

SKY-QUAD-6000A-48
SKY-QUAD-5000A-32
SKY-QUAD-4500A-24
SKY-QUAD-4000A-20
SKY-QUAD-4000SA-20
SKY-QUAD-2000A-16

NVIDIA RTX 6000 Ada
NVIDIA RTX 5000 Ada
NVIDIA RTX 4500 Ada
NVIDIA RTX 4000 Ada
NVIDIA RTX 4000 SFF Ada
NVIDIA RTX 2000 Ada



Features

- NVIDIA Ada Lovelace GPU architecture
- Fourth-generation Tensor Cores
- Third-generation RT Cores
- GDDR6 memory with ECC
- AV1 Encode and Decode Support
- NVIDIA GPUDirect Remote Direct Memory Access (RDMA) support
- NVIDIA Mosaic Technology
- Graphics bus: PCI-E 4.0 x16
- Thermal solution: Active

Introduction

The new NVIDIA RTX Ada Generation series is the ultimate workstation graphics card designed for professionals who demand maximum performance and reliability to deliver their best work and breakthrough innovations across industries. Built on the NVIDIA Ada Lovelace architecture, which combines third-generation RT Cores, fourth-generation Tensor Cores, and new CUDA® cores with error correction code (ECC) graphics memory. All help to deliver the next generation of AI graphics and petaflop inferencing performance for unprecedented speed-up in rendering, AI, graphics, and compute workloads. NVIDIA RTX professional graphics cards are certified with a broad range of professional applications, tested by leading independent software vendors (ISVs), and backed by a global team of support specialists.

Specifications

Product Name	NVIDIA RTX 6000 Ada	NVIDIA RTX 5000 Ada	NVIDIA RTX 4500 Ada	NVIDIA RTX 4000 Ada	NVIDIA RTX 4000 SFF Ada	NVIDIA RTX 2000 Ada
Part Number	SKY-QUAD-6000A-48	SKY-QUAD-5000A-32	SKY-QUAD-4500A-24	SKY-QUAD-4000A-20	SKY-QUAD-4000SA-20	SKY-QUAD-2000A-16
GPU Memory	48 GB GDDR6 with ECC	32 GB GDDR6 with ECC	24 GB GDDR6 with ECC	20 GB GDDR6 with ECC	20 GB GDDR6 with ECC	16 GB GDDR6 with ECC
Memory Interface	384-bit	384-bit	320-bit	160-bit	160-bit	128-bit
Memory Bandwidth	960 GB/s	576 GB/s	432 GB/s	360 GB/s	320 GB/s	224 GB/s
NVIDIA CUDA Cores	18,176	12,800	7,680	6,144	6,144	2,816
Tensor Cores	568	400	240	192	192	88
RT Cores	142	100	60	48	48	22
Single-Precision Performance	91.1 TFLOPS	65.3 TFLOPS	39.6 TFLOPS	26.7 TFLOPS	19.2 TFLOPS	12.0 TFLOPS
System Interface	PCI Express 4.0 x16	PCI Express 4.0 x 16	PCI Express 4.0 x 16	PCI Express 4.0 x 16	PCI Express 4.0 x 16	PCI Express 4.0 x 8
Max Power Consumption	300 W	250 W	210W	130W	70 W	70W
Thermal Solution	Active	Active	Active	Active	Active	Active
Form Factor	4.4 inches (H) x 10.5 inches (L), dual slot, full height	4.4 inches (H) x 10.5 inches (L), dual slot, full height	4.4 inches (H) x 10.5 inches (L), dual slot, full height	4.4 inches (H) x 9.5 inches (L), single slot, full height	2.7 inches (H) x 6.6 inches (L), dual slot, low profile	2.7 inches (H) x 6.6 inches (L), dual slot, low profile
Display Connectors	4 x DisplayPort 1.4a	4 x DisplayPort 1.4a	4 x DisplayPort 1.4a	4 x DisplayPort 1.4a	4 x MDP 1.4a	4 x MDP 1.4a
Max Simultaneous Displays	4 x 4096 x 2160 @ 120 Hz 4 x 5120 x 2880 @ 60 Hz 2 x 7680 x 4320 @ 60 Hz	4 x 4096 x 2160 @ 120 Hz 4 x 5120 x 2880 @ 60 Hz 2 x 7680 x 4320 @ 60 Hz	4 x 4096 x 2160 @ 120 Hz 4 x 5120 x 2880 @ 60 Hz 2 x 7680 x 4320 @ 60 Hz	4 x 4096 x 2160 @ 120 Hz 4 x 5120 x 2880 @ 60 Hz 2 x 7680 x 4320 @ 60 Hz	4 x 4096 x 2160 @ 120 Hz 4 x 5120 x 2880 @ 60 Hz 2 x 7680 x 4320 @ 60 Hz	4 x 4096 x 2160 @ 120 Hz 4 x 5120 x 2880 @ 60 Hz 2 x 7680 x 4320 @ 60 Hz
Graphics APIs	DirectX 12.07 Shader Model 6.6 OpenGL 4.68 Vulkan 1.3	DirectX 12.07 Shader Model 6.6 OpenGL 4.68 Vulkan 1.3	DirectX 12.07 Shader Model 6.6 OpenGL 4.68 Vulkan 1.3	DirectX 12.07 Shader Model 6.6 OpenGL 4.68 Vulkan 1.3	DirectX 12.07 Shader Model 6.6 OpenGL 4.68 Vulkan 1.3	DirectX 12.07 Shader Model 6.6 OpenGL 4.68 Vulkan 1.3
Compute APIs	CUDA, DirectCompute, OpenCL™	CUDA, DirectCompute, OpenCL™	CUDA, DirectCompute, OpenCL™	CUDA, DirectCompute, OpenCL™	CUDA, DirectCompute, OpenCL™	CUDA, DirectCompute, OpenCL™