# NVIDIA Professional Graphics Solutions

You need to do great things—create and collaborate from anywhere, on any device, without the distractions of slow performance, poor stability, or application incompatibility. With NVIDIA RTX™, you can unleash your vision and enjoy ultimate creative freedom.

NVIDIA RTX 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 to 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

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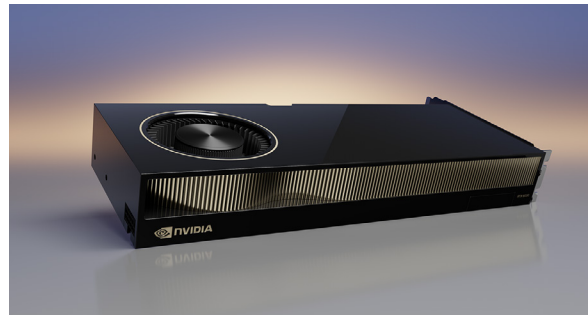
NVIDIA Desktop Workstation GPUs

NVIDIA RTX and Quadro-powered desktop workstations are designed and built specifically to drive the most challenging workloads of artists, designers, and engineers. Large GPU memory boosts application performance to accelerate your workflows and tackle the heaviest of workloads. NVIDIA RTX workstation solutions deliver significant business impact for demanding industries like manufacturing, media and entertainment, and energy.
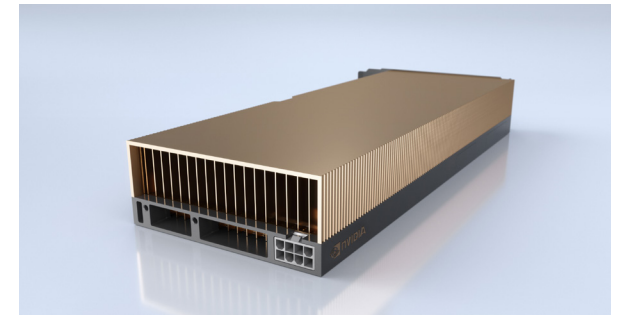
NVIDIA Data Center GPUs

Demand for visualization, rendering, data science, and simulation continues to grow as businesses tackle 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

New		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]	Error-Correcting Code (ECC) Memory	Maximum Active Displays	DisplayPort 1.4[4]	HDMI via Adaptors, HDMI	NVIDIA SLI[5]	High-Dynamic Range (HDR)[6]	NVIDIA Mosaic Technology	VR Ready[7]	Variable Rate Shading	GPUDirect for Video	Graphics Synchronization with Sync II	3D Stereo*	Encode/Decode[8]
Laptop GPUs																							
•	NVIDIA RTX 5000 Ada Generation	9,728	76 (3rd Gen)	304 (4th Gen)	16 GB	576 GB/s		42.6		681.8	•[9]	4*	Yes*	Yes*		•	•	•	•			•	•
•	NVIDIA RTX 4000 Ada Generation	7,424	58 (3rd Gen)	232 (4th Gen)	12 GB	432 GB/s		33.6		538.0	•[9]	4*	Yes*	Yes*		•	•	•	•			•	•
•	NVIDIA RTX 3500 Ada Generation	5,120	40 (3rd Gen)	160 (4th Gen)	12 GB	432 GB/s		23.0		368.6	•[9]	4*	Yes*	Yes*		•	•	•				•	•
•	NVIDIA RTX 3000 Ada Generation	4,608	36 (3rd Gen)	144 (4th Gen)	8 GB	256 GB/s		19.9		318.6	•[9]	4*	Yes*	Yes*		•	•	•				•	•
•	NVIDIA RTX 2000 Ada Generation	3,072	24 (3rd Gen)	96 (4th Gen)	8 GB	256 GB/s		14.5		231.6		4*	Yes*	Yes*		•	•	•				•	•
	NVIDIA RTX A5500	7,424	58 (2nd Gen)	232 (3rd Gen)	16 GB	512 GB/s		24.7		197.8	•[9]	4*	Yes*	Yes*		•	•	•	•			•	•
	NVIDIA RTX A4500	5,888	46 (2nd Gen)	184 (3rd Gen)	16 GB	512 GB/s		18.5		148.4	•[9]	4*	Yes*	Yes*		•	•	•	•			•	•
	NVIDIA RTX A3000 12GB	4,096	32 (2nd Gen)	128 (3rd Gen)	12 GB	336 GB/s		14.1		113.0	•[9]	4*	Yes*	Yes*		•	•	•	•			•	•
	NVIDIA RTX A2000 8GB	2,560	20 (2nd Gen)	80 (3rd Gen)	8 GB	224 GB/s		9.3		74.3		4*	Yes*	Yes*		•	•	•[10]	•			•	•
•	NVIDIA RTX A1000 6GB	2,560	20 (2nd Gen)	80 (3rd Gen)	6 GB	168 GB/s		9.3		74.6		4*	Yes*	Yes*		•	•	•[10]	•			•	•
	NVIDIA RTX A1000	2,048	16 (2nd Gen)	64 (3rd Gen)	4 GB	224 GB/s		7.5		59.7		4*	Yes*	Yes*		•	•	•[10]	•			•	•
	NVIDIA RTX A500	2,048	16 (2nd Gen)	64 (3rd Gen)	4 GB	112 GB/s		7.0		56.0		4*	Yes*	Yes*		•	•		•			•	•
	NVIDIA T600	896			4 GB	192 GB/s		3.0				4*	Yes*	Yes*		•	•					•	•
	NVIDIA T550	1,024			4 GB	112 GB/s		3.7															•
Desktop GPUs																							
•	NVIDIA RTX 6000 Ada Generation	18,176	142 (3rd Gen)	568 (4th Gen)	48 GB	960 GB/s		91.1		1,457.0	•[9]	4	4	4		•	•	•	•	•	•	•	•
•	NVIDIA RTX 4000 SFF Ada Generation	6,144	48 (3rd Gen)	192 (4th Gen)	20 GB	320 GB/s		19.2		306.8	•[9]	4	4	4		•	•	•	•	•	•	•	•
	NVIDIA RTX A6000	10,752	84 (2nd Gen)	336 (3rd Gen)	48 GB	768 GB/s	•	38.7		309.7	•[9]	4	4	4	•	•	•	•	•	•	•	•	•
	NVIDIA RTX A5500	10,240	80 (2nd Gen)	320 (3rd Gen)	24 GB	768 GB/s	•	34.1		272.8	•[9]	4	4	4	•	•	•	•	•	•	•	•	•
	NVIDIA RTX A5000	8,192	64 (2nd Gen)	256 (3rd Gen)	24 GB	768 GB/s	•	27.8		222.2	•[9]	4	4	4	•	•	•	•	•	•	•	•	•
	NVIDIA RTX A4500	7,168	56 (2nd Gen)	224 (3rd Gen)	20 GB	640 GB/s	•	23.7		189.2	•[9]	4	4	4	•	•	•	•	•	•	•	•	•
	NVIDIA RTX A4000	6,144	48 (2nd Gen)	192 (3rd Gen)	16 GB	448 GB/s		19.2		153.4	•[9]	4	4	4		•	•	•	•	•	•	•	•
	NVIDIA RTX A2000 | NVIDIA RTX A2000 12GB	3,328	26 (2nd Gen)	104 (3rd Gen)	6 GB or 12 GB	288 GB/s		7.9		63.9	•[9]	4	4	4		•	•	•	•				•
	NVIDIA T1000 | NVIDIA T1000 8GB	896			4 GB or 8 GB	160 GB/s		2.5				4	4	4		•	•					•	•
	NVIDIA T400 4GB	384			4 GB	80 GB/s		1.0				4[13]	3	3		•	•					•	•
	Quadro GV100	5,120		640	32 GB[11]	870 GB/s	•	14.8	•	118.5	•[12]	4	4	4	•	•	•	•		•		•	•
Data Center GPUs																							
•	NVIDIA L40	18,176	142 (3rd Gen)	568 (4th Gen)	48 GB	864 GB/s		90.5		724.1	•[9]	4	4	4	•	•	•	•	•	•	•	•	•
	NVIDIA A40	10,752	84	336	48 GB	696 GB/s	•	37.4		299.4	•[9]	4	3	3	•	•	•	•	•	•	•	•	•
	NVIDIA A10	9,216	72	288	24 GB	600 GB/s		31.2		249.9	•[9]							•		•			•
	NVIDIA A16	4x 1,280	4x 10	4x 40	4x 16 GB	4x 232 GB/s		4x 4.5		4x 17.8	•[9]							•					•

* Check with OEM manufacturer for specific display topology. Laptop GPU display support varies by system configuration. Check with your OEM system vendor to confirm which specification is supported.

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. Effective TFLOPS using the sparsity feature. NVIDIA Ada Lovelace architecture using FP8 matrix multiply with FP16 or FP32 accumulate; NVIDIA Ampere architecture using FP16 matrix multiply with FP16 or FP32 accumulate.
4. Feature support varies by system-level implementation. Check with your workstation OEM vendor for system specific configurations
5. SLI functionality is provided via NVLink.
6. Supported adaptors are required for HDMI.
7. Supports multi-view rendering (MVR) feature.
8. For more details on GPU-specific video encode/decode format support, refer to: **https://developer.nvidia.com/video-encode-and-decode-gpu-support-matrix-new**
9. Ensures data integrity and reliability by eliminating soft errors on direct random-access memory (DRAM) only.
10. Support for configuration at 60W TGP and above. NVIDIA RTX A1000 and RTX A2000 8GB Laptop GPUs, VR-Ready at 60W TGP and above.
11. NVIDIA GV100 uses HMB2 as its primary memory type.
12. Ensures data integrity and reliability by eliminating soft errors on both GPU cache and on-board DRAM.
13. NVIDIA T400 4GB 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**