NVIDIA RTX A5000
PERFECTLY BALANCED.
BLAZING PERFORMANCE.

Amplified Performance for Professionals

The NVIDIA RTX™ A5000 delivers the power, performance, capabilities, and reliability professionals need to bring their boldest ideas to life. Built on the NVIDIA Ampere architecture, the RTX A5000 combines 64 second-generation RT Cores, 256 third-generation Tensor Cores, and 8,192 CUDA® cores with 24 GB of graphics memory to supercharge rendering, AI, graphics, and compute tasks. Connect two RTX A5000s with NVIDIA NVLink[1] to scale memory and performance with multi-GPU configurations[2], allowing professionals to work with memory intensive tasks such as large models, ultra-high resolution rendering, and complex compute workloads. Support for NVIDIA virtual GPU software increases the versatility for enterprise deployments.

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 virtual GPU (vGPU) software support
> NVIDIA Quadro® Sync II[3] compatibility
> NVIDIA RTX Experience™
> NVIDIA RTX Desktop Manager software
> NVIDIA RTX IO support
> HDCP 2.2 support
> NVIDIA Mosaic[4] technology

To learn more about the NVIDIA RTX A5000, visit **www.pny.com/nvidia-rtx-a5000**

SPECIFICATIONS

PNY Part Number	**Retail: VCNRTXA5000-PB** **Single Bulk: VCNRTXA5000-SB** **Education: VCNRTXA5000-EDU** **Bulk: VCNRTXA5000-BLK**
GPU memory	**24 GB GDDR6**
Memory interface	**384-bit**
Memory bandwidth	**768 GB/s**
Error-correcting code (ECC)	**Yes**
NVIDIA Ampere architecture-based CUDA Cores	**8,192**
NVIDIA third-generation Tensor Cores	**256**
NVIDIA second-generation RT Cores	**64**
Single-precision performance	**27.8 TFLOPS[5]**
RT Core performance	**54.2 TFLOPS[5]**
Tensor performance	**222.2 TFLOPS[6]**
NVIDIA NVLink	**Low profile bridges connect two NVIDIA RTX A5000 GPUs[1]**
NVIDIA NVLink bandwidth	**112.5 GB/s (bidirectional)**
System interface	**PCI Express 4.0 x16**
Power consumption	**Total board power: 230 W**
Thermal solution	**Active**
Form factor	**4.4" H x 10.5" L, dual slot, full height**
Display connectors	**4x DisplayPort 1.4a[7]**
Max simultaneous displays	**4x 4096 x 2160 @ 120 Hz, 4x 5120 x 2880 @ 60 Hz, 2x 7680 x 4320 @ 60 Hz**
Power connector	**1x 8-pin PCIe**
Encode/decode engines	**1x encode, 2x decode (+AV1 decode)**
VR ready	**Yes**
vGPU software support[7]	**NVIDIA vPC/vApps, NVIDIA RTX Virtual Workstation**
vGPU profiles supported	**See the Virtual GPU Licensing Guide**
Graphics APIs	**DirectX 12 Ultimate, Shader Model 6.6, OpenGL 4.68[9], Vulkan 1.3[8]**
Compute APIs	**CUDA 11.6, DirectCompute, OpenCL 3.0**

1 NVIDIA NVLink sold separately. | 2 Connecting two RTX A5000 cards with NVLink to scale performance and memory capacity to 48GB is only possible if your application supports NVLink technology. Please contact your application provider to confirm their support for NVLink. | 3 Quadro Sync II card sold separately. | 4 Windows 10 and Linux. | 5 Peak rates based on GPU Boost Clock. | 6 Effective teraFLOPS (TFLOPS) using the new sparsity feature. | 7 Display ports are on by default for RTX A5000. Display ports are not active when using vGPU software. | 8 GPU supports DX 12.0 API, hardware feature level 12 + 1. | 9 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**