



NVIDIA PROFESSIONAL GRAPHICS SOLUTIONS

You want to do great things, but don't have the tools to do your best work.

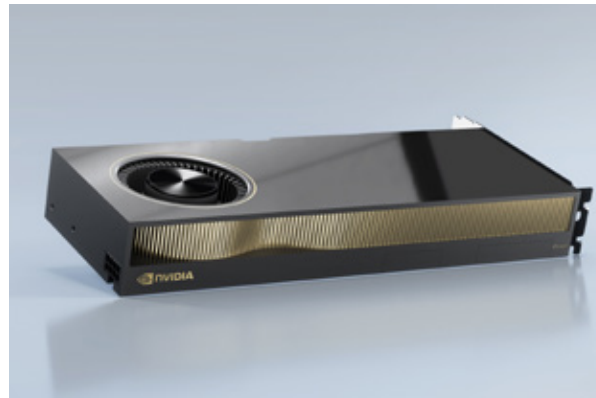
Now you can create and collaborate from anywhere, on any device, without battling slow performance, poor stability, or application incompatibility. With NVIDIA RTX™, you can unleash your vision and enjoy ultimate creative freedom.

NVIDIA RTX and NVIDIA Quadro professional visualization products power a wide range of laptop, desktop, and data center solutions. Leverage the latest advancements in real-time ray tracing, AI, virtual reality (VR), and interactive, photorealistic rendering. You can develop revolutionary products, tell vivid visual stories, and design groundbreaking architecture like never before. Support for advanced features, frameworks, and SDKs across all of our products gives you the power to tackle the most challenging visual computing tasks, no matter the scale.



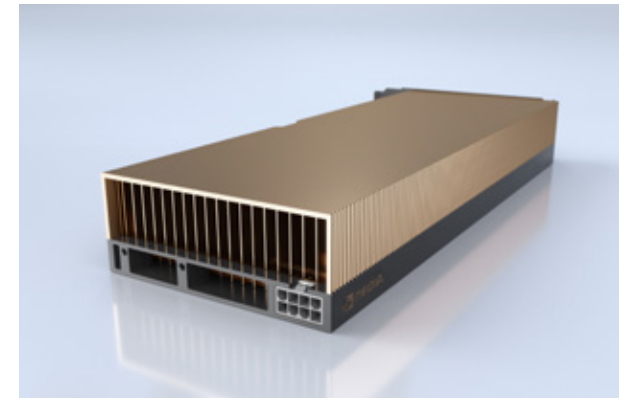
NVIDIA Professional Laptop GPUs

Today, professionals increasingly need to work on complex workflows like VR, 8K video editing, and photorealistic rendering on the go. NVIDIA RTX and Quadro RTX laptop GPUs deliver world-class performance in a portable form factor. NVIDIA RTX laptop GPUs combine the latest advancements in real-time ray tracing, advanced shading, and AI-based capabilities, so professionals can tackle demanding workflows from anywhere.



NVIDIA Desktop Workstation GPUs

NVIDIA RTX and Quadro RTX powered desktop workstations are designed and built specifically for artists, designers, and engineers, to take on their most challenging workloads. Connect multiple high-end desktop GPUs to scale up to 96 GB of GPU memory and performance to speed up your workflow. This delivers significant business impact across industries like manufacturing, media and entertainment, and energy.



NVIDIA Data Center GPUs

Demand for visualization, rendering, data science, and simulation continues to grow as businesses handle larger, more complex workloads. Scale up your visual compute infrastructure and tackle graphics-intensive workloads, complex designs, photorealistic renders, and augmented and virtual environments at the edge with NVIDIA GPUs. Optimized for reliability in enterprise data centers, NVIDIA GPUs feature both active and passive thermal solutions to fit into a variety of servers.

NVIDIA PROFESSIONAL GRAPHICS SOLUTIONS

| GPU SPECIFICATIONS | | | | | | PERFORMANCE | | | | DISPLAY TECHNOLOGY | | | | | | VIRTUAL REALITY | | OPTIONS | | | | |
|--|-----------------|--------------|------------|-----------------------|-----------------|--|------------------------------|--|------------------------------------|-------------------------|---|-------------------------|--------------------------|-----------------|---------------------------------------|----------------------------------|-----------------------|-----------------------|----------------------|---------------------------------------|-------------------------|-----------------------------|
| NVIDIA CUDA® Processing Cores ¹ | NVIDIA RT Cores | Tensor Cores | GPU Memory | Peak Memory Bandwidth | NVIDIA® NVLink® | Floating-Point Performance, Single Precision (TFLOPS, Peak) ² | Accelerated Double Precision | Tensor Performance (TFLOPS, Peak) ^{3,4} | Error-Correcting Code (ECC) Memory | Maximum Active Displays | DisplayPort 1.2 ⁵ and 1.4 ⁶ | HDMI via Adaptors, HDMI | NVIDIA SLI® ⁷ | USB-C Connector | High-Dynamic Range (HDR) ⁸ | NVIDIA Quadro® Mosaic Technology | VR Ready ⁹ | Variable Rate Shading | GPUDirect® for Video | Graphics Synchronization with Sync II | 3D Stereo ¹⁰ | Encode/Decode ¹¹ |

LAPTOP GPUs

| NEW | GPU Model | CUDA Cores | RT Cores | Tensor Cores | GPU Memory | Peak Memory Bandwidth | FP16 Performance (TFLOPS, Peak) | Accel. DP | Tensor Perf. (TFLOPS, Peak) | ECC | Max Displays | DP 1.2/1.4 | HDMI | SLI | USB-C | HDR | Mosaic | VR Ready | VR Shading | GPU Direct | Sync II | 3D Stereo | Encode/Decode | |
|-----|------------------|------------|--------------|---------------|--------------|-----------------------|---------------------------------|-----------|-----------------------------|-----|--------------|------------|------|-----|-------|-----|--------|----------|------------|------------|---------|-----------|---------------|--|
| | NVIDIA RTX A5000 | 6,144 | 48 (2nd Gen) | 192 (3rd Gen) | 16 GB | 448 GB/s | 21.7 | | 174.0 | | 4* | Yes* | Yes* | | | | | | | | | | | |
| | NVIDIA RTX A4000 | 5,120 | 40 (2nd Gen) | 160 (3rd Gen) | 8 GB | 384 GB/s | 17.8 | | 142.5 | | 4* | Yes* | Yes* | | | | | | | | | | | |
| | NVIDIA RTX A3000 | 4,096 | 32 (2nd Gen) | 128 (3rd Gen) | 6 GB | 264 GB/s | 12.8 | | 102.2 | | 4* | Yes* | Yes* | | | | | | | | | | | |
| | NVIDIA RTX A2000 | 2,560 | 20 (2nd Gen) | 80 (3rd Gen) | 4 GB | 192 GB/s | 9.3 | | 74.7 | | 4* | Yes* | Yes* | | | | | | | | | | | |
| | NVIDIA T1200 | 1,024 | | | 4 GB | 192 GB/s | 3.7 | | | | 4* | Yes* | Yes* | | | | | | | | | | | |
| | NVIDIA T600 | 896 | | | 4 GB | 160 GB/s | 2.5 | | | | 4* | Yes* | Yes* | | | | | | | | | | | |
| | NVIDIA T500 | 896 | | | 2 GB or 4 GB | 80 GB/s | 3.0 | | | | | | | | | | | | | | | | | |
| | Quadro RTX™ 6000 | 4,608 | 72 | 576 | 24 GB | 672 GB/s | 14.9 | | 119.4 | | 4* | Yes* | Yes* | | | | | | | | | | | |
| | Quadro RTX 5000 | 3,072 | 48 | 384 | 16 GB | 448 GB/s | 9.4 | | 75.2 | | 4* | Yes* | Yes* | | | | | | | | | | | |
| | Quadro RTX 4000 | 2,560 | 40 | 320 | 8 GB | 448 GB/s | 8 | | 63.9 | | 4* | Yes* | Yes* | | | | | | | | | | | |
| | Quadro RTX 3000 | 1,920 | 30 | 240 | 6 GB | 336 GB/s | 5.4 | | 42.9 | | 4* | Yes* | Yes* | | | | | | | | | | | |
| | Quadro T2000 | 1,024 | | | 4 GB | 128 GB/s | 3.5 | | | | 4* | Yes* | Yes* | | | | | | | | | | | |
| | Quadro T1000 | 896 | | | 4 GB | 128 GB/s | 2.6 | | | | 4* | Yes* | Yes* | | | | | | | | | | | |
| | Quadro P620 | 512 | | | 4 GB | 96 GB/s | 1.5 | | | | 4* | Yes* | Yes* | | | | | | | | | | | |
| | Quadro P520 | 384 | | | 2 GB or 4 GB | 48 GB/s | 1.1 | | | | | | | | | | | | | | | | | |

DESKTOP GPUs

| NEW | GPU Model | CUDA Cores | RT Cores | Tensor Cores | GPU Memory | Peak Memory Bandwidth | FP16 Performance (TFLOPS, Peak) | Accel. DP | Tensor Perf. (TFLOPS, Peak) | ECC | Max Displays | DP 1.2/1.4 | HDMI | SLI | USB-C | HDR | Mosaic | VR Ready | VR Shading | GPU Direct | Sync II | 3D Stereo | Encode/Decode | |
|-----|------------------|------------|--------------|---------------|------------|-----------------------|---------------------------------|-----------|-----------------------------|------|--------------|------------|------|-----|-------|-----|--------|----------|------------|------------|---------|-----------|---------------|--|
| | Quadro GV100 | 5,120 | | 640 | 32 GB | 870 GB/s | 14.8 | | 118.5 | •#11 | 4 | 4 | 4 | | | | | | | | | | | |
| | NVIDIA RTX A6000 | 10,752 | 84 (2nd Gen) | 336 (3rd Gen) | 48 GB | 768 GB/s | 38.7 | | 309.7 | •#12 | 4 | 4 | 4 | | | | | | | | | | | |
| | NVIDIA RTX A5000 | 8,192 | 64 (2nd Gen) | 256 (3rd Gen) | 24 GB | 768 GB/s | 27.8 | | 222.2 | •#12 | 4 | 4 | 4 | | | | | | | | | | | |
| | NVIDIA RTX A4000 | 6,144 | 48 (2nd Gen) | 192 (3rd Gen) | 16 GB | 448 GB/s | 19.2 | | 153.4 | •#12 | 4 | 4 | 4 | | | | | | | | | | | |
| | NVIDIA RTX A2000 | 3,328 | 26 (2nd Gen) | 104 (3rd Gen) | 6 GB | 288 GB/s | 8.0 | | 63.9 | •#12 | 4 | 4 | 4 | | | | | | | | | | | |
| | NVIDIA T1000 | 896 | | | 4 GB | 160 GB/s | 2.5 | | | | 4 | 4 | 4 | | | | | | | | | | | |
| | NVIDIA T600 | 640 | | | 4 GB | 160 GB/s | 1.7 | | | | 4 | 4 | 4 | | | | | | | | | | | |
| | NVIDIA T400 | 384 | | | 2 GB | 80 GB/s | 1.0 | | | | 4#13 | 3 | 3 | | | | | | | | | | | |
| | Quadro RTX 8000 | 4,608 | 72 | 576 | 48 GB | 672 GB/s | 16.3 | | 130.5 | •#12 | 4 | 4 | 4 | | | | | | | | | | | |
| | Quadro RTX 6000 | 4,608 | 72 | 576 | 24 GB | 672 GB/s | 16.3 | | 130.5 | •#12 | 4 | 4 | 4 | | | | | | | | | | | |
| | Quadro RTX 5000 | 3,072 | 48 | 384 | 16 GB | 448 GB/s | 11.2 | | 89.2 | •#12 | 4 | 4 | 4 | | | | | | | | | | | |
| | Quadro RTX 4000 | 2,304 | 36 | 288 | 8 GB | 416 GB/s | 7.1 | | 57 | | 4 | 3 | 4 | | | | | | | | | | | |
| | Quadro P1000 | 640 | | | 4 GB | 80 GB/s | 1.8 | | | | 4 | 4 | 4 | | | | | | | | | | | |
| | Quadro P620 | 512 | | | 2 GB | 80 GB/s | 1.3 | | | | 4 | 4 | 4 | | | | | | | | | | | |
| | Quadro P400 | 256 | | | 2 GB | 32 GB/s | 0.6 | | | | 4#13 | 3 | 3 | | | | | | | | | | | |

DATA CENTER GPUs

| NEW | GPU Model | CUDA Cores | RT Cores | Tensor Cores | GPU Memory | Peak Memory Bandwidth | FP16 Performance (TFLOPS, Peak) | Accel. DP | Tensor Perf. (TFLOPS, Peak) | ECC | Max Displays | DP 1.2/1.4 | HDMI | SLI | USB-C | HDR | Mosaic | VR Ready | VR Shading | GPU Direct | Sync II | 3D Stereo | Encode/Decode | |
|-----|--|------------|----------|--------------|------------|-----------------------|---------------------------------|-----------|-----------------------------|------|--------------|------------|------|-----|-------|-----|--------|----------|------------|------------|---------|-----------|---------------|--|
| | NVIDIA A40 | 10,752 | 84 | 336 | 48 GB | 696 GB/s | 37.4 | | 299.4 | • | 4 | 3 | 3 | | | | | | | | | | | |
| | NVIDIA A10 | 9,216 | 72 | 288 | 24 GB | 600 GB/s | 31.2 | | 249.9 | • | | | | | | | | | | | | | | |
| | NVIDIA A16 | 4x 1,280 | 4x 10 | 4x 40 | 4x 16GB | 4x 232 GB/s | 4x 4.45 | | 4x 17.82 | • | | | | | | | | | | | | | | |
| | Quadro RTX 8000 (Passive Thermal Solution) | 4,608 | 72 | 576 | 48 GB | 624 GB/s | 14.9 | | 119.4 | •#12 | | | | | | | | | | | | | | |
| | Quadro RTX 6000 (Passive Thermal Solution) | 4,608 | 72 | 576 | 24 GB | 624 GB/s | 14.9 | | 119.4 | •#12 | | | | | | | | | | | | | | |

* Check with OEM manufacturer for specific display topology.

1. CUDA parallel processing cores cannot be compared between GPU generations due to several important architectural differences that exist between streaming multiprocessor designs.

2. Peak rates are based on GPU Boost clock.

3. FP16 matrix multiply with FP16 or FP32 accumulate.

4. Effective TFLOPS using the Ampere sparsity feature.

5. NVIDIA Turing™, Volta™ and Pascal architectures support DP1.4. Adaptors available for DVI-SL, DVI-DL, HDMI, and VGA.

6. NVIDIA RTX Turing and Ampere architecture-based active desktop GPUs support display stream compression (DSC).

7. SLI functionality is provided via NVLink.

8. Supported adaptors are required for HDMI.

9. Supports multi-view rendering (MVR) feature.

10. For details on GPU-specific video encode/decode format support, refer to: <https://developer.nvidia.com/video-encode-and-decode-gpu-support-matrix-new>

11. Ensures data integrity and reliability by eliminating soft errors on both GPU cache and on-board DRAM.

12. Ensures data integrity and reliability by eliminating soft errors on direct random-access memory (DRAM) only.

13. NVIDIA T400 and Quadro P400 desktop GPUs can drive four displays via multi-stream transport (MST).

For more information on NVIDIA professional graphics solutions, visit: www.nvidia.com/en-us/design-visualization/rtx/

© 2021 NVIDIA Corporation. All rights reserved. NVIDIA and the NVIDIA logo, CUDA, GPUdirect, NVLink, Pascal, Quadro, Quadro RTX, RTX, SLI, Turing, and Volta are trademarks or registered trademarks of the respective owners with which they are associated. Features, pricing, availability, and specifications are subject to change without notice. AUG21

