

OCTOBER 2023

ExaGrid With Commvault: Maximum Deduplication Savings With Easy Management

Tony Palmer, Principal Analyst, Validation Services

Abstract

This report documents a validation by TechTarget’s Enterprise Strategy Group of ExaGrid testing that demonstrated both the capacity savings and ease of use available from a combined ExaGrid and Commvault backup solution.

The Challenges

Storing and protecting continually growing amounts of data increases stress on both infrastructure and IT staff. According to recent Enterprise Strategy Group research, nearly three-quarters (73%) of survey respondents reported having more than 500 TB of on-premises backup data, with 12% reporting that they have 5 PB or more.¹ In another survey, 74% of respondents reported that backup data volumes were growing more than 20% annually, with 32% reporting that they were growing more than 50% annually (see Figure 1).² Protecting this data is a critical assignment in today’s world in which organizations depend on highly available applications and data, but backing up that data strains IT budgets.

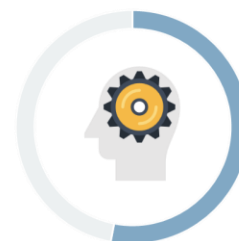
Figure 1. IT and Data Protection Challenges



73%
report having more than 500 TB of backup data



74%
report backup data volumes growing more than 20% annually



53%
report that their IT environments are more complex than 2 years ago

Source: Enterprise Strategy Group, a division of TechTarget, Inc.

In addition, while technology innovations are transforming IT in positive ways, organizations continue to struggle with IT complexity. According to recent Enterprise Strategy Group research, 53% of survey respondents reported that their IT environments are more complex than they were two years ago.³ Further complicating IT is the trend

¹ Source: Enterprise Strategy Group Research Report, [Cloud Data Protection Strategies at a Crossroads](#), August 2023.

² Source: Enterprise Strategy Group Research Report, [The Strategic and Evolving Role of Data Governance](#), July 2023.

³ Source: Enterprise Strategy Group Research Report, [2023 Technology Spending Intentions Survey](#), November 2022.

This Enterprise Strategy Group Technical Review was commissioned by ExaGrid and is distributed under license from TechTarget, Inc.

toward having more generalists on the IT staff handling a range of tasks, rather than experts dedicated to specific areas like data protection.

The Solution: ExaGrid Storage With Commvault Backup Software

Individually, Commvault and ExaGrid each offer data deduplication, but together they offer a highly efficient backup solution that can reduce the storage footprint by up to 300%, saving storage capacity and cost. This reduction also ensures that minimal WAN bandwidth is required for remote replication. No operational changes are needed to the Commvault configuration, making it a simple task for administrators to add an ExaGrid target to an existing Commvault environment. While a complete review of all ExaGrid and Commvault features is beyond the scope of this paper, a brief description is provided below.

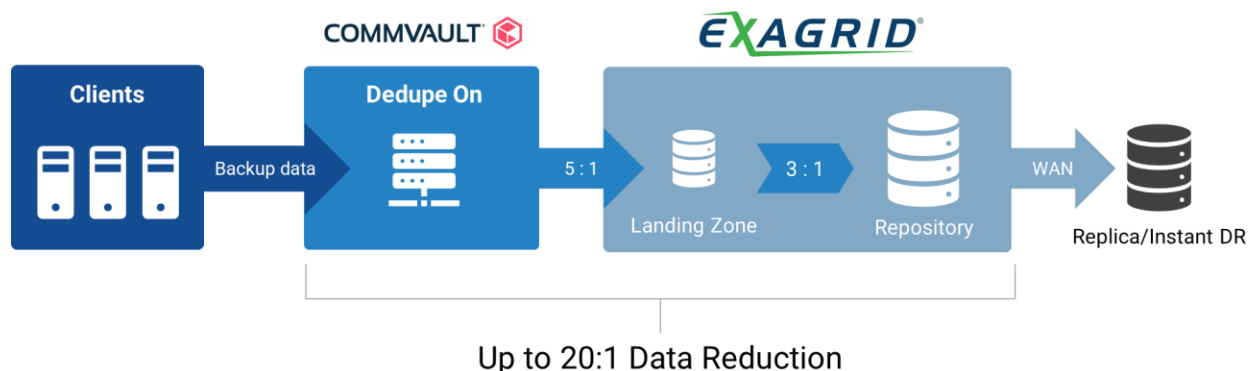
ExaGrid's data deduplication works with Commvault data deduplication and Commvault compression, and both can be either enabled or disabled. When Commvault compression is enabled, ExaGrid performs decompression prior to deduplication and always returns backup data to Commvault in the form it was originally written. The testing results in this technical review were achieved with Commvault data deduplication enabled and compression disabled. Enterprise Strategy Group did not test ExaGrid with Commvault compression enabled.

ExaGrid Tiered Backup Storage is designed specifically for data protection. A key feature is its Landing Zone tier that ensures fast backups and restores. Data sent to ExaGrid goes to the Landing Zone before being deduplicated and stored in the repository (see Figure 2). This eliminates the storage bottleneck of typical inline deduplication and minimizes the backup window. It also enables fast restores, since no deduplicated data rehydration is needed.

The ExaGrid Repository Tier provides global deduplication, long-term data retention, and replication to additional ExaGrid appliances over the WAN, or to the public cloud (Amazon AWS and Microsoft Azure), for disaster recovery. It is a scale-out system that supports different capacity-sized models and uses RAID6 protection with hot-swappable components. Other features include:

- Adaptive deduplication, enabling deduplication and replication during the backup window.
- A non-network-facing tier, plus delayed deletes and immutable data objects, which enables ransomware recovery.
- Data encryption at rest and during WAN replication.
- Integration with Commvault compression and deduplication enabled.
- Support for heterogeneous backup application environments.

Figure 2. ExaGrid and Commvault Deduplication



Source: Enterprise Strategy Group, a division of TechTarget, Inc.

Commvault has long been a leading provider of backup and data protection solutions. Commvault's single platform provides enterprise-class protection and recovery of on-premises and cloud-based physical and virtual files, applications, and databases. Features include:

- High-performance backups.
- Fast, granular recovery.
- End-to-end encryption.
- Built-in ransomware protection.
- Flexible copy data management that enables multiple uses of backup data.
- Global deduplication.

Enterprise Strategy Group Tested

ExaGrid + Commvault = A Reduction in Storage Footprint Through Enhanced Deduplication

Enterprise Strategy Group first reviewed ExaGrid testing that demonstrated how a combined ExaGrid/Commvault solution can increase data deduplication rates, saving customers money on storage capacity. Commvault deduplicates backups inline and sends them to the ExaGrid appliance. The deduplicated data first hits the ExaGrid landing zone and is then further deduplicated before being stored on the ExaGrid non-network-facing repository.

It should be noted that this testing was designed to demonstrate the additional deduplication on the ExaGrid target. It emulates a real-world scenario but does not use the large amount of data that customers typically would be backing up; these large data volumes by nature increase the deduplication ratio.

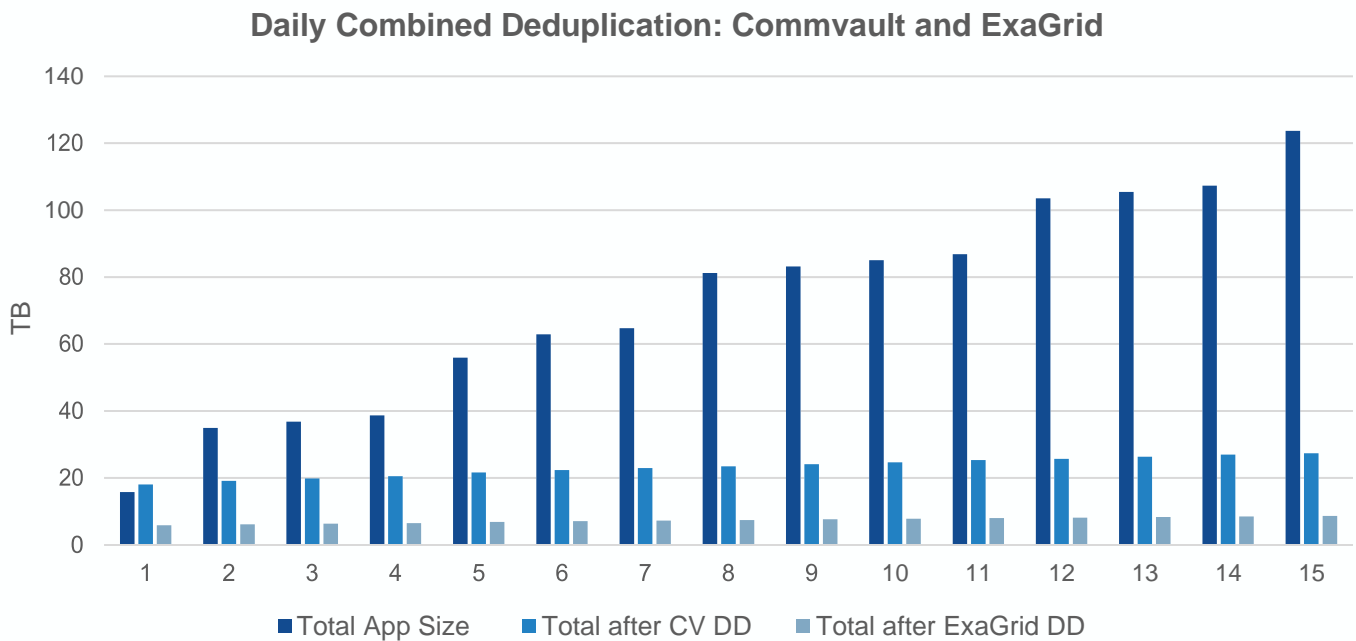
First, we reviewed the test bed setup and data-generating tool.

- **Data set.** A 15.5 TB data set was created using ExaGrid's FileMod data generation tool. The data was 10 TB of file system data; five Windows Server 2019 VMware virtual machines (VMs), each with 1 TB of file system data; and a 500 GB SQL Server database. File sizes ranged from 4 KB to 500 MB, with an average of 10 MB, and were spread across more than 110K directories. Directories were assigned a seed value to ensure that successive backups included unique data.
- **Backup process.** For the file system and VM data, the testing ran five full backups with four to six incremental backups in between. All the SQL backups were full. Backups were run twice a day to simulate—but accelerate—a typical work week, with data changing 1% for file systems, and 0.025% for SQL, with each backup. Data changes included growing, shrinking, changing, renaming, deleting, reordering, etc.
- **Backup resources.** Backups were executed using Commvault CommServer version 11.24.7 and were sent to an ExaGrid EX84 appliance via the CIFS/Samba protocol. The ExaGrid Landing Zone was sized for 84 TB.

The chart in Figure 3 shows the results of the 15 days of testing, with Commvault compression disabled. It's important to note that shortly after this testing was completed, ExaGrid announced that they support Commvault compression enabled along with Commvault deduplication enabled. Over that time, the complete application data set grew, while both Commvault and ExaGrid deduplication reduced the amount of data on disk.

- Over the course of 15 days, the 15.5 TB data set had five full backups, plus about 25 incrementals in between the fulls, for a total data set size of about 123.76 TB.
- Of that 123.76 TB, Commvault delivered only 27.25 TB to ExaGrid—about a 4.5:1 dedupe ratio.
- ExaGrid then deduped that 27.25 TB down to 8.66 TB, an additional 3:1 ratio.
- The total combined deduplication was more than 14:1.

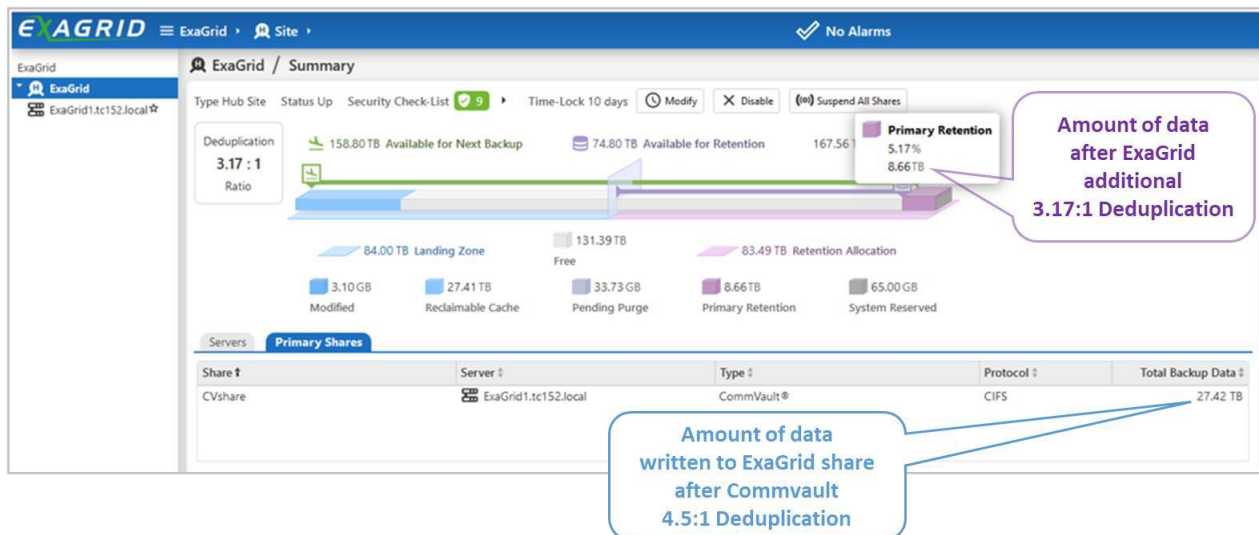
Figure 3. Data Reduction by Commvault and ExaGrid



Source: Enterprise Strategy Group, a division of TechTarget, Inc.

Next, we logged into the ExaGrid appliance to view the final day details from the ExaGrid GUI.⁴ Figure 4 shows the details of total deduplication.

Figure 4. Total Combined Deduplication



Source: Enterprise Strategy Group, a division of TechTarget, Inc.

⁴ Note that the screenshot in Figure 4 was taken just as the final backups were completing. As a result, the total data stored on ExaGrid shows as 8.63 TB, while the actual final data amount was 8.66 TB.

Adaptive Deduplication

ExaGrid has a feature called adaptive deduplication that enables ExaGrid to start deduplication only when appliance resource usage is low. This enables ExaGrid to promptly get the back-end, additional deduplication accomplished without slowing down any backup or restore processes. ExaGrid also offers graphs showing write rates, modified bytes, deduplication rates, and the replication queue over time, making it easy for administrators to understand how data is being managed. Administrators can zoom in to view specific dates or zoom out for an overview. Figure 5 shows that ExaGrid adaptive deduplication (in purple) increased as the write rate of the Commvault backup (in green) began to slow.

Figure 5. ExaGrid Adaptive Deduplication



Source: Enterprise Strategy Group, a division of TechTarget, Inc.

Why This Matters

Storing backup data is critical for ensuring maximum business productivity, but continual data growth strains budgets. Organizations can save money by deduplicating backup data to reduce the amount of storage they need.

Enterprise Strategy Group validated that the combined ExaGrid/Commvault solution reduced 123 TB of data down to 8.66 TB, reducing storage capacity requirements by 14:1. It should be noted that the ExaGrid test setup was conservative, and the methodology delivered deterministically random data. Customers in real-world production environments are likely to see even higher deduplication rates with this joint solution.

Ease of Deploying and Managing Commvault and ExaGrid

Next, we viewed the ease of deploying and managing backups using Commvault and ExaGrid.

Adding ExaGrid to an existing Commvault environment is as simple as adding a new Commvault Library and Storage Policy for ExaGrid and selecting them for use. This is a quick and easy task with which Commvault administrators are familiar. Enterprise Strategy Group viewed a demo of the one-time initial deployment, which is

also simple and involves a few tasks on the ExaGrid side and then on the Commvault side. Then, we viewed a demo of the ExaGrid share creation process.

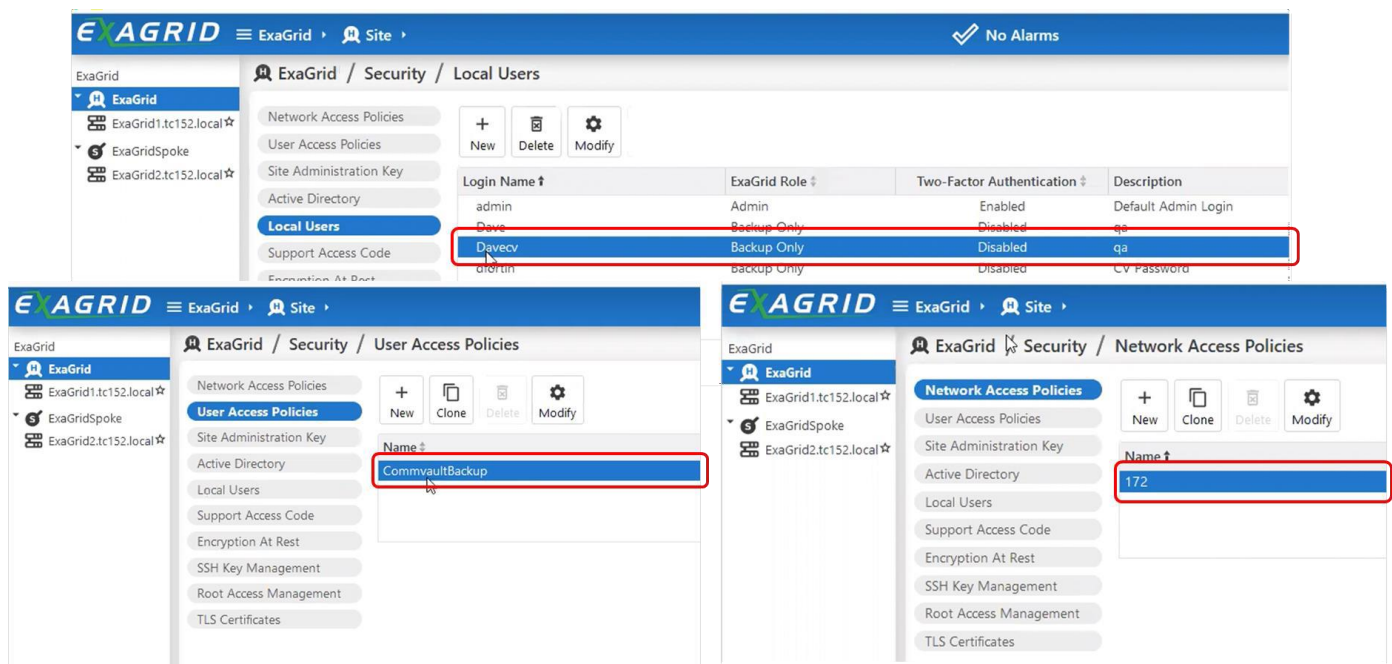
Initial Deployment

ExaGrid Tasks

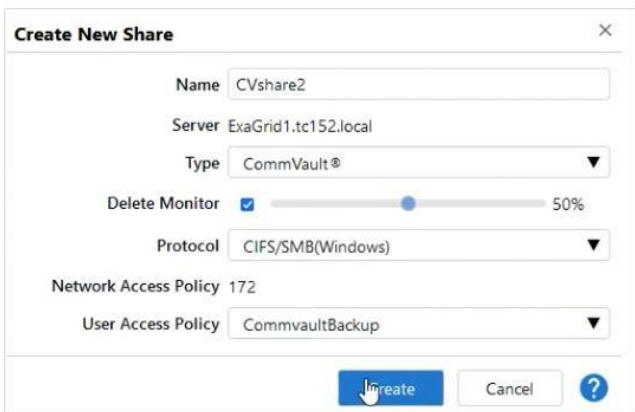
After logging into the ExaGrid EX84, we added users and created access policies before creating an ExaGrid share for backups (see Figure 6). You can have up to 32 EX84 appliances in a single scale out system which can protect petabytes of data.

- From the **Security/Local Users** tab, we added a user called DaveCV and assigned backup-only privileges.
- Next, we clicked on **Security/User Access Policies** and created a new policy called *CommvaultBackup*.
 - With a click of the **Modify** button, we added DaveCV to that policy; users can be added singly or in groups. (While it is not required to assign users to shares, it is a good security practice.)
- Next, we created a Network Access Policy with an open 172 IP address. These policies can assign specific hosts, IP addresses, and ranges of IP addresses.

Figure 6. Creating ExaGrid Access



Source: Enterprise Strategy Group, a division of TechTarget, Inc.



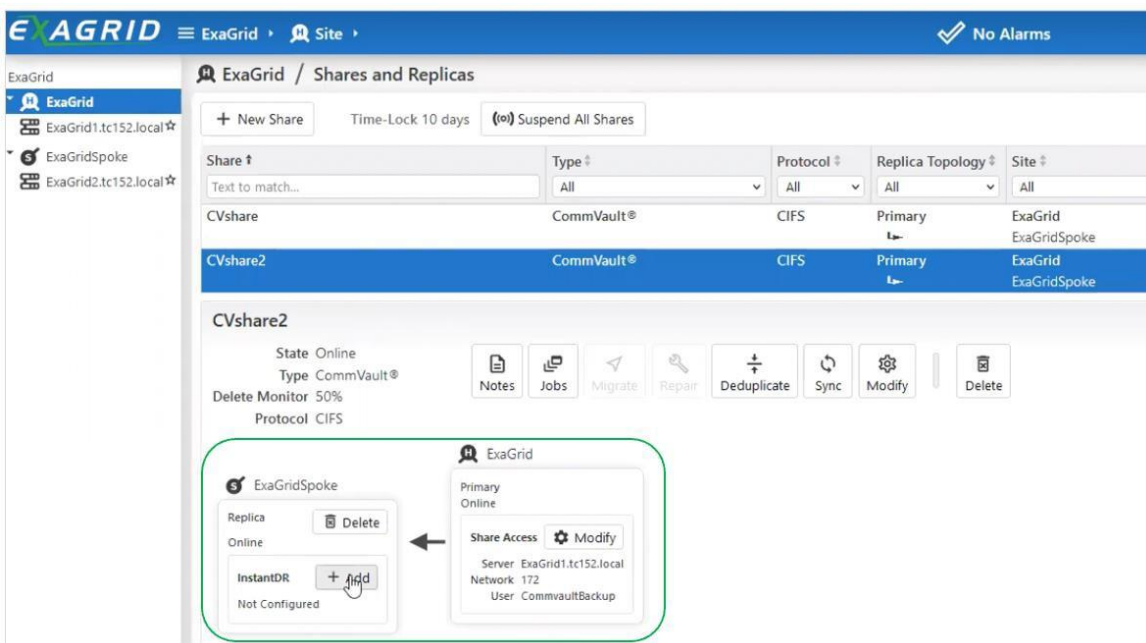
After these policies were set up, we created a share on the ExaGrid EX84.⁵ From the **Shares and Replicas** tab, we clicked **+New Share**, named it **CVshare2**, and selected **Commvault** as the type. The tight integration between Commvault and ExaGrid ensures that data sent to Commvault shares is optimized by ExaGrid features.

Next, we selected the **CIFS/SMB** protocol, the **Network Access Policy**, and the **CommvaultBackup User Access Policy** created previously. A Delete Monitor is also available, which will alert the administrator in case a selected percentage of the share is deleted within 24 hours as a protection against ransomware or other unauthorized data

deletion.

Next, we added an ExaGrid EX40000E appliance to serve as a remote ExaGridSpoke and initiated replication from the first appliance. The Commvault backup data was then backed up to the ExaGrid Landing Zone, deduplicated and stored on the ExaGrid share, and then replicated to the remote site. We also enabled ExaGrid InstantDR on the remote server so that data could be exposed using the same User Access Policy (see Figure 7).

Figure 7. Creating ExaGrid Primary and InstantDR Shares



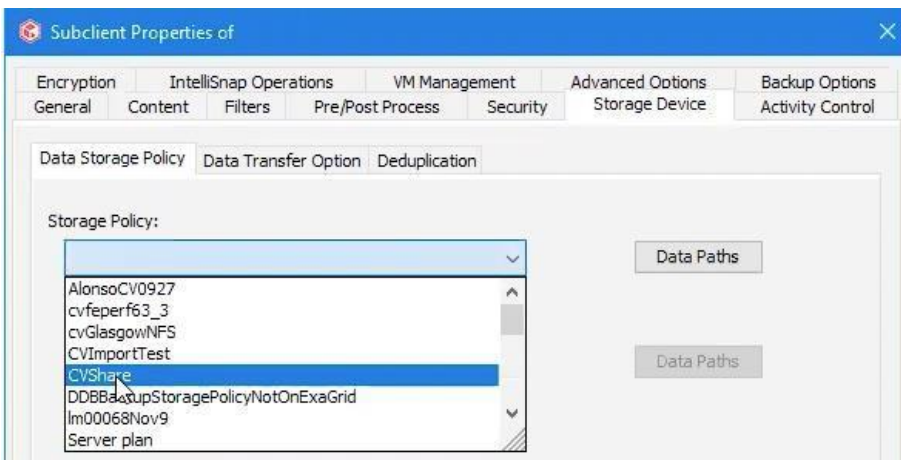
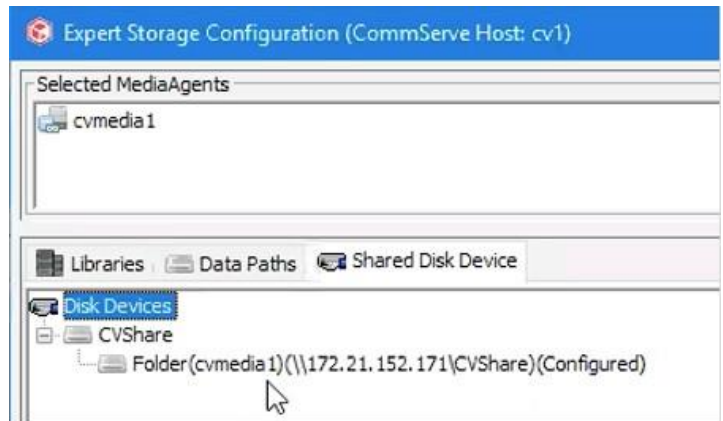
Source: Enterprise Strategy Group, a division of TechTarget, Inc.

⁵ Note that the share creation screenshots were complete after the deduplication testing. The new share we created was *CVshare2*, while the actual share used in the testing was *CVshare*.

Commvault Tasks

On the Commvault side, we created a Library and Storage Policy that, together, tell Commvault how to access the ExaGrid share.

First, we created the Commvault Library. From the Commvault CommCell Console, we clicked on **Libraries/Storage/Expert Storage Configuration** and added the previously created *cvmedia1* media server to the CVShare library to access the share. Next, we clicked **Shared Disk Device** and added the path and a base folder for backups to the ExaGrid share.



Then, we created the Commvault Storage Policy. We scrolled through the Commvault directory of hosts and clusters to find the file data, VMware VMs, and SQL database that we wanted to back up. From the **Storage Policies** tab, we selected **CVShare**, clicked on **Properties**, and added the 24 subclients containing that data.

Next, we added a retention time of 90 days and enabled creation of a media server deduplication database to track application size and data reduction.

From the **Properties** menu, we could add content as well as filters, pre- and post-process tasks, security, and other properties. From this menu, we clicked **Storage Device/Data Storage Policy** and selected **CVShare**. Choosing this storage policy, which contains the library that we created to write to the ExaGrid share, is what linked our selected backup data set to the target. We disabled compression and encryption in Commvault, since ExaGrid does disk-level encryption. (It bears repeating that shortly after this testing was completed, ExaGrid announced that they support Commvault compression enabled along with Commvault deduplication enabled.) Finally, we created a schedule to execute backups every 12 hours.

Why This Matters

As IT complexity increases, organizations are looking carefully at new technology solutions. Optimal technology innovations consider the need for simplicity and ease of use. Complexity breeds inefficiency and cost, while simplicity ensures faster time to value.

Enterprise Strategy Group validated how easy and fast it was to set up an initial ExaGrid/Commvault deployment to securely back up data, replicate it, and make it available for instant restore. The ExaGrid tasks are simple to complete using the intuitive GUI, and the Commvault tasks are familiar to Commvault administrators.

Conclusion

Increasing efficiency never goes out of style for obvious reasons: Saving money is a key business objective. According to Enterprise Strategy Group's latest *Technology Spending Intentions Survey*, becoming more operationally efficient remains the most-cited objective for organizations' digital transformation efforts, as it has been for the past four years.⁶

It is well known that both ExaGrid and Commvault offer backup solutions that are easy to use and can help organizations store data more efficiently with deduplication. What may be less well known is that a combined ExaGrid and Commvault solution can achieve even greater deduplication: up to 20:1 in many cases. A 20x reduction in backup storage costs can be a significant boost to any budget.

Enterprise Strategy Group validated that:

- The combined ExaGrid/Commvault solution reduced a 123 TB backup data set that included 15 days of file data, VMware VMs, and SQL database backups down to 8.66 TB, a 14:1 reduction, in a realistic yet intentionally conservative test environment.
- Deploying and managing the combined solution was simple, intuitive, and fast.

We also reviewed an ExaGrid total cost of ownership (TCO) calculator that shows significant savings using Commvault and ExaGrid compared to Commvault with standard primary storage disk.

Any solution should be tested and planned with an organization's specific needs and objectives in mind. Taking this into consideration, if your organization is looking to spend less on backup storage without adding complexity to its data protection scheme, Enterprise Strategy Group recommends looking closely at the combined ExaGrid and Commvault solution.

⁵ Source: Enterprise Strategy Group Research Report, [2023 Technology Spending Intentions Survey](#), November 2022.

©TechTarget, Inc. or its subsidiaries. All rights reserved. TechTarget, and the TechTarget logo, are trademarks or registered trademarks of TechTarget, Inc. and are registered in jurisdictions worldwide. Other product and service names and logos, including for BrightTALK, Xtelligent, and the Enterprise Strategy Group might be trademarks of TechTarget or its subsidiaries. All other trademarks, logos and brand names are the property of their respective owners.

Information contained in this publication has been obtained by sources TechTarget considers to be reliable but is not warranted by TechTarget. This publication may contain opinions of TechTarget, which are subject to change. This publication may include forecasts, projections, and other predictive statements that represent TechTarget's assumptions and expectations in light of currently available information. These forecasts are based on industry trends and involve variables and uncertainties. Consequently, TechTarget makes no warranty as to the accuracy of specific forecasts, projections or predictive statements contained herein.

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 TechTarget, 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 Client Relations at cr@esg-global.com.

About Enterprise Strategy Group

TechTarget's Enterprise Strategy Group provides focused and actionable market intelligence, demand-side research, analyst advisory services, GTM strategy guidance, solution validations, and custom content supporting enterprise technology buying and selling.

✉ contact@esg-global.com

🌐 www.esg-global.com