HPE GreenLake

Accelerate your AI outcomes with HPE ProLiant Gen11 AI Inference Solutions

Speed time to value with the new AI inference Solutions. These systems are certified and optimized for inference at the edge and in the data center. The solutions are built on HPE's new ProLiant Gen11 servers and integrate NVIDIA® accelerated computing. HPE ProLiant Gen11 servers are designed from the ground up to support advanced accelerators.

HPE ProLiant Gen11 AI Inference Solutions

Computer vision AI	Generative visual AI	Natural language processing AI
Purpose-built for AI at the edge Loss prevention Smart spaces	**Optimized for visual apps** 3D animation Image/video generation	**Powering large language models** Speech AI Fraud detection

Computer vision AI at the edge

Analyzes video footage from cameras to detect and track objects, people, and behaviors for enhanced operations, security, and smart space experiences.
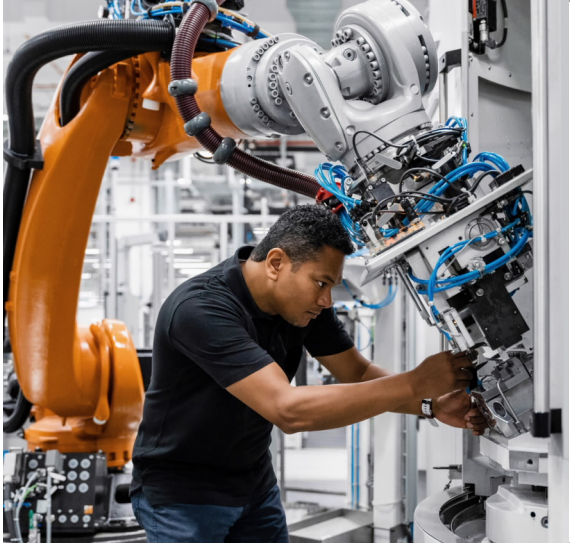
New!

HPE ProLiant DL320 Gen11
Up to four NVIDIA L4 GPUs

NVIDIA Metropolis

Purpose-built for AI at the edge

Generative visual AI

- Generate lifelike and dynamic 3D animations including character movements, physics simulations, and environmental effects.
- Curate new visual content, such as images or videos, that mimic real-world data.

New!

HPE ProLiant DL380a Gen11
Up to four NVIDIA L40 GPUs

NVIDIA AI Enterprise

Optimized for visual apps

Natural language processing AI

Develop and deploy end-to-end AI to power natural language processing for:
- Speech AI
- Fraud detection
- Predictive maintenance

New!

HPE ProLiant DL380a Gen11
Up to four NVIDIA L40S GPUs

NVIDIA AI Enterprise

Powering large language models

Top reasons to choose HPE ProLiant Gen11 for AI workloads

1
Innovate with more graphical capabilities than ever before
Empower your innovators with advanced GPU accelerators on HPE ProLiant Gen11 ultra-scalable architecture, with up to 33% more GPU density[1] and flexibility for AI workloads of all sizes at the edge and in the data center.

2
Expect more from your infrastructure
HPE ProLiant Gen11 compute delivers the efficiency, scalability, and economics to accelerate business outcomes while lowering TCO. Consolidate more workloads and increase ROI with breakthrough performance.

3
Control cost by rightsizing and scaling as needed
Deploy AI workloads on GPUs and high-performance clusters with predictable, transparent costs. You can quickly scale as needed with next-generation architecture using continuous monitoring. You can rightsize capacity and even enable capacity bursting on-site, on demand.

4
Get an intuitive cloud operating experience
Simplify the way you control compute from edge to cloud with a cloud operating experience. Transform business operations and pivot your team from reactive to proactive with global visibility and insight through a unified console with HPE GreenLake for Compute Ops Management. You can automate tasks for efficiency in deployment, simplified support, and lifecycle management.

5
Provide trusted security by design
Protect your infrastructure, workloads, and data from threats to hardware and risks from third-party software with a trusted edge-to-cloud security posture built on an HPE compute core hardened through a proven, zero trust approach to security.

[1] Compared to previous generation

Learn more at
HPE.com/ProLiant/AI

Visit **HPE GreenLake**

Chat now (sales)

Hewlett Packard Enterprise