University of South Dakota achieves high-availability sharing and long-term offsite storage with Spectra solution

Spectra T380 Tape Library, BlackPearl and Spectra NAS

> " A core requirement of our solution was its shareability across all stakeholders. BlackPearl's hybrid ecosystem easily manages our research data, moving it seamlessly to tiers of storage, while greatly improving both data accessibility and usability for USD researchers and administrators alike. "

Douglas Jennewein, Director of Research Computing at the University of South Dakota

UNIVERSITY OF SOUTH DAKOTA

About HPC at USD

University of South Dakota (USD) operates two supercomputer systems, the Lawrence Supercomputer and the Legacy Supercomputer. Coupled with advanced data platforms and high-speed research networks, and managed by the USD Research Computing Group (USDRCG), these advanced digital resources are employed to accelerate computational and data-driven research and scholarship. Their mission is to leverage this cyberinfrastructure to reduce time to discovery – growing the number of affiliated research practicioners and increasing USD's competitiveness for research funding. Access is free to all non-commercial entities in South Dakota.

The Challenge

Located in Vermillion, South Dakota, the University of South Dakota Research Computing Group advances discovery by making science, engineering, and medicine more productive. The USDRCG serves the University of South Dakota academic community, as well as collaborators both nationally and internationally. To that end, USD operates two supercomputers, named Lawrence and Legacy, that boast over 2000 and 680 cores respectively. Lawrence entered production in early 2018 and has an estimated performance of over 100 TFLOPS.

USD's supercomputing user base has grown from a small cohort of bioinformatics faculty in 2006 to virtually all Science, Technology, Engineering and Mathematics disciplines, with emerging cases in the Humanities. Their existing cyberinfrastructure had a maximum capacity of 470TB across both supercomputers. Additionally, individual laboratories managed their own data storage capabilities locally, posing a significant obstacle to cross-campus data accessibility and collaboration.

To manage a growing supercomputing infrastructure and expanding user base, USD launched the South Dakota Data Store (SDDS) project. The SDDS proposed two service tiers to serve all faculty, staff, postdocs, students and graduate students in South Dakota – the Sharing Tier and the Archival Tier.

*The Legacy Supercomputer is named after the 'Legacy' sculpture on USD's Vermillion campus*

The Sharing Tier needed to provide high-re-liability, high-availability, network-accessible data storage for research requiring persistent access to large quantities of data. The Archival Tier would offer long-term offsite archival-grade data storage. To accomplish this, the organization required a solution that would, at a minimum, double their system's data storage capacity, as well as provide the ability to scale over time.

The Solution

To store their research data, USD deployed a Spectra NAS solution, a BlackPearl® Hybrid Storage Platform and a T380 Tape Library with LTO drives. By installing the NAS with 8TB SAS drives as the local file system on one site, and the BlackPearl and T380 as the research archive platform at another site, USD implemented a multi-tier, storage-diverse architecture that enabled the high data reliability and availability they desired. USD's complete Spectra solution balances the accessibility of network-attached storage with the long-term protection, cost benefits and scalability of a hybrid storage system leveraging tape.

BlackPearl's tight integration with the Globus research data management platform allows the USD Research Computing Group to talk to all storage targets through one single interface. The Spectra NAS is used with Globus as NFS/CIFS file storage, and also as a landing zone to stage archival data. Cyberduck, and Spectra's free and open source EON Browser, are used to send any additional unmanaged data to the Archive Tier through BlackPearl. Their new system should easily increase capacity tenfold, and enables accessibility to a shared pool of storage from multiple sites.

Finally, the organization needed to move tens of terabytes onto the new system. The Spectra solution enabled USD to migrate off of a disk-based archive. BlackPearl's migration capabilities allow data to be moved from one storage type to another quickly and transparently, transferring data in the background while users continue to access assets without disrupting operations.
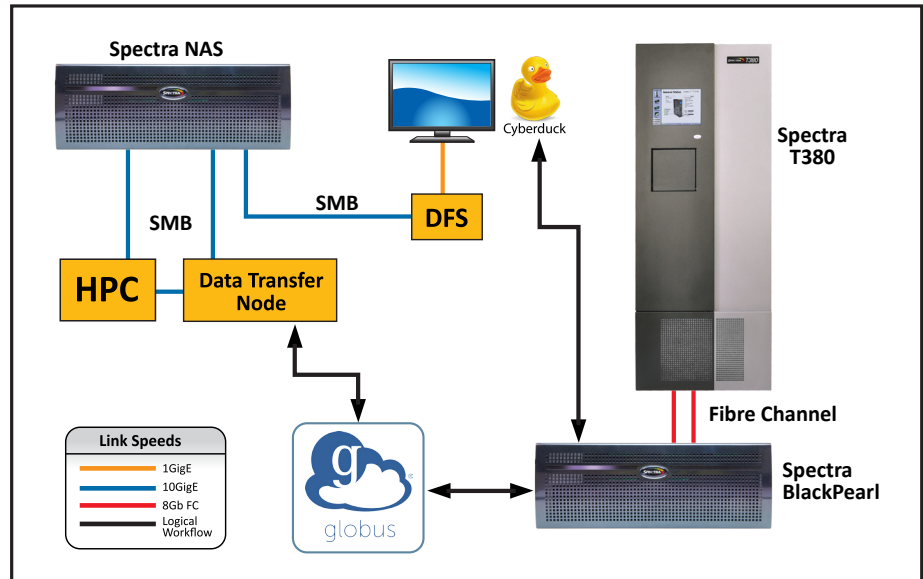
Why Spectra?

- Storage diversity
- Accessiblity
- Ease-of-use
- NFS to object storage pathway
- Integration with Globus

Environment Snapshot

- Spectra BlackPearl Multi-Purpose Hybrid Storage Platform
- Spectra 4U NAS Solution
- Spectra T380 Tape Library
- Globus research data management platform
- Cyberduck client software
- Spectra EON Browser
- Spectra Certified Media

Solution Recap

Spectra BlackPearl – BlackPearl provides an object-based, hybrid storage platform that easily and cost effectively scales up as object and file volumes grow; adapts as operational requirements change; and enables easy synchronization of data between on-prem and cloud storage. Designed for numerous concurrent workflows, BlackPearl reduces the need for expensive third-party data movers by

integrating with a range of certified clients and simple file movers.

Spectra T380 Tape Library – Designed to be easy to use and manage, the Spectra T380 is scalable from 50 to 380 LTO slots. More than 4.5PB of data can be stored in a single library. Scale to 12 LTO drives with the ability to transfer data up to 15.5TB/hr. (32.4TB/hr. compressed*) using LTO-8 tape technology. The Spectra T380 Tape Library provides maximum flexibility by supporting LTO tape technology and IBM® TS11X0 Tape technology, enabling users to select the tape technology that is the perfect for their business.

Spectra NAS Solution – The Spectra NAS Solution is the optimal disk platform for the storage of mid-tier data, including primary storage offload, data staging, backup and archiving. Flexible, simple and affordable, the Spectra NAS delivers file storage for as low as 7.5 cents per gigabyte. The expandable NAS disk solution provides raw storage capacities from 48TB to 10.7PB. Designed for a variety of workloads, a single NAS solution supports three different disk drive types, including 4TB, 8TB and 12TB enterprise drives; 8TB archive drives; and high-performance SSD drives. Reliable, economical and archive-ready, Spectra NAS simplifies the data storage process.

* Assumes 2:1 compression for LTO-5 and 2.5:1 for LTO-6, LTO-7, LTO-7 Type M, LTO-8, TS1150, TS1155 and TS1160.

V1-020923