COHESITY

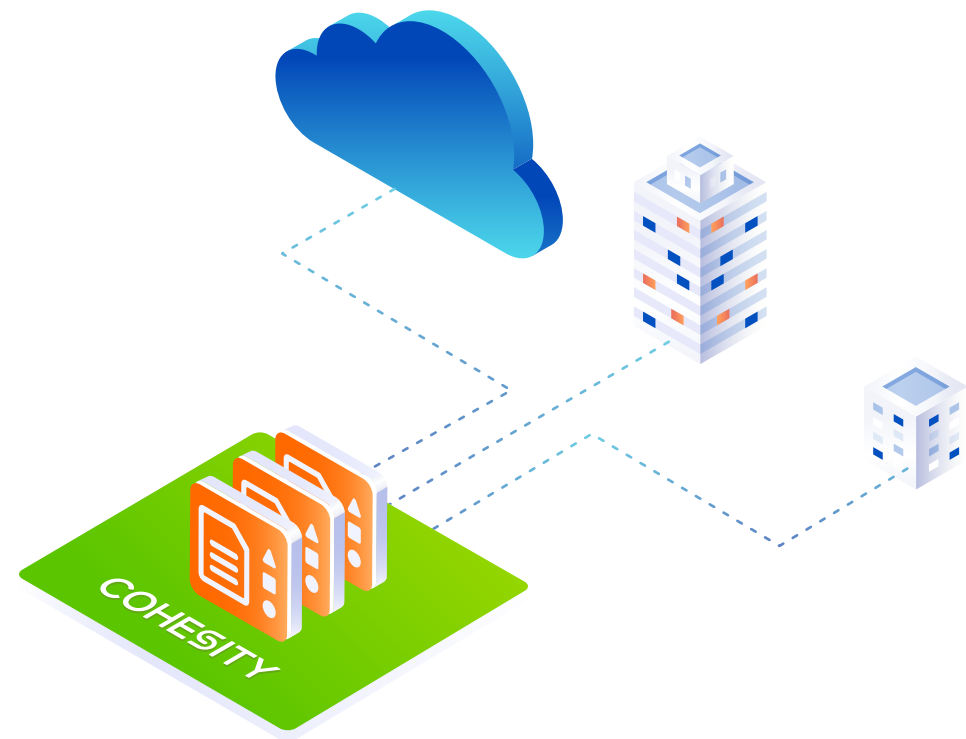# Get more **mileage** from your NAS investment

From Vintage to Vogue

Older vehicles become less efficient and effective over time due to a combination of factors, particularly compared to newer, more advanced models. Likewise, legacy NAS is challenged when it comes to being cost-effective to own and operate while keeping up with the demands of modern applications.

Auto manufacturers can debut vehicles that go farther for the price paid at the pump every year because they design their newest models to be increasingly fuel efficient. Think about all the ways brand new cars are purpose-built for efficiently navigating both old streets and newly paved highways as well as changing traffic and safety patterns. Now consider how much enterprise network-attached storage (NAS) has changed in the last decade, despite the unprecedented growth of unstructured data and the rapid adoption of public and hybrid clouds. The short answer is, it hasn't.

Modern, scale-out NAS is built for digital business in a hybrid cloud world—with efficient capabilities that boost the value of storage investments. The result: lower capital investment, operational support, and data center footprints even as IT environments change. This ebook describes the evolution of NAS and how newly introduced approaches are helping to overcome cost and operational challenges while putting organizations on the road to infrastructure modernization and competitive advantage.

The Journey to NAS

Looking in the rearview mirror, it's easy to track how IT storage and data access evolved. It had a common engine—the block, named from the chunking of data and distribution of it evenly and arbitrarily across volumes.

At the time, servers and storage were tightly coupled. Each hard disk stored only the information read and written locally by that machine. The size of the file consumed the entire slot allocated to storage and at best, only the administrator responsible for the machine knew what was stored on it and where. Yet as digital files and sizes grew throughout the 1970s, the inextricable server–storage link became impractical.
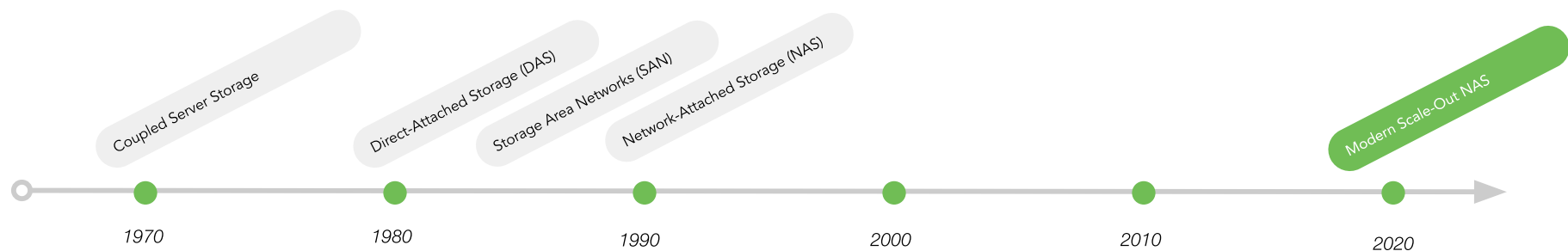
In the early 1980s as digital business accelerated, the need to decouple servers from storage systems became clearer, and a new approach rolled into the market. The introduction of direct-attached storage (DAS) appliances—deployed as hard-disk drives (HDDs) or solid-state drives (SDDs)—brought the delivery of computer storage connected to one computer and not accessible to other computers. Groups of drives, external to the server but directly attached through Small Computer System Interface (SCSI), Serial Advanced Technology Attachment (SATA), Serial-Attached SCSI (SAS), Fibre Channel (FC) or iSCSI, were standalone. No systems were yet networked through Ethernet or FC switches.1

Although fast, DAS required significant IT administrator time and faced usability challenges because it only supported one user at a time and lacked an interface for concurrent access across organizational teams. Yet companies trusted it for their highest performance workloads.

At the same time, the networking industry was evolving and enterprises welcomed another option to overcome key DAS shortcomings while boosting storage-space sharing between several servers: the storage-area network (SAN). Yet, it too had drawbacks, including the inability to read or write simultaneously from several machines.

In the mid-1980s, several companies—Novell, Sun Microsystems, 3COM, and Auspex Systems among them—began to promote a completely different computer data storage system: network-attached storage (NAS).

Why NAS? Because digital collaboration was skyrocketing and workforces were editing and sharing documents and folders. NAS would replace arbitrary blocks of storage with hierarchies of files. For users, file-based storage quickly became a very popular and practical operational approach. For IT organizations, it created a new model of data storage, access, and management.

Coupled Server Storage

Direct-Attached Storage (DAS)

Storage Area Networks (SAN)

Network-Attached Storage (NAS)

Modern Scale-Out NAS

1970   1980   1990   2000   2010   2020

¹Tech Target: Definition of DAS. https://searchstorage.techtarget.com/definition/direct-attached-storage

Under the Hood of Traditional NAS

NAS today is very similar to NAS from the 1980s. It's preconfigured storage software installed on dedicated hardware with a data transfer protocol, either:

Network File Systems (NFS) – Designed for Linux and UNIX systems, NFS works across hardware, OS, and network architecture

or

Server Message Blocks (SMB) – Designed for Microsoft, SMB is a derivative of the common internet file sharing (CIFS) protocol

In today's world, traditional NAS systems like older automobiles are less efficient than more contemporary offerings. And that makes them more costly to own in two critical ways:

More parts (disks) means greater capital expense

To keep up with exponential data growth, storing files on traditional NAS (even scale-out NAS) requires adding resources—new hardware with expanded RAM footprints and improved CPUs. Every time IT teams do that, they add cost to the business. This quickly becomes a significant part of valuable IT budget, particularly when additional systems are purchased to house files that teams haven't touched in weeks or months—or for which the business must store for extended time periods in order to address government or industry compliance requirements.

Every NAS device added also is limited by the physical capabilities of its hardware. That means as data grows, servers sprawl. This issue is compounded by traditionally inefficient NAS deduplication capabilities. Deduplication (aka dedupe) is one of the most valued technologies in enterprise storage products because it's designed to eliminate the need to store multiple copies of identical files. Dedupe techniques reduce the amount of storage capacity required to store any given amount of data. The challenge with traditional NAS is that even when products can dedupe data across data center volumes (which not all can), legacy dedupe isn't variable sliding window, and often incurs storage amplification penalties when handling small files. These inefficiencies all contribute to the rate at which new hardware investment becomes required to keep up with ongoing growth.

Moreover, because NAS hardware is siloed from other storage and devices, organizations often have multiple copies of data siloed across their data centers and clouds, resulting in mass data fragmentation. And of course, there are costs associated with having multiple copies of data stored in varying locations.

For traditional NAS offerings, introducing hybrid cloud into the equation requires yet another investment because teams must purchase a separate object gateway.

And all this is just to store the data. If operational teams also want to enhance security, additional infrastructure is typically required to perform antivirus scans, compliance checks, file auditing, analytics, and the like. None of these investments is free and all further stretch already tight IT budgets.

More elbow grease (labor) adds operational expense

In addition to traditional NAS capital costs (CapEx), organizations incur the operational expenses (OpEx) of running legacy NAS systems. Beyond comparing per GB disk costs, infrastructure spending must be done with a focus on lowering OpEx. For example, if it takes one person to manage 100TB of storage, it ideally takes one person to manage 100PB. This level of operational efficiency is only possible with modern architectures. Using a traditional architecture to manage across data center silos, clustered NAS pairs, and even multiple sites and/or clouds requires extra effort and time—taking away from innovation opportunities.

Because they were designed over a decade ago and for traditional file use cases, traditional NAS devices don't include serious architectural considerations for the cloud and hybrid IT environments where object storage predominates. Traditional NAS is therefore a rockier road to ongoing IT modernization initiatives.

Teams need dedicated staff resources to install and manage all of these disparate parts:

- Separate object gateways

- Separate cloud gateways

- And perhaps once again, separate application infrastructure (e.g., AV, compliance, analytics, and data management/tiering)

If the organization has a heterogeneous environment, NAS also requires specialists who understand each of the different management user interfaces (UI) that may have accumulated over the years. These resources must be on hand for ongoing troubleshooting and updating of existing and growing numbers of systems storing files.

Lastly, legacy NAS was designed for data traversing country roads but today enterprises are connected via networked superhighways which support modern data sources—from Kubernetes to containers, Hadoop to NoSQL. Administrators typically don't have the time nor resources to tune traditional NAS for new workloads and deployment models which they weren't originally designed to support.

" Cohesity is unique in that it plugs into all of our automation products and integrates seamlessly with our backup and NAS strategies, increasing our efficiency and unifying our environment. "

- **Matthew Dartez,** Director of DevOps and Infrastructure, Kyriba

The Right Files Model for Digital Business

Modern, scale-out NAS is an efficient architecture purpose-built for hybrid cloud, supporting:

- Today's performance-intensive workloads (e.g., video and images stores, corporate file shares)
- High capacity-intensive workloads (e.g., high file-count applications and home directories)
- Unmatched storage efficiency—lowering equipment, power, and cooling costs
- Boosting security with integrated antivirus, file audit, and search capabilities

The right modern NAS provider will offer software-defined, web-scale flexibility for enterprise file and object services in an approach that simplifies management and lowers operational costs wherever workloads are deployed.

	Modern, Scale-Out NAS	Traditional NAS	Impact
Number of files per directory	3.2 million	1 million	Although 1 million is stated, 100,000 files in any one directory is recommended as best practice for traditional NAS.
Max-path length	131092	1023	Lower traditional NAS limits force customers to reconsider directory structure.
Max file size (non-sparse)	70 TiB	4 TiB	With the modern approach, the max file size is based on the physical size of the cluster. If doubled, the max file size would scale, too.
Max number of views/shares	800,000	80,000 (SMB) 40,000 (NFS)	In a high-share environment (e.g., home directories), more deployed traditional NAS clusters would create more silos compared to modern approaches.
User quotes per directory/view	100,000	20,000	In a home directory environment (e.g., a university), traditional NAS requires admins to closely monitor user quotas.
Maximum number of Snapshots	1 million	20,000	The ability to perform nightly Snapshots via a policy on a directory with traditional NAS would be limited to fewer years than modern, scale-out NAS.

From Vintage to Vogue | The Journey to NAS | Under the Hood of Traditional NAS - 1 | Under the Hood of Traditional NAS - 2 | The Right File Model for Digital Business | Lower TCO Starts with Improving Data Reduction | Eliminating Hardware, Reducing Data Center Footprints | Convergence and Expanding Use Reduce Operating Expenses | Enjoy the Efficiency and Luxury of Cohesity SmartFiles

Lower TCO Starts with Improving Data Reduction

It's more important for IT organizations to show rapid return on investment (ROI) in times of economic uncertainty. As file repositories continue to rapidly grow, maintaining the status-quo with legacy NAS can be negatively impacting business. There's a better, up-to-date way.

Modern, scale-out NAS is software-defined. It features more efficient data reduction technologies for impressive storage efficiency. Just as a newer vehicle is more efficient, modern NAS reduces TCO today and tomorrow with key capabilities that boost ROI:

- **Advanced data deduplication** – Based on advanced sliding window variable dedupe technology, modern NAS ensures duplicate data that has been slightly changed is still deduped—going beyond traditional dedupe capabilities. And files and objects can be deduplicated across data center silos.

- **High compression** – High compression ratios come from modern NAS's embedded Zstandard compression algorithms.

- **Small file optimization** – Modern NAS eliminates the small file amplification penalty common to traditional NAS file systems. Small files can be stored with near 100% efficiency and without the need for triple-mirroring inefficiency.

- **Cross-volume dedupe** – Unlike many traditional NAS appliances, modern NAS 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 removed.

Eliminating Hardware, Reducing Data Center Footprints Cuts Capital Costs

In contrast to traditional NAS, a modern approach to files with advanced data reduction allows for more data in the same hardware space—at less cost. The engine powering the advantage is deduplication, or how storage can be used most efficiently, and by doing so, reduce infrastructure costs.

The most powerful and flexible global deduplication architecture is variable-length data deduplication technology that spans an entire cluster rather than simply a single node, resulting in significant savings across the entire storage footprint. With variable-length deduplication, the size is not fixed. Instead, the algorithm divides the data into chunks of varying sizes based on the data characteristics. The chunks are cut in a data-dependent way that results in variable sized chunks and results in greater data reduction than fixed-size deduplication. The efficiency benefit of variable-length deduplication compounds over time, as additional data is retained.

Integrated data compression adds a boost. Compression works well on a single file, but across files, there is a need for some macro-level data compression. Why? Because when two identical copies of a file are stored, compression can individually compress the files while deduplication can completely eliminate the need to store any data for the second copy. So adding compression to the deduped data further reduces data size. This works by finding small byte patterns common between the deduplicated blocks. Based on the type of data being ingested, compression can provide no benefit for encrypted or random data or up to 5–10x compression for common log files. Deduplication ratios for VMs, databases, file shares all lie somewhere in between that range.

Modern scale-out NAS provides tremendous data reduction benefits to enterprises—for both dedupe only and dedupe combined with compression.

For organizations modernizing data centers through infrastructure consolidation, deploying a capacity efficient architecture based on modern data reduction technology can accelerate the achievement of sustainability goals by reducing data center footprints. In addition to cutting carbon emissions, IT teams can boost security and better defend against ransomware attacks by making their attack surfaces smaller.

> " *Cohesity's cloud-native approach to merging file and object services promises to be more scalable, manageable, and cost-effective than traditional NAS storage. We believe this approach is the way that storage needs to evolve to support today's increasingly complex, multi-cloud data management requirements.* "

- **Steven Hill,** Senior Analyst, Applied Infrastructure and Storage Technologies, 451 Research

Convergence and Expanding Use Reduce Operating Expenses

Modern, scale-out NAS is converged for greater efficiency. It features a single pane of glass for complete data visibility no matter where files are stored—in the data center, at the edge, or in the cloud. It also eliminates time-consuming and costly forklift upgrades. With a modern approach, IT staff can easily embrace cloud storage and eliminate downtime while more efficiently addressing regulatory compliance and governance requirements.

An ideal way to future-proof investments, modern, scale-out NAS like some of today's most efficient cars also extends hybrid engineering. Uniquely architected to support hybrid cloud and on-premises deployment, a modern, scale-out NAS platform is:

- **Software-defined for maximum flexibility** – It goes beyond NAS, enabling data management use cases from backup and recovery to target storage using one platform while supporting all modern and existing workloads.

- **Capable of running apps on the same platform** – It runs anti-virus protection, file audit, and content search on the same foundation with no additional infrastructure required.

- **Stronger data protection** – It integrates cybersecurity through multi-layered, built-in security and compliance with unified permissions to prevent, detect, and analyze threats, lowering risk profiles.

> " *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. 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.* "

\- **John White,** Chief Innovation Officer, Expedient

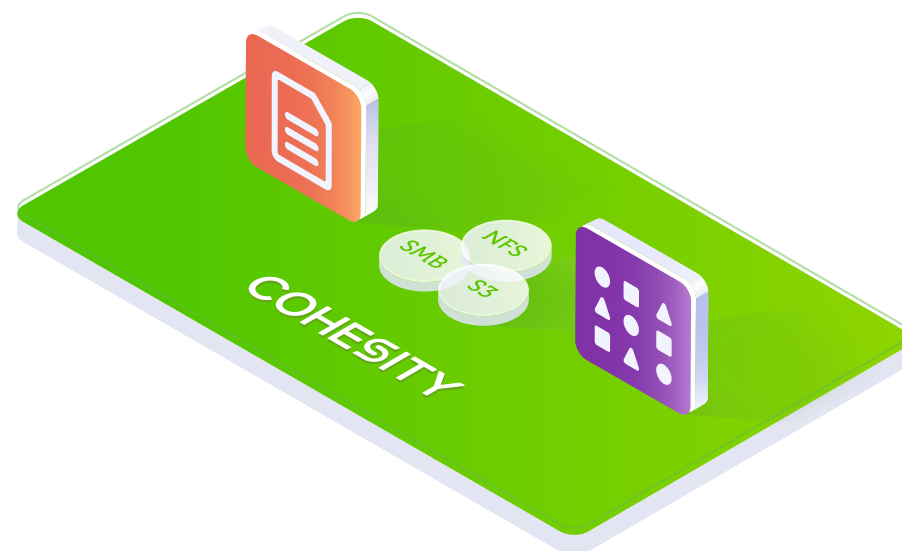| From Vintage to Vogue | The Journey to NAS | Under the Hood of Traditional NAS - 1 | Under the Hood of Traditional NAS - 2 | The Right File Model for Digital Business | Lower TCO Starts with Improving Data Reduction | Eliminating Hardware, Reducing Data Center Footprints | Convergence and Expanding Use Reduce Operating Expenses | Enjoy the Efficiency and Luxury of Cohesity SmartFiles |

Enjoy the Efficiency and Luxury of Cohesity SmartFiles

When organizations don't care how much they pay to go far or fast, an inefficient vehicle works. But in a fast-paced, ever-changing business environment, efficiency matters. Modern, scale-out NAS isn't a nice to have. It's an IT efficiency must have. That's why there's Cohesity SmartFiles.

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

Learn more about how efficient data reduction can save your organization time and money, freeing your IT staff for innovation.