

NVIDIA A100 80GB GPU 卡参数

生产厂家	NVIDIA（原厂，非改装卡）
GPU 型号	NVIDIA A100 80GB PCIe
CUDA Cores	6912
Streaming Multiprocessors	108
Tensor Cores Gen 3	432
GPU Memory	80GB HBM2e ECC on by Default
Memory Interface	5120-bit
Memory Bandwidth	1555 GB/s
NVLink	2-Way, 2-Slot, 600 GB/s Bidirectional
MIG (Multi-Instance GPU) Support	Yes, up to 7 GPU Instances
FP64	9.7 TFLOPS
FP64 Tensor Core	19.5 TFLOPS
FP32	19.5 TFLOPS
TF32 Tensor Core	156 TFLOPS 312 TFLOPS*
BFLOAT16 Tensor Core	312 TFLOPS 624 TFLOPS*
FP16 Tensor Core	312 TFLOPS 624 TFLOPS*
INT8 Tensor Core	624 TOPS 1248 TOPS*
INT4 Tensor Core	1248 TOPS 2496 TOPS*
Thermal Solutions	Passive
vGPU Support	NVIDIA Virtual Compute Server (vCS)
System Interface	PCIE 4.0 x16
Maximum Power Consumption	300 W
GPU Power Cable	兼容 NVIDIA A100 80G 的专用供电线

* 带 * 的数值为启用稀疏化（Sparsity）后的性能。