PULP: 10 Years of Open Source Hardware

Integrated Systems Laboratory (ETH Zürich)

Frank K. Gürkaynak  kgf@iis.ee.ethz.ch

PULP Platform
Open Source Hardware, the way it should be!
Our latest design Occamy: 0.75 TFLOP/s, 400+ cores

- Chiplet based design
- 2x Compute chiplets (Occamy)
  - 216+1 RISC-V cores
  - GF12LPP
  - Running at 1 GHz
- 2x 16GByte HBM memories
- Silicon Interposer (Hedwig)
- Finished in less than 15 months

How did we manage this?

- Taped out on 1\textsuperscript{st} of July 2022

More on Occamy – talk by Gianna – Wed 14:15 – Chiplets for AI
Open source hardware is for us a necessity

- Modern IC design like Occamy is complex and expensive
  - We need partners to help and collaborate
  - We need support (IPs, donations) to realize designs

Open Source to the rescue

- Makes it easy to collaborate with external partners, build teams (both industrial and academic)
  - Less paperwork/NDAs to get started
  - Partners see/are aware of what we provide
- What we do can be re-used (permissive licensing) by our partners
- Results can be more easily verified
We started almost exactly 10 years ago (April 2013)

- Investigating new computing architectures
  - Efficient over a wide range from IoT applications to HPC systems

- Key points
  - Parallel processing
  - Near threshold computing
  - Efficient switching between operating modes
  - Making best use of technology
  - Heterogeneous acceleration

- Parallel Ultra Low-Power (PULP) platform was born
Today the PULP team has grown to more than 70 people

- Headed by Luca Benini
- Teams in both ETH Zürich and University of Bologna
Our research focus: cluster-based many-core accelerators

Innovation factors

Extensions to processor cores
- Explore new extensions
- Efficient implementations

Shared-memory Accelerators
- Domain specific
- Local memory

Multiple computing clusters
- Communication
- Synchronization

High-speed on-chip interconnect (NoC, AXI, other..)

Support Infrastructure

Additional accelerators

Tightly coupled data memory interconnect

Support core

L2 memory

External Memory Controller

L2 Accelerator #1

L2 Accelerator #2

L2 Accelerator #M

DMA

mem bank

mem bank

mem bank

mem bank

mem bank

mem bank

compute core

compute core

compute core

compute core

compute core

ACC #1

ACC #2

Instruction Cache

Peripherals

EXT

EXT

EXT

EXT

Frank K. Gürkaynak - PULP: 10 Years of Open Source Hardware
In the last 20 years IC Design has changed a lot

What used to be a complete chip is now a small part of a SoC!
There is so much that makes up a modern SoC

User-Space Software

Kernel-Space Software

LINUX KERNEL

PULP DRIVER

VIRTUAL MEMORY MANAGEMENT LIBRARY

HW ABSTRACTION LIBRARY

HOST DOMAIN

L2 SPM

Mem

AXI interconnect

LLC + Mem ctrl

I/O

Ext Mem

SENSORS

RISC-V core

MMU

I$ D$

PULP CLUSTER

L1 SPM

Mem

Interconnect

DMA

HW ACC

RV 32

RV 32

RV 32

RV 32

I$
In a typical design, innovation is only in a limited scope

In a typical design, innovation is only in a limited scope.

Open-source silicon-proven SoC template helps concentrate work where it counts.

User-Space Software

Kernel-Space Software

Hardware

HETEROGENEOUS APPLICATION

ACCELERATED KERNEL

LINUX KERNEL

PULP DRIVER

VIRTUAL MEMORY MANAGEMENT LIBRARY

HW ABSTRACTION LIBRARY

ETH Zürich

ALMA MATER STUDIORUM

UNIVERSITAS BOLOGNENSIS

Frank K. Gürkaynak - PULP: 10 Years of Open Source Hardware
What PULP provides is a box of building blocks

**RISC-V Cores**
- RI5CY 32b
- Ibex 32b
- Snitch 32b
- Ariane 64b

**Platforms**
- **Single Core**
  - PULPino
  - PULPissimo
- **Multi-core**
  - OpenPULP
  - Mr. Wolf

**Interconnect**
- Logarithmic interconnect
- APB – Peripheral Bus
- AXI4 – Interconnect

**Peripherals**
- JTAG
- UART
- SPI
- I2S
- DMA
- GPIO
- AXI4

**IOT**
- Neurostream (ML)
- HWCE (convolution)
- HWCrypt (crypto)
- PULPO (1st order opt)

**HPC**
- Occamy
- Hero
Why is RISC-V so special: Freedom to Explore and Fail!

- The ISA provides a contract between HW and SW
  - As long as you stick to the ISA, you can develop HW and SW independently
  - All RISC-V research in HW can continue to rely on growing SW ecosystem for RISC-V
- RISC-V comes with plenty of options for extensions
  - There are reserved encoding spaces for instruction set extensions
- Being able to change everything gives great flexibility
  - Do you want 33 registers, or a 48 bit accumulator.. No problem
  - You need to bring the SW support for your additions.
What if we had a tiny 32b core

Introducing SNITCH

- Start with a simple RISC-V core
- Focus on key features:
  - Lightweight microarchitecture
  - Extensibility: Performance through ISA extensions
  - Latency tolerant
  - Competitive frequency
- Around 15-25 kGE
What if we had a tiny 32b core and add a big 64b FPU

Introducing SNITCH

- Start with a simple RISC-V core
- Focus on key features:
  - Lightweight microarchitecture
  - Extensibility: Performance through ISA extensions
  - Latency tolerant
  - Competitive frequency
- Around 15-25 kGE
- Capable 64b FPU with many extensions

Frank K. Gürkaynak - PULP: 10 Years of Open Source Hardware
What if we add a Floating-point Repetition Buffer? (FREP)

Remove control flow overhead

• Programmable micro-loop buffer
• Sequencer steps through the buffer, independently of the FPU
• Integer core free to operate in parallel: Pseudo-dual issue
• High area- and energy-efficiency

```
mv r0, zero
loop:
  addi r0, 1
  fmad r2, ssr0, ssr1
  bne r0, r1, loop
  frep r1, 1
  loopt: fmad r2, ssr0, ssr1
  ...
```

Allows custom instruction set extensions
What if we could stream data to from FPU directly? (SSR)

• Intuition:
  High FPU utilization ≈ high energy-efficiency
  • Idea: Turn register R/W into memory loads/stores.
  • Extension around the core’s register file
  • Address generation hardware

• Increase FPU/ALU utilization by ~3x up to 100%

• SSRs ≠ memory operands
  • Perfect prefetching, latency-tolerant
We have a processor that maximizes FPU efficiency

In an 8-core cluster

- **Integer core** uses 2% of power
- **SSR/FREP** hardware uses 5% of power
- **L1 Memory** uses 27% of power (47.19mW)
- **FPU** uses 50% of power (87.44mW)
- **FPU**
- **ICACHE** uses 3% of power (4.82mW)
- **Miscellaneous** uses 14% of power (25.26mW)

Inevitable to have local memory (e.g., GPU/GPU L1 cache, vector register file)

The flexibility of open ISA made it easy for us to explore such an approach.
PULP uses a permissive open source license

• All our development is on GitHub
  • HDL source code, testbenches, software development kit, virtual platform

https://github.com/pulp-platform

• Allows anyone to use, change, and make products without restrictions.
Open Source Hardware licensing still a critical issue

• Two main flavors, divided opinion
  • Permissive (Apache, MIT, BSD..): Favored by the industry, minimum obligations
  • Reciprocal (GPL, LGPL,..): Feared by industry

• In theory, it should be possible to have reciprocal licensing for open hardware
  • For example text of LGPL problematic for IC Design use.
  • Cern OHL (https://cern-ohl.web.cern.ch/), comes in many flavors (reciprocal, permissive)
  • Still more work needed, not many people understand issues of IC Design
  • Lawyers (in companies) prefer well-known licenses (less work for them).

• PULP uses Solderpad (http://solderpad.org/licenses/)
  • Permissive license based on Apache
  • Clarifications for hardware use added by Andrew Katz
  • Had no issues (so far) neither with academic nor industrial collaborations
At the moment most of our Open Hardware is still *only* RTL
# State of Open Source for Hardware: Rapid Developments

<table>
<thead>
<tr>
<th>TYPE</th>
<th>EXAMPLES</th>
<th>STATUS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open Specifications</td>
<td>RISC-V</td>
<td>Established</td>
</tr>
<tr>
<td>Architectures</td>
<td>PULP</td>
<td>Quite mature</td>
</tr>
<tr>
<td>Implementations in RTL</td>
<td>Snitch, Hero</td>
<td>Many</td>
</tr>
</tbody>
</table>
## State of Open Source for Hardware: Rapid Developments

<table>
<thead>
<tr>
<th>TYPE</th>
<th>EXAMPLES</th>
<th>STATUS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open Specifications</td>
<td>RISC-V</td>
<td>Established</td>
</tr>
<tr>
<td>Architectures</td>
<td>PULP</td>
<td>Quite mature</td>
</tr>
<tr>
<td>Implementations in RTL</td>
<td>Snitch, Hero</td>
<td>Many</td>
</tr>
<tr>
<td>Open source Hard IP</td>
<td>FLL, DDR PHY..</td>
<td>Very Limited</td>
</tr>
</tbody>
</table>
# State of Open Source for Hardware: Rapid Developments

<table>
<thead>
<tr>
<th>TYPE</th>
<th>EXAMPLES</th>
<th>STATUS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open Specifications</td>
<td>RISC-V</td>
<td>Established</td>
</tr>
<tr>
<td>Architectures</td>
<td>PULP</td>
<td>Quite mature</td>
</tr>
<tr>
<td>Implementations in RTL</td>
<td>Snitch, Hero</td>
<td>Many</td>
</tr>
<tr>
<td>Open source Hard IP</td>
<td>FLL, DDR PHY..</td>
<td>Very Limited</td>
</tr>
<tr>
<td>Process Design Kits</td>
<td>Skywater 130nm</td>
<td>Just Started</td>
</tr>
</tbody>
</table>
# State of Open Source for Hardware: Rapid Developments

<table>
<thead>
<tr>
<th>TYPE</th>
<th>EXAMPLES</th>
<th>STATUS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open Specifications</td>
<td>RISC-V</td>
<td>Established</td>
</tr>
<tr>
<td>Architectures</td>
<td>PULP</td>
<td>Quite mature</td>
</tr>
<tr>
<td>Implementations in RTL</td>
<td>Snitch, Hero</td>
<td>Many</td>
</tr>
<tr>
<td>Open source Hard IP</td>
<td>FLL, DDR PHY..</td>
<td>Very Limited</td>
</tr>
<tr>
<td>Process Design Kits</td>
<td>Skywater 130nm</td>
<td>On its way</td>
</tr>
<tr>
<td>Open Source Tools</td>
<td>Open Lane</td>
<td>Quite usable</td>
</tr>
</tbody>
</table>
What is PULP doing to maintain our cores?

• **We (ETHZ and University of Bologna) are research groups**
  - Motivated to develop new architectures and systems
  - We needed efficient RISC-V cores (and peripherals) for our work
  - Not so good (or interested) in providing industrial level support for these cores

• **We need help to**
  - Provide support
  - Develop industrial verification
  - Governance of open source repositories

• **Happy to receive this help from**
  - Open HW group (Ariane -> CVA6, RI5CY -> CV32E40P)
  - LowRISC (ZeroRiscy -> Ibex)
  - Others?
Academic open source → Industrial open source

• OpenHW Group is a not-for-profit, global organization (EU, NA, Asia) driven by its members and individual contributors where HW and SW designers collaborate in the development of open-source cores, related IP, tools and SW such as the Core-V family of cores.

• OpenHW Group provides an infrastructure for hosting high quality open-source HW developments in line with industry best practices.
PULP has been used in all our publicly funded research

Swiss Funding (nano-tera)
ICYSoC
Near threshold computing

EU funded research projects
OPRECOMP
Approximate computing
Multi-precision arithmetic

European FPA
EPI / EUPilot
Snitch based accelerators

International projects
MITACS / Polara
Vector extensions for RISC-V processors
Industrial Collaborations

**PULPv1,2,3 (ST28 FD)**
Demonstrators of 28nm FD-SOI capabilities
Various publications ‘15 – ‘18

**Arnold (GF22)**
IoT SoC combining eFPGA with RISC-V core
Schiavone et al TVLSI ‘19

**Vega (GF22)**
IoT Processor with ML acceleration
Rossi et al ISSCC ‘21
Rossi et al JSSC ‘22

**Marsellus (GF22)**
IoT Processor with low power modes and AI Accelerators
Conti et al ISSCC ‘23

Currently working with Meta, Intel, GF, IHP, PragmatIC, IIT
Open source collaboration scheme explained

Private

resources

GAP8

commits

PULP
Parallel Ultra Low Power

Mr. Wolf

commits

GitHub

Custom Designs

clone

issues

commits

PULP: 10 Years of Open Source Hardware

Frank K. Gürkaynak
The PULP-Based Commercial Nanodrone Platform

ULP heterogeneous model [1]

PULP-Shield [2]

AI-deck

- ~ 5 g – 30x28 mm
- PULP GAP8 SoC
- Off-chip DRAM/Flash
- QVGA ULP Camera
- GIS engine
- WiFi module

- ~ 8 g – 40x28 mm
- PULP GAP8 SoC
- Off-chip DRAM/Flash
- QVGA ULP Camera
- WiFi module

More on our work on Drones – talk by Lorenzo – Today 16:30
So proud to have supported others in their research
PULP chips until now

2013  2014  2015  2016  2017

2018  2019  2020  2021

2022

Frank K. Gürkaynak - PULP: 10 Years of Open Source Hardware
PULP chips until now – major publications
PULP chips until now – 64b/32b cores
PULP chips until now – 55 manufactured – 5 on the way

<table>
<thead>
<tr>
<th>Year</th>
<th>(No.)</th>
<th>Chip Name</th>
<th>Technology/Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>2013</td>
<td>(3)</td>
<td>PULPv1</td>
<td>STMicroelectronics 28FD SOI Multi-core processor with approximate FPUs</td>
</tr>
<tr>
<td>2014</td>
<td>(5)</td>
<td>Diana</td>
<td>UMC 65 4-core system with ML and Crypto accelerators</td>
</tr>
<tr>
<td>2015</td>
<td>(10)</td>
<td>Fulmine</td>
<td>UMC 65 4-core system with ML and Crypto accelerators</td>
</tr>
<tr>
<td>2016</td>
<td>(3)</td>
<td>VivoSoC 2.001</td>
<td>SMIC 130 Mixed signal system for biosignal acquisition</td>
</tr>
<tr>
<td>2017</td>
<td>(2)</td>
<td>Mr. Wolf</td>
<td>TSMC 40 8+1 core IoT processor</td>
</tr>
<tr>
<td>2018</td>
<td>(6)</td>
<td>Poseidon</td>
<td>GF 22FDX Dual 64bit RISC-V core, 32bit Microcontroller system, ML accelerator</td>
</tr>
<tr>
<td>2019</td>
<td>(7)</td>
<td>Baikonur</td>
<td>GF 22FDX Dual 64bit RISC-V core, 3x 8core snitch clusters, Body biasing test vehicle</td>
</tr>
<tr>
<td>2020</td>
<td>(3)</td>
<td>Dustin</td>
<td>TSMC 65 IoT processor with 16 cores and QNN enhancements</td>
</tr>
<tr>
<td>2021</td>
<td>(7)</td>
<td>Kraken</td>
<td>GF 22FDX IoT processor with Spiking Neural and Ternary Inference Engines</td>
</tr>
<tr>
<td>2022</td>
<td>(7)</td>
<td>Occamy</td>
<td>GF 12LPP ML accelerator with 216 + 1 cores and HBM interface</td>
</tr>
</tbody>
</table>

Check http://asic.ethz.ch for all our chips.
Coming soon from the PULP team

<table>
<thead>
<tr>
<th>FlexIBEX</th>
<th>Iguana</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>IoT processor with a twist</em>&lt;br&gt;a kHz range design</td>
<td><em>Going all the way in</em>&lt;br&gt;open source</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Carfield</th>
<th>Occammy</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Cars (and cats) can also use</em>&lt;br&gt;a bit of PULP</td>
<td><em>Bringing up the beast</em></td>
</tr>
</tbody>
</table>

Frank K. Gürkaynak - PULP: 10 Years of Open Source Hardware
A bit of a public service announcement

Join us in Barcelona – Registration open - https://riscv-europe.org/index.html

Monday-Friday 5-9 June 2023
There is much more to come ...