AI: scale from Edge to Server with RISC-V and Linux

3rd RISC-V Israel Virtual Meetup

JUNE 8TH 2020, 6PM ISRAEL TIME

Florian Wohlrab
RISC-V Ambassador & Sales Manager
8. June 2020
• Who is Andes Technology?
• Where is RISC-V used, sample applications
• V-Series: Vector for everyone
• Closer look on our RISC-V Cores
At A Glance

Who We Are

- Pure-play CPU IP Company
- Major Open-Source Contributor/Maintainer

RISC-V Founding Premier Member
Taiwan Stock Exchange Listed

Running Task Groups

Quick Facts

- 15 years old company
- 200+ Licensees Worldwide
- 80% R&D employees
- 5B+ accumulated Andes-embedded SoC shipped
- 17K+ AndeSight IDE installation
Active Roles in RISC-V International

- **RISC-V Board Member**
- Vice Chair, Technical Steering Committee
- **RISC-V Ambassador**
RISC-V Adoption: Applications

Edge to Cloud:
- ADAS
- AIoT
- Blockchain
- FPGA
- MCU
- Multimedia
- Security
- Wireless (BT/WiFi)

1 to 1000+ core
40nm to 7nm
Many in AI

General-Purpose
- ARM Dominates
  - GP MCU
  - Big
  - Little

Cost-sensitive
- House Keeping
  - Offload FSM
- DSP
  - Vector datapath

Performance-driven
- RISC-V Stronghold

-confidential-
Example of Edge Computing – Vision Processing

Pre-ISP Processing & Post Processing

**Image Processing**

- Vector Processor NX27V
  - VLEN=256b

**NN Processing**

- Vector Processor NX27V
  - VLEN=512b
- Accelerator

Legend:

- Blue: Andes
- Gray: Third Party

AxI BUS / NOC

- ISP
- Control Core A25MP
- DMA
- SRAM
- DDR Interface
- Other I/O
- ISP Interface
Andes RISC-V on Audio product

- D25F on LE Audio(BLE 5.2) for True Wireless Earbuds and Hearing Aids
  - Customer Tape out already!

This product will be the one of the first SoCs supporting both of LE audio and Classic Audio

Bluetooth Baseband and Audio Subsystem

- LC3 codec - high quality, low power
- LE Audio/BLE 5.2 – Multi-Stream, Broadcast Audio

Application CPU

- **Power Efficient** – Excellent PPA
- **Mature Ecosystem** – Support various RTOSs
- **RISC-V DSP Extension** – outstanding performance
- DSP library

DSP/AI Accelerator

- DSP Algorithm for Noise and Echo Cancellation
- Machine Learning for Key Word Spotting
Example of Cloud Computing - Datacenter

Data Processing

- Main CPU Multicore AX27MP/45MP
- Vector Processor NX27V
- Coordinator Core A27
- Controller NX27V
- Accelerator
- Controller NX27V
- Accelerator

Legend:
- Blue: Andes
- Gray: Third Party

Legend:
- Blue: Andes
- Gray: Third Party

SRAM

PClx Interface

DDR Interface

PCIx Interface

DDR Interface

Legend:
- Blue: Andes
- Gray: Third Party
RISC-V

Added Value and contributed extensions
Andes Added Value in RISC-V

Andes extensions to RISC-V

- Baseline ISA extension to speed up memory access and branches
- CoDense to reduce code size (12% better measured by GCC)
- PowerBrake to save power by stalling pipeline
- StackSafe HW stack protection
- vPLIC vectored dispatch and preemption (reduce 57% of latency)

- Powerful features to differentiate your products
- Create competitive edge for your systems
RISC-V DSP Extension (Packed SIMD/DSP)

- Andes contributed market-proven DSP(SIMD) as P-Extension
- Designed to accelerate slow video, audio/voice and low data rate DSP workloads

Real world speedup using P-Extension

- Increase power efficiency to your DSP applications
Andes Custom Extension

- ACE unlocks RISC-V’s Potential of DSA
  - Define ACE instructions to handle time critical codes
  - Another approach to co-processor or accelerators

- All-in-one COPILOT development environment
  - Automation tool and ease of use
  - Extensions are easy to re-use, can be used as a library
AndesCore™ RISC-V Families

- **Fast and Compact /Slim and Efficient**
  - **N(X)45**: V5, Dual Issue, Superscalar, FPU, PMP
  - **D(X)45**: N(X)45, FPU, DSP
  - **A(X)45**: D45, MMU
  - **A(X)45MP**: L2$, L1/IO coherence
  - **NX27V**: VPU, FPU, PMP
  - **N(X)25F**: V5/32&64b, FPU, PMP
  - **D25F**: N25F, DSP
  - **A(X)27MP**: MemBoost, MMU, DSP
  - **A(X)25MP**: L2$, L1/IO coherence
  - **A(X)25**: N(X)25F, MMU, DSP
  - **A(X)27**: MemBoost, MMU, DSP
  - **A(X)45**: L2$, L1/IO coherence

- **Linux with FPU/DSP Extension**
  - **N22**: V5[e], 32/16 GPR

- **Cache-Coherent Multicores**
  - **D(X)45**: N(X)45, FPU, DSP
  - **A(X)25**: N(X)25F, MMU, DSP
  - **A(X)27MP**: MemBoost, L2$, L1/IO coherence

**Features/Performance**

- **In Production**
  - 8-stage 1.2GHz*
  - 5-stage 1.1GHz*
  - 2-stage 700MHz*

- **In Development**

* X represents 64-bit CPU core
* TSMC 28HPC+ RVT, SS, 0.81V, OC, with I/O constraints.
V-Series

Cray style, scalable vector processor
Why Andes Vector Processor?

Open
Open ISA and ecosystem creates the collaborative RISC-V community

Extensibility
Andes Vector supports bfloat16 and INT4 data types for AI training and inference and Andes Custom Extension

Scalability
Scalable Vector Register to support implementations from MCU to supercomputer

Performance
57x faster performance in parallel computing and real-time processing

Efficient
Vector processing reduces instruction issue bandwidth and starts dependent instruction sooner

Visualization
CPU pipeline visualization tool for performance optimization and stall bubble analysis
NX27V One Vector for All Implementations

Configurable compute data width (VLEN)

- **512b**
  - Cloud Computing – Datacenter, HPC, Server

- **256b**
  - Edge Computing – 5G, Vision Processing

- **128b**
  - Endpoint Computing – Voice
First RISC-V Vector Engine Shipped in Industry!

RVV v0.8 support*
AI optimized with BFloat16 & INT4
1GHz, 0.3mm² in TSMC 7nm FF+
Configurable & scalable
- Vector length 128-bits to 512-bits
- Licensee configurable ALUs

Low power, simple to use
- Multi-level clock-gating
- In-order, 1R/W SRAM, cell based

> 50 VPU in < 10W Open Compute card
Core Overview

Some details

From 2-stage over 5-stage to 8-stage
AndesCore™ 25 Series

- 32-bit and 64-bit cores
- AndeStar V5 architecture:
  - RV-IMACF + Andes V5 Extensions
  - Optional: F/D and S-mode/MMU
- 5-stage pipeline, single-issue
- Configurable multiplier
- Optional branch prediction
- I/D caches and Local Memory
  - Optional parity or ECC protection
  - Hit-under-miss caches
  - HW unaligned load/store accesses
- Bus interface
  - A master port (AHB, AXI, AXIx2)
  - An optional slave port (AHB)

Andes V5
- PLIC, PMU, Debug
- Interrupt Intf., WFI, HW Bkpt
- 25-series uCore, PMP
- ACE, MMU (A*25), FPU/DSP
- I/D Cache, Mul/Div, Br. Pred
- ILM, BIU, DLM

AndesCore™ A25MP/A25MP Multicore
- Platform-Level Interrupt Controller
- Debug Support
- JTAG
- Trace Port x4
- I/D LM Slave Port x4 (AHB-32/64)
- Coherence Slave Port (AXI-64)
- ACU Engine/
  L2 Cache Controller
- Bus Master Interface (AXI-128)

1~4 A25/AX25 CPUs:
- RV-IMACFD ISA + V5 extensions
- P-extension draft
- Supporting SMP Linux

Bus Interfaces
- LM slave port
- Coherence slave port
- AXI bus master interface
  - N:1 synchronous clock ratio

PLIC for interrupt handling

Debug/trace support

Andes Coherence Unit
- MESI cache coherence protocol
- Duplicate L1 dcache tags
- IO coherence for cacheless masters

L2 Controller
- Size: 128KB to 2MB

CPU Subsystem
- AXI/AHB IP
- APB
- Customer’s or Partner’s IP’s

Taking RISC-V® Mainstream
A27 and AX27
- RV*GC-N-P
- 5-stage single-issue
- Programmable PMA table
- MMU for Linux
- Leveraging the mature 25-series; same performance on Local Memory

MemBoost
- Higher memory throughput for Vector
- Performance over 25-series:
  - 200% higher bandwidth
  - 50% lower latency
AndesCore™ 45 Series

- **8-stage in-order dual-issue**
- **AndeStar™ V5 ISA:**
  - RV*GCN (S/D FPU)
  - RV*P-ext (DSP/SIMD)
  - MMU: for Linux Applications
  - ALL have Andes extensions
- **Dual-issue most instruction pairs**
  - Except for 2 MUL/FPU/DSP/LD_ST and some special dependent ALU instruction pairs
  - Late ALUs enable 0-cycle load-use penalty
- **MemBoost** for memory subsystem
- **Low power dynamic branch prediction**
- **Unaligned data accesses**
- **Fast or small multiplier**

**AndesCore™ 45 Series Pipeline**

- F1: Instruction Scheduler
- F2: Instruction Fetch
- ID: Instruction Decode
- II: Instruction Issue
- EX: Execution
- MM: Memory Management
- LX: Load/Store
- WB: Write Back

- **ALU0**
- **ALU2**
- **ALU1**
- **ALU3**

- I$/ILM: DEC: ISS: AG

- **D$/DLM:** Multiplier

- **DSP:** Floating Point Unit

- ($/LM: 2^{nd} cycle for alignment)

- 45-Series Pipeline
Time-to-Market
Get the whole set, IDE, Debug Probes, BSP’s and Core IP
Complete Development Environment

- **AndeSight™ Feature-Rich IDE**
  Free Evaluation on SID and ICE target

- **AndeShape™ Development Boards**
  Full-Featured ADP-XC7K
  Corvette-F1 Amazon FreeRTOS-qualified

- **AndeSoft™ Software Stack**
  Bare metal demo projects
  FreeRTOS ver10
  Linux

- **Debugging Hardware**
  AICE-MINI+, AICE-MICRO
AndesCore™ Ecosystem - Tools

- IAR Embedded Workbench®
  - Support RISC-V
  - Support P-Extension (DSP/Packed-SIMD)
  - excellent optimization technology
  - static code analysis
  - Extensive Debugging via I-jet probe
Summary

- RISC-V is fast growing → The Future of SoC
  - Efficient and extensible architecture for all computing devices
  - From number-crunching vector processors to Linux ready Cores
  - More flexibility in RISC-V

- Andes commits to serve emerging RISC-V demands
  - Most matured offerings for RISC-V processor IP’s
  - Strong development tools and SW from Andes and partners
  - V5 cores already in AIoT, FPGA, MCU, Security, Storage, Wireless

AndesCores For Your Next SoC Projects!
תודה רבה לך
THANK YOU