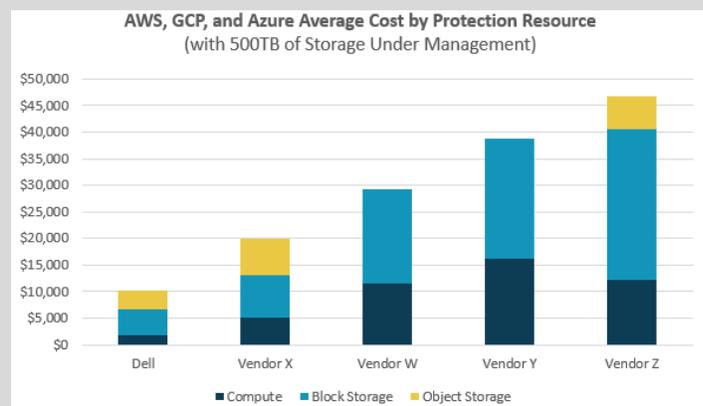ESG Economic Validation

Understanding the Economics of In-cloud Data Protection:

A Dell Technologies Data Protection Solution Designed with Cost Optimization in Mind

By Vinny Choinski, Senior Validation Analyst; and Christophe Bertrand, Senior Analyst
September 2021

Executive Summary

Implementing and operating data protection technology in public cloud environments can be complex and costly. IT teams often struggle to manage backup solutions designed for on-premises environments that are "shoe-horned" into public cloud deployments and it is not uncommon for organizations to incur large cost overruns when using public cloud infrastructure. ESG validated that Dell Technologies data protection solutions are extremely cost-efficient when running in any public cloud environment. Dell EMC PowerProtect Data Manager in tandem with Dell EMC PowerProtect DD Virtual Edition delivers enterprise-grade data protection features and performance that organizations are looking for to support their public cloud production environment protection needs. The solution provides a combination of features such as file-level recovery,

AWS, GCP, and Azure Average Cost by Protection Resource (with 500TB of Storage Under Management)

next-generation backup services, near-zero RTO restores, disaster recovery capabilities, and backup data management while managing cloud resources very effectively to reduce costs. ESG compared a deployment of the Dell Technologies solution to four top competitors in Amazon Web Services (AWS), Microsoft Azure, and Google Cloud Platform (GCP) to validate its cost advantages. A reduction of cost of more than 80% was confirmed during ESGs analysis of multiple modeled scenarios.

Introduction

This ESG Economic Validation quantifies the savings and benefits organizations can achieve when utilizing Dell EMC PowerProtect Data Manager (Data Manager) as their cloud IaaS data protection solution. ESG audited Data Manager cost model metrics and leveraged a scenario that factored in all the storage and compute costs of a typical data protection schema at 12 different storage levels between 10TB and 500TB production levels.

Background

The economics of cloud computing has helped drive massive adoption. However, there are many factors, both economic and technical, that can derail an organization's cloud migration strategy. In fact, an ESG research study showed that 32% of respondents cited cost overruns as a major reason for moving an application back on premises while another 29% cited unpredictable costs. From a technical perspective, 26% of respondents identified difficulty providing adequate protection of applications or data.[1]

Figure 1. Reasons for Moving Cloud Workloads back on Premises

Which of the following led your organization to move workloads running on cloud infrastructure back on premises? (Percent of respondents, N=875, multiple responses accepted)

Cost(s) exceeded expectations	32%
Difficulty implementing security measures	32%
Unpredictable costs	29%
Limited access to new technologies	29%
Inability to support scalability requirements	27%
Difficulty providing adequate protection of applications or data	26%
Inability to meet availability expectations	26%
Experienced a security breach/incident	26%
Inability to meet functionality or usability expectations	25%
Difficulty maintaining or proving compliance	25%
Inability to meet elasticity expectations	24%
Poor or unpredictable application performance	22%

Source: Enterprise Strategy Group

Organizations of all sizes continue to leverage the public cloud to give them the speed and agility they need to drive digital transformation. Two popular cloud use cases are backing up on-premises and edge workloads to the public cloud and protecting cloud-based applications in the cloud.

It is critically important for organizations to pay close attention to all the factors, technical as well as financial, prior to migrating applications and data to the cloud in order to attain the desired outcomes.

---

[1] Source: ESG Custom Research Survey Commissioned by Dell Technologies, *Dell EMC VMware Intel Hybrid Multi-cloud Cornerstone Survey*, February 2020.

Dell Technologies Solution

Dell EMC PowerProtect Data Manager with Dell EMC PowerProtect DD Virtual Edition

Data Manager offers unified data protection for the enterprise that centralizes, automates, and accelerates backup and recovery across the entire IT environment. This includes providing the same quality of data protection for cloud-hosted applications as for applications running in on-premises data centers. Deployment of Data Manager is accomplished by using virtualized appliances, both for the centralized server controlling policies and schedules and the storage repository.

This integrated solution also includes a client-based feature called *DD Boost,* which integrates with the file system and application backup processes running on Linux and Windows clients. DD Boost enhances client direct backup to the PowerProtect DD Virtual Edition (DDVE) storage repository. This reduces the overall solution cost by distributing the data movement workload across the client virtual machines, which reduces the need to provision dedicated data mover or backup appliance node virtual machines.

Figure 2. PowerProtect Data Manager with PowerProtect DD Virtual Edition Workflow
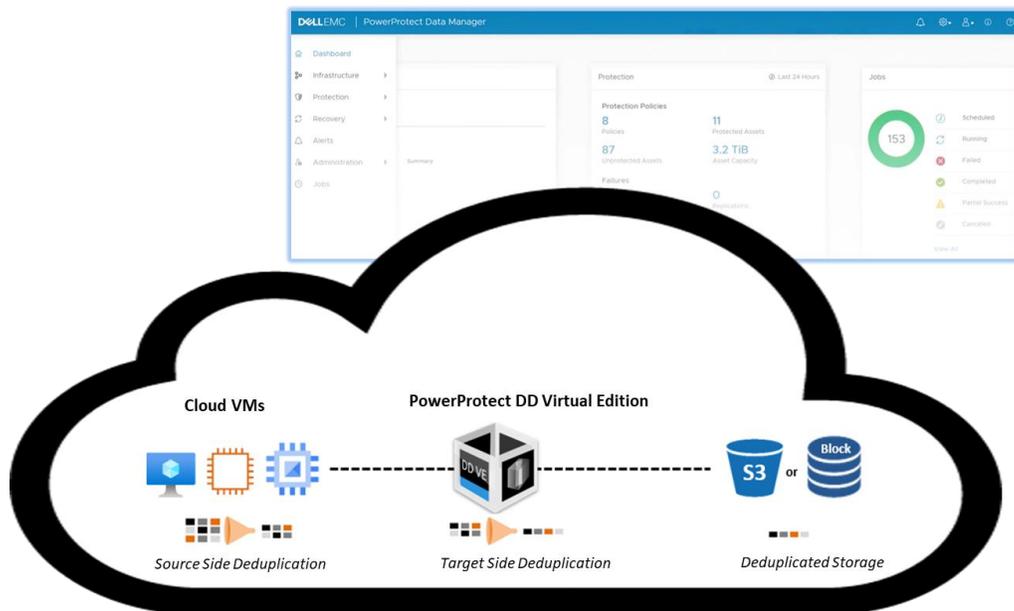
Source: Enterprise Strategy Group

Figure 2 demonstrates the workflow for Data Manager running in AWS, Azure, and GCP with data flowing from a VM to DDVE, where it is then stored on object or block cloud storage. A key part of the workflow efficiency is deduplication at all phases, which results in a very low TCO, as demonstrated in this ESG Economic Validation. Some key architectural features include:

- **PowerProtect Data Manager:** This is the core component of the solution and supports the policy, scheduling, and catalog functions to manage backup and recovery and provide source-based deduplication.

- **PowerProtect DD Virtual Edition Server:** This is a software-defined data protection appliance running on a virtual machine leveraging object and block storage. DDVE provides data management functions, including variable length deduplication, data integrity, and integration with Data Manager Software.

- **DD Boost:** DD Boost is a client-based software component that enables the client to efficiently back up and retrieve deduplicated data directly to and from DDVE.

ESG Economic Validation

ESG conducted an economic analysis of the use of Data Manager as an IaaS data protection solution in three public clouds. Testing focused on the economic benefits organizations can expect with Data Manager compared to four competitive solutions operating in AWS, Azure, and GCP cloud environments.

The economic validation process leverages ESG's core competencies in market and industry analysis, forward-looking research, and technical/economic validation. ESG started the process by conducting in-depth research to better understand and quantify how Data Manager and competitors operate in public clouds. The findings were used as the basis for a simple economic model comparing the costs of operating Data Manager to its competitors operating in the top three public clouds.

Economic Value Overview

As more organizations move production applications to the cloud, they often find that they need to run the same data protection workflows in the cloud as they do on-premises, for consistency and ease of management. These workflows commonly include application-aware backups with application-consistent restores that support other requirements such as disaster recovery, business continuance, test, and development. Organizations, however, will likely not be able to achieve these types of recoveries or meet SLAs using the native data protection services provided by cloud providers. These services typically rely on volume-based snapshot technologies, which often cannot meet the consistency and granularity of data recovery required by key business applications, as shown in Figure 2.

AWS Cost Model

Using research data and publicly available pricing, ESG first modeled Data Manager compared to four competitors running in AWS. From an architectural perspective, each of these solutions can be deployed in the cloud as virtual appliances or virtual machines, leveraging block and object storage where applicable.

Figure 3. AWS Data Protection Analysis
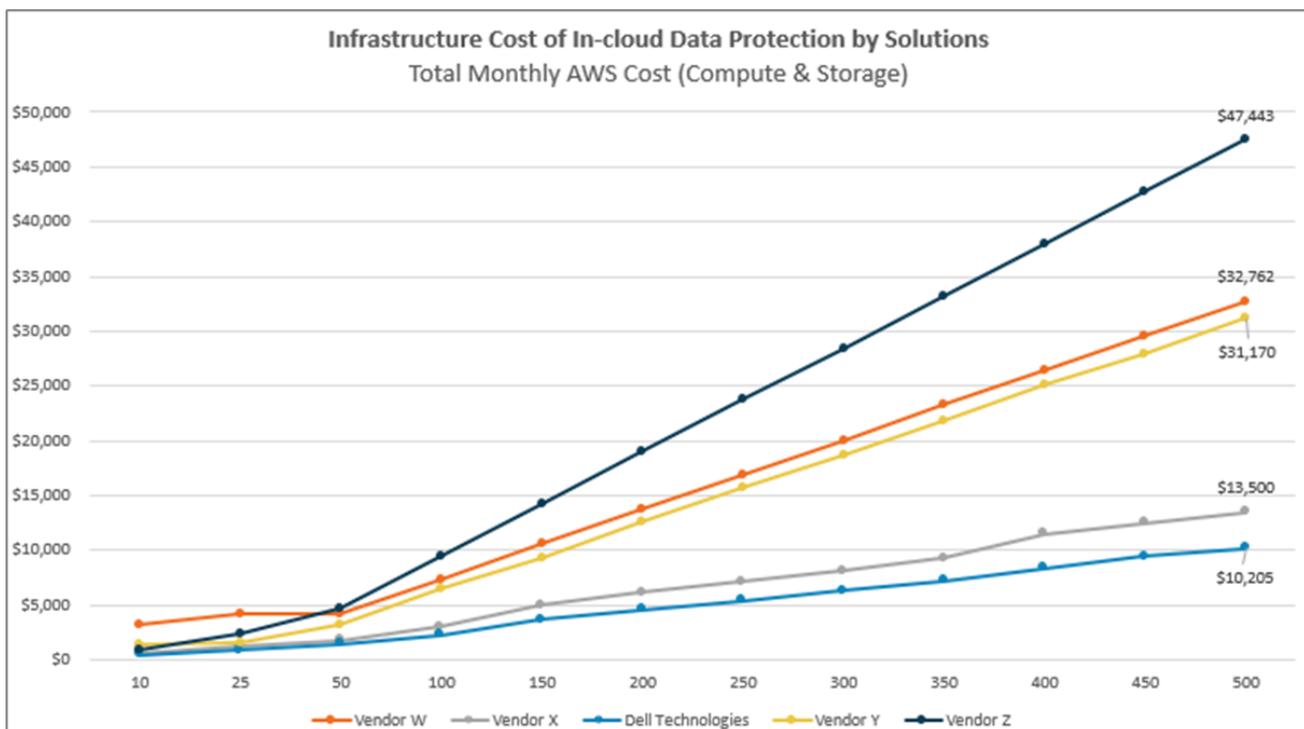
Source: *Enterprise Strategy Group*

The modeling is based on the amount of production data to be protected and the resources required by each solution. The analysis was done for each vendor and includes three critical cost components: compute, block storage, and object storage for each vendor running as a complete virtual appliance or as data protection software on virtual machines. Overall data protection resource efficiency for each solution is heavily dependent on each vendor's architecture design concepts.

As shown in Figure 3, ESG's analysis compares the Data Manager solution against Vendor W and Vendor Y, which do not efficiently handle storage repository requirements in their designs, and Vendor Z, which has the highest cost due to its very high compute requirements. For Vendor X, Data Manager started off equal in costs, but as data grew, the Dell Technologies solution developed a slight cost advantage.

Figure 3 shows that the Dell Technologies solution delivers up to a 78% reduction in total monthly costs for in-cloud data protection infrastructure over its least efficient competitor. This represents a savings of $37,237 a month and $446,850 annually at the 500TB protection level. The effect of each component (compute and storage) on the overall cost of the infrastructure is explained in more detail in the subsequent ESG Economic Analysis section of this report.

Azure Cost Model

Next, as shown in Figure 4, ESG continued its data protection analysis based on the same solutions deployed in the Azure cloud. In the Azure-modeled scenario, ESG implemented the vendor solutions based on their suggested architecture recommendations to compare the cost of each vendor when running in Azure.

Figure 4. Azure Data Protection Analysis
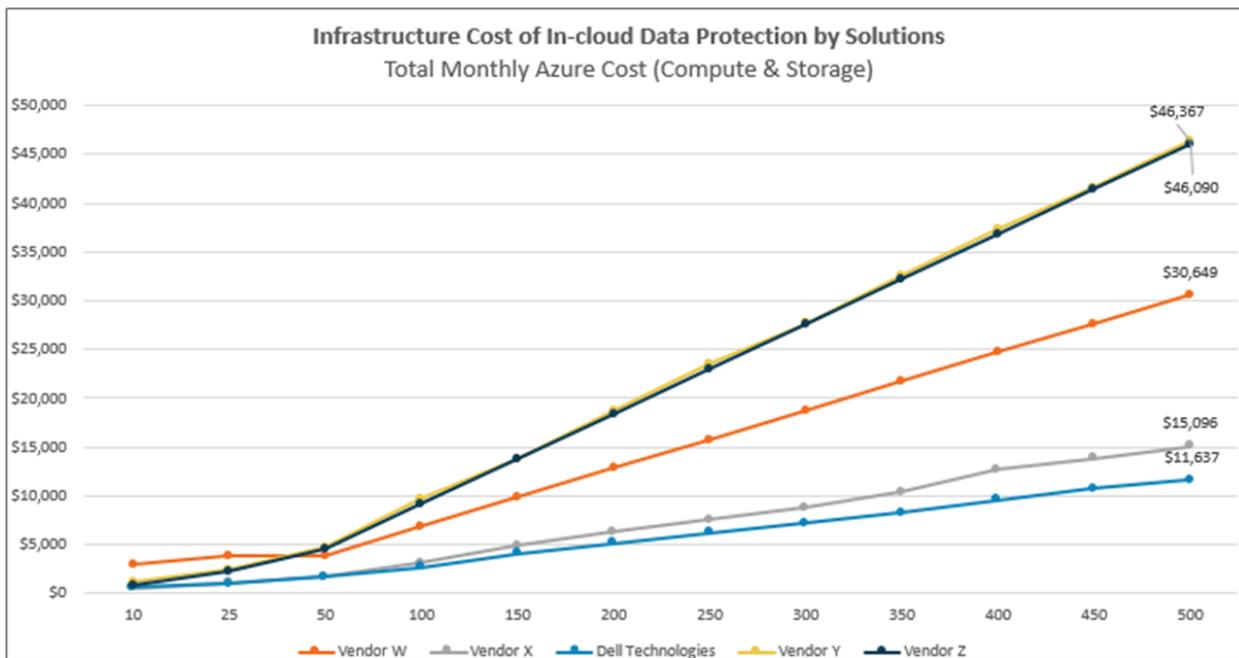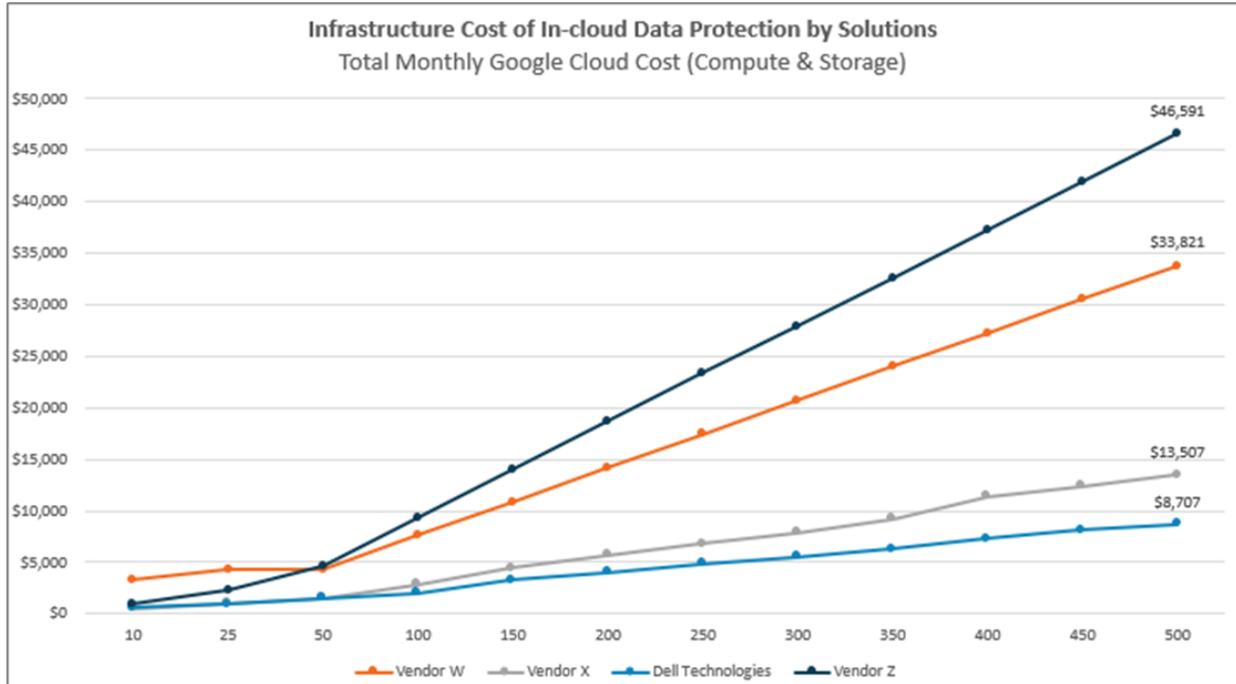

Source: Enterprise Strategy Group

Here, the Dell Technologies solution continues to incur the lowest overall costs as more data is managed over time. In the AWS model, Vendor Z had the overall highest costs. However, we now see that Vendor Y has nearly the same cost structure. This is due to higher compute requirements on Azure than they had on AWS. Vendor W also continues to have higher costs due to less efficient storage utilization but is more efficient in compute than Vendors Y and Z. Vendor X remains closer to Data Manager, but again, as data growth increases, the cost gap continues to widen. It is anticipated that this trend would continue with more saving from the Dell Technologies solution as more data is stored. The overall savings

for Data Manager in Azure is 74.9% and represents a cost benefit at 500TB of $34,730 per month with an annual savings of $416,762.

Google Cloud (GCP) Cost Model

As shown in Figure 4, ESG applied the same analysis of Data Manager with three of the competitors when deployed on the Google Cloud Platform (GCP). At the time of this analysis, Vendor Y did not support GCP deployment.

Figure 5. GCP Data Protection Analysis

Source: Enterprise Strategy Group

In the GCP-modeled scenario, ESG implemented vendor solutions based on their suggested architectures to compare the cost of each vendor. The Dell Technologies solution continued to be the most cost-effective and efficient solution, with a cost savings of 81.31%, which equates to a monthly savings of $37,884 and annual savings of $454,610 at the 500TB level.

ESG Economic Analysis

In order to further validate the cost savings of the Data Manager solution, ESG leveraged the economic modeled scenarios created on AWS, Azure, and GCP to determine where the cost savings were being derived. In each environment, the vendors utilized complete virtual appliances or data protection software running on virtual machines with the options of using block or object storage. In some cases, vendors did not support lower cost object storage. The vendors also supplied a variety of recommendations pertaining to what servers were required and the amount of compute, as well as when to increase compute to handle different levels of protection data processing. Another major factor is the use of features within the vendors' software such as deduplication technologies. These factors were all considered in the economic model to determine each vendor's efficiency in each of the three cloud environments.

Compute Efficiency

ESG analyzed the compute resources required to deliver enterprise-class data protection at scale for cloud-based production environments. The analysis included the cost of the compute resources required for a data protection solution to handle the load of backup jobs as the production environment was scaled from 10 to 500TB. The backup load was

modeled on a 70% filesystem and 30% database workload distribution with a 28-day retention period. A daily change rate of 1% was used for filesystem backups and 3% was used as the daily change rate for database backups. The resource modeling was based on reference architecture guides and publicly available pricing.

Storage Efficiency

A major component of any in-cloud, hybrid, or on-premises data protection solution is the backup repository. This is where the backup images of protected data are stored. If not managed efficiently, the backup repository can quickly grow very large as more backup jobs complete over time, new systems are added to the protection schema, and the amount of production data naturally grows over time. Object storage has become an important cost savings element as organizations leverage more cloud-based features in their data protection solutions. The Dell Technologies solution leverages its proprietary deduplication technology between virtual machines, DDVE, and block or object storage to deliver backup repository efficiency. For enterprise-level data protection, each virtual machine runs client software, which includes DD Boost agent libraries, for client to DDVE data movement and deduplication efficiency.

As shown in Figure 6, ESG broke down the major deployment cost components. A blended average for AWS, Azure, and GCP were created for compute, block storage, and object storage for Data Manager in comparison to the other four vendors at a mid-storage utilization point of 250TB.

Figure 6. Three Cloud Average Data Protection Resource Cost Analysis
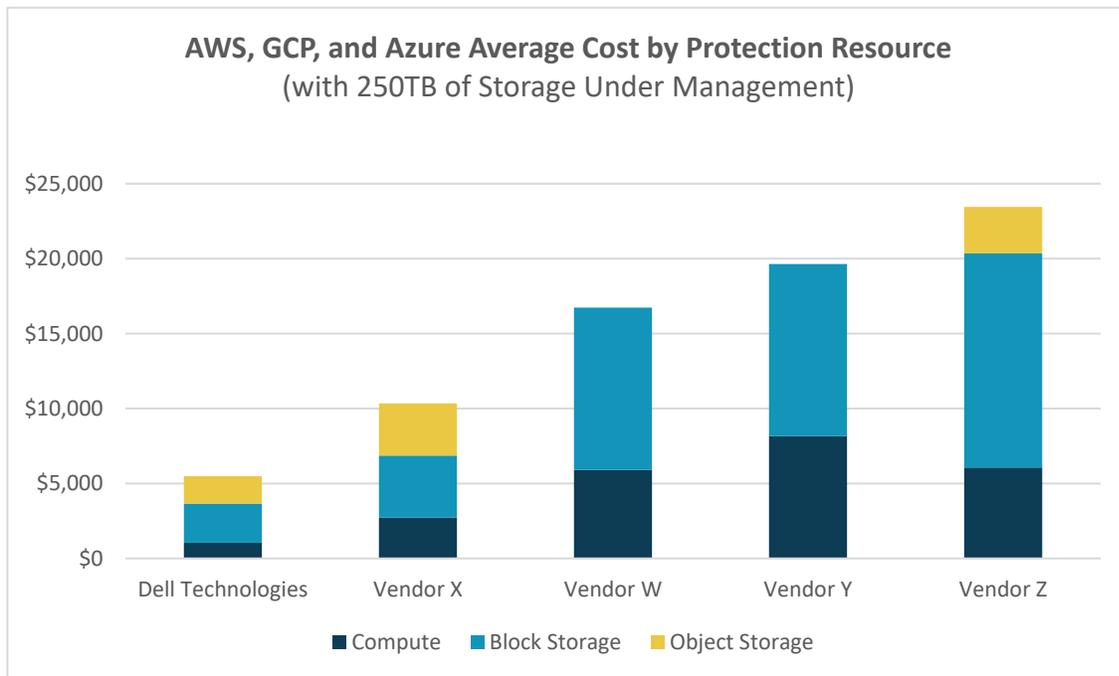
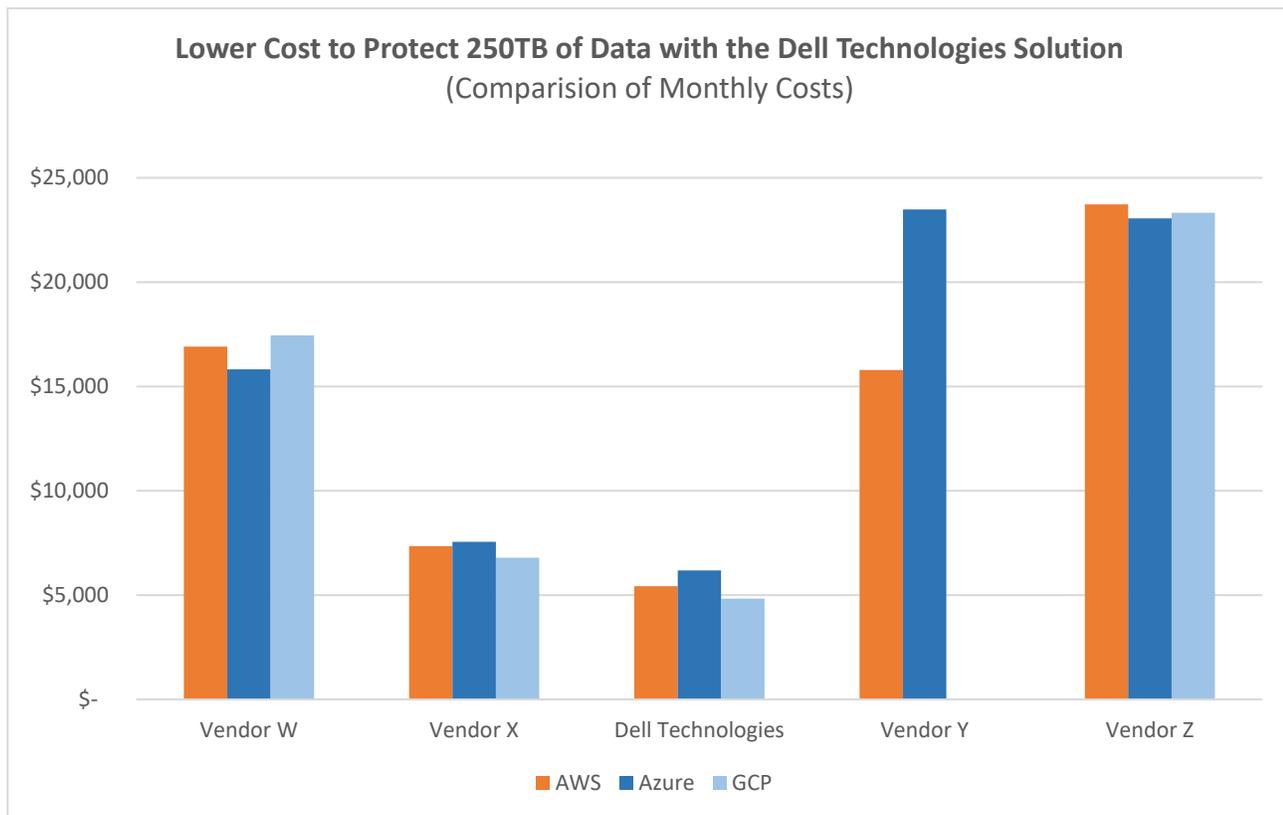

Source: Enterprise Strategy Group

As Figure 6 shows, the Dell Technologies solution effectively utilized the available resources in all three cloud environments and proved to be the most efficient vendor on compute, block storage, and object storage across the majority of backup data sets, including the 250TB data set level shown here.

The Data Manager distributed schema delivers compute efficiency. As a data protection solution, its architecture is able to process more workloads on lower cost servers with its direct client backup feature. Many of the other vendors require more compute resources and larger servers and have lower breakpoints where additional servers are required as the environment scales. This enables organizations using Data Manager to keep compute costs low. As shown in Figure 6, Data Manager is 62% lower in compute costs than Vendor X, its closest competitor at the 250TB mid-point in the data analysis.

Regarding storage utilization, the Dell Technologies solution leverages deduplication at the source side and the target side back-end storage to keep costs down, while also utilizing lower cost object storage without sacrificing performance. The solution enables each EC2 client to perform its own data deduplication to efficiently send only the unique data blocks directly to the DDVE appliance. Backup processing metadata, such as deduplication chunk indexing, is sent directly to the backup server. This schema decouples backup data transport from backup processing tasks for improved performance and storage efficiency. For block storage efficiency, Data Manager is 37% less than Vendor X, its closest competitor, and 48% less on Object storage usage.

Another view of efficiency for Data Manager with DDVE solution over the four competitors in the three public clouds is shown in Figure 7. In this ESG-modeled scenario, we looked at the overall cost of each vendor to operate in these environments based on their deployment specifications. This view is a snapshot at 250TB of data under management. Similar to the other models, Dell Technologies has a clear cost advantage in all three clouds. Compared to the highest cost competitor, Dell Technologies demonstrates a 77% monthly cost reduction. Compared to its closest cost competitor, the Dell Technologies solution is 24% less expensive on a monthly basis across all three cloud environments.

Figure 7. Overall Cost Analysis for each Cloud with 250 TB of Data Under Management

Source: Enterprise Strategy Group

What the numbers mean:

- At 250TB of protection data under management, roughly the mid-point of the model, the Data Manager with DDVE solution delivers a 77% cost reduction over Vendor Z. This is approximately $17,900 a month and $214,800 annually in savings based on the total data protection application resources required.

- These costs savings are due to the decoupling of metadata from the backup data transport process and the efficient metadata processing provided by the DD Boost running on the client systems. In addition, highly efficient

deduplication processing (client-side via DD Boost and target-side on DDVE) resulted in a much higher deduplication ratio on S3 storage for the Dell Technologies solution.

The Bigger Truth

These days, there is a lot of pressure on data protection professionals to deliver seamless access to business data and applications. And the recent increase of cyber-threats has only added to the pressure on data protection teams. It doesn't matter if the information exists on-premises or in the cloud, organizations expect uninterrupted access to their business assets, and, if data needs to be restored, they expect quick and complete recovery from a cost-efficient data protection solution.

ESG validated the many benefits of the Data Manager with DDVE solution for public cloud IaaS environments by reviewing data protection architecture requirements, analyzing capacity utilization capabilities, and auditing detailed pricing data with a competitive comparison focus. ESG's modeled scenario shows a drop of more than 80% in cost to protect data in the cloud using the Dell Technologies solution. This cost savings is highly driven by the Dell Technologies solution's flexible deployment capabilities, client direct backup features, mature source and target deduplication technology, and efficient use of object and block cloud storage. ESG also found the solution delivered these cost benefits while still providing a solution that offers application-aware backup and recovery.

If you are looking for a cost-efficient data protection solution for your public cloud IaaS environments that can also meet your business SLAs and recovery consistency requirements, we highly recommend that you consider exploring the features and capabilities of the Data Manager with DDVE solutions from Dell Technologies.