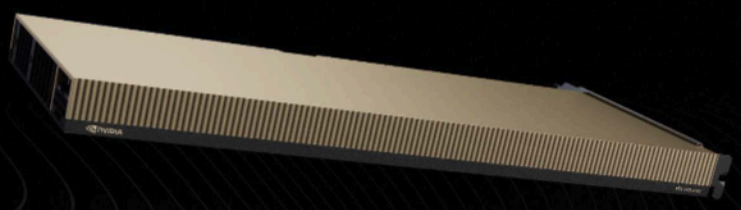
Datasheet



NVIDIA RTX PRO 4500 Blackwell Server Edition

Power-efficient NVIDIA Blackwell performance for any enterprise workload.

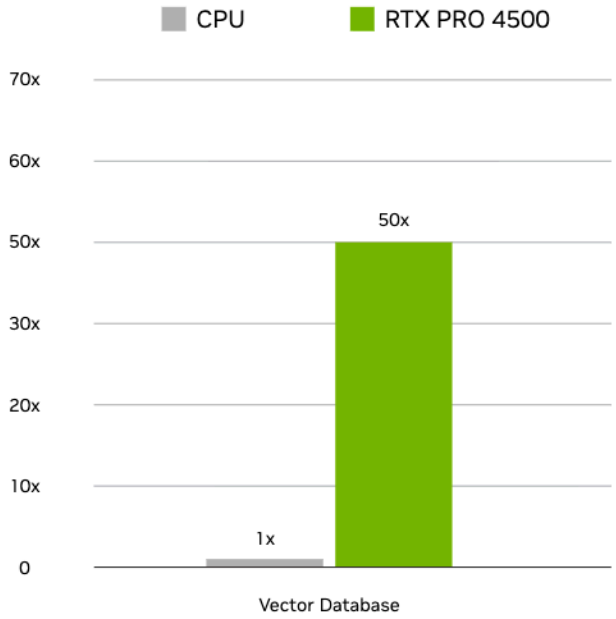
PNY Part Number: NVRTXPRO4500TCGPU-KIT

The NVIDIA RTX PRO™ 4500 Blackwell Server Edition GPU is an energy-efficient, multi-workload accelerator designed to deliver breakthrough performance in a compact, single-slot form factor. Based on the revolutionary NVIDIA Blackwell architecture, the RTX PRO 4500 Blackwell offers flexible capabilities for data processing, AI video, data, video streaming, and high-end visual computing across data center, edge, and cloud deployments.

Breakthrough Acceleration Across Data Center and Edge

The RTX PRO 4500 Blackwell delivers a comprehensive set of compute, graphics, and video acceleration capabilities in a compact design. With a power-efficient 165-watt power envelope and single-slot form-factor, the RTX PRO 4500 Blackwell can be deployed in almost any data center or edge platform, making it an efficient, cost-effective solution for any server or cloud instance from NVIDIA's partner ecosystem.

Up to 50x Higher Performance for Vector Databases



AMD 9654 192 vCPU vs. 8x RTX PRO 4500 server comparison. 33M vectors indexed in Milvus v2.5.25. CPU (HNSW) | GPU (cuVS CAGRA)

Key Features

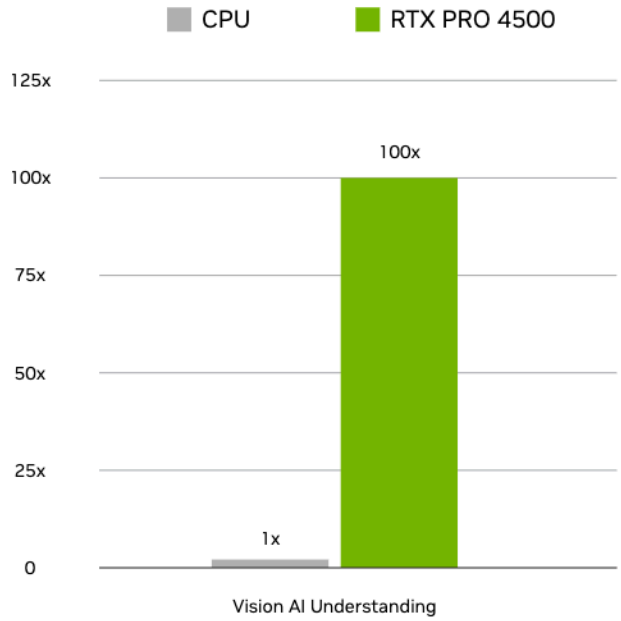
- Fifth-gen Tensor Cores with support for FP4 precision
- Fourth-gen Ray Tracing Cores
- 32 GB of GDDR7 memory
- 800 GB/s of memory bandwidth
- Ninth-gen NVENC and sixth-gen NVDEC with 4:2:2 support
- PCIe Gen 5

Workloads

- Data processing with NVIDIA CUDA-X Libraries
- Vision AI with NVIDIA Metropolis
- Small-model inference
- Graphics and visual computing
- Video processing and streaming

With the RTX PRO 4500 Blackwell Server Edition, massive, fragmented, multimodal enterprise data can be transformed into AI-ready data. Vector index builds and vector search applications are accelerated and optimized with cuVS for existing databases and vector search libraries, delivering up to 50x performance over CPU-only systems.

Up to 100x Higher Performance for Vision AI Understanding



Dual-socket AMD 9654 vs. 8x RTX PRO 4500 server comparison, running Qwen3-VL-8B for video captioning.

With fifth-generation Tensor Cores and 32 GB of GDDR7 memory, NVIDIA RTX PRO 4500 Blackwell GPUs paired with the NVIDIA Metropolis and CV-CUDA platform take vision AI applications and video-content-understanding performance to new heights. Compared to CPU-only systems, the RTX PRO 4500 Blackwell Server Edition can deliver up to 100x higher performance for AI video understanding, enabling enterprises to gain real-time insights and accelerate intelligent vision applications from the edge to the data center.

NVIDIA Blackwell Architecture Innovations

Fifth-Generation Tensor Cores

The RTX PRO 4500 Blackwell delivers up to 2x the performance of the previous generation. It supports FP4 precision for faster AI model processing with reduced memory usage, enabling higher throughput, more concurrent users, and lower latency for agentic and generative AI applications.

Fourth-Gen Ray Tracing Cores

Fourth-generation ray-tracing cores deliver 2x the ray-triangle intersection rate versus the previous generation. This enables up to 100x more ray-traced triangles to create photoreal, physically accurate scenes and immersive 3D designs with NVIDIA RTX™ Mega Geometry.

Next-Gen Video Engines

Video conferencing, video production, and streaming workflows are enhanced with real-time AI processing. Each RTX PRO 4500 Blackwell has 3x ninth-generation NVENC and 3x sixth-generation NVDEC engines with support for 4:2:2 encoding and decoding to explore a new realm of high-resolution video workflows.

GDDR7 Memory

New and improved GDDR7 memory significantly boosts bandwidth and capacity, empowering applications to run faster and work with larger, more complex datasets. Thirty-two GB of GPU memory and 800 gigabytes per second (GB/s) of bandwidth support larger datasets, power larger AI models, and accelerate complex simulations, immersive 3D scenes, and compute-intensive video tasks.

PCIe Gen 5

Support for PCIe Gen 5 provides 2x the bandwidth of PCIe Gen 4, improving data-transfer speeds from CPU memory and unlocking faster performance for data-intensive tasks like AI, data science, and 3D modeling.

Universal NVIDIA Multi-Instance GPU (MIG)

With MIG, a single RTX PRO 4500 Blackwell can be split into two isolated instances, each with 16 GB of memory to enable concurrent execution of multiple workloads, optimized GPU utilization, and secure isolation of different applications or users.

Enterprise-Ready

The RTX PRO 4500 Blackwell is optimized for 24/7 enterprise data center operations and is designed, built, tested, and supported exclusively by NVIDIA to ensure maximum uptime. The RTX PRO 4500 Blackwell is available within leading data center servers, edge platforms, and cloud instances from NVIDIA's partner ecosystem. These include a broad range of NVIDIA-Certified Systems™ for data center and edge.

NVIDIA-Certified Systems are the essential platform for building high-performance, scalable, and secure infrastructure for AI and accelerated computing workloads. NVIDIA-Certified Systems are tested and validated to deliver optimal performance for a wide range of workloads, including NVIDIA AI Enterprise, NVIDIA Omniverse™, graphics, visualization, and high-performance computing.

Technical Specifications

GPU Architecture	NVIDIA Blackwell Architecture
CUDA® Parallel Processing cores	10,496
NVIDIA RT Cores	82
FP4 Tensor Core	1.6 PFLOPS
FP8 Tensor Core	811 TFLOPS
FP16 | BF16 Tensor Core	406 TFLOPS
TF32 Tensor Core	203 TFLOPS
Tensor Single-Precision Performance (FP32)	51 TFLOPS
Peak RT Core Performance	154 TFLOPS
GPU Memory	32 GB GDDR7
Memory Interface	256-bit
Memory Bandwidth	800 GB/s
Power Consumption	165 W
Multi-Instance GPU	Up to 2x MIG at 16 GB each
NVENC | NVDEC	3x | 3x
Confidential Compute	Supported
Interconnect	PCI Express 5.0 x 16
Form Factor	Single-slot, FHFL (4.4” H x 10.5” L)
Thermal Solution	Passive
Power Connector	1x PCIe CEM5 16-pin

Ready to Get Started?

Ready to Get Started?
To learn more about RTX PRO Servers in the data center,
visit www.pny.com/nvidia-rtx-pro-4500-blackwell