Modalix MLSoC

Product Brief

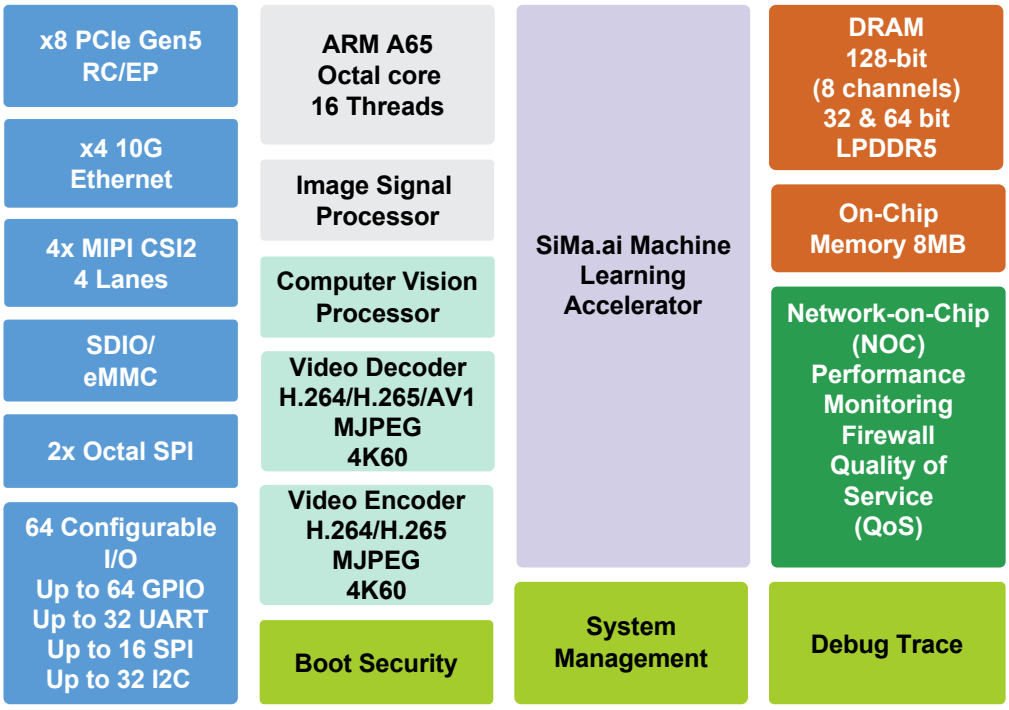
Overview

The **Modalix MLSoC™** is SiMa.ai's second-generation multi-modal, machine-learning system-on-a-chip designed for scalable Physical AI. Modalix delivers high-performance local AI processing without cloud connectivity.

Modalix is **purpose-built** for executing generative **artificial intelligence (GenAI)**, **computer vision**, and **machine learning (ML)** inference. Its pipelined architecture combines a highly programmable ML accelerator assisted by an on-chip Arm™ application processor (APU) and a Digital Signal Processor (DSP)-based Computer Vision Unit (CVU). The Modalix chip delivers **50 TOPS** (tera operations per second) in an incredibly compact and low-power 25mmx25mm package, delivering GenAI capabilities at under 10 watts.

The Modalix system architecture seamlessly integrates multiple compute engines and peripherals, delivering **best-in-class performance per watt** for advanced workloads, including multimodal transformers, Large Language Models **(LLMs)**, Large Multimodal Models **(LMMs)**, and Generative AI **(GenAI)**. It also supports legacy convolutional neural networks **(CNNs)** and traditional computer vision **(CV)** algorithms, ensuring compatibility across modern and established ML pipelines.

Modalix Architecture



All architectural building blocks are seamlessly integrated via an internal, secure Network-on-Chip (NoC)

Key Features

- 8 x Arm Cortex-A65 @ 1.4GHz
- 4 x LPDDR5 memory interfaces
- 128-bit LPDDR5/4x/4, 4x32-bit or 2x64-bit up to 6,400 Mbps (LPDDR5)
- x8 PCIe Gen5 root-complex and endpoint, with bifurcation
- 4 x 10G Ethernet interfaces
- 4 x MIPI CSI-2 interfaces (4 x 4 lanes)
- Video encode/decode engines for H.264/H.265 up to 4K60
- Arm Mali-C71AE operating at 1.2GHz and support for 24-bit Wide Dynamic Range
- Computer Vision Unit (CVU) – Quad-core Synopsys ARC EV74 video processor, operating at 1GHz quad-core, delivering up to 720 16-bit GOPS
- Network-on-Chip (NOC)
- Boot and security unit (BSU)
- HW secure boot engine ensuring "chain of trust", decrypting and authenticating of the boot image with OTP secure keys

Target Applications

Developing AI solutions with Modalix is streamlined through the SiMa.ai ONE Platform, which includes the Palette software suite. Palette enables unified programming of the APU, CVU, and ML accelerators utilizing end-to-end pipelines, tailored for applications such as:

- Smart Vision
- Conversational Interface
- Drones
- Robotics
- Industry 4.0
- Automotive
- Smart Retail / Healthcare
- Aerospace & Defence

Accelerated Development

SiMa.ai's ONE Platform enables fast AI development and deployment of on-device AI. With its user-friendly Palette™ Software Suite, including the Palette Neat™ Agentic Development Environment, users enjoy efficient, streamlined workflows. Design from scratch or bring your ONNX model, tune it with SiMa's Model Compiler, and develop applications using popular frameworks, like OpenCV quickly and efficiently across various AI applications.

Ordering Information

Product Description	Part Number
Modalix Chip - ComTemp (0 - +70C) 25mm x 25mm, FCBGA (1369 balls) package	SM2-201B-C0AA0-000
Modalix Chip - IndTemp (-40 - +85C) 25mm x 25mm, FCBGA (1369 balls) package	SM2-201B-I0AA0-000

SiMa.ai

About SiMa.ai

SiMa.ai is a leader in Physical AI, delivering a purpose-built, software-centric platform that brings best-in-class performance, power efficiency, and ease of use to Physical AI applications. The company is focused on scaling Physical AI across robotics, automotive, drones, industrial automation, aerospace and defense, smart vision, and healthcare. Learn more at www.sima.ai.