Datasheet

NVIDIA RTX 2000 Ada Generation

Performance for endless possibilities.

Powering the Next Era of Innovation

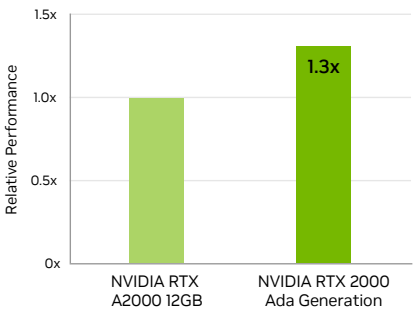
The NVIDIA RTX™ 2000 Ada Generation is a power-efficient, compact GPU that brings the power of RTX into the reach of more professionals. With real-time ray tracing, AI-accelerated compute, and high-performance graphics, the RTX 2000 empowers users to tackle complex tasks, from content creation and design to data analysis and AI-driven applications, with incredible speed and precision. Featuring the NVIDIA Ada Lovelace GPU architecture, it combines 22 third-generation RT Cores, 88 fourth-generation Tensor Cores, 2,816 CUDA® cores, and 16GB of GDDR6 graphics memory with ECC support. The RTX 2000 delivers breakthroughs in speed, efficiency, and power for everyday workflows, enabling creators, designers, and engineers to achieve new levels of productivity and innovation from the desktop.

NVIDIA RTX professional graphics cards are certified for a broad range of professional applications, tested by leading independent software vendors (ISVs) and workstation manufacturers, and backed by a global team of support specialists. Get the peace of mind to focus on what matters with the premier visual computing solution for mission-critical business.

Key Features

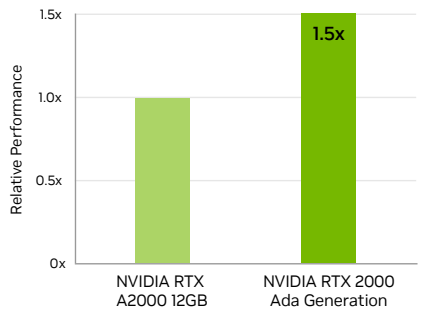
- Four Mini DisplayPort 1.4a
- AV1 encode and decode support
- DisplayPort with audio
- NVIDIA RTX Experience™
- NVIDIA RTX Desktop Manager software
- NVIDIA RTX IO support
- HDCP 2.2 support
- NVIDIA Mosaic[1] technology

Graphics



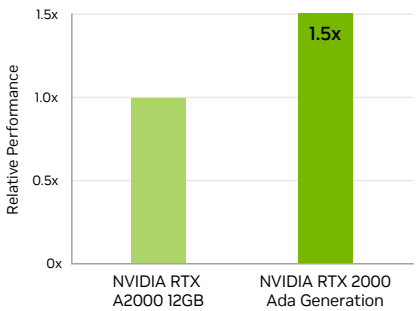
Tests run on an Intel Core i9-12900K Processor @ 3.2GHz (5.2GHz Turbo), 64GB RAM, Windows 11 Enterprise x64, SPECviewperf 2020, NVIDIA Driver 550.76. Relative speedup for 4K geomean score. Performance based on pre-release build, subject to change.

Rendering



Tests run on an Intel Core i9-12900K Processor @ 3.2GHz (5.2GHz Turbo), 64GB RAM, Windows 11 Enterprise x64, NVIDIA Driver 550.76. Average relative speedup for 1080p resolution Arnold, Blender, Cinebench, V-Ray 5.0, and Keyshot render tests. Performance based on pre-release build, subject to change.

Generative AI



Tests run on an Intel Core i9-12900K Processor @ 3.2GHz (5.2GHz Turbo), 64GB RAM, Windows 11 Enterprise x64, Stable Diffusion WebUI v1.6.0, NVIDIA Driver 550.76. Average relative speedup for 512x512 and 1024x1024 image generation. Performance based on pre-release build, subject to change.

Specifications	
GPU memory	16GB GDDR6
Memory interface	128-bit
Memory bandwidth	224 GB/s
Error-correcting code (ECC)	Yes
NVIDIA Ada Lovelace architecture-based CUDA Cores	2,816
NVIDIA fourth-generation Tensor Cores	88
NVIDIA third-generation RT Cores	22
Single-precision performance	12.0 TFLOPS[2]
RT Core performance	27.7 TFLOPS[2]
Tensor performance	191.9 TFLOPS[3]
System interface	PCIe 4.0 x 8[4]
Power consumption	Total board power: 70 W
Thermal solution	Active
Form factor	2.7" H x 6.6" L, Dual Slot
Display connectors	4x Mini DisplayPort 1.4a
Max simultaneous displays	4x 4096 x 2160 @ 120 Hz 4x 5120 x 2880 @ 60 Hz 2x 7680 x 4320 @ 60 Hz
Encode/decode engines	1x encode, 1x decode (+AV1 encode and decode)
VR-ready	Yes
Graphics APIs	DirectX 12, Shader Model 6.6, OpenGL 4.6[5], Vulkan 1.3[5]
Compute APIs	CUDA 11.6, OpenCL 3.0, DirectCompute

Ready to Get Started?

To learn more about NVIDIA RTX 2000, visit
www.nvidia.com/rtx-2000/

1 Windows 10, 11, and Linux are supported. This configuration does not offer framelock synchronization or display overlap functionality. I 2 Peak rates based on GPU Boost Clock. I 3 Effective FP8 teraFLOPS (TFLOPS) using the sparsity feature. I 4 RTX 2000 Ada Generation utilizes a full-length PCIe x8 interface. I 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

2999417. JAN24