HPE

Brochure

AI ON A BUDGET: unlocking AI for SMBs

Unlocking AI for SMBs

Artificial intelligence (AI) is no longer confined to massive datacenters or organizations with substantial budgets. As AI continues to expand into new areas, businesses are discovering innovative ways to harness its power without the need for complex infrastructure. For small and medium-sized businesses (SMBs), AI has become a strategic advantage—a tool that drives innovation, improves productivity, and provides a competitive edge through cost-effective solutions. IDC predicts that by 2027, half of all SMBs will significantly shift their IT budgets to prioritize AI, as technology advances, vendors refine pricing, and AI becomes essential for staying competitive.*

Imagine SMBs unlocking the full potential of AI to accelerate decision-making, enhance creativity, and boost efficiency without breaking the bank. With the HPE ProLiant ML110 Gen11 and NVIDIA® L4 GPU Generation GPU, this vision becomes reality. These technologies make cutting-edge AI accessible and practical for businesses of all sizes.

Built with flexibility, scalability, and performance in mind, this platform is tailored to support AI workloads at the edge or within compact IT environments. From data analytics and machine learning to advanced graphics and content creation, AI is transforming how businesses operate. The HPE ProLiant ML110 Gen11 and NVIDIA® L4 GPU empower SMBs to deploy impactful AI solutions at a fraction of the cost typically associated with high-end systems.

The future of AI is here, and it's within reach for businesses of all sizes.

*IDC FutureScape: Worldwide Small and Medium-Sized Business 2025 Predictions

Solution components

HPE ProLiant ML110 Gen11

Figure 1. HPE ProLiant ML110 Gen11 Server

The HPE ProLiant ML110 Gen11 is a 4.5U tower server engineered to meet the demanding needs of SMB workloads, providing a balance between performance, storage and expansion. The tower form factor simplifies deployment without requiring dedicated rack space, making it ideal for offices or edge locations with limited IT infrastructure.

Powered by 4th/5th Generation Intel® Xeon® Scalable Processors, supporting up to 1.5 TB of memory capacity, and flexible storage options including multiple SAS/SATA drive options and adequate PCIe Gen5 expansion slots. HPE ProLiant ML110 Gen11 delivers exceptional compute capability and memory bandwidth for both compute and data workloads.

With the multiple expansion slots available, HPE ProLiant ML110 Gen11 can accommodate up to two NVIDIA® L4 GPU, unlocking accelerated AI inferencing, machine learning, and advanced visualization capabilities, enabling SMBs to deploy cutting-edge AI solutions on-premises.

Additionally, integrated HPE management tools such as HPE iLO provide robust remote monitoring and control, ensuring reliability and ease of administration. The HPE ProLiant ML110 Gen11 offers SMBs a cost-effective and versatile platform to drive innovation and maintain a competitive advantage.

NVIDIA® L4 24GB PCIe GPU

Figure 2. NVIDIA® L4 GPU

The NVIDIA® L4 24GB PCIe GPU is a versatile, energy-efficient accelerator designed for modern AI inference, video processing and visualization workloads in edge environments. It delivers exceptional performance for tasks such as AI inference, intelligent video analytics, cloud graphics, and virtual desktop acceleration.

Key specifications include:

— 24 GB GDDR6 GPU ECC Memory for reliable processing

— 7680 CUDA Cores for parallel processing, accelerating tasks like video processing, AI and graphics

— 233 4th Gen Tensor Cores for AI acceleration, enabling faster deep learning training and Inference

— 192-bit Memory Interface for efficient data transfer between the GPU and memory

— 432 GB/s Memory Bandwidth for smooth handling of large files and demanding applications

— 75W Max Power Consumption, delivering excellent performance-per-watt efficiency

— Low profile, single-slot form factor design for compatibility with a wide range of servers

Supported on HPE ProLiant ML110 Gen11 servers, the NVIDIA® L4 GPU's full specifications and HPE compatibility details are available in the HPE QuickSpecs for NVIDIA Accelerator.

HPE Smart Choice offers

Finding the right solution can often feel complicated and overwhelming, especially when navigating lengthy processes or choosing the ideal configuration. That's where HPE Smart Choice offers come in. They simplify the journey by offering pre-configured HPE systems available through authorized resellers. This streamlined approach provides SMBs with a fast and hassle-free way to purchase and deploy IT infrastructure, empowering businesses to focus on what matters most.

Benefits of HPE Smart Choice offers include:

— **Streamlined process.** From purchase to deployment, every step is simple and efficient.

— **Pre-configured solutions.** The program offers a selection of HPE's most popular and best-selling HPE ProLiant servers that are fully configured and ready to ship.

— **Simplified ordering.** Order these pre-configured solutions using a single SKU, making the purchasing process faster and easier.

— **Fast shipping.** HPE Smart Choice products are prioritized within the supply chain to ensure quick and predictable delivery, often in days rather than weeks.

— **Competitive pricing.** The program is designed to offer attractive pricing.

— **Easy deployment.** Because the products are pre-configured, they are ready for immediate deployment, simplifying on-site integration and saving time.

— **Growth enablement.** This offering helps businesses grow by providing enterprise-grade performance and flexibility at an affordable price, with solutions that can scale with their needs.

Learn more on the HPE Smart Choice offers product page.

Software components

Management

HPE offers two powerful solutions to simplify and secure server management—from individual systems to large-scale environments.

HPE Integrated Lights Out (HPE iLO) has long been the cornerstone of HPE system management. Integrated across all HPE servers, HPE iLO delivers out-of-band management capabilities and forms the foundation of HPE's server security. Learn more in the HPE Integrated Lights out Product page.

Building on that foundation, HPE Compute Ops Management takes system oversight to the cloud. This platform streamlines the entire server lifecycle, providing a centralized, unified view of distributed environments to improve control, visibility, and efficiency. Explore more on the HPE Compute Ops Management product page.

Figure 3. HPE Integrated Lights Out (HPE iLO)

Together, HPE iLO and HPE Compute Ops Management give you the flexibility to manage your infrastructure—your way, whether on-premises or from the cloud.

Security

HPE has a long-standing reputation for delivering secure, trusted infrastructure—a legacy that continues with the HPE ProLiant ML110 Gen11. Built with a robust set of security features, these systems give businesses the confidence to operate and scale securely from day one.

At the heart of this protection is HPE Integrated Lights Out 6 (HPE iLO 6), which provides advanced security features such as silicon root of trust, firmware verification, and secure recovery. Together, these capabilities safeguard your server against firmware-level attacks and ensure that only validated code is executed. Learn more in the HPE iLO 6 Security Technology Brief.

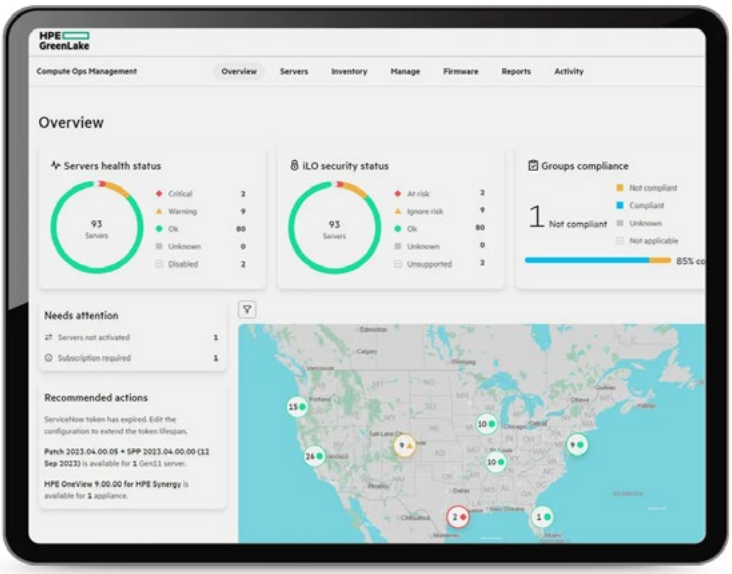


Figure 4. HPE Compute Ops Management

Complementing this, the Unified Extensible Firmware Interface (UEFI) enhances system startup and configuration security. UEFI provides a more modern, flexible BIOS alternative with secure boot capabilities that help protect against malware and unauthorized firmware changes. Explore the HPE UEFI User Guide for Gen11.

Together, these integrated features strengthen your security posture—from the hardware foundation up—ensuring every server starts, runs, and scales with confidence.

Operating systems

A wide range of server operating systems have been certified for use with the HPE ProLiant ML110 Gen11, ensuring flexibility and compatibility for diverse IT environments. Supported operating systems include Microsoft Windows Server, Red Hat® Enterprise Linux® (RHEL), SUSE Linux Enterprise Server (SLES), VMware ESXi™, and Canonical Ubuntu, providing businesses with the flexibility to choose the platform that best suits their workloads and operational requirements. Visit the HPE Servers Support and Certification Matrices site at HPE Servers Support & Certification Matrices to ensure your desired OS version is supported.

Workloads

HPE ProLiant ML110 Gen11 is a versatile platform for general-purpose and emerging AI workloads, providing solid foundation for growing businesses, when partnered with an AI independent software vendor. As AI becomes essential to modern operations, equipping HPE ProLiant ML110 Gen11 with a GPU accelerator unlocks new capabilities, enabling organizations to enhance traditional workloads with AI-driven automation and insights to boost efficiency, productivity, and innovation.

For entry level AI workload—such as AI Inference, computer vision, predictive maintenance, natural language processing—the HPE ProLiant ML110 Gen11 combined with the NVIDIA® L4 GPU provides an ideal entry point. A recommended configuration features 12-core CPU, 64 GB of system memory and approximately 2 TB of storage, paired with a 10GbE network interface. This serves as a strong starting point for most AI workloads, allowing organizations to scale GPU, memory or storage resources as data volumes or model complexity grows.

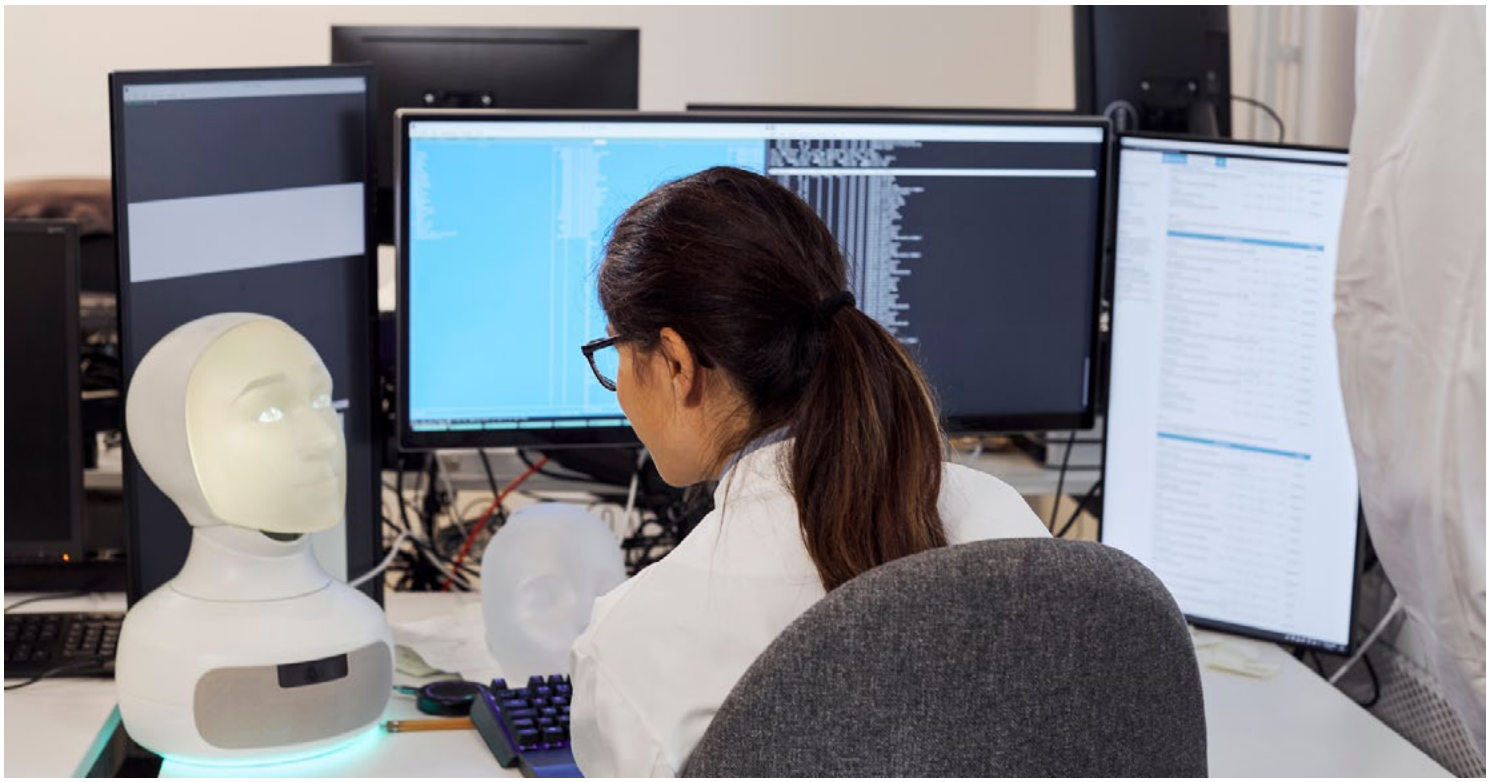
This configuration strikes the optimal balance of performance, scalability, and efficiency—helping organizations accelerate AI adoption while maintaining operational simplicity and cost control. It features GPU ECC memory for reliable handling of large datasets, along with CUDA and Tensor cores that deliver parallel processing and AI acceleration for faster training and inference. All of this runs on a PCIe 4.0 bus, ensuring high-speed data transfers throughout the system.

Data analytics at the edge

HPE ProLiant ML110 Gen11 combined with NVIDIA® L4 GPU enables fast, secure, and scalable data analytics directly where data is generated. With GPU-accelerated parallel processing and ECC memory for reliability, insights can be derived locally, reducing latency, minimizing cloud dependency, and maintaining full control over sensitive business data.

Edge analytics empowers organizations to act on information in real-time rather than waiting for cloud-based aggregation. A small retailer can use the HPE ProLiant ML110 Gen11 to process point-of-sales (POS) data locally, identifying customer purchasing trends, peak hours, and location-based buying patterns to refine promotions and optimize inventory. In industrial or logistics environments, the same confirmation can process high-frequency sensor data to detect anomalies, predict equipment issues, or optimize production efficiency.

The NVIDIA® L4 GPU offers exceptional power efficiency and performance density, making it ideal for continuous, multi-site analytics at the edge, where space and energy resources are limited.



Machine learning and predictive analysis at the edge

In the large arena of AI, one of the most common uses is machine learning (ML). ML is the process by which a computer system learns from data collection. Through ML, a computer can learn to identify patterns in data and images over a period of time. This type of learning can then be used, through Inference, for customer personalization allowing the business to become more familiar with customers buying preferences and trends.

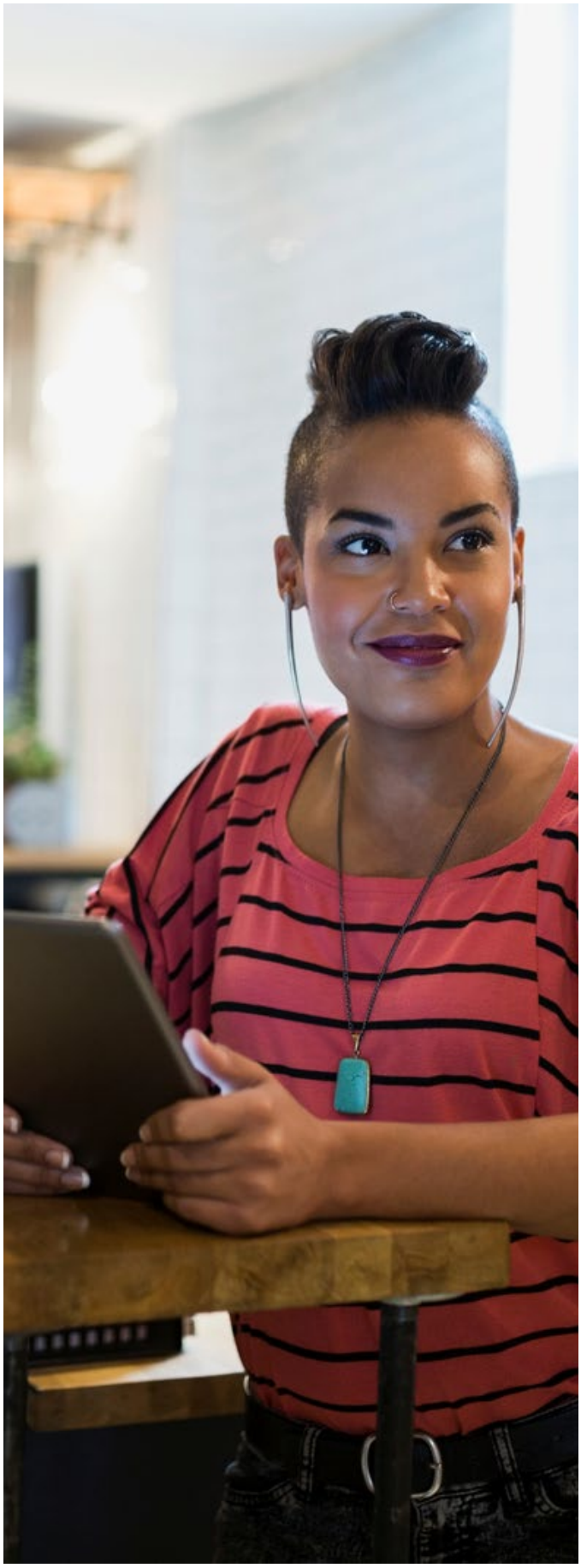
ML and Inference may seem like a very complicated and expensive undertaking reserved only for the largest of enterprises; on the contrary, with the HPE ProLiant ML110 Gen11 and the NVIDIA® L4GPU, ML and Inference have arrived at the edge, learning at the edge where the data originates. Let's explore three use cases where this solution can be deployed.

Personalization, looking back at our example of the retail business using the solution to run their business and personalize the interactions with customers is a prime example of ML and Inference at the edge. In the example, the data is collected and analyzed, providing analyses to help the manager build targeted sales campaigns by uncovering customer trends.

Another use at the edge is to analyze video. Integration of cameras into this solution can help businesses with tasks like data collection for analysis or surveillance. Cameras and off-the-shelf software applications can be integrated into this solution to provide several additional capabilities like video analytics, video management systems, retail video analytics, and industrial video applications. In addition to this HPE's wide range of resellers and partners may provide the required integration services of video software onto the platform.

For more complex environments or heavier workloads, for example, in a large industrial environment, predictive maintenance is a transformative approach, enabling companies to anticipate equipment failures before they occur and significantly reduce downtime and maintenance costs. In complex facilities such as automotive plants, chemical refineries, and heavy machinery operations, equipment generates massive amounts of sensor data—vibration, temperature, pressure, and acoustic signals—that require powerful processing and analytics at the edge.

By deploying predictive maintenance at the edge, large industrial operators can improve equipment reliability, increase operational efficiency, enhance safety, and ultimately reduce costs—turning raw sensor data into actionable intelligence in near real time.



AI-driven video analytics and computer vision

HPE ProLiant ML110 Gen11 combined with the NVIDIA® L4 GPU delivers a versatile edge platform for processing, enhancing and interpreting video data in real time. This solution supports a full range of video-centric AI workloads -from super-resolution and denoising to event detection, object recognition and scene understanding at premises.

In retail, it can improve visual clarity in camera feeds while simultaneously analyzing customer flow or shelf activity. In security environments, it enhances and interprets surveillance footage to detect unusual behavior or recognize license plates. In industrial settings, it monitors production lines for quality issues or safety hazards.

By processing video locally, organizations reduce latency, bandwidth consumption, and cloud dependency while maintaining data privacy. The NVIDIA® L4 GPU accelerates decoding, AI inference, and transcoding in parallel, providing continuous, power-efficient performance. Together with the HPE ProLiant ML110 Gen11 scalability this configuration enables intelligent, real-time video analytics across diverse edge deployments from smart stores to secure facilities.

Language intelligence at the edge: natural language processing (NLP)

Natural language processing (NLP) and computer vision are two of the most transformative domains in AI, enabling businesses to extract actionable insights from images, videos, and text.

The HPE ProLiant ML110 Gen11 and NVIDIA® L4 GPU, enables SMBs to leverage AI-driven solutions for a wide range of language-centric tasks such as customer sentiment analysis, which helps businesses quickly understand and respond to customer feedback from reviews, social media, and surveys. Document classification automates the organization and tagging of large volumes of emails, reports, and support tickets, significantly reducing manual workload and improving operational efficiency. Intelligent chatbots can enhance customer service by delivering natural, context-aware conversations, freeing up human agents to focus on more complex issues. Automated summarization of long-form documents allows businesses to rapidly extract key insights from reports, contracts, or research materials, saving valuable time.

For SMBs, deploying NLP can up level business efficiencies, enabling real-time inference, unlocking insights from unstructured text data, personalize customer interactions at scale, and automate repetitive tasks, supporting dynamic, data-driven decision-making.

AI workloads can be processed locally at the edge—reducing latency and maintaining full control over both data and models. This setup supports data privacy by ensuring sensitive information does not leave the local environment. HPE ProLiant ML110 Gen11's exceptional expandability also allows for the addition of storage drives and PCIe-based components, including both the NVIDIA® GPU and HPE Storage controller for hardware RAID. This ensures not only powerful AI performance but also data redundancy and reliability through robust storage configurations.

Sensor fusion and edge data correlation

Sensor fusion and edge data correlation enable organizations to combine and analyze data from multiple sources—such as environmental sensors, IoT devices, and operational systems directly at premises. On platforms like the HPE ProLiant ML110 Gen11 with NVIDIA® L4 GPU, multiple data streams could be processed in parallel to generate actionable insights in real time.

For example, temperature, humidity, and energy consumption readings from a building or manufacturing facility could be correlated to detect anomalies, optimize resources usage or support predictive maintenance. Performing this analysis locally reduces latency, increases responsiveness and ensures sensitive operational data remains secure on-site.

The NVIDIA® L4 GPU provides local processing of diverse sensor data, allowing organizations to analyze and correlate information in real time while keeping power consumption low.

Summary

This solution brief demonstrates how small and medium-sized businesses can deploy powerful, cost-effective AI solutions using HPE ProLiant ML110 Gen11 paired with the NVIDIA® L4 GPU. It emphasizes that AI is no longer exclusive to large enterprises with deep pockets—SMBs can now harness AI at the edge for a variety of workloads including data analytics, machine learning, computer vision, natural language processing, video content creation, scientific computing, and medical imaging.

The document outlines:

Hardware components. Detailed specs and use cases for HPE ProLiant ML110 servers, and the NVIDIA® L4 GPU.

Software and management tools. Including HPE HPE iLO and HPE Compute Ops Management for secure and simplified server administration.

AI workloads at the edge. Real-world examples such as retail analytics, predictive maintenance, and video enhancements.

Deployment guidance. Encourages SMBs to work with HPE resellers to tailor solutions to their unique needs.

HPE Smart Choice offers. A streamlined purchasing and deployment model for pre-configured systems.

Ultimately, the brief positions HPE's edge-ready solutions as a gateway for SMBs to enter the AI space affordably, without sacrificing performance or scalability.

Learn more or purchase your products at the HPE Smart Choice offers product page.

Figure 5. HPE ProLiant ML110 Gen11 Server and a NVIDIA® L4 GPU

QuickSpecs and product information

— HPE ProLiant ML110 Gen11 QuickSpecs link

— NVIDIA Accelerators for HPE QuickSpecs link

Visit HPE.com

Chat now

a00155883enw, Rev. 2

HEWLETT PACKARD ENTERPRISE

hpe.com