

SKY-TESL-H100

NVIDIA® Tesla® H100



Features

- NVIDIA Hopper GPU architecture
- Compute-optimized GPU
- 14592 NVIDIA® CUDA® Cores
- 456 NVIDIA® Tensor Cores
- 80GB HBM2e memory with ECC
- Up to 2TB/s memory bandwidth
- Max. power consumption: 350W
- Graphics bus: PCI-E 5.0 x16
- Thermal solution: Passive

Introduction

NVIDIA® Tesla® H100 (SKY-TESL-H100-80P) PCIe card is compute-optimized GPU built on the NVIDIA Hopper architecture with dual-slot 10.5-inch PCI Express Gen5 interface in a passive heatsink cooling design suitable for data centers. Combining NVIDIA Gen4 tensor cores and HBM2e memory, they provide a high-performance computing solution. Supporting a Gen2 multi-instance GPU (MIG) feature, which guarantees quality of service (QoS) with secure, partitioned hardware, they allow maximum utilization of GPU resources. The NVIDIA NGC™ catalog provides software, libraries, and optimized AI models and applications to complete data center solutions. With cutting-edge features and technologies, NVIDIA Tesla H100 is perfect for AI deep learning training and inference, data analytics, and high-performance computing (HPC) applications. NVIDIA Tesla is the first choice for high-standard computing solutions in enterprise and science deployments.

Specifications

Product Name	Tesla H100
Part Number	SKY-TESL-H100-80P
GPU Architecture	Hopper
GPU Memory	80GB HBM2e
Memory Bandwidth	2TB/s
NVIDIA CUDA Cores	14592
Tensor Cores	456
Single-Precision Performance	48 TFLOPS
Double-Precision Performance	24 TFLOPS
Fast FP64	Yes
System Interface	PCI Express 5.0 x16
Max Power Consumption	350W
Power Connector	16-Pin PCIe
Thermal Solution	Passive
Multi-Instance GPU	Up to 7
Form Factor	4.4 inches H x 10.5 inches L dual slot, full height
NVLink Support	3 NVLINK Bridges for 2 GPUs, 600GB/s
Media Acceleration	7 JPEG Decoder, 7 Video Decoder
Display Connectors	Headless Design