Innovation Unleashed: Solutions Enabling Embedded Intelligence

Aniket Saha
Sr. Director, Product Marketing
aniket.saha@sifive.com

SiFive
Agenda

01 Global Trends
02 SiFive Core IP Portfolio
03 Domain Specific Features
04 Commercial Products
05 Future
Global Trends

- **Internet users by 2022**: 60% of the population will be using the Internet, up from 45% in 2017.
- **Internet traffic by 2022**: 85 GB of Internet traffic per month, per user, up from 29 GB in 2017.
- **Devices/connections by 2022**: 3.6 networked devices and connections per person, up from 2.4 in 2017.

Exabytes per Month

Source: Cisco VNI Global IP Traffic Forecast, 2017-2022
SiFive Core IP
Embedding Intelligence Everywhere

**Consumer**
- AR/VR/Gaming devices
- Smart Home
- Imaging/Wearables

**Storage/Networking/5G**
- SSD, SAN, NAS
- Base Stations, Small cells, APs
- Switches, Smart NICs, Offload cards

**ML/Edge**
- Sensor Hubs, Gateways
- Autonomous machines
- IoT devices
U Cores
64-bit Application Processors

S Cores
64-bit Embedded Processors

32-bit Embedded Processors

Embedding Intelligence from the Edge to the Cloud
SiFive Core IP 2 series:

SiFive’s smallest and most efficient RISC-V processor IP

- Higher Performance
- Configurable Core and Memory System
- Ultra low-latency interrupts
- 32-bit Embedded Processors
- Efficient RISC-V MCU
- Low latency interrupts
SiFive Core IP 3 and 5 series:
The world's most deployed RISC-V processor IP

Efficient Performance
Coherent, Heterogenous, Multicore
Hard Real-time capabilities

32-bit Embedded Processors
64-bit Embedded Processors
64-bit Application Processors

Configurable
Efficient
Mature
SiFive Core IP 7 series:

The highest performance commercial RISC-V processor IP

- 32-bit Embedded Processors (E7 Series)
- 64-bit Embedded Processors (S7 Series)
- 64-bit Application Processors (U7 Series)

Common Feature sets
Hard Real-time capabilities
Unprecedented scalability

- ~60% increase in CoreMarks/MHz*
- ~40% increase in DMIPS/MHz*
- 10% increase in Fmax*

*Compared to SiFive 5 series
Scalable throughput provided by 8+1 cores per cluster

Extensible design via custom instructions

Configurable memory architecture for application specific tuning

Tightly integrated memory for low latency access

64-bit addressability for real time latency sensitive applications

Mixed-precision arithmetic for efficient compute of ML workloads

Tightly integrated memory for low latency access

Enhanced determinism for hard real-time constraints

Functional safety provided by in-built fault tolerance mechanisms

A single pre-integrated and verified deliverable

Cache lock capability for mission-critical computing

In-cluster coherent combination of real-time and application processors
# Product Map

| 7 Series | 32-bit embedded cores | E Cores | 32-bit embedded cores  
MCU, edge computing, AI, IoT |
| --- | --- | --- | --- |
|Highest performance: 8-stage, dual-issue superscalar pipeline | E7 Series | 32-bit embedded cores | E76-MC | Quad-core 32-bit embedded processor  
Compare to Cortex-R8 |
| | | E76 | High performance 32-bit embedded core  
Compare to Cortex-M7 |
| 3/5 Series | Efficient performance: 5–6-stage, single-issue pipeline | E3 Series | 32-bit embedded cores  
E31 features + single-precision floating point  
Compare to Cortex-R5 |
| | | E31 | Balanced performance and efficiency  
Compare to Cortex-R5 |
| 2 Series | Power & area optimized: 2–3-stage, single-issue pipeline | E2 Series | 32-bit embedded cores  
E24 | Compare to Cortex-M4F  
E21 + single-precision floating point  
Compare to Cortex-M4 |
| | | E21 | E20 + User Mode, Atomics, Multiply, TIM  
Compare to Cortex-M0+ |
| | | E20 | Our smallest, most efficient core  
Compare to Cortex-M0+ |
| S7 Series | 64-bit embedded cores | S Cores | 64-bit embedded cores  
Storage, AR/VR, machine learning |
| | S76-MC | Quad-core 64-bit embedded processor  
Compare to Cortex-R8 |
| | S76 | High-performance 64-bit embedded core  
Compare to Cortex-R8 |
| U7 Series | 64-bit application cores | U Cores | 64-bit application cores  
Linux, datacenter, network baseband |
| | U74-MC | Multicore: four U74 cores and one S76 core  
Compare to Cortex-A55 |
| | U74 | High performance Linux-capable processor  
Compare to Cortex-A55 |
| U5 Series | | U Cores | 64-bit application cores  
Linux, datacenter, network baseband |
| | U54-MC | Multicore application processor with four U54 cores and one S76 core  
Compare to Cortex-A53 |
| | U54 | Linux-capable application processor  
Compare to Cortex-A53 |
Coherent in-cluster combination of application processors and real-time processors

Configurable memory maps and coherent accelerator ports for tightly coupling storage specific accelerators

Deterministic mode for FAST DATA applications with hard real-time constraints

Tightly integrated memories and Cache lock capability for critical real time workloads

Optional FPU for applications which don’t need floating point capability

Storage, ML, Cryptography specific custom instructions

64-bit real-time addressability for BIG DATA applications
## 5G/Networking

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complex arithmetic</td>
<td>Capability for accelerating baseband functions</td>
</tr>
<tr>
<td>High bandwidth accelerator ports</td>
<td>For enabling intelligent offload processing</td>
</tr>
<tr>
<td>Configurable memory maps</td>
<td>For optimizing QoS</td>
</tr>
<tr>
<td>In-cluster coherence of application and real-time processor</td>
<td>Enables 5G latency (&lt;1ms) requirements</td>
</tr>
<tr>
<td>Hard real-time capabilities</td>
<td>For scheduling baseband protocol layers</td>
</tr>
<tr>
<td>High throughput</td>
<td>Processing for next gen 5G stacks</td>
</tr>
<tr>
<td>Tightly Integrated Memories and Cache lock capability</td>
<td>For critical real time workloads</td>
</tr>
</tbody>
</table>
AR/VR/Sensor Fusion

- **Low Latency peripheral access** and coherent accelerator port
- **Coherent in-cluster combination** of application processors with real time processors
- **Simple caching hierarchy** for ease of application optimization
- **Combine** with SiFive 2, 3 or 5 series for designs with tight power constraints
- **Workload specific customizations** (AR/VR/MR/CV)
- **Mixed precision arithmetic** for accelerating machine learning compute
Core IP 7 Series
Standard Cores

01 E76, E76-MC
02 S76, S76-MC
03 U74, U74-MC

Standard Cores represent pre-configured implementations of a Core Series which are available for free RTL and FPGA evaluations.
Enterprise SSD

- FADU Annapurna SSD Controller  
  - World’s first RISC-V SSD controller
- FADU Bravo Series Enterprise SSD
- 3.5GB throughput and 800K IOPS at less than 1.8W
- Powered by SiFive E51

“SiFive’s RISC-V Core IP was 1/3 the power and 1/3 the area of competing solutions, and gave FADU the flexibility we needed in optimizing our architecture to achieve these groundbreaking products.” J. Lee, FADU CEO
Intelligent Edge

• Microsemi’s PolarFire SoC

• World’s first RISC-V SoC FPGA architecture bringing Real-time to Linux

• Targeted for real-time Linux applications at the Edge

• Defense-grade security features
  • Secure boot
  • DPA safe crypto core
  • SECDED on all memories
  • Physical memory protection/PMP

• Powered by SiFive U54-MC and SiFive E51
Wearable AI

• **Huangshan No. 1** (MHS001) from Huami using Upbeat Tech

• **Integrated biometric signal processor** with 4 dedicated AI engines and built-in CNN based inference engine

• 38 percent more efficient than the Arm Cortex-M4

• Powered by **SiFive E31**

“The world’s first artificial intelligence powered wearable chipset”
SiFive Core IP: Embedded Intelligence Everywhere

- Efficient Performance
- Scalability
- Compelling Feature Set

Embedding intelligence for a world of a Trillion Connected Devices