



NVIDIA GeForce RTX 4070 SUPER 12GB

Ada Lovelace PCIe 2024 4nm

OVERVIEW

35.5 FP32 TFLOPS	12 GB VRAM GDDR6X	— Memory Bandwidth	220 W TDP
----------------------------	-----------------------------	------------------------------	---------------------

PERFORMANCE METRICS

Precision	Peak (dense)	2:4 Sparse	Bits	Type
FP64	0.6	N/A	64	Standard
FP32	35.5	N/A	32	Standard
TF32	35.5	not published	32	Standard
BF16	71.0	not published	16	Standard
FP16	71.0	not published	16	Standard
FP8	141.9	not published	8	Standard
FP6	—	not published	—	Standard
FP4	—	not published	—	Standard
INT8	283.9	not published	8	Standard

MEMORY SPECIFICATIONS

Capacity	12 GB	TDP	220 W
Type	GDDR6X	Max Power	220 W
Bandwidth	—	FP32 Efficiency	0.16 TFLOPS/W
Interface Width	192-bit	FP16 Efficiency	0.32 TFLOPS/W

POWER & EFFICIENCY

HARDWARE DETAILS

Vendor: NVIDIA	Architecture: Ada Lovelace	Process Node: 4nm
Form Factor: PCIe	Launch Year: 2024	Variant: 12GB

COMPUTE ENGINE

CUDA Cores: 7,168	Tensor Cores: 224 (4th Gen)	SMs: 56
Base Clock: 1.98 GHz	Boost Clock: 2.48 GHz	

INTERCONNECT

PCIe: 4.0

CHIP DESIGN

Transistors: 35.8 billion Die Size: 295 mm²

ENTERPRISE FEATURES

Sparsity: Yes