tinyML® Talks

Enabling Ultra-low Power Machine Learning at the Edge

“Unlocking TinyML: A Cost-Efficient Approach to Custom Microcontroller Design for AI Applications”

Mohamed Shalan – Head of Design and EDA, Efabless Corporation

October 31, 2023

www.tinyML.org

TINY ML TALKS webcast

Thank you, **tinyML Strategic Partners**, for committing to take tinyML to the next Level, together

TINY ML TALKS webcast

Executive Strategic Partners

Qualcomm
AI research

Advancing AI research to make efficient AI ubiquitous

Power efficiency

Model design, compression, quantization, algorithms, efficient hardware, software tool

Personalization

Continuous learning, contextual, always-on, privacy-preserved, distributed learning

Efficient learning

Robust learning through minimal data, unsupervised learning, on-device learning

A platform to scale AI across the industry

Perception

Object detection, speech recognition, contextual fusion

Reasoning

Scene understanding, language understanding, behavior prediction

Action

Reinforcement learning for decision making

Edge cloud

Cloud

IoT/IIoT

Automotive

Mobile

Accelerate Your Edge Compute

SYNTIANT

Making Edge AI A Reality

www.syntiant.com

TINY ML TALKS webcast

Platinum Strategic Partner

AITRIOS

DEPLOY VISION AI AT THE EDGE AT SCALE

SONY

TINY ML TALKS webcast

Gold Strategic Partners


Build the
Future of tinyML
on arm

TINY ML TALKS webcast

EDGE IMPULSE

The Leading Development Platform for Edge ML

edgeimpulse.com

Driving decarbonization and digitalization. Together.

Infineon serving all target markets as
Leader in Power Systems and IoT

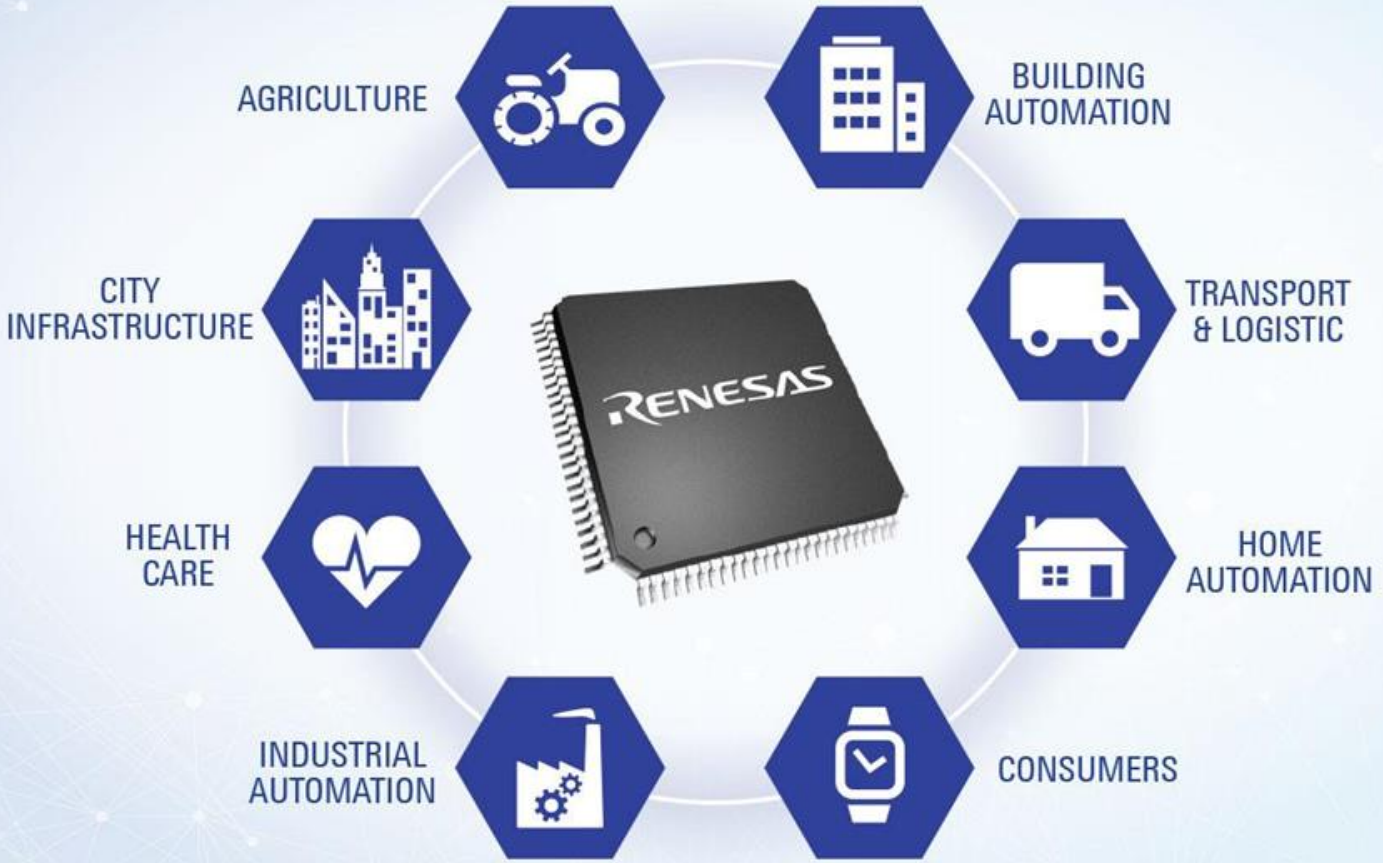
www.infineon.com

innatera

NEUROMORPHIC INTELLIGENCE FOR THE SENSOR-EDGE

www.innatera.com

Renesas is enabling the next generation of AI-powered solutions that will revolutionize every industry sector.



renesas.com

STMicroelectronics provides extensive solutions to make tiny Machine Learning easy

life.augmented

www.st.com/ai

synaptics®

ENGINEERING EXCEPTIONAL EXPERIENCES

We engineer exceptional experiences for consumers in the home, at work, in the car, or on the go.

www.synaptics.com

TINY ML TALKS webcast

Silver Strategic Partners

Join Growing tinyML Communities:

17.4k members in
49 Groups in 41 Countries

tinyML - Enabling ultra-low Power ML at the Edge

https://www.meetup.com/tinyML-Enabling-ultra-low-Power-ML-at-the-Edge/

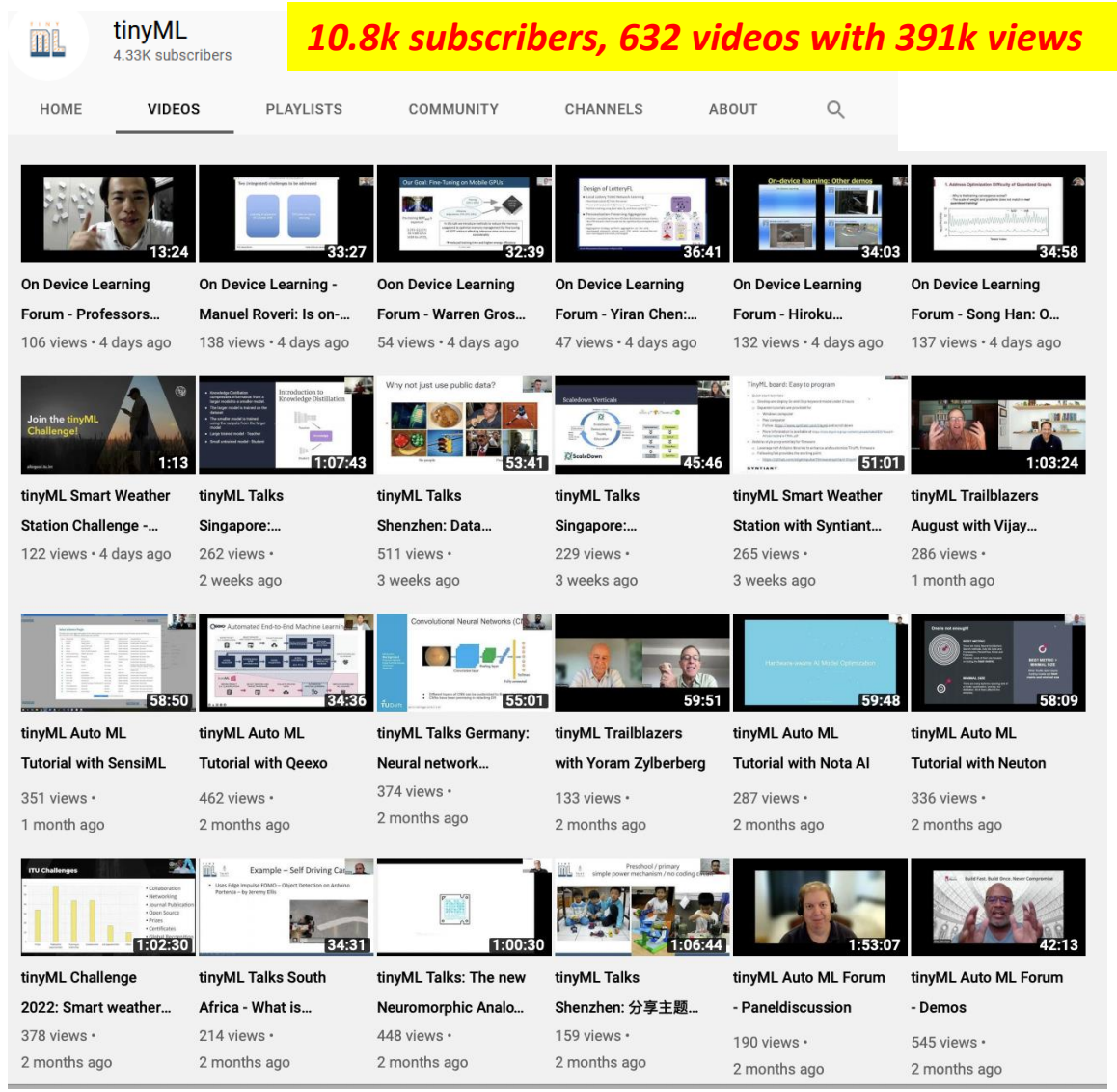
4k members
&
13k followers

The tinyML Community

https://www.linkedin.com/groups/13694488/

TINY ML TALKS webcast

YouTube



Subscribe to tinyML YouTube Channel for updates and notifications

(including this video)

www.youtube.com/tinyML

TINY ML FOUNDATION

Register now
https://www.tinyml.org/event/asia-2023/

WEVOLVER

https://www.wevolver.com/article/2023-edge-ai-technology-report

TINY ML FOUNDATION

2023 Edge AI Technology Report

The guide to understanding the state of the art in hardware & software in Edge AI.

TINY ML TALKS webcast

Reminders

Slides & Videos will be posted tomorrow

PDF

tinyml.org/forums

youtube.com/tinyml

Please use the Q&A window for your questions

TINY ML TALKS webcast

Mohamed Shalan

Mohamed Shalan is the Head of Design and EDA at Efabless Corporation. Before his role with Efabless, Mohamed held various engineering positions at several organizations, including Mentor Graphics, Motorola SPS, and Mindspeed. Mohamed holds a Ph.D. degree in computer engineering from the Georgia Institute of Technology. His research interests are multifaceted, encompassing open-source EDA, hardware, and software codesign for embedded systems, with a strong emphasis on low-power computing. His contributions to these areas have resulted in over 40 peer-reviewed publications and three US patents.

TINY ML TALKS webcast

Unlocking TinyML

A Cost-Efficient Approach to Custom Microcontroller Design for AI Applications

TINY ML TALKS webcast

Microcontrollers

- It is estimated that around 300 billion microcontrollers are in the world today
- Around 38 billion units to be shipped in 2023 alone (estimated)
- Strong microcontroller growth in recent years was driven by IoT applications
- The current growth will be driven by TinyML

TINY ML TALKS webcast

Are microcontollers ready for TinyML wave?

ISA Extensions

Embedded FPGA Fabric

Fixed/Prog. Accelerators

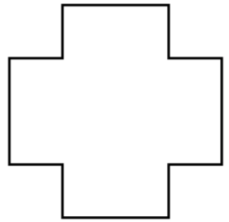
But, microcontrollers are sub-optimal!

Desired functionality

```
total = 0
for i = 1 to N
loop
    total += M[i]
end loop
```

General-purpose processor

Application-specific processor



Single-purpose processor

microcontrollers

Credit: Frank Vahid

TINY ML TALKS webcast

Why Custom SoCs?

- Optimized Performance
- Energy Efficiency
- High Integration ➔ Lower cost
- Protecting your IP
- Create the differentiator
- Cost Efficiency in the Long Run
- Controlling your supply chain
- …

TINY ML TALKS webcast

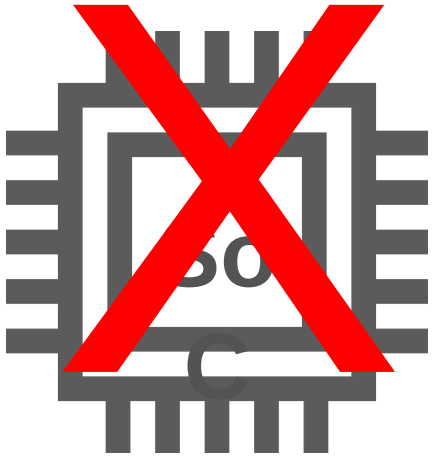
Why not Custom SoCs?

- Cost, high NRE
- Foundry PDK Access
- IP Access
- Complexity
- High Risk
- …

TINY ML TALKS webcast

1. Market Access — $$
2. Expertise — $$+Time
3. IP — $$+NDA
4. EDA Tools — $$+NDA
5. PDK — NDA
6. AFFORDABLE MFG/ASSY/TEST — $$

At least these **6** things must exist to create useful chips

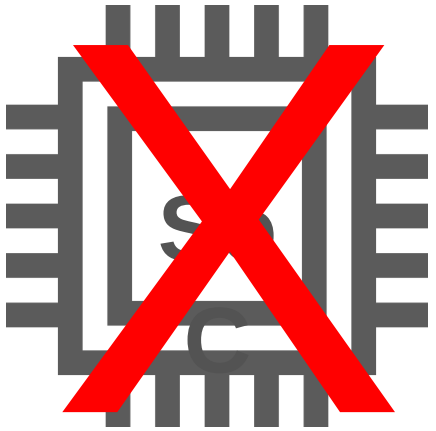


YOU **CAN'T** BUILD A USEFUL **CHIP**

TINY ML TALKS webcast

1. Market Access — $$
2. Expertise — $$+Time
3. IP — $$+NDA
4. EDA Tools — $$+NDA
5. PDK — NDA
6. AFFORDABLE MFG/ASSY/TEST — $$

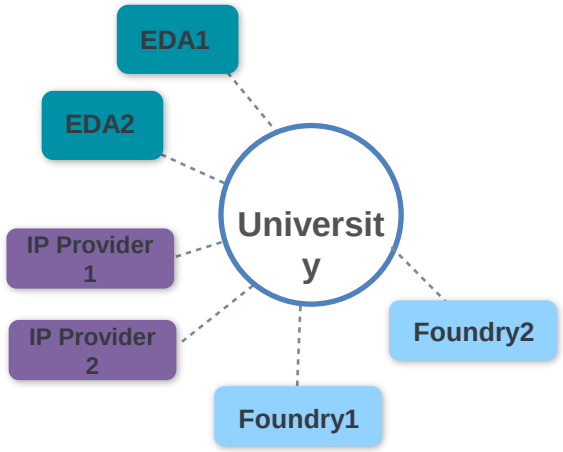
Partial solutions **don't work**

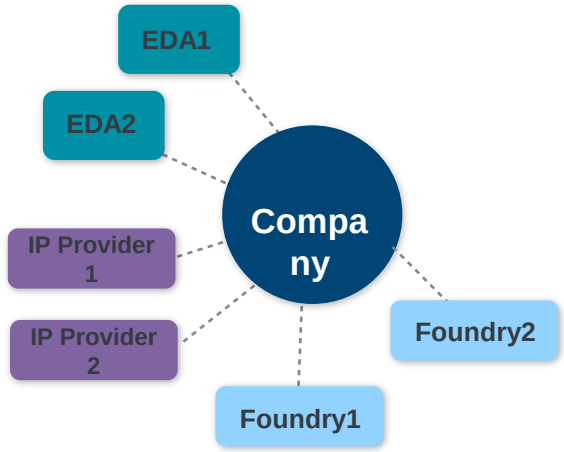


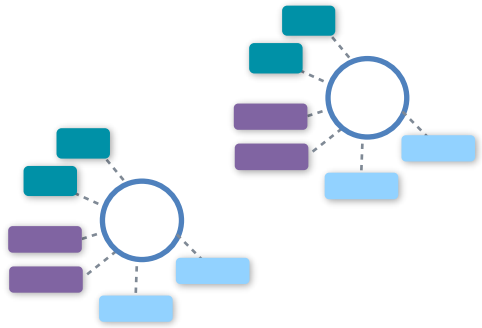
STILL **CAN'T** BUILD A USEFUL **CHIP**

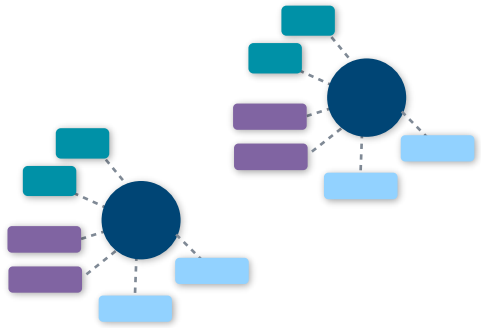
TINY ML TALKS webcast

TODAY'S APPROACH TO CUSTOM SILICON









Every design entity needs to establish **independent** business and contractual (including NDA's) relationships with **multiple** IP providers, EDA vendors, and foundries

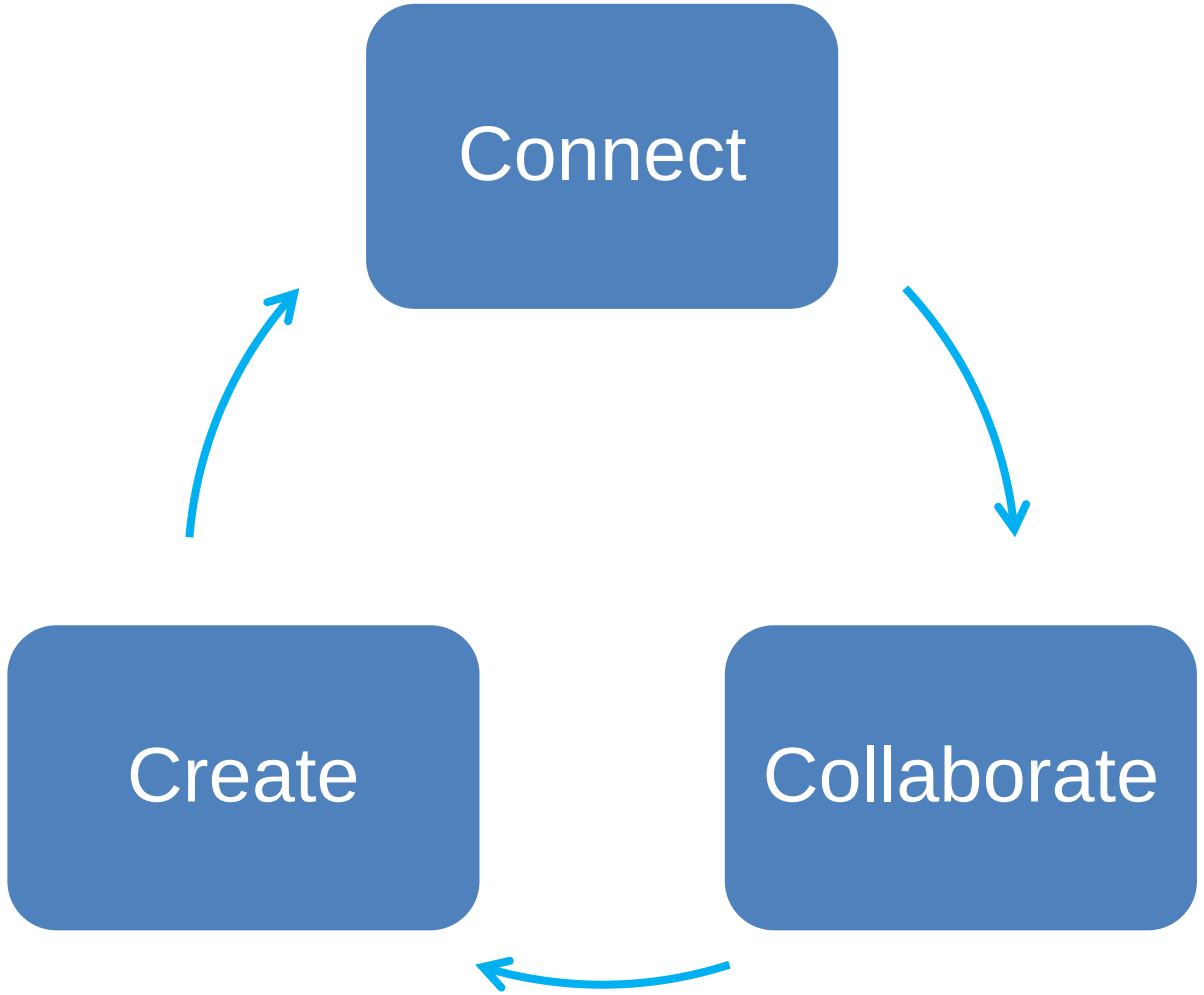
TINY ML TALKS webcast

efabless.com

Simplifying Chip Creation and Opening it to Everyone

TINY ML TALKS webcast

The largest community of chip creators



Users and experts co-design custom chips for smart products

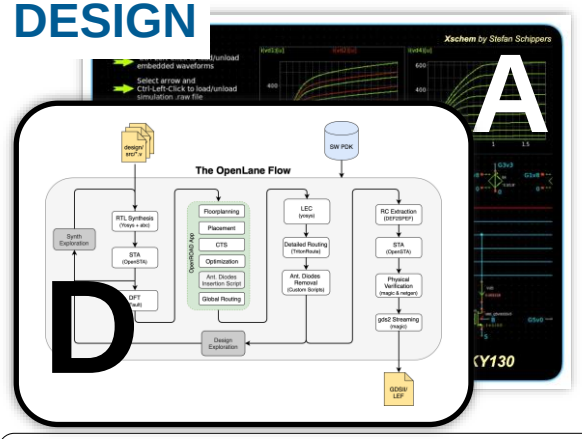
Design, manufacture and test your chips for under $10K.

TINY ML TALKS webcast

Complete & Path from Design to Validation

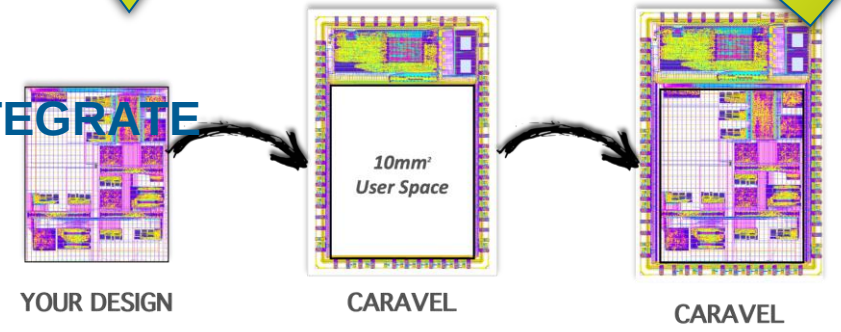
chipIgnite Rapid IC Creation

1 DESIGN



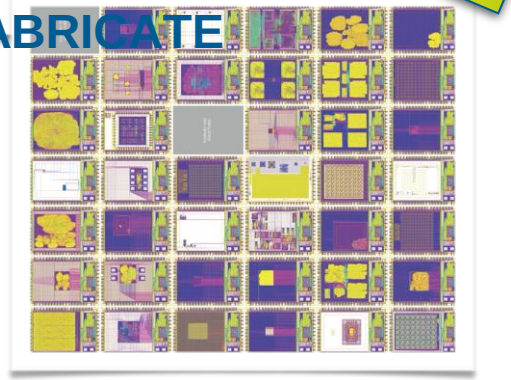
OS or Proprietary EDA Digital & Analog

2 INTEGRATE



3 SUBMIT

4 FABRICATE



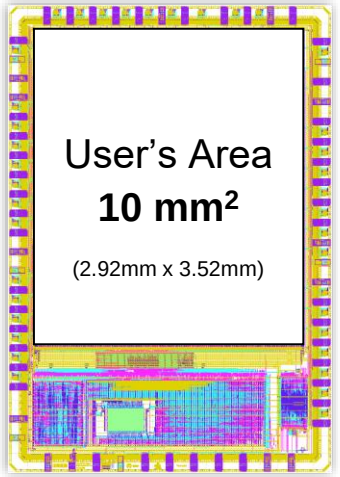
5 TEST/Validate

TINY ML TALKS webcast

Start from the 18th floor to reach the 20th

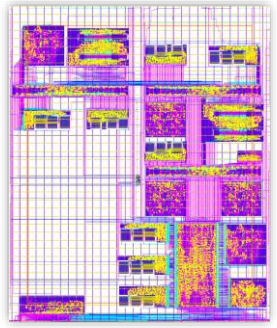
*Build on **existing foundational** work by others*

*You only need to know your design … or **your code***

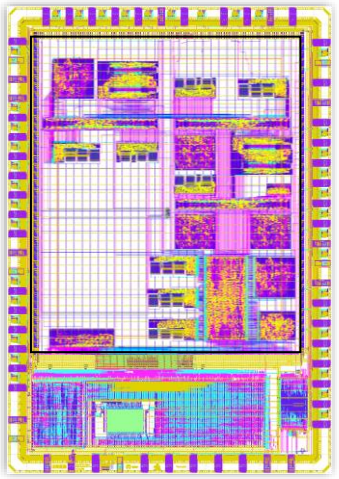


CARAVEL

+



Your Design

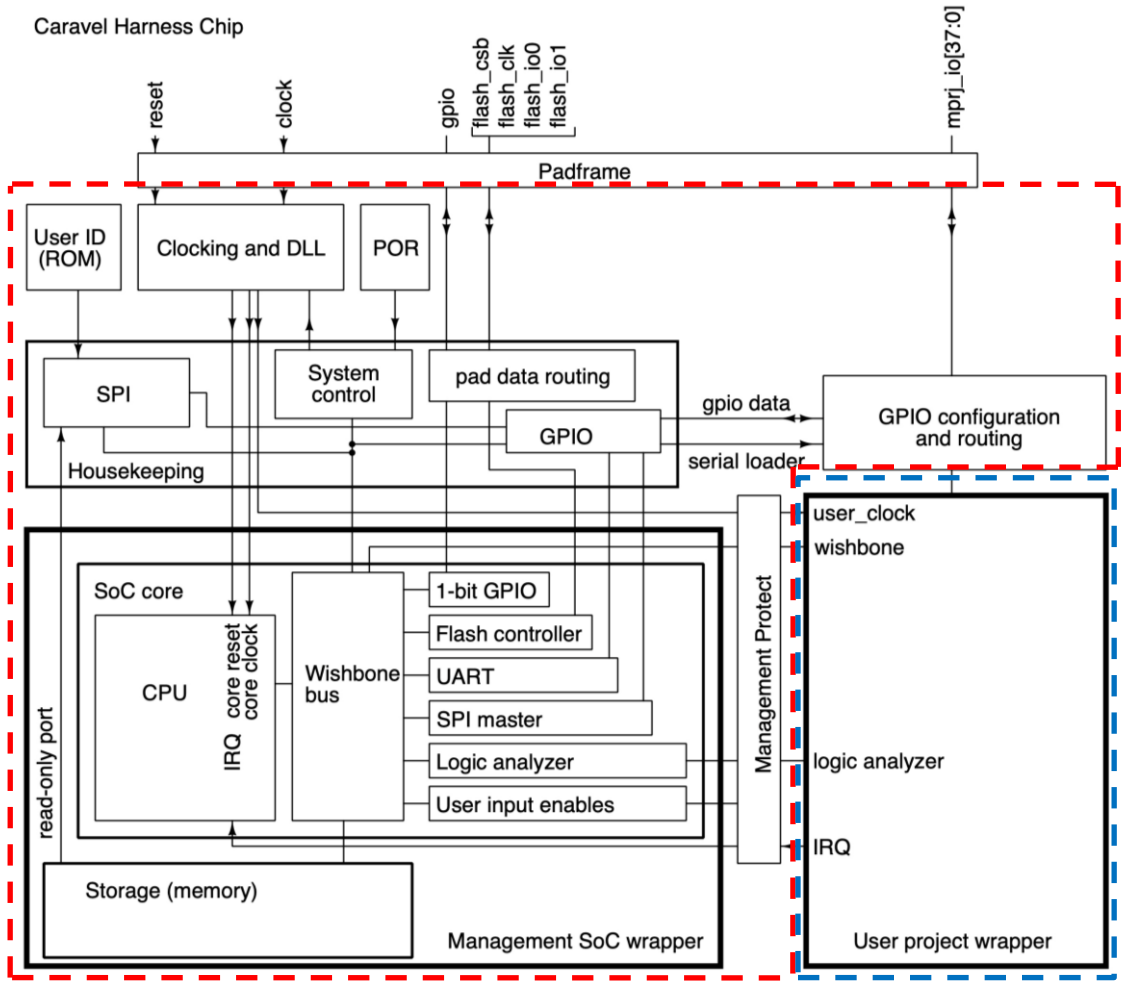


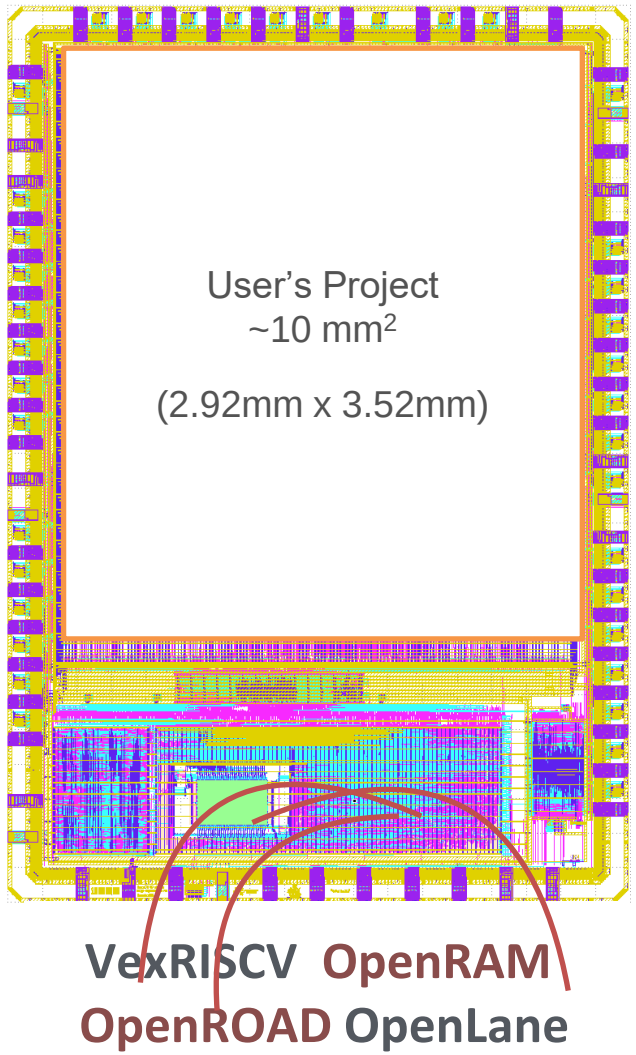
CARAVEL + Your Design

Plug & Play Dev Board

TINY ML TALKS webcast

Caravel SoC Platform





Uniform Open Source File Structure to ease reuse & modifications

- .github/workflows
- .travisCI
- def
- docs
- gds
- irsim
- lef
- lvs
- macros
- mag
- maglef
- ngspice
- oas
- openlane
- qflow
- scripts
- signoff
- spef
- spi/lvs
- utils
- verilog
- xyce

https://github.com/efabless/caravel

TINY ML TALKS webcast

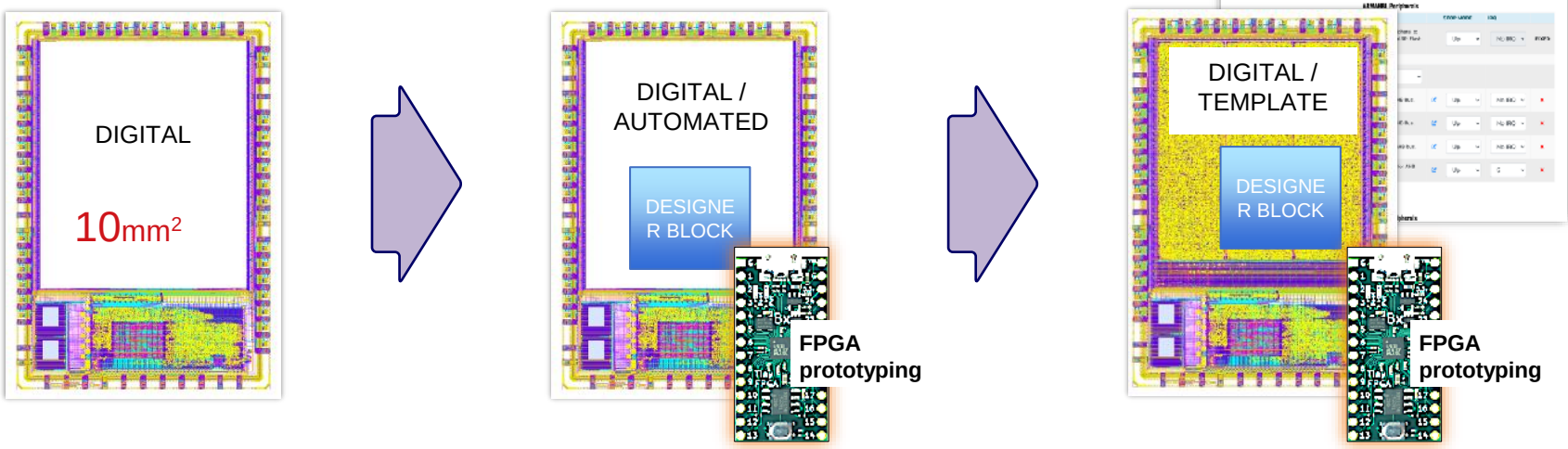
One Size Fits All

LOW COMPLEXITY

IP Development

Digital & low frequency analog

Enabling larger designer base



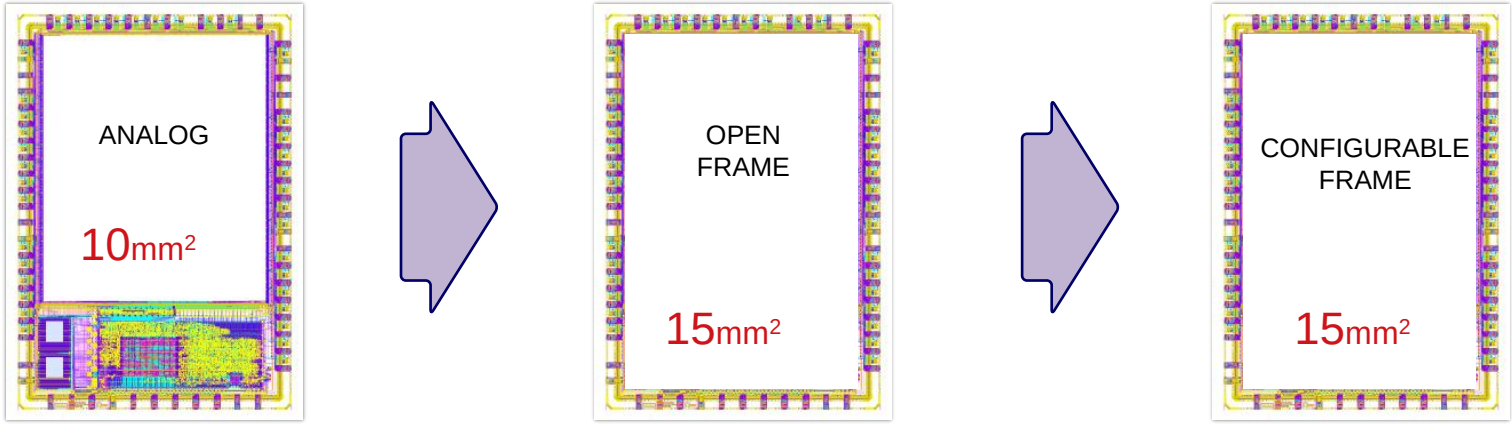
HIGH COMPLEXITY

IP Development

Complete Custom ASIC

Analog & Digital

Expert designer base

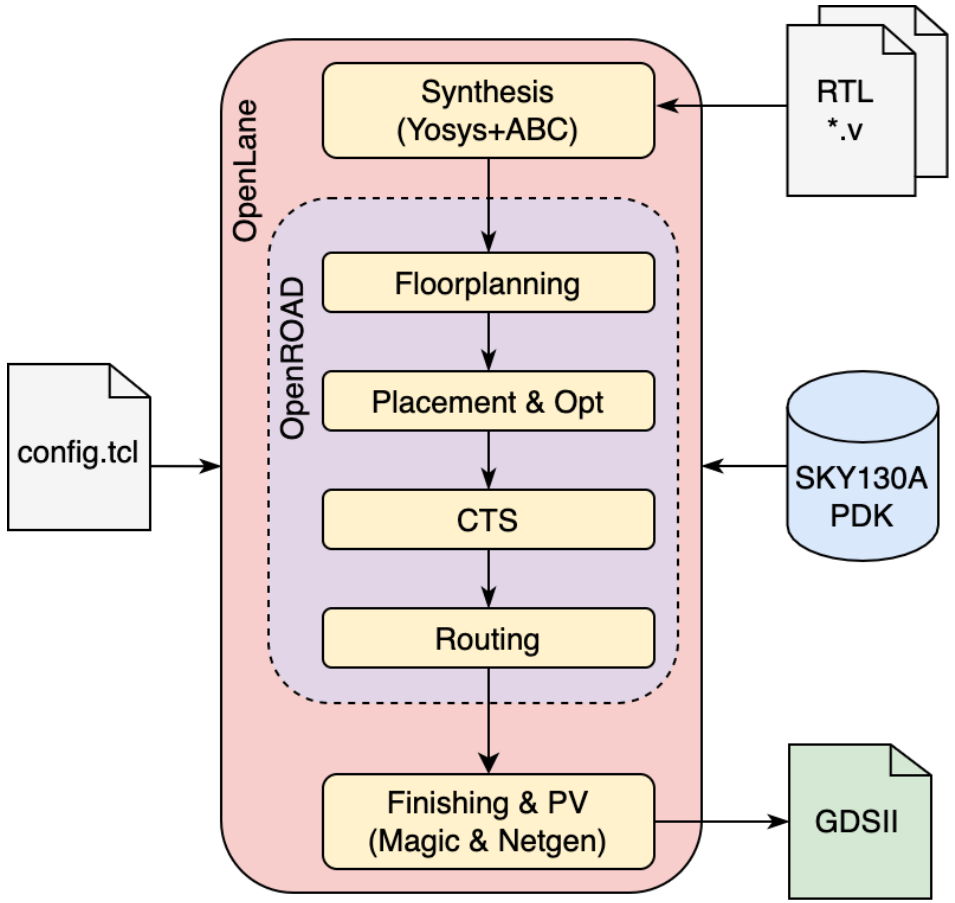


TINY ML TALKS webcast

OpenLane RTL2GDS Flow

Automate code-to-chip like a ***GNU software compiler -*** *with trade-offs in area and performance.*

It opens the door for software developers to generate hardware That's at least a 1000x more potential designers!



OpenLane

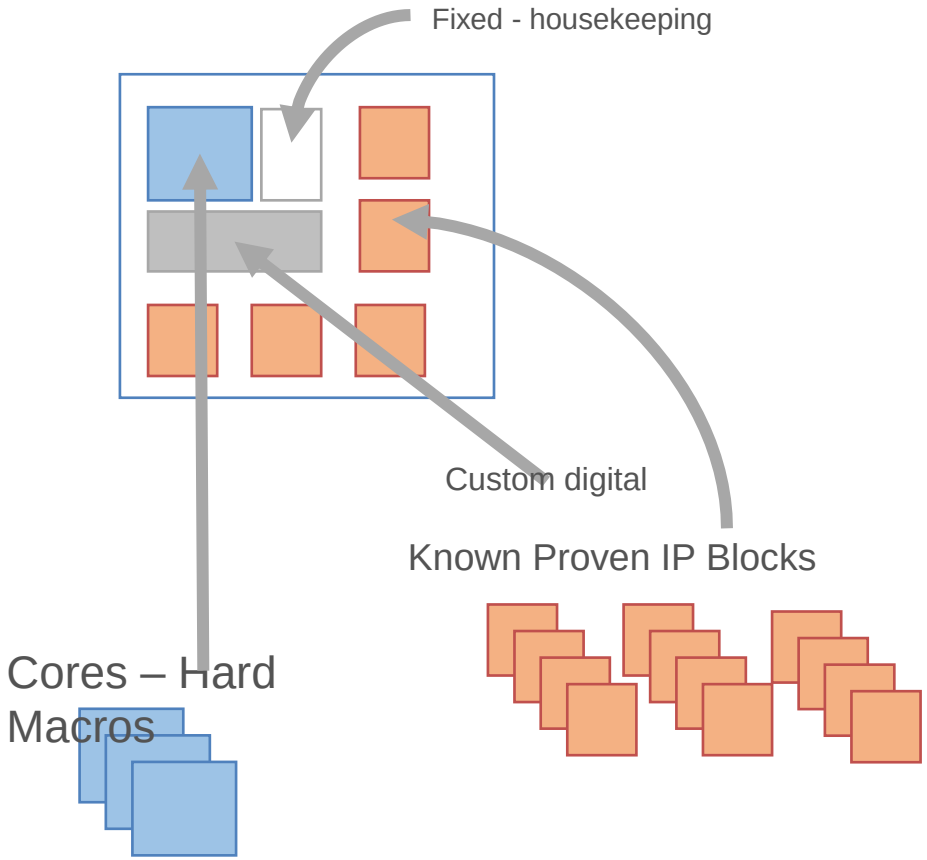
TINY ML TALKS webcast

A library of IP blocks

Open Source IP

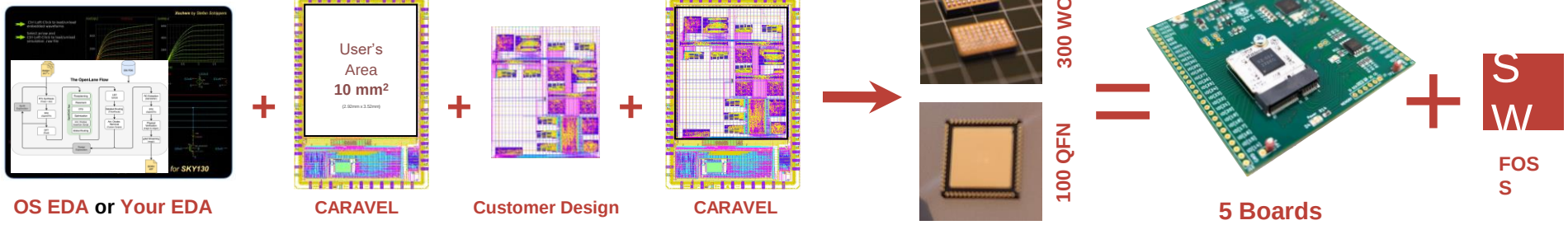
Obfuscated Proprietary IP

- Silicon proven / verified functions
- Highly leveraged known proven IP blocks
- Scope of customization is constrained
- Standard IP packaging to enable reuse



chipIgnite

Rapid IC Creation



OS EDA or Your EDA + CARAVEL + Customer Design + CARAVEL → = + SW

- Design using SkyWater SKY130 process + std cells + IO cells
- Bring your Own EDA, use Open Source Flows or Use EF Cloud Synopsys EDA
- Start from the Caravel Harness - book your slot for **$200**
- You *choose your licensing terms* for your design - by *default it's proprietary*
- Design Area (10mm² or 15mm²) is not just silicon area -

  It is a framed house with plumbing and electricity - just bring your design!
- Cycle Time 95 days +/- depending on the package choice

skywater TECHNOLOGY FOUNDRY

Rapid Prototyping	**$9,750**	**100 QFN**	**parts + dev boards**
Low Volume Production	**$20,000**	**1000 QFN**	**parts + dev boards**

TINY ML TALKS webcast

10,000 community members

1,000 designs

500 chips

2 years!

TINY ML TALKS webcast

Community Developed Designs

Selected Open Source IP Blocks

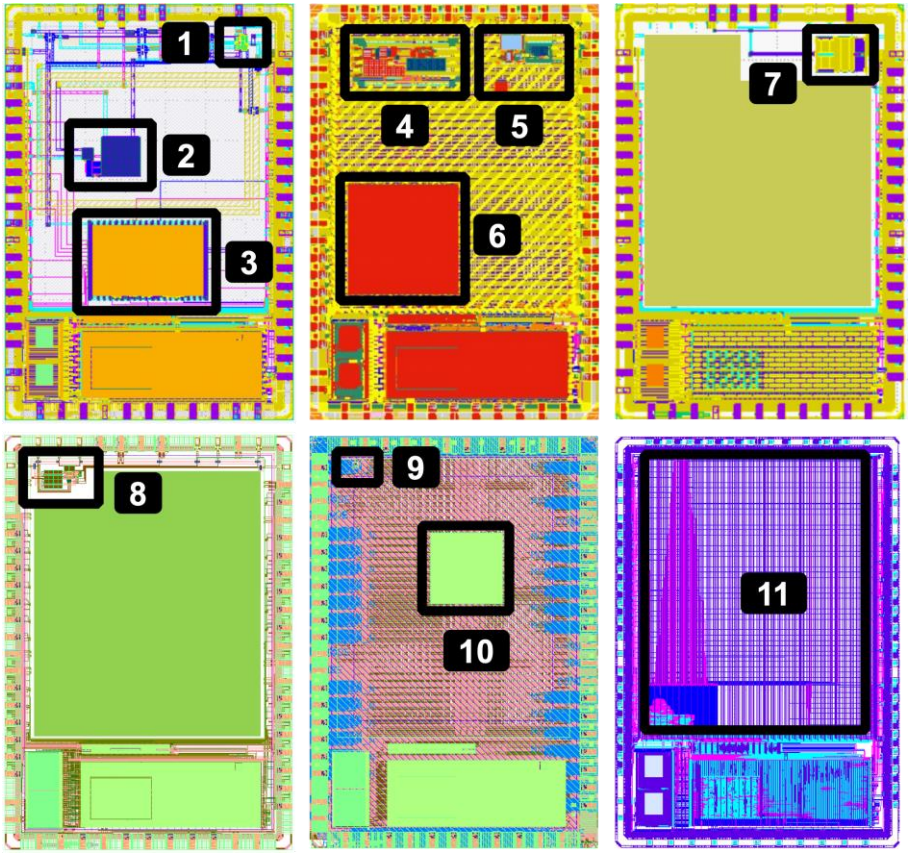
- A pseudo random number generator for critical real-time processors.
- A General Purpose Bandgap generating constant voltage at output, independent of Temp and Supply variations.
- Convolutional Neural Network Accelerator on a wishbone slave for Raven Core in Caravel SoC.
- Maverick 603 Radio
- Chaos automaton
- Arduino pin compatible Single RISCV 32 Bit core Project
- HSV to RGB color conversion accelerator
- CMOS Bandgap
- USB for RISC-V microcontroller
- Analog Spiking Circuit and 10-bit DAC
- Auditory perception acoustic front end
- GPS Baseband
- 1V8 LDO
- NAND Flash

Select Open Source SoCs

- Sleep Apnea Detection System
- Electro Mechanical Water Quality Monitoring
- Influenza detector
- Image Detection at Edge of Neural Networks
- Satellite Radio
- Epileptic Seizure Detection
- LED Lighting for Bangladesh Solar Home System
- Nanopore DNA Sequencing
- Smart Garden
- Industrial Motor Controller
- Reconfigurable Flight Controller SoC for UAVs
- TinyML Image Detection at Edge with DNN
- Donkey Kong Bongo PS2 Keyboard

2021 IEEE SSCS Chipathon

IEEE SOLID-STATE CIRCUITS SOCIETY™ IC Innovation



	Function	Team	Chip URL
1	5G bidirectional amplifier	Pakistan3 (FAST National University)	https://efabless.com/projects/560
2	Wireless power transfer unit	Pakistan2 (FAST National University)	
3	Variable precision fused multiply-add unit	Pakistan1 (FAST National University)	
4	Oscillator-based LVDT readout	India2 (Anna University)	https://efabless.com/projects/474
5	Temperature sensor	India1 (Anna University)	
6	GPS baseband engine	India3 (Anna University)	
7	Ultra-low-power analog front-end for bio signals	Brazil2 (U. Federal de Santa Catarina)	https://efabless.com/projects/476
8	TIA for quantum photonics interface	USA4 (University of Virginia)	https://efabless.com/projects/470
9	Bandgap reference	Egypt (Cairo University)	https://efabless.com/projects/473
10	Neural network for sleep apnea detection	USA2 (University of Missouri)	
11	SONAR processing unit	Chile (University of the Bio-Bio)	https://efabless.com/projects/540

Link to Article [SSCS Magazine Article - NOV 2021]

Paid runs via Efabless chipIgnite (130 nm SkyWater)

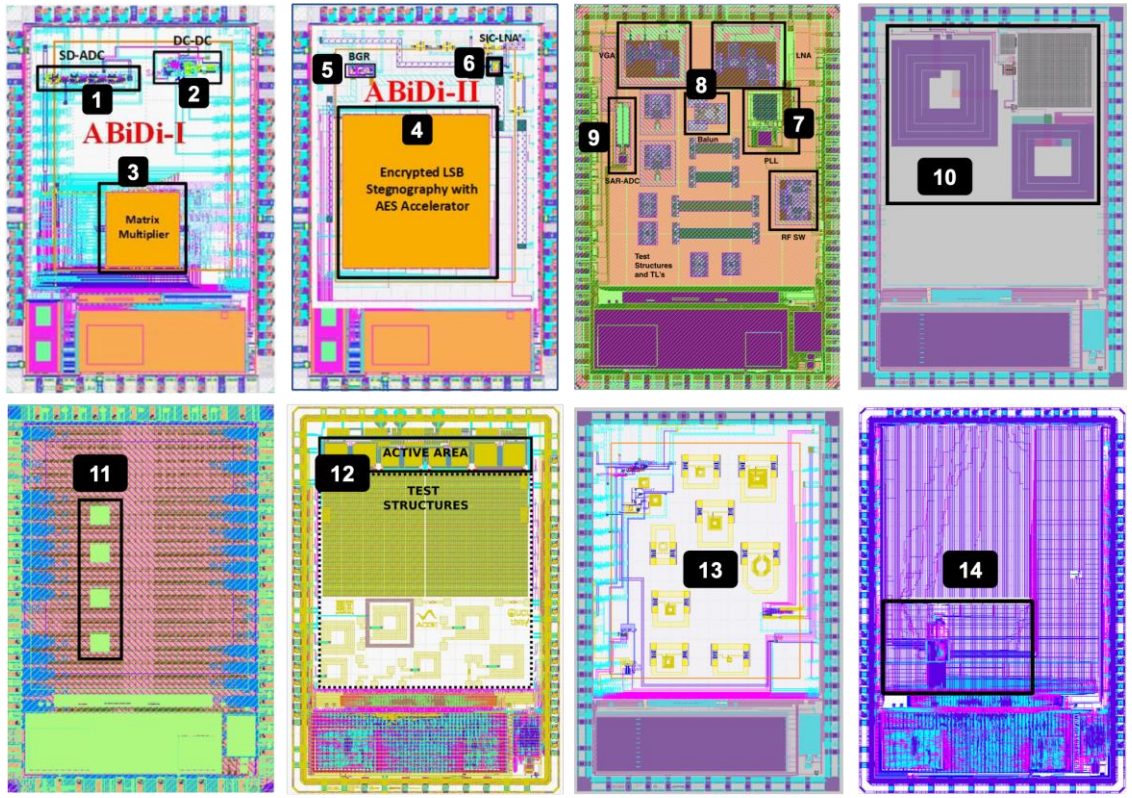
All designs are open source

Made possible with Efabless' **chipIgnite**

efabless.com

TINY ML TALKS webcast

2022 IEEE SSCS Chipathon



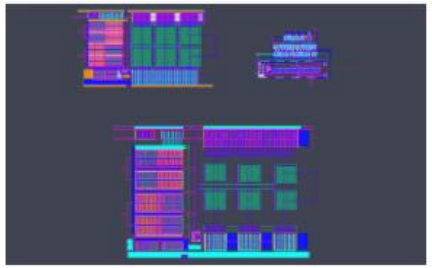
Link to Article [SSCS Magazine Article - JAN 2022]

	Function	Team	Chip URL
1	Spatial Sigma-Delta ADC	Pakistan1 (FAST National University)	https://platform.efabless.com/projects/1486
2	On-Chip DCDC Converter with Fast Transient Response	Pakistan4 (FAST National University)	
3	Matrix Multiplier for AI at the Edge	Pakistan7 (FAST National University)	
4	Encrypted LSB Steganography with AES Accelerator	Pakistan2 (FAST National University)	https://platform.efabless.com/projects/1443
5	CMOS Bandgap Reference	Pakistan3 (FAST National University)	
6	Self-Interference Cancellation LNA	Pakistan4 (FAST National University)	
7	Sub-Sampling PLL for SerDes Applications	Austria (Johannes Kepler Univ., Linz)	https://platform.efabless.com/projects/1431
8	60 GHz Demonstrator Chip	Brazil (University of São Paulo)	
9	Low-Power 10-bit SAR ADC	USA1 (University of Alabama & MIT Lincoln Lab)	
10	Boost Converter for Battery-Powered IoT Applications	Greece (Aristotle University of Thessaloniki)	https://platform.efabless.com/projects/1457
11	Radiation-Hardened ALU	USA2 (North Carolina A&T State University)	https://platform.efabless.com/projects/1593
12	DC-DC Buck Converter for CubeSat	Chile[1]/Argentina[2]/Uruguay[3] [1]Universidad Técnica Fed. Santa María [2]Universidad Nacional del Sur & Instituto Nacional de Tecnología Industrial [3]Universidad Católica	https://platform.efabless.com/projects/1427
13	Electrochemical Water Quality Monitoring	USA5 (University of Tennessee)	https://platform.efabless.com/projects/1469
14	Mix-Pix - A Mixed-Signal Circuit for Smart Imaging	Chile (Universidad del Bío-Bío)	https://platform.efabless.com/projects/1494

Made possible with Efabless' **chipIgnite**

TINY TALKS webcast

efabless.com/projects

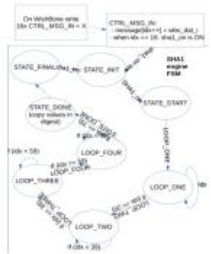


caravel_fulgor_opamp public

Diego Hernando | http://www.fundacionfulgor.org.ar/sitio/index.php

Operational amplifier (opamp) based on the Miller topology designed in Skywater SKY130 CMOS process.

MPW-1 SKY130 2.2k

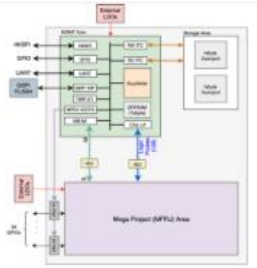


SHA1 engine public

Konrad Rzeszutek Wilk

The SHA1 engine, while not the most secure nowadays is still used by git commits and TPM PCR...

MPW-2 SKY130A 731

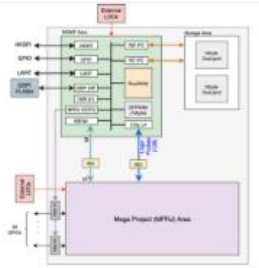


Caravel-SOFA-HD public

Xifan Tang | https://sites.google.com/site/pegaillardon/home

SOFA-HD (Skywater Opensource FPGAs)

MPW-1 SKY130 1.6k

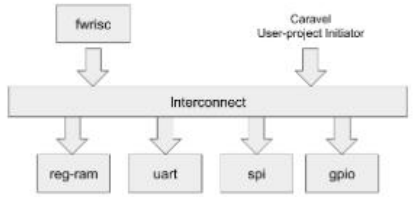


Caravel public

Sylvain Munaut | https://github.com/PyFive-RISC-V

Peripherals tests for future SoC targeting Micro/Circuit Python

MPW-1 SKY130 1.8k

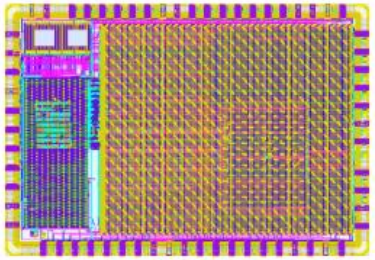


FWPayload public

Matthew Ballance | http://github.com/mballance

A simple RISC-V core+peripherals subsystem for the Google-sponsored Open MPW shuttles for SKY130.

MPW-1 SKY130 962

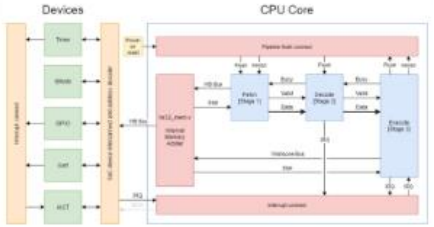


Caravel_Astria_Testchip public

Astria Nur Irfansyah | http://www.its.ac.id

Test circuits consisting of synthesizable comparators for a stochastic ADC, to be submitted for...

MPW-1 SKY130 1.2k

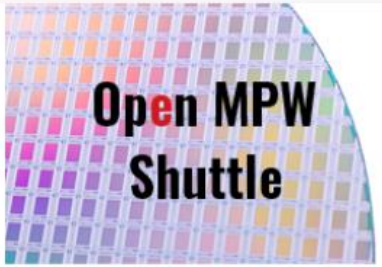


HS32Core public

Kevin Mack Baragona | https://github.com/hsc-latte

Open Source Hardware Processor

MPW-1 SKY130 1.4k



10_bit_potentiometric_DAC public

Sameer S Durgoji | https://www.vlsisystemdesign.com/

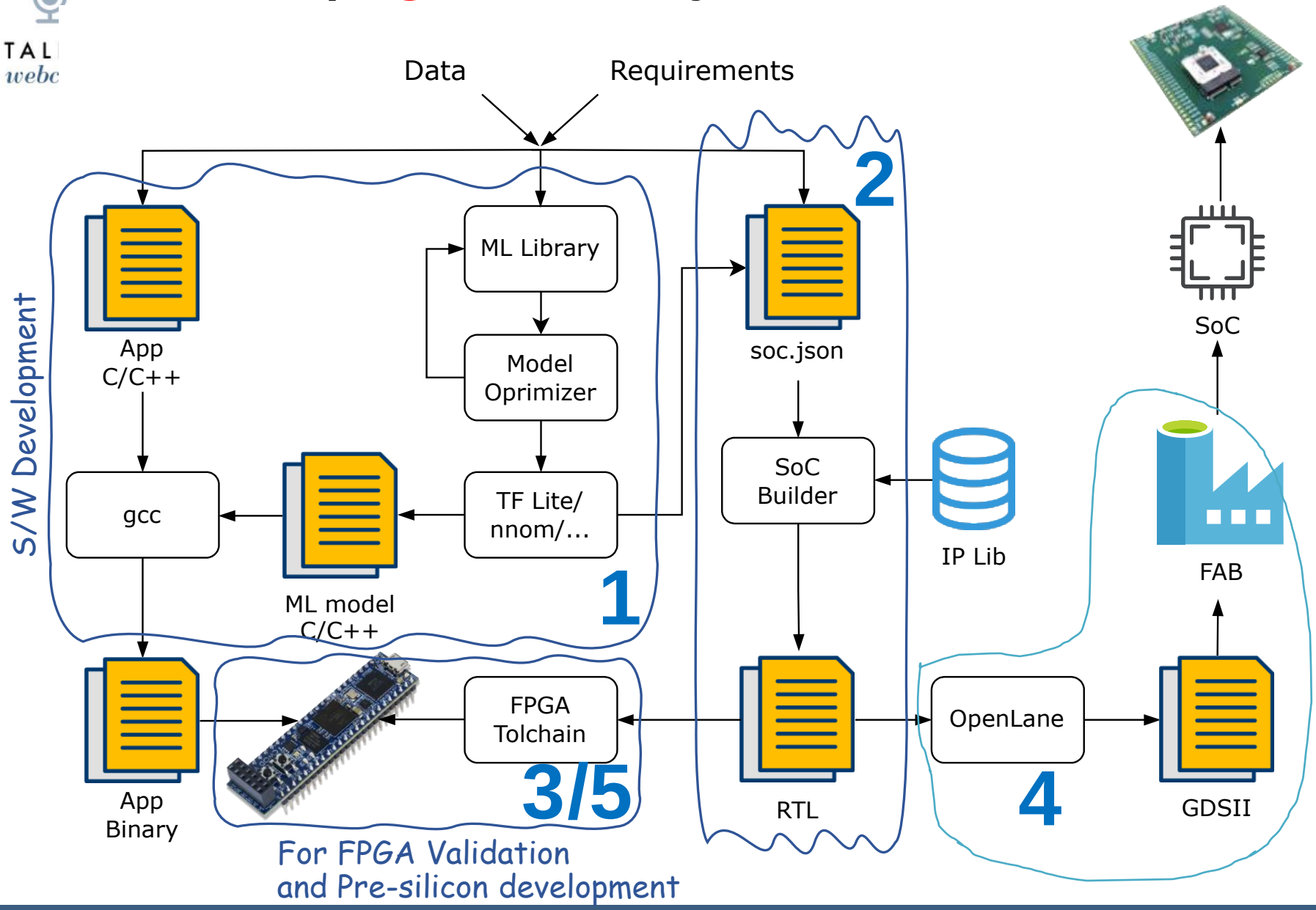
Design of a 10 Bit Potentiometric Digital to Analog Converter with 3.3V analog voltage, 1.8V...

MPW-2 SKY130A 1.1k

TINY ML TALKS webcast

But, Where is TinyML?

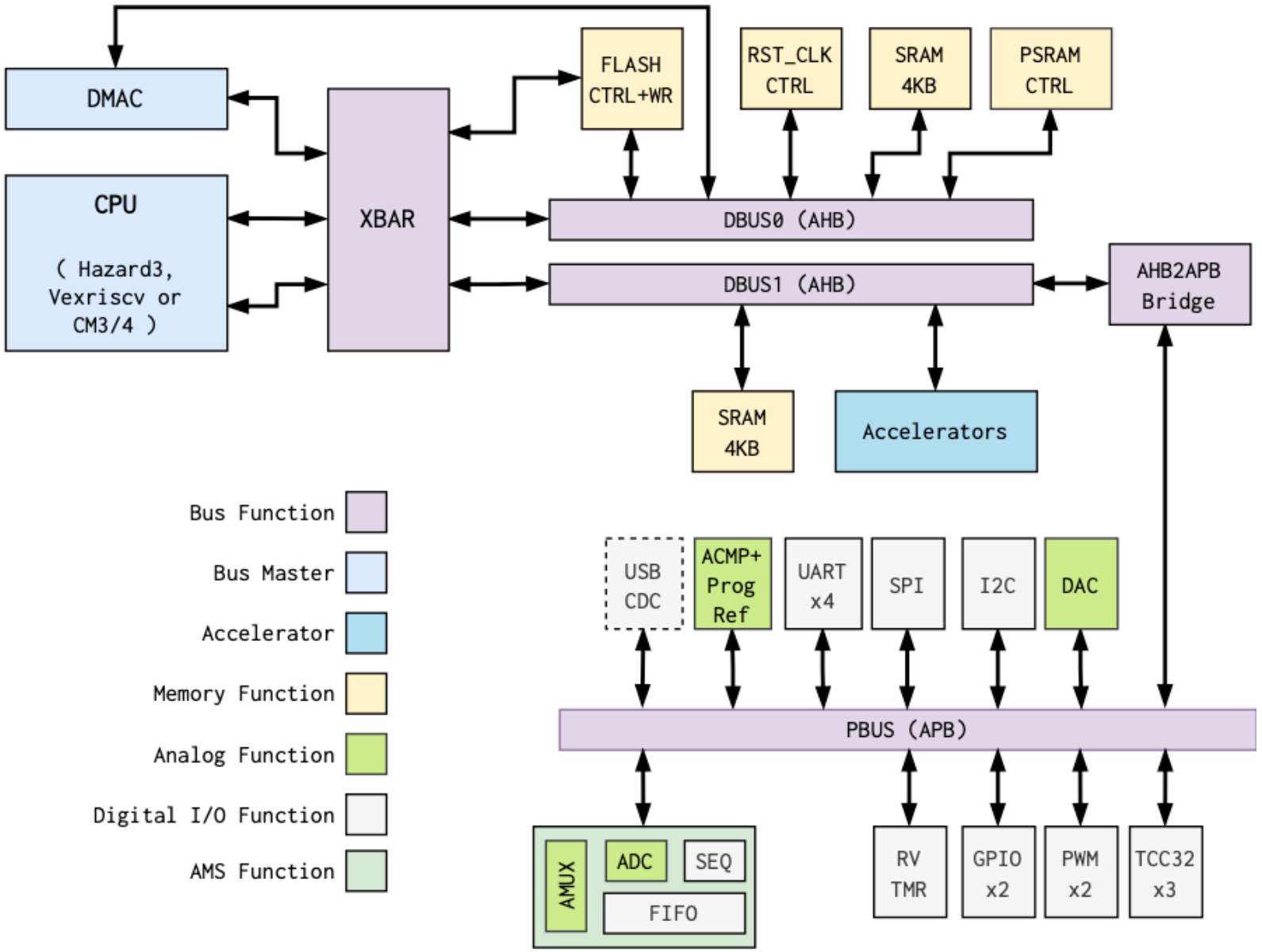
ChipIgnite TinyML Reference Flow



TINY ML TALKS webcast

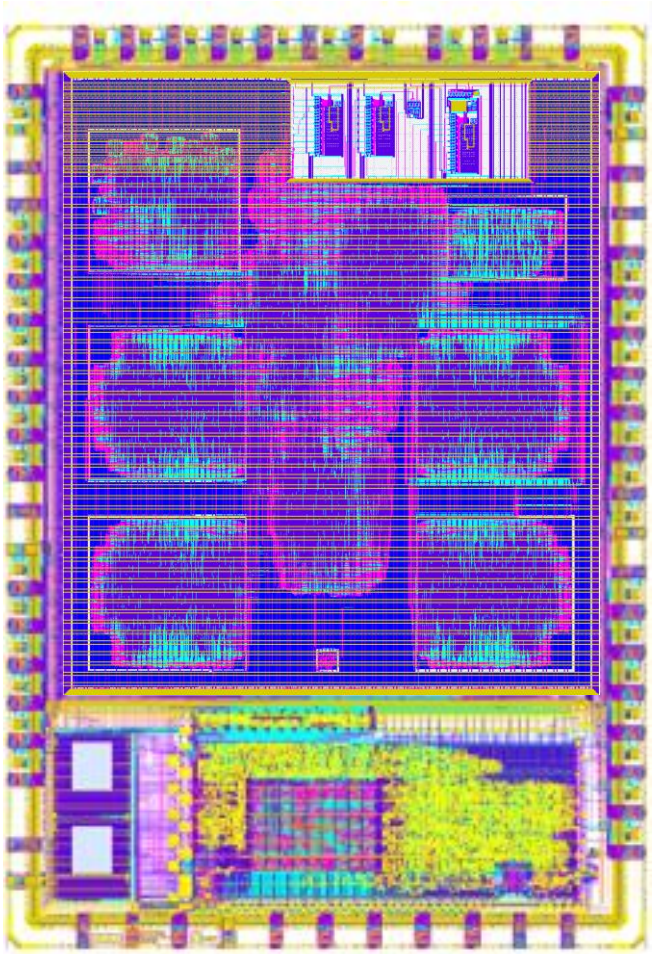
Cheetah SoC

- Targets applications that involve a trained ML models to process low speed sensor data/signal including audio.
- Has ML and Security accelerators

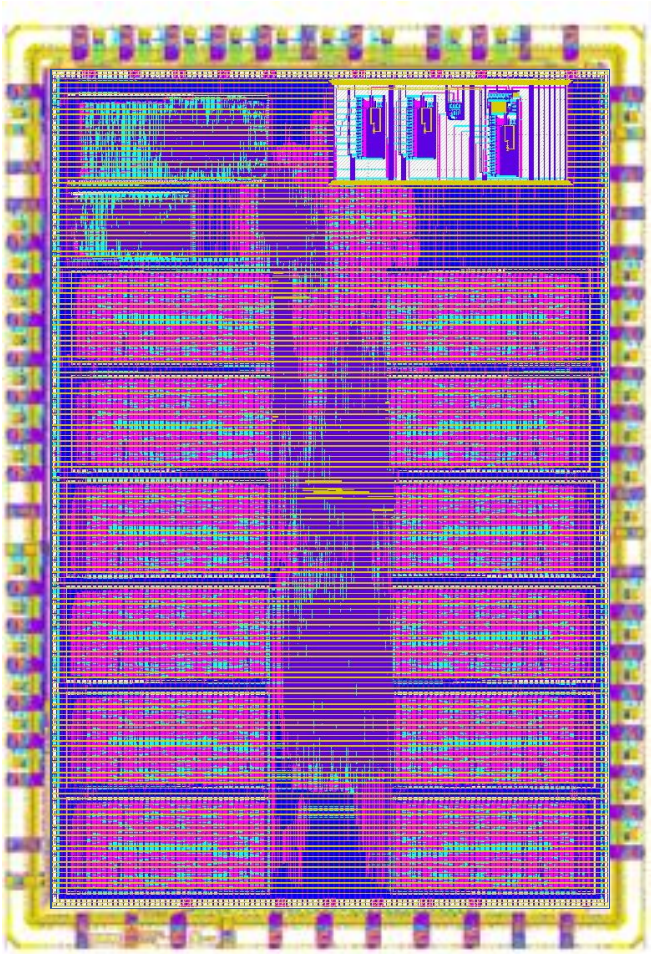


TINY ML TALKS webcast

Cheetah SoC Iterations



Cheetah SoC V1
Taped out: Sept. 23
Caravel



Cheetah SoC V2
Tape out: Nov. 23
Open Frame

TINY ML TALKS webcast

More Information about Efabless and chipIgnite

https://efabless.com/

Join the Open Source Silicon Community

https://open-source-silicon.dev/

Thank you!

Copyright Notice

www.tinyml.org

Copyright Notice

www.tinyml.org