TACC
TEXAS ADVANCED COMPUTING CENTER

How Versity and Spectra Logic Built an Exascale Archive for the Texas Advanced Computing Center

The Texas Advanced Computing Center (TACC) needed a scalable storage solution capable of managing an exabyte of data generated by Horizon, its new AI supercomputer. Here's why TACC selected Spectra, Versity, and Dell for its exascale archive.

INDUSTRY
RESEARCH, HPC

THE CHALLENGE

As TACC prepared to launch the Horizon supercomputer, it faced a major storage challenge: How to manage the massive, AI-driven datasets the system would generate? Existing infrastructure couldn't scale to exabyte levels. TACC needed a long-term archive that ensured speed, reliability, and flexibility.

THE SOLUTION

TACC deployed Ranch, a next-generation archive system designed to scale with Horizon's demands. Through a joint solution with Versity, Dell , and Spectra Logic, TACC gains seamless multi-tier performance, rapid data access, and future-ready scalability to support Horizon's unprecedented AI and simulation workloads.

KEY RESULTS

01 Delivered a future-ready, exabyte-scale archive to support Horizon's AI and simulation workloads.

02 Achieved fast data retrieval through Versity's flash and object extended cache design.

03 Ensured complete disaster recovery with metadata stored on media and read-only remote replication.

04 Seamless open integration across POSIX and S3 for diverse research applications.

05 Direct, expert-driven support from Versity's engineering team

SPECTRA

V Versity

TACC at the University of Texas at Austin has long led the nation in high-performance computing. Since its founding in 2001, TACC has operated some of the United States' most powerful academic supercomputers — from Ranger in 2007 to today's 40 petaFLOP/s Frontera.

Now, TACC is entering the next era of this legacy with **Horizon, a 400 petaFLOP/s high-performance supercomputer** that is a cornerstone of the U.S. National Science Foundation's Leadership-Class Computing Facility (NSF LCCF).

Horizon isn't just an incremental upgrade. It delivers a tenfold improvement in simulation performance and **a 100x increase in artificial intelligence capabilities** compared to Frontera.

Horizon will become the **largest supercomputer in the NSF portfolio dedicated to open, academic research** — empowering scientists to explore new questions in fields ranging from climate modeling to quantum mechanics.

But with transformative capability comes a challenge: How to handle the enormous, AI-driven volumes of data that Horizon will generate?

TACC needed a storage strategy that could reliably, affordably store an exabyte of data at a single site.

ABOUT THE TEXAS ADVANCED COMPUTING CENTER (TACC)

The Texas Advanced Computing Center (TACC) at The University of Texas at Austin is the leading academic supercomputing center in the country. TACC delivers world-class and innovative systems, tools, software, and expertise to researchers who seek to make an impact in the world, and advance discovery across disciplines.

Since 2001, TACC has provided access to a wide range of computational resources — from high-performance computing and AI at scale, to storage and visualization — and offered staff expertise, thought leadership, support, training, and education and outreach.

TACC is also home to the U.S. National Science Foundation Leadership-Class Computing Facility (NSF LCCF), which is poised to revolutionize computational research and discovery.

TACC's approach is people-centered, its funding ensures sustainability, and its staff are dedicated to making the power of computing available to all.

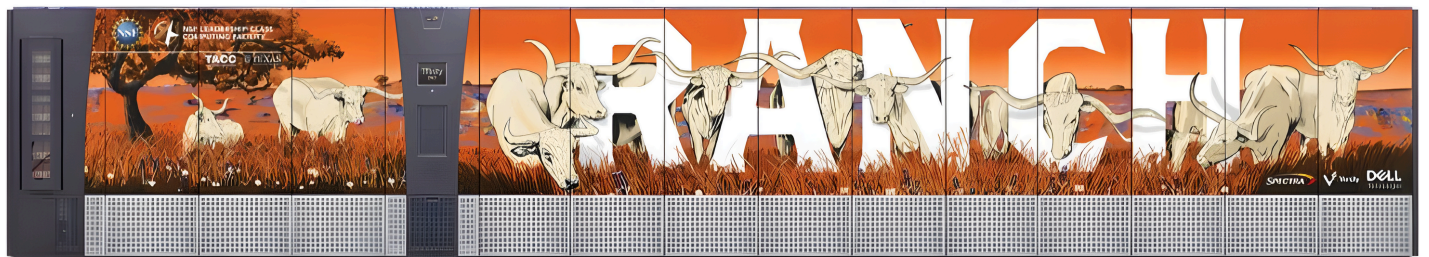## MANAGING EXASCALE DATA QUANTITIES

When you're dealing with data at the petabyte or exabyte scale, one storage medium stands out: **tape**.

"We've always had pretty big tape archives at TACC," says David Cooper, Senior Systems Administrator of HPC Large-Scale Systems. "Not only does tape provide more capacity for the money, but it's also more stable and secure for long-term storage. **When data is written to tape, we don't have to worry about it quite as much as other mediums like disk or flash.**"

To support the Frontera supercomputer, TACC uses a Quantum Scalar i6000 tape library with LTO-9 drives. While that system met Frontera's needs, it could not scale to the massive influx of data Horizon would generate each year, nor did it provide the openness, parallelism, or performance required for exascale AI workloads.

Recognizing this, TACC chose not to pursue incremental fixes but instead made a proactive shift to a fully modern archive purpose-built for Horizon's scale. To meet these next-generation demands, the team made a strategic shift to implement a Spectra Logic TFinity tape library paired with Versity ScoutAM software, Dell PowerEdge R760 servers, and five Dell ME5 storage arrays.

Called Ranch, this integrated system is installed in a prominent, high-traffic area of TACC to help educate and inspire a diverse range of visitors — from students to industry leaders. As a nod to its Texas roots, the TFinity library is wrapped in a custom graphic featuring longhorn cattle, blending technical excellence with local pride.

A MATCH MADE FOR HIGH-VOLUME DATA MANAGEMENT

When planning Horizon's next-generation archive, TACC knew it needed more than just a bigger tape library. They needed a complete, integrated solution that could scale to an exabyte and beyond. After evaluating multiple vendors, the team chose the combined strength of the Spectra Logic TFinity library with Versity's ScoutAM software.

Together, Spectra and Versity provide the scalability, openness, and reliability that align with TACC's priorities.

The Spectra TFinity is already trusted by some of the world's leading high-performance computing sites to support multi-exabyte environments.

Versity ScoutAM complements that strength with a modern, open, and vendor-agnostic software layer designed specifically for exabyte-scale archives. Its transparent, modular design integrates cleanly with TACC's Dell front-end systems and delivers both S3 and POSIX access, ensuring compatibility with current research workflows while future-proofing the archive.

This combined proven track record of secure storage and intelligent data management gave TACC the confidence to move forward with a solution that could meet Horizon's considerable demands for decades to come.

> "Versity's software is quite transparent and simple. Unlike other vendors that tie solutions to specific filesystems or tape formats, Versity's platform isn't locked down. **It gives us more flexibility.**"
>
> **Junseong Heo**
> Storage Manager
> Advanced Computing Systems at TACC

"**Scalability was the most crucial factor for us, followed by price — we usually try to get as much storage 'bang for the buck' as we can,**" says Cooper. "**Compatibility was also important because we didn't want to deal with vendor lock-in.**"

ScoutAM also introduces a key performance advantage for TACC's researchers: **an extended caching system that accelerates data access.**

Horizon's users will be generating and analyzing far larger datasets than before, so minimizing wait times for data retrieval is critical.

To address this, the Ranch archive includes a 4-petabyte flash tier and a 16-petabyte object storage tier as a high-speed cache in front of the tape library. ScoutAM intelligently stages active data across these tiers, **allowing researchers to retrieve files much faster than with a tape-only archive.**

This ensures Horizon's immense output remains not only preserved but readily accessible — a key differentiator for AI-driven and simulation-heavy workloads.

For researchers, ScoutAM also preserves a familiar user experience. TACC's teams can continue browsing archived data through a hierarchical filesystem structure while benefiting from far greater simplicity and reliability than previous systems.

SUPERIOR TECHNICAL SUPPORT

Horizon will play a key role in the U.S. National Artificial Intelligence Research Resource Pilot (NAIRR) — an initiative launched in 2024 to provide researchers across the country with access to **advanced AI infrastructure.**

With so much agency-backed investment riding on Horizon's success, even brief system downtimes are unacceptable. That made reliable, responsive technical support a top priority for TACC.

"For a system of such substantial size, inevitably there will be technical problems," Cooper explains. **"We really don't want to be down or have catastrophic data loss. A high level of support and software recovery systems are very important for us."**

That's where the combined Spectra and Versity solution truly shines.

> "We needed a new system that could keep up with the demands of our users and compute clusters.
>
> The Spectra TFinity and Versity ScoutAM solution **allows us to continue the high level of service our users expect."**
>
> **David Cooper**
> Sr. Systems Administrator
> HPC Large-Scale Systems
> Texas Advanced Computing Center (TACC)

On the hardware side, TACC's technical team has already tapped into the Spectra Assisted Self Maintenance (ASM) Kit, an industry-first support solution that minimizes downtime by keeping a curated set of spare parts on-site for immediate use — with remote guidance available when needed.

On the software side, Versity's ScoutAM reinforces resilience with a disaster recovery architecture embedded directly into the system. All data elements can be fully recovered from the mass storage media itself, without the need for proprietary or specialized recovery software. Every piece of metadata required to restore a collection resides with the data on physical media, ensuring independence and long-term accessibility.

For even greater protection, ScoutAM's policy engine can automatically create multiple tape copies of critical data, ensuring local redundancy. It also supports asynchronous read-only remote replication, allowing metadata, cache data, and archival data to be mirrored to one or more disaster recovery or secondary sites, ensuring seamless continuity of operations.

The deployment process further highlighted the partnership in action. Spectra delivered a seamless deployment.

"We are very happy with the service we received from Spectra. We were truly impressed with how the team methodically assembled the library frame by frame, and aligned the robotics to validate the system," says Heo. **"That's when we truly realized that the wonderful power of this system is in its simplicity."**

In parallel, Versity engineers worked hands-on with software configuration and tuning, rapidly responding to feedback.

THE NEXT CHAPTER OF AI COMPUTING

The deployment of Horizon marks a bold new chapter in TACC's commitment to advancing open, academic research at unprecedented computational scales. With Spectra's TFinity library providing durable capacity and Versity's ScoutAM software ensuring intelligent, transparent data management, TACC has built a storage foundation that is **scalable, resilient, and future-ready.**

Horizon now has an archive as advanced as its compute power — one that keeps pace with rapid AI-driven workloads, protects data integrity, and ensures discoveries remain accessible for decades to come.

As Horizon powers scientific breakthroughs across disciplines, the storage foundation built with Spectra and Versity ensures that no discovery is lost — only unlocked.

"Versity has a **very elegant and simple disaster recovery solution.**"

David Cooper
Sr. Systems Administrator
HPC Large-Scale Systems
Texas Advanced Computing Center (TACC)