

Google TPU v4 32GB vs Google TPU v6e 32GB

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

OVERVIEW

	Google TPU v4 32GB	Google TPU v6e 32GB
FP32 TFLOPS	—	—
VRAM	32 GB	32 GB
Memory Bandwidth	1.2 TB/s	1.6 TB/s
TDP	—	—

PERFORMANCE BY PRECISION

Precision	Google TPU v4 32GB peak TFLOPS (dense)	Google TPU v6e 32GB peak TFLOPS (dense)
INT8	275.0	1836.0
BF16	275.0	918.0

SPECIFICATIONS

	Google TPU v4 32GB	Google TPU v6e 32GB
Architecture	TPU	TPU
Process Node	—	—
Launch Year	2020	2024
Form Factor	—	—
Memory Type	HBM2	HBM
Interface Width	—	—
Interconnect	—	—
TDP	—	—
FP32 Efficiency	—	—