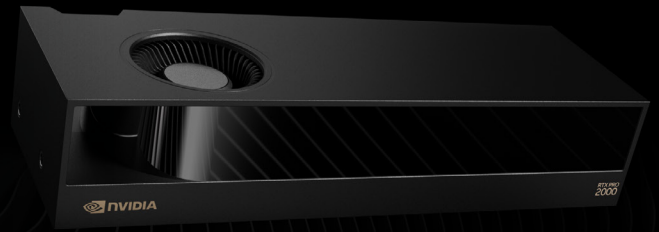




NVIDIA RTX PRO 2000 Blackwell

Powering the next era of AI.



Unlock the Power of NVIDIA Blackwell in a Compact Form Factor

As AI continues to advance at an incredible pace, industries face mounting pressure to harness its transformative power and adopt tools capable of handling agentic AI and hyper-realistic rendering. Enterprises require solutions that combine performance, scalability, and versatility to address the growing complexity of everyday workloads—from deploying domain-specific AI models with edge AI solutions to rendering billion-polygon engineering designs or simulating real-world physics with increased fidelity and precision.

The NVIDIA RTX PRO™ 2000 Blackwell GPU delivers exceptional performance in a power-efficient, compact form factor. Equipped with fifth-generation Tensor Cores, fourth-generation RT Cores, CUDA® Cores, and 16 GB of ultra-fast GDDR7 memory, this powerhouse accelerates demanding graphics and AI workflows. With enterprise-grade support and certification for major ISV applications, focus on mission-critical work with ease. Whether you're creating stunning 3D assets in SOLIDWORKS Visualize or a Siemens NX expert seeking speed, power, and seamless multi-app workflows, the NVIDIA RTX PRO 2000 Blackwell features next-level performance, power efficiency, and cutting-edge GPU hardware architecture.

Unlock unparalleled levels of productivity by integrating local AI inference into complex workflows. With improved bandwidth over the previous generation and the new FP4 data format, run complex AI models at unprecedented speeds. Revolutionize your daily tasks with extreme productivity boosts and edge inference, enabling you to innovate and iterate at breakthrough speeds while future-proofing your desktop for the new era of agentic AI. Personal AI assistants work alongside you to take actions, automate routine tasks, and evolve your daily workflows. RTX PRO 2000 brings local agentic AI to every desktop: run efficient NVIDIA® Nemotron™ Nano-class reasoning models and lightweight agents on-device, keeping your data, IP, and prompts on your workstation.

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
- 16 GB of GDDR7 memory
- 288 GB/s of memory bandwidth
- Ninth-generation NVENC and sixth-generation NVDEC with 4:2:2 support
- PCIe Gen 5
- Four Mini DisplayPort 2.1b connectors
- AI Management Processor

Breakthrough Innovations

The NVIDIA Blackwell architecture combines AI, ray tracing, and neural rendering technology, with massive performance and memory improvements to drive cutting-edge professional creative, design, and engineering workflows and power the next decade of innovation.

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

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

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

Next-Gen Video Engines: Enhance video conferencing, video production, and streaming workflows with real-time AI processing. Ninth-generation NVENC and the sixth-generation NVDEC engines provide 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 your applications to run faster, and work with larger, more complex datasets. With 16 GB of GPU memory and 288 GB/s bandwidth, tackle intricate design projects, leverage AI inference locally, and drive larger multi-app workflows.

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

PCIe Gen 5: Support for PCIe Gen 5 provides 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 2.1b: Achieve unparalleled visual clarity and performance, driving high-resolution displays at up to 8K at 165 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.

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.

Specifications

GPU Architecture	NVIDIA Blackwell
NVIDIA® CUDA® Cores	4,352
Tensor Cores	Fifth generation
Ray Tracing Cores	Fourth generation
AI TOPS	545 AI TOPS ^{1,2}
Single-Precision Performance	17 TFLOPS ¹
RT Core Performance	52 TFLOPS ¹
GPU Memory	16 GB GDDR7 with ECC
Memory Interface	128-bit
Memory Bandwidth	288 GB/s
System Interface	PCIe 5.0 x 8 ³
Display Connectors	4x Mini DisplayPort 2.1b ³
Max Simultaneous Displays	Up to four simultaneous displays ⁴
Video Engines	1x NVENC (ninth generation) 1x NVDEC (sixth generation)
Power Consumption	Max power consumption: 70 W
Thermal Solution	Active
Form Factor	2.7" H x 6.6" L, dual slot
Graphics APIs	DirectX 12, Shader Model 6.7, OpenGL 4.6 ⁵ , Vulkan 1.4 ⁵
Compute APIs	CUDA 12.8, OpenCL 3.0, DirectCompute

Ready to Get Started?

To learn more, visit: [nvidia.com/rtx-pro-2000](https://www.nvidia.com/rtx-pro-2000)

1. Peak rates based on GPU Boost Clock.
2. Effective FP4 TOPS with sparsity.
3. Uses full-length PCIe interface.
4. Multi-monitor: Four independent displays at 4K 165 Hz using DisplayPort.Two independent displays at 4K 360 Hz or 8K 100 Hz with DSC using DisplayPort. Other display configurations may be possible based on available bandwidth.
5. Product is based on a published Khronos specification and is expected to pass the Khronos conformance testing process when available. Current conformance status can be found at www.khronos.org/conformance

