VMware and NVIDIA to Enable Next-Gen Hybrid Cloud Architecture and Bring AI to Every Enterprise

300,000-plus VMware Customers Can Benefit from Unified Management of NVIDIA AI Software with All Apps, Tap into Security, Acceleration Capabilities of NVIDIA BlueField-2 DPU

VMworld — At VMworld 2020, VMware and NVIDIA today announced a broad partnership to deliver both an end-to-end enterprise platform for AI and a new architecture for data center, cloud and edge that uses NVIDIA® DPUs (data processing units) to support existing and next-generation applications.

Through this collaboration, the rich set of AI software available on the NVIDIA NGC™ hub will be integrated into VMware vSphere, VMware Cloud Foundation and VMware Tanzu. This will help accelerate AI adoption, enabling enterprises to extend existing infrastructure for AI, manage all applications with a single set of operations, and deploy AI-ready infrastructure where the data resides, across the data center, cloud and edge.

Additionally, as part of Project Monterey separately announced today, the companies will partner to deliver an architecture for the hybrid cloud based on SmartNIC technology, including NVIDIA's programmable BlueField®-2 DPU. The combination of VMware Cloud Foundation and NVIDIA BlueField-2 will offer next-generation infrastructure that is purpose-built for the demands of AI, machine learning, high-throughput and data-centric apps. It will also deliver expanded application acceleration beyond AI to all enterprise workloads and provide an extra layer of security through a new architecture that offloads critical data center services from the CPU to SmartNICs and programmable DPUs.

"We are partnering with NVIDIA to bring AI to every enterprise; a true democratization of one of the most powerful technologies," said Pat Gelsinger, CEO of VMware. "We're also collaborating to define a new architecture for the hybrid cloud—one purpose built to support the needs and demands of the next generation of applications. Together, we're positioned to help every enterprise accelerate their use of breakthrough applications to drive their business."

"AI and machine learning have quickly expanded from research labs to data centers in companies across virtually every industry and geography," said Jensen Huang, founder and CEO of NVIDIA. "NVIDIA and VMware will help customers transform every enterprise data center into an accelerated AI supercomputer. NVIDIA DPUs will give companies the ability to build secure, programmable, software-defined data centers that can accelerate all enterprise applications at exceptional value."

View today's VMworld 2020 CEO discussion featuring Pat Gelsinger and Jensen Huang, and join us at GTC 2020 on October 5 to learn more.

UCSF Advances Healthcare with NVIDIA and VMware

Among the organizations integrating their VMware and NVIDIA ecosystems is the UCSF Center for Intelligent Imaging. A leader in the development of AI and analysis tools in medical imaging, the center uses the NVIDIA Clara™ healthcare application framework for AI-powered imaging, and VMware Cloud Foundation to support a broad range of mission critical workloads. The center provides the University of California San Francisco community and academic and industry partners a critical resource for discovering, innovating and adopting AI to improve patient care.

"AI can be used to detect disease in large patient imaging studies more rapidly than the human eye, and, with further research, this technology will enable doctors to provide the fastest, most accurate and safest diagnoses and treatments for patients," said Christopher Hess, chair of Radiology and Biomedical Imaging at UCSF. "Bringing our NVIDIA Clara AI application frameworks and VMware Cloud Foundation together will help us expand our work in AI using a common data center infrastructure for activities such as training and research, and to help support time-sensitive urgent care diagnostics."

Enterprise-Ready Platform for AI

The first aspect of NVIDIA and VMware's collaboration – the integration of NVIDIA NGC with VMware vSphere and VMware Cloud Foundation – will simplify the deployment and management of AI for the most demanding workloads. Industries ranging from healthcare to financial services, retail and manufacturing will be able to easily develop and deploy AI workloads using containers and virtual machines, on the same platform as their enterprise applications, at scale across the hybrid cloud.

VMware customers will be able to accelerate data science and AI workloads building on existing infrastructure, resources and toolsets – helping to broaden adoption of AI and ML technologies. Data scientists, developers and researchers will gain immediate access to the wide array of NGC's cloud-native, GPU-optimized containers, models and industry-specific software development kits. NGC software is supported on a select set of pre-tested NVIDIA A100-powered servers expected from leading system manufacturers, including Dell Technologies, Hewlett Packard Enterprise and Lenovo.

Delivering New Hybrid Cloud Architecture for Next Gen Apps

The second element of VMware and NVIDIA's collaboration recognizes that, as next-generation workloads grow in complexity, SmartNICs and DPUs are critical technologies for securely accelerating a wide range of enterprise applications where the data resides.

VMware and NVIDIA are delivering a new architecture for the hybrid cloud that will help organizations evolve their infrastructure and operations and introduce a new security model that offloads hypervisor, networking, security and storage tasks from the CPU to the DPU. This new architecture will also extend the VMware Cloud Foundation operating model to bare metal servers.

The architecture is the cornerstone of VMware's Project Monterey, a technical preview announced at VMworld 2020 today. Leveraging the NVIDIA BlueField-2 DPU with VMware Cloud Foundation, customers will be able to speed up a wide range of next-gen and general-purpose applications, deliver programmable intelligence and operate a distributed, zero-trust security model across data centers, the edge and telco clouds.

Early Access for Visionary Enterprises

Extensive software engineering collaboration on the NVIDIA and VMware enterprise AI and accelerated computing platforms is underway. Companies seeking to operationalize AI and securely accelerate applications on their hybrid clouds can sign up for updates on availability. Tune in to VMware sessions at GTC 2020 next week to find out more.

About NVIDIA

NVIDIA's invention of the GPU in 1999 sparked the growth of the PC gaming market, redefined modern computer graphics and revolutionized parallel computing. More recently, GPU deep learning ignited modern AI — the next era of computing — with the GPU acting as the brain of computers, robots and self-driving cars that can perceive and understand the world. More information at http://nvidianews.nvidia.com/.

About VMware

VMware software powers the world's complex digital infrastructure. The company's cloud, app modernization, networking, security, and digital workspace offerings help customers deliver any application on any cloud across any device. Headquartered in Palo Alto, California, VMware is committed to being a force for good, from its breakthrough technology innovations to its global impact. For more information, please visit https://www.vmware.com/company.html.

are solely responsible for the content on such websites. The information in this press release is for informational purposes only and may not be incorporated into any contract.

Shannon McPhee
+1-310-920-9642
smcphee@nvidia.com
Eloy Ontiveros
VMware
+1-650-427-6145
eontiveros@vmware.com