Rackspace Technology Launches On-Demand GPU-as-a-Service Powered by NVIDIA Accelerated Computing

November 4, 2024

GPUaaS provides customers on-demand access to powerful accelerated resources for AI, machine learning, data analytics, and graphics rendering workloads

SAN ANTONIO, Nov. 04, 2024 (GLOBE NEWSWIRE) -- Rackspace Technology® (NASDAQ: RXT), a leading hybrid, multicloud, and AI technology services company, today announced the expansion of Rackspace Spot with a new geographic location and an on-demand GPU-as-a-Service powered by NVIDIA accelerated computing. The expansion comes as the demand for compute power is expected to continue to grow at a rapid pace. According to IDC, AI datacenter capacity is "projected to have a compound annual growth rate (CAGR) of 40.5% through 2027."[1] Rackspace Spot addresses that demand as a premier platform for computation-intensive applications like artificial intelligence, machine learning, and data analytics by providing on-demand, fully managed Kubernetes clusters delivered through a unique open market auction.

"Rackspace's GPUaaS will give customers on-demand access to powerful accelerated resources optimized for AI, machine learning, data analytics, and graphics rendering workloads," said Brian Lillie, President of Private Cloud for Rackspace Technology. "With Spot GPUaaS, you can harness high-performance GPUs without substantial upfront investments in hardware, achieving both cost-efficiency and scalability."

The Rackspace Spot GPU-enabled platform is hosted in our newest next-generation Rackspace SJC3 data center in Silicon Valley. The state-of-the-art facility offers enhanced performance and reliability, providing customers with advanced infrastructure from the ground up.

"Rackspace is excited to unveil significant Spot enhancements that deliver increased customer value and performance. Gen-2 provisioning is now available, significantly reducing deployment time and increasing reliability. In addition, Spot is now in seven global locations with the addition of our latest site in San Jose, CA (SJC3)," said Lance Weaver, Chief Product and Technology Officer of Rackspace Private Cloud. "The new data center, constructed from the ground up with the latest OpenStack technology, enhances performance and reliability. This strategic location in the US-West region reduces latency and provides faster access for our West Coast customers, improving their overall user experience. Moreover, this new data center caters to higher-density infrastructure demands, enabling our customers to scale resources seamlessly as their business needs evolve."

Introducing a New Server Class
Rackspace Spot is introducing a new server class equipped with NVIDIA accelerated computing, specifically the NVIDIA H100 Tensor Core GPU.

- GPU H100 Virtual Server v2.Mega Extra-Large offers one NVIDIA H100 GPU, an accelerator-optimized Intel 8568Y CPU with 48 hyperthreaded cores, 128 GB of memory, and advanced storage and networking capabilities.

This advanced GPU offering allows customers to harness the power of NVIDIA accelerated computing to speed up complex workloads such as artificial intelligence, machine learning, and data analytics. Customers can achieve faster processing times and improved operational efficiency with enhanced computational capabilities. Access to cutting-edge technology equips businesses to innovate rapidly and maintain a competitive advantage in their respective markets.

Rackspace Spot is available on-demand at spot.rackspace.com. Detailed specifications of the Spot GPU instances are available by clicking here.

About Rackspace Technology
Rackspace Technology is a leading end-to-end, hybrid, multicloud, and AI solutions company. We can design, build, and operate our customers' cloud environments across all major technology platforms, irrespective of technology stack or deployment model. We partner with our customers at every stage of their cloud journey, enabling them to modernize applications, build new products, and adopt innovative technologies.

Media Contact: Natalie Silva, publicrelations@rackspace.com

[1] *AI Datacenter Capacity, Energy Consumption, and Carbon Emission Projections* (Doc #US52131624)