

NVIDIA A100 SXM4 40GB vs NVIDIA A2 16GB

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

OVERVIEW

	NVIDIA A100 SXM4 40GB	NVIDIA A2 16GB
FP32 TFLOPS	19.5	4.5
Memory	40 GB	16 GB
Memory Bandwidth	1.6 TB/s	200 GB/s
TDP	400 W	60 W

PERFORMANCE BY PRECISION

Precision	NVIDIA A100 SXM4 40GB peak TFLOPS (dense)	NVIDIA A2 16GB peak TFLOPS (dense)
FP64	9.7	—
FP32	19.5	4.5
TF32	156.0	9.0
BF16	312.0	18.0
FP16	312.0	18.0
FP8	—	—
FP6	—	—
FP4	—	—
INT8	624.0	36.0
INT4	1248.0	72.0

SPECIFICATIONS

	NVIDIA A100 SXM4 40GB	NVIDIA A2 16GB
Architecture	Ampere	Ampere
Process Node	7nm	—
Launch Year	2020	2021
Form Factor	SXM	PCIe
Memory Type	HBM2e	GDDR6
Interface Width	4096-bit	128-bit

	NVIDIA A100 SXM4 40GB	NVIDIA A2 16GB
Interconnect	—	—
TDP	400 W	60 W
FP32 Efficiency	0.05 TFLOPS/W	0.07 TFLOPS/W