CSS N2: Arm Neoverse N2 Platform, Delivered to Partners as a Fully Verified, Customizable Subsystem

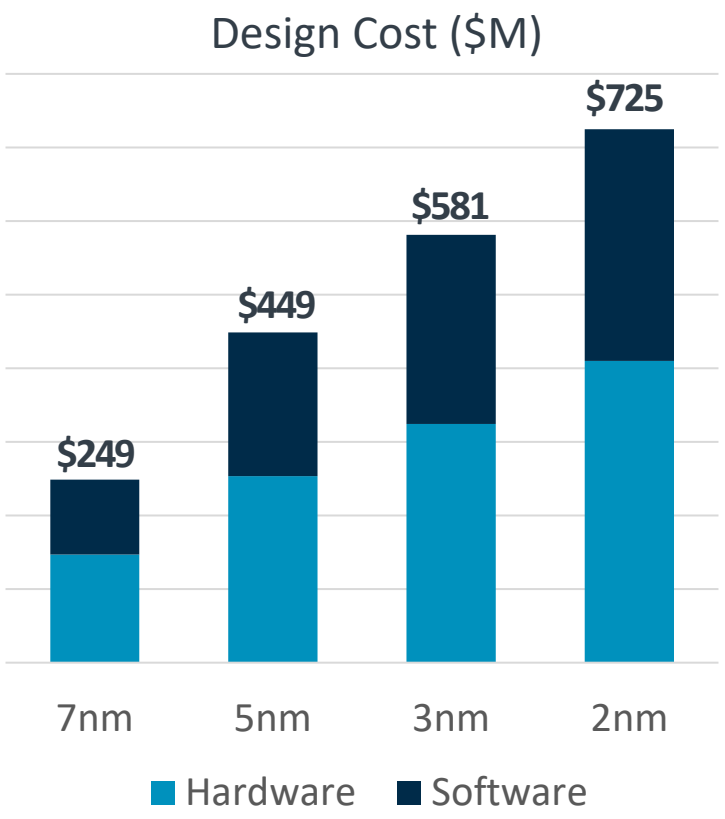
Hot Chips 2023

Anitha Kona, Lead System Architect & Fellow, Arm

August 28th, 2023

The Pressure on Silicon Development

Advanced Node Cost



Source: IBS Global Semiconductor Industry Service Report – July 2022

Time to Market Demands

Design Talent
HELP WANTED

Performance
MORE, MORE, MORE

Validation
NO SILICON SPINS

Cost, Revenue
REDUCE IT, RAMP IT

Need for Specialized Silicon

Moore's Law
SLOWING

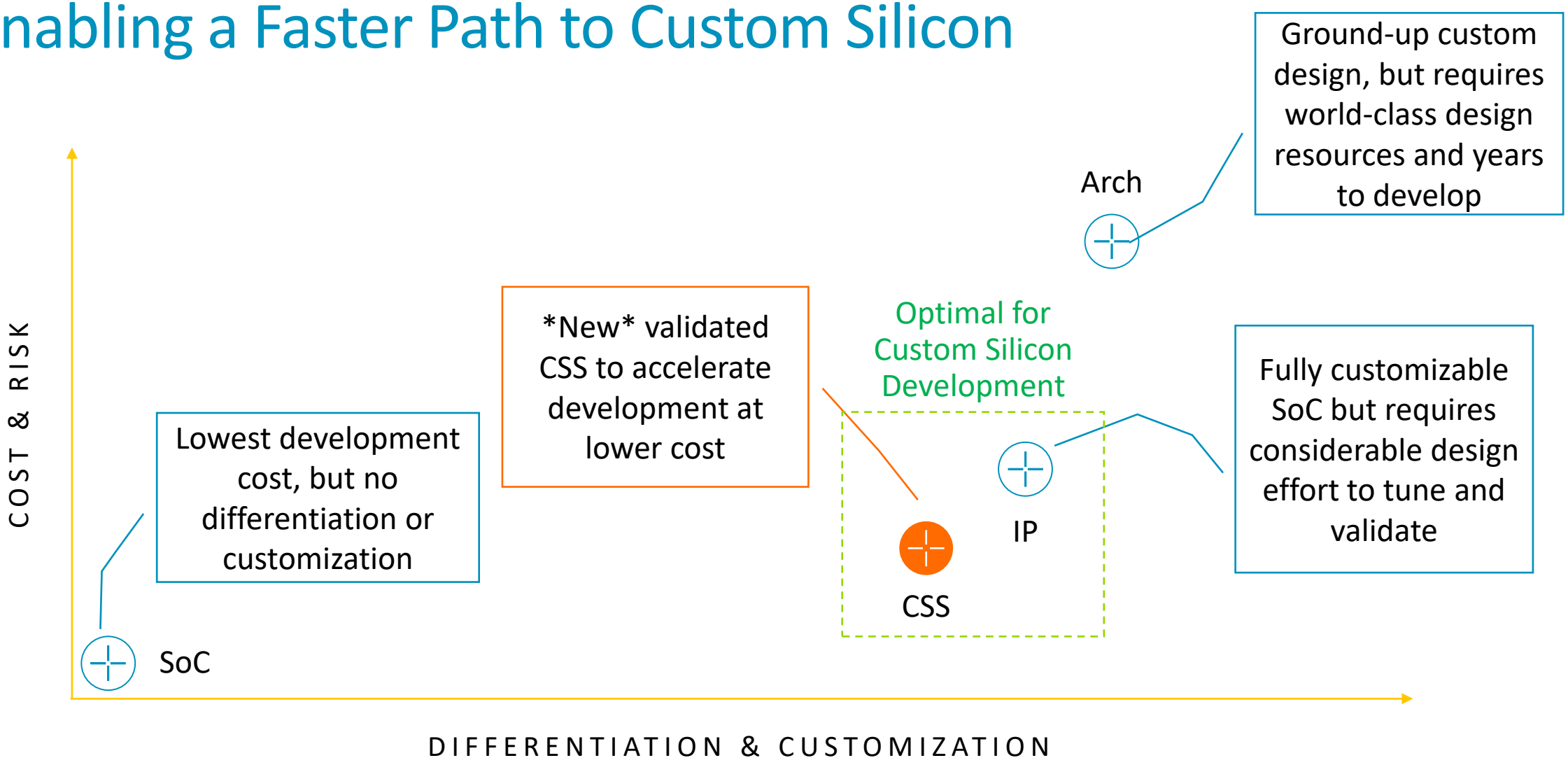
1B
5G SUBSCRIBERS

120 ZB
DATA GENERATED

Sustainability
NET 0 BY 2050

Source: 2023 data. Cisco,Statistica/IDC

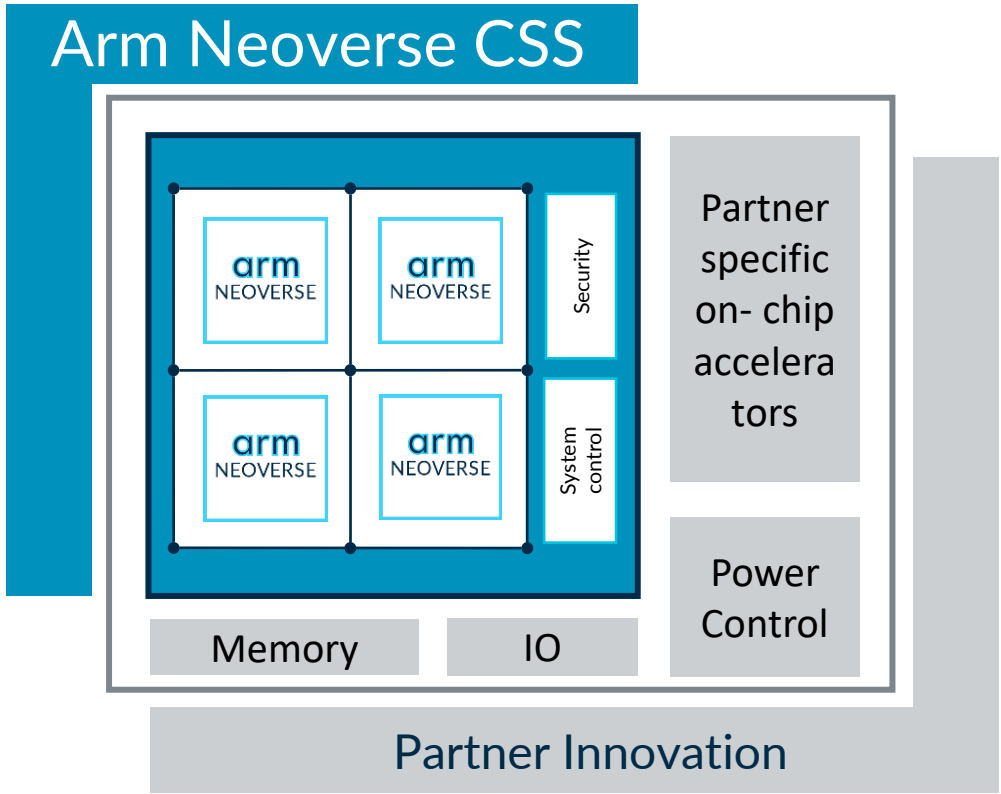
Enabling a Faster Path to Custom Silicon



The Relief: Arm Neoverse Compute Subsystem (CSS)

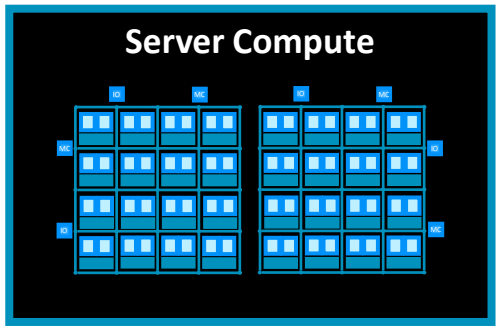
- Fastest Path to Production Silicon
- World Leading Performance
- Leading Edge Technology

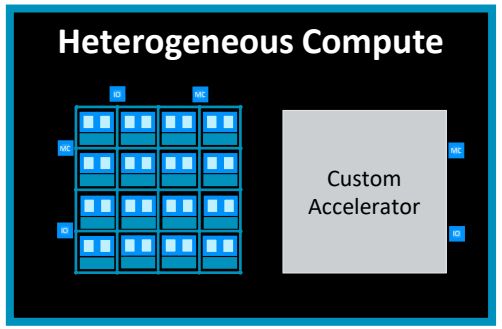
- Neoverse CSS delivers PPA optimized compute in leading edge technology node – fully validated RTL, pre-tuned implementation
- Deployment ready for emerging technologies – PCIe Gen5, CXL-based memory-pooling
- Enables custom or heterogenous integration – accelerators or specialized compute
- Can support chiplet-based designs
- SystemReady with reference SW stack



Arm Neoverse CSS N2 Cloud-to-Edge Differentiated Solutions

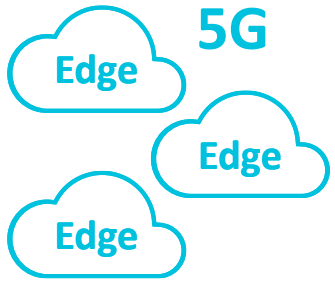
Scale-Out Cloud

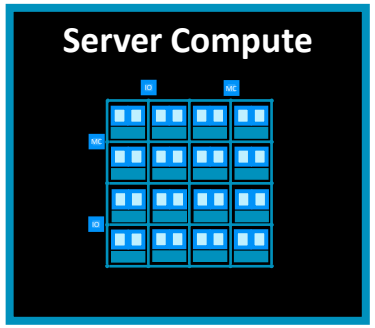


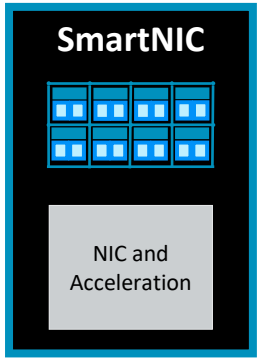


Performance Option for
Hyperscale
Up to 256C in a system

Networking and Infrastructure Edge

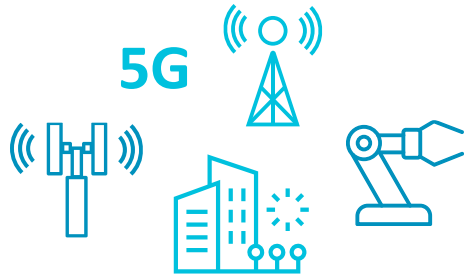


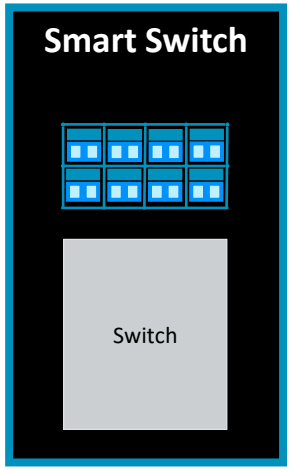


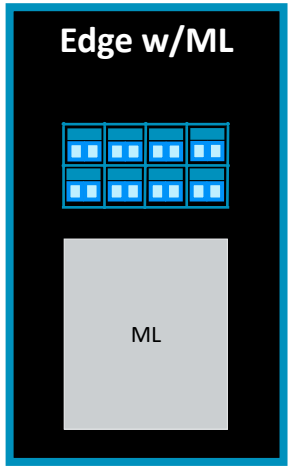


- DPU/ SmartNIC
- Smart Switching, Routing
- Edge telco server
- 5G Control Unit
- Edge compute w/ML

5G Infrastructure





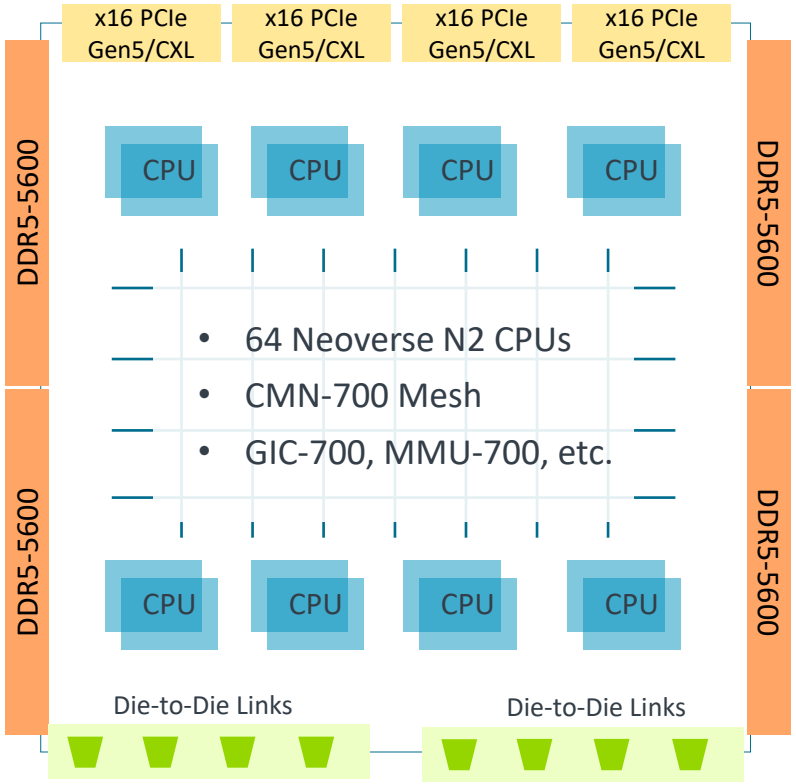


Broad market optimized
24~64C

arm NEOVERSE

Arm Neoverse CSS N2 ("CSS Genesis")

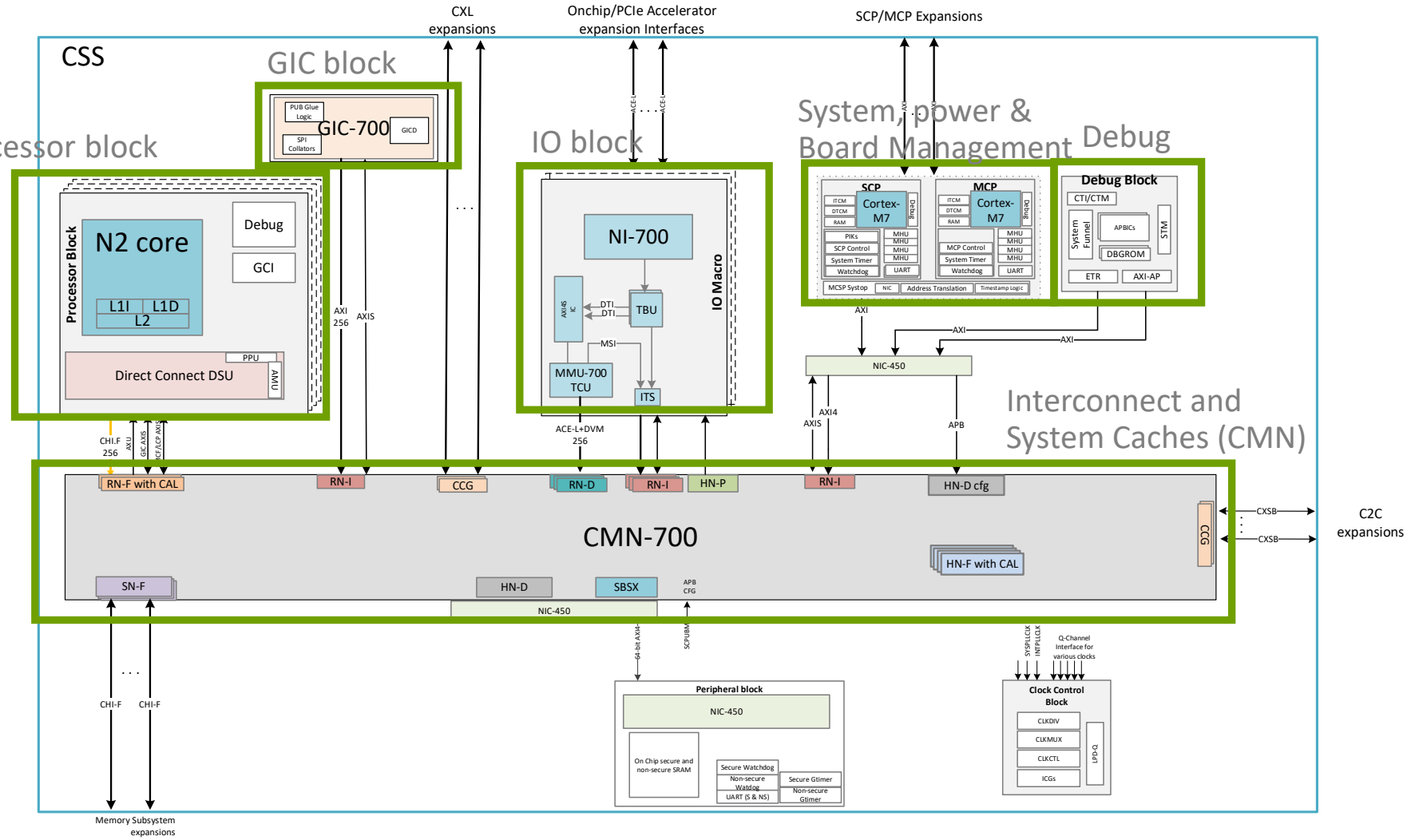
Proven in Silicon



Feature	5nm Specification
# of CPU Cores	24/32/64 N2 CPUs, 2.1-3.6GHz, 5nm advanced process SPECint2k17_Rate(64T) est: ~250 (3.0GHz)
Cache hierarchy	L1: 64KB I-cache (parity), 64KB D-cache (ECC) Configurable up to 1MB L2 private cache (private, ECC) Configurable up to 64MB system-level cache (shared, ECC)
Memory/IO Connectivity	Provides standard AMBA CHI/AXI interfaces to support Up to 8x 40b DDR5 or LPDDR5 channel Up to 4x x16 PCIe/CXL Gen5 lanes Bifurcation support up to x4 for each x16 PCIe lanes
Die-to-Die connectivity	Support for SMP compute or accelerator attach with UCIe or partner-specific D2D PHY
General I/O	Expansion interface to connect slower speed peripheral like USB/ I3C/QSPI/etc.
System management	System Control Processor(SCP)/ Manageability Control processor(MCP)
CSS area (24/32/64c)	~53/61/198 mm sq

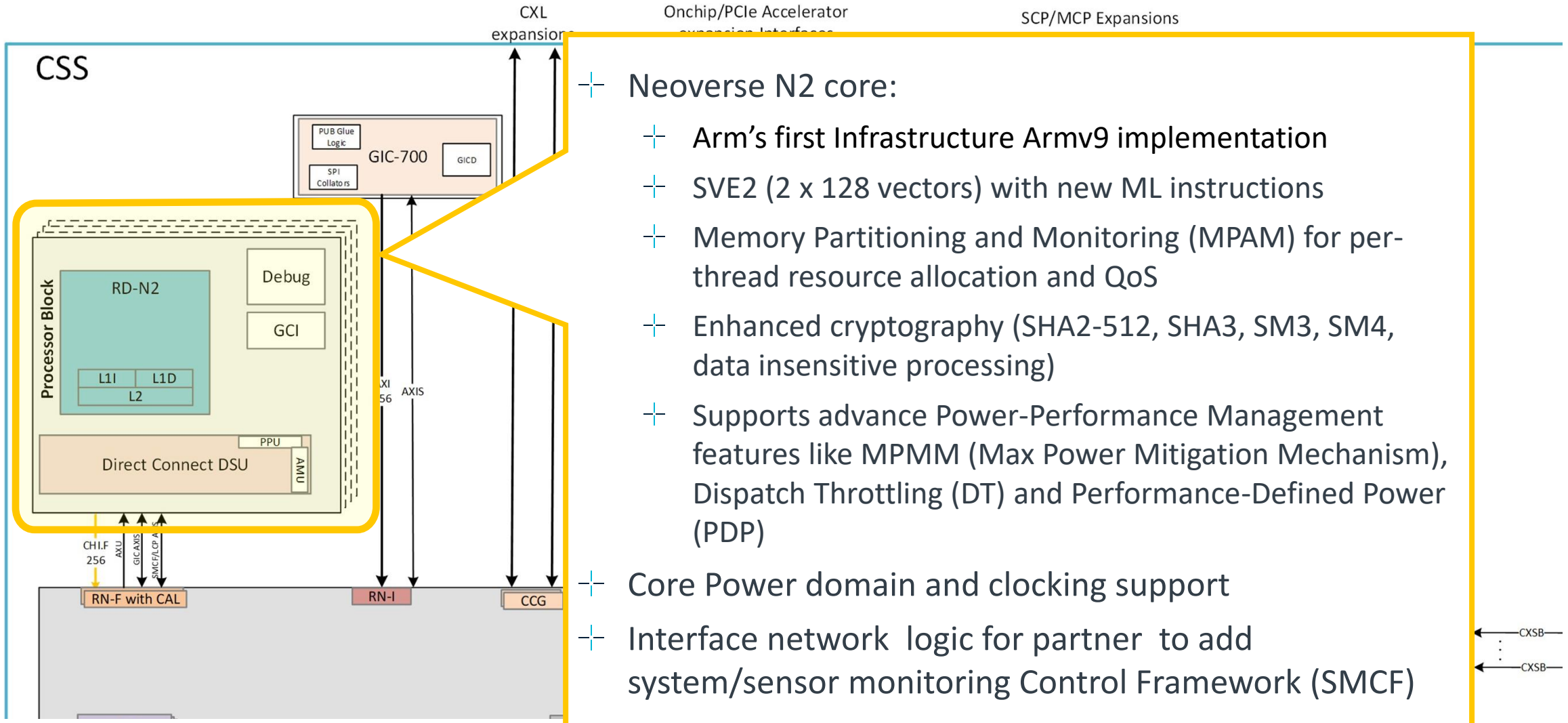
Arm Neoverse CSS N2 Block Diagram

CSS is SystemReady **c**ompliant to:
Arm Base System Architecture 1.0
Arm Server Base System Architecture 6.1
Arm Server Base Boot Requirements 1.2



arm NEOVERSE

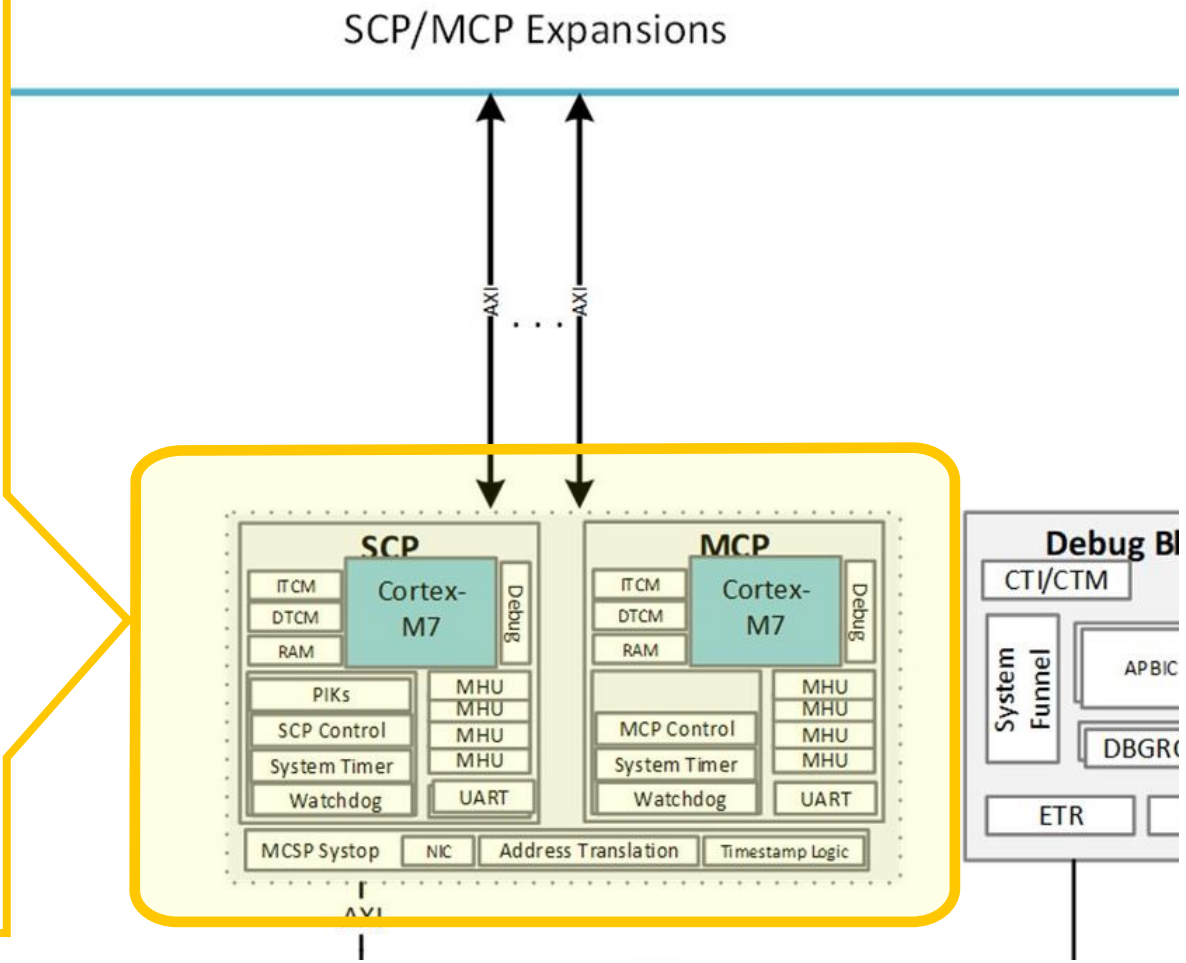
Processor Block



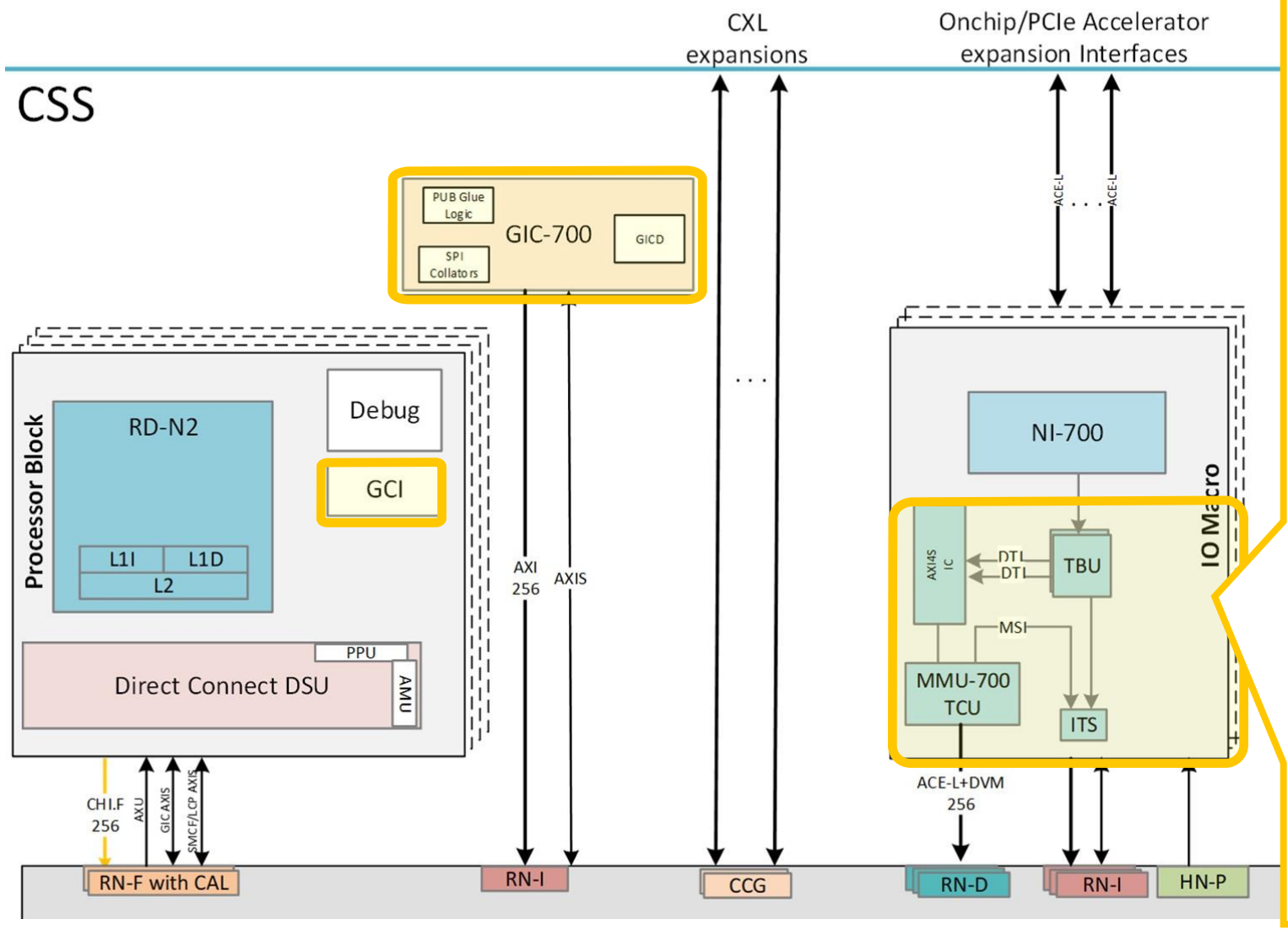
- Neoverse N2 core:
  - Arm's first Infrastructure Armv9 implementation
  - SVE2 (2 x 128 vectors) with new ML instructions
  - Memory Partitioning and Monitoring (MPAM) for per-thread resource allocation and QoS
  - Enhanced cryptography (SHA2-512, SHA3, SM3, SM4, data insensitive processing)
  - Supports advance Power-Performance Management features like MPMM (Max Power Mitigation Mechanism), Dispatch Throttling (DT) and Performance-Defined Power (PDP)
- Core Power domain and clocking support
- Interface network logic for partner to add system/sensor monitoring Control Framework (SMCF)

System Control and Management

- System Control Processor (SCP):
  - SCP is trusted core responsible for controlling system-related functions
  - Controls resets, clock control, power and voltage domains
  - Responsible for all power management control
- Manageability Control Processor (MCP):
  - MCP is a satellite controller communicating with the external Baseboard Management Controller (BMC) for the following functions:
    - On-chip management
    - System reliability, accessibility, and serviceability (RAS) handling
    - Event logging and communication alerts

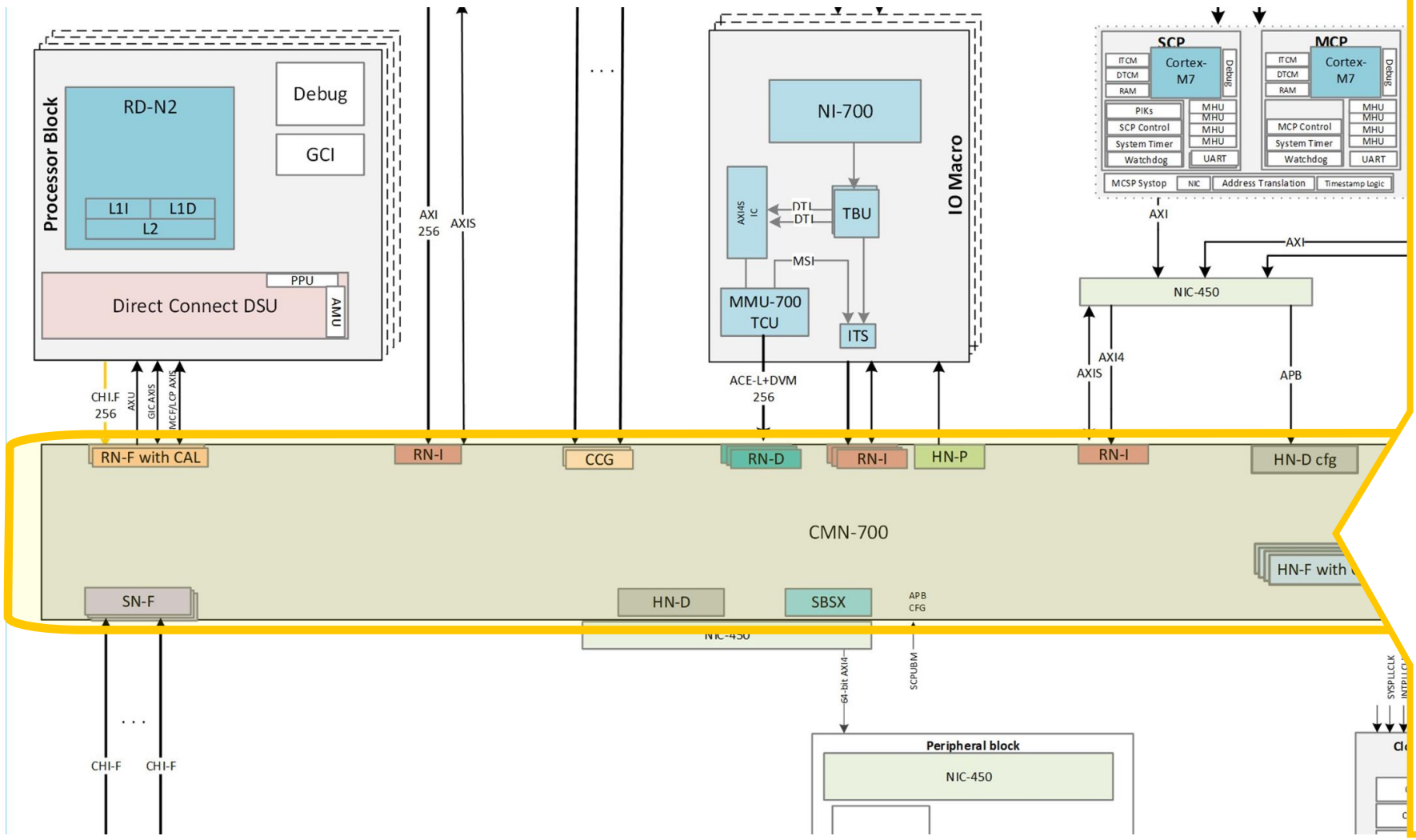


System MMU and Interrupt Controller

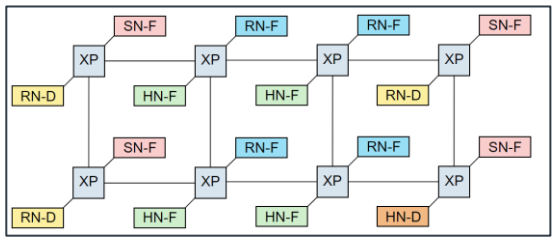


- System MMU (MMU-700):
  - Compliant with SMMU architecture version 3.2
  - Provides address translation and virtualization support
  - Armv9 & Armv8.4 support:
    - MPAM, Secure-EL2 for multiple secure OS's and TLB range invalidate
  - Single or two-stage address translation for PCIe and on-chip accelerators/devices
- GIC:
  - The GIC-700 (GICv4.1) is a generic interrupt controller that handles interrupts from peripherals to core and between cores

System Interconnect



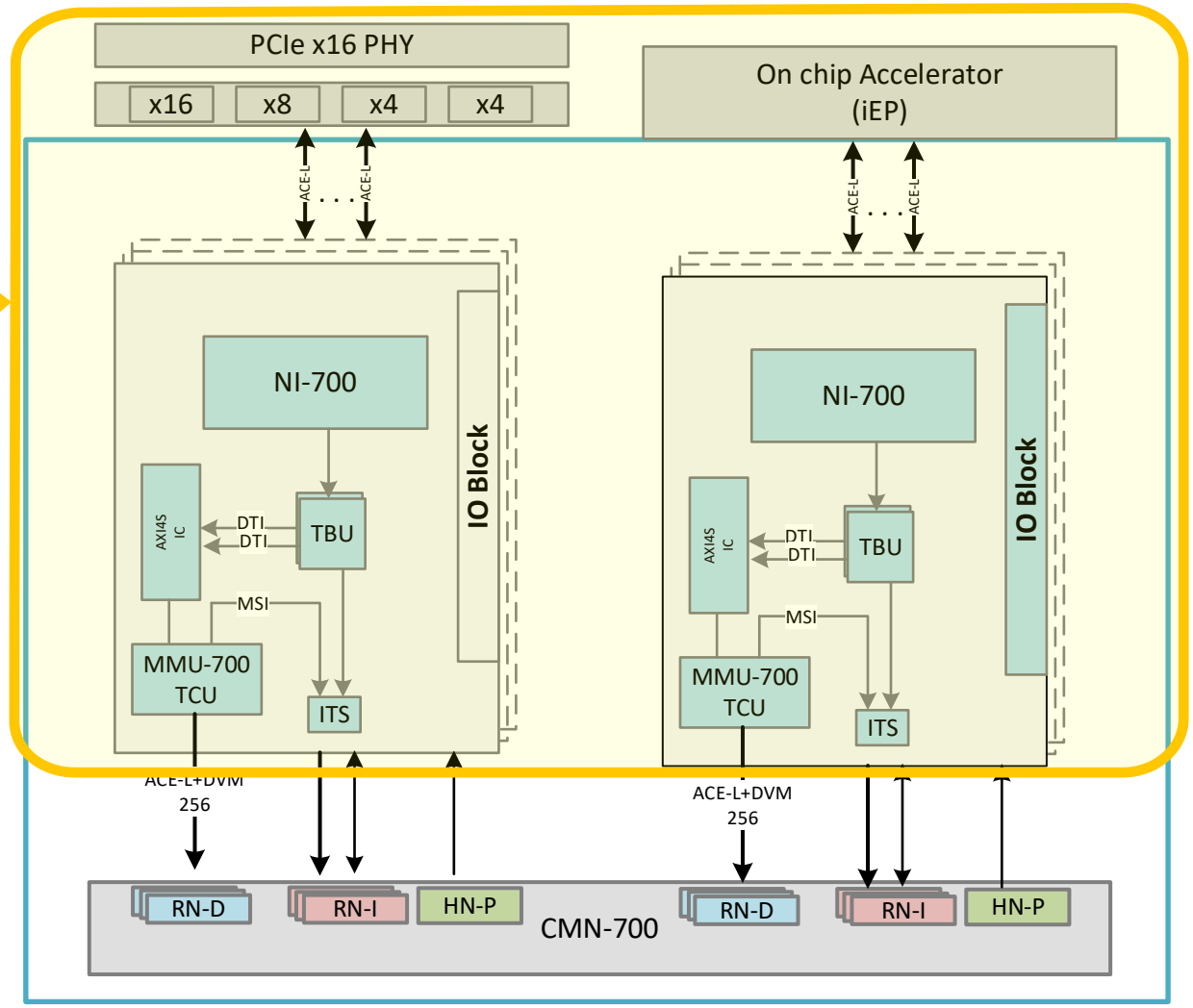
- Neoverse CMN-700:
  - Mesh based coherent interconnect
  - Shared system level cache (SLC) for CPUs and IO
  - Latest Armv9.0-A and AMBA 5 features (Atomics, Stashing, and RAS)
  - System Cache Groups for affinity and isolation
  - Fully coherent multi-chip support



arm NEOVERSE

Accelerator Attach

- CSS N2 provides an IO block with all the interrupt and address translation logic required for partners to add their specific on-chip accelerators or PCIe controllers for external devices

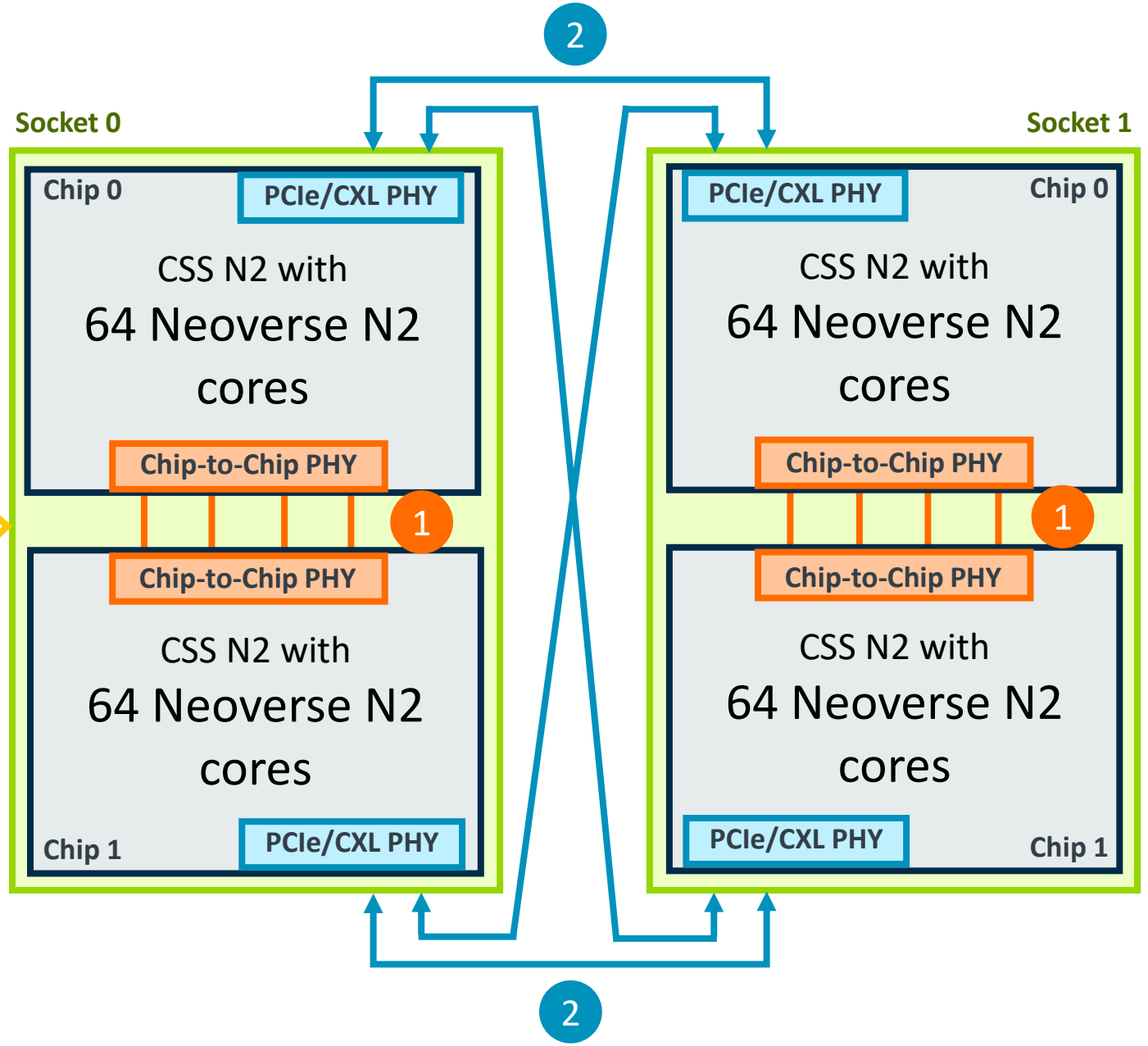


Core Scaling

- CSS N2 allows for connecting
  - 2 chiplets per socket
  - 2 sockets per system
- Scale up to 256 Neoverse N2 cores, delivering more performance in a fully coherent fashion
- SMP protocol for both chip-to-chip (C2C) and socket-to-socket (S2S) communication

1 C2C links, can be UCIe or partner specific C2C PHY

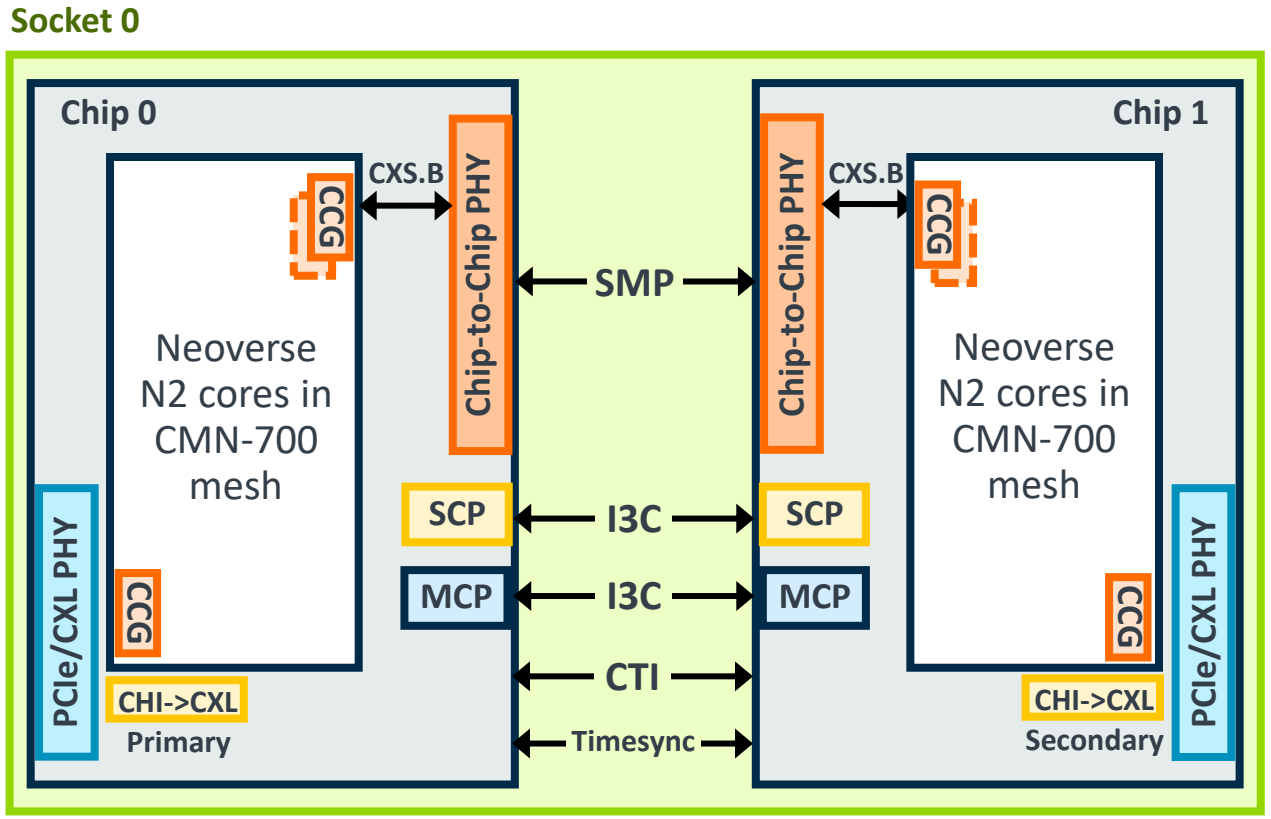
2 Standard CXL PHY for S2S



arm NEOVERSE

Multichip Interfaces

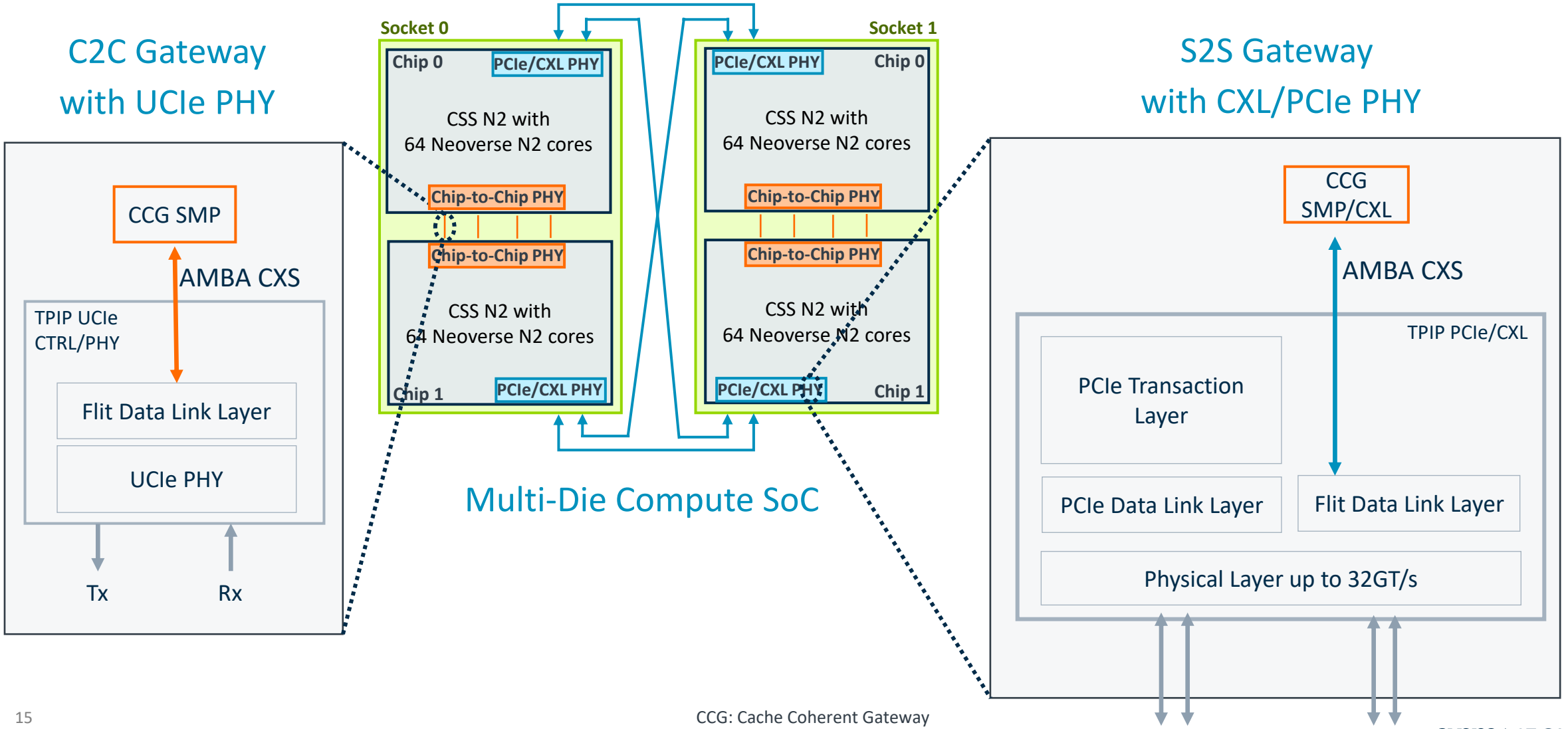
- CSS N2 provides interfaces required for multichip solutions:
  - Coherent SMP interfaces
  - Standard serial interfaces
  - Cross Trigger CTI interfaces
  - Timer synchronization interfaces



CCG: Cache Coherent Gateway

arm NEOVERSE

CMN Gateway for SMP Connections

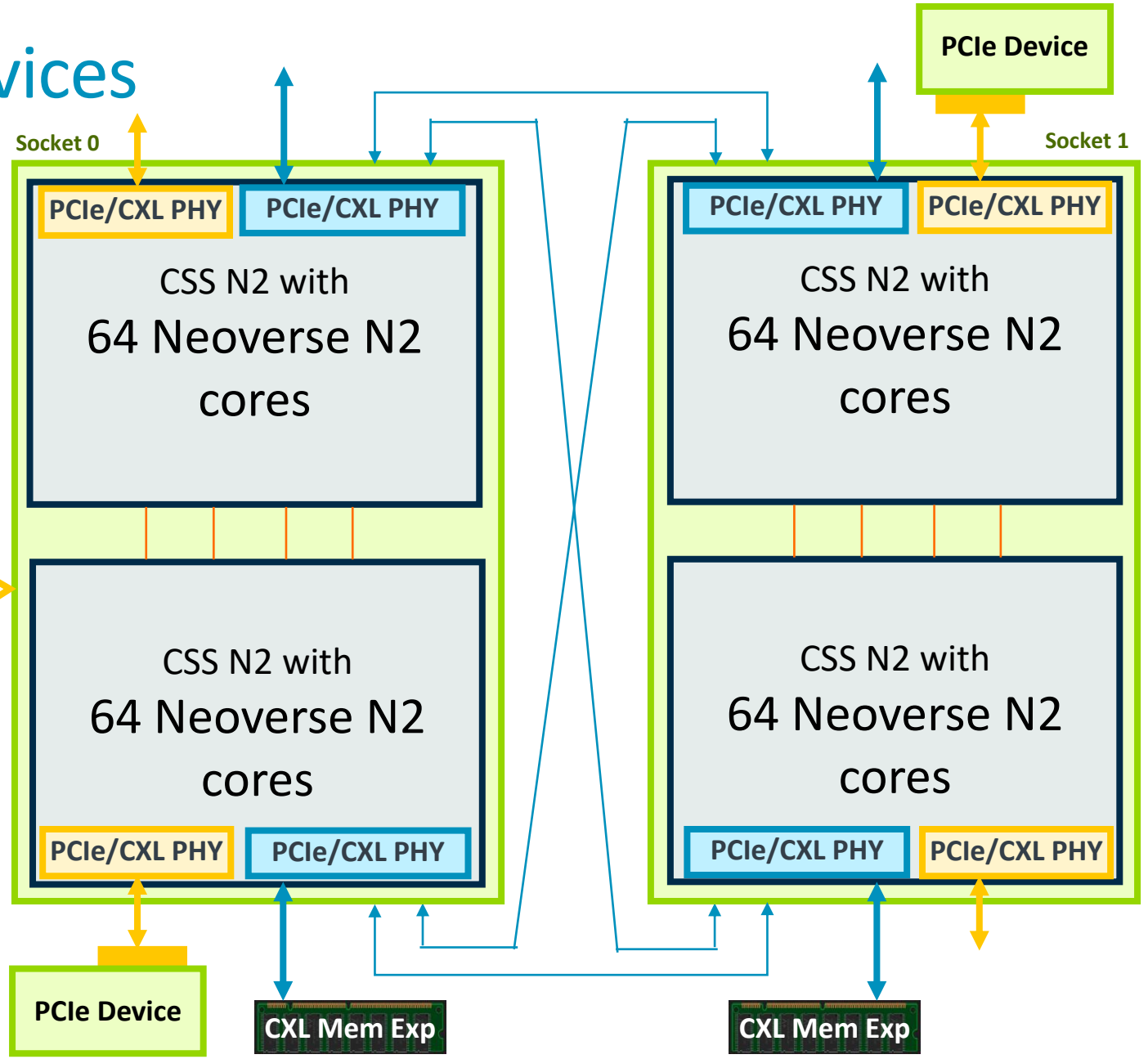


CCG: Cache Coherent Gateway

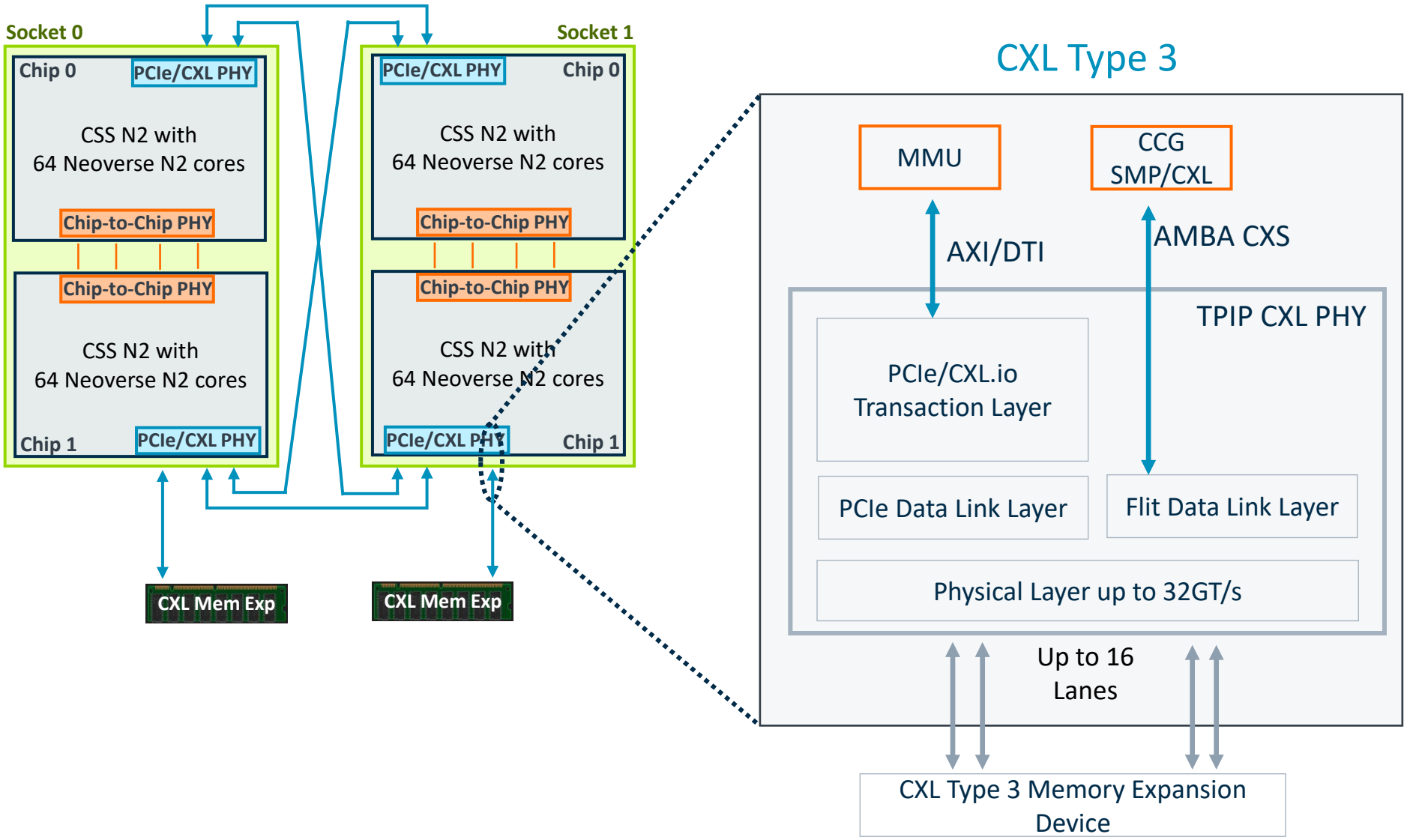
arm NEOVERSE

External CXL and PCIe Devices

- CSS N2 allows partners to expand memory capacity through industry standard CXL1.1 and connect PCIe Gen5 accelerators and devices externally



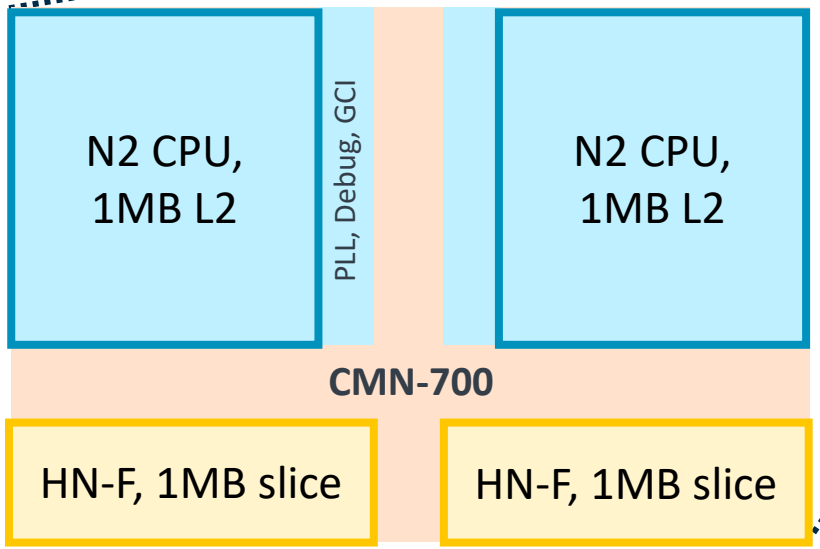
CMN Gateway for CXL Type 3 Memory Expansion

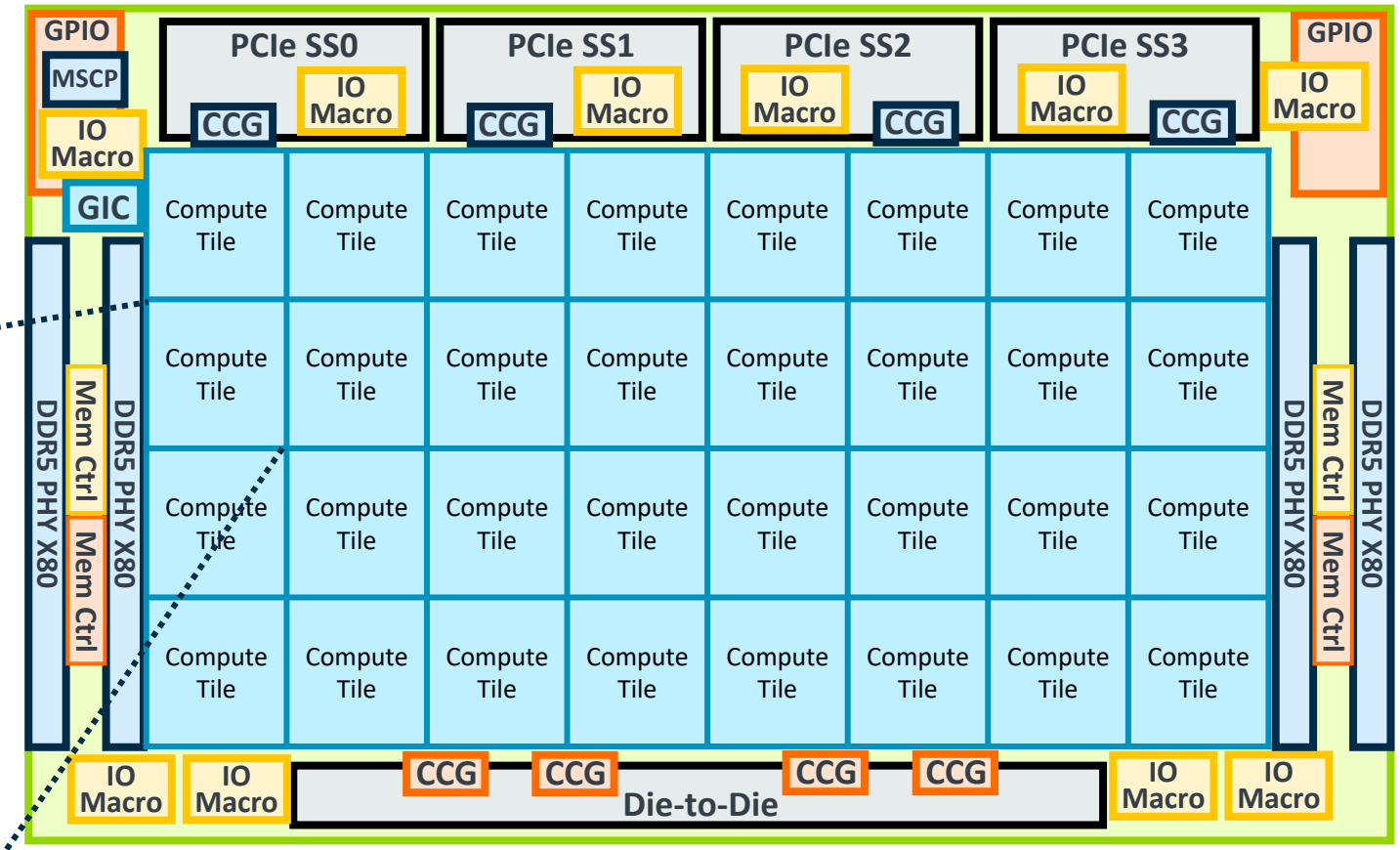


arm NEOVERSE

Arm Neoverse CSS N2 Platform Floorplan Example

- 5nm process, 64 Cores and CMN SLC are laid out in tiled fashion
- Each tile has Dual N2 cores + two HN-F[1] slices
  - Tile Size : 2.54*2.44 mm
- Neoverse CSS N2 Estimated area ~198mm2

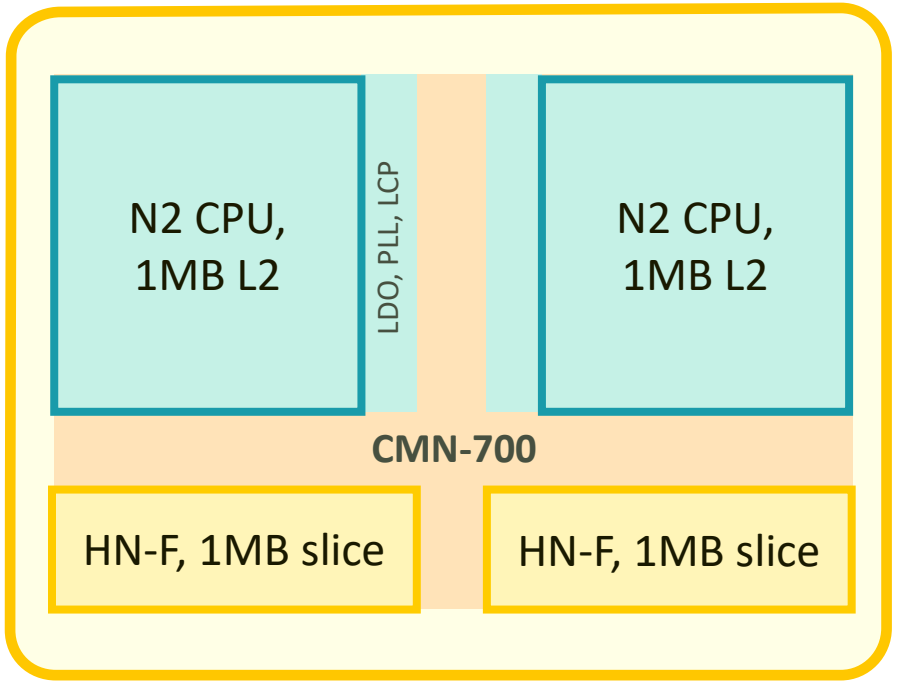




1. HN-F: Home Node, Fully Coherent

arm NEOVERSE

Arm Neoverse CSS Platform Physical IP

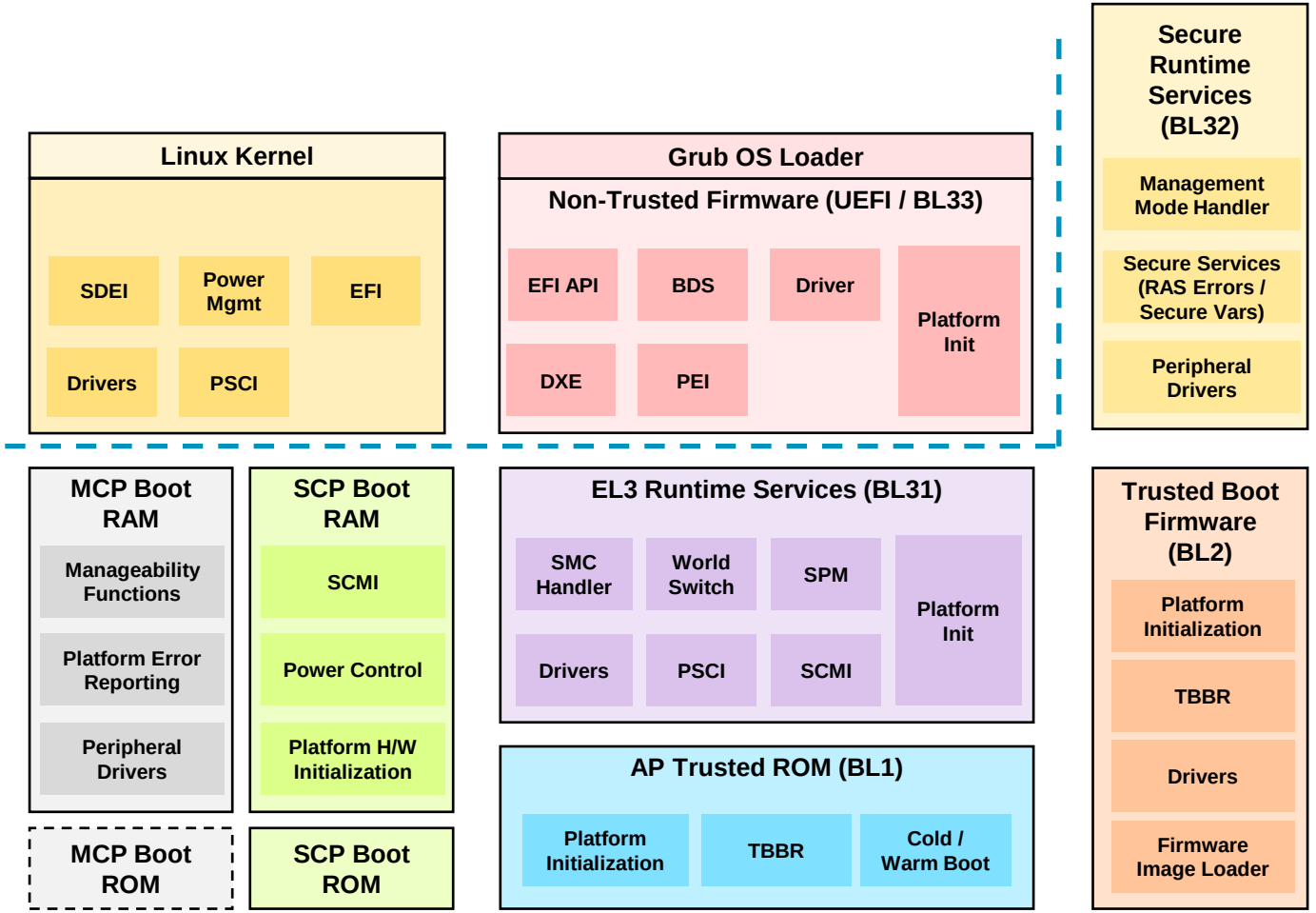


- Arm builds physical IP including standard cell libraries, fast memory cache instances(FCIs) optimized for CSS
- N2 POP – core implementation with optimized PPA
  - Core and mesh configuration
  - Metal stack
  - Floorplan
  - EDA flow
  - Sign-off
- N2 POP is delivered with reference flow scripts, physical IP, supporting utilities and documentation
- CMN-700 and HN-F FCI packages optimized for configuration and PPA

Neoverse Reference Design Platform Software

Integrated reference stack targeting a Fixed Virtual platform (FVP) model of the system

- Designed to provide an extensible reference, bootstrapping partner engineering development. RTL for FPGA build.
- Reference stack to demonstrate:
  - Platform firmware ports for subsystem topology
  - Integrating multiple components (firmware, OS, services)
  - Software features (boot flows, security, power management, etc.)
- Arm Neoverse Platform Software documentation
  - https://neoverse-reference-design.docs.arm.com/en/latest/index.htm



arm NEOVERSE

The Future of Infrastructure is Built on Arm

- Neoverse CSS **enables specialized silicon development** – customers can differentiate memory, IO, acceleration and physical topology on top of Neoverse CSS
- Neoverse CSS **improves time-to-market and reduces development cost** by delivering a pre-integrated, pre-validated, PPA-optimized compute subsystem to customers
- CSS N2 is the **first generation** of Neoverse CSS products, offering customization while leveraging over a decade of platform and software ecosystem standards development

PARTNER A

13

Months from Kick-off to Working Silicon

PARTNER B

80

Engineering Years Saved

arm

Thank You
Danke
Gracias
Grazie
谢谢
ありがとう
Asante
Merci
감사합니다
धन्यवाद
Kiitos
شكرًا
ধন্যবাদ
תודה

arm

Thank You
Danke
Gracias
Grazie
谢谢
ありがとう
Asante
Merci
감사합니다
धन्यवाद
Kiitos
شكرًا
ধন্যবাদ
תודה