VPX-REDI
(OpenVPX)

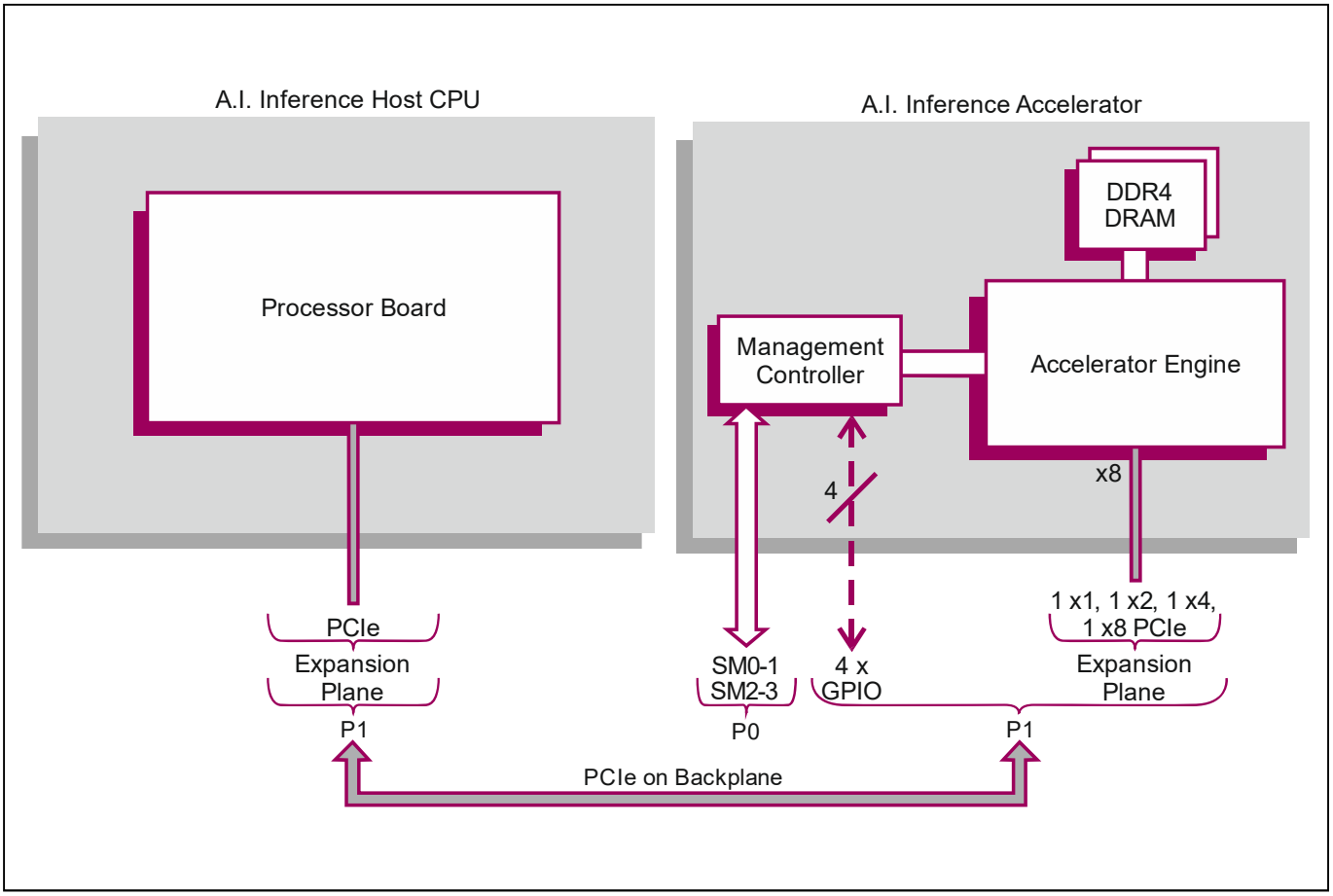
TR AEx/3sd-RCx
RCR - Series

3U VPX-REDI™ Artificial Intelligence Accelerator Board

Key Features

TR AEx/3sd-RCx is a rugged, conduction-cooled 3U VPX artificial intelligence accelerator board. Paired with a Concurrent Technologies processor board, TR AEx/3sd-RCx is designed to boost performance of inference at the edge applications in the defence, exploration and transportation markets.

- Companion to TR H4x/3sd-RCx
- Supported by the Intel® OpenVINO™ toolkit
- Includes pre-trained models
- Supports popular frameworks like Caffe, TensorFlow, MXNet



CONCURRENT TECHNOLOGIES

Concurrent Technologies Plc 4 Gilberd Court, Colchester, Essex, CO4 9WN, UK
Tel: +44 (0)1206 752626

Concurrent Technologies Inc. 400 West Cummings Park, Suite 1300, Woburn, MA 01801, USA
Tel: (781) 933 5900

email:info@gocct.com **www.gocct.com**

powerBridge Computer

Ehlbeek 15a
30938 Burgwedel
fon 05139-9980-0
fax 05139-9980-49

www.powerbridge.de
info@powerbridge.de

TR AEx/3sd-RCx Inference Accelerator Engine

The TR AEx/3sd-RCx acts as a complementary accelerator for TR H4x/3sd-RCx rugged compute cards. Differing neural network algorithms prefer different hardware architectures and the TR AEx/3sd-RCx can significantly increase inference throughput by providing a dedicated solution to offload neural network inference for many algorithms. Below are its key advantages:

- High throughput versus CPU only AI inference for many models
- Significant performance per watt advantage versus CPU/GPU only AI inference
- Dedicated processing hardware frees up CPU resources
- Low latency architecture suitable for real time applications
- 3U VPX form factor enabling a compact embedded solution
- Reprogrammable hardware allows for changing requirements and adaptation to new neural network models

Deep Learning

Historically, machine learning would require researchers and domain experts to manually design complex filters to extract trends and features from data. Today, deep learning algorithms and accelerators can be deployed to rapidly and effectively train models to recognise new and differing input data.

Deep learning is a natural evolution of machine learning which enables ever complex neural network models to analyse and evaluate real world problems. Deep learning models have multiple internal layers of neurons which can be trained to solve problems such as:

- Object recognition
- Object detection
- Feature segmentation

Neural network models can often garner a higher accuracy than human judgement, making them a valuable tool in mission critical applications.

Inference at the edge

Inference is the process of using a trained neural network to sense, reason and act upon outcomes based on given stimuli. Traditionally, inference takes place in large datacenters, where data captured in the field must be transported to the datacentre in order for it to be processed. This incurs penalties such as:

- Significant latency from data collection to inference resulting in inaccurate situation reports and assessments
- Unnecessary loading of servers and connections to download and upload data
- Security and privacy concerns where data must be transported in order for it to be processed

Inference at the edge enables near instant output from trained neural network models from within deployed hardware, to provide excellent quality actionable intelligence.

Key factors in this are:

- Low latency; due to local processing
- Less concern about connectivity, bandwidth or loading
- Secure; data can only be accessed on the device and deployed platform

OpenVINO™ Toolkit

The Open Visual Inference and Neural Network Optimisation (OpenVINO) toolkit enables TR AEx/3sd-RCx to process Convolutional Neural Networks (CNN) quickly and efficiently. It provides a collection of tools to easily optimise, deploy and analyse the neural network model even with little knowledge of the hardware architecture and AI itself.

At the heart of OpenVINO is the Deep Learning Deployment Toolkit which consists of:

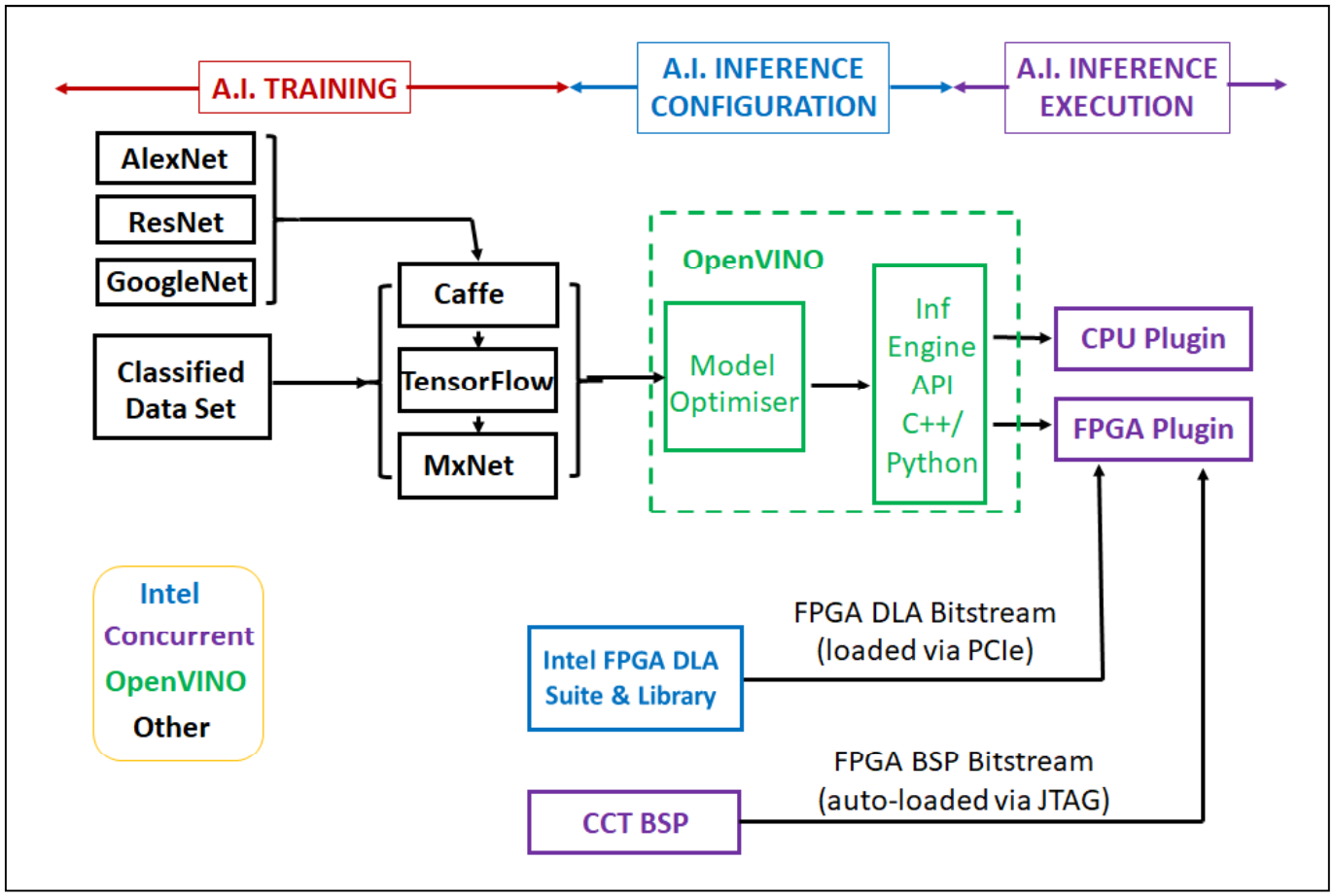
1. Model Optimiser; Converts and optimises trained models into Intermediate Representation binaries such that they are understood by and performant on the Accelerator Engine
2. Inference Engine; Loads models onto the Accelerator Engine and manages Inference heterogeneously between CPU and Inference Accelerator Engine using an OpenCL backbone.
3. Output; Presents results in text, image and video formats with OpenVX™ and OpenCV

OpenVINO supports models trained in popular frameworks such as:

- TensorFlow
- Caffe
- MXNet

With a wide range of supported topologies including:

- AlexNet
- MobileNet
- ResNet
- GoogleNet
- Inception
- SSD
- SqueezeNet
- YOLO



VPX-REDI Embedded Computer Board

- conduction-cooled 3U VPX-REDI AI Accelerator Engine utilizing the Intel Arria 10 GX FPGA family
- OpenVPX profile PCIe on Expansion Plane compatible with:
  - SLT3-PAY-1F1U1S1S1U1U2F1H-14.6.11-0
  - SLT3-PER-1Q-14.3.4
  - based on ANSI/VITA 65.0-2019 and ANSI/VITA 65.1-2019

Processing Element

- Intel Arria 10 GX FPGA

DRAM

- 8 Gbytes soldered DDR4 ECC DRAM:
  - dual channel architecture

Optional Maintenance Serial Port

- 1 x maintenance port configurable via P1 utility plane:
  - supports RS232 Tx/Rx signals
  - 16550 compatible UARTs
  - option: pins configurable as 2xGPIO

Other Peripheral Interfaces

- 2 x GPIO signals via P1 utility plane
- option for GPIO to be connected to GND

VPX Expansion Plane, PCI Express

- PCIe VPX Expansion Plane interface (VITA 65) supports:
  - 1 x8 PCIe port
  - compatible with OpenVPX module profiles
- PCIe interface supports Gen 1, Gen 2 and Gen 3

Management

- VITA 46.11 IPMC on board controller:
  - SM0-1 and SM2-3
  - temperature, voltage and current sensors accessed via System Management interface

Board Security Features

- option for proprietary board-level security features

Software Support

- support for Linux
- for other operating systems contact Concurrent Technologies for further information

Safety

- PCB (PWB) manufactured with flammability rating of UL94V-0

Electrical Specification (Estimated)

- typical current consumption for AI Accelerator Engine, 8 Gbytes DRAM:
  - +12V VS1 @ 4.5 A
  - +3.3V AUX @ 0.5A

Environmental Specification

- conduction-cooled (VITA 48.2)
- operating temperature at card edge:
  - VITA 47 Class CC4, -40 C to +85 C
- non-operating temperature:
  - VITA 47 Class C4, -55 C to +105 C
- operating altitude:
  - -1,000 to 50,000 feet (-305 to 15,240 meters)
- 5% to 95% Relative Humidity, non-condensing

Mechanical Specification

- 3U VPX form-factor (VITA 46.0, VITA 48.0): 3.9 inches x 6.3 inches (100mm x 160mm)
- slot width (VITA 48.0):
  - 1.0 inch VPX-REDI Type 1, RCR-Series Type 1 Extended Covers Two Level Maintenance (VITA 48.2)
- connectors to VITA 46.0 for P0 and P1
- captive screws available to secure front handles
- operating mechanical:
  - shock - VITA 47 Class OS2, 40g
  - random vibration - VITA 47 Class V3, 0.1g /Hz

Please contact your local Concurrent Technologies sales office for further details on board build options and accessories.

Datasheet Code 1830/0120