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 NVLink1 to scale memory and performance with multi-GPU configurations2, 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

Arrow Contact Information

Email: IntelligentSolutions@arrow.com, Online: www.arrow.com/AIS/nvidia

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

Copyright © 2021 NVIDIA Corporation. All rights reserved. NVIDIA, the NVIDIA logo, CUDA, GPUDirect, NVLink, Quadro, RTX Experience, and RTX are trademarks and/or registered trademarks of NVIDIA Corporation in the U.S. and other countries. Other company and product names may be trademarks of the respective companies with which they are associated. All other trademarks are property of their respective owners. NVIDIA RTX A5000 | Datasheet | APR21

---

**SPECIFICATIONS**

- **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 second-generation RT Cores**: 64
- **Single-precision performance**: 27.8 TFLOPS5
- **RT Core performance**: 54.2 TFLOPS5
- **Tensor performance**: 222.2 TFLOPS6
- **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
- **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 supported**: NVIDIA vPC/vApps, NVIDIA RTX Virtual Workstation, NVIDIA Virtual Compute Server
- **vGPU profiles supported**: See the Virtual GPU Licensing Guide
- **Graphics APIs**: DirectX 12.09, Shader Model 5.17, OpenGL 4.69, Vulkan 1.27
- **Compute APIs**: CUDA, DirectCompute, OpenGL™