

NVIDIA GH200 144GB HBM3e vs NVIDIA GH200 96GB HBM3

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

OVERVIEW

	NVIDIA GH200 144GB HBM3e	NVIDIA GH200 96GB HBM3
FP32 TFLOPS	67.0	67.0
VRAM	144 GB	96 GB
Memory Bandwidth	4.9 TB/s	4.0 TB/s
TDP	1.0 kW	1.0 kW

PERFORMANCE BY PRECISION

Precision	NVIDIA GH200 144GB HBM3e peak TFLOPS (dense)	NVIDIA GH200 96GB HBM3 peak TFLOPS (dense)
FP64	34.0	34.0
FP64_TC Tensor Core	67.0	67.0
FP32	67.0	67.0
TF32	494.0	494.0
BF16	990.0	990.0
FP16	990.0	990.0
FP8	1979.0	1979.0
FP6	—	—
FP4	—	—
INT8	1979.0	1979.0

SPECIFICATIONS

	NVIDIA GH200 144GB HBM3e	NVIDIA GH200 96GB HBM3
Architecture	Hopper	Hopper
Process Node	4nm	4nm
Launch Year	2024	2023
Form Factor	Superchip	Superchip

	NVIDIA GH200 144GB HBM3e	NVIDIA GH200 96GB HBM3
Memory Type	HBM3e	HBM3
Interface Width	—	—
Interconnect	NVLink-C2C	NVLink-C2C
TDP	1.0 kW	1.0 kW
FP32 Efficiency	0.07 TFLOPS/W	0.07 TFLOPS/W