Technical Research Study

PROWESS

Cost Considerations for Placing Enterprise Workloads in Public Clouds

Analysis conducted by Prowess Consulting indicates that a private cloud solution from Dell Technologies could save money compared to the public cloud for a variety of workloads.

Executive Summary

The rapid adoption of public cloud services has brought flexibility and scalability to organizations. It has also introduced challenges around cost predictability and infrastructure control. As organizations seek to regain visibility and governance over their IT environments, private cloud solutions offer a compelling alternative. These solutions provide not only cost efficiency but also the ability to tailor infrastructure to specific operational needs.

Private cloud infrastructure can provide the highest levels of control, security, and customization for IT organizations. At the same time, private clouds can provide the greatest cost efficiency for a variety of workloads. While the public cloud remains an excellent solution for certain workloads, the security and control of on-premises IT infrastructure, combined with the flexibility of disaggregated infrastructure built on robust compute and storage technologies, enables organizations to lower their total costs. This approach can help businesses precisely align resources with their specific operational needs, which, in turn, can help ensure scalability, compliance, and performance without compromising governance. By leveraging the adaptability and tailored capabilities of dedicated infrastructure, organizations can unlock greater agility, reliability, and cost efficiency.

To assess the price differential of a private cloud environment versus the public cloud, Prowess Consulting investigated the costs of running various workloads on public clouds versus private cloud solutions. In this study, commissioned by Dell Technologies, we examined a matchup of Amazon Web Services® (AWS®) and Dell Technologies solutions across five different common workloads, as representative examples of two as-a-service options. The unified Dell Technologies approach to disaggregated infrastructure, with software-driven automation and intrinsic data security, offers independent scaling for compute and storage and optimized server utilization, which can help reduce costs. Indeed, we discovered that private cloud deployments with Dell Technologies solutions on average cost 26–64% less than public cloud–native options run on AWS, including accounting for factors such as managed-service and co-location fees.

Highlights

Compared to public cloud options, Dell Technologies solutions cost up to:[1]

26%
less for VMs

64%
less for AI

40%
less for containers

37%
less for high-throughput storage

58%
less for balanced-performance storage

Study Overview

To assess the price differential of a private cloud implementation versus public cloud–native solutions, we investigated the unit cost of running different workloads on public clouds through hyperscalers and with hybrid or private cloud solutions. This study supports a cloud-smart approach—one that considers the optimal deployment model for each workload rather than assuming a one-size-fits-all solution.

Deployments Considered in This Study

For the public cloud, we examined the representative cost of running several different workloads in Amazon® Elastic Compute Cloud™ (Amazon EC2®). We chose AWS because it is one of the most widely used public cloud services.[2] Moreover, by examining infrastructure-as-a-service (IaaS) instances, we can create a good baseline for comparison with on-premises alternatives.

For private cloud solutions, we examined the workloads in integrated and more layered environments, which would be representative of a private cloud layout. Dell Technologies solutions offer as-a-service pricing, which makes comparison with deployments on Amazon EC2 IaaS instances practical.

To make the comparison, we evaluated both exemplars using the following workload scenarios:

- Virtual machines (VMs)
- AI
- Containers
- High-throughput storage
- Balanced-performance storage

Study Details

For each of the use cases examined, we analyzed both cost and non-financial considerations. For each use case, we briefly describe the assumptions we made in the analysis; you can find more details about the underlying configurations for the various use cases in the appendices.

Virtual Machine (VM) Workload

The global virtualization market was valued at USD 96.1 billion in 2024 and is projected to reach USD 386.8 billion by 2030, growing at a CAGR of 26.1%.[3] Virtualization remains foundational to enterprise IT, but as workloads scale, so do the challenges of managing cost and control—particularly in public cloud environments. Organizations are increasingly seeking infrastructure strategies that offer predictable pricing, governance over data and workloads, and the ability to optimize performance. Private cloud deployments provide these advantages, making them a compelling option for virtualized workloads.

To evaluate the matchup between the public cloud and private cloud solutions that include on-premises infrastructure, we used the case of a US-based bank employing a fleet of VMs/cloud instances across two separate US locations to manage customer data, mobile banking back-end systems, and digital customer services:

- Both the Dell Technologies and the AWS configurations supported a total of 1,200 VMs
- The Dell Technologies solution ran on 20 Dell™ PowerEdge™ R660 servers running VMware Cloud Foundation®

In addition, the Dell Technologies solution included managed services and co-location in order for the solution to be hands-off in the same way that AWS is.

Our analysis indicates that VMs **cost up to 26% less on Dell Technologies infrastructure with a Dell APEX™ subscription than on AWS**. Table 1 and Figure 1 show the results of the comparison. (For configuration details for this workload comparison, see **Appendix A**.) This cost advantage is further strengthened by the operational control the Dell Technologies infrastructure provides. This control enables organizations to manage VM placement, performance tuning, and data residency with greater precision and transparency than is typically possible in public cloud environments.

Table 1 | Breakdown of three-year costs for 1,200 VMs running on AWS® and the Dell Technologies solution, respectively

Category	Amazon Web Services® (AWS®) Cost[4]	Dell Technologies Cost[5]
Compute, storage, snapshots, and virtualization	$278,224.75	$157,502.00
Connectivity and data egress	$10,667.14	Not applicable (N/A)
Co-location services or enterprise support plan	$26,624.29	$32,740.00
Managed services	$19,492.86	$54,580.00
Monthly subtotal	**$335,009.04**	**$244,822.00**
Three-year total	**$12,060,325.44**	**$8,813,592.00**

Note: No discounts (such as public-cloud committed-spend discounts) were applied in this analysis.

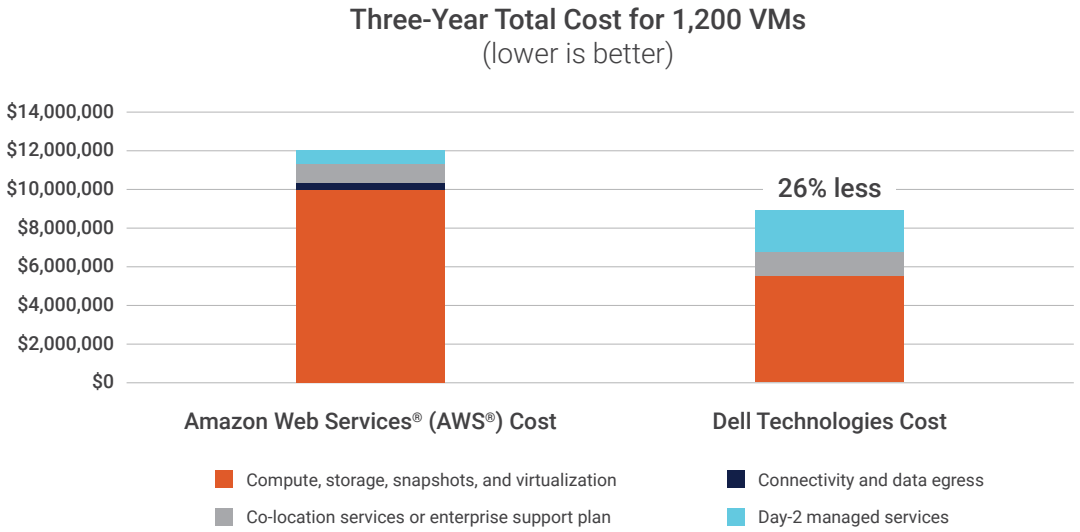

Figure 1 | Comparison of three-year costs for 1,200 VMs running on AWS® and the Dell Technologies solution

Analysis: VMs

Data sovereignty, co-tenancy risks, and unpredictable data-transfer fees are all critical considerations when deploying VMs in the public cloud. These factors can introduce compliance challenges and cost variability. The Dell Technologies private cloud solutions mitigate these risks by offering dedicated infrastructure, centralized control over data and workload placement, and a transparent cost structure. These benefits enable IT teams to maintain governance while optimizing for performance and budget.

Regional data-transfer fees must also be factored into consideration for running workloads in the public cloud. Table 1 accounts for data-egress charges only; moving data between public-cloud geographical regions and even availability zones for things like disaster recovery (DR) and high availability (HA) can also incur additional charges, which can add up quickly and significantly.

VM co-tenancy can be another issue to consider when running VMs in the public cloud. Proper security can address the possibilities of data being exposed to other public-cloud instances running on the same physical host in a public-cloud data center. But noisy neighbors—cloud instances on the same physical host, network, or storage pool that monopolize disproportionate network input/output (I/O) resources—can increase instance latency and degrade connectivity and cloud-application performance. Having dedicated VM hosts on premises or in a co-location facility can help address this by increasing your IT organization's control of load balancing and VM placement.

AI Workload

AI is an automation mega-trend set to dominate the coming decades, and it is only getting started. The global enterprise AI market is expected to reach USD 58.11 billion in 2025, and it is projected to grow to USD 474.16 billion by 2030, growing at a CAGR of 52.17%.[6] A big part of this growth comes from infrastructure costs. As organizations scale AI workloads, maintaining control over infrastructure, data locality, and GPU availability becomes essential for managing both performance and budget. Private cloud solutions offer a strategic advantage by enabling predictable cost structures and operational control across the AI lifecycle.

AI has diverse applications. One key use is training models, a complex, resource-intensive process that occurs infrequently. Another is deploying these models to generate insights through inferencing. Inferencing is fast and computationally modest, but it must happen potentially thousands of times a day for some AI applications. Meanwhile, the iterative nature of AI workloads—particularly during model training and tuning—can dramatically amplify infrastructure costs in public cloud environments, where compute, storage, and data movement are all metered and often priced at a premium.

Dell Technologies provides a model to help control AI-development costs. In the Dell™ AI Factory model, AI workloads span the full lifecycle—from early-stage model experimentation and training to scalable, production-grade inferencing. Dell Technologies provides the foundational infrastructure and taps common software ecosystems to accelerate this lifecycle and streamline model deployment.

To evaluate this full-lifecycle AI production model between the public cloud and private cloud solutions that include on-premises infrastructure, we considered the scenario of a pharmaceutical company. The company uses high-performance GPUs for AI work in genomics and drug discovery. For this comparison, we evaluated:

- Two Dell PowerEdge XE9680 servers and three PowerEdge R660 servers as part of the Dell AI Factory solution
- Two Amazon EC2 P5.48xlarge instances, two ml.r5.16xlarge instances, and 20 ml.t3.medium instances

(**Note**: The smaller servers and AWS instances were included for data processing and to host notebooks for data scientists; for this analysis, the ml.r5.16xlarge and ml.t3.medium instances correspond in role to the PowerEdge R660 servers.)

Each PowerEdge XE9680 server included eight GPUs, and each P5.48xlarge instance has eight vGPUs. Given the heavy data-analysis needs of pharmaceutical firms, we estimated constant utilization over a four-year term. Our analysis indicates that running high-end GPUs **costs up to 64% less on the Dell Technologies solution than on AWS**. Table 2 and Figure 2 show the results of the comparison. (For configuration details for this workload comparison, see **Appendix B**.) In addition to some cost savings, the Dell AI Factory model gives organizations full control over infrastructure provisioning, GPU access, and data governance, all of which are critical factors for AI workloads that are iterative, data-intensive, and sensitive to latency.

Table 2 | Four-year pricing for high-throughput storage on AWS® and the Dell™ AI Factory solution, respectively

	Amazon Web Services® (AWS®) Cost[4]	Dell™ AI Factory Cost[5]
Four-year AI instance/server cost	$2,502,754.07	$757,231.00
Server administration (four-year total)[7]	N/A	$6,663.00
Energy costs for power and cooling (four-year total)	N/A	$125,502.00
Data center costs for rack space (four-year total)	N/A	$2,080.00
Four-year total	**$2,502,754.07**	**$891,476.00**

Note: This pricing assumes that an organization would use the three-year Amazon EC2 Instance Savings Plan twice (as a three-year Amazon EC2 Instance Savings Plan paired with a one-year plan was still more expensive).

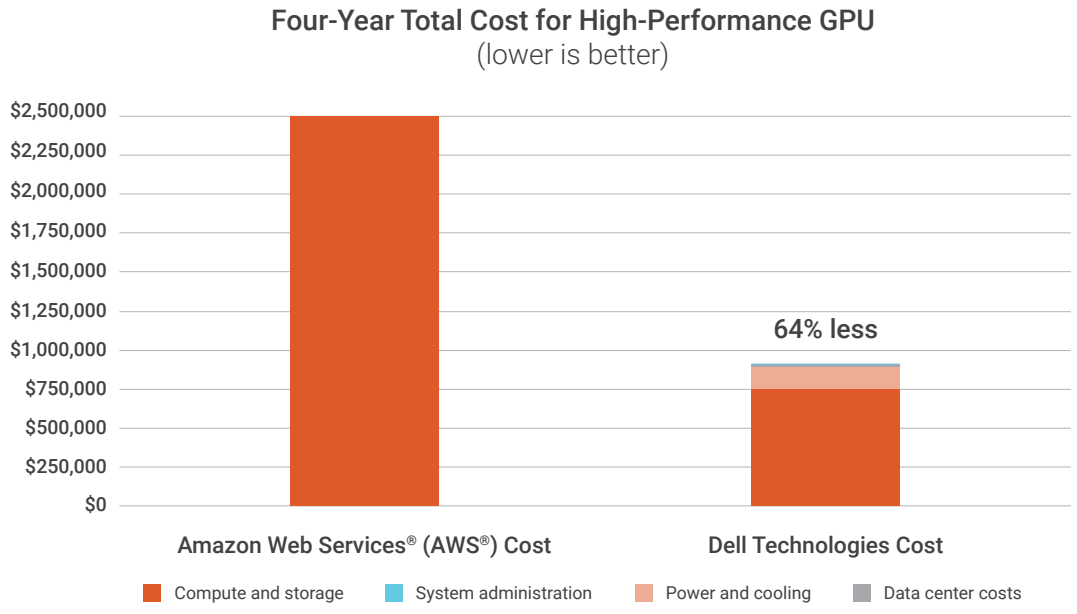

Figure 2 | Comparison of four-year costs of high-performance GPU usage on AWS® and the Dell™ AI Factory solution

Analysis: AI

As AI adoption accelerates, organizations face growing challenges in securing high-performance GPU resources, managing data movement, and maintaining security and compliance. Public cloud offerings often introduce unpredictability through regional GPU shortages, variable pricing, and data mobility and egress fees. An example of how to address these issues in a private cloud is Dell AI Factory. The Dell AI Factory architecture offers localized infrastructure, consistent performance, and centralized control, which can help enterprises scale AI workloads securely, efficiently, and with full visibility into costs. Other factors to consider when running workloads on-premises versus in the public cloud include data sovereignty, security for proprietary data, and control of data.

Container Workload

While the public cloud has long been the default for containerized workloads, organizations are increasingly seeking greater control over their infrastructures to optimize performance, security, and cost. Private cloud environments now offer the flexibility to deploy containers where they make the most strategic and financial sense.

Containerization is a vital component of modern, cloud-native app architecture, and it underpins everything from DevOps continuous integration/continuous delivery (CI/CD) pipelines to serverless applications. Hybrid infrastructures enable containers to span clouds and seamlessly extend on-premises. The rise of cloud-native development and alternatives to strictly public-cloud deployments for a variety of use cases represents another workload that might profitably move from the public cloud to a private cloud.

Comparing the deployment of containers to the public cloud and to a more flexible infrastructure, Prowess Consulting evaluated a scenario of an e-commerce platform that starts with 12,000 containers (to host microservices to manage inventory, payments, user profiles, product recommendations, and so on) and that expands its footprint to 40,800 containers over a three-year period to keep up with its expanding business. We compared:

- Dell PowerFlex™ servers consumed via Dell APEX subscriptions/the Dell APEX Data Center Utility solution using the Red Hat® Enterprise Linux® operating system (OS) and Dell™ Cloud Platform for Red Hat OpenShift®
- Amazon EC2 C6id.large instances using Amazon® Elastic Kubernetes® Service (Amazon EKS®) and Amazon® Elastic Container Registry (ECR)

Our analysis indicates that containers **cost up to 40% less on the Dell Technologies solution than on AWS**. Table 3 and Figure 3 show the results of the comparison. (For configuration details for this workload comparison, see **Appendix C**.)

Table 3 | Breakdown of annual and total costs for a deployment scaling up to 40,800 containers over three years on AWS® and the Dell Technologies solution, respectively

	Containers	AWS® Cost[4]	Dell Technologies Cost[5]
Year 1	12,000	$8,523,950.52	$4,900,000.00
Year 2	26,400	$16,712,942.52	$9,800,000.00
Year 3	40,800	$24,901,934.52	$14,700,000.00
Co-location services (three-year total)	N/A	N/A	$207,014.40
Three-year total		**$50,138,827.56**	**$29,607,014.40**

Note: No discounts (such as public-cloud committed-spend discounts) were applied in this analysis.

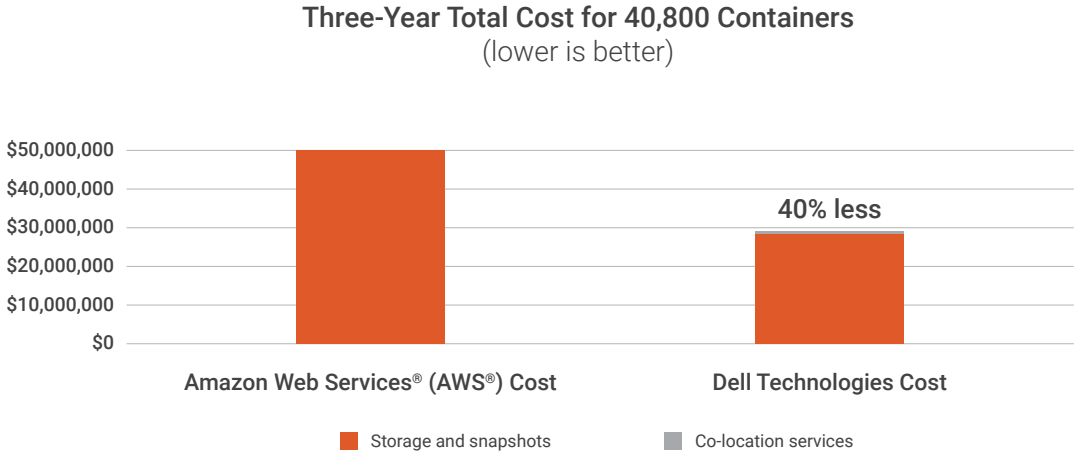

Figure 3 | Comparison of total three-year cost of a deployment scaling up to 40,800 containers on AWS® and the Dell Technologies solution

Analysis: Containers

Other factors to consider when running containers in the public cloud include tradeoffs between supporting a wide variety of cloud-instance architectures for cloud-native apps and refactoring tools for DevOps to support them. While your DevOps teams might see this as an ordinary part of their CI/CD pipeline, it still represents a draw on engineering and development resources that could be applied to other priorities.

High-Throughput Storage Workload

For enterprises managing business-critical applications, performant storage is essential—but so is maintaining control over where and how that storage is provisioned. As organizations seek to reduce unpredictability in both cost and compliance, private cloud solutions offer a compelling alternative to public cloud storage, enabling tighter governance, consistent performance, and long-term cost efficiency. Consider, for example, a large financial-services firm, with an ever-growing need for performant storage, increasing to 697 TB over three years. The company likely has large relational databases for customer transactions, in addition to several big NoSQL databases to support big data, analytics, and AI applications. All of these depend on high-throughput storage to keep the business running. However, the firm might also need to keep much of this data in a public cloud or on-premises for reasons of long-term data retention and global availability.

To evaluate this workload in the context of public cloud versus private cloud, Prowess Consulting used AWS as the representative for the public cloud and the Dell Technologies private cloud environment. For AWS, the high-throughput storage used Amazon® Elastic Block Store (EBS) io2 Block Express storage. The Dell Technologies solution ran on Dell PowerMax™ 2500 storage, having 107 TiB capacity starting in year 1 with a 100% storage commitment. Our analysis indicates that high-throughput storage **costs up to 37% less on the Dell Technologies solution than on AWS**. Table 4 and Figure 4 show the results of the comparison. (For configuration details for this workload comparison, see **Appendix D**.) In addition to cost savings, the Dell Technologies private cloud architecture provides organizations with full control over data placement, performance tuning, and compliance management—factors that are often constrained or opaque in public cloud environments, but that can be tough to model quantitatively.

Table 4 | Three-year pricing for high-throughput storage on AWS® and the Dell Technologies solution, respectively

		Amazon Web Services® (AWS®) Cost[4]		Dell PowerMax™ 2500 Cost[5]
Year 1	420 TB Amazon® EBS storage (50% of 840 TB)	$261,583.88	Dell PowerMax 2500 cost (100% utilization of 850 TiB)	$2,000,000.00
Year 2	588 TB Amazon EBS storage (70% of 840 TB)	$366,217.43	–	–
Year 3	697 TB Amazon EBS storage (83% of 840 TB)	$434,229.23	–	–
I/O operations per second (IOPS)	256,000 IOPS (three-year total)	$1,389,772.80	More than 2.7 million IOPS per node	Included
Full snapshots	1/month, Amazon EBS (three-year total)	$424,812.21	Unlimited	Included
Incremental snapshots	30/month, Amazon EBS (three-year total)	$632,509.81	Unlimited	Included
Storage for backups	Amazon® S3 (three-year total)	$13,084.22	Unlimited	Included
Co-location services	N/A	–	Three-year total	$93,006.00
Storage administration[7]	N/A	–	Three-year total	$112,500.00
Three-year total		**$3,522,209.58**		**$2,205,506.00**

Note: No discounts (such as public-cloud committed-spend discounts) were applied in this analysis. Snapshots are included as part of the Dell PowerMax 2500 pricing.

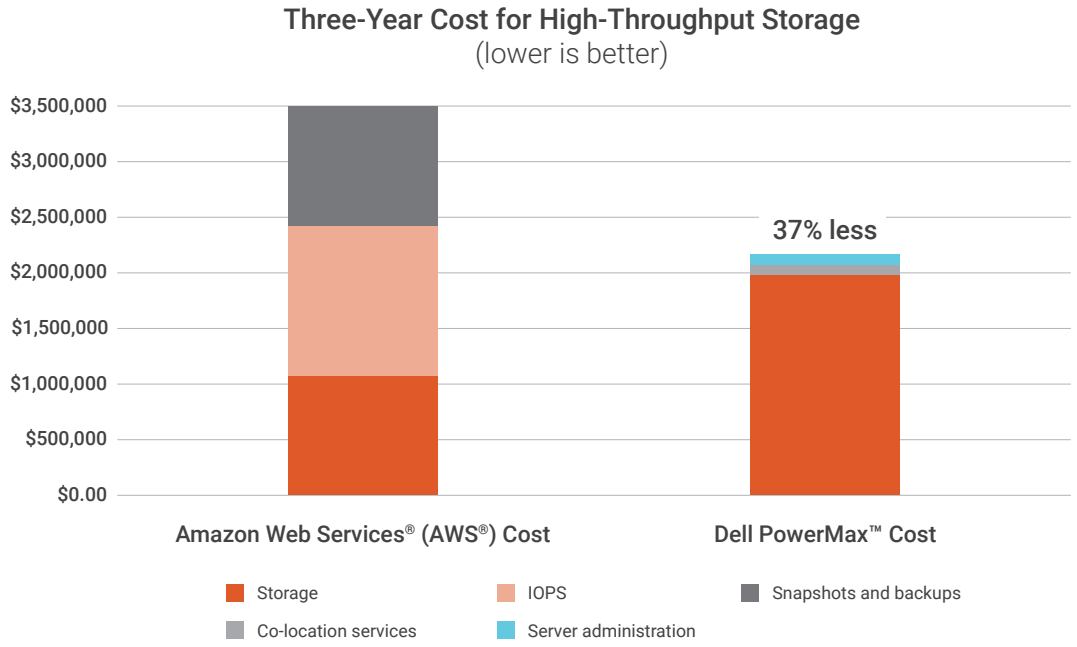

Figure 4 | Comparison of per-GiB cost of high-throughput storage on AWS® and the Dell Technologies solution

Analysis: High-Throughput Storage

Data sovereignty and regulatory compliance are increasingly critical concerns for enterprises. Public cloud deployments can introduce complexity and risk due to shared infrastructure and jurisdictional ambiguity. By contrast, private cloud solutions—such as those enabled by Dell PowerMax—offer organizations the ability to retain full control over data residency, access policies, and performance service-level agreements (SLAs).

Data control can also be an issue. Industries that require highly performant storage (such as financial services) also require full control of their data, whether for regulatory-compliance reasons or to better secure intellectual property. Co-residency in public clouds and co-locations can pose security issues for high-value data and can produce challenges for data availability. Another major issue can be server availability. Server downtime can be costly for companies in terms of both money and brand reputation. To help counter this threat, Dell PowerMax 2500 provides six-nines (99.9999%) availability,[8] versus four-nines (99.99%) with Amazon EBS.[9]

A more subtle challenge regarding data control can be government subpoenas or discovery in lawsuits interfering with access to data stored in the public cloud. Retaining data on-premises can be a solution to this challenge for organizations. The high-throughput storage analysis in this study does not include any charges for data egress or transfer, and you will need to factor in such additional fees when evaluating storage in the public cloud.

Balanced-Performance Storage Workload

While not all workloads demand high-throughput storage, organizations still require dependable, cost-effective solutions that offer control over data placement and performance. Balanced-performance storage is ideal for many enterprise applications, and deploying it in a private cloud environment enables IT teams to maintain governance, optimize resource use, and avoid the unpredictable costs often associated with public cloud storage. Consider, for example, a marketing organization that is building a new customer-facing digital platform with data needs to grow to 813 TB over three years.

To evaluate this kind of workload in the context of public cloud versus private cloud, we used AWS as the representative for the public cloud and Dell Technologies disaggregated infrastructure as the representative private cloud. For AWS, the balanced-performance storage utilized Amazon EBS gp3 storage; the Dell Technologies solution used the following:

- Dell PowerStore™ 1200T with 200 TB of NVM Express® (NVMe®) storage
- Dell™ PowerProtect™ Data Domain™ 6900 storage with 120 TB storage

Both Dell Technologies storage components run at 3.5:1 deduplication,[10] and they start with 50% committed utilization in year one. Our analysis indicates that balanced-performance storage **costs are up to 58% less on the Dell Technologies solution than on AWS**. Table 5 and Figure 5 show the results of the comparison. (For configuration details for this workload comparison, see **Appendix E**.) This cost advantage is complemented by the operational control the Dell Technologies infrastructure provides, which allows organizations to manage their data locality, retention policies, and performance tuning without the constraints or variability of public cloud environments.

Table 5 | Three-year pricing for balanced-performance storage on AWS® and the Dell Technologies solution, respectively

		Amazon Web Services® (AWS®) Cost[4]		Dell Technologies Cost[5]
Year 1	280 TB Amazon® EBS storage (50% of 980 TB)	$506,343.00	Dell PowerStore™ 1200T with backups via Dell™ PowerProtect™ Data Domain™ 6900 (50% utilization)	$306,103.00
Year 2	392 TB Amazon EBS storage (70% of 980 TB)	$708,880.19	Dell PowerStore 1200T with Dell PowerProtect Data Domain 6900 cost (70% utilization)	$379,550.00
Year 3	465 TB Amazon EBS storage (83% of 980 TB)	$840,769.37	Dell PowerStore 1200T with Dell PowerProtect Data Domain 6900 cost (83% utilization)	$442,822.00
Snapshots	30/month, Amazon EBS, saved to Amazon® S3	$994,579.22	Unlimited	Included
Co-location services	N/A	–	Three-year total	$19,877.70
Storage administration[7]	N/A	–	Three-year total	$112,500.00
Three-year total		**$3,050,571.78**		**$1,260,852.70**

Note: No discounts (such as public-cloud committed-spend discounts) were applied in this analysis.

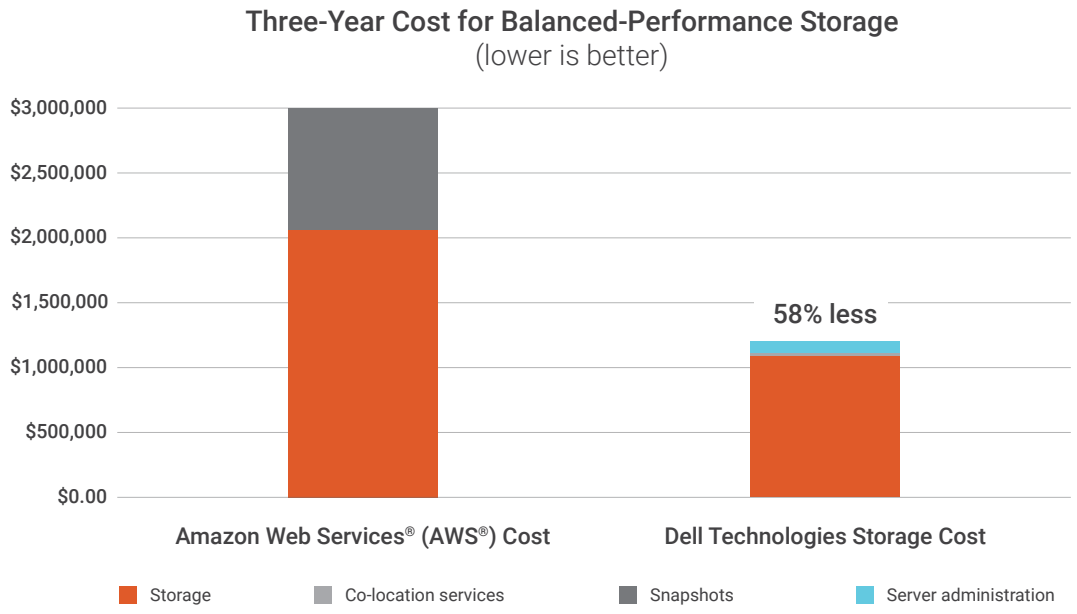

Figure 5 | Comparison of three-year costs of balanced-performance storage on AWS® and the Dell Technologies solution

Analysis: Balanced-Performance Storage

Beyond cost, organizations must weigh the governance and compliance implications of where and how their data is stored. Public cloud environments can introduce complexity through shared infrastructure and region-based limitations. In contrast, the Dell Technologies private cloud solutions offer centralized control, enabling organizations to enforce data-governance policies, reduce compliance risk, and ensure consistent performance, especially for workloads that support customer-facing platforms or sensitive business operations.

Moreover, moving data out of the public cloud, between public-cloud geographical regions, and even between availability zones can incur additional charges, all of which can quickly add up. The balanced-performance storage analysis in this study does not include charges for data egress or transfer, and you will need to factor in such additional fees when evaluating storage in the public cloud.

Other Dell Technologies Considerations

Beyond the benefits that emerge from flexible, disaggregated infrastructure solutions in general, which can provide the control afforded by private clouds, we found several additional advantages specific to Dell Technologies servers and storage. These include:

- **Interoperability**: Dell Technologies servers and storage provide organizations with an IT consumption model that can integrate tightly with their existing IT environments, allowing organizations to make use of their current investments, modernize at their own pace, and manage their infrastructures from a single console. And because Dell Technologies solutions enable organizations to manage VMs and containers within the same infrastructure, this interoperability can help organizations with IT modernization.
- **Leveraging existing team skillsets**: Dell Technologies servers and storage allow organizations to utilize their IT staff's existing knowledge base by providing unified management experiences and integration with familiar tools and platforms, enabling efficient adoption of new technology services.
- **Security and governance**: Dell Technologies servers and storage provide built-in security and compliance features that include access controls, data encryption, and regulatory certifications, which help organizations maintain a secure and compliant IT environment. By keeping sensitive data on-premises or in controlled co-location environments, the Dell Technologies infrastructure simplifies compliance with data-residency and regulatory requirements. These capabilities reduce the complexity of managing security across multicloud environments while supporting a flexible consumption model.
- **Optional managed services**: Dell Technologies offers optional managed services that allow organizations to offload some or all IT operations in a cloud-like manner. This can free internal teams to focus on strategic initiatives while maintaining operational continuity.
- **Partners and a broad ecosystem**: Dell Technologies collaborates with a wide range of partners to support hosting, outsourcing, and application-management services across diverse platforms, including public clouds. This breadth of ecosystem support can help organizations adapt to evolving workload needs.
- **Proactive and predictive support**: Dell ProSupport™ and ProSupport Plus provide 24/7 access to technology experts and proactive issue resolution through AI-driven predictive analysis. These support services help reduce downtime, help streamline maintenance, and help IT teams stay focused on innovation.

Conclusion

Some workloads might be best in the public cloud, regardless of cost considerations. A good IT strategy should take this into account and make an honest assessment of what can be moved elsewhere if doing so can bring enough benefits. These benefits could include lower costs, but they could also include other important considerations, such as workload interoperability, productivity impacts, security needs, and the convenience of having your data stored locally for restoration or DR.

Cost is an important consideration when weighing public cloud versus private cloud. Whether it is senior management telling IT to do more with less or IT billing cloud charges back to individual teams, it can feel like there is inescapable pressure to have visibility into costs and to manage those costs, particularly on public-cloud spend. You are in good company: a recent survey showed that 69% of organizations report reducing cloud spend as a goal (and of those organizations, 35% report that they aim to reduce spending on public cloud by 25% or more).[11]

Equally important are the hidden operational and compliance risks that can accompany public cloud deployments. From unpredictable billing to limited control over data residency, these challenges underscore the value of a cloud strategy that prioritizes transparency, governance, and long-term flexibility. Simplifying compliance and reducing security complexity are additional reasons why many organizations are reevaluating their cloud strategies in favor of infrastructure that offers greater transparency and control.

Prowess Consulting assessed different workloads to evaluate price differences for different use cases. These comparisons cover a wide variety of workloads common to organizations of different sizes and across industries. We found that, on average, private cloud solutions such as those from Dell Technologies can be 26–64% more cost-effective than the public cloud (as represented by AWS). These findings suggest that infrastructure solutions like those evaluated here might offer organizations a consistent and dependable foundation for modernizing IT while maintaining control over cost, performance, and compliance.

These findings support the value of a cloud-smart strategy—one that balances the strengths of both public and private cloud environments. Rather than defaulting to a single model, organizations can evaluate each workload based on its performance, compliance, and cost requirements. Private cloud solutions, such as those evaluated in this study, have evolved to offer the flexibility, scalability, and security needed to complement or even replace public cloud deployments for many use cases.

Appendix A: VM Workload Configuration Details

Dell Technologies Assumptions

- Compute, storage, and virtualization:
  - 20 x Dell PowerEdge R660 servers with 1 x Dell PowerStore 500T with 186 TB capacity
  - Three-year Dell APEX infrastructure contract with 80% commitment, storage utilization of 85%, and billing capped at 85%; actual utilization can be higher
  - 60–70 VMs per node
  - VM shapes range from 2 vCPUs with 4 GB of memory to 24 vCPUs with 128 GB of memory
- Managed services and co-location services included

AWS Assumptions

- Amazon EC2 multitenant compute on a three-year savings plan
- General-purpose solid-state drive (SSD) storage (Amazon EBS gp2) used
- Snapshots taken daily:
  - 2–4% data change rate
  - 30-day retention
- Private connectivity (10 Gbps, four connections)
- Data egress amounts to 15–30% of persistent storage
- Day-2 managed services used

Table 6 | Line-item breakdown of monthly costs for 1,200 VMs running on AWS® and the Dell Technologies solution, respectively

Category	Amazon Web Services® (AWS®) Cost[4]	Dell Technologies Cost[5]
Compute	$134,434.75	N/A
Storage	$68,287.14	N/A
Storage snapshots	$75,502.86	N/A
Compute/storage/storage snapshots/virtualization	N/A	$157,502.00/month • 20 Dell™ PowerEdge™ R660 servers • 1 Dell PowerStore™ 500T with 186 TB capacity • VMware Cloud Foundation® 60–70 VMs per node
Private connectivity	$6,570.00	Included with co-location
Data egress	$4,097.14	N/A
Co-location services	N/A	$32,740.00
Enterprise support plan	$26,624.29	Included with co-location
Managed services	$19,492.86	$54,580.00
Monthly total	**$335,009.04**	**$244,822.00**

Table 7 | Breakdown of AWS® instance families used (assuming a 30% committed-spend discount); **note**: because this workload is based on an actual enterprise deployment, specific instance types and numbers have been aggregated to help preserve confidentiality[4]

Amazon Web Services® (AWS®) Instance	Three-Year Savings Plan Monthly Subtotal[4]
C6id	$16,620.66
M6id	$73,412.69
R6id	$44,401.40
Monthly total	**$134,434.75**

Appendix B: AI Workload Configuration Details

Dell Technologies Assumptions

- Dell PowerEdge R660 and PowerEdge XE9680 servers
- NVIDIA® HGX™ H100 8-GPU SXM 80 GB, 700 W GPUs assembly
- Managed service and co-location hosting
- Five-year term ProSupport
- Dell™ ProDeploy Plus for servers

Table 8 | Assumptions for the Dell™ AI Factory on-premises solution

Workload Task	Server/Instance	GPUs per Server/Instance	Additional Purchases
Cluster management and notebooks	3 x Dell™ PowerEdge™ R660	N/A	2 x Dell PowerSwitch™ S5232-ON network infrastructure and 1 x Dell PowerSwitch N3200-ON OOB management (used by all five servers)
Data processing, model fine-tuning, and inference	2 x PowerEdge XE9680 (30 TB storage)	8 x NVIDIA® HGX™ H100	2 x Dell PowerSwitch S5232-ON network infrastructure and 1 x Dell PowerSwitch N3200-ON OOB management (used by all five servers)

Table 9 | Dell™ AI Factory on-premises solution costs

Dell™ AI Factory On-Premises Solution Line Item	Four-Year Total
Dell hardware with 5-year Dell ProSupport™ and Dell™ ProDeploy Plus (for servers)	$757,231.00
System administration	$6,663.00
Energy costs for power and cooling	$125,502.00
Data center costs for rack space	$2,080.00
Total	**$891,476.00**

AWS Assumptions

- Two P5.48xlarge Amazon EC2 instances along with two ml.R5.16xlarge instances with one 3.5 TB EBS GP2 storage each and 20 ml.t3.medium instances for notebook use would be needed to match the Dell Technologies AI workload comparison (including NVIDIA HGX H100 GPUs)
- Amazon EBS GP2 storage:
  - 2 x 3,500 GB x $0.10/GB/month = $700.00/month
- Amazon® S3 storage data egress:
  - 15 TB per month x 1,024 GB in a TB = 15,360 GB per month
  - 15,360 GB/month x $0.02/GB = $307.20/month
- Three-year Amazon EC2 Instance Savings Plan used twice (as a three-year Amazon EC2 Instance Savings Plan paired with a single one-year plan was still more expensive)
- Four-year amortization

Table 10 | Assumptions for AWS®

Workload Task	Server/Instance	GPUs per Server/Instance	Additional Purchases
Cluster management	N/A	N/A	N/A
Notebooks	20 x ml.t3.medium	N/A	N/A

Workload Task	Server/Instance	GPUs per Server/Instance	Additional Purchases
Data processing	2 x ml.r5.16xlarge	N/A	7 TB Amazon® EBS storage per month and 1 TB in and 15 TB outbound S3 data transfer
Model fine-tuning	ml.p5.48xlarge	8 x NVIDIA® HGX™ H100	N/A
Inference	ml.p5.48xlarge	8 x NVIDIA HGX H100	N/A

Table 11 | Breakdown of AWS® costs assuming 80%, 24/5 utilization

Amazon Web Services® (AWS®) Resource	Four-Year Total
1 x p5.48xlarge instance (training)	$1,644,799.66
1 x p5.48xlarge instance (inferencing)	$548,266.55
2 x ml.r5.16xlarge (data processing)	$257,079.40
20 x ml.t3.medium (notebooks)	$4,262.86
2 x 3,500 GB Amazon® EBS GP2 storage	$33,600.00
15 TB in S3 outbound data transfers	$14,745.60
Total	**$2,502,754.07**

Appendix C: Container Workload Configuration Details

Dell Technologies Assumptions

- Containers housed on integrated Dell PowerFlex racks, Red Hat Enterprise Linux OS, and Dell Cloud Platform for Red Hat OpenShift for Kubernetes, in addition to managed services
- Single rate per worker node used as the pricing structure

AWS Assumptions

- Two regions (primary and DR secondary in active/active configuration)
- Container orchestration and registry
- SUSE® Rancher for Kubernetes
- Compute with Ubuntu® OS (three-year reserved price)
- Ephemeral storage
- Persistent block SSD storage
- Snapshots taken daily:
    - 5–10% data change rate
    - 30-day retention
- Replication and data transfer
- Private connectivity (10 Gbps)
- Data egress is 15–30% of persistent storage
- Day-2 and additional managed services

Workload Assumptions for the Dell Technologies Solution and AWS

- Scale (production and DR):
    - Active/active DR strategy
    - Initial deployment: 12,000 containers
    - Year-3 end state: 40,800 containers
- Container size:
    - 2 vCPU
    - 2 GB memory
    - 47 TB ephemeral storage
    - 17 GB persistent block storage

Appendix D: High-Throughput Storage Workload Configuration Details

Dell Technologies Assumptions

- Storage provided by Dell PowerMax 2500:
  - Capacity: 850 TiB
  - IOPS: More than 2.7M per node
  - Throughput: Up to 350 GBps
- Deduplication: 3:1 to account for a wide range of data types[10]
- Capital expenditure (CapEx) purchase
- Server-administration fee allotted for managing the solution

AWS Assumptions

- Pricing:
  - Amazon EBS io2 Block Express
  - IOPS: Up to 256K per volume
  - Throughput: Up 4 GBps per volume
  - No managed services

Appendix E: Balanced-Performance Storage Workload Configuration Details

Dell Technologies Assumptions

- Powered by Dell PowerStore 1200T and a Dell PowerProtect Data Domain 6900
- Three-year term
- Server-administration fee allotted for managing the solution
- Deduplication: 3.5:1 to account for a wide range of data types[7,10]
- IOPS:
  - Dell Power Protect Data Domain 6900: 100K

AWS Assumptions

- Pricing:
  - Amazon EBS gp3
  - IOPS: Up to 16K per volume
  - Throughput: Up 1 Gbps per volume

Endnotes

[1] Based on a comparison of Dell Technologies solutions against an Amazon Web Services® (AWS®) solution as of April 2025. Source: Prowess Consulting

[2] CRN. "**AWS, Microsoft, Google Fight For $90B Q4 2024 Cloud Market Share**." February 2025.

[3] Global Industry Analysts. "**Virtualization Software**." June 2025.

[4] Pricing obtained from **https://aws.amazon.com/ec2/pricing/** as of August 2025.

[5] Pricing supplied by Dell Technologies as of August 2025.

[6] Mordor Intelligence. "**Enterprise AI Market Size and Share – Growth Trends and Forecasts (2025 to 2030)**." Accessed May 2025.

[7] 1/100 full-time employee (FTE) equivalent for a server administrator at a fully burdened annual cost of $150,000.

[8] Dell Technologies. "**Dell PowerMax data sheet**." 2025.

[9] Amazon. "**Amazon Elastic Block Store Service Level Agreement**." Updated May 2022.

[10] Prowess Consulting used 3.5:1 as a conservative figure for data reduction through deduplication even though Dell Technologies storage solutions often achieve 5:1 data reduction. Source: Dell Technologies. "**Dell PowerStore Gen 2 appliances specification sheet**." Accessed April 2025.

[11] Vega Cloud. "**2023 IT & Cloud Optimization Report**." February 2023.

0925/250025