



\$USD ▾



NEWS AI

Top 6 RISC-V Chips with Multi-core Design and AI Accelerator for AI and ML

DFRobot Dec 05 2023 7669

The adoption of RISC-V, a free and open-source computer instruction set architecture first introduced in 2010, is rapidly accelerating, driven primarily by the increasing demand for AI and Machine Learning(ML). Research firm Semico projects a staggering 73.6 percent annual growth in the number of chips incorporating RISC-V technology, with a forecast of 25 billion AI chips by 2027 generating a revenue of US \$291 billion. This article will highlight the popular RISC-V Architecture-based chip products for AI and ML currently accessible in the market.

The Advantages of RISC-V Architecture Compared to x86 and ARM for AI and ML

- 
- RISC-V is inevitable
 - RISC-V will have the best processors
 - RISC-V will have the best ecosystem



Home



Category



0

Shopping Cart



Me



\$USD



iteration and higher computational power in AI algorithms. The RISC-V instruction set can be tailored and customized according to specific application requirements, allowing for better adaptation to different AI algorithms, including deep learning and neural networks.

- **Efficiency advantage of RISC-V processors.** Licensees choose RISC-V for its higher efficiency compared to traditional microprocessors, and in comparison to ARM and x86, RISC-V exhibits approximately a 3x advantage in computational performance per power.
- **Single IP more flexible in combination or reconfiguration.** RISC-V's IP can be further 'deconstructed' compared to ARM, to address various chip design scenarios. This also brings higher scalability to the RISC-V architecture, allowing designers to freely disassemble modules like playing with LEGO bricks and combine them to create their ideal chips.
- **Short transition time between ARM and RISC-V.** Switching between ARM and RISC-V is akin to a programmer with knowledge of data structures switching between C language and Python. Only the instructions change, while the overall design philosophy remains unchanged. A designer proficient in ARM architecture may only need about two weeks to transition to RISC-V development.

RISC-V Chip Products for AI and ML

SiFive Intelligence™ X390

The Intelligence X390 processor is designed to meet the increasing demands of artificial intelligence and machine learning applications. It builds upon the foundation of the X280 with key enhancements that



Home



Category



Shopping Cart



Me



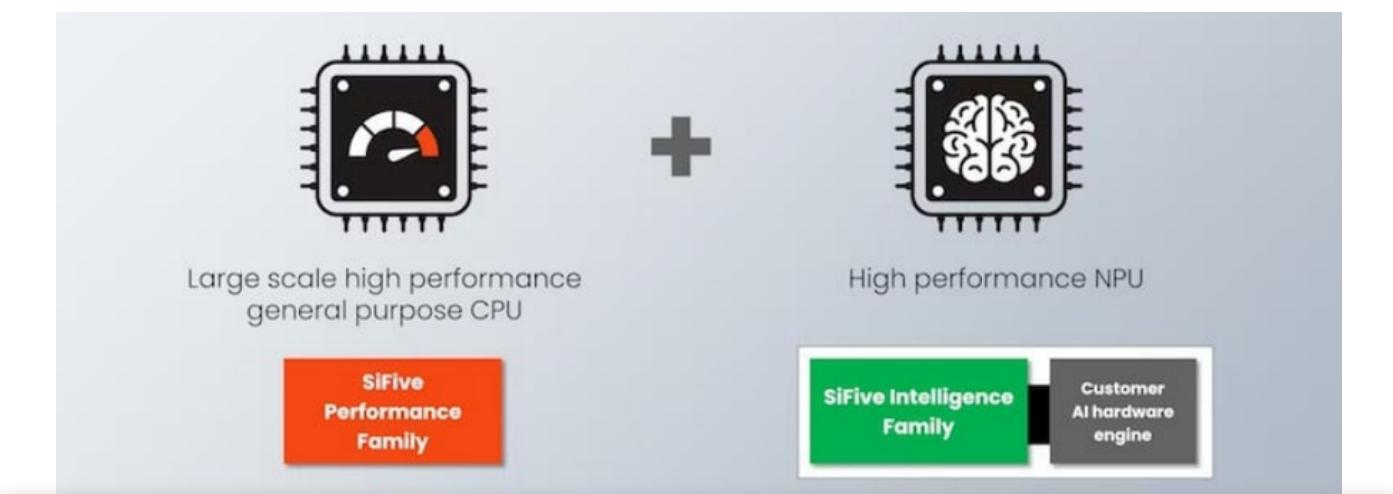
\$USD ▾



Additionally, the processor integrates SiFive's VCIX technology, allowing companies to incorporate custom vector instructions or acceleration hardware for unprecedented performance optimization flexibility. The enhanced vector computation capabilities make the X390 processor particularly suitable for neural network training and inference tasks.

Key Features

- SiFive Intelligence Extensions for ML workloads
- 512-bit vector register-length processor
- Performance benchmarks
 - 5.75 CoreMarks/MHz
 - 3.25 DMIPS/MHz
 - 4.6 SpecINT2k6/GHz
- Built on silicon-proven U7-Series core
 - 64-bit RISC-V ISA
 - 8-stage dual-issue in-order pipeline
- High performance vector memory subsystem
 - Up to 48-bit addressing
- Multi-core, multi-cluster processor configuration, up to 8 cores



Home



Category



0

Shopping Cart



Me



\$USD ▾



with up to 16 cores

Highest performance commercially licensable RISC-V processor

- 12 SpecINT2k6/GHz (P870 Processor)

2x 128b VLEN RVV

Vector crypto and hypervisor extensions

IOMMU and AIA

Non-inclusive L3 cache

Proven RISC-V WorldGuard security

- P800-Series Architectural Features

64-bit RISC-V core with extensive Virtual Memory Support

Four-issue, out-of-order pipeline tuned for scalable performance

Private L2 Caches and Streaming Prefetcher for improved memory performance

SECDED ECC with Error Reporting

These two processors differ in design objectives and application focus. The SiFive Performance P870 is primarily used for high-performance computing and data center applications, while the SiFive Intelligence™ X390 is designed for edge artificial intelligence and machine learning applications.



Home



Category



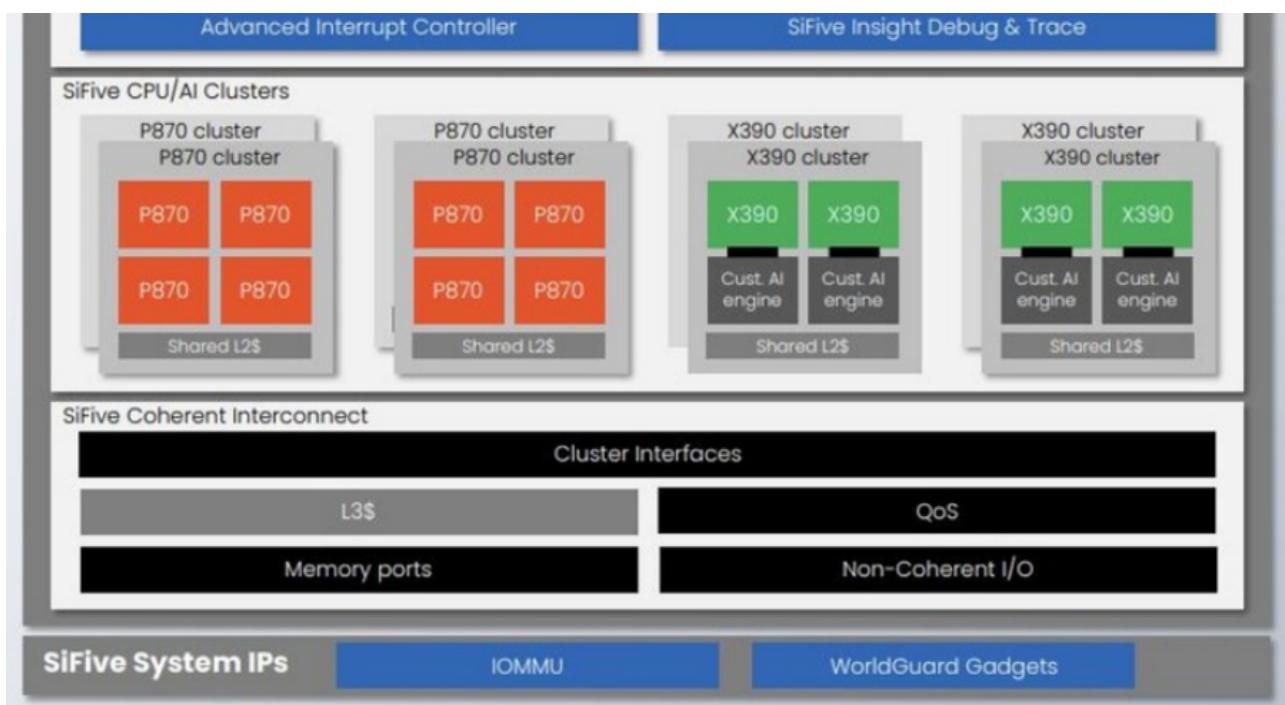
Shopping Cart



Me



\$USD



T-Head XuanTie C910

T-Head XuanTie C910 delivers industry-leading performance in control flow, computing and frequency through architecture and micro-architecture innovations. The C910 processor is based on the RV64GC instruction set and implements the XIE (XuanTieInstruction Extension) technology. C910 adopts a state-of-the-art 12-stage out-of-order multiple-issue superscalar pipeline with high frequency, iPC, and power efficiency. C910 supports hardware cache coherency. Each cluster contains 1-4 cores. The C910 supports the AXI4bus interface and includes a device coherence port. The C910 uses the SV39 virtual address system with XMAE (XuanTie Memory AttributesExtension) technology. In addition, C910 includes the standard CLINT and interrupt controllers and supports RV-compatible debug interface and performance monitors.

Key Features



Home



Category



0

Shopping Cart



Me



\$USD ▾



- Non-blocking inclusive/non-inclusive cache
 - Open-source at <https://github.com/OpenXiangShan/HuanCun>
 - Some design choices inspired by SiFive Block Inclusive Cache

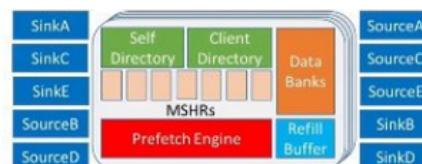


Figure. Overview of HuanCun

- Configurable parameters
 - Size, #ways, #MSHRs
 - Inclusion policy, prefetch policy, replacement policy
- Optimized design and improved performance
 - Better timing and higher frequency
 - Up to 30% IPC increase on sensitive benchmarks



Figure. Performance improvements on some memory-sensitive benchmarks



中国科学院计算技术研究所 (ICT, CAS)

Esperanto ET-SoC-1 chip

The Esperanto ET-SoC-1 chip integrates over 1000 RISC-V processor cores and 24 billion transistors, including 1088 energy-efficient ET-Minion 64-bit RISC-V in-order cores and 4 high-performance ET-Maxion 64-bit RISC-V out-of-order cores. Each core is equipped with a vector/tensor unit, with expected operating frequencies ranging from 500MHz to 2GHz. Additionally, the chip incorporates 1 RISC-V service processor, over 160 million bytes of on-die SRAM for caches and scratchpad memory, and interfaces supporting large external memory, including LPDDR4x DRAM and eMMC flash, PCIe x8 Gen4, and other common I/O interfaces. At peak rates, the ET-SoC-1 is capable of achieving 100 to 200 Tera Operations per Second (TOPS), while typically consuming less than 20 watts of power. What sets Esperanto's solution apart is how each board uses multiple low-power SoC chips instead of a giant SoC. It could be used as a compelling energy-efficient solution for Machine Learning recommendation in large data centers.



Home



Category



0

Shopping Cart



Me

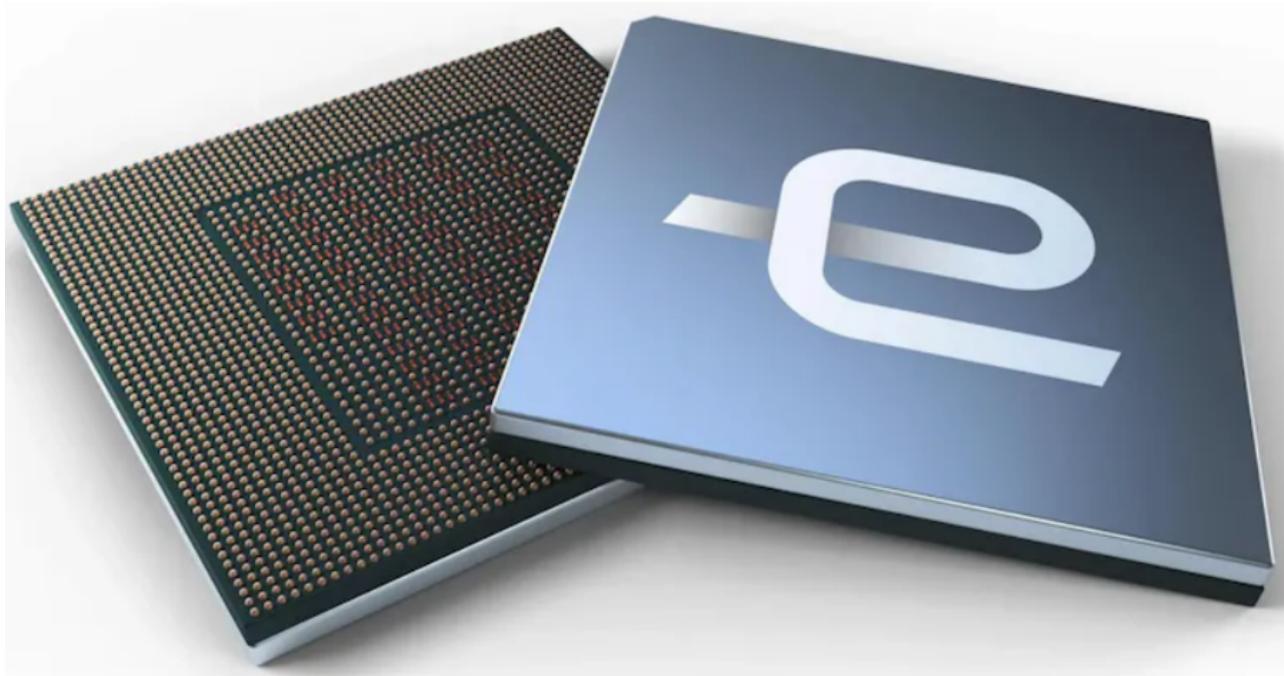


\$USD ▾



cores

- Inclusion of 4 high-performance ET-Maxion 64-bit RISC-V out-of-order cores
- Consists of approximately 24 billion transistors
- Specifically designed for AI and machine learning applications
- Offers exceptional parallel processing capabilities



Meta Training Inference Accelerator (MTIA) Chip

MTIA is designed by Meta to handle their AI workloads more efficiently. The processor cores are based on the RISC-V open instruction set architecture (ISA). The chip is a custom Application-Specific Integrated Circuit (ASIC) built to improve the efficiency of Meta's recommendation systems, e.g. Content understanding, Facebook Feeds, generative AI, and ads ranking all rely on deep learning recommendation models (DLRMs), and these models demand high memory and computational resources.



Home



Category



0

Shopping Cart



Me

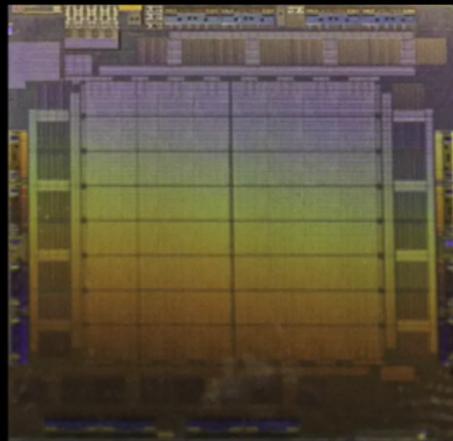


\$USD ▾



Specification

TECHNOLOGY	TSMC 7nm
FREQUENCY	800 MHz
DIMENSIONS	19.34 x 19.1 mm (373 mm ²)
TDP	25 Watts
GEMM TOPS (MAC)	102.4 (INT8) 51.2 (FP16)
MEMORY BANDWIDTH	800 GB/s (on-chip SRAM) 176 GB/s (off-chip DRAM)



The first-generation MTIA ASIC was designed in 2020 specifically for Meta's internal workloads. The chip was fabricated using the TSMC 7nm process and runs at 800 MHz, providing 102.4 TOPS at INT8 precision and 51.2 TFLOPS at 16-bit floating-point precision. It also features a thermal design power (TDP) of 25 W. The MTIA chip is part of a full-stack solution that includes silicon, PyTorch, and recommendation models; all co-designed to offer a wholly optimized ranking system for Meta's customers. The release of their first AI chip, MTIA, is a significant development. It further fuels the AI hardware race and contributes to the evolution of hardware designed specifically for AI applications.

Conclusion

The chips listed in the article feature a multi-core design, with each core offering high performance and energy efficiency. They support multi-threaded operations, enabling simultaneous execution of multiple tasks. Additionally, they all support Single Instruction Multiple Data (SIMD) instruction sets, which can accelerate parallel data



Home



Category



0

Shopping Cart



Me



\$USD ▾



innovations and applications.

REVIEW

0 Comments

1 Login ▾



Start the discussion...

LOG IN WITH

OR SIGN UP WITH DISQUS



Name



Share

Best Newest Oldest

Be the first to comment.

Subscribe

Privacy

Do Not Sell My Data

Recent Blogs



Home



Category



0

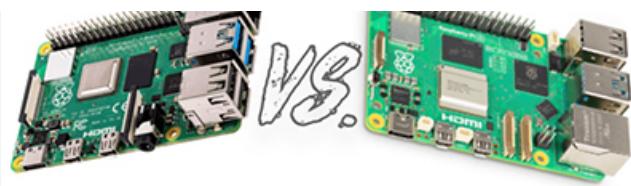
Shopping Cart



Me



\$USD ▾



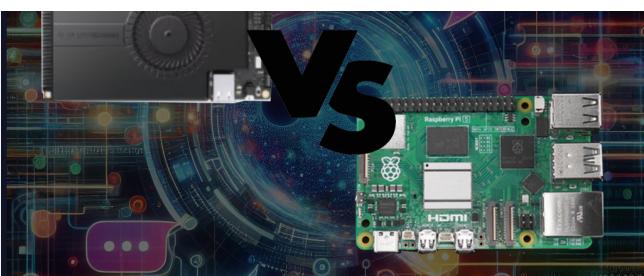
RASPBERRY PI 5

Essential Compatibility Insights for Raspberry ...

Discover how to fully harness the Raspberry Pi 5's...

SELECTION GUIDE

Apr 19 2024



How to Select SBC (Lattepanda/Raspberr...

Discover why Lattepanda Sigma is well-suited for loca...

SELECTION GUIDE AI

Mar 29 2024



Selection Guide of Linux Systems Compatible wi...

This article will provide a detailed introduction to...

SELECTION GUIDE

Mar 26 2024



Home



Category



0

Shopping Cart



Me