Data Archiving Using Enhanced MAID

Aloke Guha
COPAN Systems Inc.
aloke.guha@copansys.com

Abstract

This paper discusses archive data and its specific attributes, as well as the requirements for storing, cataloging, managing, and retrieving archive data. We discuss how MAID (Massive Array of Idle Disks), and particularly our enhanced implementation of MAID, are particularly well suited for the storage of long-term archive data. Finally, we present an overview of the Revolution 220A file archive product with MILLENNIA ARCHIVE™ software and how it is in use today in some typical applications.

1. Introduction: Archive Data

An archive is a collection of data that is maintained as a long-term record of a business, an application, or an information state. Archives are kept for auditing, regulatory, analysis, or reference purposes, rather than for application or data recovery. Although much of the emphasis on archiving has historically been on risk mitigation, there is an increasing focus on maximizing the business value of information by providing more direct access to the archive data.

The storage infrastructure and the management of an archive depend on the type of data in the archive and its frequency of access. Some information will only be accessed very infrequently, if at all, and can tolerate the long access delays inherent in tape storage. However, most archive data should be readily accessible for either regulatory or business reasons, and requires a storage platform with fast response times and data integrity. Another important requirement is the economics of archive storage. Because archives are retained for long periods of time and may accumulate to large volumes of data, storage costs must be low.

2. Attributes of Archive Data

Archive data has an additional set of attributes than those of transactional data or data protection. In addition, the management of archive data has specific requirements.

2.1. Retention and Disposition

Most archives are meant to be retained for a long period of time, and the required retention period may extend well beyond the life of the storage device. The data must be guaranteed to exist, and must be maintained with integrity, even as the storage devices fail and/or are replaced after obsolescence. Compliance requirements specify how archived data should be disposed of once the retention period expires. Some regulations mandate that all copies of the data be verifiably destroyed.

2.2. Performance

Archive performance includes the time it takes to locate or find the correct data file or object, the access time for retrieval of the file, and the transfer rate at which data is migrated into the archive and/or retrieved. For small objects, access time becomes more important, while transfer rates become significant in storing or retrieving large objects such as multimedia files.

2.3. Content Retrieval

Since the size of an archive can be expected to significantly grow over time, the ability to easily retrieve specific information, or a particular file, is extremely important. An effective archive management system allows data objects to be retrieved using many methods, for example, by name or by specifying multiple query parameters. This includes file or content attributes or user-defined attributes that tag the information during the archival process.

2.4. Regulatory and Compliance Guidelines

Archive records are increasingly dictated by regulatory guidelines such as SEC-17a-4, Sarbanes Oxley, and DoD 5015.2-STD. These regulations specify requirements on retention period, disposition, data integrity, security, access controls, audit trails, and data availability.

2.5. Data Immutability and Data Integrity

A common characteristic of archive data is immutability, often driven by regulatory requirements. The archive management must ensure that stored records cannot be changed, whether accidentally or intentionally. All archive records have to be read-only until the expiration date, at which point in time they can be deleted

3. MAID and Enhanced MAID

The term MAID stands for Massive Array of Idle Disks and was first introduced in 2002 [1]. Since 2003, MAID storage has been gaining interest for both data protection and archive data storage. SNIA [2] defines MAID as *"a storage system comprising an array of disk drives that are powered down individually or in groups when not required. MAID storage systems reduce the power consumed by a storage array."*

We have developed an application-driven, enhanced MAID platform in which a maximum of 25% of the drives are powered on at any one time. This platform is specifically designed for the challenges of storing long-term data, and the infrequent access requirements of archive data. This approach results in a much lower cost than traditional disk-based systems, and is competitive with the cost of automated tape libraries. In addition, the higher density packaging of the drives provides a smaller footprint and consumes less power for operations and cooling.

Our enhanced MAID platform incorporates the following technologies.[1] Overviews of our enhanced MAID platform architecture, POWER MANAGED RAID software®, and DISK AEROBICS software® can be found in published references [3, 4].

3.1.1. Three-Tier Architecture

Three levels of processing separate functionality, simplify system management of long-term data, and allow performance to scale with capacity.

[1] COPAN Systems has multiple patents pending on these technologies.

3.1.2. POWER MANAGED RAID Software

All disk partitions are RAID protected using a scheme that generates parity, even as drives are power managed. The controlled power management of the drives increases the service life of the drives, which significantly increases overall data reliability.

3.1.3. DISK AEROBICS Software

Drive health is proactively monitored and managed by periodically exercising all disks, verifying data integrity, and detecting potential drive failures before they occur. Data is copied from a "suspect" or potentially failing drive to a spare healthy drive, thereby avoiding putting the RAID set into a critical or vulnerable state.

4. Archiving with Enhanced MAID: The Revolution 220A

We have developed an enhanced MAID platform designed to be the repository of choice for long-term archive data for the following reasons:

4.1. Tuned to Archive Workload

Drives are powered on only when needed, consistent with the infrequent access workload of archive data.

4.2. Access or Read Performance

Archive data needs to be accessible in minutes, not hours or days, in order to obtain maximum business value and meet regulatory guidelines. The performance for read access to MAID depends on whether the data resides on a spinning or idle drive. With our enhanced MAID platform, the nominal first access delay is 13 to 14 seconds when the access is to a powered-down disk partition. After the initial spin-up latency, all subsequent accesses to the same disk partition are at the same performance as all-spinning disk (milliseconds). This is especially beneficial in archive data access where multiple retrieve requests to the same or related data can occur in burst fashion long after the archive process. Therefore, archive data access on MAID is closer in performance to all-spinning RAID disk than tape.

4.3. Migrate or Write Performance

Because many archives are created by migrating data from primary storage systems to an archive storage system, write performance is an important metric unlike the read access throughput for individual file access. Our enhanced MAID platform comprises multiple storage

shelves, each with its own dedicated RAID hardware. Therefore, system bandwidth or performance scales with capacity.

4.4. Drive Reliability and Data Integrity

As described in Section 3, the combination of POWER MANAGED RAID software and DISK AEROBICS software significantly increases drive reliability. COPAN Systems has been tracking drive reliability since the introduction of its enhanced MAID product in April 2004, and is experiencing drive reliability in excess of four times that specified by SATA drive manufacturers. In addition, DISK AEROBICS software checks to ensure the integrity and accuracy of all data at least once every 30 days.

4.5. Scale and Capacity

Since archive data continuously grows over time, it is best maintained on a high-density storage system that can provide a large capacity in a small footprint. Our disk drive packaging and cooling and vibration management allow 896 drives to reside in a single cabinet, providing for 22.4 TB/sq ft with 250 GB drives and 44.8 TB/sq ft with 500 GB drives. The Revolution 220A scales up to a billion files.

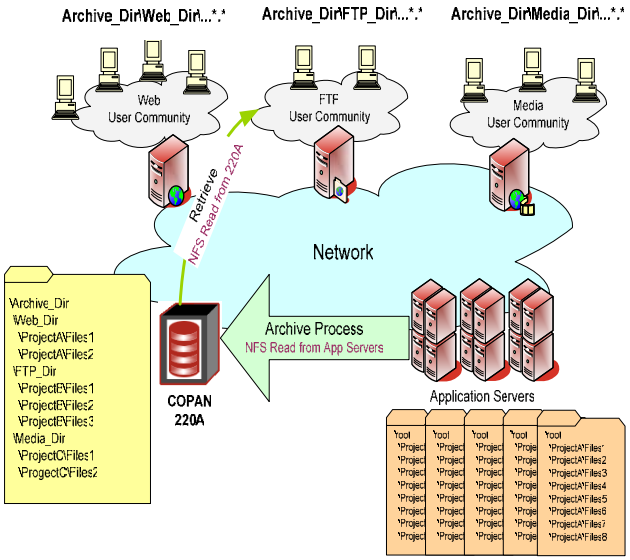

Figure 1. Archiving and Retrieval in Revolution 220A

5. MILLENNIA ARCHIVE™ Software

The Revolution 220A [6] is based on the enhanced MAID platform described earlier, and incorporates MILLENNIA ARCHIVE software, which was developed to specifically address archiving needs such as automated data movement, data immutability and data integrity, security, and audit trails.

The Revolution 220A appears to application servers as a NFS file access storage device, and the MILLENNIA ARCHIVE software manages the migration of files to and from application servers. When files are migrated to the Revolution 220A, the software stores, catalogues, manages, and protects all files. Any client on the network with adequate access rights can retrieve a file from the Revolution 220A by executing an NFS mount of the appropriate directory. Figure 1 schematically depicts the archiving and retrieving to and from the 220A.

In addition to the archive benefits provided by the enhanced MAID platform, MILLENNIA ARCHIVE software adds the following archive related benefits.

5.1. Scheduled Archiving

The software provides for either on-demand or batch-processed archiving based on a specific date, time period, or event as shown in Figure 2.

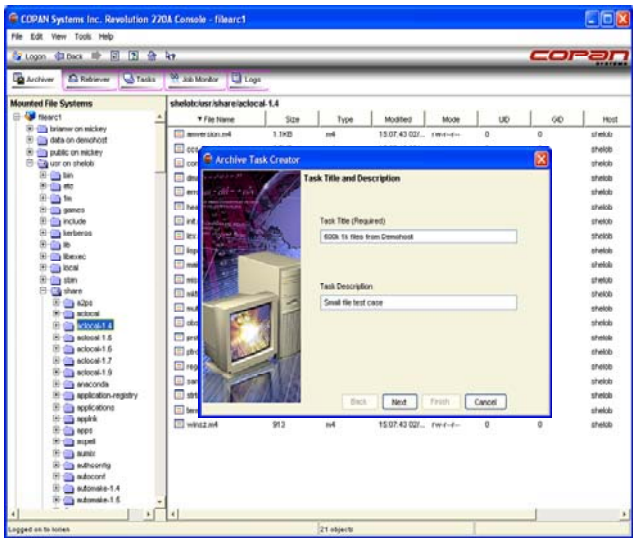

Figure 2. Scheduling an archive task in the Revolution 220A

Multiple desired attributes of the archive can be specified using a simple and intuitive API in the MILLENNIA ARCHIVE software such as Retention Period, Number of Copies Retained, Number of Versions Maintained, and optional user-defined attributes specified for future retrieval of the files archived.

5.2. Retention and Disposition

Users can designate a specific retention period and expiration date for each file. Unlike removable tape media, where tracking and updating many floating tape copies can be arduous, the Revolution 220A with the MILLENNIA ARCHIVE software provides a closed mechanism to track and delete all copies of the archive data. This ensures electronic shredding, which is very difficult to achieve in any removable media storage system such as tape or optical archives.

5.3. Content Retrieval

Flexible query-based search and indexing schemes in the software ensures that users can retrieve the specific file they need based on inherited file metadata (e.g., system from which the file was archived, name and type of file, user and/or group that migrated the file) as well as user-defined file metadata or other attributes that were specified at the time of archiving to the Revolution 220A. The retrieve process works as simply as retrieving a file from a traditional networked file system server such as a standard NFS server (Figure 3).

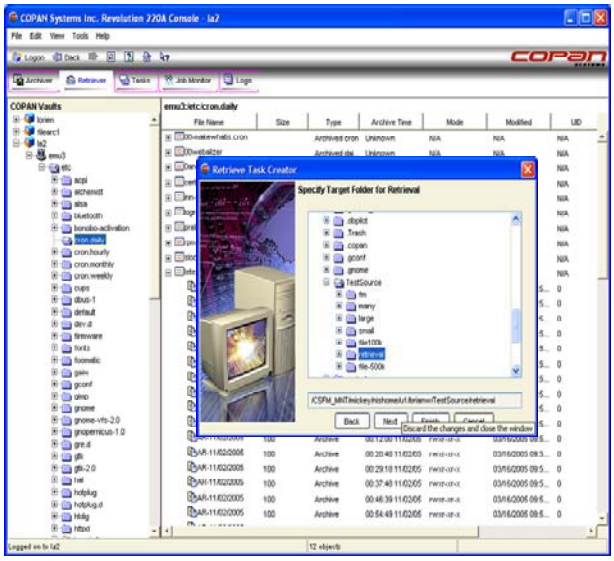

Figure 3. Retrieving an archived file in the Revolution 220A

5.4. Data Immutability and Data Integrity Checking

Each archived file is protected against deletion or modification until the specified expiration date. Every file retrieved from the Revolution 220A is verified using the SHA 256 hash.

5.5. Automated Versioning

Users can specify the number of versions to maintain of a file, and easily retrieve a specific version when required.

5.6. Security

The software imposes a security layer to prevent unauthorized access to the archive. In addition, because the disks can be powered off and kept offline, there is an additional security mechanism since data on those disks is not accessible. Furthermore, using a disk-based archive means that the data cannot be physically removed, preventing data breaches that can occur when tapes are in transit.

5.7. Audit Trails

Identifiable audit trails are provided for all archive and retrieval events.

6. Applications for the Revolution 220A

The Revolution 220A is in use by several customers for archive management. Here are examples of a few applications:

6.1. Archive of historical stock trading data for market analysis

A brokerage firm had been storing stock tick data that reflect daily trading fluctuations for stock market analysis. This data had been earlier stored on tape and is now being archived on the Revolution 220A using the MILLENNIA ARCHIVE software. Previously, retrieving a specific stock trade data sometimes took two to four days given the operational complexity of locating and handling tape media. With the archive on the Revolution 220A, all stock tick data is effectively online. A trader can now retrieve trade data directly from the archive using NFS mounts and reads.

6.2. Archive digital assets to enable easy reuse for new business purposes

A film/animation company is currently using the Revolution 220A for the long-term storage of its digital project assets (sound, movie clips, footage, animation, special effects, etc). When a movie is completed, the MILLENNIA ARCHIVE software migrates the files from expensive primary storage to the Revolution 220A. Search capabilities enable developers to browse the archive for specific footage to be re-used in a sequel, commercials, DVDs, promos, etc. Prior to the Revolution

220A, the archive data was stored on tape, and specific files were often so difficult to find that the data was recreated rather than retrieved.

6.3. Store archive data in easily accessible file formats for regulatory/customer support

A ticketing company that provides ticket sales and distribution through one of the largest e-commerce sites on the Internet is required to keep transactional history for many years for customer support and potential litigation. The company is deploying the Revolution 220A as a way to make all its digital data readily available to applications, instead of using expensive primary storage or inaccessible off-line media.

7. Conclusions

With applicability of MAID storage to long-term data, COPAN Systems has developed the Revolution 220A based on an enhanced MAID platform and using specialized archiving software – MILLENNIA ARCHIVE software – to specifically meet the needs of affordable, highly reliable, disk-based storage for archive data.

8. References

[1] D. Colarelli and D. Grunwald, Massive Arrays of Idle Disks for Storage Archives, 2002 ACM/IEEE conference on Supercomputing, November 2002, Baltimore, Maryland.

[2] The Dictionary of Storage Networking Technology, Storage Networking Industry Association (SNIA), 2005/2006, pp. 121-122.

[3] A. Guha, "A New Approach to Disk-Based Mass Storage Systems," 12th NASA Goddard - 21st IEEE Conference on Mass Storage Systems and Technologies, April 2004, College Park, Maryland.

[4] A. Guha, "Tuning MAID Storage for Backup and Archive Data," 13th NASA Goddard – 22nd IEEE Conference on Mass Storage Systems and Technologies, April 2005, Monterey, CA.

[5] COPAN Systems Revolution 220A: http://www.copansys.com/products/specifications_220A.htm