VPX3U-AD5000E-CX7

HPEC with NVIDIA Ada RTX 5000 GPU and ConnectX-7

Key Features

- NVIDIA RTX™ 5000 (AD104) GPU with 9728 CUDA Cores, 304 Tensor Cores, 76 RT Cores
- NVIDIA® ConnectX®-7 provides the module with up to 100GbE Ethernet and a PCIe Gen5 switch
- 16 GB GDDR6 256-bit memory with up to 576 GB/s
- Module power: 100W - 140W, configurable

GPU Features

- Ada GPGPU parallel processing:
  - ☐ CUDA Toolkit 12, CUDA Compute capability 8.9
  - ☐ 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 (4th Gen)
- 46 Ray Tracing cores (3rd Gen)
- NVENC (8th Gen) and NVDEC (5th Gen) with up to 8K video encoding and hardware decoding support

Connectivity / System Management

- On-board IPMI controller for system management
- PCIe Gen5 configurable switch
- Switching is offloaded from the CPU to the ConnectX with NVIDIA ASAP^2 technology
- Daisy Chain options
- Support for 40/100GBASE-KR4 protocols
- 10GBASE-KR Data and Control planes
- GPUDirect RDMA and RoCE support
- Block-level hardware encryption and the use of dedicated encryption keys per user
- 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): 1.5 kg
- SOSA Aligned options with support for 14.6.11 or 14.6.13 slot profile

Overview

The VPX3U-AD5000E-CX7 HPEC module includes an NVIDIA RTX 5000 Ada embedded GPU and a ConnectX SmartNIC. The NVIDIA RTX 5000 Ada embedded GPU provides the advanced processing capabilities for high performance embedded computing (HPEC) and artificial intelligence (AI) processing. The ConnectX-7 provides the Ethernet and PCIe connectivity needed to move large datasets efficiently.

The NVIDIA Ada architecture includes CUDA cores for HPEC, 4th generation Tensor cores for AI and data science computations, and 3rd generation Ray Tracing (RT) cores for visually accurate rendering. The Ada GPU uses a new TSMC 4N NVIDIA Custom Manufacturing Process which to increased efficiency. The denser Ada GPUs have more CUDA and Tensor cores operating at higher clock frequencies at the same power, delivering significantly more performance per watt compared to WOLF's previous generation product. For example, comparing the 144L to the 153L, the 153L can achieve two to three times the GFLOPS/W (depending on power and temperature limits).

The NVIDIA ConnectX-7 SmartNIC provides PCIe and Ethernet connectivity. ConnectX-7 is ideal for the high-speed, secure, data transfer capabilities required for data-heavy tasks such as sensor data processing and other C5ISR tasks. The ConnectX-7 also provides support for RDMA over Converged Ethernet (RoCE), enabling the fastest method for transferring data across the network to the GPU.

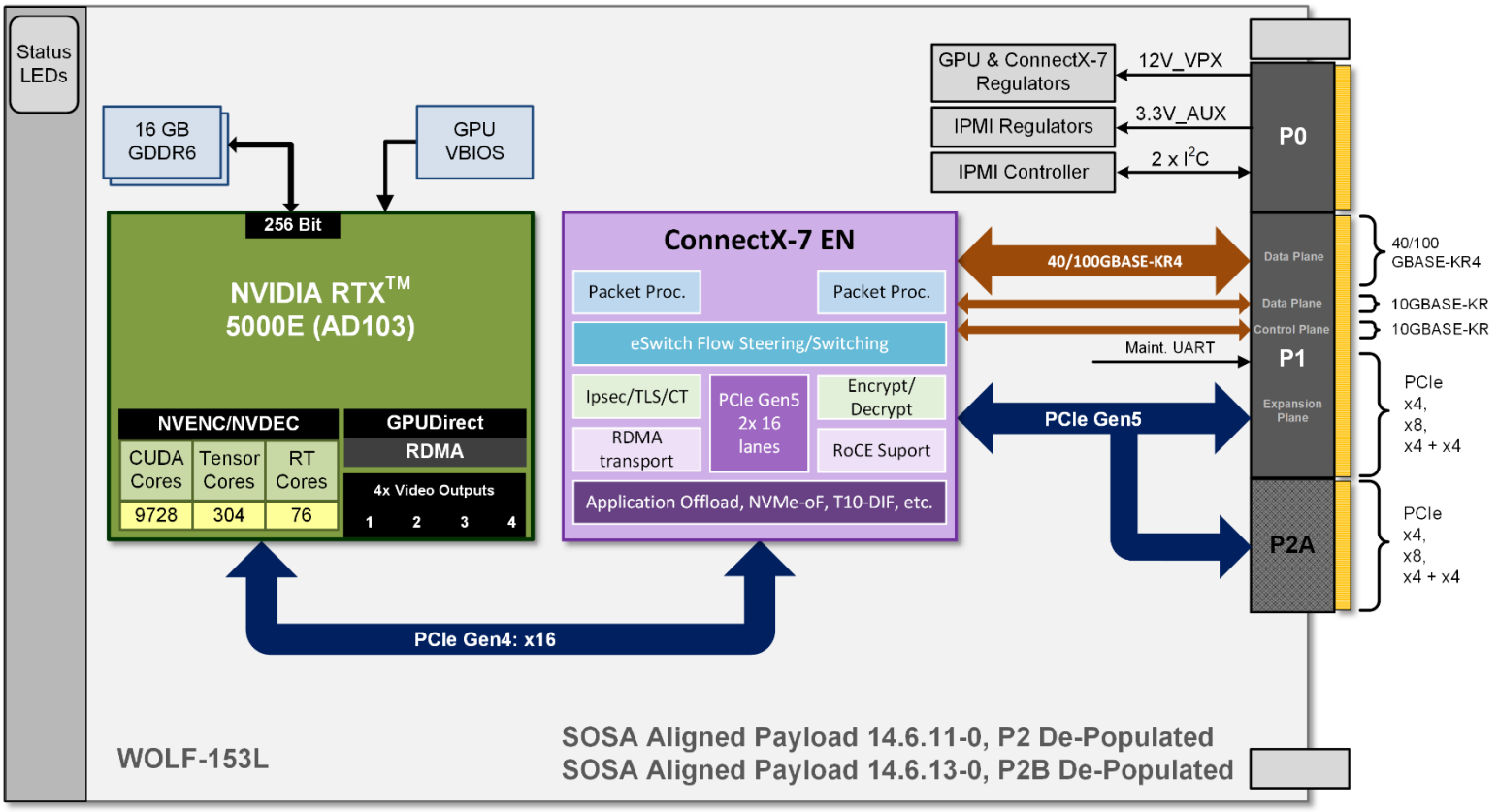
NVIDIA RTX

This information is subject to change

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

Chip-Down Design

The VPX3U-AD5000E-CX7 module uses a WOLF chip-down design to provide NVIDIA's advanced Ada architecture GPU and ConnectX SmartNIC technology on an extremely rugged board, making it an excellent choice for aerospace and defense applications. WOLF designs and manufactures modules in North America with full component traceability.



Power and Performance

An NVIDIA GPU will operate at the GPU clock speed available at the set TGP (total GPU subsystem power). The highest clock speeds are available at the highest TGP power allowed by the GPU. When the TGP setting is decreased the clock speed will also decrease resulting in a decrease in processing speed. The GPU base clock speed will also decrease if the GPU temperature exceeds 89°C to protect the GPU from heat damage. If the GPU temperature is below 87°C the GPU can operate at full boost clock speeds when higher processing is required.

The Ada AD103 GPU typically operates at TGP power levels from 80W to 115W. At 115W the base clock of 1425 MHz provides 27.7 TFLOPS FP32 performance while the boost clock of 2115 MHz provides 41.1 TFLOPS. At 80W the base clock of 930 MHz provides 18.1 TFLOPS while the boost clock of 1680 MHz provides 32.7 TFLOPS.

This information is subject to change

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

NVIDIA ADA GPU

NVIDIA Ada GPUs are manufactured using a new TSMC 4N NVIDIA Custom Manufacturing Process, an enhanced version of the N5 (5nm) node process. This allows a higher transistor density and lower voltage requirements, which provides increased efficiency. As a result, Ada GPUs have many more CUDA cores at the same die size as the previous generation, and higher clock speeds at the same power level, which leads to greatly increased processing/watt compared to the previous generation. The new Ada architecture also provides a big increase in the GPU's memory cache size, providing a boost to memory subsystem handling at the same bandwidth. With the increased performance and memory handling abilities, and improved next gen Tensor cores and RT Cores, the Ada GPUs are able provide significant performance increases compared to the previous generation.

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 Ada architecture GPUs include the fourth-generation Tensor Core design which supports many data types for improved performance, efficiency, and programming flexibility, including a sparsity feature, a Tensor Float 32 (TF32) precision mode, and a new FP8 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-7 PCIE GEN5 AND ETHERNET 100GBE

Getting large amounts of data into and out of a module is an important system design consideration. The WOLF-153L module includes a ConnectX-7 SmartNIC, which provides a configurable PCIe Gen5 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 Ada GPU includes the NVENC video encode and NVENC decode hardware acceleration engine. Using the GPU for video encoding provides an efficient, high quality method to achieve real time 8K and 4K encoding without burdening the system CPU. The Ada encoding engine includes support for several popular codecs and is the first GPU to include AV1 hardware encoding and 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-144L modules, which allows the WOLF-153L to be a plug-in upgrade for previous WOLF products.

The following SOSA aligned profiles are supported:

- 14.6.11-0 Payload Slot Profile, P2 depopulated
- 14.6.13-0 Payload Slot Profile, P2B depopulated

This information is subject to change

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

Ordering Codes

The following table defines series of common order codes for the VPX3U-AD5000E-CX7 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 Ada RTX5000 with ConnectX-7 SmartNIC Single Slot Configurations	
153L33-F***-***VPX3vA0	3U VPX, Conduction Cooled, NVIDIA Ada RTX5000, ConnectX-7, SOSA Payload profile with P2 depopulated
153L33-F***-***VPX3vA0	3U VPX, Conduction Cooled, NVIDIA Ada RTX5000, ConnectX-7, SOSA Payload profile with P2B depopulated
Related Product:	
153833-F005-***VPX3vA0	3U VPX, Conduction Cooled, 1”, NVIDIA Ada RTX5000, PCIe switch, SOSA Payload profile with P2 depopulated

* 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

This information is subject to change

Datasheet Rev.2