NVIDIA H100 NVL GPU

Product Brief

PB-11773-001_v01 | March 2024

Document History

PB-11773-001_v01

Version	Date	Authors	Description of Change
01	March 14, 2024	VNK, SM	Initial release

Table of Contents

List of Figures

List of Tables

Overview

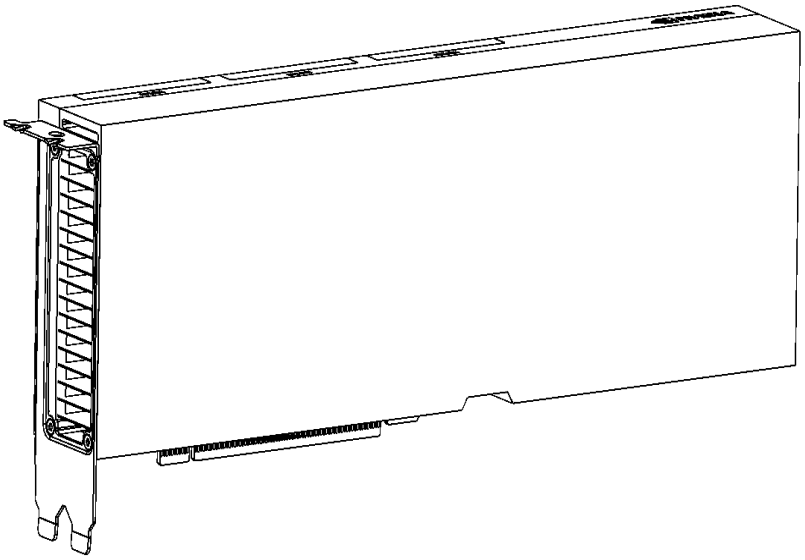
The NVIDIA® H100 NVL Tensor Core GPU is the most optimized platform for LLM Inferences with its high compute density, high memory bandwidth, high energy efficiency, and unique NVLink architecture. It also delivers unprecedented acceleration to power the world's highest-performing elastic data centers for AI, data analytics, and high-performance computing (HPC) applications. NVIDIA H100 NVL Tensor Core technology supports a broad range of math precisions, providing a single accelerator for every compute workload. The NVIDIA H100 NVL supports double precision (FP64), single-precision (FP32), half precision (FP16), 8-bit floating point (FP8), and integer (INT8) compute tasks.

The NVIDIA H100 NVL card is a dual-slot 10.5 inch PCI Express Gen5 card based on the NVIDIA Hopper™ architecture. It uses a passive heat sink for cooling, which requires system airflow to operate the card properly within its thermal limits. The NVIDIA H100 NVL operates unconstrained up to its maximum thermal design power (TDP) level of 400 W to accelerate applications that require the fastest computational speed and highest data throughput. The NVIDIA H100 NVL debuts the world's highest PCIe card memory bandwidth of nearly 4,000 gigabytes per second (GBps). This speeds time to solution for the largest models and most massive data sets.

The NVIDIA H100 NVL card features Multi-Instance GPU (MIG) capability. This can be used to partition the GPU into as many as seven hardware isolated GPU instances, providing a unified platform that enables elastic data centers to adjust dynamically to shifting workload demands. As well as it can allocate the right size of resources from the smallest to biggest multi-GPU jobs. NVIDIA H100 NVL versatility means that IT managers can maximize the utility of every graphics processing unit (GPU) in their data center.

NVIDIA H100 NVL cards use three NVIDIA® NVLink® bridges. They are the same as the one used with NVIDIA H100 PCIe cards. This allows two NVIDIA H100 PCIe cards to be connected to deliver 600 GB/s bidirectional bandwidth or 10x the bandwidth of PCIe Gen4, to maximize application performance for large workloads.

Figure 1. NVIDIA H100 NVL with NVLink Bridge Volumetric



Specifications

Product Specifications

Table 1 through Table 3 provide the product, memory, and software specifications for the NVIDIA H100 NVL card.

Table 1. Product Specifications

Specification	NVIDIA H100 NVL
Product SKU	P1010 SKU 210 NVPN: 699-21010-0210-xxx
Total board power	PCIe 16-pin cable strapped for 450 W or 600 W power mode: > 400 W maximum (default) > 310 W power compliance limit[1] > 200 W minimum PCIe 16-pin cable strapped for 300 W power mode: > 310 W maximum (default) > 310 W power compliance limit > 200 W minimum
Thermal solution	Passive
Mechanical form factor	Full-height, full-length (FHFL) 10.5”, dual-slot
PCI Device IDs	Device ID: 0x2321 Vendor ID: 0x10DE Sub-Vendor ID: 0x10DE Sub-System ID: 0x1839
GPU clocks	Base: 1,080 MHz Boost: 1,785 MHz
Performance states	P0
VBIOS	EEPROM size: 8 Mbit UEFI: Not supported
PCI Express interface	PCI Express Gen5 x16; Gen5 x8; Gen4 x16 Lane and polarity reversal supported
Multi-Instance GPU (MIG)	Supported (seven instances)

Specification	NVIDIA H100 NVL
Secure Boot (CEC)	Supported
Zero Power	Not supported
Power connectors and headers	One PCIe 16-pin auxiliary power connector (12v-2x6 auxiliary power connector)
Weight	Board: 1,214 grams (excluding bracket, extenders, and bridges) NVLink Bridge: 20.5 grams per bridge (x 3 bridges) Bracket with screws: 20 grams Enhanced straight extender: 35 grams Long offset extender: 48 grams Straight extender: 32 grams

Table 2. Memory Specifications

Specification	Description
Memory clock	2,619 MHz
Memory type	HBM3
Memory size	94 GB
Memory bus width	6,016 bits
Peak memory bandwidth	3,938 GB/s

Table 3. Software Specifications

Specification	Description[1]
SR-IOV support	Supported - 32 VF (virtual functions)
BAR address (physical function)	BAR0: 16 MiB[1] BAR2: 128 GiB[1] BAR4: 32 MiB[1]
BAR address (virtual function)	BAR0: 8 MiB, (256 KiB per VF)[1] BAR1: 128 GiB, 64-bit (4 GiB per VF)[1] BAR3: 1 GiB, 64-bit (32 MiB per VF)[1]
Message signaled interrupts	MSI-X: Supported MSI: Not supported
ARI Forwarding	Supported
Driver support	Linux: R535 or later Windows: R535 or later
Secure boot	Supported
CEC firmware	Version 00.02.0134.0000 or later
NVFlash	Version 5.816.0 or later
NVIDIA® CUDA® support	x86: CUDA 12.2 or later

Specification	Description[1]
Virtual GPU software support	Supports vGPU 16.1 or later: NVIDIA Virtual Compute Server Edition
NVIDIA AI Enterprise	Supported with VMWare
NVIDIA certification	NVIDIA-Certified Systems™ 2.8 or later
PCI class code	0x03 – Display controller
PCI subclass code	0x02 – 3D controller
ECC support	Enabled
SMBus (8-bit address)	0x9E (write), 0x9F (read)
IPMI FRU EEPROM I2C address	0x50 (7-bit), 0xA0 (8-bit)
Reserved I2C adresses[2]	0xAA, 0xAC, 0xA0
SMBus direct access	Supported
SMBPBI SMBus Post-Box Interface)	Supported

Note:

[1]The KiB, MiB, and GiB notations emphasize the "power of two" nature of the values. Thus,

- **256 KiB = 256 × 1024**
- **16 MiB = 16 × 1024^2**
- **64 GiB = 64 × 1024^3**

[2]See the "CEC Hardware Root of Trust" section in this product brief.

The operator is given the option to configure this power setting to be persistent across driver reloads or to revert to default power settings upon driver unload.

Environmental and Reliability Specifications

Table 4 provides the environment conditions specifications for the NVIDIA H100 NVL card.

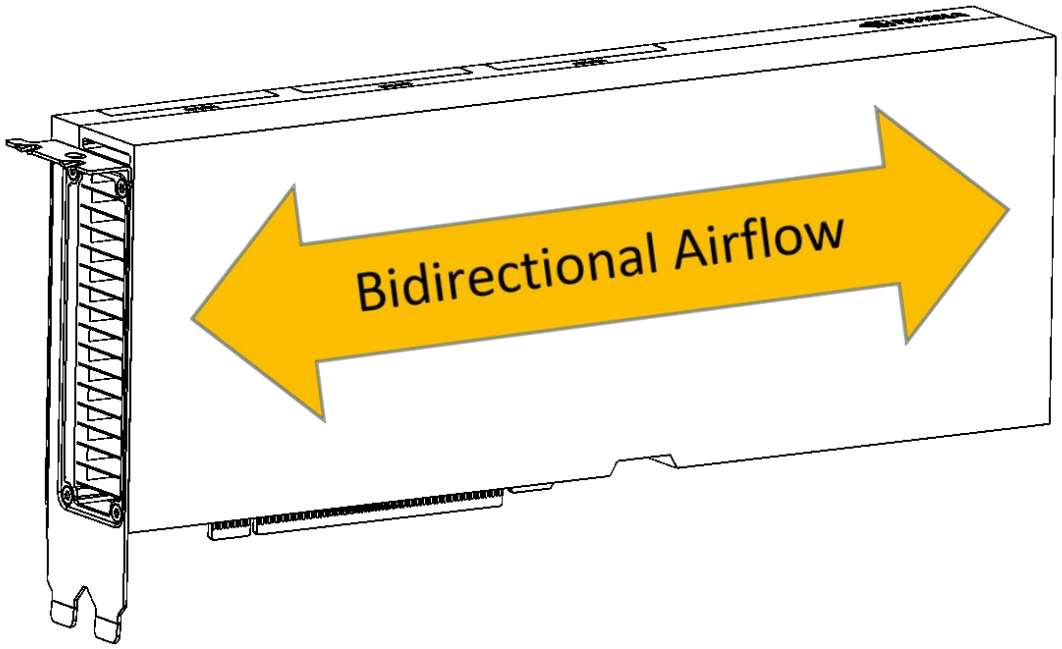
Table 4. Board Environmental and Reliability Specifications

Specification	Description
Ambient operating temperature	0°C to 50°C
Ambient operating temperature (short term)	-5°C to 55°C
Storage temperature	-40°C to 75°C
Operating humidity (short term)	5% to 93% relative humidity
Operating humidity	5% to 85% relative humidity
Storage humidity	5% to 95% relative humidity
Mean time between failures (MTBF)	TBD

Airflow Direction Support

The NVIDIA H100 NVL card employs a bidirectional heat sink, which accepts airflow either left-to-right or right-to-left directions.

Figure 2. H100 NVL Airflow Directions



Product Features

PCI Express Interface Specifications

The following subsections describe the PCIe interface specifications for the NVIDIA H100 NVL card.

PCIe Support

The NVIDIA H100 NVL GPU card supports PCIe Gen5. Either a Gen5 x16, Gen5 x8, or Gen4 x16 interface should be used when connecting to the NVIDIA H100 NVL card.

Single Root I/O Virtualization Support

Single Root I/O Virtualization (SR-IOV) is a PCIe specification that allows a physical PCIe device to appear as multiple physical PCIe devices. Per PCIe specification, each device can have up to a maximum of 256 virtual functions (VFs). The actual number can depend on the device. SR-IOV is enabled in an NVIDIA H100 NVL card with 32 VFs supported.

For each device, SR-IOV identifies two function classes:

- Physical functions (PFs) constitute full-featured functionality. They are fully configurable, and their configuration can control the entire device. Naturally, a PF also has full ability to move data in and out of the device.
- Virtual functions (VFs), which lack configuration resources. VFs exist on an underlying PF, which may support many such VFs. VFs can only move data in and out of the device. They cannot be configured and cannot be treated like a full PCIe device. The OS or hypervisor instance must be aware that they are not full PCIe devices.

The NVIDIA H100 NVL requires that SBIOS and software support in the operating system (OS) instance or hypervisor is configured to enable support for SR-IOV. The OS instance or hypervisor must be able to detect and initialize PFs and VFs.

Interrupt Messaging

The NVIDIA H100 NVL card only supports the MSI-X interrupt messaging protocol. The MSI interrupt protocol is not supported.

CEC Hardware Root of Trust

The NVIDIA H100 NVL provides secure boot capability using CEC. Implementing code authentication, rollback protection and key revocation, the CEC device authenticates the contents of the GPU firmware ROM before permitting the GPU to boot from its ROM.

It also provides out-of-band (OOB) secure firmware update, secure application processor recovery, and remote attestation.

The Hardware Root of Trust feature occupies up to two I2C addresses (in addition to the SMBus addresses). I2C addresses 0xAA and 0xAC should therefore be avoided for system use.

Multi-Instance GPU Support

The NVIDIA H100 NVL card supports Multi-Instance GPU (MIG) capability by providing up to seven GPU instances per NVIDIA H100 NVL GPU. MIG technology can partition the NVIDIA H100 NVL GPU into individual instances, each fully isolated with its own high-bandwidth memory, cache, and compute cores, enabling optimized computational resource provisioning and quality of service (QoS).

For detailed information on MIG provisioning and use, consult the *Multi-Instance GPU User's Guide*: https://docs.nvidia.com/datacenter/tesla/mig-user-guide/index.html

Programmable Power

The Programmable Power feature provides partners the general ability to configure the power cap of the card for system power and thermal budget or performance-per-watt reasons.

The power cap can be modified using either of these two NVIDIA tools:

- In-band: nvidia-smi (power cap adjustment must be reestablished after each new driver load)
- Out-of-band: SMBPBI (power cap adjustment remains in force across driver loads and system boots)

Power limit specifications for the NVIDIA H100 NVL are presented in Table 1.

nvidia-smi

nvidia-smi is an in-band monitoring tool provided with the NVIDIA driver and can be used to set the maximum power consumption with driver running in persistence mode. An example command to reduce the power cap to 310 W is shown:

```
nvidia-smi -pm 1
nvidia-smi -pl 310
```

To restore the NVIDIA H100 NVL back to its default TDP power consumption, either the driver module can be unloaded and reloaded, or the following command can be issued:

```
nvidia-smi -pl <TDP>
```

SMBPBI

An out-of-band channel exists through the SMBus Post-Box Interface (SMBPBI) protocol to set the power limit of the GPU. This also requires that the NVIDIA driver is loaded for full functionality. The power cap can be adjusted through the following asynchronous command:

Table 5. SMBPBI Commands

Specification	Value
Opcode	10h – Submit/poll asynchronous request
Arg1	0x01 – Set total GPU power limit
Arg2	0x00

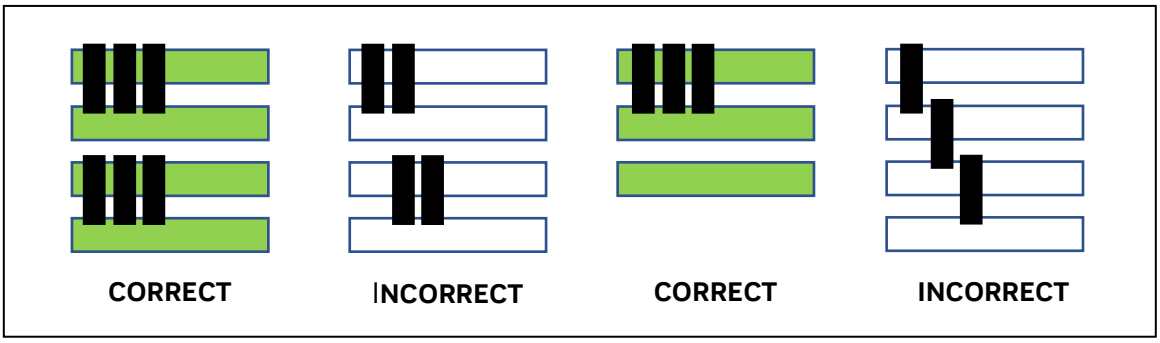
Using SMBPBI, the configured power limit setting can be made persistent across driver reloads. Refer to the *SMBus Post-Box Interface (SMBPBI) Design Guide* (DG-06034-002) for full implementation details.

NVLink Bridge Support

NVIDIA NVLink is a high-speed point-to-point (P2P) peer transfer connection. Where one GPU can transfer data to and receive data from one other GPU. The NVIDIA H100 NVL card supports NVLink bridge connection with a single adjacent NVIDIA H100 NVL card.

Each of the three attached bridges spans two PCIe slots. To function correctly as well as to provide peak bridge bandwidth, bridge connection with an adjacent NVIDIA H100 NVL card must incorporate all three NVLink bridges. Wherever an adjacent pair of the NVIDIA H100 NVL cards exists in the server, for best bridging performance and balanced bridge topology, the NVIDIA H100 NVL pair should be bridged. Figure 3 illustrates correct and incorrect NVIDIA H100 NVL NVLink connection topologies.

Figure 3. NVLink Topology – Top Views



For systems that feature multiple CPUs, both NVIDIA H100 NVL cards of a bridged card pair, should be within the same CPU domain. That is, under the same CPU's topology, and ensuring this benefits workload application performance. There are exceptions, for example, in a system with dual CPUs wherein each CPU has a single NVIDIA H100 NVL card under it. In that case, the two NVIDIA H100 NVL cards in the system may be bridged together. See the "PCIe and NVLink Topology" section.

NVIDIA H100 NVL card, NVLink speed, and bandwidth are given in the following table.

Table 6. H100 NVL Card NVLink Speed and Bandwidth

Parameter	Value
Total NVLink bridges supported by NVIDIA H100	3
Total NVLink Rx and Tx lanes supported	48
Data rate per NVIDIA H100 NVL NVLink lane (each direction)	100 Gbps
Total maximum NVLink bandwidth	600 Gbytes per second

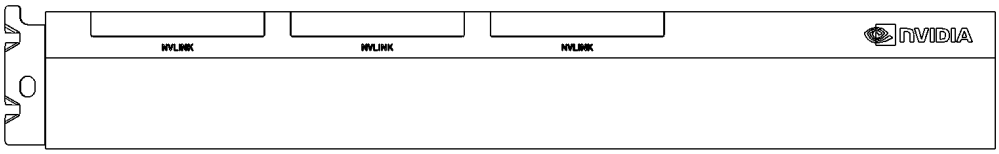
NVLink Bridge

The 2-slot NVLink bridge for the NVIDIA H100 NVL card (the same NVLink bridge used in the NVIDIA Ampere architecture generation, including the NVIDIA A100 PCIe card), has the following NVIDIA part number: 900-53651-0000-000.

NVLink Connector Placement

Figure 4 shows the connector keep-out area for the NVLink bridge support of the NVIDIA H100.

Figure 4. NVLink Connector Placement – Top View



Sufficient clearance must be provided both above the card's north edge and behind the backside of the card's PCB to accommodate NVIDIA H100 NVL NVLink bridges. The clearance above the north edge should meet or exceed 2.5 mm. The backside clearance (from the rear card's rear PCB surface) should meet or exceed 2.67 mm. Consult the *NVIDIA Form Factor 5.5 Specification for Enterprise PCIe Products Specification* (NVOnline: 1063377) for more detailed information.

NVLink bridge interfaces of the H100 NVL card include removable caps to protect the interfaces in non-bridged system configurations.

PCIe and NVLink Topology

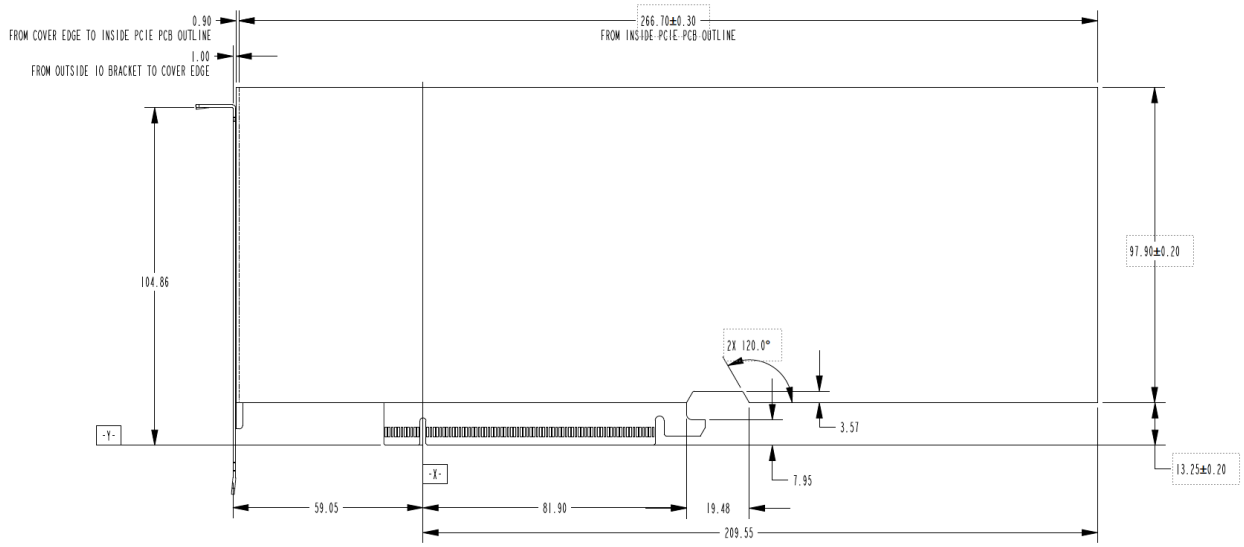
As stated, it is strongly recommended that both NVIDIA H100 NVL cards of a bridged card pair should be within the same CPU topology domain. Unless a dual CPU system has only two H100 NVL cards each of which is under its own CPU. Full NVLink connection topology guidance is as follows:

- Best NVLink Topology (Recommended):
  - Bridge two GPUs under the same CPU or PCIe switch
  - GPU count in a system should be in powers of two (1, 2, 4, 8, and so on)
  - Locate the same (even) number of GPUs under each CPU socket
  - Maintain a balanced configuration: same count of CPU:GPU:NIC for each grouping
- Good NVLink Topology:
  - Bridge two GPUs under different PCIe switches but under the same CPU
  - Same number of GPUs and NICs under each CPU socket, but not powers of two
- Allowed but Not Recommended:
  - Bridge two GPUs under two different CPUs
  - Odd number of GPUs under each CPU
  - Unbalanced configurations: Different ratios of CPU:GPU:NIC for each grouping

Form Factor

The NVIDIA H100 NVL card conforms to NVIDIA Form Factor 5.5 specification for a full-height, full-length (FHFL) dual-slot PCIe card. For details refer to the *NVIDIA Form Factor 5.5 Specification for Enterprise PCIe Products Specification* (NVOnline:1063377). In this product brief, nominal dimensions are shown.

Figure 5. NVIDIA H100 NVL PCIe Card Dimensions



Power Connector

This section details the power connector for the NVIDIA H100 NVL card.

Power Connector Placement

The board provides a PCIe 16-pin power connector on the east edge of the board.

Figure 6. PCIe 16-Pin Power Connector

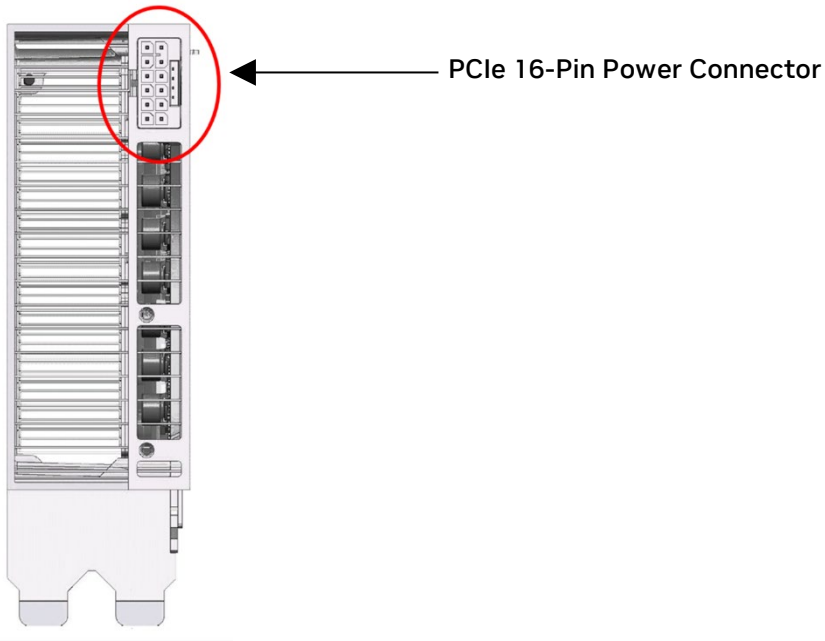
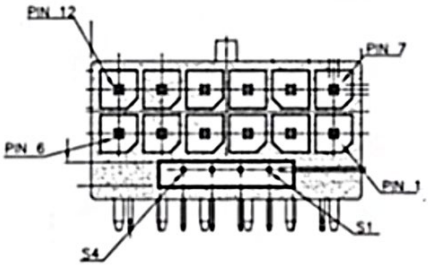


Figure 7 shows the pin assignments for the PCIe 16-pin power connector, per PCIe CEM 5.1 specification.

Figure 7. PCIe 16-Pin Power Connector Pin Assignment



12V PWR Cable Plug Pin Assignment

Pin	Signal
1 to 6	+12 V
7 to 12	Ground
S1	CARD_PWR_STABLE
S2	CARD_CBL_PRES#
S3	SENSE0
S4	SENSE1

Table 7 lists the power level options identifiable by the PCIe 16-pin power connecter per CEM5 PSU, and the corresponding Sense0 and Sense1 logic. The NVIDIA card senses the Sense0 and Sense1 levels and recognizes the power available to the NVIDIA card from the power connector. If the power level identified by Sense0 and Sense1 is equal to or greater than what the NVIDIA card needs from the 16-pin connector, the NVIDIA card operates per normal. If the power level identified by Sense0 and Sense1 is less than the default power cap of the NVIDIA card, the card will not boot.

The NVIDIA H100 NVL requires up to 400 W from the 16-pin auxiliary power connector. Table 7 shows the supported auxiliary power connector Sense pin logic and maximum supported TGP per power level.

Table 7. PCIe CEM 5.1 1-Pin PCIe PSU Power Level vs. Sense Logic

Power Level	Sideband 3 (Sense0)	Sideband 4 (Sense1)	Maximum TGP
451 - 600 W	0	0	400 W
301 - 450 W	1 (float)	0	400 W
151 - 300 W	0	1 (float)	310 W
Up to 150 W	1 (float)	1 (float)	Not supported. Insufficient power

Table 8 lists supported auxiliary power connections for the NVIDIA H100 NVL GPU card.

Table 8. Supported Auxiliary Power Connections

Board Connector	PSU Cable
PCIe 16-pin	PCIe 16-pin
PCIe 16-pin	CPU 8-pin to PCIe 16-pin

CPU 8-Pin to PCIe 16-Pin Power Adapter

A CPU 8-pin to PCIe 16-pin power adapter is available for testing the H100 NVL in systems that do not have native PCIe 16-pin power connectors. Figure 8 illustrates the power adapter. The power adapter provided by NVIDIA can only support 310 W TGP operation. To enable full 400 W TGP operation of the H100 NVL card, partners are advised to build their own power adapters (as needed) which support the 301 - 450 W power sense option.

- NVPN: 030-1546-000 – CPU 8-pin to PCIe 16-Pin Power Adapter
- Astron MFN: DAMAF01041-H

Figure 8. CPU 8-GPU Pin to PCIe 16-Pin Power Adapter

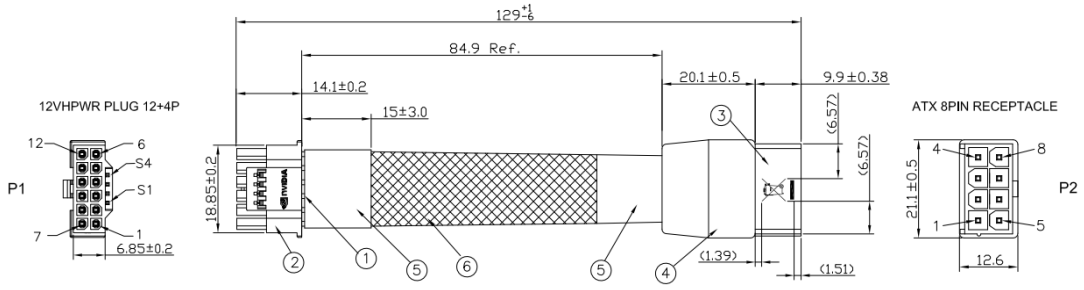
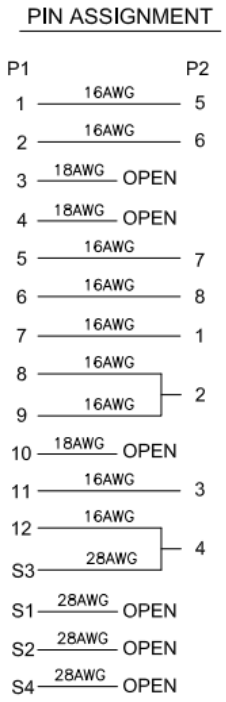


Figure 9 shows the CPU 8-pin to PCIe 16-pin power adapter pin assignments.

Figure 9. CPU 8-Pin to PCIe 16-Pin Power Adapter Pin Assignments



Note: The power adapter supports the four Sideband signals, hardware-strapped per Row 3 of Table 7. This strapping corresponds to the "151 – 300 W" power level PCIe CEM 5.1 specification (shown in Row 3 of Table 4 3). As a result, it supports only a 310 W TGP for the H100 NVL card. To support a TGP of 400 W, a power adapter or cable strapped as in Row 1 or Row 2 of Table 7 should be used.

Power Adapter Availability

The power adapter is provided with sample NVIDIA H100 NVL cards only and is intended to enable initial testing and qualification activities in systems that provide older power connectors.

For production cards, native system auxiliary power cabling that uses the PCIe 16-pin power connector, provides at least 400 W, and features Sense0 and Sense1 pins strapped for the 301 W - 450 W power class (as shown in Table 7) is recommended.

If only 300 W is available on the native auxiliary power cable from the system, Sense0 and Sense1 pins should be strapped for the 151 W - 300 W power class (as shown in Table 7), in which case the H100 NVL card will consume a maximum of 310 W. Consult NVIDIA for details.

Extenders

The NVIDIA H100 NVL card provides two extender options, shown in Figure 10 and Figure 11.

- NVPN: 151-0398-000 -- Enhanced Straight Extender
  - Card + extender = 312 mm
- NVPN: 682-00007-5555-001 – Straight extender
  - Card + extender = 312 mm
- NVPN: 682-00007-5555-000 – Long offset extender
  - Card + extender = 339 mm

Using a standard NVIDIA extender ensures greatest forward compatibility with future NVIDIA product offerings.

If the standard extender will not work, OEMs may design a custom attach method using the extender-mounting holes on the east edge of the PCIe card.

Figure 10. Enhanced Straight Extender

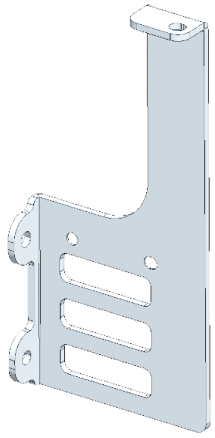
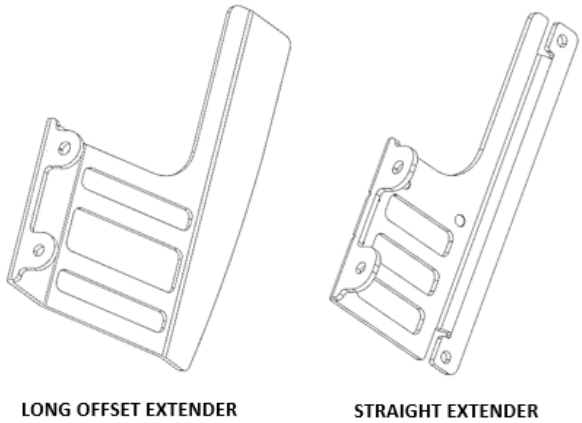


Figure 11. Legacy Long Offset and Straight Extenders



Support Information

Certification

- Windows Hardware Quality Lab (WHQL):
  - Windows 10, Windows 11
  - Windows Server 2019, Windows Server 2022
- Ergonomic requirements for office work W/VDTs (ISO 9241)
- EU Reduction of Hazardous Substances (EU RoHS)
- Joint Industry guide (J-STD) / Registration, Evaluation, Authorization, and Restriction of Chemical Substance (EU) – (JIG / REACH)
- Halogen Free (HF)
- EU Waste Electrical and Electronic Equipment (WEEE)

Agencies

- Australian Communications and Media Authority and New Zealand Radio Spectrum Management (RCM)
- Bureau of Standards, Metrology, and Inspection (BSMI)
- Conformité Européenne (CE)
- Federal Communications Commission (FCC)
- Industry Canada - Interference-Causing Equipment Standard (ICES)
- Korean Communications Commission (KCC)
- Underwriters Laboratories (cUL, UL)
- Voluntary Control Council for Interference (VCCI)

Languages

Table 9. Languages Supported

Languages	Windows[1]	Linux
English (US)	Yes	Yes
English (UK)	Yes	Yes
Arabic	Yes	
Chinese, Simplified	Yes	
Chinese, Traditional	Yes	
Czech	Yes	
Danish	Yes	
Dutch	Yes	
Finnish	Yes	
French (European)	Yes	
German	Yes	
Greek	Yes	
Hebrew	Yes	
Hungarian	Yes	
Italian	Yes	
Japanese	Yes	
Korean	Yes	
Norwegian	Yes	
Polish	Yes	
Portuguese (Brazil)	Yes	
Portuguese (European/Iberian)	Yes	
Russian	Yes	
Slovak	Yes	
Slovenian	Yes	
Spanish (European)	Yes	
Spanish (Latin America)	Yes	
Swedish	Yes	
Thai	Yes	
Turkish	Yes	

Note:

[1]**Microsoft Windows 10, Windows 11, Windows Server 2019, and Windows Server 2022 are supported.**

NVIDIA Corporation | 2788 San Tomas Expressway, Santa Clara, CA 95051
http://www.nvidia.com