Open. Together.

OCP
SUMMIT

EW-Server

Gen-Z Technology: Enabling Memory Centric Architecture

Greg Casey, Server Strategist, Gen-Z & Dell/EMC

THE GEN - Z CONSORTIUM

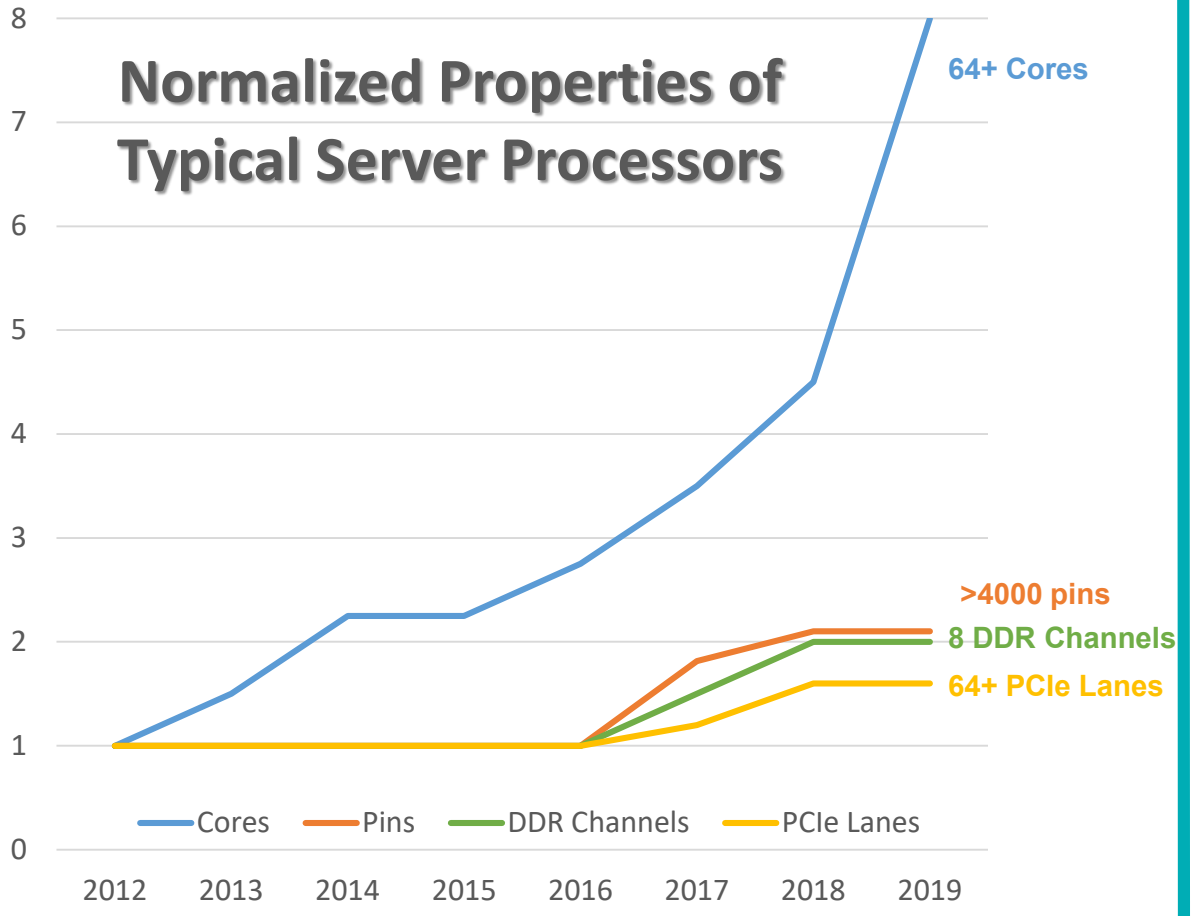
- Launched in October of 2016 to create an open, industry standard for a high speed, low latency, scalable, memory centric fabric
- Demos of memory pooling with multiple servers shown over the 2 years
- Members have released design IP and silicon vendors have started detailed designs for Gen-Z devices
- Released and draft GenZ Documentation is available for public review and comment: **www.GenZConsortium.org**

GEN Z Consortium Members

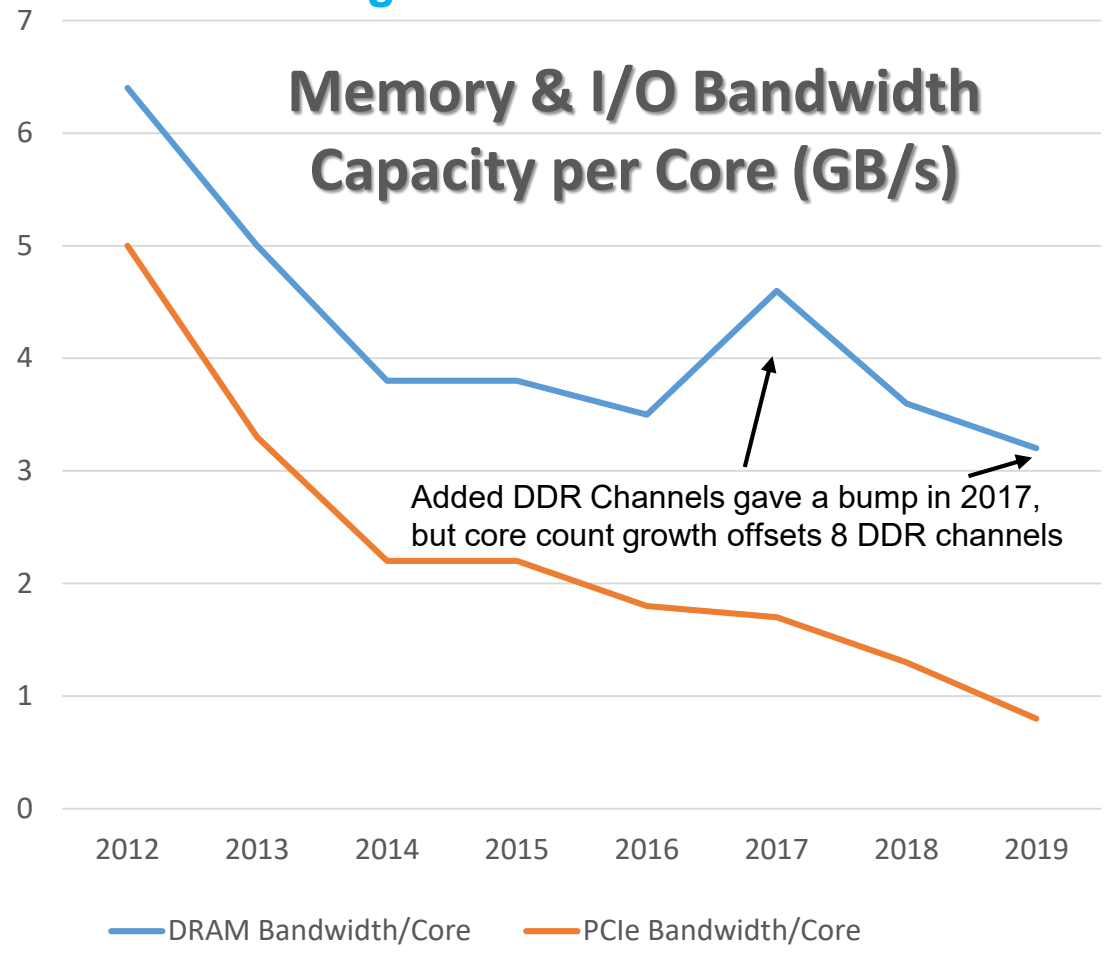
- Aces
- Allion Labs
- AMD
- Amphenol
- Arm
- Avery Design Systems
- Broadcom
- Cadence
- Cisco
- Cray
- Dell EMC
- Everspin
- ETRI
- FIT
- Genesis Connected
- Google
- H3C
- Hitachi
- HP
- HPE
- Huawei
- IBM
- IDT
- IntelliProp
- ITT Madras
- Jess Link
- Keysight
- Lenovo
- Liqid
- Lotes
- Luxshare-ICT
- Marvell
- Mellanox
- Mentor
- Micron
- Microsemi
- Microsoft
- Mobiveil
- Molex
- NetApp
- Node Haven
- Nokia
- Oak Ridge Natl Labs
- PLDA Group
- Qualcomm
- Red Hat
- Samsung
- Samtec
- Seagate
- Senko Advanced Comp
- Simula Research Lab
- SK hynix
- Smart Modular
- Sony Semi
- Spin Transfer Tech
- Teledyne LeCroy
- TE
- Toshiba Memory Corp
- Univ. New Hampshire
- VMware
- Western Digital
- Xilinx
- Yadro
- Yonsei University
- 3M

Computer-Memory Balance is Degrading

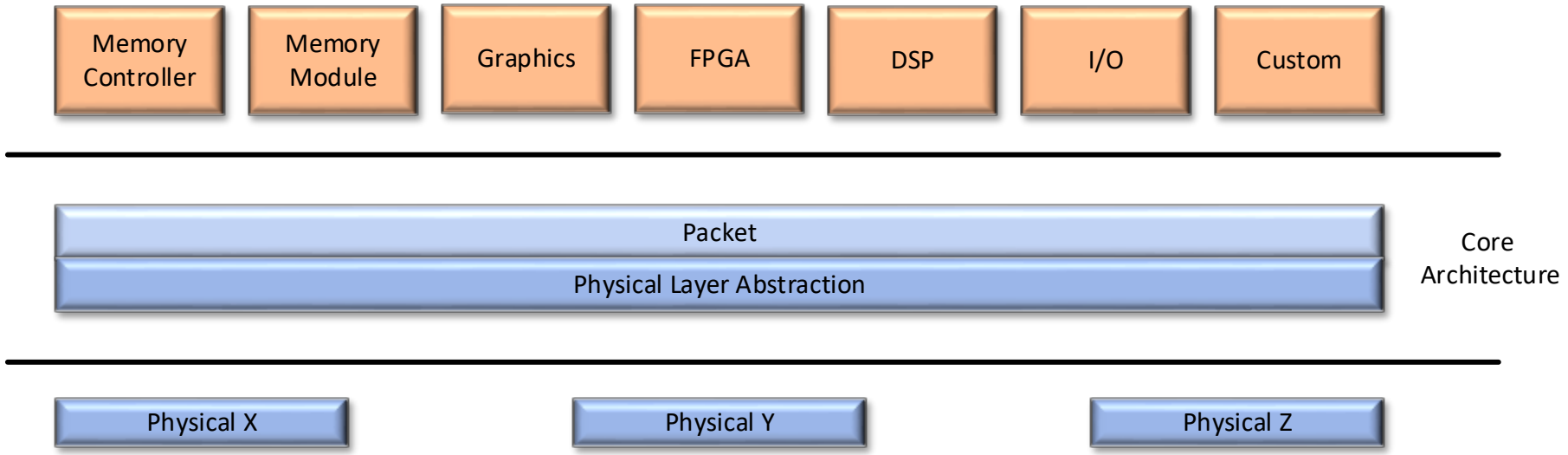
Processor memory and I/O technologies …



… are being stretched to their limits



Layered Architecture



- Core architecture defines operations, protocol, and physical layer abstraction

10s-100s GB/s to TB/s per link bandwidth

Multiple physical layers and signaling rates specified per market

Leverage existing IEEE 802.3 electrical standards with Gen-Z-specific optimizations

Supports PCIe electrical, logical, and LTSSM at all signaling rates

Gen-Z Architecture Attributes

- Feature-scalable packetized transport
- Scalable and power-proportional link, physical layers, and underlying memory media access.
- Split memory controller and media controller paradigm
  - Breaks processor-memory interlock—numerous benefits, e.g.,
    - Abstracts media to enable memory controller to transparently support multiple media types and media generations
    - Accelerate solution innovation and industry agility (eliminates “big bang” events)
  - Transparently integrate performance acceleration techniques to reduce load-to-use latency and increase aggregate bandwidth, mitigate NVM latencies, etc.
- Supports processor-centric and memory-centric architectures
  - Processor-centric provides solution evolution path
  - Memory-centric provides enables new solution architectures not possible / practical with processor-centric

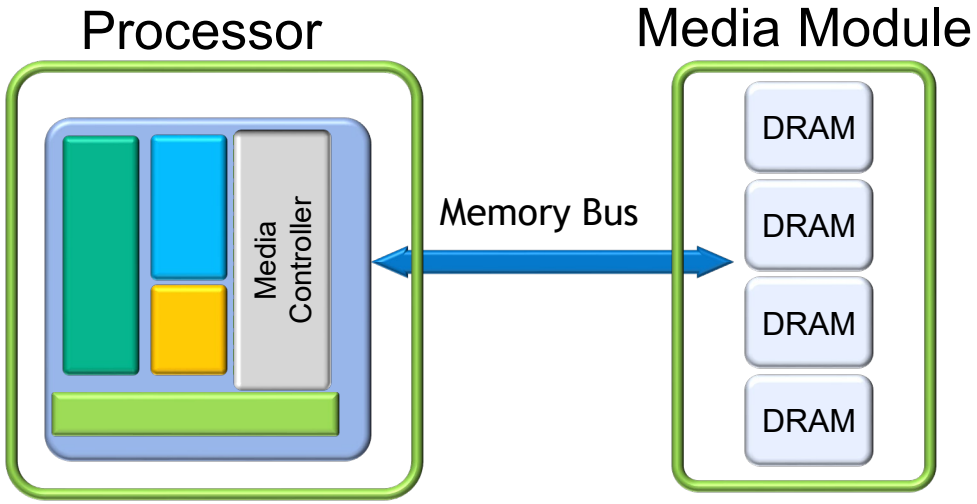
Gen-Z Architecture Attributes (continued)

- Supports unmodified OS and unmodified applications
  - MMU memory mapping to directly access Gen-Z-attached memory
  - Supports logical PCI / PCIe devices
- Abstract physical layer interface supporting multiple physical layers and media
  - Easily tailored to market-specific needs.
  - Rapid evolution or replacement without waiting for entire ecosystem to move in lock-step
- Market-driven packaging and fabric topologies
  - Co-packaged and discrete components
  - Single or multi-link point-to-point topologies
  - Switched fabric topologies—component-integrated switch logic or discrete switch components
  - Single enclosure (client, server, storage, network, etc.) to multi-enclosure / rack scale
- Supports legacy connectors and mechanical form factors
- Supports a new, scalable connector and new modular mechanical form factors
- Common protocol enables democratized communications among all component types

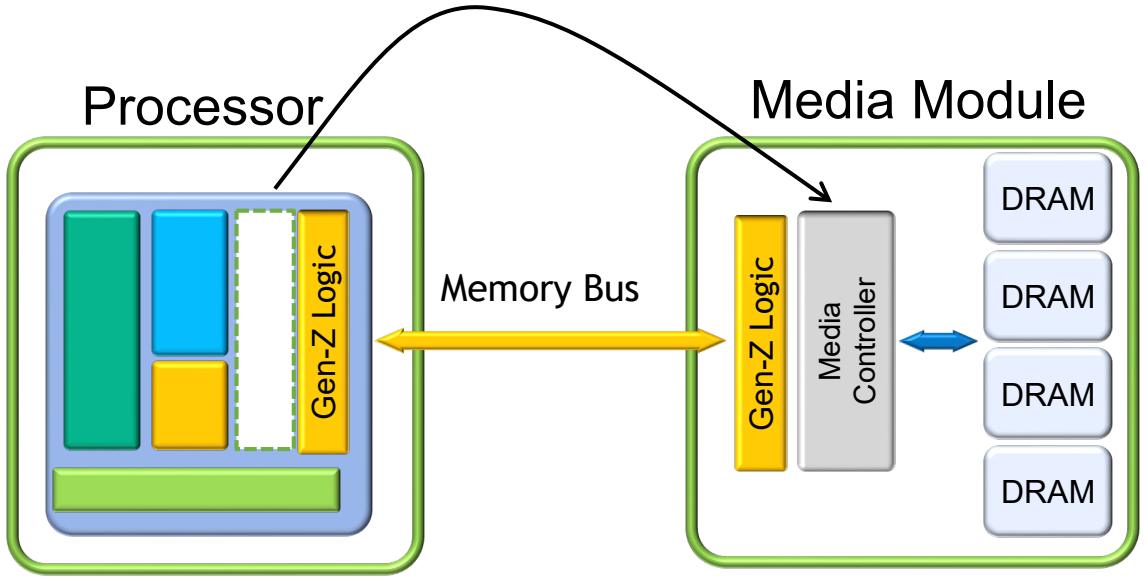
Datagram Packets

- Datagram packet model
  - Requesters ensure reliability (if required)
  - Responders simply execute requests and generate responses (if required)
- Datagrams operate over:
  - Point-to-point and switch topologies
  - Multipath options to provide aggregate bandwidth and resiliency
- Optional encapsulation and strong-ordering domain for specialized communications
  - For example, transparently augment communications without changing primary / third-party protocols
  - For example, transparently tunnel third-party protocols without end-component modifications

Gen-Z Allows Memory Innovation



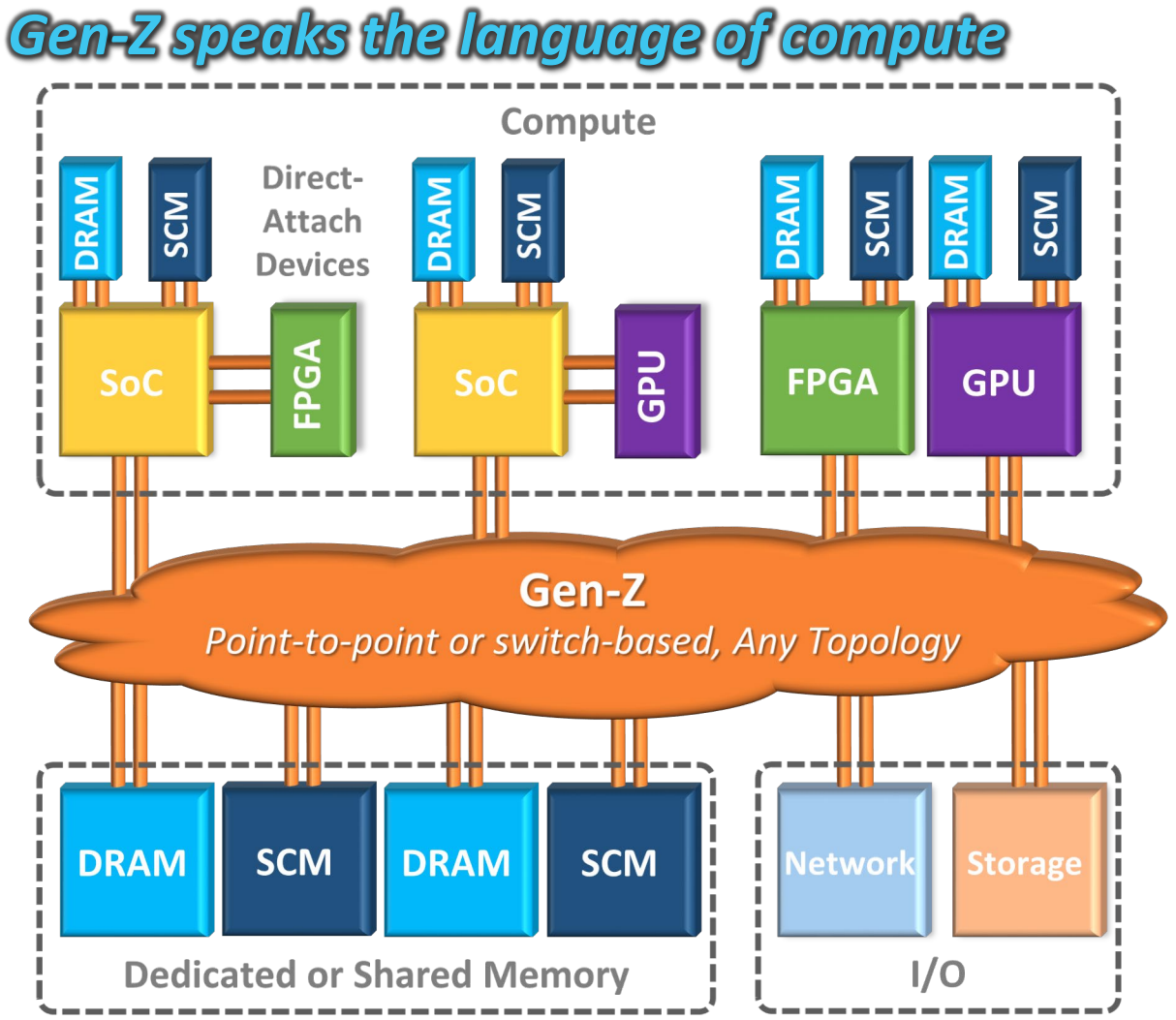
288pins / DIMM
Synchronous Interface

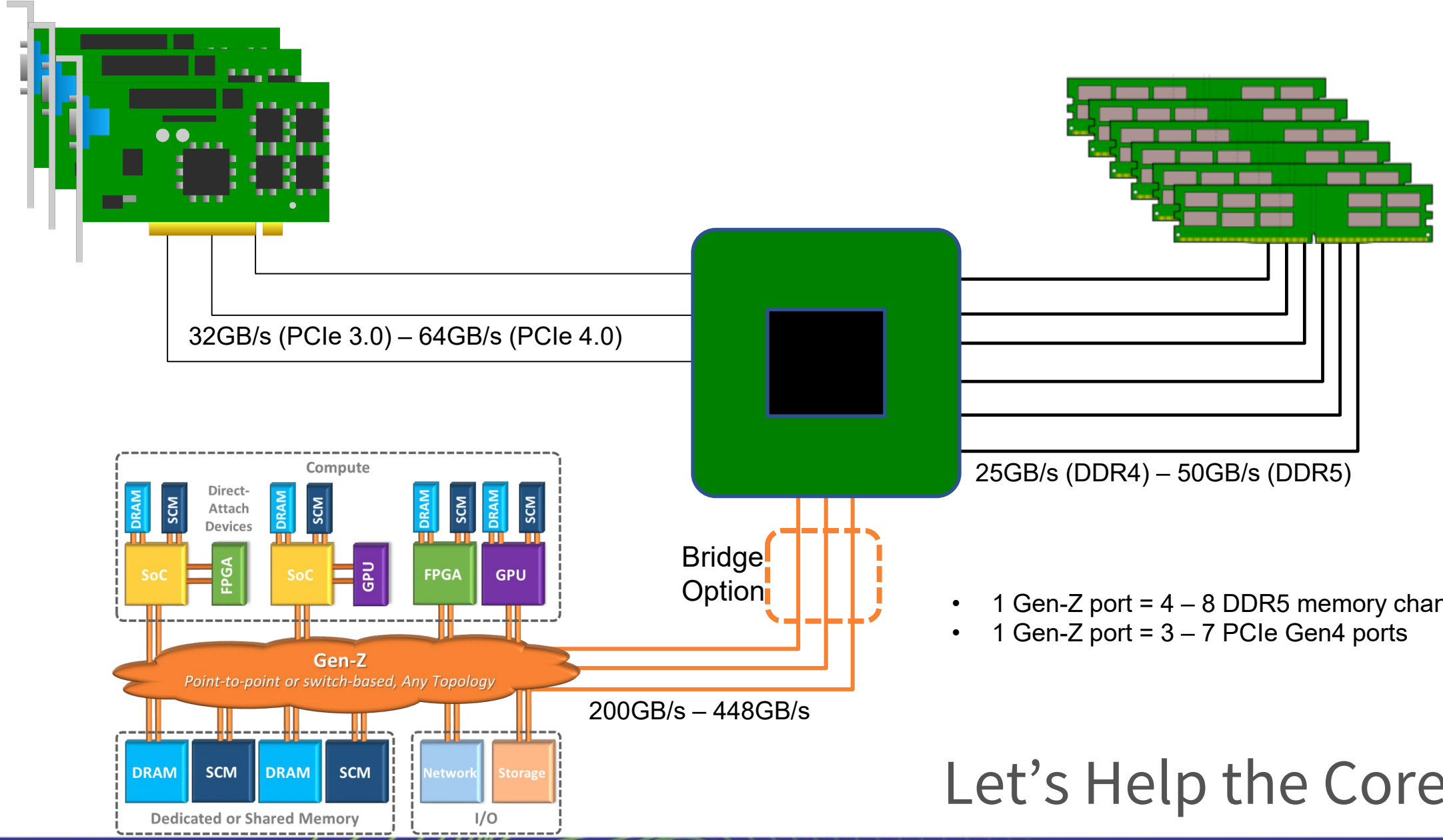


Split Memory Controller
Asynchronous Interface
Processor is media agnostic

Gen-Z Connects Disaggregated Components

- High Performance
  - High Bandwidth, Low Latency, Scalable
  - Eliminates protocol translation cost / complexity / latency
  - Eliminates software complexity / overhead / latency
- Reliable
  - No stranded resources or single-point-of-failures
  - Transparently bypass path and component failure
  - Enables highly-resilient data (e.g., RAID / erasure codes)
- Secure
  - Provides strong hardware-enforced isolation and security
- Flexible
  - Multiple topologies, component types, etc.
  - Supports multiple use cases using simple to robust designs
  - Thorough yet easily extensible architecture
- Compatible
  - Use existing physical layers, no OS modifications required
- Economic
  - Lowers CAPEX / OPEX, unlocks / accelerates innovation



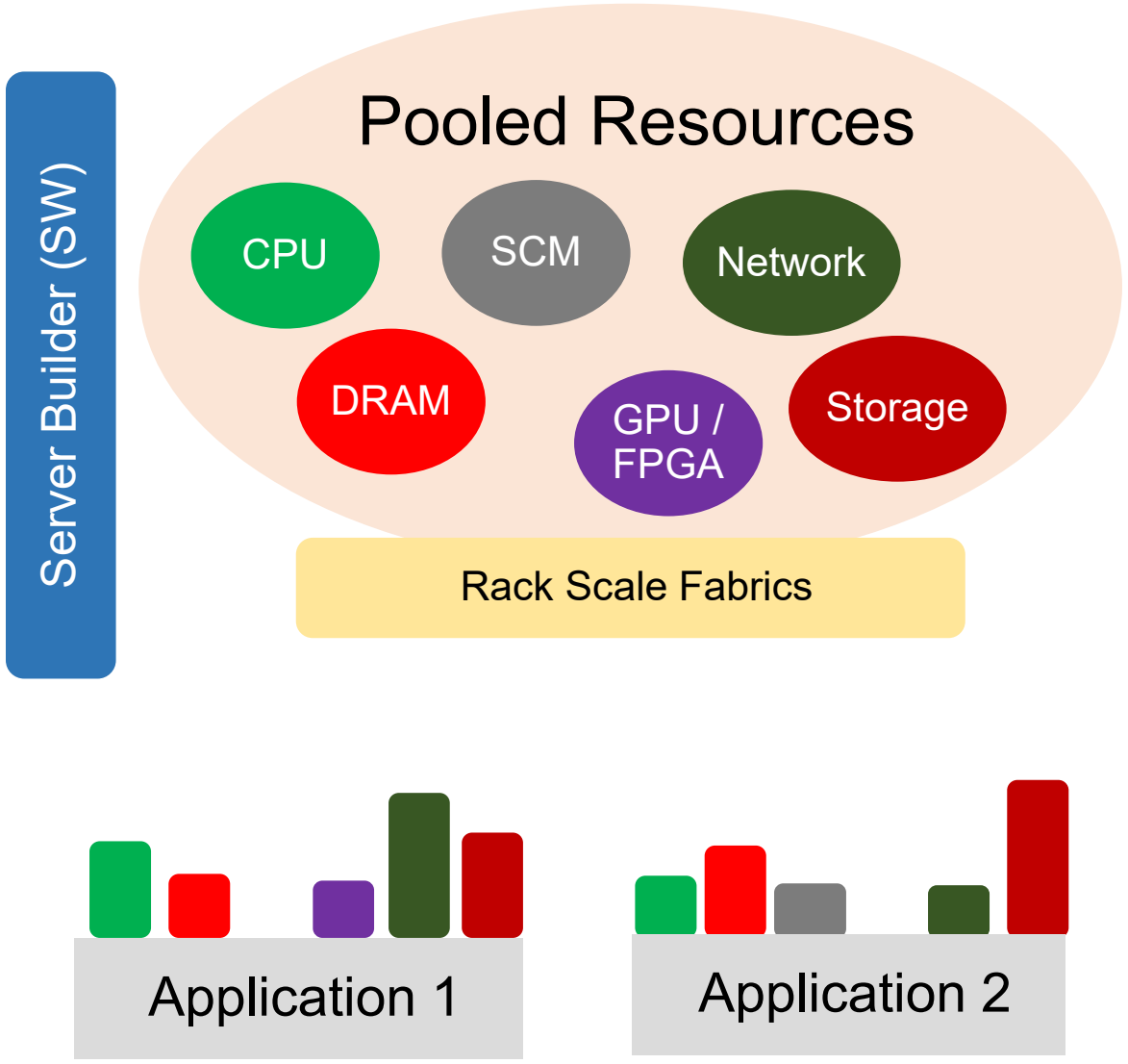


- 1 Gen-Z port = 4 – 8 DDR5 memory channels
- 1 Gen-Z port = 3 – 7 PCIe Gen4 ports

Let's Help the Cores

Server Disaggregation

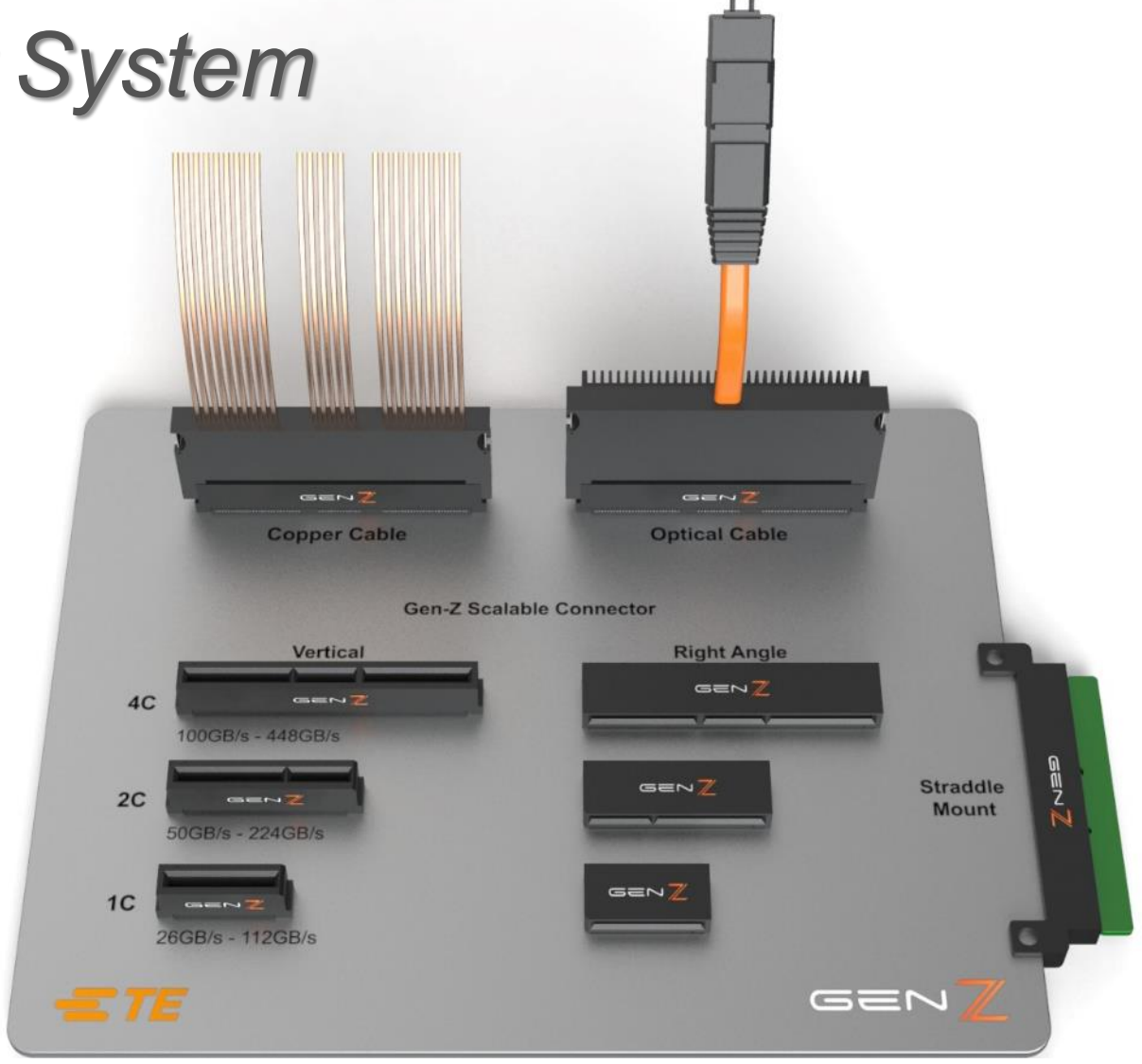
- All resources are collected into shared pools
- High-speed, low-latency fabric connects pools
- Management software:
  - Configures network to connect components
  - Assigns resources
- Result:
  - Disaggregated server
  - True bare-metal bootable server
  - Ready for installation of any OS and application



Flexible: *Universal Connector System*

- Vertical, horizontal, right angle, straddle mount
- Same connectors for memory, I/O, storage, etc.
- Cabled solutions: for copper & optical
- Eliminates “hard choices”
  - Universal connector eliminates industry fragmentation
    - Simplifies supply chain—drives volume and lowers cost
  - Any component, any slot, any time
    - Any mix of static and hot-plug form factors
  - Multi-connector option to provide added scalability
    - 80W incremental power
    - Incremental bandwidth
  - Supports internal and external cable applications
    - Enables modular system design
    - Enables system disaggregation
    - Eliminates expensive board materials
- Multiple technologies—Gen-Z, PCIe, etc.

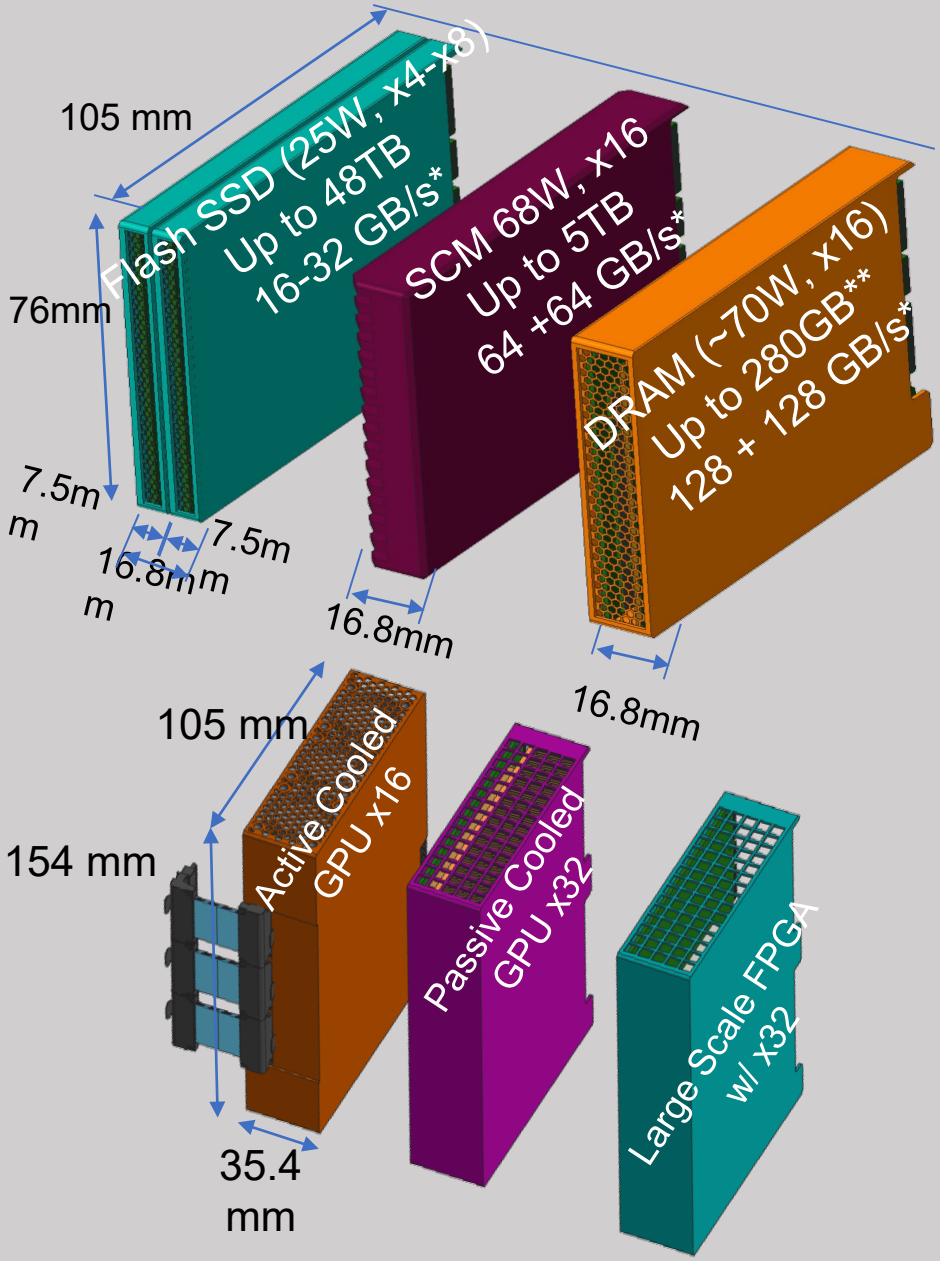
- OCP NIC 3.0 Spec uses the 4C+ Connector



Gen-Z members contributed mechanical & electrical specification to SNIA—see SFF-TA-1002 Gen-Z Scalable Connector specification (final version is publicly available) covers remaining functionality.

Scalable Form Factor[1]

- Supports any component type
  - Flash, SCM, DRAM, NIC, GPU, FPGA, DSP, ASIC, etc.
- Supports multiple interconnect technologies—Gen-Z, PCIe, etc.
- Single and double-wide—scale in x-y-z directions
  - Increased media, power, performance, and thermal capacity
  - Double-wide can be inserted into pairwise single slots
- Supports 1C, 2C, and 4C scalable connectors
  - Larger modules can support multiple connectors—scale power & performance
- Scalable Form Factor Benefits:
  - Simplifies supply chain
  - Lower customer CAPEX / OPEX
  - Consistent customer experience
  - Increases solution and business agility @ lower dev cost
  - Eliminates Potential ESD Damage
    - Can safely move modules from failed / old to new enclosure
  - Eliminates SPOF or stranded resources
    - Multiple links per connector, multiple connectors per module
  - Scalable thermal plus improved airflow across components



[1] Draft specification publicly available—see www.genzconsortium.org
* Bandwidth calculated using 32 GT/s Signaling
** DRAM module provides 3.5x the highest-capacity DDR5 DIMM

2018 Deliverables

- Gen-Z Core Spec 1.0 was released in February 2018
- The Gen-Z Phy 1.0 spec release and includes
  - 25G NRZ fabric & local requirements
  - PCIe G4/G5
- Gen-Z Scalable Connector 1.1 spec release and adds
  - Internal cables, 48V power delivery, new 4C-HP version
- SFF 8639, SFF 8639 Compact, and SFF8201 1.0 contain form factor requirements based on industry specs with exceptions for Gen-Z
- ZSFF 1.0 contains requirements for form factors unique to Gen-Z
- PECFF 1.0 contains a Gen-Z form factor that has a CEM compatible outline
- Working on deliverables for Management/SW, Test/Compatibility & Design Guides

Call to Action

- Now is the time to engage with Gen-Z. Do it now while there is the opportunity to influence first products.
- Gen-Z embraces OPEN.
  - ✓ We shared the Gen-Z connector with the industry SFF-TA1002
  - ✓ We embraced EDSFF recommended form factors
  - ✓ All of our released specs are publicly available
- Gen-Z allows companies to focus on innovation in their area(s) of expertise and provides the interconnect for these innovative products
- Gen-Z is pleased to welcome CXL (Compute Express Link) to the industry and sees opportunity for solutions that bridge between these two interconnects.
  See computeexpresslink.org for more information
- How can OCP and Gen-Z collaborate ? The Gen-Z and OCP organizations must work together to continue driving common elements into designs that benefit our members and the industry.

www.GenZConsortium.org

linkedin.com/company/gen-z-consortium/

Twitter.com/ GenZConsortium

Gen-Z Consortium Channel


Open. Together.
OCP Global Summit | March 14–15, 2019
OCP
SUMMIT