Quantum

CASE STUDY

Preserving Cutting-Edge Research at TACC with Quantum Archive Solutions

The Texas Advanced Computing Center (TACC) at The University of Texas at Austin provides researchers with powerful supercomputing resources that enable breakthrough research in a wide range of fields. By replacing an aging archive environment with a new solution from Quantum, TACC has gained a reliable, scalable environment that can deliver strong performance and help enforce strict usage policies for the large user community.

TACC
TEXAS ADVANCED COMPUTING CENTER

FEATURED PRODUCTS

StorNext®

"We recognized that Quantum could deliver the reliability and restorability we needed over the long term."

Frank Douma
Senior Systems Administrator,
Large-Scale Systems, Texas Advanced Computing Center (TACC)

> With Quantum, I don't have to try to make things work. Instead I can focus on ways to better support cutting-edge research.

Frank Douma - Senior Systems Administrator, Large-Scale Systems, Texas Advanced Computing Center (TACC)

TACC
TEXAS ADVANCED COMPUTING CENTER

SOLUTION OVERVIEW

- Quantum StorNext® storage platform
- Quantum Xcellis® workflow storage
- Quantum QXS™ hybrid storage
- Quantum Scalar® i6000 tape libraries

KEY BENEFITS

- **Improved longevity and reliability** using a solution with a long-term roadmap.
- **Optimized archive utilization** by implementing user quotas.
- **Accelerated archiving performance** with parallelized workflow storage.
- **Gained scalable capacity** for fast-growing research data volumes.

From astrophysics and biomedicine to materials research and earth sciences, researchers in a broad range of disciplines capitalize on the immense resources of TACC to answer the world's most complex questions. Many of those researchers use the organization's Frontera supercomputer, which was ranked #5 on the November 2019 TOP500 list.[1]

To preserve data and enable long-term access to that data, TACC offers researchers a large-scale archive environment, called Ranch. "We currently have 55,000 defined users, with about 1,500 active users who are keeping long-term data," says Frank Douma, senior systems administrator of large-scale systems at TACC. "Of the active users, about 150 are power users who retrieve data three or four years after it is archived. They might run that data through an improved process or validate a process against a previous set of results."

TACC was using a hierarchical storage manager (HSM) solution for the archive environment, which held 1.8 billion files and offered a capacity of 160 PB of tape. But the HSM was reaching the end of its life. "The vendor discontinued one of its tape drives, software updates were slowing down, and the solution used an old, proprietary operating system that posed a support problem," says Douma. "The whole environment was beginning to atrophy. We needed a new environment that could deliver the data reliability and restorability promised in our service-level agreement."

Douma also wanted to change the way people used the archive. "Researchers were using the environment more like nearline storage, where they might store their frequently accessed files and current data sets. But it's not designed for that. It's not backed up or replicated, and it doesn't have the performance of nearline storage," says Douma. "We wanted to implement policies that would make sure researchers would use the environment for long-term archiving rather than as another place to throw any kind of data."

[1] *https://www.top500.org/lists/2019/11/*

BUILDING A NEW ARCHIVE WITH QUANTUM

After evaluating several possible archive solutions, TACC decided to move forward with a Quantum solution powered by the StorNext platform. "We recognized that Quantum could deliver the reliability and restorability we needed over the long term," says Douma.

The newly designed environment uses 30 PB of user-facing disk storage from another vendor, which is connected by InfiniBand to six Quantum Xcellis workflow storage nodes—four of which serve as distributed data movers (DDMs). The DDMs help preserve system performance during archive operations by taking data migration duties from Quantum StorNext metadata controllers and clients. The nodes are then connected to a Quantum Scalar i6000 tape library with 24 LTO-8 drives. TACC has licensed 5,000 tape cartridge slots.

TACC initially purchased a Quantum QXS hybrid disk environment with 4.7 PB thinking the center would use it for the user-facing storage. But they ultimately deployed it to support the HSM. "I wanted to give the HSM a little more life," says Douma. "By integrating the QXS solution, TACC can continue to support the legacy environment until everyone has migrated their data from it."

ENFORCING USER QUOTAS WITH STORNEXT

StorNext 6 gives Douma control over the allotment of storage to researchers. "The StorNext platform enables us to enforce policies that prevent researchers from using the environment like nearline storage," says Douma. "We've put rock-solid quotas in place for directories—each researcher gets 2 TB and 100,000 files. If someone needs more than that, they can request a project space environment. With control over quotas, we can make sure we are optimizing utilization of the archive."

ACCELERATING ARCHIVING PERFORMANCE USING WORKFLOW STORAGE

"The previous HSM environment was using a single server for 40 tape drives, which created a real bottleneck," says Douma. "Quantum workflow storage with DDMs provides parallelization of I/O. We concurrently run multiple tape drives across multiple backplanes to get much higher performance than what we had in the past.

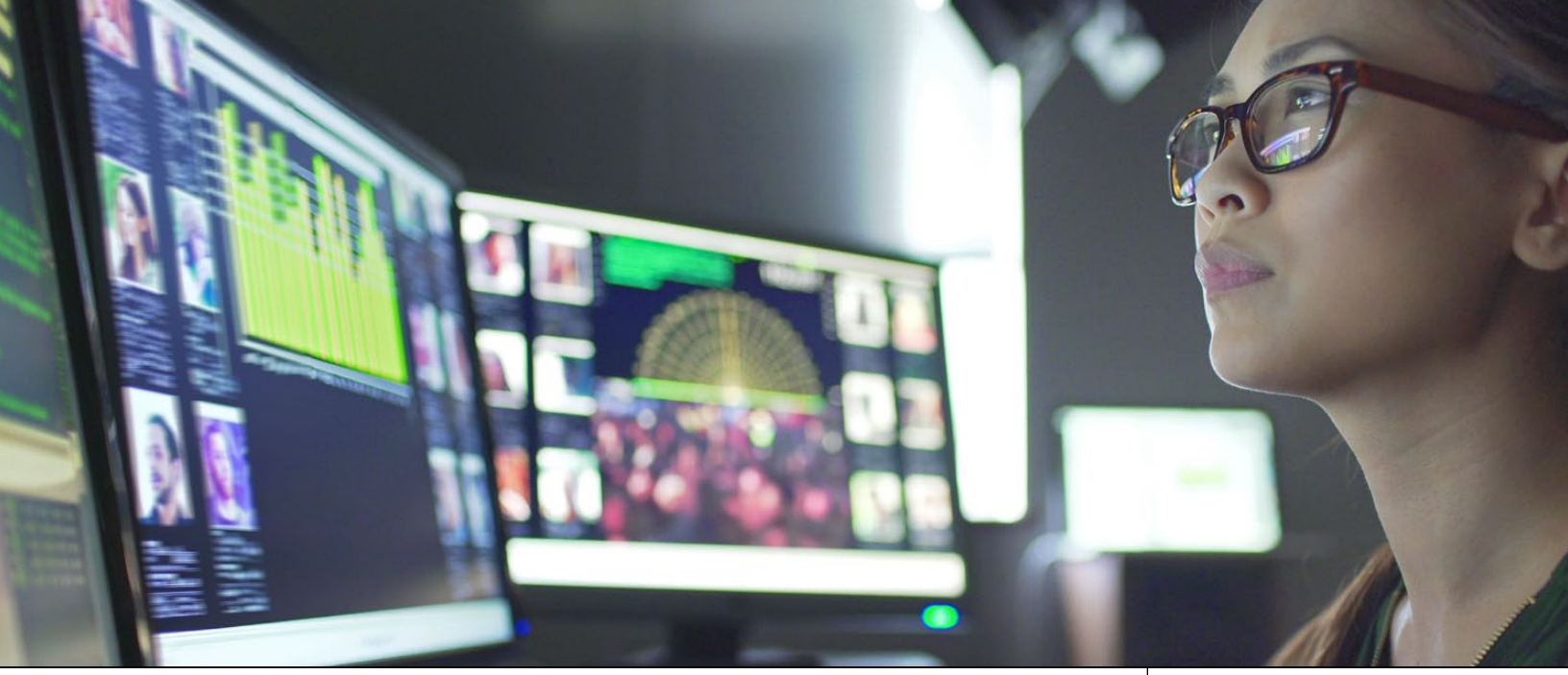
"Performance wasn't our highest priority, but the Quantum solution is delivering very strong throughput from disk to tape," says Douma. "We have 16 LTO-8 drives that write at speeds of more than 500 MB per second as they archive our primary user file system. The aggregate is about 5.5 to 6 GB per second. That level of performance is a great value-added advantage for us."

PLANNING TO SCALE THE ARCHIVE FOR GROWING DATA VOLUMES

The new environment has a licensed capacity of 5 billion user files. But in the future, TACC plans to expand the Quantum environment to

"Performance wasn't our highest priority, but the Quantum solution is delivering very strong throughput from disk to tape...the aggregate is about 5.5 to 6 GB per second. That level of performance is a great value-added advantage for us."

Frank Douma
Senior Systems Administrator, Large-Scale Systems, Texas Advanced Computing Center (TACC)



handle even more research data. With that scalability, TACC can continue to support a growing number of researchers, running more sophisticated workloads to solve increasingly complex problems.

"TACC's design, implementation, and rollout into production of a Quantum-based long-term archive has been very successful. We have no reason to believe we won't continue to enjoy that success over the coming years."

MOVING BEYOND FIGHTING FIRES

The Quantum environment enables Douma to stay focused on new initiatives and future planning rather than constantly worrying about keeping the lights on. "Ideally, the gear should just run because our design is sound, our rule sets are solid, and our user community is being held to quotas," says Douma. "With Quantum, I don't have to try to make things work. Instead I can focus on ways to better support cutting-edge research."

ABOUT TACC

Founded in 2001 at The University of Texas at Austin, the Texas Advanced Computing Center (TACC) provides powerful, leading-edge computing technologies, innovative software, and deep expertise to enable researchers to answer some of the world's most complex questions. The organization's technology ecosystem includes high-performance computing (HPC) resources, data analysis and visualization tools, cloud resources, algorithms, and more. Scientists, engineers, and other researchers tap into these resources for more than 3,000 projects each year.

Quantum

Quantum technology and services help customers capture, create, and share digital content—and preserve and protect it for decades at the lowest cost. Quantum's platforms provide the fastest performance for high-resolution video, images, and industrial IoT, with solutions built for every stage of the data lifecycle, from high-performance ingest to real-time collaboration and analysis and low-cost archiving. Every day the world's leading entertainment companies, sports franchises, research scientists, government agencies, enterprises, and cloud providers are making the world happier, safer, and smarter on Quantum. See how at **www.quantum.com**.

www.quantum.com

CS00484A-v01