

Cohesity SmartFiles: Beyond Scale-Out NAS

Unstructured Data Growth Challenges

According to Gartner, data will grow by 800% over the next five years, of which 80% will be unstructured in the form of file shares, backups, archives, logs, media files, dev/test and analytics. Traditional network attached storage (NAS) was designed over a decade ago for traditional use cases, and without serious architectural consideration for cloud and hybrid IT environments. The reality today is that there are multiple data silos resulting in mass data fragmentation across the data center and cloud.

Challenges are not limited to the cost and management of multiple silos, or delivering hybrid cloud seamlessly. Storage costs often destroy storage budgets. Addressing compliance, regulatory, and governance requirements add to complexity and cost. That's why we created Cohesity SmartFiles: to go beyond traditional file and object platforms and address real data management pain with real on-premises, hybrid, and cloud solutions.

What is Cohesity SmartFiles

Cohesity SmartFiles is the industry's first software-defined, data-centric, multiprotocol file and object solution for the enterprise. SmartFiles is part of the Cohesity DataPlatform and goes beyond traditional scale-out NAS in manageability, scale, storage efficiency, integrated applications, cybersecurity, and multi-tiered data management.

Why Is SmartFiles Smart?

Unlike traditional NAS, the Cohesity DataPlatform with SmartFiles is data-centric, not storage or infrastructure centric. This means there is no need to move data to apps as apps and data are integrated in the same platform. This removes the complexity and cost of running file ecosystem apps for file environments. Cohesity Marketplace apps and machine-learning assisted management provide intelligence and ease-of management that separates SmartFiles from traditional scale-out NAS appliances. It's smart and with virtually effortless management regardless of scale.

SmartFiles uniquely provides freedom of choice, broad compatibility, stress-free management, integrated security, search of the entire geographic enterprise, and lower storage cost.

KEY BENEFITS

- *Freedom of choice*
- *Broad compatibility*
- *Stress-free operation*
- *Eliminate security risks*
- *Lower storage cost*
- *"Google-like" search*

Freedom of Choice

Software-defined means customers have the freedom to choose the most appropriate hardware for any workload. It's the freedom to match capacity, performance, and cost with workload requirements. But this is not just about today, it's about tomorrow. SmartFiles future-proofs unstructured data workloads on the assurance that workloads are not bound to proprietary hardware – or even the data center. SmartFiles runs in the cloud with all the functionality available on-premises. And it's an easy on-ramp to the cloud through native S3 protocol – along with unified NFS and SMB protocols for on-premises and the cloud.

Broad Compatibility

SmartFiles has you covered if you have a mixed OS environment or hybrid cloud environment. File permissions across NFS and SMB environments are mapped for transparent file access from either protocol. Files and objects can be accessed simultaneously, and all data is always available as S3. Hybrid environments are easier as SmartFiles speaks the language of Windows/Linux/Unix and the cloud. Seamless data interoperability helps make complex hybrid environments easy.

Stress-Free Operation

It begins with an intuitive UI and "click-and-go" ease. It ends with a bit more time to focus on strategic matters – or just extra time to relax. Machine-learning driven management acts as a hidden smart assistant to remove complexity and save time. Data migrations are easy using Cohesity's backup solution – which removes the complexity of other migration tools. Automate operations through rich APIs. Cold data on costly Tier 1 storage automatically moves itself to a cost-optimized tier or the cloud. Utilize policies to ensure consistent management. Manage across multiple clusters and locations from a single management pane. Cloud or on-premises. Core or edge. Physical or virtual. Data movement is always easy and transparent to applications. Cohesity SmartFiles alleviates the stress and complexity of traditional scale-out NAS.

Integrated Apps

Integrate an entire NAS ecosystem on a single platform. That means anti-virus protection, file audit, and content search all run on the Cohesity DataPlatform with no additional infrastructure required. Detect anomalous file accesses from file logs. Index and search contents inside of files for compliance and security reasons. The SmartFiles solution eliminates the cost of infrastructure for a NAS ecosystem while removing the management effort required to manage NAS ecosystem hardware.

Eliminate Security Risks

Cybersecurity is not an option for unsecured data. It's not a want – it's a need. SmartFiles integrated cybersecurity makes security easy to deploy and manage. It's a multi-layered security approach to prevent, detect, and analyze threats:

- **Prevent** – Immutable file system, software encryption, over the wire encryption, multi-factor authentication, DataLock (WORM), and adherence to FIPS 140-1 and 140-2 standards.
- **Detect** – Anti-virus: Detect infected files and prevent files from being compromised
- **Analyze** – Observe user and behavior analytics for security and compliance reasons

“Google-like” Enterprise Search

It's not just simple, it's deep and data-centric. Get fast search results across file metadata and file content as both are indexed. Unlike traditional NAS, use results to drive data-driven services and make decisions on actual data contents, and not just data about the data. Data is no longer dark. Rather, it's an asset you need for smart operations. Conduct a single simple search across the entire enterprise. Search any data – anywhere. Across silos, VMs, backups, data centers, remote sites, and multiple clouds.

Low storage cost

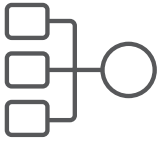
Being software-defined, SmartFiles customers have a variety of hardware price points to choose from. SmartFiles reduces hardware storage requirements through data deduplication, compression, small file optimization, and dedupe across data center volumes. Storage costs can be reduced even further through tiering from Tier 1 storage to a cost-optimized Cohesity tier. Let's explore this technology.

- **Data deduplication** – Based on advanced sliding window variable dedupe technology. This means that duplicate data that has been slightly changed is still deduped – which is beyond traditional dedupe capabilities.
- **Compression** – Zstandard compression algorithm for high compression ratios.
- **Small file optimization** – Eliminated the small file amplification penalty common to traditional NAS file systems. Small files are stored efficiently and without the need for triple-mirroring inefficiency. Put simply, 50TB of small files under SmartFiles consumes about 50TB of capacity.
- **Cross-volume dedupe** – Unlike many traditional NAS appliances, data is deduped across storage volumes and application silos across the data center. Duplicate data that can't be reduced by traditional silo-by-silo dedupe is eliminated by the Cohesity DataPlatform.

SmartFiles supports the migration of older inactive data from Netapp, Dell/EMC Isilon, or Pure Storage transparently and seamlessly to Cohesity or to a cloud tier. Apps won't know the difference and storage costs are instantly and drastically reduced.

“At Expedient we see 2X growth year-over-year, and operating at such a fast pace requires us as a service provider to push our partners and engineering teams to continuously innovate,” said John White, chief innovation officer, Expedient. “We provide the back end for leading companies all over the world, and as an example, just one of our customers has more than 2.5 billion files, so it was imperative to enable a scale-out, software-defined file and object solution across our environment.”

There are dozens of workloads and use cases in which SmartFiles provides benefit. Amongst these use cases are:



Content Management

- Document management
- Scalable libraries
- Files and objects



Splunk Cold Buckets

- Cost-effective cold archiving
- on-premises, cloud or both
- Frozen data to cloud



Corporate Video

- Corporate communication
- In-house productions
- Education and training



Digital Archives

- Digital documents
- Audio and video files
- Active archives
- Passive archives



Video Surveillance

- Security cameras
- Body and mobile cams
- Rapid retrieval from large amounts of data

There are many industries that depend on Cohesity for file and object services. These industries include:



Financial Services

- Images and documents
- Call logs and correspondence



Law Firms

- Documents and research
- Collaboration
- eDiscovery and legal hold



State, Local and Education

- Collaboration and research
- Safety and security
- Innovation and technology



Healthcare

- PACS archives
- VNA archives



Life Sciences and Medical Research

- Next generation sequencing
- Genomics analysis

FEATURE	DESCRIPTION
NFSv3, CIFS, SMB2.x, SMB 3.0, and S3 APIs	Simultaneous multiprotocol access to the same data with unified permissions. Supports NFS, SMB, and S3 API
Strict consistency	Guaranteed consistent data regardless of scale
SnapTree® snapshots and clones	Limitless and fully-hydrated snapshots for granular Cohesity Views (file systems) as well as writable snapshot clones that provide instant creation, testing, and development of View-based data sets
Fully distributed shared-nothing file system	Limitless scalability, always-on availability, non-disruptive upgrades, pay-as-you-grow model
API-first design	Architecture built with an API-first design for maximum flexibility and ease of automation
Data management platform	Single platform for data protection, files, objects, dev/test, analytics, and cloud integration
Global deduplication and compression	Unparalleled storage efficiency with global deduplication, compression, and small file capacity optimization across all cluster nodes to significantly reduce data center cost and footprint
Erasur coding and replication factor	Data is protected against any individual node failure with erasure coding or replication across nodes

FEATURE	DESCRIPTION
Global indexing and search	File and object metadata is indexed upon ingest, enabling Google-like search across all files in a cluster. Indexed search of file contents through Insight integrated app. Search across geographic locations and clouds through Cohesity Helios.
Windows Active Directory and Kerberos Integration with Role-Based Access Control (RBAC)	Simplify user and group access to data utilizing credentials and permissions with Windows AD and Kerberos mechanisms. Create and manage custom Cohesity cluster administration roles for domain users and groups
MMC Integration	Snap-in for the Microsoft management console, which allows Cohesity file shares to be managed by the MMC
Quotas	Easily establish user and file system quotas with audit logs
Policy-based backup protection	Integrated data protection software and SnapTree technology is available to allow simplified data protection of objects with fully-hydrated snapshots
Quality of service (QoS)	QoS policies are provided that optimize performance for different types of workloads
Encryption	Support for data-at-rest as well as data-in-flight encryption using the industry standard 256-bit Advanced Encryption Standard (AES) algorithm. The platform is also FIPS 140-1 and 140-2 compliant.
Write Once Read Many (WORM)	Enables long-term retention of data that has compliance controls mandating a policy that objects cannot be modified during the lock time
Replication for disaster recovery	Built-in, granular, and secure replication services for geo redundancy
Cloud integration (CloudArchive, CloudTier, CloudReplicate)	Archive into public cloud services for long-term retention. Utilize cloud tiering for transparent capacity expansion into the cloud. Replicate into the cloud for disaster recovery and dev/test. Also, SmartFiles runs in the cloud of major cloud service providers with all of the functionality of on-premises deployments.