



“Ensuring that the right information is securely stored and retained in accordance with internal policies and external statutes has become one of the most pressing issues for today’s IT departments, especially in light of the phenomenal growth of digital archiving activities.”

# Assureon - Delivering More Secure, Cost-Effective, Windows-Based Storage

## Executive Summary

*The critical nature of digital information in today’s business environment requires organizations to carefully evaluate their alternatives for storing and archiving data. Information such as e-mail, legal records, intellectual property, and healthcare data often must be stored and protected for years—yet has to be available when you need it most.*

*This paper provides an overview of the factors that organizations need to consider when evaluating storing such valuable data. It also provides a discussion of Assureon, a solution from Nexsan Technologies whose architecture simplifies and addresses the many complex requirements for cost-effective, scalable, and secure data storage.*

Included in this paper:

### THE NEED FOR SECURE STORAGE

- Business Challenges
- Key Secure Storage Scenarios
- Storage Planning

### ASSUREON: MEETING THE NEED FOR MORE SECURE STORAGE

### ASSUREON: IMPROVED TOTAL COST OF OWNERSHIP

### CONCLUSION

## The Need for Secure Storage

The importance of digitized information for today’s enterprise cannot be overstated. E-mail archive, Office documents, Portable Document Format (PDF) files, digital images, scanned document or check images, financial reports (ERM/COLD data), ECM data, medical PACS images and scanned patient records, audio and video files—these and other kinds of electronic data are the lifeblood of today’s organizations.

As such, managing, securing, and effectively scrubbing vast amounts of information is a critical part of IT operations. Ensuring that the right information is securely stored and retained in accordance with internal policies and external statutes has become one of the most pressing issues for today’s IT departments, especially in light of the phenomenal growth of digital archiving activities. As the Enterprise Strategy Group notes, “Total worldwide digital archive capacity in the commercial and government sectors is expected to grow from 3,000 petabytes (PB) in 2005 to more than 27,000 PB by 2010.”

To manage this growth in secure data storage needs, organizations need to develop efficient, consolidated file storage infrastructures that can help to reduce overall costs while improving the processes of protecting and sharing information.

## Business Challenges

The manner in which vital business information is managed and stored can have both short-term and long-term consequences. When evaluating storage needs, every

“The ability to gather, organize, and categorize millions or billions of files that must be manageable and usable over long periods. In some industries, this means storing and managing files for 6–10 years or longer.”

“Businesses of all sizes face many challenges in managing their growing repositories of e-mail. These include compliance concerns, storage management limits, the ability to find archived e-mail when needed...”

organization will have its own specific requirements and challenges based on the types of information that the business generates. A financial firm may work exclusively with contracts and loans, a law firm with intellectual property such as technical manuals and litigation documents, and a health organization with a range of images and patient records such as radiology, medical histories, and billing data. Without a doubt, almost every organization from small to large has a need to manage an ever increasing volume of e-mail.

Although specific details will vary by company, most businesses will face some common challenges when looking at their storage options. These include:

- The ability to gather, organize, and categorize millions or billions of files that must be manageable and usable over long periods. In some industries, this means storing and managing files for 6–10 years or longer.
- The ability to meet and prove compliance to internal corporate requirements, including mandates from legal counsel, as well as external regulations such as the Sarbanes-Oxley Act (SOX), Securities and Exchange Commission’s SEC 17a and the Health Insurance Portability and Accountability Act (HIPAA).
- Immutability, business continuity -- long-term security and protection of corporate data, including preventing electronic theft; protecting against data loss due to natural disaster such as fire or earthquake; guarding against Internet-based threats such as viruses and worms; protecting the integrity and accuracy of data; protecting the confidentiality of customer information; and taking preventative measures against hardware failure.
- Implementing efficient “e-discovery” technology in order to quickly locate data and, if needed, recover lost information.
- Providing better, faster, and more reliable services to customers.
- Implementing technologies that can scale cost-effectively to accommodate growing data storage demands.

## Key Secure Storage Scenarios

Almost any medium or large organization operating today has the need for some kind of secure, fixed-content data storage capability. But there are three scenarios in which cost-effective, secure, and scalable storage is paramount:

**E-mail archiving.** For most organizations, e-mail is a mission-critical application. The use of and volume of e-mail has exploded in recent years, with storage requirements increasing exponentially. Businesses of all sizes face many challenges in managing their growing repositories of e-mail. These include compliance concerns, storage management limits, the ability to find archived e-mail when needed, and efficient retrieval of e-mail from traditional tape or optical archives (a process that can cost many thousands of dollars).

**Managing financial data.** Safe, effective storage of financial data over long periods is important to any company. The need is particularly acute in the financial services sector where immutability, secure retention, and secure deletion at the end of the retention period are all vitally important. Financial institutions have to make significant investments in core technologies, including storage, to address the business challenges and regulatory requirements of protecting business-critical information. Check imaging, mortgage contracts, loan agreements, securities trades, customer statements—these and other kinds of information that are central to financial services demand secure and scalable storage technologies.

**Protecting healthcare information.** Healthcare organizations are rapidly

“...while storage solutions that are based on commodity components provide some protection by using a redundant array of independent disks (RAID) component and replication, there is no guarantee that data in such a solution is healthy and in a usable format.”

“Selecting a CAS archive that transparently integrates with your applications enables easier data migrations as you upgrade or change applications.”

digitizing their practices, and some have reached the point of being virtually paperless. Putting healthcare information into electronic format poses major challenges, in terms of both the size and the critical nature of what needs to be retained, including large files such as radiology images and other diagnostic imagery, lab tests, and other patient data. HIPAA and other regulations, along with the demands of providing high-quality customer care, require healthcare organizations to carefully evaluate their storage options to select the optimum solution for ensuring integrity, privacy, and immediate access to this data.

## Storage Planning

Given the conditions described here, organizations need to carefully plan for their storage technology options to accommodate today's needs and tomorrow's demands. Again, while every company is going to have its own specific requirements, there are some general considerations to keep in mind when planning for enterprise fixed content data storage. Here is a brief overview of each.

**Protection.** Many application and archive vendors give their customers the option to use commodity hardware along with their own add-on software and will offer this as a “complete archiving solution.” Such “solutions” are typically offered at an attractive price and may look like a desirable alternative for companies that are seeking to control costs. However, while storage solutions that are based on commodity components provide some protection by using a redundant array of independent disks (RAID) component and replication, there is no guarantee that data in such a solution is healthy and in a usable format. In addition, security features can easily be compromised because the files are not protected from theft, modification, or deletion.

Because the archive software is not integrated with the commodity storage hardware, it would be fairly simple for someone to gain direct access to the storage solution and potentially access, modify, or destroy the stored information. For example, an attacker could bypass the application security, directly access the commodity RAID storage system, and then reformat the entire system, destroying all the archived data. Or worse—the attacker could delete a single file that would be almost impossible to detect until too late. A good archiving solution should provide strong management of both the hardware and the software layers. It should provide a layer of security for the storage so that it cannot be accessed without going through the management layer.

**Migration.** An accurate evaluation of fixed-content storage technologies requires an understanding of traditional content-addressable storage (CAS), which is used for storing information so that it can be retrieved based on its content rather than on its storage location. When a typical CAS archive receives a file, it creates a CAS address at the application level—not at the storage level. As a result, the application is the only mechanism that knows how to find the file.

If an organization later wants to change the application or to upgrade the archive system, it must migrate all of the data to the new system and must use the old application to perform the migration. This creates what is called “vendor lock-in” because the migration process can require vendor expertise and can be time-consuming, especially if the archive that is being migrated is large. It can also temporarily prevent the use of mission-critical enterprise assets such as an e-mail system.

Selecting a CAS archive that transparently integrates with your applications enables easier data migrations as you upgrade or change applications.

**Energy costs.** Storage systems are among the most energy-intensive of all IT components. Much of the cost of running or sharing part of a data center goes to the power that is needed to operate the disk drives, which consume about 60 percent of the power in a storage system.

A major cause of this energy consumption lies in the design of typical CAS storage systems: They have a CAS single object store and its related database that is distributed across multiple drives. Because the location of specific files is unknown, a simple request for a single file can result in the spin up of multiple drives, consuming a lot of electricity for a small task. Today, many organizations have sustainability guidelines that include directives to reduce energy consumption, and IT departments need to consider the environmental and cost effects that each storage alternative will have.

“Because the location of specific files is unknown, a simple request for a single file can result in the spin up of multiple drives, consuming a lot of electricity for a small task.”

Multiple virtual archives and corresponding drive power saving technologies are critical factors when selecting a CAS archive for fixed-content data.

**Scalability.** This is a feature of the storage architecture that is frequently misunderstood—often to the dismay of IT departments after they have committed to a storage vendor. When evaluating storage options, administrators need to understand that scalability is very different from the total amount of usable storage that a particular storage technology offers.

When planning for fixed-content CAS storage, IT departments need to remember that “object count”—the number of storage objects that are put into a CAS storage system—can have a dramatic and often diminishing effect on the storage space that is supposedly available. For example, a company might decide that 5 terabytes of storage is adequate to archive its e-mail traffic for several years. However, while the majority of the company’s e-mail messages may be quite small—perhaps just a few kilobytes each—typical storage systems will count each e-mail message as an object. Thus the company may reach the object limit much sooner than it reaches the actual storage limit.

Some CAS architectures have the limitation that once the object limit is reached, no more data can be stored on those nodes’ storage. This renders the remaining available disk space on those nodes permanently unusable. The only option is to invest all over again in more storage nodes in order to add more objects. But those objects can only be stored on the new storage, not the old, unused storage.

“Some CAS architectures have the limitation that once the object limit is reached, no more data can be stored on those nodes’ storage. This renders the remaining available disk space on those nodes permanently unusable.”

In sharp contrast, other CAS architectures independently scale object count and storage capacity, so the full capacity of the system can always be fully utilized. Choosing a CAS archive solution with a scalable architecture is essential to reducing long-term fixed-content storage costs.

**Performance.** Performance and scalability are related. A common complaint with storage systems is that as they fill up—or more accurately, as storage objects accumulate—the speed at which an archive database operates can degrade dramatically. The result can be significant time lost and immense frustration for people needing to find and retrieve specific bits of information, particularly if they are working under a project or compliance-driven deadline.

What’s important to remember is that a typical archive within a storage system contains a single object store and database that cannot be modified to accommodate multiple file systems. The ability to generate multiple file systems, each of which can manage its own object database, is key to improving performance. It provides for better organization of storage objects and makes it possible for server processors to better manage smaller, more discrete groups of data.

**Integration.** A secure fixed-content storage system should be able to integrate seamlessly with an organization’s existing enterprise and departmental applications and other IT infrastructure such as Active Directory and Domain Name Service (DNS).

IT departments already have many demands and the addition of a new IT component should not lead to resources being taken away from other projects. In addition to the seamless integration as noted above, an ideal scenario would be where all components of a secure storage solution are tightly integrated and

share a similar look and feel throughout the management interface for the entire solution.

“Assureon is an integrated solution: it has all the processing nodes, storage nodes, switches, cables and software installed and fully configured into a unified secure archiving system.”

## Assureon: Meeting the Need for More Secure Storage

Assureon from Nexsan Technologies can satisfy today’s most demanding fixed-content storage requirements. Assureon is an easily managed, scalable storage appliance that provides numerous services that ensure privacy, integrity and longevity.

Assureon is an integrated solution: it has all the processing nodes, storage nodes, switches, cables and software installed and fully configured into a unified secure archiving system. Simply connect the Assureon solution to an enterprise network, plug it into standard 110/220V AC UPS power distribution units (PDUs), and it’s ready to go.

Internally, the Assureon architecture makes possible highly-flexible scalability through the creation of multiple virtual storage archives combined with being able to independently scale processing nodes (servers) and storage nodes (high-density RAID storage chassis). The result is virtually limitless scalability, fast performance, and secure data separation that provides separate secure archives for different divisions or for SaaS (Software-as-a-Service) hosted archive customers.

Assureon can be configured to be as simple as a single server (processing node) attached to a RAID box (storage node), or as complex as many dozens of servers attached to redundant GigE network and storage area network (SAN) switches, numerous high-density RAID boxes, racks, and a KVM switch. Data to be archived can be ingested from corporate application or file servers or from Assureon Edge NAS Heads.

Assureon provides key features that meet customer needs for storage that is more secure, scalable, and cost-effective:

**Windows features and integration.** Assureon’s architecture is tightly integrated with the Windows® Server 2008 64-bit operating system providing a fast, reliable, and scalable environment . Because it shares the same features as other Windows-based products, Assureon—unlike open-source alternatives—is easily and seamlessly integrated into Windows environments that are used in enterprises worldwide. In addition to Assureon’s robust Windows support, Assureon also provides full LINUX/UNIX application support through the Assureon Edge NAS Heads (NFS/CIFS support).

Assureon is fully integrated with key Windows technologies such as Active Directory® to eliminate duplicate account administration, the Active Directory integrated Domain Name Service (DNS), SQL Server, NIST validated cryptographic modules, IIS for web services, .NET, AJAX, WSS NAS, and WUDSS NAS. Assureon utilizes familiar management tools and provides an easy to use web-based administration interface that manages the entire Assureon system., making it quick and effortless for IT administrators to begin using and managing Assureon.

### Protection

#### **Immutable disk-based archive with WORM Disk Technology.**

Assureon’s fully-integrated hardware and software solution provides secure data protection that is simply not possible with software running on block storage. Assureon’s data integrity begins with a strong foundation of Write Once, Read Many (WORM) technology that protects files and volumes from accidental or deliberate modification or deletion after secure archiving.

**Digital file fingerprinting to ensure and validate data integrity.** Assureon’s dual cryptographic hashing algorithms (both SHA1 and MD5) create a “fingerprint”

“Assureon can be configured to be as simple as a single server (processing node) attached to a RAID box (storage node), or as complex as many dozens of servers attached to redundant GigE network and storage area network (SAN) switches, numerous high-density RAID boxes, racks, and a KVM switch.”

“Assureon audits all archived files for long-term data integrity to ensure each serial numbered file exists in each WORM store and that each file continues to match its fingerprint.”

“It self-diagnoses and self-heals any discrepancy, replacing a bad file with a good copy.”

of each file’s content prior to archiving. After the file is transferred to Assureon, this fingerprint validates that every bit of the file is authentic (and was not corrupted in transfer). If one bit is different, it won’t match its fingerprint. It’s revalidated just before the file is written to the secure WORM store.

**Data integrity auditing and file availability auditing.** Periodically throughout the duration of each file’s retention period, Assureon audits all archived files for long-term data integrity to ensure each serial numbered file exists in each WORM store and that each file continues to match its fingerprint. It self-diagnoses and self-heals any discrepancy, replacing a bad file with a good copy.

**Configurable retention times.** Organizations can choose to create compliant-level retention periods for data so that it cannot be deleted or modified for a set period of time, or they can use flexible retention settings for files that do not require regulatory compliance or corporate governance.

**Encryption of files at rest.** Assureon can protect archive files from a privacy and confidentiality standpoint by encrypting files at rest (within Assureon) with AES256 Encryption. The U.S. government has validated both AES192 or AES256 for protecting Top Secret Classified information. Assureon uses the stronger of the two encryption algorithms. At the end of a retention period, if permanent deletion is approved, the encryption key is securely destroyed and all copies of the file are deleted. Even if a backup copy was made on tape, the file becomes unreadable and cannot be recovered regardless of its location.

**Embedded replication to business continuity or disaster recovery site.** Assureon data ingested at a primary site can easily be replicated to another Assureon at a business continuity or disaster recovery site (either locally or geographically dispersed). Assureon configs can provide active-passive or active-active failover to secondary site in the event that a primary site experiences a planned outage or disaster.

**Application Integration.** Assureon integrates with enterprise applications including e-mail archiving, Enterprise Content Management (ECM)/Document Management, Enterprise Report Management (ERM/COLD), check imaging, document imaging, and many other fixed content applications.

**Management and integrations.** Assureon Edge NAS Heads can provide a Common Internet File System (CIFS) and Network File System (NFS) interface. Files that are written to these shares can be automatically archived to Assureon based on pre-described policies that the user configures.

Assureon Clients (File System Watcher (FSW)) can archive directly from data directories on Win2003 and Win2008 application and file servers.

**Powerful Archiving Policies.** Policies can be configured for different file folder locations, specify data category and subcategory, whether the original file is left after being archived, replaced with a shortcut or deleted.

**Migration.** Your data is securely archived within Assureon. With appropriate permissions, you can use Assureon Explorer to restore files to the original or to an alternate directory – or restore shortcuts to the original or an alternate directory.

**Reporting.** Assureon provides a suite of extensible reporting tools that provide information on storage utilization, de-duplication efficiency, performance and alerts. Error conditions are written to event logs and can also be configured to HTML e-mails to the administrator immediately or delayed by minutes or number of alerts.

**Access Log Reporting.** A secure record is also kept of both successful and unsuccessful attempts to access archived data.

**Legal Hold / Block Disposition.** A file, file(s) or folder(s) may be put on legal hold to prevent a file in litigation from reaching the end of its retention period and then the file

accidentally being deleted.

**Metadata searches.** Search through hundreds of terabytes of data and find and/or retrieve a single file—or thousands of files—within a few seconds based on file metadata such as category, subcategory, server, path, filename, etc.

**Rapid full-text content indexing and search.** Assureon’s optional full-text content indexing and search makes it possible to search through hundreds of terabytes of data and retrieve a single file—or thousands of files in your search results —within seconds based on file content, phrases, or keywords inside Office documents, spreadsheets, presentations, PDF files, etc., and more file metadata such as path, filename, extension, date, author, and type, etc. An upgrade is available to include OCR of scanned images. Users can simultaneously use multiple search criteria and combine this with metadata searches to narrow down results with drill-down content search. This functionality can be used to support legal tasks, special projects, and other finding file content customer needs.

“Assureon provides features and benefits that not only meet technical, corporate, and regulatory requirements, but also help control the total cost of ownership of a storage system.”

## Assureon: Improved Total Cost of Ownership

A common question when organizations are evaluating storage is, “How much will it cost?” Assureon provides features and benefits that not only meet technical, corporate, and regulatory requirements, but also help control the total cost of ownership of a storage system. Assureon accomplishes this through a Windows heritage that facilitates ease of use and ease of integration; with “self-healing” functionality that provides automated auditing, self-diagnosis and self-healing of archives; and with energy savings that are delivered through the AutoMAID (Automatic Massive Array of Idle Disks) technology that is a core feature of the Assureon solution.

**Windows heritage.** As noted earlier, Assureon was designed on—and built to seamlessly integrate with—IT environments that are based on the Windows operating system. Not only does this help IT administrators to quickly come up to speed on using and managing Assureon, but it also minimizes potential issues and long-term resource drains that often result when a storage system’s operating environment is different from the rest of the enterprise. Assureon also supports Linux and Unix environments through the Assureon Edge NAS Head’s NFS support. Assureon Edge supports CIFS and NFS.

“Assureon can save between 20-60% of disk space for storing departmental files, email archiving as compared to storing those high-duplication types of files on traditional storage.”

**Efficient storage de-duplication.** Assureon can save significant amounts of disk space required to store high-duplication archiving files where many files will have the same content, but different metadata (e.g. different path, filename, etc.). Assureon can save between 20-60% of disk space for storing departmental files, email archiving as compared to storing those high-duplication types of files on traditional storage. It stores the data by preserving each file’s metadata but deduplicating saving duplicate file content more than once.

Each metadata instance points to the same Content Addressable Storage (CAS) object in the Assureon CAS object store – however each metadata instance can have its own retention policies and category and subcategory. Each file’s retention date can be separate – based upon the metadata. They share the same single instance in the CAS object store.

**Self-diagnosing, self-healing.** The storage provided in Assureon is designed to provide “lights-out” archiving. The self-diagnosis and self-healing features that are built into Assureon automatically scan stored files to ensure that file-level integrity is intact, that files match their digital fingerprints, that each serialized file is in each store and that any missing or corrupted files are self-healed with a good copy of the file. These built-in audit and repair features of Assureon help dramatically reduce the amount of time that IT personnel need to spend on manual administration of archives.

**AutoMAID.** With the AutoMAID technology in Assureon, organizations can save up to 60% off traditional storage energy costs. This is accomplished with a sophisticated

range of AutoMAID features which cause drives with reduced levels of disk activity (older archives) to begin spinning slower, then yet slower, then unload heads, to further reduce power consumption. Assureon's virtual archive technology combines with AutoMAID to maximize savings by enabling creating a new virtual archive every x months or x years. Then as the older archives are accessed less, they go into deeper levels of AutoMAID – thereby further reducing energy consumption. AutoMAID can also help organizations to meet sustainability initiatives that target specific “green” metrics, such as reducing energy consumption and related carbon footprints.

## Conclusion

Today's data-intensive business activities, combined with regulatory environments, require organizations to carefully evaluate and implement secure storage systems that can archive and protect critical information such as e-mail, office documents, scanned documents, financial documents and reports, legal contracts, and healthcare files.

Assureon addresses the needs of organizations by providing highly scalable, high-performance secure archiving – enforced from the software all the way to the firmware on the hardware to ensure data protection. It protects data by offering features such as self-healing and self-auditing technology. It is easy to manage and integrates seamlessly into Windows environments and can also integrate into LINUX/UNIX environments. It is cost-effective and helps organizations to manage the total cost of ownership for their storage system by providing energy-reducing technology and low-maintenance operations.

With Assureon, organizations can address the complex challenges of managing their data over the long term by using an affordable, high-performance “safe deposit box” for the digital age.

“Assureon addresses the needs of organizations by providing highly scalable, high-performance secure archiving – enforced from the software all the way to the firmware on the hardware to ensure data protection.”