Scalable, High Performance S3 Object Storage on Tape Media

Horst Schellong

MSST - May 22, 2023
Santa Clara University, CA

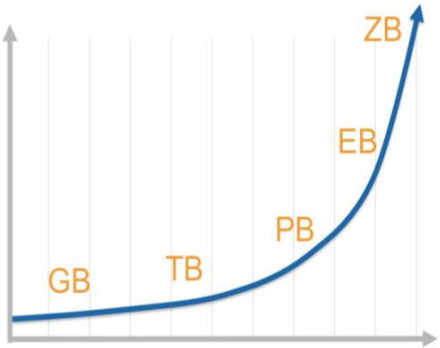
PoINT Software & Systems

Overview:

- Motivation: PoINT Archival Gateway (PAG) - why Object Storage on Tape
- Advantages
- System components
- Architecture
- Erasure Code options
- PoINT Archival Gateway – Unified Storage: storage classes
- Supported hardware
- AWS compatible Lifecycle Policies
- Direct access (S3) and restore (Glacier API) support
- Use cases
- Benefits
- Case study
- Q&A

Tape-based Object Storage

Motivation



Exponential data growth
in unstructured file data

80% of corporate data is unstructured file data

Hot	Warm	Cold
Active 10%	Less Active 20%	**Inactive 70%**

Archiving required for legal and compliance reasons

Tape-based Object Storage

Motivation

Storage classes at AWS, Google etc.

AWS	AWS cost/month	Retrieval	Google Cloud
S3 Standard	$23/TB	ms	Multi-Regional / Regional
S3 Standard-IA	$12.5/TB	ms	Nearline
S3 Glacier	$3.6/TB	1 min to 12 hrs	Coldline
S3 Glacier Deep Archive	$0.99/TB	Within 12 hrs	"Ice cold archive storage"

Tape !

Mark Russinovich, Microsoft Azure CTO:
*"After evaluating various technologies including Blu-ray and magnetic disk, Microsoft came to the conclusion that **tape** was still the way to go for Azure's **archival storage tier**"*

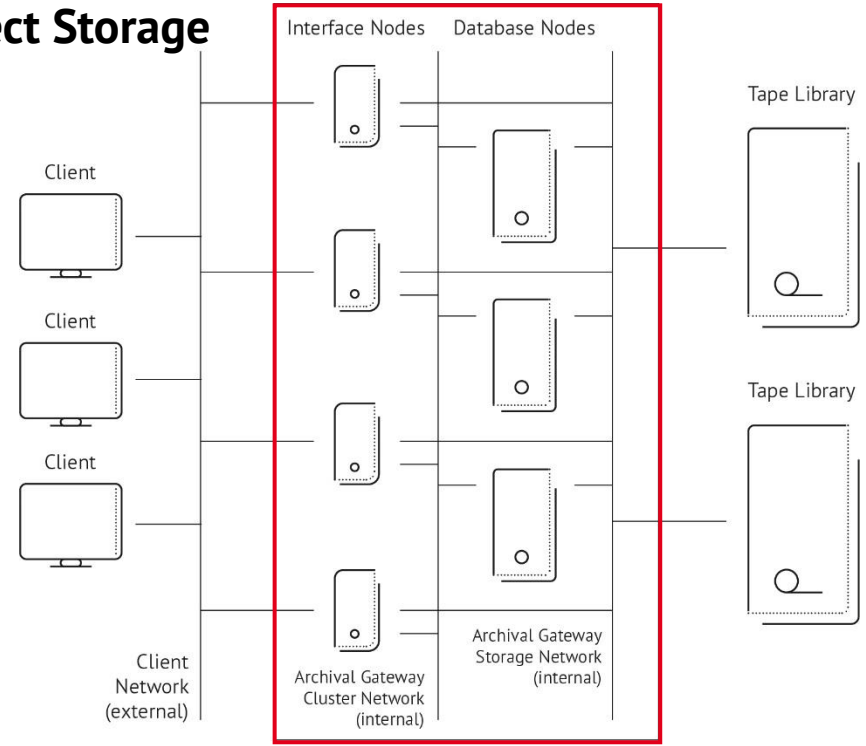
Advantages of Object Storage Approach for Tape

- Designed for massive amounts of data without decrease of efficiency
- Ingest of complete objects - perfectly supports sequential character of tape
- Scalability and system security by adding nodes
- Rich custom metadata (in contrast to file systems)
- Immutability (Versioning, Object Lock) – fits well with tape
- Redundancy by Erasure Coding – can be realized with multiple tape media
- Standardized S3/HTTP application interface
- HTTP can easily be routed and connected to any networked system
- HTTP concept simplifies handling of tape-specific challenges like high access times and timeouts

- **Don't tier or replicate data from object storage to file system**
- **Don't put S3 service on top of a file system**

Tape-based Object Storage

System Components



Interface Nodes (IFN)

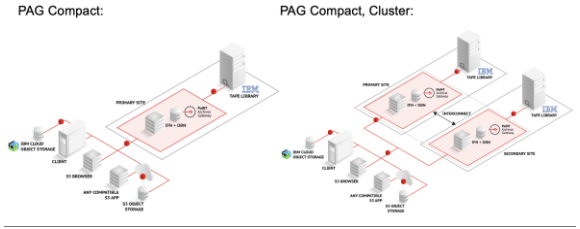
- S3 REST API
- Library Drives Control
- Up to 8 drives
- Multiple active nodes

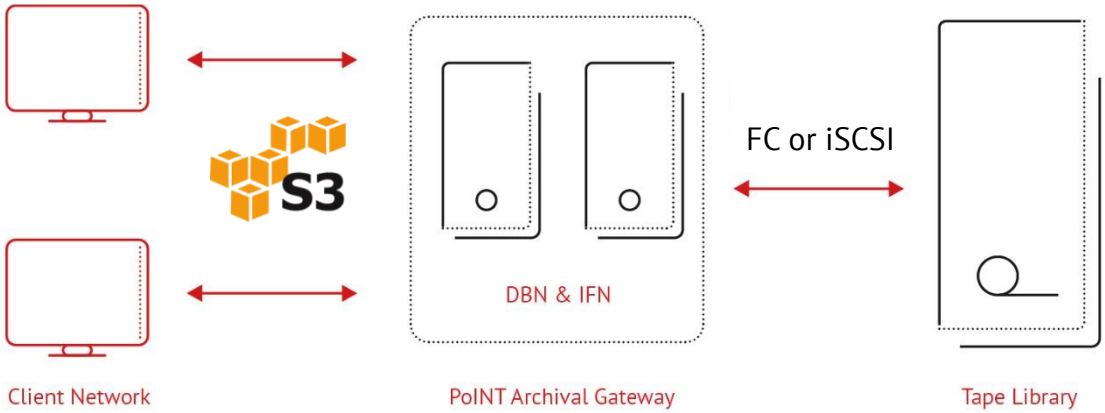
Database Nodes (DBN)

- Databases and Logs
- HTML admin interface
- Library Robotics Control
- Up to 4 nodes possible

Single Node Solution (PAG Compact)

- DBN and IFN services on one node
- Easy deployment





Tape-based Object Storage

Architecture

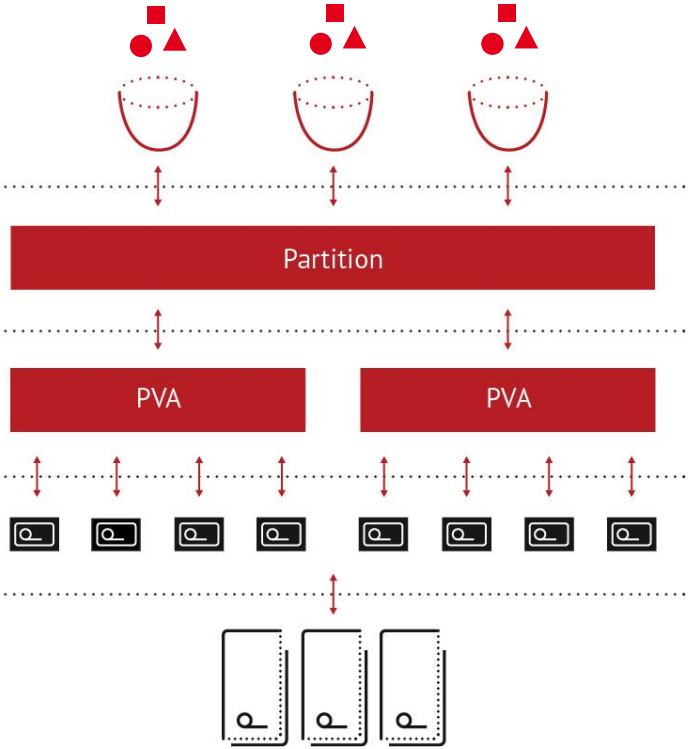
Object Repository (Bucket)

Archival Storage Partition
- Space provided by PVAs
- Expands automatically

Protected Volume Array (PVA)
- 1 - 4 volumes
- Protected by Erasure Codes

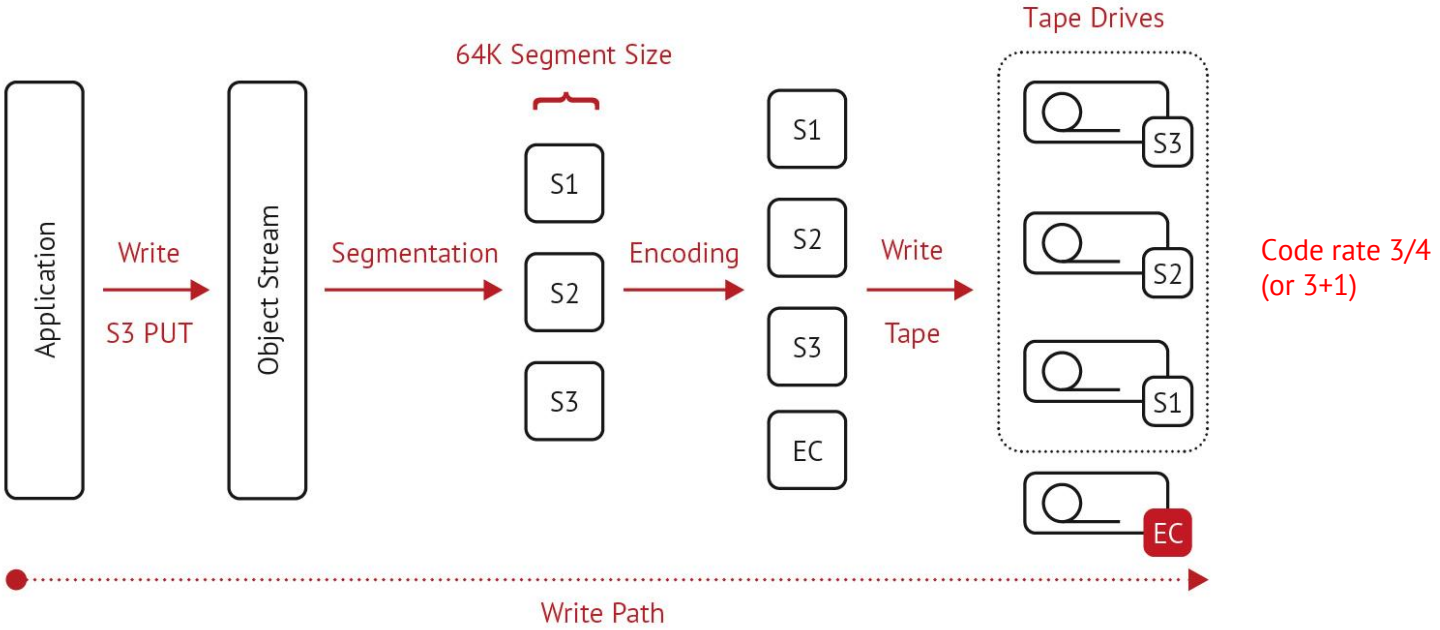
Archival Storage Volume
- Medium

Archival Storage Device
- Group of up to 8 tape loaders or library units



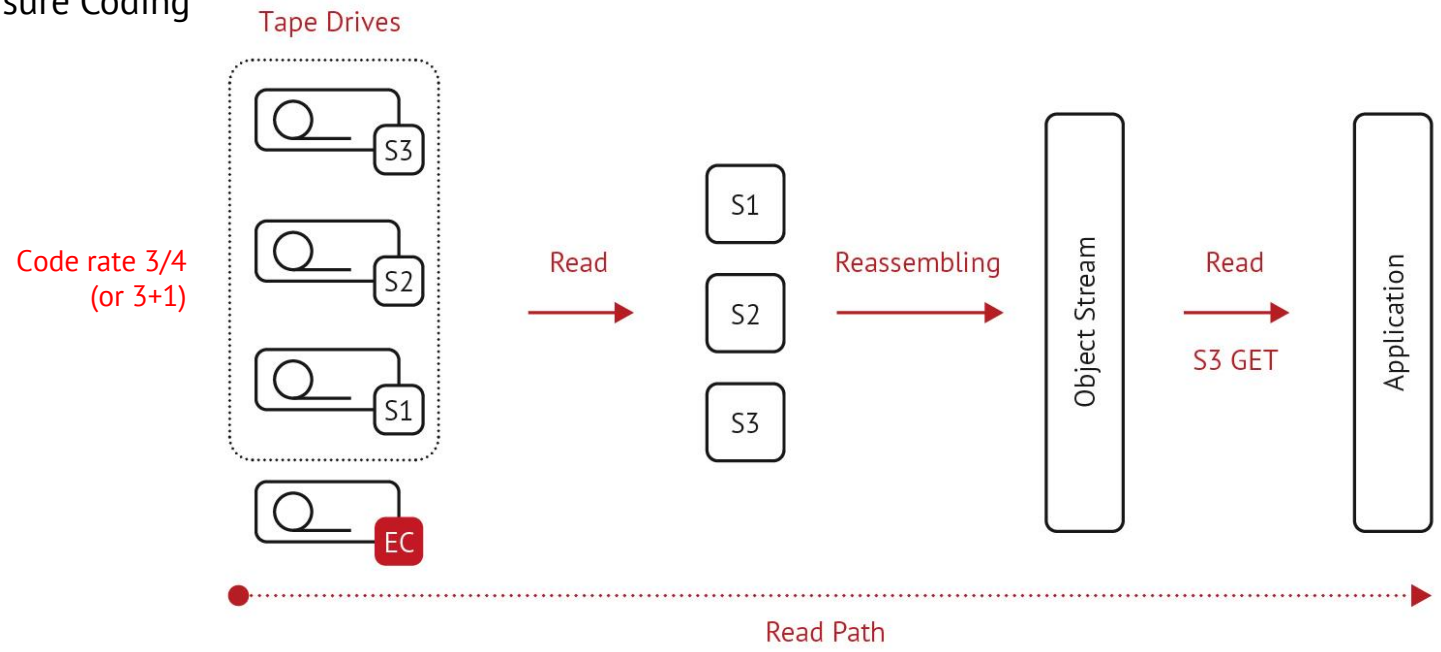
Tape-based Object Storage

Erasure Coding



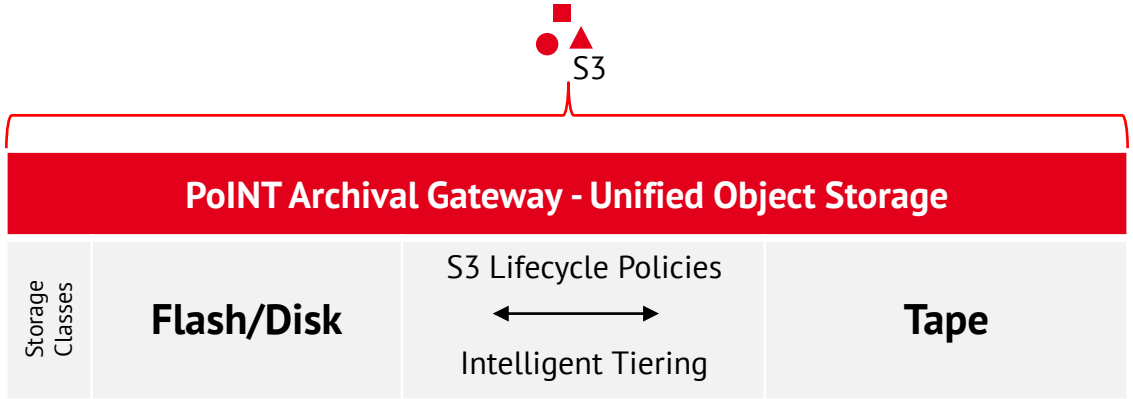
Tape-based Object Storage

Erasure Coding



Tape-based Object Storage

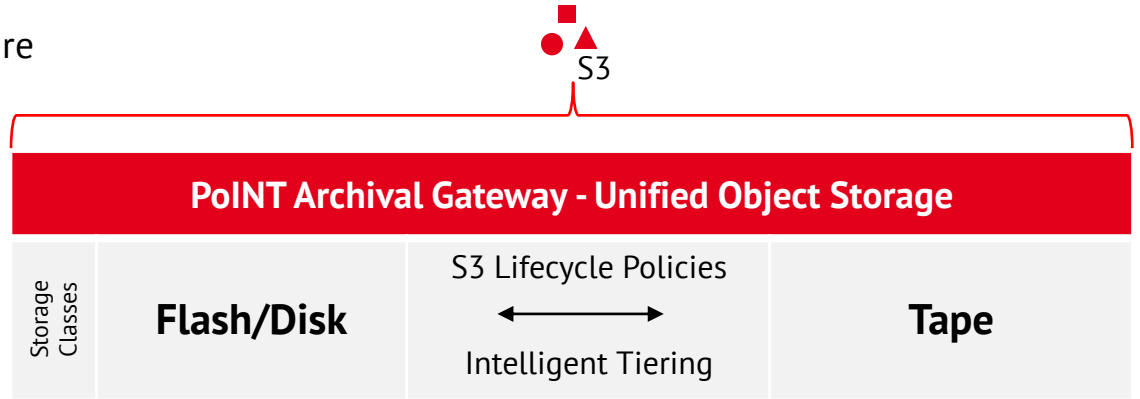
Storage Classes



Tape-based Object Storage

Features

- Software-defined
- Multiple storage classes
- Single namespace across flash/disk and tape
- AWS S3 compatible
- Free choice of storage hardware
- Flexible configurations
- Automatic replication



Tape-based Object Storage

Storage Classes

- Flash/Disk and Tape

Storage Class	Purpose
Flash/Disk	Warm, frequently accessed data (low access times, within msec)
Tape	Cold, long-lived, archive data (**direct access**, within sec to minutes)

Tape-based Object Storage

Storage Classes

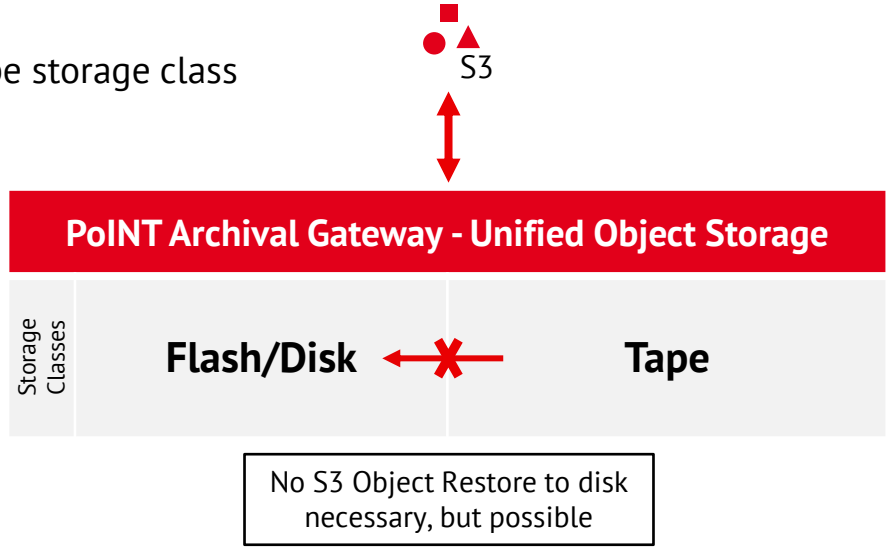
- Flash/Disk and Tape
- AWS S3 compatible

PoINT Storage Class	*Corresponding AWS Storage Class*
Flash/Disk	*S3 Standard*
Tape	*S3 Glacier*

Tape-based Object Storage

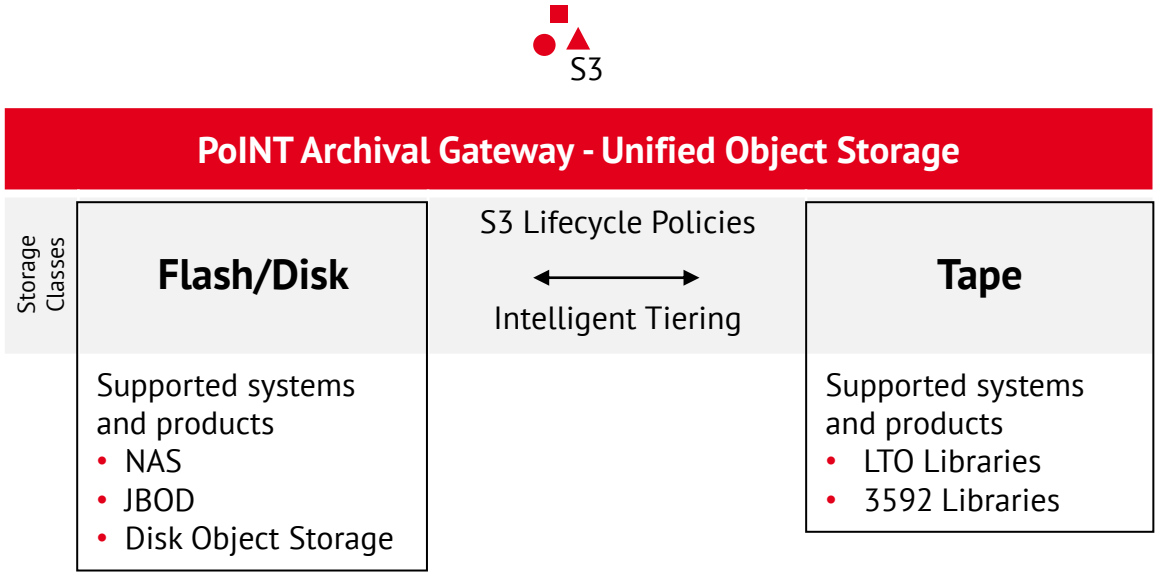
Storage Classes

- Flash/Disk and Tape
- AWS S3 compatible
- Transparent access to flash/disk and tape storage class



Tape-based Object Storage

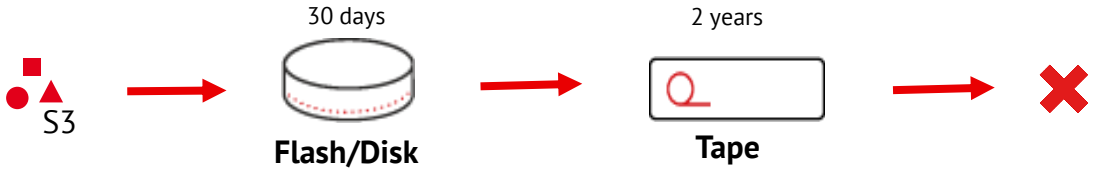
Supported Storage Hardware



Tape-based Object Storage

Lifecyle Policies

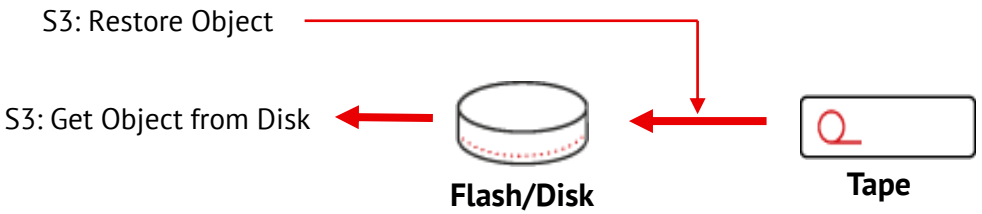
- AWS S3 compatible
- Transition policies
  - E. g. migrate objects from disk to tape 30 days after creation
- Expiration policies
  - E. g. delete objects one year after creation
- Example
  - Storage of periodic logs in disk storage class for one month for analysis purposes
  - After one month migrate to tape storage class for archiving purposes
  - After further 24 months delete objects



Tape-based Object Storage

Restoring Objects from Tape to Disk

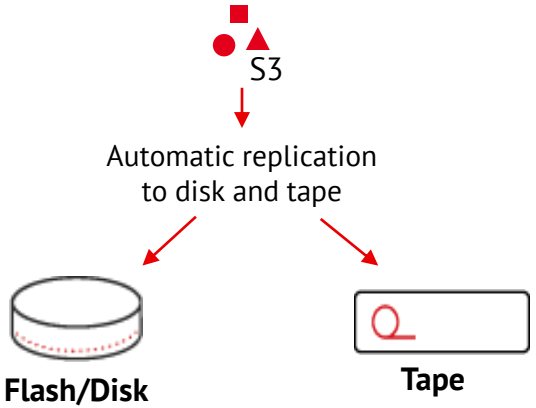
- AWS S3 compatible command: Restore Object (Glacier API)
- Objects can temporarily be restored on disk
- Number of days how long a restored object is stored on disk can be specified



Tape-based Object Storage

Automatic Replication

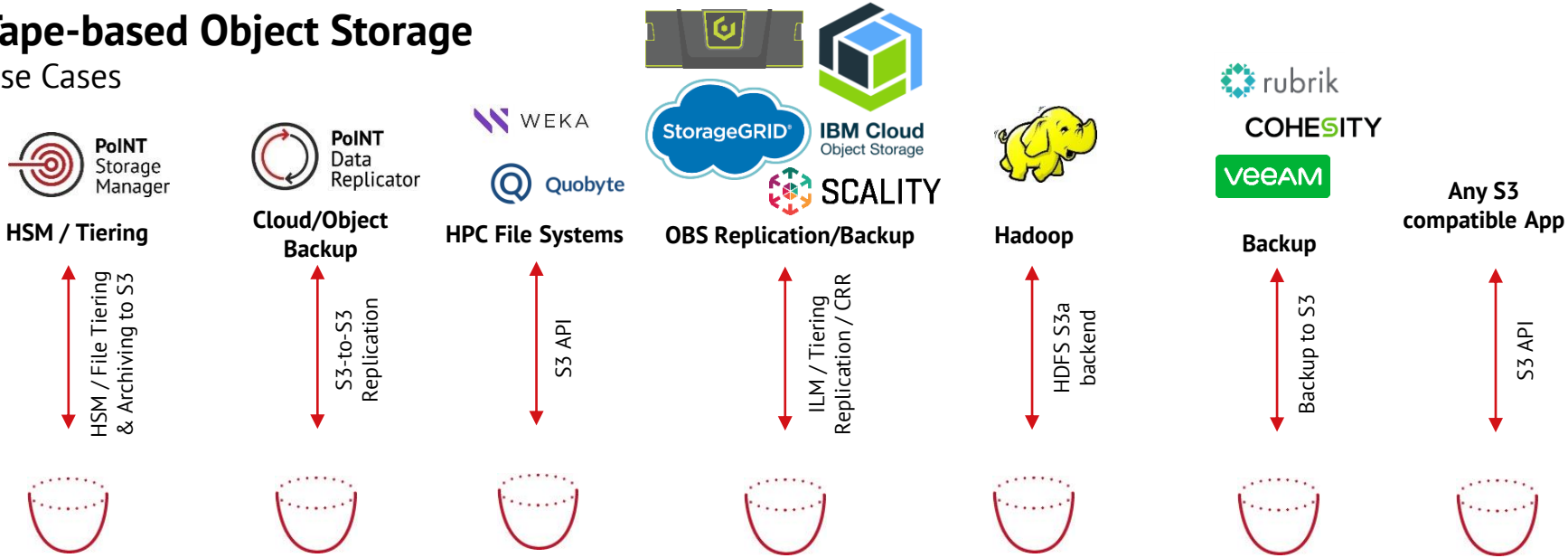
- Automatic replication to flash/disk and tape



PoINT Software & Systems

Tape-based Object Storage

Use Cases



Tape-based Object Storage

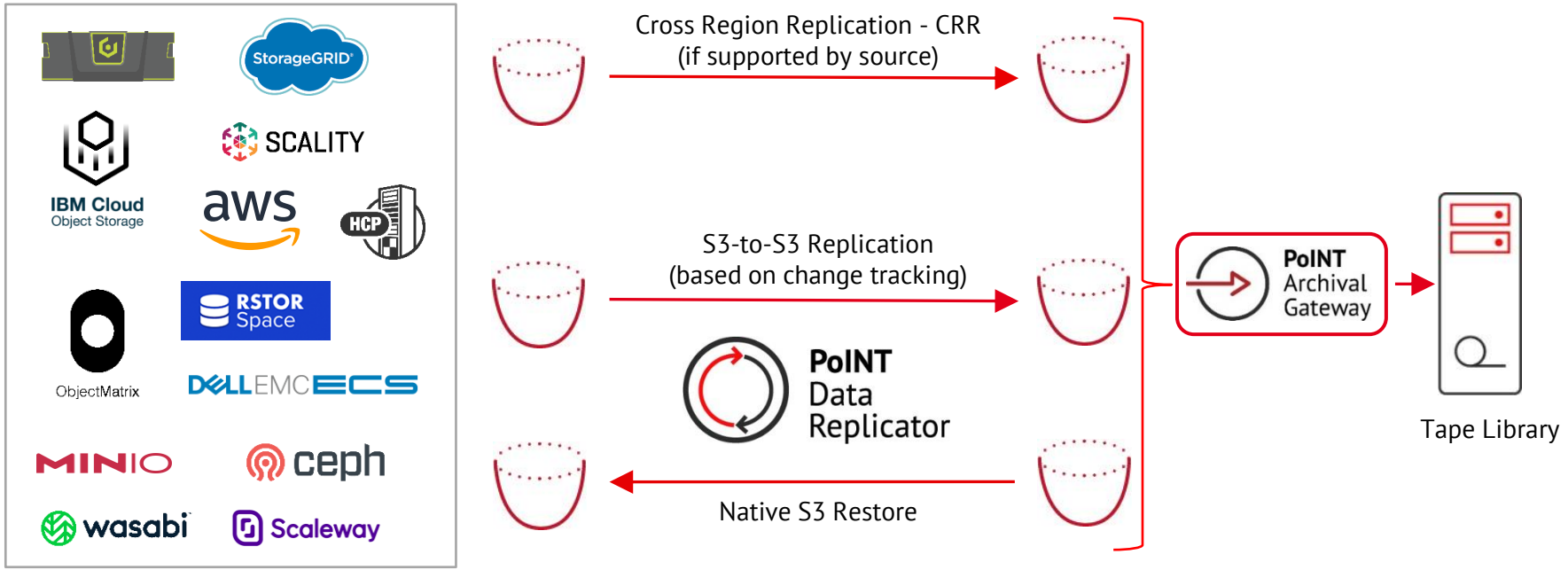
Use Case – Object/Cloud Backup to Tape

How secure is cloud/object storage?

Data loss scenarios	Solutions
• Hard drive failure • Node failure • Rack failure • Site failure	✓ Erasure coding ✓ Replication, CRR ✓ Multiple sites
• Software errors • Accidental or malicious deletion • Ransomware • The human factor	✓ **Backup**

Tape-based Object Storage

Use Case – Object/Cloud Backup to Tape



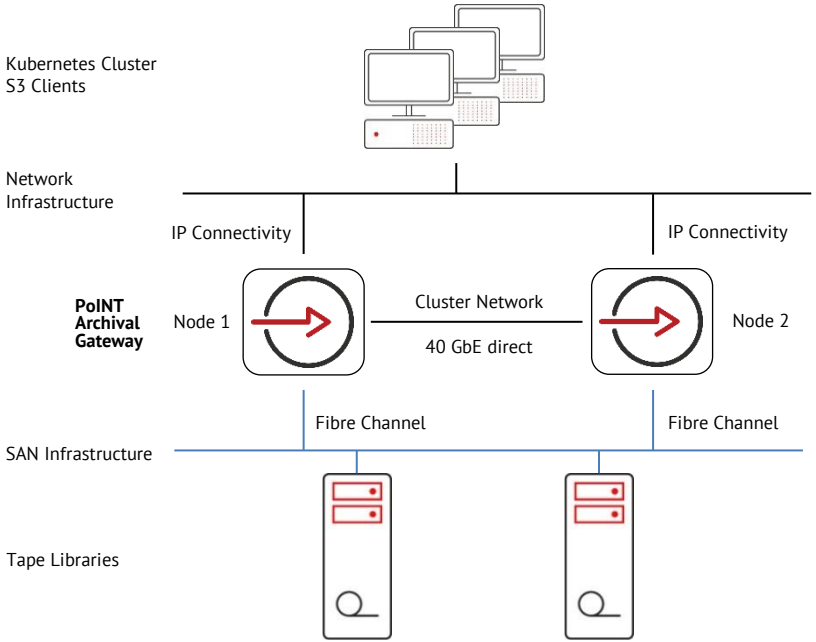
Tape-based Object Storage

Benefits

- Savings in storage costs (compared to all data on HDD or CSP tape)
- No ingest, egress and transaction costs (compared to public cloud)
- No vendor lock-in regarding storage hardware
- Data sovereignty to meet national laws/governance structures (on-premises private cloud)
- Fulfillment of archiving and compliance requirements (by long-term data preservation)
- Protection from ransomware and cybercrime ("Air-gap" by tape)
- Investment protection by standardized interface (AWS S3)
- Sustainability by less energy consumption and lower CO_2 emissions (compared to all data on HDD)

Tape-based Object Storage

Case Study – EMBL European Bioinformatics Institute



The challenges

- Archiving workloads from Kubernetes Cluster via S3
- Storage capacity in high three-digit PB range
- Long-term preservation and high availability
- Multiple vendor support incl. support for LTO and 3592 drives
- High performance for read and write access (1 PB per week)
- Low TCO

The solution – PoINT Archival Gateway

- Scale-out architecture and unlimited storage space
- Native S3 interface including versioning
- Logical partitioning for workload requirements
- WORM and retention management
- Read and write performance up to 3 PB per week

The benefits

- Compatibility with any S3 client
- Workload specific configuration on bucket or partition level
- Scalable by nodes, drives, slots and media
- Software-defined for independence and sustainable planning
- No vendor and technology lock-in

www.point.de