



NVIDIA HGX B100

8x B100

Blackwell

baseboard

2024

OVERVIEW

28.0 FP8 PFLOPS Dense	1.5 TB Aggregate Memory	64.0 TB/s Aggregate Bandwidth	5.6 kW Aggregate Power
---------------------------------	-----------------------------------	---	----------------------------------

PERFORMANCE METRICS

Precision	Peak (PFLOPS)	Per GPU (TFLOPS)	TFLOPS/W
FP64	0.2	30.0	—
FP32	0.5	60.0	—
TF32	7.0	875.0	—
FP16	14.0	1750.0	—
BF16	14.0	1750.0	—
FP8	28.0	3500.0	—
FP4	56.0	7000.0	—
INT8	28.0	3500.0	—

All figures are dense. Vendors commonly headline the sparse number, which is twice the dense one.

CONFIGURATION

Accelerator	B100 SXM 192GB
Count	8
Memory per GPU	192 GB
Bandwidth per GPU	8.0 TB/s

POWER & EFFICIENCY

Aggregate Power	5.6 kW
TDP per GPU	700 W
FP8 Efficiency	5.00 TFLOPS/W
Aggregate Bandwidth	64.0 TB/s

SYSTEM DETAILS

Vendor: NVIDIA	Family: HGX	Architecture: Blackwell
Form Factor: baseboard	Release Year: 2024	Process Node: 4nm

INTERCONNECT

NVSwitch + NVLink