QCT Supports NVIDIA Enterprise AI Factory with Optimized QuantaGrid System

Latest QCT offering powers the next wave of enterprise AI accelerated by NVIDIA RTX PRO 6000 Blackwell Server Edition GPUs

SAN JOSE, CA, UNITED STATES, August 27, 2025 /EINPresswire.com/ -- Quanta Cloud Technology (QCT), a global data center solution provider, today announced they are leveraging the NVIDIA Enterprise AI Factory validated design with its latest AI-optimized QuantaGrid server. The QuantaGrid D75E-4U with NVIDIA RTX PRO 6000 Blackwell Server Edition GPUs is purpose-built to accelerate the most demanding enterprise AI workloads, from generative, agentic, and physical AI to robotics, simulation, and digital twin technologies.

> With the QuantaGrid D75E-4U accelerated by NVIDIA RTX PRO 6000 Blackwell Server Edition GPUs, we're not only empowering our customers—we're also transforming our own manufacturing processes."
>
> *Mike Yang, President of QCT*

As an NVIDIA system partner for building AI factories, QCT delivers infrastructure guidance for deploying and managing AI workloads on-premises using the NVIDIA Blackwell accelerated computing platform. The QuantaGrid D75E-4U is an NVIDIA RTX PRO Server designed to meet enterprise IT needs with accelerated computing, networking, storage, and software—all optimized to reduce deployment risks and speed up time-to-value.

"Enterprises are rapidly evolving their AI strategies, and they need infrastructure that can keep pace with the complexity and scale of modern AI workloads," said Mike Yang, President of QCT. "With the QuantaGrid D75E-4U accelerated by NVIDIA RTX PRO 6000 Blackwell Server Edition GPUs, we're not only empowering our customers—we're also transforming our own manufacturing processes through AI-driven automation and simulation."

Beyond enabling customers, QCT also leverages its own AI-ready infrastructures internally to transform manufacturing operations. In its smart factories, QCT deploys these NVIDIA RTX PRO

Servers to power digital twins that simulate and optimize production lines, enabling predictive maintenance, robotics coordination, and vision-based inspection. These real-world deployments demonstrate how QCT's AI infrastructure drives tangible business outcomes.

Built on the groundbreaking NVIDIA Blackwell architecture, the NVIDIA RTX PRO 6000 Blackwell GPU brings next-generation AI capabilities to enterprise environments. Equipped with 96GB of ultra-fast GDDR7 memory, the NVIDIA RTX PRO 6000 Blackwell GPUs provide unparalleled performance and flexibility to accelerate model fine-tuning, real-time inference, and a broad range of enterprise use cases including:

 Computer Vision at Scale – From intelligent video analytics in smart cities to defect detection in advanced manufacturing, enterprises can deploy the QuantaGrid D75E-4U powered by the NVIDIA RTX PRO 6000 Blackwell GPUs to build high-performance computer vision pipelines. With 96GB memory capacity, it ensures that high-resolution image and video streams can be processed in real time, supporting mission-critical insights and automation.

 LLM Inference and Generative AI – With optimized support for NVIDIA NIM microservices and frameworks like NVIDIA NeMo, the QuantaGrid D75E-4U enables low-latency inference for large language models. Enterprises can run secure, on-premises LLM deployments for applications such as enterprise knowledge assistants, customer service automation, and domain-specific generative AI.

 Digital Twin Simulation – NVIDIA Omniverse with NVIDIA RTX PRO 6000 Blackwell GPUs enables physics-accurate digital replicas of factories to simulate, test and optimize production. Validated on QuantaGrid D75E-4U systems, AI workloads such as predictive maintenance, vision-based inspection, robotics coordination and generative optimization run seamlessly to reduce risk and improve efficiency.

In addition to the readily available QuantaGrid D75E-4U, QCT also offers a comprehensive portfolio of NVIDIA MGX and NVIDIA HGX systems from 1U to 10U. These systems integrate diverse NVIDIA GPUs with QCT's advanced system design, including optimized airflow, redundant power, and tool-less serviceability.

To learn more about QCT AI-ready infrastructures accelerated by NVIDIA , visit:
https://go.qct.io/nvidia/qct-servers-powered-by-nvidia-gpus/

About QCT
Quanta Cloud Technology (QCT) designs, manufactures, integrates, and services cutting-edge offerings for 5G Telco/Edge, AI/HPC, Cloud, and Enterprise infrastructure via its global network. Product lines include hyper-converged and software-defined data center solutions as well as servers, storage, and network switches from 1U to entire racks with a diverse ecosystem of

hardware components and software partners to fit a variety of business verticals and workload parameters.

Jean Ko
QCT
+ +886912025348
jean_ko@quantatw.com
Visit us on social media:
LinkedIn
Facebook
X
