

Analyzing the Economic Benefits of Datrium Cloud DVX

By Kerry Dolan, Senior IT Validation Analyst

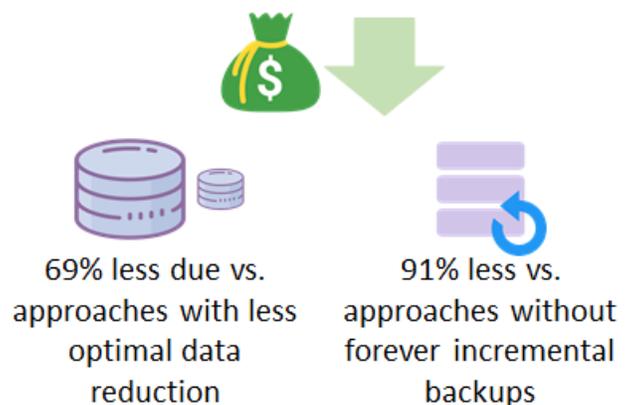
May 2018

Executive Summary

It can be a juggling act for IT administrators to keep data fully protected, make it quickly available when needed, and still remain affordable. Some organizations are looking to cloud backup to simplify data protection operations and facilitate fast restore, but storage costs can be high. Datrium Cloud DVX offers a backup solution to AWS for VMs stored in on-premises DVX “open converged” systems; Cloud DVX reduces storage capacity and AWS costs based on its data reduction and incremental backup methods. ESG audited Datrium TCO calculators that compared Cloud DVX with alternative cloud backup deployments, demonstrating Cloud DVX advantages.

ESG’s audit revealed that Cloud DVX offers a distinct advantage based on its on-premises and in-cloud global deduplication and compression; forever incremental backups after the initial seed; and elimination of secondary site and on-premises hardware, floor space, power, cooling, networking, and administration.

Cloud DVX Cost Advantages:



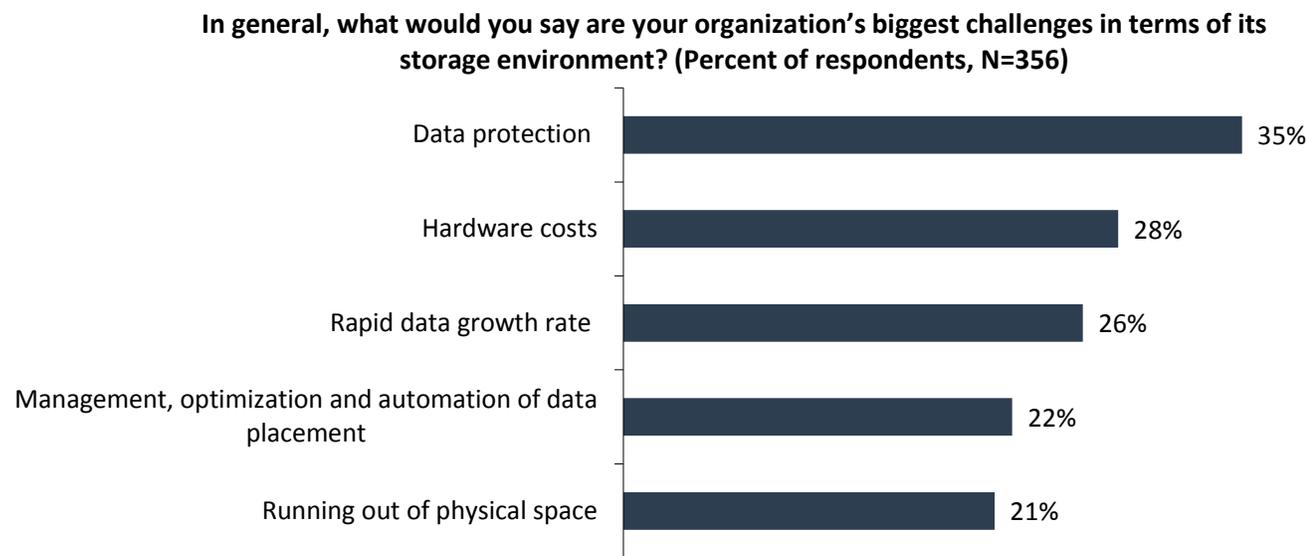
Introduction

This ESG Economic Value Validation focuses on the quantitative and qualitative benefits derived by deploying Datrium’s Cloud DVX, a software-based version of the Datrium DVX converged system, versus traditional backup or other cloud backup deployments. The Datrium DVX combines a tier 1 hyperconverged system with scale-out backup and cloud disaster protection in a single system. Cloud DVX is designed to be used in conjunction with on-premises DVX systems for simple backup and archiving to the public cloud with low TCO.

Challenges

While numerous innovations have altered the data storage landscape in recent years, challenges remain. According to ESG research, North American mid-market and enterprise IT managers cited data protection, hardware costs, and rapid data growth as their top storage challenges (see Figure 1).¹ These are intertwined; the more data grows, the higher the cost to store and protect it. Complicating matters—and increasing costs—are two key trends: 1) more stringent compliance requirements that result in not only keeping more data, but also keeping it longer; and 2) the need for high data availability so mobile workers can create, consume, share, and collaborate on more data.

Figure 1. Top Five Storage Environment Challenges



Source: Enterprise Strategy Group

Some organizations are beginning to use public cloud object storage such as Amazon Web Services (AWS) for backup and long-term data retention in order to simplify operations and reduce costs, and to keep data more available globally with faster restore time. Upfront costs are replaced by “pay as you go” pricing, and there are no disk or tape systems to manage and refresh; restore can be done more quickly from the cloud than having to wait days for tapes to be brought onsite and recovered. There can be tradeoffs, however. Cloud storage costs may be difficult to predict as data grows over time, and cloud backup solutions can add AWS management complexity that administrators must deal with.

The Solution: Datrium Cloud DVX

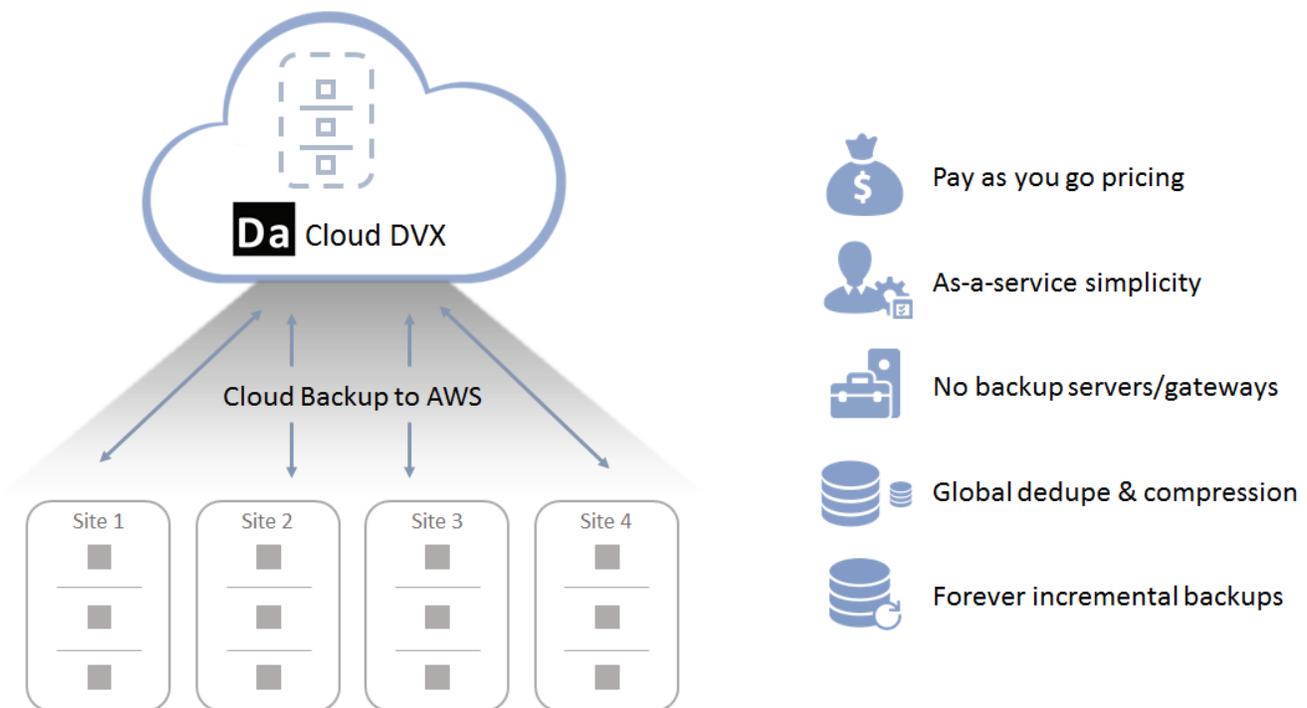
Datrium, a provider of “open-converged” infrastructure, recently released its Cloud DVX, a backup-as-a-service platform which runs natively in AWS (using EC2 compute and S3 object storage) to protect VMs running in on-premises DVXs. Cloud DVX offers low TCO and “as-a-service simplicity” with automation so administrators don't have to manage AWS complexity.

¹ Source: ESG Master Survey Results, [2017 General Storage Trends](#), December 2017.

It is simple to set up and self-healing, and upgrades occur automatically. Virtual machines (VMs) residing in on-premises DVXs are securely backed up to Cloud DVX, with data reduction and backup methods that minimize storage capacity needs, reduce costs, and enable fast restore.

Datrium DVX was built with features that make it easily transferable to the cloud, allowing for near effortless hyper-scalability. The Datrium DVX file system uses a key value store and object interface, rather than a traditional block interface. While many file systems would need extensive reworking to function in AWS, the Datrium file system was ported with very little change. In addition, Datrium DVX splits the compute and data functions into two nodes; the compute nodes deliver I/O services, and data nodes provide data durability. This architectural approach mirrors the AWS cloud, which uses a similar construct: compute happens in EC2, and data services in S3, making the infrastructure highly scalable.

Figure 2. Datrium Cloud DVX



Source: Enterprise Strategy Group

ESG Economic Validation

ESG completed a quantitative economic analysis of Datrium Cloud DVX with a focus on storage capacity and AWS costs.

ESG’s Economic Validation process leverages ESG’s core competencies in market and industry analysis, forward-looking research, and technical/economic validation. ESG audited TCO models that Datrium created to help customers understand the differences among cloud backup solutions. These models focus on the impacts of data reduction and backup processes.

Datrium Cloud DVX Economic Value Overview

ESG’s economic analysis revealed that Cloud DVX can significantly shrink the amount of storage required in the cloud when compared to alternative cloud-ready backup solutions, and therefore minimize AWS storage costs. ESG found that Cloud DVX provides its customers with significant cloud storage savings and benefits due to:

- **Extensive data reduction** – Cloud DVX includes compression and end-to-end global deduplication across all sites and the cloud, reducing data volumes and cloud storage costs.
- **Forever incremental backups** – Cloud DVX backs up only differential data to the cloud after the initial seed, requiring less storage and reducing costs.
- **Reducing hardware and complexity** – Cloud DVX eliminates the need for either secondary data protection sites or on-premises systems to connect to AWS. Backups are done from an on-premises DVX directly to the Cloud DVX, with no on-premises backup server or gateway. This reduces the hardware, power, cooling, networking, and management required for data protection. In addition, end-to-end encryption eliminates the need for AWS VPN and its associated costs.



Data Reduction

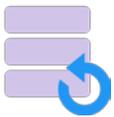
Data that is backed up is usually retained for at least 90 days, with some data retained for a year, five years, and even permanently. Every organization will have different retention requirements, and these may change over time due to compliance statutes and internal corporate governance. Organizations want the benefits of having data in public clouds, such as offloading data protection management tasks, and having data restored more quickly than through traditional backup tape processes. However, as production data grows, so does backup data, and the cost of these growing data stores can make cloud storage prohibitively expensive.

Datrium Cloud DVX:

- *Compression.*
- *Global deduplication.*
- *Across all sites and cloud.*

With Datrium Cloud DVX, compression and deduplication are always on across all sites replicating to the cloud, so no duplicate data is stored in the cloud. This global deduplication provides greater capacity savings when compared to other dedupe methods such as local source dedupe or target-based dedupe. This is a significant benefit for backups; every backup created stores additional copies of the same data until it is expired, so backup is often the single biggest contributor to storage growth. According to Datrium, Cloud DVX customers average 4.5:1 local deduplication and an additional 1.5X in-cloud global deduplication.

Forever Incremental Backups/Fast Restore



While there are many backup methods, a common one is weekly full backups and nightly incremental backups, resulting in four full backups and 24-27 incrementals per month (depending on the number of days in the month). This data grows significantly as backups are piled on one another for months until the prescribed retention time is reached, at which time a backup is deleted.

Datrium Cloud DVX uses a different method. The first backup is a full backup to seed the cloud, but after that all backups are incremental, only sending changed data to the cloud; the incrementals are stored as synthetic fulls with the incremental and previously stored data. Combined with the always-on global deduplication, Cloud DVX keeps the amount of data sent to the cloud as low as possible. The savings are significant over time.

In addition, data is restored from AWS to the on-premises DVX, where VMs are restarted instantly. Only the data that does not already exist on the target DVX is restored; the Cloud DVX retains an inventory of all indices on all DVXs being protected. As a result, if 80% of the data you want to restore is already on the DVX, it will only push the data you don't have. Egress out of AWS is expensive, but Cloud DVX has the intelligence in the cloud to reduce restore data and keep costs down.

“Datrium Cloud Backup was so simple to use. . . I was blown away – [restore] was virtually instantaneous because Datrium took advantage of the fact that most of the VM data being recovered was already on my local system.”

- University of Maryland ARE

The University of Maryland Agricultural and Resource Economics department was looking for cost-effective, offsite data protection for their research data, and tried Datrium Cloud DVX. They were impressed with the speed of restore, as the quote indicates.



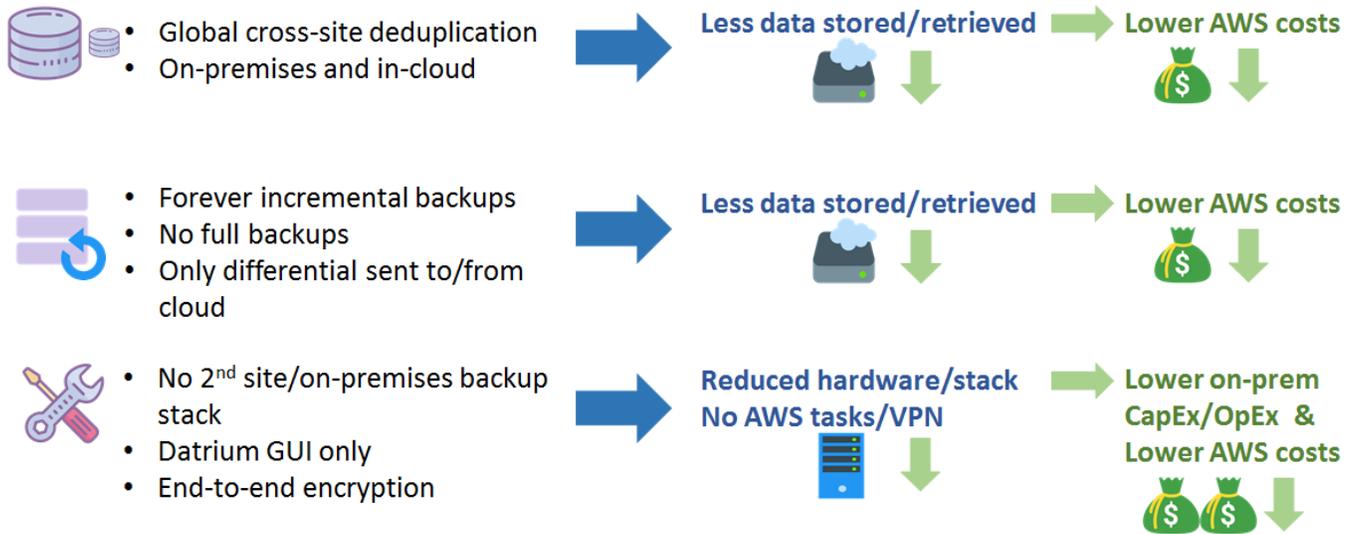
Eliminating Hardware and Complexity

To provide for disaster recovery, traditional backup solutions require a second site to store backup data, and most cloud backup solutions require additional on-premises infrastructure (beyond the appliance itself) in order to send backups to AWS. This may include backup servers and software, cloud gateways, etc., which also add the costs of data center racks and floor space, power, cooling, networking, and management. In addition, many of these cloud backup solutions require that backup administrators learn to use the AWS Dynamo database, AWS S3 object storage Buckets, and other tools that they would otherwise not need to know.

Datrium Cloud DVX requires no additional on-premises backup stack; the on-premises DVX backs up directly to the Cloud DVX. Setup is simple with Cloud DVX and AWS; once Cloud DVX users enter AWS credentials into the Datrium GUI (which is integrated with vCenter), the DVX cloud instance is automatically created in EC2, as is the S3 Bucket, and both are updated automatically and supported proactively. No AWS interface is used, so administrators have a single GUI to handle the on-premises DVX for primary workloads as well as backup to the Cloud DVX. Datrium manages Cloud DVX as a service, so problems with the EC2 instance or S3 Buckets are handled by Datrium, and regular customer telemetry enables them to proactively fix any problem before they impact customers. (It should be noted that these cost savings were not included as part of this analysis; they would be in addition to the cloud cost savings represented in this paper. In addition, end-to-end encryption eliminates the need—and the cost—of AWS VPN services.

Figure 3. Datrium Cloud DVX Advantages

Datrium Cloud DVX Advantages



Source: Enterprise Strategy Group

ESG Analysis

ESG audited Datrium’s TCO calculators, leveraging not only vendor-provided material but also publicly available pricing and industry knowledge of economics and technologies. The primary objective of these models was to demonstrate two key factors that impact both the amount of data stored and the AWS cost:

- the impact of improved data reduction rates.
- the impact of forever incremental backups versus occasional full backups combined with incrementals.

The Datrium calculators were designed with AWS pricing available at the time. Our audit used an i3xlarge, three-year reserved, no upfront costs, hourly-charge EC2 instance; S3 object storage charges were specified for GB/month/container, Gets, Puts, Egress, and VPN. The expected solution-dependent effects of write amplification, key value store merges, and get multiplication factors were also considered across all considered deployments.

For each model, inputs included the number of source sites, and then per site:

Why This Matters

Cloud backup can simplify data protection, speed restore, and shift costs to monthly OpEx. But costs can skyrocket as data grows over time.

Datrium Cloud DVX, with always-on deduplication and forever incremental backups, shrinks the amount of data stored in AWS, and subsequently the cost. This makes the cloud a cost-effective resource for protecting data over the long term.

- Total logical capacity.
- Annual data growth rate.
- Daily data change rate.
- Data reduction rate, on-premises and in the cloud.
- Retention period for data in the cloud.
- Annual amount of restore data.
- Number of full and incremental data transfers to cloud per month.

Customer Inputs	
Number of sources (sites/systems)	4
Total logical capacity per source (GB)	20000
Annual data growth rate% per site	10.00%
Local (on-premises) data reduction ratio	4.50
Logical change rate /day	2.000%
In-cloud data reduction ratio	1.50
Retention period for data in the cloud (Days)	90
Restore data size per year (GB)	10000
EC2 Type (1-On-demand, 2-Reserved 1YR, 3-Reserved 3YR)	3

With these details in place, the model calculated the amount of data that would be stored in the cloud over five years, the AWS costs per month and year, and the average AWS cost over three years. Thirty-day months were assumed in the calculators. They also provided a detailed breakdown of individual AWS costs resulting from the expected number of Gets, Puts, Data, Egress, and standard business support. *It should be noted that no costs are included in this model for Datrium Cloud DVX software itself or other deployment costs; these would be in addition to the AWS costs calculated here. This analysis focuses on AWS costs only. The TCO cost savings from elimination of backup hardware, gateways, etc. are also in addition to the AWS cloud cost savings.*

Impact of Deduplication

The first comparison quantified the impact of data reduction rates, comparing two deployments that were identical except for the on-premises and in-cloud deduplication rates. Each deployment included four sites, each replicating 20 TB to Cloud DVX, for a total of 80 TB. Data grew 10% annually, with 2% daily change rate, and a 90-day retention rate in the cloud. Each restored 15 TB per year, and each used forever incremental backups after the initial full backup to seed the cloud.

- The Datrium DVX deployment had an on-premises data reduction rate of 4.5:1, and in-cloud data reduction rate of 1.5:1. These are average rates for Cloud DVX customers.
- Deployment 2, based on known rates of other cloud backup solutions, had an on-premises reduction rate of 2:1, and in-cloud reduction rate of 1:1.
- Deployment 2 lacked end-to-end encryption, and therefore required AWS VPN services.

With this data, the model first calculated the amount of data stored and required storage capacity. Data amounts were increased by the growth and change rates and were reduced by the deduplication rates and deletion of expired data. The Cloud DVX with higher deduplication rates stored 69% less data (and needed 69% less storage capacity) than Deployment 2.

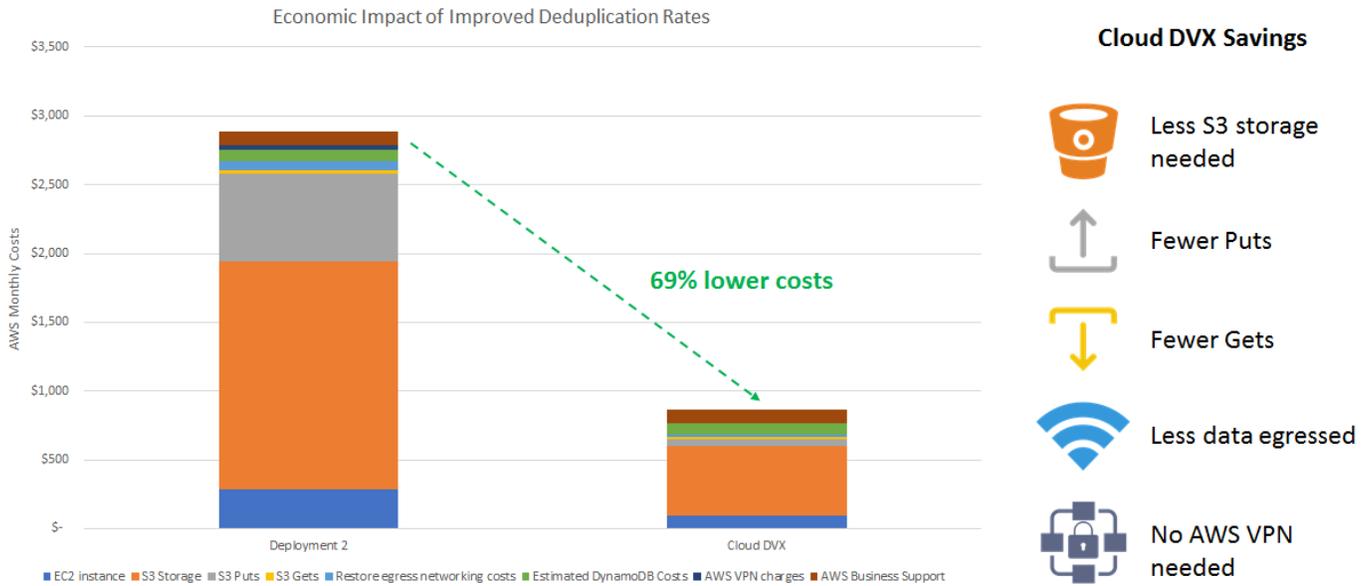


**Cloud DVX:
69% Less storage
capacity**

Next, we looked at AWS costs, where the amount of data stored has a big impact. Because Cloud DVX deduplicated and stored and restored less data, the costs were lower for S3 Storage. Transactional traffic, including Puts, Gets, and Egress, were also reduced. In addition, Deployment 2 required AWS VPN charges, while the DVX did not because of its end-to-end encryption. EC2 instance costs grow as capacity grows; to accommodate data growth and performance, we added one instance for every 30 TB of S3 storage used. Cloud DVX required a single instance that grew to two over five years, while Deployment 2 required three EC2 instances, growing to four instances over five years.

Figure 4 shows the monthly AWS costs for month 1: The Cloud DVX solution started at \$867, while for Deployment 2 it started at \$2,887; for Cloud DVX the three-year average AWS cost was \$13,164, while for Deployment 2 it was \$40,996. Cloud DVX costs were 69% or 3.1X lower because of its extensive data reduction, including S3, Put, Get, Egress, and no VPN charges. These differences remained over five years.

Figure 4. Datrium Cloud DVX Cost Advantages from Improved Deduplication Rates



Source: Enterprise Strategy Group

Table 1 details annual AWS capacity stored and cost for Cloud DVX and Deployment 2 over five years.

Table 1. Five-year AWS Capacity (GB) and Cost Comparison

	Year 1	Year 2	Year 3	Year 4	Year 5
Cloud DVX Capacity	28,087	29,537	30,988	32,439	33,889
Dep 2 Capacity	92,933	97,733	102,533	107,333	112,133
Cloud DVX Cost	\$12,289	\$12,932	\$14,271	\$15,040	\$15,524
Dep 2 Cost	\$38,775	\$41,207	\$43,004	\$44,802	\$46,599

Impact of Forever Incremental Backups versus Occasional Full Backups

The next model calculated the amount of data stored/capacity required and AWS costs with Datrium Cloud DVX forever incremental backup, compared with an almost identical deployment (Deployment 3) that required five monthly full backups and 25 incrementals. Both deployments used four sites, each with 20 TB of data, for a total of 80 TB, and included 10% annual growth rate and 2% daily change rate. Both deployments had 4.5:1 deduplication on-premises, and both retained data for 365 days before expiring. Differences between the deployments:

- Cloud DVX used forever incremental backups after the initial seed, while Deployment 3 required five full backups and the rest daily incrementals each month. This delivers significant capacity and cost savings for DVX.

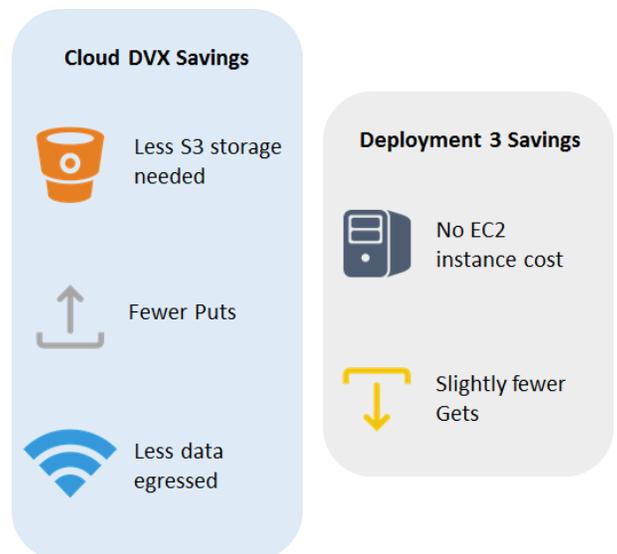
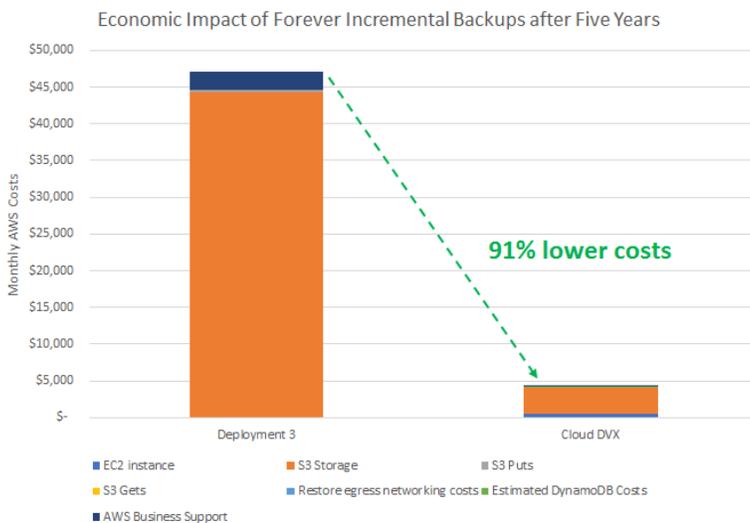
- Data reduction in the cloud for Cloud DVX is 1.5:1, while for Deployment 3 it is 1:1. This results in additional capacity savings for Cloud DVX.
- While Cloud DVX used an EC2 instance, Deployment 3 did not, eliminating the monthly EC2 costs. This results in AWS cost savings for Deployment 3.

First, ESG looked at the amount of S3 storage needed. As before, data amounts were increased by the growth and change rates and were reduced by the deduplication rates and deletion of expired data. The significant difference is in backup data; because Cloud DVX uses forever incremental backups and performs no full backups after the initial seed, it stores 91% less data than Deployment 3, requiring 91% less storage capacity.



Next, we looked at AWS costs. Figure 5 shows that after five years, Cloud DVX cost 91% or 10X less than Deployment 3. Table 2 shows details of the capacity and cost comparison over five years.

Figure 5. Datrium Cloud DVX Cost Advantages from Forever Incremental Backup Process



Source: Enterprise Strategy Group

What the numbers mean:

- The impact of full backups is significant. Deployment 3 needed capacity for five full backups per month, plus 25 incrementals; DVX only stored the initial full, and then all incrementals.
- The higher in-cloud deduplication rate increased the data reduction advantage for Cloud DVX.
- Even with Cloud DVX’s additional charges for EC2, Cloud DVX delivered 91% lower cost over five years.

Table 2. Five-year AWS Capacity (GB) and Cost Comparison

	Year 1	Year 2	Year 3	Year 4	Year 5
Cloud DVX Capacity	103,701	112,648	121,352	130,056	138,760
Dep 3 Capacity	1,236,889	1,354,222	1,471,556	1,588,889	1,706,222
Cloud DVX Cost	\$24,655	\$41,247	\$44,184	\$47,881	\$50,628
Dep 3 Cost	\$216,287	\$429,181	\$468,515	\$507,848	\$547,182

One Cloud DVX beta customer keeps data for 25-30 years and has been using tape as the only cost-effective storage for the long term, leaving the company without a fast restore option.

After Hurricane Irma, the company was unable to restore data for days.

The savings available with Cloud DVX, based on data reduction and forever incremental backups, is providing this customer with fast restore options without breaking the bank on storage costs.

The Bigger Truth

IT administrators are asked to juggle conflicting priorities all the time; it seems to be a requirement for the job. One daily reminder of that is the conflict of keeping data fully protected and at the same time highly available, while keeping costs down. Doing any two of these simultaneously is fairly simple—for example, organizations can fully protect data with tape backups and keep costs down. But with tape, data can take days to restore, which is disruptive to business when you really need your data. Cloud backup is a new phenomenon that can help keep data protected, and have it quickly restored—but the costs of cloud backup can quickly add up as backup data accumulates over time.

It is clear that cost reduction remains a key priority for most organizations. In ESG’s annual spending intentions surveys, cost reduction has been at or near the top of the list of business initiatives driving the most technology spending since 2009.² It’s one of the reasons that the HCI market is growing—HCI can make your VM stack simpler and cheaper to manage. With data protection, it’s important to look at the whole stack when considering costs. Traditional backup solutions require you to invest in a full storage stack (whether disk or tape) at a secondary location, requiring upfront expense but having resources sitting idle. Even cloud backup solutions usually require an on-premises stack of hardware, power, cooling, networking, and management.

Datrium, provider of the DVX “open-converged” system, has recently released its Cloud DVX to back up VMs to a cloud-based DVX instance in AWS. This shifts DVX data protection costs from upfront CapEx to monthly OpEx and makes backup simpler to manage and faster to restore than tape. Cloud DVX provides a built-in, as-a-service benefit by removing the burden of managing the AWS data store; customers want a solution for data protection, not another infrastructure to learn and manage.

But an essential feature for making cloud backup affordable is the ability to reduce storage capacity needs to bring costs down. ESG analyzed TCO calculators that demonstrate the economic impact of minimizing cloud storage costs with the Datrium Cloud DVX. Datrium Cloud DVX offers two key features that keep costs down when compared with other cloud-ready backup solutions: higher on-premises, and in-cloud global deduplication rates, and a capacity-friendly, forever incremental backup policy. Our analysis showed:

² Source: ESG Research Report, [2018 IT Spending Intentions Survey](#), February 2018.

- Datrium Cloud DVX deduplication rates can result in up to 69% lower AWS costs than an alternative backup solution that provides half of the deduplication savings.
- Cloud DVX forever incremental backups can provide up to 91% lower AWS costs compared with alternative on-premises solutions (with no EC2 instance required) that rely on storing more full backup copies in the cloud.
- In addition, what is not included in these calculations are the elimination of on-premises (or secondary site) backup systems and accompanying costs for power, cooling, floorspace, networking, management, etc. These add even more savings.

Every organization's requirements are different, and your particular savings may vary based on several factors, including your daily data change rate, growth rate, and data retention requirements, the amount of data you expect to restore, and AWS pricing. However, ESG's TCO analysis clearly indicates that customers should ask several questions when evaluating a cloud backup solution. These include:

- Does the solution use deduplication and compression for data reduction? What are the on-premises and in-cloud deduplication rates?
- Does the solution require periodic full backups, or does it use forever incrementals?
- Is encryption built-in, or do I need to purchase AWS VPN?
- Are an on-premises backup server/gateway and associated data center costs such as power, cooling, networking, and management required?
- How does recovery work? How many hops are required to get data back up in the primary site? How fast is it?

ESG recommends that each organization perform its own analysis to better understand the potential savings.

The cloud can be a great place for backup with Datrium Cloud DVX, where customers get both a cost-effective place to store data for as long as it's needed, as well as faster recovery time than traditional tape-based backup solutions. It's also great for customers who are looking for simple, scalable offsite backups that they don't have to worry about.

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.



Enterprise Strategy Group is an IT analyst, research, validation, and strategy firm that provides market intelligence and actionable insight to the global IT community.

© 2018 by The Enterprise Strategy Group, Inc. All Rights Reserved.

