



# Google TPU v4 32GB vs Google TPU v5e 16GB

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

## OVERVIEW

|                  | Google TPU v4 32GB | Google TPU v5e 16GB |
|------------------|--------------------|---------------------|
| FP32 TFLOPS      | —                  | —                   |
| VRAM             | 32 GB              | 16 GB               |
| Memory Bandwidth | 1.2 TB/s           | 800 GB/s            |
| TDP              | —                  | —                   |

## PERFORMANCE BY PRECISION

| Precision | Google TPU v4 32GB peak TFLOPS (dense) | Google TPU v5e 16GB peak TFLOPS (dense) |
|-----------|----------------------------------------|-----------------------------------------|
| INT8      | 275.0                                  | 393.0                                   |
| BF16      | 275.0                                  | 197.0                                   |

## SPECIFICATIONS

|                 | Google TPU v4 32GB | Google TPU v5e 16GB |
|-----------------|--------------------|---------------------|
| Architecture    | TPU                | TPU                 |
| Process Node    | —                  | —                   |
| Launch Year     | 2020               | 2023                |
| Form Factor     | —                  | —                   |
| Memory Type     | HBM2               | HBM                 |
| Interface Width | —                  | —                   |
| Interconnect    | —                  | —                   |
| TDP             | —                  | —                   |
| FP32 Efficiency | —                  | —                   |