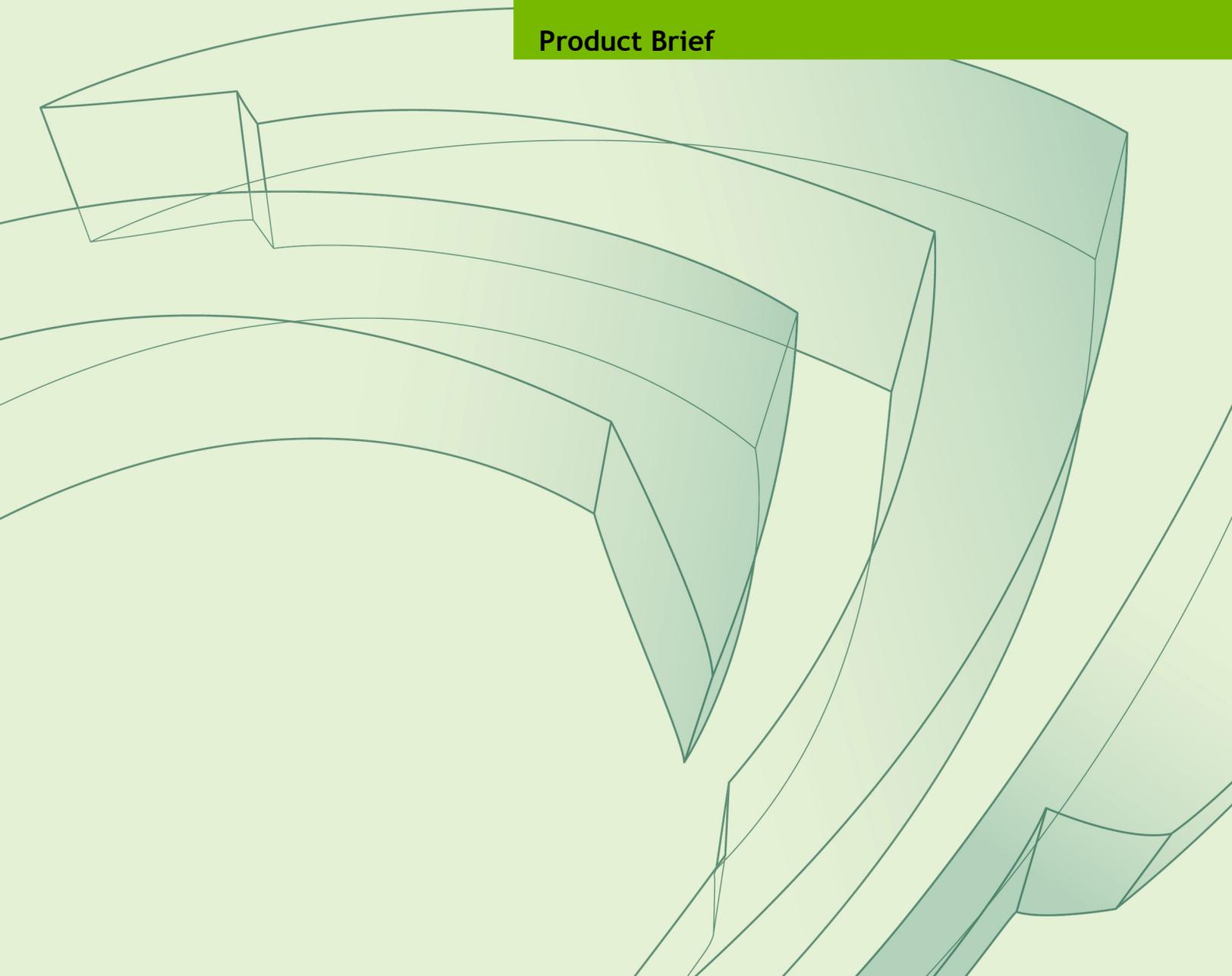




TESLA P100 PCIe GPU ACCELERATOR

PB-08248-001_v01 | October 2016

Product Brief



DOCUMENT CHANGE HISTORY

PB-08248-001_v01

Version	Date	Authors	Description of Change
01	October 3, 2016	GG, SM	Initial Release

TABLE OF CONTENTS

- Overview 1**
- Specifications..... 2**
 - Product Specifications 2
 - Thermal Specifications 4
 - System Airflow Requirements 5
 - Power and Electrical Specifications 6
- Design Discussion 7**
 - Form Factor 7
 - Power Connector Placement 8
 - CPU 8-Pin to PCIe 8-Pin Dongle 9
 - Extenders 9
- Support Information..... 11**
 - Certificates and Agencies 11
 - Certifications 11
 - Agencies 11
 - Languages 12

LIST OF FIGURES

Figure 1.	Tesla P100 PCIe Airflow Directions with Optional I/O Bracket	5
Figure 2.	Tesla P100 PCIe Board Dimensions with Optional I/O Bracket.....	7
Figure 3.	CPU 8-Pin Power Connector with Optional I/O Bracket	8
Figure 4.	CPU 8-Pin to PCIe 8-Pin Dongle.....	9
Figure 5.	Long Offset Extender	10
Figure 6.	Straight Extender.....	10

LIST OF TABLES

Table 1.	Product Specifications	2
Table 2.	Memory Specifications	3
Table 3.	Software Specifications.....	3
Table 4.	Board Environment and Reliability Specifications	4
Table 5.	Thermal Specifications	4
Table 6.	Minimum CFM Requirements per Inlet Air Temperature.....	6
Table 7.	Input EDP Continuous Specifications	6
Table 9.	Supported Auxiliary Power Connections	8
Table 10.	Languages Supported	12

OVERVIEW

The NVIDIA® Tesla® P100 GPU Accelerator for PCIe is a dual-slot 10.5 inch PCI Express Gen3 card with a single NVIDIA® Pascal™ GP100 graphics processing unit (GPU). It uses a passive heat sink for cooling, which requires system air flow to properly operate the card within its thermal limits. The Tesla P100 PCIe supports double precision (FP64), single precision (FP32) and half precision (FP16) compute tasks, unified virtual memory and page migration engine.

For performance optimization, NVIDIA GPU Boost™ feature is supported. By adjusting the GPU clock dynamically, maximum performance is achieved within the power cap limit.

Tesla P100 PCIe boards are shipped with ECC enabled by default to protect the GPU's memory interface and the on-board memories. ECC protects the memory interface by detecting any single, double, and all odd-bit errors. The GPU will replay any memory transaction that have an ECC error until the data transfer is error-free. ECC protects the DRAM content by fixing any single-bit errors and detecting double-bit errors. There is no replay associated with ECC. The Tesla P100 PCIe with HBM2 memory has native support for ECC and has no ECC overhead, both in memory capacity and bandwidth.

For more information on compute capabilities, HBM2, unified virtual memory, and page migration engine visit NVIDIA official website.

The Tesla P100 for PCIe is available in two memory configurations:

- ▶ Tesla P100 PCIe with 16GB HBM2
- ▶ Tesla P100 PCIe with 12GB HBM2

As a board design, both the memory configurations are identical and this product brief covers both configurations.

SPECIFICATIONS

PRODUCT SPECIFICATIONS

Table 1 provides the product specifications for the Tesla P100 PCIe GPU Accelerator.

Table 1. Product Specifications

Specification		Tesla P100 PCIe 16GB	Tesla P100 PCIe 12GB
Product SKUs		NVPN: 699-2H400-0201-XXX	NVPN: 699-2H400-0202-XXX
Total board power		250 W	
GPU SKUs		GP100-892-A1	GP100-893-A1
PCI Device IDs		Device ID: 0x15F8 Vendor ID: 0x10DE Sub-Vendor ID: 0x10DE Sub-System ID: 0x118F	Device ID: 0x15F7 Vendor ID: 0x10DE Sub-Vendor ID: 0x10DE Sub-System ID: 0x11DA
NVIDIA® CUDA® cores		3584	
GPU clocks	Base	1189 MHz	
	Maximum boost	1328 MHz	
VBIOS	EEPROM size	4 Mbit	
	UEFI	Supported	
PCI Express interfaces		PCI Express 3.0 ×16 Lane and polarity reversal supported	
Power connectors and headers		One CPU 8-pin auxiliary power connector	
Weight	Board	1177 Grams	
	Bracket with screws	21 Grams	
	Long offset extender	52 Grams	
	Straight extender	42 Grams	

Table 2 provides the memory specifications for the Tesla P100 PCIe board.

Table 2. Memory Specifications

Specification	Tesla PCIe 16GB	Tesla PCIe 12GB
Maximum memory clock	715 MHz	715 MHz
Memory size	16GB HBM2	12GB HBM2
Memory bus width	4096-bit	3072-bit
Peak memory bandwidth	Up to 732 GB/s	Up to 549 GB/s

Table 3 provides the software specifications that apply to both the Tesla P100 PCIe 16GB and the Tesla P100 PCIe 12GB.

Table 3. Software Specifications

Specification	Description
Compatibility mode supported	Compute (default) Graphics
Base address	BAR0: 16 MB BAR1: 16 GB (Compute) or 256 MB (Graphics) BAR3: 32 MB I/O BAR: 4 KB (Graphics only)
PCI class code	0x03 - Display Controller
PCI sub-class code	0x00 - VGA-Compatible Controller (Graphics) 0x02 - 3D Controller (Compute)
ECC support	Supported (Enabled by default)
SMBus (8-bit address)	0x9E (write), 0x9F (read)
SMBus direct access	Supported
SMBPBI (SMBus Post Box Interface)	Supported

Table 4 provides the environment conditions specifications to both the Tesla P100 PCIe 16GB and the Tesla P100 PCIe 12GB.

Table 4. Board Environment and Reliability Specifications

Specification	Condition
Operating temperature	0 °C to 45 °C
Storage temperature	-40 °C to 75 °C
Operating humidity	5% to 90% relative humidity
Storage humidity	5% to 95% relative humidity
Mean time between failures (MTBF)	Uncontrolled environment: 769566 hours at 35 °C Controlled environment: 1523813 hours at 35 °C

THERMAL SPECIFICATIONS

Table 5 provides the thermal specifications that apply to both the Tesla P100 PCIe 16GB and the Tesla P100 PCIe 12GB.

Table 5. Thermal Specifications

Parameter	Value	Units
Total board power	250	W
GPU shutdown temperature	85	°C
GPU slowdown temperature (50% clock slowdown)	82	°C
GPU maximum operating temperature	80	°C

System Airflow Requirements

The Tesla P100 PCIe board employs a bidirectional heat sink, which accepts airflow either left-to-right or right-to-left directions.

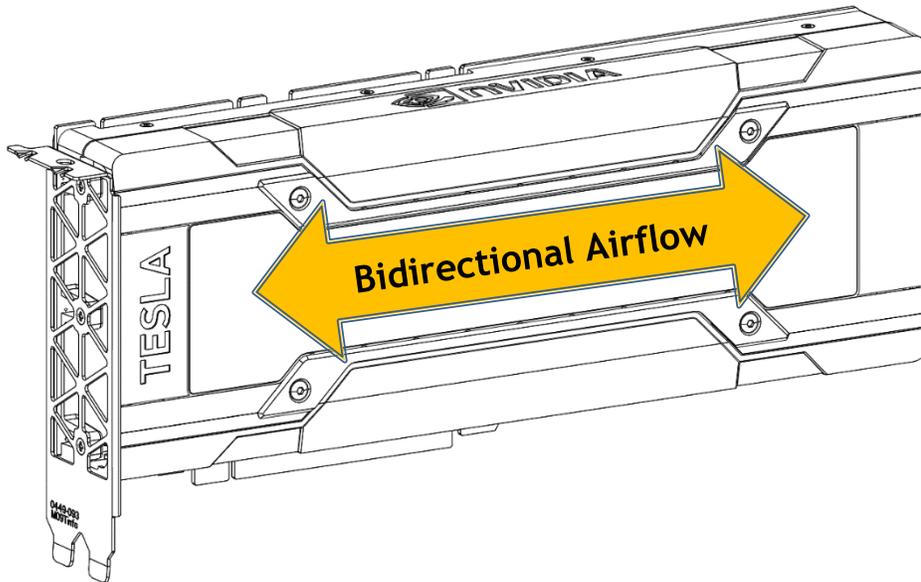


Figure 1. Tesla P100 PCIe Airflow Directions with Optional I/O Bracket

Airflow Rate vs. Heat Sink Inlet Temperature

Table 6 specifies the minimum volumetric airflow rate for the Tesla P100 PCIe boards at the specified air temperature points measured at the card's inlet. These data points represent the conditions under which the board was qualified.

The following table applies to both the Tesla P100-PCIe 16GB and the Tesla P100-PCIe 12GB.

Table 6. Minimum CFM Requirements per Inlet Air Temperature

Heat Sink Inlet Temperature (°C)	Minimum Airflow Rate (CFM) ^{1,2}	Qualified by NVIDIA
30	12	
35	15	Yes
40	19	Yes
45	26	Yes
50	37	
Any ³	12	

Notes:

¹The 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.

²The 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 the system airflow design guidance.

³ This is the minimum required airflow when input power is provided to the card regardless of heat sink inlet temperature or the system's operating mode. For example, if the system is in idle mode but the 12V input rails to the card are active, the system needs to provide the specified minimum airflow.

POWER AND ELECTRICAL SPECIFICATIONS

Table 7 provides the Input EDP Continuous specifications for input rails. The Input EDP Continuous is specified over a 1 second time scale. The peak input power over a 1 second moving average will not exceed the specified Input EDP Continuous value.

The following table applies to both the Tesla P100-PCIe 16GB and the Tesla P100-PCIe 12GB.

Table 7. Input EDP Continuous Specifications

Input Rails	Maximum Value	Moving Average
PCI Express edge connector (12V ± 8%)	66 W	1 s
	5.5 A	1 s
CPU 8-pin power connector (12V ± 8%)	240 W	1 s
	20 A	1 s
PCI Express edge connector (3V3 ± 9%)	10 W	1 s
	3 A	1 s
Total graphic power ¹	250 W	1 s

Note:

¹Individual rails are specified at their absolute maximum values, but the total power consumption of the board during operation will never exceed the specified total graphic power.

DESIGN DISCUSSION

FORM FACTOR

The Tesla P100 PCIe board conforms to NVIDIA Form Factor 3.0 specification. In this product brief, nominal dimensions are shown; for tolerances, see the attached 2D mechanical drawings.

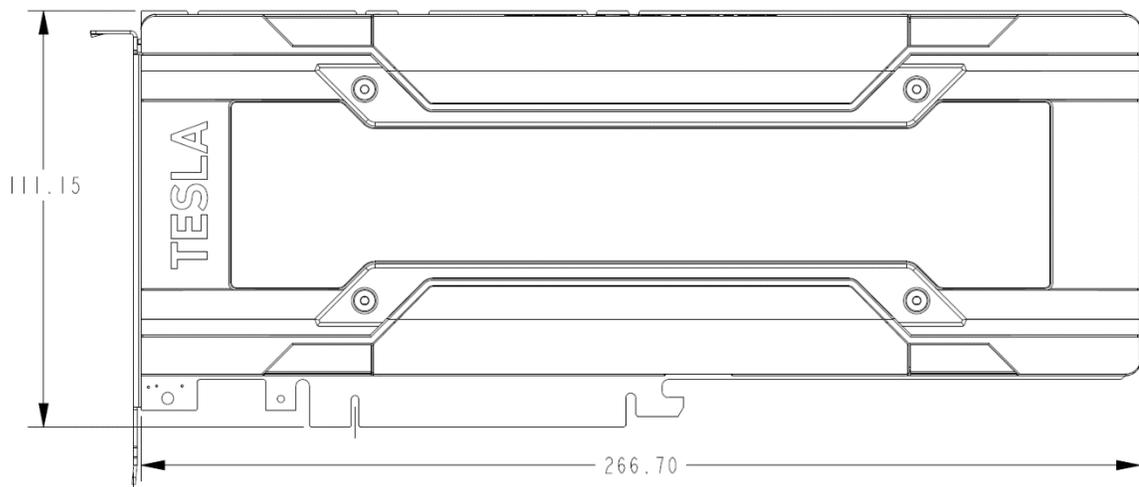


Figure 2. Tesla P100 PCIe Board Dimensions with Optional I/O Bracket

POWER CONNECTOR PLACEMENT

The board provides a CPU 8-pin power connector on the East edge of the board.

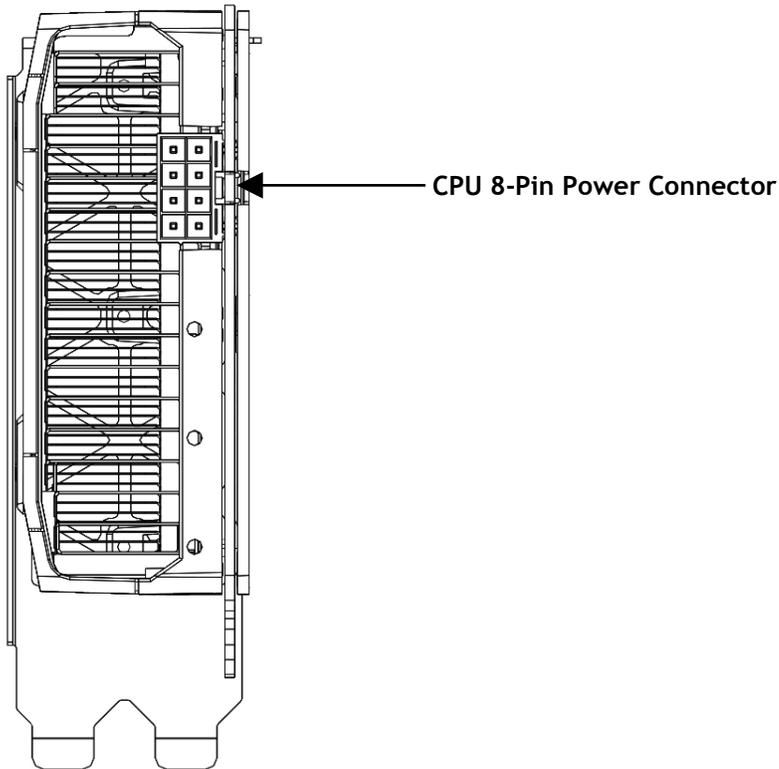


Figure 3. CPU 8-Pin Power Connector with Optional I/O Bracket

Table 8. Supported Auxiliary Power Connections

Board Connector	PSU Cable
CPU 8-pin	1x CPU 8-pin cable
CPU to PCIe 8-pin dongle	2x PCIe 8-pin cable 2x PCIe 6-pin cable ¹ 1x PCIe 8-pin cable and 1x PCIe 6-pin cable ¹

Notes:

¹The PCIe 6-pin cable must be capable of carrying up to 120 W.

CPU 8-Pin to PCIe 8-Pin Dongle

Figure 9 lists the pin assignments of the dongle. The part number for the dongle is:

- ▶ NVPN: 030-0571-000

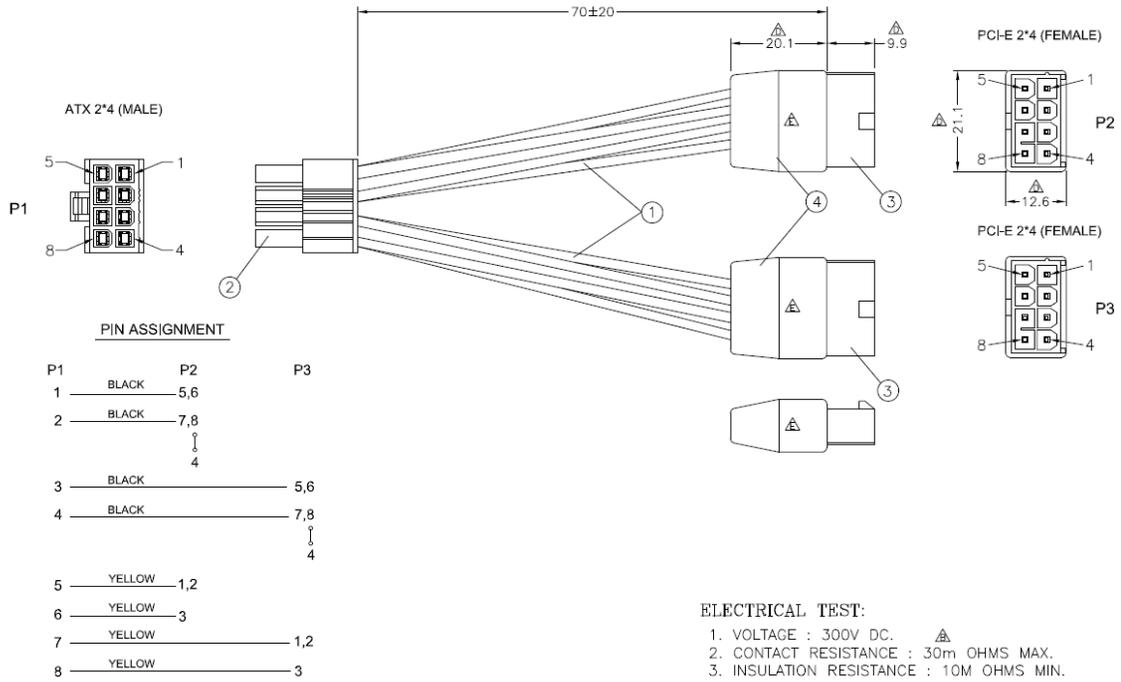


Figure 4. CPU 8-Pin to PCIe 8-Pin Dongle

EXTENDERS

The Tesla P100 PCIe board provides two extender options as shown in the following figures.

- ▶ NVPN: 682-00003-5555-002 – Long offset extender (Figure 5)
 - Card + extender = 339 mm
- ▶ NVPN: 682-00003-5555-000 – Straight extender (Figure 6)
 - Card + extender = 312 mm

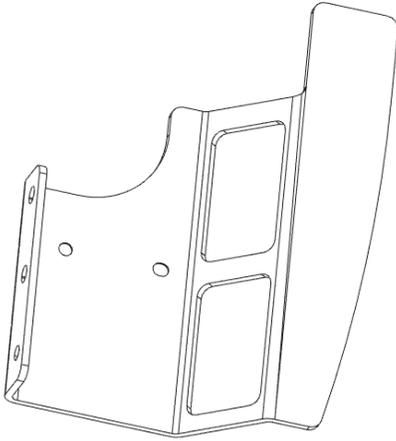


Figure 5. Long Offset Extender

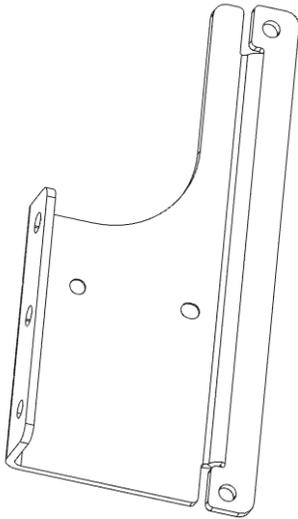


Figure 6. Straight Extender

- ▶ Using the 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 heat sink baseplate. The extender mounting holes will vary among NVIDIA products, so designing for flexibility is recommended.

SUPPORT INFORMATION

CERTIFICATES AND AGENCIES

Certifications

- ▶ Windows Hardware Quality Lab (WHQL):
 - Certified Windows 7, Windows 8.1, and Windows 10
 - Certified Windows Server 2008 R2 and Windows Server 2012 R2
- ▶ 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:

¹Windows 7, Windows 8, Windows 8.1, Windows 10, Windows Server 2008 R2, and Windows Server 2012 R2 are supported.

Notice

The information provided in this specification is believed to be accurate and reliable as of the date provided. However, NVIDIA Corporation (“NVIDIA”) does not give any representations or warranties, expressed or implied, as to the accuracy or completeness of such information. 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 publication supersedes and replaces all other specifications for the product that may have been previously supplied.

NVIDIA reserves the right to make corrections, modifications, enhancements, improvements, and other changes to this specification, at any time and/or to discontinue any product or service without notice. Customer should obtain the latest relevant specification 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. NVIDIA hereby expressly objects to applying any customer general terms and conditions with regard to the purchase of the NVIDIA product referenced in this specification.

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 these specifications will be suitable for any specified use without further testing or modification. Testing of all parameters of each product is not necessarily performed by NVIDIA. It is customer’s sole responsibility to ensure the product is suitable and fit for the application planned by customer and to do 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 specification. NVIDIA does not accept any 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 specification, 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 specification. 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 specification is permissible only if reproduction is approved by NVIDIA in writing, is reproduced without alteration, and is accompanied by all associated conditions, limitations, and notices.

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. 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 NVIDIA terms and conditions of sale for the product.

Trademarks

NVIDIA, the NVIDIA logo, CUDA, NVIDIA GPU Boost, Pascal, and Tesla 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

© 2016 NVIDIA Corporation. All rights reserved.