COHESITY

WHITE PAPER

Cohesity Data Cloud: A Unified Platform for Superior Cyber Resilience and Economic Outcomes

An executive overview

TABLE OF CONTENTS

Executive summary

The modern era of data protection and security—defined by cyber resilience and AI—demands a modern platform. The ideal modern platform is built to spec with common enterprise requirements in mind. These include speed of cyber recovery, scale to support hundreds of data sources across hybrid and multiple clouds, adherence to Zero Trust Security and other principles, ease of use, and native capabilities that encourage reuse of enterprise data for AI scenarios. All of these requirements should be part of a platform that minimizes capital expenses and ongoing operational costs.

The Cohesity Data Cloud delivers on all these requirements today. Further, its near-term evolution promises an even more compelling value proposition with the addition of NetBackup and its support for over 1,000 workloads to Cohesity's industry-leading hyperconverged platform.

This white paper describes the evolution of the Cohesity Data Cloud, specifically as it relates to the integration of NetBackup. We also detail how the portfolio offers enterprises unmatched data protection, advanced cyber resilience, and groundbreaking AI-based analytics—all at exabyte scale.

We have created a true "sum of the parts" solution. NetBackup customers will retain everything they love about that application, while also benefiting from the unique capabilities of the Cohesity Data Cloud.

History has examples of quintessential moments where innovative elements are combined to create a best-in-the-world platform. The P51 Mustang, arguably the best piston-based fighter aircraft of all time, was a combination of the P51 Airframe (with its laminar flow wing and Meredith effect radiator) and the Rolls-Royce Merlin engine (with best-in-class power/performance). As you will see in this paper, we are using the file system of the Cohesity Data Cloud as the proverbial Merlin engine of the evolved platform. It will bring world-class capabilities, which we will describe in detail.

Introduction

> **“Cohesity is committed to futureproofing our customers’ investments . . . this means ongoing support for all Veritas NetBackup, NetBackup Appliances and Alta Data Protection . . . for many years to come”.**
>
> Sanjay Poonen, Cohesity CEO

The data protection industry evolved from its roots in compliance and auditing to support disaster recovery and business continuity. Now, market demands have matured further. We are in a new era defined by two themes: cyber resilience and AI.

In this white paper, we describe how Cohesity—the pioneer in hyperconverged data protection and now the leader in AI-powered data security—continues to innovate to help organizations meet the moment.

We will detail the enduring architectural advantages of the Cohesity Data Cloud, our flagship data platform. Today, this platform delivers world-class cyber resilience for the largest brands, including 85% of the FORTUNE 100. We’ll also examine what the future holds for enterprise IT leaders as the NetBackup application becomes more deeply integrated into Cohesity Data Cloud. Finally, we’ll review the platform attributes that enable advanced AI scenarios by reusing high-quality enterprise backup data stored in the platform.

Desired Outcomes

What compels IT leaders to adopt a hyperconverged platform like the Cohesity Data Cloud? A key driver is the concern and anxiety about the risk and cost associated with their status quo. A fragmented, siloed data estate is costly to manage daily and is notoriously difficult to protect and secure in the face of changing cyber threats.

Organizations that modernize with the Cohesity Data Cloud often achieve superior outcomes in five key areas that we call the 5 S's:

- **Speed** - They can recover from cyberattacks much faster than their previous systems.
- **Security** - They improve their security posture, detect threats, protect data, and rapidly recover from cyberattacks.
- **Scale** - They can secure and protect their entire data estate on a single platform, even at petabyte scale.
- **Simplicity** - They can run their data estate and perform backup and recovery workflows from a unified control plane and set of APIs.
- **Smarts** - They gain business and operational insights from their data, with advanced AI capabilities.

The Anatomy of the Cohesity Data Cloud: Today and Tomorrow

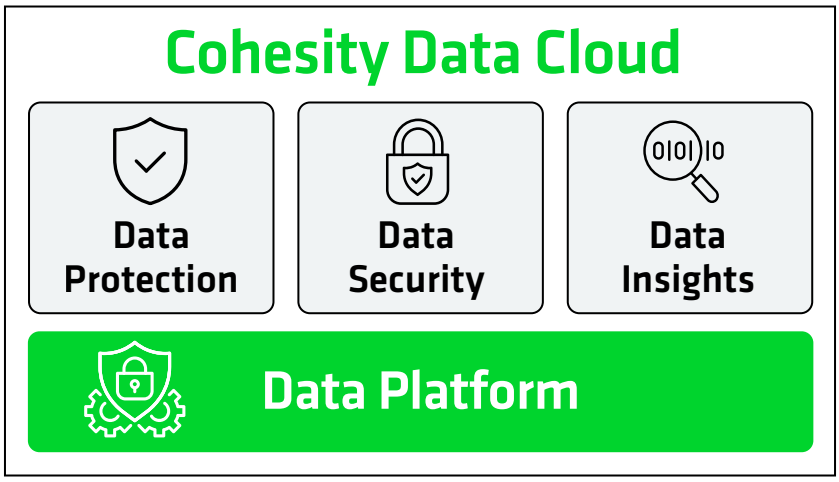
Supporting these aforementioned business outcomes was the goal when Cohesity's founding engineers built the industry's first hyperconverged, web-scale platform.

The Cohesity Data Cloud is designed to effectively consolidate and manage all secondary data, including backups, files, objects, test/dev, and analytics data, on a web-scale platform that spans from the edge to the cloud.

The most noteworthy aspect of this new platform is its file system: SpanFS®

As the name suggests, SpanFS is designed to span everything:

- **Scale**: SpanFS provides unlimited scale across multiple nodes. SpanFS is completely distributed and does not have a single point of failure. It scales linearly and dynamically rebalances data as nodes are added or removed. It provides always-on availability, nondisruptive upgrades, and a pay-as-you-grow consumption model.
- **Speed of cyber recovery:** SpanFS helps deliver rapid recover from cyberattacks, up to 10x faster than other solutions. The Cohesity Data Cloud always maintains an unlimited number of fully-hydrated backup snapshots that can be instantly mounted, making your data readily available (via direct mount) when you need it, while restoring data to production locations in the background. This near-instant data recovery ensures business operations are restored quickly, after successfully responding to a cyber attack.
- **Private and public cloud**: SpanFS manages data across private data centers and public cloud sites. Its web-scale approach allows for deployment in the public cloud, which can be used for archival, tiering, or replication. For replication, SpanFS is deployed in the data center or public cloud to manage data in supporting multiple use cases.
- **Storage**: 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 dedupe, encryption, replication, global indexing and search, and good performance for both sequential and random operations.
- **Tenants**: SpanFS supports multiple tenants with strong QoS capabilities, data isolation between tenants, separate encryption keys, and role-based access control.



- **Media tiers**: SpanFS spans across SSD and HDD media tiers and uses the most appropriate tier based on IO profiles.

Our engineering teams are adding an OST interface to SpanFS to complement these capabilities. OST (Open Storage Technology) is a protocol within NetBackup that opens the broadest array of storage options in the industry, in a native fashion without special data handling. This OST interface to SpanFS will support its integration with NetBackup.

The SpanFS system also includes several built-in security capabilities, including:

Encryption of data at rest and in motion

Cohesity Data Cloud encrypts all data at rest and data flows within the platform. Encryption prevents unauthorized users from viewing data outside of the platform. Data stored in the platform is unintelligible unless accessed and decrypted by an authorized user or process.

Immutable data storage

Data backed up by the Cohesity Data Cloud will never change from its saved state. Our underlying file system provides immutable backup snapshots to prevent modification, and the premature or accidental deletion of data. Based on a hyperscale architecture, Cohesity stores backed-up data in our secured file system that is inaccessible from outside a Cohesity cluster. The backup snapshots are stored in a read-only state. No external application or unauthorized user can modify the snapshot.

Access control: Based on Zero Trust principles

As defined by the National Institute of Standards and Technology (NIST), Zero Trust is "… the term for an evolving set of cybersecurity paradigms that move defenses from static, network-based perimeters to focus on users, assets, and resources." Zero Trust principles in the context of the Cohesity Data Cloud focuses on validating the authenticity and authorization of users for any access or changes to the platform.

Multifactor authentication (MFA)

MFA provides strong authentication of users to thwart unauthorized changes to the platform settings or data. MFA improves platform security by requiring users to identify themselves by more than a username and password. Passwords and usernames are susceptible to brute force attacks and can be stolen. MFA requires the user to authenticate login requests with a response only they can provide (such as a mobile phone challenge) or Time-based One-time Password (TOTP). Cohesity supports native MFA or third-party MFA providers such as Ping, Duo, Okta, and more

Role-based access controls (RBAC)

Granular role-based access control in the Cohesity Data Cloud enables organizations to grant the least privilege required for users to execute their job requirements, minimizing risk and keeping areas outside their responsibilities unreachable. Organizations can restrict Cohesity user roles to specific applications, capabilities, or workflows in the platform, limiting what users do based on their role and responsibilities. For example, organizations can restrict specific users to only perform backups or data discovery.

Quorum

Cohesity Data Cloud uses quorum features to prevent unilateral changes to the platform within administrative accounts. This crucial control protects against unintentional user error, rogue admins, or compromised accounts. With quorum, user requests to change settings or administrative functions require multiple approvals.

Auditing

The Cohesity Data Cloud maintains an audit trail for all actions performed on the Cohesity cluster. These records provide proof of compliance and operational integrity. Audit trails can also identify areas of noncompliance by providing information for audit investigations. Audit logs capture user activity for login/logout, changes to data or the data's properties, and job scheduling. The platform organizes logs by categories, such as Active Directory or Cluster, for rapid analysis.

A Look to the Future

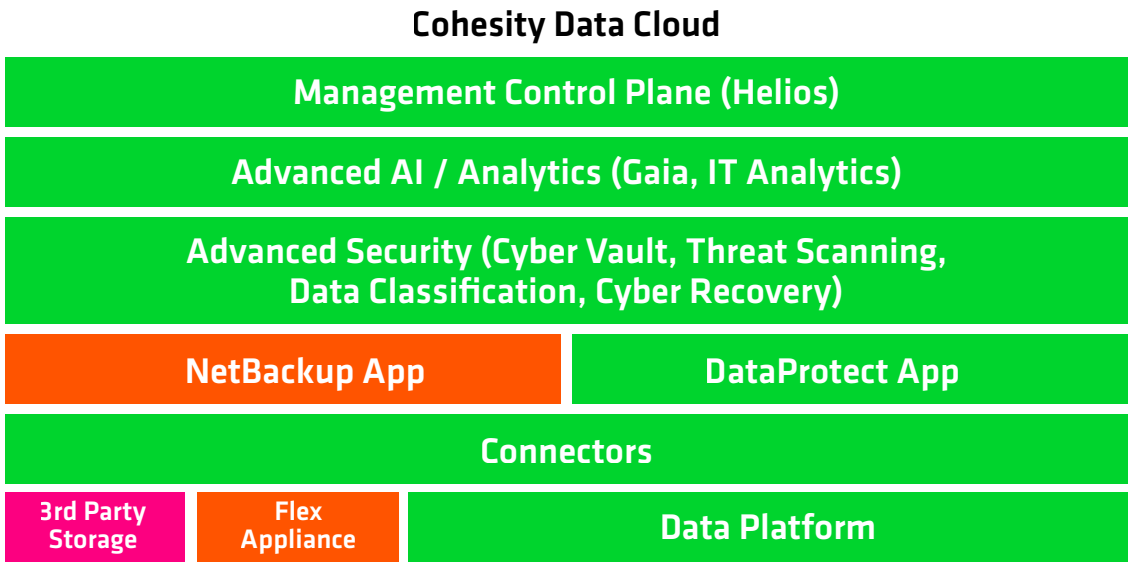
With the business outcomes in mind—as well as the file system and foundational security capabilities, it's worth unveiling what the Cohesity Data Cloud will look like when NetBackup is fully integrated into its substrate and shared services.

Data Protection Applications

NetBackup was the pioneering backup and recovery application, initially launched in 1987. Since then, the product has won praise from thousands of customers and earned "Leadership" status in the Gartner Magic Quadrant 19 times. The application supports over 1,000 data sources and can write to an additional 1,000+ targets. In fact, it is the only application that supports many crucial data sources inside the world's largest brands.

Meanwhile, Cohesity created the hyperconverged, web-scale data protection industry with its world-class file system. Its DataProtect product features unmatched speed of cyber recovery and performance at petabyte scale. That's part of the reason why the hyperconverged approach is used in so many of the FORTUNE 500.

> **"Veritas wrote a world-class backup application, while Cohesity built a world-class platform around its unique file system, which also happened to perform backup and recovery. This provided us with the opportunity to combine them, using the Cohesity as the underlying storage system for the NetBackup application."**



The management control plane (Helios) governs both NetBackup and DataProtect. Further, you can see that data insights modules can draw from data stored in both places. Finally, advanced security capabilities can be used with the data protection apps to provide a robust cyber resilience strategy.

At the data protection tier, you'll notice that NetBackup customers have multiple options. First, NetBackup can continue to write to disaggregated storage devices like Data Domain. Additionally, Cohesity continues to invest in Flex appliances, so existing customer investments are protected. Finally, with the upcoming OST integration, NetBackup can write to SpanFS and use all the benefits of the Cohesity Data Cloud.

Note the "API External Integration" box on the left-hand side. Cohesity Data Cloud has an API-first approach to help customers integrate their Cohesity deployment with other enterprise systems, such as upstream applications (provisioning, infrastructure-as-code tools, CDM) and downstream apps like ITSM, SIEM/SOAR, and billing.

Advanced Data Security

Foundational data protection capabilities were sufficient for compliance and auditing scenarios. They continued to be valuable for disaster recovery and business continuity situations as well. But in the age of destructive cyberattacks, organizations need more than the basics—they need full-featured cyber response and recovery features. Cohesity and the newly-added Veritas team have a strong track record of innovation here. Our work is focused on ensuring security workflows can be applied atop both DataProtect and NetBackup applications.

There is so much to say about our combined security capabilities that we'll create a white paper and evolution presentations focused solely on this. Cohesity and Veritas bring a complementary set of security features. We plan to integrate them into a single set of standard security functions across DataProtect and NetBackup, as the combined platform evolves.

Both DataProtect and NetBackup customers will enjoy an industry-leading portfolio of security applications that cover the gamut of required capabilities. This includes platform security, cyber vaults, proactive ("peace time") and reactive ("war time") threat scanning, clean room, orchestration, data classification, and Data Security Posture Management (DSPM). Moving to a common file system and data structure will allow us to provide these security capabilities on a single code base. Our common file system approach is a key component of our security approach.

AI & Analytics

Cohesity now manages over 100 exabytes of data—far more than every other vendor in the category combined. In this era of AI, all this data is in high demand from enterprise development teams. So, how does Cohesity help?

The Cohesity Data Cloud is a modern platform ready-made for the AI era. Your high-quality backups are indexed and stored in a powerful file system, purpose-built for the AI services that are taking the world by storm. Cohesity Gaia, our AI product built atop our platform, is a RAG (retrieval augmented generation) service that lets you have a conversation with your data.

In addition to these data insights, Cohesity Data Cloud features "AI-powered operational insights." Here, the platform uses embedded AI to continually improve performance and efficiency. Your infrastructure and InfoSec teams can use these features to run your data estate more intelligently and use AI to investigate and remediate potential threats faster.

Finally, IT Analytics offers observability with deep insights through a single console to locate, analyze, and correlate data across your multicloud environments. Use this reporting module to optimize cost, keep tabs on the health of your storage environments with monitoring and alerting, and reduce your risk by identifying unprotected data.

The Superior Economics of Cohesity Data Cloud

Cohesity Data Cloud delivers superior total cost of ownership compared to all other alternatives. Adding NetBackup to the platform compounds the economic upside. Among the key advantages you will enjoy:

- Greater operational efficiency, because you can protect and secure more petabytes of data per operator, thanks to Cohesity's scale and performance;
- Lower data center expenses with respect to space, cost, and power, because of Cohesity's highly efficient hyperconverged hardware;
- Retire superfluous add-on point products, due to the platform's built-in security capabilities;
- Optimize your storage costs with the platform's best-in-class data deduplication and compression;
- Ongoing lower operational costs with flexible deployment models that match your business requirements.

A recent IDC study found that a cohort of Cohesity customers achieved these business value metrics:

- Three-year ROI of 233%
- 6-month payback period
- 52% reduction in tool-related costs
- 39% more efficient IT infrastructure teams
- 36% more efficient security and backup teams
- 6% more productive compliance teams

The Road Ahead

Architectural Evolution

So far, we have described the resulting value proposition of the unified platform. The next logical question: "How do we get there?"

We aim to perform this evolution with minimal disruption, so that it's nearly invisible to you. After all, you still have a data estate to protect, and you must continue to support your business with backup and recovery operations.

To address this, we've created distinct, thoughtful approaches to give you a seamless path to the unified platform. We call these approaches ***Evolution*** and ***Revolution***. Several principles guide our effort to deliver the most compelling outcome to you.

Let's step through these one at a time.

Principle 1: Enduring Hardware Support: No backup and recovery hardware systems will be "stranded" or "orphaned" by this evolution. Every piece of hardware that you've purchased or will purchase will be usable for NetBackup, DataProtect, or the unified platform until the end of life of the hardware. This is true for heritage Veritas hardware (appliances and OST partner hardware), Cohesity white label, and partner-branded hardware for data nodes.

Principle 2: Deployment of Choice: If you like your deployment model, you can keep your deployment model. Many customers have embraced the simplicity of appliances. Some prefer the build-your-own approach, while others opt for the hyperconverged deployment approach. We continue to offer a choice of deployment to meet the needs of any data center–on-prem or cloud–in a cost- effective and elastic model. ***We do note, however, that those storage options underpinned by SpanFS will provide customers with enhanced capabilities over those without it.***

Principle 3: Evolution Peace of Mind: Both data center and cloud needs evolve over time, and we are evolving along with them. We want to bring you along with us rather than force customers into change on a specific timeline. A NetBackup customer might want to introduce YARA-based scanning as an example, or perhaps they'd like to deploy a Digital Jump Bag™ in a cyber vault. They may also want to experiment with hyperconverged architectures, using node-based clusters, as opposed to heritage Veritas appliances, or distributed topological approaches. For these customers, we will publish a phased evolution plan, with measured and thoughtful migration steps. We will couple this with operational support and services capabilities, working with you every step of the way to ensure that this evolution adheres to the demanding SLAs associated with your deployments.

A side benefit of the evolutionary approach is that we'll provide a single operational interface, analytics tools, and data security capabilities that will be available across both product platforms. Over time, we'll merge the data and finally the control planes for these discrete platforms, and they'll represent a single deployment of the combined platform.

Principle 4: Revolutionary Acceleration: Other customers are intrigued by the Cohesity hyperconverged approach and are eager to try it. They can lower their TCO with this architecture, achieve greater scale, and experience smoother growth and upgrades. For these customers, we will offer a set of tools and services to support a revolutionary deployment approach. We'll enable these customers to deploy net-new workloads on DataProtect, and/or to migrate existing workloads and data to DataProtect.

Evolution Focus Areas

This unification will be done in five architectural focus areas, most of which are actively happening in parallel. Customers will start realizing the benefits early. Throughout the process, Cohesity will continue to support you and empower you to meet your SLAs to your stakeholders. While the work is happening in parallel, our customers have asked us for a phasing description to help them understand what major changes happen in what order. We present that here, with the caveat that these "phases" will overlap dramatically with each other. With that in mind, here are the elements of the **Cohesity Five Phase Evolution Plan:**

1. Unified Management Plane

The first step is the most visible one—to provide a single management interface. DataProtect administrators already know and love the familiar Cohesity UX ("Helios"). In this phase, we bring NetBackup functions into Helios. Upon completing this phase, you will enjoy a single management console for both NetBackup and DataProtect deployments.

2. Unified Data / Storage Plane

SpanFS integration is key to unlocking the strengths of NetBackup with the web-scale architecture of SpanFS. All the functionality described above can use SpanFS scale and resilience. Unlike InfoScale storage management, SpanFS will offer additional operational benefits. Data virtualization in SpanFS ("Views") is inherent to the architecture. It will dramatically improve the scale of operations (like instant access and universal shares) by intelligently balancing the cost of these features across an entire SpanFS cluster. Further, it will increase concurrent operations at scale. In addition, NetBackup will enumerate both NetBackup views and **SmartFiles** views of the same backup sets. (SmartFiles is a discrete mechanism for interacting with SpanFS.) This way, the solution can process NetBackup data using DataProtect security analysis capabilities.

SpanFS storage integration with NetBackup OST provides unique capabilities when compared to traditional integrations with third-party systems. To describe this, we'll highlight the key features of OST, followed by a description of the differentiated attributes of the integration.

NetBackup Open Storage Technology (OST) is the industry-standard for writing deduplicated backup data and performing optimized copy operations to third party storage arrays. Without OST, these arrays can only replicate (FIFO) data between arrays. OST provides the granularity to use native dedupe data movers to make backup copies efficiently and manage retention independently. This copy service alone drives down backup storage costs. (Secondary copies are often retained longer than the primary copy. Such requirements cannot be accomplished with Volume Replication Policies.)

In addition to optimizing copy services, OST provides:

- Accelerator – Patented Dedupe Preprocessor that avoids dedupe processing backup selection data that have not changed since last processed.
- Accelerator for VMware – Patented Dedupe Preprocessor the operated on VMDK Snapshot data.
- Optimized Synthetic Protection – Leverages the knowledge of the dedupe store to enumerate backup images from existing stored dedupe data without data movement. Optimized Synthetic Protection uses a full selection from the protected source, and assembles a backup image using only deduplicate data in the backup storage and without any backup data read or write costs.
- Instant Recovery for VMware – presents VMDK for Startup or Incrementally recovers a VMDK to a previous point in time.
- SDK - enables third-party vendors to write OST "Plugins" to enhance their value with NetBackup capabilities.
- In addition to OST functionality, the SpanFS-OST implementation will provide these exclusive capabilities:
- Client Direct Deduplication everywhere NetBackup Agents are present. Here, Client Direct establishes a connection from the agent directly to the storage server. This minimizes the need for Media Server Resources to process backup data, thereby lowering cost.

- Intelligent Stream Handling for repeatable deduplication preprocessing efficiency. This will apply across workload and backup stream types including VMware, Hyper-V, AHV, NDMP, non-encrypted and encrypted databases, and different file system types. Intelligent Stream Handlers unravel a stream of data and determine the boundaries unique to the stream. This ensures that the dedupe engine works efficiently and consistently each time the same data type is processed.
- Instant Access leverages SmartFiles views of backup images, via file system NAS shares. Instant Access provides a direct way to explore protected data in a backup image without recovering the data with a NAS share of the content.
- Universal Shares use "Thin Provisioned" SmartFiles scale-out NAS shares for application dump and sweep operations. Application owners use this mechanism to control, protect, recover, and validate their protection requirements without any knowledge of the backup application. This, in turn, drives down NAS costs.

The integration with SpanFS will be 100% OST compliant and will add MSDP unique value and functionality. **This is a superset of OST that is superior to every other third-party OST compliant dedupe solution**.

The combined capabilities of NetBackup Data Reduction Intelligence and the virtualization and scale capabilities of SpanFS will provide an even stronger solution for customers that can host both Data Protect and NetBackup workloads on the same data storage plane.

3. Unified Security Protection

Here, you will enjoy 100% consistent security capabilities that work on both NetBackup and DataProtect deployments. Let's detail a few use cases that will be fulfilled in this phase. You search for Indicators of Compromise (IoCs) with a set of YARA rules across both applications. You can orchestrate dev/test functions for both products in a similar manner. Finally, you can send alerts from both applications downstream to your SIEM/SOAR systems to expand your threat detection capabilities. All of this will be delivered as a part of Phase 3.

4. Capability Convergence - Basic

In this phase, we converge DataProtect and NetBackup capabilities onto the combined end-state platform. This will be mostly invisible to many of our customers, as they'll already be using both NetBackup and DataProtect applications from a single UI. Still, this is an important step to aggregating all the key assets from both product portfolios.

This is the first point where a customer can operate both NetBackup and DataProtect applications on a single backup target if the target supports SpanFS.

5. Capability Convergence - Advanced

In this phase, we bring the remaining NetBackup microservices to the integrated platform and allow you to run a single, unified platform for all your deployments. There are architectural considerations we are still working through for this phase. For example, our engineering teams are still designing the NetBackup archive/restore from tape capability. This is critical for organizations that have decades worth of tape storage. We need to offer a way for them to access this tape from the combined platform.

Conclusion

The world of data is diverging, not converging. Complexity will continue to grow for a few familiar reasons: an exploding volume of data across data centers, clouds, and edge locations; there are more data sources than ever; the regulatory landscape continues to add more requirements; and the threat landscape of destructive cyberattacks continues to morph in unpredictable ways.

Cohesity Data Cloud is a popular enterprise platform to deliver superior outcomes despite this complexity. The platform's value proposition gets even more compelling in the coming months with the addition of NetBackup: greater speed of cyber recovery, greater scale and performance for the most data sources, advanced security capabilities built in from the ground up, industry-leading simplicity and ease of use, and embedded AI capabilities to streamline operations and accelerate data reuse scenarios across your organization.

A Note about Forward Looking Statements

This document includes forward-looking statements that are subject to risks, uncertainties, and assumptions. You should not rely upon forward-looking statements as predictions of future events. All statements other than statements of historical fact could be deemed forward-looking. Forward-looking statements include statements concerning new or planned products and features or service availability, and technological developments.

Although we believe that the expectations reflected in the forward-looking statements are reasonable, We cannot guarantee that the future results, performance or events reflected in the forward-looking statements will be achieved.

Any unreleased services or features referenced in this document are not currently available and may not be made generally available on time or at all, as may be determined in our sole discretion. Any such referenced services or features do not represent promises to deliver, commitments, or obligations of Cohesity, Inc. and may not be incorporated into any contract. Customers should make their purchase decisions based upon services and features that are currently generally available.

About the authors

Tim Burlowski

VP of Product Management

Tim Burlowski serves as a Vice President of Product Management at Cohesity and is responsible for data protection and resilience, including roadmap & strategy for NetBackup. Tim joined Veritas in 1998 and has spent most of his career focused on protecting the world's data by improving product quality, simplifying customer experience, and increasing scale and resilience for NetBackup and NetBackup Appliance customers.

Jim Tavares

Senior Director, Security Solutions

Jim Tavares leads the Security Center of Excellence at Cohesity—a world-class team with diverse backgrounds in Security, Backup/Recovery, and IT. Previously, he worked for VMware, where he deployed 5G public and private cloud-based networks in the U.S., Europe, and Asia. Before that, he was a longtime Cisco veteran, where he held leadership roles in Product Management, Solutions Development, Services, Strategy, and Channel Management. Jim has undergraduate and graduate degrees in engineering from the University of Pennsylvania and an MBA from Rutgers University.

Learn more at Cohesity

COHESITY

cohesity.com
1-855-926-4374
2625 Augustine Drive, Santa Clara, CA 95054

2000060-001-EN 5-2025