

NVIDIA GeForce RTX 5070 12GB vs NVIDIA GeForce RTX 5090 32GB

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

OVERVIEW

	NVIDIA GeForce RTX 5070 12GB	NVIDIA GeForce RTX 5090 32GB
FP32 TFLOPS	30.9	104.8
VRAM	12 GB	32 GB
Memory Bandwidth	672 GB/s	1.8 TB/s
TDP	250 W	575 W

PERFORMANCE BY PRECISION

Precision	NVIDIA GeForce RTX 5070 12GB peak TFLOPS (dense)	NVIDIA GeForce RTX 5090 32GB peak TFLOPS (dense)
FP64	0.5	1.6
FP32	30.9	104.8
TF32	30.9	104.8
BF16	61.7	209.5
FP16	123.5	419.1
FP8	246.9	838.2
FP6	—	—
FP4	493.9	1676.0
INT8	246.9	838.2

SPECIFICATIONS

	NVIDIA GeForce RTX 5070 12GB	NVIDIA GeForce RTX 5090 32GB
Architecture	Blackwell	Blackwell
Process Node	4nm	4nm
Launch Year	2025	2025
Form Factor	PCIe	PCIe
Memory Type	GDDR7	GDDR7
Interface Width	192-bit	512-bit

	NVIDIA GeForce RTX 5070 12GB	NVIDIA GeForce RTX 5090 32GB
Interconnect	—	—
TDP	250 W	575 W
FP32 Efficiency	0.12 TFLOPS/W	0.18 TFLOPS/W