



NVIDIA Quadro RTX 5000 16GB

Turing PCIe 2018 12nm

OVERVIEW

11.2 FP32 TFLOPS	16 GB VRAM GDDR6	448 GB/s Memory Bandwidth	265 W TDP
---------------------	---------------------	------------------------------	--------------

PERFORMANCE METRICS

Precision	Peak (dense)	Bits	Type
FP64	0.3	64	Standard
FP32	11.2	32	Standard
TF32	—	—	Standard
BF16	—	—	Standard
FP16	89.2	16	Standard
FP8	—	—	Standard
FP6	—	—	Standard
FP4	—	—	Standard
INT8	178.4	8	Standard
INT4	356.8	4	Standard

MEMORY SPECIFICATIONS

Capacity	16 GB	TDP	265 W
Type	GDDR6	Max Power	265 W
Bandwidth	448 GB/s	FP32 Efficiency	0.04 TFLOPS/W
Interface Width	256-bit	FP16 Efficiency	0.34 TFLOPS/W

POWER & EFFICIENCY

HARDWARE DETAILS

Vendor: NVIDIA	Architecture: Turing	Process Node: 12nm
Form Factor: PCIe	Launch Year: 2018	Variant: 16GB

COMPUTE ENGINE

CUDA Cores: 3,072	Tensor Cores: 384 (Turing)	SMs: 48
Base Clock: 1.62 GHz	Boost Clock: 1.81 GHz	

INTERCONNECT

GPU-to-GPU: NVLink	Bandwidth: 50 GB/s	Links: 2
PCIe: 3.0 x16		

CHIP DESIGN

Transistors: 13.6 billion	Die Size: 545 mm ²
---------------------------	-------------------------------

ENTERPRISE FEATURES

ECC: Yes	Sparsity: No
----------	--------------