

Making Size Matter: Spectra Storage Solutions for Petascale (and Beyond)

Addison Snell

November 2010

White paper

MARKET DYNAMICS

Even as children, we love to build things big. Watch a toddler playing with blocks, and chances are good you'll witness an attempt to build the tallest possible tower. Carefully stacking the blocks end on end, the tower gradually moves a foot off the ground, then two, and as it reaches tabletop height and beyond the stability falters, and even the child can see it beginning to sway. The eventual demise may be hastened by (or blamed on) an unfortunate sibling or pet, but we know the project was doomed to eventual failure. Size matters, but you have to know how to get there.

As adults we still love things that are big, and not just as status symbols. Bigness represents a new frontier of discovery or achievement, whether we are engineering the world's tallest skyscraper or finding the largest known prime number. Individual electronic components may get smaller, but their capabilities are measured in giga this and tera that, each generation of technology introducing us to another obscure prefix. (For the uninitiated, mega, giga, and tera are succeeded by peta, exa, zetta, and yotta.)

Nowhere is this more pronounced than high performance computing (HPC) applications, where these unlikely sounding stats stay a leap or two ahead of mass market adoption. Terabytes sound like tasty HPC snacks these days, as our first petaflop systems approach exascale in their data volumes. For many data-intensive HPC application areas – such as oil and gas exploration, medical imaging, finance, and scientific research – the ability to store, access, and process vast amounts of data in a timely fashion is critical to new generations of discovery.

The Spectra Logic T-Finity library is designed with Size Matters in mind. T-Finity's scalability is its obvious feature, with up to 100 frames, 120,000 slots, and eight tape robots storing and retrieving hundreds of petabytes of data today, and new generations of tapes on the horizon that will extend this capability into exabytes.

But scalability of this nature cannot come



Spectra Logic T-Finity

Source: Spectra Logic

without thoughtful engineering. Like the child's block tower that needs more thought and planning than straight-up-and-down design, a tape library must have performance, reliability, and facilities features to make it relevant, stable, and useful at scale.

The first pillar of scalability for the T-Finity line is bandwidth, with dual robotic libraries and high drive counts that support high rates of data transfer. With LTO 5 tape drives (the T-Finity line also supports LTO 4, and IBM TS1130 "Jaguar" drives), the T-Finity line delivers 24.2 TB/hour¹ of throughput for a single, 24-drive frame, expandable to 120 TB/hour for a 25-frame, 120-drive library. T-Finity can place up to 25 frames in a single linear row without pass-throughs, in the frames or between the frames, that would have the potential to slow down access times to cartridges and data.

Reducing failures, and recovering from failures gracefully, are critical for managing the sheer numbers of components in petascale systems. To enable reliability at scale, Spectra targeted enterprise-class features for its HPC deployments. The high-availability architecture includes dual, redundant robotics and redundant control and communication paths that provide automatic failover when required, but Spectra has also invested in reducing the failures themselves. The library tracks tape media health and is capable of predictive failure analysis, so that tapes can be replaced before failures occur. T-Finity also supports storage virtualization and integrated data encryption as part of its BlueScale software system for the entire family of products.

At the same time, scalability to hundred of petabytes and eventually exabytes is moot if you lack a physical infrastructure that can house the system. The first issue is a simple matter of floor space. Each T-Finity media expansion frame (approximate nine square feet) stores up to 1,300 slots (about 4 petabytes of data), a compelling value proposition for replacing older, less dense offerings.

One final important consideration is the long-term outlook for your scalability investment. Tape libraries can stay active for 10 years, 15 years, or longer, incorporating new technologies as they become available. In 2009 Spectra celebrated its corporate 30-year anniversary, and the stability of Spectra's commitment to scalable archive solutions can be a comfort to HPC users whose data continue to grow exponentially along with their quest for new engineering processes or scientific discoveries.

INTERSECT360 RESEARCH ANALYSIS

The HPC industry is in no danger of running out of problems to solve, and a simple corollary to this tenet is that data sizes – and organizations' need to access them – will continue to expand rapidly. Spectra's assertion that "SIZE MATTERS" is at once both trivial and substantive. Of course it pays to be able to build the tallest tower, but building one that will stand requires attention to the myriad technologies that support scale.

The scalability of T-Finity is enabled by an architectural and software stack that integrates performance, reliability, and facilities features. T-Finity is designed not only for storing a maximum number of bytes, but also for maintaining timely access to data in a stable environment that leaves room in the datacenter for other aspects of the HPC solution. Spectra's long-term commitment to scalable tape archives for high-performance markets gives the company the opportunity to stay at the forefront of archive technologies for its customers.

¹ Compressed data.