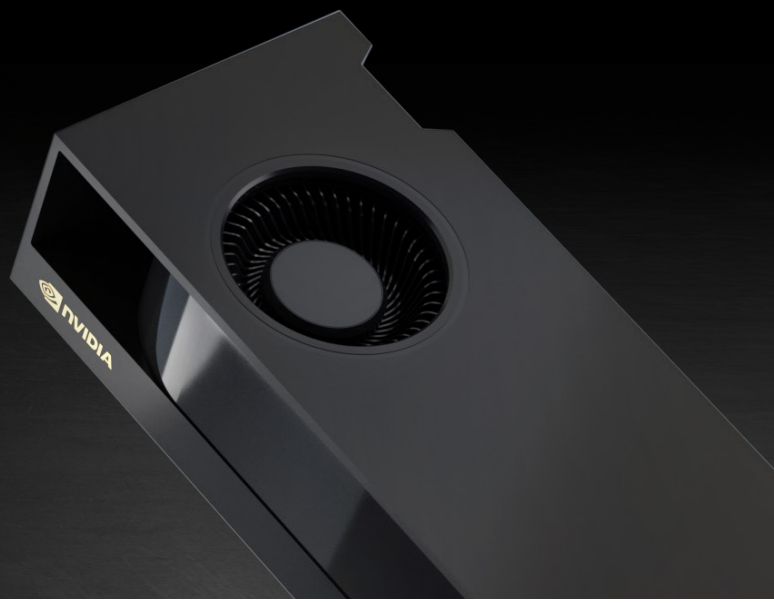
DATASHEET



NVIDIA RTX A2000 12GB

Compact Design. Unmatched Performance.

Accelerate Your Workflow

The NVIDIA RTX™ A2000 brings the power of NVIDIA RTX technology, real-time ray tracing, AI-accelerated compute, and high-performance graphics to more professionals. Built on the NVIDIA Ampere architecture, the VR ready RTX A2000 combines 26 second-generation RT Cores, 104 third-generation Tensor Cores, and 3,328 next-generation CUDA® cores and 6 or 12GB of GDDR6 graphics memory with error correction code (ECC) support for error free computing. RTX A2000 GPUs feature a power-efficient low profile, dual-slot PCIe form factor that fits into a wide range of small form factor workstations, and the RTX A2000 12GB doubles memory for even larger models and datasets. Design bigger, render faster, and work smarter than ever before with RTX A2000 12GB GPUs.

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 Mini DisplayPort 1.4a connectors with latching mechanism
- AV1 decode support
- DisplayPort with audio
- NVIDIA RTX Experience™
- NVIDIA RTX Desktop Manager software
- HDCP 2.2 support
- NVIDIA Mosaic[1] technology

Learn More

To learn more about the NVIDIA RTX A2000 12GB, visit **www.pny.com/rtx-a2000-12gb**

SPECIFICATIONS

PNY Part Number	**Retail: VCNRTXA200012GB-PB** **Single Bulk: VCNRTXA200012GB-SB** **Education: VCNRTXA200012GB-EDU** **Bulk: VCNRTXA200012GB-BLK**
GPU memory	**12 GB GDDR6**
Memory interface	**192-bit**
Memory bandwidth	**288 GB/s**
Error-correcting code (ECC)	**Yes**
NVIDIA Ampere architecture-based CUDA Cores	**3,328**
NVIDIA third-generation Tensor Cores	**104**
NVIDIA second-generation RT Cores	**26**
Single-precision performance	**8.0 TFLOPS[2]**
RT Core performance	**15.6 TFLOPS[2]**
Tensor performance	**63.9 TFLOPS[3]**
System interface	**PCI Express 4.0 x16**
Power consumption	**Total board power: 70 W**
Thermal solution	**Active**
Form factor	**2.7" H x 6.6" L, dual slot**
Display connectors	**4x mDP 1.4a** **with latching mechanism**
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 decode)**
VR ready	**Yes**
Graphics APIs	**DirectX 12 Ultimate,** **Shader Model 6.6,** **OpenGL 4.6[5], Vulkan 1.34[5]**
Compute APIs	**CUDA 11.6, DirectCompute,** **OpenCL 3.0**

1 Windows 10 and Linux. | 2 Peak rates based on GPU Boost Clock. | 3 Effective teraFLOPS (TFLOPS) using the new sparsity feature. | 4 GPU supports DX 12.0 API, hardware feature level 12 + 1. | 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**