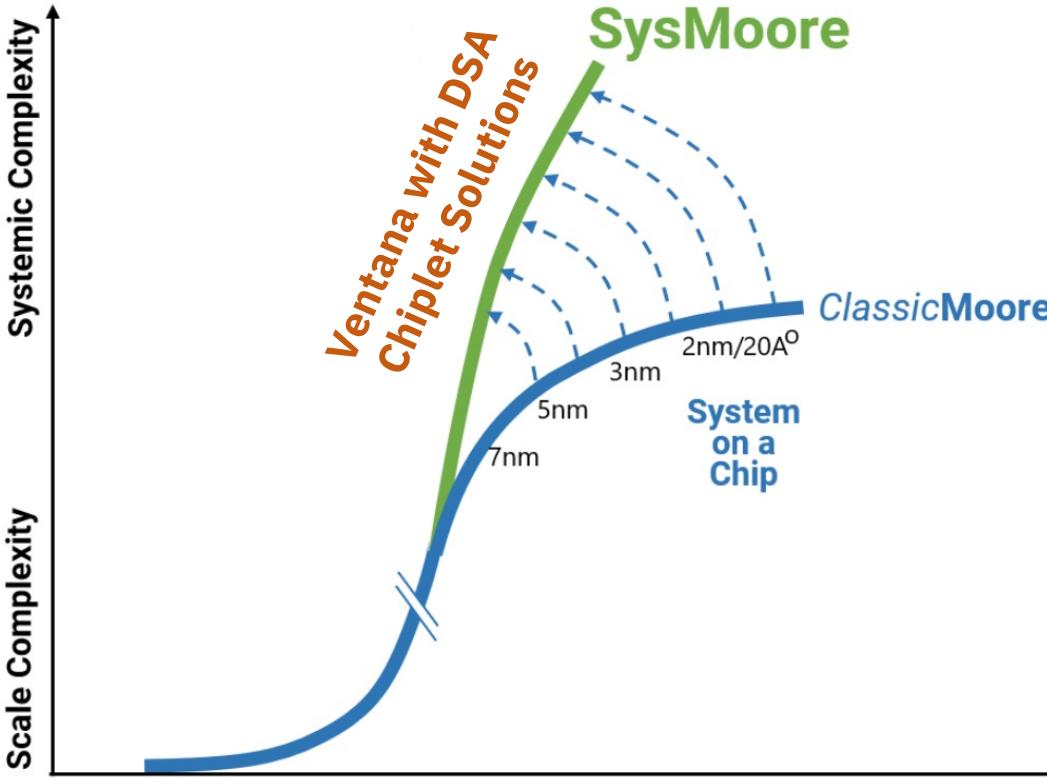




## High Performance RISC-V Chiplets for Cloud and Edge Computing

Jeff Maguire, Director of Product Management  
August 24, 2023

# Chiplets Are Key To Extend Moore's Law



Source: Synopsys, <https://www.synopsys.com/glossary/what-is-sysmoore.html>

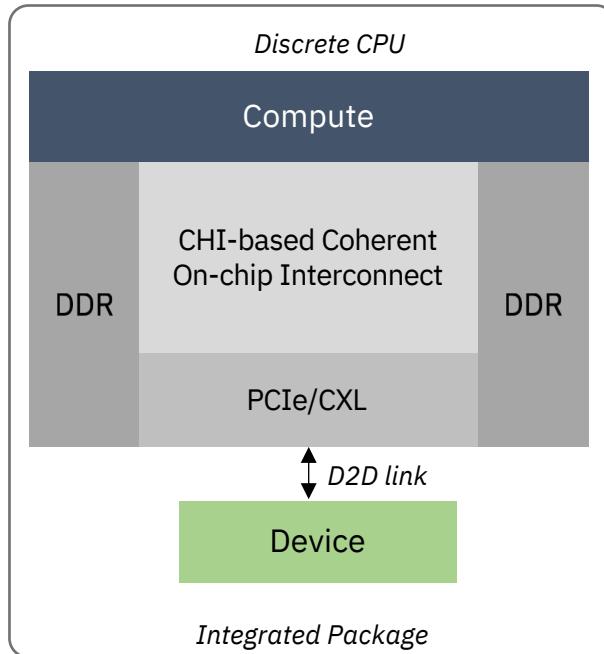
*“It may prove to be more economical to build large systems out of smaller functions, which are separately packaged and interconnected<sup>1</sup>. ”*

-Gordon E. Moore

<sup>1</sup>: “Cramming more components onto integrated circuits”, Electronics, Volume 38, Number 8, April 19, 1965

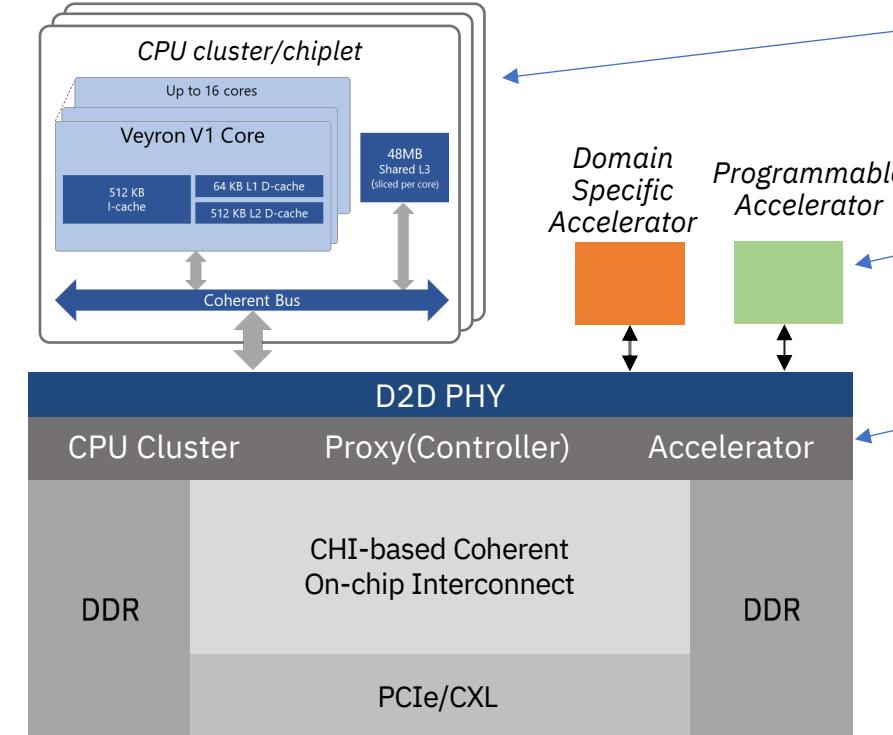
# Die Disaggregation Unlocks Innovation

**Traditional PCIe/CXL Attached Accelerator**



**Device Attach:** brings devices in-package as chiplets, connected over D2D with device protocols (CXL)

**Ventana's Chiplet Architecture**



Ventana chiplets are CPU cluster interfaced over a D2D link with AMBA® CHI protocol

Accelerators typically interface with AXI or ACE-Lite

Proxies represent the chiplets, unpacking D2D transport streams into protocol transactions on the SoC block interface

**Die Disaggregation:** breaks a monolithic SoC into chiplets, which can then be composed into systems using SoC protocols over D2D (CHI,AXI)

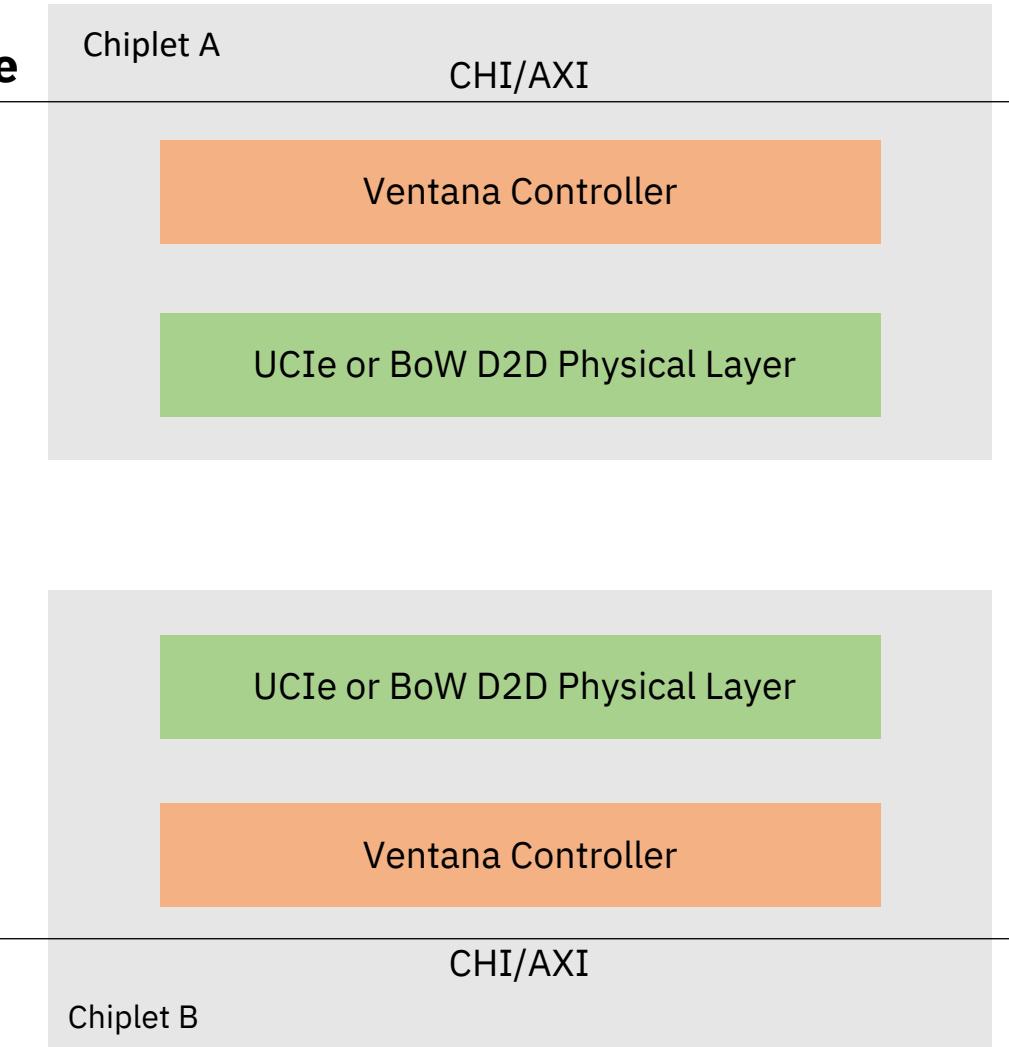
Ventana's Chiplet Architecture enables Chiplet vs IP integration to be interchangeable

# Pioneering Efficient Chiplet Interconnect Controller

## Requirements

- Low D2D latency
- Predictable memory performance
- Standard SoC buses and networks must map easily to D2D transport
  - CHI, AXI, ...
- HW memory coherency for efficient support of accelerators

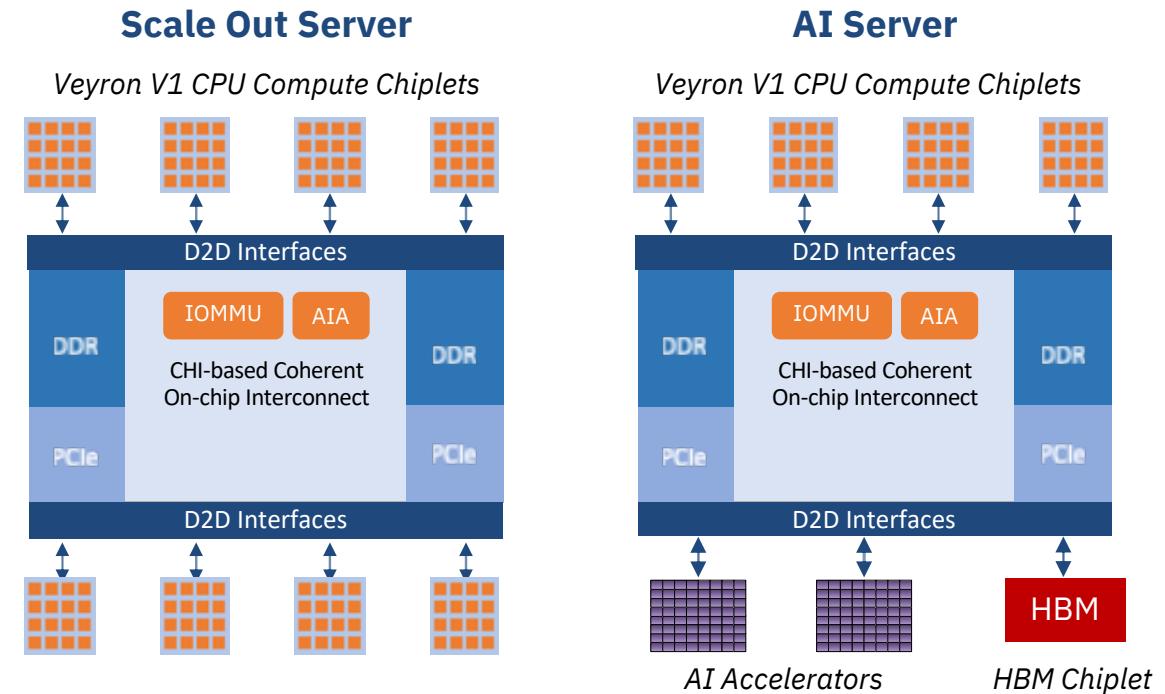
< 7ns latency  
vs 100ns+ for PCIe



Ventana has developed a low latency controller IP which transports standard SoC buses over D2D links

# Chiplets Enable Composability and Late Binding

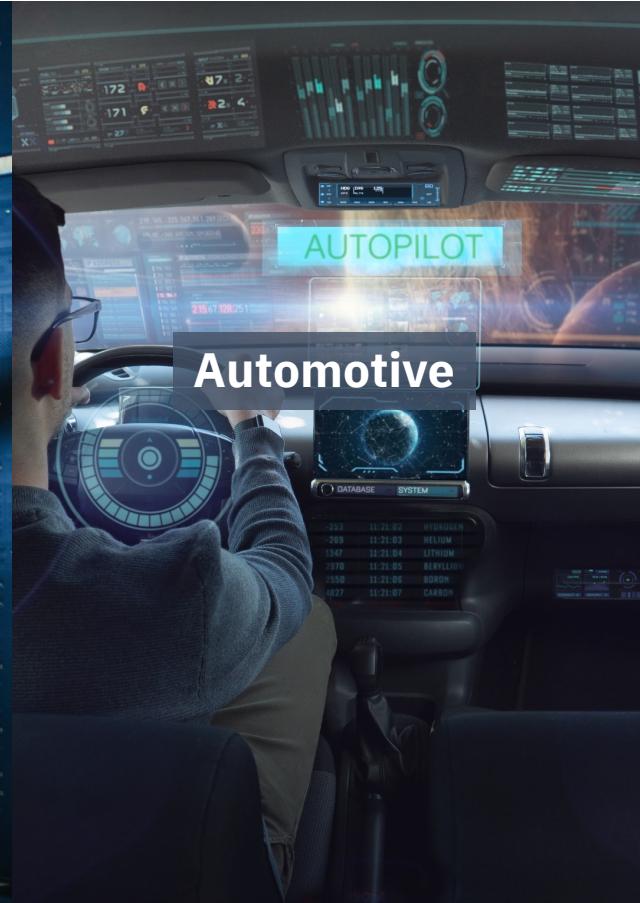
- Standard chiplets composable into desired solution with short lead time
- Switch between desired configuration of CPUs, Memory, and Accelerators
- Option of Standard IO Hub or Custom IO Hub with integrated features
- IO Hub and Custom Accelerators can be designed in N-1/N-2 process node for cost efficiency
- **Development Time:** ~1 year
- **Development Cost:** < \$25M



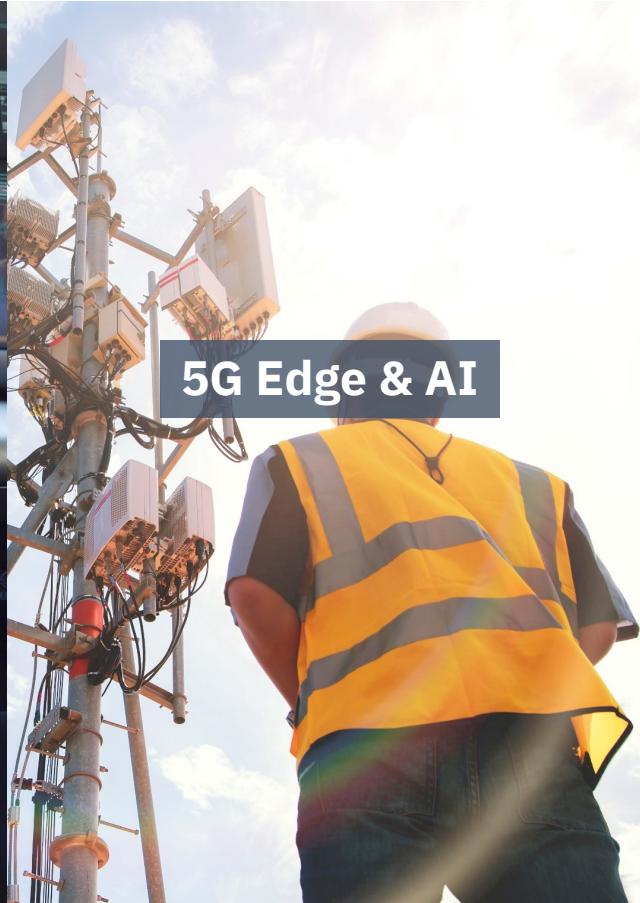
# Target Markets



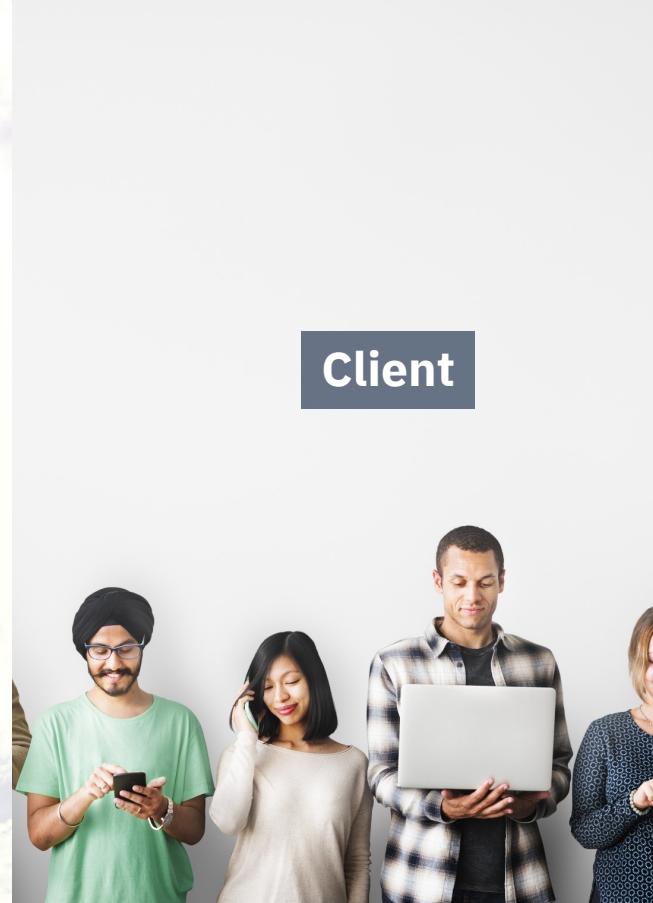
Data Center



Automotive

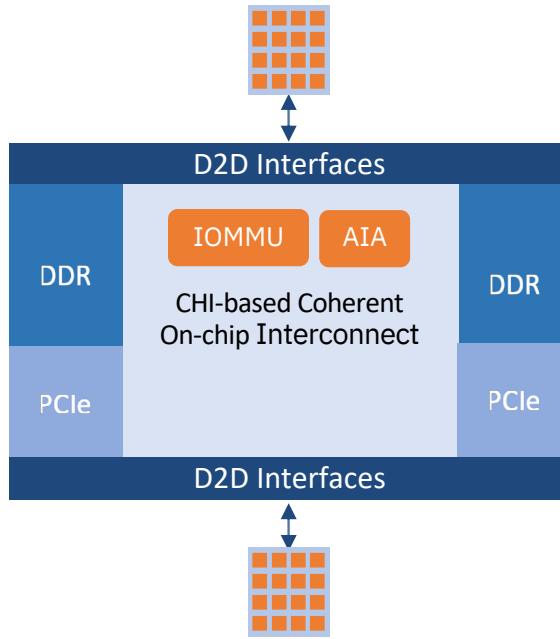


5G Edge & AI



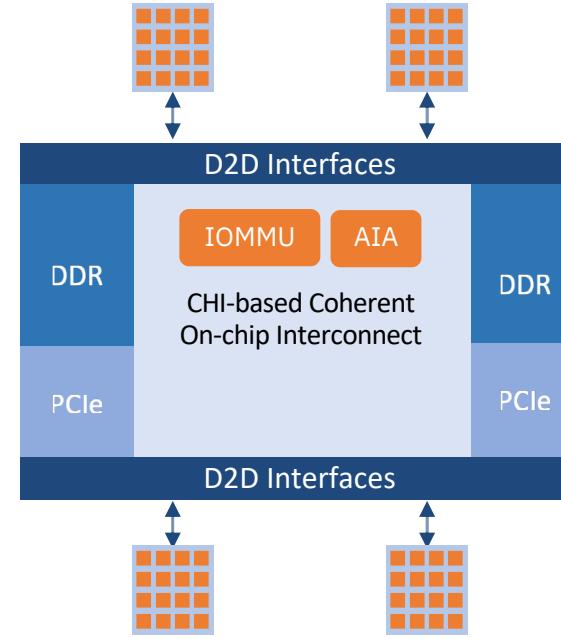
Client

# Scalable Architecture for Server-class Compute



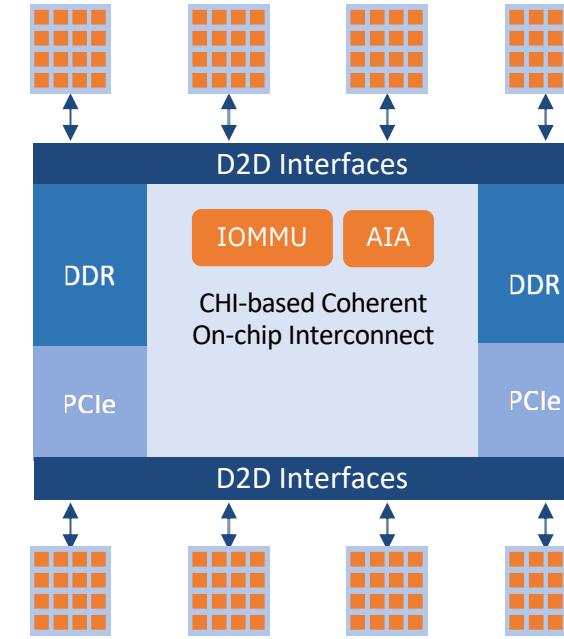
2 chiplets (32 cores)

**Entry level server**



4 chiplets (64 cores)

**Mid level server**

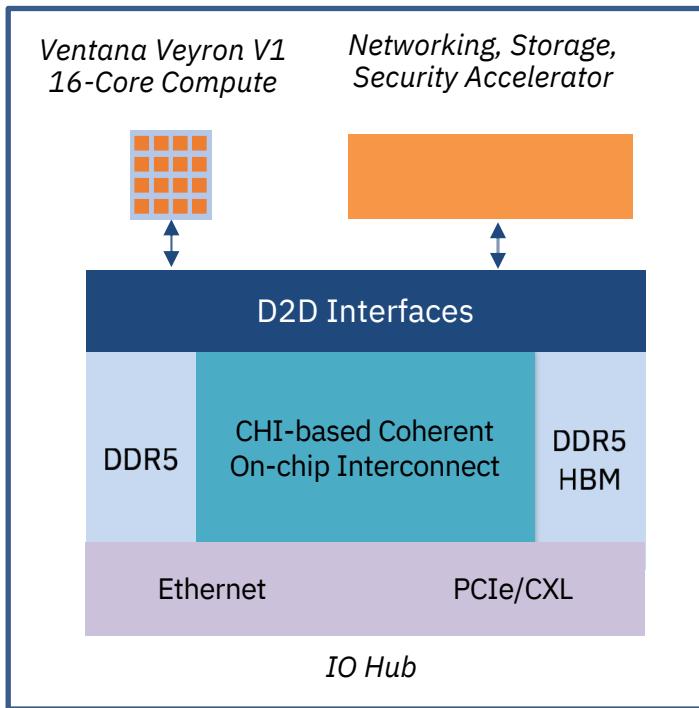


8-12 chiplets (128-192 cores)

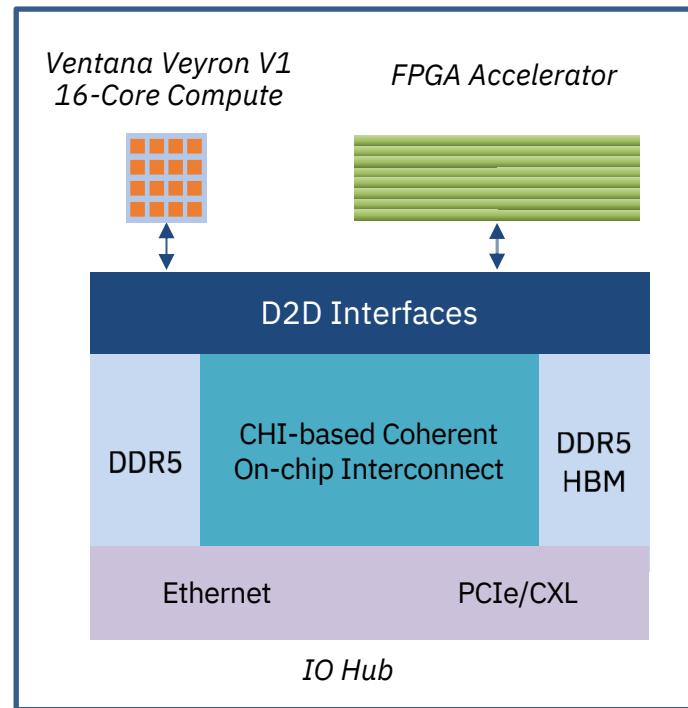
**High end server**

# DPU and Edge AI Compute

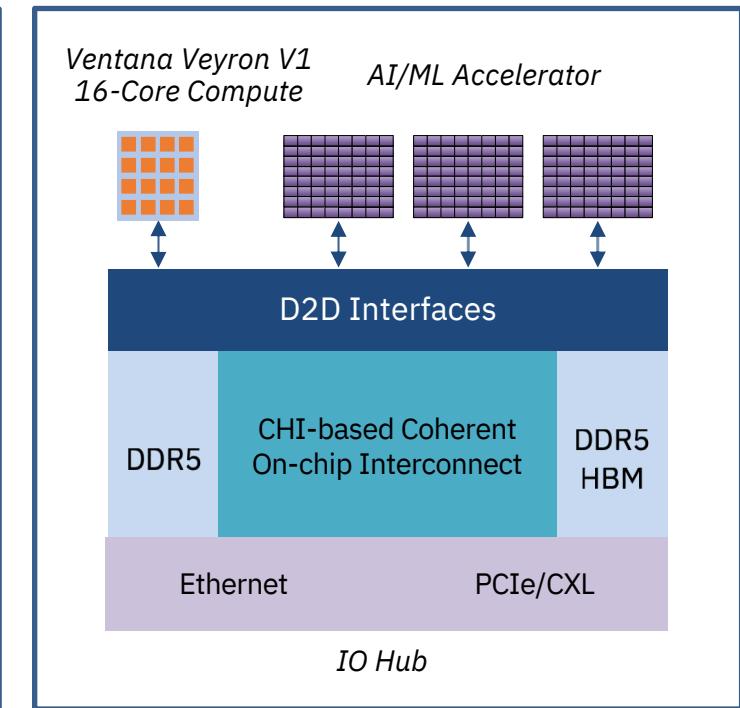
**DPU Accelerated**



**DPU + FPGA**

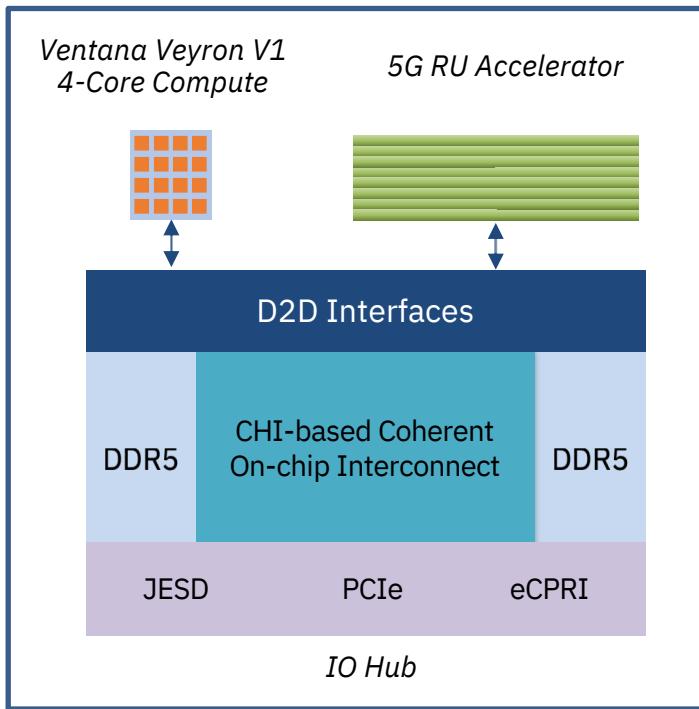


**Edge AI Compute**

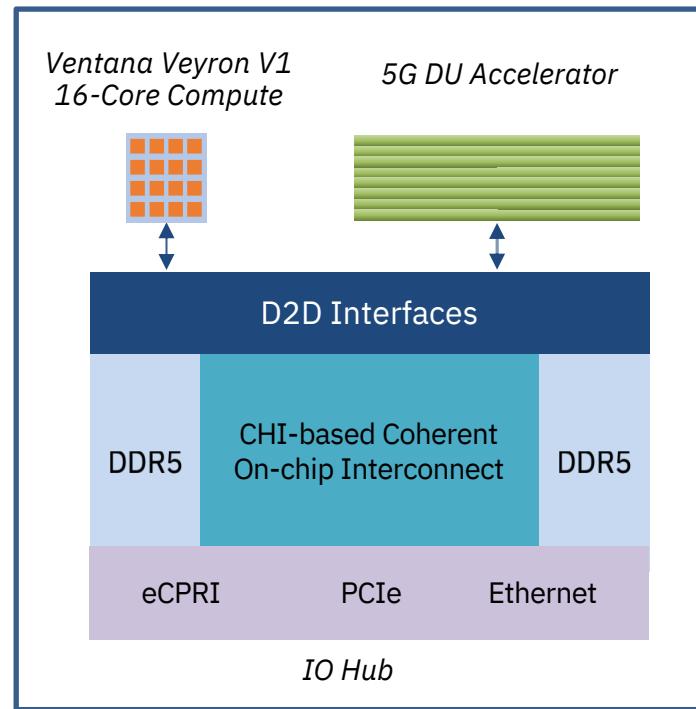


# 5G Open RAN Solutions

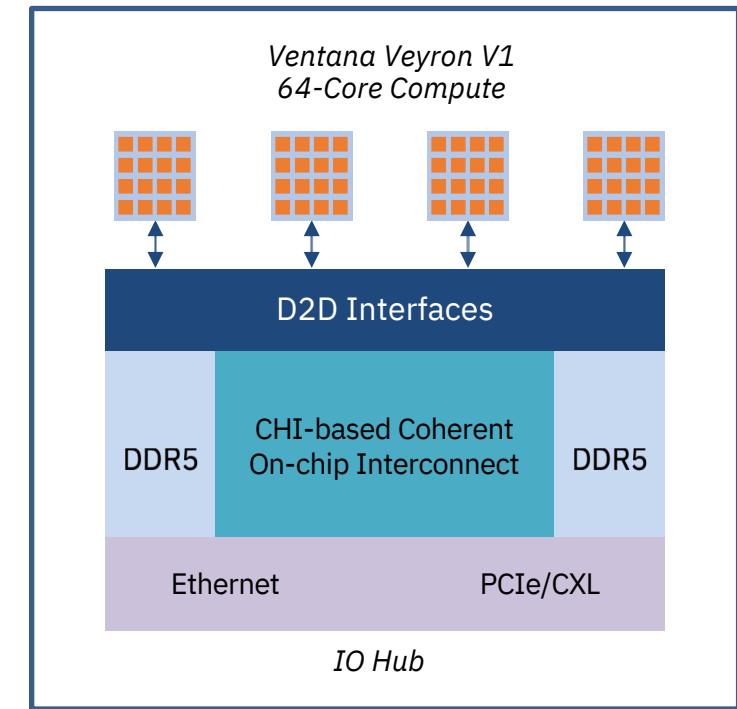
**Open RAN 5G RU**



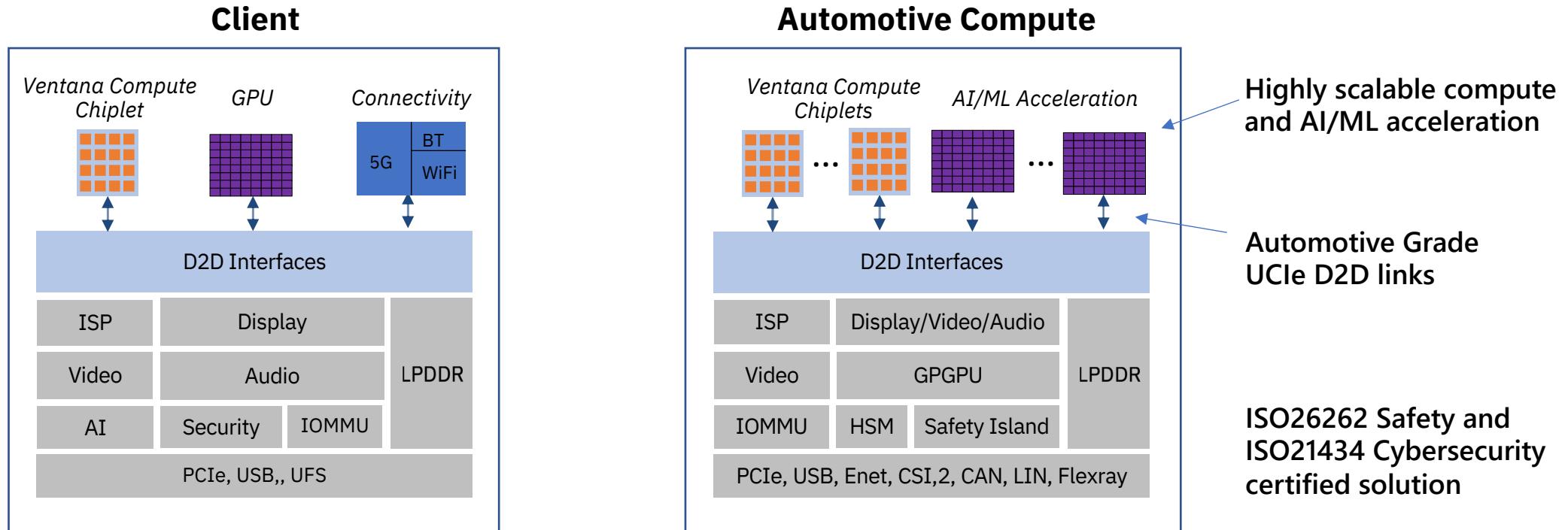
**Open RAN 5G DU**



**Open RAN 5G CU**



# Client and Automotive Processors

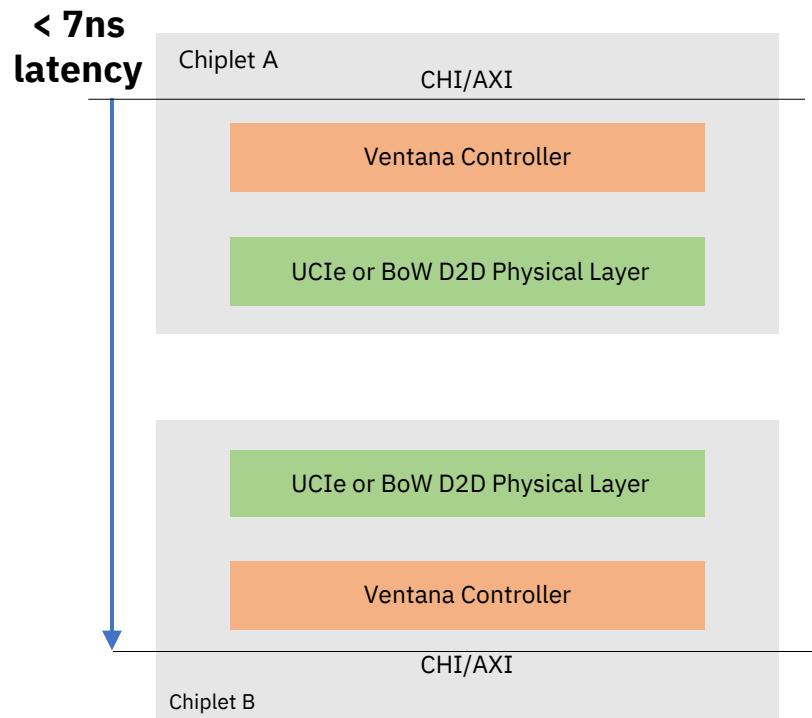


- Cost tends to drive domain specific SoC hub integrations
- Chiplets can then be used to provide a high degree of scalability across product tiers

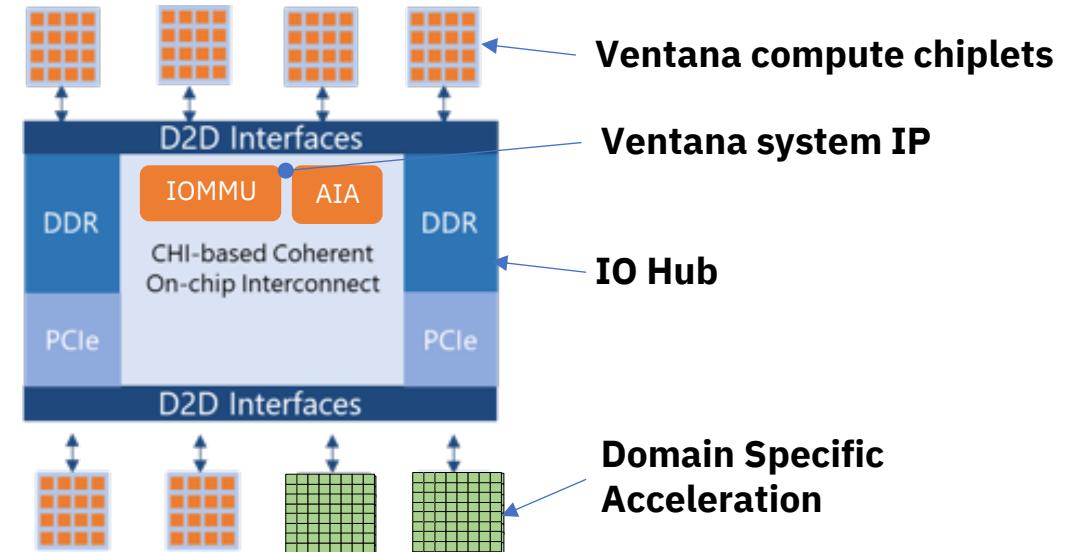
# Veyron: Putting It Together With Chiplets

## Solving the D2D Latency Problem

- Ventana has developed low-latency controller IP which maps standard SoC buses such as CHI & AXI to D2D PHY



## Veyron Chiplet Solution



- Veyron compute chiplets
  - In latest process node technology
  - Scalable CPU performance/count
- IO Hub
  - Implemented in process node of choice
  - Customized for application requirements
- Custom Domain Specific Acceleration
  - Low-cost process node



THE RISC-V PERFORMANCE LEADER

Thank You