ACTIVE Archive

THE ACTIVE ARCHIVING ECOSYSTEM

Building a Flexible Archival Repository Your Way

Take control over your archival data's accessibility, cost, and security

July 2022

FOREWORD

Organizations want to ease the financial and operational burden of buying storage for fast-growing data and supporting legacy systems. Archiving is a popular method, but it has its drawbacks. Identifying aging and inactive files is time-consuming for storage admins. If busy IT teams even get that far, they archive data to tape or cloud-based cold storage, where it languishes on its cheap storage tier with little value or lifecycle intelligence.

Active archiving is different. The multi-vendor ecosystem accommodates various storage systems and platforms, from disk and tape to on-premises and the cloud depending on client needs. The software intelligence layer creates a virtual archival repository and may automate data movement to media, balanced for performance and cost.

In healthcare, active archiving supports multiple use cases and reduces IT operational overhead. Patient care improves with an active archive that enables the healthcare organization to consolidate patient records from multiple providers. Clinicians use a single sign-on to launch the patient's complete digitized historical record to provide continuity of care. The active archive system also satisfies regulatory requirements with strong user access, activity logging, and integrity protection.

No matter what vertical you're in, active archiving keeps your archived data immediately accessible so you and your end-users can find it when you need it. Broad integration across storage systems and platforms saves money and time, and cybersecurity measures like air gap and strong user authentication keep your data safe.

And one of the best things?

With the active archive's scalability, ease of use, and integration tools, you can build on to it incrementally, on your schedule. What are you waiting for?

Foreword by:

Dr. Kel Pults, DHA, MSN, RN-BC, NREMT
Chief Clinical Officer, MediQuant
Board Member, Active Archive Alliance

2

TABLE OF CONTENTS

With the active archive's scalability, ease of use, and integration tools, you can build on to it incrementally, on your schedule.

What are you waiting for?

EXECUTIVE SUMMARY

Active archiving addresses the relentless rate of data growth and how to manage it for accessibility, cost, and security.

The heart of the active archive solution is an intelligent management layer, which provides views of data stored on multiple platforms as a virtual data repository. The software monitors and may migrate archival data across cost-effective storage tiers while maintaining accessibility and cybersecurity.

The three main drivers behind active archiving's fast market growth are:

1. *Keep active archives accessible*. As active archive's storage intelligence layer automates data movement to lower storage tiers, it also monitors data locations and enables lifetime accessibility. Users and IT easily access archived data.

2. *Keep active archives low-cost*. By automatically moving archival data to lower-cost storage tiers, active archiving provides massive storage capacities in a small footprint and low cost/GB. And organizations can usually leverage existing systems and cloud subscriptions to help implement an active archive strategy.

3. *Keep active archives secure*. Active Archive solutions monitor data integrity, which protect long-term archives from bit rot, corruption, or failed media. Tape's air gap defense is impregnable to ransomware and other cyberattacks. In addition, moving disparate legacy systems to a single active archive reduces the IT footprint, thereby reducing access points and risk.

Additional advancements, such as multi-vendor active archiving systems, sustainability, analytics, and AI/ML, all contribute to active archiving's fast marketplace growth and continual innovation.

> *"Active archive solutions solve the issue of utilizing a more cost-effective storage tier, making data available and searchable, and taking advantage of the cloud and on-prem solutions in a unified platform."*
>
> — **Betsy Doughty**, VP Marketing, Spectra Logic

ACTIVE ARCHIVING:
WHAT IT IS, WHERE IT'S HEADING

Data is multiplying at a colossal rate of growth, with common forecasts coming in for 64 ZB to 175ZB and even higher. The good news is that most of that data are not stored. Only a fraction, around 16 zettabytes (ZB), will live in storage by 2025.

However, the sobering news is that this 16ZB still represents an enormous amount of data that the world needs to find a way to store in about three years' time.

Look at it this way: a single zettabyte equals 1000 exabytes, 1 exabyte equals 1000 petabytes, and 1 petabyte equals 1000 terabytes. That is a lot of data.

Massively increasing media density and storage systems are only part of the solution. Energy usage and materials costs rise along with greater density, as does expanding data center real estate.

And storing to the cloud, while an excellent option for cloud customers, has major ramifications for cloud providers. Even major cloud providers are data centers that cannot simply absorb zettabytes-worth of customer data storage—at least, not without charging extremely high ongoing costs.
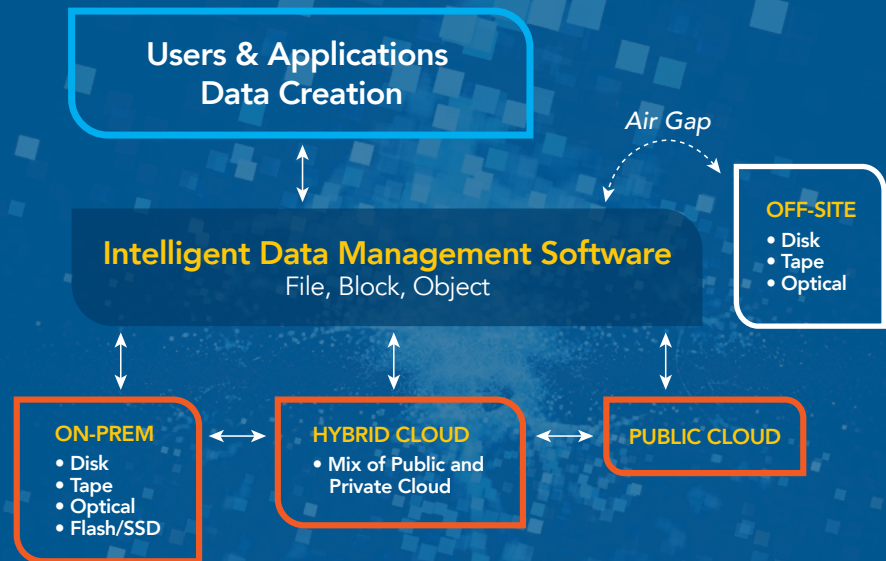
Enter Active Archiving

Active Archiving enables data owners to build intelligent archival storage systems that combine disk and tape, in the data center and in the cloud, which are cost-effective, accessible, and secure.

These smart and flexible active archiving systems are the enabling technology that enables users to cost-effectively access, manage, and derive value from massively growing archives.

THE ACTIVE ARCHIVE

Integrates Intelligent Software and Scalable Storage for the Optimal Archive Solution

Users & Applications Data Creation

Air Gap

OFF-SITE
• Disk
• Tape
• Optical

Intelligent Data Management Software
File, Block, Object

ON-PREM
• Disk
• Tape
• Optical
• Flash/SSD

HYBRID CLOUD
• Mix of Public and Private Cloud

PUBLIC CLOUD

THE BIG THREE: ACCESS, COST, SECURITY

Archival data may not require high performance computing, but it does need to stay accessible for productivity, legal, business value, analytics, and compliance. IT needs to secure that data against ransomware and other types of cyberattack. And IT needs to do all this at reasonable costs, even given extraordinary projected data growth.

"For users who need remote access to on-premises active archives, solutions that offer an object storage interface allow the archive to be securely shared by remote users and other facilities."

— **Philip Storey**, CEO, XenData

Accessibility

Accessibility and speed of access are two sides of the same coin. Accessibility means that archival data is searchable and retrievable. Speed of access is how long the search and retrieval process takes.

Active archiving's unified repository dramatically simplifies discovery and access, while enabling IT to store archives across cost-effective storage tiers. IT uses policies to direct the software intelligence to automate data movement within the transparent and easily accessible file structure.

Active archiving uses innovative technologies to protect accessibility, such as object storage interfaces for tape that simplifies tape-to-cloud movement and remote access.

Cost

Controlling costs is a given in any business, but how to control them can be a challenge. Active archiving enables IT to create efficient and cost-effective archival storage systems. IT lowers archive expenses by consolidating storage, simplifying archive management, and deploying intelligent data movement to balanced storage tiers.

> *"Retiring and consolidating legacy systems into a smaller footprint reduces risks to stored data."*
>
> — **Dr. Kel Pults**, Chief Clinical Officer, MediQuant

Active archiving enables IT to efficiently consolidate storage systems into a virtual data repository, which improves accessibility. IT can save even more money and energy costs by using active archiving to physically consolidate archived files from legacy systems to highly balanced storage tiers on disk, tape, and/or cloud.

Consolidating also lowers the risk of managing and protecting data on disparate storage systems scattered across different locations.

Simplified retention management with intelligent tiering supports lean IT teams and tight budgets by automating data movement and storage tier placement, while searchability and analytics increases the value of archival data for business intelligence and file management.

Cybersecurity

Fast-growing data stores give ransomware hackers more targets than ever. Each megabyte of data increases risk exposure and complicates cybersecurity defense. Unstructured data is especially challenging to defend because it makes up at least 80% of stored data and is easier to encrypt than structured data.

Tape systems' air gap defense remains the best way to repel hackers, because they simply cannot access the data. Air gap refers to disconnecting stored data from any network access.

Since active archiving also includes disk and cloud storage, it is critical to protect archival data on these platforms as well. IT should practice strong user authentication and 3-2-1-1-0 data defense, as well as deploying zero-trust networks.

> *"Moving data copies to offline and secure off-site locations establishes an air gap against ransomware hackers."*
>
> — **Rich Gadomski**, Head of Tape Evangelism, FUJIFILM Recording Media U.S.A. Inc.

CREATING THE ECOSYSTEM

The important thing to remember is that active archiving is not a single vendor solution. An active archiving system easily includes data center, remote locations, and the cloud; SSD, HDD, tape, and optical; block, object, and file storage.

Users can start small and leverage storage systems they already own. Active archiving enables efficient scaling by adding additional storage and platforms, such as deploying cloud to an existing disk-to-tape archive workflow. The flexible ecosystem is efficient, secure, scalable, and highly accessible.

Let's look at the main archival storage environments of tape, disk, and cloud; and how they integrate with each other in the active archiving ecosystem.

"The question with active archiving isn't cloud or not, but what data needs to be in the cloud, on-prem, or both, and when."

— **Betsy Doughty**, Vice President Marketing, Spectra Logic

Tape

Tape systems have always been the most popular storage platform for archived files. Media longevity and low cost, minimal energy usage, and resistance to cyberattacks make it a natural fit.

Active archiving builds on those benefits by taking advantage of tape's growing integration with disk and cloud, creating a policy-driven ecosystem that efficiently drives archives data to optimal storage environments.

Object storage front-ends are an enabling technology for tape's efficient integration with object-based disk and cloud storage, such as AWS S3. This simplifies replicating and tiering data from disk-based object storage to tape.

Disk

When users need fast access to archives, storing archives on economical disk-based tiers balances performance and reduces costs over primary storage. Active archiving software monitors access and keeps archived files on disk, or as they grow more inactive, moves them onto tape or the cloud.

Active archiving also uses disk caching to grant fast access and to efficiently move large amounts of archival data to tape or cloud systems. For example, a petabyte-scale disk cache for archival data enables users to access archived data at high speeds. When the active archive moves less active archive data to tape, the cache acts as a buffer that protects performance when moving large amounts of data.

"Large, long-term active archive solutions will leverage the low TCO and active access benefits of disk. New solutions will enable storage resource consolidation for even greater utilization and cost efficiency. Tape will continue to address colder storage, maintaining its long-term standing in this space."

— **Mark Pastor**, Director, Platform Product Management at Western Digital

ACTIVE ARCHIVING IN THE ENTERPRISE

"IT managers are looking for solutions addressing the needs of all storage tiers, whether NAS, Parallel File Systems, Cloud or Tape, wherever it resides. Secured active archives can use data analytics to boost archiving efficiency, lower storage TCO and improve enterprise-wide decision making."

—**Ferhat Kaddour**, VP Sales & Alliances, Atempo

TAPE AS CLOUD

"Tape as Cloud is very economical and can be used as part of a multi-cloud solution, where organizations send data to two or more Cloud Storage providers for strong data protection."

—**Dave Thomson**, Senior VP Sales, and Marketing, QStar Technologies

Optical

Optical storage has an important place in the active archive's mixed media. More advanced types of optical technologies, such as photonics, drive long-term data storage innovation even farther. Photonic disk stores extremely dense data at a low cost. It has a 100-year life span, uses minimal energy, and resists cyber-attacks.

"Active archive solutions frequently utilize multiple forms of media. We envision photonic storage media as an ideal complement to hard drives and tape media, creating an ideal active archive to satisfy any workload while remaining cost-effective and adhering to stringent cybersecurity and compliance requirements."

— **Steven Santamaria**, CEO of Folio Photonics

"Unstructured data is growing exponentially. It becomes inactive very quickly but will nevertheless need to be archived due to compliance and legal requirements."

— **Thomas Thalmann**, Managing Director, PoINT Software & Systems GmbH

Cloud

Cloud-based archiving, especially hybrid cloud in concert with tape and disk, offers cost/performance tiering choices for active archiving customers. Active archiving in the cloud also supports multiple clouds, such as storing archival data to AWS and replicating copies to Azure.

New cloud classes like "Tape as Cloud" enable IT to archive data to a remote cloud storage provider via a cloud API protocol, such as AWS S3.

Also note that in an active archive system, the cloud's cold tiers may not be archival data's final resting place. Tape is easily more economical even then the cloud, especially if the organization is facing egress fees. In this case, an active archiving system may start at production systems on SDD and fast HDD, move from there to a lower speed HDD or disk-based cache, and from there to a cloud tier to remain accessible to users and cloud-based applications. As the file continues to age and grow inactive, policies move it from the cloud to on-premises tape.

EVOLVING ACTIVE ARCHIVING

Active Archiving technology is constantly evolving as marketplace demand increases. Top innovations include Artificial Intelligence and Machine Learning (AI/ML), sustainability, analytics, and compliance.

AI/MLOps

Artificial Intelligence/Machine Learning Operations (AI/MLOps) integrates AI and ML into managing large-scale data storage, including large active archive ecosystems. AI/MLOps monitors and applies predictive analytics on system operations such as storage system utilization, archival data monitoring, and policy-driven data movement.

Active archive makers may develop and build in AI/ML into their operations, or partner with vendors specializing in archival AI/ML across multi-vendor systems.

"Organizations ingesting massive amounts of unstructured data struggle with quickly deciding which files to keep on what media. These businesses need policy- and AI-driven storage solutions that can automatically identify, separate, and store data to the appropriate storage tiers."

— **Bruce Kornfeld**, Chief Marketing & Product Officer, StorMagic

Sustainability

When it comes to sustainability, tape is particularly suited to saving energy costs.

By adding a tape system to the active archiving environment, IT can reduce energy consumption and CO2 emissions over other types of storage media. Tape can reduce energy consumption by 87% and CO2 emissions by 95% compared to spinning disk on-premises and in the cloud.

Analytics

Analytics software is critical for any organization seeking to understand, manage, and gain value from their data. In archiving, analytics drives data awareness and cost implications of moving data to optimal storage throughout its lifetime.

Analytics also plays a mission-critical role for mining business intelligence from archives. For example, more healthcare organizations want to mine archived data to improve patient safety and to improve treatments and patient outcomes.

"IT leaders want to effectively protect and manage their data throughout its lifecycle. Active archiving is a critical part of that work, since it enables IT to smoothly combine sustainable, scalable storage solutions to protect and ensure active archiving operations."

— **Tina Brown**, COO, Overland-Tandberg

Compliance and Governance

Compliance is not only about retention schedules and data protection, but also about protecting data accessibility and compliance despite disruptions like changing regulations, acquisitions, or storage system replacements.

Active archiving supports this level of compliance with its virtual repository and close monitoring and analytics. Active archiving also supports legal defensibility by acting as the system of record.

"Organizations can deploy active archiving solutions within their own data center or remote facility, which enables them to maintain data within in-house security parameters and to meet data residency requirements."

— **Tim Sherbak**, Enterprise Products and Solutions Manager, Quantum

"Interoperability between current systems and archival data is expected to promote adoption and satisfaction across stakeholder personas. Hospitals expect their migration and archive partners to advise and play a role in their overall data management and data governance strategy."

— **Marc Probst**, Chief Information Officer, ELLKAY

SUMMARY
Active archiving is:

SIMPLE
- Automated, policy-based migration efficiently frees up primary storage
- Data stays accessible wherever it is stored
- Compatible with flash, disk, tape, optical, or cloud; file, block or object storage

SCALABLE
- Simply manage massive data across multiple storage tiers
- Sceamlessly add new storage capacity or technologies
- Data resiliency over simple data protection

COST-EFFICIENT
- Reduce TCO by aligning data class with storage type and SLAs
- Optimize IT infrastructure without added cost or complexity
- Automatically balance storage tiers by frequency of access or policy

RELIABLE
- Flexible storage includes on premise, off premise, online or offline
- Storage diversity protects access and minimizes risk
- Meet regulatory compliance and governance requirements

In a world of storage-busting data growth, active archiving saves the day—and your data.

For more information on the ACTIVE ARCHIVE ALLIANCE or to learn more about an industry-specific active archive solution to meet your needs, contact:
PR@activearchive.com

in | y | f

activearchive.com

THE ACTIVE ARCHIVE ALLIANCE

No single vendor has a lock on an end-to-end active archive solution set. This is why the Active Archive Alliance's founders and members gather to champion an integrated, powerful, flexible archiving system.

The Active Archive Alliance serves as a vendor-neutral, trusted source for providing end users with technical expertise and guidance to design and implement modern active archive strategies that solve data growth challenges through intelligent data management. Active archives enable reliable, online, and cost-effective access to data throughout its life and are compatible with flash, disk, tape, optical, or cloud as well as file, block or object storage systems. They help move data to the appropriate storage tiers to minimize cost while maintaining ease of user accessibility.

Active Archive Alliance Members & Sponsors

FUJIFILM   SPECTRA

MediQuant®   IRON MOUNTAIN®   OVERLAND TANDBERG

ATEMPO   PoINT software & systems   Western Digital.

IBM   Quantum.

Folio Photonics   QSTAR TECHNOLOGIES SECURING THE FUTURE OF YOUR DATA®

Legacy Data Access   XenData   StorMagic

ELLKAY   Sullivan | Strickler