

ESG Lab Review

ExaGrid-Veeam Accelerated Data Mover

Date: October 2016 **Author:** Kerry Dolan, Lab Analyst and Brian Garrett, Vice President, ESG Lab

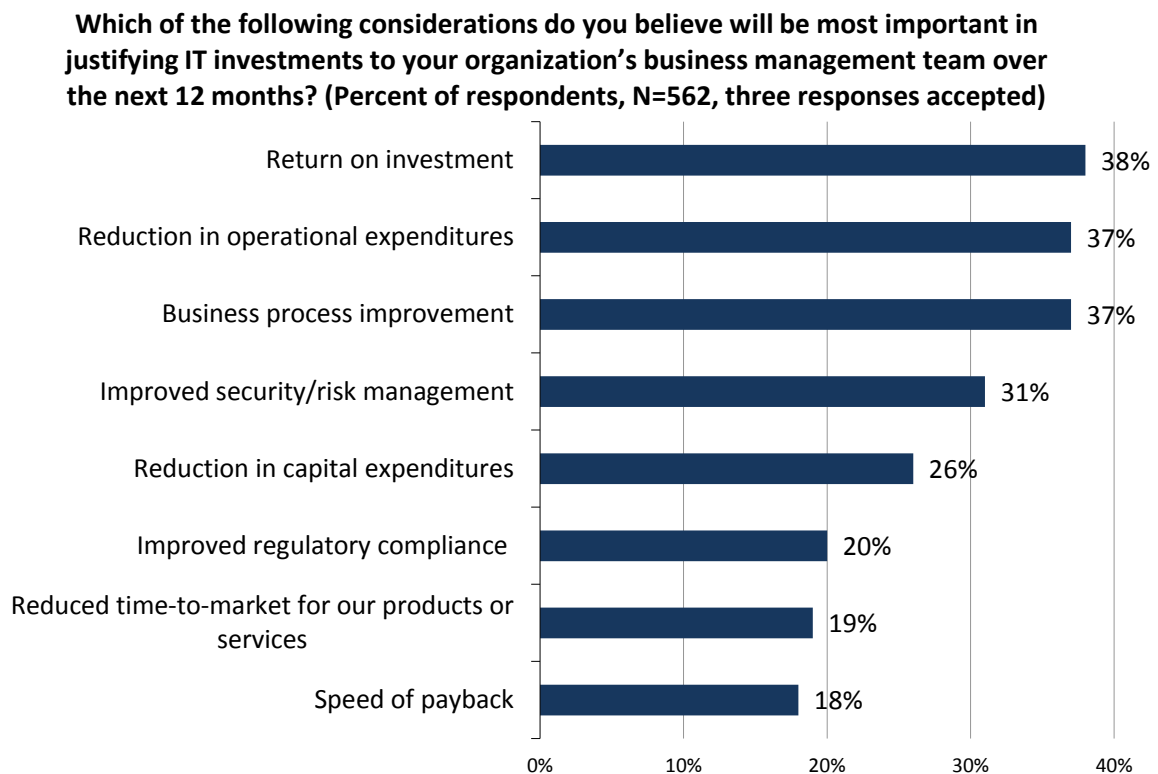
Abstract: This ESG Lab review documents hands-on testing of the ExaGrid-Veeam Accelerated Data Mover, with particular focus on its ability to improve the performance and efficiency of both full and synthetic full backups.

The Challenges

Today, organizations have come to count on the economic benefits that they can gain by consolidating virtualized applications. But as virtualization proliferates and data volumes continue to grow, data protection performance can suffer. It is not surprising, then, that in ESG's annual spending intentions research, the same few IT priorities continue to hold the top spots. Increasing use of server virtualization, improving data backup and recovery, and managing data growth remain in the top four most-cited IT priorities in 2014, where they have been since 2010.¹

In addition, when asked to name the top considerations for justifying IT investments, respondents to ESG's 2014 research survey cited return on investment, reduction in OPEX, and business process improvement most often. Whatever their challenges, organizations recognize that it is worth investing in IT solutions that can rein in operating costs and improve business processes.

Figure 1. Most Important Considerations For Justifying IT Investments



Source: Enterprise Strategy Group, 2014.

¹ Source: ESG Research Report, [2014 IT Spending Intentions Survey](#), February 2014.

Given these customer priorities, ExaGrid and Veeam make an ideal pairing. Veeam’s singular focus on data protection and availability for virtualized applications, combined with ExaGrid’s commitment to reducing backup data storage and speeding the backup process, make a powerful combination that can save both time and infrastructure resources with improved data protection.

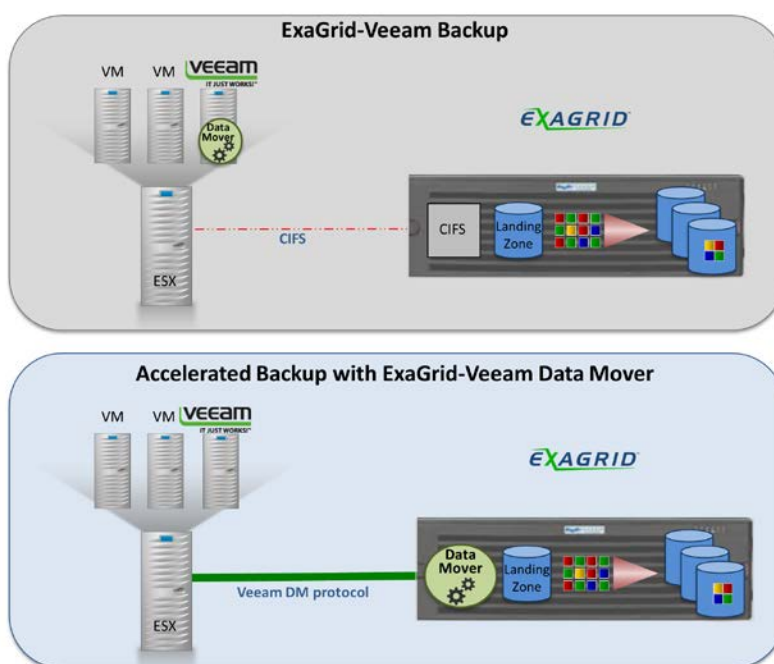
ExaGrid-Veeam Accelerated Data Mover

ExaGrid and Veeam have worked closely together for some time. The ExaGrid backup appliance is an affordable, scalable disk target for Veeam’s virtual machine (VM) backup software. ExaGrid appliances are high-performance backup storage devices with built-in deduplication. ExaGrid’s unique landing zone retains a non-deduplicated copy of the most recent backup for fast and easy restore and instant VM recovery. The scale-out grid architecture adds processors, memory, bandwidth, and disk with each added appliance, keeping your backup window fixed regardless of data growth. Veeam has made a name as a premier provider of VM data protection and availability solutions for organizations of all sizes.²

The most recent product of their collaboration is the ExaGrid-Veeam Accelerated Data Mover, a feature integrated into ExaGrid appliances with version 4.7 of the ExaGrid software. In essence, the Veeam Data Mover execution engine moves from the Veeam server to the ExaGrid appliances, freeing up both network and server resources; now, the tasks required for executing full, incremental, and synthetic backups are offloaded to the ExaGrid appliances. Two key benefits result from this. First, it enables Veeam-to-Veeam communications, ensuring faster backups and restores than when Veeam communicates via CIFS or NFS. Second, many tasks perform faster inside ExaGrid than when using Veeam backup server and network resources. In addition, when you add the feature to one ExaGrid appliance in a grid, it becomes available to all of them.

Figure 2 makes this easy to see. The top shows how ExaGrid-Veeam backups are performed without the new feature. The Data Mover tasks required to execute a Veeam backup are performed on the server, and data is sent over a standard CIFS interface to the ExaGrid appliance. The ExaGrid Landing Zone retains an un-deduplicated copy of the latest backup for fast restore, and ExaGrid’s Adaptive Deduplication occurs in parallel with backups, with only changed data sent to the back-end disk.

Figure 2. ExaGrid-Veeam Accelerated Data Mover



² For additional information, see our previous ESG Lab Validation [ExaGrid with Veeam: Virtual Machine Backup without Compromise](#), July 2014.

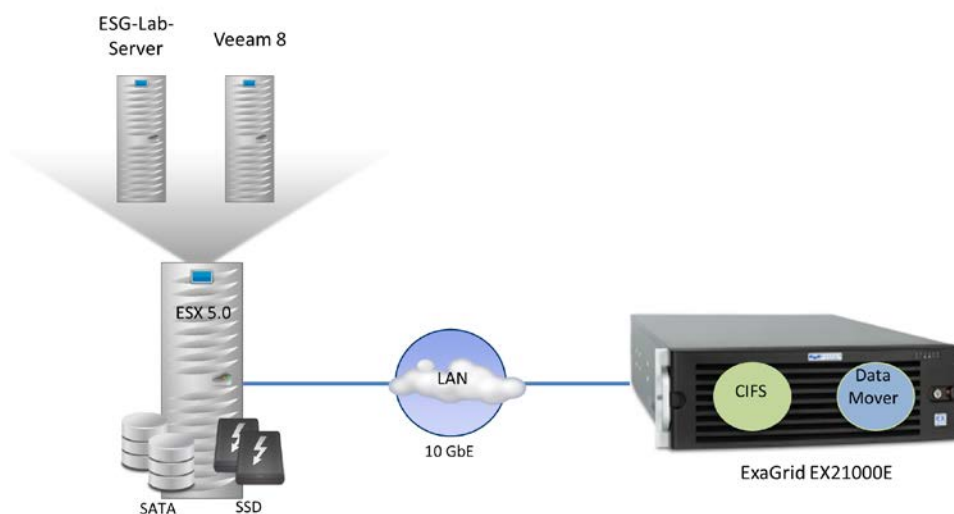
The bottom of Figure 2 demonstrates the process *with* the ExaGrid-Veeam Accelerated Data Mover installed. The Veeam server still initiates the backup, but the Data Mover has been re-located to the ExaGrid appliance. The Veeam server communicates “Veeam to Veeam” over an enhanced protocol instead of “Veeam to CIFS,” for a faster backup; CIFS is a chatty protocol, requiring acknowledgements between communications that can increase latency, and uses I/O sizes not optimized for streaming of backup data. The backup processing occurs on the ExaGrid array, freeing up network and server CPU resources for other tasks.

ESG Lab Tested

ESG Lab tested the ExaGrid-Veeam Accelerated Data Mover capability with a focus on comparing the performance of ExaGrid-Veeam backups using CIFS with the performance of backups using the Data Mover. Testing was designed to mimic a common weekly backup workflow: starting with a full backup, running daily incremental backups, and then pulling those backups together into a synthetic full backup. Instant VM recoveries were also tested. ESG used a combination of onsite testing and audits of previously run tests.

The test bed included an ExaGrid EX21000E appliance running version 4.7 of the ExaGrid firmware. A single ExaGrid share was configured, with two separate repositories: one for CIFS and one for the Data Mover. The appliance was connected by 10GbE LAN to a VMware ESXi 5.0 host configured with two six-core Xeon processors, 16GB RAM, 43GB of SATA storage, and one SSD for Veeam vPower temporary storage to host changes during instant VM recoveries. This server hosted two virtual machines: 1) a simulated 150GB Exchange database with 40GB system files, named **ESG-Lab-Server**, and 2) a Windows 2008 Veeam 8 virtual backup server. Figure 3 provides a high-level overview of the test bed.

Figure 3. ESG Lab Test Bed



Improving Backup Performance using the ExaGrid-Veeam Accelerated Data Mover

Initial testing began with a full backup run on the CIFS repository, followed by an identical full backup using the Data Mover repository. These backup jobs were run with both ExaGrid and Veeam defaults and best practices configured. This included Veeam server-side deduplication, Veeam dedupe-friendly compression, and LAN optimization. Details regarding both duration and distribution of resource consumption were captured about each backup. It should be noted that ExaGrid supports Veeam deduplication and dedupe-friendly compression; ExaGrid further deduplicates data above and beyond the Veeam algorithm. This enables customers to gain the maximum capacity savings.

Next, the Exchange database and system files that were harvested from a real-world IT environment were modified to simulate a typical, real-world 2% daily change rate, and an incremental backup was run using both the CIFS and Data Mover repositories. Over time, eight incremental backups were executed to both repositories with 2% of the data changing between each incremental. This process simulated a common weekly backup workflow.

Once a simulated week of incrementals was completed, the next step in testing was to create synthetic full backups using both the CIFS and Data Mover repositories. Figure 4 shows the benefit of the ExaGrid-Veeam Accelerated Data Mover for speeding synthetic fulls. The top of Figure 4 shows a regular synthetic full, in which the Veeam Data Mover operates on the backup proxy server, pulling data from the ExaGrid array over CIFS, stitching together the last full backup with multiple incrementals, and sending the new synthetic full back over the network to the landing zone. (ExaGrid has the only product with a landing zone, where an un-deduplicated copy of data is available so that synthetic fulls and restores can be more quickly executed.)

The bottom of Figure 4 shows the synthetic full using the Data Mover. The Veeam server initiates the backup and communicates with the Data Mover on ExaGrid over an optimized protocol. With the Data Mover inside the ExaGrid appliance, all the processing required to pull together the most recent full and subsequent incrementals occurs within the appliance. The new synthetic full is then written to the landing zone.

Figure 4. Synthetic Full Backup Comparison, CIFS vs. Data Mover

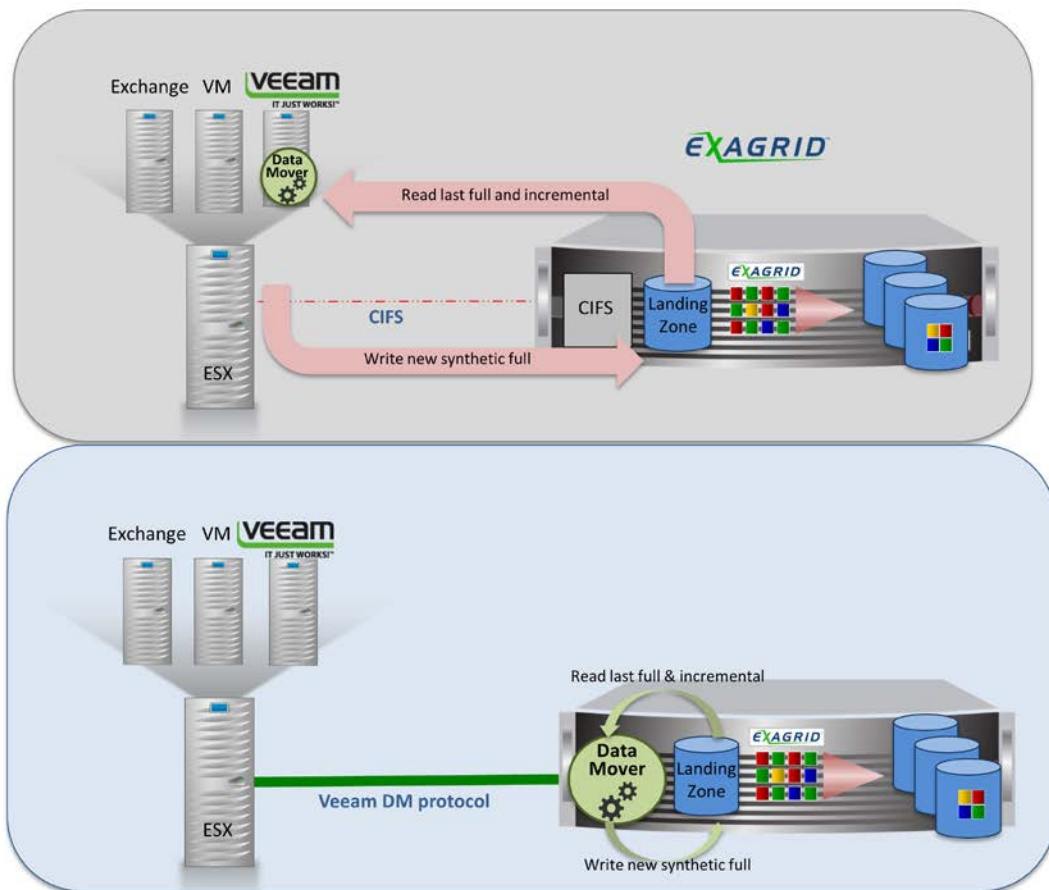
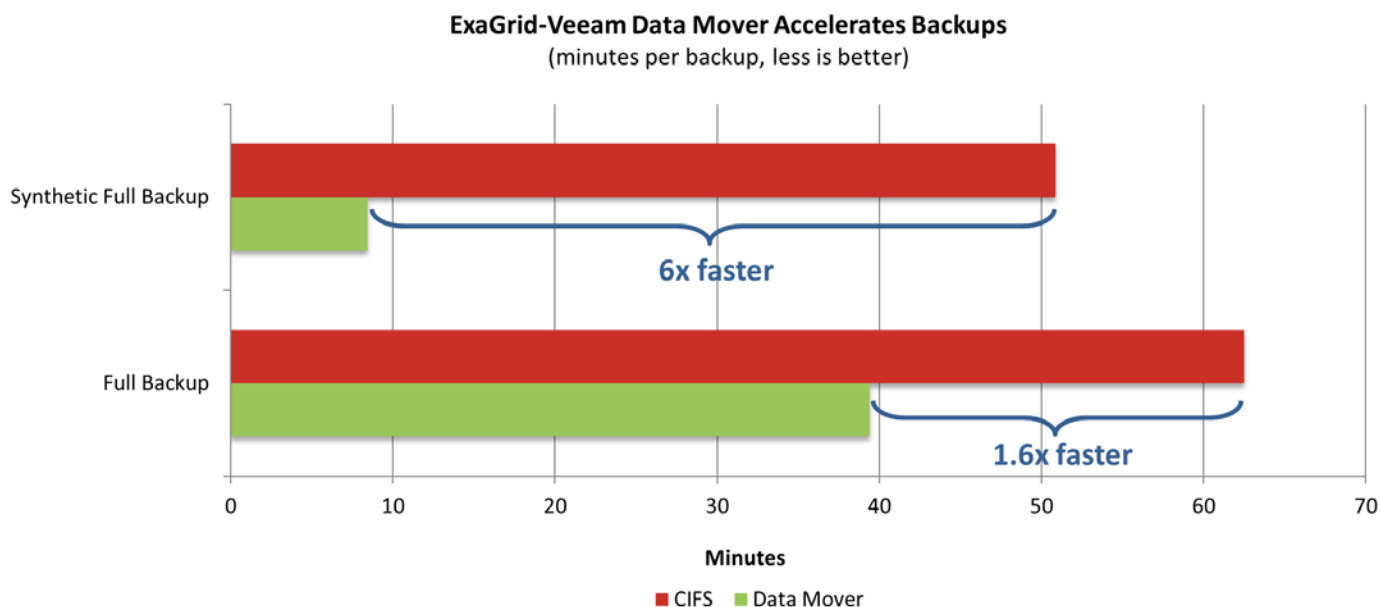


Figure 5 shows the benefit of the ExaGrid-Veeam Accelerated Data Mover in terms of duration of the backup process. For our testing with a 190GB Exchange and system files data set, the full backup required 62 minutes and 29 seconds (6:29) for CIFS, but only 39:25 using the Data Mover, for a 1.6x faster backup. The comparison for the synthetic full was even more significant. The CIFS synthetic full required 50:53 for this data set, while the Data Mover synthetic full took only 8:26, for a 6X faster backup.

Figure 5. *Faster Synthetic Full Backups*



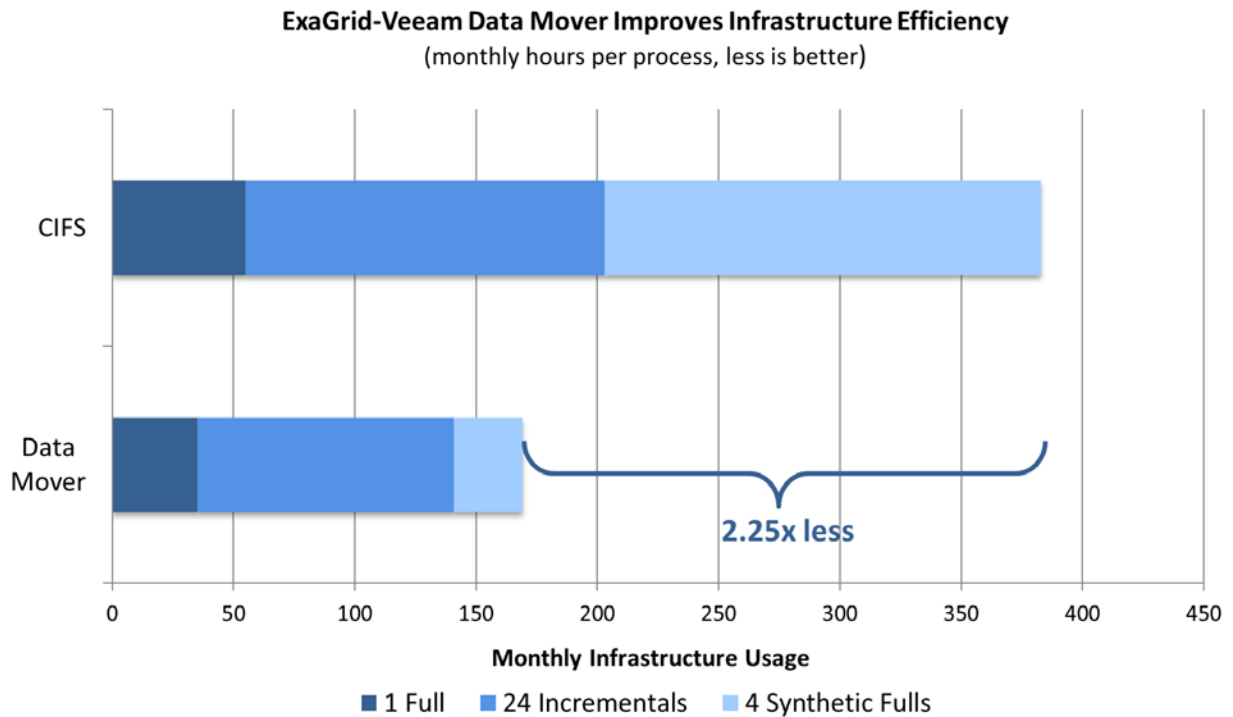
What the Numbers Mean

- The full backup using CIFS took more than one hour: 62:29 (62.5 minutes).
- The full backup using the Data Mover took 39:25 (39.42 minutes), or 1.6X faster.
- The synthetic full backup using CIFS took 50:53 (50.88 minutes).
- The synthetic full backup using the Data Mover took 8:26 (8.43 minutes), or 6X faster.

ESG Lab also looked at the efficiency of infrastructure usage for multiple backup tasks over the course of a month. Infrastructure (and IT administrator time) is used for multiple data protection tasks each day, week, month, and year. The time required for full backups, daily incrementals, and synthetic fulls consumes precious server and network production cycles during these tasks. Most organizations recognize that the importance of protecting their data far outweighs the disruption to the infrastructure, but the fact remains that data protection is an overhead task that does not contribute to production activities and bottom-line revenue. Any opportunity to reduce the time and wear on the infrastructure benefits the company by returning production cycles to the end users.

Figure 6 provides a visual demonstration of this benefit. Using the data set and backup duration times generated during our testing, ESG Lab extrapolated how much time would be consumed by full, incremental, and synthetic full backups using both CIFS and the Data Mover over four weeks (28 days), for a 10TB data set. Using a typical workflow, this would include an initial full backup, 24 daily incrementals (one/day for six days) and four synthetic fulls (done on the seventh day, and including a final incremental). Those times have been added together and charted by number of hours in Figure 6.

Figure 6. Improving Infrastructure Efficiency



What the Numbers Mean

The total number of hours in four weeks is 672. Based on ESG Lab testing:

- CIFS backup tasks would consume 383 hours, or approximately 57% of the hours in that month.
- Data Mover backup tasks would consume 169 hours, or 25% of the total hours in that month.
- Data Mover would reduce the amount of infrastructure time spent on these tasks by 2.25X.
- The Data Mover would reduce the time for synthetic full backups by 84%, from 180 hours to 28 hours.
- The Data Mover would reduce the time for incremental backups by 28%, from 148 hours to 106 hours.
- The Data Mover would reduce the time for synthetic full backups by 36%, from 55 hours to 35 hours.

Another metric ESG Lab measured was the consumption of resources during backup tasks. Table 1 outlines the key areas of bottleneck for each backup task, including the source, the Veeam proxy server, and the network. The most notable detail is in the network column. For each backup task, the network bottleneck dropped significantly between the CIFS version of the task and the Data Mover version.

Table 1. Reducing Network Bottlenecks

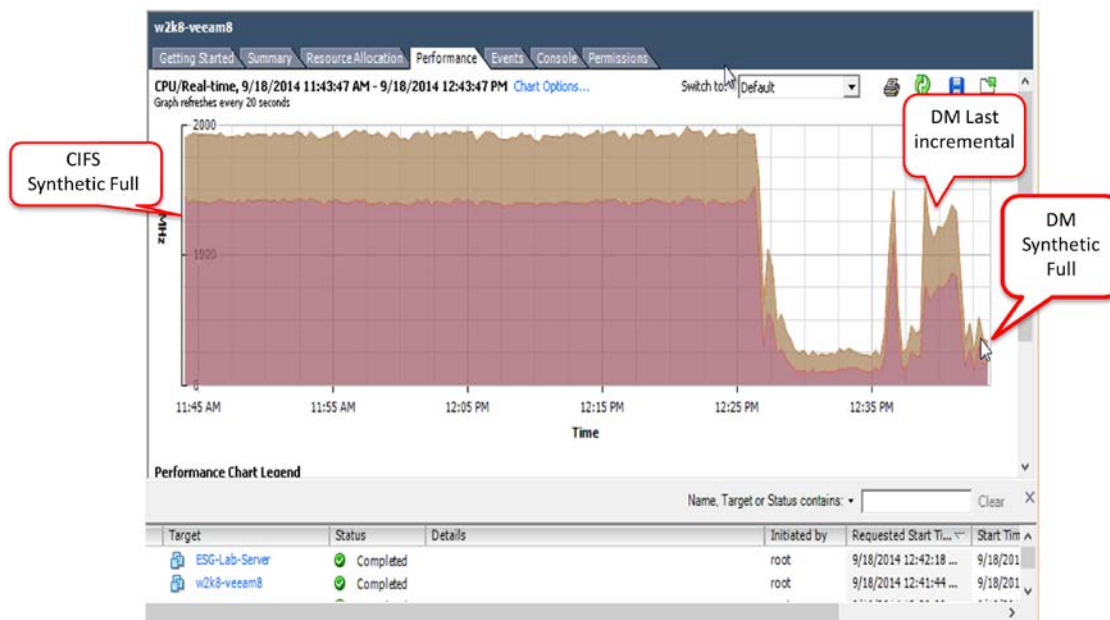
Backup Task	Source	Proxy	Network
CIFS Full	48%	95%	83%
DM Full	81%	85%	8%
CIFS Incremental	67%	70%	47%
DM Incremental	99%	52%	1%
CIFS Synthetic Full	61%	50%	31%
DM Synthetic Full	92%	62%	5%

What the Numbers Mean

- The network bottleneck was 83% for CIFS full backup, and dropped to 8% using the Data Mover.
- The network bottleneck was 47% for the CIFS incremental backups, and dropped to 1% using the Data Mover.
- The network bottleneck was 31% for the CIFS synthetic full, and dropped to 5% using the Data Mover.

In addition, ESG Lab looked at the CPU usage on the Veeam server during both CIFS and Data Mover synthetic full backups. Figure 7 shows the vSphere view of CPU usage. With the CIFS synthetic full, CPU consumption was high, consuming almost 2Ghz. Just after the 12:25 time stamp it dropped significantly, when that task was complete. When the Data Mover synthetic full started, the first task was to run a final incremental, and CPU usage reached 1Ghz-1.5Ghz briefly. When the Data Mover synthetic full process began, CPU usage dropped well below 1Ghz. Since the ExaGrid array had taken over the Data Mover processing, the Veeam server had gained significant CPU resources.

Figure 7. Server CPU Usage



Finally, ESG Lab tested the duration of Veeam Instant VM Recoveries with both CIFS and the Data Mover after a synthetic full was completed and during a synthetic full creation. These Instant VM Recoveries enable a user to boot a VM directly from the backup on the ExaGrid landing zone so that users can maintain productivity while IT resolves a problem with a VM. While the VM runs on this “spare tire,” writes are retained on the Veeam server. In addition, we tested a vMotion Instant VM recovery with both CIFS and the Data Mover; this is used to migrate the VM back to production storage once the issue has been resolved.

In every case, the instant recovery took less than 3.5 minutes. What is important to note is that even while the Data Mover is in action creating a synthetic full, ExaGrid has enough horsepower to complete an instant recovery in minutes, while other solutions may take hours. As a result, administrators can be assured of fast return to productivity even during a disaster.

Why This Matters

Backup can interfere with production operations, and therefore the faster the backup is completed, the faster regular operations can be resumed. As further evidence of the need to keep data protection from disrupting production, Veeam 8 includes the ability to throttle backups as necessary to prevent interruption with production. Synthetic full backups are a common method of creating a complete backup data set on a regular basis, but without the disruption and time to create a regular full backup.

ESG Lab validated that the ExaGrid-Veeam Accelerated Data Mover reduced the duration of full backups by 1.6X, and reduced the duration of synthetic fulls by 6X. Extrapolated to a 10TB data set using a typical workflow over four weeks, ESG Lab validated that backup tasks would use 2.25X less infrastructure time with the Data Mover than with CIFS. ESG Lab also validated a significant reduction in CPU processing on the Veeam server during a synthetic full, when Data Mover processing is conducted on the ExaGrid array.

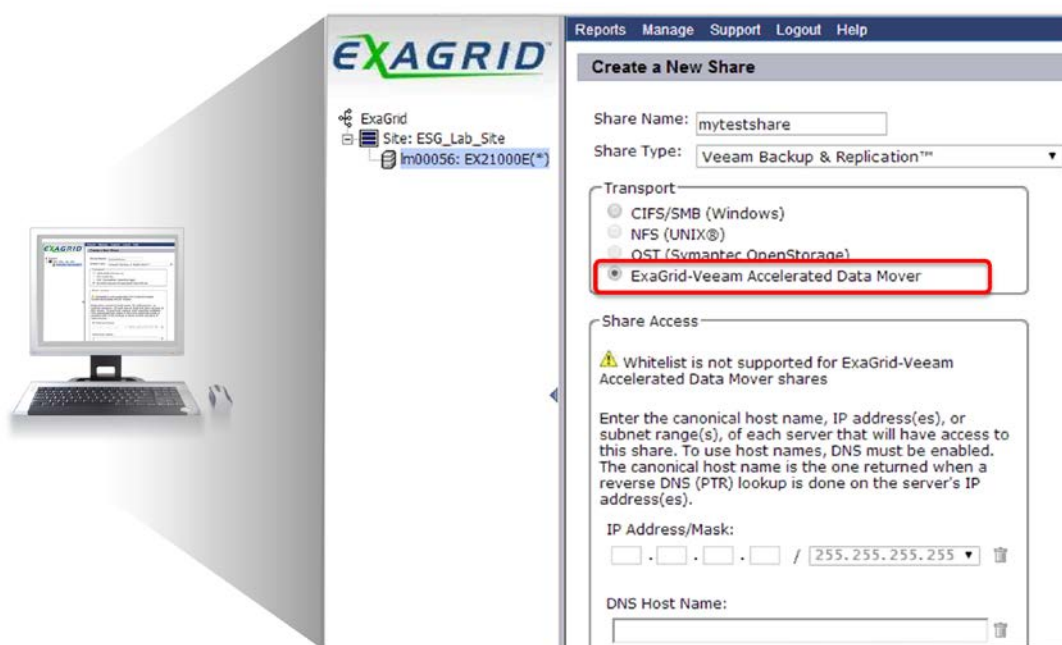
Finally, ESG Lab confirmed that instant recoveries with the Data Mover were just as fast as with CIFS, taking less than 3.5 minutes. ExaGrid’s landing zone and Adaptive Deduplication enable instant recovery to be accomplished in *seconds to minutes*. This is in stark contrast to in-line deduplication solutions that often take hours for an “instant recovery” because the data must be re-hydrated. This differentiator enables the ExaGrid solution to deliver multiple benefits: 1) makes synthetic full backups a data resource that can actually be used; 2) enables Veeam SureBackup, which verifies backup/restore viability; and 3) enables Veeam VirtualLab for creating isolated test environments.

Ease of Deploying ExaGrid-Veeam Accelerated Data Mover

Enabling the ExaGrid-Veeam Accelerated Data Mover is extremely easy. The feature is available to all customers with ExaGrid version 4.7 installed and a valid maintenance and support agreement. For organizations with multiple ExaGrid arrays, installation on one ExaGrid system will deliver the Data Mover to all ExaGrid arrays in the grid.

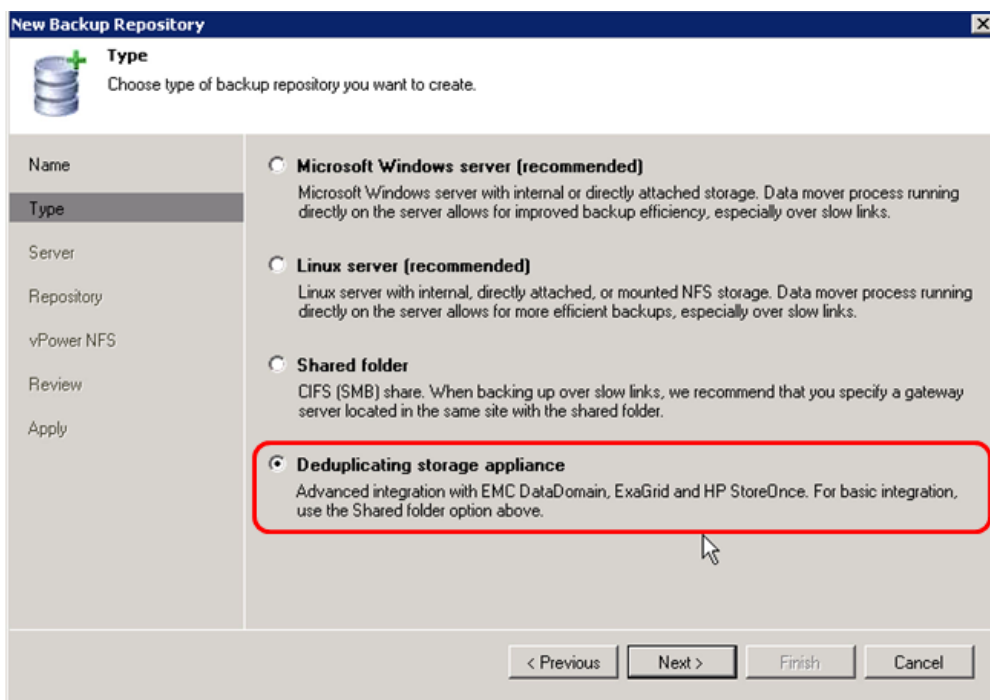
ESG Lab started by inserting security credentials on the ExaGrid EX21000E through the management console. Next, in the Manage screen we selected Create New Share, with the share name *mytestshare* and the share type Veeam. Since credentials had already been inserted, the list of Transport options included the ExaGrid-Veeam Accelerated Data Mover (Figure 8).

Figure 8. Create New ExaGrid Share



Next, the Lab accessed the Veeam management console. In the Backup Repository menu we selected Add Backup Repository, and when presented with Type options we selected Deduplication Storage Appliance (Figure 9). It should be noted that ExaGrid is one of only three disk deduplication appliances that have completed an advanced integration with Veeam to be included in this option. ExaGrid is the only vendor with integration to the Veeam Data Mover; others are proprietary technology integrations. Next, we selected the radio button for ExaGrid, added the DNS Name of the ExaGrid array (*esg-lab-test-accel-dm*), entered the same security credentials as on the ExaGrid console, completed SSH settings, and clicked Finish. The new repository is displayed in the Veeam GUI in the bottom of Figure 9.

Figure 9. Create New Veeam Repository



	Name	Type	Host
Backup Proxies			
Backup Repositories	Default Backup Reposit...	Windows	This server
WAN Accelerators			
Service providers			
SureBackup			
Application Groups			
Virtual Labs			
	esg-lab-test-accel-dm	ExaGrid	10.0.0.100
	esg-lab-test-cifs	CIFS	
	Win2012withDedup	Windows	msvm-srv37.apps...

Why This Matters

While most backup targets support the leading backup applications, not all targets are actually *integrated* with backup applications. That requires a level of partnership between the vendors, including engineering coordination and collaboration that results in additional features and simplicity of deployment. ExaGrid and Veeam have been engaged from early on, and as a result, the ExaGrid appliances are deeply integrated with Veeam solutions.

Their ongoing partnership has made it possible to shift some of the burden of backup processing to the disk-based backup storage using the ExaGrid-Veeam Accelerated Data Mover. ESG Lab validated the ease of deploying the Data Mover functionality. With security credentials and just a few mouse clicks, ESG Lab deployed the Data Mover on an ExaGrid EX21000E array, making it available to all arrays in the grid. This enables IT managers to simply “set it and forget it.”

The Bigger Truth

As your data grows and changes, and your data center evolves, so must your data protection processes. Virtualization and data growth continue to require that organizations pay attention to their backup and recovery schemes to ensure that they are reliable, scalable, and able to be completed with as little disruption as possible. Until recently, backup applications delivered the client, media server, and storage node in a consolidated bundle on a physical server; when that became overloaded, organizations needed to reproduce it to handle additional loads. Today, more modular, service-oriented architectures are available that distribute pieces of the process among various components.

The ExaGrid-Veeam Accelerated Data Mover is a great example of distributing various parts of the backup process to improve efficiency, scalability, and resource utilization. By placing the Veeam Data Mover on ExaGrid appliances, this solution shifts the burden of backup processing to the storage device, freeing up your production environment for production tasks. The result is faster backups, greater infrastructure efficiency, and less network congestion. By improving your backup process, you improve ROI and reduce operating costs.

Veeam is a leader in virtual machine backups with its simple management, deduplication, and compression, in many ways becoming the standard for VM backups. But every backup application needs a disk target, and ExaGrid provides not only a scale-out backup storage solution for growing data stores, but also built-in, Adaptive Deduplication to further reduce the amount of storage needed, and the unique landing zone that retains an un-deduplicated copy of the latest backup.

The landing zone and Adaptive Deduplication give ExaGrid a leg up on the competition. In-line deduplication solutions shrink the data going to the disk target over the network, but the “Achilles heel” of these solutions is the need to re-constitute the data during restore—that takes time, when you can least afford it. When you have a failure and need to access your data, in-line dedupe solutions will suffer a time penalty because of this rehydration. ExaGrid will not, as ESG Lab has validated more than once. In-line vs. Adaptive Deduplication is certainly every customer’s choice, but many don’t realize the significance. This solution gives back time and infrastructure resources to do more productive tasks. Data protection is absolutely essential—but the less bite it takes out of your infrastructure and time, the better.

ESG Lab validated the benefits of this solution. Our testing with a simulated Exchange data set demonstrated 1.6X faster full backups and 6X faster synthetic fulls using the Data Mover vs. traditional CIFS. Our testing also demonstrated that using the ExaGrid-Veeam Accelerated Data Mover could save significantly on infrastructure usage; extrapolating to a 10TB data set, using the Data Mover would reduce the hours needed for a month’s worth of typical full backups, incrementals, and synthetic fulls by more than 2X. And that’s without adding in the typical once- or twice-a-week restores/recoveries that backup administrators must do. ESG Lab also demonstrated significant network savings and reduction in server CPU usage when comparing CIFS backups to Data Mover backups. With ExaGrid’s scale-out architecture, these benefits accrue to *every appliance in a grid*, so as you scale you gain not only processor, memory, bandwidth, and disk resources, but also the backup speed and infrastructure efficiency benefits of the Data Mover.

ESG Lab is intrigued by the possibility of what else might be enabled by having a Veeam foothold inside the ExaGrid scale-out grid—could the future include some kind of data protection analytics, reporting, even managing compliance or e-discovery? Instead of only moving data around, could Veeam’s intelligence and the ExaGrid landing zone work together for other purposes? These tasks would be nearly impossible with in-line dedupe solutions, but ExaGrid could do it. Surely, if you want to look inside data sets for additional information, you’d rather do that on the backup target than on your production system.

If this validation has piqued your interest, then you should definitely give this new, easy-to-use technology a test drive. And if you are considering a competitive solution, give that a test drive too, so you can evaluate full and synthetic full backups, instant VM recovery, and restore. If you are looking for time and resource savings as well as truly instant recovery, the ExaGrid-Veeam Accelerated Data Mover may be your answer.

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. This ESG Lab report was sponsored by ExaGrid.

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 from time to time. 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.