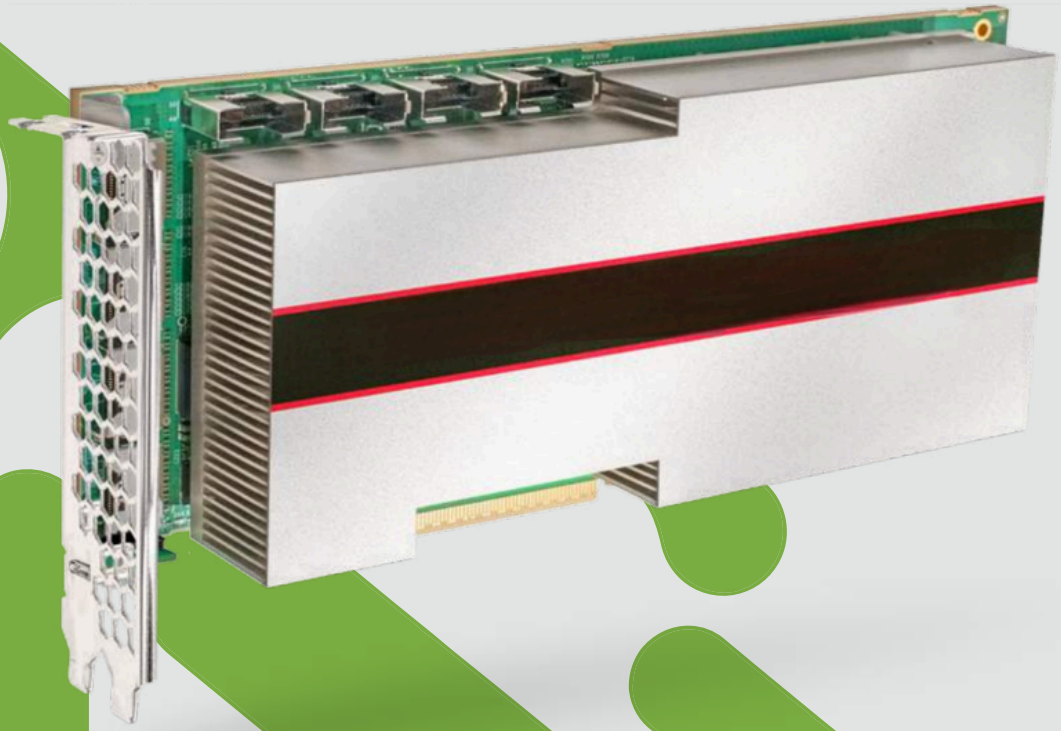
MACx GPU Accelerator

MAC01 | Multimedia AI Accelerator Card

Multi-modal AI Processing

Hardware-level acceleration for CV, NLP and speech models.



Why MACx?

- FP32 Tensor Core: 40 TFLOPS
- FP16 Tensor Core: 120 TFLOPS
- INT8 Tensor Core: 240 TFLOPS
- Fast memory: 64 GB

MACx (MAC01) is a multimedia AI accelerator card with hardware-level acceleration for CV, NLP and speech models.

MAC01 combines AI acceleration with full-stack development support and is presented for deployment across various industries.

Development & Integration

MACA software development platform

Hardware abstraction layer, programming model, acceleration libraries, runtime libraries, and toolchain.

Supported AI frameworks

PyTorch | TensorFlow | MXNet

OpenCL programming | Model compression toolchains

MACx / MAC01 Specifications

Compute

FP32 Tensor Core: 40 TFLOPS

FP16 Tensor Core: 120 TFLOPS

INT8 Tensor Core: 240 TFLOPS

Video Encoding

2 encoders

H.264, H.265, JPEG

Video Decoding

14 units

H.264, H.265, VP9, JPEG

Memory & Interface

Fast memory: 64 GB

PCIe Gen5: 128 GB/s

Form Factor

Full-height, full-length

Dual-slot card

Power

TDP: < 300 W

MACA Software Platform

Frameworks: PyTorch, TensorFlow, MXNet

Programming: OpenCL; libraries and development tools

Runtime: firmware, driver and runtime libraries

Engines: AI inference and video codec acceleration

MACx Development Stack

MACA - Multimedia AI Computing Architecture

MACA is the software development platform for MAC accelerator cards. The stack focuses on collaborative software and hardware optimization.

AI Frameworks

PyTorch | TensorFlow | MXNet

Programming & Tools

OpenCL, acceleration libraries, debugger, profiler and porting tools.

System Stack

Hardware abstraction, firmware, driver, runtime, AI engine and codec engine.

Applications: SAN | NAS | DAS

