



Active Archives for Media and Entertainment

High-throughput, single-tier disk archive for large media files

AmpliStor Solutions Paper
March 2012
V 3.0

Active Archives for petabyte-scale Unstructured Data

While the adoption of disk storage systems for backups has successfully occurred, long-term electronic data archives are still largely maintained on magnetic tape. Tape's advantages as a medium for long-term archiving are quite clear as well: very low media cost, low power cost (zero for offline tapes), extended archive life and removability. For these reasons, tape still is the incumbent technology for long-term data archiving of large-scale data, including the critical digital assets managed in media and entertainment. Many storage architectures will store 3 copies of all data on 2 media, of which one will be kept remotely. Typically, that is one copy on Tier 1 NAS or SAN and two copies on tape.

It is clear that tape does have challenges that must be carefully considered. Among them are access performance, accessibility, reliability, data integrity assurance, technology compatibility, scalability and management cost. Archives maintained on tape are typically offline, with high retrieval latencies, and hence are impractical and inconvenient for use as an "active archive". The



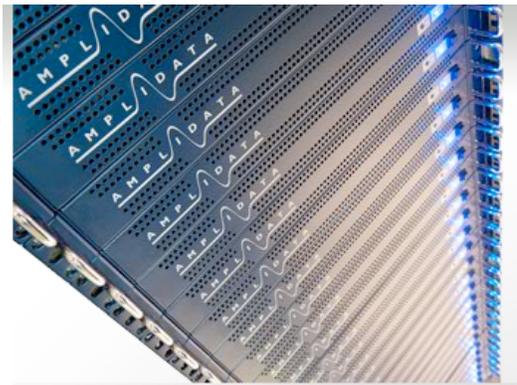
The management of large tape libraries also increases the overall cost of ownership beyond what is promised by the media by itself. Guaranteeing data integrity is another key challenge for electronic archives, and maintaining data integrity over time requires trusting the tape media. The problem of tape media refresh and migration will continue to become worse as capacities increase to hundreds of petabytes and beyond. The problem of managing, migrating and refreshing tapes at this scale becomes analogous to painting the Golden Gate Bridge from end-to-end while the bridge is growing in length – it's by definition a never ending problem!

Access to archival data is a key requirement for applications in industries such as media & entertainment, where large digital media assets are frequently created as original content, and then re-purposed in the future for other purposes (for example, distribution on new media such as Blue-Ray, or released in different markets or other channels). A system that can provide an alternative media for these "active archives", with the performance and convenience of disk, but at a new low-cost of ownership, would provide a tremendous advantage for these applications.

AmpliStor Optimized Object Storage

With the rapid advancements in disk storage technology, an opportunity is now arising to maintain petabyte scale data archives on very inexpensive, high-density disk media, and to do so in a highly reliable manner. The problems that need to be solved to make this a reality are to provide ultra-reliable long-term archive life (archives are essentially a “primary” copy of data without an additional backup), assured data integrity, seamless scalability, automated management and low operational costs.

For this purpose, Amplidata has created AmpliStor, the first purpose-built Optimized Object Storage (OOS) system specifically designed to address the challenges of reliable, low-cost active archive for very large scale unstructured data. AmpliStor is a scalable solution, that enables the maintenance of tens of petabytes of data and beyond, and is specifically optimized to provide reliability and efficiency for large-scale unstructured data, including media assets



(images, audio, video) as well as large documents, backup streams and other binary assets. AmpliStor solves the key issue of reliability on multi-terabyte disk drives, in order to provide an unbreakable storage platform. Moreover, it does so with the following key advantages over tape:

- **Online performance:** AmpliStor provides instantaneous access to large unstructured data objects. The system can scale-out performance to tens of Gigabytes per second.
- **Auto-management:** the system can be easily scaled to petabytes by adding more storage nodes, with the system self-managing capacity, monitoring system health and self-healing failures within the system, such as disk drives or storage nodes.
- **Data migration:** the system provides auto-migration capabilities to new storage nodes, with new higher disk densities. This enables very long-term archives and eliminates the problems of tape media obsolescence as well as complex data migration problems.

All of these advantages are combined into a system that provides up to 90% power savings versus other disk based storage systems, and provides a new low cost point, and total cost of ownership. In many cases, the TCO of AmpliStor is much closer to tape than to other disk based storage solutions. AmpliStor provides very low costs both from a capital and operational perspective. For large data archive applications, this provides a number of distinct advantages:

- **Low-latency, online access to data:** AmpliStor provides a fully online archive storage system, with all data immediately available on spinning disk media. The system enables immediate access to any object in the system through native object interfaces.
- **Unbreakable Storage:** AmpliStor provides any desired level of data reliability of availability – by tolerating any (user-specified) number of failures. It solves the reliability issues of RAID on multi-terabyte disk drives. Data integrity is pro-actively monitored, verified and assured through checksums and disk scrubbing. Component level failures such as disk failures are automatically resolved by the self-healing design.
- **Long-Term Archive Life:** the system combines ultra-reliability, assured data integrity, plus an architecture that enables organic growth to incorporate newer technologies. After the useful-life of disks is exceeded, new generations of (higher-density) storage nodes can be incorporated into the system, with automatic, non-disruptive, migration of data to the newer storage capacity. This capability will “future-proof” large-scale data archives across leaps in technology in the coming decades.
- **Scalability with low management effort:** AmpliStor can be scaled with additional AmpliStor storage nodes on the fly. The system auto-detects and utilizes new capacity, to scale the system without reduced manual intervention. Self-monitoring of disk & node-health plus self-healing reduces the need for intervention as the system grows.
- **The lowest storage overhead:** For petabyte-scale data archives composed from millions of big files, AmpliStor utilizes BitSpread to provide super-reliable protection from data loss without the lowest possible overhead. AmpliStor will require 50-70% fewer disk drives and data center rack / floor space than RAID and Replication based storage systems.
- **90% less power consumption:** AmpliStor has the lowest storage overhead, leverages low-power disk drives, plus efficient storage enclosures to provide the most reliable storage for only 7 Watts per TB. Future versions of the AmpliStor software will enable disk drive and node-level power-down, to further decrease power cost to the level of offline systems such as tape.
- **AmpliStor provides 50-70% lower Total Cost of Ownership:** all of these capabilities are provided at very low Capex and Opex(data center floor space, power and cooling

costs), and lowest management overhead to provide an incredibly low TCO.

The AmpliStor system architecture consists of AmpliStor storage nodes and AmpliStor controller appliances, connected over 10 Gigabit or one Gigabit Ethernet networks. The system can be scaled from a small configuration of a few storage nodes up to thousands of nodes, accessed by hundreds of controllers to serve large groups of concurrent users. This enables the system to be deployed in a range of backup and archival applications with a broad spectrum of requirements. For additional information on the AmpliStor system, see us at www.amplidata.com, email info@amplidata.com or contact your local sales office.

Amplidata European Headquarters

Antwerpse Steenweg 19
9080 Lochristi, Belgium
Tel: +32 9 324 25 90
Fax: +32 9 324 25 99

Amplidata US Headquarters

2100 Seaport Blvd, Suite 400
Redwood City, CA 94063 USA
Tel: 650-367-9683

Addendum: Comparing Disk Storage versus Tape for data archiving

From a performance perspective, access latency is a primary concern for tape systems for use within active archives. The best-case scenario for tape systems is a latency of seconds, which happens if the file or object that is being accessed is on a tape already loaded into a tape drive. If the tape is not loaded, the tape will have to be located and loaded, which can take minutes (in the case of a robot), or hours in the case of a human operator. Ultimately, the time it takes to seek a tape drive is much more than the time it takes a disk system to position the head of the disk drive.

Tape reliability has several concerns, including the issue of MTBF (Mean Time Between Failures). As tapes are accessed more frequently, the more they are physically manipulated, the worse the MTBF becomes. This is termed the shoe-shining effect: the more the heads rub over the tape, the sooner the tape will become less reliable and eventually unusable. Symantec NetBackup considers a tape that was mounted 32 times as unreliable. The tape can be re-used but only if all data on it is completely rewritten.

Reliability also relates to the ability for tape systems to actively verify the media for integrity. Here, there is a significant difference between data stored on disks and data on tape. With tape, users must trust vendor claims for media integrity, since there are no active processes like active data verification & assurance (data scrubbing, healing). As an alternative, tapes can be periodically loaded and rewritten to verify data integrity. More recent high-quality tape media such as LTO5 have a very long life span rating, with some tape formats specifying a 30-year lifespan at 20-degrees centigrade and 40% non-condensing humidity. These ranges of temperature and humidity might sound acceptable, but they are actually quite narrow. The more the storage conditions depart from those ideals, the shorter the life of the tape.

Global access is the ability to access data from anywhere. By default, tape is only accessible on location, where the tape is mounted. To enable remote access, extra devices will be required such as a remote tape mounting and access solution. These are typically very cumbersome to install and maintain. Alternatively, making copies of the same tape and distributing these copies to multiple locations can also help. Users also face the problem of technology evolution: LTO1 tapes created just a few years ago are no longer readable by LTO5 tape drives. This forces users to continually migrate, or otherwise maintain a collection of obsolete tape drive hardware just to read older generation tape media.

Tape systems are innately scalable from a media perspective. More tape cartridges can be acquired at will, and with the latest LTO tapes the individual tape capacities run into terabytes per

cartridge. There is a downside however: the size of the cartridges and the number of cartridges in a library affect the availability and reliability parameters. More tapes means more time needed to locate a specific one; bigger tapes capacities means longer random access times.

As stated earlier, one of the central advantages of tape versus disk is the very low cost of media. Regardless of how disk prices have declined, tapes are still dramatically lower in cost. However, building a reliable tape archive that tackles the challenges mentioned above requires very expensive tape drives and robots. For a passive archive, most challenges can be addressed at an acceptable management cost. As soon as the archive needs to provide remote access, low latencies, and automated management, the required costs will increase.

Given the low cost of tape, it is also tempting to make additional copies of data on tape. Copies are typically made at the same time as the original, so the copy faces exactly the same challenges. Fresh backups can be created every so many years, but that again causes the shoe-shining effect. This makes the overall management of the tape archive very complex, unless once again expensive robots or silos are acquired and configured to automate these tasks.