



NVIDIA GB200 Grace Blackwell Superchip 372GB

Blackwell

Superchip

2024

4nm

OVERVIEW

160.0 FP32 TFLOPS	372 GB VRAM HBM3e	16.0 TB/s Memory Bandwidth	— TDP
-----------------------------	-----------------------------	--------------------------------------	-----------------

PERFORMANCE METRICS

Precision	Peak TFLOPS	Bits	Type
FP6	10000.0	—	Standard
INT8	10000.0	8	Standard
FP8	10000.0	8	Standard
FP4	10000.0	4	Standard
FP16	5000.0	16	Standard
BF16	5000.0	16	Standard
TF32	2500.0	32	Standard
FP32	160.0	32	Standard
FP64	80.0	64	Standard

MEMORY SPECIFICATIONS

Capacity	372 GB	TDP	—
Type	HBM3e	Est. Max Power	—
Bandwidth	16.0 TB/s	FP32 Efficiency	—
Interface Width	—	FP16 Efficiency	—

POWER & EFFICIENCY

HARDWARE DETAILS

Vendor: NVIDIA	Architecture: Blackwell	Process Node: 4nm
Form Factor: Superchip	Launch Year: 2024	Variant: Grace Blackwell Superchip 372GB

INTERCONNECT

GPU-to-GPU: NVLink-C2C	Bandwidth: 900 GB/s
------------------------	---------------------

ENTERPRISE FEATURES

Sparsity: Yes
