



PRODUCT ANALYSIS:

Creating an Object Storage System that Bridges the gap to Tape



by George Crump, Lead Analyst

All data is not equal. It needs to be stored on different types of medium depending on the use case. Even within the archive dataset, all data is not created equal. Just like there are tiers of storage for primary data, archive data, for the maximum balance of cost and organizational efficiencies, should support multiple storage targets. The problem is archive storage is missing a middle tier of disk storage that bridges the gap to tape.

The Types of Archive Data

The reason multiple tiers of archive hardware are needed is that there are three types of data within the archive data set. First there is data that has just qualified for archiving because it has not been accessed for more than six months. This data has the highest probability of being recalled, and users will want access to that data almost as fast as if that data were still on primary storage. A scalable NAS, leveraging high-capacity hard disks, can service this class of archive data well.

There is also data within an archive that is multi-years old and may never be accessed. It is a long-term archive data set. Economically storing this data, but making sure it is accessible, is critical. While it still needs to be accessible it does not need to be recovered instantly; a few hours or even days may be an acceptable recovery time. A tape library is an ideal solution for this category of archive data. It requires no power when not in use and can scale to very high capacities. If this data is needed restoration from tape should take less than an hour.

In between the recently archived and long-term archived data sets is the middle archive. This intermediate data has not been accessed within a one-year to seven-year time frame. Object storage is an ideal target for this data set, but most object storage vendors have not optimized their solutions for this use case. These solutions have a similar four-year lifecycle of a NAS storage system. Spectra Logic recently announced ArcticBlue which increases the endurance of archive disk systems and can be the bridge from NAS to Tape.



What is ArcticBlue?

The goal of ArcticBlue is to deliver a storage solution that is half the cost of an archive NAS solution but intended to store data twice as long. If they can achieve that goal, Spectra Logic may deliver an ideal bridge to tape.

Delivering on the Full Promise of SMR Drives

The key for Spectra Logic is to manage the lifecycle of the drive in the same way that the archive software manages the lifecycle of the data. At the heart of ArcticBlue are Seagate hard disk drives that leverage Shingled Magnetic Recording (SMR) technology to deliver 8TBs of capacity.

Managing the SMR drive is an essential component of the ArcticBlue system. SMR drives are unlike standard hard disk drives that have independent tracks. The tracks on an SMR drive overlap and to make sure data remains intact, require intelligent updates to the drive. For optimal use, SMR drives need to be written to sequentially in large blocks. They don't respond well to data deletion or random I/O. The special handling of an SMR drive is very similar to tape media, which of course Spectra Logic has been doing for decades.

Not Your Father's MAID

Another critical capability in surpassing the typical life of disk archive system is powering down the drives. Most disk archive systems power their drives all the time. Doing so wastes power and shortens drive life. The concept of powering down drives is not new; laptops have done it for years, and MAID (Massive Array of Inexpensive Disks) tried to provide an enterprise solution. While spin down hard disk drive technology worked well in the laptop use case, MAID never actually worked for the enterprise.

Spectra Logic's take on powering down hard disk drives has less to do with power efficiency, although there is some gain there, and more to do with prolonging the life of the disk drive itself. The system has four large bands of drives set in a 20+3 configuration for data protection. When not in use, groups can be independently powered down. But unlike a MAID storage system, ArcticBlue can power all the bands making them all active at the same time. This power management reduces drive failure 1100% over a seven-year period.

Of course there will be times where data is requested from a powered down band. The first differentiator for ArcticBlue stems from its ability to power all bands simultaneously. The system does not have to wait to power down a band so another band can be powered up. Another advantage that ArcticBlue has over legacy MAID systems is that it is built on an object foundation. It is driven by a RestFul API which is perfectly capable of sustaining the latency of a power up cycle.



The Multi-Tier Archive

ArcticBlue provides the middle ground for an archive that requires large capacity and retention measured in decades. Organizations with this type of archive need it to be cost-effective while ensuring data integrity. To meet these requirements an archive strategy will need to include three tiers of storage. A front-end NAS tier that delivers reliable performance and high capacity. This tier will hold 1-2 years of data. Spectra Logic's Verde disk system can be the solution for this tier. The middle tier will hold data for 1 to 7 years and because of the reasons described above, and ArcticBlue fills that requirement. Then tape will be the final tier containing data for seven years to decades, and obviously, Spectra Logic provides a suite of tape libraries that meet that requirement. Ideally if disk and tape are used in the archive there should be the ability to dual copy, meaning that Data is written to both disk and tape at the same time. This not only provides protection but also pre-populates the tape tier, so when it is time to remove the data from the disk tier it does not need to be copied yet again.

StorageSwiss Take

While much of the hype in the industry is focused on all-flash arrays and production data, the significant challenge is storing the massive amount of secondary data that organizations are creating as well as meeting its unique demands. [Spectra Logic](#) has brought to market an impressive array of secondary storage solutions that promise to lower the cost of storing and retaining this data while preserving it and making it easier to access.