# Why private cloud is the key to AI success

4 best practices for maximizing your private cloud with NVIDIA AI Computing by HPE

Artificial intelligence (AI) has captured the attention of businesses worldwide, promising to completely transform industries with its ability to automate processes, analyze data, and predict outcomes with remarkable precision. But as AI — and especially generative AI (GenAI) — take on more prominent roles, the strain on traditional infrastructure is increasing. To avoid buckling under that strain, organizations are turning to private cloud, a compelling solution that enables smarter and more efficient workload deployment.

Why is private cloud becoming so compelling? Private cloud combines the agility of cloud computing with the control of on-premises infrastructure as explored in Shenkan & Associates white paper "The Private Cloud Advantage for AI".[1] Think of it as your AI command center: secure, scalable, and custom-fit to your needs. According to IDC, 61% of organizations already use private cloud for AI[2] and it is rapidly becoming the go-to approach for other businesses looking to take advantage of the technology.

"Public cloud alone isn't the answer for AI workloads," says Michael Corrado, senior worldwide marketing manager for AI and private cloud at Hewlett Packard Enterprise. "Organizations find that private cloud offers superior control and scalability over public cloud — so long as they follow certain best practices."

## What makes private cloud different?

Unlike public cloud, which is shared across multiple users and may be more vulnerable to data breaches, private cloud is owned and managed by a single organization. It offers virtualized computing and storage resources tailored to specific workload needs, ensuring optimal performance and control.

Such precision matters in AI, where workloads are both resource-intensive and data-sensitive. In a private cloud environment, organizations can customize hardware, software, and security configurations to support high performance tasks such as GenAI model training or real-time analytics. The ability to scale seamlessly is another key factor that sets apart private cloud, enabling businesses to adapt as AI demands grow.

Specifically, private cloud offers:

## 1. Stronger security and data control

AI thrives on data, but that data is often proprietary or sensitive. Private cloud ensures it stays protected. Unlike public cloud, private environments let organizations enforce strict access controls and comply with regulations like the European Union's General Data Protection Regulation or the recently enacted EU AI Act. According to a Forrester survey, 86% of high-maturity organizations using private clouds report better security outcomes.[3] This isn't just about keeping hackers out; it's about safeguarding intellectual property and ensuring AI models deliver trustworthy results. For many organizations, making certain private data doesn't end up as part of a public AI model is critical.

## 2. Customizable performance

AI workloads demand more than just computing power. You also need the ability to fine-tune all the technology supporting those workloads, including GPUs, storage, and bandwidth. Private cloud offers the computational power and ideal controls that organizations need to optimize costs, improve performance, and ensure more predictable workloads. In addition, the Forrester survey found that 81% of mature organizations that embraced private data centers experienced improved collaboration or fewer data silos.

## 3. Better cost predictability

AI's appetite for resources can make public cloud costs spiral out of control. Private cloud offers a more predictable financial model, especially for organizations with consistent, high-demand workloads. One study calculated that private cloud infrastructure can cost 45% less than comparable public cloud services for the most intensive tasks.[4] And that's just the start: One U.S. retailer, for example, said it slashed cloud expenses by up to 90% after migrating to private cloud, opening new pathways to reinvest those savings.[5]

[1] "The Private Cloud Advantage for AI," Shenkan and Associates, October 2024

[2] "Essential Elements for Private Cloud Strategies," IDC, August 2024

[3] "Cloud Maturity Drives Business Success," Forrester, June 2024

[4] "The Economic Case for Hybrid Cloud," 451 Research

[5] "Private cloud makes its comeback, thanks to AI," CIO, May 14, 2024

# 4. Seamless scalability

As AI adoption grows, infrastructure must scale to keep up. But scaling AI projects is tough. A recent survey found that 51% of AI projects currently in production are yet to be delivered at scale.[6] Private cloud makes this type of scaling easier, offering the flexibility to add resources without the exponential cost increase often seen in public cloud. Hybrid models, where workloads can shift between private and public clouds as needed, further enhance flexibility and scalability.

# 5. Low latency and higher bandwidth

Speed isn't negotiable for certain applications, such as operating autonomous vehicles or conducting real-time fraud detection. Private cloud keeps data processing close to its source, minimizing latency and maximizing throughput. These capabilities are critical for organizations that require split-second insights.

## Best practices for maximizing private cloud

While private cloud offers distinct advantages, realizing its full potential requires a smart strategy. Here are four ways to improve how private cloud works for you:

1. **Start with clear goals.** Jumping into AI without a plan is a recipe for disaster, resulting in wasted time and money. Define the specific inefficiencies you want to address and the outcomes you're aiming to achieve. Whether automating customer support or optimizing supply chains, a focused use case ensures your AI investment will deliver the desired results.

   HPE AI Services experts can help you define your AI initiative plans, assess required data for your use case, establish AI security and data privacy approaches, define your technology stack, and figure out how you'll show proof of value to your key stakeholders.

2. **Choose a full-stack solution.** AI isn't plug and play. Success requires seamless integration of hardware, software, and support. Look for solutions that are designed to reduce complexity.

   HPE Private Cloud AI is one such solution. This scalable, pre-configured, full stack private cloud option is part of the NVIDIA AI Computing by HPE portfolio. A turnkey option codeveloped with NVIDIA, it delivers quick, easy, out-of-the-box AI productivity with enterprise-grade confidence, control, and security. Organizations get everything they need to immediately develop, deploy, and monitor AI applications, models, and workloads seamlessly across hybrid cloud and multicloud environments. "HPE Private Cloud AI lets you deploy enterprise-grade AI workloads for your business in hours instead of days or months," says Corrado. "It's enterprise-grade and profit-ready."

3. **Prioritize security and compliance.** Data is the lifeblood of AI and protecting it should be nonnegotiable. Unfortunately, a recent Deloitte survey found that 55% of organizations are so worried about data-related problems that they avoid embracing GenAI for many use cases.[7]

   To overcome these concerns, organizations should look for private cloud environments that include robust governance tools that allow them to monitor and control data usage. Such tools ensure compliance with industry regulations and trust in AI outputs. HPE Private Cloud AI can also help thanks to its enterprise-grade controls for data privacy, security, transparency, and other governance needs.

4. **Focus on scalability.** AI demands evolve rapidly, so it's smart to start small so your private cloud environment can grow with your organization's needs. This reduces up-front costs while allowing you to scale efficiently as projects expand. The challenge comes when that growth really starts to ramp up.

   HPE GreenLake cloud helps organizations scale their AI infrastructure by offering on-demand access to high performance computing resources, AI software, and a unified control plane for seamless management across hybrid environments, regardless of size. It integrates on-premises and cloud platforms, helping ensure flexibility, scalability, and support for modern AI applications. This streamlined approach enables businesses to manage and fine-tune their AI operations over time. "Nothing can help you scale like HPE GreenLake," says Corrado. "As your AI use increases, we come in with hardware, software, and expertise that adjusts to your needs."

[6] "2024 Global AI Trends Report," WEKA.io, August 2024

[7] "Now decides next: Moving from potential to performance," Deloitte, August 2024

## Private cloud is a competitive imperative

The rise of AI is reshaping how businesses operate, but the infrastructure supporting it needs to keep pace. Private cloud offers the security, performance, and flexibility required to make AI a sustainable success. Most important, it enables organizations to adapt as technology evolves.

By embracing NVIDIA AI Computing by HPE turnkey solutions and following these four best practices, businesses can turn AI from a challenging experiment into a competitive advantage.

"Private cloud isn't just a better option for AI — it's a smarter one," Corrado says. "But it takes the right strategy and solutions to unlock its full potential."

## Learn more at

[HPE.com/ai](HPE.com/ai)

**Visit HPE.com**

**Chat now**

a50011918ENW, Rev. 2

HEWLETT PACKARD ENTERPRISE

hpe.com