# NVIDIA A100 SXM4 40GB vs NVIDIA GeForce RTX 3070 8GB

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

## OVERVIEW

| | NVIDIA A100 SXM4 40GB | NVIDIA GeForce RTX 3070 8GB |
|---|---|---|
| FP32 TFLOPS | **19.5** | **20.3** |
| Memory | **40 GB** | **8 GB** |
| Memory Bandwidth | **1.6 TB/s** | **448 GB/s** |
| TDP | **400 W** | **220 W** |

## PERFORMANCE BY PRECISION

| Precision | NVIDIA A100 SXM4 40GB peak TFLOPS (dense) | NVIDIA GeForce RTX 3070 8GB peak TFLOPS (dense) |
|---|---|---|
| FP64 | 9.7 | 0.3 |
| FP32 | 19.5 | 20.3 |
| TF32 | 156.0 | 20.3 |
| BF16 | 312.0 | 40.6 |
| FP16 | 312.0 | 40.6 |
| FP8 | – | – |
| FP6 | – | – |
| FP4 | – | – |
| INT8 | 624.0 | 162.6 |
| INT4 | 1248.0 | 325.2 |

## SPECIFICATIONS

| | NVIDIA A100 SXM4 40GB | NVIDIA GeForce RTX 3070 8GB |
|---|---|---|
| Architecture | Ampere | Ampere |
| Process Node | 7nm | 8nm |
| Launch Year | 2020 | 2020 |
| Form Factor | SXM | PCIe |
| Memory Type | HBM2e | GDDR6 |
| Interface Width | 4096-bit | 256-bit |

| | NVIDIA A100 SXM4 40GB | NVIDIA GeForce RTX 3070 8GB |
|---|---|---|
| Interconnect | – | – |
| TDP | 400 W | 220 W |
| FP32 Efficiency | 0.05 TFLOPS/W | 0.09 TFLOPS/W |