

NVIDIA RTX A4000

Sleek Design. Powerful Performance.



Amplified Performance for Professionals

The NVIDIA RTX™ A4000 is the most powerful single-slot GPU for professionals, delivering real-time ray tracing, AI-accelerated compute, and high-performance graphics to your desktop. Built on the NVIDIA Ampere architecture, the RTX A4000 combines 48 second-generation RT Cores, 192 third-generation Tensor Cores, and 6,144 CUDA® cores with 16GB of graphics memory with error-correction code (ECC) so you can innovate with uncompromised computing accuracy and reliability. The RTX A4000 also features a power-efficient, single-slot PCIe form factor that fits into a wide range of workstation chassis, so you can do exceptional work without limits.

NVIDIA RTX professional graphics cards are certified with 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 needed to focus on what matters with the premier visual computing solution for mission-critical business.

Features

- > PCI Express Gen 4
- > Four DisplayPort 1.4a connectors
- > AV1 decode support
- > DisplayPort with audio
- > 3D stereo support with stereo connector
- > NVIDIA GPUDirect® for Video support

- » NVIDIA Quadro® Sync II¹ compatibility
- > NVIDIA RTX Experience™
- » NVIDIA RTX Desktop Manager software
- > NVIDIA RTX IO support
- > HDCP 2.2 support
- > NVIDIA Mosaic² technology

1 Quadro Sync II card sold separately. | 2 Windows 10, Windows 11, and Linux. | 3 Peak rates based on GPU Boost Clock. | 4 Effective teraFLOPS [TFLOPS] using the new sparsity feature. | 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

SPECIFICATIONS

GPU memory	16GB GDDR6
Memory interface	256-bit
Memory bandwidth	448 GB/s
Error-correcting code (ECC)	Yes
NVIDIA Ampere architecture- based CUDA Cores	6,144
NVIDIA third-generation Tensor Cores	192
NVIDIA second-generation RT Cores	48
Single-precision performance	19.2 TFLOPS ³
RT Core performance	37.4 TFLOPS ³
Tensor performance	153.4 TFL0PS ⁴
System interface	PCIe 4.0 x16
Power consumption	Total board power: 140 W
Thermal solution	Active
Form factor	4.4" H x 9.5" L, single slot
Display connectors	4x DisplayPort 1.4a
Max simultaneous displays	4x 4096 x 2160 @ 120 Hz, 4x 5120 x 2880 @ 60 Hz, 2x 7680 x 4320 @ 60 Hz
Power connector	1x 6-pin PCle
Encode/decode engines	1x encode, 1x decode (+AV1 decode)
VR ready	Yes
Graphics APIs	DirectX 12 Ultimate , Shader Model 6.6, OpenGL 4.6 ⁵ , Vulkan 1.3 ⁵
Compute APIs	CUDA 11.6, DirectComput OpenCL 3.0

Learn more

