COHESITY

The Essential Guide to Modern Data Management

Learn How Cohesity Can Accelerate Your Digital Transformation Journey

Table of Contents

The Roadblocks to Innovation

Data is a critical element of innovation, but in practice, few organizations are using their data as a strategic asset. Many IT teams struggle simply to meet basic SLAs for protection, and availability, let alone leverage their data for competitive advantage.

Surveys have revealed an alarming range of issues associated with managing enterprise data that impact both business and IT: budget overruns, poor customer service, security, compliance exposures, and even sinking morale within overworked operations teams.

The underlying cause of these issues is a phenomenon we call mass data fragmentation, and it has several profound aspects:

Infrastructure Fragmentation

Data has exploded in volume and been scattered across a myriad of locations from multiple public cloud environments and data centers to remote offices and the edge, often with little global oversight. At each of these locations, data is isolated in specialized infrastructure for functions such as backup, disaster recovery, network storage, archiving, dev/test and analytics, often from multiple vendors.

To make matters worse, there are silos within silos. For example, a single backup solution can require several dedicated infrastructure components, such as backup software, master and media servers, target storage, deduplication appliances and cloud gateways—each of which may hold a copy of a given data source. Each infrastructure component may come from different vendors with their own user interface and support contracts. It is not unusual to find four or more separate configurations simply to perform backup for different data sources.

Operational Inefficiencies

These infrastructure silos have a knock-on impact on operational efficiency. There is typically no sharing of data between functions, so storage tends to be overprovisioned for each silo rather than pooled. Likewise, multiple copies of the same data are propagated between silos, taking up unnecessary storage space.

Operational efficiency is compromised by the need to manage multiple proprietary systems with different UIs, where each may require specialist administrators to maintain rather than operational generalists. Given today's world of "no downtime", tighter SLAs, and slimmer budgets, it is no wonder that IT teams are reporting high levels of stress and burnout as they grapple with the increasing complexity.

Dark Data

In an era where data is acknowledged as a critical asset, it is perhaps surprising that most organizations don't have detailed knowledge about the majority of their data's contents, location, owner, access history, or whether it contains sensitive information. Their data is dark. With the majority of enterprise data—up to 80 percent—residing in backup, file services and object storage, petabytes of data is routinely exempt from being classified, indexed, or tracked.

Clearly this adds considerable risk to the business. How do you prove compliance in handling PII, in light of regulations such as GDPR and CCPA? How do you detect anomalous user behavior or programmatic ransomware attacks? And from an operational point of view, how can you optimize expensive storage by deleting or archiving unneeded data when you don't know which items to keep?

The Public Cloud Misalignment

And, what about the public cloud? Yes, the rise of the public cloud has accelerated the pace of innovation. Organizations can now take advantage of the elasticity and economics of the public cloud, while reducing time to market for new applications. However, the public cloud is not a cure-all. In fact, public cloud environments don't just carry over the data fragmentation, operational inefficiencies, and dark data challenges from on-premises environments but also introduce several new challenges.

With respect to storage, the public cloud introduces new infrastructure silos for data created by your IaaS, PaaS, or SaaS apps. And protecting your cloud data is your responsibility—not the provider's. Inadequate protection can lead to missed SLAs or data loss.

Also, an organization's data is never situated entirely on just one cloud. For the majority of enterprises, their footprint straddles multiple public clouds and on-prem environments. The result is a need for data mobility between these various environments. A lack of which can lead to dramatic inefficiencies. An example is modern test data management. The advent of dev/test in the cloud adds additional roadblocks, including the misalignment of formats among on-prem and public cloud VMs. This schism between on-prem and public cloud environments, leads to manageability strains, presents a severe impediment to application mobility, and often leads to dramatic cost challenges—as more copies of data need to be created and stored in the public cloud. It's clear: the public cloud is not a cure-all, but rather introduces several modern challenges for the enterprise.

How Did We Get Here?

There has been almost no fundamental innovation from the data management industry in decades. Basic techniques for performing backup and recovery, disaster recovery, cloud mobility, NAS storage, and copy data management have evolved only incrementally and within proprietary vendor silos. Until recently, IT teams have had no option but to perpetuate the traditional hardware-centric approach which only adds more complexity, risk, and cost rather than solving underlying issues.
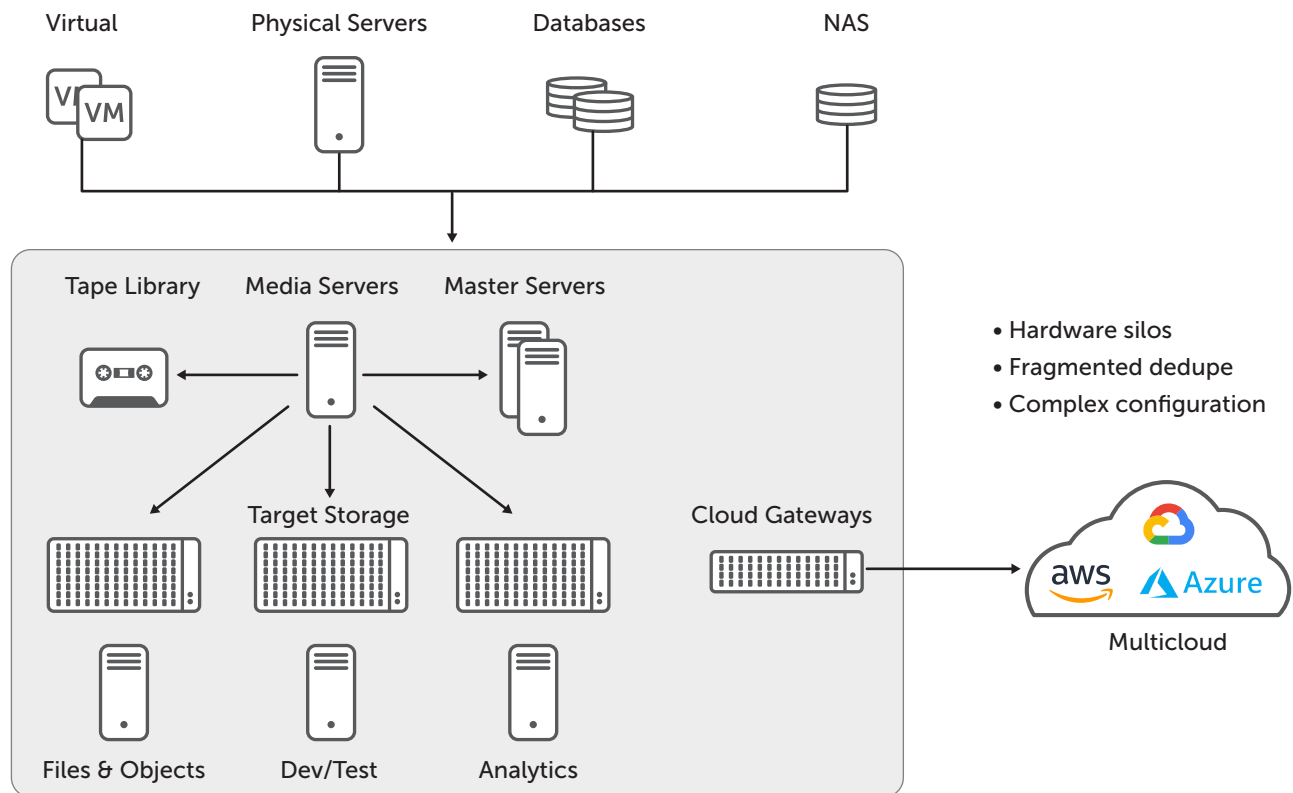
Cohesity realized that a fundamentally new approach is needed in order to break the cycle of technical debt, liberate IT and business teams from outdated architectures, and free trapped data to become an asset instead of a costly liability.

The breakthrough idea was to leverage the approach taken by the hyperscalers such as Google, Amazon, and Facebook—who manage exabytes of the world's consumer data highly efficiently and scalably—and apply the same architectural principles to managing enterprise data.

Cohesity DataPlatform: Redefining Data Management

The secret to the hyperscalers' success lies in their architectural approach, which has three major components: a distributed file system—a single platform—to store data across locations, a single logical control plane through which to manage it, and the ability to run and expose services atop this platform to provide new functionality through a collection of applications.

Cohesity DataPlatform: A Software-Defined Platform Powered by a Distributed Web-Scale File System

The Cohesity solution takes this same three-tier hyperscaler architectural approach and adapts it to the specific needs of enterprise data management. The foundation of the Cohesity DataPlatform is Cohesity SpanFS®, a 3rd generation web-scale distributed file system. SpanFS enables the consolidation of all data management services, data, and apps onto a single software-defined platform, eliminating the need for the complex jumble of siloed infrastructure required by the traditional approach.

Predicated on SpanFS, Cohesity DataPlatform's patented design allows all data management infrastructure functions— including backup and recovery, disaster recovery, long-term archival, file services

Run Apps

One UI

One Platform

and object storage, test data management, and analytics—to be run and managed in the same software environment at scale, whether in the public cloud, on-prem, or at the edge. Data is shared rather than siloed, stored efficiently rather than wastefully, and visible rather than kept in the dark—simultaneously addressing the problem of mass data fragmentation while allowing both IT and business teams to holistically leverage its value for the first time.

Table of Contents

A Single Global Management User Interface

The second major component of the Cohesity architecture is its unified control plane: a SaaS-based management interface known as Cohesity Helios, which enables all aspects of the solution to be managed globally through a single GUI. This dramatically simplifies operations by replacing multiple UIs consoles with a single global dashboard that allows policies to be set and actions taken as needed in one place. Built-in machine learning reduces administrative burden by providing proactive actionable recommendations, what-if analyses, anomaly detection, and health checks that would otherwise go unnoticed or require additional manual effort.

Uniquely, Helios also enables global search across all data/workloads/locations, cutting through previously disconnected silos to provide a single view access to your entire data estate. This allows operators to quickly locate files or VMs of interest, as well as regulated or sensitive information to ensure compliance.
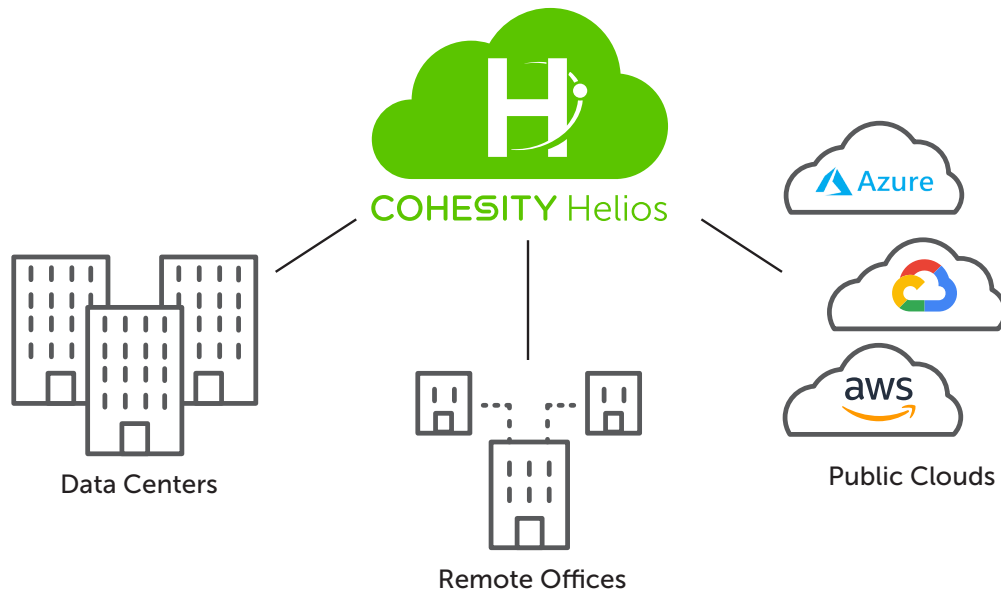
Helios: Unified Management Across Locations

The Ability to Run Applications on the Same Platform

The third major architectural component is the capability to run applications directly on the Cohesity DataPlatform. Traditionally, if organizations wanted to get more value out of their data they would make a copy of it and move that to separate compute infrastructure compounding the problem of mass data fragmentation. Based on the principle of moving compute to the data rather than the data to compute, this unique innovation amongst data management vendors flips the traditional approach on its head and enables new levels of efficiency in extracting value.

Table of Contents

Cohesity users can simply download and install certified applications from the Cohesity Marketplace and immediately run them against data previously captured and indexed from a backup or through Cohesity SmartFiles, a file and object services solution—another platform-native capability. Important data-related operations such as eDiscovery, scanning for cyber exposures, conducting compliance checks, or performing analysis through third-party apps, can now be conducted in the same environment, under the control of the same UI, with access to the full consolidated data estate in a single pass. Such an approach is an efficiency breakthrough compared to segmenting and copying data subsets to new instances of infrastructure to perform similar functions.

These innovations are made possible by the SpanFS distributed file system.

SpanFS: A Unique File System that Powers the Cohesity DataPlatform

At the heart of the Cohesity DataPlatform is a fully distributed, shared-nothing file system. Inspired by web-scale principles, Cohesity SpanFS, is meticulously architected to address the challenge of mass data fragmentation. To effectively consolidate data, enterprises need a file system that is simultaneously able to handle the requirements of multiple use cases. In order to meet modern data management requirements, Cohesity SpanFS provides the following:
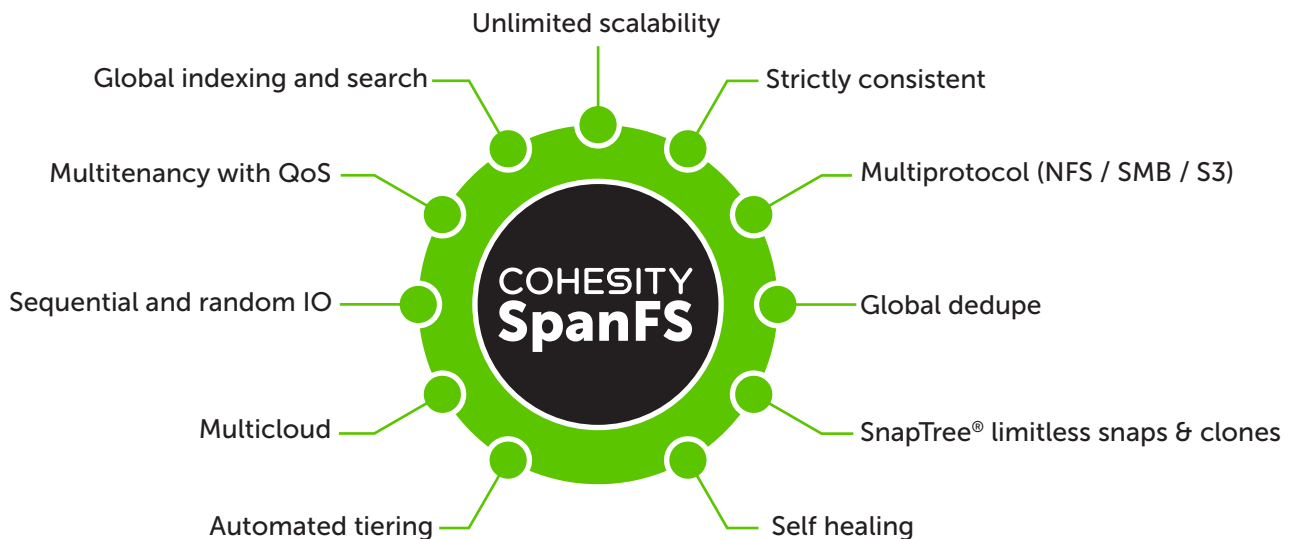
SpanFS Attribute	Implication
Unlimited Scalability	Start with as little as three nodes and grow limitlessly on-prem or in the cloud with a pay-as-you-grow model.
Strictly Consistent	Ensure data resiliency with strict consistency across nodes within a cluster.
Multi Protocol	Support traditional NFS and SMB based applications as well as modern S3-based applications. Read and write to the same data volume with simultaneous multiprotocol access.
Global Dedupe	Significantly reduce data footprint by deduplicating across data sources and workloads with global variable-length deduplication.
Unlimited Snapshots and Clones	Create and store an unlimited number of snapshots and clones with significant space savings and no performance impact.
Self-Healing	Auto-balance and auto-distribute workloads across a distributed architecture.
Automated Tiering	Automatic data tiering across SSD, HDD, and cloud storage for achieving the right balance between cost optimization and performance.
Multi Cloud	Native integrations with leading public cloud providers for archival, tiering, replication, and protect cloud-native applications.
Sequential and Random IO	High I/O performance by auto-detecting the IO profile and placing data on the most appropriate media.
Multitenancy with QoS	Native ability to support multiple tenants with QoS support, data isolation, separate encryption keys, and role-based access control.
Global Indexing and Search	Rapid global search due to indexing of file and object metadata.

AutoNation

AutoNation manages over 1.2PB of data. With Cohesity, AutoNation saved more than 60% compared with SAN, owing to a dramatically smaller data footprint as a direct result of global, variable-length deduplication and compression from Cohesity. They were able to manage a large and growing volume of data using <50% of the storage footprint required by other solutions.

# Backup and Recovery: Leveraging Data Starts by Protecting It

Running on Cohesity DataPlatform, Cohesity DataProtect is a platform-native backup and recovery application. DataProtect provides a modern backup and disaster recovery solution for a wide range of data sources and environments, including multiple public cloud providers, on-prem data centers, as well as edge locations.

DataProtect takes advantage of all the underlying DataPlatform benefits when it comes to scale, performance, and efficiency. In addition, the distributed metadata structure and Cohesity SnapTree™ combine to provide unique benefits to backup and recovery operations.

DataProtect provides some of these core benefits:

- Simplified backup and recovery operation by eliminating multiple legacy point products on a single, web-scale solution
- Supports a broad range of data sources, including leading hypervisors, relational databases, NoSQL, Kubernetes, and SaaS applications
- Reduce downtime with the ability to instantly restore files, objects, and VMs at scale with parallelized recovery and fully hydrated snapshots
- Defend backup data against ransomware with immutable backups and write once read many capabilities (WORM)

Cohesity SnapTree's B+ tree structure, combined with the ability to parallel ingest and recover, gives users the ability to recover hundreds of files, objects, and VMs instantly. Unlike legacy and other modern backup solutions, Cohesity uniquely maintains fully hydrated snapshots that can be restored within minutes. With Cohesity, recovery is rapid. Instead of waiting for the restore to complete, Cohesity mounts the data using NFS and SMB for the application to access it, while the restore happens in the background.
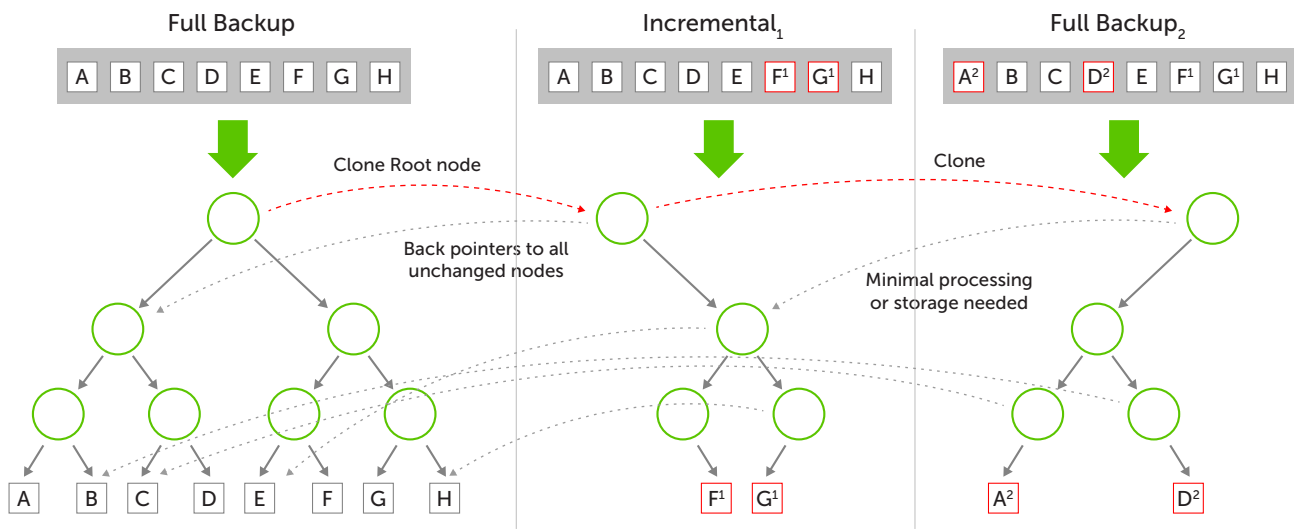
✓ **Fully hydrated snapshots instantly available with 'zero cost'**

Highly Efficient SnapTree™ Structure

Supporting All Your Backup and Recovery Data Sources

Cohesity's vision to tackle mass data fragmentation and increase infrastructure efficiency begins with a tactical focus on backup and recovery. Vestigial inefficiencies begin there and it's also a practical way to gain access to the breadth of data managed by any given organization from which to derive further value. Cohesity DataProtect supports a gamut of data sources, from virtual environments to modern databases.

Virtual and Physical

Database/Apps

Storage

Virtual Workloads

Workload Type	Supported Data Sources
Hypervisors	VMware
	Microsoft Hyper-V
	Nutanix AHV
Containers	Kubernetes

Cohesity treats hypervisors as just another service to be backed up. This means that the administrative experience remains largely the same whether you're backing up SQL server, VMware or Hyper-V.

-Brian Posey, Contributing Editor, TechTarget

Relational and Modern Databases

Workload Type	Supported Data Sources
Relational Databases	Oracle
	Microsoft SQL Server
	SAP HANA
Distributed Databases	NoSQL and Hadoop

With Cohesity, Brown University saw CapEx savings of 50% and >$75K per year in network licensing and additional maintenance cost savings.

Physical Workloads

Workload Type	Supported Data Sources
Physical	Windows, RHEL, CentOS, AIX, Solaris
NAS	Isilon, NetApp ONTAP, Pure Storage FlashBlade
Storage Arrays	Pure Storage FlashBlade, Cisco HyperFlex, HPE Nimble Storage, HPE Primera

Table of Contents

Native Cloud Architecture

Cloud adoption continues to accelerate, yet it doesn't solve all your enterprise IT problems. In fact, cloud adoption can contribute to mass data fragmentation by generating data in more locations and in different silos and formats. A common misconception is that cloud providers take care of everything for you, including backing up your data. This is far from the truth, in fact all major cloud providers employ a shared responsibility model.

Customer	Data (Protection and Security)			
	Backup	Encryption	File Systems	Applications
	Systems and Network			
	Security	OS	Patches	Network / Firewall

Cloud Provider	Infrastructure Management			
	Provisioning	Configuration	Scaling	Platform Services
	Physical Infrastructure			
	Regions	Availability Zones	Data Centers	Servers / Storage / Network Hardware
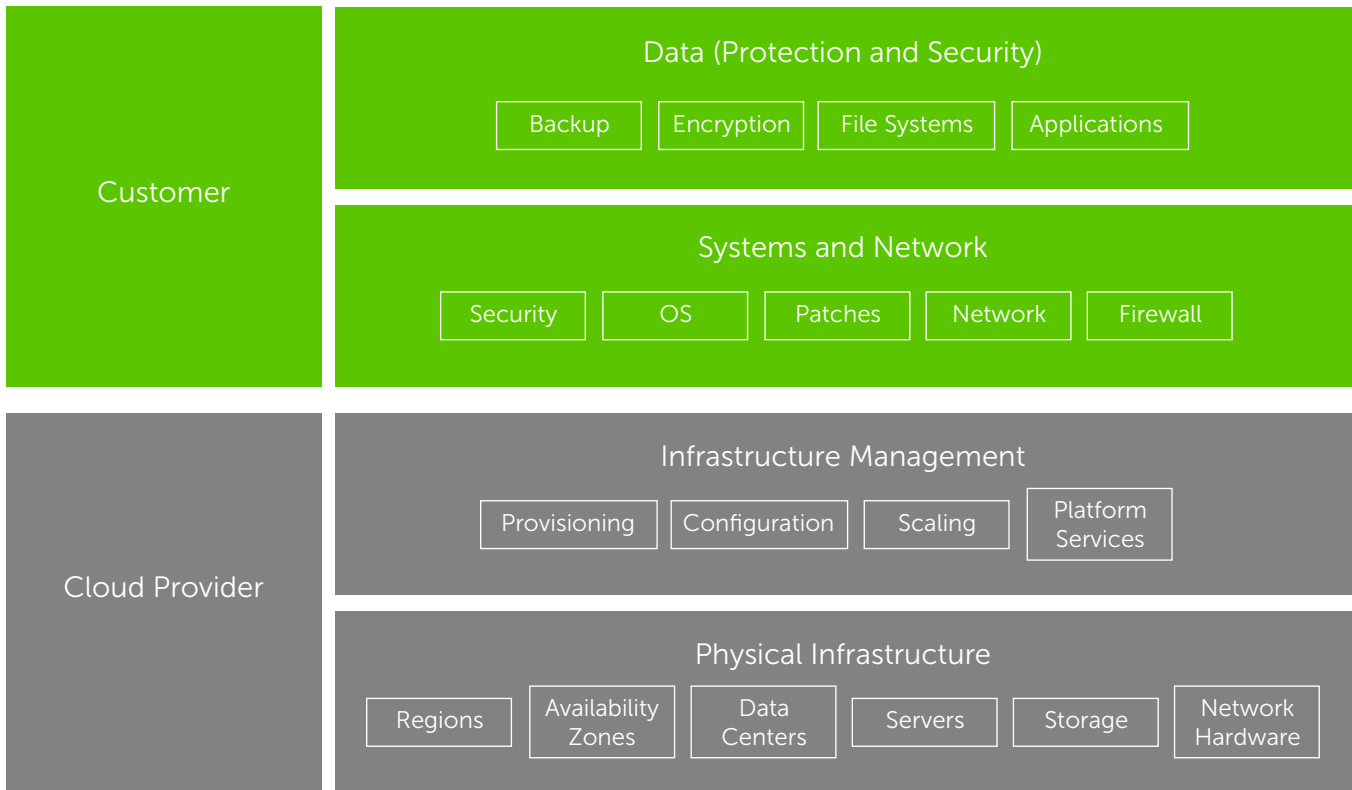
A shared responsibility model implies that they'll take care of physical cloud infrastructure, as in their data centers and servers. But when it comes to your data, it's almost solely your responsibility to secure, manage, and back up your data. If data was accidentally deleted or if a bad actor employs the latest ransomware scheme to get a hold of your cloud account, you may be out of luck if you didn't properly back up your cloud data.

In addition to running on-prem, Cohesity's software-defined platform also runs on major cloud providers including Amazon Web Services (AWS), Microsoft Azure, and Google Cloud Platform (GCP). It provides native cloud integration without the costs or complexity of bolt-on gateways. Cohesity DataPlatform not only allows organizations to protect their cloud data sources but can serve as a disaster recovery option or archival target for on-prem data. Alternatively, one can also archive data directly to cloud storage for long-term retention from on-prem with Cohesity.

Table of Contents

Easy Cloud On-Ramp: Replace Tape with Cloud Archival for Long-Term Retention

With multiple providers and cloud regions to choose from, it is easy to replace tape backups and archives with Cohesity to simplify operations and reduce costs. Cohesity CloudArchive features provide native integration to all major cloud providers using their cloud-native APIs as well as industry standard S3-compliant APIs. This allows on-prem data to be automatically and efficiently archived to cloud storage for long-term retention based on policy-based backups and data retention rules. Cohesity supports a wide range of cloud storage options from Amazon S3, Amazon Glacier, to Microsoft BLOBs, to Google Nearline and Coldline enabling you to best match your costs, redundancy, and retrieval requirements.

> " *For a film production environment, it was key to quickly and easily retrieve and restore files, and to seamlessly archive to the cloud. In our business, if we cannot restore data for our production teams fast enough, then that may mean we miss production deadlines or our film isn't launched on Netflix when we want. The Cohesity solution was key to our bottom line.* "
>
> **-Tyson Clark**, Technical Director, Air Bud Entertainment

Bringing Data Protection Everywhere You Go

It's imperative that you back up all your cloud data. Cohesity provides comprehensive backup of cloud data including IaaS, PaaS, and SaaS workloads:
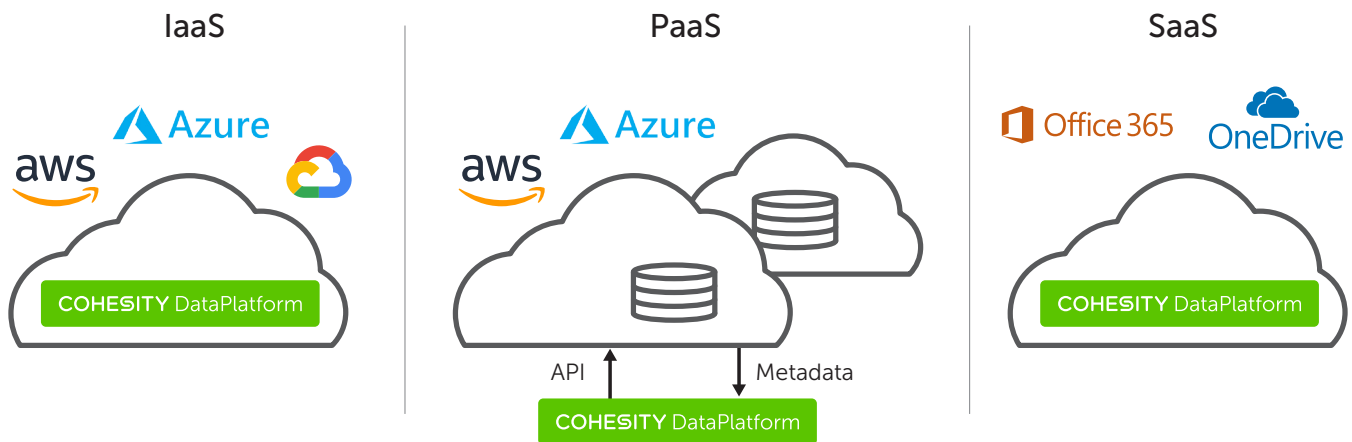
IaaS (Infrastructure as a Service) Backup

Homegrown IaaS-based applications are often VMs running in the cloud, commonly with internal or external attached cloud storage which may require additional backup. In most cases, they look much like on-prem VMs and require the operating system and application software to be set up and installed when recovering from an outage. Therefore backup should address the OS, apps, and any additional attached data storage.

Cohesity provides three different methods to back up your IaaS cloud VMs: 1) agent-based, 2) cloud-native backup, and 3) cloud snapshot manager.

1. **Cohesity protection service backup** requires a software agent to be installed on your cloud VMs. Data is protected, indexed, and efficiently stored in a Cohesity instance running in the cloud.

2. **Native snapshot backup** is a more lightweight solution that uses cloud provider APIs instead of agents to manage the backup process and to minimize the impact on production VMs. Data is protected, indexed, and efficiently stored in a Cohesity instance running in the cloud.

3. **Cloud snapshot manager** is an even more lightweight solution that uses cloud provider APIs and snapshots to manage the backup process remotely from an on-prem Cohesity instance and one running in the cloud. Data is stored directly in cloud storage to minimize data transfer costs and metadata is stored and indexed in Cohesity.
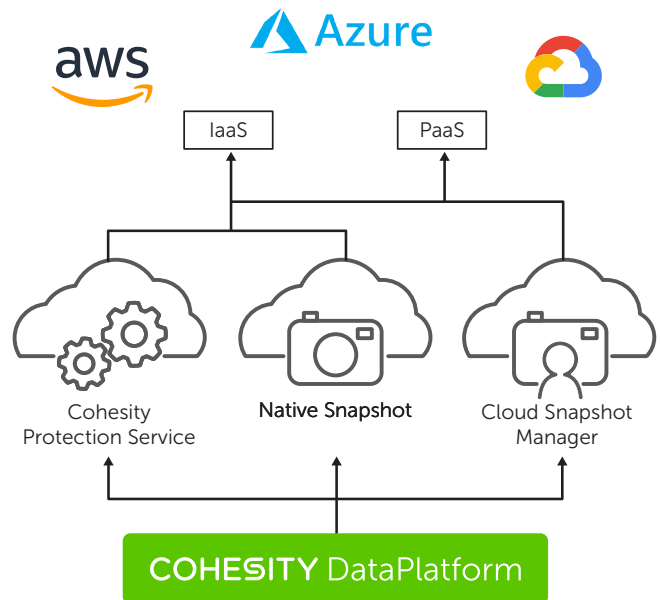
> **❝** *eSilicon was not only managing billions of files, we were also doing complete chip design and manufacturing operations in the cloud, something no other company has done before. Cohesity enabled eSilicon to complete our move to the Google Cloud for true elasticity, helping us deliver optimized solutions to our customers.* **❞**
>
> - **Naidu Annamaneni**, CIO and Vice President of Global IT

PaaS (Platform as a Service) Backup

PaaS includes platform or middleware services including analytics, serverless databases, and others. For these services, cloud providers may manage the OS and application software but the application data still needs to be backed up.

Cohesity supports the backup of PaaS and database services including Amazon RDS (Relational Database Services). Using Cohesity's Cloud Snapshot Manager capability you have a lightweight solution to back up PaaS services and leverage enterprise features like global search either from Cohesity running in the cloud or on-prem so you can remotely back up the cloud without the need of additional cloud resources.

SaaS (Software as a Service) Backup

SaaS provides a fully managed stack from the infrastructure, platform, and complete application components. Some SaaS services even provide a basic one-size-fits-all approach to backing up your data— typically for 30 days. The problem with this is you have little control over your backups and may not be able to meet your unique enterprise or compliance requirements.

Cohesity supports flexible and configurable backup and recovery of Microsoft 365, Exchange Online, and OneDrive SaaS applications and services. You gain the ability to manage your data retention beyond 30 days, perform instant global search and granular recovery for the mailboxes or files you need quickly.

Disaster Recovery in the Cloud Doesn't Have to Be Hard

On paper, cloud's pay-as-you-go model is a great fit for disaster recovery where resources typically sit idle for months on end waiting for an outage to happen or to be spun up for the occasional six month or annual test. In reality, using cloud for disaster recovery is much harder in practice due to different network environments, different data formats, and complex manual process steps required to move and spin up workloads in the cloud.
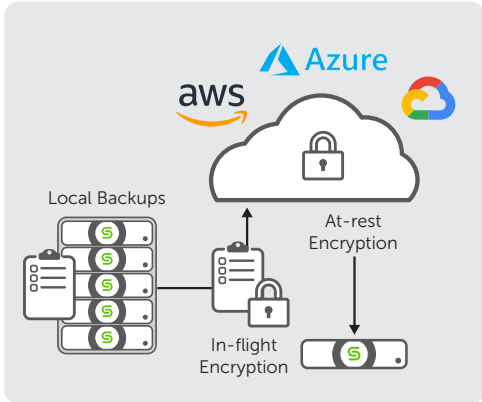
Cohesity solves these challenges with a comprehensive set of data management solutions to address your disaster recovery needs:

1. Cohesity DataPlatform runs directly on major cloud providers. This provides a single platform to unify, move, failover, and failback data between on-prem and cloud for disaster recovery. Data can be recovered and failover from an active Cohesity cloud instance and restored and spun up from archived data in cloud storage.

2. Cohesity CloudSpin provides a quick and seamless way to bridge the compatibility gap and convert on-prem VM formats from VMware and Hyper-V into cloud-native formats for Amazon EC2 and Azure VMs.

3. Cohesity Runbook ties everything together with any easy to use automation and orchestration engine that removes manual steps and complexity from the disaster recovery process. It provides a graphic design canvas to quickly design process workflows, a validation check to ensure the process works every time, and the orchestration engine that executes the entire DR process from end to end.
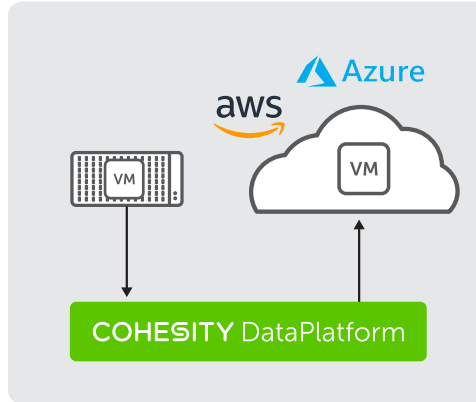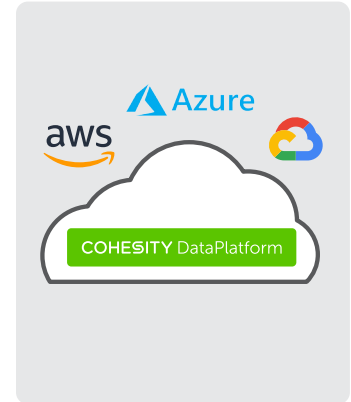
COHESITY

Cohesity Runbook – Automation and Orchestration

DR with **CloudArchive**   DR with **CloudSpin**   DR with **DataPlatform**

Days ——————————— Hours ——————————— Minutes

Recovery Time

VansonBourne

“ *34% of Organizations have experienced downtime as a result of not being able to failover to a cloud-based disaster recovery copy.* **”**

- Vanson Bourne Survey 2019

Hybrid and Multi-Cloud Mobility and Data Replication

Most organizations are choosing to implement both hybrid and multi-cloud strategies. This is a good starting point, but your hybrid or multi-cloud implementations would be limited and siloed if you were not able to easily move and exchange data between all your providers and on-prem.
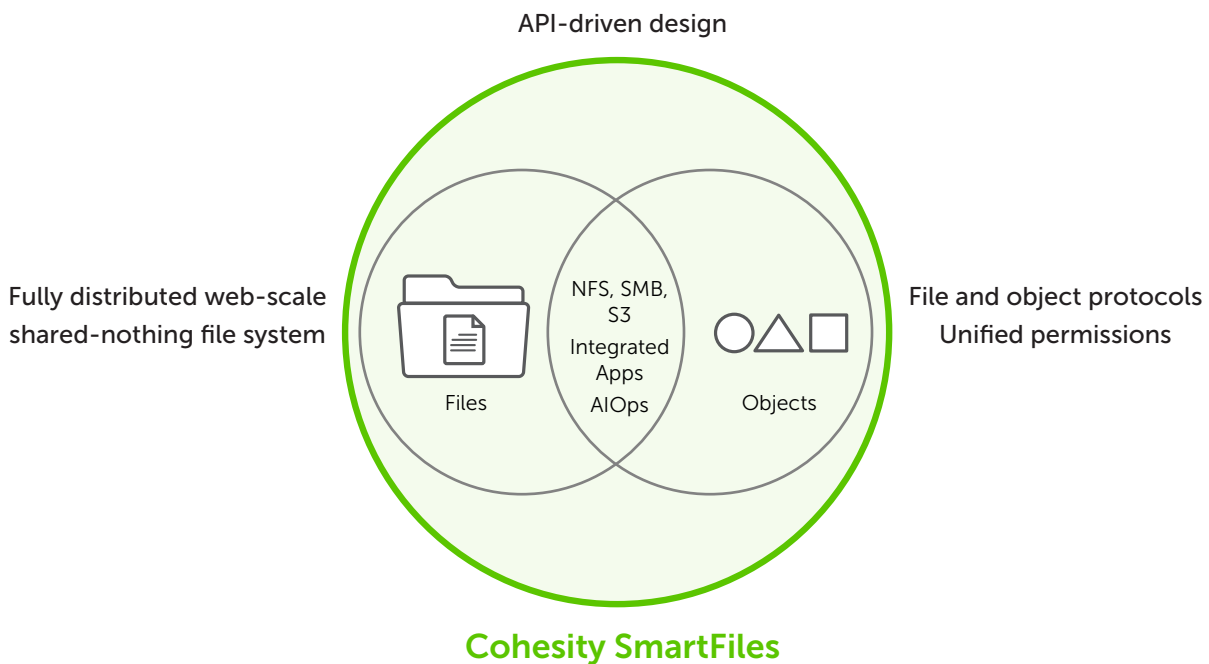
Cohesity's software-defined platform accelerates the adoption of hybrid and multi-cloud. Data can be stored and unified on one platform providing a central place to manage all your data. Data mobility is made easy using our fast and efficient Cohesity replication engine between all your cloud providers and on-prem clusters.

SmartFiles: Beyond Scale-Out NAS

Another manner in which Cohesity aims to address mass data fragmentation is by folding in direct support for unstructured file and object services. Part of the Cohesity DataPlatform, Cohesity SmartFiles provides multiprotocol file and object services with unified permissions via NFS, SMB, and S3. This combination of protocols can be used to facilitate hybrid and cloud deployment. It is built on the aforementioned SpanFS for scale, performance, fast-ingest, and support for limitless snapshots and clones.

In addition to providing general file and object services, SmartFiles includes multiple data reduction technologies to maximize the amount of data stored in each raw TB of storage hardware.

Additionally, SmartFiles uniquely takes advantage of integrated apps from the Cohesity Marketplace to support a file ecosystem without separate hardware infrastructure. This includes antivirus protection, file audit with the ability to detect anomalous file accesses, and content search inside of files. As the foundation for a data-centric approach, such search results can be used to drive eDiscovery, compliance, and other business processes.

API-driven design

Fully distributed web-scale shared-nothing file system

NFS, SMB, S3 Integrated Apps AIOps

Files

Objects

File and object protocols Unified permissions

Cohesity SmartFiles

Software-Defined Approach

As a software-defined solution, SmartFiles customers can choose the most appropriate hardware for specific file and object workloads. SmartFiles is not bound to proprietary hardware, allowing a selection of approved hardware platforms. SmartFiles also runs in the cloud with all the functionality of that of on-prem deployments.

Broad Compatibility

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 accessible via the S3 protocol, regardless of the original file protocol.

Policy-Based Management

Policies are used to automate file and object services. Machine learning is also utilized to assist with management tasks. One example is predictive capacity planning for proactive capacity management. There is also support for data migration from existing filers. The data from the existing filer can be backed up and the completed backup can be presented to users and applications as a storage volume. Operations can be automated through RESTful management APIs. Cold data on costly tier 1 storage can be automatically moved by policies to a cost-optimized tier or to the cloud. Policies can be utilized to ensure consistent management across multiple clusters and locations using a single management pane. Cloud or on-prem. Core or edge. Physical or virtual. Data movement is transparent to applications.

Google-like Enterprise Search

Data is indexed for fast search results. This applies to search of file metadata and search of file contents. Search results can be used to drive data-driven decisions and services based on actual data contents. SmartFiles search capabilities can be used to shed light on dark data, or used to drive compliance and governance requirements.

Integrated Cybersecurity

SmartFiles integrates cybersecurity with 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:** Detect infected files through Cohesity Helios' machine-driven anomaly detection.
- **Analyze:** Observe user and behavior analytics for security and compliance reasons.

SmartFiles Use Cases

There are dozens of file and object workloads and use cases that SmartFiles is designed to address. This may include the following:

- Corporate file shares
- Corporate video
- Video surveillance
- Splunk® cold buckets
- Digital archives

Table of Contents

Leverage Backups for Test Data Management: Speed Up Modern Application Development

In order to build and release software at a competitive pace, developers need high-quality test data. However, provisioning test data fast enough is a key challenge for many enterprises, taking them up to weeks to deliver test data to developers. For many companies, agile development is blocked by a lack of simplified and agile infrastructure.

Cohesity solves these challenges directly and enables IT operators to make their backup data productive by instantly and securely provisioning quality data to development and test teams through zero-cost clones. By repurposing backup up data as test data, Cohesity reduces data sprawl, costs, and accelerates the development of high-quality applications within Cohesity's runtime environment. Agile Dev and Test is a key example of getting more out of your existing data under management.

Combined with automated and on-demand data refresh, Cohesity helps development and test teams with high-quality, relevant data.

Cohesity accelerates application development and enterprise productivity by consolidating previously disparate solutions: test data management and backup and recovery. By using one platform and repurposing backups for test data, Cohesity helps:

- **Reduce development roadblocks** and allows more accurate testing—earlier in the release cycle (effectively applying shift-left principles to data management practices.)
- **Simplifies infrastructure and access** to production-quality test data, without impacting production.
- **Makes security and compliance simple** with data masking, and greater visibility and control on data.

# Helios: Manage Your Data Globally with a Single UI

Clearly, the underlying DataPlatform allows for many different data management functions traditionally found in point solutions—backup and recovery, test data management, file and object services—to be converged. At this stage you may wonder how easy it is to manage them all. So let's revisit the Cohesity approach to data management:

1. One platform to consolidate workloads and data

2. One UI to manage this data—globally

3. The ability to run apps and services on top of this platform

This second step, the ability to manage data globally, is important. If the first step to tackling mass data fragmentation is consolidating multiple workloads across different locations, gaining streamlined visibility and the ability to control data is critical. It's simple to see why. On the path to leveraging data, gaining control and visibility are foundational steps.
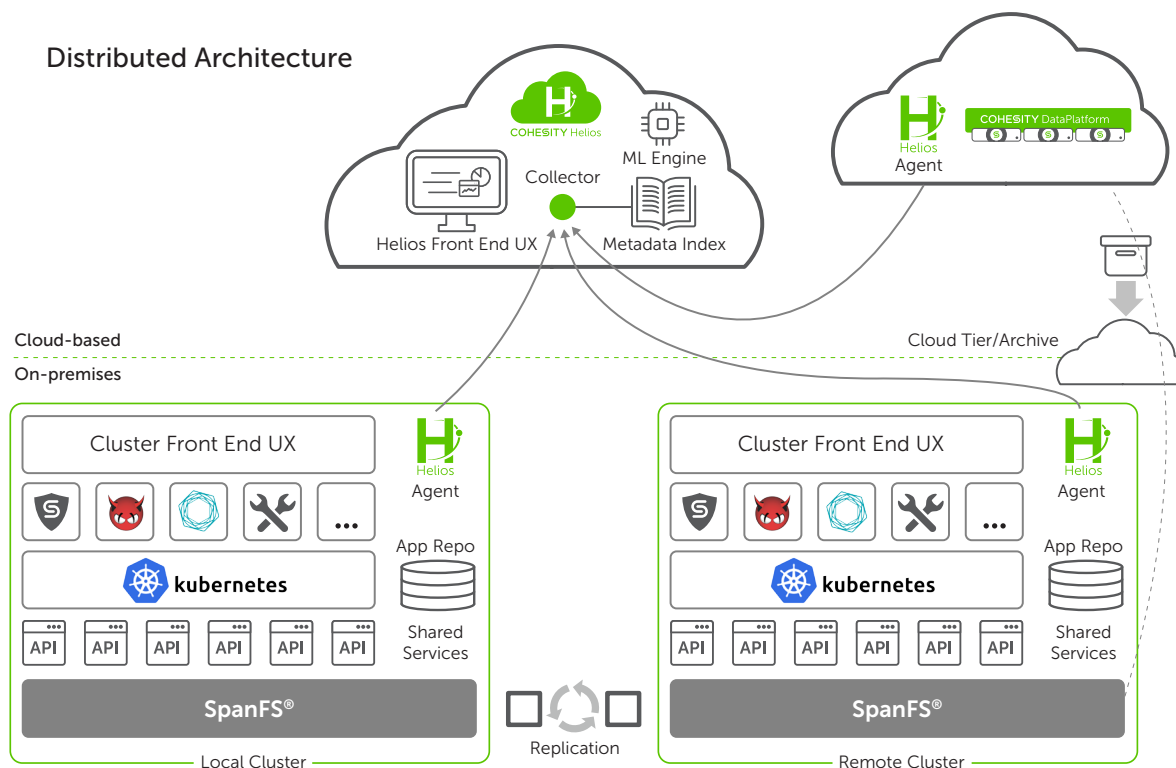
Cohesity Helios, the second major component of the Cohesity architecture, is a SaaS-based administrative console that enables all aspects of the solution to be managed globally through a single GUI. This is a single dashboard that replaces multiple dashboards and the need to manage, monitor, and operate multiple clusters. Helios is not a passive dashboard, allowing users to manage policies and clusters globally, and proactively detect and avoid downtime with built-in machine learning.

Table of Contents
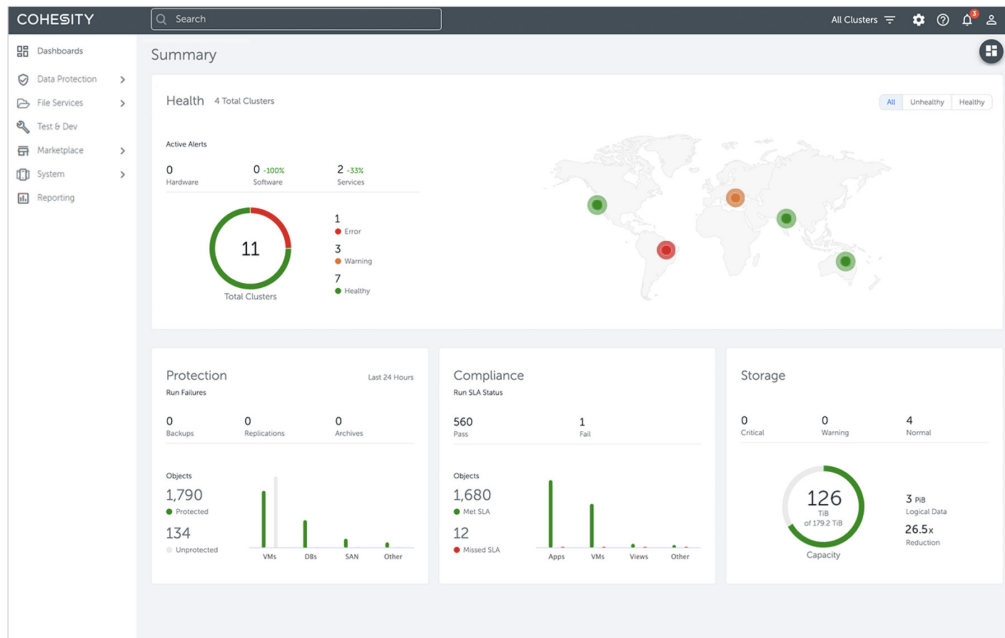
Helios provides global access and control of data across locations and predictive and actionable insights to help meet SLAs and improve business continuity. In this way, tackling mass data fragmentation doesn't stop with data consolidation. It's critical to streamline data control and management, as well.
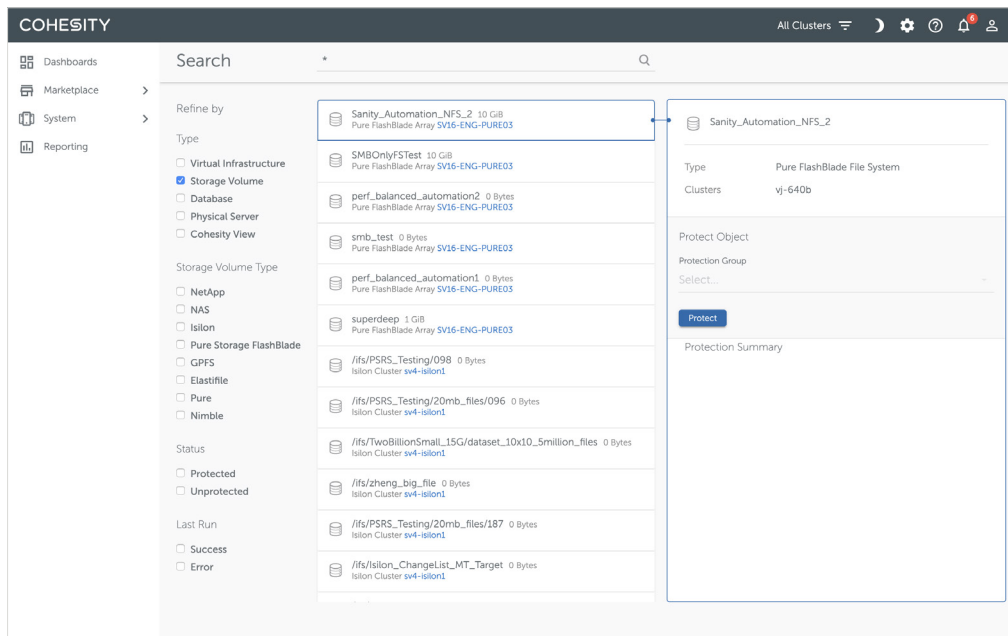
Given Helios collects metadata from distributed Cohesity instances, it also makes Global Actionable Search possible, allowing organizations to search and recover any data across any cluster, irrespective of its physical location—including native cloud snapshots. Search results return instantaneously, but also offers an actionable recommendation that puts the user in her desired workflow, giving them the keys to take action without leaving the page they're on.

Table of Contents

Further Unlock Your Data's Value: Run Apps

Instead of creating multiple copies and shipping data around, Cohesity DataPlatform makes backup and other unstructured data productive by allowing you to run Cohesity and third-party applications directly on the Cohesity cluster. The philosophy of *bringing compute to data* empowers IT operators to derive insights from previously dormant data to meet their security, compliance, and analytics requirements. Cohesity's unique architecture allows applications to access and use the data within each Cohesity cluster, providing visibility and insights into the data that previously was dark.

Users can download Cohesity and third-party applications directly from Cohesity Marketplace or build their own apps using the Cohesity App SDK to address their unique business requirements.

The overarching benefit of Marketplace and apps is that the value of your data—and your ROI—grows with time. New applications continuously allow users to unlock more value from their data, in the form of insights as an example, or more effectively secure data in the face of evolving threats and compliance regulations. Here are just some of the capabilities enabled with Marketplace:

- **Enable eDiscovery** and compliance monitoring effortlessly.
- **Uncover cyber exposures** and blind spots within your production environment by running on-demand and automated scans on backup snapshots against known vulnerabilities.
- **Automate data migration and disaster recovery** with a single click and migrate workloads much more effectively, without ever needing to leave the Cohesity experience.

New capabilities are continuously added, unlocking further value. Keep in mind, however, these are not product enhancements akin to those of other companies. Value here is not incremental and each new app dramatically helps reduce administration overhead, allows data insights to help with business competitiveness, and makes compliance and security simple.

Simplifying Evolving Compliance Challenges

With growing regulatory mandates and focus on giving more control to people over their data, the need to make data easier to find and determine its level of sensitivity or risk is increasingly important. Fragmentation of data across multiple silos makes it difficult to meet complex and evolving compliance requirements.

Cohesity provides a data classification and governance solution that makes the process of defining classification policies and discovering data both simple and automatic. Cohesity's solution incorporates the ability to choose a pre-built compliance template, for GDPR, as an example, or search and define custom classification policies directly from the same Cohesity user interface.

Cohesity's solution works across multiple data sources and clusters. With rapid search, complex and large global environments can automatically meet evolving compliance requirements. After data discovery, users have options to control data access, movement, alerting, tracking, deletion, and redaction.
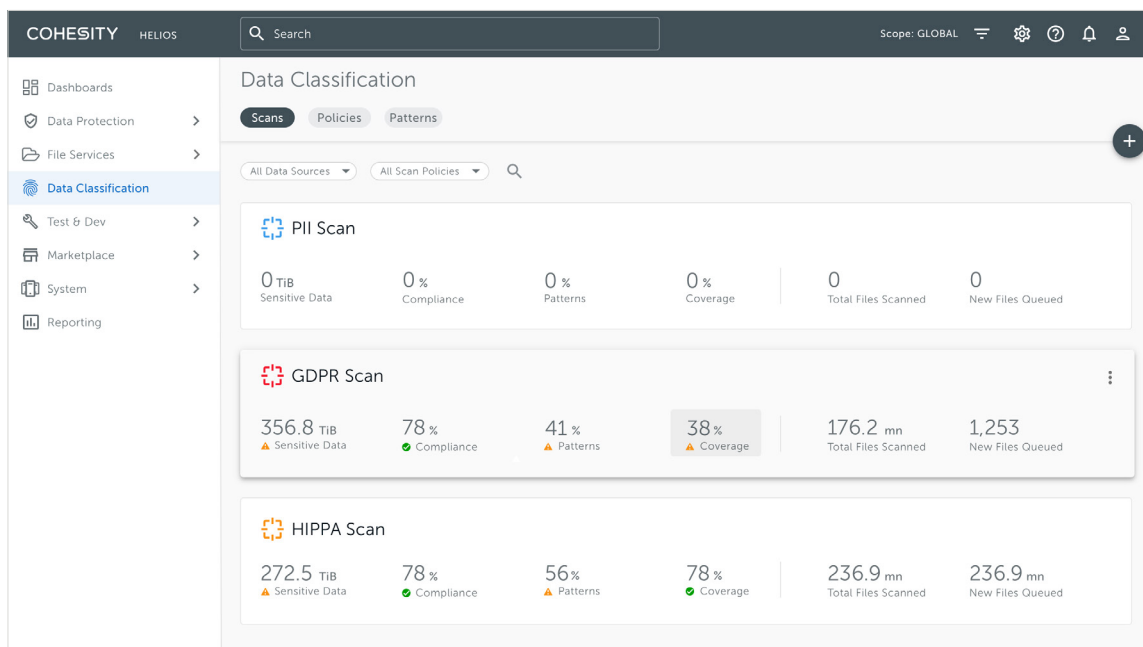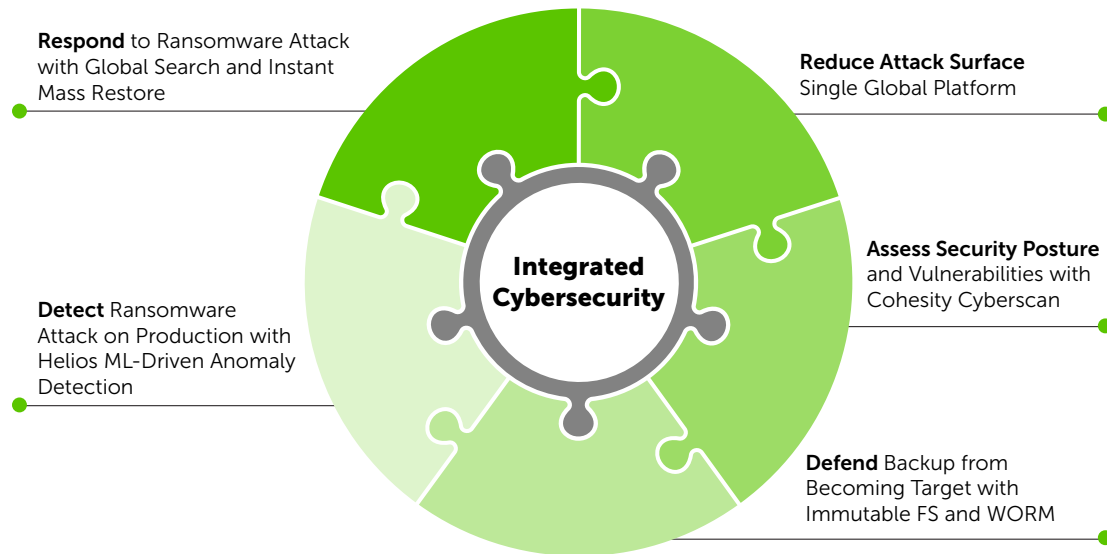
Why a Platform Approach Matters: Comprehensive Data Security

It should not come as a surprise that the impact of cyber breaches is projected to top [$2.1 trillion in 2019](#), quadrupling since 2015. Cohesity's integrated security offers a holistic approach in:

1. Protecting your backup data against ransomware attacks

2. Leveraging data to discover cyber vulnerabilities and detect attacks in progress

3. Reducing downtime with instant restore at scale

This unique set of capabilities is delivered through a combination of platform attributes combined with the ability to run Marketplace applications and is a great example of the value of taking a holistic platform approach to data management.

Respond to Ransomware Attack with Global Search and Instant Mass Restore

Reduce Attack Surface Single Global Platform

Detect Ransomware Attack on Production with Helios ML-Driven Anomaly Detection

Integrated Cybersecurity

Assess Security Posture and Vulnerabilities with Cohesity Cyberscan

Defend Backup from Becoming Target with Immutable FS and WORM

Reduce the Attack Surface: Cybersecurity is one of the key tenants of Cohesity's architecture, starting with a single software-defined data management platform that consolidates multiple point solutions and reduces data footprint. This helps enterprises to reduce their attack surface.

Capital One data theft that impacted 106M people was due to a configuration vulnerability.

Defend Backup Against Ransomware Attack: According to Cybersecurity Ventures, $11.5 billion ransom payout keeps the criminal actors motivated that are now increasingly targeting backup data. Cohesity SpanFS, an immutable file system, and WORM stops malware from modifying or deleting the backup. The industry's first WORM for backup creates time-bound protection to the snapshot that even your administrator or security officer role can't overrule.

In the last 10 months, 140 local governments, police stations and hospitals have been held hostage by ransomware attacks.

Machine Learning-Powered Detection and Actionable Recommendations: In the case that attackers hold your primary environment as hostage, Cohesity is not just able to detect a potential attack, but also point to the last known clean snapshot to initiate a response.

Deep Visibility: Recover with confidence. Cohesity helps to assess the health and recoverability status of your backup snapshot to ensure that no known cyber security vulnerabilities are reinfected into the production environment by means of restore.

Rapid Response: With unlimited fully hydrated snapshots, combined with a distributed architecture, global actionable search, and live mount, Cohesity uniquely allows IT to instantly recover hundreds of files, objects, and VMs rapidly. You can bounce back from even the most severe attacks, instantly and effortlessly.

Table of Contents

Enabling Modern Business by Redefining Data Management

IT organizations are facing unprecedented demands to not simply support business operations efficiently, but to also act as a source of innovation and competitive advantage. We believe that mass data fragmentation is often a significant roadblock to digital transformation and that more effective management of data is key to enabling IT to deliver against those expectations.

Cohesity has built a unique solution based on the same architectural principles employed by cloud hyperscalers managing consumer data, but optimized for the enterprise world. Our solution is based on a 3rd generation web-scale distributed file system, with a single UI that enables global search and control of all your data, combined with the ability to download and run apps to extract more value to help the business.

The unique capabilities of SpanFS at the heart of the Cohesity DataPlatform allows all data management functions and workloads—including backup and recovery, target storage, DR, archiving, NAS, cloud tiering and dev/test provisioning and data analytics—to be run and managed in a software-defined environment rather than in isolated silos. This effectively curtails the damaging impacts of mass data fragmentation and dramatically streamlines IT's ability to support the digital business.

If this paper has piqued your interest, we encourage you to learn more at www.cohesity.com or speak to one of our representatives at www.cohesity.com/contact/.