Datasheet

NVIDIA RTX PRO 5500 Blackwell Workstation Edition

Powering the next era of AI.



Transform Workflows With Scalable Workstation Performance

As AI moves from experiment into production, industries face mounting pressure to scale the work behind it. Agentic AI and physical AI demand more computing resources—simulations that run for hours, models that must stay resident in memory, and agent pipelines that keep working long after someone steps away from a desk. Enterprises need breakthrough performance they can deploy to entire teams, not just individual desks.

Unlock next-level AI performance with the NVIDIA RTX PRO™ 5500 Blackwell Workstation Edition. Built on the NVIDIA Blackwell architecture and equipped with 84 GB of ultra-fast GDDR7 memory, it accelerates agentic and physical AI, simulation, and graphics—from synthetic data generation to large language model (LLM) inference, neural rendering, and complex 3D modeling. Designed for rack-mounted workstation deployment with a choice of air- or liquid-cooled thermal solutions, it lets organizations centralize, share, and scale workstation GPU power across the enterprise.

That capacity lets you work with larger models, longer context windows, and higher-fidelity scenes without offloading. Build and deploy agentic AI applications more efficiently with the FP4 data format, which reduces memory demands. Design hyperdetailed products, immersive environments, and lifelike renders faster, and streamline video pipelines with three encode and three decode engines supporting 4:2:2 color and multi-stream workflows.

Multi-Instance GPU partitions a single card into isolated instances, giving more users predictable, guaranteed performance from the same hardware while IT provisions, secures, and monitors every GPU from one place. With the RTX PRO 5500, securely build and run local AI agents with NVIDIA NemoClaw™, craft photorealistic visuals with neural rendering, and optimize precision-critical simulations for engineering and scientific research—all with stability and speed.

Key Features

- Enhanced streaming multiprocessors (SMs) built for neural shaders
- Fifth-generation Tensor Cores support FP4 precision, DLSS 4 multi-frame generation
- Fourth-generation Ray Tracing Cores built for detailed geometry
- 84 GB of GDDR7 memory
- 1.4 TB/s of memory bandwidth
- Ninth-generation NVIDIA Encoder (NVENC) and sixth-generation NVIDIA Decoder (NVDEC) with 4:2:2 support
- PCIe Gen 5
- Up to four DisplayPort 2.1b connectors
- NVIDIA RTX PRO Sync support
- Multi-Instance GPU (MIG) support
- Air- or liquid-cooled thermal solutions
- AI Management Processor

Breakthrough Innovations

NVIDIA Blackwell Streaming Multiprocessor: The new SM delivers increased processing throughput and neural shaders that integrate neural networks inside programmable shaders, driving the next decade of AI-augmented graphics.

Fifth-Generation Tensor Cores: Deliver up to 3x the performance of the previous generation and support for FP4 precision for faster AI processing with reduced memory usage, enabling local fine-tuning of LLMs and generative AI.

Fourth-Generation Ray Tracing Cores: Double the ray-triangle intersection rate of the previous generation to create photorealistic, physically accurate scenes and 3D designs with NVIDIA RTX™ Mega Geometry, which enables up to 100x more ray-traced triangles.

Next-Gen Video Engines: Ninth-generation NVENC and sixth-generation NVDEC engines add 4:2:2 encoding and decoding support, accelerating video production and streaming workflows with real-time AI processing.

GDDR7 Memory: Improved GDDR7 memory boosts bandwidth and capacity so applications run faster on larger, more complex datasets. With 84 GB of GPU memory and 1.4 TB/s bandwidth, tackle massive 3D and AI projects, fine-tune models locally, and drive larger multi-app workflows.

DLSS 4: Multi Frame Generation ensures ultra-smooth frame pacing for lifelike simulations, with up to 3x faster frame rates and stunning image quality in supported game engines and 3D rendering applications.

PCIe Gen 5: Double 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.

DisplayPort 2.1: Drive high-resolution displays at up to 8K at 240 Hz and 16K at 60 Hz. Increased bandwidth enables seamless multi-monitor setups, ideal for multitasking and collaboration, while HDR and higher color depth support ensures superior color accuracy for precision work, such as video editing, 3D design, and live broadcasting.

Universal MIG: Divide a single RTX PRO 5500 Blackwell into multiple isolated instances, each with dedicated resources, allowing for concurrent execution of multiple workloads, optimized GPU utilization, and secure isolation of different applications or users.

Enterprise Reliability

Designed for professionals who demand the best, NVIDIA RTX PRO solutions deliver unparalleled performance, reliability, and support. Every GPU is rigorously tested for a wide range of design, engineering, and AI workflows and continually optimized through enterprise drivers. With extensive ISV certifications, robust IT management tools, and enterprise-grade support, RTX PRO workstations are the trusted choice for enterprise and mission-critical work. Support for NVIDIA NemoClaw allows organizations to enforce strict privacy and security guardrails when running local, always-on AI agents, ensuring data remains protected during complex AI workflows.

Technical Specifications

GPU Architecture	NVIDIA Blackwell
NVIDIA CUDA™ Cores	21,760
Tensor Cores	Fifth generation
Ray Tracing Cores	Fourth generation
AI TOPS	2,376 AI TOPS[1, 2]
Single-Precision Performance	105 TFLOPS[1]
RT Core Performance	320 TFLOPS[1]
GPU Memory	84 GB GDDR7 with ECC
Memory Interface	448-bit
Memory Bandwidth	1,398 GB/s
System Interface	PCIe 5.0 x16
Display Connectors	Up to 4x DisplayPort 2.1b
Max Simultaneous Displays	>4x 3840 x 2160 @165 Hz >2x 7680 x 4320 @100 Hz
NVIDIA RTX PRO Sync	Supported[3]
Video Engines	>3x NVENC (ninth generation) >3x NVDEC (sixth generation)
MIG Support	>Up to 2x 42 GB[4] >Up to 1x 84 GB[4]
Power Consumption	Total board power: Up to 600 W
Power Connector	1x PCIe CEM5 16-pin
Thermal Solution	Active, air-cooled[5]
Form Factor	4.4” H x 11.2” L, dual slot, FHFL[5]
Graphics APIs	DirectX 12, Shader Model 6.6, OpenGL 4.6[6], Vulkan 1.4[6]
Compute APIs	CUDA 12.8, OpenCL 3.0, DirectCompute

Ready to Get Started?

To learn more, visit nvidia.com/rtx-pro-5500

1. Peak rates based on GPU Boost Clock.
2. Effective FP4 TOPS with sparsity.
3. Requires RTX PRO Sync card, sold separately.
4. MIG is supported on Linux only. See the NVIDIA MIG User Guide for details.
5. Liquid-cooled solution is available in the RXM form factor.
6. Product is based on a published Khronos specification and is expected to pass the Khronos conformation testing process when available. Current conformance status can be found at www.khronos.org/conformance

© 2026 NVIDIA Corporation. All rights reserved. NVIDIA, CUDA, NemoClaw, RTX, RTX PRO, and the NVIDIA logo are trademarks and/or registered trademarks of NVIDIA Corporation in the U.S. and other countries. All other trademarks and copyrights are the property of their respective owners. 5513311. OCT26

NVIDIA