All Quadra VPU Products

Smart VPU | Codensity Quadra G5

	2nd Generation Smart VPUs			
	Server	**Modules**		
	G5 Smart VPU **Quadra**	Smart VPU **T1U**	Smart VPU **T1A**	Smart VPU **T2A**
Performance				
ASIC Codensity chip	G5, T1Us (10x)	G5	G5	G5 (2x)
Price	starting at $19,000	$1,500	$1,500	$2,750
Form Factor	1RU Server	U.2	AIC, HHHL	AIC, HHHL
Power Consumption	~500W	17W	20W	40W
Real-time Throughput Up to:	320x 1080p30 80x 4Kp30 20x 8Kp30	32x 1080p30 8x 4Kp30 2x 8Kp30	32x 1080p30 8x 4Kp30 2x 8Kp30	64x 1080p30 16x 4Kp30 4x 8Kp30
Latency	8 ms	8 ms	8 ms	8 ms
Encode Codecs	H.264, HEVC, JPEG, YUV, AV1			
Decode Codecs	H.264, HEVC, JPEG, YUV, VP9			
Audio Codecs	MP3, AAC-LC, HE-AAC			
Features				
Artificial Intelligence	150 TOPS	15 TOPS	18 TOPS	36 TOPS
New Capped CRF	●	●	●	●
Flexible GOP	●	●	●	●
Scaling	●	●	●	●
Cropping and Padding	●	●	●	●
Video Overlay	●	●	●	●
YUV / RGB Conversion	●	●	●	●
Configurable throughput	●	●	●	●

● Feature supported on VPU

NETINT Technologies
www.netint.com
sales@netint.com

Artificial Intelligence Specs

AI Deep Neural Network Inference Engines

INT8 Trillion Operations Per Second (TOPS)

- T1U: 15 TOPS
- T1A: 18 TOPS
- T2A: 36 TOPS

AI Deep Learning Frameworks

Support models trained with these major deep learning frameworks:

- Caffe
- Darknet
- Keras
- ONNX
- PyTorch
- TensorFlow
- TensorFlow Lite

Applications for Quadra AI Inference Engine include

- ROI-Encoding
- Scene Detection
- Background Removal
- Video Enhancement
- Facial Recognition
- Object Detection

Deployment Workflow for Pre-trained AI models

AI deep learning models are imported to Quadra VPUs with the NETINT AI Toolkit then processed (Import, quantization, validation and optimization), exported, and executed on Quadra Neural Processing Units (NPUs).

Supported Frameworks

- ONNX
- TFLite
- PyTorch
- Darknet
- TensorFlow
- Keras

Features

- 8 /16 bit quantization
- Hybrid quantization
- Accuracy validation
- Graph optimization
- Pre-processing integration

Features

- Hardware-aware optimization
- Execution graph generator
- Performance profiling
- Python/C Inference API

Specification	Input Size	Performance FPS @ 1 GHz
Yolov5s	640x640	78
Yolov5s	320x320	231
Yolov4-tiny	416x416	276
ResNet 50	224x224	228
MobileNetv2	224x224	1234
FSRCNNx3	360x640	36
DeepLabv3	257x257	452
BiSeNetV1	512x512	51
HrNet	256x192	72

Specifications

- Test hardware: T1A
- Test firmware version: 3.1
- AI capability per G5 ASIC: 18 TOPS
- Datatype for evalutation: INT8
- Batch size: 1
- Performance based on original model without pruning, sparsity or modification
- Quadra supports multiple AI modes (Full, Eco, Off) depending on the power/performance requirement

NETINT

Quadra T2A
Smart VPU

AI powered Video Processing Unit | Codensity G5

Form Factor	AIC (HH HL)
ASIC	2x Codensity G5
Interface	PCIe 4.0 x4x4
Power Consumption (Typ)	40W
Usage	24/7 Operation
Operation Temperature	0 - 50°C
RoHS Compliance	European Union (EU) ROHS Compliance Directives
Product Health Monitoring	Self-Monitoring, Analysis, and Reporting Technology (SMART) commands Temperature Monitoring and Logging
Video Encoding Standards/Formats	AVC/H.264 Baseline, Main, High, High 10 HEVC/H.264 Main, Main 10 JPG YUV 420 8 bit/10 bit encoding AV1 Main
Video Decoding Standards/Formats	AVC/H.264 Baseline, Main, High, High 10 HEVC/H.265 Main, Main 10 VP9 Profile 0, 2 JPEG YUV 420 8 bit/10 bit decoding
Throughput Capacity	Up to 64x 1080p30, 16x 4Kp30, 4x 8Kp30
Audio Codecs	MP3, AAC-LC, HE-AAC
Level	1 to 6.2 Main Tier
Resolution	32 x 32 to 8192 x 5120
Scan Type	Progressive
Bitrate	64kbit/s to 700Mbit/s
Software Integration	FFmpeg SDKs, GStreamer, LibXcoder API integration
AI Deep Neural Network Engines	36 TOPS AI Assisted Encoding
Region of Interest (ROI)	ROI enables the quality of some regions to be improved at the expense of other regions
Closed Captioning	EIA CEA-708 for H.264 and HEVC encode/decode
High Dynamic Range (HDR)	HDR10, HDR10+, HLG for H.264 & HEVC encode/decode
Low Latency	Sub-frame latency
IDR Insert	Forced IDR frame inserts at any location
Flexible GOP Structure	8 presets plus customizable GOP structure
Video 2D Processing Engine	Crop & Padding/Scaling/Overlay/YUV & RGB Conversion

NETINT

NETINT Technologies
www.netint.com
sales@netint.com

Quadra T1A
Smart VPU

AI powered Video Processing Unit | Codensity G5

Form Factor	AIC (HH HL)
ASIC	1x Codensity G5
Interface	PCIe 4.0 x4
Power Consumption (Typ)	20W
Usage	24/7 Operation
Operation Temperature	0 - 50°C
RoHS Compliance	European Union (EU) ROHS Compliance Directives
Product Health Monitoring	Self-Monitoring, Analysis, and Reporting Technology (SMART) commands Temperature Monitoring and Logging
Video Encoding Standards/Formats	AVC/H.264 Baseline, Main, High, High 10 HEVC/H.264 Main, Main 10 JPG YUV 420 8 bit/10 bit encoding AV1 Main
Video Decoding Standards/Formats	AVC/H.264 Baseline, Main, High, High 10 HEVC/H.265 Main, Main 10 VP9 Profile 0, 2 JPEG YUV 420 8 bit/10 bit decoding
Throughput Capacity	Up to 32x 1080p30, 8x 4Kp30, 2x 8Kp30
Audio Codecs	MP3, AAC-LC, HE-AAC
Level	1 to 6.2 Main Tier
Resolution	32 x 32 to 8192 x 5120
Scan Type	Progressive
Bitrate	64kbit/s to 700Mbit/s
Software Integration	FFmpeg SDKs, GStreamer, LibXcoder API integration
AI Deep Neural Network Engines	18 TOPS AI Assisted Encoding
Region of Interest (ROI)	ROI enables the quality of some regions to be improved at the expense of other regions
Closed Captioning	EIA CEA-708 for H.264 and HEVC encode/decode
High Dynamic Range (HDR)	HDR10, HDR10+, HLG for H.264 & HEVC encode/decode
Low Latency	Sub-frame latency
IDR Insert	Forced IDR frame inserts at any location
Flexible GOP Structure	8 presets plus customizable GOP structure
Video 2D Processing Engine	Crop & Padding/Scaling/Overlay/YUV & RGB Conversion

NETINT

NETINT Technologies
www.netint.com
sales@netint.com

Quadra T1U
Smart VPU

AI powered Video Processing Unit | Codensity G5

Form Factor	U.2
ASIC	1x Codensity G5
Interface	PCIe 4.0 x4
Power Consumption (Typ)	17W
Usage	24/7 Operation
Operation Temperature	0 - 50°C
RoHS Compliance	European Union (EU) ROHS Compliance Directives
Product Health Monitoring	Self-Monitoring, Analysis, and Reporting Technology (SMART) commands Temperature Monitoring and Logging
Video Encoding Standards/Formats	AVC/H.264 Baseline, Main, High, High 10 HEVC/H.264 Main, Main 10 JPG YUV 420 8 bit/10 bit encoding AV1 Main
Video Decoding Standards/Formats	AVC/H.264 Baseline, Main, High, High 10 HEVC/H.265 Main, Main 10 VP9 Profile 0, 2 JPEG YUV 420 8 bit/10 bit decoding
Throughput Capacity	Up to 32x 1080p30, 8x 4Kp30, 2x 8Kp30
Audio Codecs	MP3, AAC-LC, HE-AAC
Level	1 to 6.2 Main Tier
Resolution	32 x 32 to 8192 x 5120
Scan Type	Progressive
Bitrate	64kbit/s to 700Mbit/s
Software Integration	FFmpeg SDKs, GStreamer, LibXcoder API integration
AI Deep Neural Network Engines	15 TOPS AI Assisted Encoding
Region of Interest (ROI)	ROI enables the quality of some regions to be improved at the expense of other regions
Closed Captioning	EIA CEA-708 for H.264 and HEVC encode/decode
High Dynamic Range (HDR)	HDR10, HDR10+, HLG for H.264 & HEVC encode/decode
Low Latency	Sub-frame latency
IDR Insert	Forced IDR frame inserts at any location
Flexible GOP Structure	8 presets plus customizable GOP structure
Video 2D Processing Engine	Crop & Padding/Scaling/Overlay/YUV & RGB Conversion

NETINT

NETINT Technologies
www.netint.com
sales@netint.com

Quadra
Video Server

Smart VPU | Codensity Quadra G5

CPU Options	AMD EPYC™ 7232P Server Processor (8-core)
	AMD EPYC 7543P Server Processor (32-core)
	AMD EPYC 7713P Server Processor (64-core)
Operating System	Ubuntu 20.04.05 LTS
Memory	8x 16GB DDR4-3200
Storage	400GB M.2 SSD
NVMe Support	10x
PCIe Expansion	Up to 3x PCIe slots
Network Options	Dual 10GBase-T LAN
Power Consumption	~500W
Power Supply	700W: 100 - 140Vac
	750W: 200 - 240Vac
	750W: 200 - 240Vdc (CCC only)
Transcoders	10x NETINT Quadra T1U
Encoding Capacity	Up to 20x 8Kp30, 80 4Kp30 or 320x 1080p30
Codec Support	H.264 - Encode/Decode
	HEVC - Encode/Decode
	JPG - Encode/Decode
	VP9 - Decode
	AV1 - Encode
Software Integration	FFmpeg, GStreamer

Physical Dimensions	W: 17.2” (437mm), H: 1.7” (43mm), D: 23.5” (597mm)
Rack Size	1U
Weight	39 lbs (17.69 kg) *(fully loaded with 10 T1U VPUs)*
Environmental	50 degrees F to 95 degrees F Operating Temperature, 8% to 90% Operating Relative Humidity
Power Inputs	100 - 140Vac / 8 - 6V / 50-60Hz
	200 - 240Vac / 4.5 - 3.8A / 50-60Hz
	200 - 240Vdc / 4.5 - 3.8A (CCC Only)
Certifications	RoHS Compliant, UL Approved

NETINT

NETINT Technologies
www.netint.com
sales@netint.com