ODSA: Technical Introduction

Bapi Vinnakota, Netronome

ODSA Project Workshop

March 28, 2019

ODSA: A New Server Subgroup (Incubation)

- Extending Moore's Law
  - Domain-Specific Architectures: Programmable ASICs to accelerate high-intensity workloads (e.g. Tensorflow, Network Flow Processor, Antminer…)
  - Chiplets: Build complex ASICs from multiple die, instead of as monolithic devices, to reduce development time/costs and manufacturing costs.
- Open Domain-Specific Architecture: An architecture to build domain-specific products
  - Today: All multi-chiplet products are based on proprietary interfaces
  - Tomorrow: Select best-of-breed chiplets from multiple vendors
  - Incubating a new group, to define a new open interface, build a PoC
  - Today is our first workshop as an OCP project!

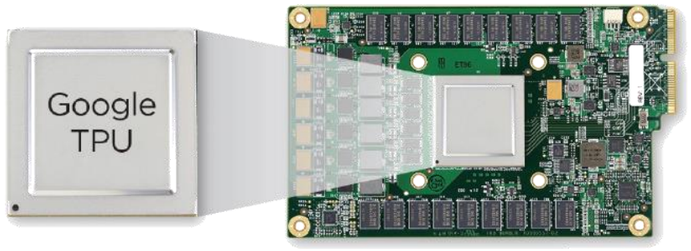
Thanks to:
Achronix: Quinn Jacobson, Manoj Roge; Aquantia: Ramin Farjad; Avera Semi: Dan Greenberg, Mark Kuemerle, Wolfgang Sauter; Ayar Labs: Shahab Ardalan; ESNet: Yatish Kumar; Kandou: Brian Holden, Jeff McGuire; Netronome : Sujal Das, Jim Finnegan, Jennifer Mendola, Brian Sparks, Niel Viljoen; NXP: Sam Fuller; OCP: Bill Carter, Archna Haylock, Dharmesh Jani, Steve Roberts, Seth Sethapong, John Stuewe, Aaron Sullivan, Siamak Tavallaei ; Samtec: Marc Verdiell; Sarcina: Larry Zu; zGlue: Jawad Nasrullah.

Domain-Specific Architectures

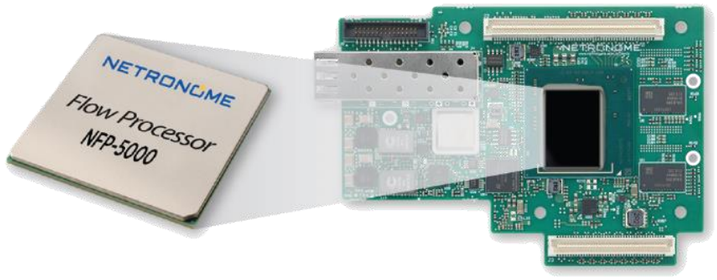
Tailor architecture to a domain*

- Server-attached devices — programmable, not hardwired
- Integrated application and deployment-aware development of devices, firmware, systems, software
- 5-10X power performance improvement

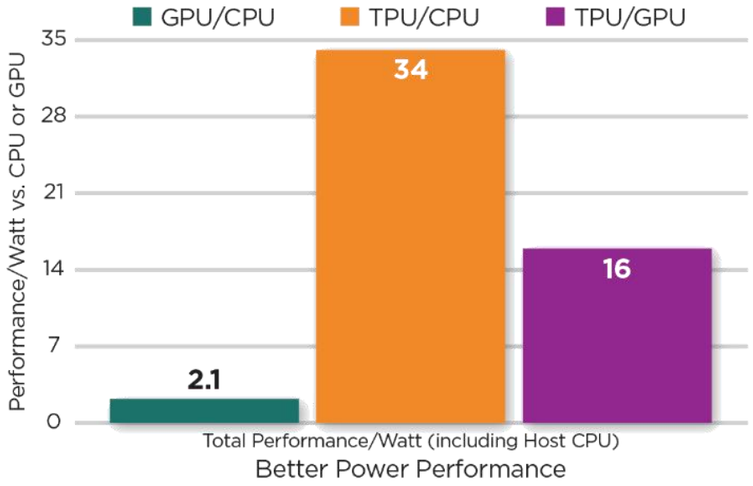
- Big - more of a processor to I/O mismatch => more memory
- Each serves a smaller market



Domain-Specific for Machine Learning and AI

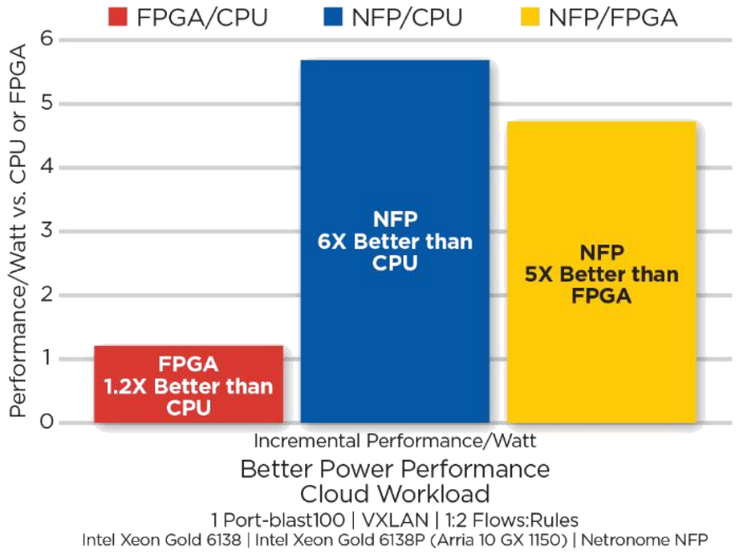


Domain-Specific for Networking and Security



Google TPU vs. CPU and GPU

Source: "An in-depth look at Google's first Tensor Processing Unit (TPU)," Google Cloud, May 2017

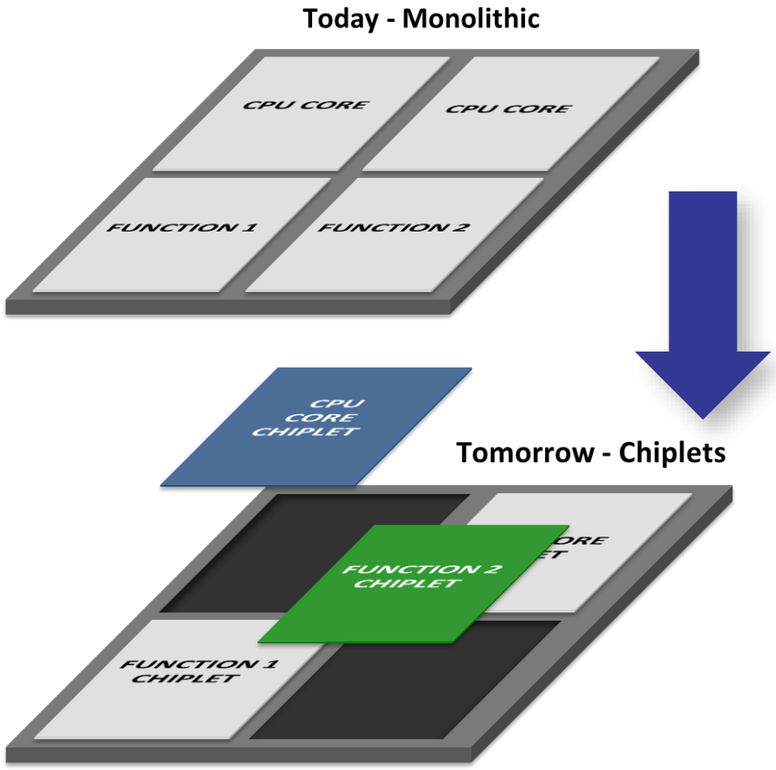


Netronome NFP vs. CPU and FPGA

Source: Netronome, based on internal benchmarks and industry reports related to Xeon CPUs and Arria FPGAs

*A New Golden Age for Computer Architecture
John L. Hennessy, David A. Patterson
Communications of the ACM, February 2019, Vol. 62 No. 2, Pages 48-60

Monolithic vs Chiplets

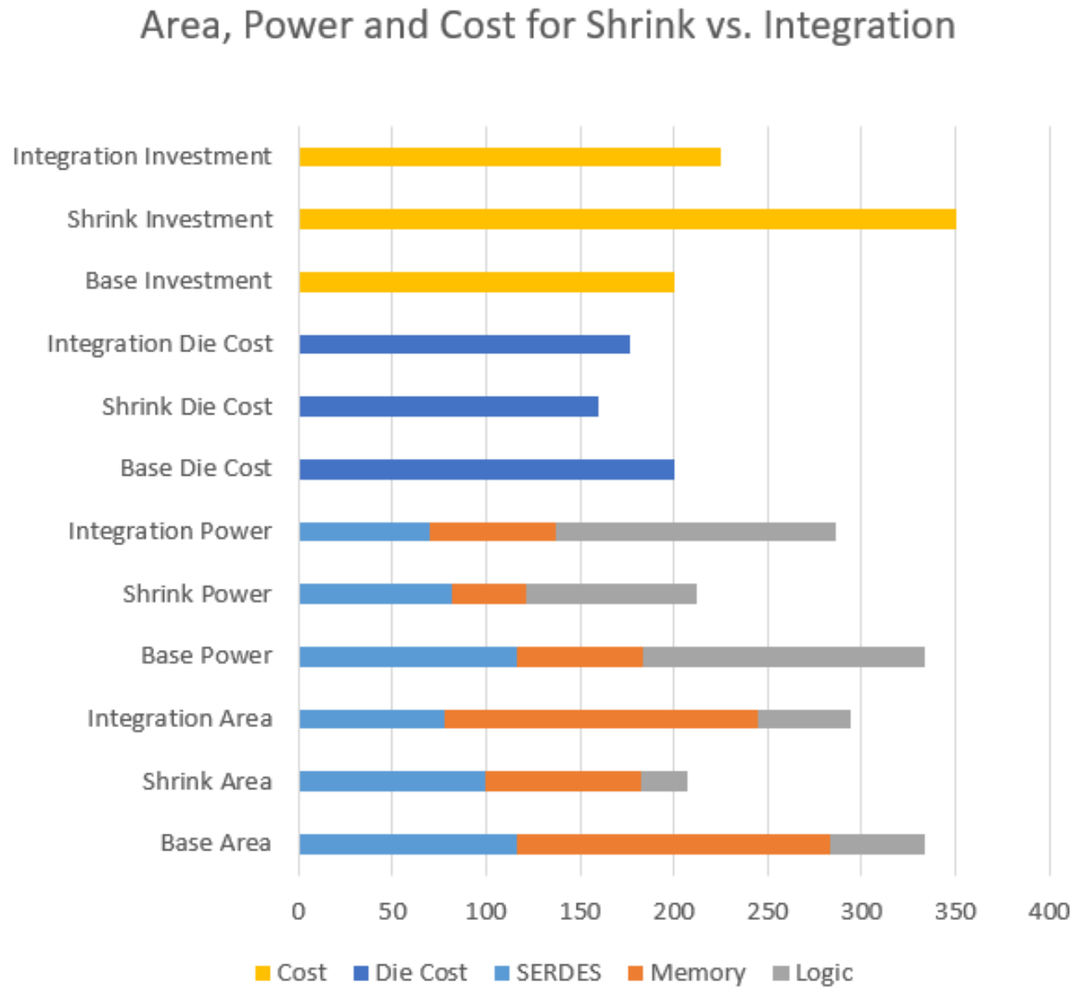


AMD Data

4 Die are ~30% cheaper than a single large die

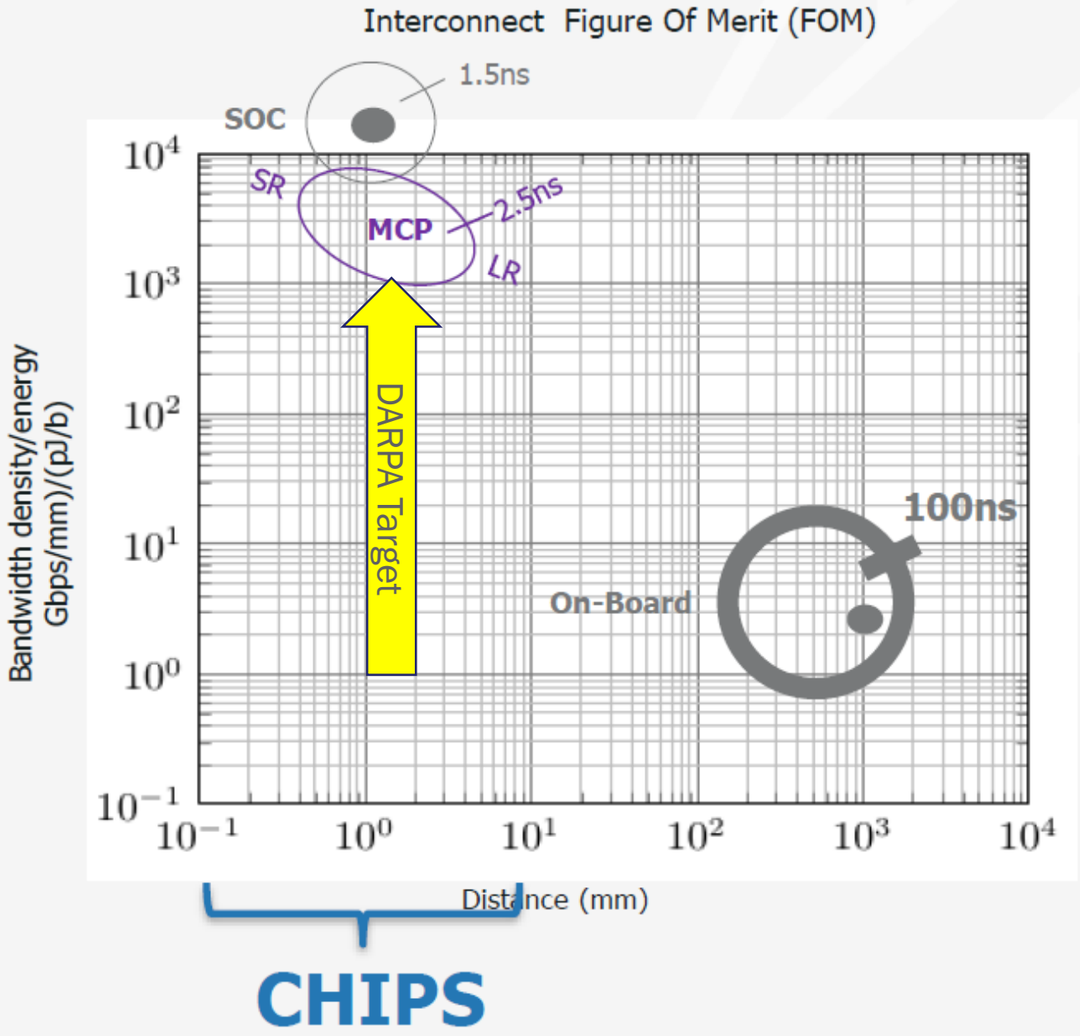
Shrink: Monolithic process shrink
Integration: Multi-chip on same process

Integration provides nearly all the benefits of a shrink at a fraction of the cost, because of efficient inter-chiplet interconnect



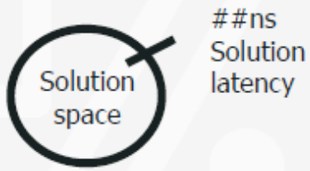
https://www.netronome.com/media/documents/WP_ODSA_Open_Accelerator_Architecture.pdf

COST & PERFORMANCE DISPARITY IN SCALING



CHIPS

Legend:

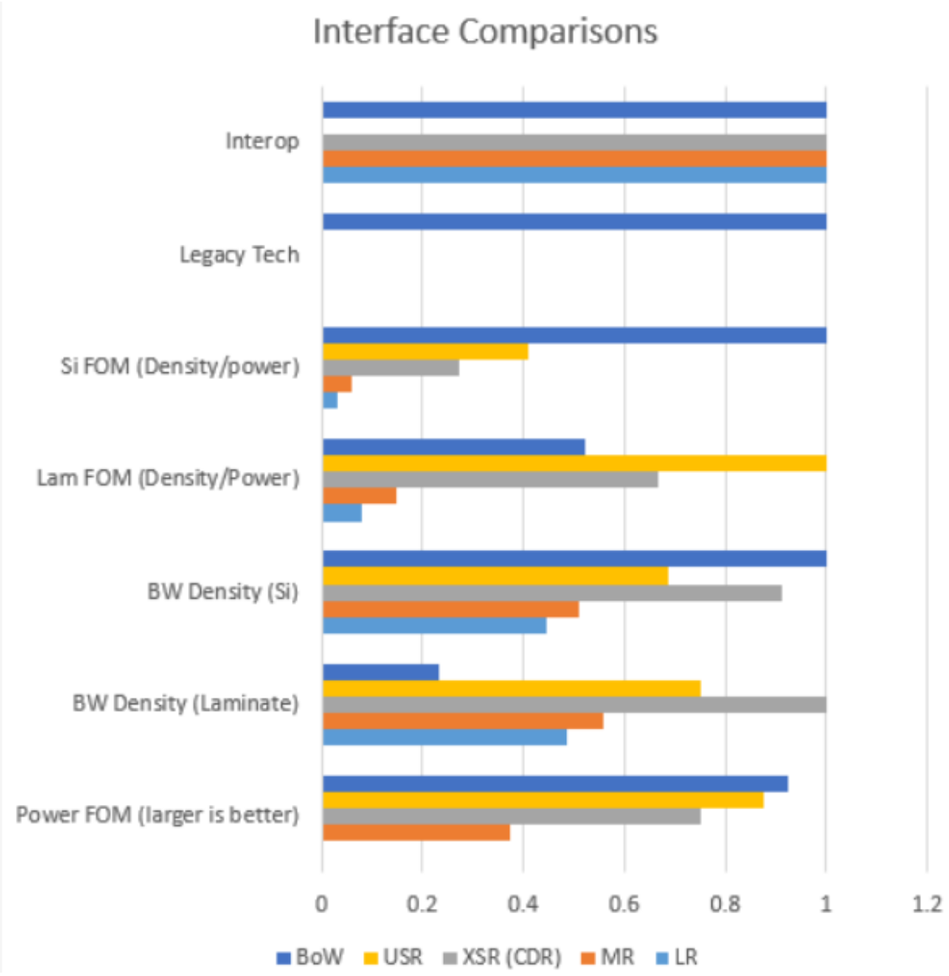


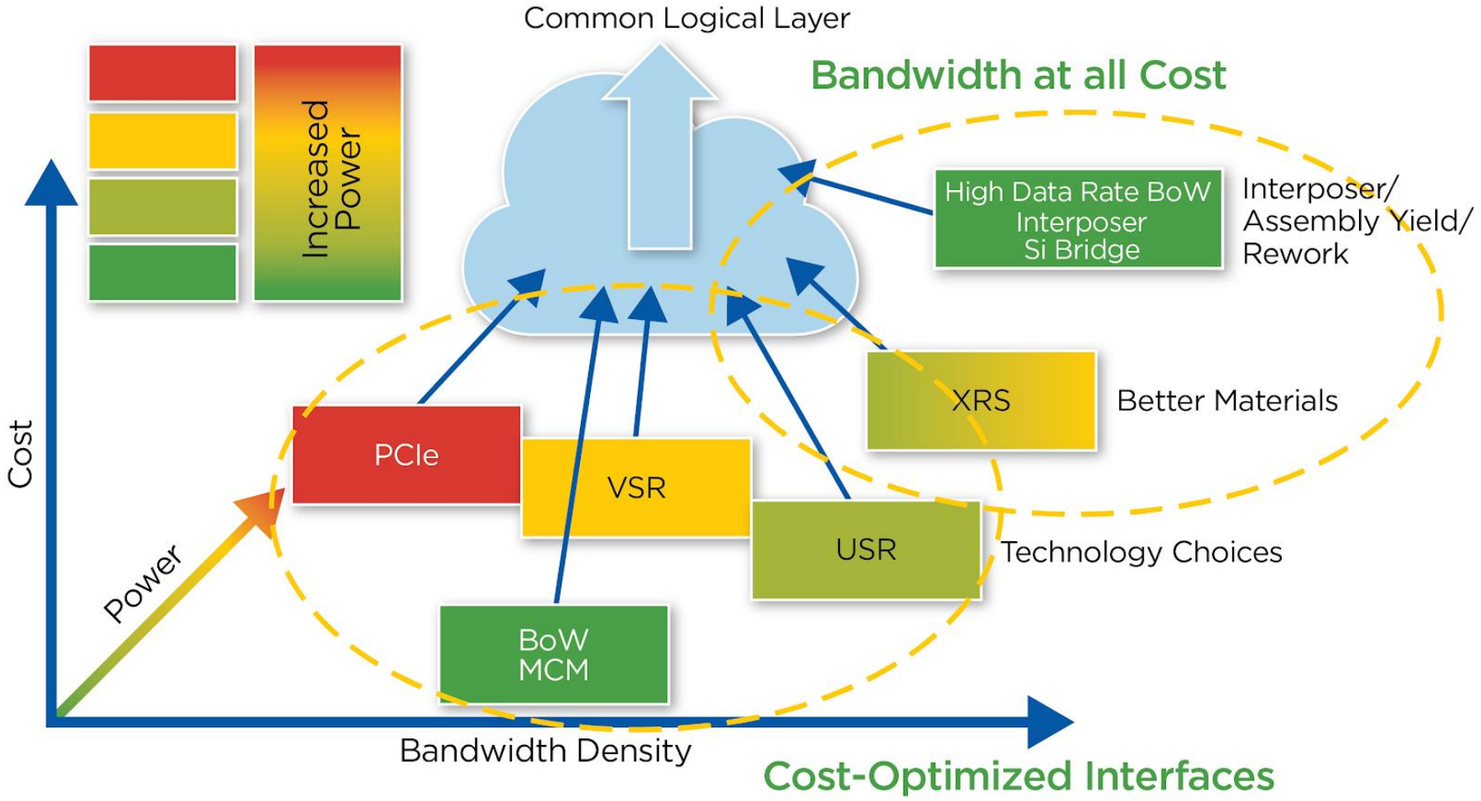
High Density Multi Chip Packaging interconnect

- 0.1-10mm; 500IO/mm shoreline; 0.1-1.0pJ/bit
- $FOM_{LongReach}$=(1Tb/mm)/(1pJ/b)=1,000
- $FOM_{ShortReach}$=(1Tb/mm)/(0.1pJ/bit)=10,000
- Latency=2.5ns

SOC-like FOM

PHY Layer Options

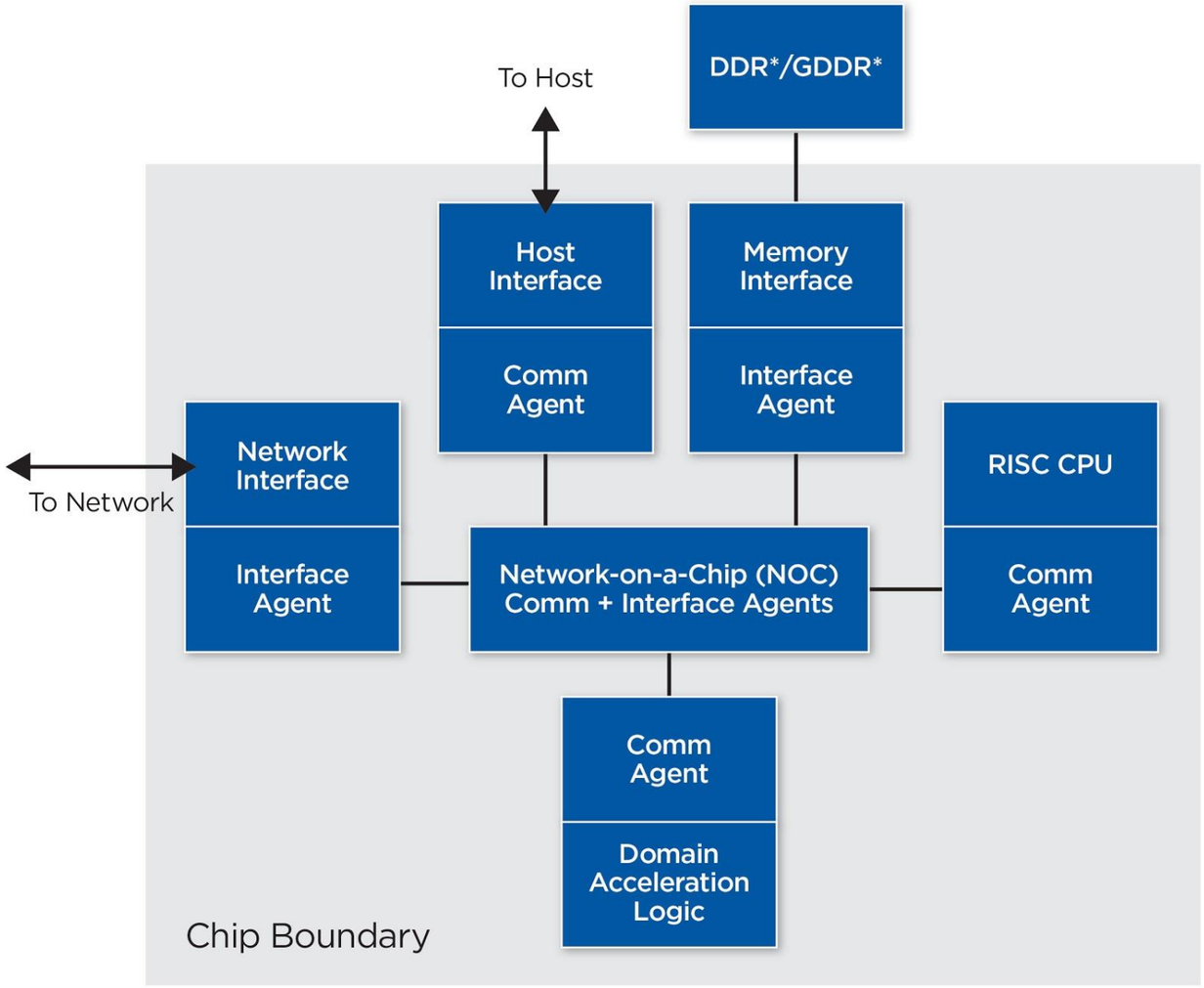




https://www.netronome.com/media/documents/WP_ODSA_Open_Accelerator_Architecture.pdf

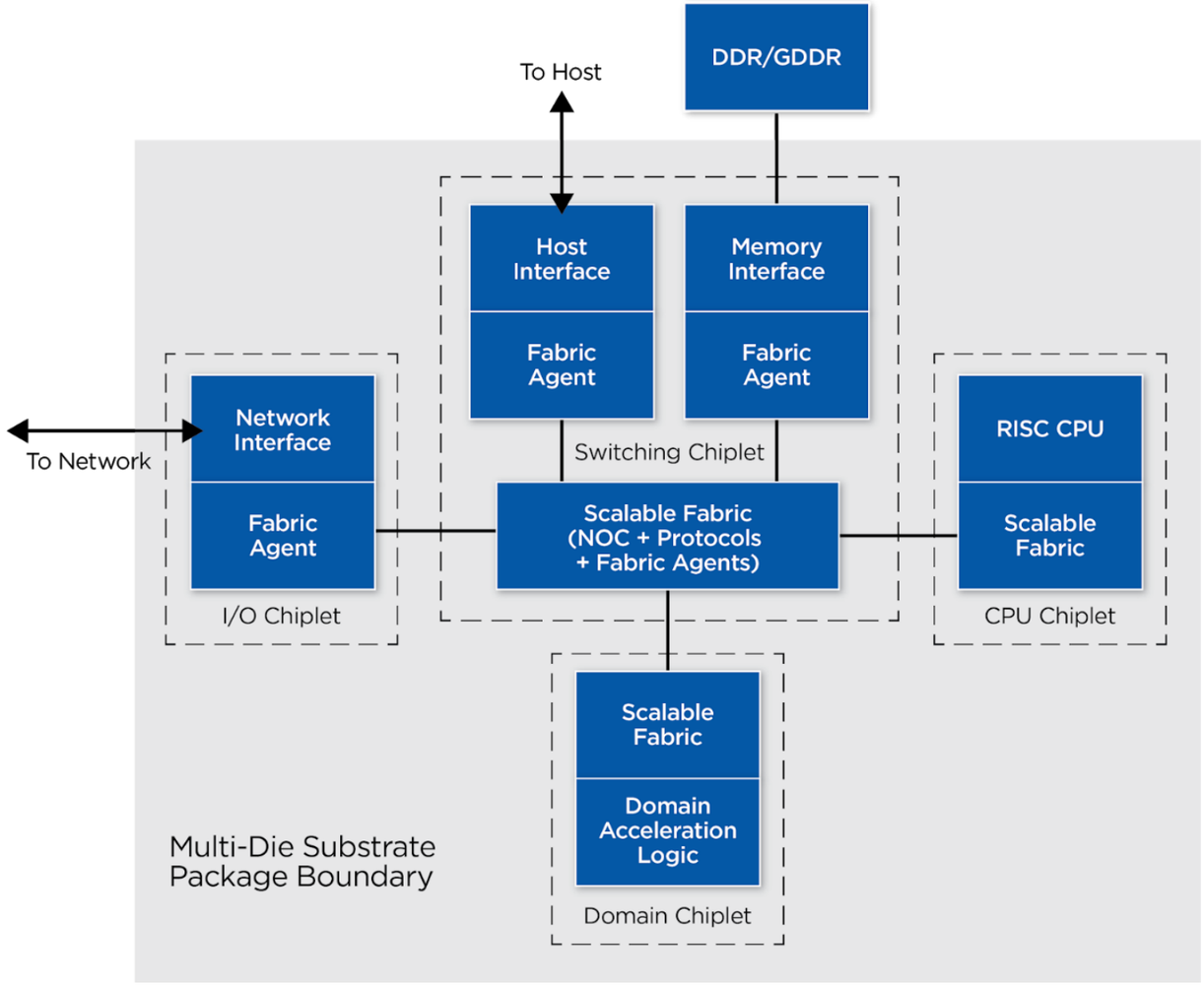
Domain-specific accelerators

- Host-attached programmable logic optimized for an application domain
  - Tensorflow, Netronome NFP, Crypto, IoT,…
- Domain-specific accelerators contain lots of generic logic ~35-45% of silicon area, development time
  - Network, Host, Memory Interfaces
  - General-purpose CPUs
  - SRAM, interconnect
  - Domain-specific logic works in coordination with host and/or CPU SW
- Ideally
  - Investment in a DSA should be limited to the domain-acceleration logic
- In reality
  - Buy IP for the “non-core” parts, spend $$’s test and integration

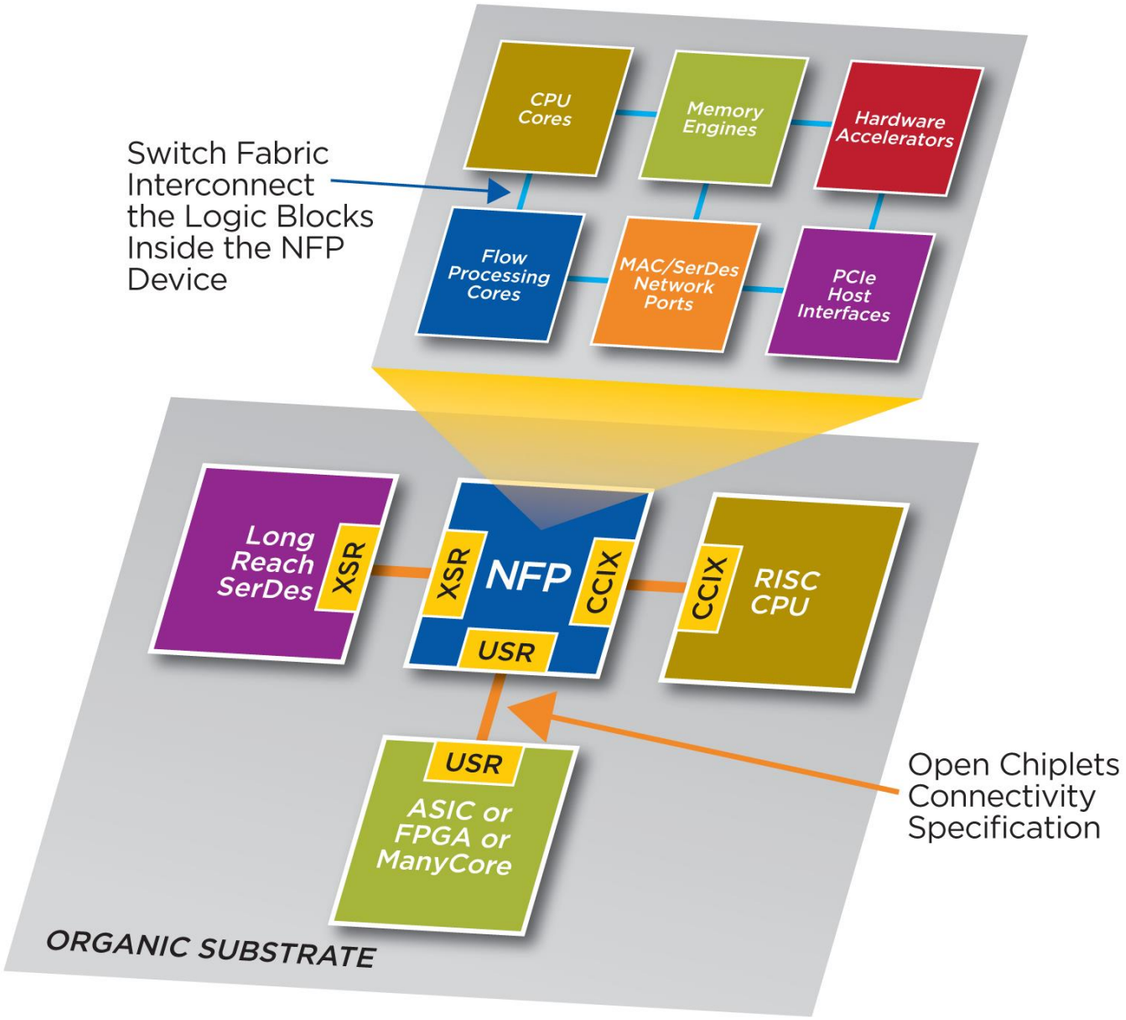


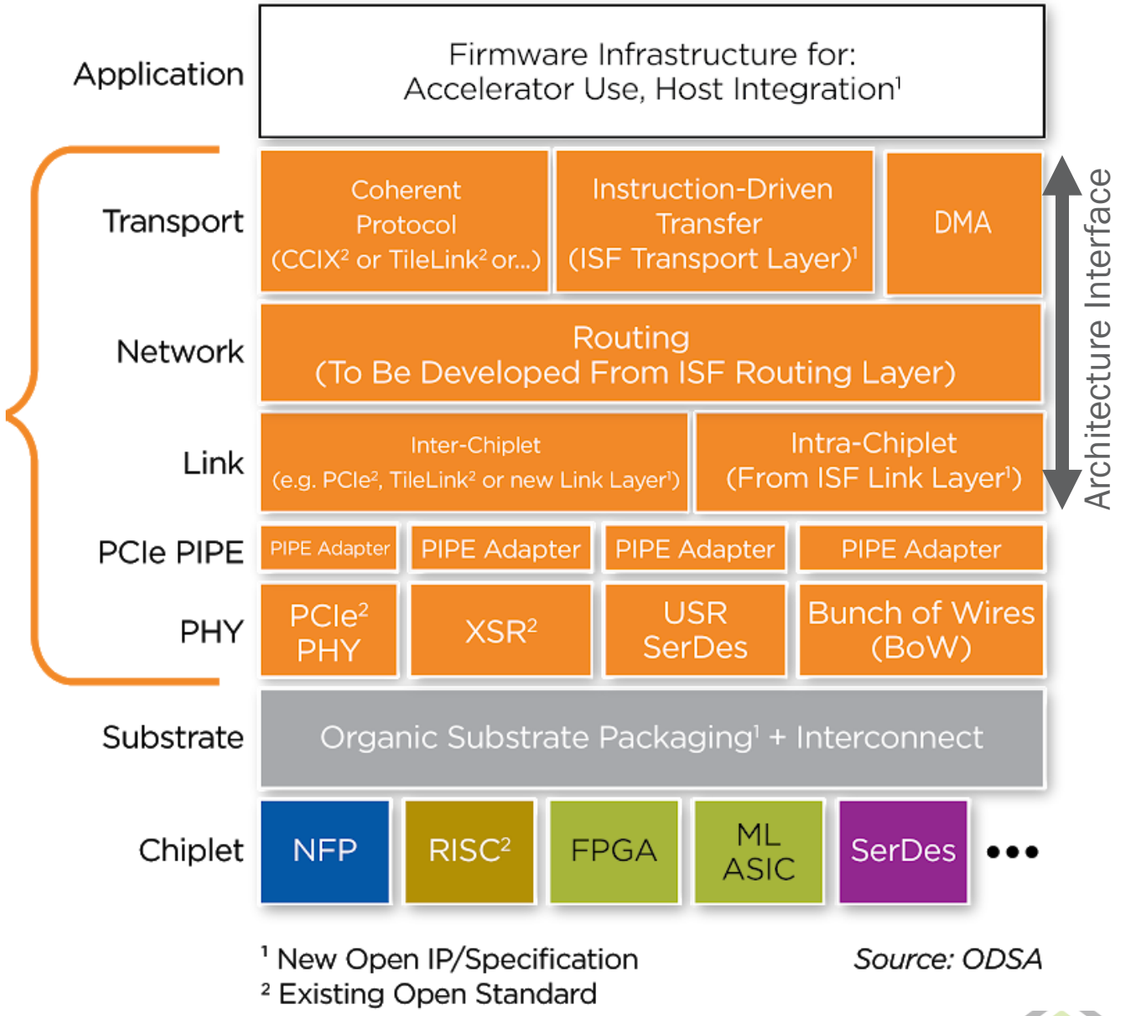
Multi-Chiplet Reference Architecture for DSA

Design Function	Value
IP Qualification	Verified IP for inter-chiplet communication
Architecture	Leverage reference architecture.
Verification	Focus investment on domain-specific logic.
Physical	Reuse chiplets instead of IP for 40% of the functions in a monolithic design
Software	Open source firmware and software for host-attached operation
Prototype	Aim for reference package design with area, power budgets and pinouts for components
Test and Validation	Develop workflow for chiplets



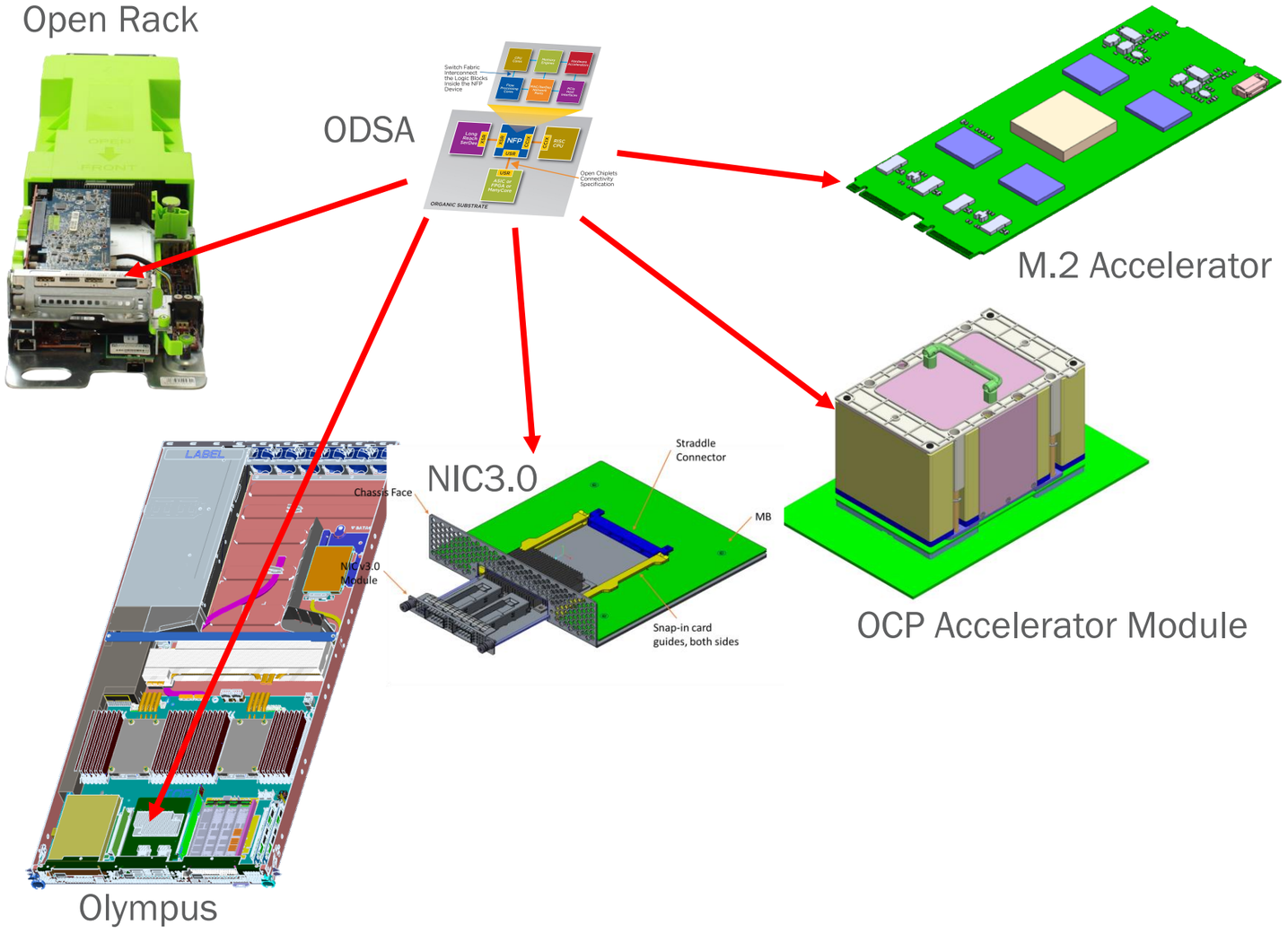
Open Interface for Chiplet-Based Design





Multiple chiplets need to function as though they are on one die

Need a Scalable Interface



Multiple OCP projects use accelerators

Open architectural interface to support accelerator designs across multiple carrier cards

Power, management, reliability requirements vary across sockets

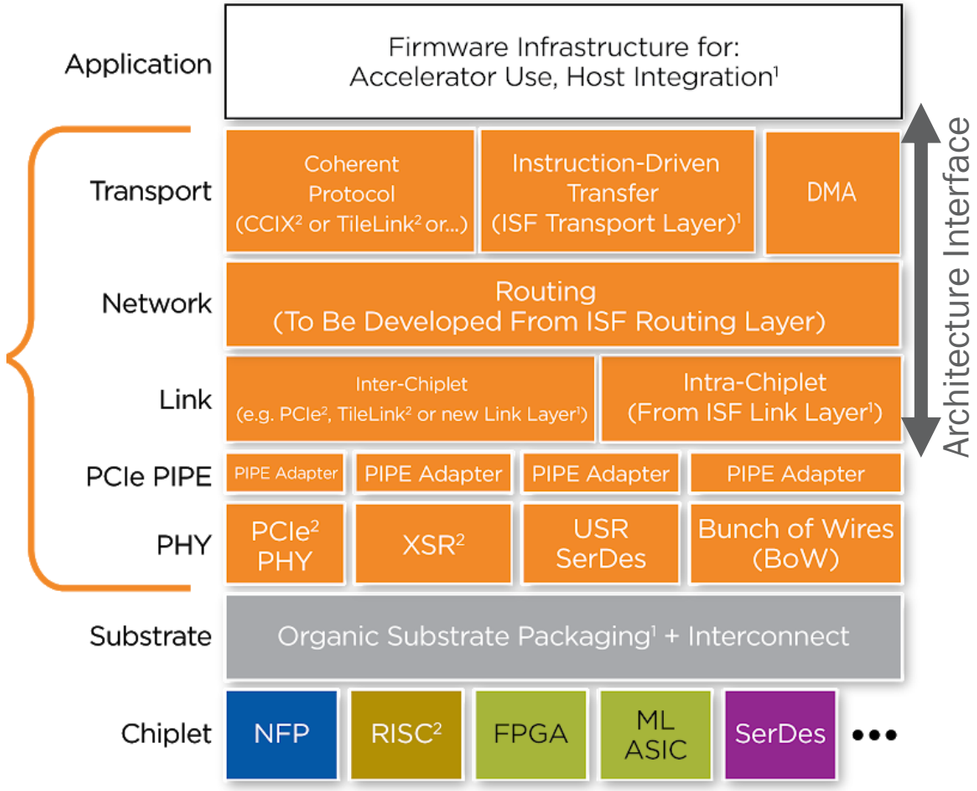
Enable a collection of ODSA-compliant chiplets, packages, sockets, in the OCP marketplace

ODSA Landing Zones

	Network I/O	Host I/O	Power	Size
NIC 2.0	Dual port x 25	X16 PCIe Gen 3	25w	
NIC 3.0	Dual port x 200	SFF: x16 PCIe Gen 4/Gen 5 LFF: x32 PCIe Gen 4/Gen 5	Small: 80w Large: 150w	Small/Large
M.2 M.2 Dual	N/A	Single: x4 PCIe Gen 3/Gen 4 Dual: x8 PCIe Gen 3/Gen 4	Single: 12w Dual: 20w	Single: 22x110 Dual: 46x110
OAM	8x16 SerDes Lanes	Typical: x16 PCIe	12V: 350w 48V: 700w	102x165
Olympus	Via x16 PCIe Cards	1x16 PCIe	75W-300W PCIe AIC	FHHL PCIe
Tioga Pass	Up to 100Gbps SH	x32 PCIe Gen3		6.5x20inch

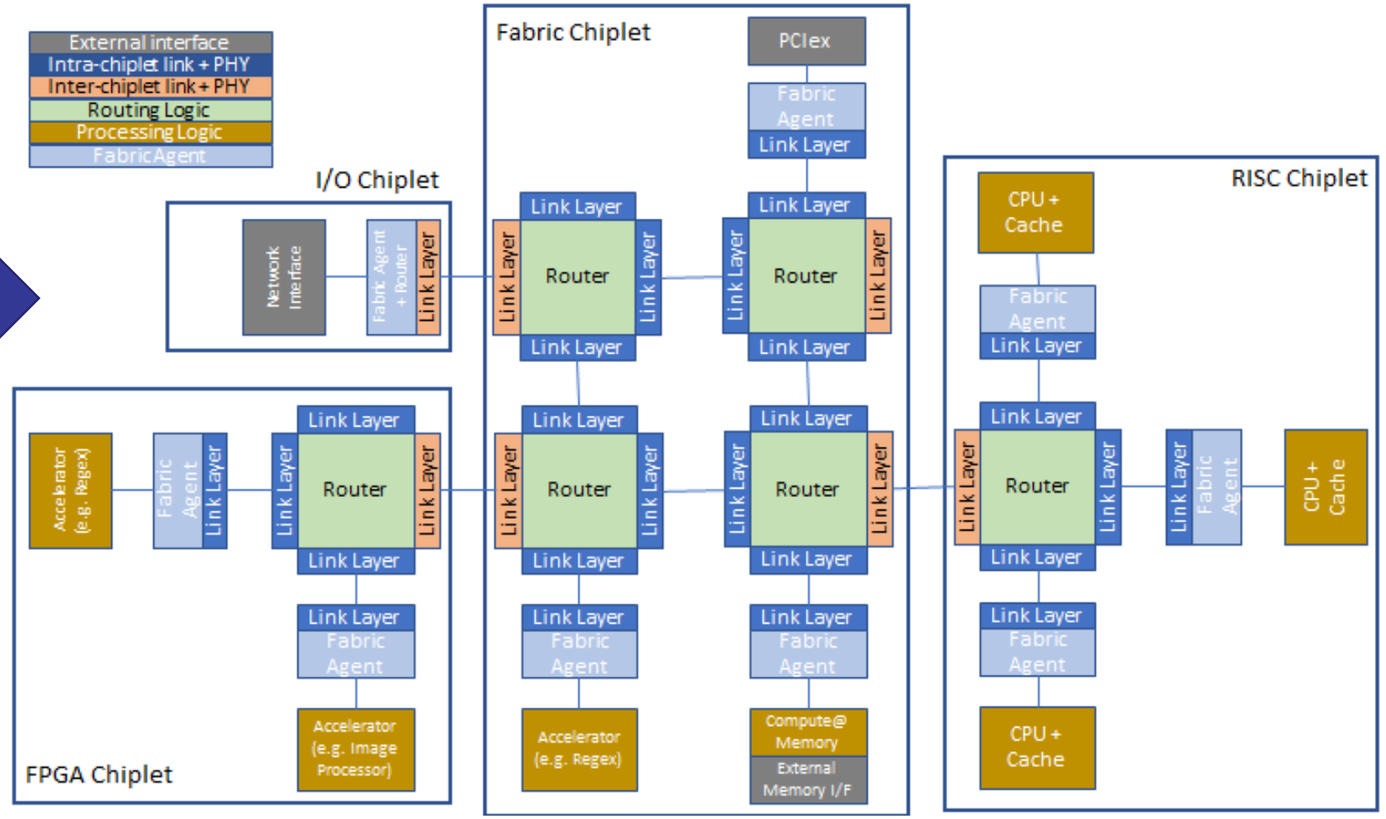
Data from Ron Renwick, John Stuewe, Siamak Tavallaei, Whitney Zhao

Cross-chiplet ODSA fabric proposal



[1] New Open IP/Specification
[2] Existing Open Standard

Source: ODSA



Progress Since the Last Workshop

- Timeline:
  - ODSA Announced 10/1/18 7 companies
  - White Paper 12/5/18 10 companies
  - First Workshop 01/28/19 35 companies
  - Joined OCP 03/15/19
  - Today 03/28/19 53 companies
- PoC
  - Identified components, use cases
- Standards
  - Characterizing PHY, new interface proposal
- Business
  - Survey, business model

TIL in the last six months

- We're solving the right problem, tbd on whether it's the right solution.
- Analog (and cache coherence) engineers have lots of opinions, likely justified, but also confusing for mere mortals.
- **How you do business drives chiplet economics and your technology choices.**
- Our interface definition must recognize this diversity while focusing our effort.
- You need a new business/workflow model that make chiplets work across this diversity

~~How to Participate~~
Please Help! : Join a Workstream

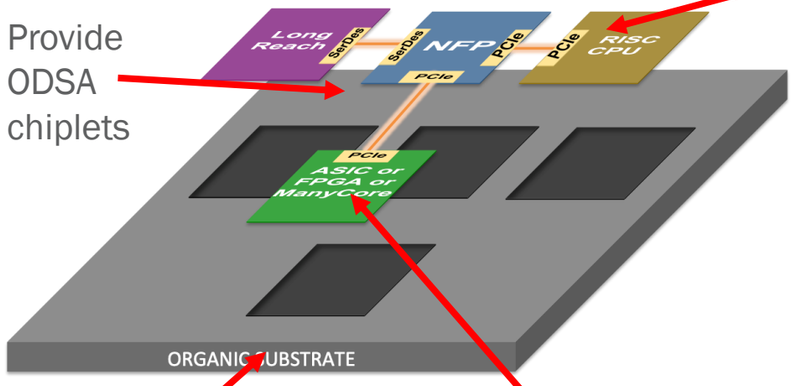
Join the PoC, Build fast:
(Quinn Jacobson/Jawad Nasrullah)

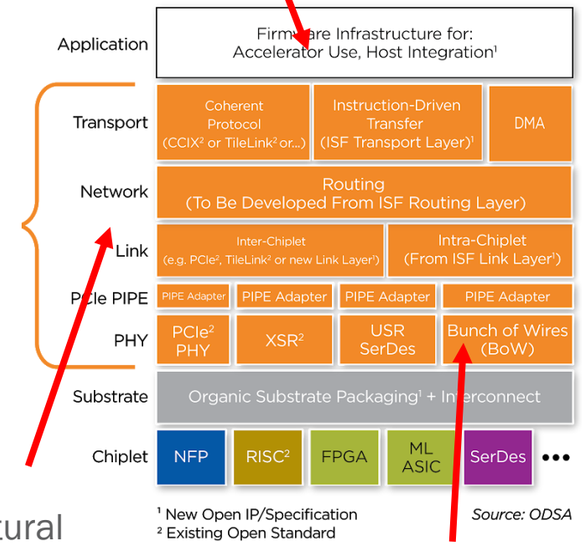
Join Interface/Standards:
(Mark Kuemerle/Aaron Sullivan)

Join Business, IP and workflow:
(Sam Fuller/Jeff McGuire)

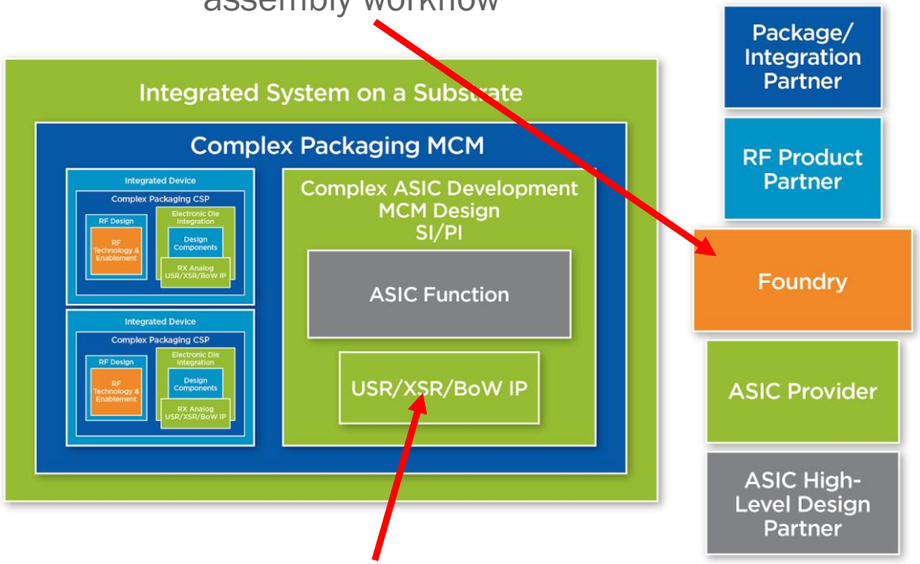
Develop software

Provide ODSA chiplets





Define test and assembly workflow



Develop Packaging + Socket, Dev Board

Provide FPGA IP

Define Architectural Interface

Provide PHY technology

Provide Chiplet IP

Workstream contact information at the ODSA wiki

Consume. Collaborate. Contribute.