

ESG Lab Review

Enterprise-class Cloud Data Protection with the NetBackup CloudCatalyst 5240 Appliance from Veritas

Date: November 2017 **Author:** Vinny Choinski, Senior Validation Analyst; and Brian Garrett, VP ESG Lab

Abstract

This ESG Lab Report documents hands-on validation and performance auditing of the Veritas NetBackup CloudCatalyst 5240 appliance with a focus on the value of extending an existing NetBackup infrastructure to the public cloud.

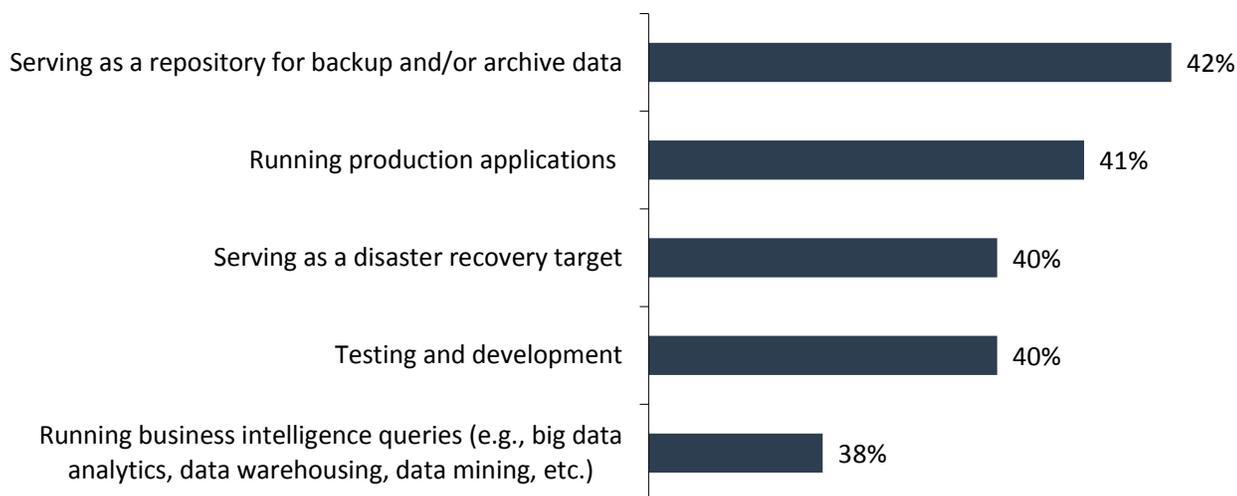
The Challenges

Today, applications and data are hyper-fragmented across multiple clouds, data centers, countries, and even continents. The result of this fragmentation can make it extremely complex and costly to tackle modern data protection challenges with traditional backup methods. Organizations that want to participate in the new digital era need to avoid cobbling together legacy and makeshift backup solutions to deal with the scale and performance challenges.

ESG research indicates that a growing number of organizations are using the public cloud to mitigate the capital and operational expenses associated with traditional IT hardware deployments.¹ And consistent with previously conducted ESG research in the area of cloud-based data protection (see Figure 1), data protection is the most commonly cited use case for cloud infrastructure services (IaaS and/or PaaS).²

Figure 1. Top Five Use Cases for Cloud Infrastructure Services

For which of the following purposes has your organization used cloud infrastructure services (IaaS and/or PaaS)? (Percent of respondents, N=430, multiple responses accepted)



Source: Enterprise Strategy Group, 2017

¹ Source: ESG Research Report, [2017 IT Spending Intentions Survey](#), March 2017.

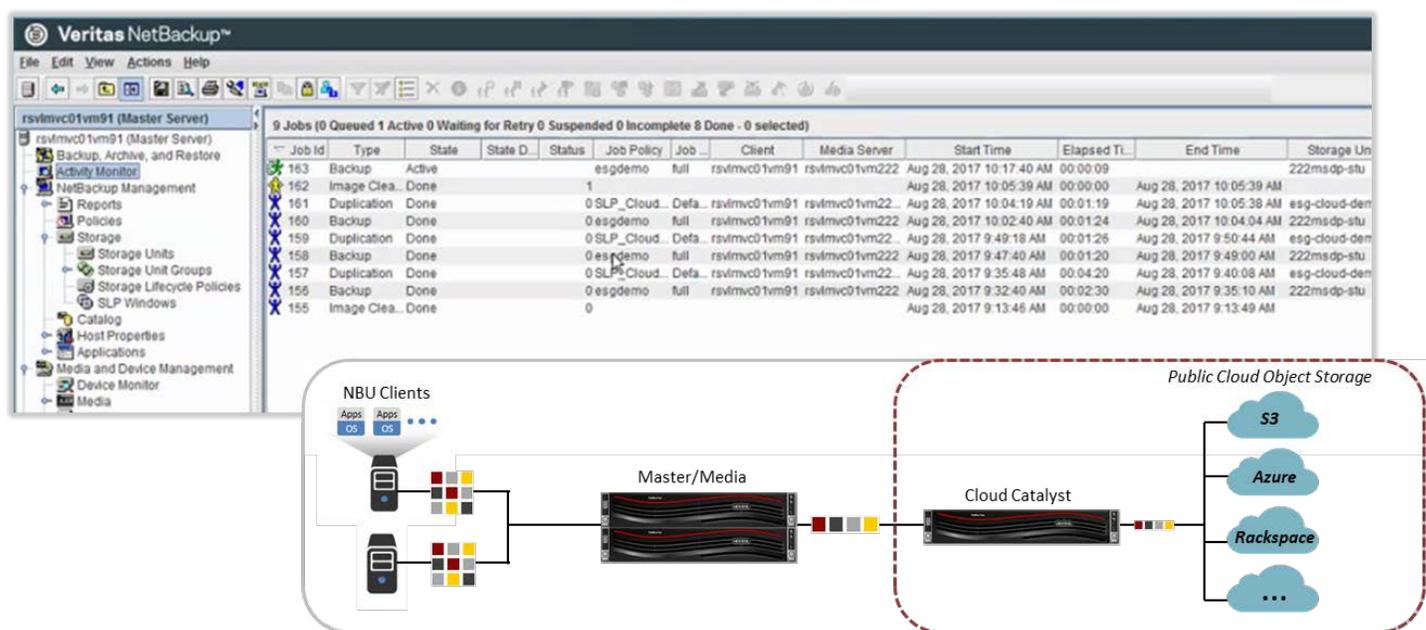
² *ibid.*

The Solution: NetBackup CloudCatalyst 5240 Appliance

The Veritas NetBackup CloudCatalyst 5240 appliance is a fully integrated enterprise backup appliance with intelligent end-to-end deduplication that extends into multi-cloud environments, effectively lowering the cost of cloud storage to make it a feasible option for long-term retention. It is a purpose-built backup appliance (PBBA) that helps customers address today's data center challenges by simplifying data protection and management.

As shown in Figure 2, the NetBackup CloudCatalyst appliance seamlessly integrates cloud storage into a customer's existing NetBackup ecosystem providing complete visibility and control so you can strike the right balance between your short-term operational and long-term recovery requirements and cost. It makes it easy to move archived and infrequently accessed data offsite. Data is moved in native NetBackup media server deduplication pool (MSDP) format to cost-efficient cloud storage for long-term retention. NetBackup automated backup lifecycle policies provide the orchestration required to move data to multiple clouds when and where you need it with complete visibility. Full integration with standard media servers and appliances enables critical data to be kept onsite on performance-optimized storage for immediate recovery of mission-critical data for operational recovery.

Figure 2. Solution Overview



Source: Enterprise Strategy Group, 2017

Key features include:

Visibility and Integration: Seamless visibility, deduplication, and cataloging of data between data center and multi-cloud environments through full integration with the NetBackup catalog.

Scalability and Performance: A modular architecture with fast end-to-end deduplication that scales into the public cloud of your choice.

Security: Encryption of data at rest in the public cloud with full auditability and control.

Investment Protection: The NetBackup CloudCatalyst appliance runs on the same field-proven 5240 platform as Veritas' flagship line of media server appliances.

ESG Lab Validated

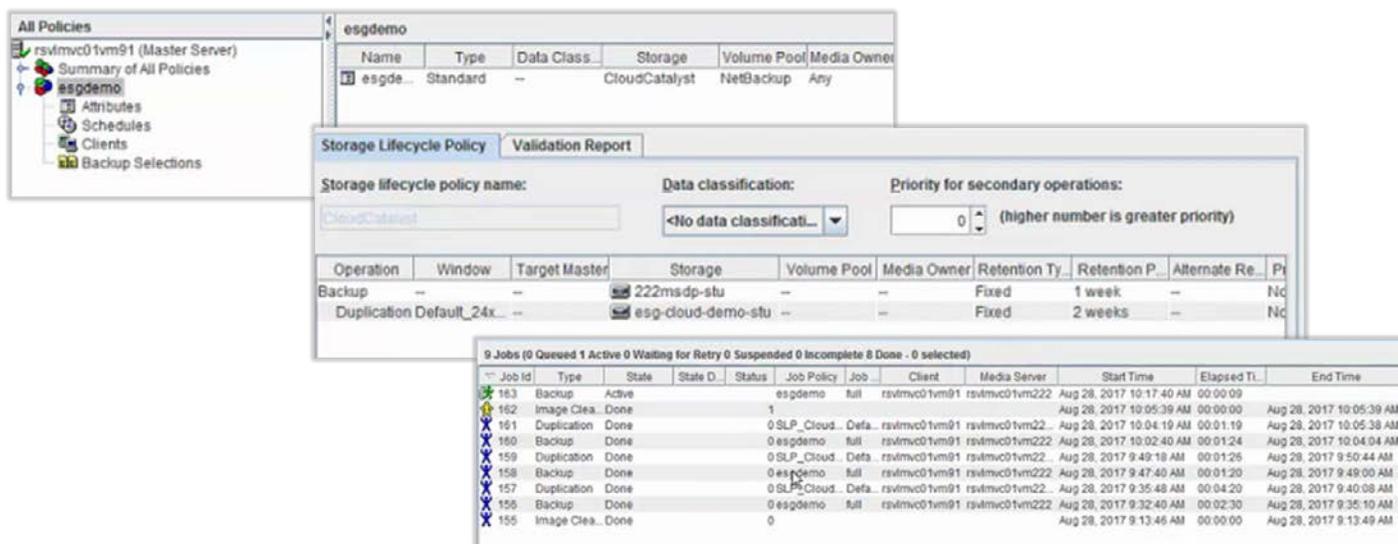
ESG Lab performed hands-on evaluation of the NetBackup CloudCatalyst 5240 appliance from our corporate office in Milford, MA by leveraging a remote Veritas demo environment. We also audited extensive performance testing results and reviewed test harness configurations supplied by the Veritas performance development team.

Easy NetBackup Integration

ESG Lab testing began with an exploration of how easy it is to extend the benefits of an existing NetBackup infrastructure to the public cloud. A remote demo environment with an existing NetBackup CloudCatalyst 5240 appliance and a standard 5240 media server appliance that was powered up and authenticated was used to review the policy schema and backup data workflow.

As shown in the upper left side of Figure 3, under the **ALL Policies** tab, we created a policy called *esgdemo* for a single Linux server in the test environment. The server had the NetBackup client agent installed and a simple directory structure full of data to be protected. We added the Linux server to the *esgdemo* policy as a client, selected the end-user file data as the Backup Selection, and created a schedule that enabled us to run ad-hoc backup jobs as needed.

Figure 3. Policy Schema



Source: Enterprise Strategy Group, 2017

As shown under the **Storage** tab in Figure 3, ESG used a storage lifecycle policy (SLP) called *CloudCatalyst* to manage the test backup data. An SLP is a storage plan for a set of backups. An SLP contains instructions in the form of storage operations to be applied to the data that is backed up by a backup policy. Operations are added to the SLP that determine how the data is stored, copied, replicated, and retained.

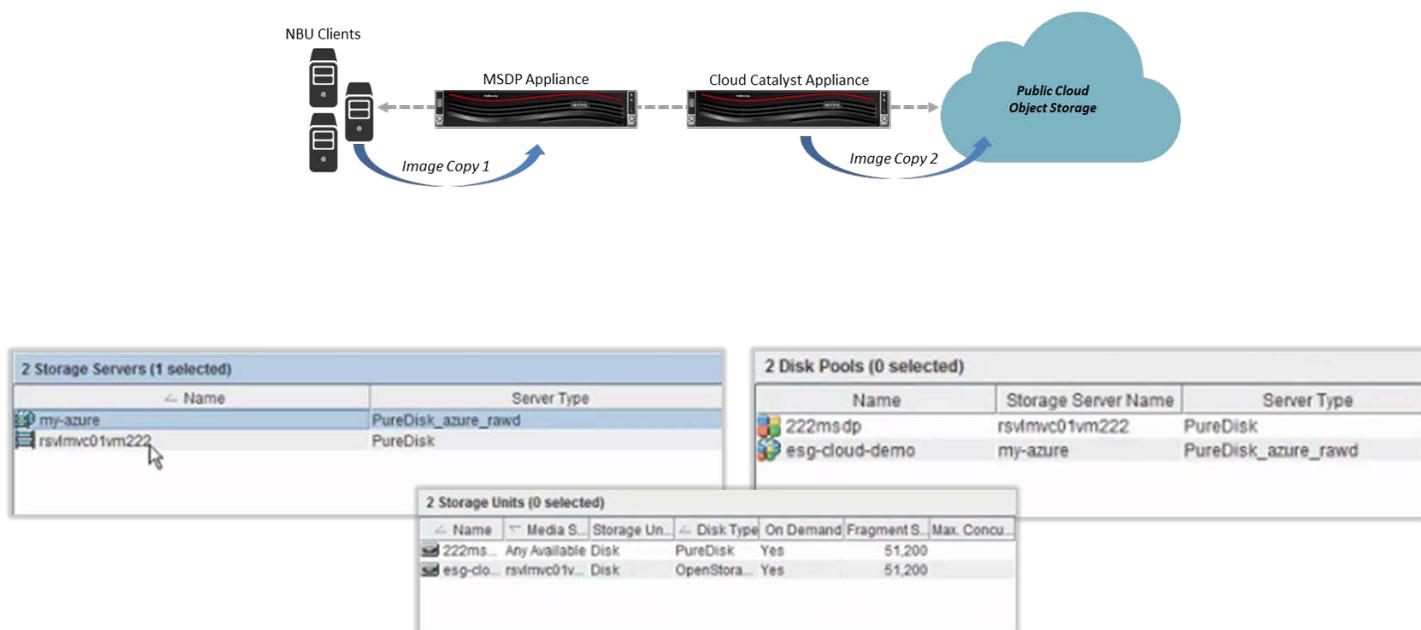
The **Storage Lifecycle Policy** tab in the middle of the figure shows the details of the *CloudCatalyst* SLP we configured for testing. With this SLP, backup data is first sent to a storage unit called *222msdp-stu* on a standard media server appliance where the data is deduplicated and stored for a retention period of one week. Then, as defined by the SLP, a duplication job is run to copy data to a storage unit called *esg-cloud-demo-stu* on the NetBackup CloudCatalyst appliance where the data is further deduplicated and then transferred in object format to a public cloud repository using a Veritas OST framework cloud connector.

As shown on the bottom right side of Figure 3, we used the jobs monitor view from the NetBackup management interface to track the status of the backup and duplication jobs. It should be noted that the jobs monitor view will not return a status for the duplication job until all the data has been moved to the cloud repository.

Next, ESG explored the storage workflow configuration including how the public cloud object storage repository was integrated. As shown at the top of Figure 4, the test environment consisted of three Linux servers with NetBackup client software, a classic MSDP appliance, a NetBackup CloudCatalyst appliance, and a Microsoft Azure Blob storage repository.

Next, we created two storage servers with a goal of exploring the backup storage workflow (see Figure 4). The first storage server, *rsvimvc01vm222*, was configured on the MSDP appliance. The second storage server, *my-azure*, was created on the NetBackup CloudCatalyst 5240 appliance. We also created two storage units with associated disk pools. The first storage unit *222msdp* was configured with a disk type of PureDisk, and the second storage unit *esg-cloud-demo* was configured with a disk type of OpenStorage.

Figure 4. Storage Workflow Configuration



Source: Enterprise Strategy Group, 2017

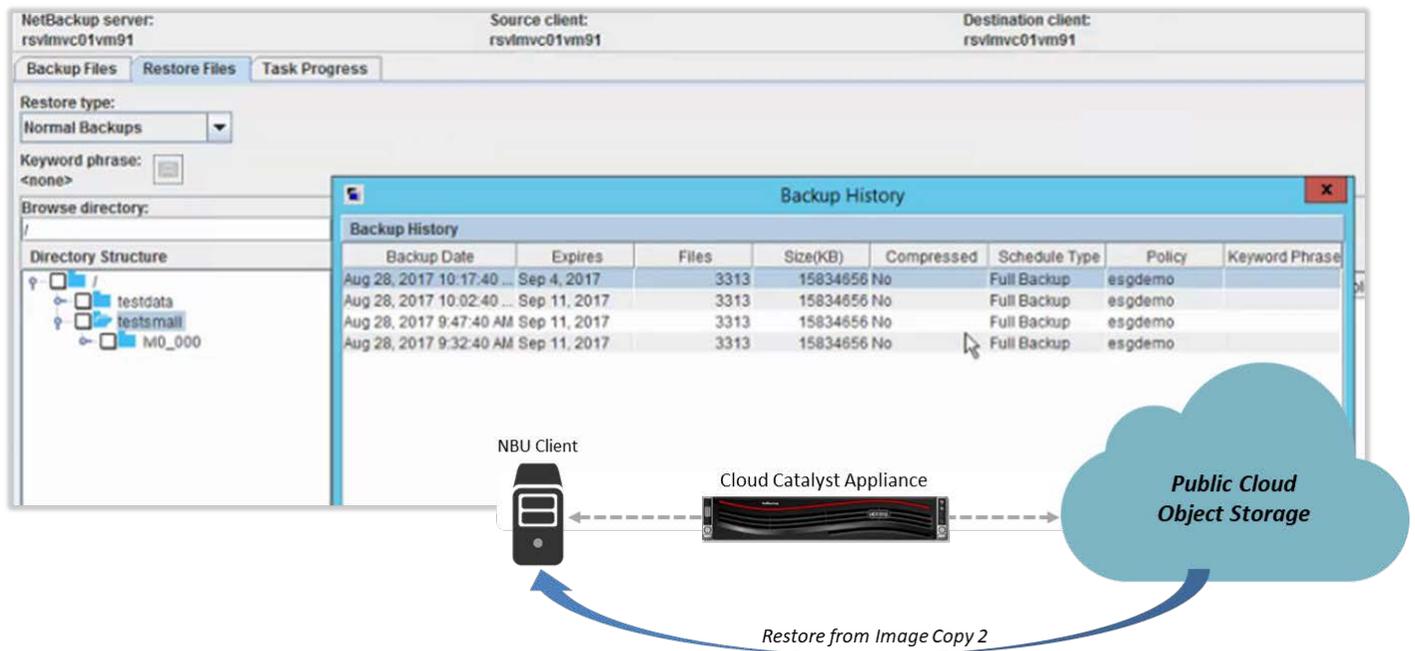
ESG configured the first storage server, storage unit, and disk pool combination in the test environment as a classic MSDP type target designed to receive, deduplicate, and store backup job data on-premises. The second storage server, storage unit, and disk pool combination was configured to receive duplication jobs from other media servers in the environment and send that data to a public cloud repository. This combination was configured with all the properties and values to communicate with a specific cloud service provider, in this case Microsoft Azure, and an OST cloud connector for data transfers.

ESG noted that the NetBackup CloudCatalyst appliance has deduplication capabilities similar to a standard 5240 appliance. Therefore, it could be leveraged for both deduplication and cloud data transfers in small environments.

Finally, as shown in Figure 5, we explored the ability to browse, select, and restore backup data stored in the public cloud directly from the NetBackup management interface. To demonstrate this process, we conducted a granular file level restore for a test environment Linux client that was backed up to the Microsoft Azure cloud via the NetBackup CloudCatalyst 5240 appliance.

Because all backup image data, including the object data that was stored in the public cloud, was recorded in the NetBackup catalogue, we could easily browse and select the files we wanted to restore. However, because the backup schema created multiple backup copies, including an onsite image, we had to delete copy number one from the catalog and flush the NetBackup CloudCatalyst cache to force a recovery from the public cloud repository. Once these steps were completed, we successfully conducted a restore directly from the Azure cloud repository to the Linux client.

Figure 5. Cloud-based Recovery



Source: Enterprise Strategy Group, 2017

Why This Matters

Things move and change fast in today’s always-on, always-up business environments. Today, you would be hard-pressed to find a data center that is not using some form of cloud compute or cloud storage to augment its on-premises infrastructure. To keep up with these changes, data protection solutions must be able to span these worlds efficiently without losing functionality.

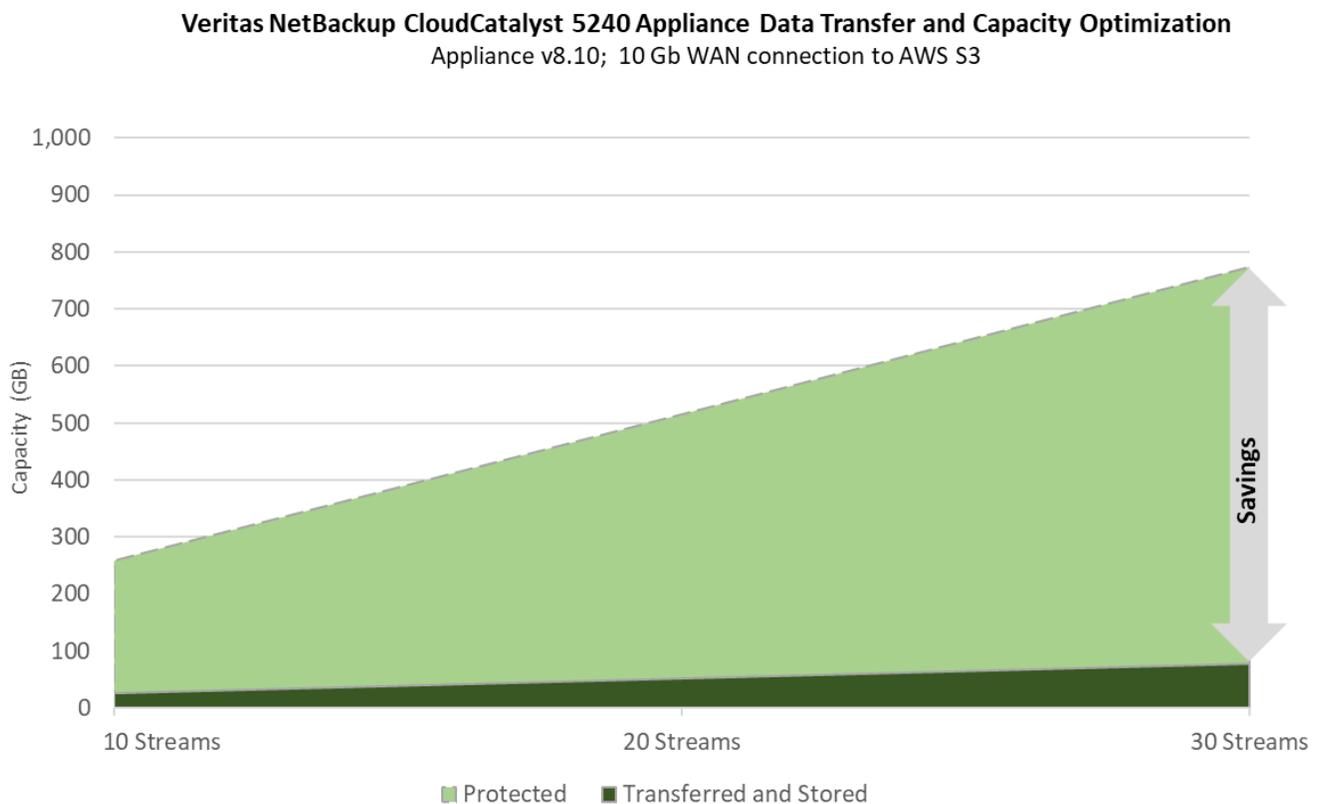
ESG confirmed that with the NetBackup CloudCatalyst 5240 appliance, you are not simply moving a bucket of data around the gameboard to minimally satisfy protection requirements. We confirmed that the solution delivers end-to-end visibility and management of both on-premises and in-cloud data. All of the policy and data lifecycle management capabilities are available for all backup images no matter where they are stored. This ensures you have all of your data, where you want it and when you need it the most.

Performance

Performance is an important component of any data protection solution that becomes even more important when applications and data centers are geographically distributed. Meeting backup and recovery requirements while dealing with the bandwidth and latency issues associated with moving data to and from a public cloud service can be very challenging.

ESG started its performance exploration by auditing the results of the NetBackup CloudCatalyst 5240 performance tests. Figure 6 shows the capacity optimization results that were reported after a series of backup jobs were run.³ The light green area shows the amount of data that was protected and the dark green area shows the data that was transferred over the WAN and stored on the public cloud after duplication and compression. Note how the savings increase as the number of backup streams increase. These savings apply to both cloud storage capacity and network bandwidth.⁴

Figure 6. Data Transfer and Capacity Optimization



Source: Enterprise Strategy Group, 2017

What the Numbers Mean

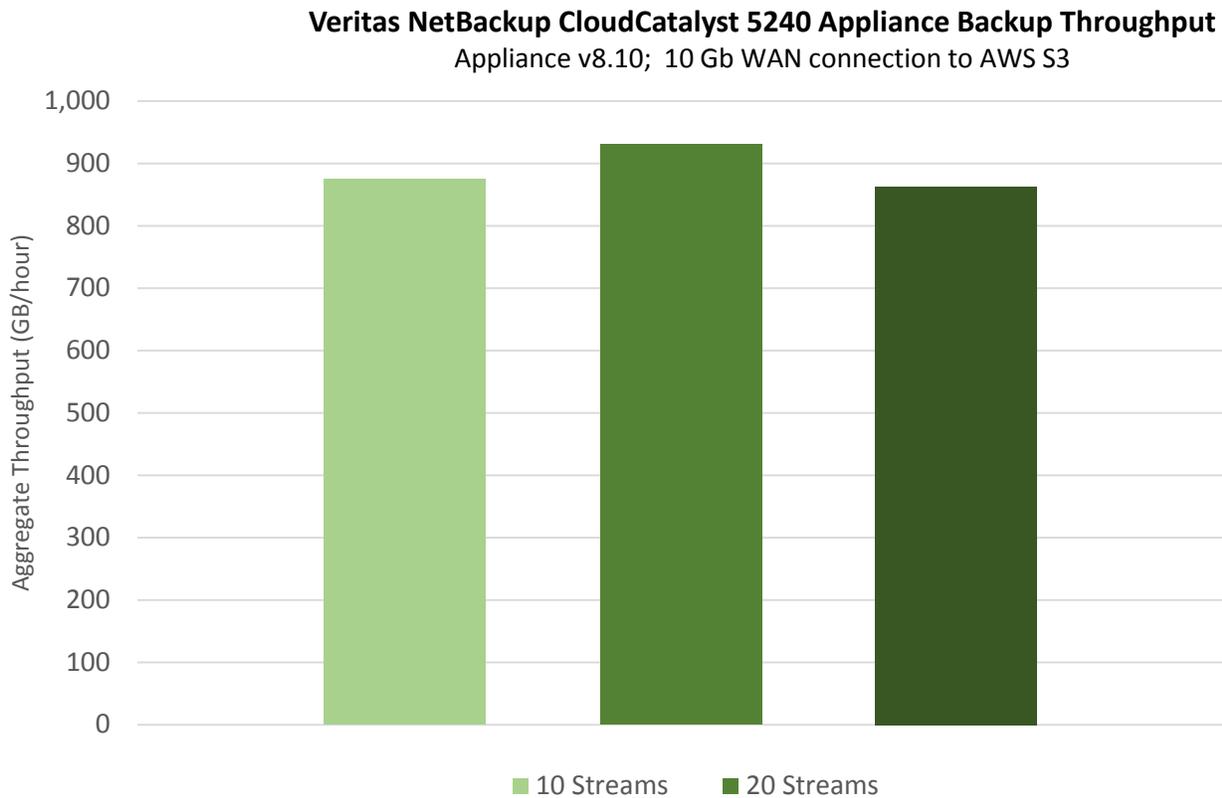
- Duplication and compression reduced the amount of data that was transferred over the WAN and stored in the public cloud by 90%.
- A 90% data reduction rate that reduces WAN bandwidth by a factor of ten to one (10:1) makes backup jobs complete up to ten times quicker with lower networking infrastructure costs.
- Based on ESG's experience with industry-leading data protection solutions, a data reduction rate of 90% is a conservative and achievable goal for most organizations.

³ A NetBackup 5240 Media Server appliance with 27 TB of storage and 192 GB of RAM was leveraged for first pass of deduplication.

⁴ A NetBackup CloudCatalyst appliance with 14 TB of storage and 192 GB of RAM was leveraged for a second pass of deduplication before transfer.

Next, as shown in Figure 7, ESG audited the throughput capabilities of the solution. Figure 7 shows the performance results as the number of backup jobs was increased from ten to 30 streams. The solution reached maximum throughput at 20 backup streams, and at 30 streams it started to show the effects of a network bottleneck.

Figure 7. Cloud Backup Throughput



Source: Enterprise Strategy Group, 2017

What the Numbers Mean

- The NetBackup solution with a NetBackup CloudCatalyst 5240 appliance scaled up to an aggregate throughput level of nearly one terabyte per hour (930 GB/hour).



Why This Matters

The throughput and cost of a WAN connection to the public cloud are vital considerations when architecting a data protection solution that extends into a public cloud. NetBackup CloudCatalyst is purpose-built with a goal of optimized WAN bandwidth utilization and reduced costs. ESG Lab confirmed that a NetBackup CloudCatalyst 5240 appliance can deliver nearly one terabyte per hour of aggregate backup performance (930 GB/hour). An audit of its utilization statistics (e.g., CPU, disk, and memory) confirmed that it was running at the peak performance potential given the amount of WAN bandwidth that was available to the public cloud. ESG also noted that if NetBackup CloudCatalyst is sending data from MSDP to the cloud, it will only process the unique bits that are sent from MSDP using the Optimized Copy functionality within NetBackup. This is different from solutions that go through a rehydration process, and then deduplicate data again in a different format prior to sending to cloud. This also means that a physical throughput rate of 930 GB/hour can translate to as much as 17 TB per hour if on-premises deduplication rates reach 95%.

The Bigger Truth

The future of data protection is data management. For organizations to grow beyond “data protection,” which is inclusive of backups, snapshots, and replication, they need something that most backup products do not have: end-to-end contextual awareness and intelligence on the information within the data being protected. A mature data management solution provides insight into the information contained within the data being protected/preserved, and then automates the determination of policies regarding retention, destruction, and resiliency/availability of the data.

Of equal importance, especially as corporate data continues to span multiple data centers and multiple clouds, is the efficient movement of protection data across the entire ecosystem. In fact, when ESG survey respondents were asked to identify the most important data protection mandates from IT leadership, two of the top three most-cited responses were increased speed/agility of recoveries and increased speed or frequency of backups.⁵

ESG confirmed that the NetBackup CloudCatalyst 5240 appliance seamlessly integrates with NetBackup with a goal of delivering end-to-end visibility and management of both on-premises and in-cloud data. We were able to easily implement the same data protection and storage lifecycle management policies across on-premises and public cloud storage repositories alike. The cataloging integration extended right through to the recovery process as individual files were restored from a Linux server backup directly from a Microsoft Azure public cloud. ESG Lab confirmed that NetBackup configuration with a single NetBackup CloudCatalyst 5240 appliance can sustain backup rates of up to 930 GB/hour with capacity and WAN bandwidth reduction of 90% or more.

For data protection professionals, keeping pace with today’s always-on, always-up business environments can be very challenging. Applications and data have become hyper-fragmented and data centers have become hyper-distributed. And some factors are outside the control of the solution, like bandwidth, latency, and data access and movement fees charged by cloud service providers. We believe Veritas has done a great job recognizing and addressing these challenges and the NetBackup CloudCatalyst 5240 appliance is a great example of this. However, even with the best technology, you cannot overlook the importance of proper planning. Knowing where your protection data lives and how and where you plan to recover that data is key. ESG looks forward to Veritas’ continued qualification of cloud connectors and the packaging of the CloudCatalyst appliance into a virtual edition. We believe these efforts will add even more agility to an already nimble solution.

⁵ Source: ESG Research Survey, *2017 Data Protection Modernization Trends*, December 2016.

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.