VPX3U-A4500E-CX6

HPEC with NVIDIA Ampere and ConnectX-6

PRELIMINARY INFORMATION

KEY FEATURES

- NVIDIA RTX™ GA104 GPU with 17.66 TFLOPS FP32 peak performance*, 5888 CUDA Cores, 184 Tensor Cores, 46 RT Cores
- NVIDIA® ConnectX®-6 providing the module with up to 100GbE Ethernet and PCIe Gen4
- 16 GB GDDR6 256-bit memory with up to 512 GB/s
- Module power: 80-170W, configurable
  *Peak performance requires the highest power configuration mode.

GPU FEATURES

- Ampere GPGPU parallel processing:
  - ☐ CUDA Toolkit 11, CUDA Compute capability 8.6
  - ☐ CUDA-X AI and CUDA-X HPEC libraires
  - ☐ OpenCL™ 3.0, DirectX® 12 Ultimate, OpenGL 4.6, OpenGL ES 3.2, Vulkan™ 1.2
- 184 Tensor Cores (3rd Gen)
- 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

- PCIe Gen4 configurable interface, up to x16
- Daisy Chain and NTB options
- Backplane Ethernet with 10/40/100 GBASE-KR4 and 10GBASE-KR data planes; supports GPUDirect RDMA
- Block-level hardware encryption and the use of dedicated encryption keys per user
- Switching is offloaded from the CPU and run on the ConnectX hardware with NVIDIA ASAP^2 technology
- On-board IPMI controller for system management
- Windows and Linux drivers

MECHANICAL / OPEN SYSTEMS ARCHITECTURE

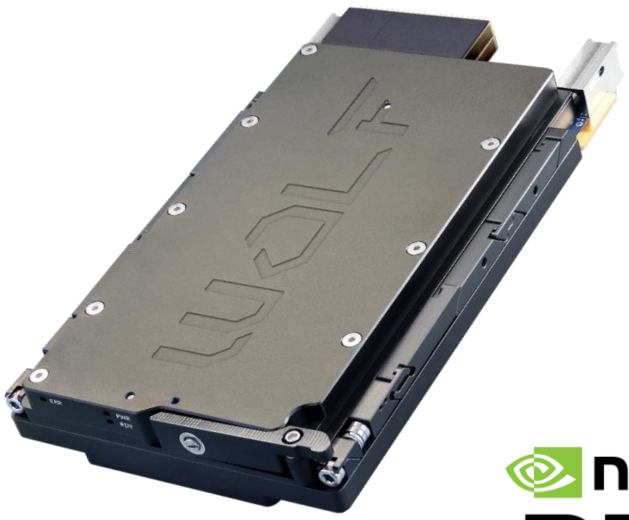
- High level of ruggedization:
  - ☐ Rugged conduction cooled
  - ☐ Operating temperature: -40° to +85°C
  - ☐ Vibration (sine wave): 10G peak, 5 - 2000Hz
  - ☐ Shock: 40G peak
- Dimensions: 160mm x 100mm x 25.4mm
- Weight (approximately): TBD
- SOSA Aligned options with depopulated P2 or P2B
- Support for the SOSA legacy payload slot profile

OVERVIEW

The VPX3U-A4500E-CX6 HPEC module includes an NVIDIA RTX™ A4500 embedded GPU and a ConnectX SmartNIC. The NVIDIA RTX™ A4500 embedded GPU provides the advanced processing capabilities for high performance embedded computing (HPEC) and artificial intelligence (AI) processing. The ConnectX-6 provides the Ethernet and PCIe connectivity needed to move large datasets efficiently.

The NVIDIA Ampere architecture includes CUDA cores for HPEC, 3rd generation Tensor cores for AI and data science computations, and 2nd generation Ampere Ray Tracing (RT) cores for visually accurate rendering. The Ampere architecture introduces support for PCIe Gen 4, improvements to memory handling, and improvements to warp scheduling, which all optimize the way data is distributed for processing. The Ampere GPU is very power efficient, delivering significantly more performance per watt compared to previous generations.

The NVIDIA ConnectX-6 SmartNIC provides PCIe and Ethernet connectivity. ConnectX-6 is ideal for the high-speed, secure, data transfer capabilities required for data-heavy tasks such as sensor data processing and other C4ISR tasks. The ConnectX-6 also provides support for RDMA over Converged Ethernet (RoCE), enabling the fastest method for transferring data across the network to the GPU.

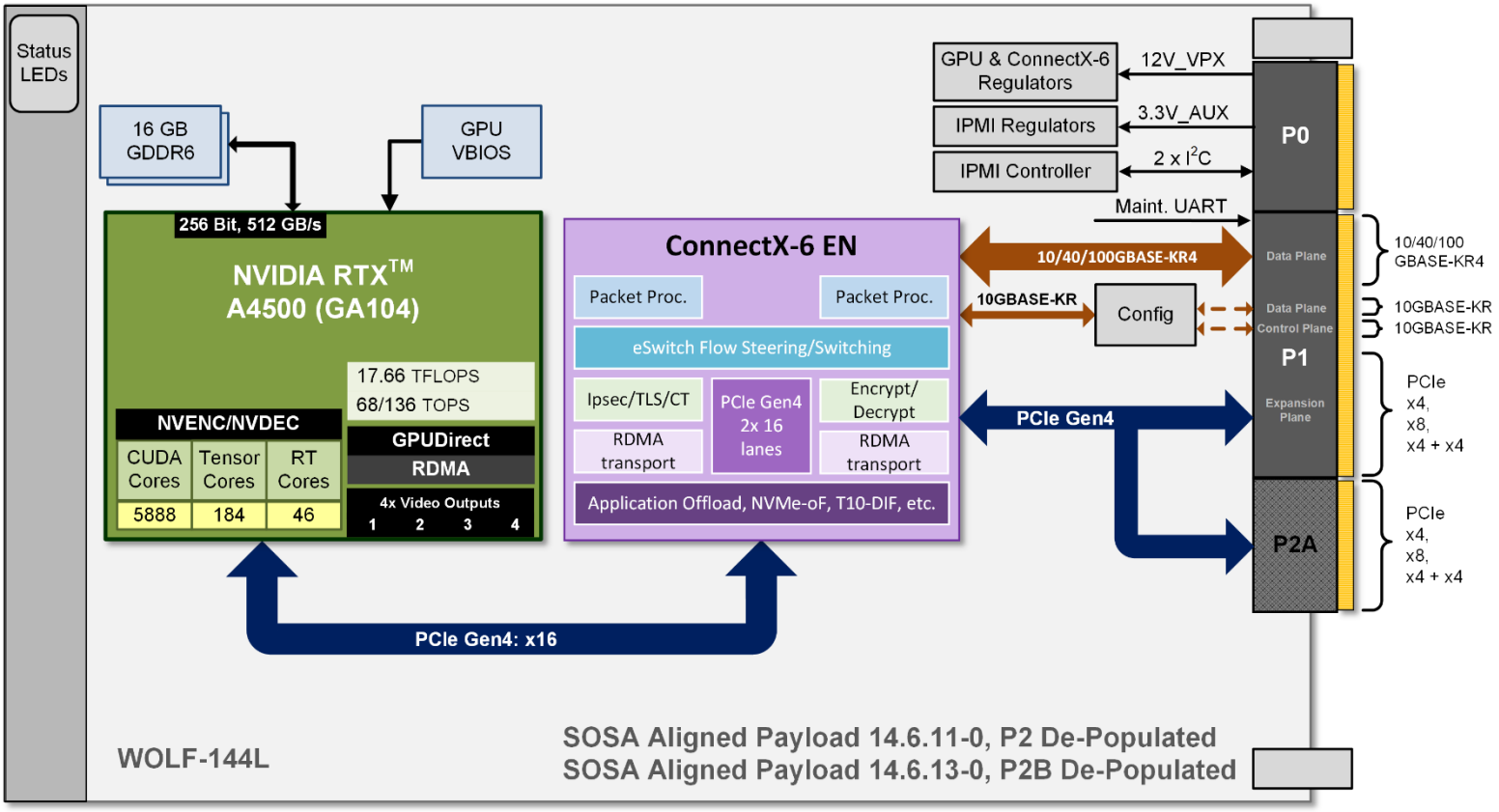


NVIDIA RTX

PRELIMINARY

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

PRELIMINARY INFORMATION



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.

PRELIMINARY

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.

ConnectX-6 PCIe Gen4 and Ethernet 100GbE

Getting large amounts of data into and out of a module is an important system design consideration. The WOLF-12T0 module includes a ConnectX-6 SmartNIC, which provides a configurable PCIe Gen4 interface. It also provides up to 100GBASE-KR4 on the data plane, RDMA over Converged Ethernet (RoCE) with support for NVIDIA GPUDirect RDMA, and enhanced security features such as hardware-verified secure boot, hardware-accelerated cryptography, and encrypted storage.

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 popular codecs and is the first GPU to include 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.

SOSA Slot Profile Support

This module's configurable switch provides support for SOSA aligned slot profiles. The module can be configured with pin mappings that are compatible with older generation WOLF-134C and WOLF-134S modules, which allows the WOLF-144L to be a plug-in upgrade for previous WOLF products.

The following SOSA aligned profiles are supported:

- SLT3-PAY-1F1U1S1S1U1U2F1H-14.6.11-0 Payload Slot Profile, P2 depopulated
- SLT3-PAY-1F1U1S1S1U1U4F1J-14.6.13-0 Payload Slot Profile, P2B depopulated

PRELIMINARY

PRELIMINARY INFORMATION

ORDERING CODES

The following table defines series of common order codes for the VPX3U-A4500E-CX6 module. The asterisks denote characters of the part number that are defined based on common configuration options. Some configuration options for this module include:

- Default Power Threshold
- Conformal Coatings
- PCIe Configuration Options
- Ethernet configuration
- Variant Locked
- P2 or P2B de-populated

Ordering Number	Description
3U VPX Ampere A4500 with CX6 Switch Single Slot Configurations	
144L33-F***-***VPX3vA0	3U VPX, Conduction Cooled, SOSA Payload profile with P2 depopulated, NVIDIA Ampere A4500, ConnectX-6
144L33-F***-***VPX3vA0	3U VPX, Conduction Cooled, SOSA Payload profile with P2B depopulated, NVIDIA Ampere A4500, ConnectX-6
Related Product:	
144833-F001-***VPX3vA0	3U VPX, Conduction Cooled, 1”, NVIDIA Ampere A4500, Rear IO: 4x DP++

* 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)

AS9100 • ISO 9001 CERTIFICATION™ Intertek

ITAR Free

GOVERNMENT-INDUSTRY DATA EXCHANGE PROGRAM GIDEP

PRELIMINARY

This datasheet is preliminary and is subject to change

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