

# PULP within the Open Source Hardware Landscape

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 

# Parallel Ultra Low Power (PULP) Platform



- Research on open-source energy-efficient computing





# We have designed and tested more than 50 PULP ICs



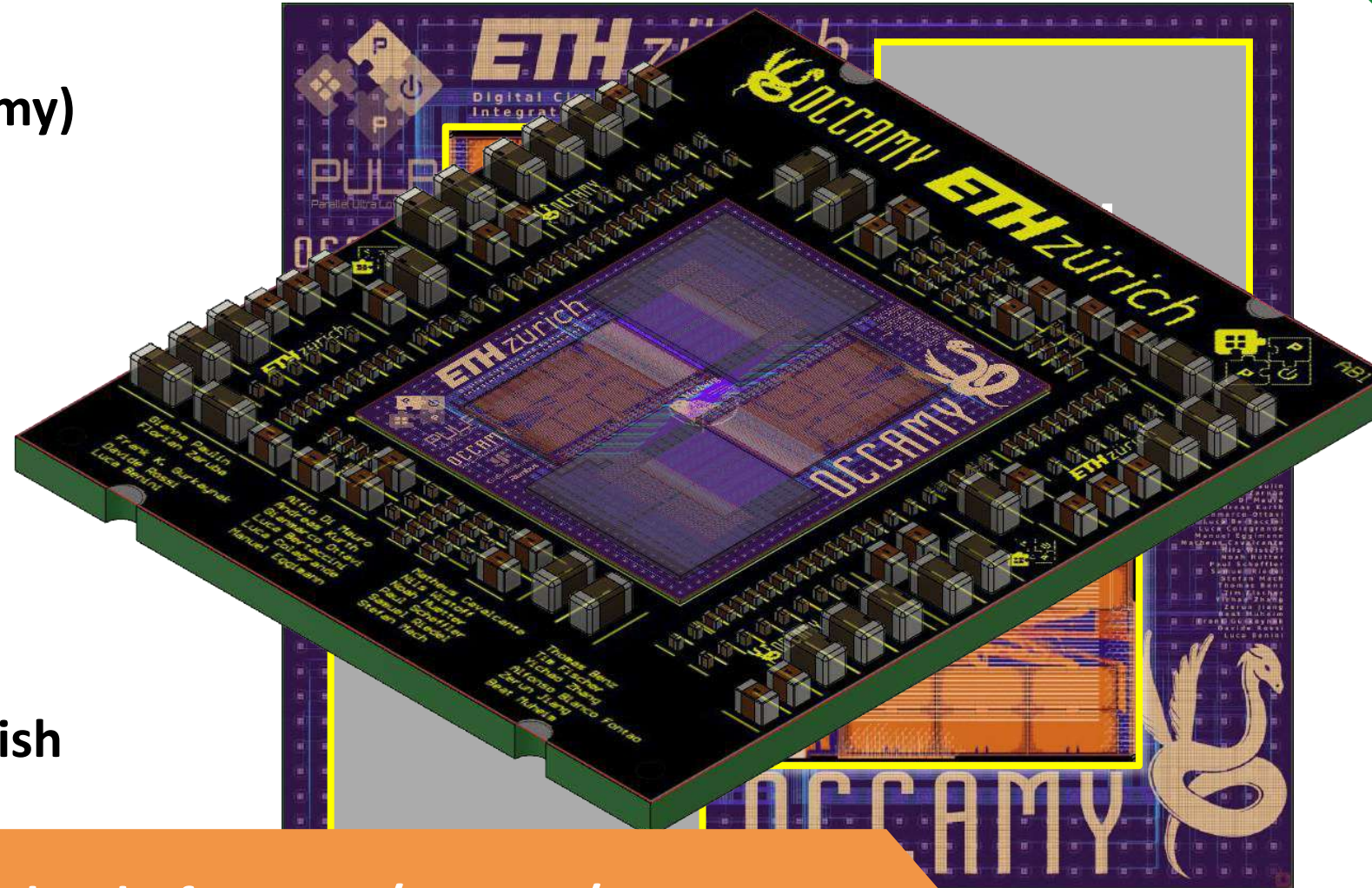
2013 (3)	2014 (5)	2015 (10)	2016 (3)	2017 (2)	2018 (6)	2019 (7)	2020 (3)	2021 (7)	2022 (9)
 	  	   	 	 	  	  	 	  	   
<b>PULPv1</b> <i>STM 28FDSOI</i> Multi-core processor	<b>Diana</b> <i>UMC 65</i> 4-core system with approximate FPU's	<b>Fulmine</b> <i>UMC 65</i> 4-core system with ML and Crypto accelerators	<b>VivoSoC 2.001</b> <i>SMIC 130</i> Mixed signal system for biosignal acquisition	<b>Mr. Wolf</b> <i>TSMC 40</i> 8+1 core IoT processor	<b>Poseidon</b> <i>GF 22FDX</i> 64bit RISC-V core, 32bit Microcontroller system, ML accelerator	<b>Baikonur</b> <i>GF 22FDX</i> Dual 64bit RISC-V core, 3x 8core snitch clusters, Body biasing test vehicle	<b>Dustin</b> <i>TSMC 65</i> IoT processor with 16 cores and QNN enhancements	<b>Kraken</b> <i>GF 22FDX</i> IoT processor with Spiking Neural and Ternary Inference Engines	<b>Occamy</b> <i>GF 12LPP</i> ML accelerator with 216 + 1 cores and HBM interface

Check <http://asic.ethz.ch> for all our chips

# Including some very complex designs like Occamy



- Chiplet based design
- **2x Compute chiplets (Occamy)**
  - 216+1 RISC-V cores
  - 0.75 TFLOP/s for the system
  - GF12LPP
  - Running at 1 GHz
- **2x 16GB HBM DRAMs**
- **Silicon Interposer**
  - GF 65nm 26mm x 23mm
- **Waiting for assembly to finish**



More on Occamy – <https://pulp-platform.org/occamy/>



# All of our designs are open source hardware



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

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



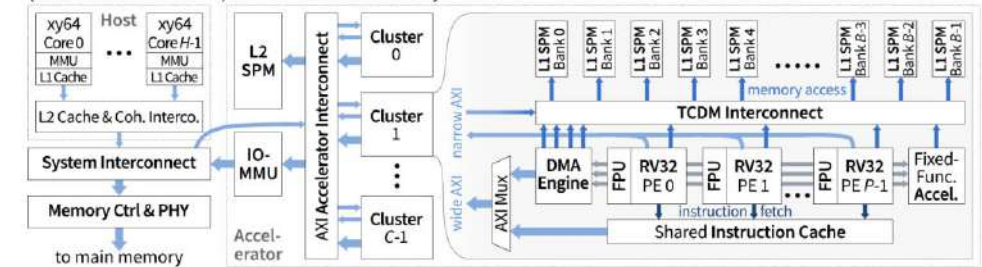
- Allows anyone to use, change, and make products without restrictions.

The screenshot shows the GitHub profile for 'pulp-platform'. It includes a repository overview with 239 repositories, 1 project, and 14 people. Under the 'Pinned' section, four repositories are listed: 'pulp' (Public), 'pulpissimo' (Public), 'snitch' (Public), and 'hero' (Public). Each repository has a brief description and statistics like stars and forks.

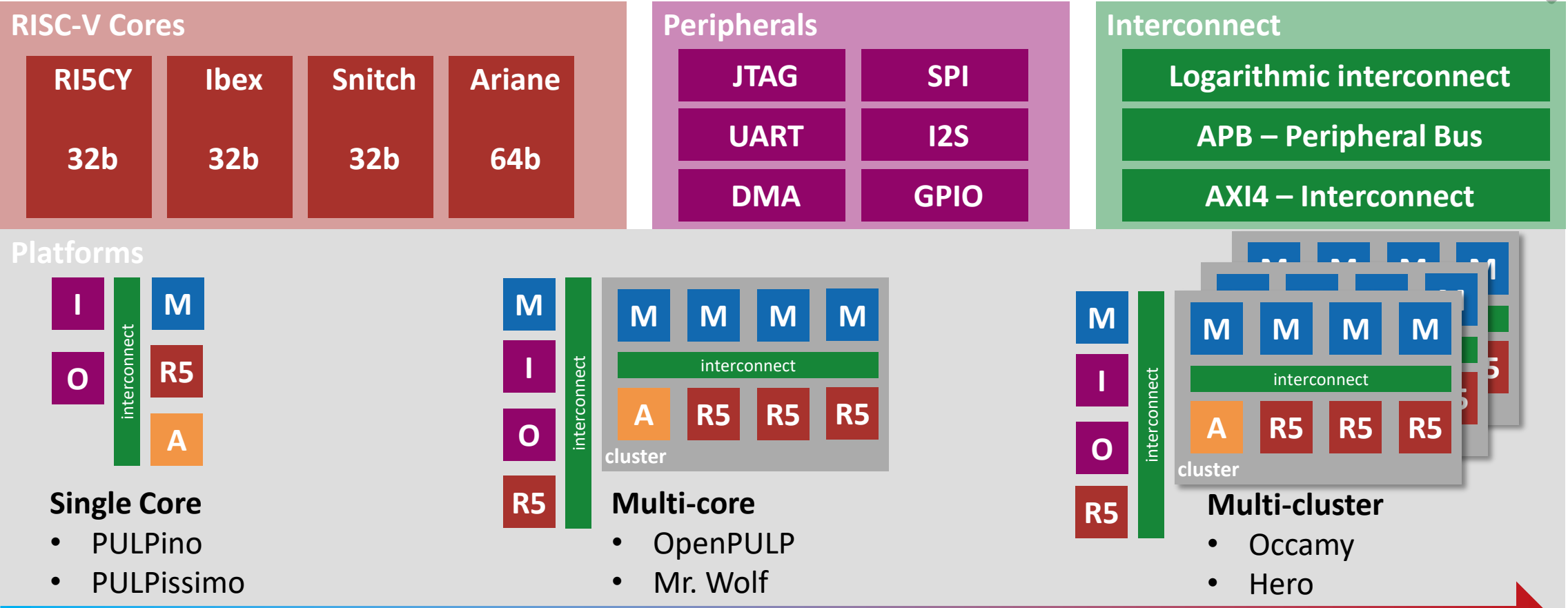
## Heterogeneous Research Platform (HERO)

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.



# What PULP provides is a box of building blocks



**IOT**

**HPC**

## Accelerators

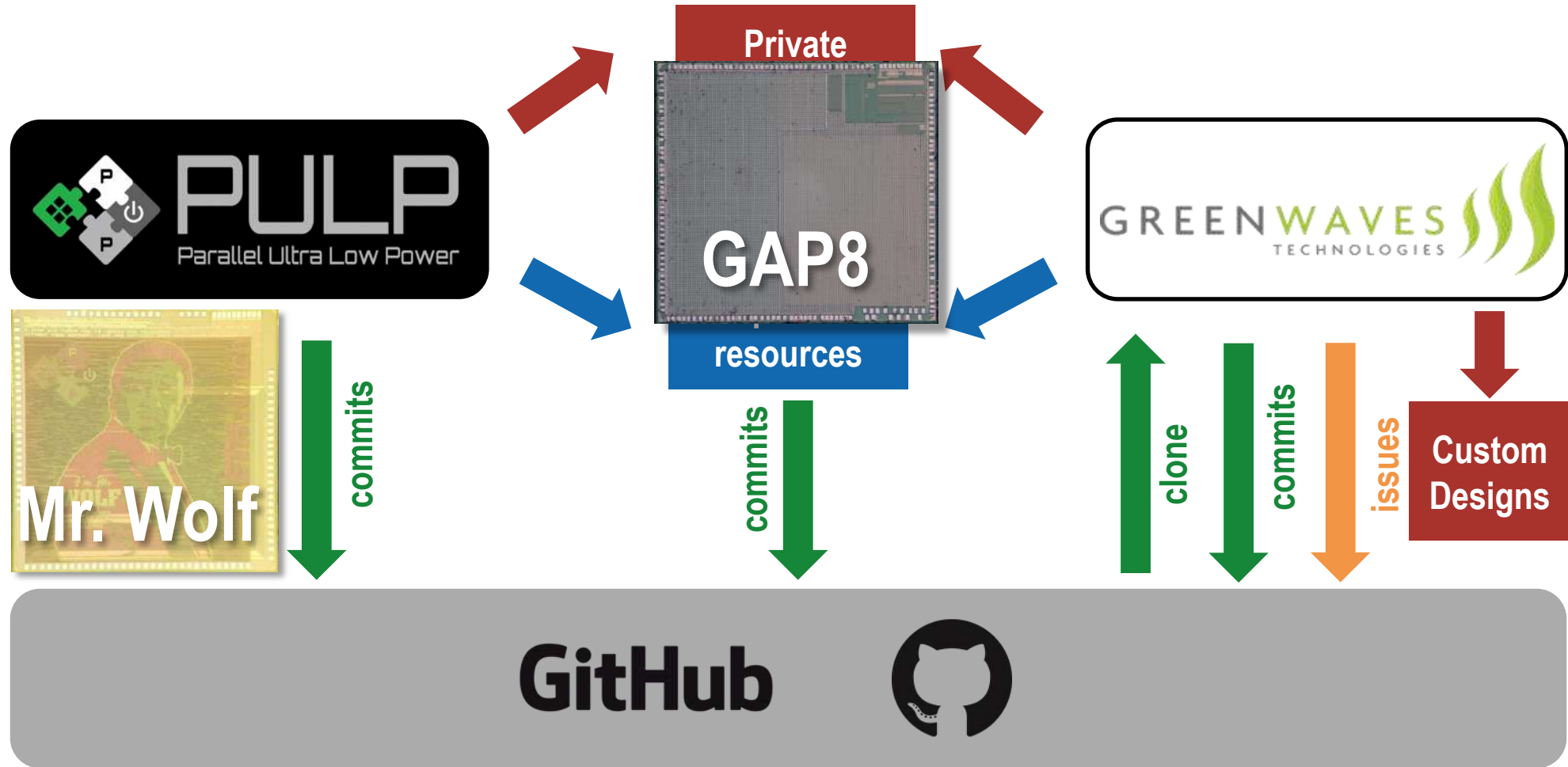
**HWCE  
(convolution)**

**Neurostream  
(ML)**

**HWCrypt  
(crypto)**

**PULPO  
(1<sup>st</sup> order opt)**

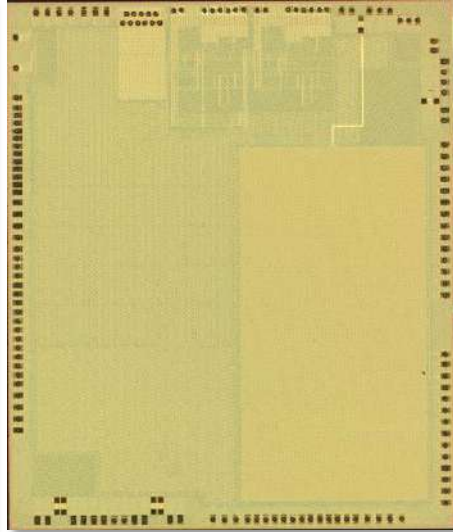
# How does PULP collaborate with 3<sup>rd</sup> parties?



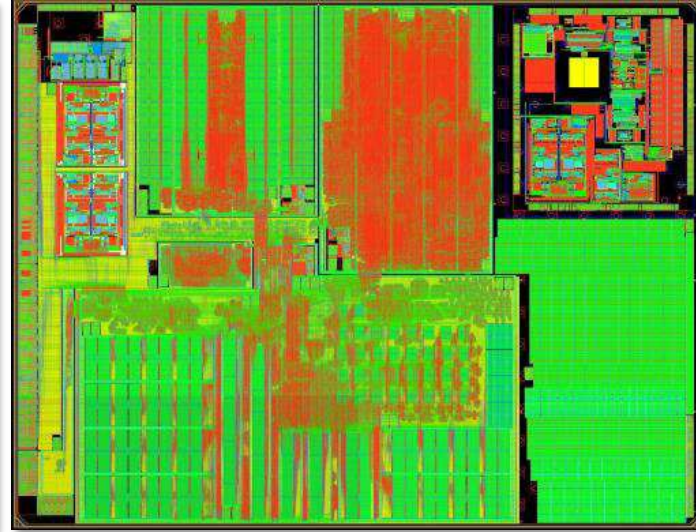
# The open model led to successful industry collaborations



**Arnold (GF22)**  
eFPGA with RISC-V core



**Vega (GF22)**  
IoT Processor with  
ML acceleration



The enabler of low-power Systems-on-Chip

**Marsellus (GF22)**  
IoT Processor with low power  
modes and event based  
computing



**Siracusa (TSMC16)**  
IoT Processor with  
NVM technology



# And many have used our work for their research



**Smallest RISC-V Device for Next-Generation Edge Computing**

**RISC-V Workshop**

**Our 1<sup>st</sup> gen. processor and 2.5D integrated device**

Processor SoC (p02)

SoC size: 300  $\mu\text{m}$  x 250  $\mu\text{m}$ , 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 Munetoh<sup>1</sup>, Chitra K Subramanian<sup>2</sup>, Arun Paidimarri<sup>2</sup>, Yasuteru Kohda<sup>1</sup>  
 IBM Research – Tokyo<sup>1</sup> & T.J. Watson Research Center<sup>2</sup>

**IBM**

*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 1GHz	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

**intel**

*VLSI Symposium 2022*

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

**Google**

Presenting the work of many people at Google

Jeff Dean  
 Google Research

**Article**  
**A graph placement methodology for fast chip design**

<https://doi.org/10.1038/s41586-021-03044-w>  
 Received: 3 November 2020  
 Accepted: 13 April 2021  
 Published online: 9 June 2021

Azalia Mirhoseini<sup>1,2\*</sup>, Anna Goldie<sup>1,2\*</sup>, Ebrahim Ghahbani<sup>1,2</sup>, Shen Wang<sup>1,2</sup>, You Arade Naz<sup>1</sup>, Jinwo Park<sup>1</sup>, Andy Tong<sup>1</sup>, Quoc V. Le<sup>1</sup>, James Laudon<sup>1</sup>, Richard

**ISSCC Keynote 2020 – Nature 2020**

**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

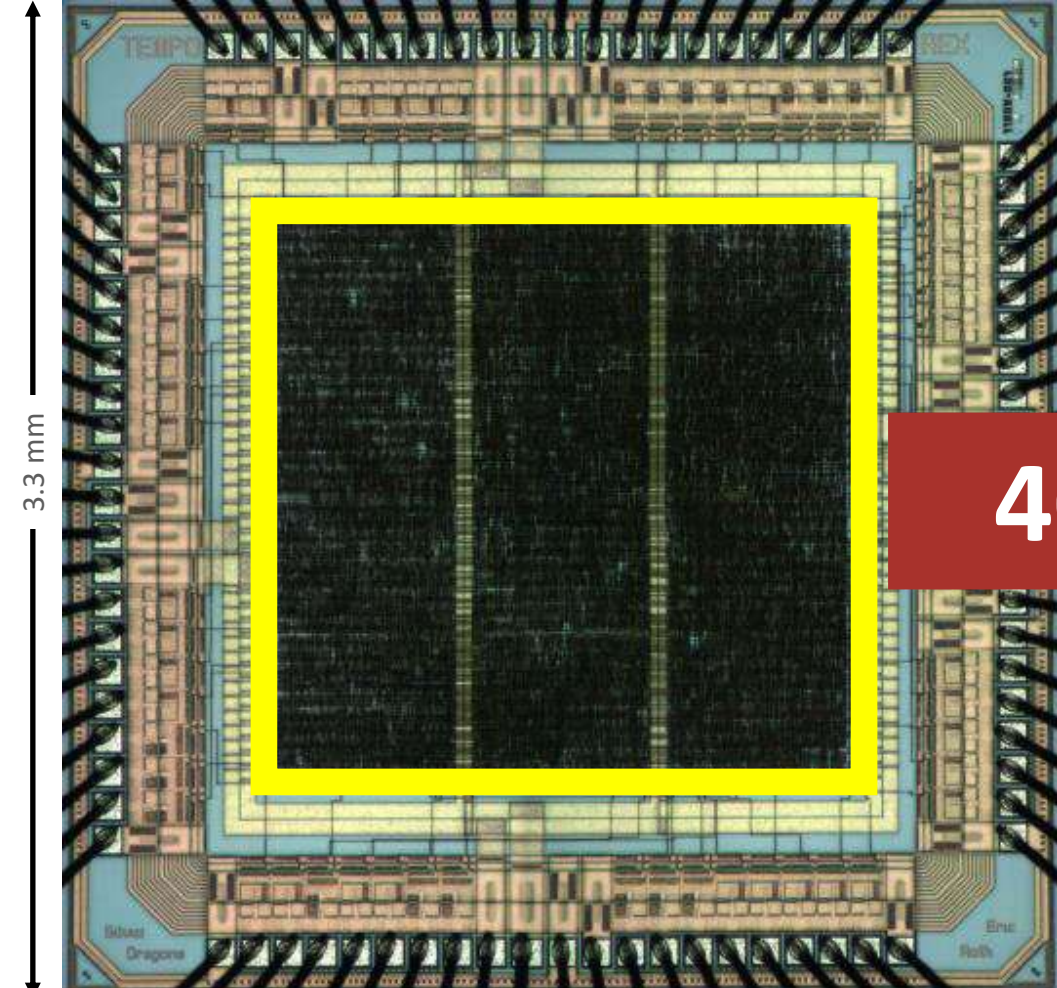
Haoxing Ren  
 haoxingr@nvidia.com  
 NVIDIA Corporation  
 Austin, TX, USA

**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**



# In the last 20 years IC Design has changed a lot



4000 x



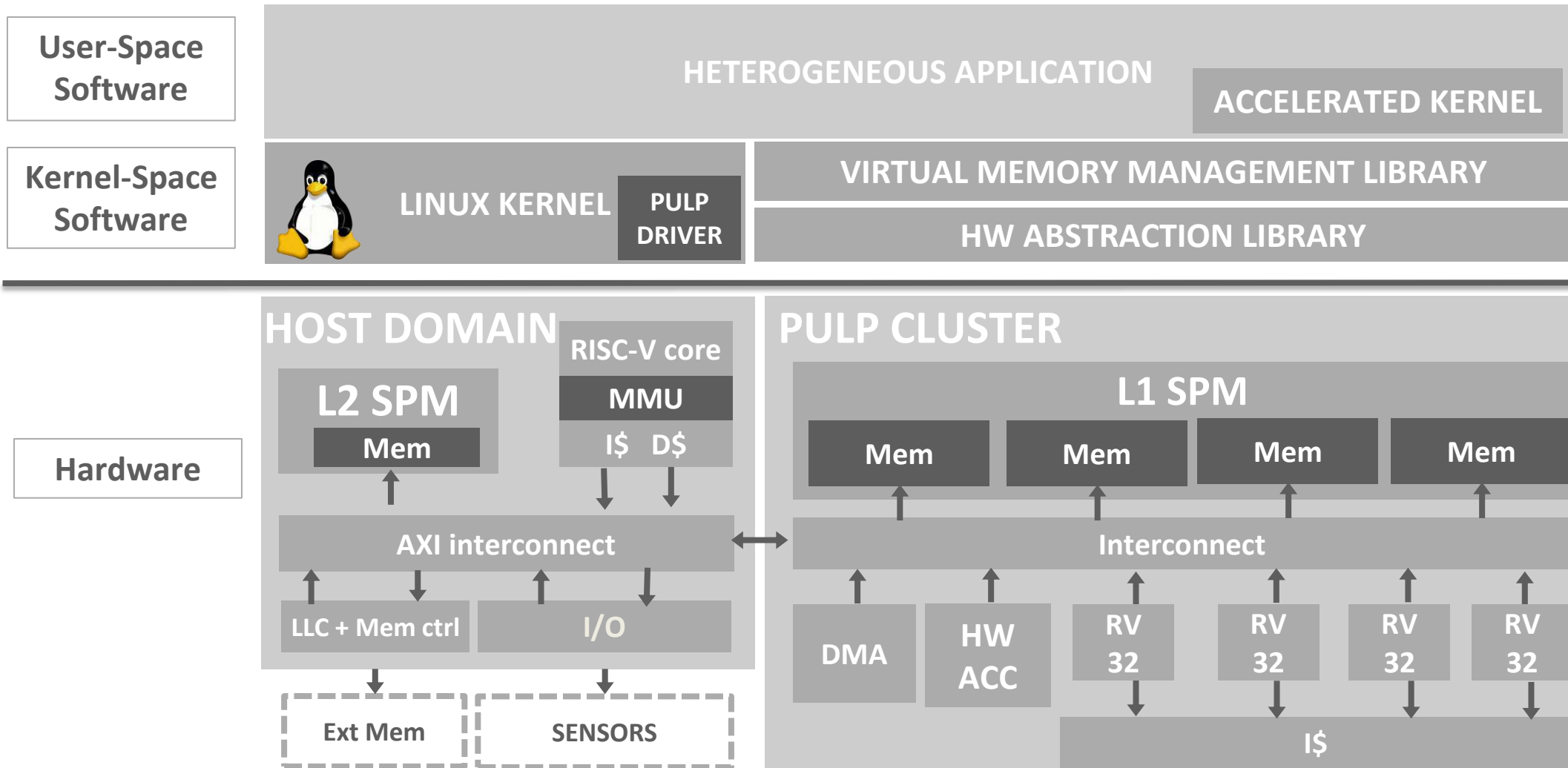
What used to be a complete chip is now a small part of a SoC !

80 MGE

the Landscape

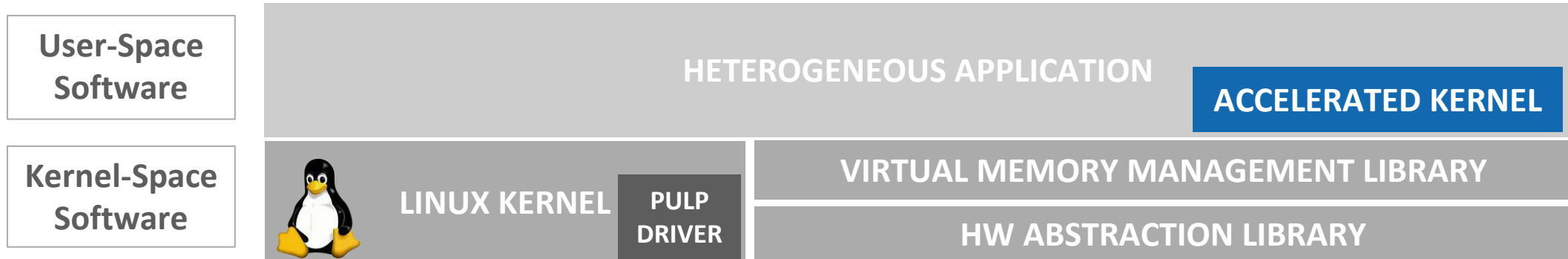
10

# There is so much that makes up a modern SoC

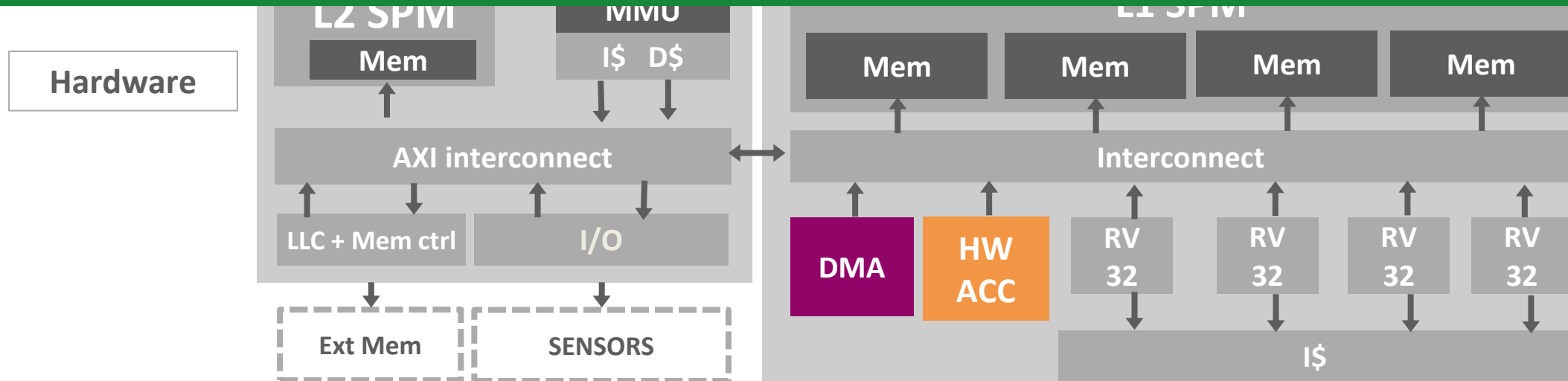




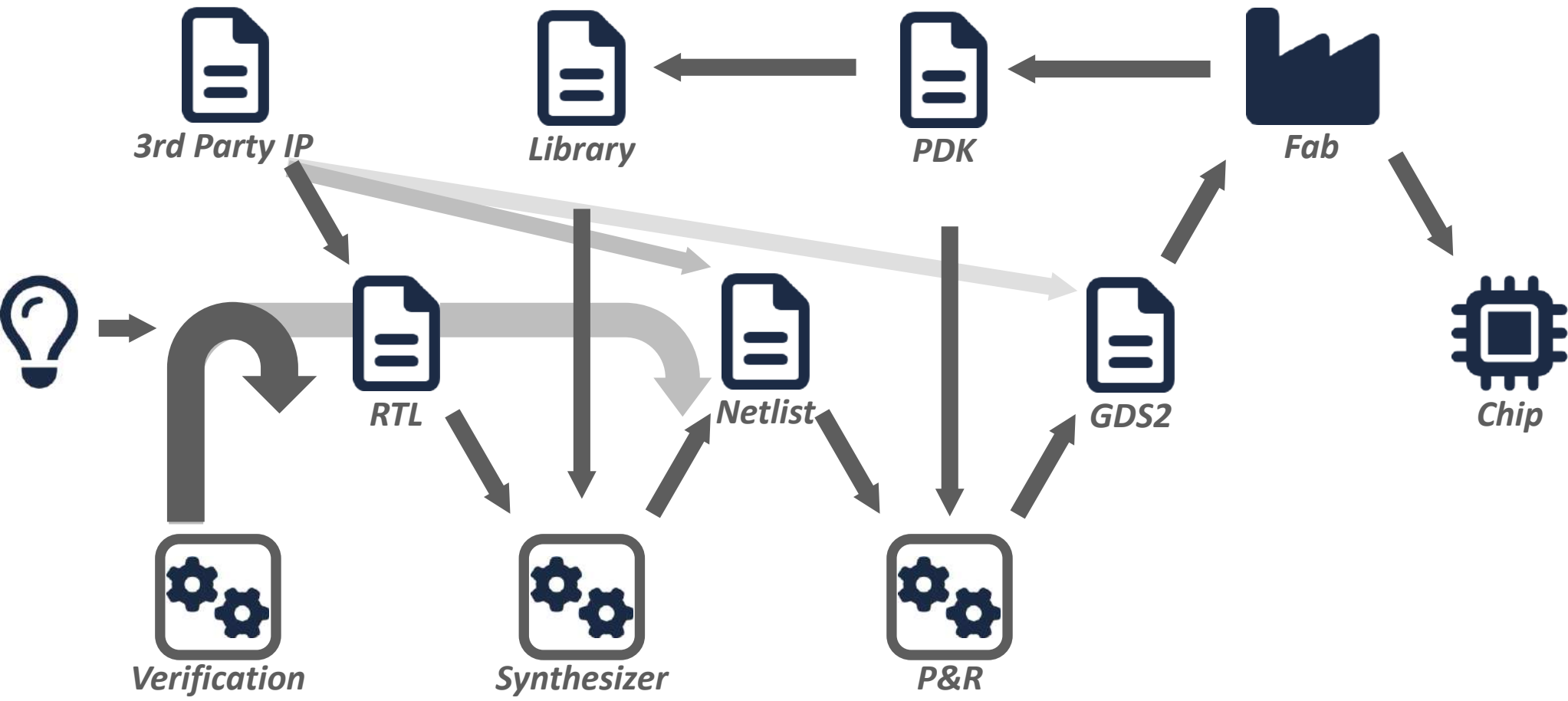
# In a typical design, innovation is only in a limited scope



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

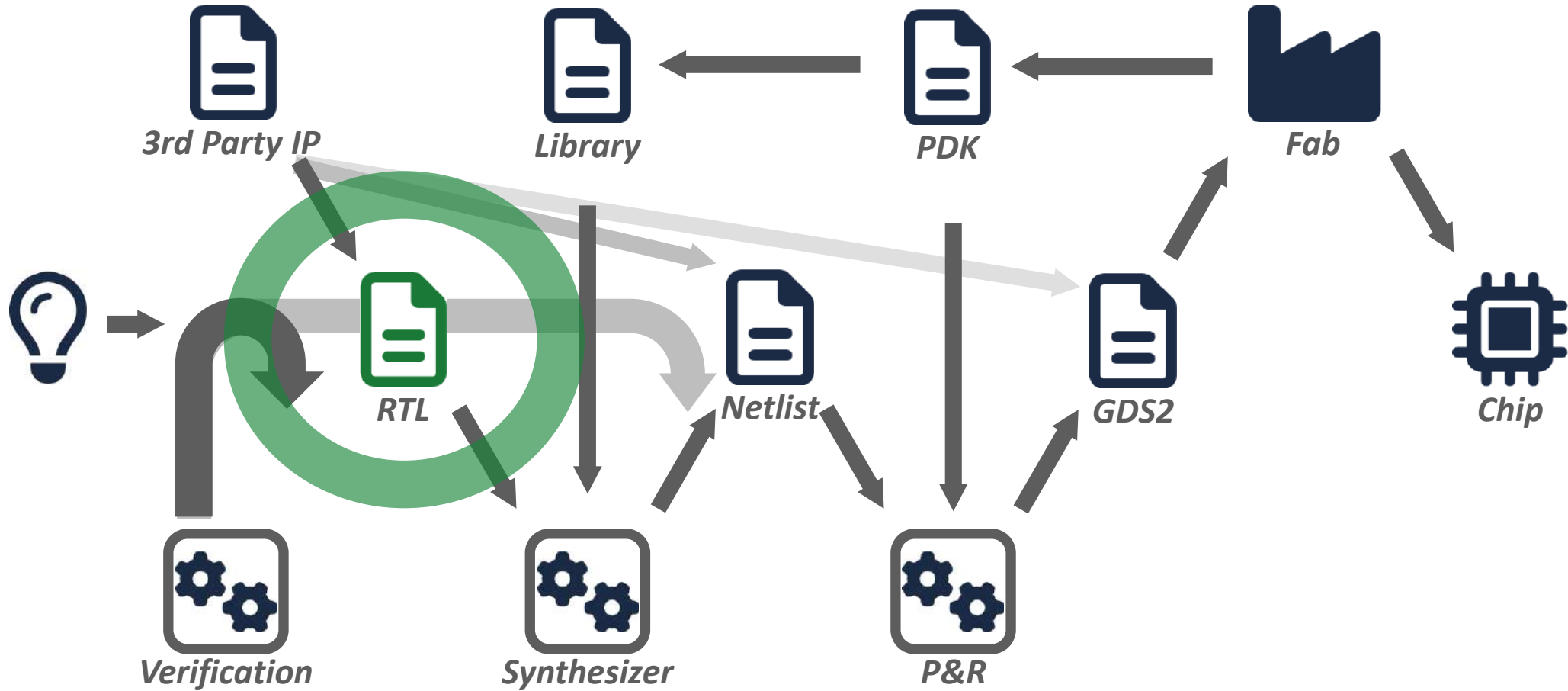


# Simplified IC design flow



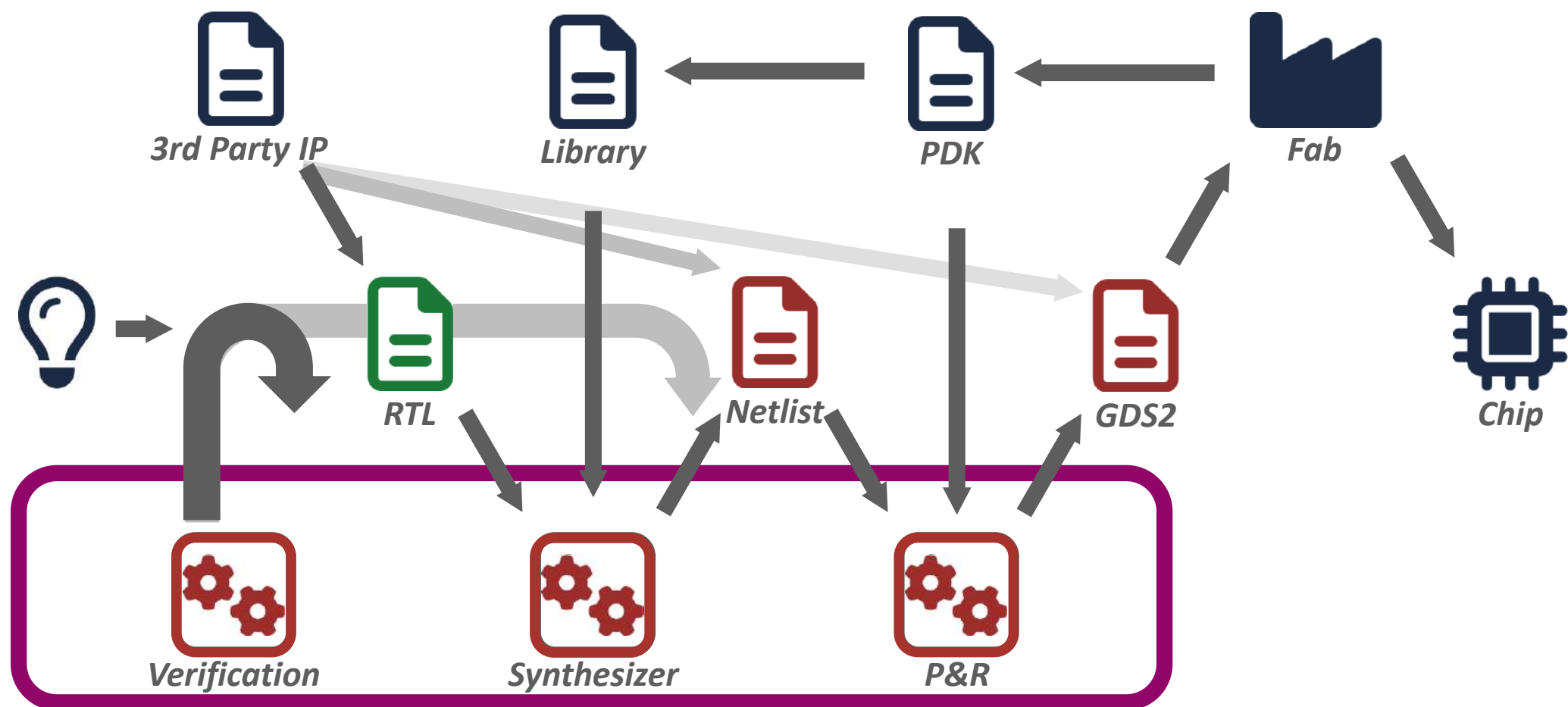
Icons taken from free icons from fontawesome.com

# Simplified IC design flow: at the moment only RTL



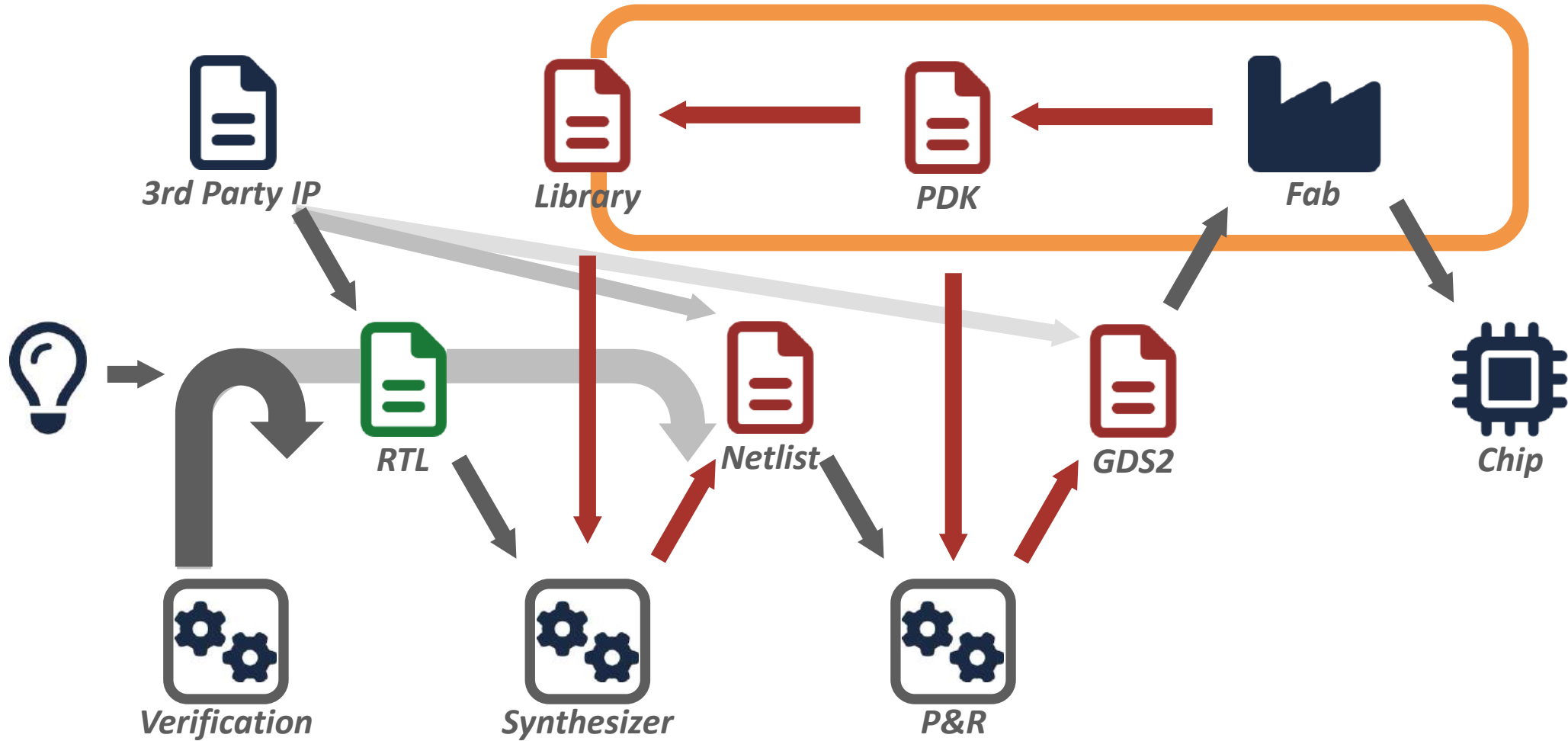


# Simplified IC design flow: proprietary tools



EDA vendors limit the output of their tools

# Simplified IC design flow: technology provider



Fabs do not make PDK information accessible

within the Open Source Hardware Landscape

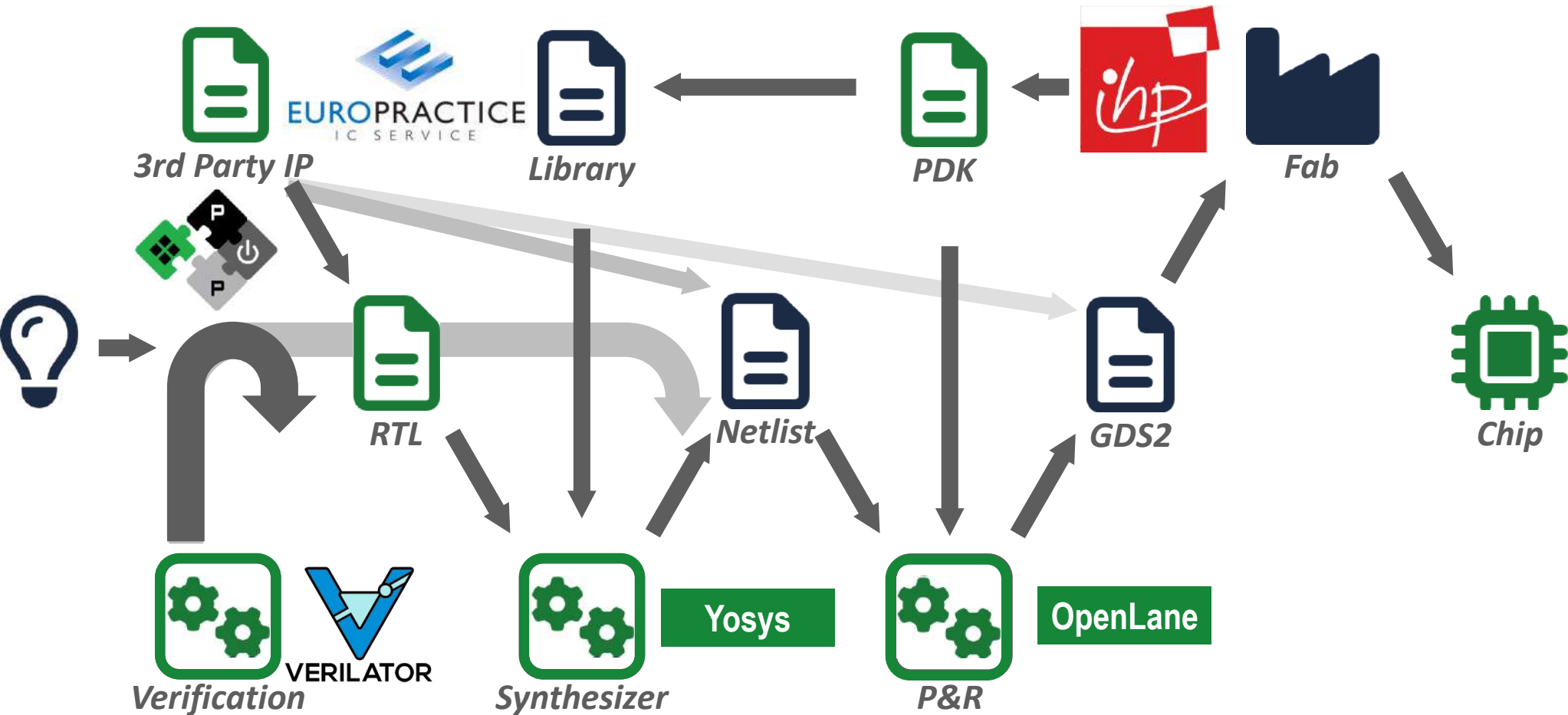
16



## The Landscape



# Simplified IC design flow: road to openness



We are getting there, first fully open chips are underway

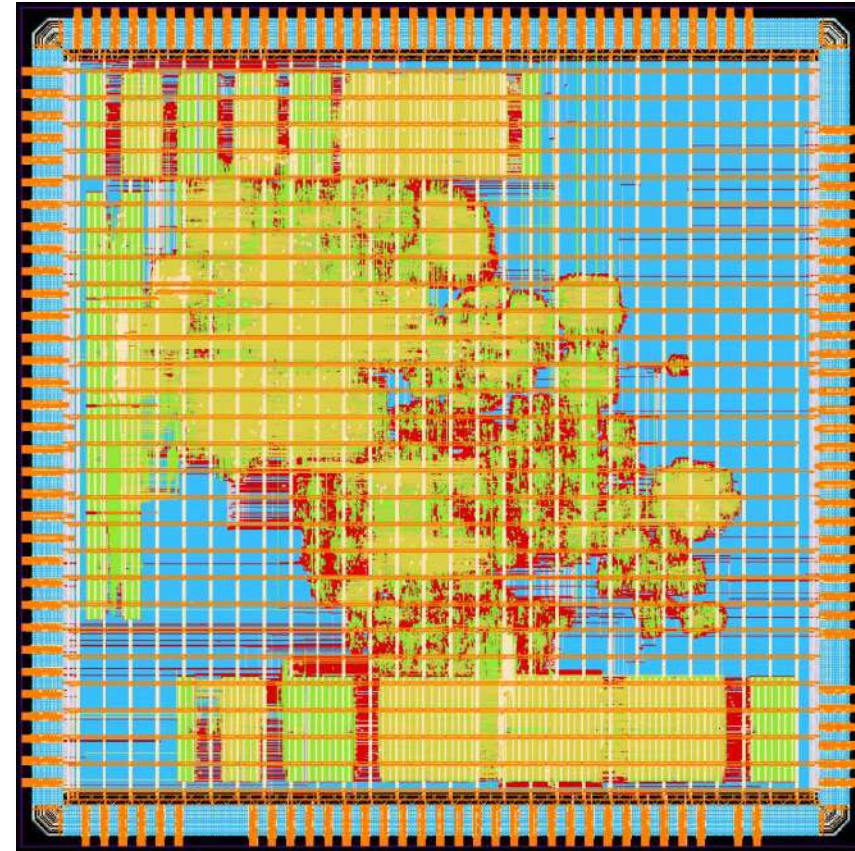
# Iguana, Linux capable fully-open RISC-V processor.



- **Designed using IHP 130nm**
  - European Open PDK
  - Tape-out July 2023
- **Mostly open flow**
  - Some parts still rely on proprietary tools/IP
  - Next version 4Q23 fully open



Hochschule RheinMain



# 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)
  - EU projects (Tristan/Isolde), Europractice





# Open source HW is making great strides



- **RTL level open source hardware is quite established**
  - (among others) PULP based solutions are widely used in Industry and Academia
- **More openness requires solutions for Technology, EDA, 3rd party IP providers**
  - Open PDKs (slowly on the way)
  - Open source EDA and/or more support from existing EDA vendors
  - Lack of IP to make it harder to design impactful systems
- **Efforts underway to support Open HW**
  - Non-profit organizations (OpenHW, RISC-V), EU projects (TRISTAN/ISOLDE), Supporting groups (Europractice)
- **IC Design is an expensive business, there will always be a financial aspect**
  - Open source HW is an enabler, lowers costs and access, but does not make IC Design free

**More work and support is needed to make open source HW viable**



<http://pulp-platform.org>



@pulp\_platform

