

NVIDIA GeForce RTX 5070 Ti 16GB vs NVIDIA GB200 Grace Blackwell Superchip 372GB

Side by side datacenter GPU specifications from the Flopper.io database.

OVERVIEW

	NVIDIA GeForce RTX 5070 Ti 16GB	NVIDIA GB200 Grace Blackwell Superchip 372GB
FP32 TFLOPS	43.9	160.0
VRAM	16 GB	372 GB
Memory Bandwidth	896 GB/s	16.0 TB/s
TDP	300 W	—

PERFORMANCE BY PRECISION

Precision	NVIDIA GeForce RTX 5070 Ti 16GB peak TFLOPS (dense)	NVIDIA GB200 Grace Blackwell Superchip 372GB peak TFLOPS (dense)
FP64	0.7	80.0
FP32	43.9	160.0
TF32	43.9	2500.0
BF16	87.9	5000.0
FP16	175.8	5000.0
FP8	351.5	10000.0
FP6	—	10000.0
FP4	703.0	20000.0
INT8	351.5	10000.0

SPECIFICATIONS

	NVIDIA GeForce RTX 5070 Ti 16GB	NVIDIA GB200 Grace Blackwell Superchip 372GB
Architecture	Blackwell	Blackwell
Process Node	4nm	4nm
Launch Year	2025	2024
Form Factor	PCIe	Superchip

NVIDIA GeForce RTX 5070 Ti 16GB		NVIDIA GB200 Grace Blackwell Superchip 372GB
Memory Type	GDDR7	HBM3e
Interface Width	256-bit	—
Interconnect	—	NVLink-C2C
TDP	300 W	—
FP32 Efficiency	0.15 TFLOPS/W	—