

# オープンソース半導体系 AIプロジェクトと可能性

---

ISHI会：今村謙之  
<https://ishi-kai.org/>  
Mail: [info@ishi-kai.org](mailto:info@ishi-kai.org)

電子情報通信学会 総合大会 2026

# アジェンダ

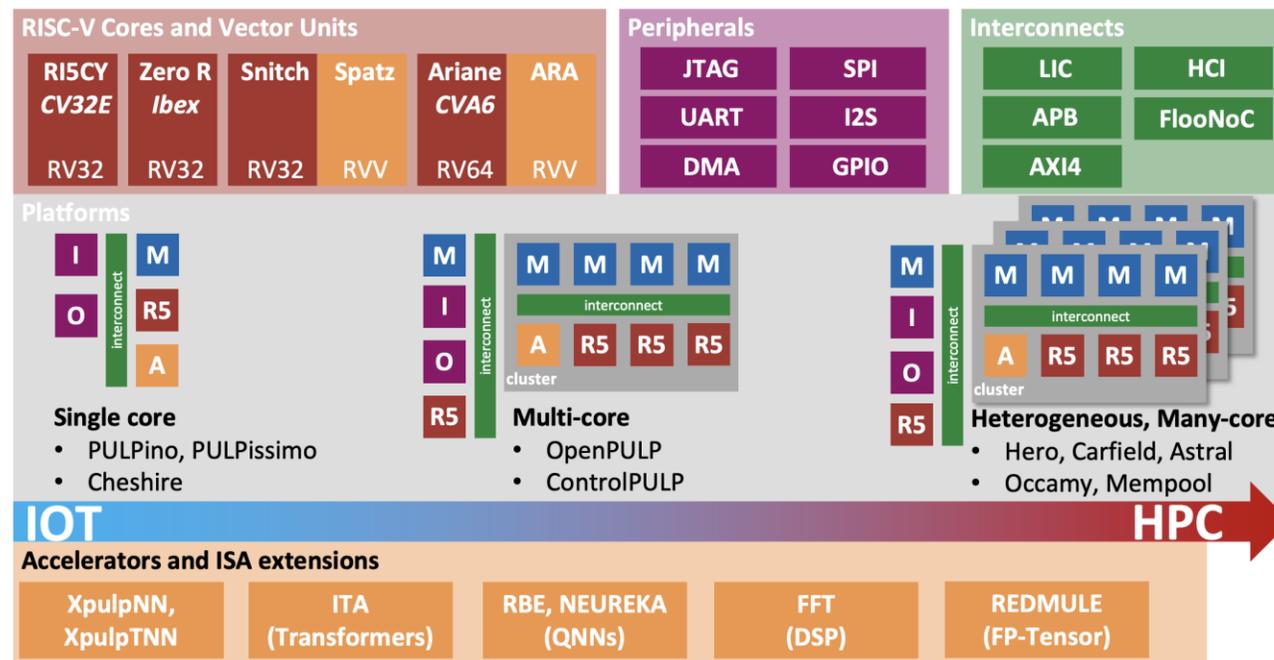
- オープンソース系AIプロジェクト
  - ロジック（デジタル）系
  - アナログ系
- 今後の展望と可能性

ロジック系  
AIプロジェクト

# PULP(Parallel Ultra-Low Power)

<https://pulp-platform.org/>

- ETH (チューリッヒ連邦工科大学) とボローニャ大学が共同で進めている「超低消費電力・高効率・オープンソース RISC-V SoC プラットフォーム」で、IoT・エッジAI・組み込み用途向けに最適化されたマルチコアアーキテクチャ



# AIコンパイラ「Deeploy」

<https://pulp-platform.github.io/Deeploy/>

- 「極低電力デバイスで最新のAIを動かす」ための次世代コンパイラ

従来のコンパイラが主に画像認識（CNN）を対象としていたのに対し、DeeployはTransformerや小規模言語モデル（SLM）を、バッテリー駆動の小さなチップで実行することに特化

## Comparing TinyML compilers

- Deeploy achieves SoA performance
  - Benchmark TinyML workloads
  - Emerging Small Foundation Models
- Deeploy supports your MCU platform
  - ARM, RISC-V, Accelerators, emerging memories, NUMA, ...

Embedded Tool Name	Supports Transformers?	Supports Accelerators?	Supports memory-aware Tiling?	Open and extensible?
TensorFlow Lite	✗	✗	✗	✓
STM Cube AI	✗	✗	✗	✗
GreenWaves NNTool	✓	✓	✓	✗
Dory	✗	✗	✓	✓
Deeploy	✓	✓	✓	✓

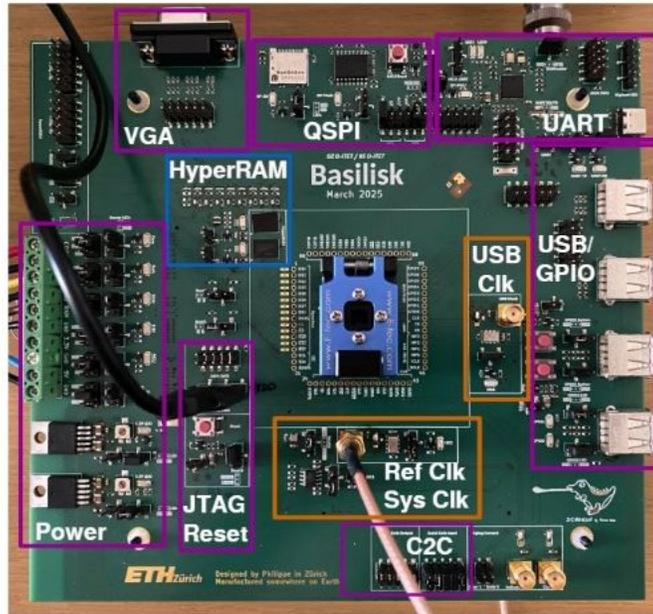
# AIアクセラレータチップ

- Kraken
  - RISC-Vコアと、AI専用のハードウェア拡張（HWPE）を多数搭載するプロジェクト
- Siracusa / Vega
  - 電圧を極限まで下げて動作させる「ニアスレッシュホールド駆動」とAIアクセラレータを組み合わせたプロジェクト
- Astral
  - 「極限環境（特に宇宙空間）」でのAI処理と、最新の「生成AIエッジ推論」の両立を目指した、2026年現在のプロジェクト

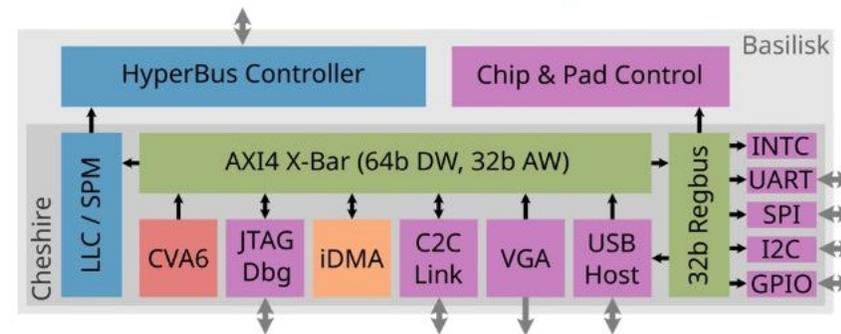
# オープンなチップ

<http://asic.ethz.ch/2024/Basilisk.html>

## Basilisk is a complete Linux-capable SoC



- **64-bit RISC-V core**
- **Rich peripherals:**
  - HyperRAM controller @154MB/s
  - C2C AXI-Link @77MB/s
- **Automatic boot via scratchpad**



[arxiv.org/pdf/2505.10060](https://arxiv.org/pdf/2505.10060)

# AI Foundry

<https://aifoundry.org/>

- Roman Shaposhnik氏らによって2023年に立ち上げられた「完全オープンソースのAIハードウェア／ソフトウェア統合基盤」で、シリコン（RTL）から推論サーバまでをモジュール化し、コミュニティ主導で構築することを目的とした新しいAIインフラストラクチャー

## ET PLATFORM ↗

ET is an open-source manycore ASIC platform for parallel computing acceleration.

- ET Platform is a RISC-V, manycore architecture
- The ET SOC1 contains 1088 compute cores (called minions)
- Each minion has two rv64imfc RISC-V HARTs with vendor-specific vector and tensor extensions
- For a full understanding of the ET SOC1 architecture check the [ET SOC1 Programmers Reference Manual]
- Licensed under Apache License, Version 2.0

## INFERENCE FRAMEWORKS ↗

ET backends for various modern inference frameworks.

- ggml/llama.cpp
- tinygrad
- onnxruntime

## RESEARCH IN QUANTIZATION ↗

A community research project to investigate data-free quantization-aware training.

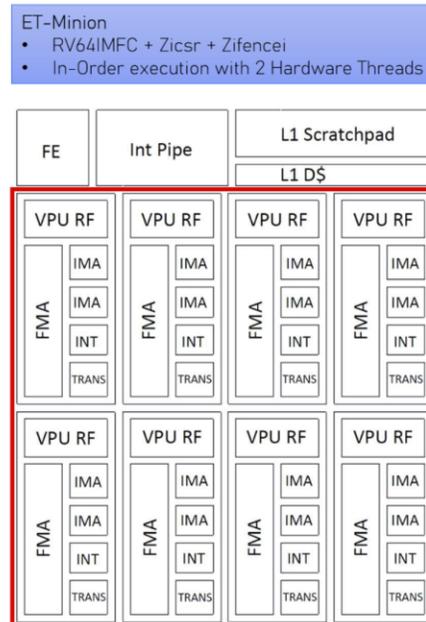
- Quantization Aware Training
- MHAQ
- tinyMHAQ

# AIアクセラレータチップ (?)

リポジトリなし

## Minion overview

- Vector Processing Unit (VPU)
  - 8 lanes, 32 bits per lane
  - 1 FMA unit, 2 IMA unit, 1 INT unit, 1 TRANS unit
  - Extends the RISC-V FP registers from 32b to 256b
  - Each FP register is like a vector with 8 elements of 32b each
- Related [custom extensions](#):
  - Packed Single (PS) → FP32, memory and compute operations
  - Packed Integer (PI) → Int32/ulInt32, compute operations
  - Atomic (AMO) → PS and PI AMO operations
  - Tensor → Memory, FMA32, FMA16A32, IMA8A32, QUANT, REDUCE operations
- Some logic implemented in MMI (hand-tuned)



## What is ET?

The fully open source ~~AI accelerator~~ **manycore platform**.

- Based on Esperanto's ETSOC-1 chip.
- 1088+ RISC-V cores with custom vector and tensor acceleration at 300-800MHz.
- 40W consumption, no external power needed.
- **Ainekko** is *open-sourcing* it under the **AI Foundry** umbrella.



# LLM

<https://github.com/aifoundry-org/llama.cpp/>

## Description

The main goal of `llama.cpp` is to enable LLM inference with minimal setup and state-of-the-art performance on a wide range of hardware - locally and in the cloud.

- Plain C/C++ implementation without any dependencies
- Apple silicon is a first-class citizen - optimized via ARM NEON, Accelerate and Metal frameworks
- AVX, AVX2, AVX512 and AMX support for x86 architectures
- 1.5-bit, 2-bit, 3-bit, 4-bit, 5-bit, 6-bit, and 8-bit integer quantization for faster inference and reduced memory use
- Custom CUDA kernels for running LLMs on NVIDIA GPUs (support for AMD GPUs via HIP and Moore Threads GPUs via MUSA)
- Vulkan and SYCL backend support
- CPU+GPU hybrid inference to partially accelerate models larger than the total VRAM capacity

The `llama.cpp` project is the main playground for developing new features for the [ggml](#) library.

### ▼ Models

Typically finetunes of the base models below are supported as well.

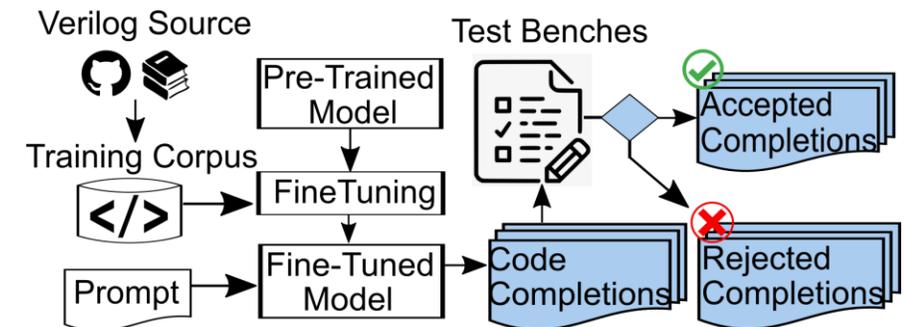
Instructions for adding support for new models: [HOWTO-add-model.md](#)

### Text-only

- LLaMA 🐶
- LLaMA 2 🐶 🐶
- LLaMA 3 🐶 🐶 🐶
- [Mistral 7B](#)
- [Mixtral MoE](#)

# チップAI生成

- Chip-Chat (NYU)
  - <https://arxiv.org/abs/2305.13243>
  - <https://github.com/MJoergen/ChipChatData>
- VeriGen: A Large Language Model for Verilog Code Generation
  - <https://arxiv.org/abs/2308.00708>
  - <https://github.com/shailja-thakur/VGen>



# ロジック系まとめ

- 
- AIチップを作るほうが主流
    - まだ、製作途中
    - まだ、オープン化されていない

アナログ系  
AIプロジェクト

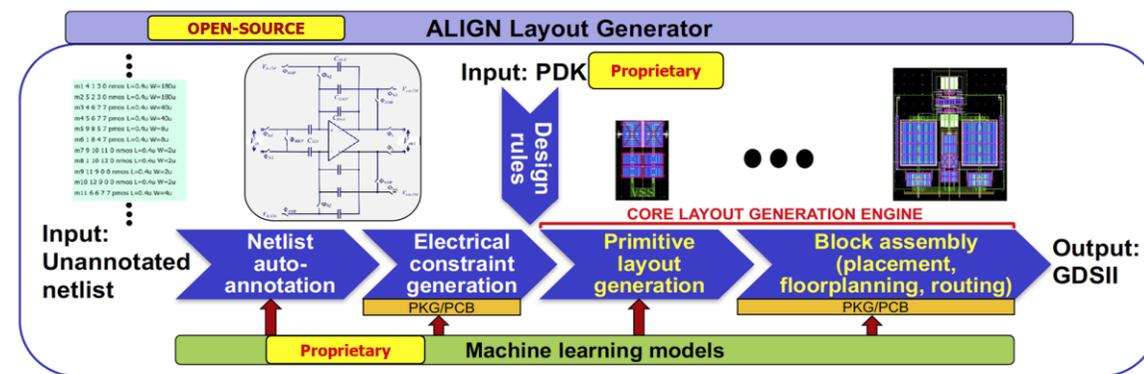
# ALIGN (Analog Layout, Intelligently Generated from Netlists)

<https://align-analoglayout.github.io/ALIGN-public/>

DARPA IDEAプロジェクト発の  
オープンソース自動アナログ  
レイアウト生成ツール



ALIGN: Layout generation from netlist → GDSII





# IEEE SSCS Chipathon

- IEEE SSCS (Solid-State Circuits Society) が主催する“オープンソース IC 設計ハッカソン”で、GF180MCU などのオープン PDK を使い、世界中の参加者がチームでアナログ／デジタル回路を設計し、最終的にテープアウトまで進める教育・実践イベント
- テーマ
  - 2024: "Automating Analog Layout"
    - OpenFASoC/Glayoutを用いたプログラミングによるレイアウト自動生成
    - 設計対象: カレントミラー、演算増幅器 (Op-amp)、電圧リファレンスなどの基本アナログブロック
  - 2025: "Blocks & Bots: An Open Chip Playground augmented with LLMs"
    - MOSbius-style Chips: アナログ版FPGAのような、プログラム可能なアナログ・ファブリック・チップの設計
    - LLM-driven Digital Building Blocks: LLM (ChatGPT等) を用いたプロンプトエンジニアリングにより、デジタル回路 (Verilog等) を自動生成・検証するフローの構築
    - AI/LLM Enhanced Analog Automation: GLayoutを活用し、自然言語による指示からアナログレイアウトを生成する技術の高度化

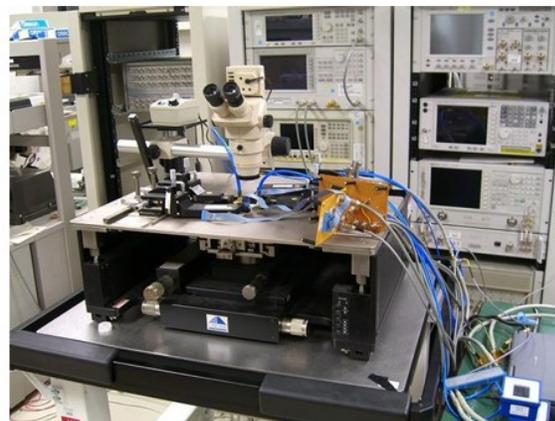
# Chipathon2023

土谷先生の呼びかけにより日本チームを結成

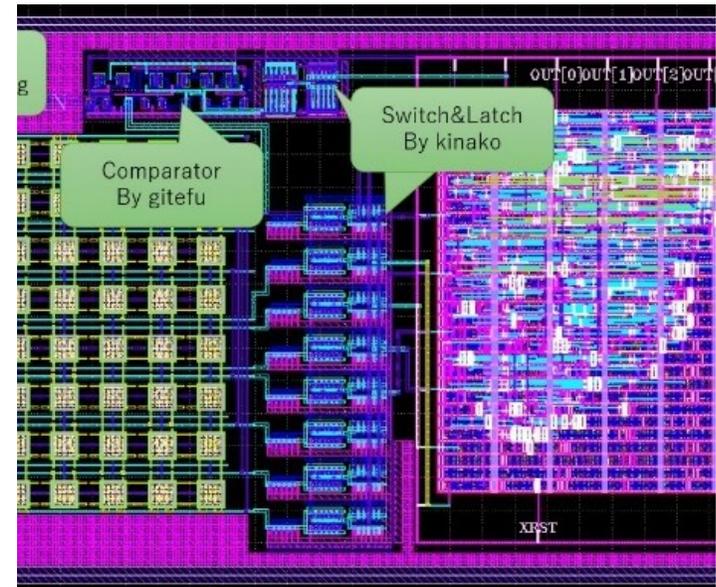
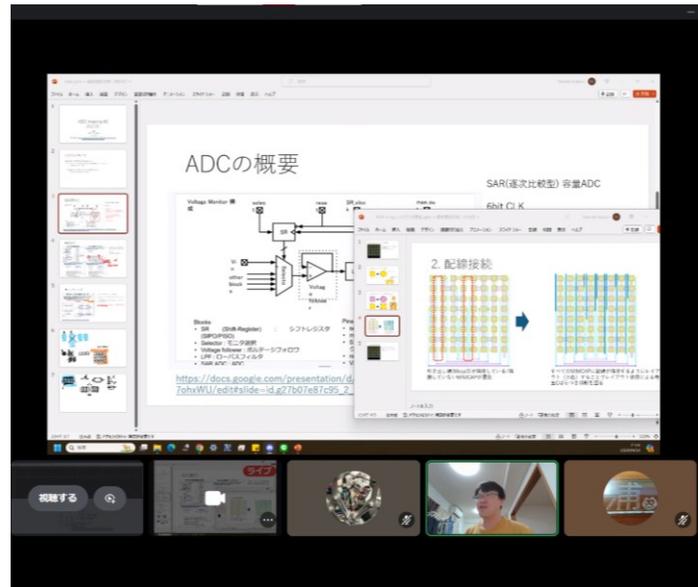
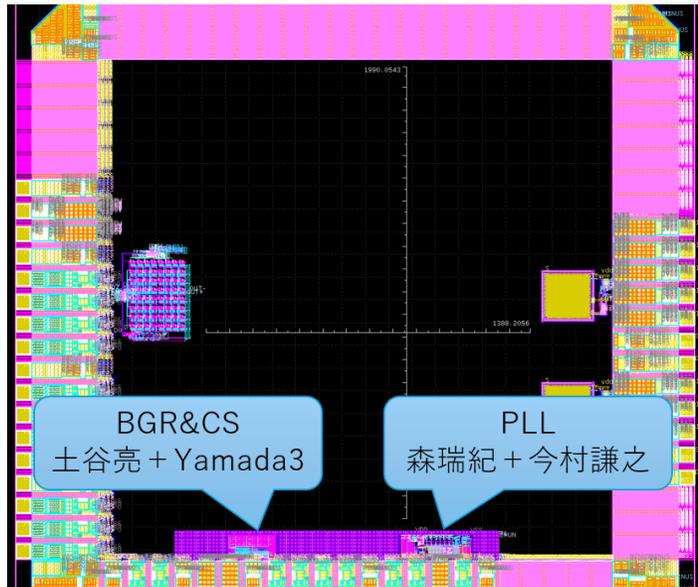
## Chipathon 2023 がやりたいこと

“lab bench on a chip”

測定器をチップ上に載せてしまおう



別の言い方をすると “AnalogDiscovery2 on a chip”



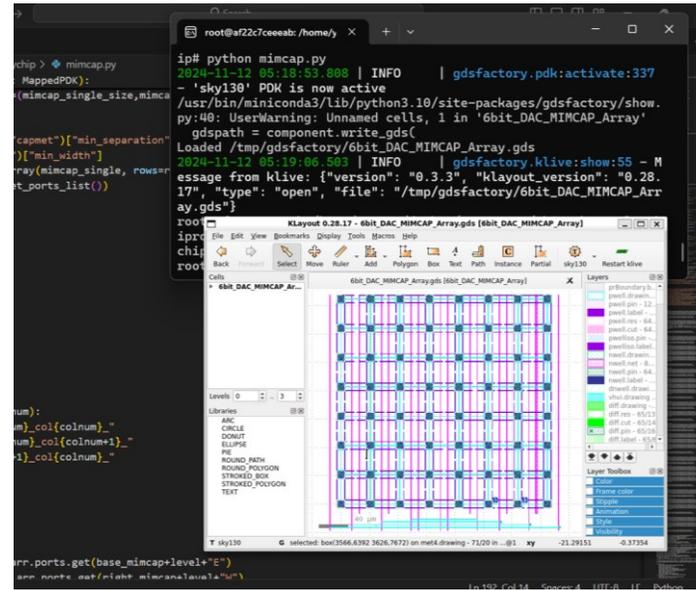
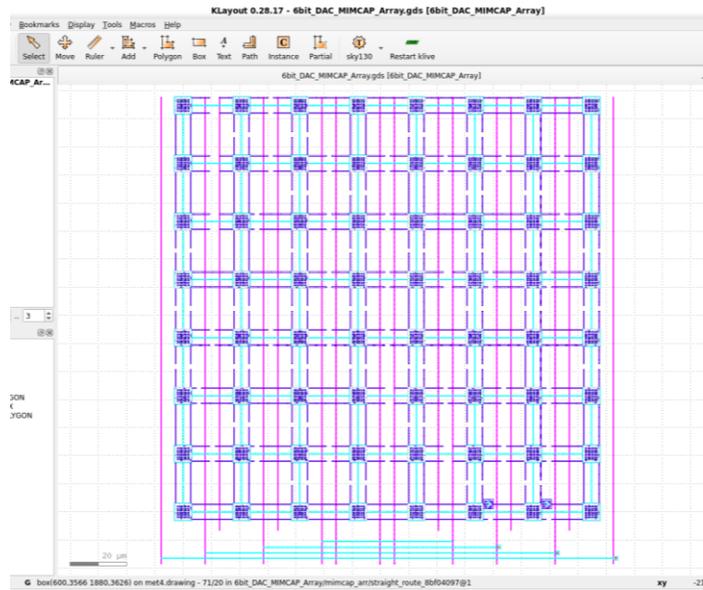
# Chipathon2023 成果

- BGR+CSグループ
  - 土谷先生とハンドルネーム：Yamada3氏
- PLLグループ
  - 慶應大学のM1森さんと私
- ADCグループ
  - 久保木先生をリーダーとしたCS系学部生やハードウェア系企業に勤める若手4名
- 現在ステータス
  - チップの製造完了

# Chipathon2024

- 目標
  - Fifty-NiftyをベースにAIによる自動生成を行う
- チーム構成
  - 2~3名の小規模チーム
- 日本チーム：2チーム
  - DAC：SaltyChipチーム
  - LNA：ISHI-Kai LNAチーム

	A	B	C	D	E	F
1		<b>Group Lead</b>	<b>Group Members</b>	<b>Project</b>	<b>Cells</b>	
9	SemiCon	A. Enes Turan	-	Transmission Gate Common-Source Amplifier		
10	SaltyChip	Yuta Shiomi	M.Yamada Bo-Yu Tseng	Transmission Gate(15) MIM Capacitor Current Mirror(8)Regulated Cascode(46)		
11	Mahowald-ers	Saptarshi Ghosh	Saptarshi Ghosh Sudansu Mishra	Regulated Cascoded Current Mirror Cascoded Common Source (12) Regulated Cascode (48) pA current Source (25) Ref: DOI: <a href="https://doi.org/10.1109/AE54730.2022.9920096">10.1109/AE54730.2022.9920096</a>		
12	Clonky	J. Clancy	J. Clancy	Current Mirror (8) Differential Pair(9) Tunable Resistor (39) Gilbert Cell (56)	x	
13	ISHI-Kai LNA	Noritsuna Imamura	-	Low noise amplifier(23) Voltage Reference & Current Source ( <a href="https://github.com/atuchiya/DC23-1-TC2/tree/master/sgnBGR">https://github.com/atuchiya/DC23-1-TC2/tree/master/sgnBGR</a> )		
14	MicroNinos	Eva Deltor	Eva Deltor Ali Abdul	Miyahara Comparator (Clocked Comparator with Offset Compensation) Requires: Diff Pair, XCP, Charge Pump, refer to this		



# Chipathon2024 : SaltyChipチーム

- 状況
  - ほぼ実装が完了し、LVSフェーズ

# アナログ系まとめ

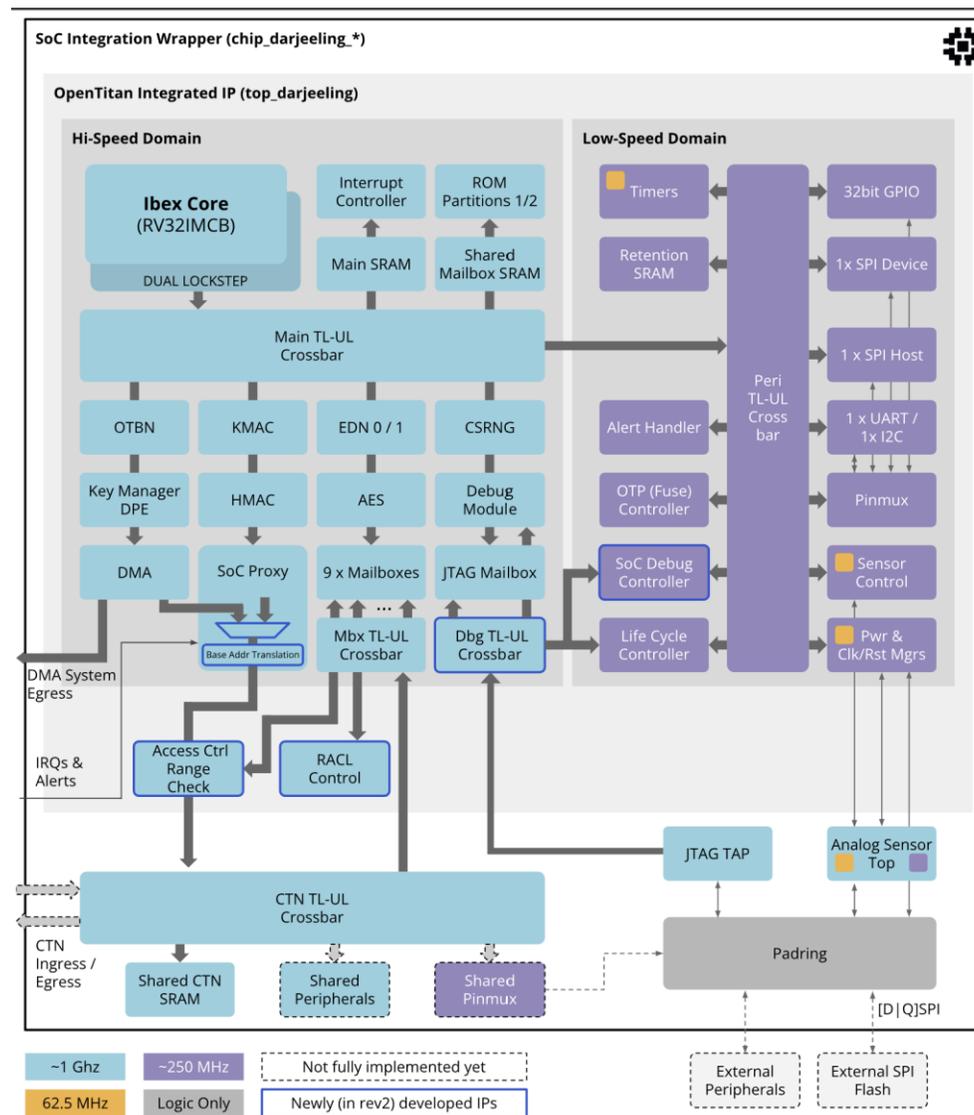
- 
- Chipathonのエコシステムを中心にサーベイすると良い
    - ぜひ、Chipathonに参加しましょう！

今後の展望と可能性

# OpenTITAN

<https://opentitan.org/>

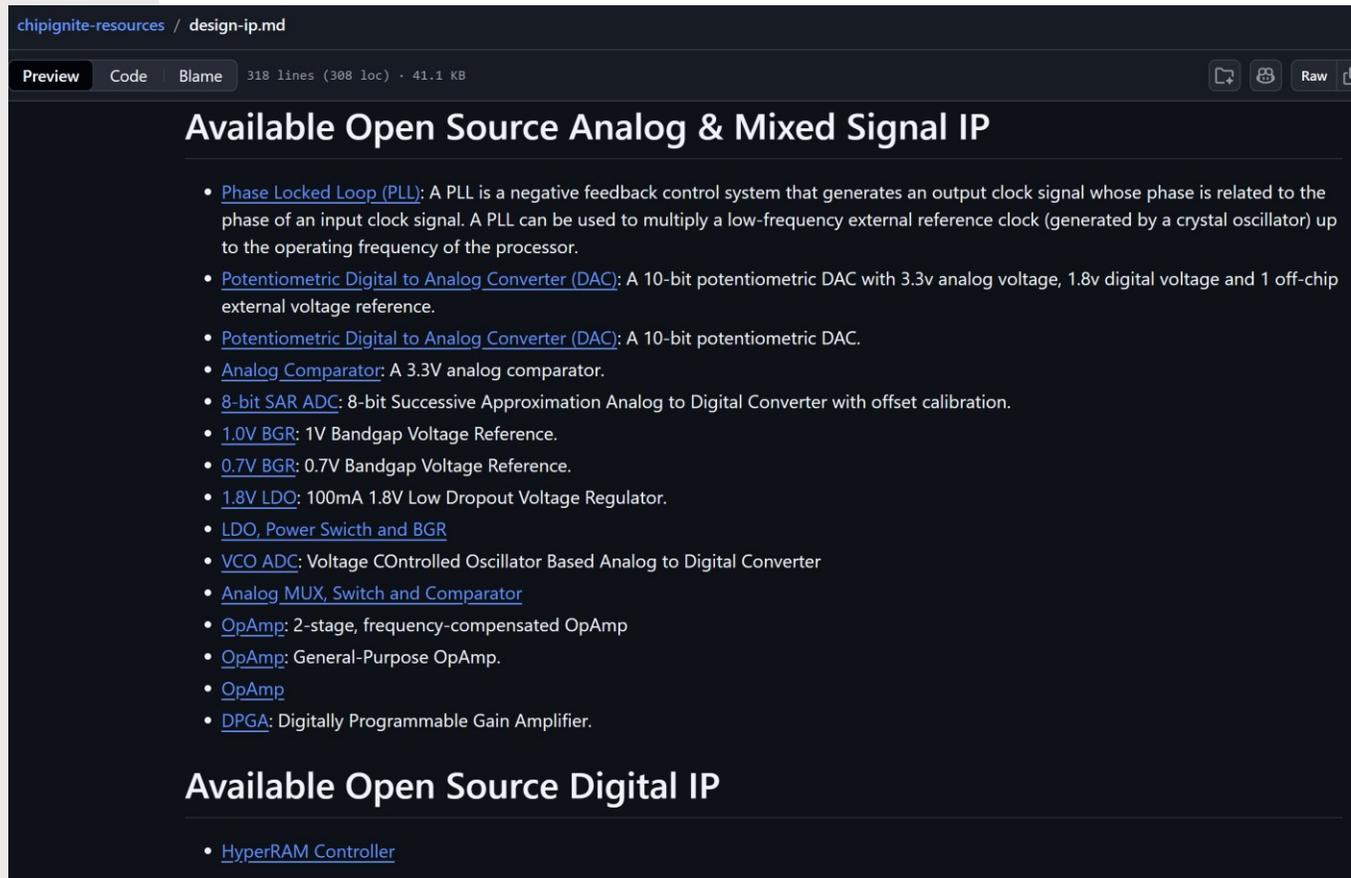
- 2019年11月にGoogle社が主導し、非営利組織「lowRISC」が管理する、世界初の「完全に透明なセキュリティ・チップ (Root of Trust: RoT)」を作るオープンソース・プロジェクト
- 2026年現在、単なる研究段階を終え、Google社のChromebookやデータセンター用サーバー、さらには次世代AIチップのセキュリティ基盤として商用利用が本格化



# Chipalooza

<https://github.com/efabless/chipignite-resources/blob/main/design-ip.md>

- eFabless社が中心となり実施した「オープンソースのアナログIPエコシステムを爆発的に増やすためのコンテスト・イベント」
  - 1回目のテープアウトまでは実施された
  - ChipFoundry.ioが引き継いでいる



chipignite-resources / design-ip.md

Preview Code Blame 318 lines (308 loc) · 41.1 KB

## Available Open Source Analog & Mixed Signal IP

- [Phase Locked Loop \(PLL\)](#): A PLL is a negative feedback control system that generates an output clock signal whose phase is related to the phase of an input clock signal. A PLL can be used to multiply a low-frequency external reference clock (generated by a crystal oscillator) up to the operating frequency of the processor.
- [Potentiometric Digital to Analog Converter \(DAC\)](#): A 10-bit potentiometric DAC with 3.3v analog voltage, 1.8v digital voltage and 1 off-chip external voltage reference.
- [Potentiometric Digital to Analog Converter \(DAC\)](#): A 10-bit potentiometric DAC.
- [Analog Comparator](#): A 3.3V analog comparator.
- [8-bit SAR ADC](#): 8-bit Successive Approximation Analog to Digital Converter with offset calibration.
- [1.0V BGR](#): 1V Bandgap Voltage Reference.
- [0.7V BGR](#): 0.7V Bandgap Voltage Reference.
- [1.8V LDO](#): 100mA 1.8V Low Dropout Voltage Regulator.
- [LDO, Power Switich and BGR](#)
- [VCO ADC](#): Voltage COnrolled Oscillator Based Analog to Digital Converter
- [Analog MUX, Switch and Comparator](#)
- [OpAmp](#): 2-stage, frequency-compensated OpAmp
- [OpAmp](#): General-Purpose OpAmp.
- [OpAmp](#)
- [DPGA](#): Digitally Programmable Gain Amplifier.

## Available Open Source Digital IP

- [HyperRAM Controller](#)

# 今後の展望と可能性

- 
- コードや回路図, レイアウトが公開され, 利用することが可能
    - AIの進化に必要なものの一つである学習用のデータとして利用可能
  - AI Foundryのように公開されているLLMやツールチェーンを自プロジェクトと連携することが可能
    - 半導体にかかわっていない人の参入が期待できる