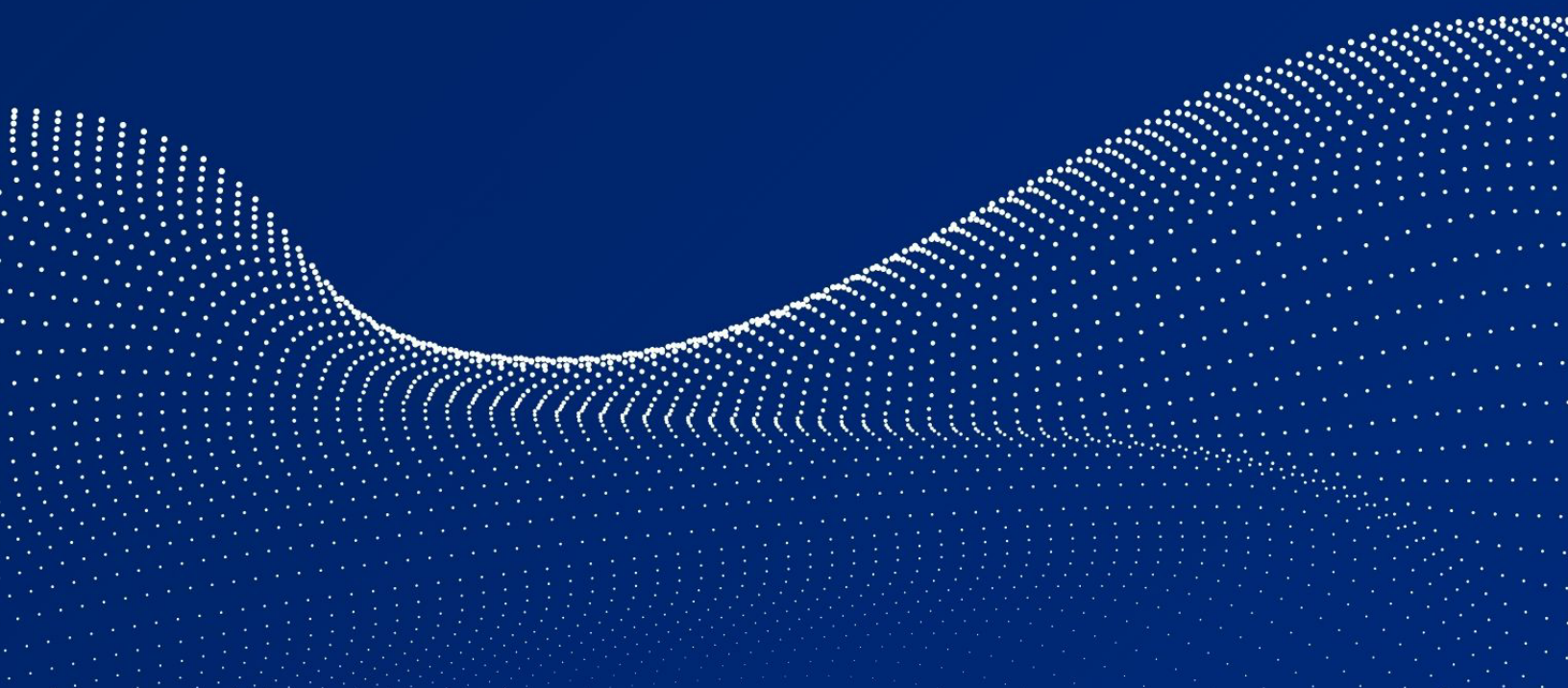
Versity

The Complete Guide to Building a Future-Proof, High-Performance S3-to-Tape Archive



Executive Summary

If your organization is archiving tens to hundreds of petabytes and evaluating how to connect the S3 API to tape, this guide is for you. It is written for storage architects and infrastructure decision-makers who need to understand not just how S3-to-tape archiving works, but where the architectural risks hide, how proprietary gateways create decades of vendor dependency, and what to look for in a gateway that will outlast multiple generations of tape hardware. Whether you are building a new archive, replacing a legacy HSM, or expanding an existing tape environment, this guide provides the technical depth and decision framework to make a gateway choice you will not regret in ten years

The S3 API has become the standard interface for archival storage. Combined with a new generation of rack-modular tape libraries capable of 40–60+ petabytes per rack, S3-to-tape archiving is now the most cost-effective architecture for preserving large-scale unstructured data on-premises.

But the gateway software that sits between S3 and tape is the most consequential decision in the architecture. Most S3-to-tape gateways are sold by the same vendors that manufacture tape libraries, bundling proprietary software with their hardware. These bundled gateways store data in proprietary catalogs that only the vendor's software can read, enforce one-to-one relationships between gateway instances and libraries, and limit throughput to single-stream patterns. The result is architectural lock-in that persists for the 10–30 year life of the archive.

This guide examines the full S3-to-tape archiving landscape: why tape is surging, how the workflow operates end to end, where the architectural risks hide, and what a modern, vendor-neutral alternative looks like in production. It draws on deployments at Pawsey Supercomputing Research Centre (150 PB), the Texas Advanced Computing Center (exabyte-scale), and the National Center for Supercomputing Applications to illustrate what hardware-agnostic S3-to-tape archiving delivers at scale.

Why S3-to-Tape Archiving Is Surging

Data Growth Is Outpacing Every Other Storage Tier

AI training pipelines, genomics, satellite imagery, autonomous vehicle telemetry, and media production are generating data at rates that make disk-only and cloud-only archiving economically unsustainable. In 2024, enterprise LTO tape shipments reached 176.5 exabytes, the fourth consecutive year of record growth. Analysts project the tape storage market will grow from approximately $6.27 billion in 2025 to nearly $11 billion by 2030.

The driver is straightforward: at petabyte scale, tape's cost per terabyte is a fraction of disk or cloud cold storage, it consumes near-zero energy when idle, and its capacity roadmap extends through LTO-

14 at a projected 913 TB per cartridge. For organizations archiving tens to hundreds of petabytes, tape is not a legacy technology. It is the only economically viable long-term storage medium.

Rack-Modular Libraries Changed the Game

Traditional enterprise tape libraries were impressive machines, but they were also imposing ones. They required dedicated floor space, custom installation, and ongoing vendor involvement for maintenance. Libraries like the IBM TS4500 and Spectra TFinity were designed for large, mature tape operations with dedicated staff. For many organizations, the operational overhead was as significant a barrier as the capital cost.

Today, a new class of rack-modular tape libraries has eliminated those barriers. These are fully self-contained libraries that fit in a standard data center rack, use field-replaceable components, and can be deployed and expanded without specialized infrastructure. They were initially conceived as a response to hyperscale operators who needed massive tape capacity with datacenter simplicity. The result is a product class that has rapidly overshot its original target market and is now displacing traditional large-footprint libraries across a wide range of deployments.

- **IBM Diamondback**
  This library is the highest-profile product in this class. A single full rack holds over 46.4 PB of native LTO-10 capacity. It deploys in under 30 minutes and is 99% customer-serviceable. It is fast, dense, and operationally simple.
- **Spectra Logic Cube**
  It holds approximately 50 PB of LTO-10 native capacity with up to 1,670 slots and 16 full-height tape drives. It can be scaled and reconfigured via software without downtime and without special tools. The Cube supports up to 16 partitions for multi-tenant environments and includes Spectra's LumOS management software. Like the Diamondback, it is engineered for simple datacenter deployment in a standalone frame roughly equivalent to a standard rack.
- **BDT Orion**
  This library fits in a standard 19" rack and holds approximately 42 PB of native LTO-10 capacity. It has gained strong traction in European markets with a similar emphasis on datacenter-friendly form factors and modular scalability.
- **Quantum Scalar i7 RAPTOR**
  Quantum's tape library delivers up to 60 PB of LTO-10 native capacity in a single rack, making it one of the densest tape libraries currently on the market. All serviceable components are customer-replaceable in under five minutes. The RAPTOR is designed for hyperscale and enterprise-class archive workloads.

All these libraries have forty to sixty petabytes of native LTO-10 capacity in a single standard rack, combined with field-replaceable components and software-driven management. With a significantly lower acquisition and maintenance costs compared to traditional large libraries, these products are changing the economics of long-term archival storage at every deployment scale.

Library	Max LTO-10 Native Capacity	Key Differentiator
IBM Diamondback	~46.4 PB	99% customer-serviceable; deploys in under 30 minutes
Spectra Logic Cube	~50 PB	LumOS software; 16 partitions for multi-tenant environments; all components customer-replaceable, and fasted tape mounts and exchanges mounts, ReST API connection to Versity
BDT Orion	~42 PB	Standard 19″ rack; strong traction in European markets
Quantum Scalar i7 RAPTOR	~60 PB	Highest density on market; all components customer-replaceable in under 5 minutes

Note that the above capacities assume LTO-10 30TB tape. With the availability of LTO-10 Poly-Aramid 40TB tape, capacities increase by 33%.

S3 as the Universal Archive Interface

The S3 API has become the universal language of object storage. Presenting a tape library as an S3 endpoint means any application that speaks S3 can write to tape without modification, no tape-specific knowledge, no custom integrations. S3 PUT and GET become the interface to archival storage, opening tape to use cases that previously required specialized HSM software: AI training datasets, scientific archives, media preservation, regulatory retention, backup targets, and cloud repatriation.

The question is no longer whether to use S3-to-tape. It is which S3-to-tape gateway to choose, and that decision has consequences that persist for the entire multi-decade life of the archive.

How S3-to-Tape Archiving Works

An S3-to-tape architecture has five core stages. Understanding the full workflow clarifies where architectural decisions create long-term risk or long-term flexibility.

1. Ingest via S3 or File Interface

Applications send standard S3 PUTs (or copy files via NFS/POSIX) to the archive platform. The S3 gateway translates object requests into file operations. Data lands in a high-speed NVMe or flash cache tier for immediate acknowledgment. From the application's perspective, the archive looks and behaves like any other S3 endpoint.

2. Policy-Driven Tiering to Tape

A policy engine evaluates incoming data against configurable rules: age, size, metadata tags, project classification, content type, access frequency, or any combination of criteria. Data that meets migration criteria is automatically staged for tape. Small files are bundled into optimized packages for tape write efficiency. Large files are split into parallel streaming segments across multiple drives.

This is where the architecture diverges sharply between bundled gateways and a full data management platform. Bundled gateways move data to tape on a schedule or when buckets fill. A policy-driven platform moves data to the right tier at the right time, across disk, tape, cloud, and object storage, without manual intervention and without being limited to a single destination.

3. Write to Any LTO Tape Library

Data is written to one or more tape libraries in open archival format. The archive platform manages drive scheduling, mount optimization, and media allocation. Multiple copies can be created simultaneously to separate libraries or sites.

This stage is where hardware agnosticism matters most. A platform that works with any LTO tape library from any manufacturer (IBM Diamondback, Spectra Cube, Quantum Scalar, BDT Orion, etc.) and manages them simultaneously from a single instance, preserves purchasing flexibility at every refresh cycle. A bundled gateway that only works with its own vendor's library locks the hardware decision for the life of the archive.

4. Retrieve via S3 GET or File Access

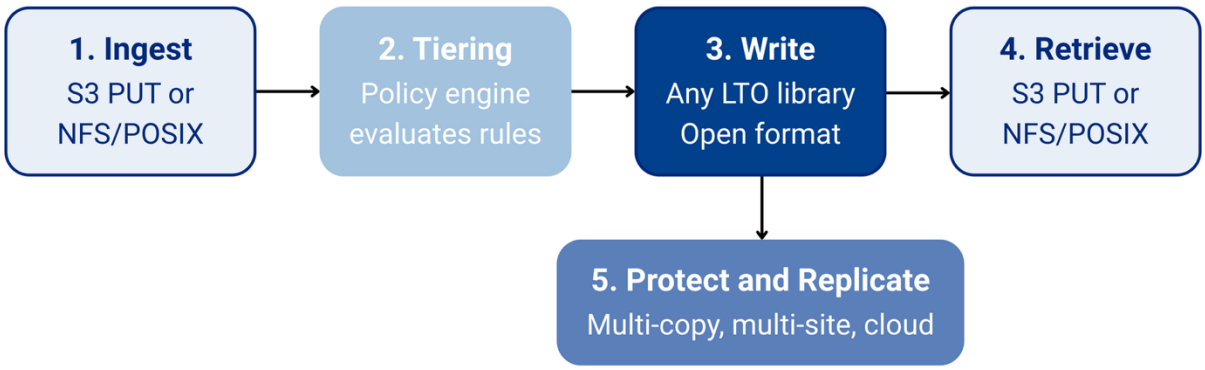
When data is needed, applications issue standard S3 GETs or access files through the POSIX mount. The platform automatically stages the requested data from tape back to the cache tier. Parallel retrieval across multiple drives ensures high throughput for large restores.

Dual-protocol access (S3 and POSIX/NFS in a single namespace) is a critical capability for environments where some applications speak S3 and others speak file. Both access the same data, the same namespace, and the same policies, with no duplication and no separate workflows.

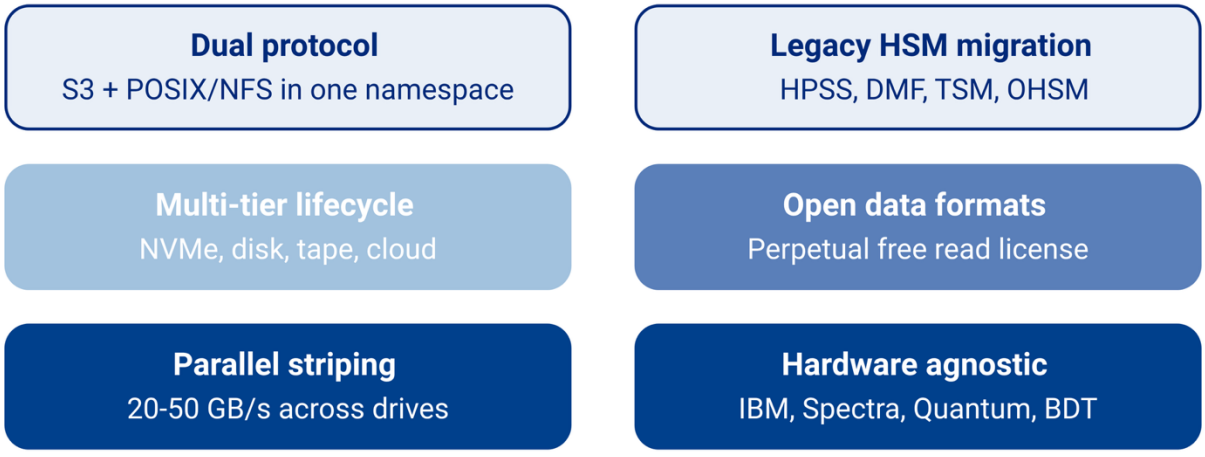
5. Protect and Replicate

Policies direct the platform to maintain copies across tape at multiple sites, on-premises object storage, and cloud storage, all governed centrally. Metadata replicates independently, enabling read-only access at secondary sites with automatic failover. This is native platform capability, not an external add-on.

How S3-to-tape archiving works with ScoutAM



What ScoutAM adds at each stage



Where S3-to-Tape Archiving Applies

S3-to-tape is not a single-use-case technology. It serves any environment where large volumes of unstructured data must be retained affordably over long-time horizons.

AI and Machine Learning Training Data

Training pipelines generate tens of petabytes that need affordable storage between runs and high-throughput retrieval when active. On-premises S3-to-tape eliminates cloud egress fees. Parallel retrieval across multiple drives delivers the throughput AI workflows demand.

Scientific Research and High-Performance Computing

Exabyte-scale archives span decades and must survive multiple hardware generations. Open data formats and hardware-agnostic gateways ensure reproducibility and long-term access regardless of vendor changes.

Media Preservation and Production

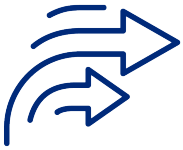
High-resolution content libraries grow indefinitely and must be preserved for the life of the content, which is often decades. S3 compatibility with media asset management platforms enables direct archive workflows without custom integrations.

Regulatory Compliance and Long-Term Retention

Financial services, healthcare, and government mandates require retention of 7–30+ years. When the retention period exceeds hardware refresh cycles, the archive software must outlast the hardware and must not lock the organization into a single vendor's pricing.

Cloud Repatriation

Organizations moving data out of public cloud cold storage (S3 Glacier, Azure Archive) to on-premises tape need an S3-compatible endpoint locally. The same application workflows continue without modification.



Legacy HSM Modernization

Organizations running IBM HPSS, IBM Spectrum Archive (TSM), HPE DMF, or Quantum StorNext can migrate to S3-to-tape workflows without a full data extraction. A platform that reads legacy formats natively enables a zero-migration conversion, importing metadata and cataloging existing media without a brute-force data migration project.

The Lock-In Problem: How Proprietary Gateways Trap Your Data

The Catalog Is the Lock

Every S3-to-tape gateway must maintain a catalog, a mapping of which S3 objects are stored on which tapes, at which physical positions. This catalog is the critical architectural component. In the bundled gateway model, the catalog is proprietary and lives exclusively inside the vendor's appliance.

The tapes themselves may use standard LTO cartridges. The physical format may be LTFS. The protocol may be standard S3. On the surface, the gateways look open. However, the logical organization of data, which object is on which tape, at which offset, in which order, is opaque and controlled by the vendor. Without the vendor's software running, the tapes are physically present but logically unreadable.

This is not a theoretical concern. Ask any bundled gateway vendor a simple question: can you read your data without the vendor's software running? For IBM Deep Archive, Spectra BlackPearl, and Quantum ActiveScale Cold Storage, the answer is no.

The "Great Re-Bundling"

In the previous era, tape vendors bundled proprietary HSM software (IBM HPSS, HPE DMF, Oracle OHSM) with their libraries and charged accordingly. Customers who needed to migrate spent years extracting data. The S3 gateway era recreates this dependency under a more modern veneer. S3 APIs and LTFS formats create an illusion of openness while the data encapsulation remains proprietary.

The difference today is that the lock-in is more opaque. Customers see familiar S3 commands and standard media and assume portability. The proprietary layer, the catalog, is invisible until the first time they need to move.

Why This Matters Over Decades

Tape archives are not three-year infrastructure investments. They are 10–30 year commitments. An organization that writes 40 PB of research data through a proprietary gateway has made an irrevocable architectural choice. When the gateway reaches end of support, when the vendor changes pricing, when the appliance fails, or when a better alternative emerges, the data cannot move without the original system to read it back.

The financial consequence is loss of pricing leverage. Once an organization has significant data locked in a proprietary catalog, the vendor controls every future pricing decision: hardware refreshes, software renewals, support contracts, and capacity expansion. This is not speculative. It is the exact dynamic that caused decades of pain for customers locked into legacy HSM platforms.

Architectural Limitations of Bundled Gateways

Beyond lock-in, the bundled gateway model imposes structural limitations that become acute at scale.

Scaling Ceilings

Most bundled gateways enforce a one-to-one relationship between a gateway instance and a tape library. Adding capacity or throughput requires deploying an entirely new appliance with its own catalog and control plane. These coarse-grained scaling steps create hard ceilings on throughput and concurrency, and each additional library becomes its own isolated silo.

A one-to-many architecture, where a single software instance manages multiple libraries from multiple vendors, eliminates this fragmentation. Capacity and throughput scale linearly with drives and media rather than in appliance-sized increments.

Performance Bottlenecks

Object gateways optimized for transactional workflows are not built for sustained high-bandwidth tape operations. Large files are typically written as single streams or multipart uploads, limiting throughput to one or a small number of drives. Massive restores become serialized operations.

Modern HSM architectures solve this with parallel file striping, splitting large files across multiple drives simultaneously. A data parallelism engine that groups small files into optimized packages for write efficiency and splits large files into parallel streaming segments can saturate multiple tape drives at once, delivering aggregate throughput of 20–50 GB/s across parallel drives. That is the difference between hours and days at petabyte ingest volumes.

Missing Enterprise Capabilities

By abstracting tape behind a simple object interface, bundled gateways trade away capabilities that become essential at scale:

Multi-copy writes with deterministic placement. Serious data protection requires copies at multiple sites simultaneously, tape at Site A, tape at Site B, and object storage in a cloud, governed by a single policy. Bundled gateways are single-destination systems.

Tape data segregation. Regulated environments need to separate data across tape pools by classification, project, tenant, or sensitivity. Fine-grained pool management with independent retention policies, audit trails, and air-gap controls is not available in bundled gateways.

Policy-driven lifecycle management. Moving data between NVMe, disk, tape, and cloud based on age, access frequency, content type, and metadata requires a policy engine that spans all tiers.

Bundled gateways typically move data to tape on a schedule or when buckets fill. They do not support multi-tier lifecycle policies.

Disaster recovery without vendor software. If the proprietary catalog is lost or corrupted, the data on tape is physically present but logically unrecoverable. Architectures that embed self-describing metadata on the media itself allow data to be recovered independently of the original software stack.

The Hardware-Agnostic Alternative: Versity ScoutAM

Versity ScoutAM represents a fundamentally different approach to S3-to-tape archiving. It is a software-defined hierarchical storage manager that runs on commodity hardware, works with any LTO tape library from any vendor, and writes data in open formats. ScoutAM is not a simple pass-through gateway. It is a complete data management platform built on two open-source foundations: the Versity S3 Gateway (Apache 2.0) and ScoutFS, a clustered filesystem optimized for massive archival namespaces.

Works with Any LTO Tape Library

ScoutAM is hardware-agnostic. A single instance manages IBM Diamondback, Spectra Cube, Quantum Scalar i7 RAPTOR, BDT Orion, and any other LTO library interchangeably, even simultaneously. Buy the best hardware for your requirements today and retain the freedom to change vendors at the next refresh cycle. Your tape hardware decision stays independent from your archive software decision.

Open Data Formats and Perpetual Read License

Data written by ScoutAM includes self-describing metadata on the media. Versity provides a perpetual free read license, meaning your organization retains full access to archived data even if you stop using Versity software entirely. Data can be recovered directly from tape without any proprietary software. No other S3-to-tape gateway offers this guarantee.

S3 and File Access in a Single Namespace

ScoutAM pairs with the open-source Versity S3 Gateway (Apache 2.0) to deliver both an S3 object interface and a POSIX file interface within the same system. Applications that speak S3 and applications that use NFS, Samba, or direct POSIX access all interact with the same data, the same namespace, and the same policies, with no duplication and no separate workflows.

Policy-Driven Data Management

ScoutAM's automated policy engine tiers data across NVMe, disk, tape, and cloud based on age, access frequency, content type, metadata tags, or any combination of criteria. Data moves to the right

tier at the right time without manual intervention, and without being limited to the single-destination model of bundled gateway appliances.

Multi-Copy Protection and Geo-Replication

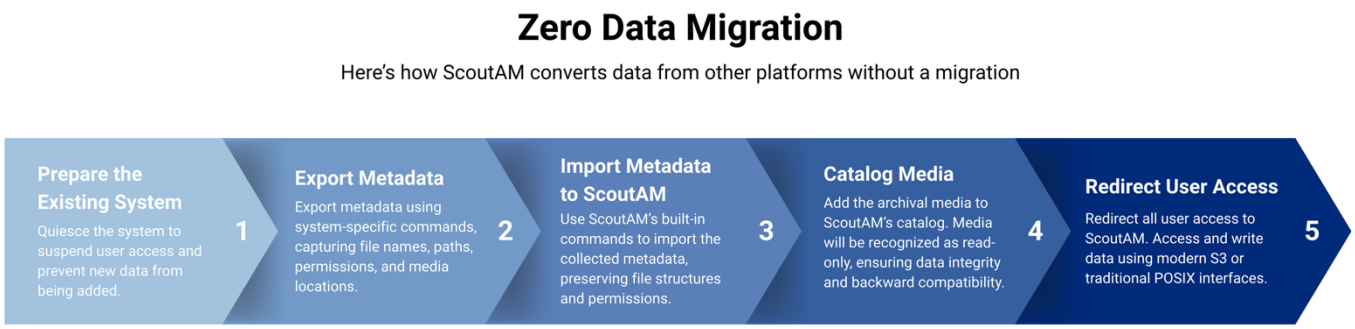
ScoutAM creates multiple copies of archived data to different destinations simultaneously: tape at Site A, tape at Site B, object storage in a private cloud, all governed by a single policy. Metadata replicates separately, enabling read-only access at secondary sites with automatic failover. This is native platform capability, not an external add-on.

Parallel Throughput for Large-Scale Ingest and Retrieval

ScoutAM's data parallelism engine groups small files into optimized 10 GB packages for efficient tape writes and splits large files into parallel 100 GB streaming segments across multiple drives simultaneously. Aggregate throughput of 20–50 GB/s across parallel tape drives enables petabyte-scale ingests and restores that single-stream gateways cannot match.

Zero Data Migration from Legacy HSM

ScoutAM natively reads data written by IBM HPSS, IBM Spectrum Archive (TSM), HPE DMF, and Oracle OHSM. Organizations moving from legacy HSM to S3-to-tape workflows can convert to ScoutAM by importing metadata and cataloging existing media, no brute-force data migration project required.



Open-Source Foundations

The Versity S3 Gateway is Apache 2.0 licensed and available on GitHub. ScoutFS is GPL licensed. This open-source foundation means the core components of the archival stack are auditable, extensible, and not dependent on a single vendor's commercial decisions.

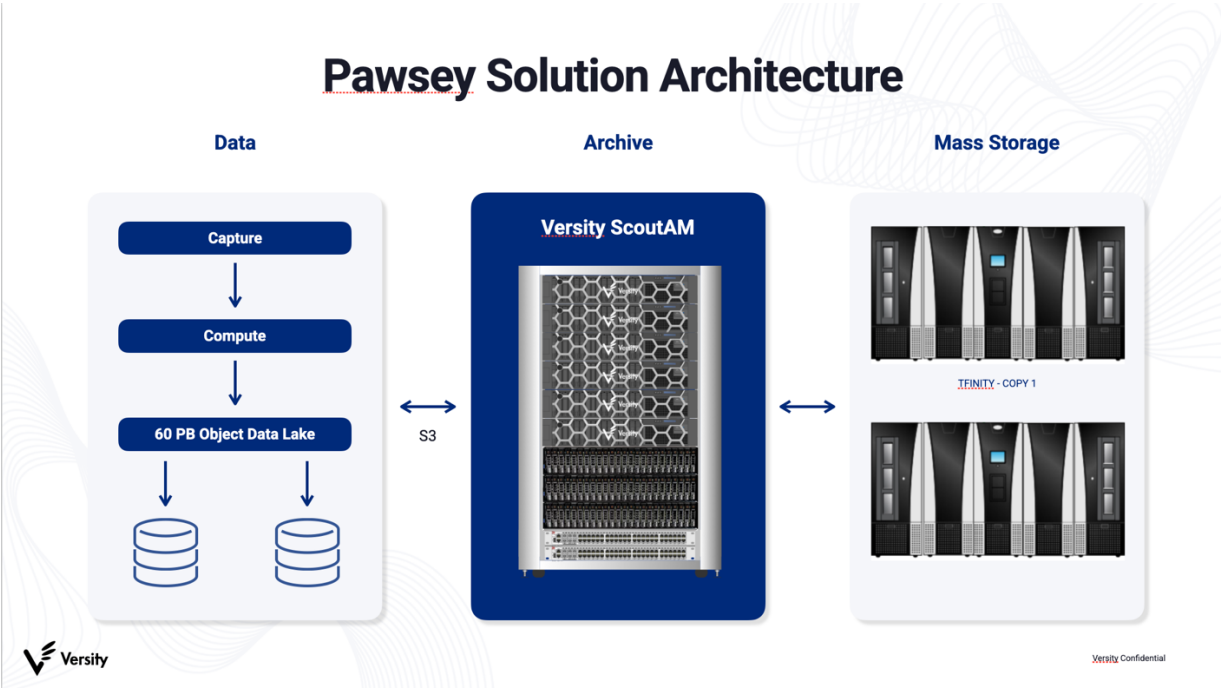
Production Deployments: S3-to-Tape at Scale

Pawsey Supercomputing Research Centre | 150 PB

Pawsey, one of Australia's premier HPC facilities, deployed Versity ScoutAM as the core of "Banksia," a 150 PB mass storage system supporting astronomy research communities. The system manages a 5 PB high-performance data cache and two large tape libraries with 68 tape drives and 74 PB capacity per library, with room to grow.

The S3 workflow challenge: Pawsey's previous storage system could not handle S3 data ingestion, limiting the Centre's ability to integrate object storage with its archive. ScoutAM resolved this by providing native S3 access to the entire archive while preserving the existing historical data collection, no data migration required.

Key outcomes: Researchers access archival data through the S3 protocol. Complex workflows are automated. Metadata searches and queries are responsive at scale. ScoutAM's zero-migration conversion allowed Pawsey to decommission its legacy system without disrupting ongoing research.

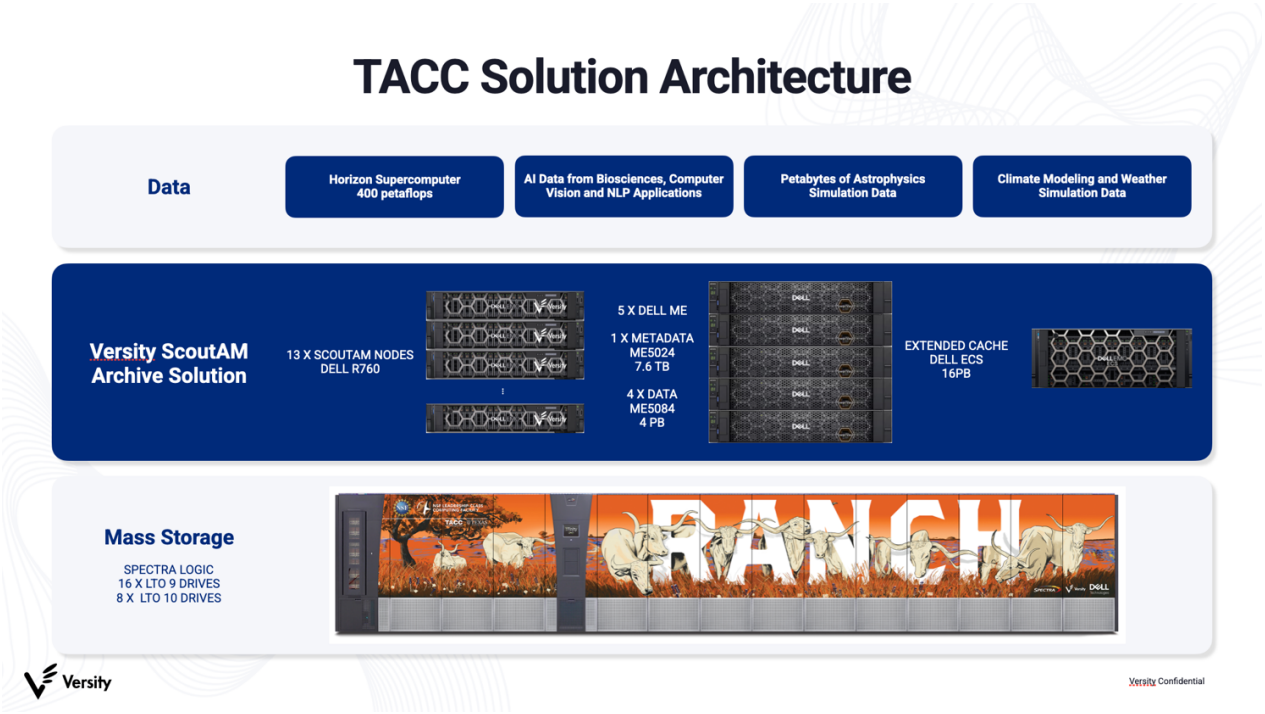


Texas Advanced Computing Center (TACC) | 1 Exabyte

TACC selected ScoutAM to power "Ranch," an exascale archive supporting Horizon, expected to become the world's largest academic supercomputer for open scientific research. The system combines Spectra Logic TFinity tape libraries with ScoutAM software on Dell servers.

Architecture decision: TACC adopted a two-tier flash-to-tape architecture, eliminating mid-tier disk in favor of a faster, more cost-efficient approach to managing up to one exabyte of data. The choice of ScoutAM over a bundled gateway was driven by the need for vendor-agnostic hardware flexibility and open data formats across a multi-decade archive lifecycle.

Why it matters: TACC's open, vendor-agnostic design gives them the flexibility to pair best-of-breed tape hardware with software that is not locked to any single library manufacturer, ensuring the archive will survive multiple hardware generations across its multi-decade lifespan while keeping pricing leverage at every refresh cycle.



Gateway Capabilities at a Glance

The S3-to-tape gateway market includes five primary products. Three are bundled with their vendor's tape hardware. Two are independent of the library. At the scale this guide addresses, petabyte and exabyte archives with multi-decade retention horizons, the architectural requirements around data readability, multi-vendor hardware flexibility, and full-spectrum lifecycle management narrow the field considerably. The following comparison shows where each product stands.

Capability	IBM Deep Archive	Spectra BlackPearl	Quantum ActiveScale Cold Storage	Versity ScoutAM
Data readable without vendor software?	No	Yes, but requires additional work	No	Yes (perpetual free read license)
Works with any tape library brand?	IBM only	Spectra only	Quantum only	IBM, Spectra, Quantum, BDT
Open source component?	No	Yes, but requires additional work	No	Yes (Apache 2.0 S3 gateway)
Multi-copy / geo-replication?	Limited	Limited	Limited	Yes, policy-driven
Tape data segregation by tenant/project?	Limited	Limited	Limited	Yes, fine-grained pools
Multi-tier lifecycle (tape + cloud + disk)?	Tape only	Tape + disk cache + cloud	Tape + disk + object	Tape + disk + cloud + NVMe
Legacy HSM migration (HPSS, DMF, TSM)?	No	No	No	Yes
Pricing model	Appliance + library	Appliance + library (CapEx or OpEx)	Appliance + library	Software subscription, no per-capacity metering
Multi-library scale pricing	Linear per appliance	Linear per appliance	Linear	One cluster, many libraries

Decision Framework: Evaluating S3-to-Tape Gateways

Step 1: Evaluate the rack-modular libraries on their hardware merits independently.

The IBM Diamondback, Spectra CUBE, and BDT Orion are excellent products. Their density, serviceability, and datacenter-friendly form factors are genuine advantages. Choose the hardware that best fits your capacity, budget, and performance requirements.

Step 2: Evaluate the S3 gateway separately as a distinct purchase decision.

Ask every gateway vendor the same question:

1. *If I stop using your software, can I still read my data?*
2. *Can I use your gateway to write to a different brand of library?*

For IBM Deep Archive, Spectra BlackPearl, and Quantum ActiveScale Cold Storage, the honest answers are no and no. For Versity ScoutAM, both answers are yes.

Step 3: Verify that the gateway supports the operational capabilities your workload requires:

- Multiple copies and geo-replication to separate destinations
- Fine-grained tape pool segregation by project, tenant, or classification
- Policy-driven lifecycle management across tiers (NVMe → disk → tape → cloud)
- POSIX and NFS access alongside S3
- Throughput adequate for your ingest and retrieval workloads

Conclusion

The new generation of rack-modular tape libraries is a genuine breakthrough. The S3 API is the right interface for modern archival storage. But the gateway that sits between S3 and tape is the most consequential decision in the architecture, and it is entirely separable from the library hardware.

Bundled gateways from tape hardware vendors create data dependencies that persist for the full multi-decade life of the archive. They impose one-to-one scaling constraints, limit throughput to single-stream patterns, and trade away enterprise capabilities in the name of simplicity. These trade-offs may be acceptable for small, single-rack deployments with short retention horizons. They are not acceptable for organizations building petabyte- and exabyte-scale archives that must survive multiple generations of hardware.

The alternative is a hardware-agnostic approach that separates the software decision from the hardware decision, writes data in open formats, provides both S3 and file access in a unified

namespace, and preserves competitive leverage at every refresh cycle. Versity ScoutAM is the only platform in the market that delivers all of these properties while also providing full-spectrum data management: parallel throughput, multi-copy protection, lifecycle tiering, legacy migration, and dual-protocol access in a single, unified solution.

Choose your tape hardware on merit. Choose your gateway on architecture. And make sure you can still read your data twenty years from now.