NVIDIA T4 NGC-Ready Platform

Design Guide
## Document History

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Authors</th>
<th>Description of Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>June 12, 2019</td>
<td>SH, SM</td>
<td>Initial Release</td>
</tr>
<tr>
<td>02</td>
<td>June 14, 2019</td>
<td>SH, SM</td>
<td>Clarified PCIe lane width for NVMe drives</td>
</tr>
</tbody>
</table>
# Table of Contents

Chapter 1. Introduction ....................................................................................................... 1
Chapter 2. Software Support ............................................................................................... 2
Chapter 3. T4 NGC-Ready Platform Specification ............................................................... 3
Chapter 4. System Topologies ............................................................................................. 4
Chapter 5. Design Discussion .............................................................................................. 7
  5.1 PCI Express Interface ....................................................................................................... 7
  5.2 CPU and System Memory ................................................................................................ 7
  5.3 Network Interface ............................................................................................................. 8
  5.4 Storage .............................................................................................................................. 8
List of Figures

Figure 4-1. T4 NGC-Ready Server Topology – Best.................................................................4
Figure 4-2. T4 NGC-Ready Server Topology – Better..............................................................5
Figure 4-3. T4 NGC-Ready Server Topology – Good...............................................................6

List of Tables

Table 3-1. T4 NGC-Ready Platform Specification..................................................................3
Chapter 1. Introduction

NGC (NVIDIA GPU Cloud) is the hub for GPU-optimized software for deep learning, machine learning, and HPC that takes care of all the software plumbing, so data scientists, developers, and researchers can focus on building solutions, gathering insights, and delivering business value.

An NGC-Ready server using the NVIDIA T4 graphics processing unit (GPU) is specified to run NGC software for deep learning (DL) training and inference, machine learning (ML), and high-performance computing (HPC) with consistent, predictable performance.

This design guide provides the platform specification for an NGC-Ready server using the NVIDIA T4 GPU. It includes the GPU, CPU, system memory, network, and storage requirements needed for NGC-Ready compliance.

In addition, some design guidance is provided for T4-based servers that may not meet the requirements specified for NGC-Ready but are viable solutions for other target markets and workloads, such as video analytics.
Chapter 2. Software Support

This chapter lists the software deliverables, versions, and target applications currently supported by NVIDIA as "NGC-Ready."

Refer to the latest list of supported Tesla Recommended Drivers (TRDs), NVIDIA® CUDA® drivers, and containerized software at NVIDIA GPU Cloud Documentation for more details and updates.

**Note:** The T4 NGC-Ready program does not currently include NVIDIA GRID™ or IVA (video analytics) applications. Check the listed NGC website to get the most up-to-date list of containerized software supported by NGC.
Chapter 3. T4 NGC-Ready Platform Specification

This chapter provides the system configuration requirements for an NGC-Ready server using T4 GPUs. NVIDIA has defined three different platform configurations for a T4 NGC-Ready server:

- **Good** – Server configuration that meets minimum requirements with limited scale-out support
- **Better** – Server configuration that meets minimum requirements with adequate scale-out support
- **Best** – Recommended server configuration for optimal performance and scale-out support

Table 3-1. T4 NGC-Ready Platform Specification

<table>
<thead>
<tr>
<th>Platform Element</th>
<th>Good</th>
<th>Better</th>
<th>Best</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPU</td>
<td>Four T4 GPUs (two GPUs per socket)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPU</td>
<td>Xeon Gold dual-socket / Skylake or Cascade Lake</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CPU cores</td>
<td>16 cores per CPU</td>
<td>18 or more cores per CPU</td>
<td></td>
</tr>
<tr>
<td>CPU speed</td>
<td>2.1 GHz base clock (minimum)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>System memory</td>
<td>192 GB per socket / 6 DRAM channels / dual DIMM per channel</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Networking</td>
<td>One 10 Gbit NIC per server</td>
<td>One 25 Gbit NIC per socket</td>
<td>One 50 Gbit NIC per socket</td>
</tr>
<tr>
<td>Storage</td>
<td>One NVMe drive per system (either socket)</td>
<td>Two NVMe drives per system (one per socket)</td>
<td></td>
</tr>
</tbody>
</table>

Notes:
1. All GPUs should be connected to the host via PCIe Gen3 x16.
2. For best RDMA performance between the GPU and NIC, choose an NIC that supports NVIDIA® GPUDirect™, such as the Mellanox ConnectX-4 EN.
3. Best NIC performance is achieved with a PCIe x16 Gen3 connection, but if PCIe lanes are limited a x8 Gen3 connection is sufficient.
4. NVMe connections to PCIe can be x4 Gen3.

Note: See the system diagrams in Chapter 4 for the system topology details.
Chapter 4. System Topologies

The recommended system topologies for a T4 NGC-Ready server are shown in the following figures.

Figure 4-1 shows the “Best” NGC-Ready configuration with full multi-node support with 50 Gbit NICs and maximum storage.

- Dual-socket Xeon Gold with 192 GB system memory per socket.
- Two T4 GPUs per CPU socket.
- One 50 Gbit NIC per CPU socket.
  - Note that if a x8 PCIe connection is used for the NIC, then a PCIe switch is not needed, and the GPUs and NICs can be connected directly to the CPU. Using a x16 PCIe connection to each NIC may increase performance but will require a PCIe switch (not shown) due to a limitation in the number of PCIe lanes supported by the Xeon CPU.
  - For best NIC performance, choose an NIC that supports GPUDirect, such as the Mellanox ConnectX-4 EN.
- One NVME driver per socket.

Figure 4-1.  T4 NGC-Ready Server Topology – Best
Figure 4-2 shows the “Better” NGC-Ready configuration with adequate multi-node support using 25 Gbit NICs and maximum storage.

- Dual-socket Xeon Gold with 192 GB system memory per socket.
- Two T4 GPUs per CPU socket.
- One 25 Gbit NIC per CPU socket.
  - For best NIC performance, choose an NIC that supports GPUDirect, such as the Mellanox ConnectX-4 EN.
- One NVME driver per socket.

**Figure 4-2. T4 NGC-Ready Server Topology – Better**
Figure 4-3 shows the “Good” NGC-Ready configuration with minimal multi-node support and maximum storage. With minimal multi-mode NIC support, a server with this configuration may have performance issues when running larger DL and RAPIDs workloads.

- Dual-socket Xeon Gold with 192 GB system memory per socket.
- Two T4 GPUs per CPU socket. Note that while a x16 PCIe connection to each GPU is preferred, some workloads can operate with acceptable performance when using a x8 PCIe connection to each T4 GPU.
- One 10 Gbit NIC per server.
- One NVME driver per socket.

Figure 4-3. T4 NGC-Ready Server Topology – Good
Chapter 5. Design Discussion

5.1 PCI Express Interface

Most NGC-Ready platforms should be able to support direct x16 Gen3 PCIe connections from each CPU socket to the two T4 GPUs for best performance.

- A x16 Gen3 connection to each T4 GPU is recommended for optimal performance. Use of a x8 PCIe connection to the T4 GPU may result in performance loss for some workloads. More performance analysis may be needed to quantify the degree of performance degradation for some workloads.

- A x8 PCIe Gen3 connection to each 50 Gbit NIC is a minimum requirement for the best NGC-Ready server configuration. Using a x16 PCIe connection may result in better performance, but it would require use of a PCIe switch to accommodate all the PCIe devices (two GPUs, one NIC and one NVMe drive per CPU).

- If there are insufficient PCIe lanes to support direct connections to the GPUs and NICs, then a PCIe switch should be used. In this case, the GPUs and NICs should be located downstream of a common PCIe switch for optimal P2P and RDMA performance.

- The NVMe drives can connect to the host via a x8 Gen3 link.

5.2 CPU and System Memory

Currently, NVIDIA is only specifying a T4 NGC-Ready platform using Intel Xeon CPUs. A single CPU server configuration (for example with AMD ROME) may be released at a future date once more testing has been completed.

- For best performance, a minimum of 18 CPU cores for every two T4 GPUs is preferred. For better or good performance, 16 CPU cores for every two T4 GPUs is the minimum requirement. Xeon Gold CPUs can meet this requirement.

- The minimum system memory configuration for NGC-Ready workloads is 192 GB when using four T4 GPUs.

- NVIDIA’s minimum system memory configuration is 192 GB per socket, because of the target application requirements, and because memory bandwidth is maximized with six DRAM channels populated.
5.3 Network Interface

In general, the preferred NIC:GPU ratio is one 33 Gbit NIC for every pair of T4 GPUs. For the optimal (Best) system configuration, this results in one 50 Gbit NIC per socket. For the “Better” system configuration, one 25 Gbit NIC per socket is adequate, but multi-node workload performance may be impacted. For the “Good” system configuration, one 10 Gbit NIC per server is needed.

For best throughput between the NICs and GPUs, choose a NIC that supports GPUDirect, such as the Mellanox ConnectX-4 EN. If the server design uses PCIe switches to support its topology, the NIC and GPUs should be located under the same PCIe switch for best performance.

5.4 Storage

One NVMe SSD drive for each socket connected via x4 Gen3 PCIe is the best option for an NGC-Ready server using T4. One NVMe per server is good enough but performance may be impacted for some servers. Consult with NVIDIA directly if the use of SSD drives is desired.
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 regards 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, GPUDirect, NVIDIA GRID, 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

© 2019 NVIDIA Corporation. All rights reserved.