



NVIDIA T4 70W LOW PROFILE PCIe GPU ACCELERATOR

PB-09256-001_v05 | April 2020

Product Brief



DOCUMENT CHANGE HISTORY

PB-09256-001_v05

Version	Date	Authors	Description of Change
01	October 24, 2018	SL, SM	Initial Release
02	November 1, 2018	SL, SM	Updated ECC information
03	December 19, 2018	AN, SM	General product update
04	March 13, 2020	SL, SM	Updated “Overview” section with T4 qualified servers statement
05	April 15, 2020	SL, SM	Updated part numbers in Table 1

TABLE OF CONTENTS

- Overview 1
- Specifications..... 3
 - Product Specifications 3
 - Environmental and Reliability Specifications 5
 - Airflow Direction Support..... 5
- Design Discussion 6
 - Form Factor..... 6
 - Extenders 6
- Support Information..... 7
 - Planned Certifications..... 7
 - Agencies..... 7
 - Languages 8

LIST OF FIGURES

Figure 1.	NVIDIA T4 Tensor Core GPU.....	2
Figure 2.	NVIDIA T4 Board Dimensions	6

LIST OF TABLES

Table 1.	Product Specifications	3
Table 2.	Memory Specifications	4
Table 3.	Software Specifications.....	4
Table 4.	Board Environmental and Reliability Specifications.....	5
Table 5.	Languages Supported	8

OVERVIEW

The NVIDIA® T4 is a single-slot, low-profile, 6.6-inch PCI Express Gen3 Universal Deep Learning Accelerator based on the TU104 NVIDIA graphics processing unit (GPU). The T4 has 16 GB GDDR6 memory and a 70 W maximum power limit. The T4 is offered as a passively cooled board that requires system air flow to operate the card within its thermal limits. NVIDIA supports x8 and x16 PCI Express for the T4.

NVIDIA T4 is a universal deep learning accelerator ideal for distributed computing environments. Powered by NVIDIA Turing™ Tensor Cores, T4 provides revolutionary multi-precision performance to accelerate deep learning and machine learning training and inference, video transcoding, and virtual desktops. As part of the NVIDIA AI Platform, T4 supports all AI frameworks and network types, delivering dramatic performance and efficiency that maximize the utility of at-scale deployments.

NVIDIA T4 boards are shipped with ECC and page retirement enabled by default.

For general design and system qualification guidance, refer to the *System Design Guide for NVIDIA Enterprise GPU Products Design Guide* (DG-07562-001).

For best performance with NVIDIA T4, make sure that you use a server that has been qualified for T4 that meets the thermal and airflow requirements for NVIDIA T4. Refer to the following website for the latest list of qualified T4 servers:

<https://www.nvidia.com/en-us/data-center/tesla/tesla-qualified-servers-catalog/>

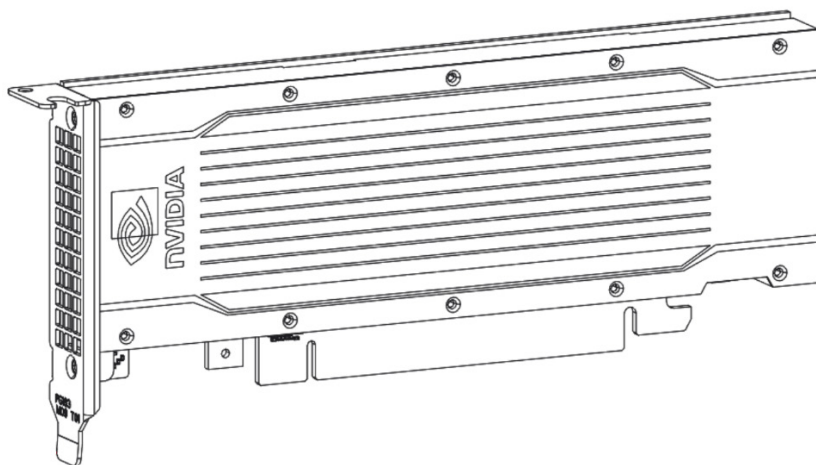


Figure 1. NVIDIA T4 Tensor Core GPU

SPECIFICATIONS

PRODUCT SPECIFICATIONS

Table 1 provides the product specifications for the NVIDIA T4.

Table 1. Product Specifications

Specification		Description
Product SKU		PG183 SKU 200 NVPN: 699-2G183-0200-xxx Order part numbers: (See Notes) 900-2G183-0000-000 900-2G183-0000-001
Total board power		70 W (default)
GPU SKU		TU104-895
PCI IDs		Device ID: 0x1EB8 Vendor ID: 0x10DE Sub-Vendor ID: 0x10DE Sub-System ID: 0x12A2
GPU clocks	Base	585 MHz
	Maximum Boost	1590 MHz
VBIOS	EEPROM size	8 Mbit
	UEFI	Supported
PCI Express interface		PCI Express 3.0 ×16 x8
Thermal cooling solution		Passive
Weight	Board	301 Grams

Specification		Description
	Full Height Bracket with screws	17 Grams
	Half Height Bracket with screws	10 Grams
Notes: T4 70W, LP bracket, P/N 900-2G183-0000-000 (LP bracket installed, ATX included as an accessory) T4 70W, ATX bracket, P/N 900-2G183-0000-000 (ATX bracket installed, LP included as an accessory) -000 has low profile bracket preinstalled while -001 has ATX bracket preinstalled. The brackets are not single wide full size. The brackets are single wide low profile.		

Table 2 provides the memory specifications for the NVIDIA T4.

Table 2. Memory Specifications

Specification	Description
Maximum memory clock	5001 MHz
Memory size	16 GB
Memory bus width	256 bits
Peak Memory bandwidth	Up to 320 GBytes/s

Table 3 provides the software specifications.

Table 3. Software Specifications

Specification	Description
SR-IOV support	Supported; 16 VF (virtual functions)
Base address (physical function)	BAR0: 16 MB BAR1: 256 MB BAR3: 32 MB
Base address (virtual function)	BAR0: 4 MB, 32-bit (16 VF x 256K) BAR1: 4 GB, 64-bit (16 VF x 256M) BAR3: 512 MB, 64-bit (16 VF x 32M)
Message Signaled Interrupts	MSI-X Supported MSI Not Supported
PCI class code	0x03 - Display Controller
PCI sub-class code	0x02 - 3D Controller
ECC support	Configurable (Enabled by default)
SMBus (8-bit address)	GPU 0: 0x9E (write), 0x9F (read)
SMBus direct access	Supported
SMBPBI (SMBus Post Box Interface)	Supported
Zero Power	Supported

Environmental and Reliability Specifications

Table 4 provides the environment conditions specifications for the NVIDIA T4.

Table 4. Board Environmental and Reliability Specifications

Specification	Condition
Operating temperature	0 °C to 50 °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: TBD at 35 °C Controlled environment: TBD at 35 °C

AIRFLOW DIRECTION SUPPORT

The T4 supports bi-directional airflow either from left to right, or from right to left. CFM requirements are identical for both airflow directions.

DESIGN DISCUSSION

FORM FACTOR

The NVIDIA T4 board is a half-height, half-length card. The dimensions are shown in Figure 2.

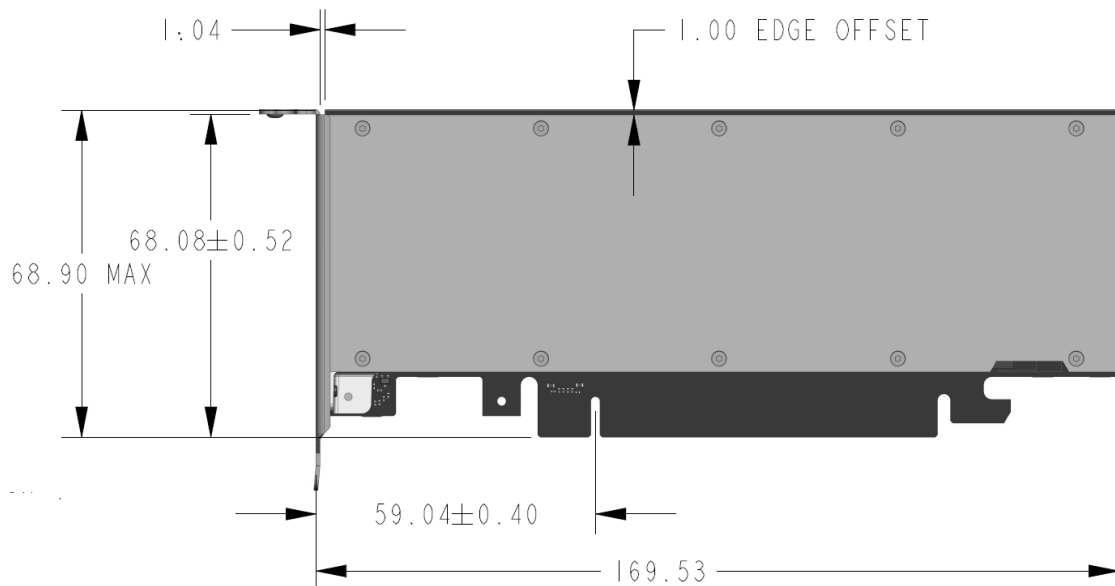


Figure 2. NVIDIA T4 Board Dimensions

EXTENDERS

NVIDIA does not offer extenders for the Tesla T4.

SUPPORT INFORMATION

PLANNED CERTIFICATIONS

- ▶ Windows Hardware Quality Lab (WHQL): Windows 10
- ▶ 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 5. 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:

¹Microsoft Windows 7, Windows 8, Windows 8.1, Windows 10, Windows Server 2008 R2, Windows Server 2012 R2 and Windows 2016 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, and NVIDIA Turing 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

© 2018, 2020 NVIDIA Corporation. All rights reserved.