

Unlocking Benefits of Sustainable Archiving

Strategies for Cost Reduction, Sustainability,
and Long-term Data Preservation

DATA MANAGEMENT

An Atempo Paper

Table of Content

1. Introduction	3
1.01 Why is Archiving Important?	3
1.02 Overview of Archiving Types	4
A. User-driven Archiving	4
B. Automated Archiving	4
C. Integrated Application-driven Archiving	4
2. Types of Archiving: Why & When	6
2.01 User-driven Archiving	6
A. Benefits of User-driven Archiving:	6
B. Use Cases of User-driven Archiving:	7
C. Implementation Methods of User-driven Archiving:	7
D. Illustrative Example	8
2.02 Automated Archiving	8
A. Benefits of Automated Archiving:	9
B. Use Cases of Automated Archiving:	9
C. Implementation Methods of Automated Archiving:	9
D. Illustrative Example	10
2.03 Integrated Application-Driven Archiving	10
A. Benefits of Integrated Application-Driven Archiving:	10
B. Use Cases of Integrated Application-Driven Archiving:	11
C. Implementation Methods of Integrated Application-Driven Archiving:	11
D. Illustrative Example	12
3. What Types of Storage for Your Archived Data	13
3.01 Active Local Archive	13
A. Criteria for selecting an Active Local Archive	13
B. Storage Mediums for Active Local Archive	14
C. Guiding The Choice of Storage Layer Supporting the Archive	15
3.02 Cold Archives for Long Term Storage	16
A. Business Reasons for Implementing a Cold Archive	16
B. Criteria for selecting a Cold Archive for Long Term Storage	16
C. Cold Archiving and Associated Media Storage	17
D. Guiding the Choice of Storage to Implement a Cold Archive	18
4. Real Miria Archiving Use Cases	20
4.01 What is Miria Archiving	20
4.02 Use case #01: Centralized Cold Archive - Minimal Egress	22
4.03 Use case #02: Collaborative Archive for Geo-Remote Sites	23
4.04 Use case #03: Multi-Sited Local Archiving	25
5. Conclusion	26
6. Appendix	27
6.01 About Miria, the Data Management Platform	27
6.02 About Atempo	27



1. Introduction

Welcome to the world of archiving! As data continues to grow exponentially, organizations face challenges managing data effectively, ensuring compliance with regulations, and optimizing storage costs. Archiving is a critical solution that addresses these challenges by providing cost-effective and sustainable ways to preserve data for the long term.

Archiving is the practice of identifying and relocating data that is no longer frequently accessed, or deemed critical and needs to be kept unchanged. As soon as the item has been identified, it is deposited at a separate storage location for intermediate or long-term preservation. It involves various methods, including user-driven archiving, automated archiving, and integrated application-driven archiving, each with its own benefits and use cases.


In this chapter, we will explore the concept of archiving and its significance in modern data management. We will also discuss the different types of archiving and their implementation methods.

Let's dive in and unlock the modern benefits of evergreen archiving strategies!

1.01 Why is Archiving Important?

Archiving is more than just storing data in a virtual attic. It is a strategic approach to data management that offers several benefits to organizations, including:

- **Cost Reduction:** Archiving enables organizations to optimize storage costs by transferring data that is accessed less frequently to more cost-effective storage solutions, such as a fast local archive or a distant cold archive. As a result of archiving inactive data, the storage footprint of active systems is reduced, offering the possibility of cost savings. These savings are two-fold, directly reducing the need for storage space and indirectly simplifying the task and workload associated with data protection and its associated storage needs.
- **Compliance and Governance:** Archiving helps organizations meet compliance and governance requirements by retaining data for regulatory or legal purposes. Archived data can be easily retrieved and audited, ensuring organizations comply with industry regulations and legal mandates.
- **Sustainability:** Archiving serves as a sustainable solution by reducing the environmental impact of data storage. Organizations can minimize their carbon footprint and contribute to a greener IT ecosystem by implementing energy-efficient storage options like active local archive, tape archive with low power consumption and minimal physical space requirements or cloud-based collaborative centralized archive.

- 
- **Long-term Data Preservation:** Archiving ensures the long-term preservation of data, protecting it from loss or corruption. With proper metadata enrichment and data immutability options, organizations can preserve data for future use, historical analysis, or business continuity.

1.02 Overview of Archiving Types

Archiving can be implemented through different methods, depending on data access dynamics and the organization's requirements. Let's explore some of the common types of archiving:

A. User-driven Archiving

- Users can manually initiate archiving through a user-friendly web interface or scripts leveraging command-line interfaces (CLI).
- **Use cases:**
Researchers archiving data for future analysis, batch archiving of coherent file sets, or project owners archiving data sets at defined project milestones and wanting to have full control over their archiving process.
- **Benefits:**
Flexibility and user control over the archiving process.

B. Automated Archiving

- Automated Archiving can be implemented in multiple ways. It is usually performed using scheduled archiving tasks or watch folders. A watch folder is a specific folder designated for automatic scanning by an archiving solution.
- **Use cases:**
Capturing metadata from invoicing and finance applications, archiving data produced by IoT devices, or automatically archiving data transferred from digital storage.
- **Benefits:**
High efficiency and simple integration with existing workflows.

C. Integrated Application-driven Archiving

- Through REST APIs, archiving can be seamlessly integrated into business applications.
- **Use cases:**
Data can be captured directly from business processes by automating the archiving of all financial transaction data without any human intervention, or by integrating archiving into specialized healthcare applications for long-term storage of patient information.

- **Benefits:**

Streamlined processes and the ability to enrich archived data with business metadata with minimal effort.

In the next chapter, we will delve into each type of archiving in more detail, explore the benefits, use cases, and implementation methods.



2. Types of Archiving: Why & When

Archiving is a critical part of data management, but not all archiving strategies are created equal. In this chapter, we'll explore the different types of archiving and their benefits, so you can choose the right strategy for your organization.

From user-driven archiving to automated archiving and integrated application-driven archiving, we'll examine the advantages of each approach and when they're most appropriate. Whether you're looking for a streamlined process, cost-effectiveness, or compliance, we'll help you find the archiving solution that fits your needs.

2.01 User-driven Archiving

User-driven archiving is a type of archiving that puts the power in the hands of the users, allowing them to initiate the archiving process when suited to their work through a user-friendly web interface or scripts using command-line interface (CLI). This type of archiving is ideal for users who need to quickly and easily archive files and data, and who want control over the archiving process.

This approach to archiving offers benefits to both the IT team and business users :

- By ensuring **IT retains control over data storage** while providing users with control over their data archiving, this approach reduces the workload on IT resources and frees up time for them to focus on new projects. Best-in-class archiving solutions enable storage administrators to pre-define archiving policies tailored to the specific needs and constraints of each user project, ready to use by the users.
- This eliminates the need for end-users to rely on IT resources and reduces bottlenecks in IT services, increasing efficiency and productivity.

For example, user-driven archiving can be implemented for compliance long-term archiving that requires users to add mandatory descriptive metadata with each archiving and automatically creates three copies of the data. Two copies are made on WORM tape media, with one media being stored externally in a remote safe, while the second copy is stored on a local object storage platform using immutable data storage.

A. Benefits of User-driven Archiving:

- **Flexibility:** Users have control over the archiving process and can choose which data to archive and when to archive it, providing flexibility in managing data based on their needs and coherent project cycles.




- **Customization:** Archive administrator can predefine archiving settings to meet specific requirements of each business team, such as metadata enrichment to include the description that make sense for each project and make them mandatory when required, retention periods, and storage locations like one copy to disk, a second to tape, a third to cold cloud archive.
- **Improved metadata and data usability:** User-driven archiving empowers users to trigger the archiving process and select coherent sets of data, resulting in more precise and meaningful metadata descriptions. This, in turn, greatly improves the capacity for reusing and retrieving data over time, as the data sets are better understood and can be more easily located and utilized.
- Additional benefits for the IT team:
 - Reduces bottleneck on IT resources
 - Reduces operational costs of workflows
 - Frees up time for IT to work on other projects

B. Use Cases of User-driven Archiving:

- **Biomedical Research Archiving:** Store genomic data, research studies, and clinical trials data in an archive to facilitate collaboration, data sharing, and data reuse for future analysis or reference across research teams.
- **Manufacturing Project Archiving:** Selecting all the files related to a completed stage of a project, including contracts, agreements, CAD drawings, bill of materials, and reports, in an archived folder for future reference, and being able to add project related descriptive metadata such as:
- **Energy - Nuclear:** Store sensitive data from nuclear facilities including sensor metrics, logs, quality control metrics, maintenance and repair reports, and more, in a centralized cold archive to ensure regulatory compliance, long-term risk management, and data security.
- **Include your business reason:** For performing Batch Archiving over large sets of coherent files or data, such as project files or historical data, in a structured and organized manner.
- And many more.

C. Implementation Methods of User-driven Archiving:

- **User-friendly Web Interface:** Users can initiate archiving through a web-based interface that provides simple drag-and-drop option for selecting and archive data.

- 
- **Command-line Interface (CLI):** Power users knowledgeable in the building of scripts can utilize CLIs to initiate archiving commands, providing more advanced options for customization and integrating the archiving in the business workflow.

D. Illustrative Example

You are a researcher working on a groundbreaking study. You have a vast amount of data collected over the years, and you want to ensure its long-term preservation for future analysis. With user-driven archiving, you can easily select the data you want to archive, enrich it with relevant metadata, and initiate the archiving process through a simple drag and drop of your data to Miria Archiving's web interface.

The archived data is securely stored in a designated location, as determined with the IT storage team for the project, from among various options.

The archiving solution also provides ongoing monitoring of the archived data, ensuring that the predetermined number of copies defined for the project is consistently maintained.

In conclusion, user-driven archiving offers businesses the flexibility and control to manage their data archives to meet their specific needs. Archiving policies can be pre-defined to ensure compliance, while providing users with the ability to add rich business metadata, improving data usability over time.

This archiving approach offers benefits across all industries. Implementation can be achieved through a user-friendly web interface or CLI, and provides ongoing monitoring to ensure data is safely and securely stored. Overall, user-driven archiving empowers businesses to manage their data archives in a way that is tailored to their unique requirements, enabling them to make better use of their data and extract greater value from it.

2.02 Automated Archiving

Automated Archiving is a type of archiving that offers businesses a streamlined and efficient way to collect, organize, and archive their files as soon as they are created. It consists of a software solution that monitors predefined folders for new files or changes to existing files, automatically archiving them to a designated location without any user intervention. This approach not only saves time and effort but also ensures data is consistently and accurately captured, making it easier to manage and retrieve in the future.

In this chapter, we will explore the benefits and use cases of Automated Archiving, and how it can be implemented in various industries to improve data management and enhance overall business operations.



A. Benefits of Automated Archiving:

- **Improved efficiency and near to real-time archiving:** Automated Archiving simplifies the archiving process by collecting and organizing files automatically upon creation, saving time and effort. This approach optimizes storage and reduces storage TCO.
- **Better data governance:** By consistently and accurately capturing data in real-time, businesses can better manage and govern their data, ensuring compliance with regulations and standards.
- **Enhanced security:** Automated Archiving ensures that data is securely stored and protected from unauthorized access or modification.
- **Increased productivity:** With the archiving process automated, employees can focus on more important tasks and projects, boosting overall productivity.
- **Accuracy:** Files are automatically archived based on pre-defined rules, reducing the risk of human error in file organization.

B. Use Cases of Automated Archiving:

- **Autonomous Vehicle Simulation:** Archive simulation and computational data from autonomous vehicle testing and development to enhance algorithm development and improve overall vehicle safety.
- **Media Management:** Automatically archive large video or audio files as they are created or added to the work storage, ensuring they are preserved for future use or monetization.
- **Industrial Manufacturing:** Archive sensor data from production lines in real-time, enabling businesses to analyze the data for quality control, performance optimization, and predictive maintenance.
- **Oil and Gas Production:** Collect and archive data from remote sensors monitoring oil rigs or pipelines, providing real-time insights into operational efficiency, safety, and environmental compliance.
- **Transportation:** Archive video surveillance footage from tube transport systems, ensuring data is captured and stored for future investigations or incident response.
- And many more.

C. Implementation Methods of Automated Archiving:

- **File System Watchers:** Use built-in file system watchers or third-party tools that monitor designated folders or complete filesystems to trigger archiving actions based on predefined rules.
- **Best in class archive solutions** comes with built-in capability to monitor folders and automatically triggering archiving actions as part of a larger workflow or business process.



D. Illustrative Example

You are a finance manager working for a large retail company, responsible for generating invoices for thousands of customers every month. With an automated archiving solution, you can designate a folder for your generated invoices.

As soon as a new invoice is created, the solution automatically generates a PDF version and archives it in a designated location, with metadata such as the customer contract number and the invoice number.

This ensures that all invoices are consistently and accurately archived, making it easier to manage and retrieve them for future reference or audits.

To conclude on this section describing Automated Archiving, it is a powerful tool that offers businesses a streamlined, efficient and yet simple way to collect, organize, and archive their files as soon as they are created. In this chapter, we explored the benefits and use cases of Automated Archiving, and how it can be implemented in various industries to improve data management and enhance overall business operations.

The automated archiving process not only saves time and effort but also ensures data is consistently and accurately captured, making it easier to manage and retrieve in the future. By leveraging the right software solutions and implementation methods, businesses can improve efficiency, enhance data governance, and boost overall productivity.


2.03 Integrated Application-Driven Archiving

Integrated Application-Driven Archiving is a type of archiving that integrates with the applications used by businesses to create and manage their data. It consists of a software solution that automatically archives data generated by applications, enabling businesses to streamline their data management processes and ensure data is consistently and accurately captured for its preservation.

In this chapter, we will explore the benefits and use cases of Integrated Application-Driven Archiving, and how it can be implemented in various industries to enhance overall business operations.

A. Benefits of Integrated Application-Driven Archiving:

- **Simplified Workflow & Improved efficiency:** Integrated application-driven archiving automates the archiving process, saving time and effort. Streamline your file management process by integrating archiving directly into your applications, eliminating the need for manual operations.
- **Enhanced data security:** The archiving process can be integrated with security protocols and access controls, ensuring data is protected from unauthorized access or modification.


- 
- **Compliance:** Integrated application-driven archiving can be configured to comply with regulations and standards, ensuring compliance with data retention and privacy laws.
 - **Improved data accessibility:** This method of archiving streamlines the process of adding large volumes of descriptive technical or business metadata to archived data, making it easier to locate and retrieve data. This not only improves data management but also enhances employee productivity by enabling quick and effective access to necessary information through a variety of relevant search criteria.
 - **Consistency:** Ensure that all relevant data from different applications is consistently archived according to predefined rules, maintaining data integrity and compliance.

B. Use Cases of Integrated Application-Driven Archiving:

- **Finance Compliance:** Archive financial records, reports, and transactions in accordance with regulatory compliance standards, ensuring financial data is accessible for future audits.
- **Healthcare:** Archiving medical imaging data, including MRIs, CT scans, and X-rays, enables easy retrieval and comparison of images over time. The goal is to move the high volume of data to a less expensive storage tier while ensuring the appropriate level of data security.
- **CRM Archiving:** Automatically archive customer data from your CRM (Customer Relationship Management) system, based on customer status, activity, or other criteria, for organized customer data management and historical reference.
- **Media & Entertainment:** Media asset management archiving automatically stores digital content (e.g. video, audio, images) upon creation or at a specific project phase. This offloads data to less expensive storage tiers while ensuring future retrieval and distribution.
- **Nuclear Energy:** Archiving sensor data from nuclear facilities, including radiation levels, temperature, and pressure readings, to ensure compliance with regulatory requirements and provide historical data for future analysis and problem-solving.

C. Implementation Methods of Integrated Application-Driven Archiving:

- **Application Plugins:** Utilize application-specific plugins or extensions to integrate archiving functionalities directly within the application's interface, streamlining the process and enabling automation. These plugins can be developed by leveraging the API (Application Programming Interfaces) of the archiving solution to seamlessly connect with the application. The use of plugins is common among applications that offer multiple integrations and have consolidated their integration process using this type of layer.

- 
- **Direct API Integration:** Utilize the APIs provided by the archive applications to develop custom archiving integrated with the application's data and functionalities in the core of its functions.

D. Illustrative Example

You are a post-production project manager in charge of adjusting the colors of a video sequence for a feature film. With integrated application-driven archiving, you can use a plugin that allows you to automatically archive your edited video sequence to a designated location within your media asset management system. This offloads the high volume data from your storage system, making it easier to switch to your next project and ensuring easy retrieval and distribution of the final video sequence in the future.

In conclusion, Integrated Application-Driven Archiving provides businesses with an efficient and streamlined method of archiving data generated by applications. This type of archiving can be configured to comply with regulations and standards, ensuring data is consistently and accurately captured for its preservation. With benefits such as improved efficiency, enhanced data security, compliance, and improved data accessibility, Integrated Application-Driven Archiving can be implemented in various industries to enhance overall business operations.

By utilizing application plugins or extensions, or direct API integration, businesses can seamlessly integrate archiving functionalities into their applications, eliminating the need for manual operations. Overall, Integrated Application-Driven Archiving provides businesses with an effective solution for improving data management processes and increasing productivity.



3. What Types of Storage for Your Archived Data

Archiving your data is not just about storing it, but also ensuring that it is easily accessible and available when needed. In this chapter delves into the different types of storage options available for archived data. The storage of archived data is crucial as it determines the accessibility, durability, and overall preservation of the data. We will explore three different types of storage:

- Active Local Archive,
- And two forms of Cold Archives.

Each type of storage has its own advantages and disadvantages, and it is essential to understand which type of storage is best suited for a particular data archiving strategy.

3.01 Active Local Archive

The Active Local Archive is a proximity storage solution that enables quick and efficient access to frequently accessed data. It is ideal for data that is still in use and requires high-performance access, while also ensuring secure long-term retention.

In this chapter, we will discuss the reasons for selecting an Active Local Archive and explore the various technologies that can be leveraged for optimal results.


A. Criteria for selecting an Active Local Archive

- **Quick Access:** This is probably the most important criteria for businesses or applications that requires immediate access to their archived data, such as financial institutions or healthcare providers that need to quickly retrieve patient records or transactional data.

Archived data is stored in a storage medium that allows for fast retrieval and access, reducing the time and effort required to locate and retrieve archived data.

- **Scalability:** As the volume of data grows, the storage solution must be able to accommodate the increased capacity.

A scalable local archive accommodate data growth and prevent storage space issues, which is crucial for businesses that expect rapid data growth or need to keep archives for extended periods.


- 
- **Data Security:** Ensuring the security and integrity of archived data is critical, particularly for organizations dealing with sensitive or confidential data. To ensure the security and integrity of archived data, especially for sensitive or confidential data, a local archive should have strong security features that prevent unauthorized access, modification, or deletion. These features may include encryption, access controls, audit trails, and other security measures that ensure data protection throughout its lifecycle, from archiving to retrieval and deletion.

B. Storage Mediums for Active Local Archive

Active local archive is a cutting-edge approach to data storage that keeps your archived data readily accessible while ensuring its security and preservation.

- **Nearline storages:** Two types of storage technologies are commonly used for active local archives as described below.
 - **NAS-like storage** utilizes Solid State Drives (SSDs) and provides traditional file system access to the archived data.
 - **Nearline object storage** is based on SSDs, and offers fast access to the archived data, making it suitable for scenarios where quick retrieval is important. Object storage is a data storage technology that stores data in a non-hierarchical structure without a traditional file system. It provides faster access to archived data and allows for additional capacity without the limitations of file systems, resulting in time saved during data access.
- **Tape library:** Tape libraries are a popular choice for archive storage due to several reasons: high capacity and low cost per gigabyte, longer shelf life, scalability, low energy consumption, security, and reliability. Another advantage of tape libraries is the capability to externalize data and have a disaster recovery plan in place for the archive content, which is particularly useful for businesses having a strong requirement for long term preservation of their data. Tape library is also a cost-effective option for long-term data retention, as the cost per terabyte is often lower than other storage options.
- **Cloud Storage:** While cloud storage may not seem like the most obvious choice for a "local" archive, it is often used as a fast and cost-effective option for collaborative nearline archive storage. With the installation of a specific fast connection to the cloud, called ExpressRoute with Azure or Direct Connect with AWS, cloud storage providers can offer quick access to data sets. This makes it possible for businesses to use it as a collaborative central archive accessed between several locations.

Additionally, cloud storage is scalable, reducing the need for upfront investments in hardware. Cloud storage providers also offer robust data security measures, ensuring that archived data is protected from unauthorized access or loss.



However, businesses should exercise caution regarding the cost of data egress, which can be expensive. Additionally, ExpressRoute is not a cheap option.

C. Guiding The Choice of Storage Layer Supporting the Archive

When choosing the type of storage, it's important to consider the specific needs and use cases of the organization and identify the key criteria guiding the choice.

Storage Type Criteria	Nearline NAS	Nearline Object Storage	Tape Library	Cloud Storage
Quick Access	4	5	2	3
Scalability	2	3	4	5
Security	4	5	5	4
Cost Effectiveness	3	4	5	3
Sustainability	2	3	5	4

Note: The scale used for the scoring ranges from 1 to 5, with 5 being the highest score.

For example, if quick access to large volumes of data is a top priority, nearline object storage would be the best choice. On the other hand, if scalability is a major concern and the business allows the time to perform a data retrieve from tape, a tape library might be the best solution.

Finally, if storage sustainability and scalability are the main decision factor, tape library would be the best choice, followed closely by Cloud storage.

In conclusion, the Active Local Archive is an excellent storage solution that provides quick and efficient access to frequently accessed data. Scalability and data security are key criteria to consider when selecting an Active Local Archive. Nearline NAS and nearline object storage are popular options due to their speed and efficiency. Tape libraries are a cost-effective and reliable option for long-term data retention, while cloud storage is a counter-intuitive yet fast option for collaborative nearline archive storage. However, businesses should exercise caution regarding the cost of data egress and the expensive ExpressRoute. When choosing the type of storage, it's essential to consider the specific needs and use cases of the organization and identify the key criteria guiding the choice, and our above scoring table can be a useful tool in making an informed decision.



3.02 Cold Archives for Long Term Storage

Data storage is a critical aspect of modern business operations, and as data volumes continue to grow at an unprecedented rate, organizations are constantly seeking innovative ways to store and manage their data efficiently.

One such approach is cold archiving, which provides a cost-effective storage solution for long term archived data that is infrequently accessed. Cold archiving comes in two types: the **Cold Local Archive** and the **Deep Cold Archive**.

In this chapter, we will delve into these two types of cold storage and explore their benefits, features, and criteria for selecting the most suitable cold storage platform to support a cold archive.

A. Business Reasons for Implementing a Cold Archive

Before delving into the technical aspects of cold archives, it's important to understand the business drivers leading to the creation of a cold long-term archive. This solution has become popular for secure and durable preservation of data over extended periods of time.

- **Regulatory compliance requirements:** Many industries are required to retain data for extended periods to comply with legal or regulatory obligations.
- **Historical record-keeping:** Organizations that need to preserve historical records or information for future reference may require a cold long-term archive.
- **The need for future analysis and the search for cost efficiency:** Archived data can provide valuable insights for future analysis, making a cold long-term archive a valuable resource for data-driven businesses. Compared to other storage options, cold long-term archives are typically more cost-effective, enabling businesses to save on storage costs.
- **Intellectual property preservation:** Businesses may need to preserve intellectual property or proprietary information for extended periods of time, making a cold long-term archive a viable option.

B. Criteria for selecting a Cold Archive for Long Term Storage

When it comes to long-term storage of data, businesses need to carefully consider the criteria for selecting a cold archive solution. In this section, we will discuss the two key criteria for selecting a cold archive: scalability and cost effectiveness, as well as data preservation.



- **Scalability and Cost Effectiveness:** As the volume of data grows, the storage solution must be able to accommodate the increased capacity at a cost that is adequate to the volume of data stored and the security level associated.

A scalable long term archive ensures that businesses can continue to store and manage their growing data archives without having to worry about running out of storage space or having to switch to a new storage solution.

This is particularly important for businesses that anticipate rapid data growth or need to keep archives for extended periods of time.

- **Data Preservation:** The preservation of data is crucial for cold archive storage, especially for organizations dealing with sensitive or confidential data.

A scalable long term archive must have robust security features to protect archived data from unauthorized access, modification, or deletion. This can include encryption, access controls, audit trails, and other security mechanisms to ensure that data is protected throughout its lifecycle, from archiving to retrieval and deletion. The long-term preservation of data is essential for businesses that require compliance with regulations or preservation of historical records.

C. Cold Archiving and Associated Media Storage

The two types of cold archive to consider for long term data preservation are **Cold Local Archive** and **Deep Cold Archive**.

- **Cold Local Archive** provides a local and cost-effective storage option for long-term archived data, often leveraging tape storage for its durability and affordability.


Tape libraries are the ideal choice for implementing a Cold Local Archive due to their high capacity and low cost per gigabyte, scalability, energy efficiency, security, and reliability.

Another strong argument for using tape libraries is the capability to externalize data and have a very accessible disaster recovery plan in place for the archive content, which is particularly useful for businesses having a strong requirement for long term preservation of their data.

- **The Deep Cold Archive** is a remote and cost-effective storage option for archived data. This type of archive, typically cloud-based, is designed for long-term data retention at a low cost. It's ideal for data that is rarely accessed, but still needs to be stored for regulatory or compliance purposes.

Examples of "deep cold" archive solutions include Amazon Glacier and Microsoft Azure Archive Blob Storage.

These solutions offer very low-cost storage for data that is accessed less than once per year, making them suitable for archiving large amounts of data like files, historical records, and other types of data that do not require frequent access but still need to be stored for extended periods.



New Gen2 and Gen3 cloud storage offers are also available that offer different approach: Wasabi for instance does not price differently cold tiers, Storj offers a distributed almost collaborative approach to cloud storage. It is recommended to take the time to study the offer available.

D. Guiding the Choice of Storage to Implement a Cold Archive

When choosing the type of storage to support your cold archive, it's important to consider the specific needs and requirements of your business use case in order to identify the key criteria guiding the choice:


Storage Type

Storage Type		
Criteria	Tape Library	Cloud Storage
Data Sovereignty	4	3
Data Durability	4-5	4
Data Security	4	4
Cost-effectiveness	5	3
Cost predictability	5	4
Risk of vendor lock-in*	3	2

Note: The scale used for the scoring ranges from 1 to 5, with 5 being the highest score.

Detailed description of criteria in the above table:

- **Data Sovereignty:** The degree to which the storage solution guarantees control and ownership over the data to the user, including compliance with data regulations and privacy laws.
- **Data Durability:** The storage solution's capacity to guarantee the long-term retention and preservation of data is essential, including its resistance to hardware failure, data corruption, and natural disasters. It's worth noting that while tape media has good durability, it may require some management to migrate data from one LTO generation to the next. However, modern solutions like Miria Archiving can manage this process seamlessly.
- **Data Security:** The degree to which the storage solution protects the data from unauthorized access, modification, or deletion, including encryption, access controls, and other security mechanisms.
- **Cost predictability:** The level of certainty regarding the cost of storage over time, including transparency of pricing and potential hidden fees or charges.

- 
- **Vendor lock-in:** The degree to which the storage solution is proprietary or interoperable with other systems, including the ability to migrate data to another solution without major disruptions or costs.

It should be noted that a modern archiving solution like Miria Archiving mitigates this risk by managing data and integrating with various storage solutions, ensuring that data can be easily moved when and where needed.

In conclusion, cold archiving provides a cost-effective and efficient solution for long-term storage of infrequently accessed data. It's important to carefully consider the business reasons for implementing a cold archive, such as compliance obligations, preservation of historical records, and cost efficiency, as well as the key criteria for selecting the appropriate storage solution, including scalability, data preservation, cost-effectiveness, and vendor lock-in.

Tape libraries and cloud storage are both viable options, and businesses should choose based on their specific needs and requirements.

Finally, modern archiving solutions like Miria Archiving can help mitigate the risk of vendor lock-in by managing data and integrating with various storage solutions.



4. Real Miria Archiving Use Cases

4.01 What is Miria Archiving

Miria Archiving software is a comprehensive solution that offers organizations a reliable and cost-effective way to manage their long-term data storage needs. It is designed to help organizations overcome the challenges of managing growing volumes of data while addressing compliance requirements and minimizing legal and financial risks.

The main advantages of Miria Archiving software are:

- **Data-centric** - It allows organizations to enrich their archives with custom metadata, making it easier to search and retrieve data and support data governance.
- **Storage versatility** - It can be deployed on any storage medium, including tape, cloud, or disk-based solutions, depending on the organization's specific needs and constraints. This makes it an ideal solution for organizations looking to implement a hybrid archiving strategy that leverages multiple storage technologies.
- **Storage agnostic & agility** – Miria provides an adaptable archiving strategy that allows organizations to make changes over time. The archived data and metadata are not tied to a specific storage solution, providing flexibility to choose the most suitable storage for the organization's needs. Miria software enables an agile archive, allowing organizations to easily relocate the archive as needed.
- **Scalability** - Miria Archiving easily manages multi-petabyte scale archives and allows organizations to easily expand their storage capacity as their data storage needs grow. This scalability makes it a cost-effective solution for organizations looking to manage their long-term data storage needs without incurring significant costs.
- **Security** - It offers several security features, including air gap cold archiving and encryption, to help organizations protect their archived data against cyber-attacks and other security threats. These security features are particularly important in a post-COVID world where the risk of ransomware attacks and other cyber-attacks has increased significantly.

Overall, Miria Archiving software is a comprehensive solution that helps organizations meet their long-term data storage needs by providing reliable, scalable, and secure archiving capabilities.

By leveraging this software, organizations can ensure the long-term availability and security of their critical data while also minimizing the costs and risks associated with long-term data storage.



Expected Capabilities Summary	Miria Archiving
User-driven Archiving	✓
Automated Archiving	✓
Integrated application-driven Archiving	✓
Active Local Archive (Active Archive)	✓
Cold Archive for Long Term Storage	✓
Tape Library	✓
Cloud Storage	✓
Object Storage	✓
Scale-Out NAS / Parallel Filesystems, etc.	✓
Metadata enrichment	✓

In summary, Miria Archiving software is a comprehensive solution that enables organizations to meet their long-term data storage needs by offering reliable, scalable, and secure archiving capabilities. By utilizing this software, organizations can ensure the long-term availability and security of their critical data while minimizing costs and risks associated with long-term data storage.

Miria Archiving use cases span a wide spectrum, catering to diverse storage and retrieval requirements:

- **Data preservation for legal or insurance purposes** with zero data restoration (and immutable data protected against modifications). This is often in the form of a **central archive consolidating data** from multiple geographically distributed sites.
- **Active and collaborative archives serving multiple sites**, where data is enriched and shared, with frequent data restoration.
- A scenario that falls between the two extremes, with project-based or site-based archives featuring variable data restoration.

In the following sections, we will cover three specific use cases, showcasing the versatility and effectiveness of Miria Archiving solution across different industries and requirements.

4.02 Use case #01: Centralized Cold Archive- Minimal Egress

Use case:

An energy company with locations on both the east and west coasts of the US, with each site having 10+ users.

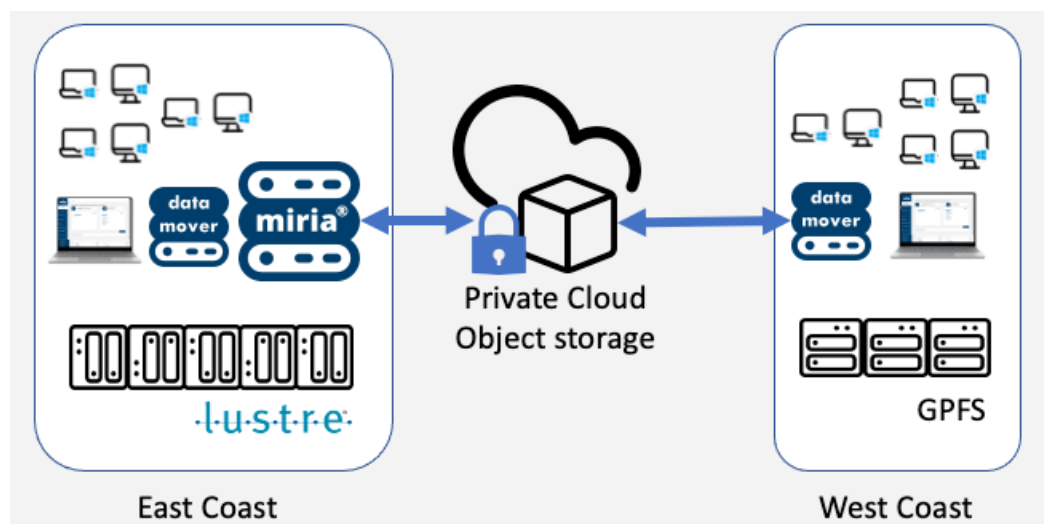
- HPC research and simulation centers at each location generate enormous amounts of data.
- The company company uses Lustre on one data center and GPFS on the other one.
- The archive currently stands at 1PB and is growing rapidly.
- Users have very few need to access data archived by the other team.

Requirements & Constraints:

The solution must:

- Support data restore by enabling cross-restore capability across different storage platforms.
- Accommodate both on-demand archiving and automated archiving to provide flexibility in archiving data.
- Ensure data security by storing sensitive data on a private cloud storage instead of public cloud storage.
- Meet legal requirements by preserving data for a minimum of 40 years.

Solution implemented using Miria:





The solution is:

- Enabling users to search, retrieve data when needed using the web interface of Miria Archiving,
- Providing automated scheduled archiving of daily work,
- Integrating with Business Applications via API for seamless archiving,
- Using object storage platforms for private cloud storage and fixed egress costs.

Benefits:

- Predictable cost
- Compliance
- Agility

4.03 Use case #02: Collaborative Archive for Geo-Remote Sites

Use case:

A Media & Entertainment post-production company,

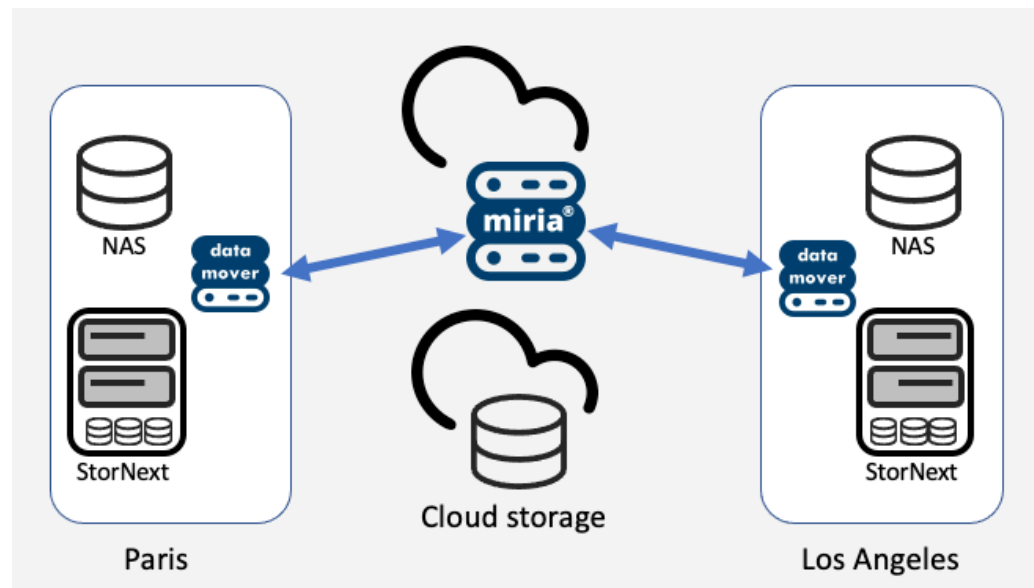
- Two sites, one located in France, and the other one on the West Coast in the US
- Each team has a local NAS storage (Qumulo or DDN depending on the site), and a Quantum StorNext shared storage for editing stations
- Users need to be able to search the data archived by the other site to be able to retrieve and use it locally

Requirements & Constraints:

The solution must:

- Have an archive with a central storage repository that can be accessible from both sites
- Support the archiving of daily jobs made of hundreds of TB of data
- Integrate with the Media Asset Management (MAM) solution used by the company

Solution implemented using Miria:



The solution is:

A central archive using a public cloud storage as the archive repository

- Enabling users to archive and retrieve data directly from their MAM.
- Offers the users the capability to use Miria's Web interface to search and preview media assets archived if needed.
- Miria also provides additional services to the IT team by doing data protection for the NAS and the StorNext servers, using Miria Backup module.

Benefits:

- Federated archive that simplifies remote team collaboration
 - Single and central archive server
 - Controlled storage costs
 - Easy integration with key business application (MAM)
 - Archive & Backup from a single solution
- Use case #03: Multi-Sited Local Archiving

4.04 Use case #03: Multi-Sited Local Archiving

Use case:

A manufacturing company operates multiple factories globally.

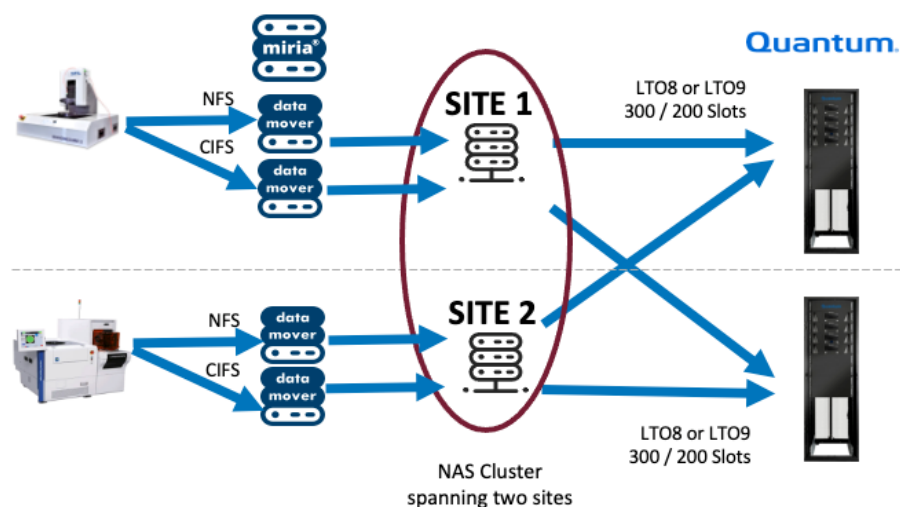
- Each factory has two production lines, each equipped with local NAS storage, and connected to the same network.
- Over 60 metrology stations exist on each production line, continuously generating data stored on the local NAS storage.
- The data collected is utilized for post-production quality control purposes.
- Each site generates 2 to 3PB of data.

Requirements & Constraints:

The solution must:

- Enable fast access to the data archived at production time for sales, mitigating hot line calls from customers and partners.
- Accommodate both on-demand data retrieval and automated archiving.
- Handle high volumes of data efficiently.
- Offer a fallback solution in case of incidents on the primary archive.
- Preserve data for a minimum of 40 years as per legal requirements.
- Support small files with an average size of 100KB.
- Handle lots of daily data retrieves.
- Ensure compliance and data security for competitive reasons.

Solution implemented using Miria:





The solution is:

- Implemented locally by each storage unit
- Automatic archiving that collects data produced by the metrology stations on each local NAS and performs an archiving job to two targets:
 - One on a local tape library
 - A second one on the tape library located in the second line of production via the same network
- Sales are able to access archived data using Miria Web UI, by creating batches of data recall based on customers PO., directly from the local tape library on each site.

Benefits:

- Predictable cost
- Very flexible search using business metadata,
- High capacity
- Built-in fallback implemented with dual archiving target
- Fast archiving of batch of small files

5. Conclusion

In conclusion, effective archiving strategies are essential for organizations to manage data in a dynamic and sustainable manner.

By implementing various archiving types and storage options, such as user-driven archiving, automated archiving, integrated application-driven archiving, active local archive, and cold archive for long-term storage, organizations can ensure that their data remains relevant, accessible, and valuable for the long term. Additionally, enriching archived data with metadata adds further value.

It is beneficial to understand that successful implementation of sustainable archiving strategies require a dynamic and adaptive approach that ensure the data remains independent of any specific storage platform. This way, the archive can adapt over time and remain relevant and usable, even if the situation or business needs change.

We hope this guide has provided valuable insights and practical tips on how to effectively manage and preserve data for the long-term.

6. Appendix

6.01 About Miria, the Data Management Platform

Miria Archiving service described above is one out of the five complementary services offered by Miria Data Management Platform:



- **Backup** for the protection of organizations files and unstructured storage against disaster.
- **Migration** for relocating files/folders with their associated user rights to another storage on premise or in the cloud.
- **Mobility** for empowering end users with secured and high-performance data movement or replicating/ synchronizing remote storage of different technology or storage.
- **Analytics** for scanning the file storages of an organization and streamlining file placement and reducing storage costs.

6.02 About Atempo

Atempo is an independent software company based in Europe with an established global presence, providing solutions to protect, store, move and retrieve all critical data for thousands of businesses worldwide. With over 30 years of experience in data protection, Atempo offers a complete range of proven solutions for backing up physical and virtual servers, workstations and migrating very large volumes of data between different storage systems.

Atempo has been recognized by the top global research firms, including Gartner as Representative Vendor in the Gartner Market Guide for Hybrid Cloud Storage 2022 (three consecutive years), by GigaOM in its 2022 Radar for Data Migration as a Challenger and Fast Mover moving into Leadership, and by Coldago Research as one of TOP 3 key players in the market in its 2022 Map for Unstructured Data Management.

Atempo's three flagship solutions, Lina, Miria and Tina, are labelled "Used by the French Armed Forces" and "France Cybersecurity". Selected as part of the French Tech 120 government program to create 25 unicorns by 2025, the company, headquartered in Paris, has a powerful network of value-added wholesalers, resellers, manufacturers, integrators and managed service providers.

For more information: www.atempo.com



Atempo Headquarters | Immeuble Iliade
23 Avenue Carnot, 91300 Massy - France
Tel: +33 164 868 300 | info@atempo.com

POWERFUL DATA PROTECTION AND DATA MANAGEMENT SOLUTIONS

atempo.com