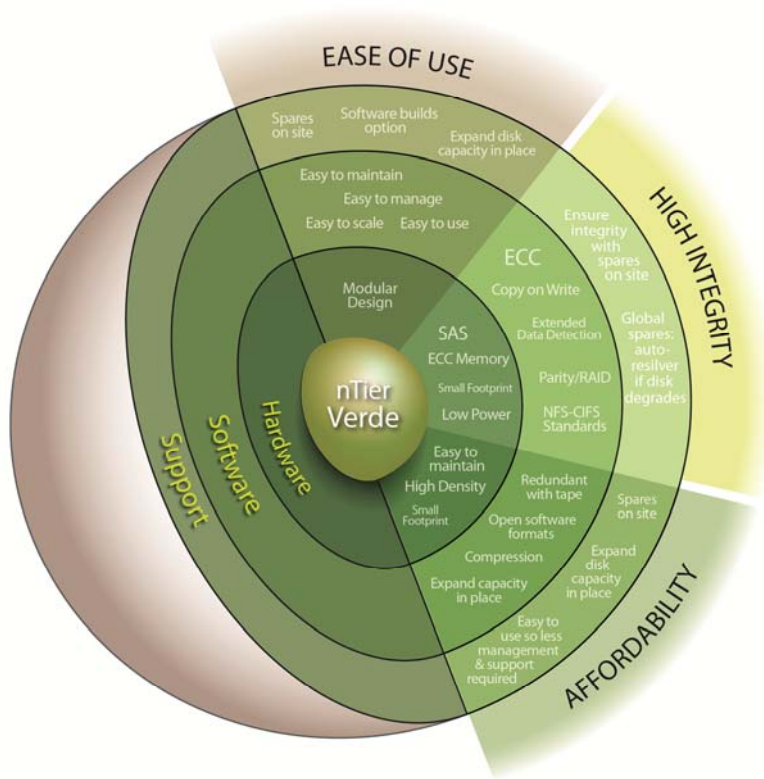


# ArchiveGrade™ Storage: Spectra® nTier® Verde Disk

---



June 2013

# Contents

- Abstract..... 3
- Introduction ..... 3
- ArchiveGrade Storage to Meet Data’s Mid-Life Demands..... 4
- Purpose-Built ArchiveGrade Disk: nTier Verde Disk ..... 4
- Making Long-term Disk Storage More Affordable Than Ever ..... 4
- Easy to Install, Maintain, Scale, and Protect..... 5
- Ensuring the Highest Levels of Data Integrity..... 6
- Conclusion..... 7

Spectra, nTier, and Spectra Logic are registered trademarks of and ArchiveGrade is a trademark of Spectra Logic Corporation. All rights reserved worldwide. All other trademarks and registered trademarks are the property of their respective owners. All library features and specifications listed in this document are subject to change at any time without notice. Copyright © 2013 by Spectra Logic Corporation. All rights reserved.

## Abstract

The massive explosion of data growth around the world is leading to new methods of analyzing data for its business value and driving new requirements for storage management. In addition, the extended lifespan of data has led to previously unimagined complexities, regulation, use cases and migration planning scenarios. To address the difficulties stemming from data growth, including minimizing errors that can be introduced as data migrates over the network and across storage technologies, a new kind of storage is required: ArchiveGrade storage.

ArchiveGrade data storage products are designed to easily and affordably handle data growth while simultaneously addressing data integrity errors that can be generated by storage media and transmission paths. The ArchiveGrade storage model, now implemented in Spectra's fourth generation nTier Verde disk (and T-Series tape libraries, not discussed in this analysis) provides easily used, scaled, and maintained storage that is also affordable. By definition, ArchiveGrade storage is designed to store data during the typically lengthy period between creation and deletion. The storage is:

- Affordable
- Easy to use, protect, and scale years into the future
- Constantly checking data integrity

## Introduction

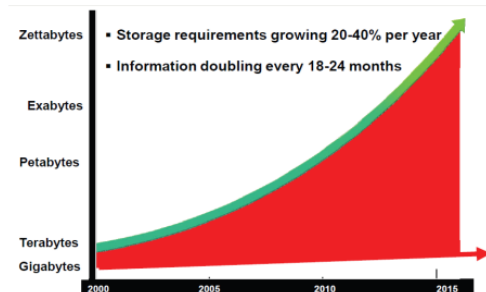
In most cases, data interacts with multiple types of hardware across its lifespan: high-performance, expensive enterprise-class and solid-state disk for data creation and active use, lower-performance disk drives for medium-term storage and tape for long-term, affordable storage. These represent the span of storage available. That leaves a lot of room in between, where requirements are less well defined. System administrators use various media with varying degrees of success in protecting data at its mid-life.

In fact, few storage platforms are built for a specific purpose. Instead, because of cost and convenience, disk and tape systems are generalists in terms of the kind of data they store and protect. These media have been appropriate for many use cases, including near-line storage and storage for less recent and less accessed data.

However, as data growth continues, the demands on data have grown, so that the generalist storage platforms are stretched beyond their capacities and capabilities. These demands, combined with shrinking IT budgets and storage longevity mandates, have driven the need for a new kind of purpose-built storage.

### Demands on data:

- E-discovery
- Regulation
- Continuity of Business
- Backup and archival storage
- Protection of personal data
- Lengthy retention while keeping data accessible



## ArchiveGrade Storage to Meet Data's Mid-Life Demands

In data's mid-life phase, the data needs to remain accessible but is too infrequently accessed to leave on high-performance, expensive disk. This longer term storage phase now warrants its own specialized storage platform. This platform needs to protect data against the challenges of longer term storage, including accessibility, affordability, scalability and concerns surrounding the risk of data corruption.

Spectra Logic is meeting this challenge with the introduction of ArchiveGrade storage, a type of storage that is designed specifically to address these requirements. This storage is affordable and easily used and scaled. With ArchiveGrade storage, sites can handle data that continues to expand at dizzying rates while remaining confident in the data's integrity.



## Purpose-Built ArchiveGrade Disk: nTier Verde Disk

Spectra Logic designed nTier Verde disk to address the specific requirements of ArchiveGrade storage, which are:

- Storing data affordably - nTier Verde disk does this through a modular design with parts that can be easily swapped and upgraded in place to preserve initial investment. The disk also reduces data center overhead with its high density and small footprint requirements.
- Easy to install, maintain, scale, and protect - nTier Verde disk provides an unparalleled ease in using, maintaining, expanding, and upgrading disk systems, all while taking advantage of industry-standard formats.
- Preserving data integrity - nTier Verde disk systems preserve data integrity by providing multiple levels of integrity checking beyond that found in typical disk systems.



*Spectra T680 library and nTier Verde, each storing 1.7 PB.*

## Making Long-term Disk Storage More Affordable Than Ever

Because IT budgets are very unlikely to ever keep pace with the growth of data creation, cost is always a consideration. nTier Verde disk meets ArchiveGrade storage requirements through its affordability coupled with its high throughput and high capacity.

---

<sup>1</sup> Chart: Stellarphoenixs blog-post of IBM graph, <http://stellarphoenixs.wordpress.com/2012/07/26/can-cloud-keep-up-with-the-rapid-increase-of-storage-requirement-of-its-clients/>. Accessed March 2013.

The nTier Verde disk system can expand capacity in place by allowing easy addition of additional chassis, and by permitting the use of larger disk drives as they become available. This makes it easy to scale while preserving your initial investment. Systems with smaller drives can be upgraded, replacing the lower capacity disk drive with higher capacity drives over time. For example, if 8 TB drives become available at a time when your system has 4 TB drives, you can move data from the 4 TB drives to a system with 8 TB drives. This doubles capacity while retaining the original chassis and without increasing rack space requirements.

nTier Verde disk supports thin provisioning, and further increases storage efficiency by virtualizing available space and allocating it as required. This greatly reduces management overhead that is required when storage space must be manually allocated.

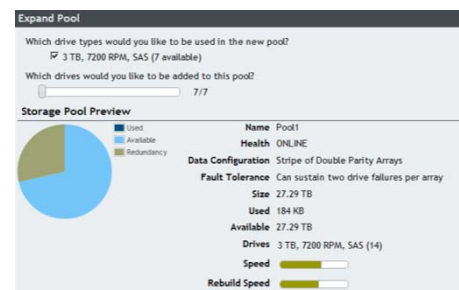
The system supports industry standard formats, such as NFS/CIFS file formats and Ethernet connectivity, so that any appropriate storage device can be used alongside nTier Verde. This prevents vendor lock-in, where systems only work with equipment from a single manufacturer.

nTier Verde disk systems support compression, a standard feature that lets you store more data onto disk, increasing disk's affordability. Verde disk is also space-efficient, which, coupled with ease of use (see next section), reduces overhead involved in managing and maintaining a data center.

## Easy to Install, Maintain, Scale, and Protect

To handle constantly accelerating data growth, ArchiveGrade disk storage needs to install and scale readily, and protect data with simplicity and ease. nTier Verde disk does this as follows:

- **Simple to Install and Use:** nTier Verde disk can be installed in an hour or less. nTier Verde disk also scales quickly, so that storage is readily integrated into and added to data center architectures. The interface is carefully designed, with extensive input from customers, so that the disk is easy to set up and use.
- **Easily Maintained - Modular and On-site Spares:** nTier Verde disk employs a modular design with parts that can be easily hot-swapped. For sites choosing self-service options, nTier Verde disk systems are available with Spectra's Assisted-Self Maintenance (ASM) program. With ASM, users store on-site many more spare components than are typically available for disk systems, including power supplies, fans, and a chassis. Using this option to store components on-site, the system and data availability is supported and the system remains affordable. Further, these options permit the user to go off the standard support grid—that is, you have the option of using in-house expertise to maintain equipment, rather than requiring that support engineers visit your site.
- **Scalable Capacity:** nTier Verde disk starts with a master node, and can grow through expansion modules. When scaled, it stores more than 1.7 PB raw data in a single 40U rack with 4 TB 3.5" disk drives. Further, with its 128-bit file system, nTier Verde disk stores an unlimited number of files and directories, so that the system can handle the anticipated flood of unstructured (that is, non-database) data.



- **Layers of Data Protection:** nTier Verde disk systems offer an option to create snapshots. This, coupled with extensive data integrity checks (see next section), help to ensure data protection and availability.

## Ensuring the Highest Levels of Data Integrity

High data integrity software and hardware, which is affordable and easy to use, is the foundation of ArchiveGrade storage. nTier Verde disk starts with an integrity level of  $10^{-16}$  through the use of SAS disk in place of the more common SATA disk. SAS ensures integrity even beyond that provided by Fibre Channel disk.<sup>2</sup> Scott Rowe, of TechRepublic's Data Center blog, says, "In reliability, SAS disks are an order of magnitude safer than [...] SATA disks."<sup>3</sup> This integrity is amplified through the nTier Verde disk system operating system use of Fletcher-4 checksums.

"SAS has a lot more intelligence onboard, and a lot of it is related to maintaining data-integrity," according to David Szabados of Seagate Enterprise's Inside IT Storage blog.<sup>4</sup> SAS provides end-to-end data internal data integrity checks, while SATA has limited checking and provides no integrity checks in memory.<sup>5</sup> SAS is also designed to work in a switched topology supporting hundreds of devices where SATA cannot expand past a single port multiplier.

nTier Verde system software supports robust data integrity through the following attributes:

- **Multi-Level Error Correction Codes (ECC):** nTier Verde disk performs checksums on every 128K block of data. The block's checksum is stored in a pointer to the data block rather than with the data block, adding a layer of protection between the corrective mechanism and the data itself. If the checksums don't match, nTier Verde disk identifies an accurate copy of that data, or rebuilds another copy through smart RAID (see below).
- **Memory with ECC and Interleaving:** An often-overlooked cause of data corruption is the memory component of the storage system.<sup>6</sup> The nTier Verde disk system addresses this by incorporating memory that incorporates integrated ECC and interleaving. ECC checksums ensure that individual errors in memory are corrected, and interleaving permits recovery from an error that affects more than a single bit within the memory of the system.
- **Internal Data Handling:** Data is never overwritten by updated versions of itself. Instead, the data remains on disk and a second copy is created. This permits the simple addition of snapshots if appropriate for the data set.

<sup>2</sup>Lowe, Scott. "How SAS, Near Line (NL) SAS, and SATA disks compare," Techrepublic.com, Feb 2012. <http://www.techrepublic.com/blog/networking/how-sas-near-line-nl-sas-and-sata-disks-compare/5323>. Accessed March 2013.

<sup>3</sup> Ibid.

<sup>4</sup> Szabados, David. "SAS Mythbusters: Data highways and SAS vs. SATA," Enterprise Media July 2011. <http://enterprise.media.seagate.com/2011/07/inside-it-storage/sas-mythbusters-data-highways-and-sas-vs-sata/>. Accessed March 2013.

<sup>5</sup> Intel, "Choosing between SAS vs. SATA Hard Disk for your Server RAID System," Server Products, July 9, 2012. <http://www.intel.com/support/motherboards/server/sb/CS-031831.htm>, Accessed March 6, 2013.

<sup>6</sup> Zhang, Yupu, Abhishek Rajimwale, Andrew C. Arpaci-Dusseau, et al. "End-to-end Data Integrity for File Systems," Computer Science Department, University of Wisconsin-Madison.

- **Copy on Write:** With nTier Verde disk, data writes precede and are completed as a step separate from parity writes, so if power is lost between the copy and the parity write, data is not silently corrupted.
- **RAID, Improved:** nTier Verde disk lets you define the number of parity disks to use. Set levels to 0, 1, 2, or 3 to indicate the number of disks that can fail without risking data loss. When using parity of greater than 0, this method of ensuring data redundancy improves on RAID by:
  - Striping data using nTier Verde system software, removing the risk associated with hardware RAID controllers
  - Intelligent re-silvering on a rebuild, restoring only allocated/written data, rather than rebuilding the entire disk, which may include unused disk space
- **Global Spare with Automatic Re-silvering:** If the system identifies a disk that is degrading, the system automatically takes advantage of global spare disks in the system—disk that is not yet assigned or in use. This automatically ensures ongoing data integrity, and as mentioned above, nTier Verde increases rebuild efficiency by restoring only the used portion of the disk.

## Conclusion

The growth of data has driven the accretion of storage in the data center, with most storage built on general-purpose platforms: SATA/JBOD disk and tape. The data growth has led to new uses of data, along with increased demands for data protection and long-term data retention. The current model of data use and storage, starting with creation on primary disk and then storage on SATA and tape for backup and disaster recovery, no longer sufficiently supports how data is accessed and used now. These uses for and demands on data require a new class of data storage that integrates specific features that ensure high data integrity and systems that are affordable and easy to use.

This new ArchiveGrade storage model is implemented in nTier Verde disk. With a focus on data integrity and long-term storage, along with flexibility, scalability, and affordability, storage can now tackle the serious issues of long-term data storage and protection while supporting easier management of steadily increasing data growth.



[www.SpectraLogic.com](http://www.SpectraLogic.com)

Spectra Logic Corporation  
6285 Lookout Road  
Boulder Colorado 80301 USA  
800.833.1132  
303.449.6400

For more information, please visit [www.SpectraLogic.com](http://www.SpectraLogic.com)