COHESITY

Version 1.0
October 2019

# Architecture Matters

## *Cohesity vs Legacy NAS Scale-Out Limits*

**ABSTRACT**

*This white paper looks at the scale-out capability of Cohesity DataPlatform and the benefits of a scale-out architecture without the design limitations of legacy NAS scale-out solutions. This is important to consider for consolidation projects, large scale deployments, and for ensuring that an architecture has scale to address future strategic plans.*

# Table of Contents



# Tables

# Overview

According to projections from the International Data Corporation (IDC), 80 percent of worldwide data will be unstructured by 2025. For many companies, the unfettered growth of data has reached that critical mass already. Unstructured data creates a unique challenge for organizations wishing to use their information for analysis. The data can't easily be stored in a database and has attributes that make it a challenge to search for, edit, and analyze. The challenges of managing unstructured data are part of the reason why this is such an important topic.

If your organization is struggling to manage their unstructured data now, the problem is going to get exponentially harder over time. Unstructured data comes in the form of file shares, backups, archives, logs, media files, test/dev, and analytics. Traditional network-attached, or scale-up, storage was designed more than 10 years ago for these use cases. As a result, existing NAS storage is optimized for only a specific subset of unstructured data workloads and has limited scale. This results in multiple storage silos, running on proprietary hardware, with different software and licensing.

With the advancement in applications, traditional NAS platforms are not able to keep up with the needs of the ever-growing data. Although some of the traditional NAS platforms can scale-out in capacity, they are still limited by their architecture. Cohesity SpanFS™ is designed for web-scale architecture, where all the components within Cohesity DataPlatform scale linearly. Cohesity SmartFiles is architected to meet your current and future data needs.

# Need for Limitless Scaling

There are countless use cases where arbitrary limits affect the design and capabilities of storage systems. There is also a need to provide for unknown scalability requirements in the future, so having greater-than-necessary scale today provides future-proofing for tomorrow. Three such cases are video surveillance, home directories, and high file-count applications.

## Video Surveillance

Companies rely on 24×7 video streams with higher and higher resolution and more frames per second to better monitor their business. Coupled with broader camera deployment, the enhanced video quality of such streams gives security personnel a more detailed, comprehensive view of the areas under surveillance.

The higher resolution and frame rates of such 24×7 video streams can significantly increase the capacity requirements. Just one continuous video stream at 30fps/1280×1024 image resolution can fill a 4 TiB file in only 32 days with MPEG-4 compression. On another scale-out NAS vendor, Vendor X, the max file size 4 TiB regardless of the size of the cluster, so one would need to account for this file limitation as part of the design.

## Home Directory

Vendor X has a limit of 20,000 user quotas per cluster. In an environment with a large number of home directories, such as a university or hospital, a new cluster will need to be deployed when nearing the 20,000-quota limit. The new vendor X cluster will add additional management overhead, as well as contribute to mass data fragmentation. In addition, the new cluster will be unable to take advantage of any storage efficiency associated with data stored on the original cluster.

## High File-count Applications

Companies have a variety of applications which result in large number of files being stored in a single directory. For instance, an engineering group could set up an NFS export for scratch space for shared development efforts. Engineers would leverage the export to store temporary files. On vendor X's solution, the best practice is to not have more than 100K files in a single directory. As a result, additional management would be required to watch the shared scratch space and to create jobs to constantly ensure the number of files stays below the limit.

# SpanFS: File System Designed for Storage at Web Scale

**SpanFS** provides the foundation for Cohesity's SmartFiles NAS and Object features. Legacy storage has evolved in two very different directions. Traditional enterprise storage focused on providing standardized file interfaces (NFS and SMB) on 'scale-up' hardware, and snapshots for resiliency. And cloud storage, developed by hyperscale companies like Google and Amazon, focused on delivering scale-out solutions on commodity hardware, with strong resilience to hardware failures, but relying on proprietary protocols and APIs for data access.

## Span Everything

To enable enterprises to take back control of their data at scale, Cohesity has built a completely new file system: SpanFS. SpanFS is designed to effectively consolidate and manage all data, including backups, files, objects, test/dev, and analytics data, on a web-scale platform that spans from the edge to the cloud.

SpanFS is designed to span everything, including physical and logical constructs:

- **Physical**:
  - **Nodes**: SpanFS provides unlimited scale across many hyperconverged nodes. SpanFS is completely distributed and doesn't have a master node. It scales linearly and dynamically rebalances data as nodes are added or removed. It provides always-on availability, non-disruptive upgrades, and a pay-as-you-grow consumption model.
  - **Storage Tiers**: SpanFS spans across tiers of persistent storage technologies, including SSD, HDD, and even remote cloud storage. SpanFS effectively utilizes the most appropriate tier, based on IO profiles.
  - **Geographic Locations and Cloud:** SpanFS interconnects remote offices, branch offices, core data centers, and public clouds into one data fabric. SpanFS can be used to build a multi-cloud data fabric and span data across multiple private and public clouds.
- **Logical**:
  - **Workloads**: SpanFS supports data protection, files, objects, test/dev copies, and analytics data. It supports all the key capabilities required by these use cases, including globally distributed NFS, SMB, and S3 storage, unlimited snapshots, global deduplication, encryption, replication, global indexing and search, and good performance, for both sequential and random IO.
  - Namespaces: SpanFS can divide physical storage pools into separate shared namespaces (or Storage Domains) that have common data reduction, availability, and/or archive policies.

- Tenants: SpanFS supports multiple tenants (or partitions) with strong QoS capabilities, data isolation between tenants, separate encryption keys, and role-based access control.

**SnapTree™:** In legacy storage, snapshots form a linked chain, with each link containing the changes from the prior snapshot. Every time a new snapshot is done, an additional link is added to the chain. As the chain grows, the performance overhead required to access the data increases proportionally because the file system must traverse the chain to access the data. Hence snapshots introduce performance overhead and are limited in scale.

SnapTree introduces a completely new approach to managing metadata at scale and enables SpanFS to provide unlimited snapshots with no performance overhead. SnapTree is based on a B+ tree metadata structure, but adds multiple innovations, including:

- Distributes the tree across nodes.
- Provides concurrent access from multiple nodes.
- Supports the creation of instantaneous clones and snapshots.
- Garbage collects unreferenced nodes in the background, using Map-Reduce.
- Ensures consistent access performance regardless of the number of snapshots and clones.
- Stores only one value per leaf node, as opposed to multiple values in traditional B+ trees. This avoids unnecessary snapshotting of multiple values.
- Supports a variable fan-out factor that increases further down the tree. This avoids making any given sub-part of the tree too hot, while at the same time keeping tree-balancing costs low.

With SnapTree, Views (volumes) and files are represented by a tree of pointers to the underlying data. The root node represents the View or individual files. The root node points to some intermediary nodes, which in turn point to the leaf nodes, which contain the location of the chunks of data. Customers can take snapshots of entire Views (volumes), or individual files within the Views. As snapshots are taken, the number of hops from the root to the leaves does not increase. **Unlike legacy NAS vendors, customers can take snapshots as frequently as desired – without ever experiencing performance degradations**.

While many traditional NAS vendors claim to be scale-out, they have limitations that contradict what it means to provide limitless scale-out.

# Data Management at Scale

Cohesity DataPlatform was created to deliver value beyond scale-out NAS in many ways:

- **Scale without boundaries**. Unlimited everything: resource pools, snapshots, capacity, open and active connections, file size/name/path length, and performance linearity at scale.
- **Storage efficiency**. Sliding-window variable deduplication, compression, small file capacity optimization, and data deduplication across multiple storage silos.
- **Multi-tiered data management**. Transparent data tiering — on-premises to the cloud. Also, tiering from other vendor storage, to SmartFiles and to the cloud.
- **Enterprise search**. A simple "Google-like" search across VMs, backups, data centers, clouds, and multiple sites.
- **Enterprise richness**. NFS, SMB, and S3 with unified permissions – same data accessed by all protocols, writable snapshots, API-driven, failover/failback, DR (disaster recovery) automation, and WORM (write once read many).
- **Powerful cybersecurity**. Software-based encryption, multi-factor authentication, FIPS 140-1 and 140-2, integrated anti-virus protection, and audit log analysis with anomalous access detection.

With this data-centric storage platform, IT teams are no longer required to deploy costly external infrastructure for anti-virus (AV) protection, as this is native within the Cohesity DataPlatform. There is also no additional management burden as the integrated AV environment automatically grows with Cohesity DataPlatform.

Cohesity DataPlatform is the future of enterprise file and object services. It's not just the ability to run anti-virus, file audit, and content search on the data platform – it's also boundless scale, leading capacity efficiency, robust cyber-security, and transparent heterogeneous data tiering that transcends traditional scale-out NAS appliances.

# Tested Limits

The following table compares the tested limits between Cohesity DataPlatform and Vendor X.

The test environment for Cohesity DataPlatform was a 4-node C2500 running Cohesity DataPlatform 6.1 or newer.

Table 1: Tested Limits — Cohesity DataPlatform vs Vendor X

| ITEM | COHESITY DATAPLATFORM | SCALE-OUT VENDOR X | LIMITATION IMPACT |
|---|---|---|---|
| Number of Files Per Directory | 3.2 Million | 1 Million<br>100,000 (Best Practice) | Vendor X states a tested limit of 1 million files per directory, but their best practices states to constrain the number of files in any one directory to one hundred thousand. |
| Max Path Length | 131092 | 1023 | Vendor X limits the path length to 1023 bytes. In the case of multi-byte character sets, the path limit will be reduced. The limit in bytes will force customers to consider directory structure. |
| Max File Size (non-Sparse) | 70 TiB | 4 TiB | On Cohesity DataPlatform, the max file size is based on the physical size of the cluster. In this test, the cluster was 4 nodes. If the cluster size was doubled to 8 nodes, the max file size would scale to 140 TiB. |
| Max Number of Views/Shares | 800,000 | 80,000 (SMB)<br>40,000 (NFS) | In a high share environment such as home directories, a customer would need to deploy multiple clusters, which create additional silos of data to match the number of shares available on a Cohesity cluster. |
| Users Quotas per Directory/ View | 100,000 | 20,000 | In a home directory environment, such as a university, the number of available user quotas for use between student and faculty would need to be monitored. |

| ITEM | COHESITY DATAPLATFORM | SCALE-OUT VENDOR X | LIMITATION IMPACT |
|---|---|---|---|
| Max Number of Snapshots | 1 Million | 20,000 | Vendor X limits the snapshots to 20,000 per cluster, but further limits the number of snapshots to 1,024 on a directory. On Vendor X, if a customer was required to preserve nightly snapshots on a directory, they would only be able to apply that snapshot policy for about 3 years |

**NOTE**: These tests were run on a Cohesity cluster with 4 nodes. On Cohesity DataPlatform, the max file size is only limited by the physical size of the cluster. If the cluster size was doubled to 8 nodes, the max file size would scale from 70 TiB to 140 TiB.

# Document Version History

| VERSION | DATE | DOCUMENT HISTORY |
|---|---|---|
| 1.0 | October 2019 | Original Document |

# Your Feedback

Was this document helpful? Send us your feedback!

# ABOUT COHESITY

Cohesity ushers in a new era in data management that solves a critical challenge facing businesses today: mass data fragmentation. The vast majority of enterprise data — backups, archives, file shares, object stores, and data used for test/dev and analytics — sits in fragmented infrastructure silos that makes it hard to protect, expensive to manage, and difficult to analyze. Cohesity consolidates silos onto one web-scale platform, spanning on-premises, cloud, and the edge, and uniquely empowers organizations to run apps on that platform — making it easier than ever to back up and extract insights from data. Cohesity is a 2019 CNBC Disruptor and was named a Technology Pioneer by the World Economic Forum.

Visit our website and blog, follow us on Twitter and LinkedIn and like us on Facebook.



*2000025-001-EN*