

ETH zürich



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

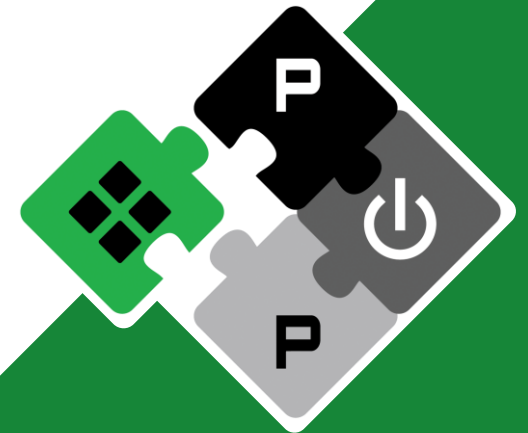
Industry Academia Collaborations on Open-Source Hardware

Integrated Systems Laboratory (ETH Zürich)

Frank K. Gürkaynak kgf@iis.ee.ethz.ch

PULP Platform

Open Source Hardware, the way it should be!



@pulp_platform 

pulp-platform.org 

youtube.com/pulp_platform 

PULP Platform by ETH Zürich and University of Bologna

OCCAMY

432 RISC-V cores

Chiptlets

GF12nm

1GHz



Occamy: A 432-Core 28.1 DP-GFLOP/s/W 83% FPU Utilization Dual-Chiplet, Dual-HBM2E RISC-V-based Accelerator for Stencil and Sparse Linear Algebra Computations with 8-to-64-bit Floating-Point Support in 12nm FinFET

How do we manage to design projects of this size at a University?

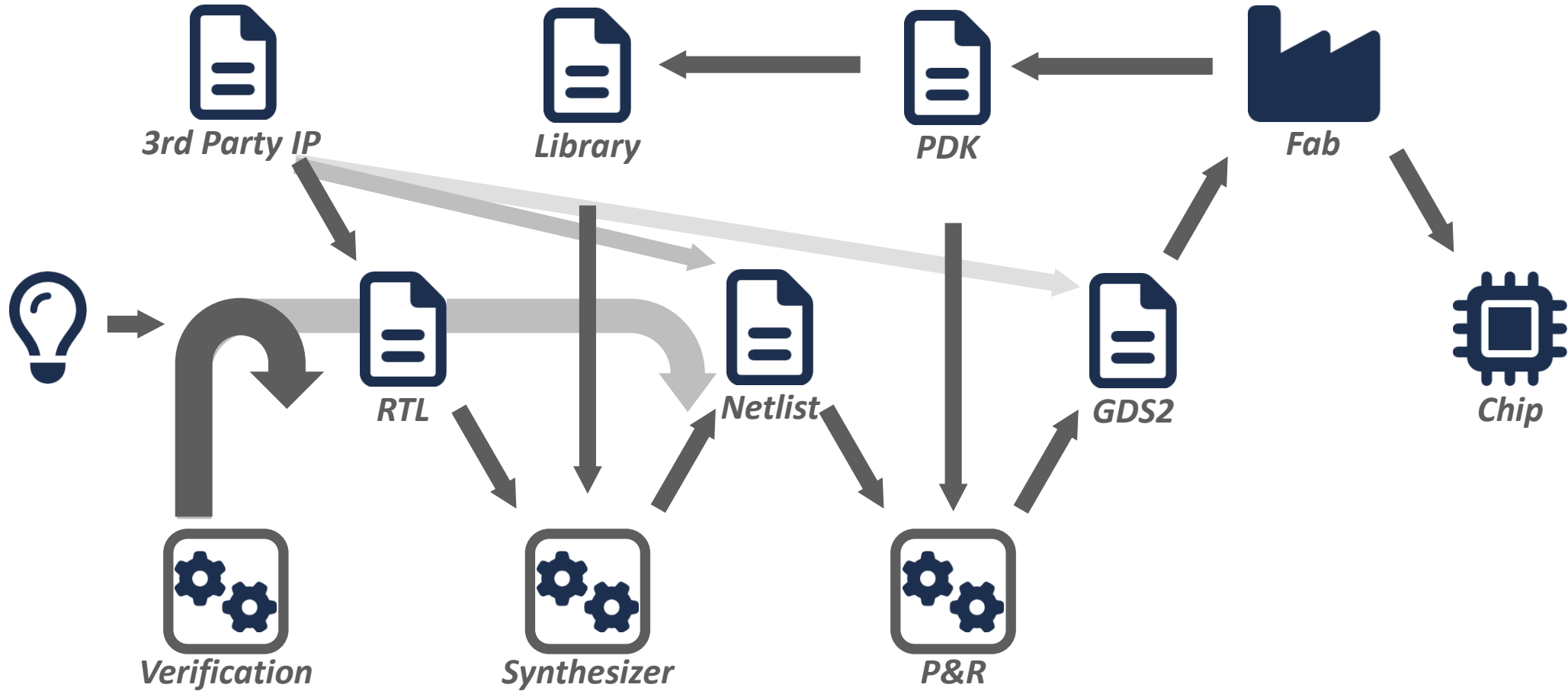
er,* Manuel Eggimann,*
Marco Ottavi,‡

In 11 years PULP team has designed more than 60 chips

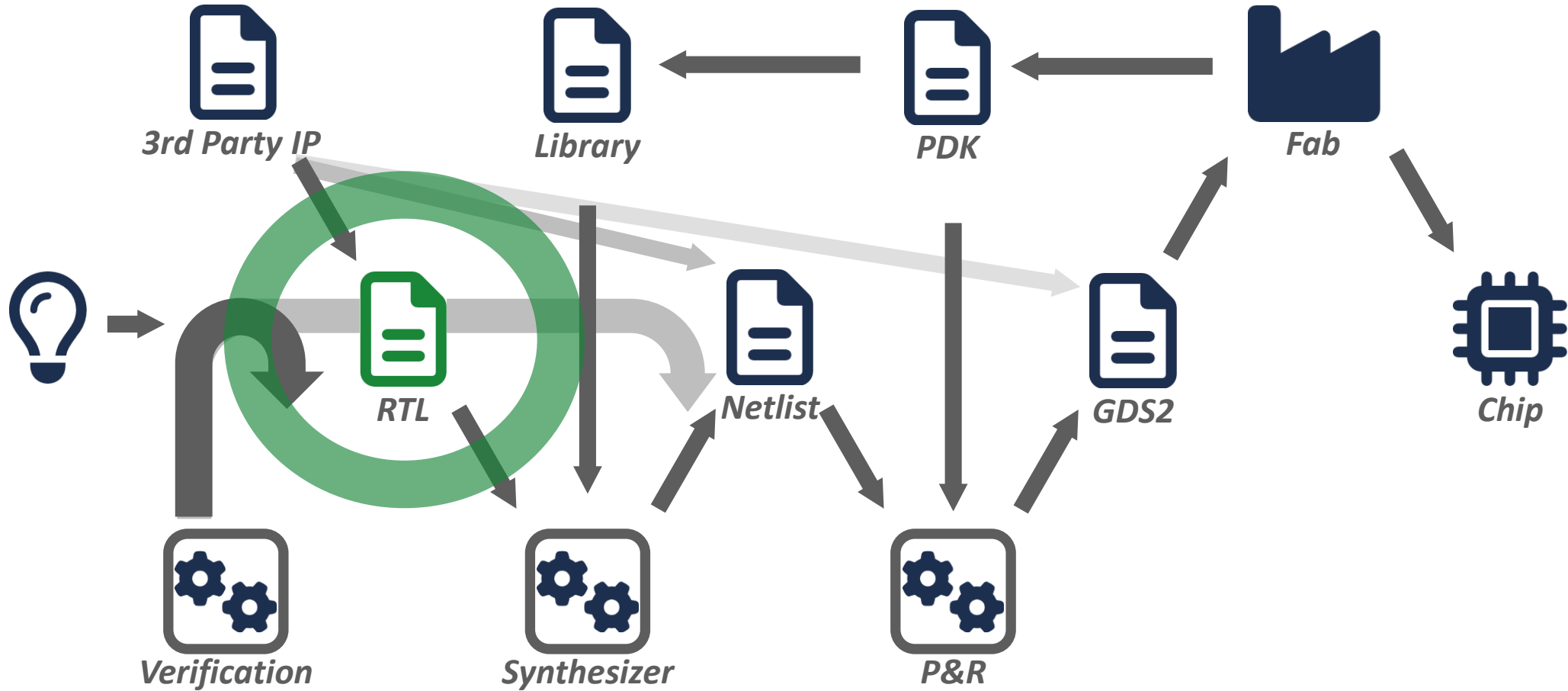


RISC-V and open-source hardware have been instrumental in our success

A simplified view of the IC design flow



Most of open source hardware is at RTL level



We have created a sandbox to design System on Chips



RISC-V Cores and Vector Units

RI5CY <i>CV32E</i>	Zero R <i>lbex</i>	Snitch	Spatz	Ariane <i>CVA6</i>	ARA
RV32	RV32	RV32	RVV	RV64	RVV

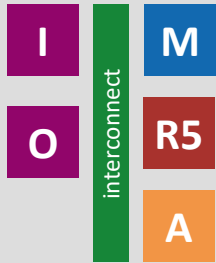
Peripherals

JTAG	SPI
UART	I2S
DMA	GPIO

Interconnects

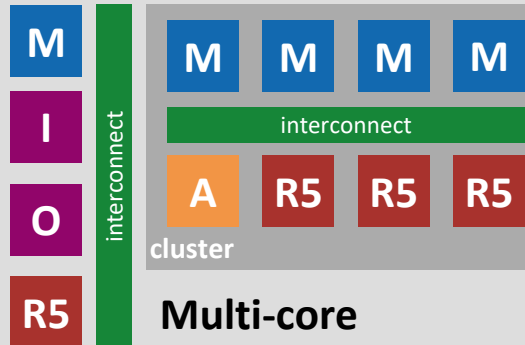
LIC	HCI
APB	FlooNoC
AXI4	

Platforms



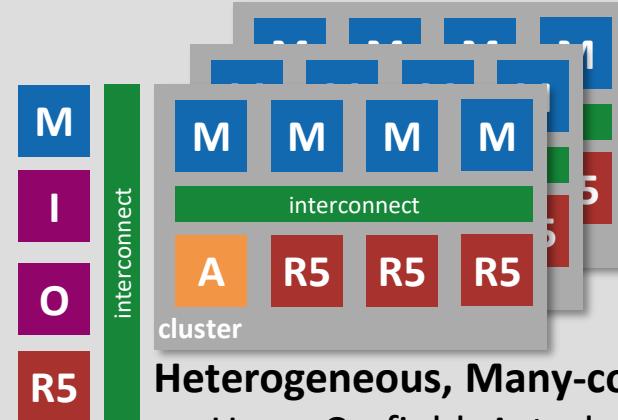
Single core

- PULPino, PULPissimo
- Cheshire



Multi-core

- OpenPULP
- ControlPULP



Heterogeneous, Many-core

- Hero, Carfield, Astral
- Occamy, Mempoool

IOT

HPC

Accelerators and ISA extensions

XpulpNN, XpulpTNN	ITA (Transformers)	RBE, NEUREKA (QNNs)	FFT (DSP)	REDMULE (FP-Tensor)
------------------------------------	-------------------------------------	--------------------------------------	----------------------------	--------------------------------------

We make everything (we can) available openly



- All our development is on GitHub using a **permissive** license
 - HDL source code, testbenches, software development kit, virtual platform

<https://github.com/pulp-platform>



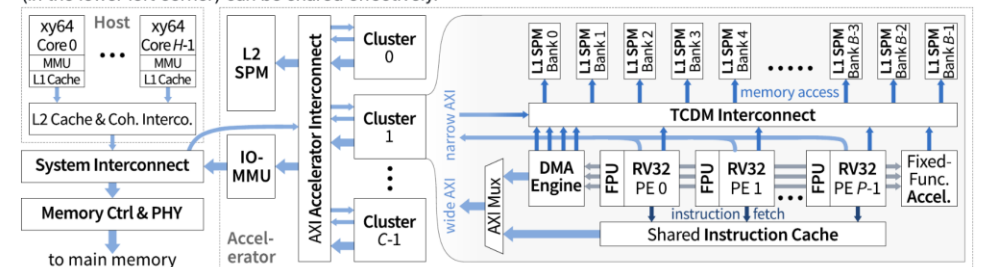
- Allows anyone to use, change, and make products without restrictions.

The screenshot shows the GitHub repository page for 'pulp-platform'. At the top, there is a repository card for 'pulp-platform' with a description: 'This is the top-level project for the PULP Platform. It instantiates a PULP open-source system with a PULP SoC (microcontroller) domain accelerated by a PULP cluster with 8 cores.' Below this, there are four pinned repository cards: 'pulp' (Public), 'pulpissimo' (Public), 'snitch' (Public), and 'hero' (Public). Each card includes a brief description and statistics like stars and forks.

Heterogeneous Research Platform (HERO)

HERO is an FPGA-based research platform that enables accurate and fast exploration of heterogeneous computers consisting of programmable many-core accelerators and an application-class host CPU. Currently, 32-bit RISC-V cores are supported in the accelerator and 64-bit ARMv8 or RISC-V cores as host CPU. HERO allows to seamlessly share data between host and accelerator through a unified heterogeneous programming interface based on OpenMP 4.5 and a mixed-data-model, mixed-ISA heterogeneous compiler based on LLVM.

HERO's hardware architecture, shown below, combines a general-purpose host CPU (in the upper left corner) with a domain-specific programmable many-core accelerator (on the right side) so that data in the main memory (in the lower left corner) can be shared effectively.



Meet Mr. Wolf (2017) in TSMC40



- **Very successful IoT processor**
 - 8+1 RISC-V cores
- **Power converter IP from Dolphin**

Win (PULP): get to use professional IP in our chips → the company to demonstrate their industry relevant design

- RTL for the entire SoC openly available

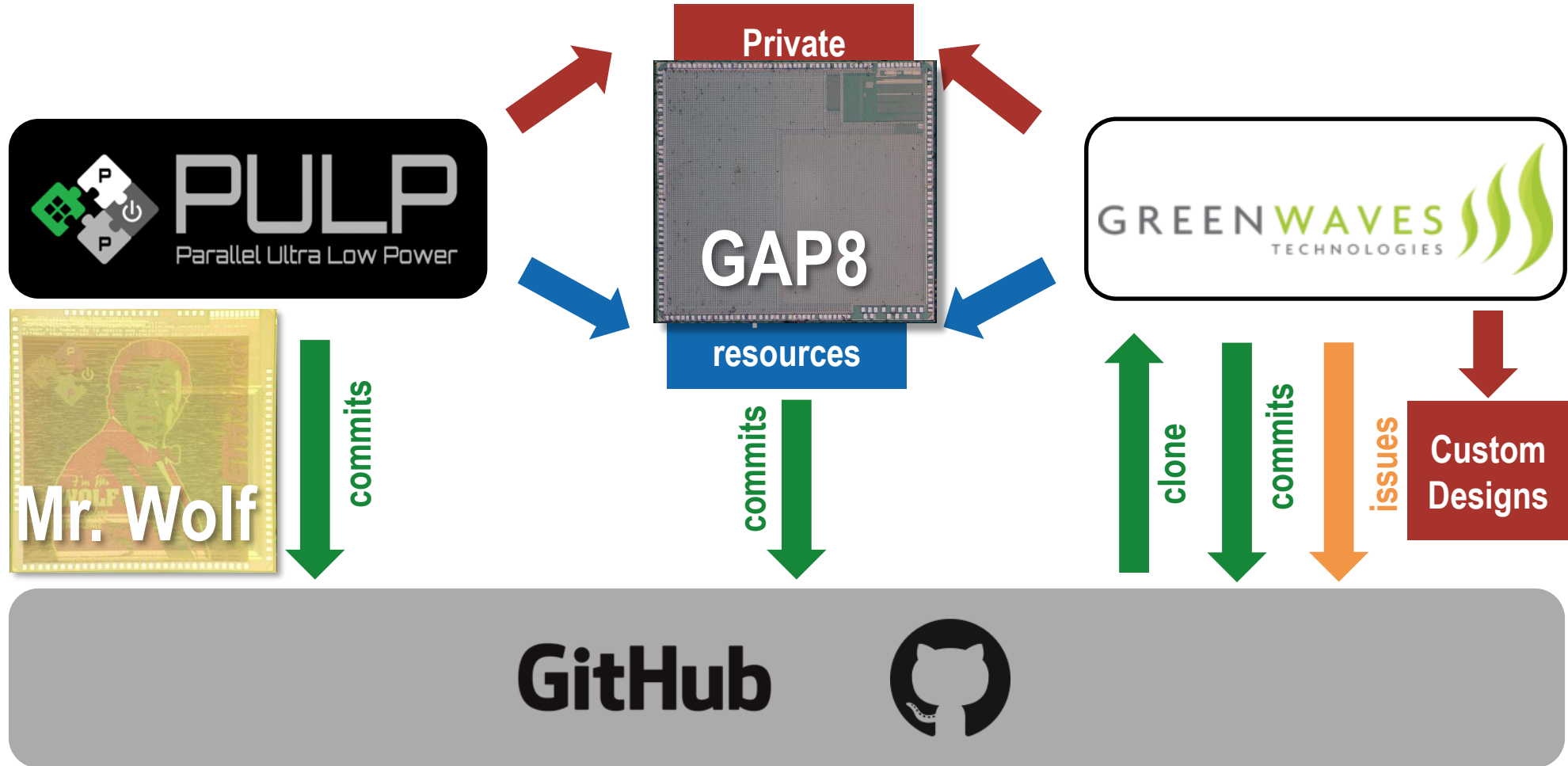
Win (Dolphin): demonstrate their IP on a SoA design →



- **Design formed the basis of GAP8/9**
 - By Greenwaves Technologies

Win (Greenwaves): SoC template that can be easily productized →

How does PULP collaborate with 3rd parties?



Designs derived from Mr. Wolf powered our nano-drones

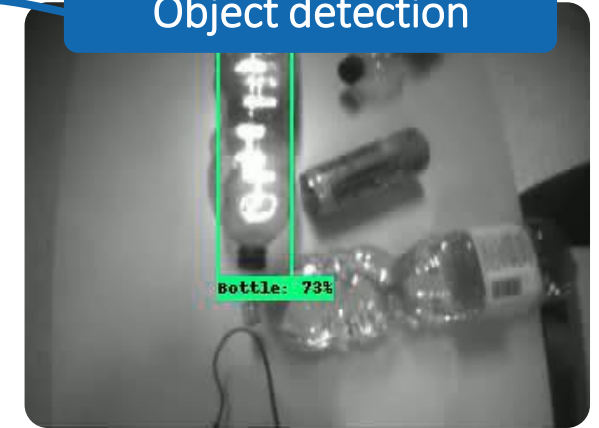


Multiple, complex, heterogeneous tasks at high speed and robustness using **only our onboard RISC-V based IoT processors**

Obstacle avoidance & Navigation



Object detection

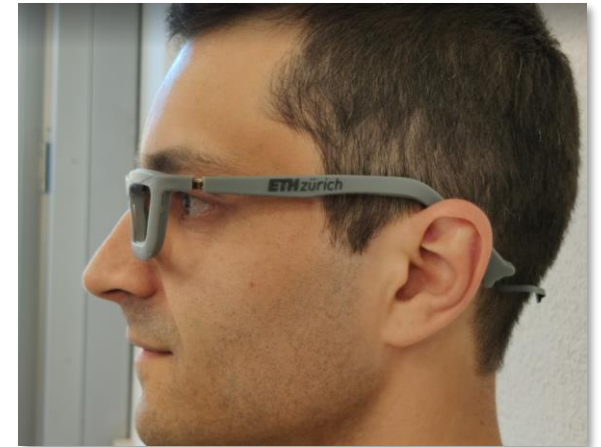
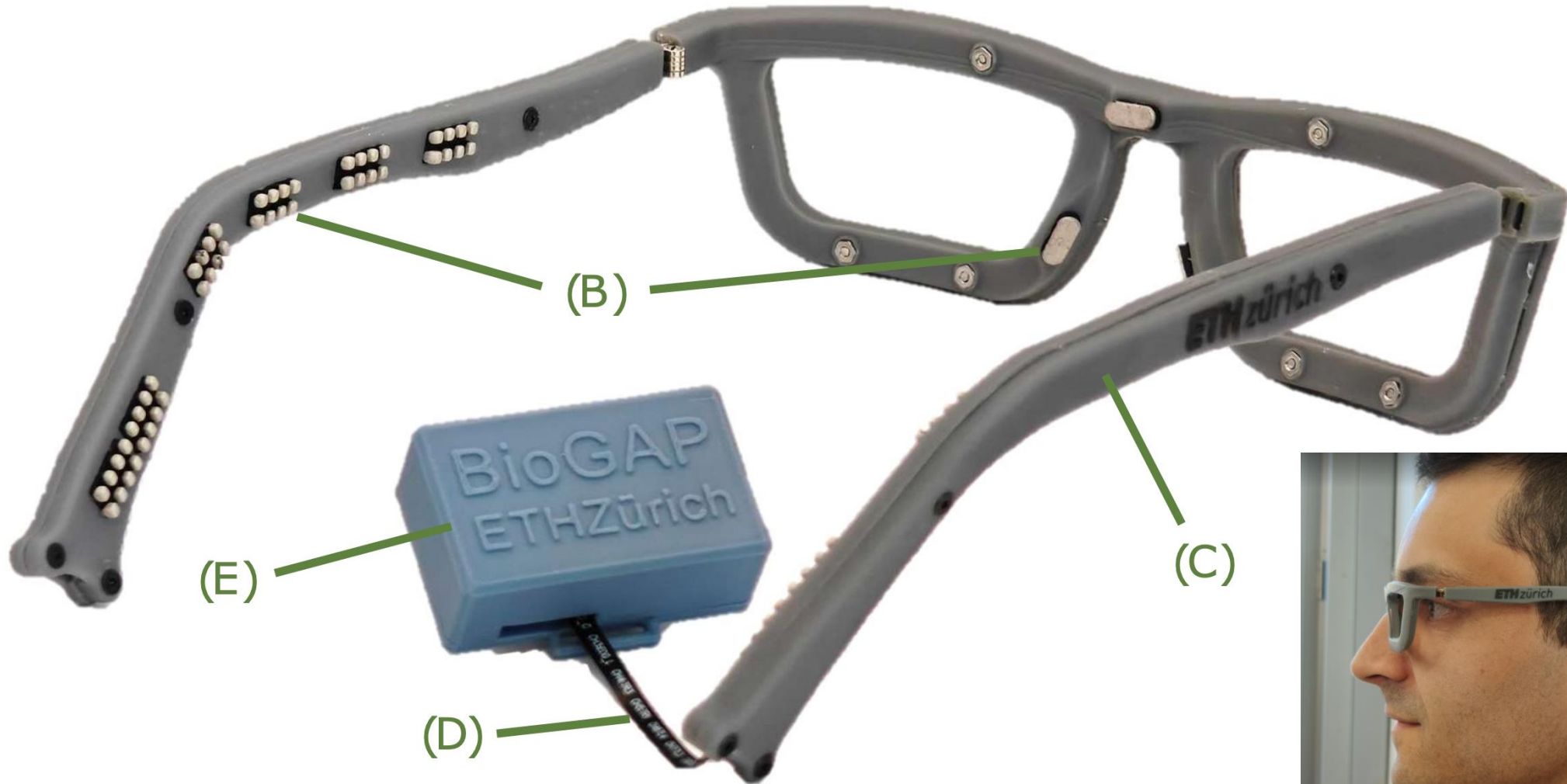


Environment exploration

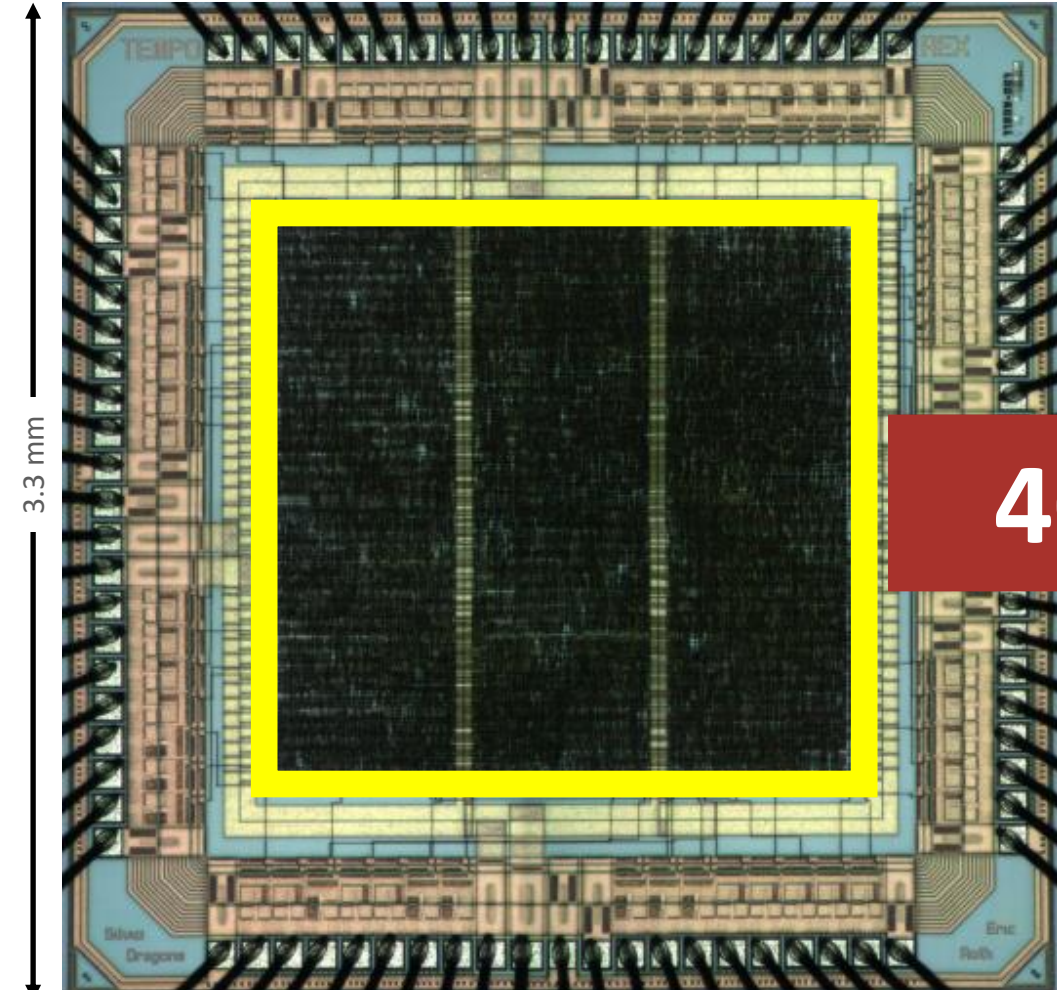


Multi-GOPS workload at extreme efficiency $\rightarrow P_{\max}$ 100mW

As well as our bio-signal acquisition systems



In the last 20 years IC Design has changed a lot



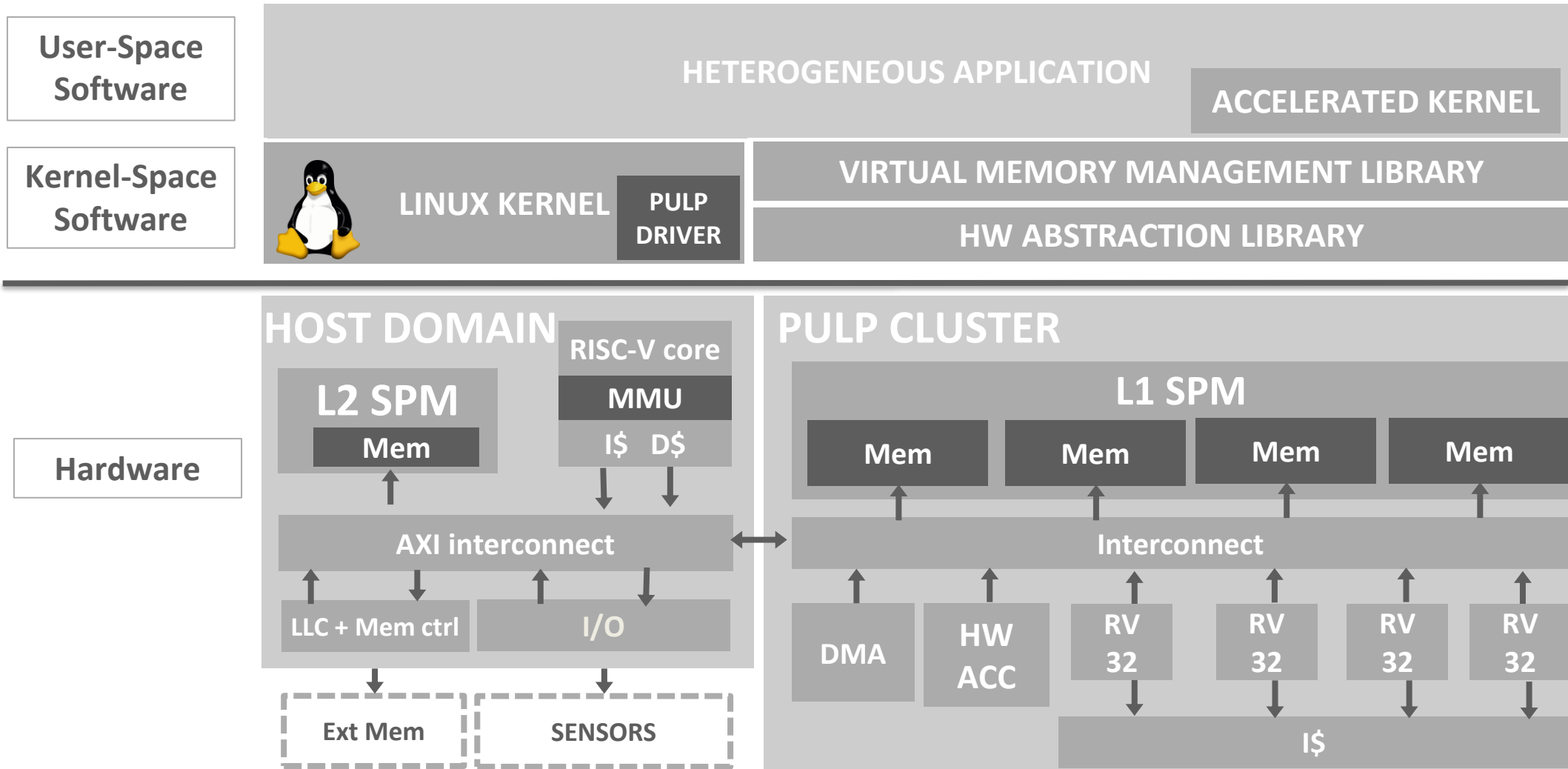
4000 x



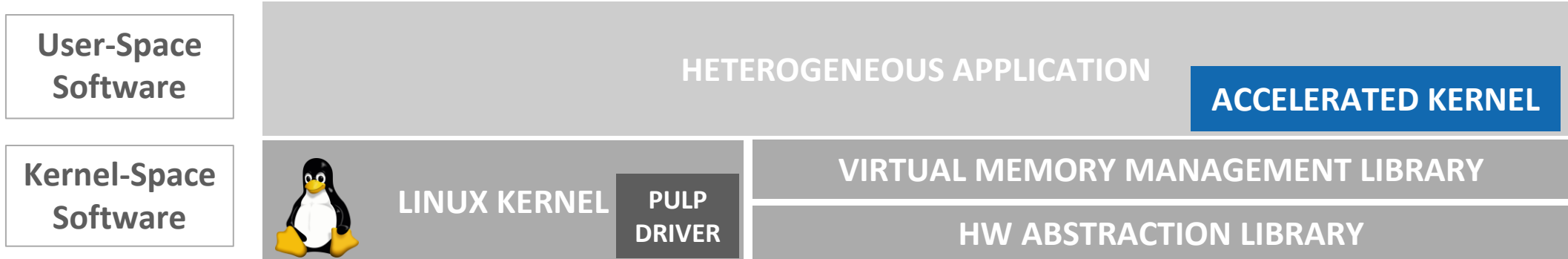
80 MGE

What used to be a complete chip is now a small part of a SoC !

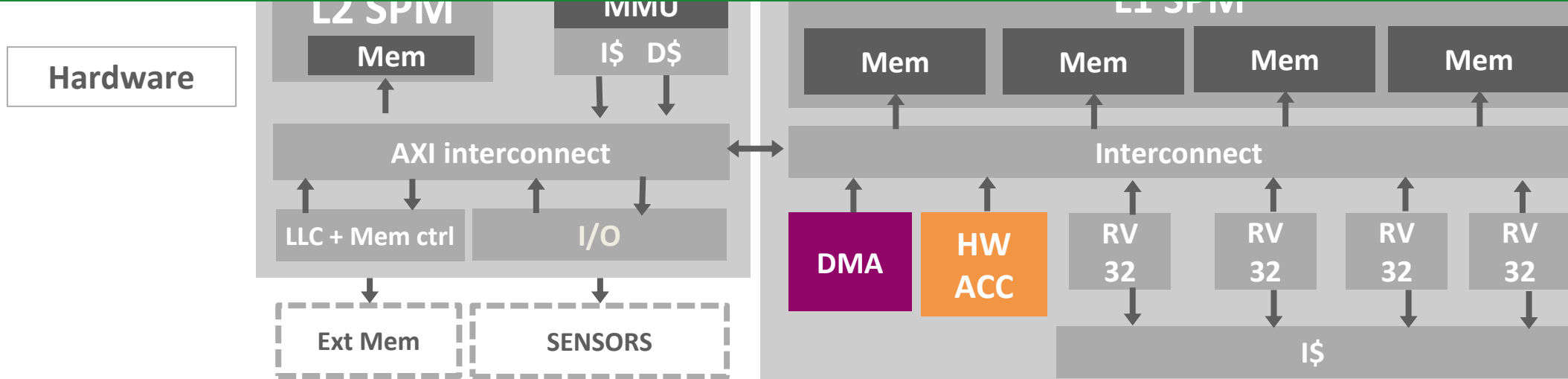
There is so much that makes up a modern SoC



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



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



Diverse set of open source based industry collaborations

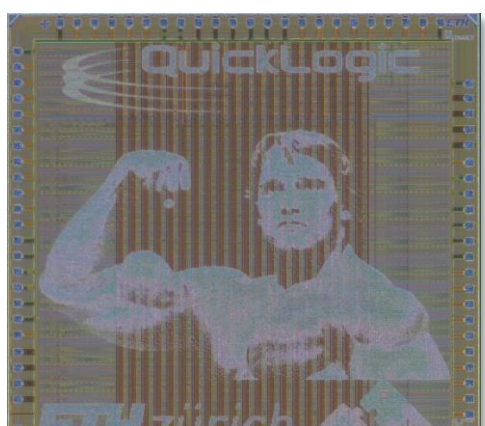


GF22 (2018)

Arnold

eFPGA coupled with a RISC-V microcontroller.

In one year from agreement to actual tapeout

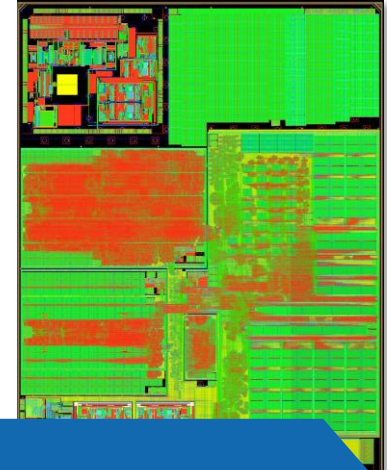


GF22 (2022)

Marsellus

Heterogeneous IoT processor
With Aggressive voltage scaling

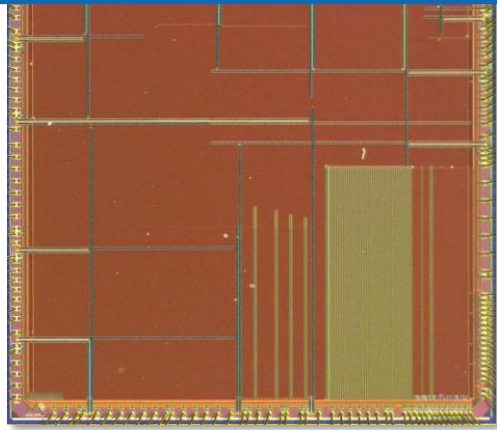
DOLPHIN



Permissive open-source licensing key to our industrial relationships

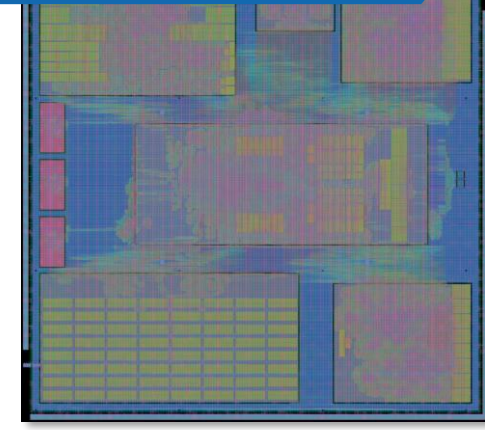
Siracusa

SoC for Extended Reality
visual processing



Carfield

Open-Research platform for
safety, resilient and
time-predictable systems



And many continue to use our work for their research



Smallest RISC-V Device for Next-Generation Edge Computing

RISC-V Workshop

Our 1st gen. processor and 2.5D integrated device

Processor SoC (D02)

SoC size: 300 μm x 250 μm , GF14LPP
 SoC arch: Based on PULPino (RV32IMC) + PULPino
 On chip memory: 2KB data SRAM
 + Authentication engine
 + Analog custom circuits (LDO, Clock/Reset, PD/LED IF)

Seiji Munetoh¹, Chitra K Subramanian², Arun Paidimarri², Yasuteru Kohda²
 IBM Research – Tokyo¹ & T.J. Watson Research Center²

RISC-V week Barcelona 2018

An 8-core RISC-V Processor with Compute near Last Level Cache in Intel 4 CMOS

Gregory K. Chen, Phil C. Knag, Carlos Tokunaga, Ram K. Krishnamurthy
 Circuit Research Lab, Intel Corporation, Hillsboro, OR, USA, gregory.k.chen@intel.com

ISA	RV64GC
Execution	Out-of-order
L1I	16kB/core, 4-way
L1D	8kB/core, 4-way
NoC	64b 2D Mesh
L2 LLC	512kB, 4-way
LLC BW	1.0 Tb/s
CNC Area Overhead	1.4%
#CNC MACs	128
CNC RF	1kB/slice
Energy Eff. 0.6V	285 GOPS/W
LLC Energy Eff. 0.6V	1.6 TOPS/W

VLSI Symposium 2022

The Deep Learning Revolution and Its Implications for Computer Architecture and Chip Design

Presenting the work of many people at Google

AutoDMP: Automated DREAMPlace-based Macro Placement

Anthony Agnesina aagnesina@nvidia.com NVIDIA Corporation Austin, TX, USA	Puranjay Rajvanshi prajvanshi@nvidia.com NVIDIA Corporation Santa Clara, CA, USA	Tian Yang tiyang@nvidia.com NVIDIA Corporation Santa Clara, CA, USA	Geraldo Pradipta gpradipta@nvidia.com NVIDIA Corporation Santa Clara, CA, USA
Austin Jiao ajiao@nvidia.com NVIDIA Corporation Santa Clara, CA, USA	Ben Keller benk@nvidia.com NVIDIA Corporation Santa Clara, CA, USA	Brucek Khailany bkhailany@nvidia.com NVIDIA Corporation Austin, TX, USA	Haoxing Ren haoxingr@nvidia.com NVIDIA Corporation Austin, TX, USA

Some smaller companies you might have heard of 😊

Accepted: 12 April 2021
 Published online: 9 June 2021

Quoc V. Le, James Liao, Richard...

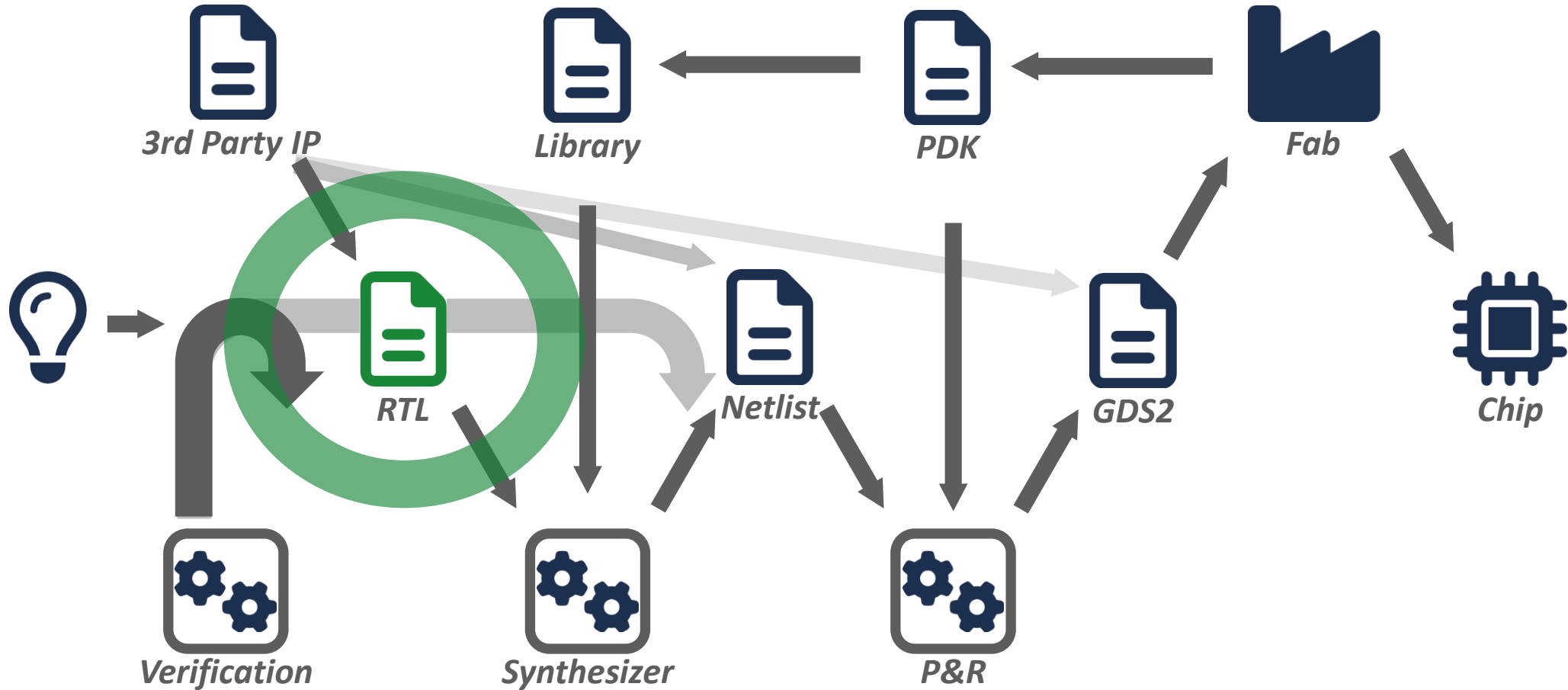
Fig. 4 | Convergence plots on Ariane RISC-V CPU. Placement cost of training a policy network from scratch versus fine-tuning a pre-trained policy network for a block of Ariane RISC-V CPU.

ISSCC Keynote 2020 – Nature 2020

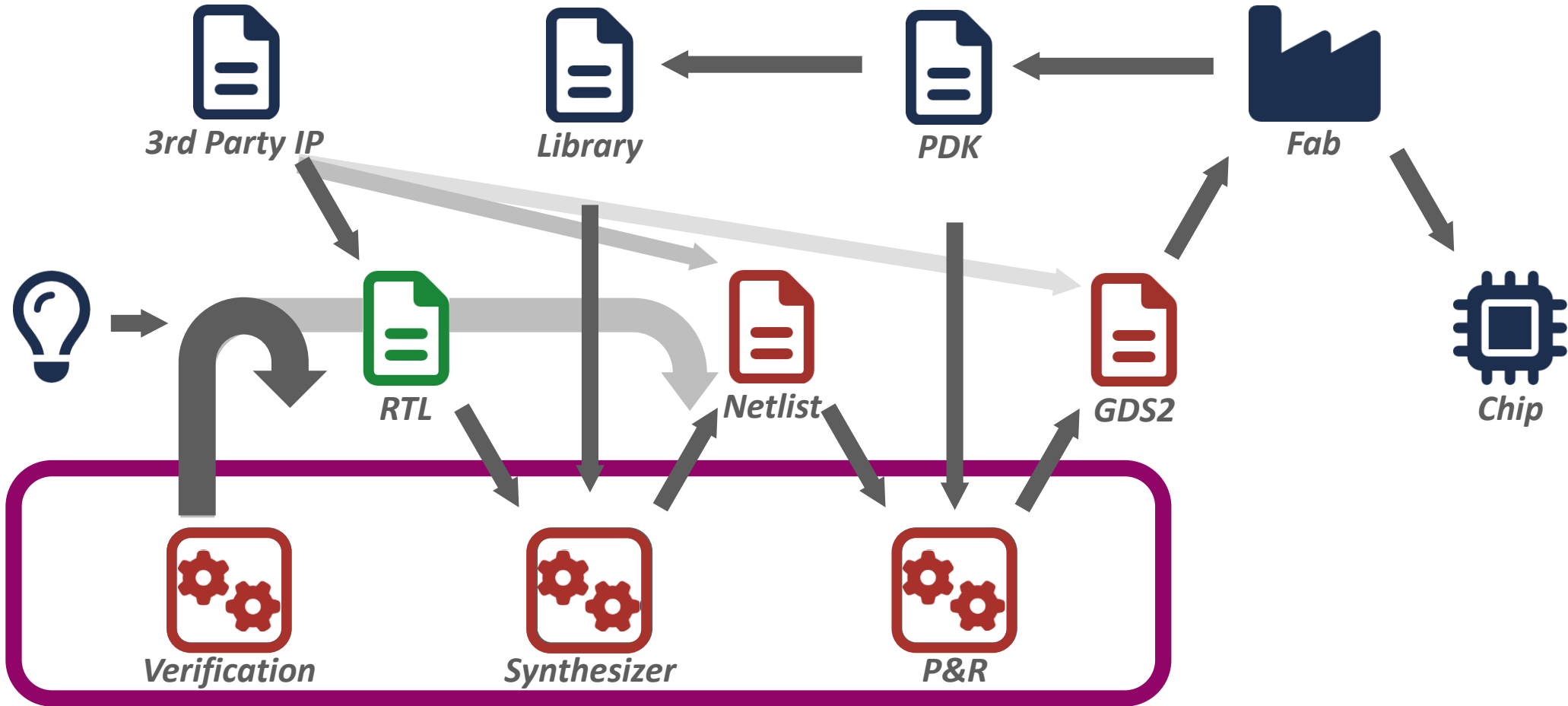
Figure 7: Pre-CTS placements of the logical groups and cell densities of the MemPool Group designs using NanGate 45nm process (freq. = 333 MHz, density = 68%). Congestion (H/V): Innovus (2.66%/1.54%), AutoDMP (3.48%/1.86%).

ISPD'23

Unlocking the rest of the design flow: Open PDK and EDA

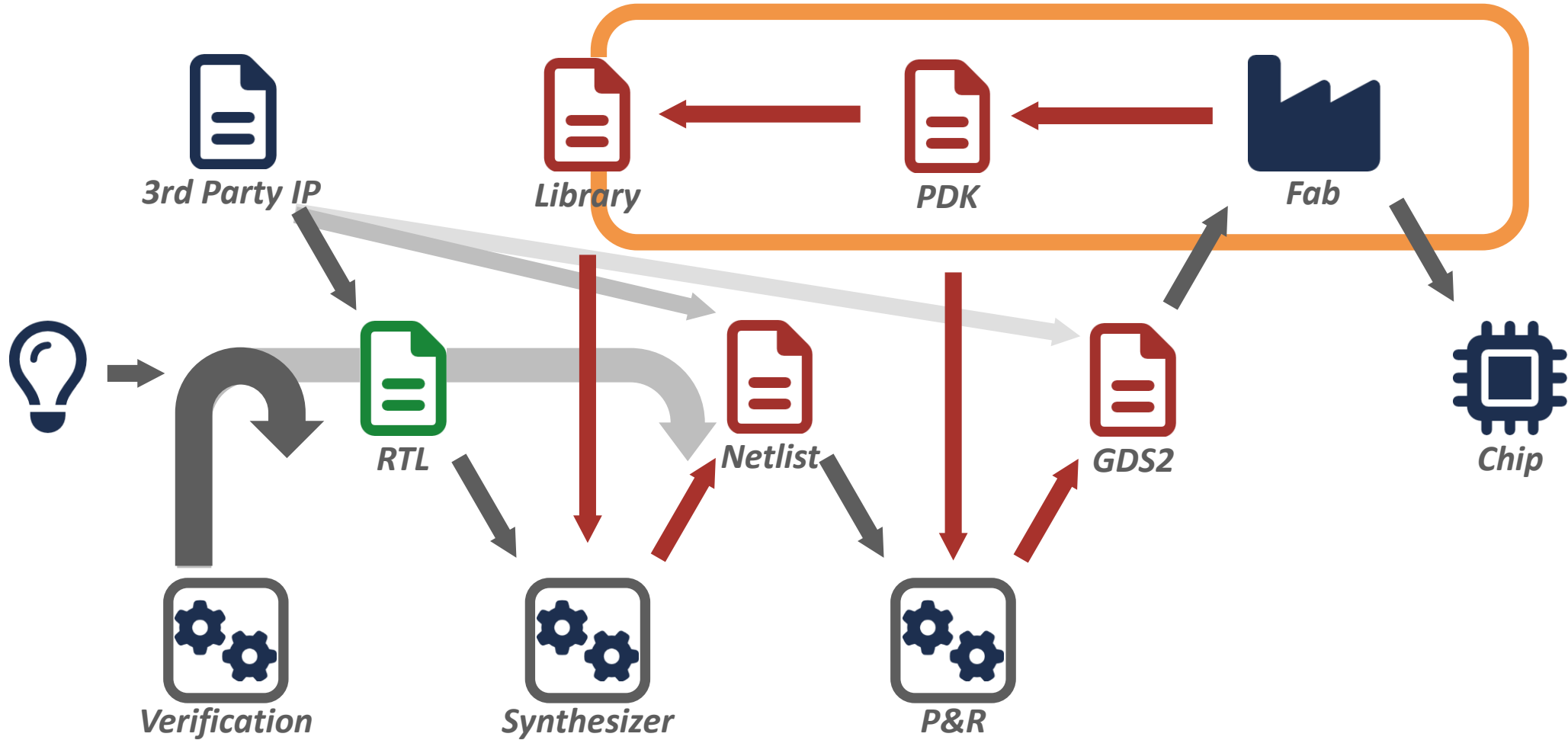


The output (and even scripts) of EDA vendors are closed



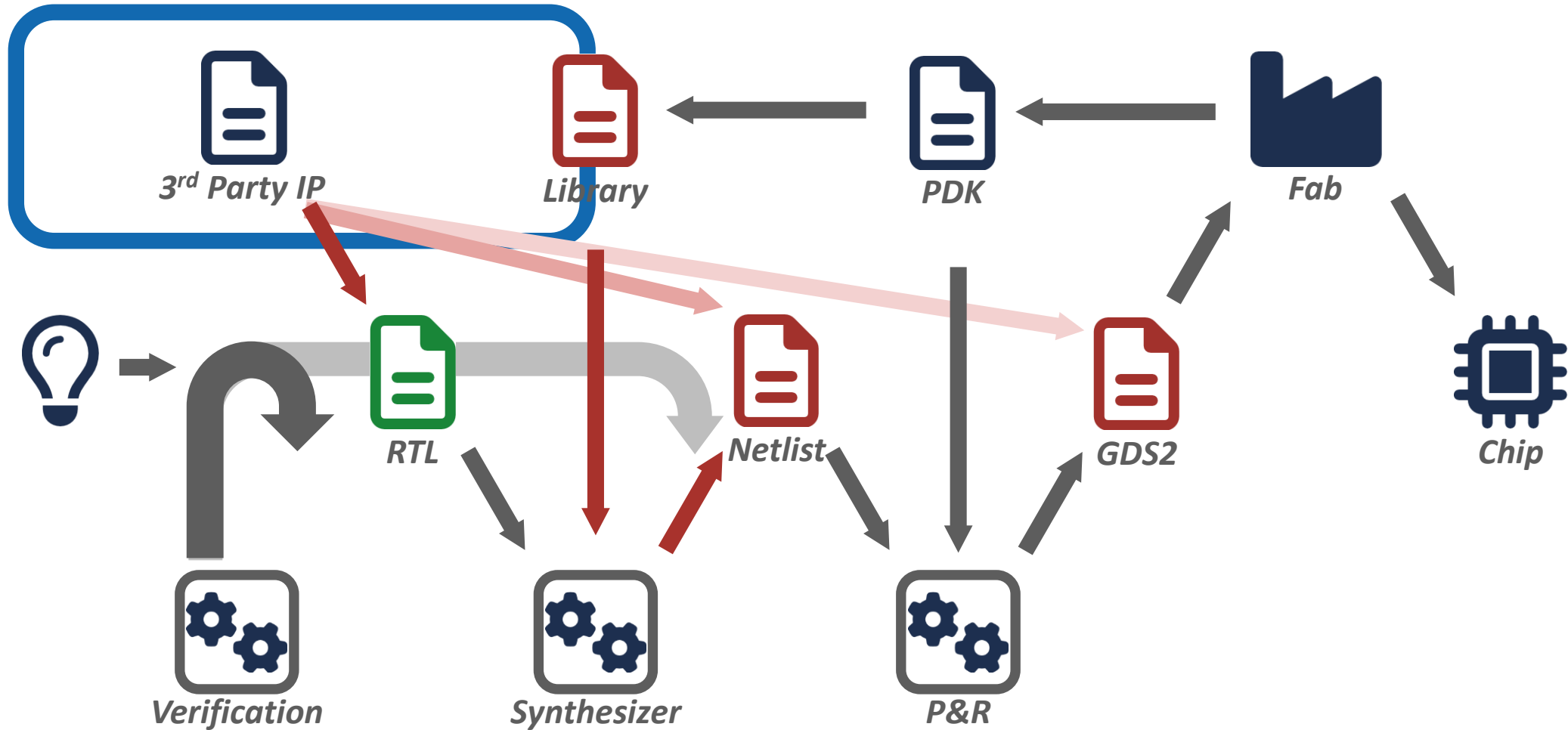
EDA vendors limit the output of their tools

The chip will contain information from the PDK of the Fab



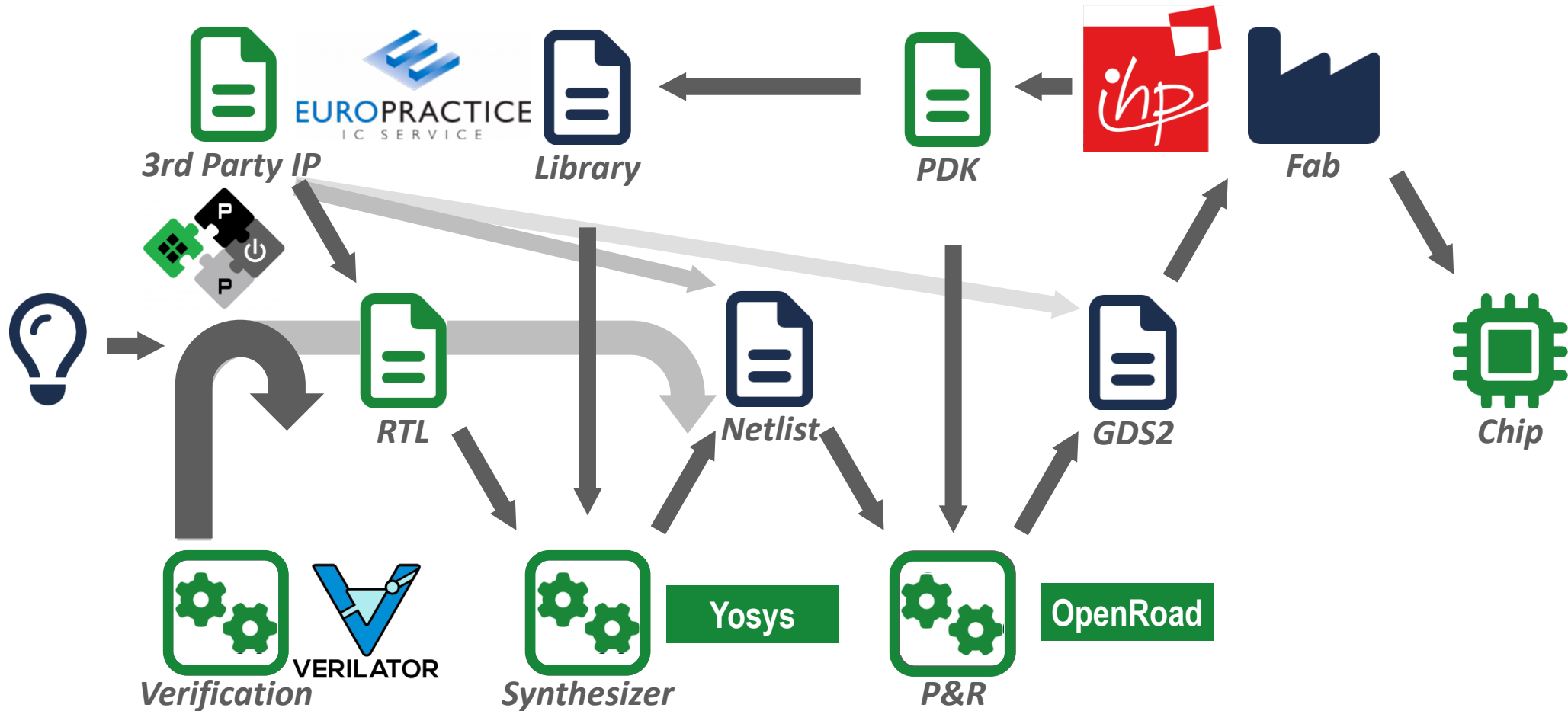
Fabs do not make PDK information accessible

Most designs will include some 3rd party IP



3rd party IP when included can limit what can be open sourced

We need openness along the whole chain: RTL, EDA, PDK



We are getting there, first fully open chips are underway

Meet Basilisk: Open RTL, Open EDA, Open PDK



- **Designed in IHP 130nm OpenPDK**

- 6.25mm x 5.50mm
- 60MHz
 - 1.08 MGE logic, 60% density
 - 24 SRAM macros (114 KiB)

- **CVA6 based SoC**

- Runs and boots Linux

- **Active collaboration with**



Open-source EDA tools already enable complex designs, it will only get better

Working with open-source EDA groups to close the gap!



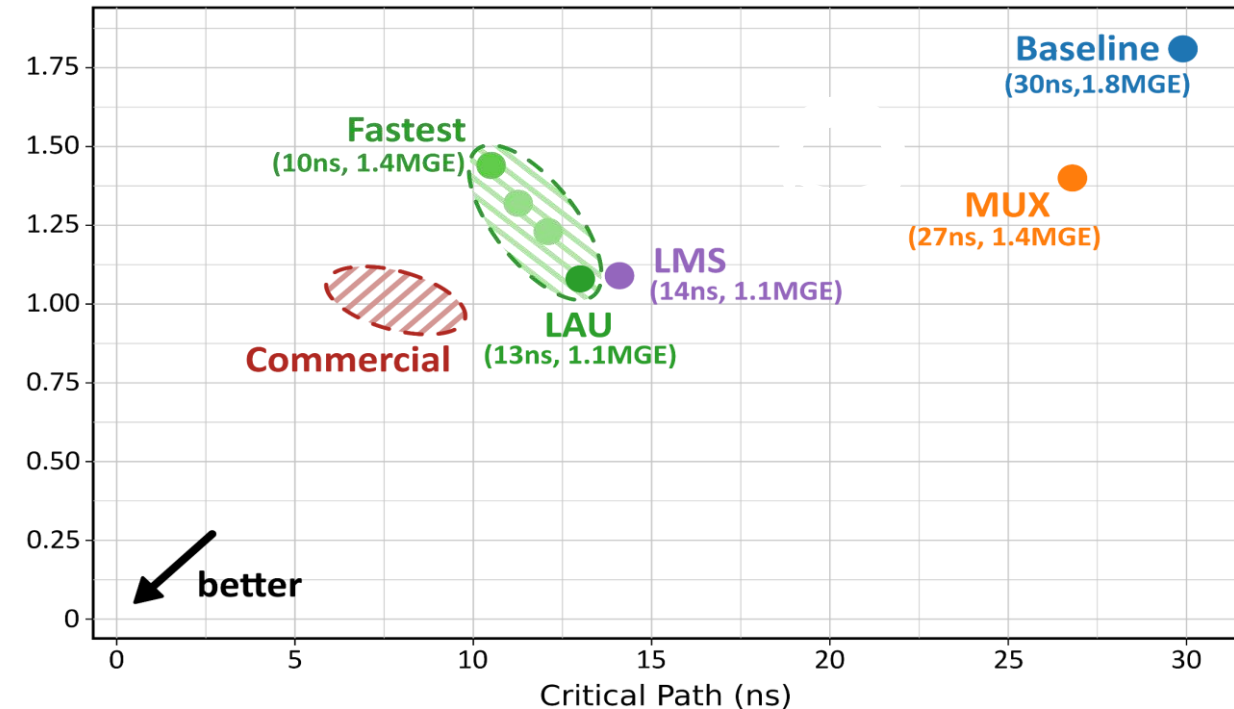
- **Basilisk is the first end-to-end open-source Linux-capable RV64 SoC**

- DRAM interface & rich IO (USB 1.1, VGA, SPI, ...)
- Silicon-proven, configurable, MGE-scale design

- **Improved FOSS EDA flow**

- SV-to-Verilog chain @ <2min runtime
- Yosys synthesis:
 - **1.1 MGE (1.6x) @ 77 MHz (2.3x)**
 - **2.5x** less runtime, **2.9x** less peak RAM
- OpenROAD P&R: tuning
 - **-12%** die area, **+10%** core utilization

Logic Area (MGE)



github.com/pulp-platform/cheshire-ihp130-o

Benefits of end to end openness



Research

- Easier collaboration (no NDAs)
- Reproducible results, benchmarking
- **Combined impact of design and design automation**

Education

- Increased accessibility
- No black boxes, full visibility
- **Experiment with flows and tools**

Industry

- Transparent chain of trust, sovereignty
- Lower initial cost
- **Faster research → product**

Our open source principles support funded projects



Successful partnerships with both industry and academia



Final words

- **We use open source because it works**
 - Allows us to manage complex designs
 - Facilitates Industry/Academia Relationships
 - Creates Auditable Designs, Reproducible Results

There is still
more to come 😊

Helps us and others concentrate work where it matters

- **Open Source sees no borders**
 - There is **no** 'European/Chinese/American Open Source',
 - There **can be** 'European/Chinese/American support for Open Source'

Open Source is global, it just can have more or less support in a region/country

