NVIDIA A100 SXM4 40GB vs NVIDIA GeForce RTX 3080 Ti 12GB

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

OVERVIEW

	NVIDIA A100 SXM4 40GB	NVIDIA GeForce RTX 3080 Ti 12GB
FP32 TFLOPS	**19.5**	**34.1**
Memory	**40 GB**	**12 GB**
Memory Bandwidth	**1.6 TB/s**	**912 GB/s**
TDP	**400 W**	**350 W**

PERFORMANCE BY PRECISION

Precision	NVIDIA A100 SXM4 40GB peak TFLOPS (dense)	NVIDIA GeForce RTX 3080 Ti 12GB peak TFLOPS (dense)
FP64	9.7	0.5
FP32	19.5	34.1
TF32	156.0	34.1
BF16	312.0	68.2
FP16	312.0	68.2
FP8	—	—
FP6	—	—
FP4	—	—
INT8	624.0	272.8
INT4	1248.0	545.6

SPECIFICATIONS

	NVIDIA A100 SXM4 40GB	NVIDIA GeForce RTX 3080 Ti 12GB
Architecture	Ampere	Ampere
Process Node	7nm	8nm
Launch Year	2020	2021
Form Factor	SXM	PCIe
Memory Type	HBM2e	GDDR6X
Interface Width	4096-bit	384-bit

	NVIDIA A100 SXM4 40GB	NVIDIA GeForce RTX 3080 Ti 12GB
Interconnect	–	–
TDP	400 W	350 W
FP32 Efficiency	0.05 TFLOPS/W	0.10 TFLOPS/W