cadence

и и о ани точи ини Л 🚽 🛄 🖼 🔗 🔞 🚺

SLOW

Tensilica DNA Processor IP For Al Inference

Industry-leading performance and power efficiency for on-device AI applications

The Cadence[®] Tensilica[®] DNA Processors are designed as standalone deep neural network accelerator (DNA) processors to serve demanding and continuously growing artificial intelligence (AI) applications spanning internet of things (IoT), smart home, smartphones, augmented reality (AR) /virtual reality (VR), drones, robotics, surveillance, and automotive – both ADAS and autonomous vehicles (AV).

Overview

On-device AI processing needs span a wide range both in the processing requirements and the types of deep neural networks. There is a need for the processors to be standalone units and scale from low-end IoT applications to high-end automotive applications. The Tensilica DNA processors, including the DNA 100 processor and the Vision C5 DSP, meet these challenges.

Our DNA 100 processor architecture incorporates a hardware engine and a Tensilica DSP. The specialized hardware compute engine inside the DNA 100 processor leverages sparsity for both compute and bandwidth reduction. A single DNA 100 processor can easily scale from 0.5 to 12 effective TMACs, and multiple processors can be stacked to achieve 100s of TMACs for use in the most computeintensive on-device neural network applications. The Tensilica DSP will accommodate any new neural network layer that is not currently supported by the hardware engines inside the DNA 100 processor, while also offering the extensibility and programmability of a Tensilica Xtensa® core using Tensilica Instruction Extension (TIE) instructions. The DNA 100 processor can run all neural network layers, including but not limited to convolution, fully connected, LSTM, LRN, and pooling..

DNA 100 Processor Features and Benefits

SELF-DRIVING

- Industry-leading performance and power efficiency
 - 2550fps for ResNet50 in a 4K MAC base array configuration
 Up to 3.4TMACs/W in 16nm
 - Up to 3.41 MACs/W in 16nm
 - Wide and configurable AXI bus width to sustain various ranges of neural network tasks
 - Generic AI processor that runs all neural network layers
 Integrated DMA
- Architected to serve wide range of compute requirements
 - Single DNA 100 processor scales from 0.5 to 12 effective 8b TMACs
 - Multiple DNA 100 processors can be stacked to achieve 100s of TMACs



Figure 1: DNA 100 Processor block diagram

- Programmable and extensible
 - Flexible and future proof
 - Support for new layers as neural networks evolve
 - Complete AI software platform
 - Tensilica Neural Network Compiler provides automatic optimized offline code generation for range of AI networks
 - Android Neural Network API (ANN) provides dynamic optimized code generation

Vision C5 DSP Features and Benefits

- Highest computational capacity in the industry
 - 1024 8b MACs and 512 16b MACs
 - 128-way 8b SIMD, 4-slot VLIW processor
 - Industry-leading 1024-bit wide memory interface with dual load/store
 - Integrated DMA
- Highly programmable and extensible
 - Flexible and future proof
 - Support for new layers as neural networks evolve
 - On-the-fly decompression
 - Flexible instruction set for quantization
- Dedicated neural network DSP that runs all neural network layers
 - Eliminates moving data between Vision C5 DSP and main processor
 - Includes normalization and max pooling layers
 - Simple programming model
 - Simplified hardware architecture
- Architected for multi-processor clusters
 - Computation capacity and scalability for all neural network applications and segments in surveillance and automotive



Figure 2: Vision C5 DSP block diagram

Toolchain

Our processors are delivered with a complete set of software tools:

- A high-performance C/C++ compiler with automatic bundling and vectorization support for the VLIW and SIMD capabilities
- Linker, assembler, debugger, profiler, and graphical visualization tools are included
- A comprehensive instruction set simulator (ISS) allows you to quickly simulate and evaluate performance
- When working with large systems or lengthy test vectors, the fast, functional TurboXim simulator achieves speeds that are 40X to 80X faster than the ISS for efficient software development and functional verification
- Tensilica Xtensa Modeling Protocol (XTMP) for system modeling in C and Xtensa SystemC (XTSC) for system modeling in SystemC[®] provide for full-chip simulations. The pin-level XTSC model offers co-simulation of the SystemC model at the pin level for fast, cycle-accurate system simulations
- Tensilica Neural Network Compiler maps any neural network trained with a framework such as Caffe, TensorFlow, and TensorFlowLite into executable and highly optimized fixedpoint code for target DSPs, leveraging a comprehensive set of hand-optimized neural network library functions
- All major back-end EDA flows are supported

Cadence Services and Support

- Cadence Tensilica application engineers can answer your technical questions, and provide technical assistance and custom training.
- Cadence-certified instructors teach a series of courses on Tensilica IP and bring their real-world experience into the classroom.
- Internet Learning Series (iLS) online courses allow you the flexibility of training at your own computer via the Internet.
- The Cadence Tensilica IP support site gives you 24x7 online access to a knowledgebase of the latest solutions, technical documentation, software downloads, and more at <u>ip.cadence.com/support</u>.

cādence[°]

Cadence software, hardware and semiconductor IP enable electronic systems and semiconductor companies to create the innovative end products that are transforming the way people live, work, and play. The company's System Design Enablement strategy helps customers develop differentiated products—from chips to boards to systems.www.cadence.com

© 2018 Cadence Design Systems, Inc. All rights reserved worldwide. Cadence, the Cadence logo, and the other Cadence marks found at www.cadence.com/go/trademarks are trademarks or registered trademarks of Cadence Design Systems, Inc. SystemC is a trademark of Accellera Systems Initiative Inc. All other trademarks are the property of their respective owners. 10963 09/18 SA/RA/PDF