

Q.ANT bringt ersten kommerziellen Photonik-Prozessor auf den Markt – für energieeffizientes High-Performance-Computing und Real-Time AI-Anwendungen

Jetzt verfügbar: Die erste Native Processing Unit von Q.ANT verspricht eine mindestens 30-fache Verbesserung der Energieeffizienz und einen erheblichen Leistungsschub, der die Nachhaltigkeit von Rechenzentren in greifbare Nähe rückt.

Stuttgart - 19. November 2024 - [Q.ANT](#), das führende Startup für photonisches Computing, hat heute die Markteinführung seines ersten kommerziellen Produkts bekannt gegeben - einer photonischen Native Processing Unit (NPU), die auf der firmeneigenen Rechenarchitektur LENA (Light Empowered Native Arithmetics) basiert. Da das Produkt auf dem Industriestandard PCI-Express aufbaut, ist es mit dem heutigen Computing-Ecosystem kompatibel. Die Q.ANT NPU kann komplexe, nicht-lineare Berechnungen nativ mit Licht statt mit Elektronen ausführen. Darüber hinaus verspricht die Technologie eine mindestens 30-mal höhere Energieeffizienz und signifikante Verbesserungen der Rechengeschwindigkeit gegenüber der herkömmlichen CMOS-Technologie. Die Q.ANT NPU wurde für rechenintensive Anwendungen wie AI-Inference, Machine-Learning und Physiksimulationen entwickelt und hat sich bei der Lösung realer Rechenprobleme bewährt, wie Q.ANT am Beispiel der Zahlenerkennung für Deep-Neural-Networks nachgewiesen hat (siehe Pressemitteilung zum [Cloud-Zugang zur NPU](#)).

„Mit unserer photonischen Chiptechnologie, die jetzt über die Standard-PCIe-Schnittstelle verfügbar ist, bringen wir die unglaubliche Leistungsfähigkeit der Photonik direkt in reale Anwendungen. Wir machen damit ein deutliches Statement: Leistung und Nachhaltigkeit können Hand in Hand gehen“, sagt Dr. Michael Förtsch, CEO von Q.ANT. „Zum ersten Mal können Entwickler KI-Anwendungen erstellen und die Möglichkeiten des photonischen Rechnens erkunden, insbesondere für komplexe, nichtlineare Berechnungen. Experten haben beispielsweise errechnet, dass eine GPT-4-Anfrage heute 10-mal mehr Strom verbraucht als eine normale Internet-Suchanfrage. Unsere photonischen Chips bieten das Potenzial, den Energieverbrauch für diese Abfrage um das 30-fache zu senken.“

Der Durchbruch von Q.ANT beruht auf der firmeneigenen LENA-Plattform, die Thin-Film Lithium Niobate (TFLN) on Insulator umfasst. Q.ANT hat dieses photonische Material seit seiner Gründung im Jahr 2018 kontinuierlich weiterentwickelt. Diese Plattform ermöglicht eine präzise Lichtsteuerung auf Chipebene. Q.ANT kontrolliert den gesamten Prozess vom Wafer bis zum fertigen Prozessor kontrolliert und nutzt sein tiefes Verständnis von Licht. Damit erreicht das Unternehmen eine mathematische und algorithmische Dichte, die die herkömmliche CMOS-Technologie übertrifft. So kann beispielsweise eine Fourier-Transformation, die in der herkömmlichen CMOS-Technologie Millionen von Transistoren erfordert, mit einem einzigen optischen Element durchgeführt werden.

„Der neuartige Ansatz von Q.ANT für die photonische Verarbeitung ist ein bahnbrechender Schritt, um den steigenden Energiebedarf der KI-Ära zu decken“, sagt Dr. Eric Mounier, Chefanalyst, Photonics & Sensing beim Analystenhaus [Yole-Group](#). „Dieser Durchbruch wird durch die Verwendung optimaler Materialien für optische Schaltkreise ermöglicht, die Q.ANT in den letzten Jahren entwickelt hat. Diese neue Prozessorgeneration öffnet endlich den Zugang zu überlegenen mathematischen Operationen, die auf herkömmlichen GPUs zu energieaufwändig waren. Erste Leistungssteigerungen werden bei AI-Inference und AI-Training erwartet, was den Weg für hocheffizientes, nachhaltiges KI-Computing ebnet.“¹

¹ Quelle: *Optical Computing report, Yole Intelligence, 2024*

Die Q.ANT NPU kann die Berechnungsanforderungen für Machine-Learning, Computer-Vision oder für das Training und die Inferenz großer Sprachmodelle (LLM) reduzieren.

- Testläufe mit dem Q.ANT NPU-Demosystem in der Cloud mit MNIST-Datensätzen ergaben, dass der Native-Computing-Ansatz von Q.ANT eine mit linearen Netzen vergleichbare Genauigkeit bei weniger Energieverbrauch erreicht.
- Auch konnte in Simulationen von Kolmogorov-Arnold-Networks (KAN) Simulationen gezeigt werden, dass 43 % weniger Parameter benötigt und die Anzahl der Operationen um 46 % reduziert werden kann, was ihn als effizientere Wahl für KI-Inferenz etabliert.
- Weitere Tests und Simulationen zur Bilderkennung zeigen, dass die Q.ANT NPU deutlich schneller trainieren kann und eine genaue Erkennung mit nur 0,1 Millionen Parametern und 0,2 Millionen Operationen erreicht. Ein herkömmlicher Ansatz hat selbst bei 5,1 Millionen Parametern und 10 Millionen Operationen Mühe, akzeptable Ergebnisse zu erzielen.

Außerdem ermöglicht sie schnellere Lösungen für partielle Differentialgleichungen in Physiksimulationen, vereinfacht die Zeitreihenanalyse und verbessert die Effizienz bei der Lösung von Problemen der Graphentheorie. Im Gegensatz zur Standard CMOS Technologie verarbeitet die Q.ANT NPU Daten über Licht und ermöglicht so energieeffizientere mathematische Operationen. Während ein herkömmlicher CMOS-Multiplikator 1.200 Transistoren benötigt, um eine einfache 8-Bit-Multiplikation durchzuführen, erreicht die Q.ANT NPU dies mit einem einzigen optischen Element.

Verfügbarkeit und Verpackung

Die Q.ANT NPU kann ab sofort bestellt werden und ist im Februar 2025 lieferbar. Die Q.ANT NPU ist als schlüsselfertiger Native Processing Server (NPS) erhältlich, der vollständig mit einer herkömmlichen Serverlandschaft kompatibel ist und in jedes HPC- oder Rechenzentrum integriert werden kann. Durch den frühzeitigen Zugang zu seiner photonischen Technologie will Q.ANT den Durchbruch im Computing beschleunigen und neue Möglichkeiten in der KI und der wissenschaftlichen Forschung in einem Bereich eröffnen, der in nur wenigen Jahren erhebliche Leistungssteigerungen und ein enormes Energiesparpotenzial bieten wird.

Integration mit bestehenden AI Software Stacks

Die intuitive Schnittstelle der Q.ANT NPU, das Q.ANT Toolkit, lässt sich nahtlos in bestehende KI-Software-Stacks integrieren und ermöglicht es Entwicklern, auf verschiedenen Ebenen zu arbeiten, von der Multiplikation bis zu optimierten neuronalen Netzwerkoperationen. Außerdem bietet es eine umfassende Sammlung von Beispielanwendungen. Für Preise und weitere Details oder eine Vorbestellung wenden Sie sich bitte an native-computing@qant.gmbh.

Images



Bildunterschrift: Photonischer Prozessor auf PCI-Express: Die Q.ANT NPU ist für rechenintensive Anwendungen wie KI-Inferenz, maschinelles Lernen und Physiksimulationen konzipiert. (Bild: Q.ANT GmbH)



Bildunterschrift: Die Q.ANT NPU ist als schlüsselfertiger Native Processing Server (NPS) erhältlich und kann in jedes HPC oder Rechenzentrum integriert werden. (Bild: Q.ANT GmbH)



Bildunterschrift: Die Q.ANT NPU lässt sich nahtlos in HPC-Umgebungen integrieren und verspricht eine höhere Energieeffizienz und signifikante Verbesserungen der Rechengeschwindigkeit gegenüber der herkömmlichen CMOS-Technologie. (Bild: Q.ANT GmbH)

Über Q.ANT

Q.ANT hat sich zum Ziel gesetzt, die digitale mit der realen Welt zu verschmelzen. Um dies zu erreichen, entwickelt das deep-tech Start-up in der Business Unit Native Sensing Quantensensorik, die subtile und feinste elektrische und magnetische Felder direkt erfassen und verarbeiten kann. In der Business Unit Native Computing entstehen photonische Prozessoren, die in der Lage sind, Informationen aus der Natur nativ zu verarbeiten. Die Native Sensing und Native Computing Technologie von Q.ANT basiert auf dem Q.ANT Para.Digm Framework zur Erzeugung, Verarbeitung und Erkennung von Licht. Damit überwindet Q.ANT die Grenzen bestehender Technologien und erschließt neue Anwendungsfelder in verschiedenen Branchen wie High Performance Computing (HPC), Künstliche Intelligenz, Medizintechnik, Luft- und Raumfahrt, Maschinenbau und Prozessindustrie. Q.ANT wurde 2018 als unabhängiges Start-up aus den Forschungslaboren von TRUMPF ausgegründet und hat seinen Sitz in Stuttgart, Deutschland.

Kontakt für die Medien

Edith Laga
Fellow PR & Public Affairs, Q.ANT
edith.laga@gant.gmbh
+49 157 830 407 51
Q.ANT GmbH
Handwerkstr. 29, 70565 Stuttgart

Johannes Manger und Veit Mathauer
Sympra GmbH Agentur für Public Relations
gant@sympra.de
+49 711 947670