

DATASHEET

NVIDIA RTX A6000

Powering the World's Highest-Performing Workstations



The NVIDIA RTX™ A6000, built on the NVIDIA Ampere architecture, is engineered to meet the demands of today's most graphics and compute-intensive workflows. Designed for **designers, engineers, scientists, and creative professionals**, it delivers the performance and scalability needed for advanced visualization, AI, rendering, and simulation.

Equipped with the latest **RT Cores, Tensor Cores, and CUDA® Cores**, the RTX A6000 offers unmatched real-time ray tracing, AI acceleration, and compute performance. With its massive GPU memory capacity, it enables professionals to handle **large-scale models, complex datasets, and immersive environments** with ease.

Certified across a wide range of **independent software vendor (ISV) applications** and validated by workstation manufacturers, the RTX A6000 ensures reliability and stability for enterprise deployments. Backed by NVIDIA's global support and AceCloud's enterprise-grade infrastructure, the RTX A6000 is the **visual computing solution of choice** for demanding workflows that cannot compromise on performance.

SYSTEM SPECIFICATIONS

GPU Memory	48 GB GDDR6
Memory Interface	384-bit
Memory Bandwidth	768 GB/s
ECC	Yes
CUDA Cores (Ampere)	10,752
Third-Generation Tensor Cores	336
Second-Generation RT Cores	84
Single-Precision Performance	38.7 TFLOPS ³
RT Core Performance	75.6 TFLOPS ³
Tensor Performance	309.7 TFLOPS ³⁵
NVIDIA NVLink	Connects two NVIDIA RTX A6000 GPUs ¹²
NVLink Bandwidth	112.5 GB/s (bidirectional)
System Interface	PCI Express 4.0 x16
Power Consumption	Total board power: 300 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 4096x2160 @ 120 Hz 4x 5120x2880 @ 60 Hz 2x 7680x4320 @ 60 Hz
Power Connector	1x 8-pin CPU
Encode/Decode Engines	1x Encode, 2x Decode (+AV1 decode)
VR Ready	Yes

Up to 40% Faster Graphics Performance¹

SPECviewperf 2020



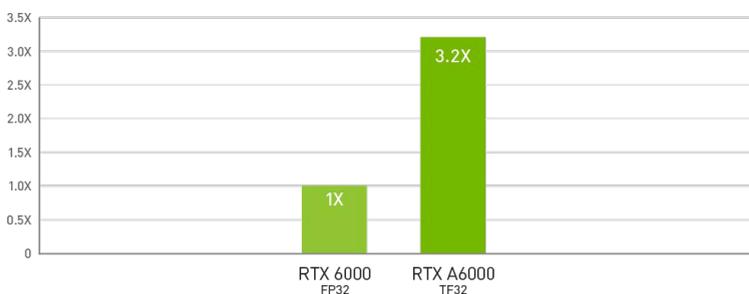
Up to 2X Faster Rendering Performance²

Autodesk VRED

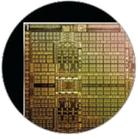


Over 3X Higher Out-of-the-Box Performance with TF32 for AI Training³

BERT Large Training

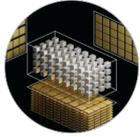


Groundbreaking Innovations



NVIDIA Ampere Architecture

The NVIDIA RTX™ A6000 is powered by the NVIDIA Ampere architecture, designed to revolutionize professional visual computing. With cutting-edge innovations in rendering, AI, graphics, and compute performance, Ampere significantly elevates RTX capabilities to meet the demands of today's most advanced professional workloads.



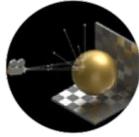
Third-Generation Tensor Cores

With TF32 precision, third-generation Tensor Cores deliver up to 5X higher training throughput for AI and data science models, without requiring code changes. They also support structural sparsity, boosting performance further. Beyond AI, Tensor Cores power features like DLSS, AI denoising, and enhanced editing for accelerated performance in select applications.



Ampere Architecture-Based CUDA® Cores

Ampere CUDA® Cores provide up to 2X performance per watt compared to Turing GPUs. This ensures massive gains for compute-heavy workloads, from 3D modeling and rendering to complex engineering simulations, making them more efficient and scalable for enterprise deployments.



Second-Generation RT Cores

The RTX A6000 introduces 2X the throughput of previous-generation RT Cores, enabling faster and more accurate real-time ray tracing. Professionals can now render photorealistic images, simulate lighting with precision, and accelerate workloads like product prototyping and motion blur rendering, achieving results with unprecedented visual accuracy.



Third-Generation NVLink

With up to 112 GB/s bidirectional bandwidth, third-generation NVIDIA NVLink® allows two GPUs to be connected for shared performance and memory—up to 96 GB combined. This enables massive scalability for rendering, AI, VR, and visual computing workloads. The slimmer NVLink connector also fits a wider range of chassis.

vGPU Software Support	NVIDIA vPC/vApps, NVIDIA RTX Virtual Workstation, NVIDIA Virtual Compute Server
vGPU Profiles Supported	1 GB, 2 GB, 3 GB, 4 GB, 6 GB, 8 GB, 12 GB, 16 GB, 24 GB, 48 GB
Graphics APIs	DirectX 12.0 ⁷ , Shader Model 5.1 ⁷¹⁰ , OpenGL 4.6 ⁸¹¹ , Vulkan 1.1 ⁸¹¹
Compute APIs	CUDA, DirectCompute, OpenCL™



PCI Express Gen 4.0

The NVIDIA RTX A6000 supports PCIe Gen 4.0, doubling bandwidth over PCIe Gen 3.0. This enhances data transfer speeds between CPU and GPU for AI, data science, and visualization tasks. PCIe Gen 4.0 also accelerates GPU Direct Storage and DMA transfers, ensuring faster video data throughput and improved I/O for demanding applications.

Why Choose L40S on AceCloud?

- Professional-grade GPU for graphics, AI, and compute
- Massive 48 GB memory for complex datasets & rendering
- Certified across leading ISV applications
- Enterprise-ready with 99.99% uptime SLA

Key Features

- PCI Express Gen 4
- Four DisplayPort 1.4a connectors
- AV1 decode support
- DisplayPort with audio
- VGA support⁴
- 3D stereo support with stereo connector
- NVIDIA GPUDirect[®] for Video support
- NVIDIA virtual GPU (vGPU) software support
- NVIDIA Quadro[®] Sync II⁵ compatibility
- NVIDIA Quadro Experience[™]
- Desktop Management Software
- NVIDIA RTX IO support
- HDCP 2.2 support
- NVIDIA Mosaic⁶ technology

Launch NVIDIA RTX A6000 on AceCloud

- ✓ Real-time ray tracing & AI acceleration
- ✓ 48 GB GDDR6 memory for large models
- ✓ vGPU-ready for enterprise deployments
- ✓ 99.99% uptime SLA

Scan to launch
your RTX
6000 VM →



Start with
[RTX 6000 now](#) 