CASE STUDY

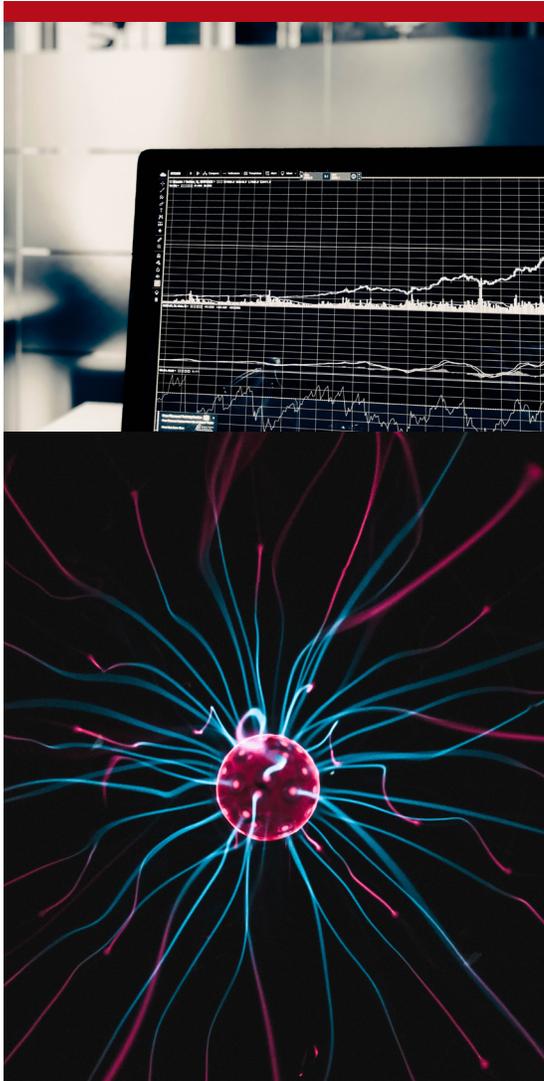## StorCycle empowers Computational Institute to store valuable research

Computational Research Institute advances the frontier of quantum mechanics, astrophysics, biology and mathematics with support from Spectra's StorCycle-based storage solution

> " Our Spectra solution has become part of the fabric of existence at the institute, allowing researchers to store copies and archives of important data sets. It works unobtrusively to provide the necessary storage management and infrastructure to support the data analysis, theory, modeling and simulation that drives our research. "

End user at Computational Research Institute

SPECTRA StorCycle®
Storage Lifecycle Management

Spectra StorCycle, T950 Tape Library & BlackPearl

The Challenge

This computational research institute on the East Coast is comprised of a community of scientists who are working to use modern computational tools to advance our understanding of science, both through the analysis of large, rich datasets and through the simulations of physical processes in astrophysics, biology, quantum physics and mathematics. The institute uses computational methods such as data analysis, theory, modeling and simulation. The institute also creates and disseminates software tools for the wider computational science community, with the goal of ensuring that technology is not a limiting factor to scientific progress. Powering their work is an onsite high-performance computer cluster comprising 7,000 cores, with plans to double the number of cores in the near future. Two offsite supercomputers managed by third parties also provide the institute's researchers with additional computational muscle.

Their research creates multi-petabyte storage requirements. The institute estimates that their data will grow at a rate of two to five petabytes per year. To safeguard the scientific data generated by the institute's users, they were looking to deploy a new research archive with an eventual backup of a 10PB CEPH file-system. The institute needed a solution that would be easy to use for end users and researchers who are highly skilled, but may not have technical large-scale storage expertise.

The Solution

The computational research institute deployed the Linux version of Spectra's StorCycle Storage Lifecycle Management software to manage multiple storage targets, enabling a complete storage lifecycle management strategy. They expanded their storage architecture by implementing a Perpetual Tier of storage with a Spectra BlackPearl Converged Storage System, BlackPearl Object Storage Disk, BlackPearl NAS solution and a Spectra T950 Tape Library with LTO-8 full-height tape drives.

The institute's new storage lifecycle management solution from Spectra Logic allows researchers to store copies and archives of important data sets using StorCycle. Researchers will archive data from CEPH filesystem to their own specific folder on a Network File System (NFS) volume via Java CLI. They will

then use StorCycle to move that data to BlackPearl based on age of the project, protecting it long-term on tape and object storage disk. The object-storage-based BlackPearl provides a portal to multiple targets in a Perpetual Tier of storage, which complements the institute's primary storage by providing lower-cost storage options. In addition, the institute also installed Spectra's free, open-source Eon Browser, which provides a simple interface that can be alternatively used for visibility into the archive and to restore access if desired.

Finally, the Globus Research Data Management Software that the institute uses integrates with BlackPearl, enabling a second workflow for Globus users to archive and restore data via a BlackPearl endpoint bucket. The research institute's new solution supports 20 to 30 users, offering the scalability they need by allowing them to easily add capacity as their archive grows. The institute is now better equipped to meet its goal of advancing scientific research through computational methods.
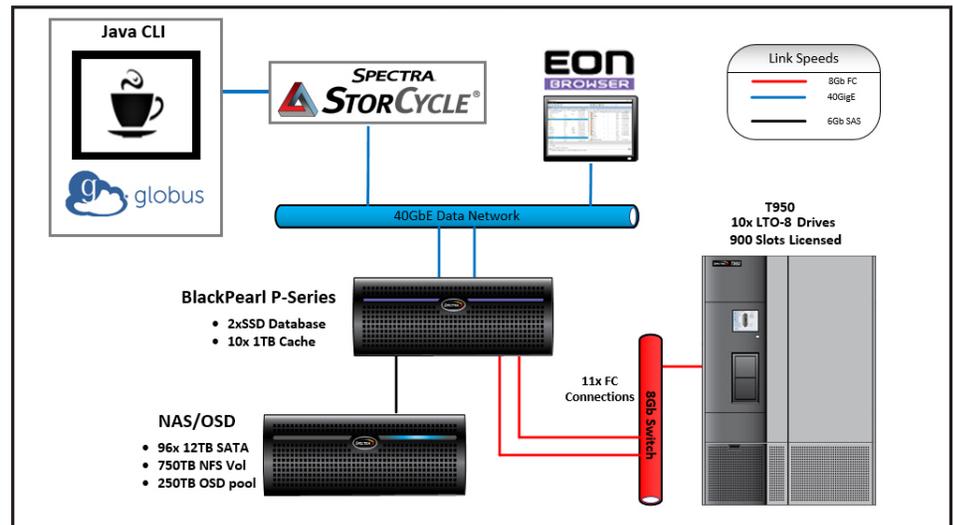
Environment Snapshot

- Spectra StorCycle® Storage Lifecycle Management Software
- EON Browser
- Globus Research Data Management Software
- Java CLI
- Spectra® BlackPearl® Converged Storage System, 4U
- Spectra® BlackPearl® Object Storage Disk Solution
- Spectra® BlackPearl® NAS Solution
- Spectra® T950 Tape Library
- LTO-8 Full-Height Tape Drives
- BlueScale® Software with Standard Encryption

Why Spectra?

- Archive architecture optimizes workflows for processing data
- Integrates seamlessly with existing infrastructure
- Open standard technology
- Scalability and integrity of object storage tape

> ❝ We've got 30 to 40PB of usable data but it doesn't all need to be touched at once. We're using StorCycle to move the data onto tape as a low-cost, low-power alternative to just expanding our storage cluster. To the users it's seamless – they just move their data into a directory and it magically ends up on tape. StorCycle automates our data movement based on policies we set. ❞

End user at Computational Research Institute

Solution Recap

Spectra StorCycle Storage Lifecycle Management Software – StorCycle is a storage lifecycle management software that automatically identifies and moves data from primary storage to a lower cost tier that includes cloud, object storage disk, network-attached storage and tape. Users can manually archive project-based directories and make additional copies for data protection. Providing seamless, familiar access and recovery of single files or entire projects, StorCycle enables organizations to manage, share and preserve their data for unparalleled long-term protection and access.

Spectra BlackPearl Converged Storage System – BlackPearl provides a modern, simple portal to a multi-tier storage architecture that breaks away from legacy solutions with capacity-based licensing models that drive greater cost when scaling. BlackPearl enables users to create a storage architecture that is cost effective, easy to manage and scalable to exabytes of data.

Spectra T950 Tape Library – The T950 library is designed and built to meet the stringent requirements of the enterprise for data integrity, data security and reliability. The T950 affordably scales in capacity and throughput, and supports multiple generations of current and future tape formats. For data archive, backup and recovery, this elite library leads the field in innovation. From the greatest storage density to proactive media management, the T950 is always protecting your data.

V1-091321