

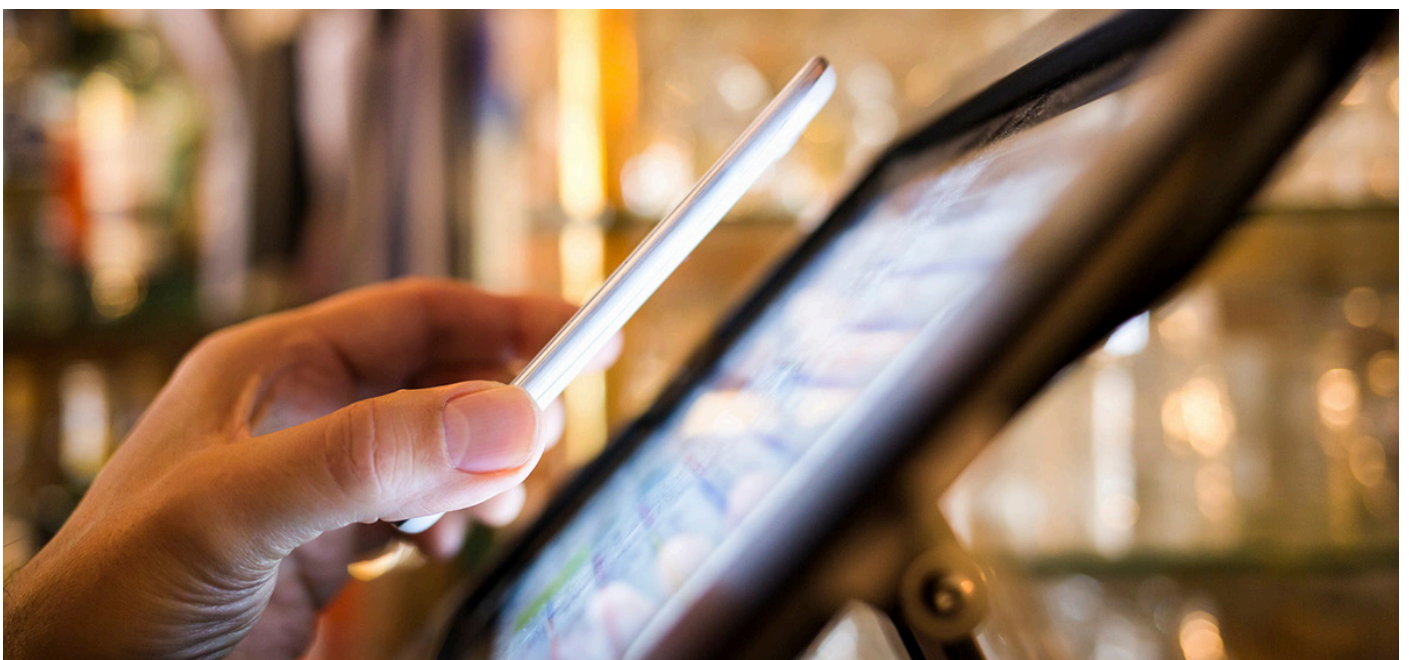


REINVENTING RETAIL WITH GROUNDBREAKING AI INNOVATION

Smart retail solutions to accelerate intelligent stores

CONTENTS

Enhancing retail operations with cutting-edge AI-driven, Intelligent Video Analytics.....	2
Accelerating retail insights with the right architecture	2
Smart retail at the edge.....	3
HPE Edgeline Converged Edge Systems.....	3
HPE ProLiant systems for a modern and optimized IT environment.....	4
NVIDIA EGX Platform for Edge Computing.....	4
Innovative solutions enabling a new wave of smart retail.....	5
Accelerating retail insights with core infrastructure.....	5
Harnessing the full power of AI capabilities.....	5
Expanding retail AI applications.....	6
Conclusion.....	6



In the retail industry, the escalating demand for highly automated environments is driving significant changes to in-store shopping. More and more retailers are leveraging the unmatched speed and precision of [artificial intelligence \(AI\)](#) to modernize in-store operations, logistics, and experiences.

With decreasing margins and inefficient processes on top of ever-increasing customer expectations, selecting and deploying the right technology is vital to success. Today's retailers face a host of challenges that can make or break their operations—these include reducing shrinkage, lowering operating costs and risk, extracting business insights, and ensuring regulatory compliance. In the area of loss prevention alone, losses due to shrinkage—the loss of inventory due to theft, fraud, errors, and damage—were estimated to reach [\\$77 billion in the U.S. & Europe in 2019](#).

The retail industry is undergoing a massive paradigm shift, leveraging a new breed of AI-powered applications to mitigate many of these issues. One area, [Intelligent Video Analytics \(IVA\)](#), is a game-changing tool that extracts rapid insights from video data streams with extreme accuracy. IVA applications use AI to reduce loss, combat shrinkage, secure customer safety, and enhance the customer experience. By taking advantage of cameras and sensors, retailers can shift from traditional technologies to disruptive technologies such as IVA.

But many retailers lack the essential infrastructure and systems to support numerous video feeds as well as the intensity of processing and analytics to deliver cost-effective, reliable outcomes. Furthermore, these companies lack in-depth knowledge and expertise to determine the optimal technologies and tools required to accelerate their applications. Without these capabilities, retailers will struggle to successfully deploy and operationalize their infrastructure and miss out on the vast potential of AI-powered IVA.

The bottom line: Retailers must innovate. Now, retailers need access to advanced resources, including state-of-the-art infrastructure, the right ecosystem of software, and expert guidance. Without the proper strategy to innovate quickly and effectively, companies will lose out on the new wave of smart retail.

ENHANCING RETAIL OPERATIONS WITH CUTTING-EDGE AI-DRIVEN, INTELLIGENT VIDEO ANALYTICS

On average, a typical retailer [loses 1.4–2% of its sales](#) to shrinkage. For mass merchants, this can add up to billions of dollars. However, the ongoing development of smart retail solutions leveraging artificial intelligence gives smart companies a critical advantage by delivering vital store and customer data from the edge to tackle the industry's most pressing concerns. According to a [recent estimate](#) by Accenture, AI has the potential to create \$2.2 trillion worth of value in retail and wholesale sectors by 2035 by driving growth and profitability.

To reach this staggering number, innovative retailers and disruptive startups use Intelligent Video Analytics (IVA) to obtain precise insights that will reduce shrinkage and boost profit margins. Intelligent stores have proven to increase cost-performance and efficiency in various processes, gaining visibility into in-store operations to streamline time-consuming and labor-intensive tasks, eliminating stockout, and optimizing merchandising.

Even in today's challenging environment, retailers are rushing to invest time, finances, and technology resources to leverage AI and IVA. However, there are several considerations to make while preparing your infrastructure for AI-powered IVA:

- What hardware and software are best suited for your IVA workloads and geographic location?
- What IVA applications provide the best fit for your infrastructure and operations?
- Can the system be tailored for the precise number and quality of video feeds as well as the intensity of processing and analytics needed?

Retailers must first answer these integral questions to implement a successful in-store retail.

ACCELERATING RETAIL INSIGHTS WITH THE RIGHT ARCHITECTURE

HPE and NVIDIA® are working closely with innovative independent software vendors (ISVs), bringing [cutting-edge AI](#) technology to retailers, the innovation that promises to reinvent the retail industry. Our combined hardware and software capabilities create highly advanced solutions that couple the performance, security, and reliability of HPE servers with the [NVIDIA EGX software stack](#) for peak speed and flexibility. With systems designed for edge environments our customers can leverage a range of applications to bolster loss prevention and significantly reduce shrinkage, while driving operational efficiencies to increase revenue.

HPE's retail solution is built on the industry-leading HPE ProLiant DL380 Gen10 servers and HPE Edgeline EL4-000 Converged Edge Systems for edge processing and HPE Apollo Gen10 servers for core data center processing. All three platforms deliver the highest levels of performance, expandability, reliability, security, serviceability, and availability in a dense form factor.

- For diverse workloads and environments—both at the edge and in the core data center—the secure HPE ProLiant DL380 Gen10 delivers world-class performance with the right balance of expandability and scalability.



- The HPE Edgeline family are purpose-built, compact, energy-efficient, ruggedized platforms for converging real-time data acquisition such as Intelligent Video Analytics—at the edge.
- The HPE Apollo 6500 Gen10 system is a platform specially engineered for machine learning and provides unprecedented performance.

These systems are tested in HPE’s global IoT and AI Labs to guarantee compatibility and functionality. Additionally, the chassis type, memory, and storage along with the number of servers, and NVIDIA T4 GPU or NVIDIA V100 Tensor Core GPU accelerators can all be tailored to fit the specific requirements of any retail environment.

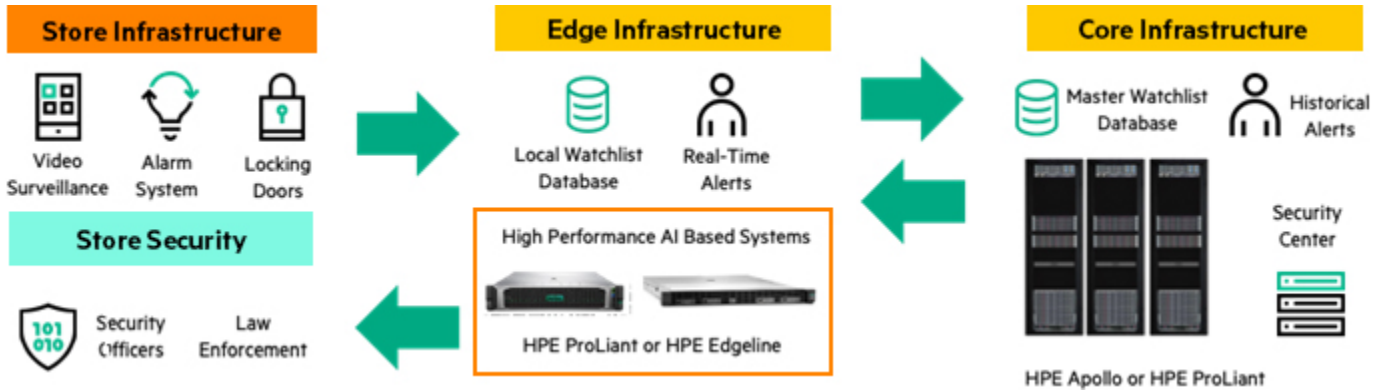


FIGURE 1. End-to-end HPE and NVIDIA IVA architecture for loss prevention

HPE’s robust portfolio offers superior processing capabilities to run the most demanding video analytics algorithms at the edge. With a high concentration of NVIDIA GPUs, systems can analyze video data as events occur in real time, enabling retailers to extract and operationalize rapid insights. The combination of HPE ProLiant or Edgeline servers, NVIDIA EGX software and NVIDIA GPUs provides dynamic support for retail IVA applications.

SMART RETAIL AT THE EDGE

Smart retail solutions from HPE at the edge starts with servers built for in-store environments.

TABLE 1. Recommended HPE server configurations for in-store IVA

	HPE ProLiant DL380 Gen10 server	HPE Edgeline EL4000 server
Number of cameras and other IoT input streams in self-checkout lanes and stations	<45 (small configuration) 45-60 (medium configuration) 60-90 (large configuration)	<45 (small configuration) 45-60 (medium configuration)
Number of NVIDIA GPUs	3 x T4 GPUs (small configuration) 6 x T4 GPUs (large configuration)	1-3 x T4 GPUs (small configuration) 4 x T4 GPUs (medium configuration)
Maximum GPUs per server	7	4
When to use	<ul style="list-style-type: none"> • Where growth is anticipated to reach large numbers of cameras (>60) across self-checkout lanes/stations. • System needed that can run a mix of in-store workloads. • If compute rack and controlled compute environment exists in store. 	<ul style="list-style-type: none"> • For smaller stores or cameras counts (up to 60 total self-checkout cameras). • If the ability to cache, process, and manage video, audio, and other rich media in one converged system is required. • If rack unit not available in-store. HPE Edgeline systems can reside where the data creation occurs, including environmentally harsh, space-constrained, or dusty environments.

HPE Edgeline Converged Edge Systems

Representing a new industry product category—converged edge systems—and pioneered by Hewlett Packard Enterprise, HPE Edgeline is the industry’s first product in this category. Chassis type, number of servers, number of NVIDIA GPUs, number of CPU cores, memory, and storage can all be tailored to fit the precise number and quality of video feeds and intensity of analytics needed.

The HPE Edgeline EL4000 Converged Edge System accommodates multiple NVIDIA GPU cards, up to four independent server cartridges, including the HPE ProLiant m510 Server, giving an impressive 64 Intel® Xeon® cores, 512 GB memory, 8 TB of SSDs, and 8 10GbE ports in a slim 1U form factor. It can accommodate NVIDIA T4 GPUs, each connecting to one server cartridge, giving us an incredible amount of



processing power to quickly run the most demanding video analytics algorithms at the edge. Dual redundant power supplies, ruggedization (up to MIL-STD through HPE partners), and the backing of industry certifications such as NEBS Level 3, make this a highly reliable system. This is on par with server equipment used by telcos and other mission-critical operators. The HPE EL4000 supports several mounting options to fit any environment—be it a rack, wall, desk, or even within customer equipment.

HPE ProLiant systems for a modern and optimized IT environment

The HPE ProLiant DL family of servers are the most flexible, reliable, and performance-optimized HPE ProLiant rack servers—ever. HPE continues to provide industry-leading compute innovations. The HPE ProLiant Gen10 portfolio, with flexible choices and versatile design, along with improved energy efficiencies, comprehensive management, and industry-leading support, ultimately lowers your TCO.

The HPE ProLiant DL380 server with NVIDIA T4 GPUs is the ideal platform for IVA and AI inference, providing unprecedented performance, security, and automation. The HPE ProLiant DL380 Gen10 server has an adaptable chassis, including new HPE modular drive bay configuration options with up to 30 small form factor (SFF), up to 19 large form factor (LFF), or up to 20 NVMe drive options along with support for up to three double-wide GPU options.

NVIDIA EGX Platform for Edge Computing

NVIDIA EGX Platform for Edge Computing is strategically engineered to accelerate AI computing at the edge. The platform features an easy-to-deploy cloud native software stack, validated on HPE Edgeline and ProLiant servers.

With the NVIDIA EGX software stack, traditional IT issues of installing software on remote systems disappear. Using Helm charts, containers, and continuous integration and continuous delivery (CI/CD), organizations can now deploy updated AI containers effortlessly in minutes.

A primary component of the NVIDIA EGX stack is the GPU Operator, which standardizes and automates the deployment of all necessary components for provisioning edge capabilities. The Kubernetes-based NVIDIA GPU Operator solves the problem of needing unique operating system (OS) images between GPU and CPU nodes; instead, the GPU Operator bundles everything you need to support a T4 GPU—the driver, container runtime, device plug-in, and monitoring with deployment by a Helm chart. Now, a single gold master image covers both CPU and GPU nodes.

All of these capabilities are hosted on the NGC container registry, a hub for GPU-optimized software that streamlines operations using pretrained models and workflows with easy-to-deploy AI frameworks and HPC application containers. NGC allows retailers to manage applications from any location—on-premises, in the cloud, and edge or using hybrid and multicloud deployments. Additionally, NVIDIA provides enterprise-grade support for NGC-ready systems with direct access to their experts, allowing retailers to transform their infrastructures with confidence.

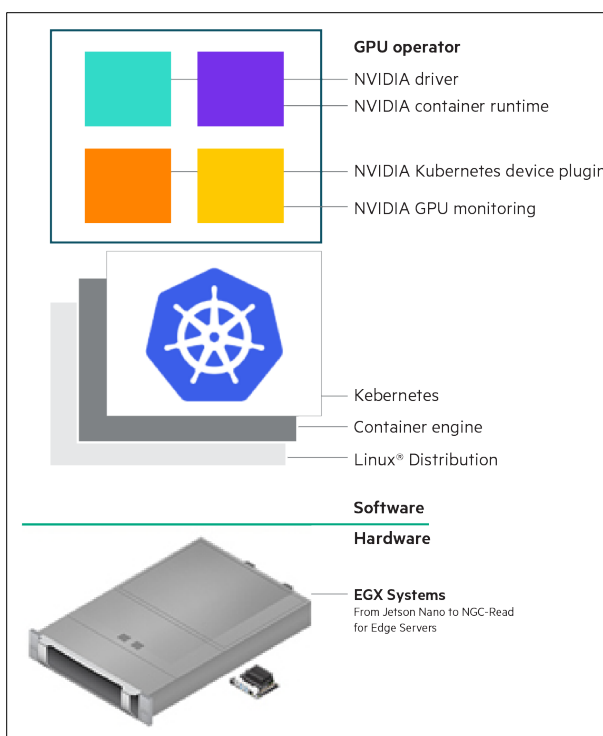


FIGURE 2. NVIDIA EGX Solution Stack (Source: NVIDIA)



INNOVATIVE SOLUTIONS ENABLING A NEW WAVE OF SMART RETAIL

HPE and NVIDIA work with a growing partner community of third-party software and hardware vendors that provide an AI analytics architecture for retail to further simplify IVA adoption. HPE's curated application programming interfaces (APIs) and independent software vendor (ISV) applications—which leverage the [NVIDIA EGX Platform](#)—have been tuned and validated to ensure compatibility with HPE and NVIDIA solutions.

The extensive list of AI-powered IVA applications supported by the solution stack includes loss prevention, in-store and online analytics, operations and logistics, as well as autonomous shopping to create an end-to-end smart retail solution.

Together, these technology leaders, HPE and HPE solution partners, deliver proven edge-to-cloud offerings that are modular, future-proof, and scalable. Now, retailers can innovate with confidence, maintaining high levels of throughput and productivity to handle their most challenging IVA applications with ease.

ACCELERATING RETAIL INSIGHTS WITH CORE INFRASTRUCTURE

Today's retail organizations are leveraging machine and deep learning—powerful components of AI—to deliver smart retail solutions. Beyond the edge, AI at the core can be used to improve forecasting, create new online customer experiences, or improve quality control. AI-powered applications utilize machine learning and deep learning, both of which are compute-intensive and data-heavy workloads that require exceptional capacity, acceleration, and durability.

But, capitalizing on the full range of AI and deep learning capabilities requires purpose-built computers capable of supporting all of the phases of AI development, including data preparation, model development, model training, and model validation. The HPE Apollo 6500 Gen10 System provides superior performance for the development and training of the advanced AI models used for the end-to-end IVA. With the HPE Apollo 6500 Gen10 System, enterprise customers receive:

- Unprecedented performance
- Economical AI and deep learning
- Rock-solid, enterprise-level reliability, availability, serviceability (RAS) features
- Easier system management
- Flexibility to support a wide range of workloads beyond AI

Backed by the processing power of [NVIDIA GPUs](#), enterprises can execute increasingly dense workloads at breakneck speeds to drastically improve model development and delivery.

HPE Apollo, ProLiant, and Edgeline systems are NGC-Ready, certified to run GPU-optimized software for deep learning, machine learning, and high-performance computing (HPC) available from the [NVIDIA GPU Cloud \(NGC\)](#). NGC GPU-optimized software saves time and money, providing easy-to-deploy, tested, ready-for-use data science, analytics, and ML/DL tools. With NGC, users can focus on building and deploying retail applications rather than managing the complexity associated with AI development environments. That way, AI is simple to develop and deploy. And with [HPE iLO and the Silicon Root of Trust](#), users can securely protect their most sensitive information.

HARNESSING THE FULL POWER OF AI CAPABILITIES

HPE and NVIDIA are empowering retailers to transform their operations with the development of leading-edge AI solutions to achieve greater performance and customer satisfaction. Together, our robust industry platforms are preparing retailers for success, backed by support from our experts around the globe as well as a broad network of software partners to deliver ready-to-use smart retail solutions or AI centers of excellence (CoE).

To get started, [HPE Pointnext Services](#) and HPE partners offer deep technical knowledge and innovation to support retailers as they transition to edge computing. HPE Pointnext Services experts provide a comprehensive portfolio of operational, advisory, and professional services that help retailers plan, build, implement, and optimize their retail IVA solutions. This approach includes accessing and eliminating roadblocks, identifying target applications, and setting workflow and performance requirements.

HPE Pointnext Services coordinates the installation, configuration, and validation of your new GPU-enabled, NGC-ready infrastructure. This process begins with service planning, which is delivered remotely and provides you with a detailed plan and schedule for their NGC Deployment and Integration Service. During service planning, HPE Pointnext Services experts communicate the list of service activities and confirm that deployment prerequisites have been addressed. Experts are available throughout the service—remotely or on-site—to ensure a seamless deployment and integration process that helps you quickly leverage the benefits of your solution and accelerate time-to-value.



This service is available for each HPE server configuration with three separation options to empower core, cluster, and core-to-edge environments:

- Container platform installation and configuration.
- Container scheduler installation and configuration.
- Cross-environment integration.

Additionally, HPE gives retailers the option to buy enterprise-grade NGC software support, serviced by NVIDIA. This provides direct access to subject matter experts to resolve issues faster. And with the ability to easily submit service requests with a clear escalation path, you can deploy their systems with confidence.

HPE Pointnext Services for NGC drastically improves time-to-production, allowing you to quickly spin-up containers to accelerate the development cycle. These services complement your IT team with globally available advisory and professional service assistance from HPE Pointnext Services, and they can be easily attached to hardware offerings and delivered alongside your new infrastructure. Working with HPE data and AI experts, retailers can take advantage of these innovations to gain a competitive edge in this dynamic open source AI ecosystem.

EXPANDING RETAIL AI APPLICATIONS

HPE and NVIDIA have the expertise and experience to design, build, and deploy cutting-edge AI capabilities. As a result, more retailers are looking to these technology leaders to address pressing business concerns, create intelligent stores, and reinvent the industry.

With [AI-enabled intelligent stores](#), retailers are driving improvements to every stage of operation. Today, retailers are implementing intelligent stores for four primary use cases: asset protection, store analytics, autonomous shopping, and store operations.

Asset protection: Retail shrinkage is a hundred-billion-dollar challenge that affects retailers globally. Theft can happen in the blink of an eye, so it requires quick detection. Using existing camera systems, retailers can install IVA applications to improve loss prevention and significantly reduce shrinkage. With AI-powered IVA, asset protection applications can detect ticket switching and mis-scans in real time, with even greater accuracy. These advanced technologies [detect shrinkage with up to 99% accuracy](#) and immediately notify sales associates for intervention.

Store analytics: Online-shopping data and in-store data generated from points of sale, cameras, and sensors are ripe with insights that can help determine customer preferences. IVA data can provide customer demographics as well as identify popular and unpopular store aisles, customer dwell time, and the number of unique visitors. With a better understanding of shopping behavior, retailers can improve store merchandising, expand overall sales, and deliver an enhanced omnichannel experience for their customers. They can even break down product interest by age and gender to group similar products for their target customer.

Autonomous shopping: The convenience of intelligent grab-and-go stores is erupting in popularity, where customers can use their mobile phones for cashier-less checkout. Grab-and-go locations are expected to [increase 4X annually](#) in the next three years. With autonomous checkout, retailers can deliver a more seamless shopping experience, while reducing costs and increasing revenue and margins. Autonomous checkout solutions range from frictionless AI-enabled shopping carts, to nano-stores and smart cabinets, to entirely autonomous stores.

Store operations: Store associates are the face of retail operations. To utilize employee resources more effectively, retailers must find ways to reduce time-consuming tasks such as performing inventory counts, replacing misplaced items, or scanning for out-of-stock products which interfere with customer service. Large retailers are now using robotics to handle these tasks, enabling store associates to manage inventory more efficiently and prioritize customer engagement. Many robotics solution providers utilize NVIDIA GPUs combined with the NVIDIA EGX software stack to build warehouse logistics machines, manufacturing robots, and smart assistants for automating repetitive tasks.

CONCLUSION

The rise of AI is enabling massive innovation to prepare retailers for the future. Implementing smart retail solutions to harness the full power of AI is a crucial step that today's companies must take to transform their operations from the ground up.

HPE and NVIDIA help you meet evolving retail industry demands for precision and performance with the industry's broadest range of computing solutions and AI leadership. HPE's robust range of servers, in conjunction with NVIDIA's GPU accelerators and GPU-optimized software offerings, ensure peak performance, efficiency, and flexibility for AI. Additionally, with a comprehensive portfolio of operational, advisory, and professional services and a vibrant ecosystem of partners experienced in the retail industry, we are dedicated to our customers' success.

Let HPE and NVIDIA reinvent your retail operations.



LEARN MORE AT

hpe.com/us/en/solutions/artificial-intelligence.html

hpe.com/us/en/servers/edgeline-systems.html

nvidia.com/en-us/industries/retail/smart-stores/

nvidia.com/en-us/industries/retail/

nvidia.com/en-us/autonomous-machines/intelligent-video-analytics-platform/

Follow us:

HPE HPC: [@HPE_HPC](https://twitter.com/HPE_HPC)

NVIDIA Data Center: [@NVIDIADC](https://twitter.com/NVIDIADC)

NVIDIA AI: [@NVIDIAAI](https://twitter.com/NVIDIAAI)

Make the right purchase decision.
Contact our presales specialists.



Chat



Email



Call



Get updates

© Copyright 2020 Hewlett Packard Enterprise Development LP. The information contained herein is subject to change without notice. The only warranties for 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.

Intel Xeon is a trademark of Intel Corporation in the U.S. and other countries. NVIDIA and the NVIDIA logo are trademarks and/or registered trademarks of NVIDIA Corporation in the U.S. and other countries. Linux is the registered trademark of Linus Torvalds in the U.S. and other countries. All third-party marks are property of their respective owners.

a50002330ENW, July 2020