

NVIDIA GH200 96GB HBM3 vs NVIDIA GH200 144GB HBM3e

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

OVERVIEW

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

PERFORMANCE BY PRECISION

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

SPECIFICATIONS

	NVIDIA GH200 96GB HBM3	NVIDIA GH200 144GB HBM3e
Architecture	Hopper	Hopper
Process Node	4nm	4nm
Launch Year	2023	2024
Form Factor	SXM	SXM
Memory Type	HBM3	HBM3e
Interface Width	—	—
Interconnect	NVLink-C2C	NVLink-C2C

	NVIDIA GH200 96GB HBM3	NVIDIA GH200 144GB HBM3e
TDP	1.0 kW	1.0 kW
FP32 Efficiency	0.07 TFLOPS/W	0.07 TFLOPS/W