

# SKY-TESL-H100-80P

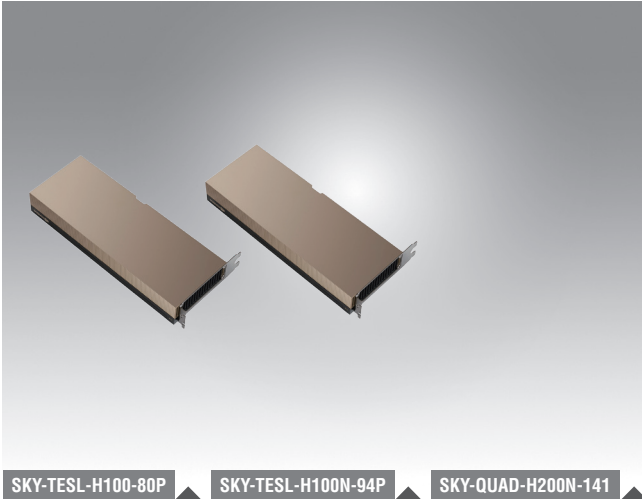
# SKY-TESL-H100N-94P

# SKY-QUAD-H200N-141

NVIDIA® Tesla® H100

NVIDIA® Tesla® H100 NVL

NVIDIA® Tesla® H200 NVL



SKY-TESL-H100-80P

SKY-TESL-H100N-94P

SKY-QUAD-H200N-141

## Features

- NVIDIA Hopper GPU architecture
- Compute-optimized GPU
- 14592 / 14592 NVIDIA® CUDA® Cores
- 456 / 456 NVIDIA® Tensor Cores
- 80GB HBM2e / 94GB HBM3 memory with ECC
- Up to 2TB/s / 3.9TB/s memory bandwidth
- Max. power consumption: 350W / 400W
- Graphics bus: PCI-E 5.0 x16
- Thermal solution: Passive

## Introduction

NVIDIA® Tesla® H100 (SKY-TESL-H100-80P) and NVIDIA® Tesla® H100 NVL (SKY-TESL-H100N-94P) PCIe Card are 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 / HBM3 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 and H100 NVL are 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	Tesla H100 NVL	Tesla H200 NVL
Part Number	SKY-TESL-H100-80P	SKY-TESL-H100N-94P	SKY-QUAD-H200N-141
GPU Architecture	Hopper	Hopper	Hopper
GPU Memory	80GB HBM2e	94GB HBM3	141GB HBM3e
Memory Bandwidth	2TB/s	3.9TB/s	4.8TB/s
NVIDIA CUDA Cores	14592	14592	14592
Tensor Cores	456	456	456
Single-Precision Performance	48 TFLOPS	67 TFLOPS	60 TFLOPS
Double-Precision Performance	24 TFLOPS	34 TFLOPS	30 TFLOPS
Fast FP64	Yes	Yes	Yes
System Interface	PCI Express 5.0 x16	PCI Express 5.0 x16	PCI Express 5.0 x16
Max Power Consumption	350W	400W	600W
Power Connector	16-Pin PCIe	16-Pin PCIe	16-Pin PCIe
Thermal Solution	Passive	Passive	Passive
Multi-Instance GPU	Up to 7	Up to 7	Up to 7
Form Factor	4.4 inches H x 10.5 inches L dual slot, full height	4.4 inches H x 10.5 inches L dual slot, full height	4.4 inches H x 10.5 inches L dual slot, full height
NVLink Support	3 NVLINK Bridges for 2 GPUs, 600GB/s	3 NVLINK Bridges for 2 GPUs, 600GB/s	2 / 4- way NVIDIA NVLink bridge, 900GB/s
Media Acceleration	7 JPEG Decoder, 7 Video Decoder	7 JPEG Decoder, 7 Video Decoder	7 JPEG Decoder, 7 Video Decoder
Display Connectors	Headless Design	Headless Design	Headless Design