUD STORAGE: WHAT IT IS AND WHY YOU NEED IT



Enterprises are generating huge volumes of data every year, with an average annual data growth of 40-50 percent. This growth has to be handled using IT budgets that are only growing at an annual average of 7 percent. Such disproportion creates a challenge for mainframe professionals. How can they store all this data cost-effectively?

Particularly challenging is deciding on the right strategy for long-term storage, also known as cold storage, for archived data that is rarely or never accessed. There can be different causes for keeping such data for the long term, which often lasts years or even decades:

- Financial data is stored for compliance and might be required in case of an audit.
- Legal information must be kept in case of legal action.
- Medical archives are stored in vast quantities and their availability is highly regulated.
- Government data has to be stored for legal reasons, sometimes even indefinitely.
- Raw data is stored by many enterprises for future data mining and analysis.

## Desired attributes of a cold storage solution

Cold storage, also referred to as “Tier 3 storage,” has different needs than Tier 0 (high-performance), Tier 1 (primary), and Tier 2 (secondary) storage. These are some of the considerations to keep in mind when designing your cold storage solution:

- **Scalability** – As the amount of generated data on average doubles in less than two years, your cold storage technology accordingly needs to be infinitely scalable.
- **Cost** – Cold storage must be as inexpensive as possible, especially because you will need a lot of it. Luckily, as it is rarely accessed it allows compromising on accessibility and performance, which can be leveraged to reduce cost.
- **Durability and reliability** – Reliability is the ability of a storage medium not to fail within its durability time frame. Both are important to check, and you will find that some cold storage options are durable but not necessarily as reliable as others, and vice versa.
- **Accessibility** – Cold storage is meant only for data that does not need to be accessed very often or very rapidly, yet the ability to access it is still important. As mentioned above, compromising on this aspect enables a lower cost.
- **Security** – The security of cold data is vital. If it is stored onsite you need to take the same security precautions as with your active data. If it is in the cloud, you must ensure the vendor has proper security mechanisms in place.

## Cold storage technology options for mainframe

Mainframe professionals have three general technology options when it comes to cold storage: tape, virtual tape, and cloud. While tapes are still the dominant cold storage media for mainframes, cloud is gaining momentum with its virtually limitless storage and pay-as-you-go model.

Here is a summary of these technologies, and their relative advantages and disadvantages:

### Tape

Tape drives store data on magnetic tapes and are typically used for offline, archival data. Despite many end-of-life forecasts, the tape market is still growing at a compound annual growth rate (CAGR) of 7.6% and is expected to reach \$6.5 billion by 2022. Tapes are considered the most reliable low-cost storage medium and, if maintained properly, can last for years.

However, they are also the most difficult to access and it can be quite an ordeal to recover from tapes in case of disaster.

### Pros of Tape:

- Often cheaper than other options, depending on the use case.
- Full control over where data is stored.
- Secure and not susceptible to malware or viruses as they are offline.
- Portable and can be carried or sent anywhere.
- Easy to add capacity.

### Cons of Tape:

- Capital investment required for large tape libraries.
- Difficult to access (slow and with bottlenecks).
- High recovery time objective (RTO).
- Requires physical access and manual handling (problematic in lockdown, for example).
- Requires careful maintenance.

# Virtual tape libraries (VTL)

A VTL is a storage system made up of hard disk drives (HDDs) that appears to the backup software as traditional tape libraries. While not as cheap as tape, HDDs are relatively inexpensive per gigabyte. They are easier to access than tape and their disks are significantly faster than magnetic tapes (although data is still written sequentially).

## Pros of VTL:

- Scalability – HDDs added to a VTL are perceived as tape storage to the mainframe.
- Performance – data access is faster than tape or cloud.
- Compatibility – works with tape software features like deduplication.
- Familiarity – behaves like traditional tape libraries.

## Cons of VTL:

- Cost varies. Infrastructure, maintenance, and skilled admins should also be considered.
- Capital investment required.
- Usually less reliable than other options.
- Less secure than offline tapes and lacks the latest security features of cloud platforms.

# Cloud Storage

Cold storage in the cloud is maintained by third-party service providers in a pay-as-you-go model. Rather than selling products, they charge for usage of storage space, bandwidth, data access, and the like.

Cloud is becoming extremely popular for cold storage, mainly because it is considerably cheaper than on-premises storage. Pay-as-you-go means that it can start at affordable prices without needing to stock up on tapes and VTLs. Also, there is no longer a need to maintain infrastructure or recruit personnel to manage data archives, as these are all handled by the cloud vendor.

The cloud provides superior agility and scalability, and although magnetic tapes are more secure, it also provides higher levels of security and compliance than many businesses can on their own. When it comes to durability, the cloud really excels by storing data redundantly across many different storage systems.

On the downside, administrators need to consider network bandwidth and the cost of uploads and restores, as using cloud is often more expensive than it appears at first glance. The leading vendors of long-term cloud storage are Amazon (Glacier and Glacier Deep Archive), Google (Cloud Storage Nearline and Cloud Storage Coldline), Microsoft (Azure Archive Blob Storage), and Oracle (Archive Storage). These vendors charge low rates for storage space but extra fees for bringing data back on-premises, which might prove costly if too much data is retrieved.

## Pros of Cloud:

- Can be cheaper, especially when being aware of hidden costs.
- Can improve cash flow thanks to an operating expenses (OpEx) financial model rather than capital expenditure (CapEx).

- Infinitely scalable.
- Accessible from anywhere.
- Advanced data management.
- High data redundancy and easy replication.
- Leading-edge security.
- Easy to integrate with mainframes.

## Cons of Cloud:

- Hidden costs (depends on use).
- Data retrieval, backup, and RTO times depend on network bandwidth.

## Cloud is rising as a mainframe cold storage choice

Cold storage in the cloud offers a unique combination of scalability, reliability, durability, security, and cost-effectiveness that on-premises options are challenged to meet.

So, in which cases cloud is preferable for cold storage over tape and VTL?

- **When data access frequency changes:** The cloud offers different cold storage tiers, based on the data access requirements, that balance between data storage cost and the data access frequency. Cold storage tiers can be cost effective, however with high data access frequency you need to be mindful of choosing a service that addresses those access needs.
- **When the data grows quickly or unpredictably:** Cloud platforms can scale to infinity with very little effort, unlike on-prem options.
- **When improving cash flow is a priority:** Predictable OpEx monthly fees can improve cash flow compared to large upfront investment in on-premises storage and infrastructure.
- **In case of mainframe skills shortage:** Attracting and retaining mainframe experts is a challenge to many enterprises. With cloud cold storage, this problem completely goes away.