AI Inference SoC Platform

SEMIFIVE SoC Platforms can quickly turn your critical IPs or winning specifications into fully functioning SoC at a fraction of risk, time, and effort. The AI Inference SoC Platform offers the best solution to build custom AI inference accelerators, enabled with silicon-proven design components on Samsung 14nm process and extensive hardware/software environments to instantly get your chip ready for systems.

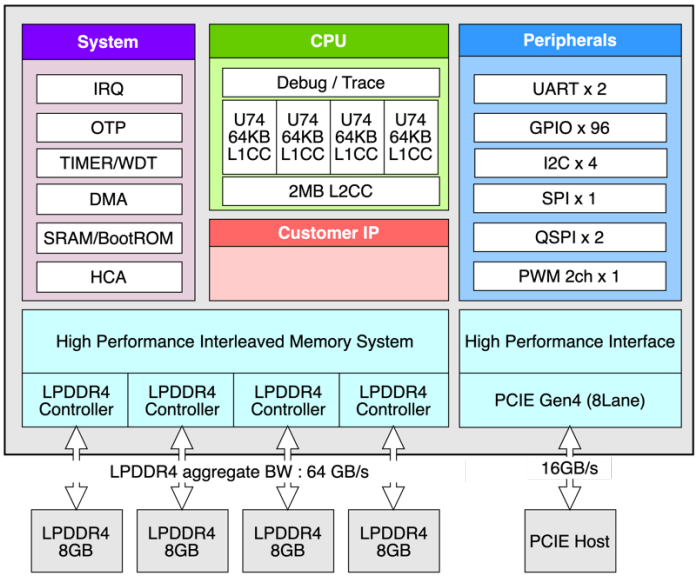
Highlights

- Silicon verified with multiple customers adoption
- Optimized for high performance, power efficient AI and inference applications
- Built using Samsung Foundry's mass production proven 14nm FinFET process technology
- Complete solution with:
    - Package design and implementation
    - Evaluation board
    - Software and drivers

Target Applications

- Data center accelerator
- AI Vision processor
- Big data analytics
- Image/Video recognition
- ADAS with real time processing

Block diagram



SEMIFIVE, Korea
info@semifive.com

SEMIFIVE, US
us@semifive.com

Key features

Process node	• Samsung Foundry 14nm
Die size	• 14 x 13 mm2 (excluding customer IPs)
Target operating frequency	• 2GHz
CPU core	• Quad-Core SiFive U74 RISC-V (up to 1GHz) • Include Debug & Trace
Memory interface	• LPDDR4 • 4GB ~ 16GB • Up-to 4266MHz
I/O Interfaces	• PCIe Gen4 X8 EP/RC • SNOR Flash memory with XIP support • Off-chip component (SPI/I2C) • SD Card & eMMC memory • Timer includes PWM • UART • GPIO
Low power features	• Clock gating controlled by S/W
System controller	• Interrupt Controller • DMA • Watch-dog Timer • PVT (Power, Voltage & Temperature) monitoring
Software	• Yocto-based Linux build system • Boot process and Linux device driver • Reference Linux BSP for the evaluation Board • DDR tuning S/W • Debug solutions (GDB/TRACE32)

SoC Platform Engagement Models

Max Efficiency Model

By reusing the platform architecture and feature subblock tailored for this domain, customers can focus on their differentiation and maximize efficiency of SoC development. Perfect for SoC prototyping or high-value applications that require super-fast time-to-market speed.

Max Flexibility Model

SEMIFIVE Platform technology offers flexibility to configure the platform architecture to your application's special requirements. SEMIFIVE can also develop new subblock design to add features necessary for the application. It's the best way to explore the trade-offs between time and cost.