NVIDIA RTX PRO 6000 Blackwell Server Edition GPU Accelerator

Product Brief

SP-12355-001_v02 | June 2025

Document History

SP-12355-001_v02

Version	Date	Authors	Description of Change
01	April 28, 2025	MD, DV	Initial release
02	June 27, 2025	MD, DV	> Added MTBF > Updated extender NVPN

Table of Contents

List of Figures

List of Tables

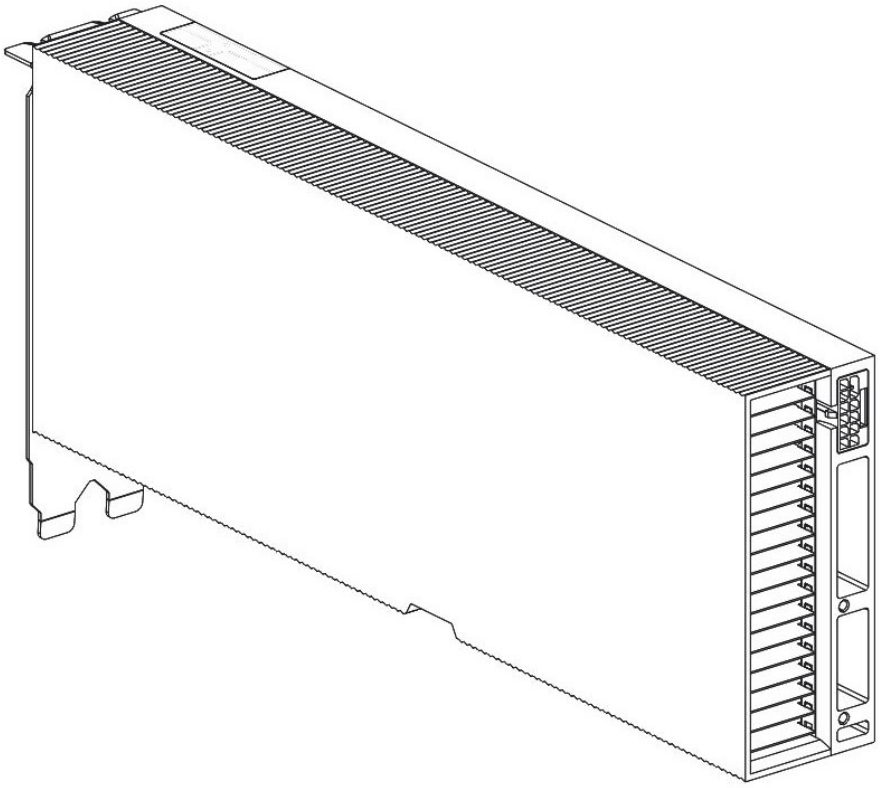
Chapter 1. Overview

The NVIDIA RTX™ PRO 6000 Blackwell Server Edition GPU Accelerator is a full-height, full-length (FHFL), dual-slot, 10.5-inch PCI Express Gen5 graphics solution based on the NVIDIA® Blackwell architecture. The card is passively cooled and is capable of 600 W maximum board power.

The NVIDIA RTX PRO 6000 supports the latest hardware-accelerated ray tracing, advanced shading, revolutionary AI features, and powerful simulation capabilities for a wide range of graphics and compute use cases in data center and edge server deployments. This includes deep learning inference as well as training workloads enabling generative AI workloads, batch and real-time rendering, virtual workstations, and cloud gaming.

RTX PRO 6000 delivers the highest level of graphics, ray tracing, and simulation performance for NVIDIA Omniverse™. With 96 GB of GDDR7 memory, even the most intense graphics applications run at the highest level of performance.

Figure 1-1. NVIDIA RTX PRO 6000 Blackwell Server Edition Volumetric



Chapter 2. Specifications

2.1 Product Specifications

Table 2-1 through Table 2-3 provide the product, memory, and software specifications for the NVIDIA RTX PRO 6000 PCIe GPU card.

Table 2-1. Product Specifications

Specification	NVIDIA RTX PRO 6000 PCIe GPU
Product SKU	PG153 SKU 210 NVPN: 900-2G153-xx00-xxx
Total board power	PCIe 16-pin cable strapped for 600 W power mode: > 600 W maximum (default) > 300 W minimum PCIe 16-pin cable strapped for 450 W power mode: > 450 W maximum (default) > 300 W minimum
Thermal solution	Passive
Mechanical form factor	Full-height, full-length (FHFL) 10.5”, dual-slot
PCI Device IDs	Device ID: 0x2BB5 Vendor ID: 0x10DE Sub-Vendor ID: 0x10DE Sub-System ID: 0x204E
GPU clocks	Base: 1,852 MHz Boost: 2,430 MHz
Performance states	P0, P8
VBIOS	EEPROM size: TBD Mbit UEFI: Supported
PCI Express interface	PCI Express Gen5 x16; Gen5 x8 Lane and polarity reversal supported
Multi-Instance GPU (MIG)	Supported (up to four instances)
NVIDIA NVLink	Not supported
Zero Power	Not supported

Specification	NVIDIA RTX PRO 6000 PCIe GPU
Connectors and headers	One PCIe 16-pin auxiliary power connector (12VHPWR auxiliary power connector) Four VESA® DisplayPort® connectors RTX PRO Sync3D Stereo
Weight	Board: 1,194 grams (excluding bracket, extender, and bridge) Bracket with screws: TBD grams Enhanced straight extender: TBD grams

Table 2-2. Memory Specifications

Specification	Description
Memory clock	12,481 MHz
Memory type	GDDR7
Memory size	96 GB
Memory bus width	512 bits
Peak memory bandwidth	1,597 GB/s

Table 2-3. Software Specifications

Specification	Description[1]
SR-IOV support	Supported – 48 VF (virtual functions)
BAR address (physical function)	BAR0: 16 MiB[1] BAR1: 128 GiB (Display Off mode; default) BAR1: 8 GiB (Display On, 8 GB BAR1 mode; resizable) BAR1: 256 MiB (Display On, 256 MB BAR1 mode; resizable) BAR2: 32 MiB[1]
BAR address (virtual function)	BAR0: 12 MiB (256 KiB per VF)1 BAR1: 192 GiB, 64-bit (4 GiB per VF)1 BAR2: 1,536 MiB, 64-bit (32 MiB per VF)1 VF BAR sizes are not applicable to Display On modes
Message signaled interrupts	MSI-X: Supported MSI: Not supported
ARI Forwarding	Supported
Driver support	Linux: R575 or later Windows: R575 or later
Secure boot[2]	Supported
NVIDIA® CUDA® support	x86: CUDA 12.9 or later
Virtual GPU software support	Supports vGPU 19.0 or later

Specification	Description[1]
NVIDIA AI Enterprise	Supported with VMWare
NVIDIA certification	NVIDIA-Certified Systems™ TBD or later
PCI class code	0x03 – Display controller
PCI subclass code	0x02 – 3D controller
Primary boot device capability	Supported
ECC support	Enabled
SMBus (8-bit address)	0x9E (write), 0x9F (read)
Reserved I2C addresses (8-bit)	0x4A, 0x40, 0x68, 0x82, 0xA0, 0xA4
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**

Table 2-4. Display Specifications

Specification	Description
DisplayPort 2.1b	Maximum raw bandwidth: 80 Gbps Example of maximum resolutions – DSC On > 3840 × 2160 × 24 bpp at 480 Hz[1] > 7680 × 4320 × 24 bpp at 165 Hz[1]
Simultaneous display support	Up to four simultaneous displays[2]
HDCP support	2.2

Notes:

[1]DisplayPort 2.1b with UHBR20: up to 4K 480 Hz or 8K 165 Hz with DSC

[2]Multi Monitor:

- **Four independent displays at 4K 165 Hz using DP**
- **Two independent displays at 4K 360 Hz or 8K 100 Hz with DSC using DP**
- **Other display configurations may be possible based on available bandwidth**

2.2 Thermal Specifications

Table 2-5 provides the PCIe reported temperatures and Table 2-6 provides the thermal specifications for the NVIDIA RTX PRO 6000 PCIe GPU card.

Table 2-5. RTX PRO 6000 PCIe Air-Cooled GPU Reported Temperatures

Specification	Units	Description
T_{AVG}	°C	Average temperature of all internal GPU sensors
T_{LIMIT}	°C	GPU temperature limit – current distance in °C from software slowdown event

Table 2-6. Thermal Specifications

Specification		Thermal Parameter Value
Thermal qualification temperature		T_{LIMIT} > +5 °C AND T_{AVG} < 80 °C
Maximum operating temperature	GPU	T_{LIMIT} = 0 °C
Hardware slowdown temperature (50% clock slowdown)	GPU	T_{LIMIT} ≤ - 2 °C
Hardware shutdown temperature	GPU	T_{LIMIT} ≤ - 5 °C

2.2.1 GPU Reported Temperatures

The following detailed definitions apply to critical temperatures that the system should monitor. The GPU reports these parameters using in-band and out-of-band management paths.

- **GPU Temperature (T_{AVG}):** T_{AVG} is the average temperature of all internal GPU thermal sensors.
- **Thermal Limit Temperature (T_{LIMIT}):** T_{LIMIT} provides the distance (in °C) from the maximum operating temperature. T_{LIMIT} can be used for temperature telemetry and fan control.

2.2.2 Thermal Qualification Temperature

Thermal qualification temperatures are used for partner thermal design and qualification. The partner thermal solution should be designed and validated to maintain GPU temperatures at or below the qualification temperature under worst-case thermal test conditions.

2.2.3 Maximum Operating Temperature

The maximum operating temperature is the thermal limit for the NVIDIA GPU Boost™ feature. NVIDIA GPU Boost will adaptively optimize the GPU clock to maximize performance while maintaining GPU temperatures below the maximum operating limits for the GPU. That is, it will try to keep T_{LIMIT} greater than zero.

It is possible for the T_{LIMIT} temperature to be equal to or less than 0 °C when the GPU temperature T_{AVG} is less than the thermal qualification temperature for some

asymmetrical workloads where the GPU power is not distributed evenly across the GPU die.

Thermal qualification temperature may be specified for T_{LIMIT} greater than zero on some products.

2.2.4 Hardware Slowdown Temperature

The hardware slowdown is the temperature above which a hardware divider will reduce the clock speed to cool the component. This is only available for the GPU. The hardware divider will further reduce core clocks by 50%, to lower the GPU temperature, at a significant impact to performance.

2.2.5 Hardware Shutdown Temperature

Hardware shutdown is defined as the temperature above which the component must shut down to prevent potential damage. The shutdown temperature should not be reached during normal operation but may occur if a failure is present in the thermal subsystem. For example, if the system fans malfunction, they are unable to cool the GPUs.

2.3 Electrical Specifications

The input voltage measured at the connectors should never exceed the voltage range defined in Table 2-7.

Table 2-7. Input Voltage Specifications

Input Rails	Minimum	Nominal	Maximum
PCI Express edge connector (12V)	11.04 V	12 V	12.96 V
PCIe 16-pin power connector (12V)	11.04 V	12 V	12.6 V
PCI Express edge connector (3.3V)	3.0 V	3.3 V	3.6 V

Note: The minimum to maximum range includes DC level and any noise or other transients. Ensure that the input rails always remain within the specified range.

Table 2-8 provides the Input EDP Continuous (EDPc) specifications for the NVIDIA RTX PRO 6000 PCIe GPU card's input rails. The Input EDPc current is specified over a 1-second moving average time scale. The peak input power over a 1-second moving average will not exceed the specified Input EDPc value.

Table 2-8. Input EDPc Specifications

Input Rails	Maximum Value	Moving Average
PCI Express edge connector (12V)	66 W	1 s
	5.5 A[1]	1 s
PCIe 16-pin power connector (12V)	600 W	1 s
	50 A[1]	1 s
PCI Express edge connector (3.3V)	10 W	1 s
	3 A	1 s
Total board power[2]	600 W	1 s

Notes:

[1]Maximum current specifications are based on nominal voltages.

[2]Individual rails are specified at their absolute maximum values. However, the total power consumption of the board during operation will never exceed the specified total board power.

Table 2-9 provides the input EDP Peak (EDPp) specifications for the NVIDIA RTX PRO 6000 PCIe GPU card's 12 V input power rails. Input EDPp values are specified over 400 µs and 50 ms moving averages. For a given power connector, the peak input current over each moving average will not exceed the specified input EDPp value. Table 2-9 also provides the total input EDPp specification for all input power sources combined, which will never be exceeded.

Table 2-9. Input EDPp Specifications

	Maximum Value[1,2]	
Specification	400 µs	50 ms
PCI Express edge connector (12 V)	2 A	1 A
PCIe 16-pin power connector (12 V)	90 A	75 A
Total (12 V)[3]	90 A	75 A

Notes:

[1]Maximum current specifications are based on nominal voltages.

[2]Maximum current specifications do not change across different default power modes. Refer to Table 4-4.

[3]For any time scale, the total EDPp specification may be less than the sum of the individual power connector EDPp values.

Table 2-10 provides the power management specifications for the NVIDIA RTX PRO 6000 PCIe GPU card.

Table 2-10. Power Management Specifications

Parameter	Value	Unit
Power capping threshold for total board power	450/600[2]	W
Power capping response time (typical)[1]	50	ms
Power capping response time (maximum)[1]	100	ms

Note:

[1]This is the length of time required to return peak power excursions to the total board power-capping threshold.

[2] This depends on the PCIe 16-pin cable strapping.

The Power Brake feature allows the system to trigger a hardware slowdown in the GPU with the external input. For example, the system can assert Power Brake during an emergency power or thermal event. Note that the Power Brake should not be used to dynamically control the power consumption of the GPU. That is, Power Brake is intended for emergency power reduction purposes only.

The system can assert Power Brake signal (**PB#**) using Pin B30 on the PCIe edge connector. The typical response time seen at the system's power supply after assertion is noted in Table 2-11. In addition, system designs must keep Power Brake asserted for a minimum duration as listed in the following table.

Table 2-11. Power Brake Specifications

Parameter	Value
PB# PCIe pin assignment	B30
Power Brake response time (maximum)	150 µs
PB# input assertion low time (minimum)	250 ms
Power Brake hardware slowdown factor	4 ×
Minimum power reduction when PB# asserted	50%

2.4 Environmental and Reliability Specifications

Table 2-12 provides the environment conditions specifications for the NVIDIA RTX PRO 6000 PCIe GPU card.

Table 2-12. Board Environmental and Reliability Specifications

Specification	Description
Ambient operating temperature	10 °C to 45 °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)	> Uncontrolled environment:[1] 864,500 hours at 35 °C > Controlled environment:[2] 1,062,500 hours at 35 °C

Notes: Specifications in this table are applicable up to 6,000 feet.

A period of not more than 96 hours consecutive, not to exceed 15 days per year.

[1] **Some environmental stress with limited maintenance (GF35).**

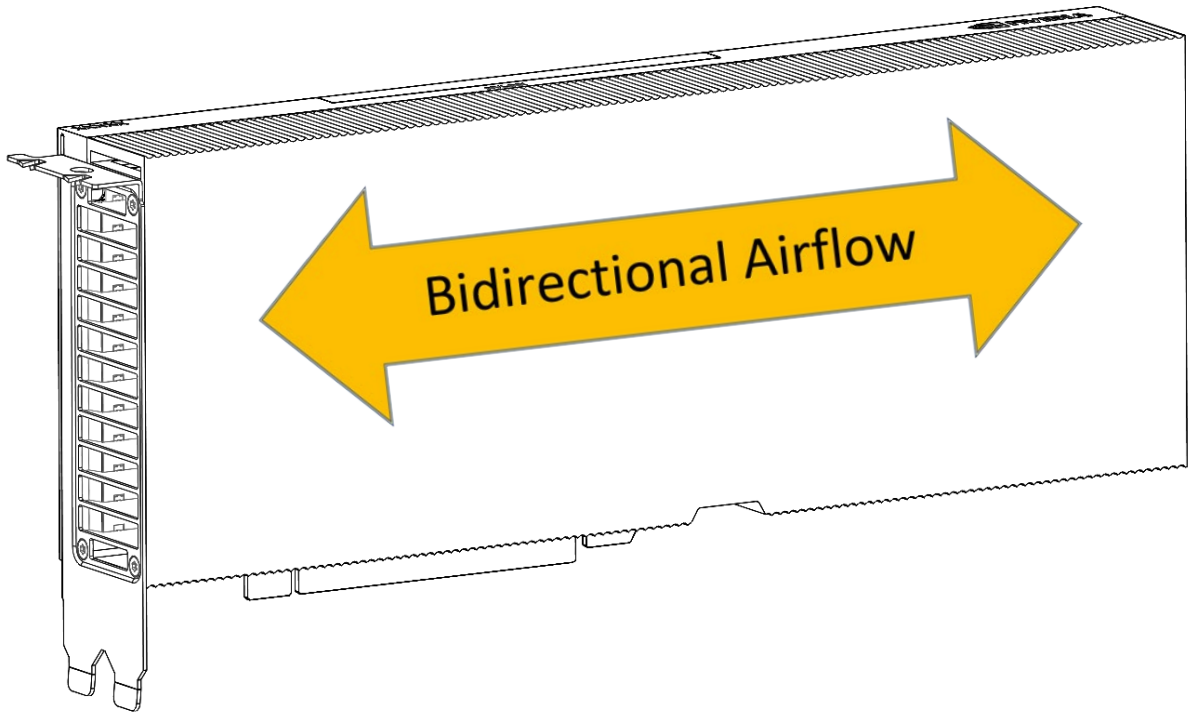
[2] **No environmental stress with optimum operation and maintenance (GB35).**

Chapter 3. System Airflow Requirements

3.1 Airflow Direction Support

The NVIDIA RTX PRO 6000 PCIe GPU card employs a bidirectional heat sink, which accepts airflow in either left-to-right or right-to-left directions.

Figure 3-1. NVIDIA RTX PRO 6000 PCIe GPU Airflow Directions



3.2 Airflow Rate vs. Heat Sink Inlet Temperature

The following table specifies the minimum volumetric airflow rate for the NVIDIA RTX PRO 6000 PCIe GPU card at the specified air temperature points measured at the card's inlet. These data points represent the conditions under which the board was qualified.

Table 3-1. Minimum CFM Requirements per Inlet Air Temperature

Heat Sink Inlet Temperature (°C)	Minimum Airflow Rate (CFM)[1,2]	Qualified by NVIDIA
45	125	Yes
40	84	Yes
35	65	Yes
30	51	Yes
25[3]	43	Yes

Notes:

[1]The minimum airflow is assumed to be a fully ducted airflow through the product's heat sink. System thermal designs must account for any air bypass the system may allow.

[2]The airflow and air inlet temperature data provided in this table is a qualification reference point, not an absolute specification. Refer to the system design guide for more details regarding guidance for the system's airflow design.

[3]This is the minimum required airflow whenever input power is provided to the card. The airflow rate must meet or exceed this value regardless of the heat sink inlet temperature or the system's operating mode. For example, even if the system is in idle mode and the heat sink inlet temperature is low, if the 12V input rails to the card are powered, then the system must still provide at least this specified minimum airflow.

Figure 3-2 shows a general relationship between air inlet temperature and the required volumetric air flow rate. An airflow rate on or above the trend line for a given air temperature at the card's inlet should meet or exceed the necessary card cooling. Airflow rates below the curve are inadequate.

Figure 3-2. System Airflow Rate vs. Inlet Temperature

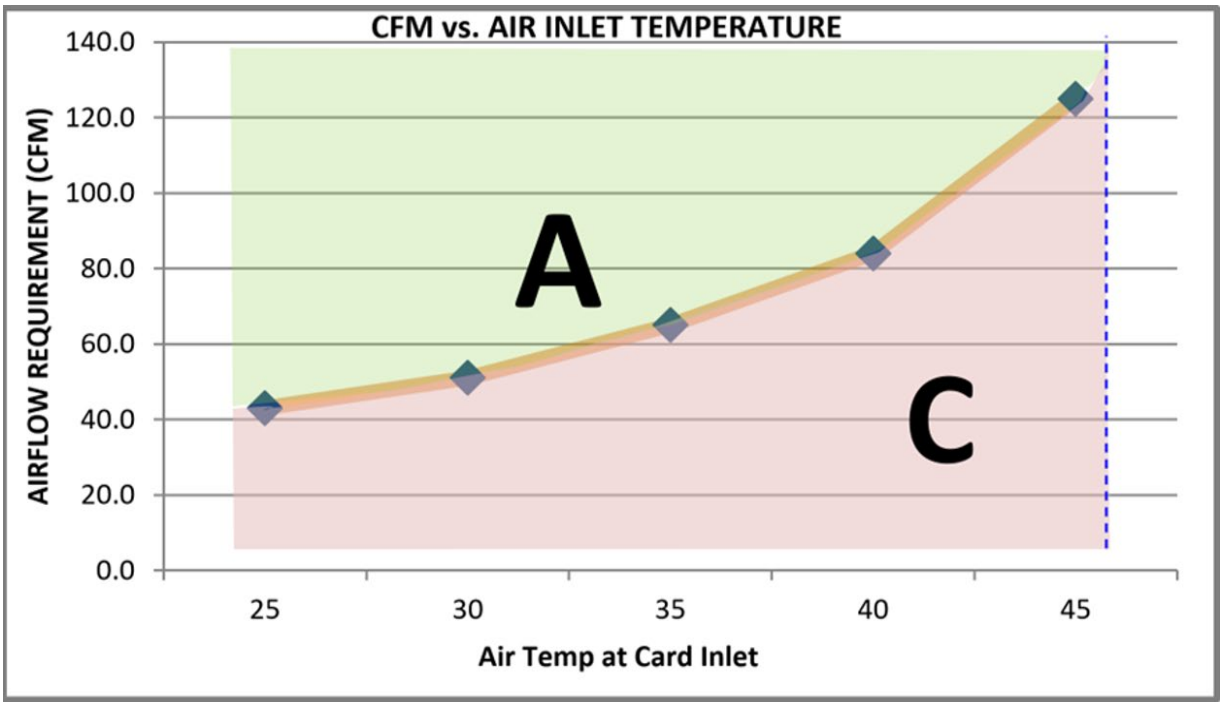


Chart Zone:

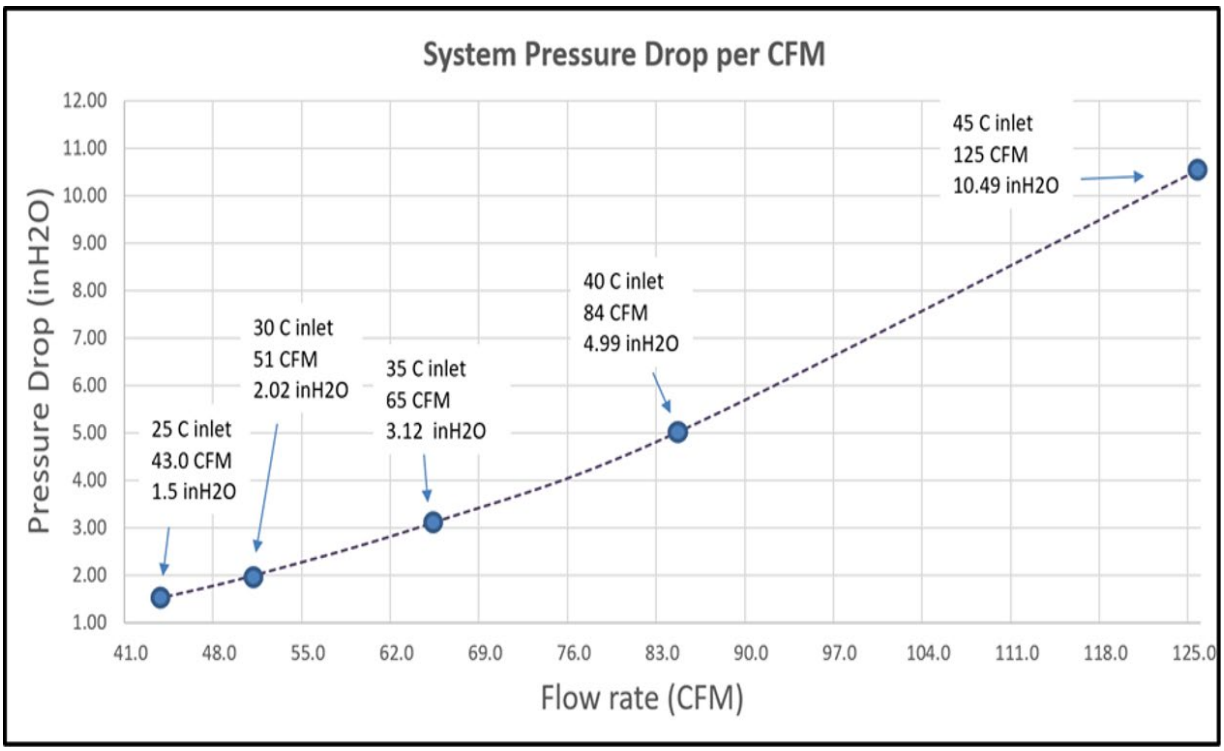
A. No thermocouples needed on critical components if qualifying in this zone. Passing GPU temperatures is sufficient to validate cooling system.

C. Not recommended; NVIDIA would not pass a qualification that falls in this zone without careful inspection of data, system, and cooling solution.

3.3 Pressure Drop vs. Airflow Rate

The relationship between CFM and air pressure drop through the heat sink is shown in Figure 3-3.

Figure 3-3. System Pressure Drop



Note: The curve in Figure 3-3 shows the nominal pressure drop. Tolerance is ± 20%.

Chapter 4. Product Features

4.1 PCI Express Interface Specifications

The following subsections describe the PCIe interface specifications for the NVIDIA RTX PRO 6000 PCIe GPU card.

4.1.1 PCIe Support

The NVIDIA RTX PRO 6000 PCIe GPU card supports PCIe Gen5. Either a Gen5 x16 Gen5 x8, or Gen4 x16 interface should be used when connecting to the NVIDIA RTX PRO 6000 PCIe GPU card.

4.1.2 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 the 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 on an NVIDIA RTX PRO 6000 PCIe GPU card with 48 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 the 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 RTX PRO 6000 PCIe GPU card 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.

4.1.3 Interrupt Messaging

The NVIDIA RTX PRO 6000 PCIe GPU card only supports the MSI-X interrupt messaging protocol. The MSI interrupt protocol is not supported.

4.1.4 ARI Forwarding

The NVIDIA RTX PRO 6000 PCIe GPU card is enabled for SR-IOV. Hence, it is an ARI device. Improvements with respect to the previous GPU family include:

- The NVIDIA RTX PRO 6000 PCIe GPU decodes all eight bits of the function number.
- The NVIDIA RTX PRO 6000 PCIe GPU ARI Linked List follows PCIe rules. Only functions that exist appear in the ARI Linked List.
- All NVIDIA RTX PRO 6000 PCIe GPU functions visible have the ARI Capability with the correct linked list pointers.
- Per the PCIe specification, VFs never appear in the ARI Linked List.
- The configuration of the host system should be as follows:
- As with any SR-IOV device, before it sets “VF Enable” to “Enable SR-IOV” in an SR-IOV Function (that is Function 0 in the case of the NVIDIA RTX PRO 6000 PCIe GPU card), the host should configure ARI Capable Hierarchy in the “SR-IOV Control” register of the SR-IOV Function. Doing so may affect the placement of that Function’s VFs. Note that while the PCIe specification supports SR-IOV devices using more than one bus number, the NVIDIA RTX PRO 6000 PCIe GPU always uses a single bus number (SR-IOV devices that use more than one bus number are extremely rare and most system software and firmware do not support them).
    - If ARI Forwarding is enabled, NVIDIA RTX PRO 6000 PCIe GPU) in the Downstream Port immediately above the NVIDIA RTX PRO 6000 PCIe GPU card, then the host should leave the ARI Capable Hierarchy bit in the NVIDIA RTX PRO 6000 PCIe GPU’s SR-IOV Control register reset to 0b. This is a suboptimal configuration of the NVIDIA RTX PRO 6000 PCIe GPU as it limits the number of VFs that can be used.

4.1.5 Polarity Inversion and Lane Reversal Support

Lane Polarity Inversion, as defined in the PCIe specification, is supported on the NVIDIA RTX PRO 6000 PCIe GPU card.

Lane Reversal, as defined in the PCIe specification, is supported on the NVIDIA RTX PRO 6000 PCIe GPU card. When reversing the order of the PCIe lanes, the order of both the Rx lanes and the Tx lanes must be reversed.

Note: PCIe pins B12 and B30 should not be connected on the host side of the PCIe connector.

If the system will not use the CLKREQ# output to gate the PCIe clock to the system slot, PCIe Pin B12 can be left unconnected (NC).

4.2 Display

This section details the operating modes for NVIDIA RTX PRO 6000.

4.2.1 Display On and Off

The RTX PRO 6000 PCIe card supports three operating modes as described in Table 4-1.

Table 4-1. Display Modes

Display Mode	BAR Address (Physical Functions)
Display Off (default)	BAR1: 128 GiB (Display Off mode; default)
Scalable visualization	BAR1: 8 GiB (Display On, 8 GB BAR1 mode; resizable)
Display enabled	BAR1: 256 MiB (Display On, 256 MB BAR1 mode; resizable)

4.2.1.1 Display Off Mode

The default Display Off mode supports SR-IOV and is required to run NVIDIA Virtual GPU software. NVIDIA RTX PRO 6000 with NVIDIA Virtual GPU software enables users to tackle massive datasets, large 3D models, and complex designs with scaled memory and performance on any device, anywhere. NVIDIA RTX PRO 6000 supports the following NVIDIA virtual GPU software:

- NVIDIA RTX Virtual Workstation (vWS)
- NVIDIA Virtual Applications (vAPPs)

  NVIDIA Virtual PC (vPC)

4.2.1.2 Display On 8GB BAR1 Mode

The Display On, 8 GB BAR1 mode is the recommended configuration for scalable visualization system deployments. In this mode, the NVIDIA RTX PRO 6000 card requires a BAR1 size of 8 GB and can drive up to four VESA® DisplayPort™ monitors through the integral DisplayPort connectors on the card's bracket.

Synchronizing content across multiple monitors driven from different RTX PRO 6000 cards is accomplished by using the NVIDIA RTX PRO Sync card.

4.2.1.3 Display On 256MB BAR1 Mode

The Display On, 256 MB BAR1 mode is the recommended configuration for professional desktop systems. In this mode, the NVIDIA RTX PRO 6000 card can drive up to four DisplayPort monitors through the integral DisplayPort connectors on the card's bracket.

Synchronizing content across multiple monitors driven from different RTX PRO 6000 cards is accomplished by using the NVIDIA RTX PRO Sync card.

4.2.2 Switching Operating Modes

The NVIDIA DisplayModeSelectorTool will be provided to switch modes. An example command to switch GPU mode is shown:

```
sudo ./displaymodeselector --gpumode <mode_name>
```

To display the list of gpumodes available, the following command may be used:

```
sudo ./displaymodeselector --gpumode
```

After switching modes, the system must be rebooted, after which the configured mode takes effect.

4.2.3 Frame Lock

The NVIDIA RTX PRO 6000 supports frame lock by using the NVIDIA RTX PRO Sync board. The frame lock and stereo connectors are on the card's north edge.

4.3 Hardware Root of Trust

The NVIDIA RTX PRO 6000 provides a primary root of trust within the GPU that provides the following:

- Secure boot
- Secure firmware upgrade
- Firmware rollback protection
- Support for in-band firmware update disable (established after each GPU reset)
- Secure application processor recovery

4.4 Multi-Instance GPU Support

The NVIDIA RTX PRO 6000 PCIe GPU card supports Multi-Instance GPU (MIG) capability by providing up to four GPU instances per NVIDIA RTX PRO 6000 PCIe GPU. MIG technology can partition the NVIDIA RTX PRO 6000 PCIe GPU into individual instances, each fully isolated with its own memory, cache, and compute cores, enabling optimized graphics and compute resource provisioning and quality of service (QoS).

For detailed information on MIG provisioning and use, consult the *Multi-Instance GPU User's Guide.*

4.5 System Management

The NVIDIA RTX PRO 6000 supports GPU out-of-band telemetry and firmware updates through SMBus and USB 2.0. The NVIDIA RTX PRO 6000 uses an onboard microcontroller (MCU) for MCTP over USB signaling for increased communication speed through the PCIe edge connector pins A5 and A6. Table 4-2 shows the pin assignments.

Table 4-2. USB Signal Pin Assignment

PCIe Pin Assignment	USB Signal
A5	D-
A6	D+

By default, the MCU is bypassed and held in reset. If the MCU is bypassed, the GPU can only be accessed by the SMBus. The MCU can be enabled by the GPIO expander at I2C address 0x20. Once the MCU is active, the GPU is accessible by the SMBus and USB. Example commands are shown below:

```
i2cset -y 0 0x20 0x03 0xC7 # configure IO expander
i2cset -y 0 0x20 0x01 0xEF # release MCU reset and disable bypass
i2cset -y 0 0x20 0x01 0xCF # Hold MCU reset and enable bypass
```

4.6 Programmable Power

The Programmable Power feature provides partners with 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 RTX PRO 6000 PCIe GPU card are presented in Table 2-1.

4.6.1 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 the driver running in persistence mode. An example command to reduce the power cap to 400 W is shown:

```
nvidia-smi -pm 1
nvidia-smi -pl 400
```

To restore the NVIDIA RTX PRO 6000 PCIe GPU card 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 600
```

4.6.2 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 4-3. SMBPBI Commands

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

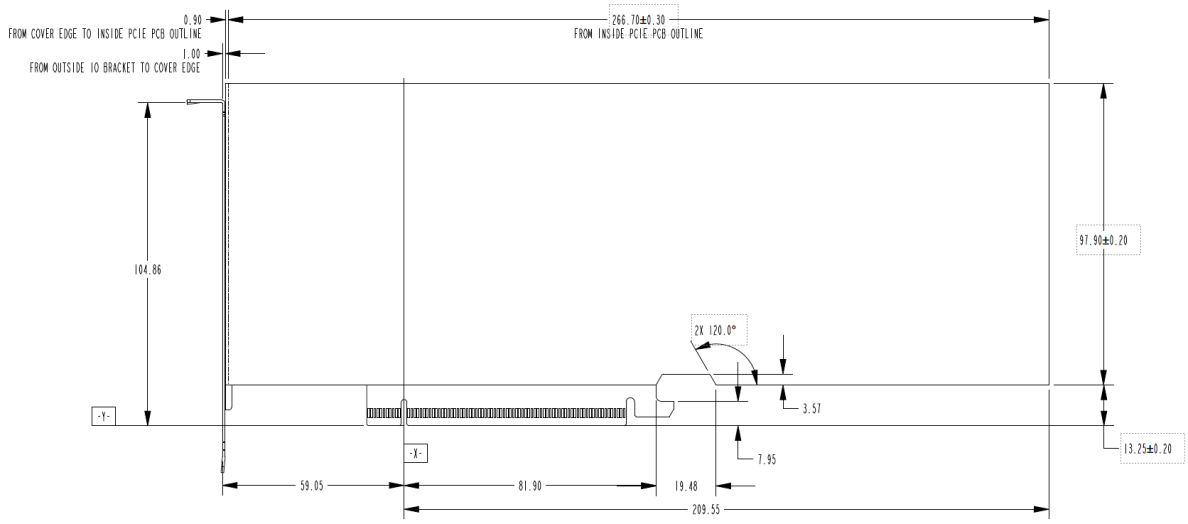
Using SMBPBI, the configured power limit setting can be made persistent across driver reloads.

4.7 Form Factor

The NVIDIA RTX PRO 6000 PCIe GPU card conforms to the NVIDIA Form Factor 6.0 specification for a full-height, full-length (FHFL) dual-slot PCIe card.

In this product specification, nominal dimensions are shown.

Figure 4-1. NVIDIA RTX PRO 6000 PCIe GPU PCIe Card Dimensions



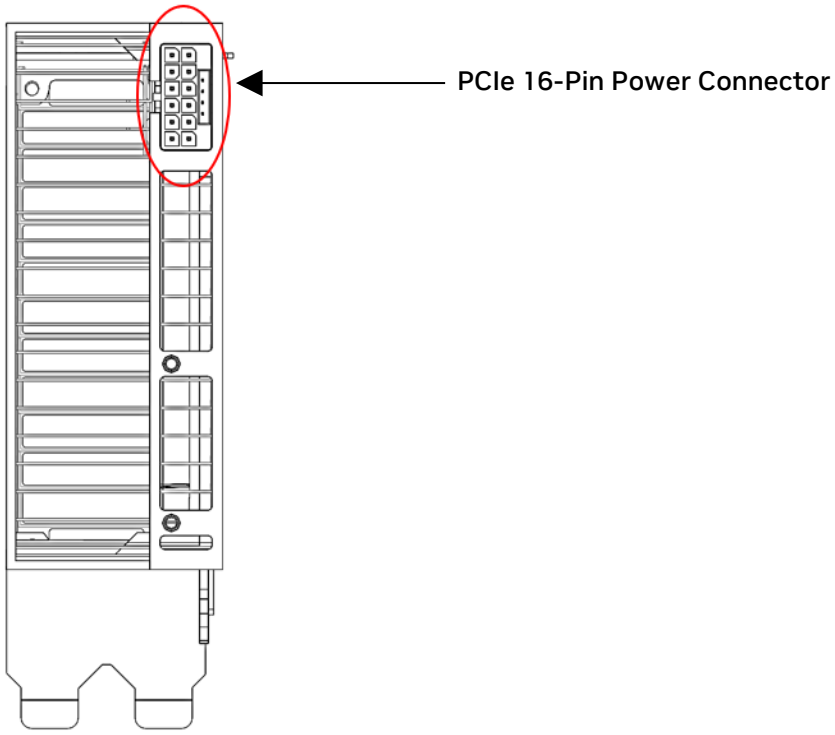
4.8 Power Connector

This section details the power connector for the NVIDIA RTX PRO 6000 PCIe GPU card.

4.8.1 Power Connector Placement

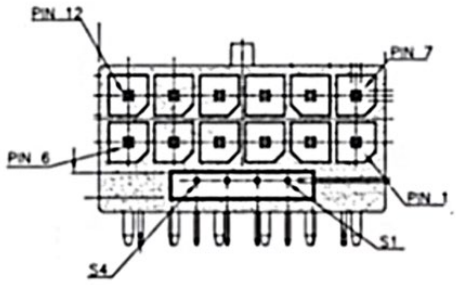
The board provides a PCIe 16-pin power connector on the east edge of the board.

Figure 4-2. PCIe 16-Pin Power Connector



The following figure shows the pin assignments for the PCIe 16-pin power connector, per the PCIe CEM 5.1 specification.

Figure 4-3. 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 4-4 lists the power level options identifiable by the PCIe 16-pin power connector 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 as 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 RTX PRO 6000 PCIe GPU requires up to 600 W from the 16-pin auxiliary power connector. Table 4-4 shows the supported auxiliary power connector Sense pin logic and the maximum supported TGP per power level. Please note, the EDPp specification is independent of the TGP power level and does not change across boot power modes. Refer to Table 2-9.

Table 4-4. PCIe CEM 5.1 16-Pin PCIe PSU Power Level vs. Sense Logic

Power Level	Sideband 3 (Sense0)	Sideband 4 (Sense1)	Maximum TGP
451 - 600 W	0	0	600 W
301 - 450 W	1 (float)	0	450 W
151 - 300 W	0	1 (float)	Not supported. Insufficient power
Up to 150 W	← SENSE PINS SHORTED →		Not supported. Insufficient power
No Power	1 (float)	1 (float)	Not supported. Insufficient power

The following table lists the supported auxiliary power connections for the NVIDIA RTX PRO 6000 PCIe GPU card.

Table 4-5. Supported Auxiliary Power Connections

Board Connector	PSU Cable
PCIe 16-pin	PCIe 16-pin

4.9 Extender

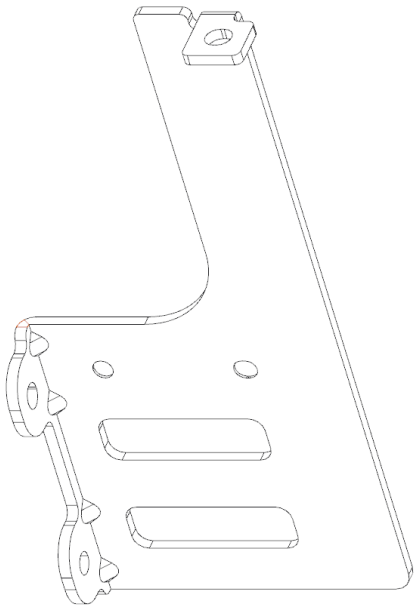
The NVIDIA RTX PRO 6000 PCIe GPU card provides the following extender option, shown in Figure 4-4.

- NVPN: 151-1346-000 - Enhanced Straight Extender
  - Card + extender = 312 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 extender compliant with the applicable PCIe CEM specification, or custom attach method using the extender-mounting holes on the east edge of the PCIe card.

Figure 4-4. Enhanced Straight Extender



Chapter 5. Support Information

5.1 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)

5.2 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)

5.3 Languages Supported

The following table lists the supported languages for the NVIDIA RTX PRO 6000 PCIe GPU card.

Table 5-1. 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.**

Notice

This document is provided for information purposes only and shall not be regarded as a warranty of a certain functionality, condition, or quality of a product. NVIDIA Corporation (“NVIDIA”) makes no representations or warranties, expressed or implied, as to the accuracy or completeness of the information contained in this document and assumes no responsibility for any errors contained herein. NVIDIA shall have no liability for the consequences or use of such information or for any infringement of patents or other rights of third parties that may result from its use. This document is not a commitment to develop, release, or deliver any Material (defined below), code, or functionality.

NVIDIA reserves the right to make corrections, modifications, enhancements, improvements, and any other changes to this document, at any time without notice.

Customer should obtain the latest relevant information before placing orders and should verify that such information is current and complete.

NVIDIA products are sold subject to the NVIDIA standard terms and conditions of sale supplied at the time of order acknowledgement, unless otherwise agreed in an individual sales agreement signed by authorized representatives of NVIDIA and customer (“Terms of Sale”). NVIDIA hereby expressly objects to applying any customer general terms and conditions with regards to the purchase of the NVIDIA product referenced in this document. No contractual obligations are formed either directly or indirectly by this document.

NVIDIA products are not designed, authorized, or warranted to be suitable for use in medical, military, aircraft, space, or life support equipment, nor in applications where failure or malfunction of the NVIDIA product can reasonably be expected to result in personal injury, death, or property or environmental damage. NVIDIA accepts no liability for inclusion and/or use of NVIDIA products in such equipment or applications and therefore such inclusion and/or use is at customer’s own risk.

NVIDIA makes no representation or warranty that products based on this document will be suitable for any specified use. Testing of all parameters of each product is not necessarily performed by NVIDIA. It is customer’s sole responsibility to evaluate and determine the applicability of any information contained in this document, ensure the product is suitable and fit for the application planned by customer, and perform the necessary testing for the application in order to avoid a default of the application or the product. Weaknesses in customer’s product designs may affect the quality and reliability of the NVIDIA product and may result in additional or different conditions and/or requirements beyond those contained in this document. NVIDIA accepts no liability related to any default, damage, costs, or problem which may be based on or attributable to: (i) the use of the NVIDIA product in any manner that is contrary to this document or (ii) customer product designs.

No license, either expressed or implied, is granted under any NVIDIA patent right, copyright, or other NVIDIA intellectual property right under this document. Information published by NVIDIA regarding third-party products or services does not constitute a license from NVIDIA to use such products or services or a warranty or endorsement thereof. Use of such information may require a license from a third party under the patents or other intellectual property rights of the third party, or a license from NVIDIA under the patents or other intellectual property rights of NVIDIA.

Reproduction of information in this document is permissible only if approved in advance by NVIDIA in writing, reproduced without alteration and in full compliance with all applicable export laws and regulations, and accompanied by all associated conditions, limitations, and notices.

THIS DOCUMENT AND ALL NVIDIA DESIGN SPECIFICATIONS, REFERENCE BOARDS, FILES, DRAWINGS, DIAGNOSTICS, LISTS, AND OTHER DOCUMENTS (TOGETHER AND SEPARATELY, “MATERIALS”) ARE BEING PROVIDED “AS IS.” NVIDIA MAKES NO WARRANTIES, EXPRESSED, IMPLIED, STATUTORY, OR OTHERWISE WITH RESPECT TO THE MATERIALS, AND EXPRESSLY DISCLAIMS ALL IMPLIED WARRANTIES OF NONINFRINGEMENT, MERCHANTABILITY, AND FITNESS FOR A PARTICULAR PURPOSE. TO THE EXTENT NOT PROHIBITED BY LAW, IN NO EVENT WILL NVIDIA BE LIABLE FOR ANY DAMAGES, INCLUDING WITHOUT LIMITATION ANY DIRECT, INDIRECT, SPECIAL, INCIDENTAL, PUNITIVE, OR CONSEQUENTIAL DAMAGES, HOWEVER CAUSED AND REGARDLESS OF THE THEORY OF LIABILITY, ARISING OUT OF ANY USE OF THIS DOCUMENT, EVEN IF NVIDIA HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. Notwithstanding any damages that customer might incur for any reason whatsoever, NVIDIA’s aggregate and cumulative liability towards customer for the products described herein shall be limited in accordance with the Terms of Sale for the product.

Trademarks

NVIDIA, the NVIDIA logo, CUDA, NVIDIA-Certified Systems, GPU Boost, Omniverse, and RTX are trademarks and/or registered trademarks of NVIDIA Corporation in the U.S. and other countries. Other company and product names may be trademarks of the respective companies with which they are associated.

Copyright

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