VPX3U-A4500-VO

NVIDIA Ampere 17.7 TFLOPS, four video outputs

IN DEVELOPMENT

KEY FEATURES

- NVIDIA RTX™ GA104, 5888 CUDA Cores, 184 Tensor Cores, 46 RT Cores, peak performance 17.66 TFLOPS*
- 16 GB GDDR6 256-bit memory with up to 512 GB/s
- Up to 4 outputs, options for DisplayPort/DVI/HDMI
- Module power: 70-130W, configurable

*Peak performance requires the highest power configuration mode.

GPU & FGX FEATURES

- Video inputs/outputs:
  - DP 1.4 output with support for 4K at 120Hz or 8K at 60Hz with 10-bit color depth
  - Support for HDMI and DVI outputs
- Ampere GPGPU parallel processing:
  - CUDA Toolkit 11, CUDA Compute capability 8.6
  - OpenCL™ 3.0, DirectX® 12 Ultimate, OpenGL 4.6, OpenGL ES 3.2, Vulkan™ 1.2
- 184 Tensor Cores (3rd Gen), 68 (dense) / 136 (sparse) Tensor TFLOP
- 46 Ray Tracing cores (2nd Gen)
- NVENC (7th Gen) and NVDEC (5th Gen) with up to 8K video encoding and hardware decoding support

CONNECTIVITY / SYSTEM MANAGEMENT

- IPMI system management
- NVIDIA GPUDirect RDMA support
- Windows and Linux drivers

MECHANICAL / OPEN SYSTEMS ARCHITECTURE

- High level of ruggedization:
  - Rugged conduction cooled
  - Operating temperature: -20° to +70°C
  - Vibration (sine wave): 10G peak, 5 - 2000Hz
  - Shock: 40G peak
- Dimensions: 160mm x 100mm x 25.4mm
- Weight (approximately): TBD
- ANSI/VITA 48, 65 (VPX-REDI, OpenVPX)
- SOSA Aligned Legacy Profile support 14.2.3

OVERVIEW

The VPX3U-A4500-VO module includes an NVIDIA RTX™ A4500 embedded GPU in a rugged 3U VPX module. The NVIDIA Ampere architecture includes CUDA cores for parallel processing, Tensor cores for dedicated AI-accelerated compute, and Ray Tracing cores for superior rendering speeds.

The NVIDIA Ampere architecture has introduced many significant improvements to the performance and efficiency of the GPU, with more flexible CUDA FP32/INT core use, more efficient third generation Tensor cores, and second generation RT cores. The Ampere GPU fabrication uses an 8nm manufacturing processing providing significant power improvements which, along with other Ampere architecture improvements, can provide up to 154 GFLOPS/W, providing almost twice the performance compared to the previous Turing generation's 86 GFLOPS/W or the Pascal generation's 62 GFLOPS/W.

Unlocking the best performance requires the best cooling capability. WOLF's advanced cooling technology is designed to move heat using a low weight, high efficiency path from the GPU die to the wedgelocks.

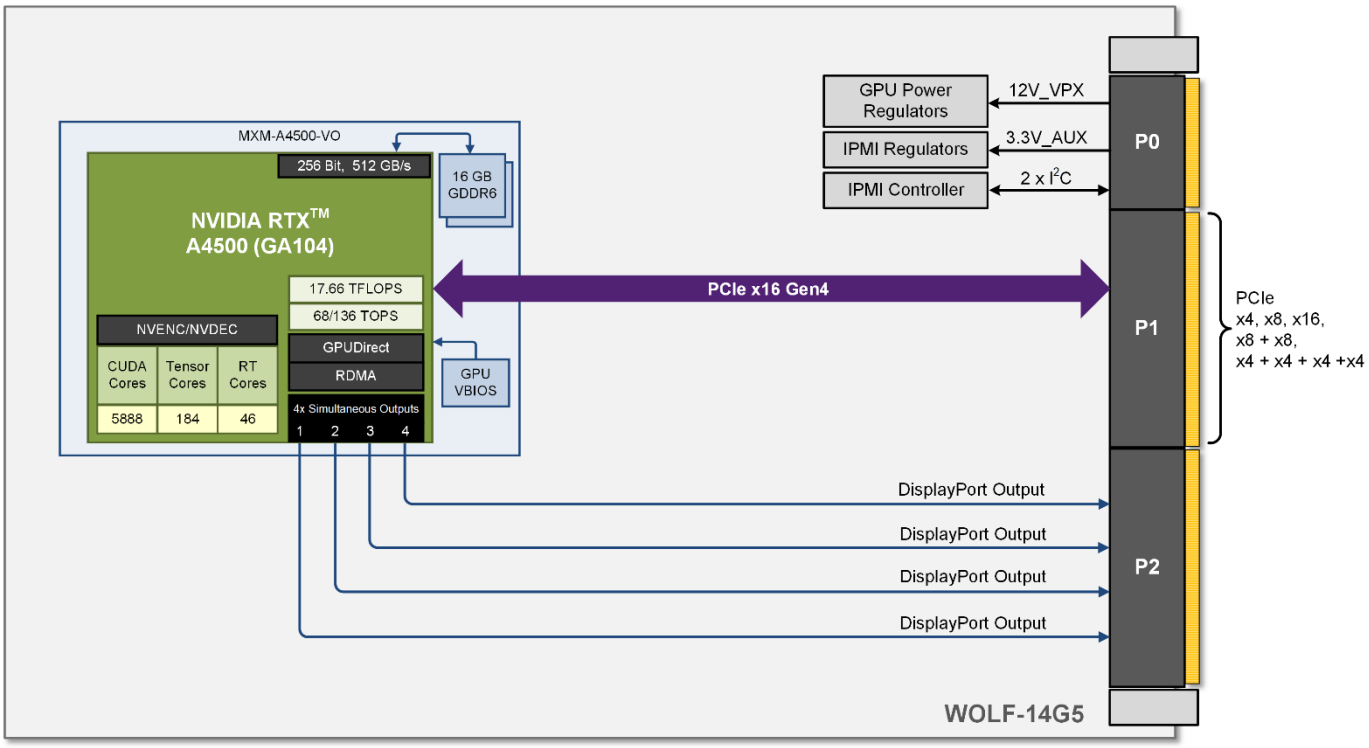
NVIDIA RTX

This information is subject to change

sales@wolf-at.com | International: +1 (905) 852-1163 | USA Toll-Free: 1 (800) 931-4114

IN DEVELOPMENT

The standard model supports DisplayPort, DVI-D or HDMI outputs.



This model integrates an industrial operating temperature NVIDIA Ampere MXM.

If you require support for a wider operating temperature of -40° to +85°C see:

- WOLF-1475, VPX3U-A2000E-WMXM-VO (with a WOLF MXM)
- WOLF-1488, VPX3U-A4500E-VO (a WOLF chip-down design, includes a PCIe switch)

This information is subject to change

IN DEVELOPMENT

NVIDIA AMPERE STREAMING MULTIPROCESSOR (SM)

Each NVIDIA Ampere architecture streaming multiprocessor (SM) partition contains CUDA cores for FP and INT operations, Tensor cores for AI, Ray Tracing (RT) cores for rendering, Texture Units, a register file, and L1/Shared Memory. Each previous generation Turing SM partition had two primary datapaths, with one able to process FP32 operations while the other was limited to integer operations. An Ampere SM partition's two primary datapaths can both process FP32 operations, with one datapath dedicated to FP32 operations and the other capable of executing either FP32 or integer operations. For operations which require only FP32 operations this doubles the number of available CUDA cores per SM. This change to the available functionality of the primary datapaths along with many other improvements to the other components in the Streaming Multiprocessor allows Ampere GPUs to provide significant performance improvements.

TENSOR CORES FOR ARTIFICIAL INTELLIGENCE AND HPEC

Tensor Cores are designed to speed up the tensor / matrix computations used for deep learning neural network training and inferencing operations. NVIDIA Ampere architecture GPUs include the third-generation Tensor Core design which supports many new data types for improved performance, efficiency, and programming flexibility, including a new sparsity feature and a new Tensor Float 32 (TF32) precision mode.

NVIDIA provides CUDA-X AI and CUDA-X HPEC libraires which have been designed to work with NVIDIA Tensor Core GPUs to provide the tools needed to accelerate development of applications for AI and HPEC.

HARDWARE ACCELERATED VIDEO ENCODE / DECODE

The Ampere GPU includes the NVENC video encode (version 7.2) and NVENC decode (version 5) hardware acceleration engine. Using the Ampere GPU for video encoding provides an efficient, high quality method to achieve real time 8K and 4K encoding without burdening the system CPU. The Ampere decoding engine includes support for several codecs, including AV1 hardware decoding support. The NVIDIA Video Codec SDK provides a complete set of APIs, samples and documentation for hardware accelerated video encode and decode.

OPENVPX AND SOSA SLOT PROFILE SUPPORT

This module's configurable PCIe interface provides support for several OpenVPX slot profiles. The module ICD is pin compatible with older generation 1116 modules, enabling a simple plug-in upgrade.

The module can also be configured to be SOSA aligned, with support for the SLT3-PAY-2F2U-14.2.3 Legacy Payload Slot Profile.

This information is subject to change

sales@wolf-at.com | International: +1 (905) 852-1163 | USA Toll-Free: 1 (800) 931-4114

VPX3U-A4500-VO

IN DEVELOPMENT

ORDERING CODES

The following table defines series of common order codes for the VPX3U-A4500-VO module. The asterisks denote characters of the part number that are defined based on common configuration options. Some configuration options for this module include:

- Display Interface configuration
- Conformal Coatings
- Variant Locked
- Default Power Threshold

Ordering Number	Description
3U VPX Ampere A2000 Single Slot Configurations, Industrial Temperature Operating Range	
14G533-F***-***VPX3vA0	3U VPX, Conduction Cooled, NVIDIA Ampere A4500, 16GB GDDR6, 4x DP/HDMI/DVI out
Related products with A2000, A1000, A500, Industrial Temperature Operating Range	
14H533-F***-***VPX3vA0	3U VPX, Conduction Cooled, NVIDIA Ampere A2000, 8GB GDDR6, 4x DP/HDMI/DVI out
14J533-F***-***VPX3vA0	3U VPX, Conduction Cooled, NVIDIA Ampere A1000, 8GB GDDR6, 4x DP/HDMI/DVI out

* Contact Sales for the latest Ordering Numbers and available options.

MANUFACTURING AND QUALITY ASSURANCE

WOLF designs modules to pass the following environmental standards:

- MIL-STD-810 (United States Military Standard for Environmental Engineering Considerations and Laboratory Tests)
- MIL-HDBK-217 (Reliability Prediction of Electronic Equipment)
- RTCA DO-160 (Environmental Conditions and Test Procedures for Airborne Equipment) on request

WOLF complies with the following management systems:

- AS9100D: Quality Management System - Requirements for Aviation, Space and Defense Organizations (certified)
- ISO 9001:2015: Quality management systems (certified)
- AS5553: Counterfeit Electronic Parts; Avoidance, Detection, Mitigation, and Disposition (compliant)
- NIST SP 800-171: Protecting Controlled Unclassified Information in Nonfederal Systems (compliant)

Boards are manufactured to meet the following standards:

- IPC-A-610 CLASS 3 (Acceptability of Electronic Assemblies)
- IPC 6012 CLASS 3 (Qualification and Performance Specification for Rigid Printed Boards, Class 3 for High Reliability Electronic Products)
- IPC J-STD-001 (Requirements for Soldered Electrical and Electronic Assemblies)

Caveat: Integrated third party modules may not meet the same standards as WOLF manufactured modules.

ITAR Free

This information is subject to change

Datasheet Rev.2

sales@wolf-at.com | International: +1 (905) 852-1163 | USA Toll-Free: 1 (800) 931-4114