

ESG Lab Spotlight

Cohesity DataPlatform: Instant VM Restore at Scale

Date: July 2017 Author: Kerry Dolan, Senior IT Validation Analyst

Abstract

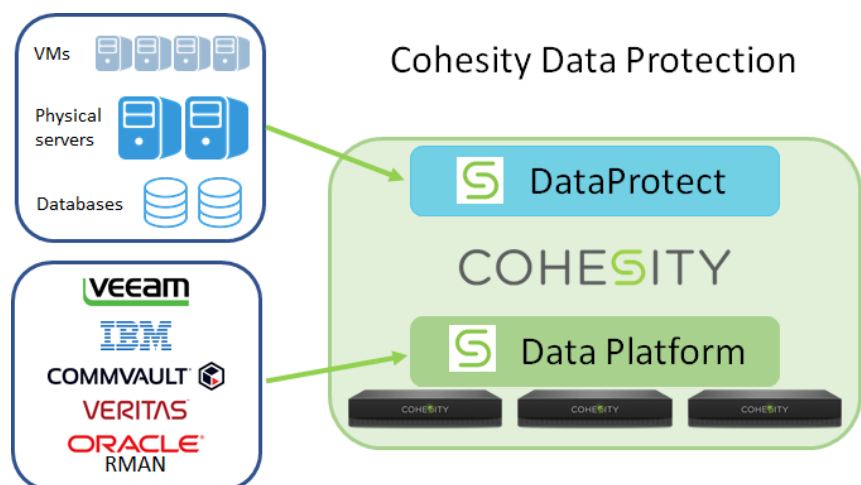
ESG Lab performed remote testing of a significant feature of the Cohesity DataPlatform: its ability to instantly restore a large number of virtual machines (VMs) to any previous recovery point, with no change as the number of VMs increases. Backup copies are always fully hydrated and instantly available. This feature enables Cohesity to deliver the web-scale protection and recovery that today's agile, dynamic organizations need to remain productive, while reducing both storage and management costs.

The Challenges: Balancing Protection, Uptime, and Cost

Respondents to an ESG survey of midmarket and enterprise organizations reported that more than half of their production workloads (52%) demand less than one hour of downtime.¹ This cannot be achieved with traditional data protection schemas, which consist of daily full backups that take hours, if not days, to recover. Incremental backups can offer shorter backup windows with more recovery points, but restores require that the backup application stitch back together the daisy-chain of incremental changes since the last full backup, a time-consuming task that leads to a performance penalty over time. Creating an occasional synthetic full backup with the most recent incremental changes provides one image ready to restore, but typically only that most recent one is available for restore. In addition, prolific secondary storage growth occurs as organizations create multiple copies for data protection, as well as discrete information silos for protection, test/dev, analytics, etc. These silos consume more storage capacity, add management complexity, and are disruptive to upgrade.

The Solution: Cohesity DataPlatform

Cohesity DataPlatform is an intelligent, scale-out, software-defined platform that is deployed on hyperconverged nodes with compute, flash, and hard disk capacity. Cohesity provides fully converged data protection with DataProtect software, a complete backup and recovery solution. Cohesity can also be used as a target for third-party backup applications including Veeam, Commvault, and NetBackup, as well as dedupe storage for database dumps and copies taken with native database tools such as Oracle RMAN. Administrators use



DataProtect to manage backup jobs, establish policies, and apply these policies to groups of servers or VMs; DataPlatform itself manages global deduplication and compression. Cohesity eliminates the complexity of deploying and scaling separate media servers, master servers, proxies, media agents and target storage; DataPlatform scales simply by adding hyperconverged nodes to the cluster, each of which brings CPU and storage capacity, expanding the global filesystem. It is tightly integrated with VMware, including the VMware APIs for Data Protection (VADP) that provide Changed Block Tracking (CBT) so only incremental changes are transferred. Cohesity DataPlatform includes unlimited snapshots that can be created

¹ Source: ESG Research Report, [The Evolving Business Continuity and Disaster Recovery Landscape](#), Feb. 2016.

on-demand; instant restore at scale; global deduplication across the cluster; sub-five-minute RPOs; a simple UI for end-to-end workflows; policy-based automation; linear scalability; granular VM-, file-, and object-level recovery; AES-256 encryption; and integration with tape libraries and the cloud for long-term archiving. Once data is backed up on Cohesity, it can be used for other use cases such as spinning up zero-cost clones for test/dev and analytics.

Snaptree

Cohesity’s SnapTree functionality enables a virtually unlimited number of snapshots that are always fully hydrated and ready to restore. With each new backup, Cohesity clones the most recent backup image and then applies the changed blocks; all backups in the catalog are immediately available as snapshots for instant restore. SnapTree provides unlimited clones because the global file system creates a series of intermediate pointers using a B-tree structure; as a result, it takes a fixed number of hops to access any block, regardless of when the snapshot was created or how many intermediate snapshots have been created.

Instant VM Restore

A significant feature of the Cohesity platform is instant restore of a virtually unlimited number of VMs and applications, to any recovery point. Because the Cohesity DataPlatform can present an NFS mount point for any file, every backup image is immediately and directly accessible to all ESX hosts. Instead of instantly restoring a few VMs like most traditional backup solutions, administrators can instantly restore an unlimited number of VMs, or even an entire vCenter. The PCIe flash in each Cohesity node ensures high performance for restored VMs.

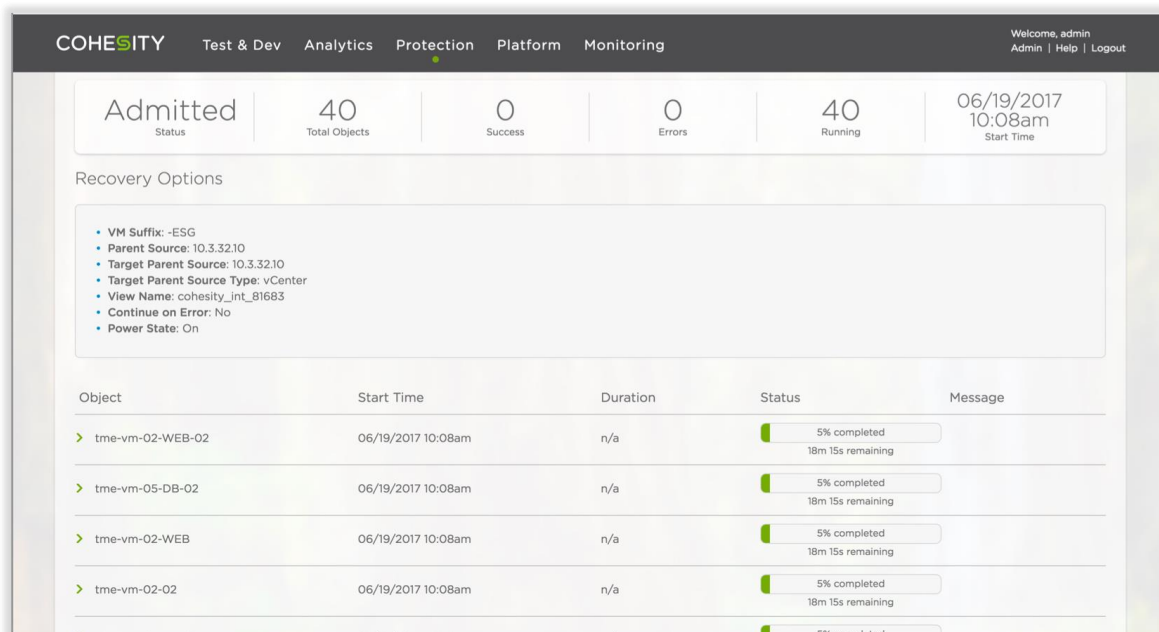
Instant restore is automated by the Cohesity console. Administrators can mount a specific image stored on Cohesity as an NFS data store, redirect the VMware ESX host to that image, and then initiate the VM boot process and application recovery. The end-user returns to full operations in the time it takes to recover the VM, without the time-consuming re-hydration and data transfer that most data protection schemas require. After application restart, administrators can use storage vMotion to transparently migrate the image back to the primary data store.

ESG Lab Testing: Instant Restore

ESG Lab tested instant restore on the Cohesity 2500 DataPlatform using the Recovery workflow, which involved cloning VMs from a Cohesity backup image, mounting them to the vCenter/ESX host, and running the VMs from the Cohesity storage.

Under the **Protection** tab of the UI, we clicked **Recovery/Recover VMs** and selected the ESG jobs detail, showing a list of completed backups of 40 Windows Server

2012 VMs. Each VM was configured with 40 GB of storage, 64 MB memory, and 2 vCPUs. We selected all the VMs, chose the 6:39 am snapshot, and clicked **Add to Cart** and **Continue**. Next, we changed the Recovery Options to append “-ESG” to



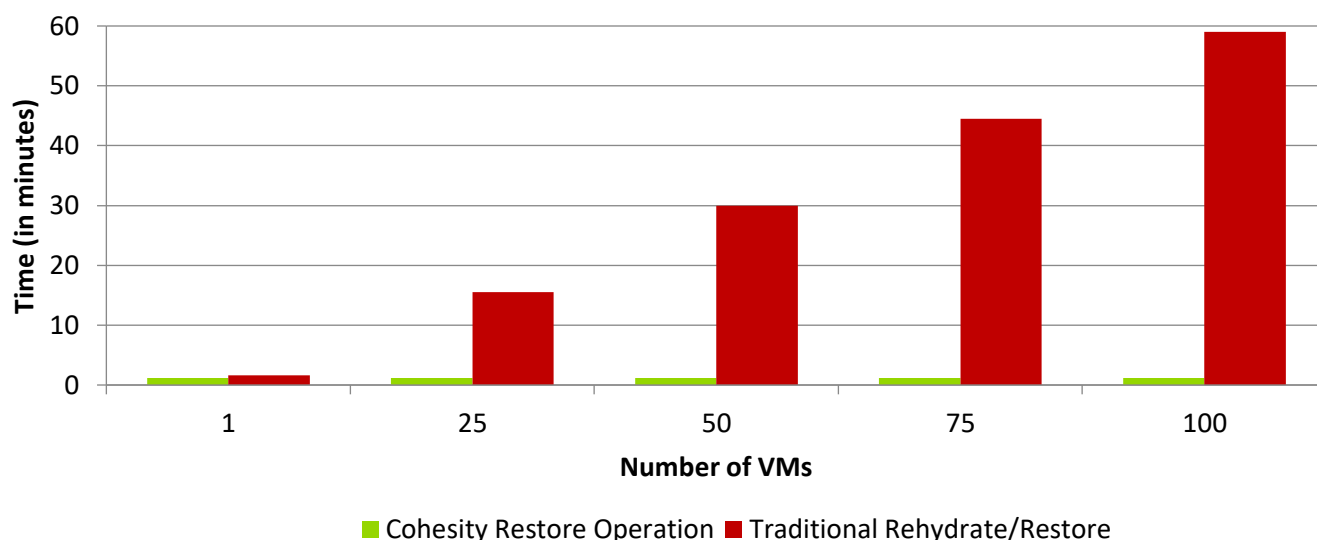
the file name and selected the options to recover to the original ESX host and vCenter, and to power up the VMs. After clicking **Finish** and **View Details**, we watched the status bars fill from 0% to 100% as each VM was booted and powered up. All of the VM data was available and VMs were ready to boot within a few seconds, and began booting on vSphere. Three VMs were booted within a minute, and all 40 VMs were powered on and ready for use within seven minutes, with full recovery completing in the background. ESG Lab also viewed a similar test using 100 x 1GB Linux VMs, demonstrating additional scale.

- Test/Dev and Recovery Workflows: In Cohesity’s test/dev workflow (used here), instant restore boots the VMs on Cohesity but does not perform the background storage vMotion back to primary storage; this enables organizations to execute tasks such as service pack testing before committing changes that impact the production data center. In the Cohesity Recovery workflow, after instant restore, storage vMotion can return the image to primary storage. ESG Lab successfully tested both scenarios.

ESG Lab also tested instant recovery of a 40GB SQL Server VM that had used both VMware VADP to capture only incremental changes and Microsoft VSS to create a consistent, point-in-time snapshot of the database. We viewed a table, deleted it, and then queried that table to prove it was no longer there; this simulated problems that plague IT organizations such as databases becoming unresponsive, corrupted, and even deleted. Next, we instantly restored that single table back to a point in time before the deletion and successfully queried the database, proving it was back up and running. This granular instant restore means that administrators can get back in operation immediately, without having to restore the whole database and disrupt other users.

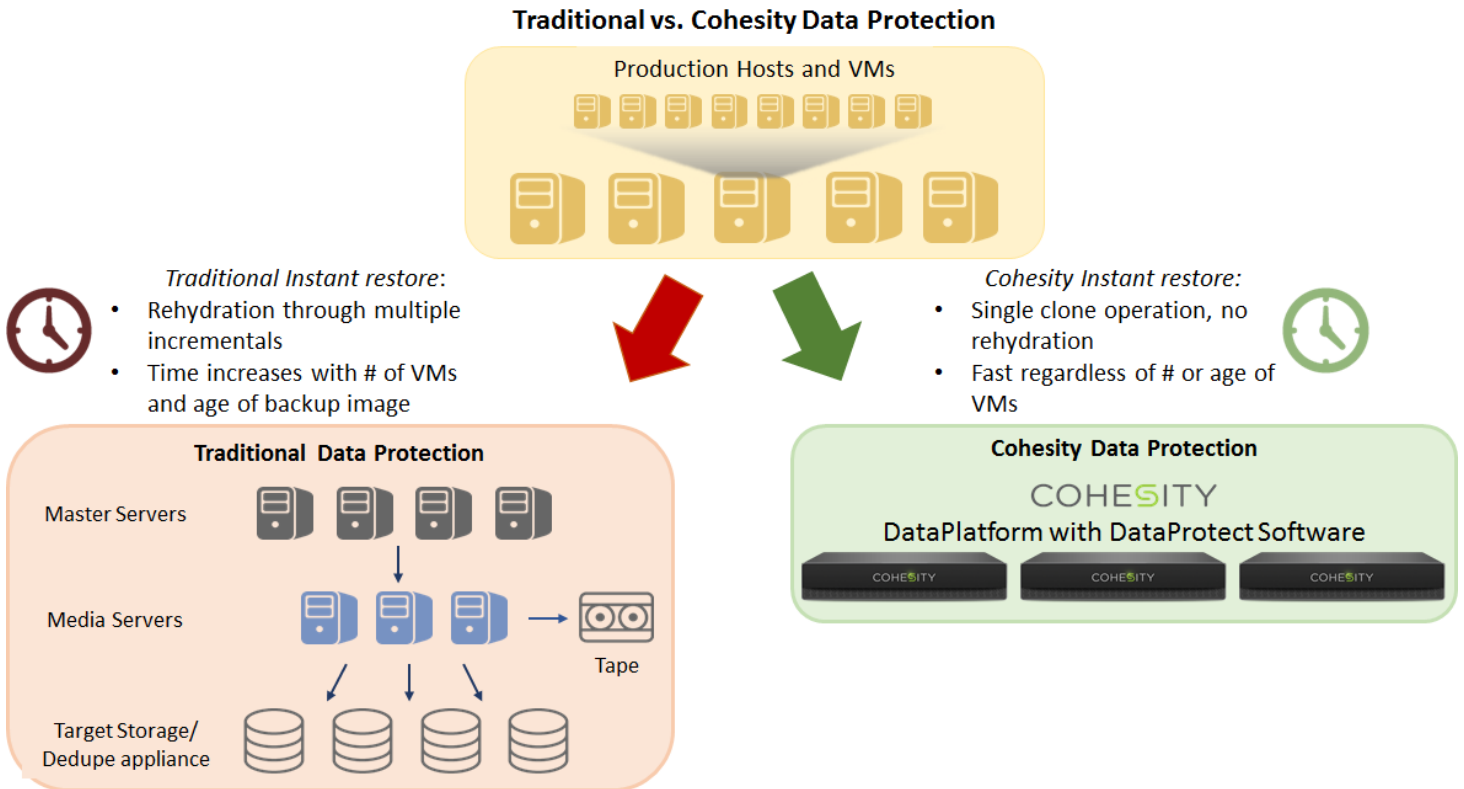
Finally, ESG Lab created a theoretical comparison of Cohesity with a similarly configured, non-flash, traditional backup software and appliance that evaluated the time for each appliance to prepare VMs to a ready-to-boot state for instant restore. All times were based on internal Cohesity test results. This comparison looked only at storage-related tasks. Non-storage-related tasks were removed from the equation, including the time to mount a data store to vSphere (a VMware task that would be the same for both appliances) and the time to actually boot the VM (a VMware task that is dependent on the VMware infrastructure). This leaves only the time for the Cohesity clone operation, and for the traditional appliance to rehydrate and restore. As the chart shows, the Cohesity time remains predictably low and constant, while the traditional solution takes significantly longer as the number of VMs scales. (Note that all data points were increased by one to enable the Cohesity data to be viewable in the chart.)

Instant Restores: Time Until Ready to Boot



What the Numbers Mean

- For a single 15GB VM, Cohesity’s elimination of the rehydration/restore process saved approximately 24 seconds compared to the traditional solution. At the 100VM count, Cohesity provided a time savings of 40 minutes compared to the traditional solution.
- At scale, the combination of flash storage and a single clone operation enabled Cohesity to restore 100 VMs in the same amount of time as one VM, while the traditional solution took progressively longer.



i Why This Matters

ESG research respondents report that for 73% of their virtualized production workloads, restoring the VMs takes more than an hour, and for 51% restores take more than 2 hours.² This is out of sync with the downtime tolerance that ESG research respondents reported, as mentioned previously. The problems with some instant recovery solutions are that they take too long, and are limited in the number of VMs that can be instantly restored at once. So, if your entire vCenter goes down, it takes a series of sequential restore operations—and significant time—to get back in operation.

ESG Lab validated the ability to instantly restore VMs at scale using the Cohesity DataPlatform. We instantly restored 40 x 40GB VMs, booting them all from the Cohesity platform: in seconds, all VM data was restored and VMs were ready to boot. Completing the boot on vSphere, three VMs were ready in one minute and the rest powered up within seven minutes, enabling users to return to productivity in a rolling fashion. We also instantly restored a single table from a 40GB SQL VM that had leveraged VADP and VSS to create an incremental, point-in-time snapshot. These capabilities can keep an organization running with minimal disruption. In addition, ESG Lab reviewed internal results that demonstrate Cohesity’s significant time advantage over traditional backup software appliances in the time it takes to get VMs ready to boot for instant restore.

² Ibid.

The Bigger Truth

Organizations have limited tolerance for production downtime, making business continuity essential. The faster you can return to production operations, the less disruption to business and revenue. Traditional backup solutions prevent data loss, but recoveries take time, prolonging business disruption. And while snapshots reduce the recovery time, incremental changes take time as you daisy chain backwards chronologically to the most recent full backup. Organizations are often forced to choose between maintaining multiple protection points and being able to restore quickly. As organizations virtualize more of their applications for greater agility and efficiency, they need data protection processes to step up to the same agility and efficiency, instead of holding them back.

Along with many other features, the Cohesity DataPlatform enables instant VM restore at scale to any recovery point, minimizing downtime and keeping organizations up and operational. And, you can restore many VMs (even an entire vCenter) from any snapshot, no matter how old or how many incremental changes have been made, in just the time it takes to boot the VMs. This takes no pre-planning and no proxy servers. Cohesity's global file system and Snaptree functionality enable every member of the cluster to access all the data immediately. For production applications, you can boot the VM and then migrate the image back to primary storage using Storage VMotion in the background, while for test/dev projects you can boot the VM and test your patches without propagating into production, and then easily destroy that clone. While other solutions limit the number of instant restores that can take place at a given time, Cohesity eliminates that bottleneck.

Key to the value is that this converged platform scales linearly, adding CPU and storage with each new node; as you add capacity, you also add restore threads, so you increase the number of VMs or physical servers to restore at the same time. Cohesity can also reduce the amount and cost of storage, since a single snapshot can serve many secondary storage purposes, from data protection to analytics to test/dev. For organizations looking to minimize downtime, maintain business continuity, and reduce costs, the Cohesity DataPlatform deserves a good look.

All trademark names are property of their respective companies. Information contained in this publication has been obtained by sources The Enterprise Strategy Group (ESG) considers to be reliable but is not warranted by ESG. This publication may contain opinions of ESG, which are subject to change. This publication is copyrighted by The Enterprise Strategy Group, Inc. Any reproduction or redistribution of this publication, in whole or in part, whether in hard-copy format, electronically, or otherwise to persons not authorized to receive it, without the express consent of The Enterprise Strategy Group, Inc., is in violation of U.S. copyright law and will be subject to an action for civil damages and, if applicable, criminal prosecution. Should you have any questions, please contact ESG Client Relations at 508.482.0188.

The goal of ESG Lab reports is to educate IT professionals about data center technology products for companies of all types and sizes. ESG Lab reports are not meant to replace the evaluation process that should be conducted before making purchasing decisions, but rather to provide insight into these emerging technologies. Our objective is to go over some of the more valuable feature/functions of products, show how they can be used to solve real customer problems and identify any areas needing improvement. ESG Lab's expert third-party perspective is based on our own hands-on testing as well as on interviews with customers who use these products in production environments.