Pushing Data from Edge to Cloud

Innovations in CPU Architecture

Caffrey Chen
Building the Chip Infrastructure of the AIoT Era

Infrastructure Provider of the AIoT Era

MCU  Security  Intelligent Computing  Industry Control  Memory Control

AliOS

Domain Specific SoC Platforms (IPs from Partners)

RISC-V Compatible Processor  Domain Specific Architecture
The Continuously Evolving Processor Architecture

First RISC-V Processor with HW TEE

Ultra High Performance Processor with AI Acceleration Engine

Coming Soon
Ultra Low Power Architecture – Xuantie902

- RISC-V RV32EMC
- 2 stage In-order Pipeline
- Machine and User Mode Support
- Configurable Hardware Multiplier
- Configurable Tightly Coupled IP
- Configurable Cache
- 10K EQG for Minimum Configuration
- Hardware Enhanced Secure Engine
• RISC-V RV64GCV
• Cluster Based Multi-core Architecture
• 64-bit, 12-stage Out-of-Order
• 3-decode, 8-issue
• Dual Issue Out-of-Order Memory Access
• High Performance Hybrid Branch Processing
• Multi-mode Dynamic Data Prefetch
• Vector Engine for AI Acceleration
Remarkable Performance

Data source:
http://www2.eecs.berkeley.edu/Pubs/TechRpts/2018/EECS-2018-151.pdf
https://www.sifive.com/cores/u74-mc
Compatible with RISC-V Specification

<table>
<thead>
<tr>
<th>Feature</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISA</td>
<td>RV64GCV</td>
</tr>
<tr>
<td>Vector</td>
<td>RISC-V Vector Extension</td>
</tr>
<tr>
<td></td>
<td>FP16/32/64, INT8/16/32/64</td>
</tr>
<tr>
<td>Privilege Mode</td>
<td>Machine + Supervisor + User</td>
</tr>
<tr>
<td>Memory Management</td>
<td>Sv39 MMU + 8/16 PMP</td>
</tr>
<tr>
<td>Interrupt Controller</td>
<td>Clint + PLIC</td>
</tr>
</tbody>
</table>
RISC-V Turbo

- Computing
- Bit operation
- Memory access
- Multi-core synchronization

- Memory management
- Cache, TLB

Extended Enhancement - RISC-V Turbo

![Bar chart comparing RISC-V Turbo and RV Std. across various benchmarks.](chart.png)
• **Front-End**
  - Fetch 8 Instructions/cycle
  - Decode 3 Instructions/cycle
  - Issue 8 Instructions/cycle

• **Back-End**
  - Out-of-Order Memory Access
  - Dedicated Branch Processing
  - Out-of-Order Vector Computing
Instruction Fetch Unit with Hybrid Prediction

• Hybrid Multi-mode Branch Prediction
  • Branch Direction Prediction
  • Branch Target Prediction
  • Return Address Prediction
  • Indirect Branch Prediction

• High-bandwidth Parallel Fetch
  • 128-bit Fetch
  • Up to 8 Instructions Packaged in Parallel
  • Instruction Cache Way Prediction
  • Loop Acceleration
• Out-of-Order Dual Issue
  • Load/store Address Pipeline
  • Independent Store Data Pipeline

• Load/store Fast Complete
  • 3 cycle Load-to-Use
  • 1 cycle Store Execution

• Powerful Prefetching Capabilities
  • Multi-mode and Multi-channel
  • Both Virtual and Physical Address Prefetching
  • Configurable Prefetch Capacity
Efficient Multi-core Interconnection

- Decoupled Processor Interface Units (PIUx)
- MOESI Coherence Protocol
- Directory-based Architecture
- Snoop Filter Supported
- Configurable L2 Cache, up to 8MB
- ECC Supported
• Compatible with RISC-V Vector Extension
• Supports FP16/32/64, INT8/16/32/64
• 256-Bit Operation Width
• Dual-issue Out-of-Order Vector Execution Pipeline
• More than 300GFLOPS of Half-precision computing power
Back-End Physical Implementation

2.5 GHz
In Conclusion

- Ultra High Performance Superscale Processor
- RISC-V Compatible plus RISC-V Turbo Technology
- Dual issue Out-of-Order Memory Subsystem
- AI Vector Acceleration Engine
Wujian100_Open – an Open Source MCU SoC Platform
Contributing to Open-Source for Building the Chip Ecosystem in the New AIoT Era