

QCT Showcases NVIDIA Accelerated Computing and Networking Technologies at the 2024 Open Compute Project Global Summit

Generative AI server solutions align with industry standards and promote an open hardware ecosystem

SAN JOSE, CA, UNITED STATES, October 15, 2024 /EINPresswire.com/ -- [Quanta Cloud Technology](#) (QCT), a leading provider of data center solutions, will showcase its accelerated computing solutions powered by NVIDIA technologies at the 2024 Open Compute Project (OCP) Global Summit.

The 2024 OCP Summit is taking place in San Jose at the McEnery Convention Center from Oct. 15-17.



QCT is on the cusp of unleashing the potential of generative AI at scale by designing and manufacturing new AI servers, AI compute racks, and liquid-cooled hardware powered by the latest NVIDIA technologies. By incorporating a combination of [NVIDIA Grace™ CPUs](#), [NVIDIA Blackwell GPUs](#), and NVIDIA NVLink™-C2C interconnect technology in NVIDIA HGX™ and NVIDIA MGX™ architecture-based systems, QCT is showcasing its commitment to modular designs and the principles of the open compute community, in collaboration with NVIDIA.

QCT also supports the NVIDIA Spectrum-X™ Ethernet platform for OCP platforms, offering AI-optimized networking for every data center. Featuring NVIDIA Spectrum-X SN5600 switches and the NVIDIA BlueField®-3 SuperNIC, Spectrum-X is an end-to-end platform built with the feature set needed for generative AI clouds and enterprises.

“At QCT, we are committed to driving innovation and enabling our customers to harness the full potential of generative AI,” said Mike Yang, President of QCT. “Our system collaborations with NVIDIA at the 2024 OCP Global Summit underscore our dedication to open standards and cutting-edge technology, empowering organizations to achieve unprecedented levels of

performance for modern data center workloads.”

QCT’s server designs based on the NVIDIA MGX architecture share the use of modular components across various products and reduce time to market while providing a compatible platform for future CPU, GPU, and DPU solutions. QCT’s latest NVIDIA MGX-based systems empower organizations to accelerate innovation, adapt to evolving requirements, and embrace the era of generative AI.

“A wide range of NVIDIA-powered AI servers, AI compute racks, and liquid-cooled hardware from our ecosystem of partners, including QCT, will be on display at OCP,” said Kaustubh Sanghani, Vice President of GPU Product Management at NVIDIA. “By providing groundbreaking compute and networking capabilities, NVIDIA and its partners are enabling enterprises to accelerate AI workloads across the data center, cloud, and network edge.”

Attendees can see these advanced design features in QCT’s product showcase during the 2024 OCP Global Summit:

□ QuantaGrid D75F-9U: An acceleration server meticulously engineered to confront the most intricate AI and HPC tasks. The D75F-9U is equipped with 8 interconnected NVIDIA Blackwell GPUs with fifth-generation NVIDIA NVLink. This dynamic configuration empowers the QuantaGrid D75F-9U to effortlessly manage diverse workloads.

□ QuantaGrid D75H-7U: The QuantaGrid D74H-7U effectively addresses the most complex AI and HPC tasks. It is enhanced with 8 NVIDIA Blackwell or NVIDIA Hopper™ architecture GPUs interconnected via NVIDIA NVLink and flexible I/O slots to support NVIDIA BlueField-3 SuperNICs and DPUs. This configuration allows the QuantaGrid D75H-7U to seamlessly handle diverse workloads, making it an ideal solution for compute-intensive tasks.

□ QuantaGrid D75E-4U: An AI server leveraging the NVIDIA MGX architecture and supporting next-gen NVIDIA PCIe GPUs, up to 8x DW AC 600W or 16x SW LC 1000W+ GPUs with the flexibility to support any AI/HPC-related workload.

□ QuantaGrid D75N-2U: The QuantaGrid D75N-2U is a scale-out, single-node NVIDIA GB200 NVL2 platform-based system equipped with two NVIDIA Grace CPUs and two NVIDIA Blackwell GPUs and built on the NVIDIA MGX architecture. Ideal for seamlessly integrating accelerated computing workloads, this system delivers exceptional performance for mainstream large language model (LLM) inference, retrieval-augmented generation (RAG), vector database search, and data processing.

□ QuantaGrid D75B-1U: An NVIDIA MGX system powered by the NVIDIA GB200 Grace Blackwell platform. Equipped with one Grace CPU and two NVIDIA Blackwell GPUs on a single platform, connected via a 900GB/s bidirectional ultra-low-latency NVIDIA NVLink C2C interconnect, it incorporates cold plate-based direct-to-chip liquid cooling and can scale up to the NVIDIA GB200

NVL72 with 18x QuantaGrid D75B-1U servers.

Visit QCT Booth #A37 during the 2024 OCP Global Summit or visit www.qct.io for more information.

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.

All other brands, names, and trademarks are the property of their respective owners.

Jean Ko

QCT

+886 912 025 348

jean_ko@quantatw.com

This press release can be viewed online at: <https://www.einpresswire.com/article/751533863>

EIN Presswire's priority is source transparency. We do not allow opaque clients, and our editors try to be careful about weeding out false and misleading content. As a user, if you see something we have missed, please do bring it to our attention. Your help is welcome. EIN Presswire, Everyone's Internet News Presswire™, tries to define some of the boundaries that are reasonable in today's world. Please see our Editorial Guidelines for more information.

© 1995-2024 Newsmatics Inc. All Right Reserved.