

ECONOMIC VALIDATION

The Economic Benefits of HPE Private Cloud AI With NVIDIA AI Computing by HPE

Assessing the Cost, Performance, and Security Advantages of an Integrated AI Platform

By Luz Andrea Vasquez, Consulting Validation Analyst, and Nathan McAfee, Principal Validation Analyst
Enterprise Strategy Group

March 2025

This Economic Validation from Enterprise Strategy Group, now part of Omdia, was commissioned by HPE and is distributed under license from TechTarget, Inc.

Contents

Introduction 3

Challenges 3

The Solution: HPE Private Cloud AI 4

 Key Features and Benefits of HPE Private Cloud AI 4

Enterprise Strategy Group’s Economic Validation of HPE Private Cloud AI 5

 Cost Savings and Efficiency 6

 Faster AI Deployment and Agility 6

 Enhanced Security and Compliance 7

Enterprise Strategy Group Analysis 8

 Cost Analysis – HPE Private Cloud AI vs. Public Cloud AI 8

 HPE Private Cloud AI vs. DIY 3-year TCO Comparison 10

 Beyond-TCO Considerations: The Business Impact Factors 11


 Considerations 13

Conclusion 14

Economic Validation: Key Findings Summary

Reviewed Benefits of
HPE Private Cloud AI

Up to 60% Lower TCO
vs. public cloud over 3 years



Up to 50% Lower TCO
vs. DIY solutions over 3 years

- **Improved Cost Efficiency and Savings:** HPE Private Cloud AI simplifies AI infrastructure management, minimizing setup effort and ongoing operational resource requirements, which helps businesses control infrastructure costs and improve overall efficiency.
- **Faster AI Deployment and Agility:** HPE Private Cloud AI accelerates AI solution development and deployment through integrated tools and a flexible platform architecture, enabling organizations to rapidly respond to market demands and maintain consistent AI performance.
- **Enhanced Security and Compliance:** HPE Private Cloud AI incorporates robust security and compliance capabilities, including a zero-trust architecture and integrated governance tools, enabling businesses to securely collaborate, protect sensitive data, and confidently meet regulatory requirements.

Introduction

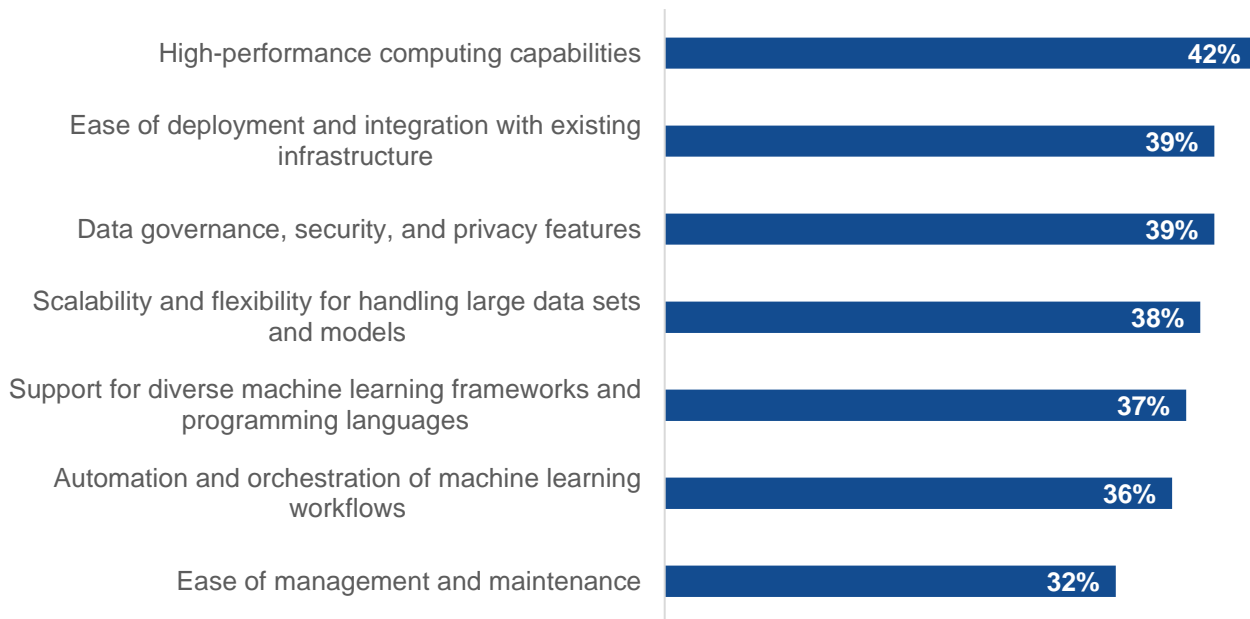
This Economic Validation from Informa TechTarget’s Enterprise Strategy Group evaluates the quantitative and qualitative benefits organizations can expect by choosing HPE Private Cloud AI (PCAI) compared to deploying and managing public cloud or traditional purchase (DIY) solutions for AI/generative AI (GenAI) infrastructure.

Challenges

As organizations expand their AI adoption, scaling from pilot projects to full production presents significant challenges. Many struggle to balance the demands of performance, security, and operational efficiency, making it difficult to fully integrate AI into business-critical operations. As shown in Figure 1, Enterprise Strategy Group research shows that 42% of organizations consider high-performance computing a top priority when selecting AI infrastructure, while 39% emphasize ease of deployment, integration, and security. This highlights the need for AI solutions that not only offer computing power but also help organizations simplify operations and compliance.¹

Figure 1. Top Priorities for AI Infrastructure: Performance, Simplicity, Security, and Scalability

When selecting AI infrastructure, which of the following capabilities are most important to your organization? (Percent of respondents, N=339, three responses accepted)



Source: Enterprise Strategy Group, now part of Omdia

Additionally, our research found that AI is being deployed across hybrid cloud (30%), public cloud (29%), edge (19%), and colocation (15%) environments, reflecting the need for scalable yet controlled infrastructure.² However, concerns around governance, security, and compliance remain a barrier to adoption. Enterprise Strategy Group research highlights that 84% of organizations report having a completely or mostly centralized AI environment, with 30% reporting that regulatory requirements are as a major factor in their choice of a physical location for their AI environment infrastructure.³

¹ Source: Enterprise Strategy Group Research Report, [Navigating the Evolving AI Infrastructure Landscape](#), December 2023.

² Ibid.

³ Ibid.

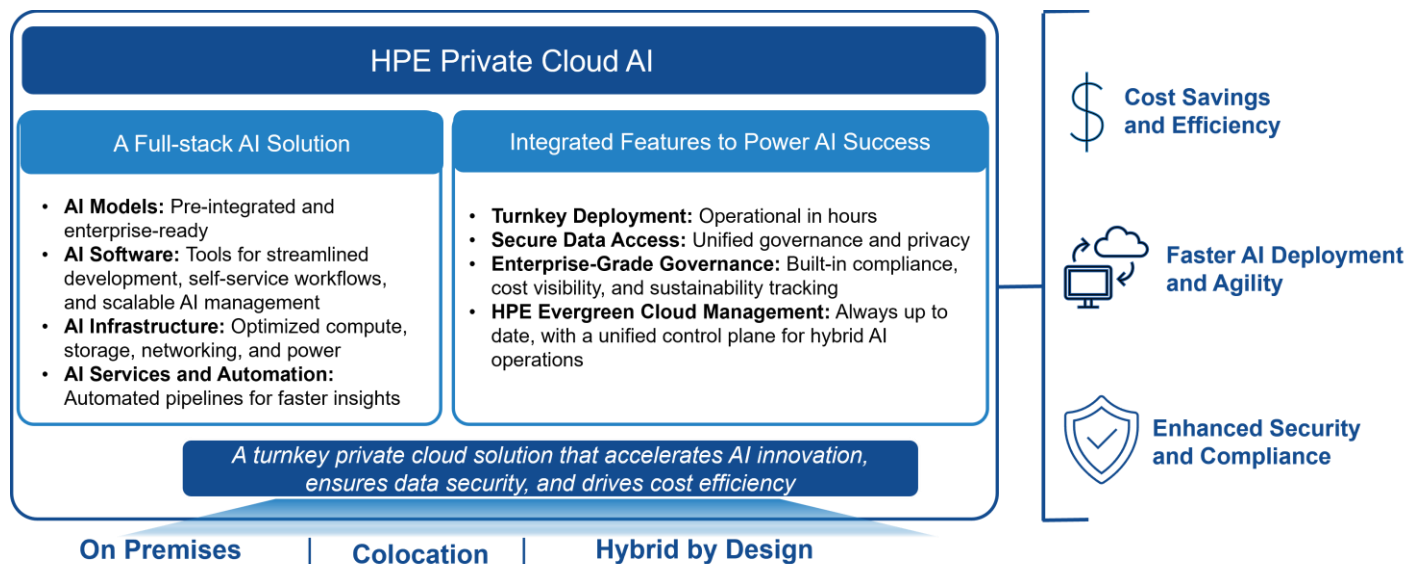
Beyond infrastructure limitations, enterprises also face operational and strategic challenges as AI adoption continues. According to Enterprise Strategy Group research, 78% of respondents report that AI adoption is in either early or mature production at their organization, with 55% reporting concerns that network security is or will be impacted by AI.⁴

Enterprises evaluating AI deployment models weigh trade-offs between DIY on-premises and public cloud AI solutions. Public cloud offers flexibility but can introduce unpredictable costs and compliance risks. DIY infrastructure provides full control but requires significant investment in hardware, security, and long-term maintenance. To succeed, organizations could benefit from a turnkey private cloud AI solution that combines scalability, security, and cost predictability without the complexity of managing everything in-house.

The Solution: HPE Private Cloud AI

To address the complexity of scaling AI securely while maintaining compliance and cost efficiency, organizations need a solution that simplifies deployment, enhances governance, and ensures high performance. NVIDIA AI Computing by HPE focuses on laying the foundational elements required for long-term AI application success with access to the processing power and scalability needed to build robust data pipelines. The first offering of this portfolio, HPE Private Cloud AI, is a fully integrated, turnkey AI solution that accelerates AI adoption by providing a pre-configured private cloud experience. It is a scalable, pre-tested AI-optimized platform that combines hardware, software, and services into a unified system, enabling organizations to efficiently build and deploy AI applications while safeguarding sensitive data.

Figure 2. HPE Private Cloud AI Architecture



Source: Enterprise Strategy Group, now part of Omdia

Key Features and Benefits of HPE Private Cloud AI

- **Instant AI productivity.** HPE Private Cloud AI provides a private cloud experience with pre-integrated hardware, software, and services, enabling rapid deployment of AI models and applications.

⁴ Source: Enterprise Strategy Group Research Report, [Architecting the Network for an AI-powered World](#), February 2025.

- HPE Private Cloud AI includes built-in AI tools such as Airflow, EzPresto, and Superset for data engineers to streamline data processing, visualization, and infrastructure management. Additional tools like Kubeflow, Ray, and MLflow enable seamless orchestration across the entire machine learning (ML) lifecycle, enhancing collaboration and automation.
- Out-of-the-box deployment ensures customers can go from setup to operation in just a few hours, accelerating time to value for AI investments.
- HPE Private Cloud AI optimizes AI workflows, improving accuracy, efficiency, and infrastructure performance.
- **Unified data access and security.** By establishing a unified, global namespace, the built-in HPE Ezmeral Data Fabric consolidates heterogeneous data assets, improving data management, governance, and security.
 - Built-in fine-grained access controls protect sensitive data and models, maintaining data integrity and privacy.
 - Multi-tenant collaboration enables teams to track experiments, iterate faster, and build on each other's work.
- **Enterprise-grade confidence and control.** Comprehensive multilayered security and compliance features protect AI models and sensitive data.
 - End-to-end AI compliance tracking enables full lineage, traceability, and auditability of AI models.
 - Single-source global support from HPE and NVIDIA ensures infrastructure reliability, reducing operational risks.
- **Flexible and scalable cloud experience.** A modular, cloud-native architecture enables organizations to start small and scale AI workloads as needed.
 - Advanced monitoring tools via the HPE GreenLake console optimize performance, sustainability, and cost.
 - HPE Private Cloud AI is deployable in secure Equinix data centers through HPE GreenLake for hybrid IT strategies.

This solution is suited to address the challenges enterprises face when transitioning AI initiatives from pilots to production, enabling faster innovation, operational efficiency, and robust data governance, while maintaining control over costs and mitigating financial risks.

Enterprise Strategy Group's Economic Validation of HPE Private Cloud AI

Enterprise Strategy Group's economic analysis suggests that, when deployed in on-premises environments, HPE Private Cloud AI provides measurable value by improving cost efficiency, accelerating AI deployments, and enhancing security and compliance. Our research identified potential savings and benefits in the following key areas:

- **Cost savings and efficiency.** Businesses deploying HPE Private Cloud AI could find it easier to manage costs by simplifying how AI solutions are set up and run. This platform reduces the need for extensive manual setup and cuts down on the IT resources needed to keep everything running, potentially resulting in more efficient infrastructure and reduced ownership costs.
- **Faster AI deployment and agility.** HPE Private Cloud AI's comprehensive design is built to speed up how quickly AI products are developed and released. With a strong focus on providing integrated tools and flexible architecture, this solution can help companies adapt quickly to changing market needs while keeping their AI applications performing reliably.
- **Enhanced security and compliance.** Organizations with a focus on security and compliance will likely benefit from the features offered in HPE Private Cloud AI. By integrating a zero-trust architecture, governance, and compliance tools, this platform enables secure AI collaboration, protecting sensitive data and ensuring regulatory compliance.

Cost Savings and Efficiency

HPE Private Cloud AI is designed to help organizations manage AI infrastructure costs more effectively by simplifying deployment, operations, and infrastructure management. This section explores how HPE Private Cloud AI can reduce costs related to setup, ongoing management, and overall infrastructure efficiency.

- **Deployment cost savings.** HPE Private Cloud AI provides its customers with pre-integrated hardware, software, AI models, and services designed for AI workloads, enabling rapid installation in less than eight hours, as reported by HPE. This streamlined process greatly reduces the need for extensive manual setup, enabling organizations to get their AI infrastructure up and running faster. In contrast, deploying a cloud-based AI infrastructure can take anywhere from a few days to several weeks, depending on many factors such as network configuration, compliance requirements, and data migration. DIY deployments often face even greater delays due to approval processes and governance policies.
- **Operational cost savings.** HPE Private Cloud AI simplifies lifecycle management with unified dashboards that offer a centralized view of AI infrastructure performance and operations. This integrated approach has the potential to reduce complexity by automating routine management tasks, letting organizations operate their AI environments more efficiently with fewer IT staff. By centralizing monitoring and controls, HPE Private Cloud AI can help lower ongoing operational costs tied to staffing and IT management. Industry trends indicate that infrastructure automation often leads to reduced operational complexity and greater cost efficiency over time.
- **Infrastructure and sustainability savings.** Organizations deploying AI infrastructure on premises with HPE Private Cloud AI can avoid egress fees commonly incurred with public cloud platforms. This approach provides more predictable expenses and optimized power consumption. Additionally, HPE's fully integrated, architected AI system minimizes the overall hardware footprint compared to DIY solutions, enabling organizations to scale AI infrastructure based on business needs. This helps reduce overprovisioning and associated upfront costs. According to HPE's internal cost analysis, organizations using HPE Private Cloud AI could achieve cost savings of up to 60% compared to public cloud solutions and up to 50% compared to traditional DIY deployments (detailed further in this paper). In terms of sustainability, HPE's focus includes energy-efficient AI infrastructure, with potential energy cost savings of up to 30% with the HPE GreenLake delivery model compared to DIY solutions. Enterprise Strategy Group found that HPE's sustainability initiatives—such as upcycling 85% of returned technology and filing over 6,000 patents for low-carbon technologies since 2004—offer sustainability benefits to organizations that use HPE Private Cloud AI.

HPE reports that HPE Private Cloud AI can be deployed in less than eight hours, offering faster setup times compared to traditional DIY and cloud solutions.

Faster AI Deployment and Agility

HPE Private Cloud AI is designed to help organizations adopt AI faster by making development smoother and speeding up deployment. By improving productivity, reducing time to market, and letting AI systems scale easily, businesses can bring AI-driven innovations to production more efficiently.

- **Accelerated AI productivity.** HPE Private Cloud AI offers a ready-to-use platform with pre-validated models, tools, and AI-optimized hardware. By simplifying model training and reducing development hurdles, this solution aims to make AI projects faster and smoother. According to HPE, these improvements could mean up to 2x better developer productivity. As a result, AI teams can make better use of their time and resources.

According to HPE, its AI-optimized infrastructure delivers up to 4x faster inference and 5x greater model throughput, accelerating time to market for AI solutions.

- **Faster time to insight and market.** Enterprise Strategy Group reviewed HPE's approach to accelerating AI deployment, which includes 4x faster inference and 5x greater model throughput and found it to be a reasonable and achievable strategy for most organizations looking to scale AI. By reducing development time and streamlining the transition from prototyping to production, HPE Private Cloud AI enables businesses to bring AI-driven solutions to market more quickly, helping them stay competitive and respond to new opportunities more effectively.
- **Seamless scalability.** Companies deploying HPE Private Cloud AI benefit from a modular, cloud-native architecture that adapts to dynamic workloads, ensuring operational flexibility and consistent performance, even during demand surges. By scaling AI infrastructure as needed, businesses can support growth and adapt to new opportunities driven by AI. Compared to public cloud, this approach offers consistent performance with predictable scaling costs, while avoiding the manual intervention and upfront investments often required with DIY setups.

Enhanced Security and Compliance

Organizations deploying AI need strong security and compliance measures to protect sensitive data and meet regulatory requirements. Many find that integrated security and governance tools help reduce operational complexity and mitigate risks as AI adoption scales.

- **Comprehensive security.** Organizations deploying AI must protect their sensitive data and maintain compliance with industry regulations. Managing security in DIY solutions requires extensive internal resources, ongoing manual configurations, and constant risk mitigation monitoring. Enterprise Strategy Group finds HPE Private Cloud AI's built-in zero-trust architecture can help customers simplify security management, reducing the need for dedicated security expertise and minimizing manual oversight. By automating security controls and integrating compliance tools, organizations can lower the risk of breaches, reduce operational costs, and maintain data governance more efficiently. Compared to the public cloud, where security responsibility is shared between providers and users, HPE Private Cloud AI enables organizations to retain full control over data protection while reducing complexity.
- **Streamlined compliance.** We believe HPE Private Cloud AI can help organizations meet their regulatory requirements with its governance and compliance tools. This approach helps organizations simplify their compliance processes, potentially lowering the risk of human error and freeing up IT resources. In contrast, public cloud platforms typically follow a shared responsibility model. While these platforms handle the security of the underlying infrastructure, businesses remain responsible for securing their own workloads and managing compliance tasks. In contrast, DIY solutions often require building compliance measures from scratch, making HPE Private Cloud AI's integrated solution a more efficient one.
- **Secure collaboration and reliability.** Enterprise Strategy Group found that HPE Private Cloud AI provides teams with integrated access to tools and pre-validated models in a secure environment. This setup fosters smoother collaboration between teams working on AI projects, reducing errors and helping ensure sensitive data remains protected. As a result, organizations can speed up development cycles and achieve better project outcomes. In public cloud environments, managing security across a shared infrastructure often requires careful configuration of access controls. DIY solutions, on the other hand, could involve building secure collaboration capabilities from scratch, which can increase time and cost. The cohesive design of HPE Private Cloud AI can help drive savings through improved productivity and reduced errors.

The HPE Private Cloud AI platform features a zero-trust architecture with built-in security and governance, helping businesses reduce risks and simplify compliance management.

Enterprise Strategy Group Analysis

Enterprise Strategy Group conducted an economic analysis to assess the potential cost implications of deploying HPE Private Cloud AI with NVIDIA AI Computing by HPE compared to public cloud AI services and DIY AI infrastructure. Using a structured Economic Validation methodology, we reviewed HPE-provided total cost of ownership (TCO) estimates, subject matter expert insights, technical documentation, and broader industry cost trends to analyze key financial factors such as hardware and infrastructure costs, operational expenses, and scalability considerations. While the HPE Private Cloud AI analysis highlights potential cost savings, it is important to recognize that the infrastructure management cost is not included in the comparison against public cloud. Although HPE Private Cloud AI is still in its early market adoption phase, its projected TCO provides a useful framework for evaluating the expected cost structure of on-premises AI infrastructure. The following analysis presents a three-year TCO comparison, highlighting cost efficiency, predictability, and operational impact relative to public cloud and DIY solutions.

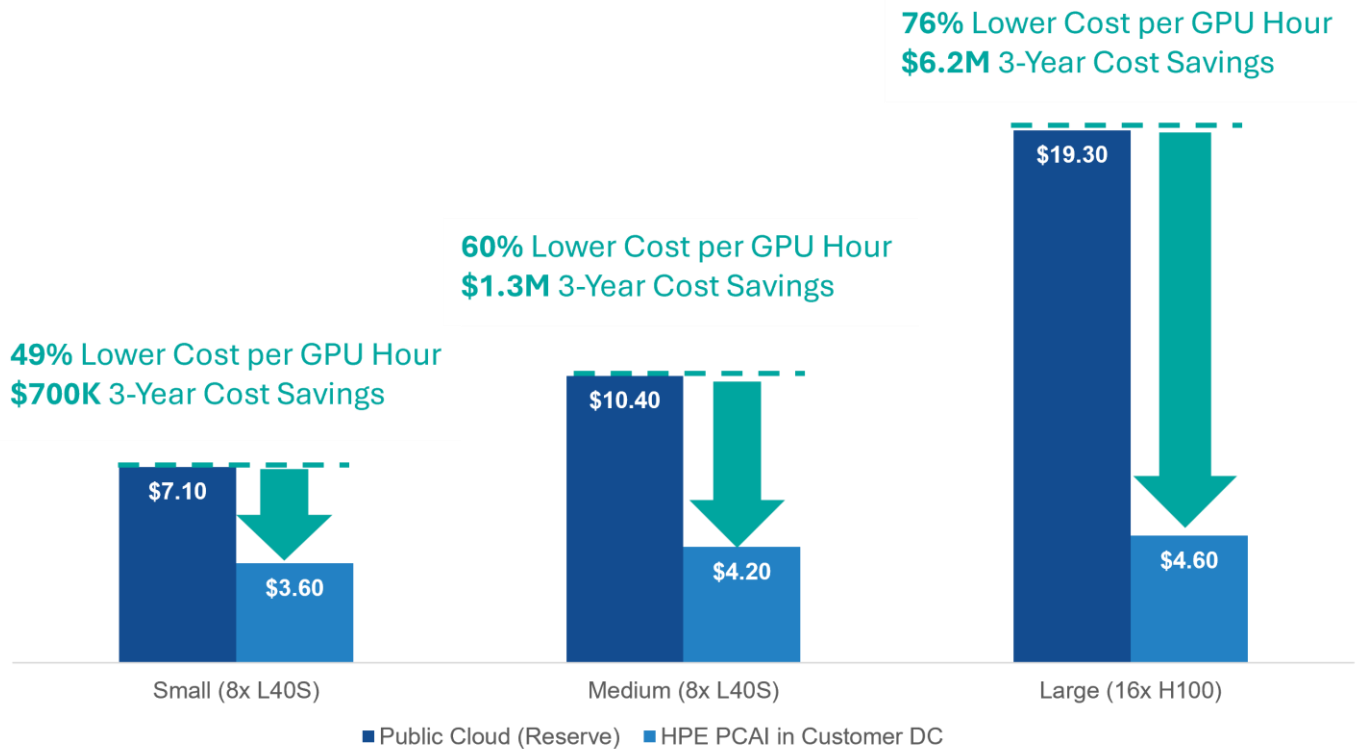
Cost Analysis – HPE Private Cloud AI vs. Public Cloud AI

The following three-year cost analysis of HPE Private Cloud AI versus public cloud AI services is based on HPE-provided cost estimates, incorporating financial factors such as hardware, software, and operational expenses. Public cloud GPU pricing assumes a 50% discount off standard on-demand rates, reflecting reserved instance pricing models. The analysis considers three HPE Private Cloud AI configurations to reflect different AI scaling needs: small and medium deployments with 8 L40S GPUs each and a large deployment with 16 H100 GPUs.

HPE Private Cloud AI hardware is treated as a one-time investment, following a three-year depreciation model with a 20% residual value at the end of the term. Software and AI tools are included in both environments, with public cloud AI bundling ML, data preparation, model monitoring, and debugging tools, while HPE Private Cloud AI provides its own AI management capabilities within a unified platform. Storage costs are modeled to match HPE Private Cloud AI storage needs, assuming a 50% split between object storage and scalable file storage, and enterprise-grade support costs are included for both environments. For HPE Private Cloud AI, the analysis assumes deployment in a customer-owned data center, where only utility costs—electricity and cooling—are factored in. This analysis does not include infrastructure management costs for HPE Private Cloud AI, whereas in a public cloud model, some of these costs are integrated into the service pricing. As a result, the organization's actual savings could be around 20-30% lower when factoring in additional infrastructure management costs.

As shown in Figure 3, the analysis compares different AI infrastructure cost models over three years, highlighting percentage cost savings per GPU hour across different HPE Private Cloud AI deployment sizes (small, medium, and large). The figure presents three pricing models—On-Demand, Reserved Pricing, and HPE Private Cloud AI—illustrating how cost structures vary based on commitment levels and consumption models. The analysis assumes a full-scale production AI operational usage model, where workloads require continuous processing and high availability to support enterprise AI applications.

Figure 3. HPE Private Cloud AI vs. Public Cloud: Percentage Cost Savings per GPU Hour
3-year TCO Savings Across Different Public Cloud AI Sizes (24/7 Operations)



*Note: Infrastructure management costs for HPE PCAI are not included. Actual savings may be ~20-30% lower

Source: Enterprise Strategy Group, now part of Omdia

Table 1. Time-to-value Phases for AI Infrastructure Comparison: Public Cloud vs. HPE Private Cloud AI

Deployment Size	Public Cloud (Reserve)	HPE Private Cloud AI in Customer DC	3-year Cost Savings	% Savings Per GPU Hour
Small (8x L40S)	\$1.5M	\$0.8M	\$700K	49% Lower
Medium (8x L40S)	\$2.2M	\$0.9M	\$1.3M	60% Lower
Large (16x H100)	\$8.1M	\$1.9M	\$6.2M	76% Lower

Infrastructure management costs for HPE Private Cloud AI are not included. Actual savings may be ~20-30% lower

Source: Enterprise Strategy Group, now part of Omdia

Key Findings: Cost Efficiency and Value of HPE Private Cloud AI

- **HPE Private Cloud AI delivers significant cost savings for AI workloads with sustained demand.** As illustrated in Figure 3 and Table 1, HPE's estimated savings indicate that HPE Private Cloud AI has the potential to deliver cost efficiencies over a three-year period when compared to public cloud reserved pricing AI services. These savings are particularly evident for workloads running at full production, with cost benefits increasing as AI infrastructure scales. Although real-world validation from deployed customers is still pending, these cost projections are consistent with broader industry trends in on-premises versus cloud TCO models. Across various deployment sizes, HPE Private Cloud AI demonstrates cost advantages:
 - Small-scale deployments (8x L40S GPUs): Up to 49% lower GPU-hour costs, saving \$700K over three years.
 - Medium-scale deployments (8x L40S GPUs): Up to 60% lower GPU-hour costs, saving \$1.3M over three years.
 - Large-scale deployments (16x H100 GPUs): Up to 76% lower GPU-hour costs, saving \$6.2M over three years.
- **Value even for lower AI workloads.** While HPE Private Cloud AI delivers the strongest cost savings for continuous AI workloads, it remains cost-effective in select configurations even at lower usage levels. Organizations with less-frequent AI compute demand should evaluate whether HPE Private Cloud AI's predictable costs outweigh the flexibility offered by public cloud.
- **Balanced cost predictability and flexibility.** HPE Private Cloud AI offers cost stability and infrastructure control, making it optimal for organizations prioritizing long-term cost efficiency. Public cloud models might be more suitable for dynamic workloads requiring on-demand scalability, despite higher costs. Selecting between HPE Private Cloud AI and the public cloud should depend on workload predictability, operational priorities, and long-term infrastructure strategy.
- **Considerations for AI infrastructure.** Public cloud pricing integrates full infrastructure and platform management, whereas HPE Private Cloud AI does not include these costs. Actual savings with HPE Private Cloud AI may be approximately 20–30% lower once these management costs are included, resulting in an expected TCO savings of approximately 60% compared to the public cloud. Decision-makers should assess workload intensity, cost predictability, and infrastructure management needs when defining their AI deployment strategy.

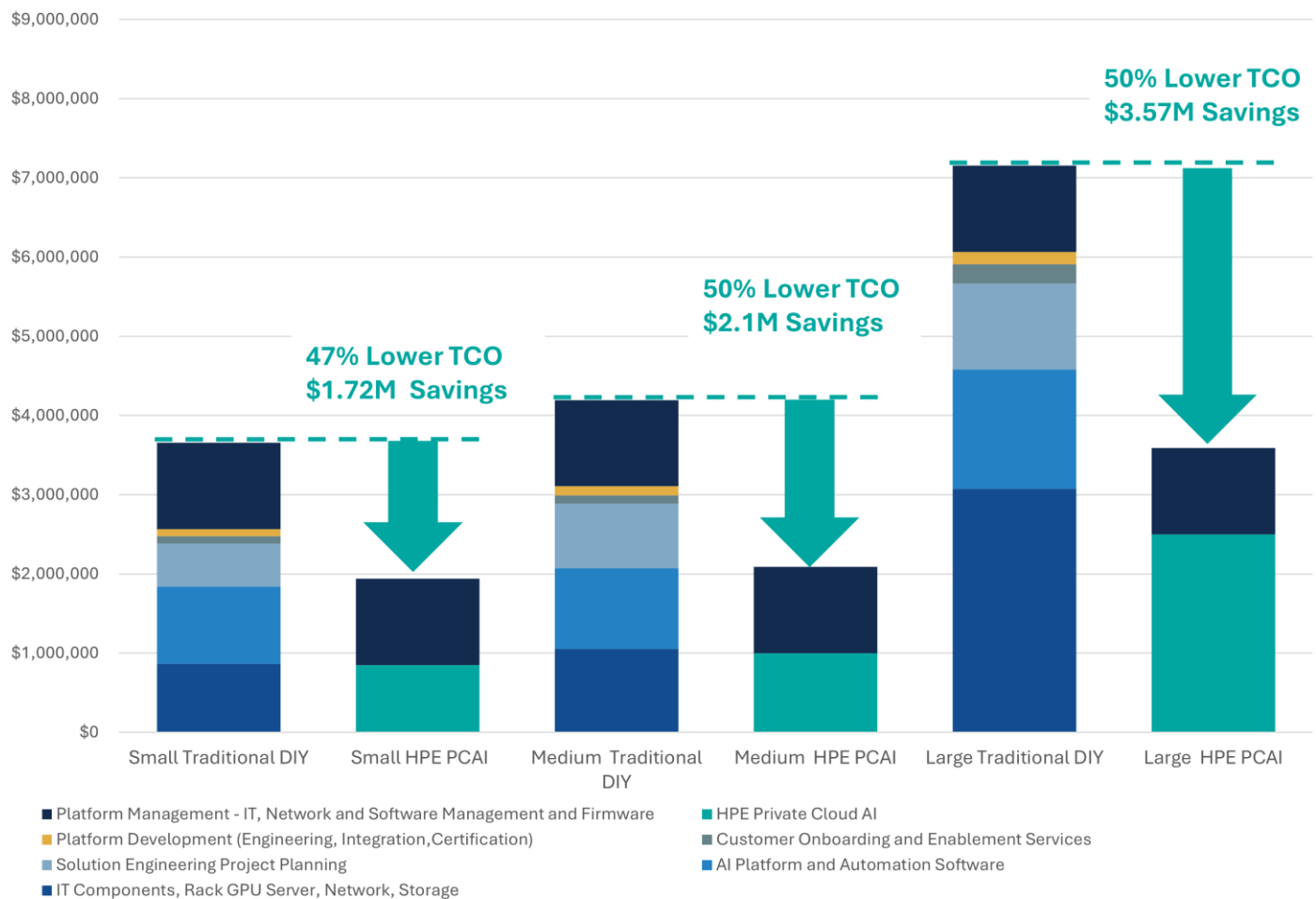
HPE Private Cloud AI vs. DIY 3-year TCO Comparison

This section presents a three-year TCO comparison between HPE Private Cloud AI and DIY infrastructure across three AI deployment sizes: small, medium, and large. The analysis includes key cost elements such as hardware, software, engineering services, and platform management.

To ensure a balanced comparison, platform management costs, including IT, network, software management, and firmware, were assumed to be identical for each deployment size and consistent across both DIY and HPE Private Cloud AI configurations, totaling \$1,087,200 for small, medium, and large setups.

Figure 4 below highlights differences in total purchase costs, as well as percentage and dollar savings for each configuration, illustrating the financial impact of deploying HPE Private Cloud AI compared to a traditional DIY approach.

Figure 4. HPE Private Cloud AI vs. DIY – 3-year TCO Savings



Source: Enterprise Strategy Group, now part of Omdia

Key Takeaways: Financial Efficiency and Cost Savings

HPE Private Cloud AI delivers up to 50% TCO savings across small, medium, and large AI infrastructure configurations compared to DIY solutions. Small configurations achieved up to **\$1.72M** in cost efficiencies, medium configurations up to **\$2.11M**, and large ones up to **\$3.57M** over three years. These savings are mainly driven by reduced engineering, operational, and software licensing costs.

- Operational efficiency savings: HPE GreenLake’s support model enables HPE Private Cloud AI to deliver up to **\$729K** for small configurations, **\$1.03M** for medium configurations, and up to **\$1.49M** for large configurations in labor efficiency alone.
- Software license savings: HPE Private Cloud AI reduces software licensing costs through their HPE GreenLake consumption-based model. Organizations can expect up to **\$974K** for small configuration sizes, **\$1.02M** for medium and up to **\$1.50M** for large sizes.

Beyond-TCO Considerations: The Business Impact Factors

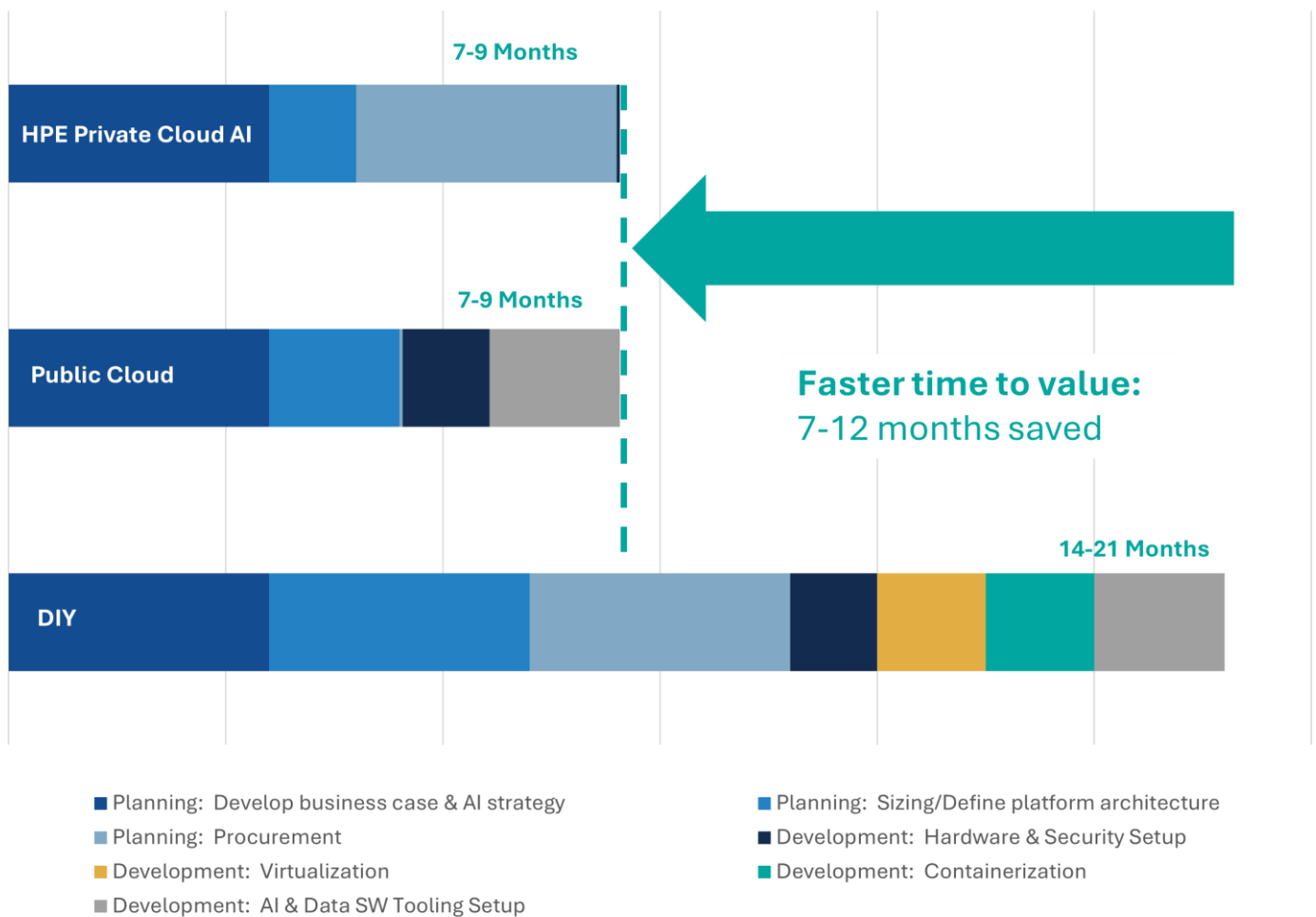
While TCO is an important factor in AI infrastructure decisions, organizations should also evaluate other business impact considerations. These include operational efficiency, time to value, risk management, and strategic control.

The following subsections outline key areas where HPE Private Cloud AI offers value that goes beyond direct cost savings.

Faster Time to Value: Infrastructure Readiness and Deployment Time

Figure 5 compares the AI configuration timelines for HPE Private Cloud AI, public cloud, and DIY solutions across different planning and development phases. This comparison highlights how each approach affects overall time to value for AI infrastructure deployment.

Figure 5. AI Platform Configuration Time: HPE Private Cloud AI vs. Alternatives



Source: Enterprise Strategy Group, now part of Omdia

Deploying a DIY AI platform requires a lengthy process of sourcing GPUs, configuring networking, and managing AI pipelines. HPE Private Cloud AI provides an integrated, out-of-the-box environment that reduces AI infrastructure and enables organizations to **deploy AI solutions seven to nine months ahead of DIY alternatives** and offers **approximately the same time to value as public cloud solutions**, accelerating business transformation.

Table 2 offers a quick comparison of key factors affecting AI infrastructure decisions across HPE Private Cloud AI, DIY, and public cloud solutions. Detailed explanations of each factor follow the table for a deeper understanding of the business impact.

Table 2. Additional Business Impact Factors for AI Infrastructure

Factor	HPE Private Cloud AI	Traditional DIY	Public Cloud
AI Project Success Rate	Integrated platform reduces deployment hurdles	Higher risk of deployment hurdles	Moderate risk due to complexity and scalability
Compliance and Security	Full data ownership, reduced compliance complexity	Requires internal compliance management	Shared responsibility can introduce complexities
Long-term Control	Supports full control over data and workflows	Full control but requires infrastructure management	Higher risk of vendor lock-in, limited customization
Cost Predictability	Fixed-cost structure avoids hidden costs	Predictable but high upfront capital costs	Potential for unpredictable costs

Source: Enterprise Strategy Group, now part of Omdia

- **AI project success rate.** Research showed that up to 80% of AI projects fail due to data complexity, scalability issues, and operational inefficiencies.⁵ HPE Private Cloud AI’s integrated, pre-validated AI platform eliminates common deployment hurdles by offering standardized AI infrastructure with built-in automation and support.
- **Compliance and security cost risk.** For highly regulated industries (e.g., finance, healthcare, defense), public cloud AI solutions introduce compliance complexities around data sovereignty, shared responsibility models, and access control policies. HPE Private Cloud AI enables organizations to maintain full data ownership and security without additional regulatory burden.
- **Long-term AI control and expertise retention.** Cloud AI providers optimize for their ecosystem, which might not align with long-term enterprise AI needs. HPE Private Cloud AI ensures enterprises retain full control over their data, models, and AI workflows, avoiding vendor lock-in while maintaining internal expertise in AI infrastructure management.
- **TCO predictability and hidden cloud costs.** While cloud-based AI solutions are often marketed as cost-efficient, they can introduce unpredictable costs, such as storage egress fees, API call charges, and the need for higher-performance instances than originally planned. HPE Private Cloud AI fixed-cost structure eliminates these variables, providing greater long-term financial predictability.

Considerations

Enterprise Strategy Group’s analysis of HPE Private Cloud AI versus public cloud and DIY solutions is built on good faith and credible, conservative, and validated assumptions provided by HPE. However, no single modeled scenario can capture the specific operational and financial factors of every organization. Each enterprise has unique needs that impact potential cost savings, including infrastructure strategy, AI workload patterns, and scaling requirements. The savings identified in this study are based on a representative deployment scenario and are intended to provide a general framework for potential outcomes. Actual savings may vary depending on factors such as hardware and software configurations, existing IT environment, and the efficiency of internal IT teams. Enterprise Strategy Group recommends that organizations evaluate their specific requirements, business goals, and current infrastructure when assessing the benefits of HPE Private Cloud AI. Consulting with their HPE

⁵ Source: James Ryseff, Brandon F. De Bruhl, Sydne J. Newberry, "[The Root Causes of Failure for Artificial Intelligence Projects and How They Can Succeed](#)," Rand.org, August 13, 2024.

representative can help enterprises explore and identify the AI deployment model that best aligns with their operational and financial objectives.

Conclusion

Choosing the right AI infrastructure is challenging for businesses looking to move from pilot projects to full-scale production. Public cloud solutions provide flexibility but can introduce unpredictable costs, compliance challenges, and data control concerns. On the other hand, DIY models offer more control but often require substantial upfront investment and extensive in-house management and specialized expertise. While HPE Private Cloud AI mitigates many of these challenges by offering a fully integrated solution, organizations must still consider infrastructure planning and ongoing operational needs.

HPE Private Cloud AI is designed to help organizations streamline AI deployment by combining infrastructure, software, and AI tools into a single platform. Enterprise Strategy Group reviewed HPE's internal analysis and found it reasonable that HPE Private Cloud AI could lower costs by up to 50% compared to DIY solutions and up to 60% compared to public cloud, especially for organizations running continuous AI workloads. Deploying a DIY AI platform typically involves a lengthy process of sourcing GPUs, configuring networking, and managing AI pipelines. In contrast, HPE Private Cloud AI provides a pre-integrated, out-of-the-box environment that reduces AI infrastructure configuration time, enabling organizations to get **AI workloads up and running seven to nine months ahead of DIY alternatives** while delivering a time to value comparable to public cloud solutions. The platform is designed to simplify AI deployment, enhance operational efficiency, and provide a more predictable cost structure. Additionally, HPE Private Cloud AI's built-in security and scalability let businesses accelerate AI adoption while minimizing vendor lock-in risks.

Beyond cost considerations, organizations must evaluate AI project success rates, security and compliance risks, and long-term control over their AI infrastructure. While public cloud solutions might offer rapid initial setup, they can introduce cost variability and governance challenges. DIY solutions provide control but require significant internal investment. HPE Private Cloud AI presents an alternative that balances cost efficiency, security, and operational flexibility. Enterprise Strategy Group finds that HPE Private Cloud AI aligns with key industry needs and offers a strong value proposition for AI deployment. Future customer validation will further confirm these findings, and organizations should engage with HPE representatives to explore how HPE Private Cloud AI fits their specific AI strategies.

©2025 TechTarget, Inc. All rights reserved. The Informa TechTarget name and logo are subject to license. All other logos are trademarks of their respective owners. Informa TechTarget reserves the right to make changes in specifications and other information contained in this document without prior notice.

Information contained in this publication has been obtained by sources Informa TechTarget considers to be reliable but is not warranted by Informa TechTarget. This publication may contain opinions of Informa TechTarget, which are subject to change. This publication may include forecasts, projections, and other predictive statements that represent Informa TechTarget's assumptions and expectations in light of currently available information. These forecasts are based on industry trends and involve variables and uncertainties. Consequently, Informa TechTarget makes no warranty as to the accuracy of specific forecasts, projections or predictive statements contained herein.

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 Informa TechTarget, 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 Client Relations at cr@esg-global.com.

About Enterprise Strategy Group

Enterprise Strategy Group, now part of Omdia, provides focused and actionable market intelligence, demand-side research, analyst advisory services, GTM strategy guidance, solution validations, and custom content supporting enterprise technology buying and selling.

 contact@esg-global.com

 www.esg-global.com