Solution brief



HPE ProLiant DL525 Gen13 Server

Performance-intensive compute for simulation, modelling, AI inference, and virtualization



* Product configuration and appearance may vary

Figure 1. HPE ProLiant DL525 Gen13 Server

HPE ProLiant Gen13 fuels the next generation of AI-driven enterprises

AI is changing how enterprises create value from data, but the pressure on infrastructure extends well beyond AI alone. As organizations modernize applications, scale analytics, virtualize more workloads, and support increasingly data-intensive business operations, they need infrastructure that can deliver greater performance, flexibility, and efficiency without adding complexity.

As these workloads grow in scale and complexity, they are exposing the limits of traditional infrastructure. Organizations face increasing pressure to deliver more performance, protect sensitive data, control rising power and operating costs, and simplify operations with constrained resources.

The organizations that succeed will be those that grow AI and modernize infrastructure without scaling cost, complexity, risk, or power consumption. HPE ProLiant Gen13 helps enterprises scale AI and modernize infrastructure with trusted security, better economics, and smarter operations.

The new HPE ProLiant DL525 Gen13 Server is a purpose-built secure, single-socket 1U platform ideal for AI inferencing, simulation, modeling, virtualization, and enterprise workloads, that delivers up to 2x greater compute performance and improved efficiency by combining exceptional core density, memory bandwidth, and next-generation I/O.

HPE ProLiant DL525 Gen13 Server is a 1U platform with AMD EPYC™ 9006 SP7 Server CPU with up to 256 cores, PCIe Gen6 IO, 16-channel DDR5 memory, and MRDIMM support up to 12800 MT/s. It combines high compute performance with expanded memory and I/O resources, enabling organizations to support larger datasets, faster data movement, and more demanding workloads within a compact footprint.



Key innovations

Exceptional core scale in a single-socket platform

Powered by the server CPU with up to 256 cores, the DL525 Gen13 delivers high compute performance and leadership density in a compact 1U design. This exceptional core density helps organizations consolidate more workloads per server, accelerate parallel processing, increase virtualization density, and improve compute economics while reducing infrastructure sprawl.

PCIe Gen6 connectivity

PCIe Gen6 I/O provides a next-generation foundation for faster data movement across networking, storage, GPUs, DPUs, and accelerators. By doubling bandwidth compared to PCIe Gen5, the DL525 Gen13 helps organizations support larger datasets, reduce bottlenecks, and adapt as AI, analytics, and data-intensive workload requirements evolve.

Next-generation memory performance

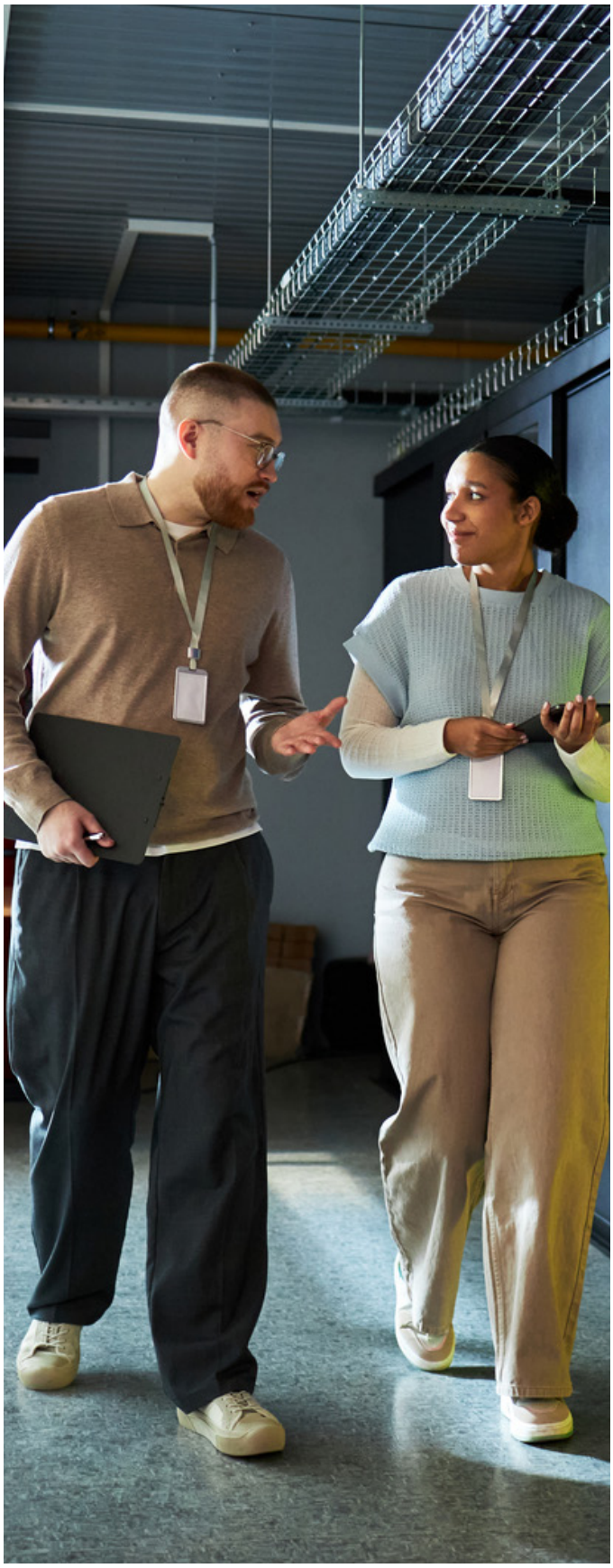
With 16-channel DDR5 memory and MRDIMM support up to 12,800 MT/s, the DL525 Gen13 helps keep processors fed with the bandwidth required for data-intensive workloads. Expanded memory performance enables organizations to support data intensive workloads such as electronic design automation (EDA), AI inference, and virtualization.

More compute from every watt

Optimized thermal design enables advanced air cooling for high-performance processors and accelerators while optional direct liquid cooling supports higher-density configurations. This helps organizations increase performance per watt, reduce operating costs, and scale demanding workloads more sustainably.

Purpose-built for demanding AI and enterprise workloads

The DL525 Gen13 combines high core scale, expanded memory bandwidth, PCIe Gen6 I/O, and flexible accelerator support to accelerate demanding AI inference, technical computing, modeling, simulation, analytics, and virtualization workloads. Its secure, efficient 1U design helps organizations deliver more performance and capacity while simplifying infrastructure and reducing total cost.



Powering the workloads shaping tomorrow's enterprise

For enterprises running demanding analytics, simulation, virtualization, AI inference, and technical computing workloads, performance is increasingly defined by how efficiently infrastructure can combine compute, memory, and I/O at scale. The HPE ProLiant DL525 Gen13 Server brings these capabilities together in a secure, balanced 1U platform that helps organizations accelerate larger datasets, move data faster, improve workload density, and solve complex problems with greater efficiency.



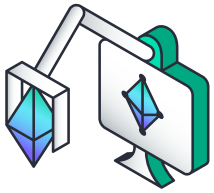
AI inference

Support CPU- and accelerator-enabled AI inference workloads that require high throughput, fast memory access, and flexible I/O. The DL525 Gen13 provides the compute density, memory bandwidth, and PCIe Gen6 expansion needed to run AI inference alongside analytics, simulation, and other data-intensive applications.



Virtualization modernization

Consolidate more virtual machines and containerized applications onto fewer servers while improving resource utilization and reducing software licensing exposure. With up to 256 cores, large memory capacity, and high-bandwidth I/O, the DL525 Gen13 helps organizations improve workload density, modernize infrastructure, and optimize cost per core and watt per core.



Engineering simulation and digital twins

Support complex product design, digital twin development, computational fluid dynamics, finite element analysis, and manufacturing simulations. These workloads require a balanced combination of compute power, memory performance, and storage throughput to process increasingly complex models and accelerate innovation cycles.



Electronic design automation

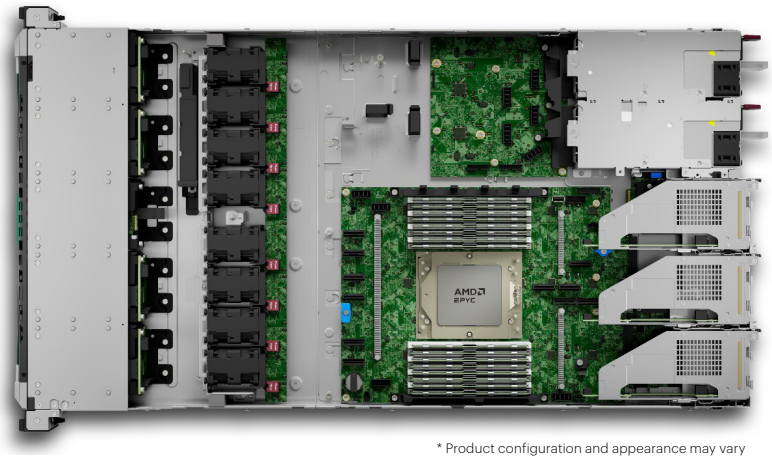
Reduce design verification and validation times for semiconductor, electronics, and embedded systems development. EDA workloads often involve highly parallel processing combined with large memory footprints. The DL525 Gen13 high core density, low-latency memory architecture, and PCIe Gen6 expansion provide the performance foundation needed to accelerate circuit verification, modeling, and emulation workloads.



Data analytics

Process larger datasets and generate faster business insights from analytics platforms, data lakes, and real-time processing environments. The combination of compute density, memory bandwidth, and fast storage connectivity helps eliminate bottlenecks that can delay analytics workloads and decision-making.

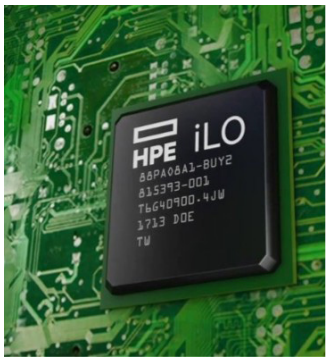
Technical specifications



* Product configuration and appearance may vary

Figure 2. Internal view of HPE ProLiant DL525 Gen13 Server

Processors	AMD EPYC 9006 SP7 Server CPUs, up to 256 cores 600W
Memory	16 channel per socket, 1DPC design, up to 4 TB DDR5 8000 MT/s; 2 TB MRDIMM 12800 MT/s
Storage	Up to 20 EDSFF E3.S 1T SSD; NVMe Up to 1 M.2 2280 SSD; NVMe
Boot	Up to 1 NS204mff boot device (front or rear)
GPU	Up to 2 DW or 4 SW FHFL at the front; up to 3 SW HHHL at the rear
I/O—front	Up to 4 PCIe Gen6 slots Up to 1 OCP 3.0 slot
I/O—rear	Up to 3 PCIe Gen6 slots Up to 1 OCP 3.0 slot
Storage Controller	PCIe and OCP options
Power	Up to 2 M-CRPS 60 or 73mm PSUs, up to 3200W
Cooling	Air, DLC
Management	HPE iLO 8 through DC-SCM, HPE Compute Ops Management, HPE OneView



* Product configuration and appearance may vary

Figure 3. HPE iLO 8

The foundation of self-securing infrastructure

As organizations accelerate AI adoption and modernize infrastructure, security challenges continue to grow. Enterprises must protect sensitive data, verify infrastructure integrity, defend against evolving cyber threats, and prepare for emerging post-quantum risks. Yet many security controls still depend on manual configuration, administrative action, and ongoing oversight. When protection depends on people doing every step right, operational gaps emerge. HPE ProLiant DL525 Gen13 Server with HPE iLO 8 helps establish the foundation for self-securing infrastructure through automatic protection, continuously verified trust, and security designed for future threats.

The HPE ProLiant DL525 Gen13 Server with HPE iLO 8 helps establish a stronger security foundation by building protection directly into the platform. From first boot through ongoing operations, the platform continuously helps verify trust, protect data, and strengthen governance without adding complexity or operational burden.

This foundation is further strengthened by the security innovations built directly into AMD EPYC processors. Built-in at the silicon level, AMD Infinity Guard offers the advanced capabilities required to defend against internal and external threats to keep your data safe.[1]

Key innovations of HPE iLO 8

Protect data automatically from first boot

Auto-SED helps protect data automatically from first boot by enabling self-encrypting drives without requiring manual configuration, external key managers, or administrative intervention. This shifts encryption from a deployment-dependent task to built-in, platform-enforced protection.

Strengthen governance for critical operations

Multi-party authorization introduces platform-enforced approval controls for high-impact administrative actions, helping reduce risk from single-user mistakes, credential misuse, unauthorized actions, and insider threats.

Prepare for tomorrow's threats today

Expanded post-quantum cryptography protection extends trust beyond firmware and into the communications and interactions that keep infrastructure running, helping organizations prepare for future cryptographic threats while strengthening trust across the platform.

Verify what you trust

Expanded component attestation helps continuously verify the authenticity and integrity of critical system components throughout the server lifecycle, strengthening confidence that trusted infrastructure remains built on trusted hardware.

[1] AMD Infinity Guard features vary by EPYC Processor generations and/or series. Infinity Guard security features must be enabled by server OEMs and/or Cloud Service Providers to operate. Check with your OEM or provider to confirm support of these features. Learn more about Infinity Guard at amd.com/en/products/processors/server/epyc/infinity-guard.html.

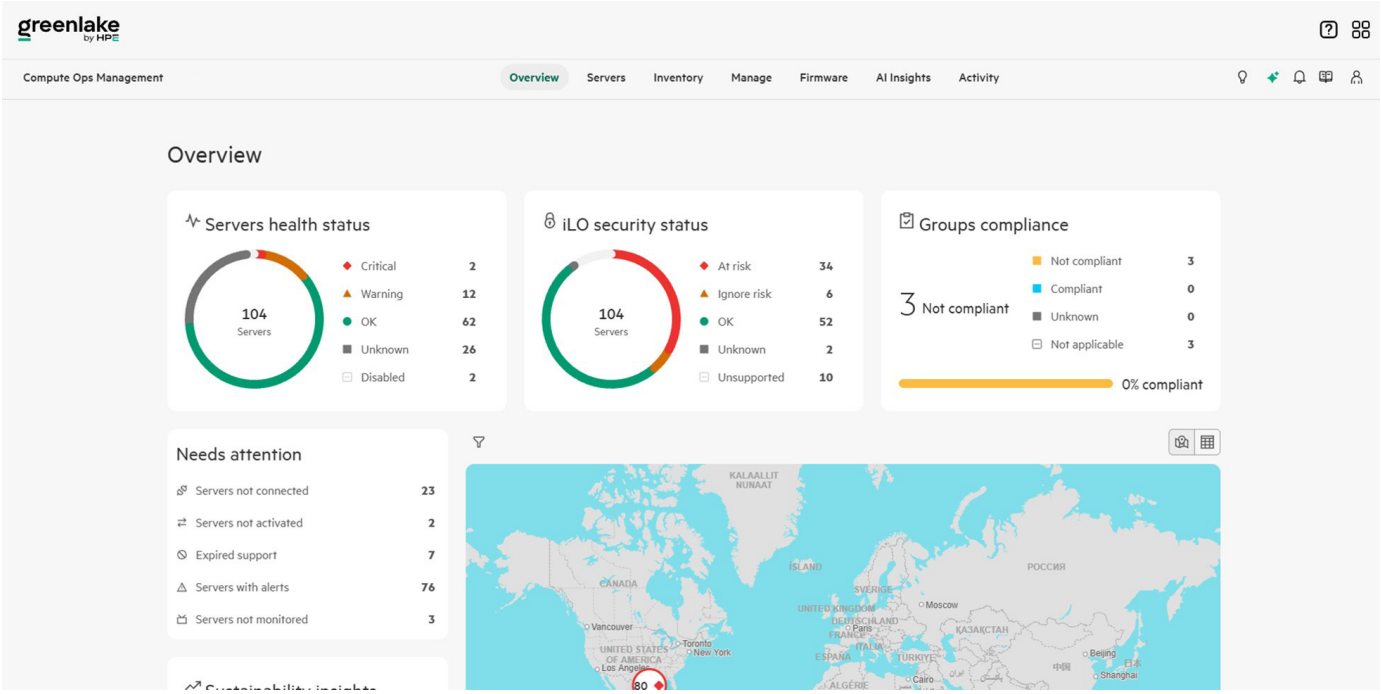


Figure 4. HPE Compute Ops Management

Advancing the next era of AI-driven server operations with HPE Compute Ops Management

As server environments scale and agentic AI and data-intensive workloads increase, manual infrastructure management cannot keep pace with the speed and complexity required. Manual processes, fragmented operational visibility, and disconnected workflows make it difficult to proactively manage infrastructure, efficiently respond to issues, and accelerate decision-making.

HPE Compute Ops Management helps advance AI-driven server operations by delivering intelligent automation, simplified lifecycle management, and visibility that drives action and decision-making while building a foundation designed for increasingly autonomous operations. The result is less time spent managing servers and more time focused on innovation and business outcomes.

HPE Compute Ops Management makes server operations simpler and smarter

Simplify infrastructure management at scale

Manage and monitor servers through a unified experience that provides visibility, actionable insights, and control across distributed environments.

Reduce manual operational effort

Simplified server onboarding, lifecycle management, and streamlined update workflows help reduce repetitive administrative tasks and improve operational consistency.

Transform insights into action

AI-driven advisory intelligence analyzes advisories against the server environment and evolves remediation from manual review to automated workflows before issues become disruptions.

Make informed power decisions

Granular power insights provide more timely visibility into power capacity, helping IT teams plan with greater confidence and make better-informed power and capacity decisions.

Accelerate support and issue resolution

Integrated support experiences help streamline issue resolution, helping organizations reduce operational risk and accelerate remediation.

Increase value with HPE Tech Care Service

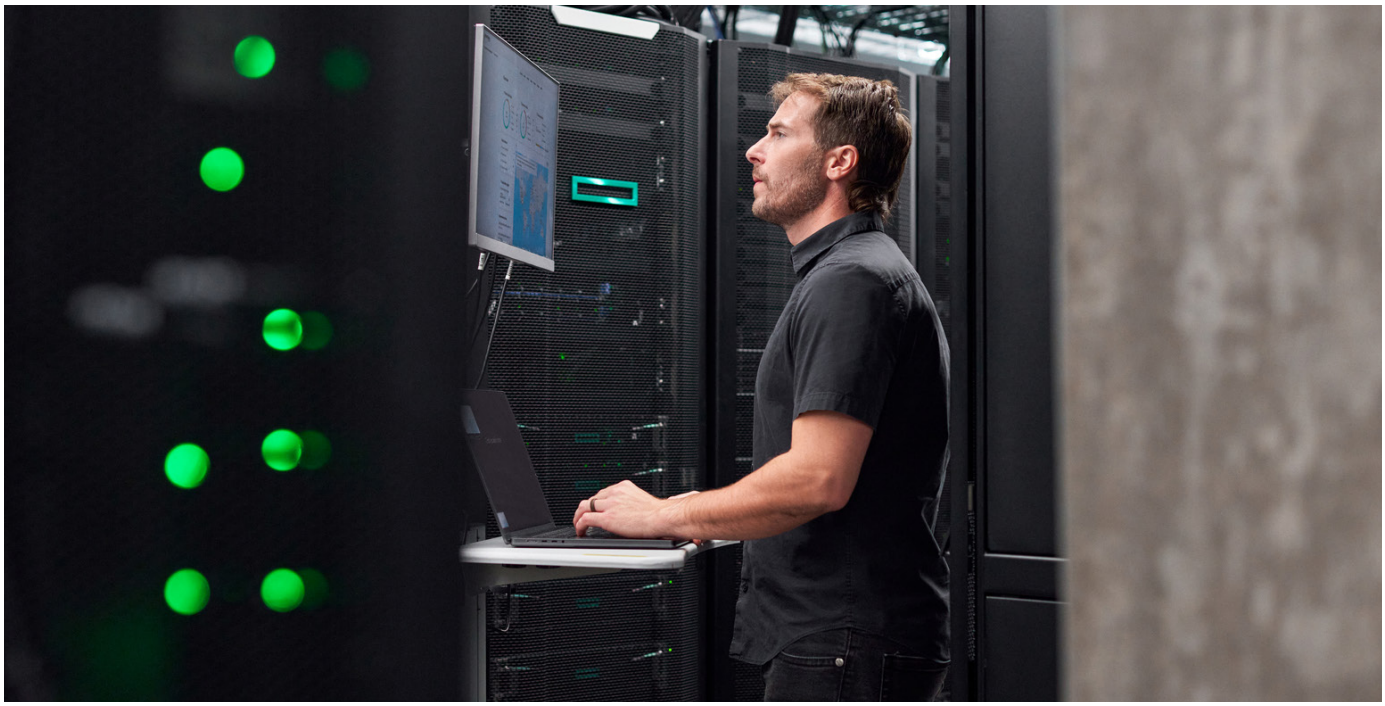
Modern infrastructure requires more than reactive support. As environments become more complex and workloads are more critical to business success, organizations need expert guidance, personalized insights, and trusted support experiences that help maximize infrastructure performance and availability throughout its lifecycle. HPE Tech Care Service helps organizations reduce operational risk, accelerate problem resolution, and optimize infrastructure health so IT teams can spend less time troubleshooting and more time driving innovation.

Purpose-built for modern enterprise demands

Enterprise infrastructure is entering a new era. Organizations must support AI inference, analytics, simulation, virtualization, and business-critical workloads simultaneously while controlling cost, risk, power consumption, and complexity.

For enterprises running performance-intensive workloads, the HPE ProLiant DL525 Gen13 Server is purpose-built to deliver greater performance and improved efficiency in a secure, single-socket platform. With AMD EPYC 9006 SP7 processor, up to 256 cores, PCIe Gen6 I/O, 16-channel DDR5 memory, MRDIMM support up to 12,800 MT/s, flexible accelerator options, and advanced thermal choices, it enables organizations to support larger datasets, faster data movement, and more demanding workloads within a compact footprint.

More performance. More efficiency. More innovation from every infrastructure investment.



Visit HPE.com

Learn more at

HPE ProLiant DL525 Gen13 Server

HPE ProLiant Compute powered by AMD EPYC Processors

Chat now

© Copyright 2026 Hewlett Packard Enterprise Development LP. The information contained herein is subject to change without notice. The only warranties Hewlett Packard Enterprise products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. Hewlett Packard Enterprise shall not be liable for technical or editorial errors or omissions contained herein.

AMD is a trademark of Advanced Micro Devices, Inc. All third-party marks are property of their respective owners.

a00162374ENW

HEWLETT PACKARD ENTERPRISE

hpe.com