



NVIDIA RTX A1000

Mighty performance. Minimal footprint.

Small Size, Big Impact—The Ultimate Compact Power

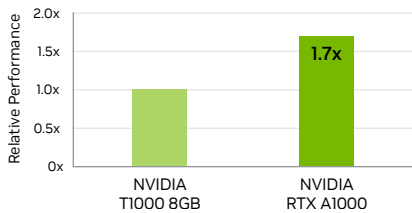
The NVIDIA RTX™ A1000 is a powerful, low-profile GPU that brings RTX-powered real time ray tracing and AI acceleration to a whole new audience. Built on the NVIDIA Ampere GPU architecture, it combines 2,304 CUDA® Cores, 72 third-generation Tensor Cores, and 18 second-generation RT Cores, and 8GB of GDDR6 graphics memory. With the RTX A1000, you can create more compelling visuals, explore new AI-powered workflows, and boost your productivity, all from a small-form-factor solution.

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

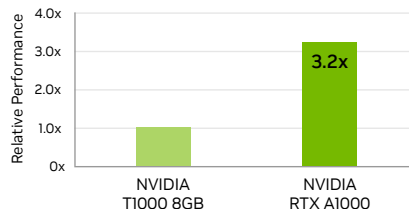
- > Second-generation RT Cores
- > Third-generation Tensor Cores
- > PCI Express Gen 4
- > Four Mini DisplayPort 1.4a
- > AV1 decode support
- > DisplayPort with audio
- > NVIDIA RTX Experience™
- > NVIDIA RTX Desktop Manager software
- > NVIDIA RTX IO support
- > HDCP 2.2 support
- > NVIDIA Mosaic¹ technology

Graphics



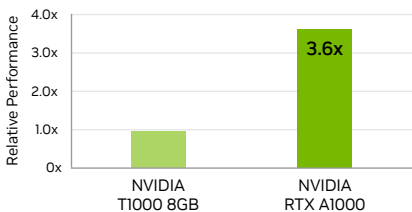
Tests run on an Intel Core i9-12900K Processor @ 3.2GHz (5.2GHz Turbo), 64GB RAM, Windows 11 Enterprise x64, SPECviewperf 2020, NVIDIA Driver 551.57. Relative speedup for 4K Maya score. Performance based on pre-release build, subject to change.

Rendering



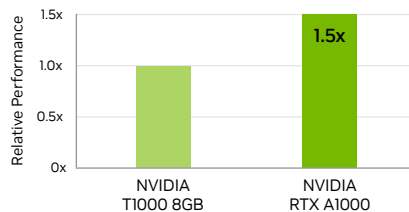
Tests run on an Intel Core i9-12900K Processor @ 3.2GHz (5.2GHz Turbo), 64GB RAM, Windows 11 Enterprise x64, NVIDIA Driver 551.57. Relative speedup for 1080p resolution Arnold render tests. Performance based on pre-release build, subject to change.

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 v1.7.0 with TensorRT extension, NVIDIA Driver 551.57. Average relative speedup for 512x512 image generation. Performance based on pre-release build, subject to change.

CAD



Tests run on an Intel i9-12900K @ 3.2GHz, 5.20GHz Turbo, 64GB RAM, Windows 11 Enterprise x64, NVIDIA driver 551.57. Results based on SOLIDWORKS 2022 FSAA On, GPU composite scores. Preliminary results on pre-production hardware and software, final performance may vary.

| Specifications | |
|--------------------------------------|--|
| GPU memory | 8GB GDDR6 |
| Memory interface | 128-bit |
| Memory bandwidth | 192GB/s |
| NVIDIA Ampere-based CUDA® Cores | 2,304 |
| NVIDIA third-generation Tensor Cores | 72 |
| NVIDIA second-generation RT Cores | 18 |
| Single-precision performance | 6.7 TFLOPS ² |
| RT Core performance | 13.2 TFLOPS ² |
| FP16 Tensor performance | 53.8 TFLOPS ³ |
| Peak INT8 Tensor performance | 107.8 TOPS ⁴ |
| System interface | PCIe 4.0 x8 ⁵ |
| Power consumption | Total board power: 50W |
| Thermal solution | Active |
| Form factor | 2.7" H x 6.4" L, single slot |
| Display connectors | 4x Mini DisplayPort 1.4a |
| Max simultaneous displays | 4x 4096 x 2160 @ 120Hz 4x 5120 x 2880 @ 60Hz 2x 7680 x 4320 @ 30Hz |
| Encode/decode engines | 1x encode, 2x decode (+AV1 decode) |
| Graphics APIs | DirectX 12, Shader Model 6.6, OpenGL 4.6 ⁶ , Vulkan 1.3 ⁶ |
| Compute APIs | CUDA 11.6, OpenCL 3.0, DirectCompute |

Ready to get started?

To learn more about the NVIDIA RTX A1000, visit:

[nvidia.com/rtx-A1000](https://www.nvidia.com/rtx-A1000)

1 Windows 10 and Linux. | 2 Peak rates based on GPU Boost Clock. | 3 Effective FP16 teraFLOPS (TFLOPS) using the sparsity feature. | 4 Peak INT8 TOPS with sparsity. | 5 RTX A1000 utilizes a full-length PCIe Gen 4 x8 interface. | 6 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

© 2024 NVIDIA Corporation and affiliates. All rights reserved. NVIDIA, the NVIDIA logo, CUDA, NVIDIA RTX, and NVIDIA RTX Experience are trademarks and/or registered trademarks of NVIDIA Corporation and affiliates in the U.S. and other countries. All other trademarks and copyrights are the property of their respective owners. APR24

