NVIDIA GB200 NVL72 GPU 186GB vs NVIDIA Rubin CPX

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

OVERVIEW

	NVIDIA GB200 NVL72 GPU 186GB	NVIDIA Rubin CPX
FP32 TFLOPS	**80.0**	–
Memory	**186 GB**	**128 GB**
Memory Bandwidth	**8.0 TB/s**	–
TDP	–	–

PERFORMANCE BY PRECISION

Precision	NVIDIA GB200 NVL72 GPU 186GB peak TFLOPS (dense)	NVIDIA Rubin CPX peak TFLOPS (dense)
FP64	40.0	–
FP32	80.0	–
TF32	1250.0	–
BF16	2500.0	–
FP16	2500.0	–
FP8	5000.0	–
FP6	5000.0	–
FP4	10000.0	–
NVFP4	–	–
INT8	5000.0	–

SPECIFICATIONS

	NVIDIA GB200 NVL72 GPU 186GB	NVIDIA Rubin CPX
Architecture	Blackwell	Rubin
Process Node	4nm	–
Launch Year	2024	2026
Form Factor	SXM	–
Memory Type	HBM3e	GDDR7
Interface Width	–	–

	NVIDIA GB200 NVL72 GPU 186GB	NVIDIA Rubin CPX
Interconnect	NVLink	–
TDP	–	–
FP32 Efficiency	–	–