



InspireSemi™

Breakthrough computing accelerator:
Blistering speed, versatility, energy
efficiency and affordability for HPC, AI
graph analytics, & blockchain applications

Investor Overview
September 2023

TSXV: INSP

© Inspire Semiconductor, Inc. – All Rights Reserved

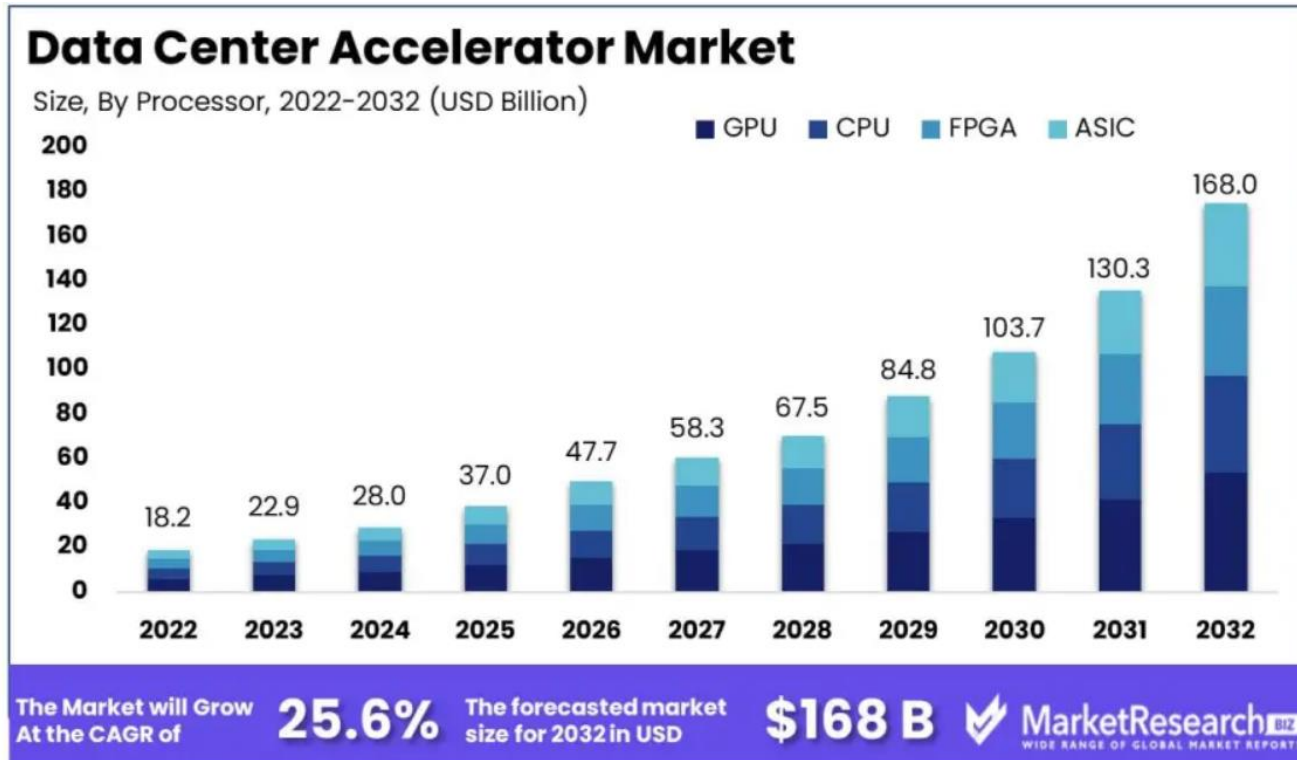
Disclaimer

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.

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

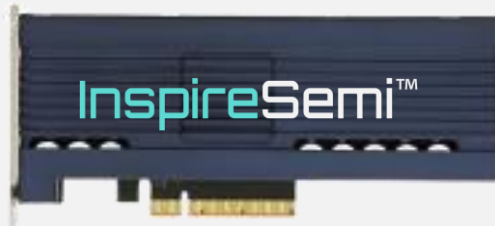


Diluted Shares Outstanding
280 Million



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



Argonne
NATIONAL LABORATORY



TACC
TEXAS ADVANCED COMPUTING CENTER



OAK RIDGE
National Laboratory

Partnerships



ASE GROUP



cādence®



RADIANT
SEMICONDUCTORS

codeplay®

Investment Opportunity in Breakthrough High Performance Computing (HPC)

Attractive Market

Addresses \$168B data center accelerated computing market with 25.6% CAGR
AI acceleration alone does not solve this problem

Leading Technology

Accelerated computing solution with >7,000 high performance 64-bit CPU cores / PCIe add-in card built on advanced RISC-V architecture. Scalable, fabless model with world class supply chain TSMC and ASE, 4nm access

Proven Team

Accomplished, hands-on semiconductor and system veterans with deep industry knowledge and operational experience

Disruptive Benefits

30-60% reduction in energy consumption (50 GLOPS/Watt); Low latency network is scalable to 256 chips and allows its use in real-time and real-safe applications where GPUs do not work (deterministic, predictable performance)

Strategic Partnerships

Strategic channels to be first mover at scale

Lenovo

PENGUIN COMPUTING
A subsidiary of SMART Global Holdings, Inc.

World Wide Technology

crsi

E4
COMPUTER ENGINEERING

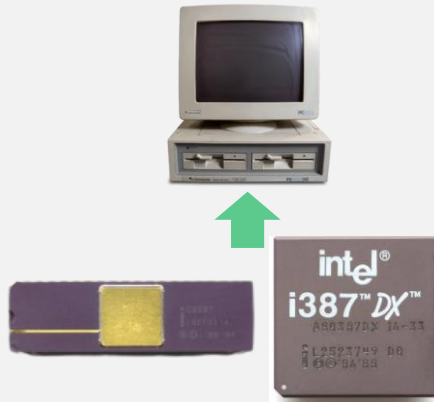
Defensibility

Upon integration, capital investment and scale-out becomes a barrier to entry due to high customer switching costs

The Third Wave of Accelerated Computing is Here

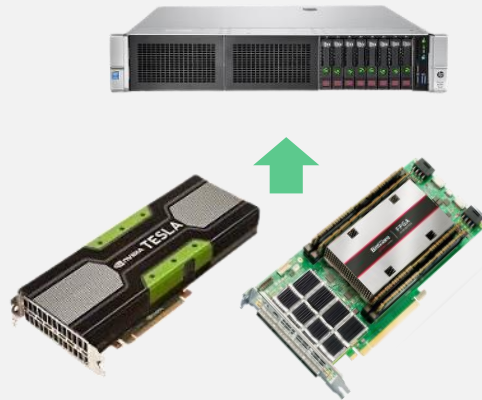
Thunderbird for HPC, AI, Graph Analytics

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

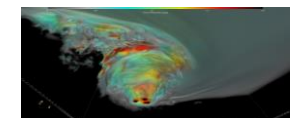
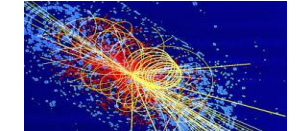
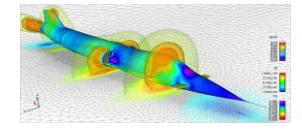
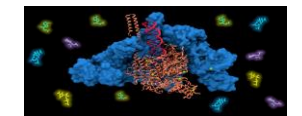
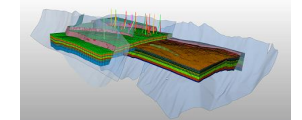
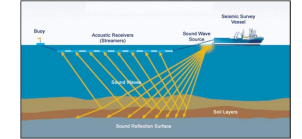


- Built for HPC
- Versatile & open software ecosystem
- Plugs into existing servers



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

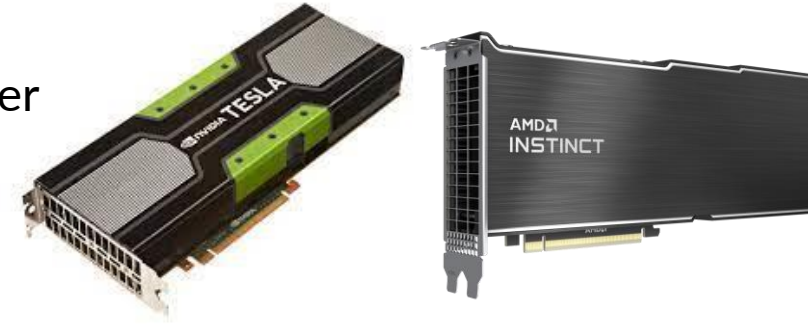
Genomics, Pharma, Life Sciences

CAE/Computational Fluid Dynamics

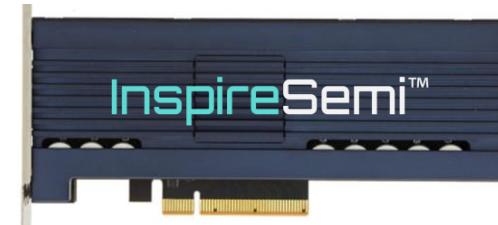
Nuclear Simulations Fission & Fusion

Climate & Weather Modeling

Datacenter GPUs



InspireSemi Thunderbird



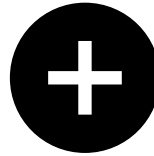
Highly differentiated “supercomputer-cluster-on-a-chip”

- Versatility as a platform across wide range of applications
- Each chip has 1,800 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 for Real-Time, Real-Safe Computing

Reproducibility

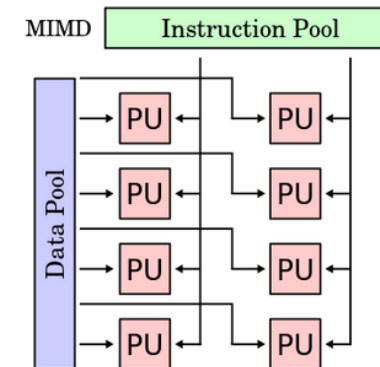
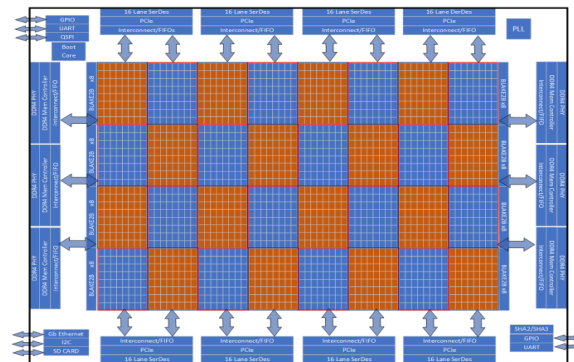
- Results verification
- Performance optimization
- Troubleshooting
- Quality assurance



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 Addresses ALL HPC & AI Customer Needs

	InspireSemi	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

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)



oneAPI



TensorFlow

PYTORCH

Glow



FREE RTOS



MLIR

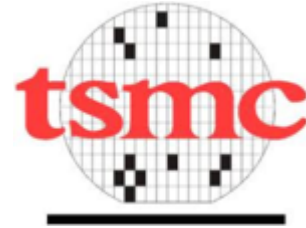


Zephyr



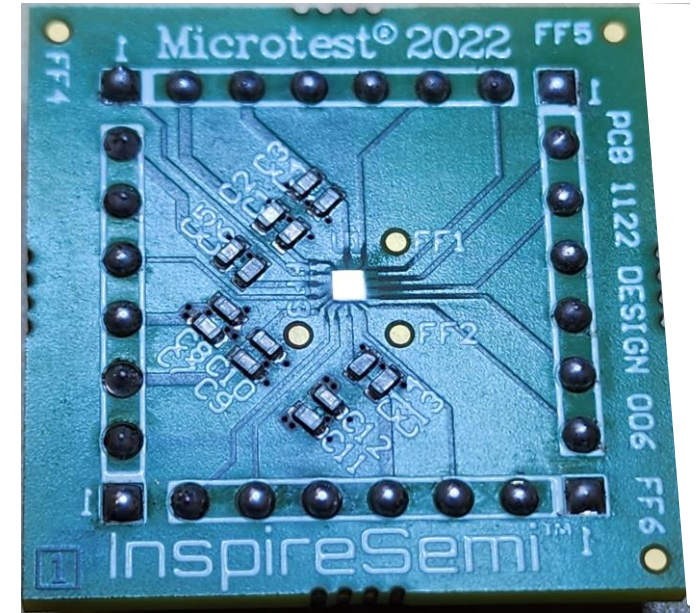
World Class Supply Chain Partners

- TSMC - Wafer fab
 - World's largest semiconductor foundry
 - Developing the most advanced process nodes
- ASE – Chip package & test
 - World's largest and highest quality OSAT (Outsourced Semiconductor Assembly and Test)
 - Leading edge package design
- Imec: Value Chain Aggregator (VCA)
 - Enable early access to tier-1 supply chain
 - Support engineering and early-prod volumes



Fundamental Capabilities Proven with TSMC 5nm Test Chip

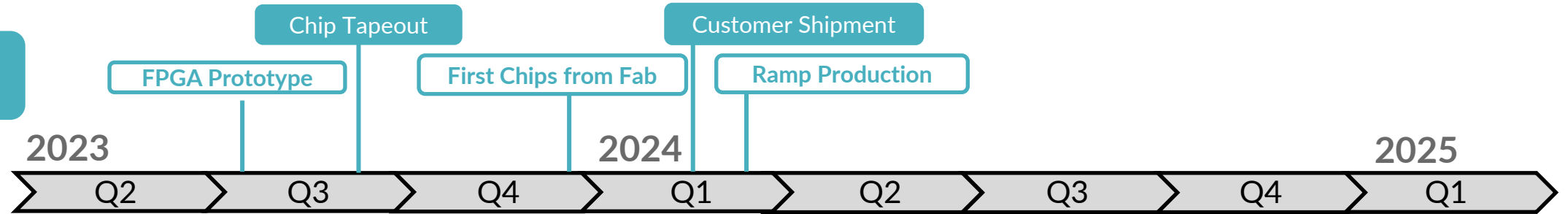
- De-risks plan for follow-on 4nm Thunderbird II
 - Validated team's ability to deliver designs on leading-edge TSMC process node
 - Worked first time, met performance and power targets
 - Including full-custom layout optimized at every level
 - Something not many companies can do, perhaps none this 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
12nm

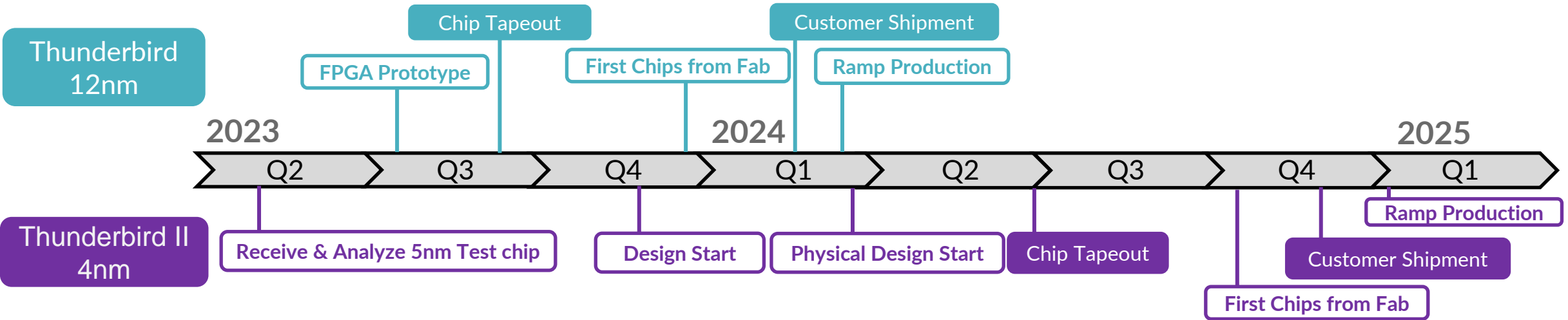


Thunderbird

- 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 Thunderbird 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

Early Adopter Program Signals Strong Market Traction

4 Committed Contracts

\$50,000 to participate

Early Customer Interest



Decision Maker Responses:

"Consider us your go-to-market partner. Tell all prospects you talk to."

"The combination of your custom designed RISC-V 'sea of cores' plus high-speed interconnect fabric is very smart, and your decision to focus on HPC (and blockchain opportunistically) rather than AI / ML was likewise very smart."

"You guys are a startup in the right place at the right time."

"This will let us run large memory simulations with AI workflows. We will take anything you can give sooner. We will provide our codes and datasets."

"We still run 7-year old Xeon Phi systems and begged Intel not to kill it, but they did. We thought we would have to buy 8-GPU systems, but really want your Thunderbird."

"We are excited to work with your RISC-V based compute accelerator on our key benchmark codes and important applications that are used across multiple DOE laboratories."

Accomplished Leadership Team



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
- 5 time CEO – 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



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



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)



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



John B. Kennedy, CFO

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



Investment Opportunity in Breakthrough High Performance Computing (HPC)

Attractive Market

Addresses \$168B data center accelerated computing market with 25.6% CAGR
AI acceleration alone does not solve this problem

Leading Technology

Accelerated computing solution with >7,000 high performance 64-bit CPU cores / PCIe add-in card built on advanced RISC-V architecture. Scalable, fabless model with world class supply chain TSMC and ASE, 4nm access

Proven Team

Accomplished, hands-on semiconductor and system veterans with deep industry knowledge and operational experience

Disruptive Benefits

30-60% reduction in energy consumption (50 GLOPS/Watt); Low latency network is scalable to 256 chips and allows its use in real-time and real-safe applications where GPUs do not work (deterministic, predictable performance)

Strategic Partnerships

Strategic channels to be first mover at scale

Lenovo

**PENGUIN
COMPUTING**
A subsidiary of SMART Global Holdings, Inc.

World Wide Technology

crsi

E4
COMPUTER
ENGINEERING

Defensibility

Upon integration, capital investment and scale-out becomes a barrier to entry due to high customer switching costs

InspireSemi™

Breakthrough computing accelerator:
Blistering speed, versatility, energy
efficiency and affordability for HPC, AI
graph analytics, & blockchain applications

Thank you!

© Inspire Semiconductor, Inc. – All Rights Reserved

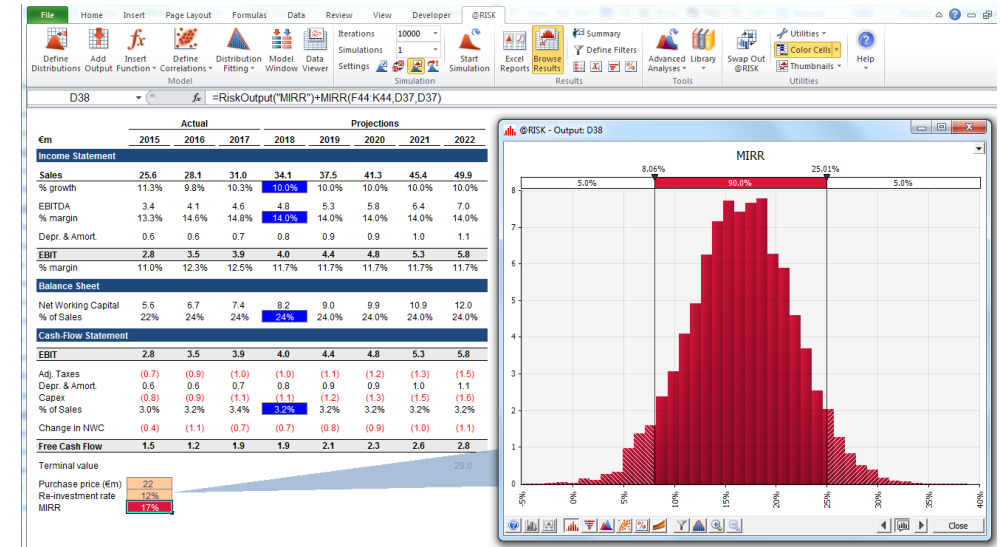
Information Contacts:

Inspire Semiconductor Holdings Inc.
John Kennedy
Chief Financial Officer
jkennedy@inspiresemi.com

KCSA Strategic Communication
Investor Relations
Phil Carlson/Scott Eckstein
inspiresemi@kcsa.com

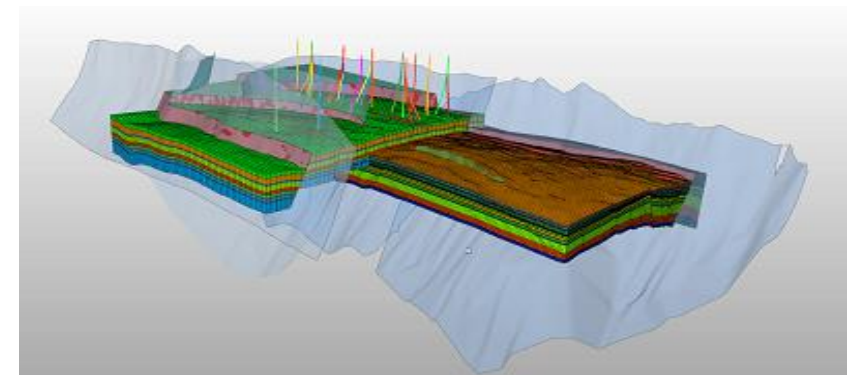
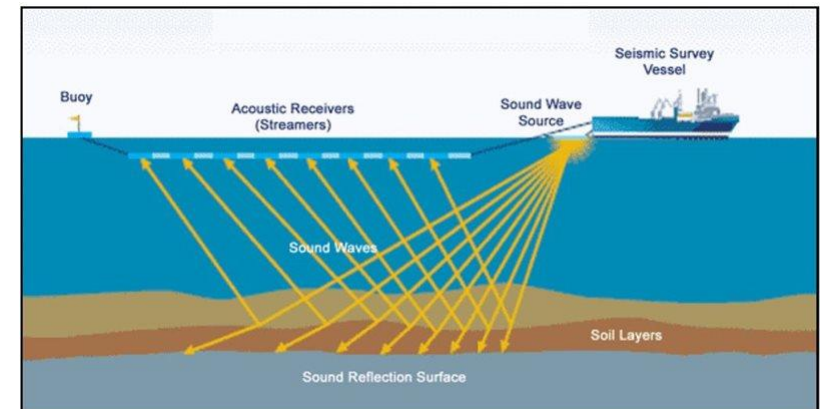
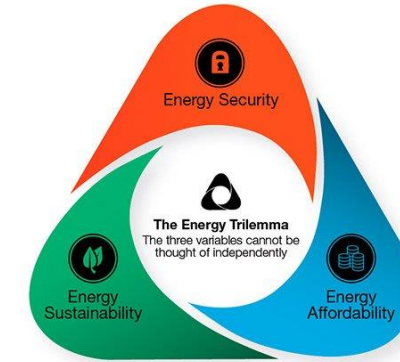
Finance Industry Opportunities

- HFT (High Frequency Trading)
 - Thunderbird can support this because of deterministic, predictable performance (GPUs cannot)
- Crypto currencies
 - Thunderbird excellent for alt-coin mining
- Large Monte Carlo Simulations have been primary workload
 - Used in corporate finance, portfolio management, options pricing....
 - Combines Monte Carlo random statistical method with simulation to give a powerful tool to obtain a distribution (array) of results for any statistical problem with numerous inputs sampled over and over again
 - Much more useful than historical approach that is just one possibility which may or may not be applicable in future
 - But cannot account for bear markets, recessions, financial crises...
 - GPU's effective, but so too will be Thunderbird
- Fraud detection and prevention a huge opportunity
 - Shell companies that hide fraudulent transactions
 - Rapid increase of electronic payments and eCommerce make this a huge exposure for banks and credit card issuers
 - Was an early AI/ML workload that had some good success
 - Graph analytics proving to be more effective and powerful since not just transactional history
 - Represents nodes with edge connections and vectors to traverse massive amount of data and find unseen relationships
 - Customer and ISV feedback - Thunderbird looks purpose-designed for graph analytics since a heavy CPU workload across massive databases

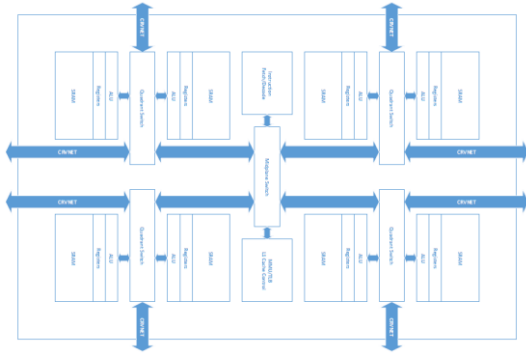


Energy Industry Opportunities

- The new imperative in the energy industry is solving the energy trilemma
 - Energy security, affordability, and sustainability
 - Energy company venture arms investing in tech companies to accelerate
- Major large-scale projects and investments in CCUS underway by all major energy companies to hit the Net Zero Scenario
 - Carbon Capture, Utilization, and Storage (aka Sequestration)
 - Process that captures CO₂ emissions from sources like coal-fired power plants and stores it underground in geologic formations so it will not enter the atmosphere
 - Project developers have announced ambitions for over 200 new capture facilities to be operating by 2030, capturing over 220 Mt CO₂ per year
- The two major HPC workloads in the energy industry continue to be their traditional upstream Oil & Gas work
 - Seismic
 - Locate and estimate O&G reserves; GPU's effective, but so too will be Thunderbird
 - **Reservoir Modeling & Simulation**
 - Predict field performance and compare economics of different recovery methods
 - **Thunderbird CPU approach perfect fit (GPUs not effective)**
 - Renewable energy is a small fraction of compute and is a mix of HPC + AI/ML
- Reservoir Modeling & Sim key HPC technology used for CCUS
 - Must ensure CO₂ pumped underground does not leak back out
 - → HPC a must since cannot afford usual industry failure rate of 70%
- Adjacent energy industry opps
 - Energy grid management and natural disaster planning
 - Energy trading



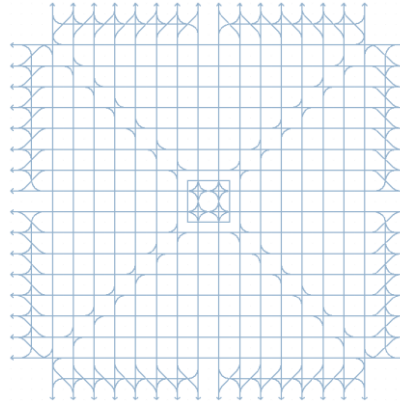
Thunderbird Technology Overview



Core

64-bit superscalar RISC-V CPU cores:

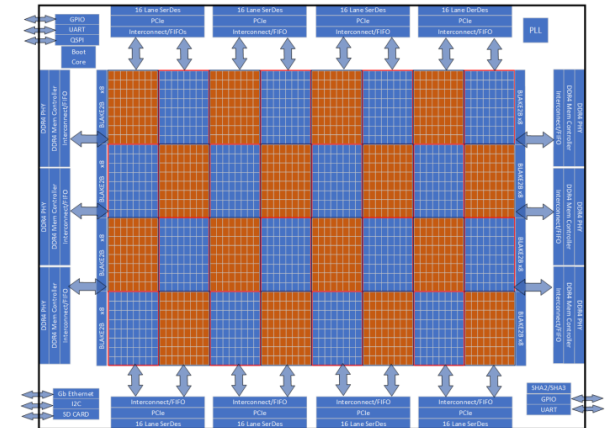
- Custom InspireSemi hi-perf design
- Multiple-issue, out-of-order, variable instruction width
- Vector, SIMD and tensor
- Mixed-precision floating point
- AI and cryptography extensions
- Tightly integrated memory and core-core network fabric
- Simple programming model
- Thriving software ecosystem



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 to arrays up to a million cores
- Six DDR4 memory controllers
- 64 lanes PCIe / Eth / chip-to-chip
- Algorithm-specific accelerators

System

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