

CASE STUDY

University achieves economic, educational and scientific breakthroughs efficiently with Spectra tape library

“ Spectra’s Data Integrity Verification features are particularly attractive due to the University’s high-performance computing environment. DIV enables customers to proactively assess each tape’s condition and check data already written to tape, so customers can confirm that hours of valuable research data are safely protected and backed up for use by future scientists. ”

Nathan Thompson, CEO of Spectra Logic Corporation



Environment Snapshot

- Spectra® T950 tape library
- Four LTO-5 drives and 410 LTO-5 slots
- HPC cluster
- IBM TSM backup and HPSS software
- Fibre Channel connectivity
- Spectra Data Integrity Verification

Background

Founded in the late 1800s, this Illinois-based University is one of the world’s premier academic and research institutions. The University’s research division has had an impact around the globe, leading to such breakthroughs as discovering the link between cancer and genetics, establishing revolutionary theories of economics, and developing tools to produce reliably excellent urban schooling.

Like many research organizations worldwide, this University’s HPC division is faced with effectively managing tremendous amounts of research data across multiple departments. Simultaneously, higher education and research institutions today have an increased need to identify staff members with cross functional skill sets that include architecting big data analytics systems, a deep understanding of math coding, and storage hardware knowledge to harmoniously operate data management solutions.

The Challenge:

Fast Access to a Growing Data Set Shared by Multiple Users

The University Research division currently manages one petabyte of research data from its science departments, and that data set is expected to double in the next year. Its director of research computing implemented and manages a centralized research computing storage environment that overcomes these roadblocks. Through his architected network environment, he consolidated massive data from multiple departments and backs it up with a centralized, shared data management solution.



CASE STUDY: Research university in Illinois

Operating one tape library for research data from multiple departments, including biology, astrophysics, sociology, centralizes and simplifies data management for the entire University. The department has been able to streamline backup processes, minimize overhead and decrease overall costs thanks to the Spectra T950 tape library.

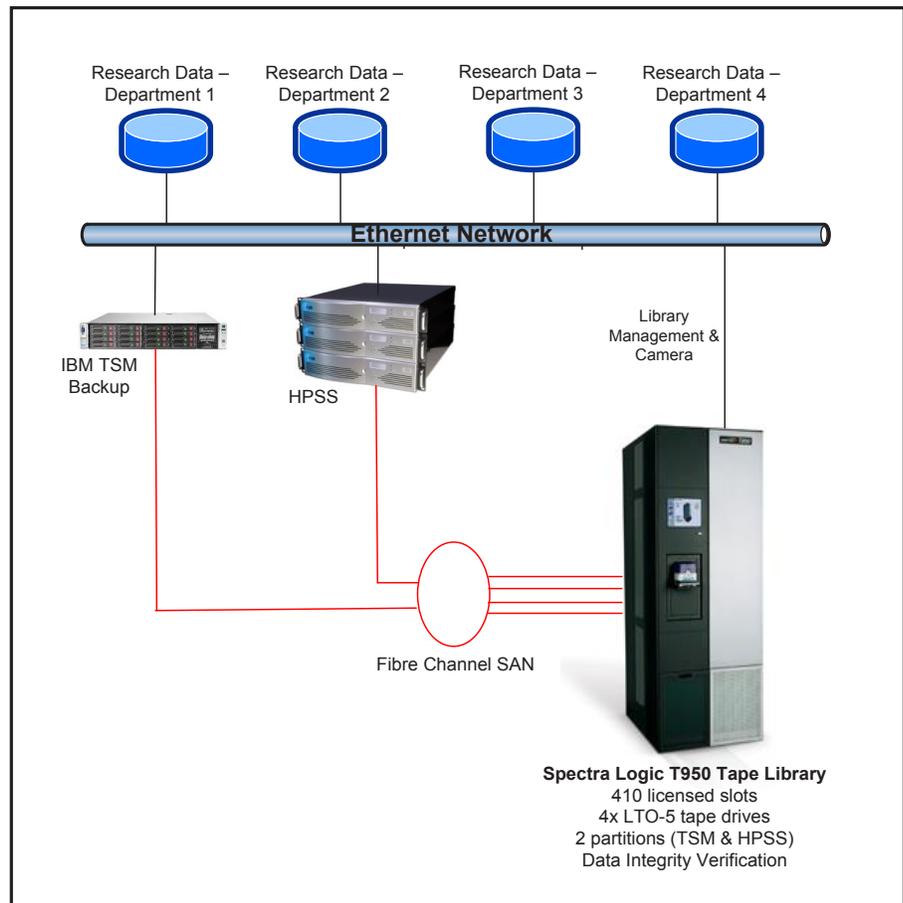
The director was a previous Spectra Logic user and already impressed with Spectra's technology for HPC research data management. When their existing tape library was decommissioned, he tasked his staff members to conduct a competitive review that ultimately validated his original instinct: Spectra tape was the right answer to address their unique data management requirements.

The Solution:

Centralized tape system consolidates data from disparate locations

Spectra tape libraries address all of the University's objectives. The library is highly reliable, so scientists from multiple departments can access data quickly as needed; expandable, to address future data growth; has the most capacity in the smallest amount of space; is easy to manage by a single staff member; and includes high availability features that no other vendor offers today.

With Data Integrity Verification (DIV), users can verify that data can be written to each tape and confirm that data has been properly written to a tape upon task completion. The Spectra library stores the "tape health" information into a database for reference purposes. DIV users can set triggers to check the health of tapes over an interval of time and/or also verify a specific tape's data integrity on request, enabling rapid spot-checks of data validity and integrity.



The University's environment is built with an HPC cluster attached to the T950 for primary backup. This configuration eliminates speed constraints when writing data from a cached storage pool in front of a tape library. Incremental backup processes are performed nightly to the T950 and a full backup is conducted to the T950 monthly. Data is currently kept inside the library indefinitely, although the University has plans in motion and grants filed to address its disaster recovery needs. Their future storage infrastructure plans includes additional frames, slots and libraries from Spectra Logic.

Why Spectra Logic?

- Consolidates massive data sets from scattered locations
- Centralized, simplified data management
- High performance, scalability and capacity
- Data Integrity Verification
- Previous experience with the product
- Spectra's HPC Expertise