

SNIA DEVELOPER CONFERENCE



*BY Developers FOR Developers*

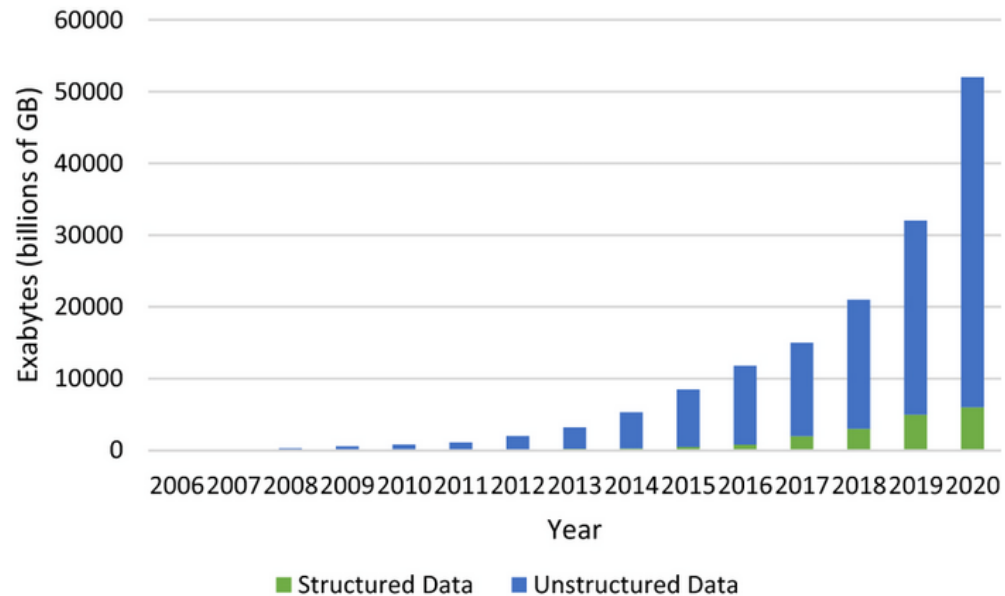
September 16-18, 2024  
Santa Clara, CA

# IBM Deep Archive

Revolutionizing Data Archiving

Jason Peipelman

# Unstructured data growth and cost of storing it in the cloud

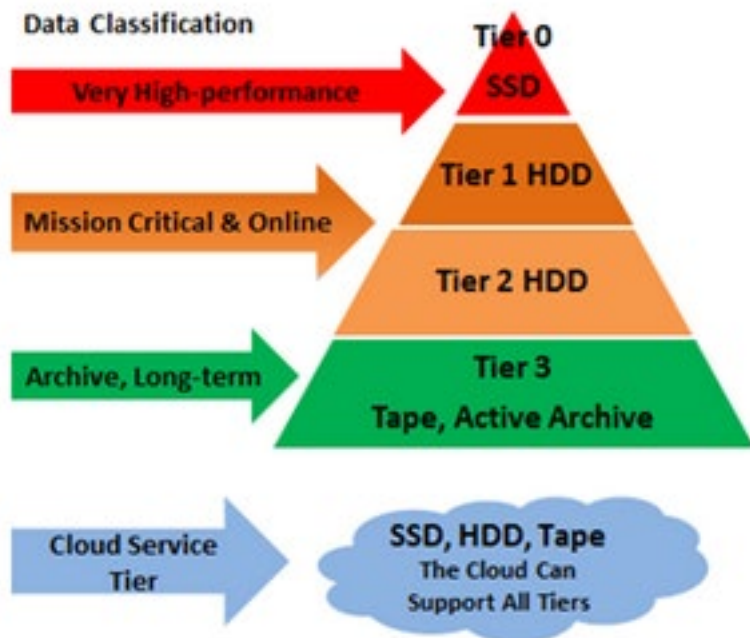


## Unstructured data

Information that is not arranged according to a preset data model or schema.

- **80-90%** of Data is Unstructured
- Unstructured Data is Growing **55-65%** Annually
- **43%** of IT Decision Makers fear their IT infrastructure won't handle future data demands

# Unstructured data growth and cost of storing it in the cloud

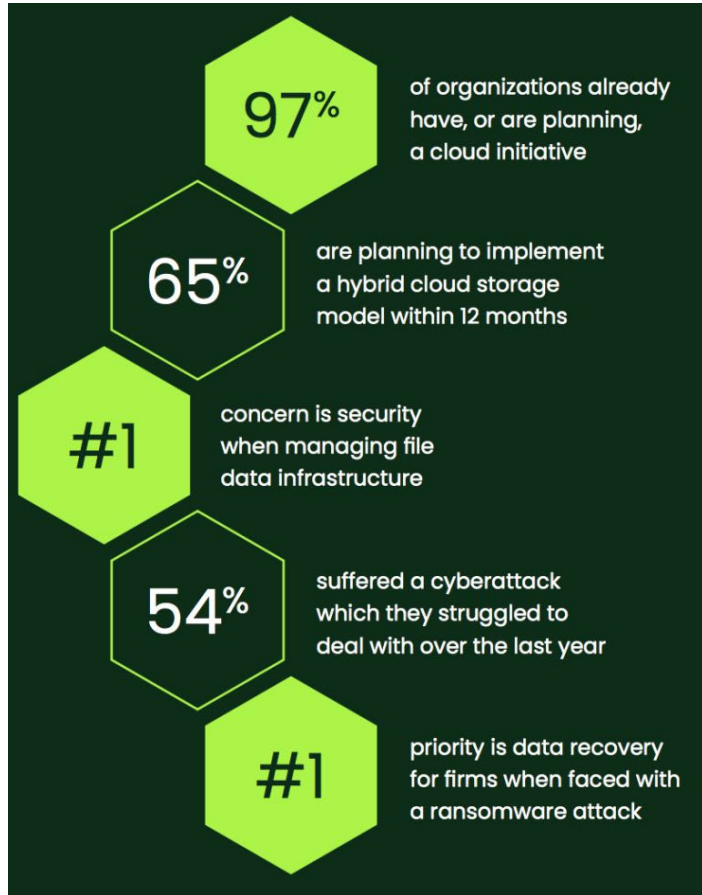


## Active archive

Moving data to a secondary storage medium that can be readily accessed if required.

- Required for compliance & security
- Integrates with data governance policies
- Allow backup & archive for high availability and disaster recovery
- S3 Glacier storage classes are being widely adopted as the industry standard interface for cloud based archival solutions.

# Unstructured data growth and cost of storing it in the cloud



## Cloud archival storage

90% of IT Decision Makers say cloud is a “cornerstone” of their digital strategy.

- 94% say their cloud storage costs are rising, and 82% say they incur unnecessary and unpredictable cloud costs like data egress fees.
- [82% of IT Decision Makers stated their biggest cloud challenge was managing cloud spending](#)

# Active Archive Expectations Survey, December 2023

## What people want in an Active Archive

- 1. Integration** Standard, S3 compliant interface allows easy integration into existing infrastructure and software.
- 2. Cost** Cost benefits that ensure stability in pricing and no unexpected fees.
- 3. Ease of use** A system that doesn't require advanced tape knowledge to run and can be configured operated by an S3 administrator.
- 4. Black box** A self-contained solution with a simple set of options that comes with everything you need.
- 5. Resiliency** Data resiliency that ensures the protection of data is prioritized over data access.
- 6. Security** Security by design is implemented from the start including encryption, monitoring, and air-gapped access control.
- 7. Sustainability** Must help to meet reduced energy consumption, carbon footprint, and cooling requirements



# IBM Deep Archive

For organizations looking for cost-effective long-term data storage in an integrated solution, **IBM Deep Archive** delivers the benefits of tape storage in an easy-to-use, easy-to-integrate solution in the footprint of a **single 19" rack with up to 27 PB of capacity and up to 4.6 TB/hour throughput.**

This solution can be used directly by your existing S3 Glacier Flexible Retrieval storage class compatible software to reduce the cost of your cold data by up to **80%**, with no data egress fees.

A black-box experience ensures complexity is hidden and internally optimized so that the entire solution is supportable by S3 administrators with **no technical tape experience.**





# Benefits

## Cost benefits

- Up to **80% lower TCO** than comparable cloud storage.
- No data egress fees.

## Standard interface

- Industry standard S3 API to tape allows leveraging of existing data movers and S3 software.
- Compatible with the **S3 Glacier Flexible Retrieval storage class**.

## Simplified management

- S3 administrators require **no specialized knowledge of tape storage**.
- Installed in less than a day.
- No tuning required, plug-and-play.

## Black-box solution

- Comes with everything you need from a single IBM seller
- Storage Expert Care provided for the entire solution.
- No FC or SAS external to the system.

## Sustainability

- Object storage energy consumption reduced by up to 97%
- 85% lower carbon footprint than comparable hard disk solutions.

## Security

- Keep your data your data by ensuring it is kept behind your firewall.
- Consistently apply security, encryption, system monitoring, and air-gapped control access.



# Configuration

## Server Hardware

- 1 Control Node

## 4 - LTO 9 Tape Drives

- 4 drives
- Up to 4.6 TB/hour.

## LTO 9 Tape Cartridges

- 1548 cartridges maximum
- Up to 27 PB of storage



**4.6 TB / hr**  
**23.34 watts/PB**

## Uses

- Deep archives
- Air gap
- Second/Third copy
- Regulatory compliance



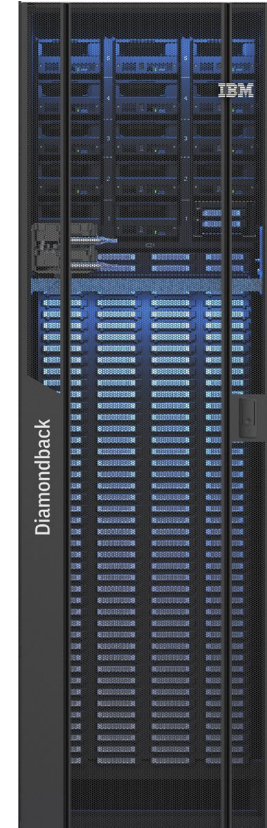
# Tape Library

## Simplified architecture

- Single-frame scale-out infrastructure
- Single robot, high availability grippers
- 1 service magazine with 10 I/O slots
- Up to 27 PB uncompressed capacity
- Up to 1548 Tape cartridges
- Up to 14 LTO-9 Tape drives

## Enhanced capabilities

- Cost kept to a minimum
- Fast deployment, <1-hr install
- Pre-loaded media option
- Storage pod design no regular media removal
- Increased data density and energy savings
- Quick Mean Time to Repair



# Capacity options

- Up to 14 PB preloaded media\*

800 cartridges maximum

- Up to 27 PB preloaded media\*

1548 cartridges maximum



\*represents maximum licensed cartridges, less can be installed if needed.



# Integration

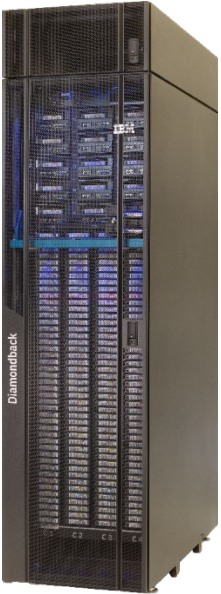
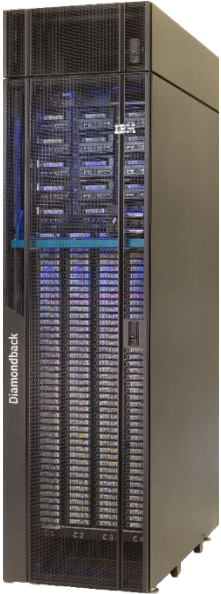
**Customer provided S3 orchestration**

**Deep Archive**  
S3 Glacier

**Deep Archive**  
S3 Glacier

**Primary Tier**  
File or S3 Standard

**Off-prem Cloud**





# Install and configure

IBM install



1. Install and configure the tape library
2. Install and cable the S3 nodes
3. Update code on the nodes
4. Update firmware on library and drives
5. Configure network settings

Cabling



1x - Management network  
2x - Data (S3) network  
2x - Power

1x - Library connection  
1x - iDRAC connection  
1x - TSSC connection

Optional

Customer configuration



6. Setup user accounts
7. Create a bucket

# S3 Glacier Flexible Retrieval

- S3 applications must use the **S3 Glacier Flexible Retrieval** storage class to read and write data.
- **Writing data** is first staged on the TapeCloud nodes. The response for the S3 API is returned and within 10 minutes the data is automatically migrated from the nodes to the tape library. Multipart upload is recommended with object sizes from 100 MiB to 5 TiB.
- **Restoring data** in an S3 Glacier Flexible Retrieval storage class are not accessible in real time. The AWS RestoreObject request is required to create a temporary copy of the object available on the nodes. Restore times include the mount and seek times for the tape cartridges that contain the data and the restore state can be monitored using the Head command.



## Supported maximums

• Object size	5 TiB
• S3 accounts	100
• S3 buckets	1,500
• Objects	1.5 billion
• Objects per S3 bucket	1,000,000
• Recommended average object size	100 MiB

# TapeCloud Manager

## Monitoring

System health  
Event notifications  
Capacity metrics  
Throughput metrics  
TapeCloud nodes  
Tape libraries  
Tape drives  
Data cartridges  
Cleaning cartridges  
Node hardware

## Management

System updates  
Library/drive firmware  
Certificates  
User accounts  
Library user accounts  
Buckets  
Software logs  
Library/drive logs  
Node logs  
Encryption settings  
NTP/time settings  
Network settings

## Service actions

Start/stop/restart nodes  
Power off/reboot nodes  
Rescan library  
Disable/enable drives  
Replace data cartridges  
Replace cleaning cartridges  
Validate cartridges



# Total cost of ownership

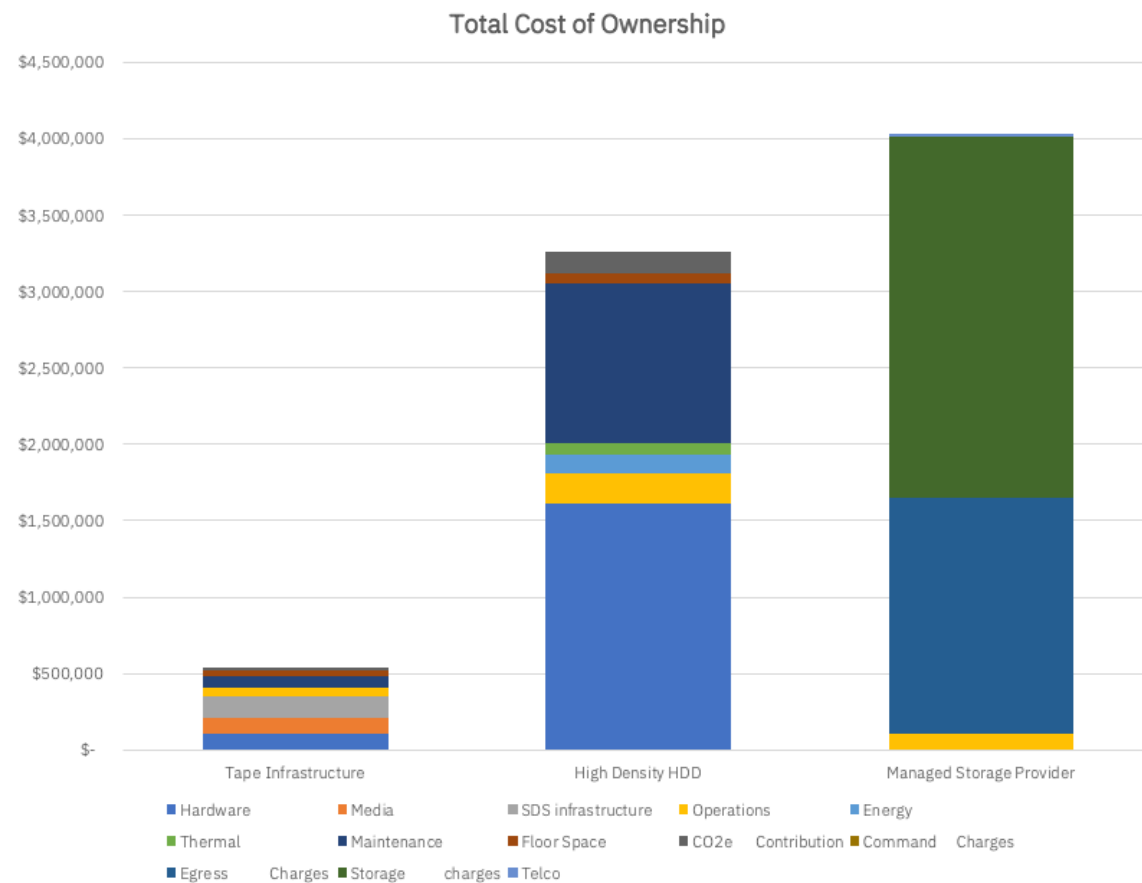
Up to  
**85%**  
Cost savings

Comparison:

- 27 PB
- 10-year retention
- AWS S3 Deep Archive w/ 1% recall/month (MSP)
- 16 TB drives w/ 5-year refresh cycle (HDD)

**AWS S3 Glacier Flexible Retrieval is 3.6x more expensive than AWS S3 Deep Archive with even greater retrieval costs**

...so, it would be off the charts.





# Positioning

## Deep Archive

- Active archive for infrequent access
- Object only with S3 Glacier storage class
- User/application accessible data archive
- Low-cost, high-capacity
- 3 PB minimum, 9+ PB optimal

## IBM Storage Scale + CES

- Transactional storage for AI and data analytics
- File and object (S3 Standard storage class) storage
- Tiered storage for user object/file frequent accessibility
- Full feature information lifecycle management platform
- Fully automated data durability platform

# Learn more!

<https://www.ibm.com/products/s3-deep-archive>

A low-cost, on-premises cloud providing secure and durable storage for data archiving and online backup

Read the data sheet



Contact us



Click here to schedule a meeting with us!



# Please take a moment to rate this session.

Your feedback is important to us.