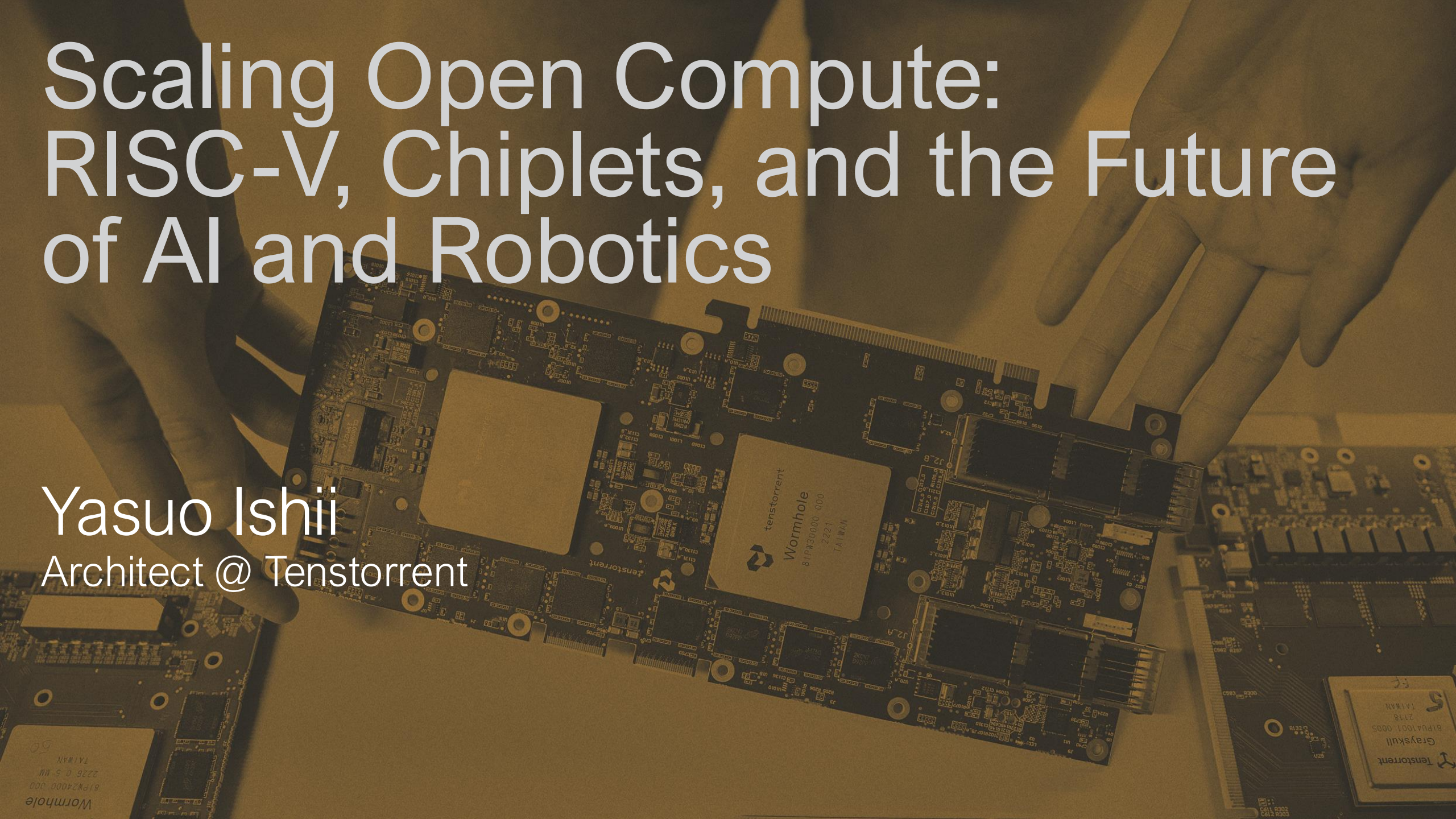# Scaling Open Compute: RISC-V, Chiplets, and the Future of AI and Robotics

Yasuo Ishii
Architect @ Tenstorrent

Golden Age of Silicon Diversity

- Silicon Heterogeneity
- Architecture Variety
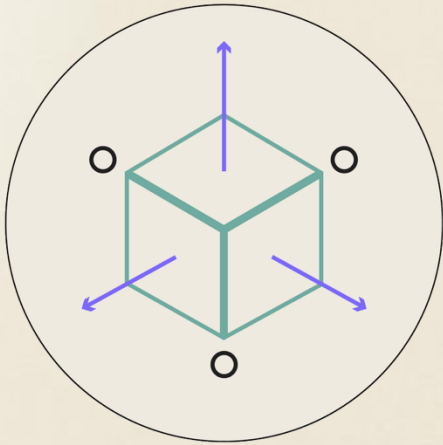- Multi-vendor Silicon Landscape
- Design Composition
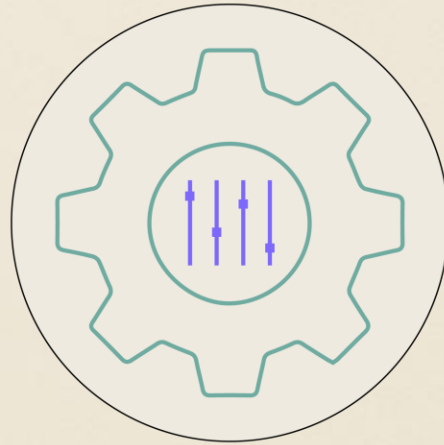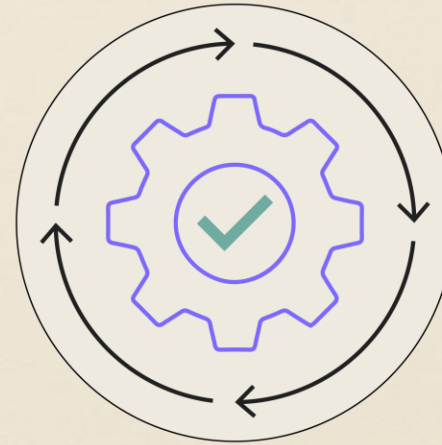

AI Personalization

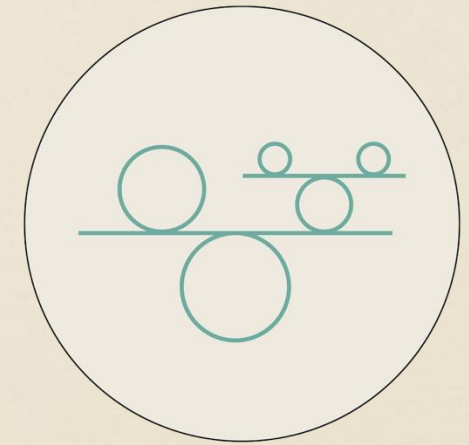## Ubiquitous AI Compute

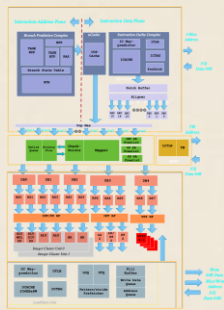# Open Source Hardware

Scalable

Extensible

Efficient

Stable

Foundational Compute IP's

CPU

Tensix NEO

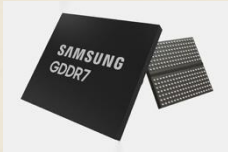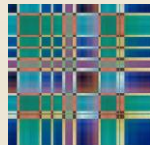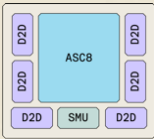**SOC Technology**
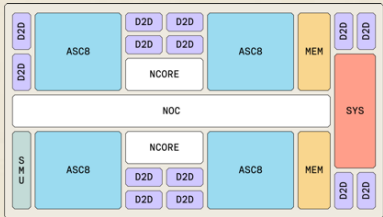
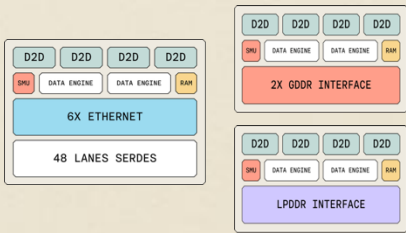PCIe   Networking   Memory   Fabric   Power/Thermal Management

RAS

Security

Debug

Chiplet Technology

CPU, AI, IO, Memory chiplets
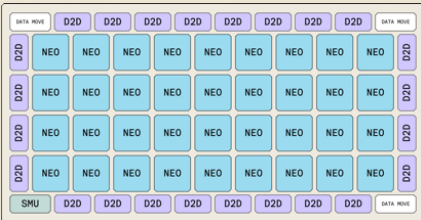
Applications
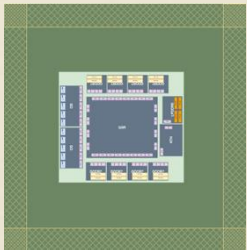
Data Center

Tenstorrent

4
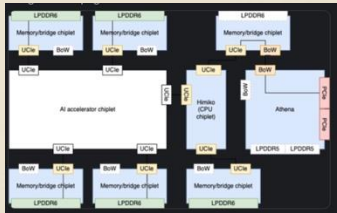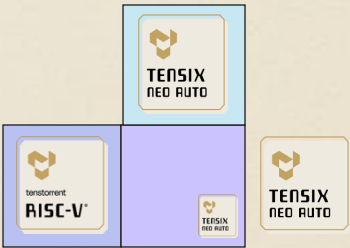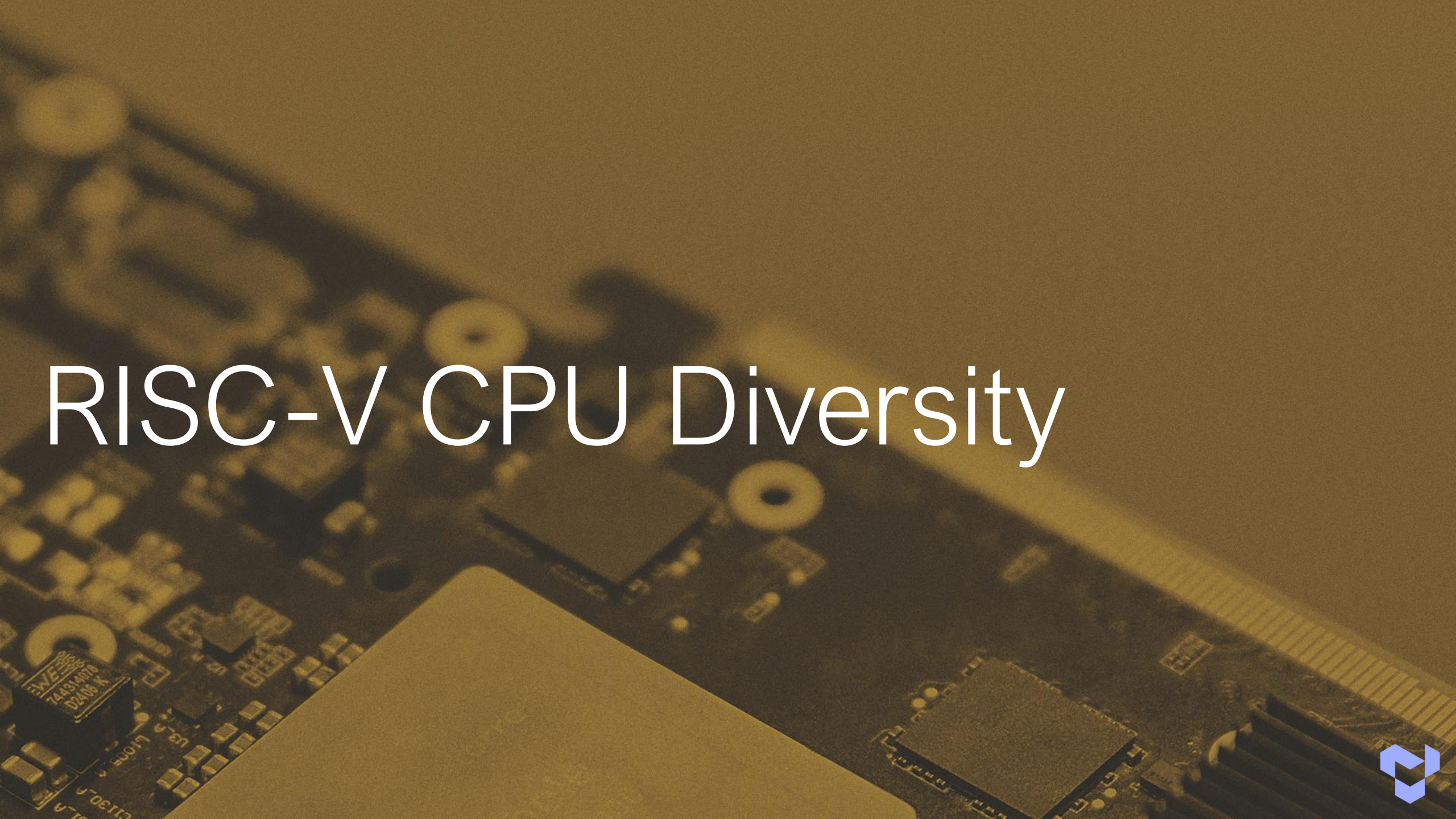
RISC-V CPU Diversity

RISC-V in the Cross-road

- RISC-V's 15-year Journey
  - An academic project in UC Berkeley
  - Mainstream adoption in embedded and industrial applications

- Sustain momentum
  - High-performance computing implementation
  - Software development

- Key Requirements
  - Advanced O-o-O CPU design with high-bandwidth memory
  - Optimized compilers, firmware, and system software

- Without high-end investment, RISC-V confined to low-cost, niche markets
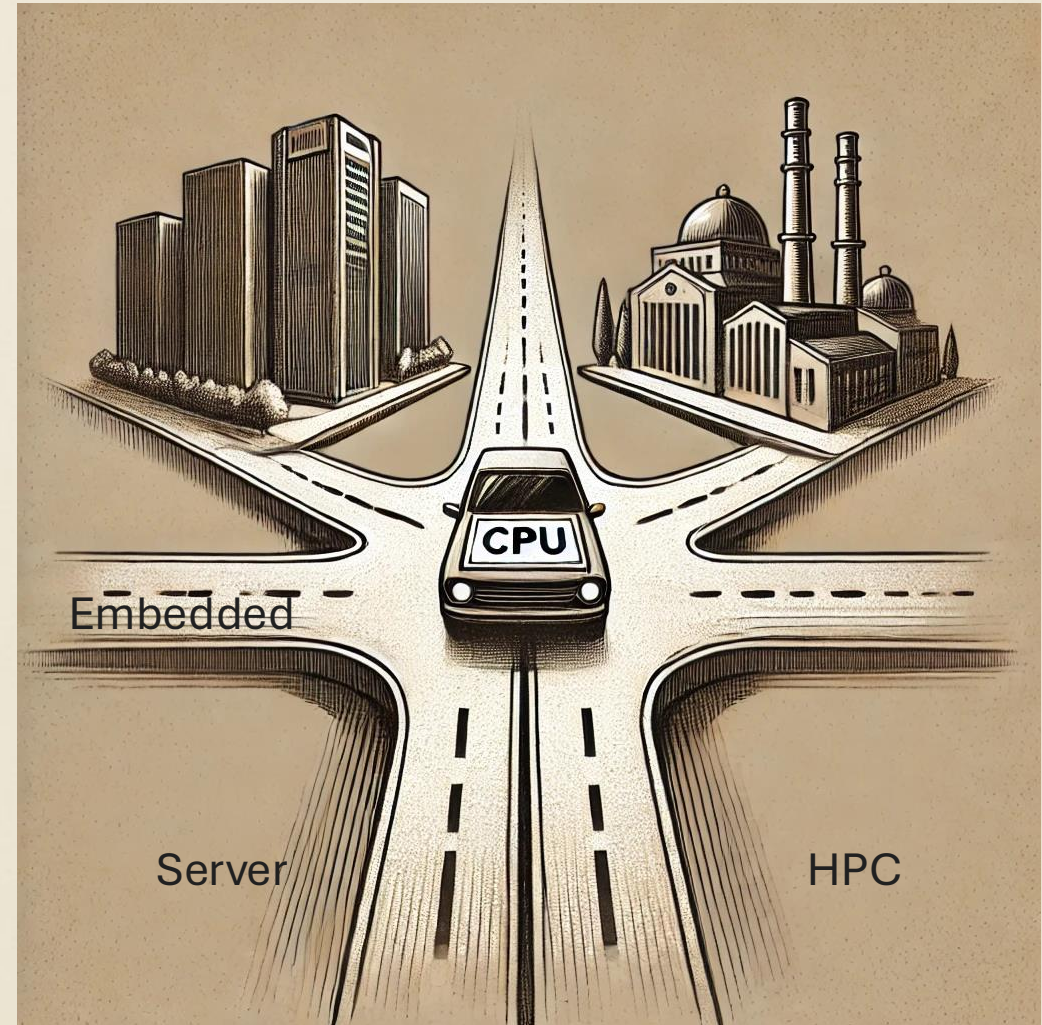
RISC-V CPU - Ascalon

- Disruptive high-performance RISC-V processor for AI and server

- 20+ SPECINT/Ghz

RVA-23

- Advanced branch predictions
- 8-wide decode
- 3 LD/ST with large load/store queues
- 6 ALU/2 BR
- 2 256-bit vector units
- 2 FPU units
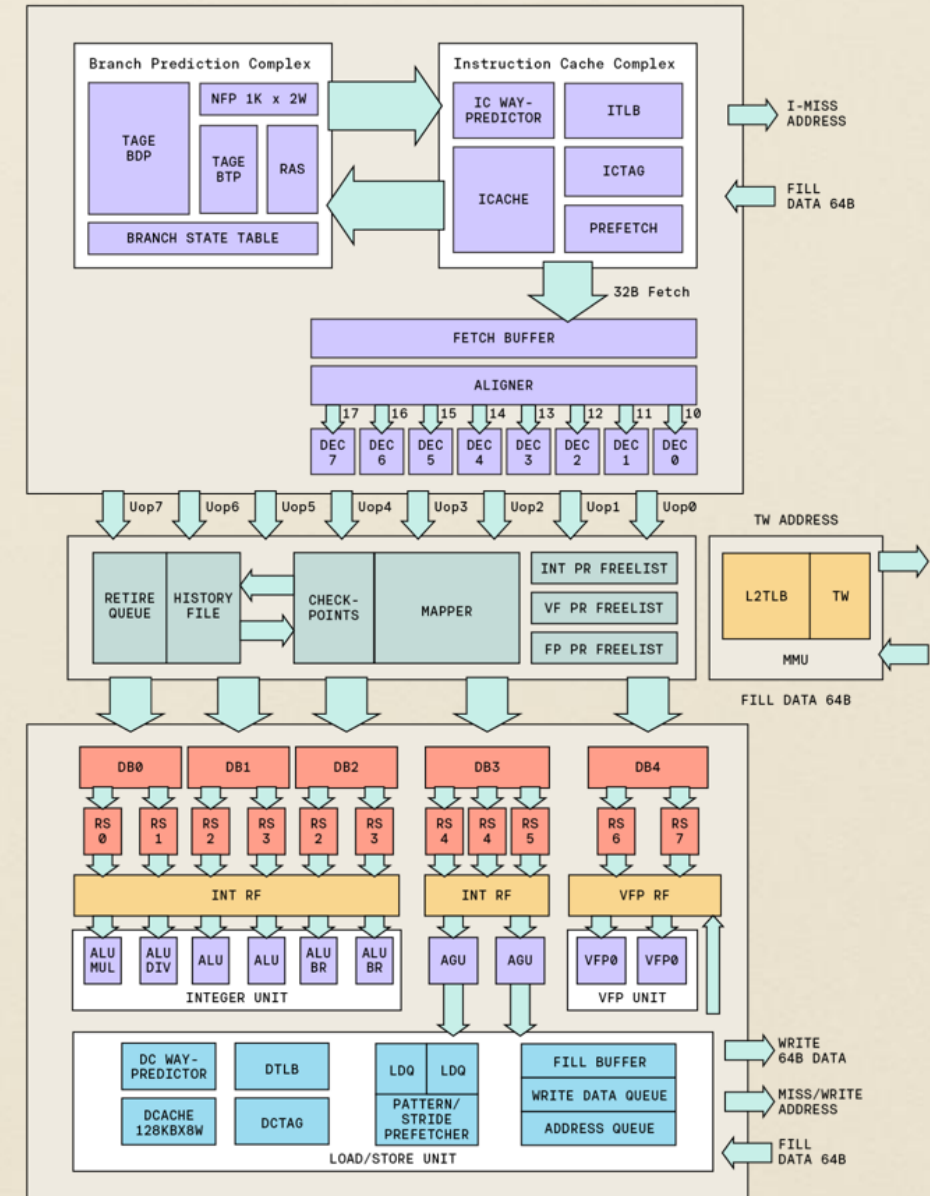
Ascalon Cluster

- Samsung SF4

- Performance correlated 20+ SPEC2k6INT/GHz

Ascalon Core

Core 4

Core 5

Core 6

Core 7

12M Share L2

Core 1

Core 2

Core 3
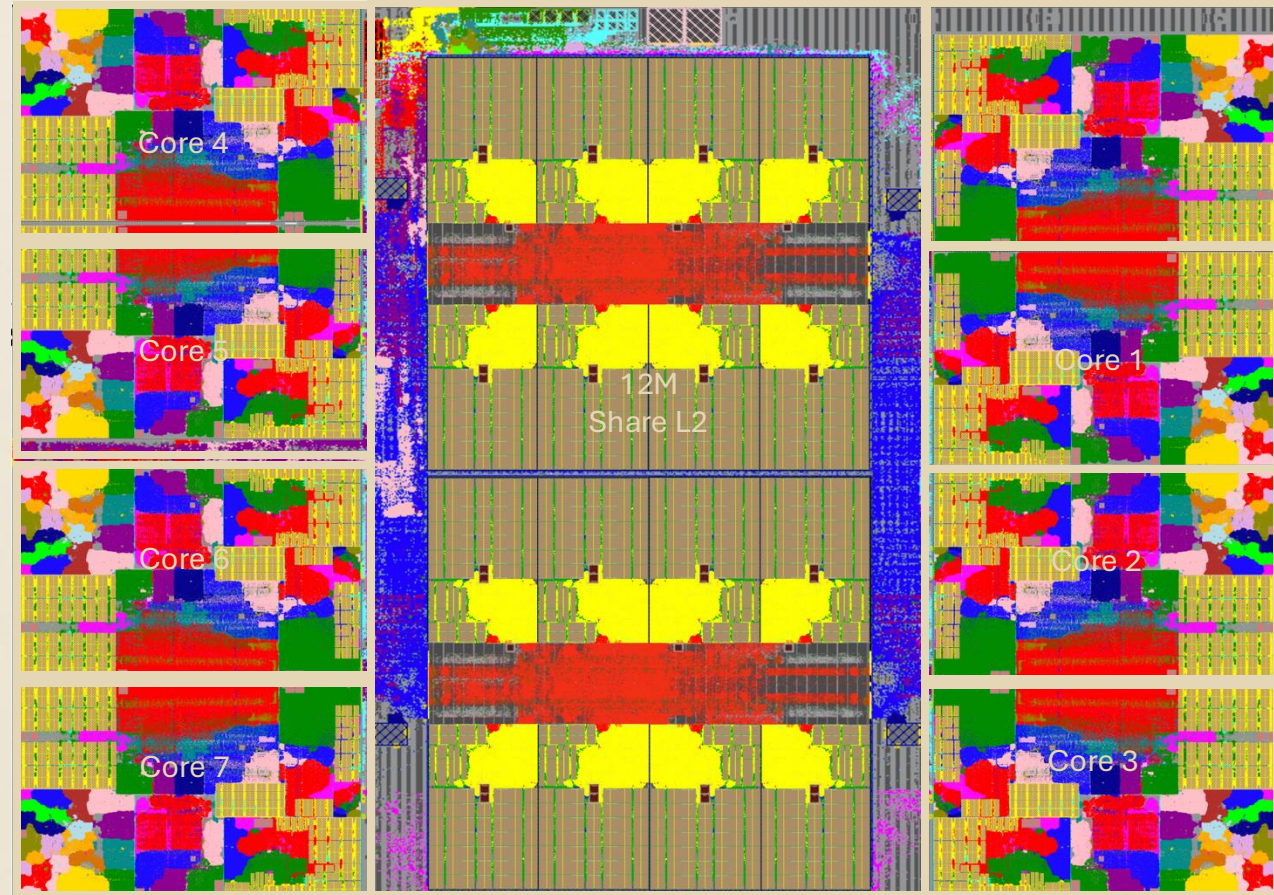
8 Ascalon Cluster
12M Shared L2

CPU Chiplet (Yayoi Project)

- Rapidus 2nm Process

- 8-core Ascalon CPU cluster

  - Interrupt Controller, Debug, etc…

- External Interface: UCIE

  - DRAM, PCIE, AI Accelerators

- System Management

  - Open Chiplet Atlas Architecture

Ascalon-Auto IP

- ISO26262 Functional Safety Features
  - Dual Core Lock Stepped with Time Disparity
  - Coherent & Non-Coherent Bus Protection with CRC or ECC
  - ECC Protection for L1 & L2 Cache
  - AXI Interface Parity Protection
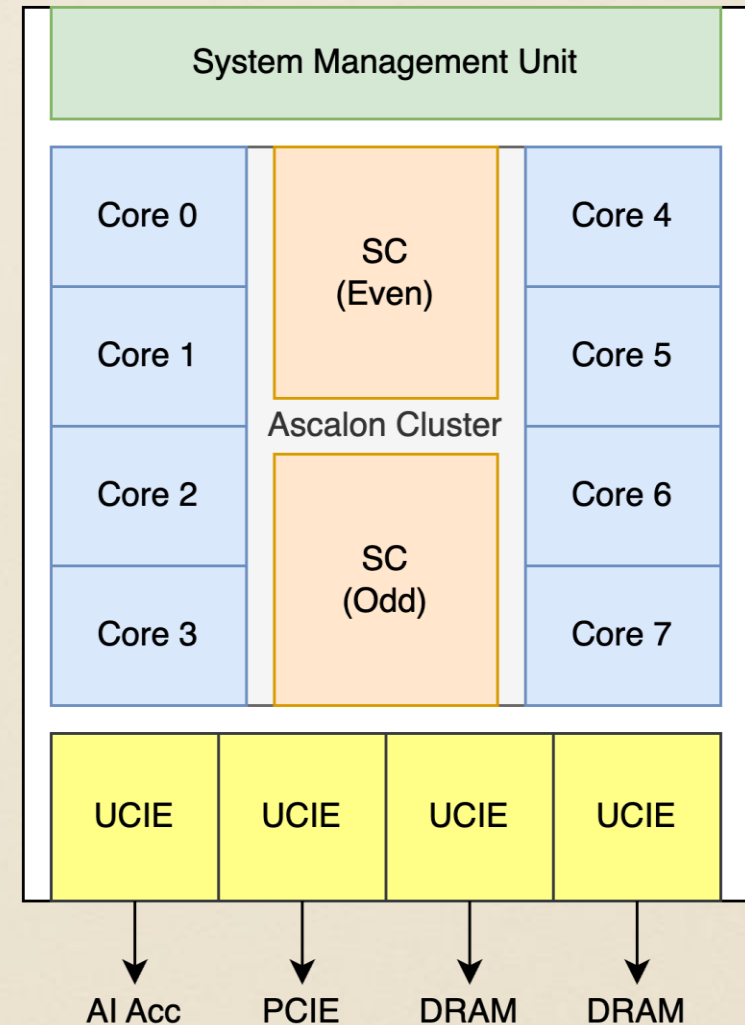 - RAS & Fault Controller
  - Safety Bus for Debug
  - Software Test Library Support

  - SPECInt2K17@2.25GHZ up to 30 per cluster
  - Dhrystone MIPS : 17.4 DMIPS Rate/ MHZ
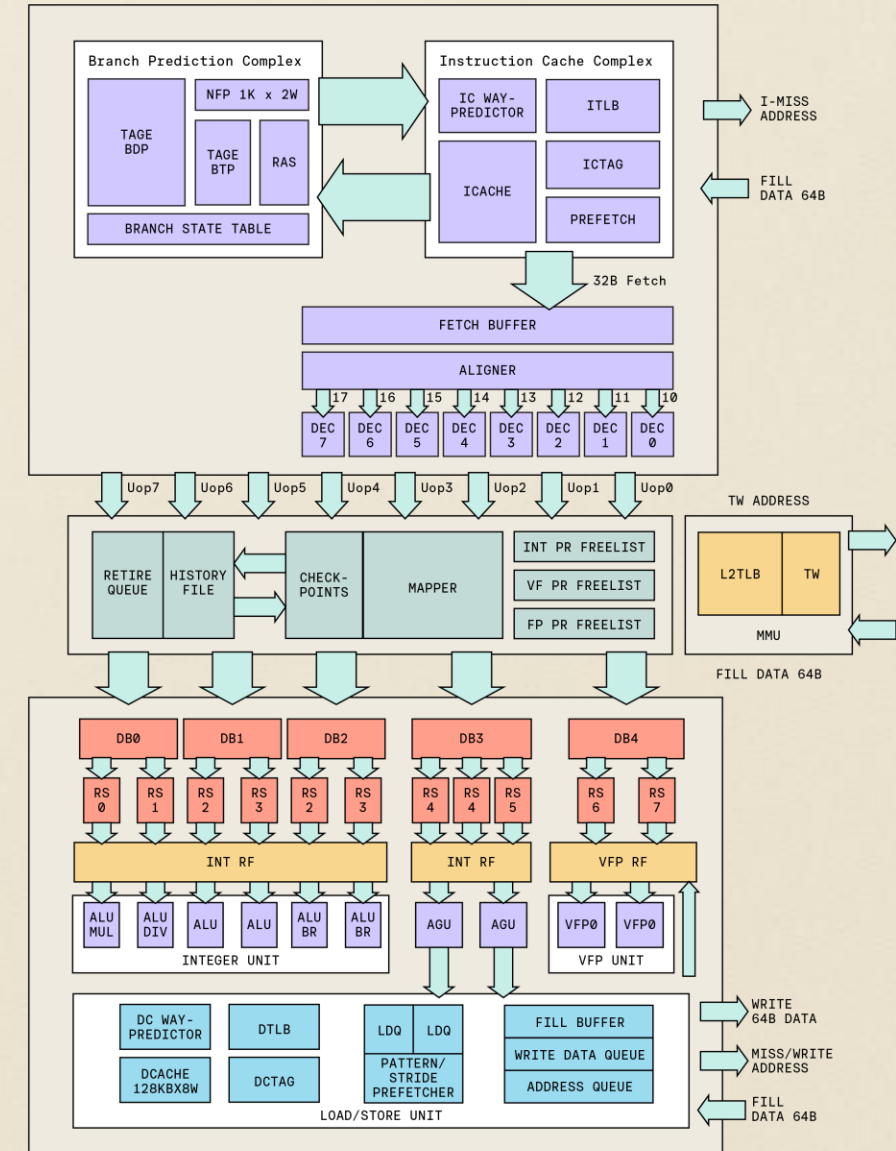
RISC-V CPU in AI
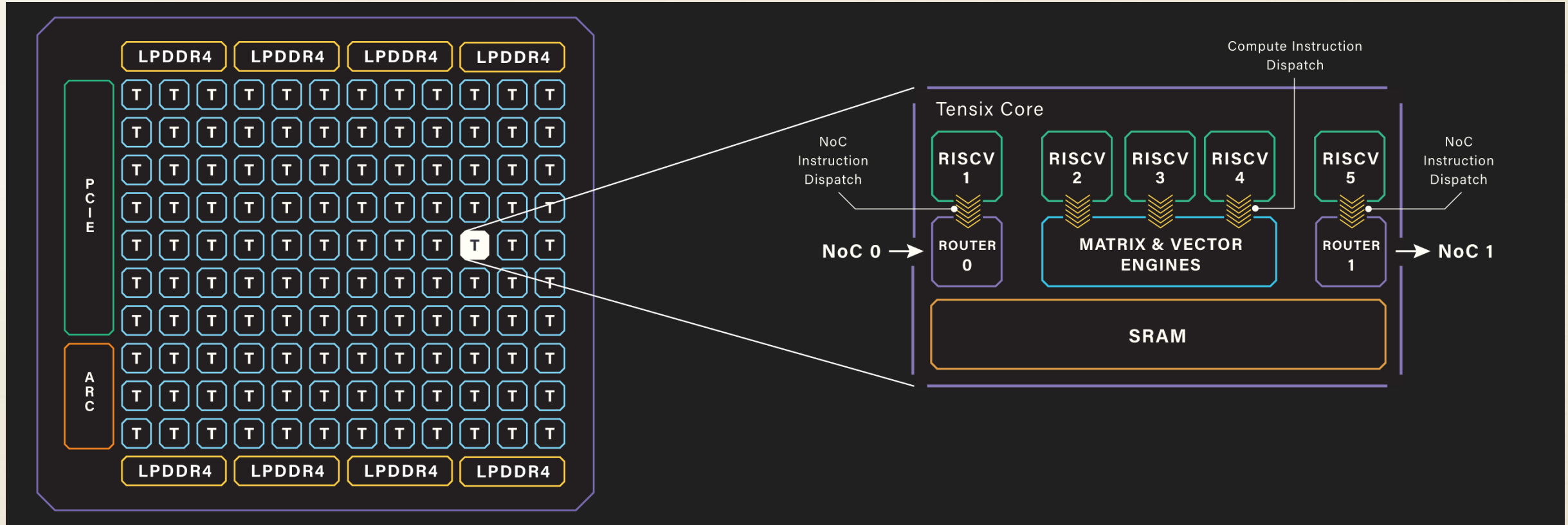
- Baby RISC-V CPU to help controlling AI acceleration engine
  - Control matrix / vector operations, data movement via custom instructions

AI Silicon Roadmap

Grayskull
AI Processor

- 120 Tensix Cores
- 12nm
- 276 TOPS (FP8)
- 16 lanes of PCIE Gen4
- 8 channels LPDDR4

GEN 1

Wormhole
Networked AI Processor

- 80 Tensix+ Cores
- 12nm
- 328 TOPS (FP8)
- 16x100 Gbps Ethernet
- 6 channels GDDR6
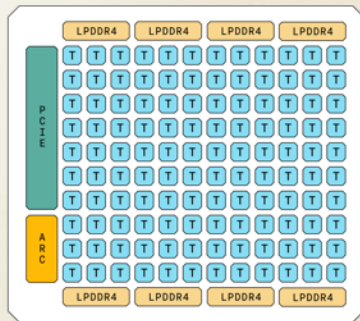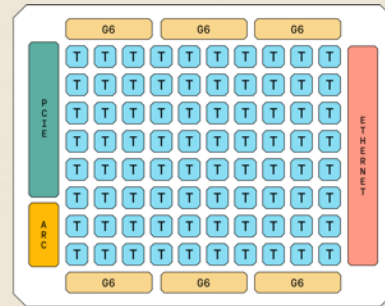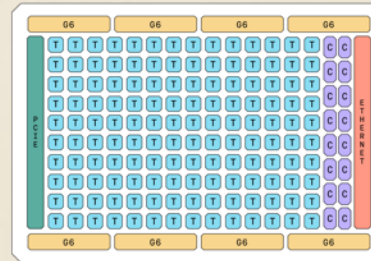- 16 lanes of PCIE Gen4

GEN 1

Blackhole
Standalone AI Computer

- 140 Tensix++ Cores
- 6nm
- 790 TOPS (FP8)
- 12x400 Gbps Ethernet
- 48 lanes of SERDES
- 8 channels of GDDR6
- 16 RISC-V CPU cores

GEN 2

Quasar
Low Power AI Chiplet

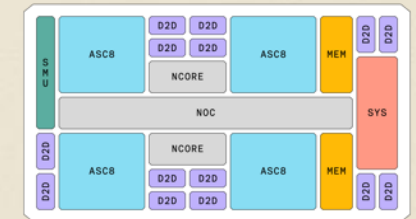- 160 Tensix Neo Cores on 4nm Chiplet
- Features incl SMC with Self-boot/Reset
- Non-blocking D2D Interfaces
- Easily stack Quasar or combine to choose your own compute

GEN 4

Aegis/Athena
Highly Performance RISC-V CPU Chiplet

- 4nm 32 RISC-V Ascalon CPU Cores
- Feature support incl SMC, IOMMU, AIA
- Non-blocking D2D Interfaces
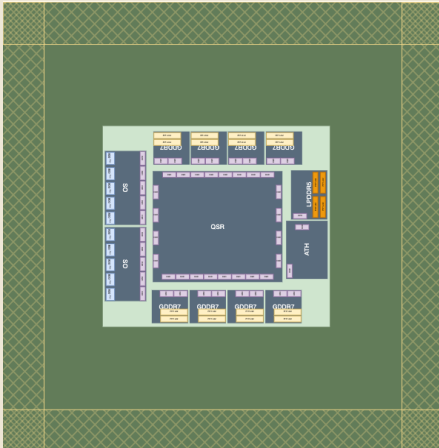- Composable IO, MEM, CPU compute

GEN 4

Chiplet Diversity

Chiplet

o Silicon SiP diversity
- o Design reuse
- o Low-cost development
- o Composability
- o Heterogeneity

o Diversity through composing chiplets from different organizational entities

Open Chiplet Architecture (OCA)

○ Open Chiplet Architecture (OCA) defines chiplet-interoperability in 5 layers:
- Physical
- Transport
- Protocol
- System
- Software

○ **Plug-Talk-Play**
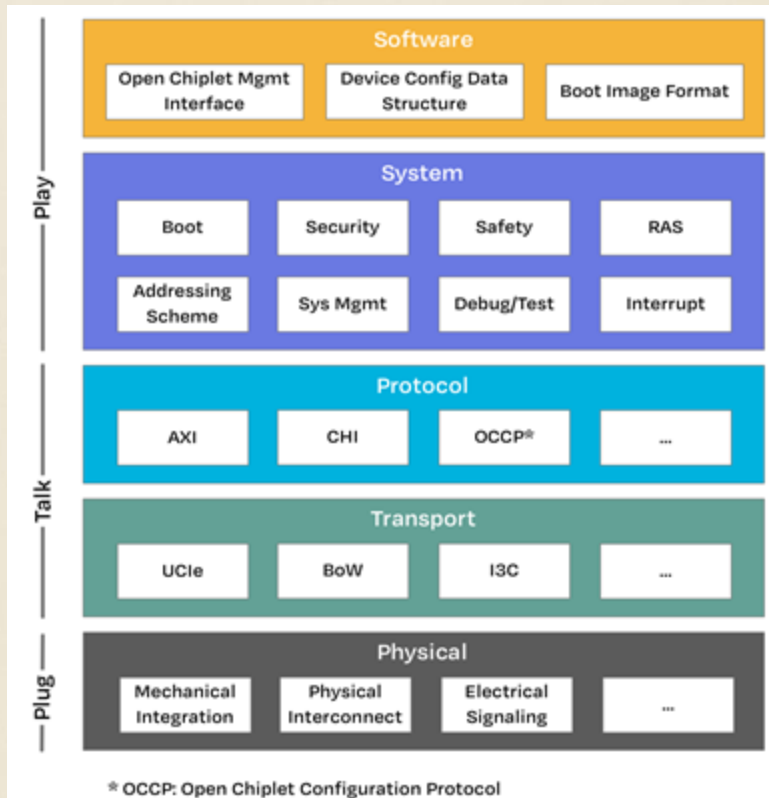
OCA
OPEN CHIPLET ATLAS

Open Chiplet Architecture Specification

Version 0.7, 09/22/2025: Under Development

Open Chiplet Architecture

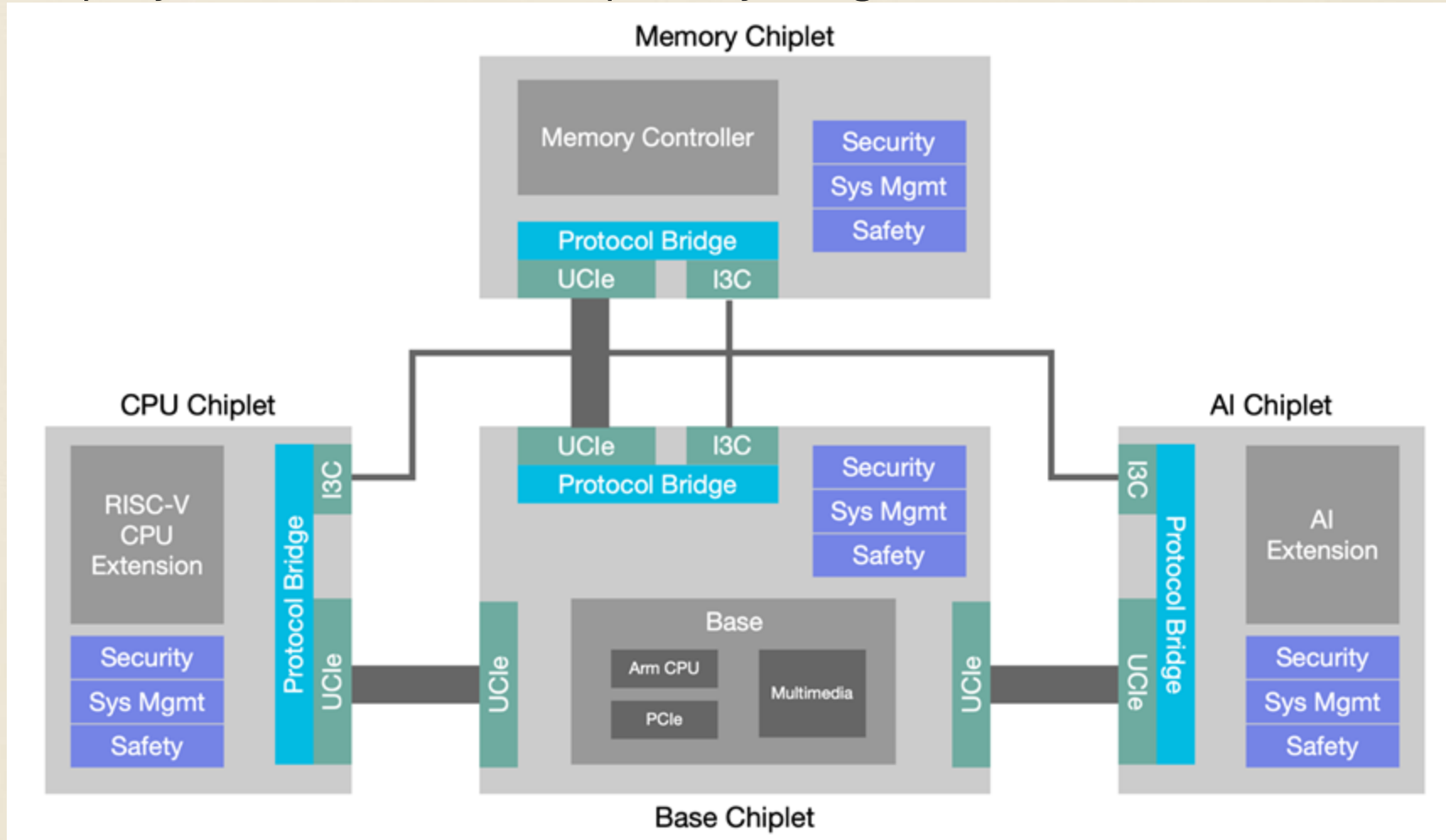o OCA takes care of chiplet compatibility

o Each company focus on its core-competency design

Robotic – Auto
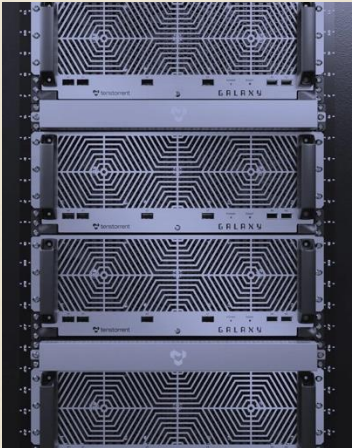
Dual SoCs
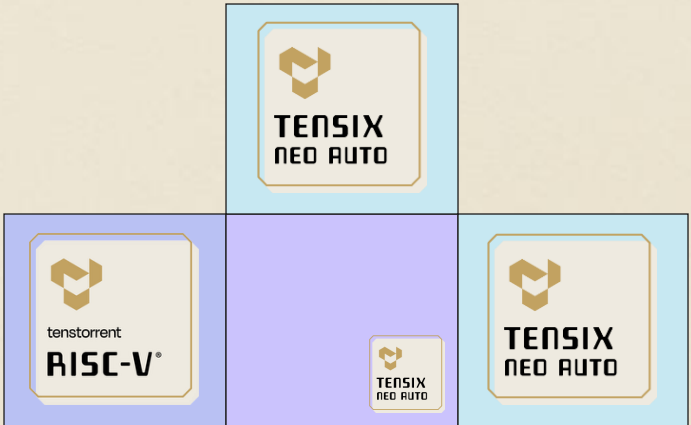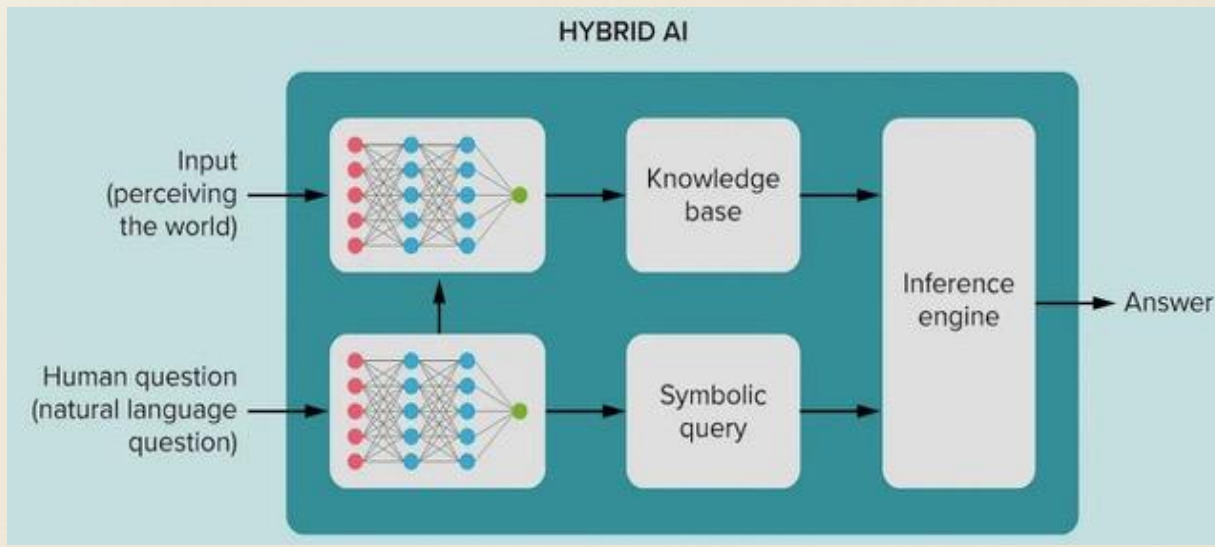
Dual SoCs + AI Co Processors

Central Compute Unit - Heterogenous

Neuro-symbolic AI Chiplet

- Heterogenous Compute with high-performance Ascalon
  - Neuro-symbolic AI
  - Optimize CPU/GPU data transfer latency

Chiplet Design Challenges

Why Chiplet 1: Performance / Cost (Yield)

256 CPU

1,000mm2

64 CPU 300 sqmm	64 CPU 300 sqmm
64 CPU 300 sqmm	64 CPU 300 sqmm
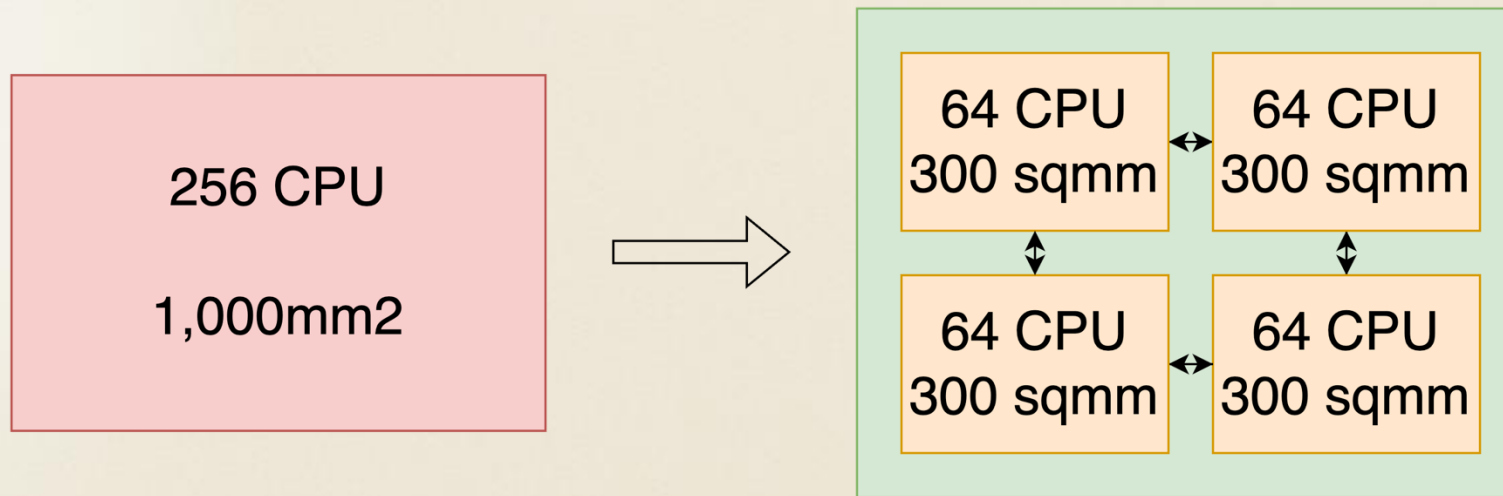
- Build (virtually) huge chip with reasonable cost

- Small chiplet cannot enjoy this (yield is already high)

Figure 2: Comparison of Unit Cost for Monolithic (Option 1) vs. Chiplets (Option 2)

M. Ahmad, J. DeLaCruz and A. Ramamurthy, "Heterogeneous Integration of Chiplets: Cost and Yield Tradeoff Analysis,"
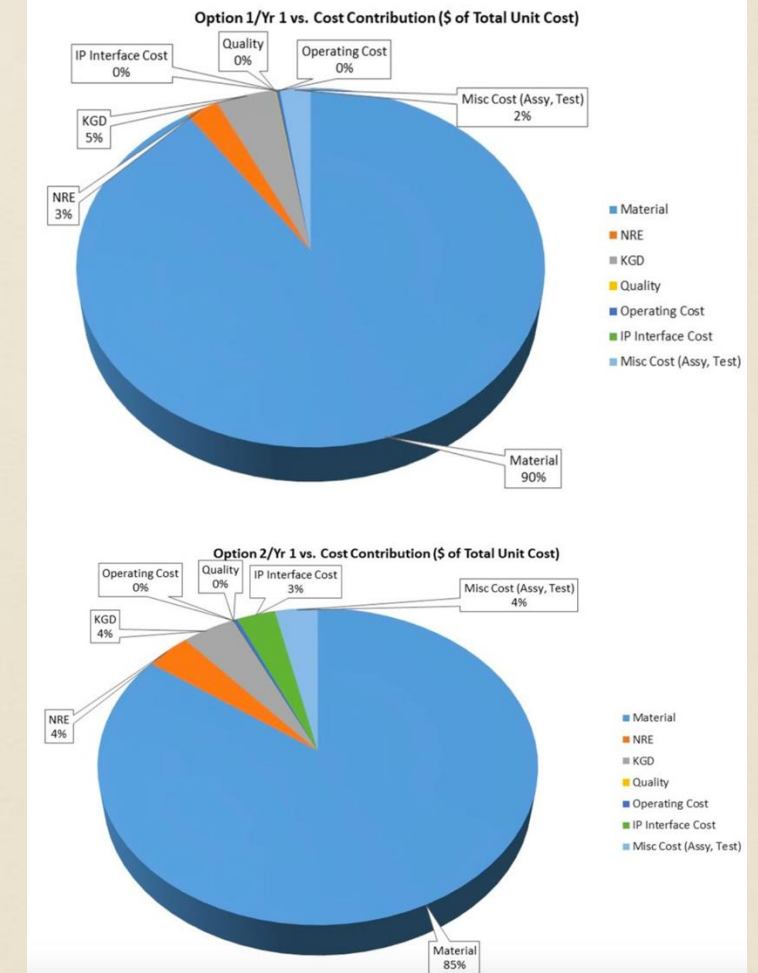
Why Chiplet 2: Reusability

- Reuse same chiplet for multiple generations
  - Building IO chiplet with 7nm, upgrading CPU chiplet from older nodes to newer nodes
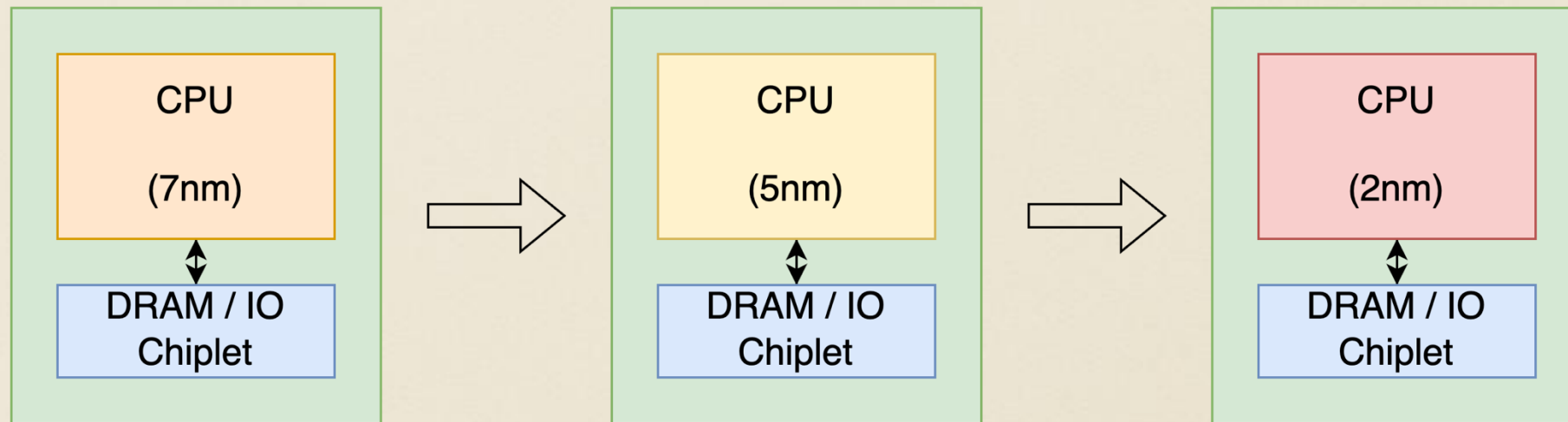
- Long term & strategic investment is required since 1st generation needs more work

Why Chiplet 3: Composability
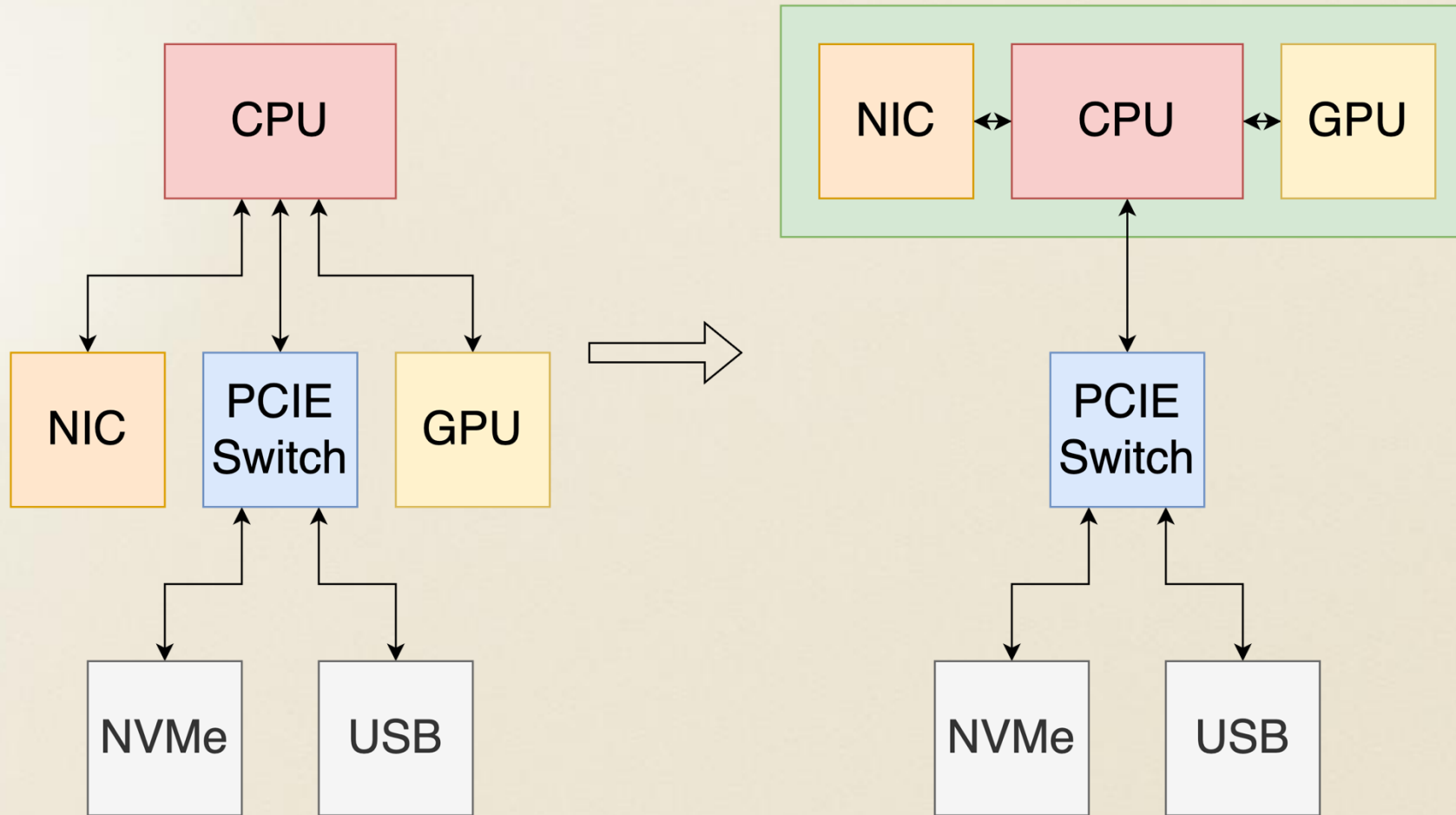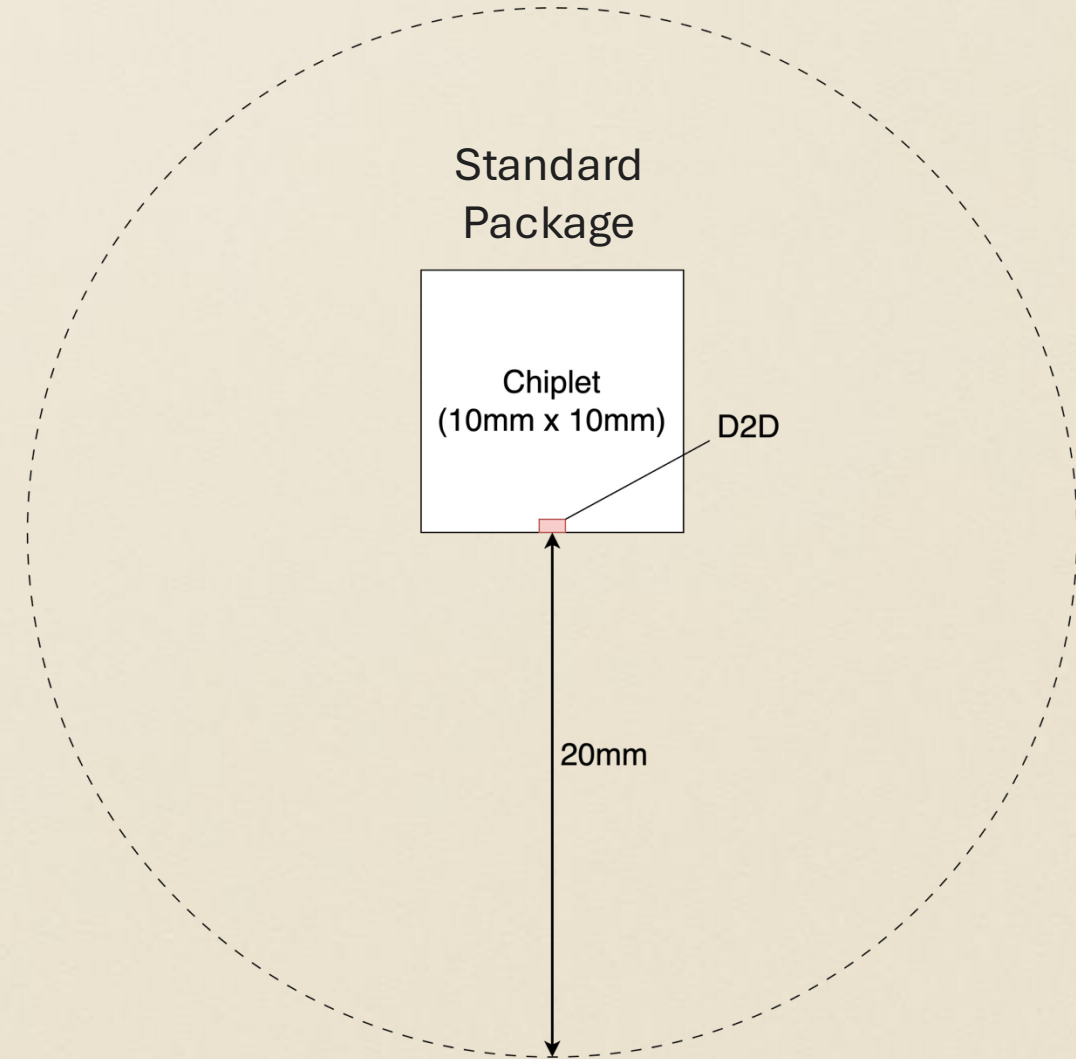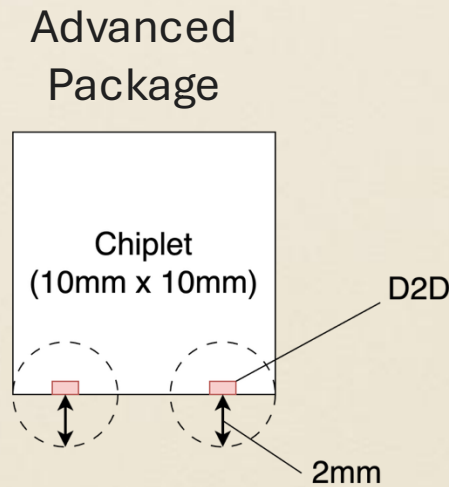

- UCIE is 10x better energy efficiency over PCIE
  - ~0.5 pJ / bit
  - ~0.5W @ 128GB/s

- Standard and ecosystem is required

Channel Reach

- UCIe-Adv has significant channel reach restriction
- Must align D2D location during design time

|	UCIe Advanced	UCIe Standard
Die Edge BW	1317 GB/s/mm	224 GB/s/mm
Bump Pitch	25 – 55 um	100 – 130 um
Channel Reach	2mm	<= 25mm
Energy Efficiency	0.25 – 0.6 pJ / bit	0.5 – 1.3 pJ / bit

Composability Challenge

- Chiplet from well designed system + chiplet from other well designed system may end up weird design.
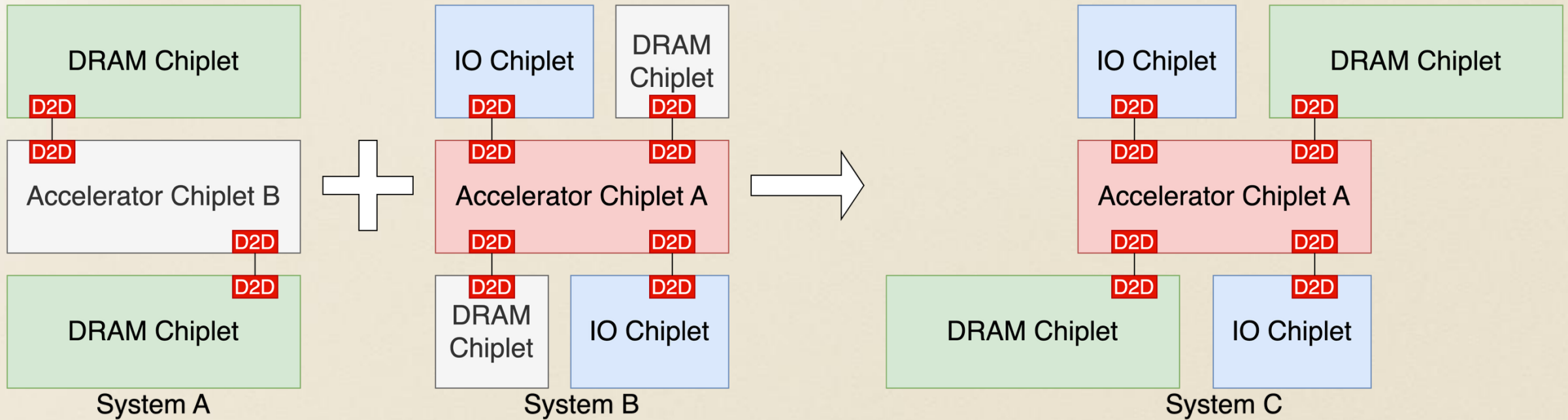
Deadlock Free Packet Routing

- Routing Strategy
  - Chiplet A is routing shortest path since it is tree
  - Chiplet B is routing X-Y dimension order

- If 2 different chiplets are combined, there is a risk of loop dependency

- There are research address this issue, but, no industry standard
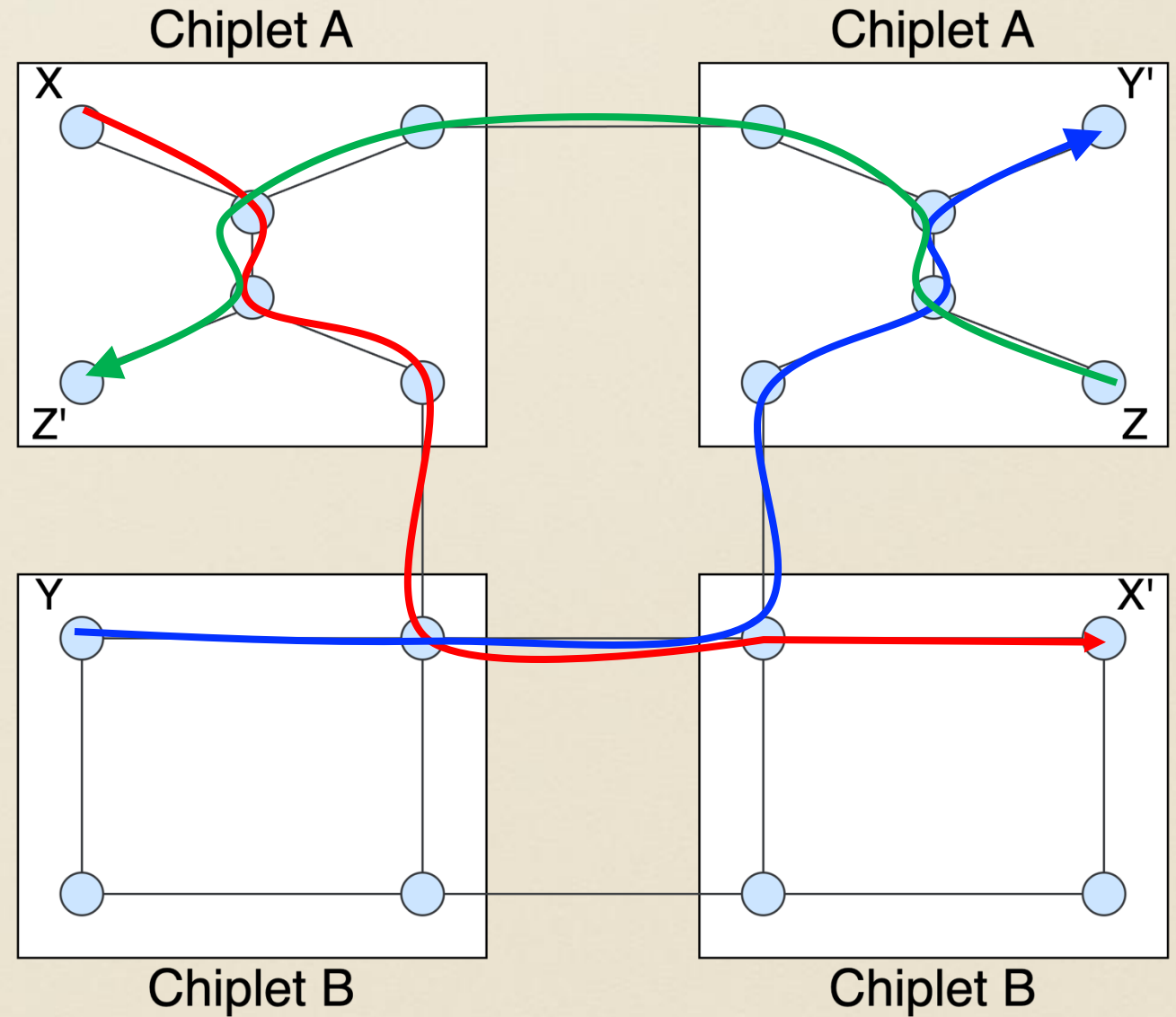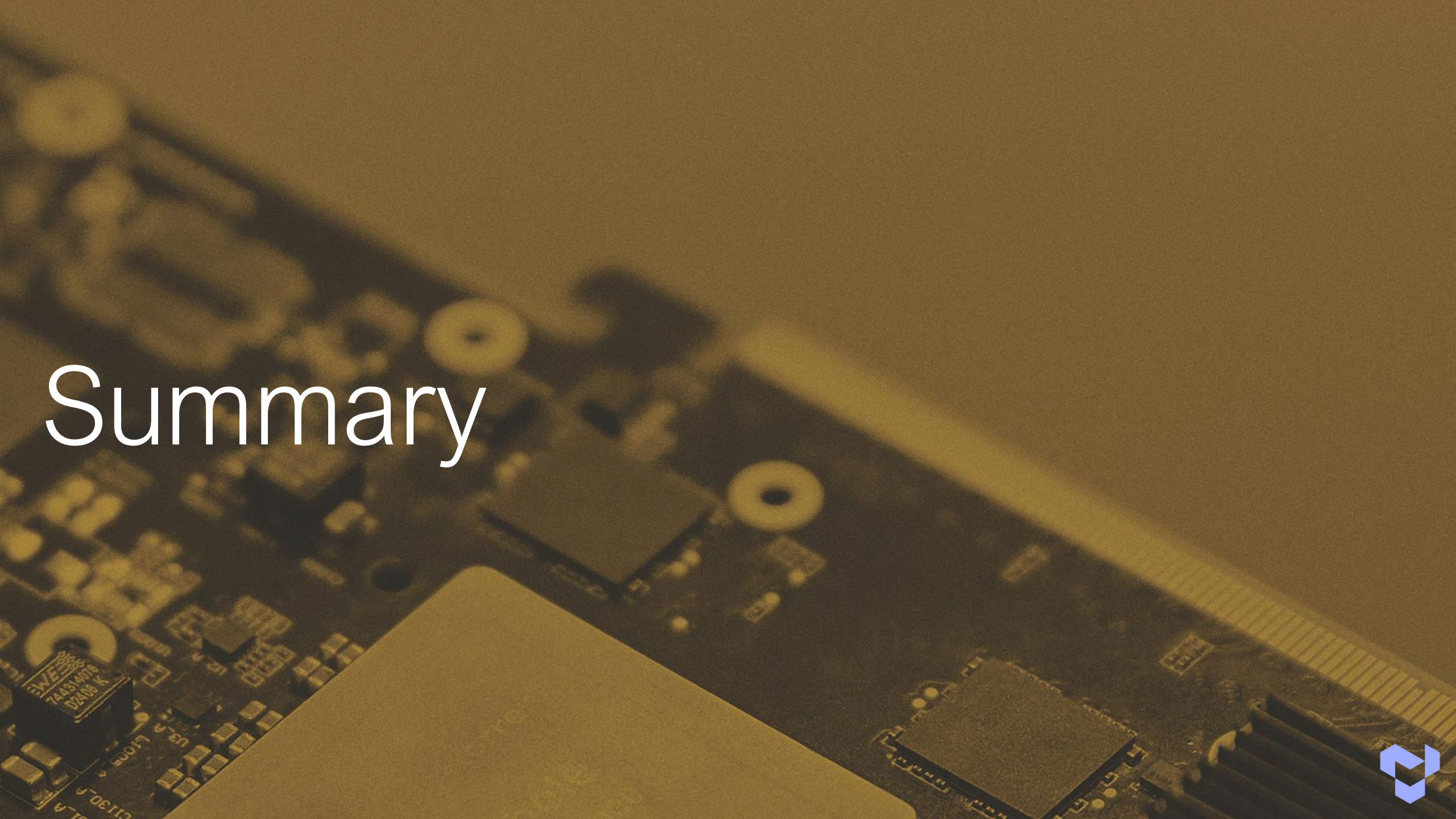
More Challenges

- Physical Layer
  - Package design

- Protocol
  - PCIE? CXL? AMBA CHI? AMBA AXI?

- System
  - Boot, Security, Safety, RAS, Power Management, Interrupt, etc…

- Software
  - System Management Software, Device Config

Chiplet is still very attractive for chip designers

- Benefit
  - Build huge chip
  - Energy efficient IO
  - Reuse same chiplet for multiple generations

- We want academic / industry combined effort to make this happen.
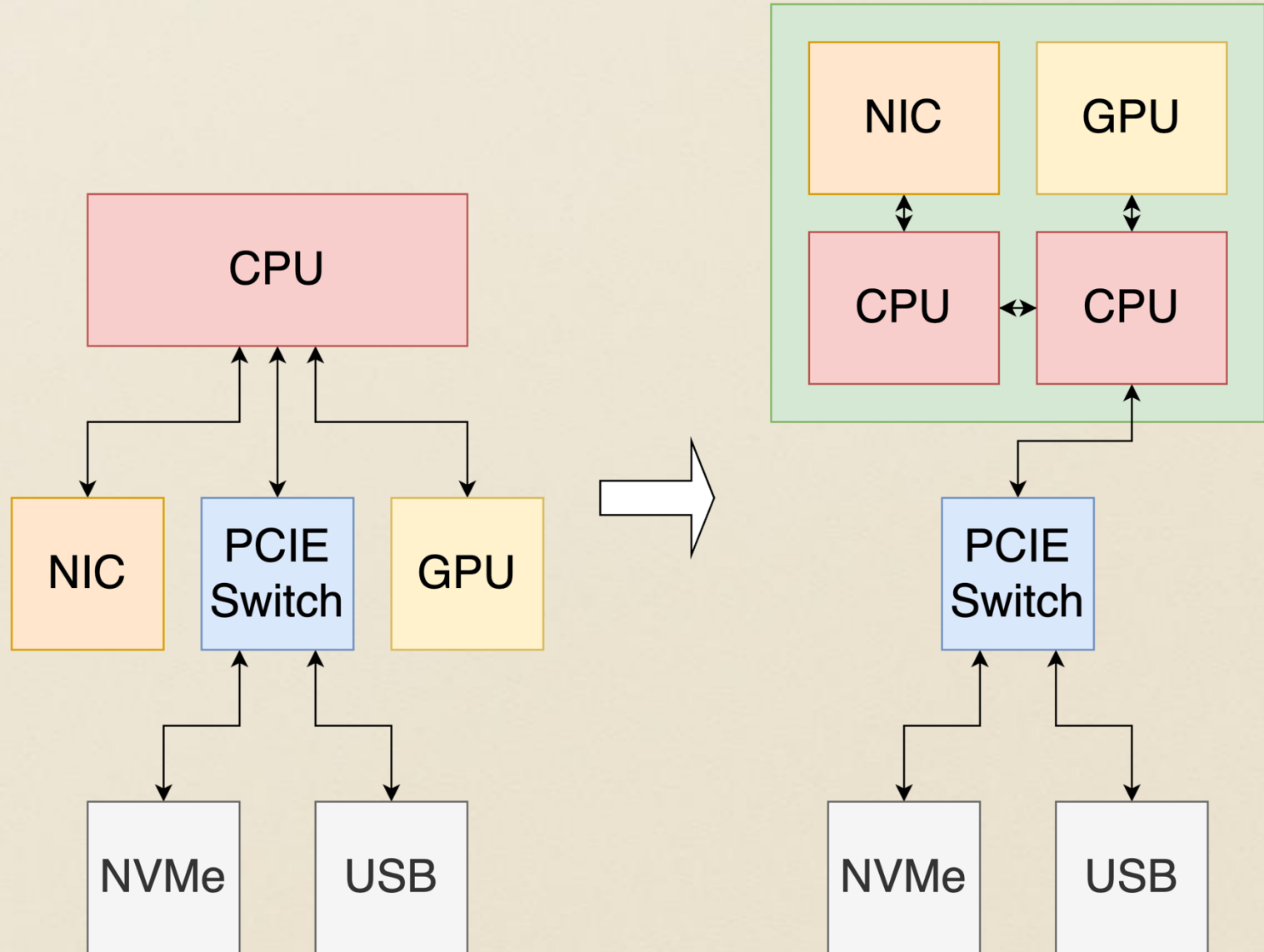
Summary

Summary

- RISC-V CPU for Application Processor / AI Accelerator
- Chiplet to address diverged market
- Open Chiplet Architecture (OCA) to address Chiplet design challenges