VPX6U-RTX5000E-DUAL-VO

NVIDIA Turing 19 TFLOPS, 768 Tensor Cores, 8 Video Outputs

Key Features

- Two NVIDIA Quadro Turing TU104 with 19 TFLOPS*, 6144 CUDA Cores, 768 Tensor Cores, 96 RT Cores
- 32 GB GDDR6 256-bit memory with up to 448 GB/s
- Up to 8 outputs, support for DP, HDMI and DVI
- Module power: 100-300W

*Peak performance requires the highest power configuration mode

GPU Features

- Eight DisplayPort 1.4 digital video outputs:
  - ☐ Support for High Dynamic Range (HDR) video
  - ☐ 4K at 120Hz or 8K at 60Hz with 10-bit color depth
  - ☐ HDMI and DVI options
- Turing GPGPU parallel processing:
  - ☐ CUDA Toolkit 10, CUDA Compute capability 7.5
  - ☐ OpenCL™ 1.2, DirectX® 12, OpenGL 4.6, OpenGL ES 3.2, Vulkan™ 1.0
- 768 Turing Cores for advanced AI inference
- 32 GB GDDR6 memory, up to 448 GB/s to each GPU
- HEVC (H.265) and AVC (H.264) Turing NVENC and NVDEC hardware acceleration with up to 8K encode resolution and B-frame support

Connectivity / System Management

- PCIe x16 Gen3; configurable PCIe switch
- Daisy Chain option supported
- NVLink® 2.0 x8 high-speed GPU interconnect option; provides 25 GB/sec peak bandwidth per direction between two GPUs (50 GB/sec bidirectional)
- Windows and Linux drivers
- On-board IPMI controller for system management

Mechanical / Open Systems Architecture

- High level of ruggedization:
  - ☐ Rugged Conduction cooled (CC) or Air cooled (AC)
  - ☐ Operating temperature: -40° to +85°C for CC, -40° to +71°C for AC
  - ☐ Vibration (sine wave): 10G peak, 5 - 2000Hz
  - ☐ Shock: 40G peak for CC, 30G peak for AC
- Dimensions: 160mm x 233mm x 25.4mm
- Weight: 1.7 kg for CC, 1.3 kg for AC
- ANSI/VITA 48, 65 (VPX-REDI, OpenVPX)
- SOSA Aligned options

Overview

The VPX6U-RTX5000E-DUAL-VO module includes two NVIDIA® Quadro® Turing™ RTX5000 embedded GPUs in a rugged 6U VPX module. The RTX5000 includes CUDA cores for parallel processing, Tensor cores for dedicated AI inference and ray tracing cores for superior rendering speeds.

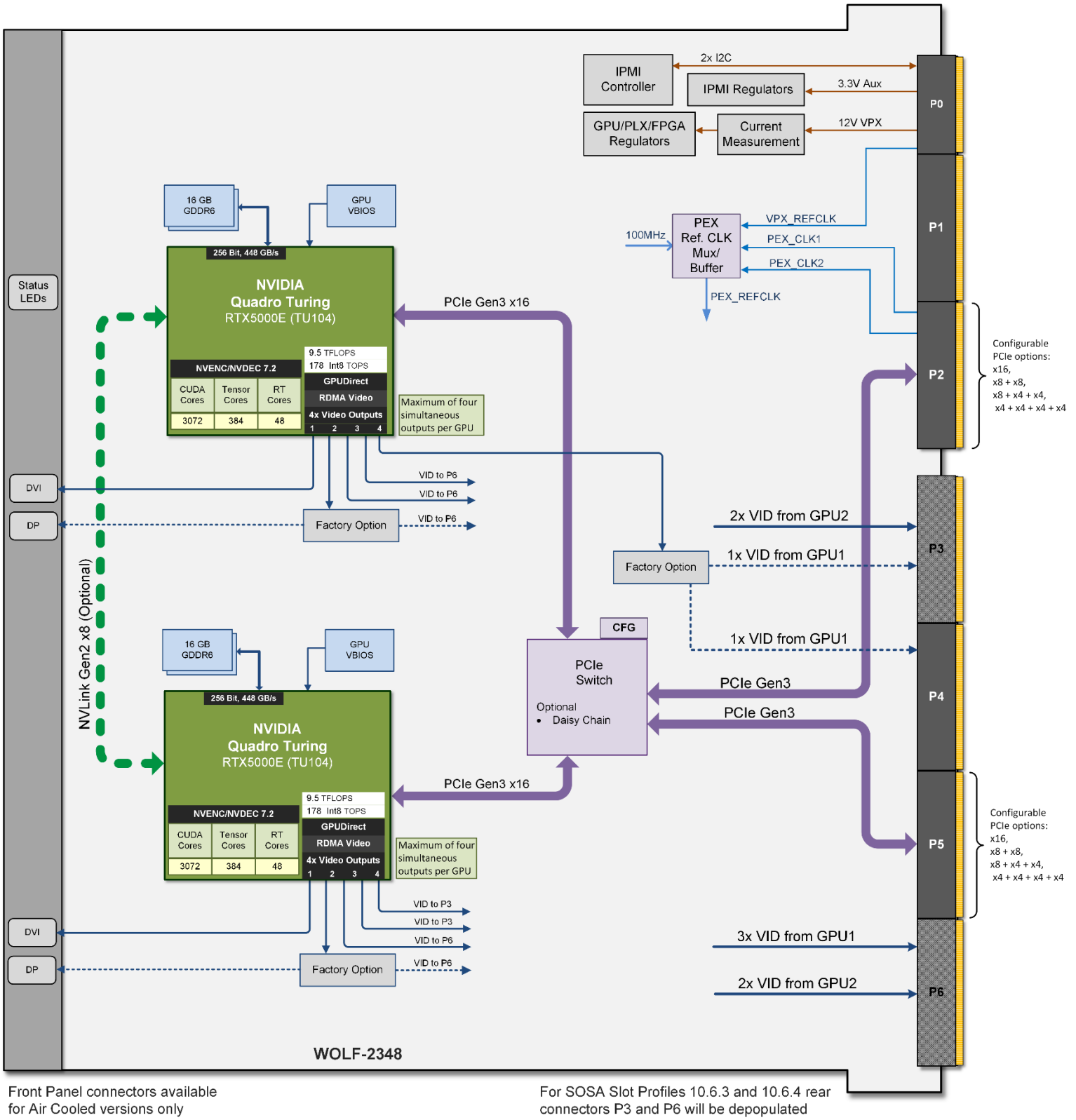
Getting data into and out of the module is an important consideration for this high-speed GPU. Support for GDDR6 memory provides twice the bandwidth of the previous generation's GDDR5 memory.

The Turing GPU with its Tensor cores provides this module with the underlying architecture required for an efficient AI inference engine. Intended to work in conjunction with TensorRT, CUDA and CuDNN, the Turing Tensor Core design adds INT8 and INT4 matrix operations, while continuing to support FP16 for higher precision workloads.

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 pipeline from the GPU die to the wedgelocks.

NVIDIA QUADRO CERTIFIED
▪ EMBEDDED

VPX6U-RTX5000E-DUAL-VO



Front Panel connectors available for Air Cooled versions only

For SOSA Slot Profiles 10.6.3 and 10.6.4 rear connectors P3 and P6 will be depopulated

NVIDIA Turing Streaming Multiprocessor (SM)

The NVIDIA Turing architecture provides a 50% improvement in delivered performance per CUDA core compared to the Pascal generation. This is due to the new Turing Streaming Multiprocessor's independent integer datapath, allowing execution of concurrent integer and floating-point instructions, and the redesigned memory path which provides two times the bandwidth and more than two times the capacity for common workloads.

Fast GDDR6 Memory

Getting data into and out of a high performance GPU requires fast graphics memory to ensure that the memory does not become a system bottleneck. In moving from GDDR5 to GDDR6 the number of data transfers per clock cycle doubled from two to four, and memory chips can be read in dual-channel modes rather than just single channel modes. The newer GDDR6 memory does all of this while also slightly reducing the memory's average power consumption compared to using GDDR5 memory.

NVIDIA also uses memory compression technology, especially data color compression for reducing the amount of graphical information that needs to be transmitted. With Turing and the GDDR6 memory the latest generation compression technology provides a 20 to 30% memory compression efficiency increase.

NVIDIA Tensor Cores for Artificial Intelligence and HPC

Tensor Cores are designed to speed up the tensor / matrix computations used for deep learning neural network training and inferencing operations. Turing GPUs include a new version of the Tensor Core design that has been enhanced for inferencing. Turing Tensor Cores add new INT8 and INT4 precision modes for inferencing workloads that can tolerate quantization and do not require FP16 precision.

NVIDIA provides CUDA-X AI and CUDA-X HPC libraires which are specialized libraires built on top of CUDA. They have been designed to work with NVIDIA Tensor Core GPUs to provide the tools needed to accelerate development of applications for AI and HPC.

Hardware Accellerated Video Encode/Decode

The RTX5000E chip includes the latest generation video encode/decode hardware acceleration engine (version 7.2). This adds support for HEVC (H.265) 8K encoding at 30 fps and B-Frame support. It also provides up to 25% bitrate savings for HEVC and up to 15% bitrate savings for AVC (H.264). Using the Turing encoding engine for video encoding provides an efficient, high quality method to achieve real time 8K and 4K encoding without burdening the system CPU.

NVENC supports CBR and VBR rate control, programmable intra-refresh for error resiliency, and a motion estimation (ME) only mode. The NVIDIA Video Codec SDK provides a complete set of APIs, samples and documentation for hardware accelerated video encode and decode on Windows and Linux.

Designed for System Integration

The VPX architecture is diverse, spanning custom backplanes, an ambiguous system specification and differing input and output methodologies. That is precisely why WOLF modules come with factory configuration options to solve system integration challenges.

The WOLF-2348 module has a configurable PCIe interface that supports several OpenVPX slot profiles. It can also be configured with pin mappings that are compatible with older generation WOLF-2118 modules, which allows the WOLF-2348 to be a plug-in upgrade for those WOLF products.

The module can also be configured for SOSA aligned slot profiles with rear connectors P3 and P6 depopulated. Please contact WOLF to discuss your system level requirements.

ORDERING CODES

The following table defines series of common order codes for the VPX6U-RTX5000E-DUAL-VO module. The asterisks denote characters of the part number that are defined based on common configuration options. Some common configuration options for this module are:

- Display Interfaces
- Conformal Coating Type
- COTS, MCOTS or Locked
- PCIe Switch Configuration
- SOSA Aligned configurations
- NVLink for dual GPU modules

	Ordering Number	Description
	6U VPX Turing RTX5000 Single Slot Configurations	
	234823-F***-***VPX6vA0	6U VPX, Air Cooled, 1”, Dual NVIDIA Turing RTX5000, 8x video outputs
	234833-F***-***VPX6vA0	6U VPX, Conduction Cooled, 1”, Dual NVIDIA RTX5000, 8x video outputs
	234823-F***-***VPX6vA0	6U VPX, Air Cooled, 1”, Single NVIDIA Turing RTX5000, 4x video outputs
	234833-F***-***VPX6vA0	6U VPX, Conduction Cooled, 1”, Single NVIDIA RTX5000, 4x video outputs

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