Huawei Ascend 960 vs NVIDIA GH200 144GB HBM3e

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

OVERVIEW

	Huawei Ascend 960	NVIDIA GH200 144GB HBM3e
FP32 TFLOPS	—	**67.0**
Memory	**288 GB**	**144 GB**
Memory Bandwidth	**9.6 TB/s**	**4.9 TB/s**
TDP	—	**1.0 kW**

PERFORMANCE BY PRECISION

Precision	Huawei Ascend 960 peak TFLOPS (dense)	NVIDIA GH200 144GB HBM3e peak TFLOPS (dense)
FP64	—	34.0
FP64_TC Tensor Core	—	67.0
FP32	—	67.0
TF32	—	494.0
BF16	1000.0	990.0
FP16	1000.0	990.0
FP8	2000.0	1979.0
FP6	—	—
FP4	4000.0	—
INT8	—	1979.0

SPECIFICATIONS

	Huawei Ascend 960	NVIDIA GH200 144GB HBM3e
Architecture	Da Vinci v4 (chiplet)	Hopper
Process Node	—	4nm
Launch Year	2027	2024
Form Factor	OAM	Superchip
Memory Type	—	HBM3e

	Huawei Ascend 960	NVIDIA GH200 144GB HBM3e
Interface Width	—	—
Interconnect	Unified Bus 2.0	NVLink-C2C
TDP	—	1.0 kW
FP32 Efficiency	—	0.07 TFLOPS/W