NVIDIA PROFESSIONAL GRAPHICS SOLUTIONS

NVIDIA professional laptop GPUs power the world's most advanced thin and light mobile workstations and unique compact devices to meet the visual computing needs of professionals across a wide range of industries. The latest generation of NVIDIA RTX professional laptop GPUs, built on the NVIDIA Ampere architecture combine the latest advancements in real-time ray tracing, advanced shading, and Al-based capabilities to tackle the most demanding design and visualization workflows on the go. With the latest graphics technology, enhanced performance, and added compute power, NVIDIA professional laptop GPUs give designers, scientists, and artists the tools they need to work efficiently from anywhere.



	GPU
sing Core	SPECIFIC
K1 Cores Cores	CATIONS
emory	
lemory Bandwidth*	
у Туре	
y Interface	
ax Power Consumption²	
yPort³	
91-9	
Model	
×	
eneration	
Precision Floating-Point mance (TFLOPS, Peak)*	PERFORM
Performance (TFLOPS,	IANCE
IIDIA FXAA™ / TXAA™ Anti- asing	OPTION
RTX Desktop Manager	IS
Support	
0ptimus®	
n MAX-Q Technology	
C/NVDEC	

Laptop GPUs

Euptop of os																					
NVIDIA RTX A5000	6,144	48 (2nd Gen)	192 (3rd Gen)	16 GB	448 GB/s	GDDR6	256-bit	80 - 165 W*	1.4	4.6	7.0	12 Ultimate	4	21.7	174.0	J	1	1	1	/	1
NVIDIA RTX A4000	5,120	40 (2nd Gen)	160 (3rd Gen)	8 GB	384 GB/s	GDDR6	256-bit	80 - 140 W*	1.4	4.6	7.0	12 Ultimate	4	17.8	142.5	J	1	1	1	/	1
NVIDIA RTX A3000	4,096	32 (2nd Gen)	128 (3rd Gen)	6 GB	264 GB/s	GDDR6	192-bit	60-130 W*	1.4	4.6	7.0	12 Ultimate	4	12.8	102.2	1	1	1	1	/	J
NVIDIA RTX A2000	2,560	20 (2nd Gen)	80 (3rd Gen)	4 GB	192 GB/s	GDDR6	128-bit	35-95 W*	1.4	4.6	7.0	12 Ultimate	4	9.3	74.7	J	1	1	/	/	1
NVIDIA T1200	1,024			4 GB	192 GB/s	GDDR6	128-bit	35-95 W*	1.4	4.6	7.0	12.1	4	3.7		1	1	1	1	1	1
NVIDIA T600	896			4 GB	160 GB/s	GDDR6	128-bit	25 W (TDP)	1.4	4.6	7.0	12.1	4	2.5		J	J	1	1		1
NVIDIA T500	896			2 GB or 4 GB	80 GB/s	GDDR6	64-bit	18 - 25 W (TDP)		4.6	7.0	12.1	4	3.0		1	1	1	1		
Quadro RTX™ 6000	4,608	72	576	24 GB	672 GB/s	GDDR6	384-bit	250 W	1.4	4.6	5.1	12.1	3	14.9	119.4	J	1	1	1		1
Quadro RTX 5000	3,072	48	384	16 GB	448 GB/s	GDDR6	256-bit	80 - 110 W	1.4	4.6	5.1	12.1	3	9.4	75.2	1	1	1	1		1
Quadro RTX 4000	2,560	40	320	8 GB	448 GB/s	GDDR6	256-bit	80 - 110 W	1.4	4.6	5.1	12.1	3	8.0	63.9	J	1	1	1		1
Quadro RTX 3000	1,920	30	240	6 GB	336 GB/s	GDDR6	192-bit	60 - 80 W	1.4	4.6	5.1	12.1	3	5.4	42.9	J	1	1	1		1
Quadro T2000	1,024			4 GB	128 GB/s	GDDR6	128-bit	40 - 60 W	1.4	4.6	5.1	12.1	3	3.5		J	J	1	1		J
Quadro T1000	896			4 GB	128 GB/s	GDDR6	128-bit	40 - 50 W	1.4	4.6	5.1	12.1	3	2.6		J	J	1	1		1
Quadro P620	512			4 GB	96 GB/s	GDDR5	128-bit	25 W (TDP)	1.4	4.5	5.1	12.1	3	1.5		J	J	1	1		1
Quadro P520	384			2 GB or 4 GB	48 GB/s	GDDR5	64-bit	18 - 25 W (TDP)		4.5	5.1	12.1	3	1.1		J	1	1	1		

*Peak possible results vary by OEM system configuration. Check with your provider for system-specific information.

- 1. CUDA parallel processing cores cannot be compared between GPU generations due to several important architectural differences that exist between streaming multiprocessor designs.
- 2. Power is TDP-based for NVIDIA T600, T500, P620 and P520. GPUs with asterisked ranges include possible maximum power consumption including the Dynamic Boost 2.0 algorithm. For system specific GPU TGP, please consult your OEM/solution provider.
- 3. Adaptors available for DVI-SL, DVI-DL, HDMI, and VGA.
- 4. 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
- 5. F16 matrix multiply with FP16 or FP32 accumulate and effective TFLOPS using the sparsity feature for NVIDIA Ampere architecture-based GPUs.



