



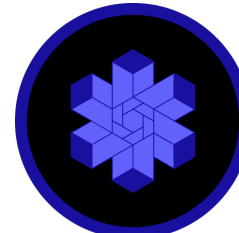
Converged RISC-V AI IP



In Order
Core



OOO
Core



OOO
Vector
Unit



Tensor
Unit

About Semidynamics



Semidynamics, founded in 2016, is a **100% European** supplier of RISC-V IP cores, HQ in **Barcelona**, specializing in **customization** of **high bandwidth high performance AI cores** for **tailored projects**

Experts in customizable AI IP

Market Trends & Challenges

Trends

- More data — Sensors upstream to cloud
- Edge Servers must process massive amount of data locally
- More AI - generative AI and LLM apps handle trillions of parameters
- New AI models emerge fast



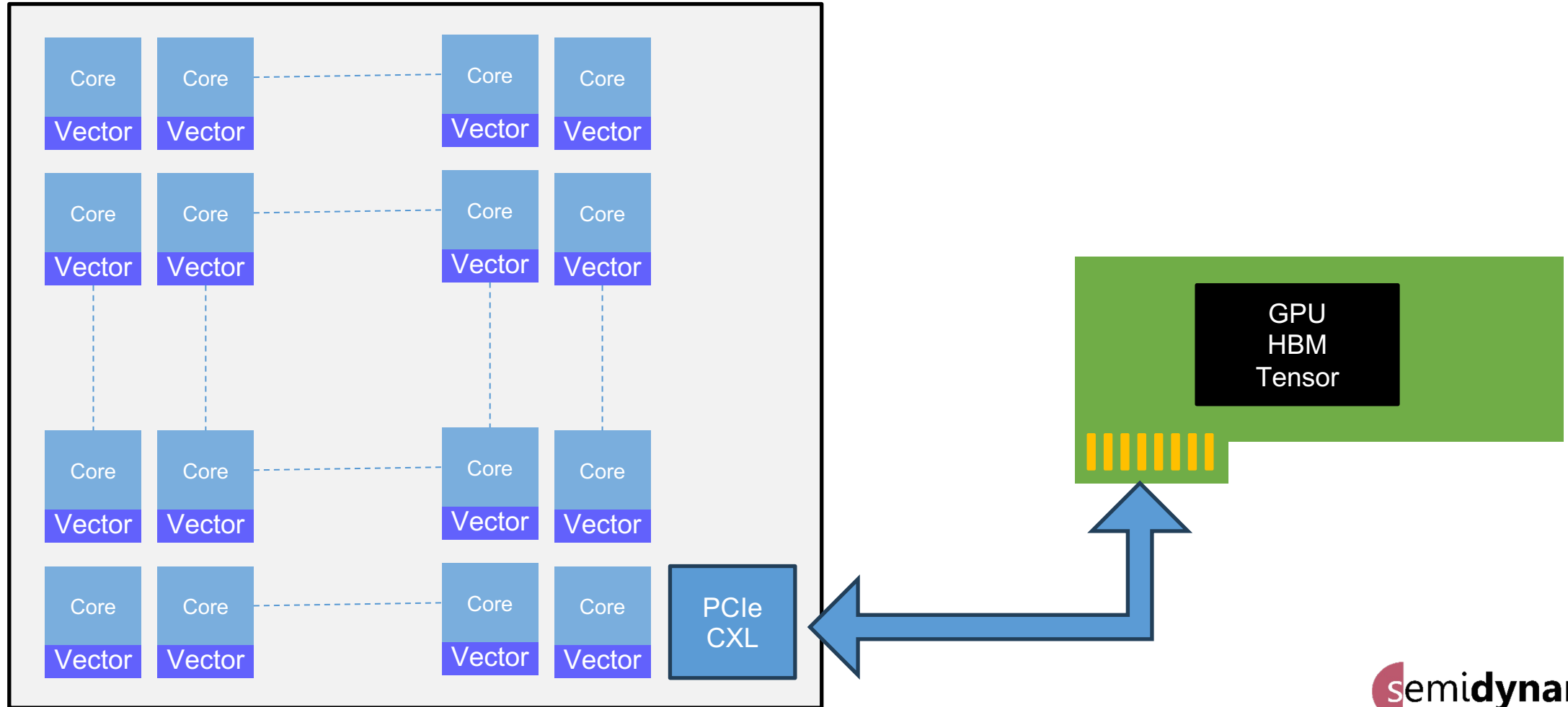
Challenges

- Increasing processing performance needs for AI workloads , increasing power consumption
- Huge amount of stored data increases likelihood of cache misses
- Increasing CPU performance needs
- Hypervisors & Containers needed for several guest OSes and domains
- Supply Chain issue: GPU cards availability becomes an issue due to ‘Nvidia Hype’

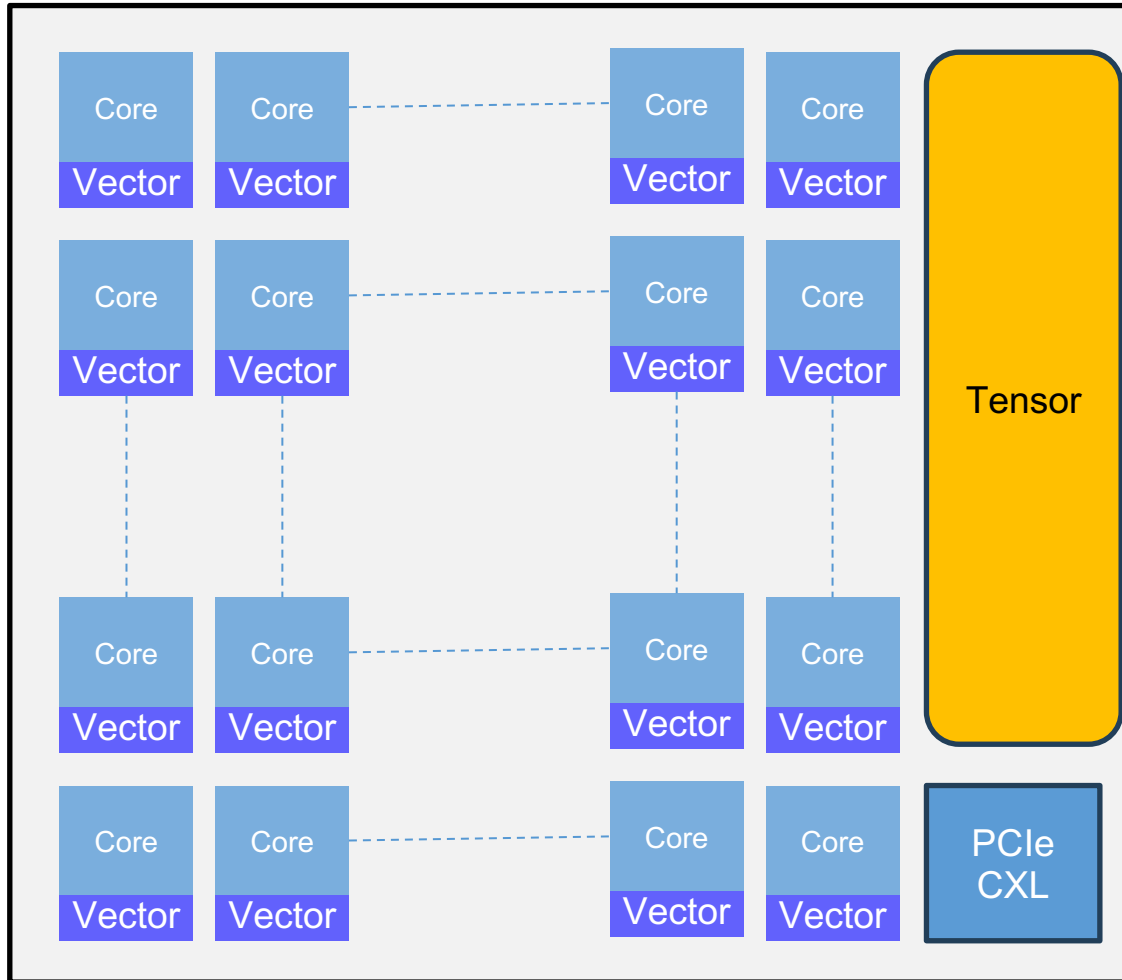
New SOC compute paradigm is required to address challenges

Classic HPC / DataCenter

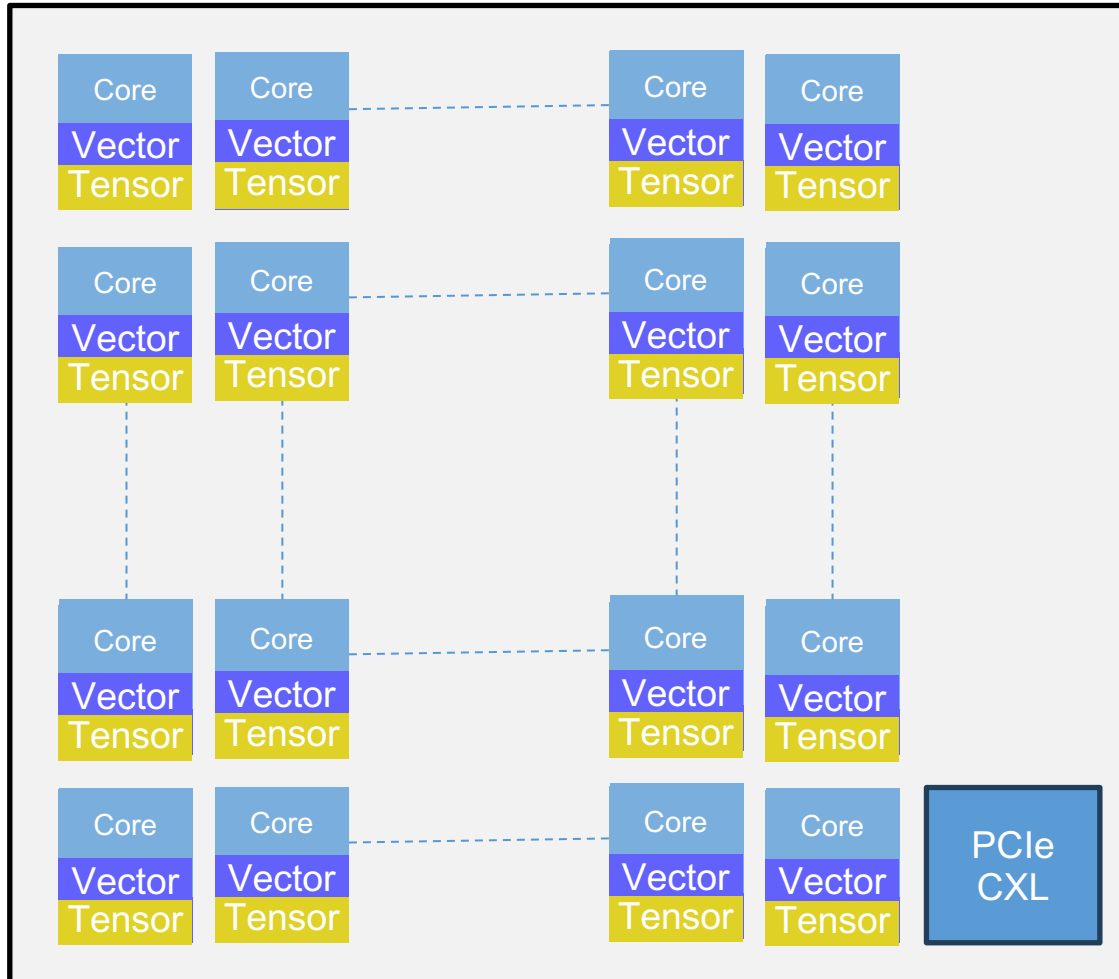
Many Cores + Vector, Tensor using GPU



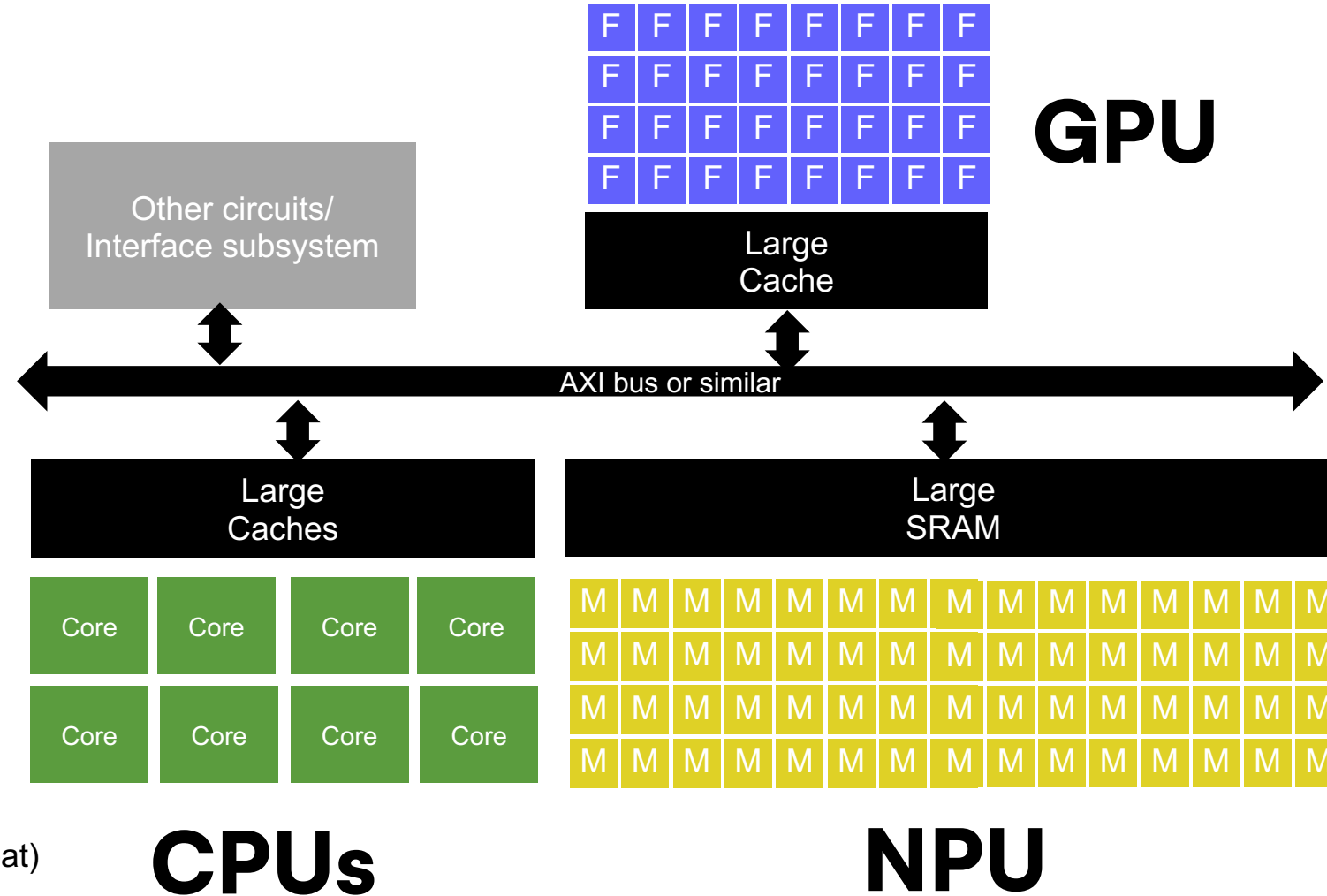
Some HPC / DataCenter proposals Bring Tensor on Silicon, as separate IP Unit



Semidynamics HPC / DataCenter proposal: Combined RISC-V CPU, Vector, Tensor

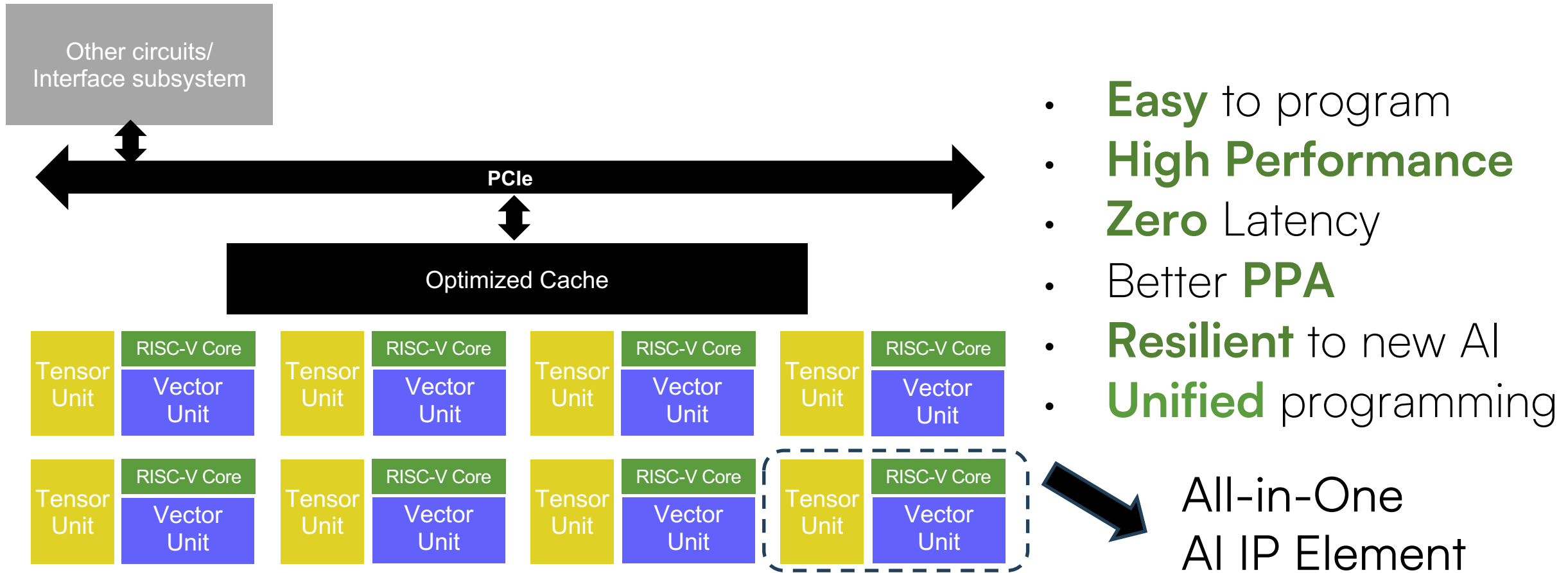


Typical AI-focused Subsystem SOC to date...



- **Hard** to program
- **High** Latency
- **Non-optimal** PPA due to caches and local storage needs
- **Obsolescence** by new AI algorithms
- **3 ISA structure** pose high change risk

Our vision: Fusing CPU, GPU, and NPU



Our IP



Atrevido 勇敢な
无所畏惧 Fearless 용감한

64b out-of-order CPU

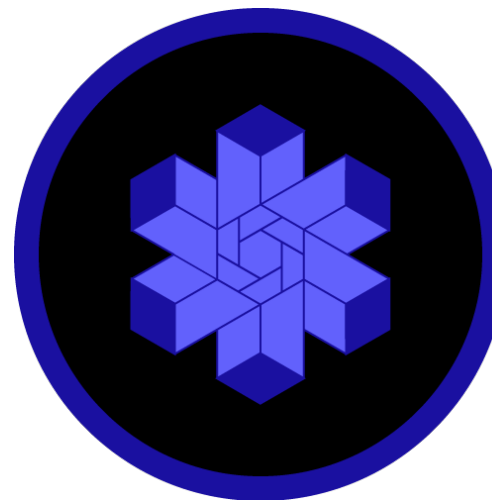
RISC-V
AXI and CHI



Avispado 賢い
机智 Smart 똑똑한

64b in-order CPU

RISC-V
AXI and CHI



Vector Unit

RVV1.0
Out-of-order



Tensor Unit

BF16, FP16, INT8

The Semidynamics Proposal

- Powerful **Out Of Order** based on Risc-V
- Combine **CPU** with **Vector** and **Tensor unit** to create powerful AI capable Compute building blocks
- Enable Hypervisor Support for Containerization
- Enable Crypto for Security / Privacy
- Easy to combine with custom logic / Unit — 3 custom instructions
- Use of **Gazzillion™ Technology** to efficiently manage large data sets

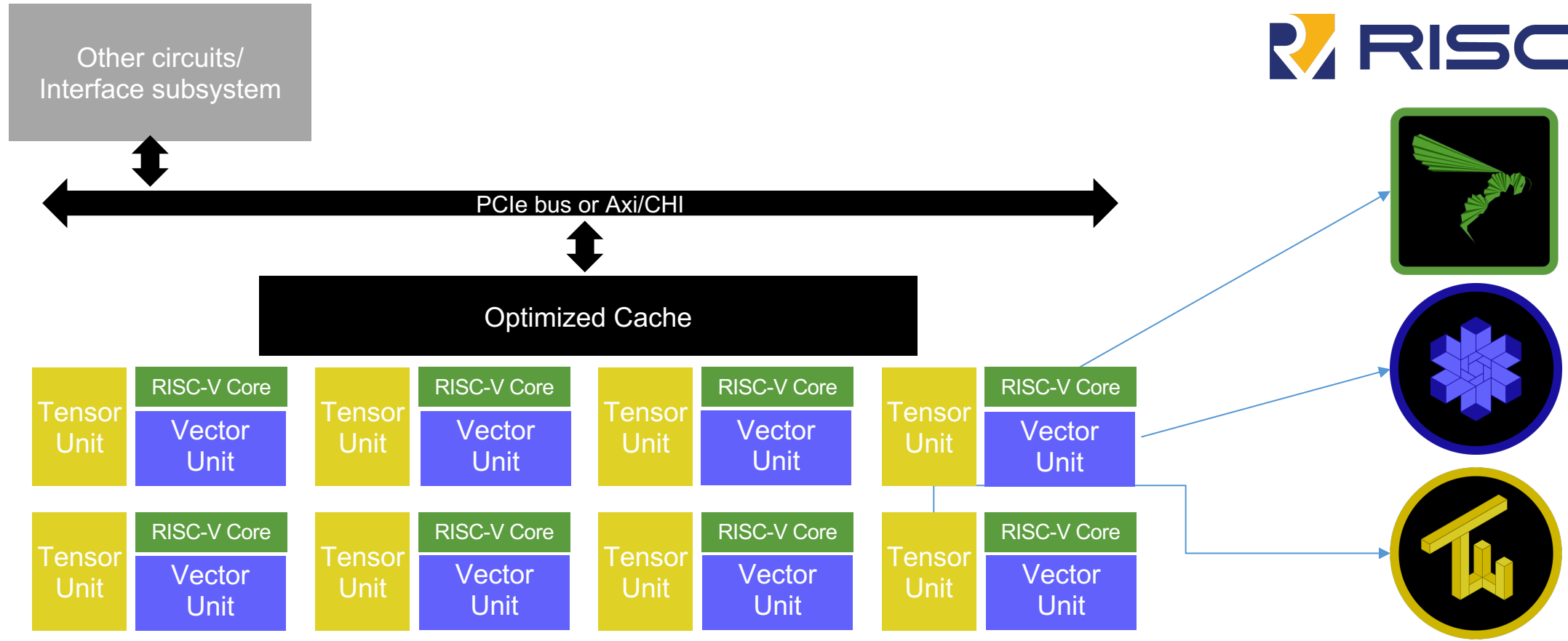


Benefits

- **Easy** to program
- **High Performance** for Parallel Codes
- **Zero** Communication Latency



Vision Delivered with unified ISA



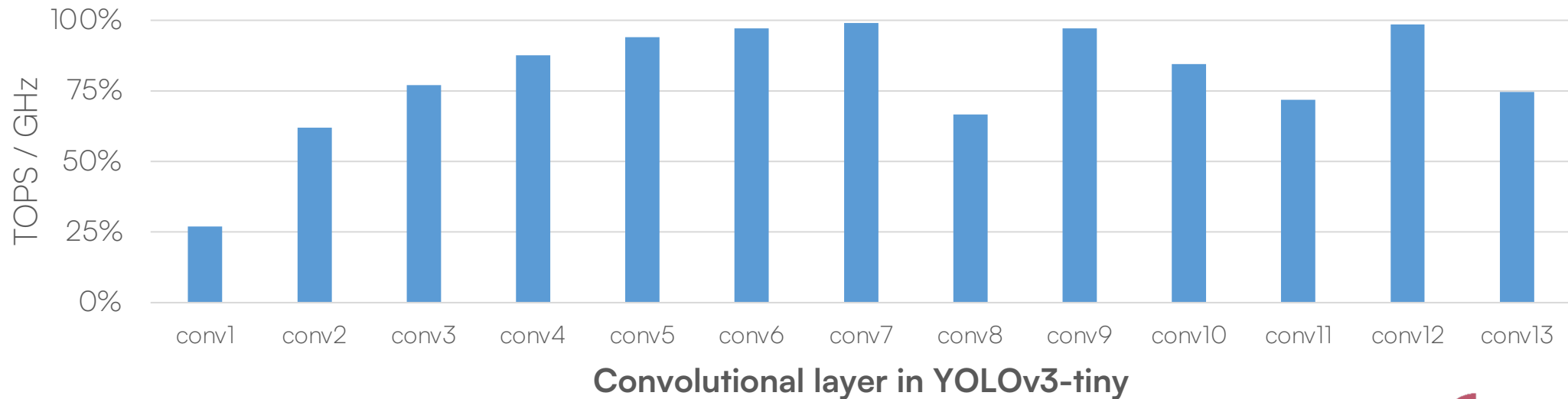
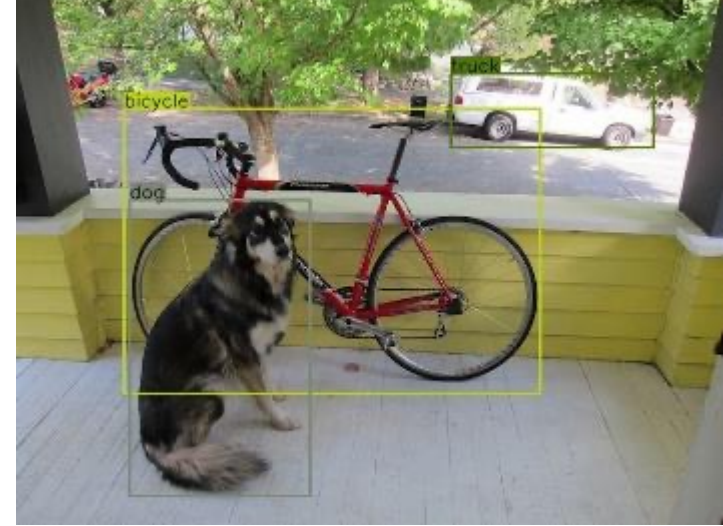
Each Vector Unit : from 4 to 32 FMAC units

Each Tensor Unit : from $\frac{1}{4}$ TOPS₈ to 2 TOPS₈

Gazzillion Technology to enable sustained DRAM access beyond 50 bytes/cycle

YOLO on our fused IP: **33 FPS**

- Performance at 1GHz
 - ATV4+Vector Unit + Tensor Unit (bf16): 33.03 FPS
 - Real-time performance with one Tensor Unit





Business Model

Flexible and customizable Business Model

Customize IP

- AXI, CHI
- Cache Sizes
- Branch predictor
- Custom instructions
- RV32
- Small Core...
- Functional Safety and Cyber Security



Evaluate

- Single Core
- Multi Core
- Vector Unit
- Tensor Unit



License

- License Fee
- Royalties



Maintenance

- Bug Fixes
- Timing fixes
- Area Fixes



Summary

- Market trend: most new SOC's from data center appliances to edge devices need AI & software flexibility
- AI workloads continuously change and require more performance
- Current designs are sub-optimal and non-resilient to AI changes
- **Our solution: fused CPU + GPU + NPU compute core**
- Available NOW

Let's build the AI future for together

Thank you