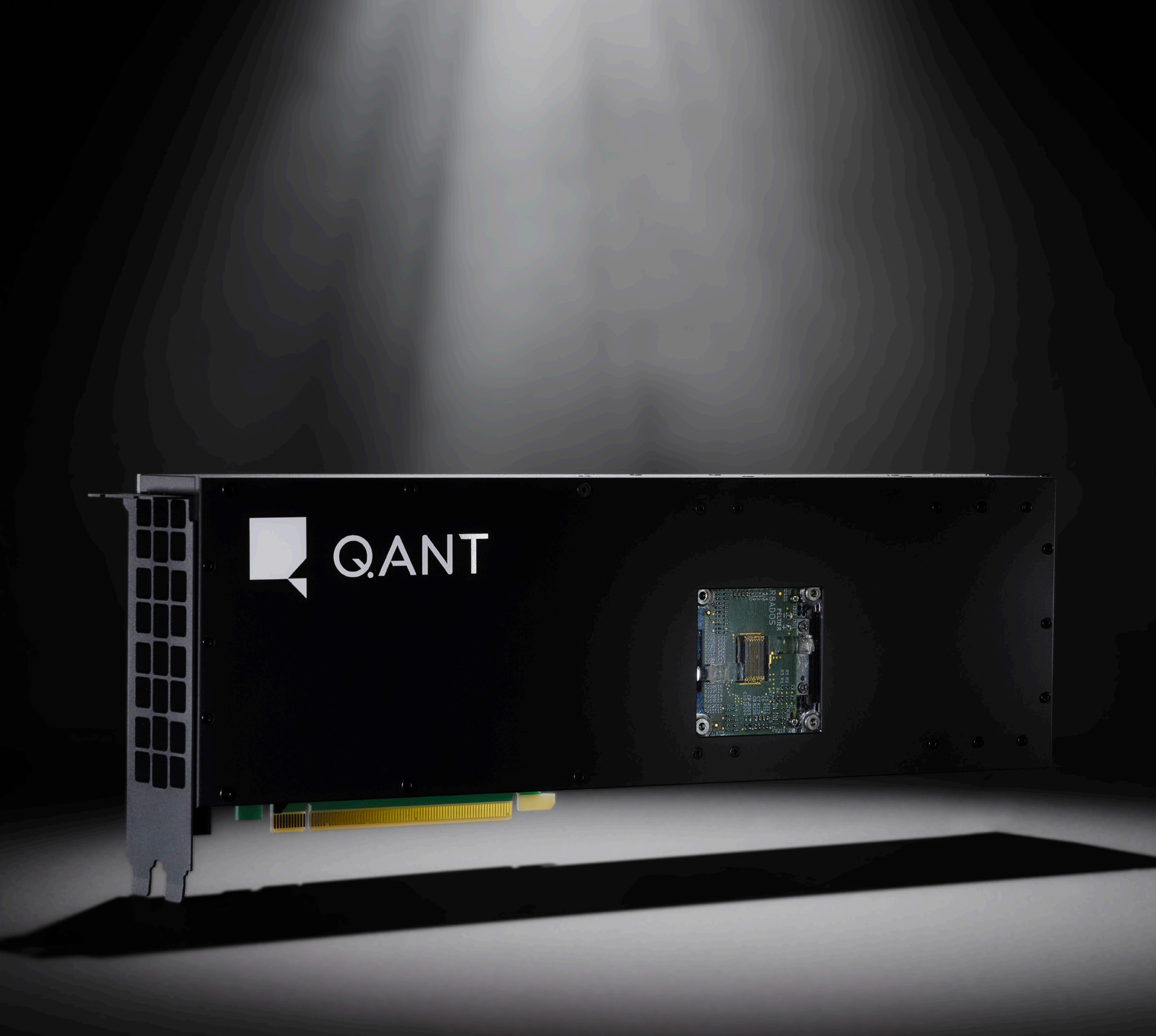
Q.ANT



Q.ANT Native Processing Unit Gen 2

The Photonic AI Accelerator

Photonic Native Processing Server NPS 2
for energy-efficient and fast HPC and AI

The Future of AI & HPC is Light

As AI scales globally, the traditional chip technology reaches its limits. Unlocking the true potential of AI demands a fundamental rethink of computing itself.

Q.ANT Native Computing ushers in a new era — computing with light instead of electricity **promising 30x energy savings and 50x higher performance**. The Q.ANT Native Processing Server (NPS) is the world's first commercially available photonic analog processor computing with light. By combining high-precision photonic processors with advanced algorithms based on non-linear math, Q.ANT Native Computing will extend today's compute architecture and enable next gen AI and HPC applications in machine learning, computer vision and physics simulations. The future of AI and HPC isn't silicon. It's light.

Overcoming today's Compute Bottlenecks with Light

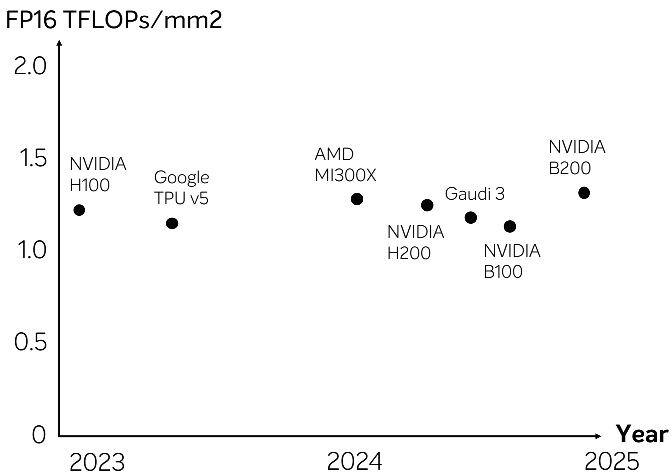
Q.ANT Native Computing tackles three fundamental challenges of today's compute ecosystem:

The progress in CMOS slows down
The miniaturization of transistors reaches its physical limits. State-of-the-art CMOS processors face performance shortcomings driven by an exponential growth in AI workloads.[1]

Energy consumption of AI data centers surges
Data center electricity consumption is set to more than double by 2030[2] limiting the upscaling of AI applications.

Limited semiconductor production capacity
AI hardware production is centralized in few regions and dominated by a small number of companies creating technological dependence.[3]

Semiconductor AI Processor Performance[4]



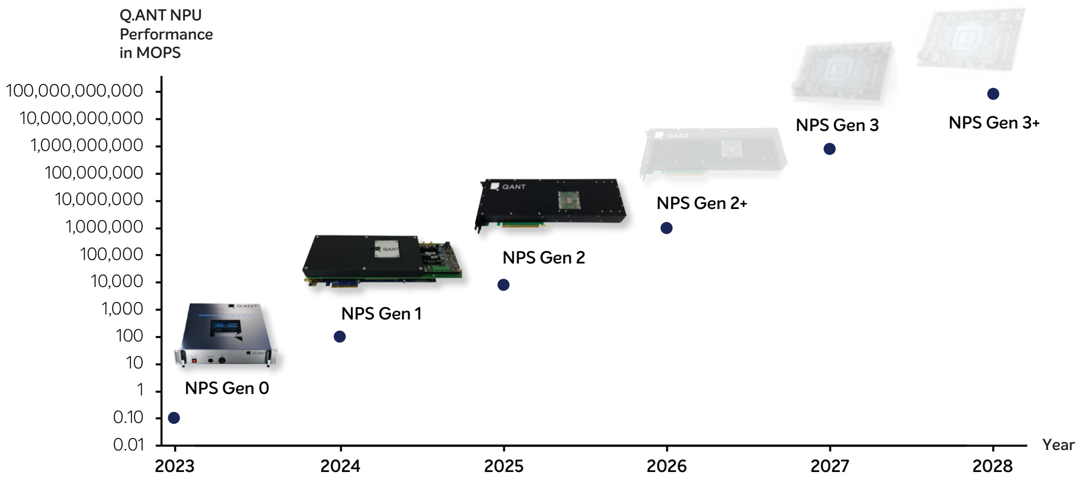
[1] EE Times (2024); AI Needs New Breakthroughs in Energy-Efficient Computing [2] International Energy Agency (2025); Energy and AI Report
[3] Semiconductor Insight (2024); AI Computing Hardware Market, Trends, Business Strategies 2025-2032 [4]Etched, NVIDIA, Future Ventures

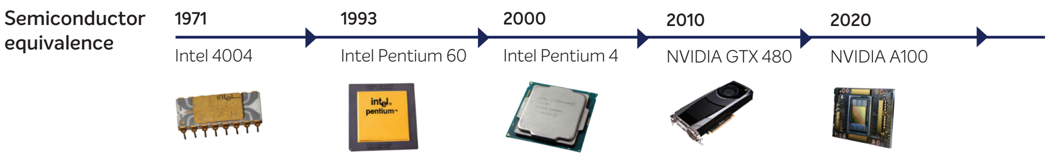
Operational at LRZ and JSC

Photonic Computing is reality: The Leibniz Supercomputing Centre (LRZ) and the Juelich Supercomputing Centre (JSC) – two of Europe's leading HPC data centers – integrate Q.ANT's NPS into their operational HPC environment. These first deployments mark a major step towards redefining how data centers approach performance, footprint and energy-efficiency.

Transistors defined the past. Light defines the future.

One Year in Photonic Processing Development equals Ten Years in CMOS Development:





This rapid advancement is driven by three key levers:

- **Physical Scaling:** Photonic processors exploit massive parallelism through multiplexing across space, time and wavelength making them ideal for large-scale networks with high-throughput.
- **Arithmetical Scaling:** Photonics naturally supports high-throughputmathematical operations essential for AI, such as matrix-vector multiplications or Fourier transforms, which can be directly implemented in the optical domain.
- **Algorithmic Scaling:** By using algorithms tailored to photonic hardware, such as nonlinear functions, results equivalent to digital can be achieved with fewer computational cycles.

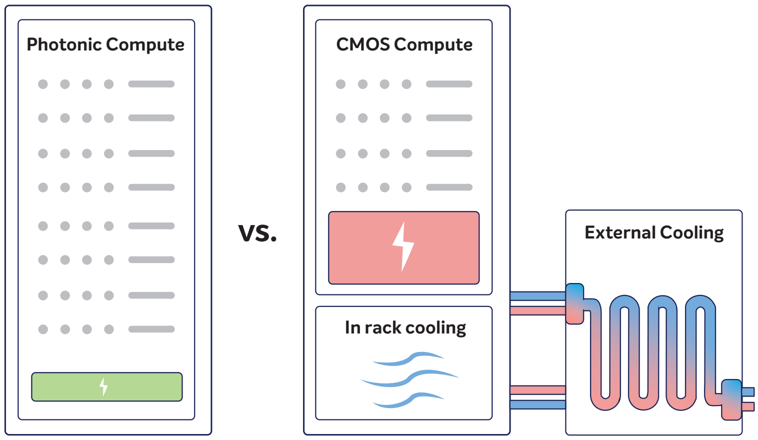
The Q.ANT Native Processing Server NPS

The Q.ANT Native Processing Server NPS is a photonic analog processor as 19“ rack-mountable server. Its core is the Native Processing Unit (NPU), a photonic AI Accelerator as PCIe card. Its Plug & Play system design enables a seamless integration into data centers and HPC environments, providing immediate access to photonic computing. The NPS is upgradable with additional NPUs to increase processing power as workloads grow.

Benefits for data centers

Native Computing by Q. ANT promises:

- **Up to 30x higher energy efficiency and 50x performance gains** per application
- **Higher computational density** for increased throughput without expanding footprint
- **Reduced infrastructure complexity** and space requirements
- **Lower cooling requirements** through removal of on-chip heat
- **Reduced operational costs** through lower energy consumption and performance gains
- **Early access to next-generation compute architecture** enables pioneering research
- **Seamless deployment** as 19” server fully compatible with standard x86 software stacks



Recognized by Gartner

> “ Photonic computing has several potential benefits over electronic computing, including increased bandwidth, processing power and storage, all while keeping energy and power consumption under control. ”

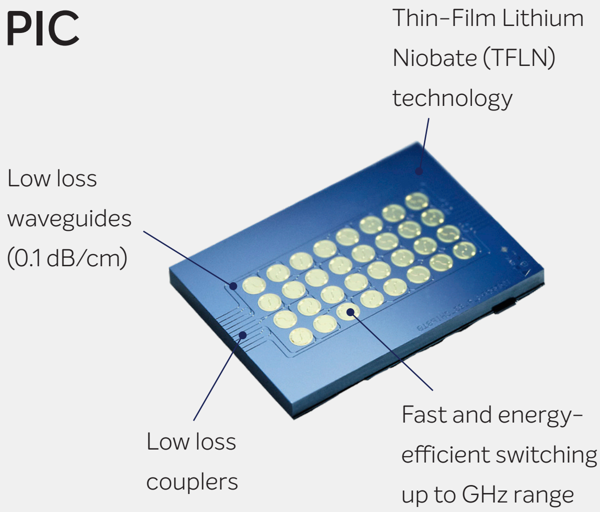
Gartner Research (2025): Emerging Tech: Emergence Cycle for Generative AI

Q.ANT is recognized as a Sample Vendor for Photonic Computing in the Gartner® Hype Cycle™ for Data Center Infrastructure Technologies report 2025.

Access the report:

Photonic Integrated Circuit PIC at the heart of NPS

Q.ANT‘s proprietary Photonic Integrated Circuit PIC based on Thin Film Lithium Niobate is at the heart of this innovation. It offers precise light control at the chip level. Q.ANT controls the entire value chain from raw materials to fully functional systems to make these processors achieve superior mathematical and algorithmic performance.



Core mathematical operations are native in optics

Mathematical operations required for a wide range of AI can be implemented directly in the optical domain using dedicated elements on the photonic chip. Photonic computing requires fewer elements than CMOS-based architectures for the same basic mathematical functions, leading to reduced circuit complexity, lower latency and higher computational density.

Mathematical operation	Required elements in optics	Required number of transistors in CMOS for 8-bit integer precision
Summation	Q.ANT photonic element γ	Approx. 200 transistors
Subtraction	Q.ANT photonic element γ	Approx. 200 transistors
Multiplication	Q.ANT photonic element Ω	Approx. 1200 transistors
Root extraction	Q.ANT photonic element γ	Approx. 7700 transistors
Squaring	Q.ANT photonic element Ω	Approx. 1200 transistors
Fourier transform	Q.ANT photonic element x	Algorithm dependent, several millions for near real-time Fast Fourier Transforms

From Wafer to Server – Q.ANT controls the full value chain

Q.ANT manages the entire value chain in photonic computing, ensuring full control over performance, quality, and scalability. The photonic chips at the heart of NPS are produced at Q.ANT’s own pilot line specifically designed for chip production using Thin Film Lithium Niobate (TFLN), the optimal material for photonic computing. By upcycling an existing CMOS production line at IMS CHIPS, this partnership establishes a blueprint for democratizing chip production capacity ensuring manufacturing resilience and reducing dependency on global supply chains.

IMS CHIPS

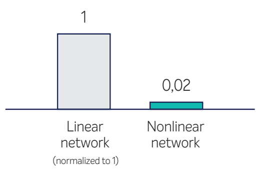
Q.ANT NPS – The choice for efficient nonlinear networks

While CMOS processors excel at linear, sequential processing, photonic processors are the natural hardware fit for large-scale nonlinear algorithms:

- Fewer parameters: Networks with nonlinear functions reduce the number of model parameters needed allowing higher accuracy per parameter and training budget.
- Fewer operations: In photonic processors one single optical element performs one nonlinear operation while CMOS requires 100-1000 transistors and multiple cycles.

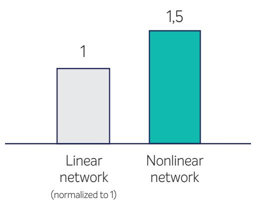
CPU

Nonlinear network runs 45x slower than linear network



NPU

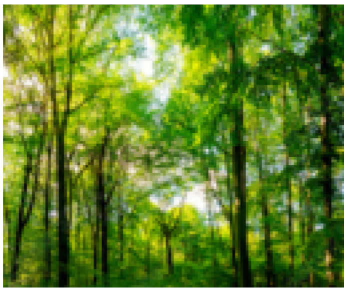
Nonlinear network runs 1.5x faster than linear network



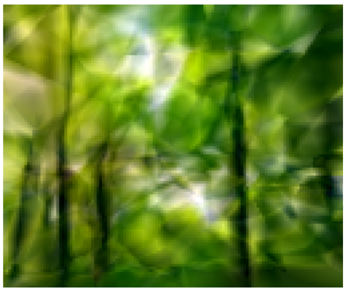
The clear picture – Image learning on Q.ANT NPS

In this example, a network using learnable nonlinear functions on Q.ANT's NPS reconstructs complex image patterns more accuretaly than a linear network on a CPU while 2x less parameters and 3x less operations are needed.

Original image



Linear network on CPU



parameters: ~20k
operations: ~670m

Nonlinear network on NPU

parameters: ~10k
operations: ~250m

Photonic Computing will benefit a broad application space

Photonic Computing has the potential to unlock new levels of performance and energy efficiency for a variety of applications through the efficient use of nonlinear functions.

	Task	Algorithm	Advantages using Photonics	Potential Use Cases
Artificial Intelligence	Function Approximation	Nonlinear Networks for high-frequency Signals	Nonlinear functions are native in optics, reducing the number of model paramenters	Financial Modeling Sales Forecasting National Grid Forecast
	Natural Language Processing	Transformer based Language Models	More efficient model architecture with fewer parameters for accelerated training & inference	Agentic AI & Chatbots Text Generation Programming
	Image Processing	Convolutional Neural Networks	Faster computations by transforming convolutions to multiplications in Fourier domain, reducing energy and memory usage	Computer Vision Medical Image Analysis Quality Control in Manufacturing
Computational Modeling & Simulation	AI-enhanced PDE Simulation	Neural Operators and Deep Operator Networks	Faster and more energy-efficient simulations remaining robust to deviations from the ideal Fourier transform	Finite-Element Analysis Computational Fluid Dynamics Process Modeling
	Molecular Property Prediction	Graph Neural Networks	Faster and more efficient networks with fewer parameters while maintaining high accuracy	Drug Discovery Materials Design Bio and Chem Informatics

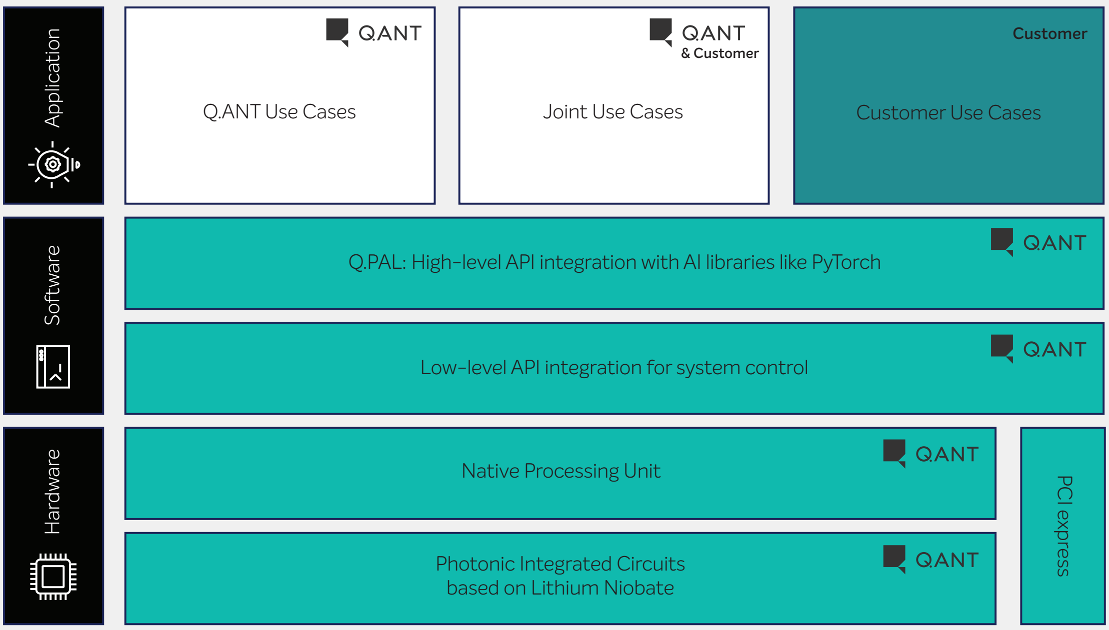
Q.PAL – The Q.ANT Photonic Algorithms Library

- The Q.ANT Photonic Algorithms Library Q.PAL is the software interface to the NPS which enables users to operate directly at the multiplication level or to leverage optimized neural network operations such as fully connected layers or convolutional layers.
- Q.PAL offers a comprehensive collection of example applications that illustrate how AI applications can be enhanced.
- These examples can be implemented directly or as a foundation for creating custom use cases.

Name	Description	Programming Language
Matrix Multiplication	Multiplication of a matrix and a vector	Python / C++
Image Classification	Classification of an image (e.g. based on ImageNet data set)	Python (Jupyter)
Semantic Segmentation	Segmentation of an image (based on a brain MRI data set)	Python (Jupyter)
Complex Line Fitting	Fitting of a high frequency line with a nonlinear network (e.g. based on simulated training data)	Python (Jupyter)

Q.ANT integrates into the established compute landscape

Q.ANT's photonic processing solution seamlessly integrates into the existing compute landscape. The Native Processing Unit at the heart of the NPS provides a PCIe interface housed in a standard 19" server, which makes the system plug-and-play. The NPU can be accessed via a software interface with C/C++ and Python APIs and will integrate into common AI frameworks such as PyTorch. Q.ANT supports customers in creating custom applications, providing the Q.ANT Photonic Algorithms Library and training resources.



Q.ANT Native Processing Server NPS 2

Technical Data Sheet

Data	Value/ designation
Form Factor	19” rack, 4U height Height: 178 mm; Width: 482 mm; Length: 595 mm
System Node	x86 processor architecture; supports multiple NPU cards per server
Operating System	Linux Debian/Ubuntu with long-term support
Network interface	2x 10 Gbit ethernet, 1x 1 Gbit service interface, Infiniband adapter optional
Software Interface	C / C++, Python API and PyTorch (Pilot integration)
API to NPU Subsystem	Linux device driver
Native Processing Unit NPU	• Full length PCIe card with 2 slot height • PCIe Gen4 x8 interface, shared memory & I/O windows • Upgradable with enhanced photonic integrated circuits • Upgradable with enhanced logic functions for performance
Throughput	8 GOPS
Photonic Integrated Circuit	Ultrafast photonic core based on z-cut Lithium Niobate
Power Supply	1600 W; AC input 115-240 V AC; 50-60 Hz
Power Consumption NPU	150 W
Operating Temperature Range	15 to 35°C
Weight	NPS (without NPU cards): 23.8 kg NPU: 2.38 kg

Join the next era of computing with light!

native-computing@qant.com | +49 711 25245-0 | www.qant.com

Q.ANT GmbH
Handwerkstraße 29
70565 Stuttgart, Germany

www.qant.com/photonic-computing