QCT Debuts Accelerated Servers Supporting Intel® Xeon® 6 processors and Intel® Gaudi® 3 AI Accelerators at COMPUTEX 2024

QCT server systems to take advantage of new processors and AI accelerators to optimize efficiency
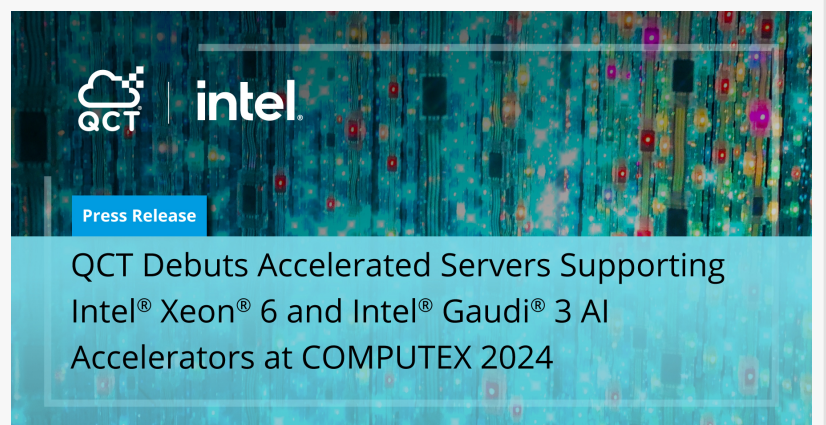
TAIPEI, TAIWAN, June 4, 2024 /EINPresswire.com/ -- Quanta Cloud Technology (QCT), a leading provider of data center solutions, is debuting accelerated servers powered by Intel® Gaudi® 3 AI accelerators and servers powered by Intel® Xeon® 6 processors at COMPUTEX 2024 in Taipei, Taiwan. QCT QuantaGrid and QuantaPlex

QCT INTEL COMPUTEX2024 Press Release

servers support for a variety of data center workloads, ensuring reliability for enterprise and edge computing applications, all the while delivering scale out power efficiencies that enable an impressive price-to-performance ratio due to the addition of the newly launched Intel® Xeon® 6 processors. QCT servers supporting Intel® Xeon® 6 processors and Intel® Gaudi® 3 AI accelerators deliver exceptional performance and efficiency for high-density and scale-out workloads.

"Today, Intel is launching its Intel Xeon 6 processors. Intel Xeon 6 processors feature microarchitecture options with new efficient cores and performance cores, delivering flexibility to meet the diverse efficiency and performance requirements of the market," said Zane Ball, Intel Corporate Vice President and General Manager of Data Platforms Engineering and Architecture. "While the upcoming Performance-cores or P-cores are architected for compute intensive workloads, the Efficiency-cores or E-cores launching now are ideal for cloud native workloads as they feature high core density and exceptional performance per watt resulting in lower TCO."

"QCT is thrilled to unveil the latest additions to our server product lines powered by Intel Xeon 6 and Intel Gaudi 3 AI accelerators at this year's COMPUTEX 2024, which are engineered to deliver exceptional efficiency and reliability," said Mike Yang, President of QCT. "By leveraging the power of Intel's latest innovations, our servers are set to empower businesses running 5G, AI and Cloud workloads."

Optimized for cloud scalability, both the Intel® Xeon 6® with E-cores and P-cores are designed for highly parallel workloads to help lower energy costs, drive sustainability, and maximize rack density, allowing customers to get more from their data center infrastructure. Intel® Xeon® 6 processors include Intel® Advanced Vector Extensions (Intel® AVX) to support both VNNI/INT8 and BF16/FP16 fast upconvert operations for image classification, object detection, natural language processing, recommendation systems, and other models to bring AI everywhere. Additionally, Intel® Gaudi® 3 AI accelerators deliver significant performance advancement over Gaudi® 2. The Intel® Gaudi® accelerator platform provides enterprises and AI developers a competitive alternative to data center GPUs, unlocking more GenAI solutions, including LLM and RAG, for more customers.

The QCT QuantaGrid D74H-7U is a flexible server system that can accommodate various vendor GPU accelerators. Its GPU sled delivers compatibility for the Intel® Gaudi® 3 AI accelerator baseboard, which accommodates eight Intel® Gaudi® 3 AI accelerator OAM mezzanine cards and provides customers with a modular subsystem that is easy to migrate into our D74H-7U system. QCT server systems supporting Intel® Gaudi® 3 AI accelerators deliver:

• Increased memory for LLM efficiency: >1TB HBM capacity per universal baseboard and 29.6 TB/s HBM bandwidth. The large HBM capacity and bandwidth allow the Intel® Gaudi® 3 accelerators to achieve state-of-the-art performance for GenAI training and inference.

• Better performance: Intel® Gaudi® 3 accelerators deliver 4x BF16 compute, 2x networking bandwidth, and 1.5x memory bandwidth over Gaudi® 2.

• Scalability: With systems built on the Intel® Gaudi® 3 accelerators, enterprise customers can address the compute demands of GenAI with flexibility and capacity to scale from 1 node to 1000s based on open, industry-standard ethernet networking.

The QCT QuantaGrid D55X-1U, QuantaGrid D55Q-2U, QuantaGrid S55R-1U, QuantaPlex S25Z-2U and QuantaPlex S45Z-2U servers powered by Intel® Xeon® 6 processors deliver:

• Increased Efficiency: Intel® Xeon® 6700 processors with E-cores support higher core density and exceptional performance per watt, 1.5x improvement in performance per watt compared to 5th Gen Intel® Xeon® processors, 2.4x improvement in performance per watt compared to 4th Gen Intel® Xeon® processors, and better energy efficiency and sustainability features to reduce one's operational carbon footprint.

• Enhanced performance: Intel® Xeon® 6700 processors with E-cores include up to 144 cores per CPU and faster memory at the same power envelope as the previous generation, up to 8 memory channels of DDR5 per CPU. The Intel® Xeon® 6 processors with Performance-cores are also designed with rich I/O capabilities.

• Greater Expandability and Flexibility: QCT servers with these processors offer extensive

expandability with up to 88 lanes of PCIe 5.0, 64 lanes of CXL 2.0, and 4 UPI links that also support up to 24GT/s, representing a 1.8x increase in bandwidth over the previous generation.

• Quality and Security: Intel® Xeon® 6 processors support more encryption keys on the latest Intel® Trust Platform Domain Extensions (Intel ® TDX) 6 firmware. These processors additionally continue support for Intel® Software Guard Extensions (Intel® SGX) with hardware-based security for your confidential computing.

Based on years of hardware manufacturing and infrastructure solution expertise, QCT has developed integrated AI and 5G solutions into their QuantaGrid, QuantaPlex, and QuantaEdge product lines. These server solutions feature toolless modular components, advanced air cooling and liquid cooling options, various high performance storage options, support for single and dual-width accelerators, and tool-less/hot swappable designs. Those in attendance at COMPUTEX 2024 are welcome to visit QCT in Booth G0042, on the 3rd floor of Hall 1, Nangang Exhibition Center, to learn more details or visit the QCT website at www.qct.io

Intel, the Intel logo, Xeon, and Xeon Inside are trademarks or registered trademarks of Intel Corporation in the US and/or other countries.

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
jean_ko@quantatw.com
Visit us on social media:
Facebook
X
LinkedIn
YouTube

This press release can be viewed online at: https://www.einpresswire.com/article/717009709