InspireSemi™

Disruptive Next Generation Accelerated Computing Platform

Blistering speed, energy efficiency, versatility, and affordability for HPC, AI and graph analytics applications

Overview for IEEE Austin

October 5, 2023

Accomplished Leadership Team

Alexander Gray, Founder, President & CTO

- 15 years experience in tech startups, entrepreneurship
- CryptoCore, SolarBridge, SunPower
- Holds 9 patents
- BSEE, University of Illinois at Urbana-Champaign (age 20)

SolarBridge Technologies SUNPOWER

Ron Van Dell, CEO

- 40 years experience and an exceptional track record of success and proven leadership skills in early-stage, turn-around and established businesses
- Former CEO of Primarion (Infineon), SolarBridge, and several other semiconductor and hardware startups
- GM Dell, VP-GM of Communication Products at Harris Semi (Intersil/Renesas)
- BSEE Michigan Technological University

SolarBridge Technologies primarion HARRIS DELL GE

John B. Kennedy, CFO

- 30+ years experience in tech startups and public companies
- Trilumina, SolarBridge, Primarion, KPMG
- BS Accounting & Finance, Elmira College, NY

TRILUMINA SolarBridge Technologies primarion KPMG

Thomas Fedorko, COO

- 35+ years hands-on technical and business leadership in semiconductor operations in both large IDM and startups
- Eta Compute, Uhnder, Bluetechnix, Black Sand (Qualcomm), Luminary Micro (TI), Oak Technology, Motorola SPS
- Technical degree from DeVry University and graduate of the Motorola Management Institute

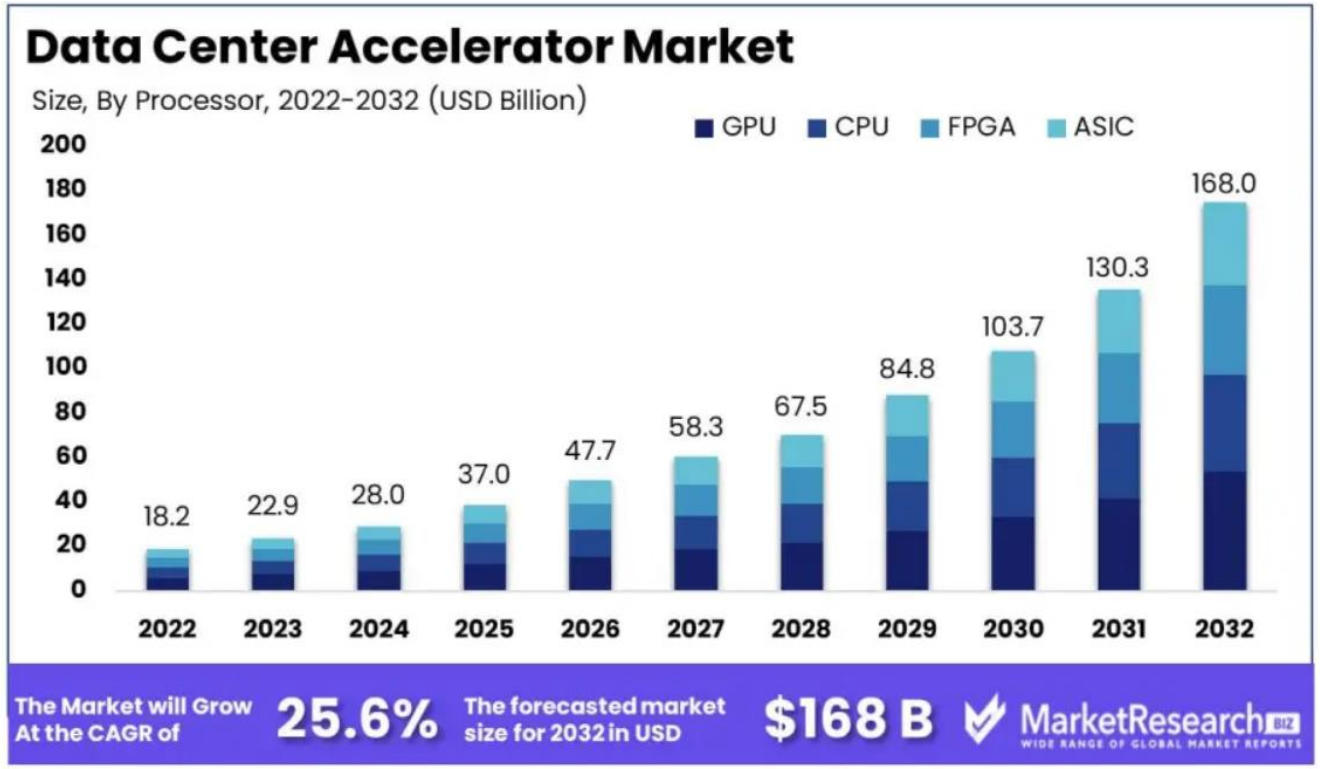
Eta Compute UHNDER Qualcomm Black Sand Technologies
Texas Instruments LUMINARY MICRO OAK MOTOROLA

Doug Norton, CMO

- 35+ years experience; enterprise, startups, Federal
- Nimbix, Newisys (Sanmina), CoWare, Cadence, IBM
- President of Society of HPC Professionals, Technology Advisors Group Austin, TEXGHS Innovation Consortium
- RISC-V International: member SIG-HPC & Marketing team
- BSEE, Missouri University of Science and Technology

NIMBIX VIRTUAL INSTRUMENTS NEWISYS IBM
CoWare cadence

InspireSemi Addresses $168B Datacenter Accelerator Market Growing at 25.6% CAGR



Source: Market.us, June 2023

Initial focus on HPC market

- Large HPC market segment
  - $5-10B TAM of early tech adopters
  - Fast-growing 25.6% CAGR, 2X in 3 years
  - Highly compute-intensive
  - Leverage open-source software
  - Install in own datacenters for fast adoption
- AI market is later upside
  - NVIDIA *de facto* AI solution – for now
  - Hyperscalers developing proprietary AI chips and apps
  - Hyperscalers and startups driving NVIDIA to add AI features at expense of HPC features

Company Overview

Year Founded
2020

Headquartered
Austin, TX

Current Valuation
~$20M

Exchange
TSXV: INSP

Diluted Shares Outstanding
280 Million

Capital Efficiency
20 Employees

Product Offering

- High performance compute accelerator on a card
- Easy-to-deploy PCIe add-in card form factor
- >7,000 high performance 64-bit CPU cores
- Innovative high speed interconnect fabric provides high bandwidth and low latency between cores
- Best-in-class for both Performance/$ & Perf/Watt
- Built on RISC-V architecture and leverages open software ecosystem

Customer Interest

Sandia National Laboratories
TACC Texas Advanced Computing Center
bp
CGG
Argonne National Laboratory

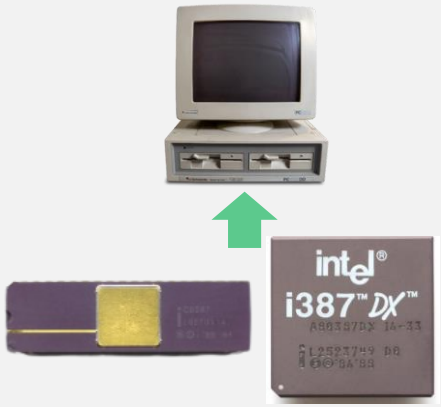
OEM Partnerships

Lenovo
Penguin Computing — A subsidiary of SMART Global Holdings, Inc.
World Wide Technology
E4 Computer Engineering
2crsi

The Third Wave of Accelerated Computing is Here

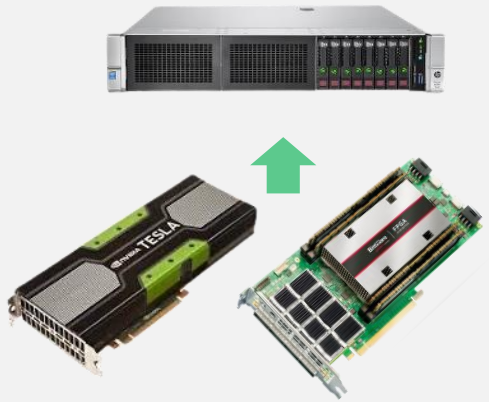
Thunderbird for HPC, AI, Graph Analytics

1980
Math Coprocessor



- Purpose-built widely applicable
- Open software ecosystem
- Plugs into existing computers

2007
GPU, FPGA



- Limited applications benefit
- Proprietary software model
- Plugs into existing servers

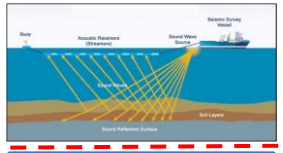
2023+
Thunderbird

- Built for HPC
- Versatile & open software ecosystem
- Plugs into existing servers

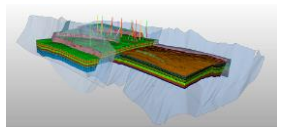
Addressing the Need to Accelerate All HPC & AI Software

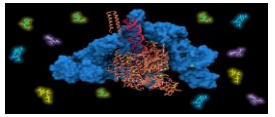
What customers always wanted…Not "yet another GPU"

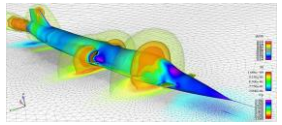

Financial simulations

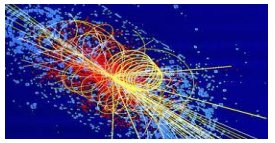

Geology: Seismic

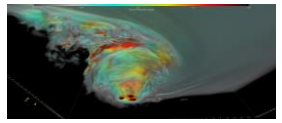

Financial Trading & Graph Analytics


Energy: Reservoir Modeling & Sim

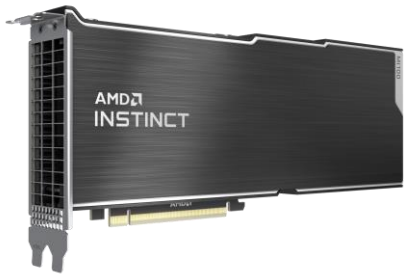

Genomics, Pharma, Life Sciences


CAE/Computational Fluid Dynamics


Nuclear Simulations Fission & Fusion


Climate & Weather Modeling

Datacenter GPUs

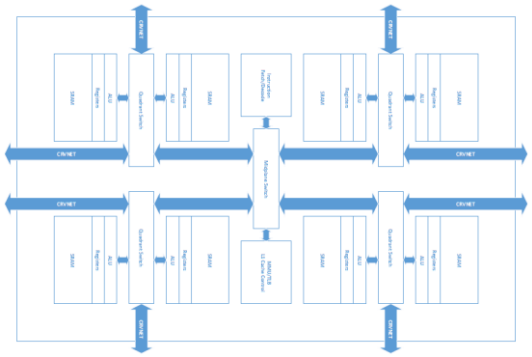


InspireSemi Thunderbird

Highly differentiated "supercomputer-cluster-on-a-chip"

- Versatility as a platform across wide range of applications
- Each chip has 1,792 CPU cores connected via high-speed network
- 4 chip PCIe card delivers >7,000 interconnected 64-bit CPU cores
- Large scale computing power, supports up to 256 chips
- Best-in-class for both Performance/$ and Performance/Watt
- Delivers unprecedented capability within an established open software ecosystem

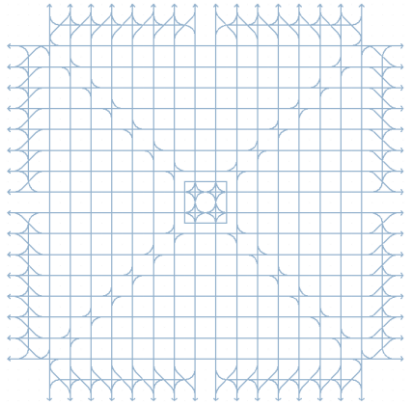
Thunderbird Technology Overview



Core

64-bit superscalar RISC-V CPU cores:

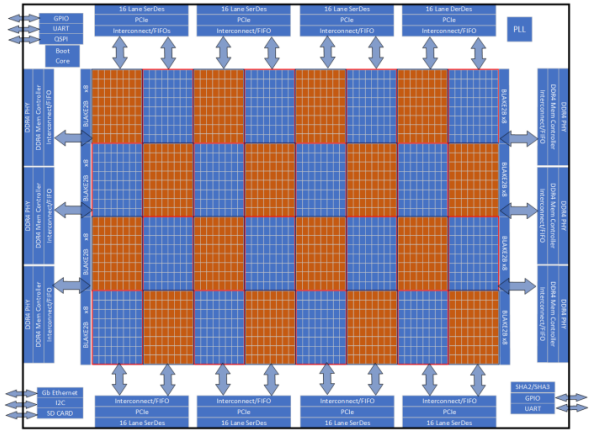
- Custom InspireSemi hi-perf design
- Multiple-issue, out-of-order, variable instruction width
- Vector, SIMD and tensor
- Mixed-precision floating point
- AI and cryptography extensions
- Tightly integrated memory and core-core network fabric
- Simple programming model
- Thriving software ecosystem



Interconnect Fabric

Manhattan street grid of 32-lane superhighways:

- Full utilization of precious routing area -> extreme bandwidth
- Flyover interchanges -> low congestion
- Express bypass lanes -> low latency
- Multiple onramps/offramps to each core
- 240TB/s local, 40TB/s global
- Uniform cellular layout



Top

- 1,792 CPU cores, SMP or HPC cluster-on-a-chip
- Network fabric extensible up 256 chips
- Six DDR4 memory controllers
- 128 lanes chip-to-chip (64 PCIe)
- Algorithm-specific accelerators

System

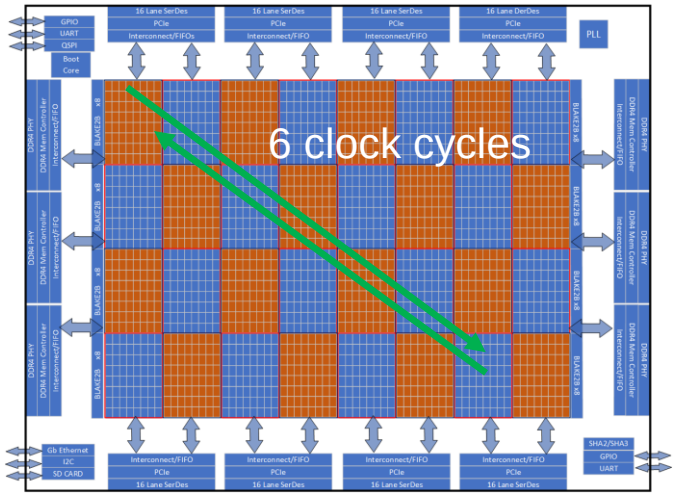
- Power conversion technology improves efficiency 10-25%
- Cooling technology maximizes density, increases efficiency

InspireSemi™

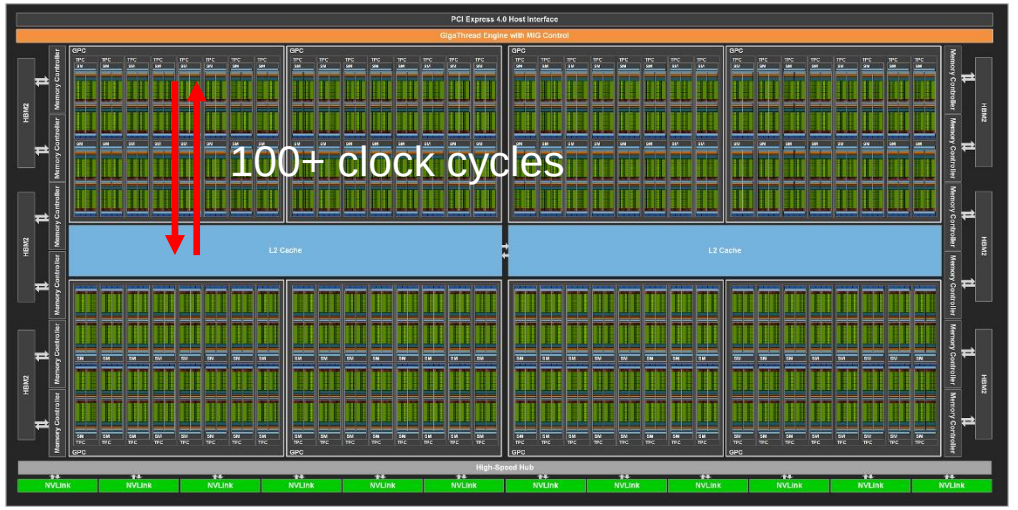
More Applications, High Utilization, and Low Latency

- Thunderbird designed to deliver real world application benefits
  - Software friendly, all CPU architecture (double precision FP64 RISC-V cores) will work with all HPC & AI software
  - High speed, low latency core-to-core communications for predictable performance
  - MIMD architecture (vs. high latency GPU SIMD)
  - Large memory – can address larger problems than fit in GPUs
  - Distributed memory – each core has its own 64KB local fast memory
- ***Result = Greater application performance with less power consumption***
- ***Deterministic + Predictable Performance addresses applications GPUs cannot***

Example – Thunderbird vs. NVIDIA GPU (A100) latency

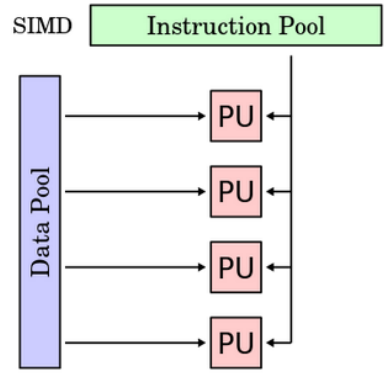


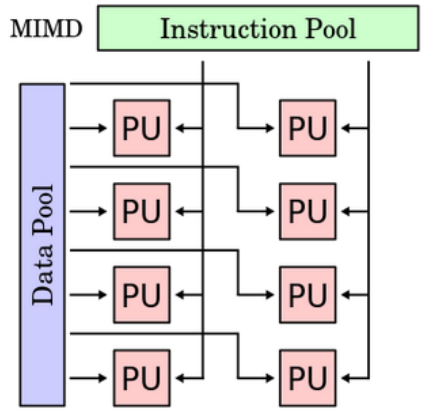
20x
greater efficiency



Thunderbird is Deterministic, GPUs are Not

- **Reproducibility of results is a must for many key applications, disqualifying GPUs**
  - E.g- high-frequency trading, self-driving cars, cryptography, healthcare imaging, smart weapons, …
- Definition:
  - A deterministic system is one that always produces the same output for a given input
  - There is no randomness or chance involved in the system's behavior
  - Instead, the system follows predictable rules that determine how the inputs will be transformed into outputs
- GPU non-determinism problems
  - GPU's use very complex hardware/software task schedulers to try to hide their latency problems (switching between tasks while waiting for others)
  - These schedulers are not deterministic, meaning users cannot know when their tasks will finish
  - They're also proprietary, obscure, and change often, frustrating any attempt to predict their behavior (as with many aspects of GPU architecture)
  - Output results can also vary, due to varying order of sub-task completion and rounding errors
- **Thunderbird solves this**
  - Each core is controlled independently by its own program thread
  - Real-time operating systems provide deterministic software scheduling
  - Easy bare-metal programming gives developers absolute control down to single clock cycles
  - Atomic instruction set provides rich, straightforward task synchronization options
  - Delivering the same results, the same time, every time





Thunderbird for Real-Time, Real-Safe Computing

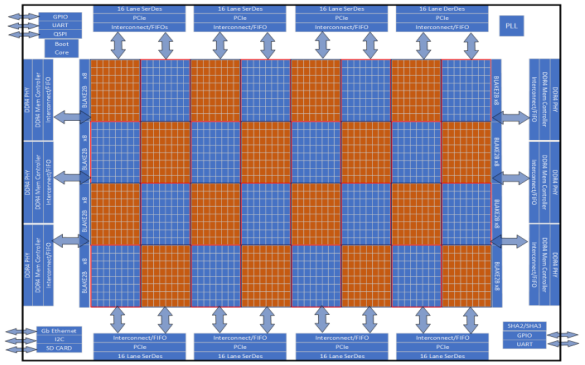
Reproducibility

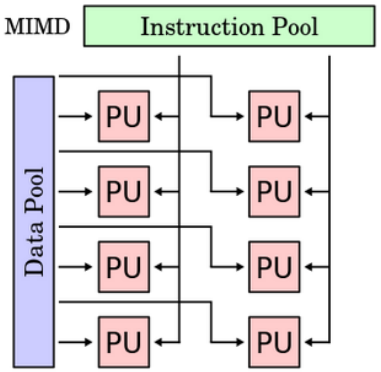
- Results verification
- Performance optimization
- Troubleshooting
- Quality assurance

Determinism/Predictability

- Safety and component functions
- Efficiency optimization
- Known timing behavior
- Quality assurance

Thunderbird supports real-time, real-safe computing applications where GPUs do not work





Thunderbird Delivers on Key HPC Application Benchmarks

8 – 12 TFLOPS
per chip (FP64)
32 – 48 TFLOPS
per card (FP64)

50 GFLOPS/ Watt
(FP64)

25x
reduction in latency

30 – 60%
lower power consumptions

Superior HPC applications benchmarks*

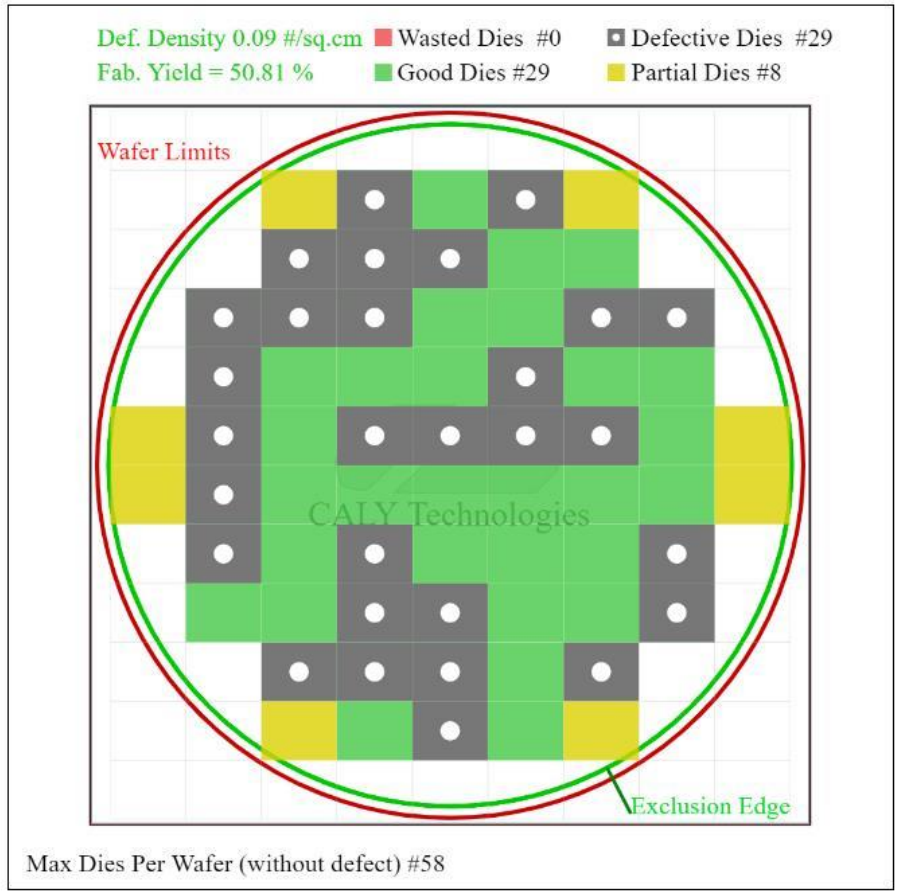
- DGEMM (Double-Precision Generalized Matrix Multiply)
- SpMV (Sparse matrix-vector multiply)

** Available under NDA*

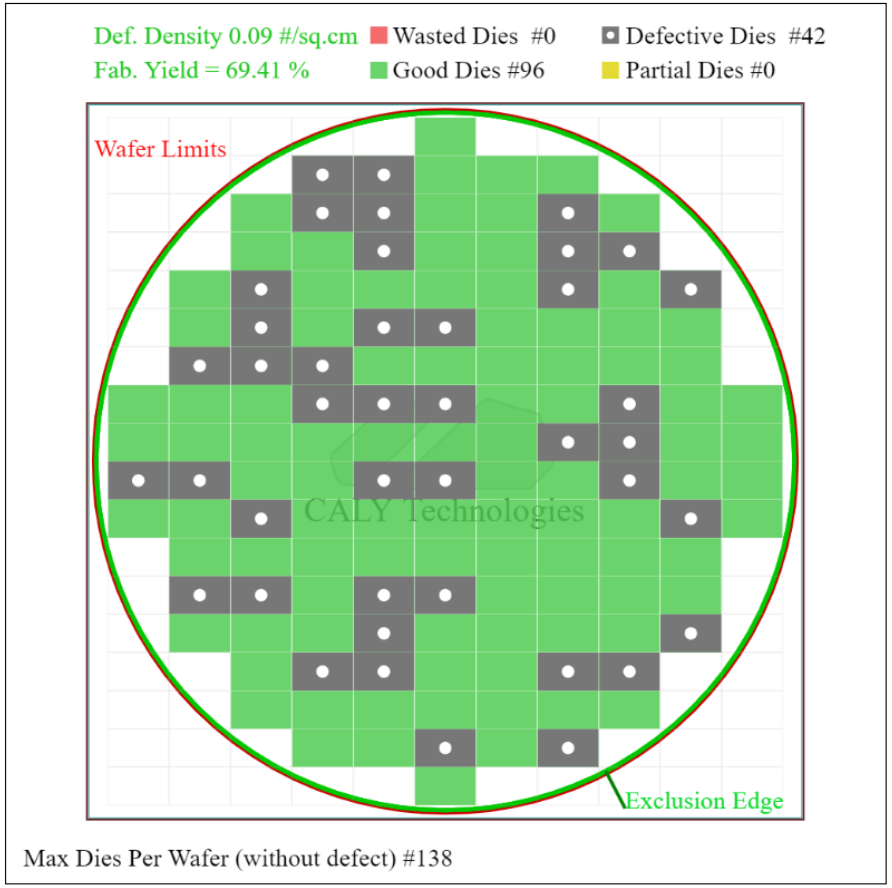
Thunderbird Smaller Die Size Benefits

Half competitors' size = 3X Advantage in Yield/Cost

Competitive chips are enormous
- At manufacturable size limits
- Exotic $$ packaging complications
- Gruesome for yield and cost



We pack comparable capability into half the area
- Higher yield, better wafer utilization
- Net 70% cost reduction
- Reduced power/latency for on-chip comms



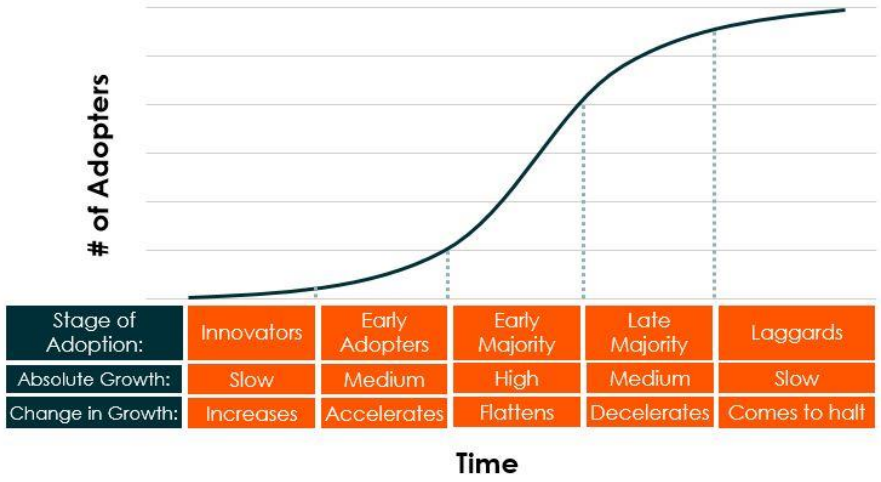
Open Software Ecosystem Solves Customer Porting Challenges

- Leverages established RISC-V software ecosystem
  - Eliminates need for proprietary software stacks
- Uses standard CPU-style programming models
  - No need for CUDA, ROCM, SYCL, etc. that GPUs require
  - Standard compiler, OpenMP, MPI, etc. approaches
- Key frameworks, compilers, & tools already exist for RISC-V
  - Standard GCC, Gfortran, GDB toolchains
  - Standard HPC libraries (e.g. – BLAS, LAPACK, FFTW)
- Key Operating Systems
  - Linux
  - Real-time kernels (RTOS)

Thunderbird Early Access Program

Thunder and Lightning

- Objectives
  - Jumpstart Thunderbird adoption and software applications
    - i.e. – accelerate the “S curve”
  - Allocate early hardware for key customers & partners
  - Key customer feedback for Thunderbird II
- Overview
  - FPGA emulator dev board w/limited number of CPU cores (“thunder”)
  - Early Thunderbird PCIe board once available (“lightning”)
  - Relevant software for FPGA and PCIe boards
  - Support, regular review meetings
  - Could be additional revenue for consulting services, more Thunderbirds
- Status
  - Strong interest, already receiving PO’s and LOIs



Thunderbird Addresses ALL HPC & AI Customer Needs

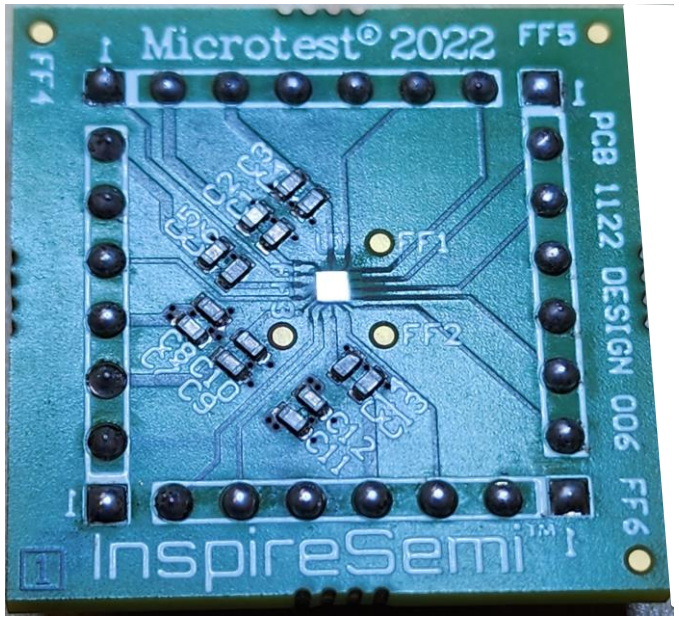
	InspireSemi Thunderbird	CPU	GPU	FPGA	AI Accelerators
Architecture	Many programs, many data streams	Few programs, few data streams	Few programs, many data streams	Programmable logic elements	Single program, many data streams
Performance	High for broad range of HPC apps	Slow, need h/w accelerators	High for AI and some HPC apps	Medium	High for AI only
Cost	Low $6,500 for 2 chip PCIe card	High ~$1K-8K (+ more servers)	High ~$7K-48K	High $8K-$10K	High ~$10K - $2.2M
Energy consumption	Low ~175W/chip	Med 240W+/chip (+ more servers)	High ~700W	High ~300W	High ~300W – 20kW
Scalability	256 chips	1-4 chips	2-8 chips	1 chip	1-2 chips
Programming model	Standard CPU-like, Any language, Full instruction set	Standard CPU, Any language, Full instruction set	Specialized C variant (CUDA, ROCM, SYCL)	Hardware description language	Proprietary, obscure
Software ecosystem	Open-source, Linux, compilers, libraries, AI frameworks, existing applications	Robust	Limited, proprietary	None	AI frameworks and proprietary software stacks

Manufacturing Partners

- **World-class, Tier 1 Supply Chain Partners**
  - Imec selected as VCA (Value Chain Aggregator)
  - TSMC is worlds largest semiconductor foundry
  - ASE is worlds largest OSAT
    (Outsourced Semiconductor Assembly & Test)
- **Wafer Capacity**
  - Leading edge node capacity tight 2022/2023
  - Large Semi companies double booked forecast
  - Global economic uncertainties now resulting in de-bookings freeing up some capacity
- **Thunderbird - 12nm HPC & AI**
  - Wafer capacity allocated to support forecast
  - ABF substrate capacity secured
- **5nm Accelerator Test Chip**
  - Key learnings for next generation
  - Working in lab, met performance and power targets

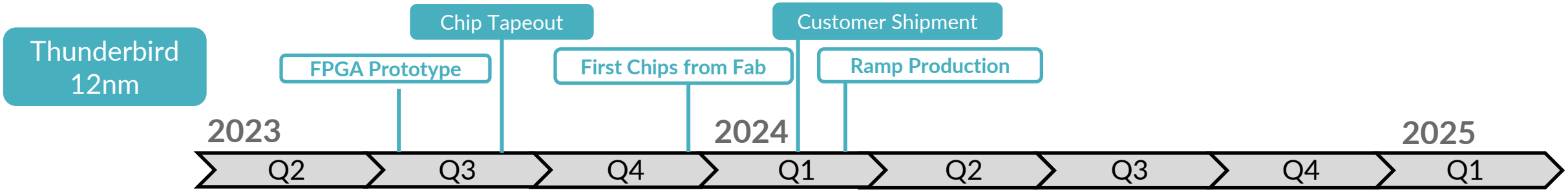
Fundamental Capabilities Proven with TSMC 5nm Test Chip

- De-risks plan for follow-on 4nm Thunderbird II
  - Validated team's ability to deliver designs on leading-edge TSMC process node
  - Worked first time, met performance and power targets
  - Including full-custom layout optimized at every level
  - Something not many companies can do, perhaps none this size
- Benchmarked chip performance results
  - Hand-crafted core can deliver ~50% higher speed than 12nm
  - Optimized core can deliver ~50% lower power than 12nm
  - Proprietary micro-architecture saves >2x power



Robust HPC & AI Roadmap

To Support Demanding Applications

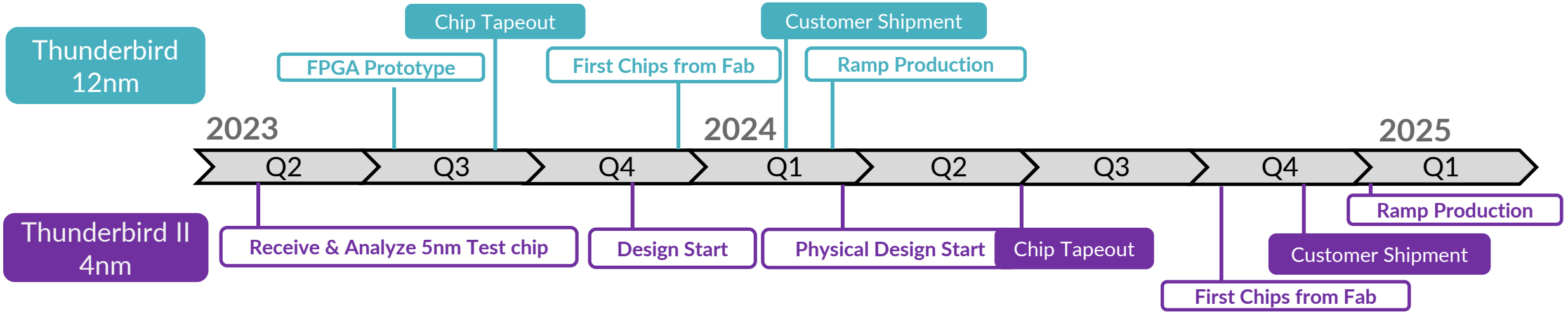


Thunderbird

- Addresses complex HPC applications
  - e.g.- CAE/CFD, energy/reservoir modeling & sim, weather, life sciences/genomics, finance, fraud detection
  - Low-cost LPDDR memory for memory-hungry HPC jobs
- Applicable for AI-augmented HPC
- Ideal for graph analytics

Robust HPC & AI Roadmap

To Support Demanding Applications



Thunderbird

- Addresses complex HPC applications
  - e.g.- CAE/CFD, energy/reservoir modeling & sim, weather, life sciences/genomics, finance, fraud detection
  - Low-cost LPDDR memory for memory-hungry HPC jobs
- Applicable for AI-augmented HPC
- Ideal for graph analytics

Thunderbird II

- TSMC 4nm – higher performance, lower power
  - Quadruples core count to 10,000/chip
- Additional features for AI
  - High Bandwidth Memory (HBM)
  - AI-specific instructions
- Enhanced vector instructions for HPC

Disclaimer

This presentation contains statements which constitute “forward-looking information” within the meaning of applicable securities laws, including statements regarding the plans, intentions, beliefs and current expectations of InspireSemi with respect to future business activities and operating performance.

Often, but not always, forward-looking information can be identified by the use of words such as “plans”, “expects”, “is expected”, “budget”, “scheduled”, “estimates”, “forecasts”, “intends”, “anticipates”, or “believes” or variations (including negative variations) of such words and phrases, or statements formed in the future tense or indicating that certain actions, events or results “may”, “could”, “would”, “might” or “will” (or other variations of the forgoing) be taken, occur, be achieved, or come to pass. Forward-looking information includes, but is not limited to, information regarding: (i) the business plans and expectations of the Company including expectations with respect to production and development; and (ii) expectations for other economic, business, and/or competitive factors. Forward-looking information is based on currently available competitive, financial and economic data and operating plans, strategies or beliefs as of the date of this presentation, but involve known and unknown risks, uncertainties, assumptions and other factors that may cause the actual results, performance or achievements of InspireSemi, to be materially different from any future results, performance or achievements expressed or implied by the forward-looking information. Such factors may be based on information currently available to InspireSemi, including information obtained from third-party industry analysts and other third-party sources, and are based on management’s current expectations or beliefs. Any and all forward-looking information contained in this presentation is expressly qualified by this cautionary statement.

Investors are cautioned that forward-looking information is not based on historical facts but instead reflect InspireSemi’s management’s expectations, estimates or projections concerning future results or events based on the opinions, assumptions and estimates of management considered reasonable at the date the statements are made. Forward-looking information reflects InspireSemi’s current beliefs and is based on information currently available to it and on assumptions it believes to be not unreasonable in light of all of the circumstances. In some instances, material factors or assumptions are discussed in this presentation in connection with statements containing forward-looking information. Such material factors and assumptions include, but are not limited to: the impact of the COVID-19 pandemic on the Transaction or the company; the ongoing conflict between Russia and Ukraine and any actions taken by other countries in response thereto, such as sanctions or export controls; and anticipated and unanticipated costs and other factors referenced in this presentation and the Filing Statement, including, but not limited to, those set forth in the Filing Statement under the caption “Risk Factors”. Although the Company has attempted to identify important factors that could cause actual actions, events or results to differ materially from those described in forward-looking information, there may be other factors that cause actions, events or results to differ from those anticipated, estimated or intended. Forward-looking information contained herein is made as of the date of this presentation and, other than as required by law, the Company disclaims any obligation to update any forward-looking information, whether as a result of new information, future events or results or otherwise. There can be no assurance that forward-looking information will prove to be accurate, as actual results and future events could differ materially from those anticipated in such statements. Accordingly, readers should not place undue reliance on forward-looking information. Should one or more of these risks or uncertainties materialize, or should assumptions underlying the forward-looking information prove incorrect, actual results may vary materially from those described herein as intended, planned, anticipated, believed, estimated or expected.

RISC-V®

Special Thanks to **Krste Asanovic**
Prof. EECS, UC Berkeley; Chairman, RISC-V Foundation

InspireSemi™

Disruptive Next Generation Accelerated Computing Platform

Blistering speed, energy efficiency, versatility, and affordability for HPC, AI and graph analytics applications

Why RISC-V?

Open Interfaces are Accepted Practice

Field	Open Standard	Free, Open Implement.	Proprietary Implement.
Networking	Ethernet, TCP/IP	*Many*	*Many*
OS	Posix	Linux, FreeBSD	Windows
Compilers	C	Gcc, LLVM	Intel icc, ARMcc, Xcode
Databases	SQL	MySQL, PostgresSQL	Oracle12C, DB2
Graphics	OpenGL	Mesa3D	DirectX
ISA	??????	----------	X86, ARM, IBM360

Why not successful open standards and multiple open-source and proprietary implementations, like other fields?

Companies and their ISAs Come and Go

- Digital Equipment Corporation, RIP! (**PDP-11, VAX, Alpha**)
- Intel's dead ISAs (**i960, i860, Itanium, …**)
- **SPARC**
  - Sun opened v8 as IEEE 1754-1994
  - Acquired by Oracle, SPARC development closed down in 2017
- **MIPS**
  - Sold to Imagination, bought by Wave
  - Opened up MIPS R6 ISA in 2018, then made not open in 2019
  - Now doing RISC-V….
- **IBM POWER**
  - Initially proprietary, opened up ISA in 2019. Crickets…
- **ARM**
  - Sold to Softbank in 2016
  - Nvidia announced acquiring ARM 2020, deal cancelled 2022
  - IPO September 2023

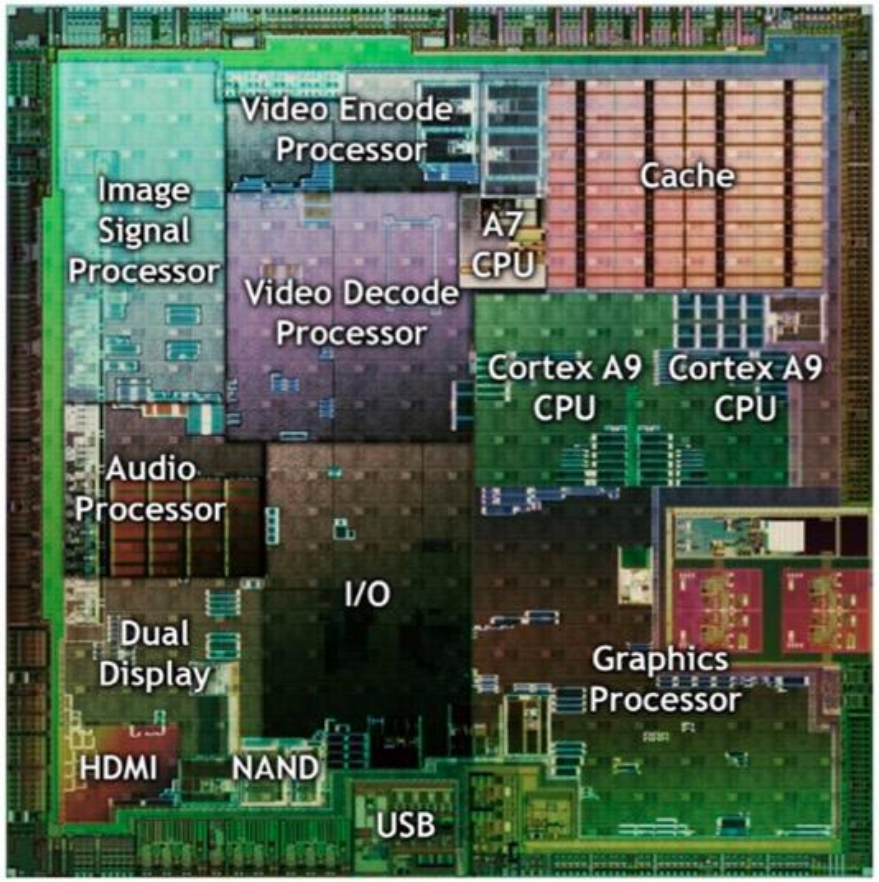
Why entrust your software investment to a proprietary ISA?

SoC ISA Balkanization

- Applications processor (usually ARM)
- Graphics processors
- Image processors
- Radio DSPs
- Audio DSPs
- Security processors
- Power-management processor
- ***>dozen ISAs on some SoCs – each with unique software stack***

Why?

- Apps processor ISA too big, inflexible for accelerators
- IP bought from different places, each proprietary ISA
- Engineers build home-grown ISA cores (don't do this!)



NVIDIA Tegra SoC

Do we need all these different ISAs?

Must they be proprietary?

Must they keep disappearing?

What if there was one stable free and open ISA everyone could use for everything?

RISC-V Background

- In 2010 at UC Berkeley, after many years and many research projects using MIPS, SPARC, and x86, time for architecture group at UC Berkeley to choose ISA for next set of projects
- Obvious choices: x86 and ARM
  - x86 impossible – too complex, IP issues
  - ARM mostly impossible – complex, no 64-bit in 2010, IP issues
- So started "3-month project" during summer 2010 to develop clean-slate ISA
  - Principal designers: Andrew Waterman, Yunsup Lee, David Patterson, Krste Asanovic
- Four years later, May 2014, released frozen base user spec
  - Many tapeouts and several research publications along the way
- Name RISC-V (pronounced "risk-five") represents fifth major Berkeley RISC ISA

RISC-V International

- **RISC-V** is a free and open Instruction Set Architecture (ISA)
- Frozen base user spec released in 2014, contributed, ratified, and openly published by RISC-V International

RISC-V International is a non-profit entity serving members and the industry

Our mission is to accelerate RISC-V adoption with shared benefit to the entire community of stakeholders.

- ✓ Drive progression of **ratified specs, compliance suite, and other technical deliverables**
- ✓ **Grow the overall ecosystem / membership**, promoting diversity while preventing fragmentation
- ✓ Deepen **community engagement and visibility**

RISC-V Ecosystem

Software

Open-source software:
Gcc, binutils, glibc, Linux, BSD, LLVM, QEMU, FreeRTOS, ZephyrOS, LiteOS, SylixOS, …

Commercial software:
Lauterbach, Segger, IAR, Greenhills, WindRiver, Micrium, ExpressLogic, Ashling, Imperas, AntMicro, …

RISC-V International

ISA specification | Golden Model | Compatibilty

Hardware

Open-source cores:
Rocket, BOOM, RI5CY, Ariane, PicoRV32, Piccolo, SCR1, Swerv, Hummingbird, WARP-V, XiangShan, BlackParrot, …

Commercial core providers:
Alibaba, Andes, Bluespec, Cloudbear, Cobham, Codasip, Cortus, Imagination, InCore, MIPS, Nuclei, Semidynamics, **SiFive**, StarFive, Syntacore, …

In-house cores:
Nvidia, WDC, Seagate, Huawei, Alibaba, …

Commercial silicon providers:
Alibaba, Bouffalo, EdgeQ, Esperanto, Espressif, Gigadevice, LeapFive, Microchip, Mythic, Renesas, Rivos, StarFive, TensTorrent, UntetherAI, Ventana, …

Why is RISC-V So Popular?

- Engineers sometimes “don’t see forest for the trees”
- Movement is not happening because some benchmark ran 10% faster, or some implementation was 30% lower power (though that might be true)
- Movement is happening because ***new business model*** changes everything
  - Pick ISA first, then pick vendor or build own core
  - Add your own extension without getting permission
  - Commercial, academic, and open-source ecosystems can coalesce around a single open standard

Comparing ISA Business Models

ISA	Chips?	Architecture License?	Commercial Core IP?	Add Own Instructions?	Open-Source Core IP?
x86	Yes, ***two*** vendors	***No***	***No***	***No***	***No***
ARM	Yes, ***many*** vendors	Yes, ***expensive and restrictive***	Yes, ***one*** vendor	***No*** *(Mostly)*	***No***
RISC-V	Yes, ***many*** vendors	Yes, ***free***	Yes, ***many*** vendors	***Yes***	***Yes, many*** *available*

RISC-V: The ISA that likes to say "Yes"

- RISC-V is inevitable
  - Industry wants this business model
- RISC-V will have the best processors
  - Inherently better ISA than others
  - More vendors competing for business with most cores designed
- RISC-V will have the best ecosystem
  - Largest number of players
  - All sockets become RISC-V over time
  - Software wants to run on the best processors available

Inherent Sustainable ISA advantage

- Simple base + modular extensions
  - Much lower area/power at equivalent performance levels over other ISAs
  - Technical reasons: reduced dynamic code size, compare+branch instruction, no complex instructions (e.g., two integer read ports max, no flags), …
- Clean-sheet design
  - Avoids legacy warts
- Designed for extensibility while retaining standard software stack
  - Variable-length ISA part of original design, supports vector extensions, and custom additions
- Community-designed
  - Input from leading experts in academia and industry
- No limit to performance levels or application domains – will surpass all other architectures quickly (2-3 years)

RISC-V and HPC, a Perfect Match

- RISC-V offers stable long-term platform for software
  - Global standard, not dependent on any single vendor
- RISC-V enables very efficient standard scalar processing
- RISC-V enables very efficient standard vector processing
- RISC-V enables tightly coupled custom extensions while maintaining compatibility with standard software
- **RISC-V enables hardware/software codesign**