



High Performance Computing (HPC) at TACC

Customer

- Texas Advanced Computing Center

Industry

- Academic and industry super computing

Applications/Use Case

- Cloud-based analytics services

Business Problem

- Aging storage infrastructure combined with a growth in customers and data
- EOL for Dell/EMC DSSD
- Annual doubling of support fees
- Vendor-lock in
- Complex manageability

Solution

- 3 Pavilion HFAs that replaced 5 EMC DSSD's
- Identical performance as the DSSD configuration
- 2X Increase in density
- 24/7 proactive support
- Standards-based NVMe-oF and OpenChoice Storage™ avoided vendor-lock in
- SAN-like manageability

The Texas Advanced Computing Center (TACC) designs and operates some of the world's most powerful computing resources. The center's mission is to enable discoveries that advance science and society through the application of advanced computing technologies.

The Texas Advanced Computing Center is more than just HPC. As a leading provider of cloud-based analytics services to firms like ConocoPhillips, Shell, Raytheon and Boeing, TACC provides the Wrangler cluster for Massively Parallel Computing (MPC) using Apache Spark and Hadoop.

This project is a collaboration between Indiana University, TACC, and University of Chicago and is supported by National Science Foundation. Wrangler is a 120 node parallel data processing cluster using Hadoop and Spark for analytics, NVMe for storage, and InfiniBand for the network transport. Existing analytics storage was based on a proprietary architecture (EMC DSSD) with an annual doubling of support fees and end-of-life notification from the supplier.

Wrangler is the most powerful data analysis system allocated in the Extreme Science and Engineering Discovery Environment, an NSF-funded virtual organization that integrates and coordinates the sharing of advanced digital services - including supercomputers and high-end visualization and data analysis resources - with researchers nationally to support science. It is used by organizations world-wide for oil and gas exploration, genomics research and aerospace designs.

Wrangler is designed for large-scale data transfer, analytics, and sharing and provides flexible support for a wide range of software stacks and workflows. Its scalable design allows for growth in the number of users and data applications.

TACC's data management and collections group lead, Christopher Jordan, wanted to replace their aging storage infrastructure as DSSD was going end of life, customers and their data were increasing, and support costs were doubling every year. TACC was also looking for a solution that met today's needs and was future-proofing. Christopher looked into various storage technologies but had difficulty finding a storage array that met all of TACC's needs.

TACC looked at DDN but disqualified it because of the inability to meet their performance needs. TACC's Dell/EMC salesperson proposed replacing the DSSD with a VMAX, but TACC found that a VMAX solution would cost over 25X more for the same performance, so they disqualified the VMAX.

Composed By Pavilion

Christopher reached out to Pavilion and worked with the sales team to design a solution that would fit their performance needs and scale easily with their plans for future growth. He found the Pavilion HFA's use of NVMe-oF gave them more performance than they got with DSSD.

Business Value

- Linear scalability that met TACC's diverse customer needs
- Capacity on-demand
- Non-invasive backup

OpenChoice Storage™ along with procurement through Dell gives us ease of deployment and technical flexibility as we leverage Pavilion's incredible performance, density, and low latencies to accelerate our massively parallel analytics workloads.

Christopher Jordan, Data Management and Collections Group Lead, Texas Advanced Computing Center

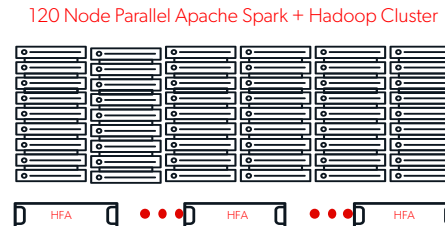
Jordan found that the Pavilion HFA provided TACC with three strategic management features:

1. On-demand reconfiguration to meet the needs of constantly changing needs of cloud subscribers
2. Thin provisioning so that cloud subscribers have capacity on-demand as expected without over-provisioning the DSSD system
3. Snapshots for zero-downtime backups

TACC saw benefit from the ability to make clones directly from snapshots, allowing them to accelerate backup operations and ease backup completion headaches. TACC also found that getting equivalent performance from DSSD cost them significantly more than the Pavilion HFA solution. They also found that the three Pavilion HFAs they needed used 50% less rack space than the original DSSD configuration. TACC also considered the Pavilion HFA because it delivered a standards-based NVMe-oF in a simple, dense and high-performance storage solution.

TACC selected Pavilion to replace their Dell/EMC DSSD storage arrays. TACC replaced their five Dell/EMC DSSD storage arrays with three Pavilion HFA's, and got the same performance. The Pavilion solution also doubled TACC's storage capacity and cut their rack space in half, saving them floor space, power, and cooling costs.

The Pavilion HFA gave TACC the increase in density needed to process assets from additional sources, improving their production workflow, and growing their customer base. The NVMe-based Pavilion HFA uses standard-based NVMe of and OpenChoice Storage™ to avoid vendor lock-in like they had with the Dell/EMC DSSD.



Summary

Pavilion worked closely with TACC to assure a smooth transition from their DSSD environment. The Pavilion HFA provided TACC with high performance, low latency, and gave them a linear price and capacity scalability. They found that the Pavilion HFA is ideal for operations while its OpenChoice Storage™ feature future-proofed their storage solution and saved them production cost. It lets them redeploy existing or purchase new NVMe SSDs as their expansion needs arise.

TACC has the Science and Technology for Academic Research (STAR) alliance program, to promote mutually beneficial exchange of ideas and information between industry and academia in advanced computing technologies, while fostering innovation and supporting economic development in the state of Texas and across the nation.

Pavilion joined the TACC STAR Alliance program where Pavilion is helping TACC promote the MPC Cloud-as-a-Service to member organizations. Pavilion participates in monthly conference calls and quarterly meeting where it advocates for the combined TACC and Pavilion HFA solution. Pavilion is also featured on the TACC

STAR web site and in the TACC STAR PR activities.

Find Out More

The Pavilion HFA is defining the future of composable disaggregated NVMe-oF. Our expertise is in simplifying and optimizing NVMe to make the impossible, possible. When storage is business-critical, there's no substitute for the guaranteed performance, functionality, high availability, and professional software support of a Pavilion HFA. We design, implement and deliver a complete solution tailored to the environment. Contact us today to get in touch with our talented extended teams of professionals.