



OPENFABRICS
ALLIANCE

13th ANNUAL WORKSHOP 2017

CCIX, GEN-Z, OpenCAPI: OVERVIEW & COMPARISON

Brad Benton

Advanced Micro Devices

March, 2017



NEWLY EMERGING BUS/INTERCONNECT STANDARDS

■ Three new bus/interconnect standards announced in 2016

- Cache Coherent Interconnect for Accelerators (**CCIX**)
www.ccixconsortium.com
- **Gen-Z**
genzconsortium.org
- Open Coherent Accelerator Processor Interface (**OpenCAPI**)
opencapi.org

■ Driving forces behind these new standards

- Tighter coupling between processors and accelerators (GPUs, FPGAs, etc.)
- Better exploitation of new and emerging memory/storage technologies
 - Streamline software stacks
 - Reduce data movement by direct access to memory
- Open standards-based solutions

■ Why 3 different standards?

- Different groups have been working to solve similar problems
- However, each approach has its differences
- Many companies involved with multiple consortia
- Possible shake outs/convergence as things move forward

CONSORTIA MEMBER COMPANIES

CCIX

AMD
Amphenol
ARM
Arteris
Avery Design
Broadcom
Bull/Atos
Cadence
Cavium
Huawei
IBM
IDT

Gen-Z

Alpha Data
AMD
Amphenol
ARM
Broadcom
Cavium
Cray
Dell
FIT
HPE
Huawei
IBM
IDT
IntelliProp
JaBil

OpenCAPI

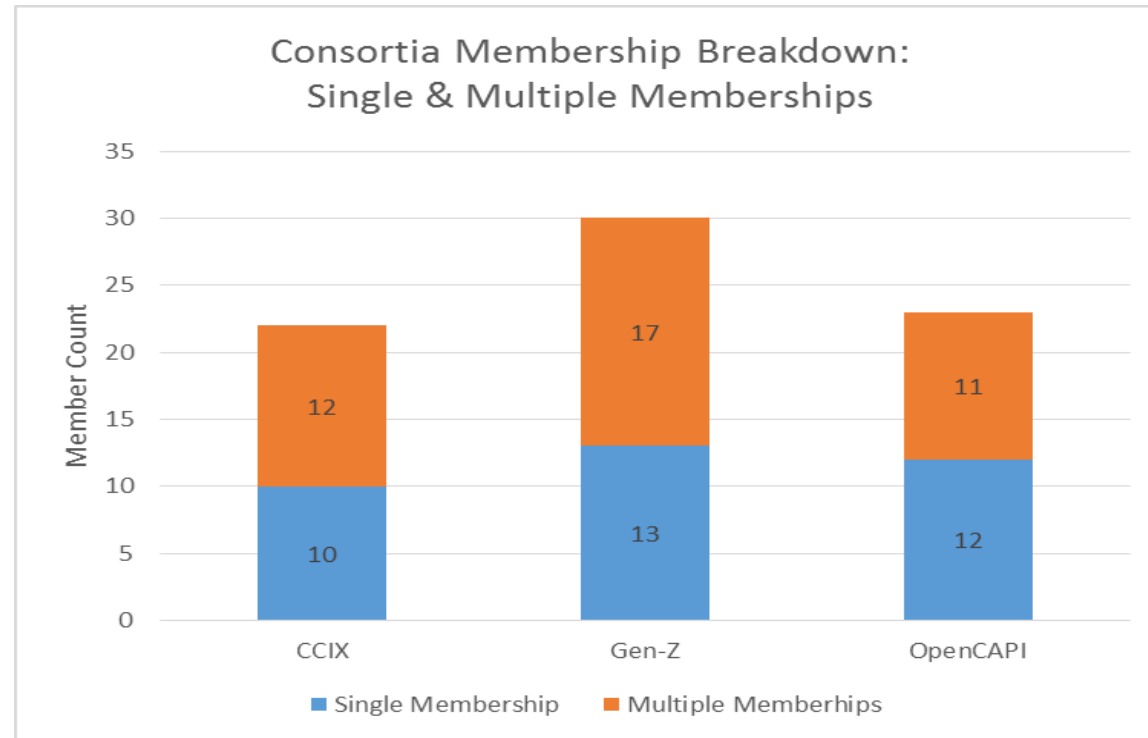
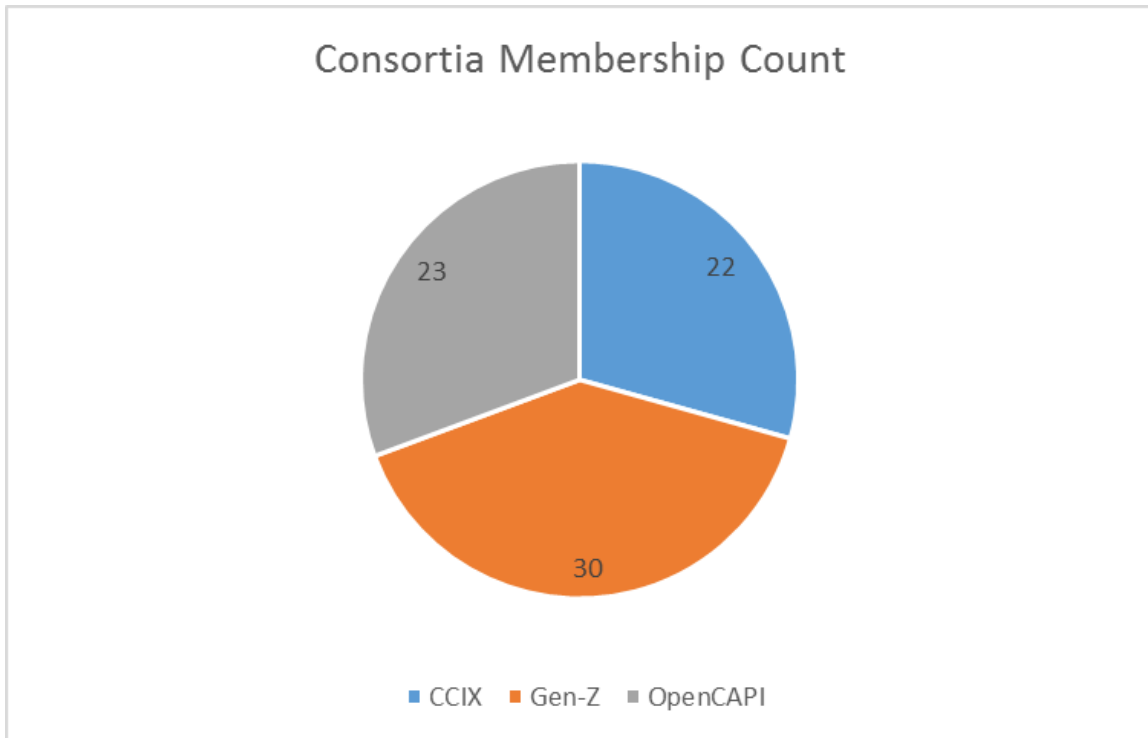
Lenovo
Lotes
Mellanox
Micron
Microsemi
Nokia
PDLA Group
Redhat
Samsung
Seagate
SK hynix
SpinTransfer
WesternDigital
Xilinx
Yadro

Achronix
AMD
Amphenol
AppliedMaterials
Dell
ELI Beamlines
Google
HPE
IBM
Mellanox
Micron
Micronsemi

NGCodec
NVIDIA
Parade
Samsung
SuperMicro
Synology
Tektronix
Toshiba
Univ. Cordoba
WesternDigital
Xilinx

CONSORTIA MEMBERSHIP DISTRIBUTION

Membership Distribution



Many companies have dual or triple memberships

CCIX: CACHE COHERENT INTERCONNECT FOR ACCELERATORS

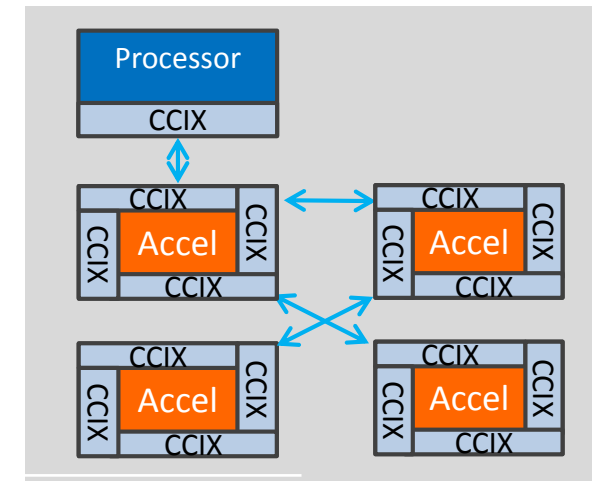
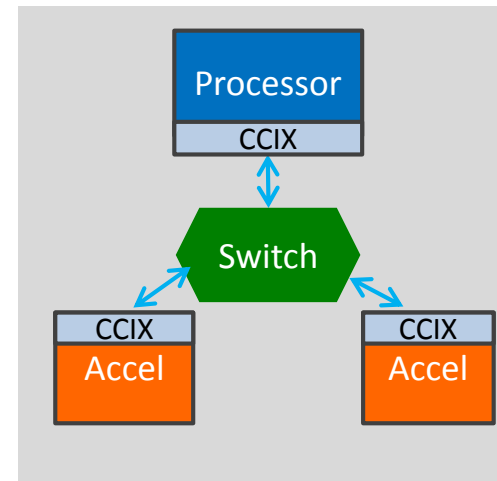
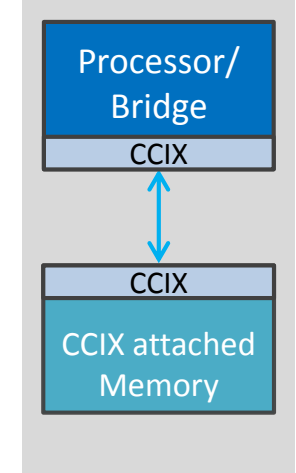
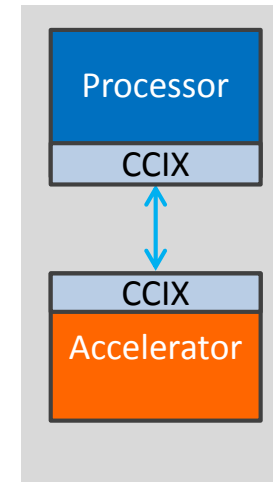
www.ccixconsortium.com

▪ **Tightly coupled interface between processor, accelerators and memory**

- Bandwidth:
 - 16Gps to 25 Gps/lane
 - Support for intermediate speeds
- Hardware cache coherence enabled across the link
- Driver-less and interrupt-less framework for data sharing

▪ **Use Cases**

- Allows low-latency main memory expansion
- Extend processor cache coherency to accelerators, network/storage adapters, etc.
- Supports multiple ISAs over a single interconnect standard



<http://www.ccixconsortium.com/>

GEN-Z: MEMORY SEMANTIC FABRIC

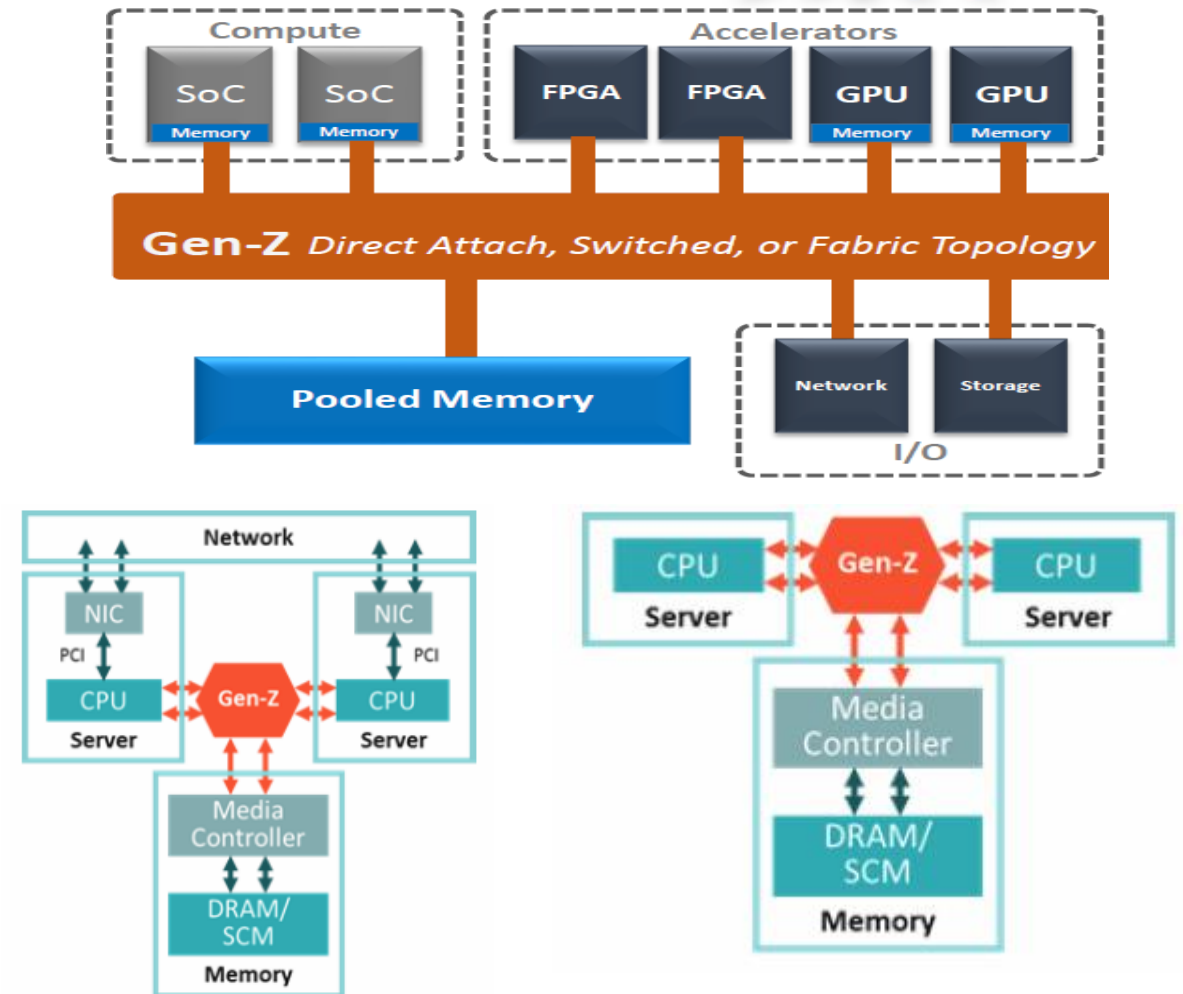
www.genzconsortium.org

▪ Scalable from component to cross-rack communications

- Direct attach, switched, or fabric topologies
- Bandwidth:
 - 32GB/s to 400+ GB/s
 - Support for intermediate speeds
- Can gateway to other networks, e.g., Enet, InfiniBand
- Unify general data access as memory operations
 - byte addressable load/store
 - messaging (put/get)
 - IO (block memory)

▪ Use Cases

- Component disaggregation
- Persistent memory
- Long haul/rack-to-rack interconnect



<http://www.genzconsortium.org/>

OpenCAPI: OPEN COHERENT ACCELERATOR PROCESSOR INTERFACE

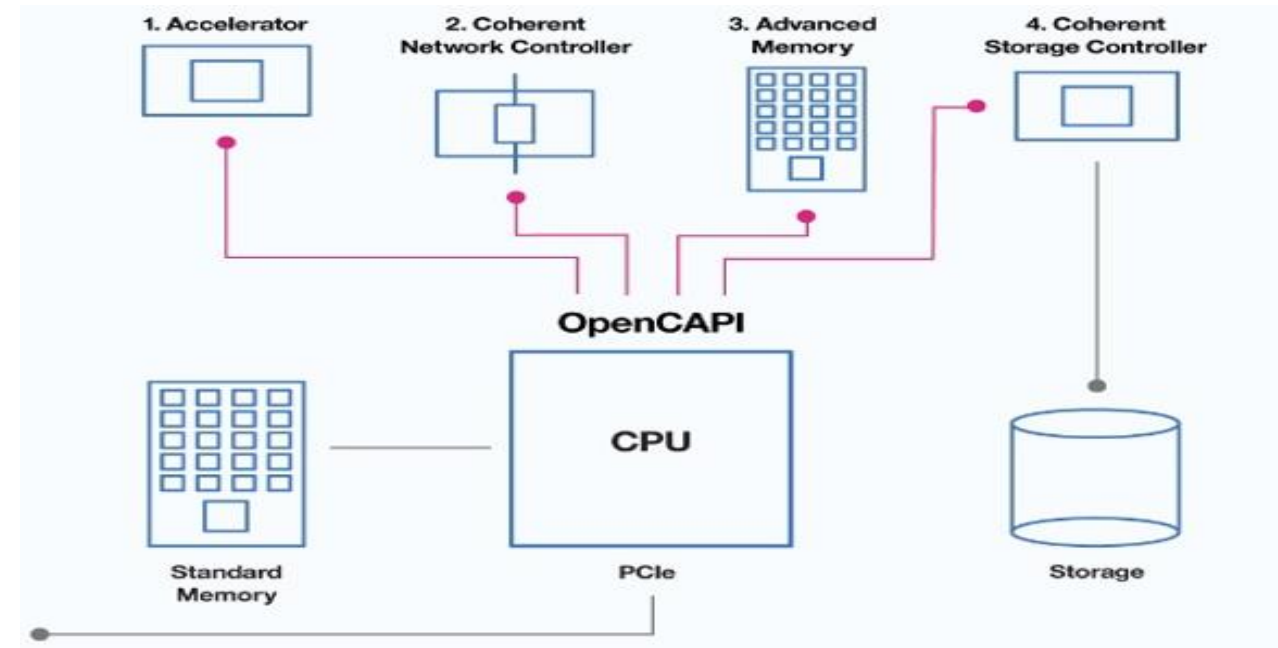
www.opencapi.org

▪ **Tightly coupled interface between processor, accelerators and memory**

- Operates on virtual addresses
- virt-to-phys translation occurs on the host CPU
- OpenCAPI 3.0:
 - Bandwidth:
 - 25 Gps/lane x8
 - Coherent access to system memory
- OpenCAPI 4.0:
 - Support for caching on accelerators
 - Bandwidth:
 - Additional link widths: x4, x8, x16, x32

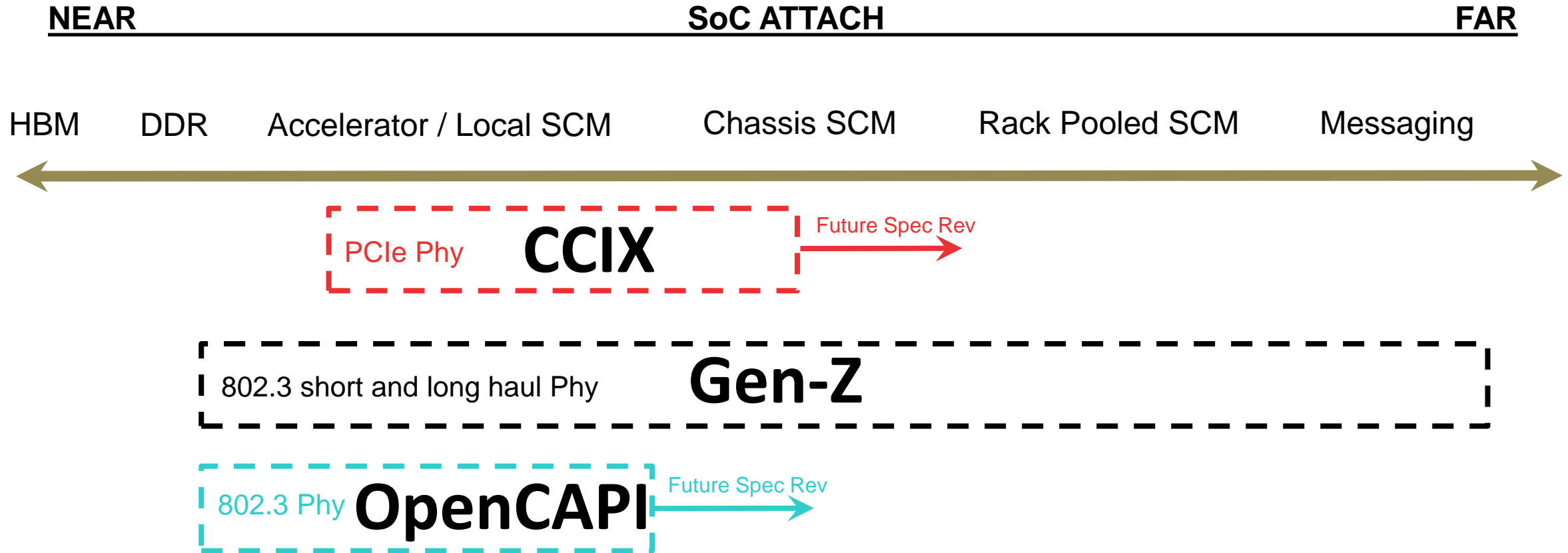
▪ **Use Cases**

- Coherent access from accelerator to system memory
- Attached devices operate on virtual addresses
- Allows low-latency advanced memory expansion
- Agnostic to processor architecture



<http://www.opencapi.org/>

BUS REACH



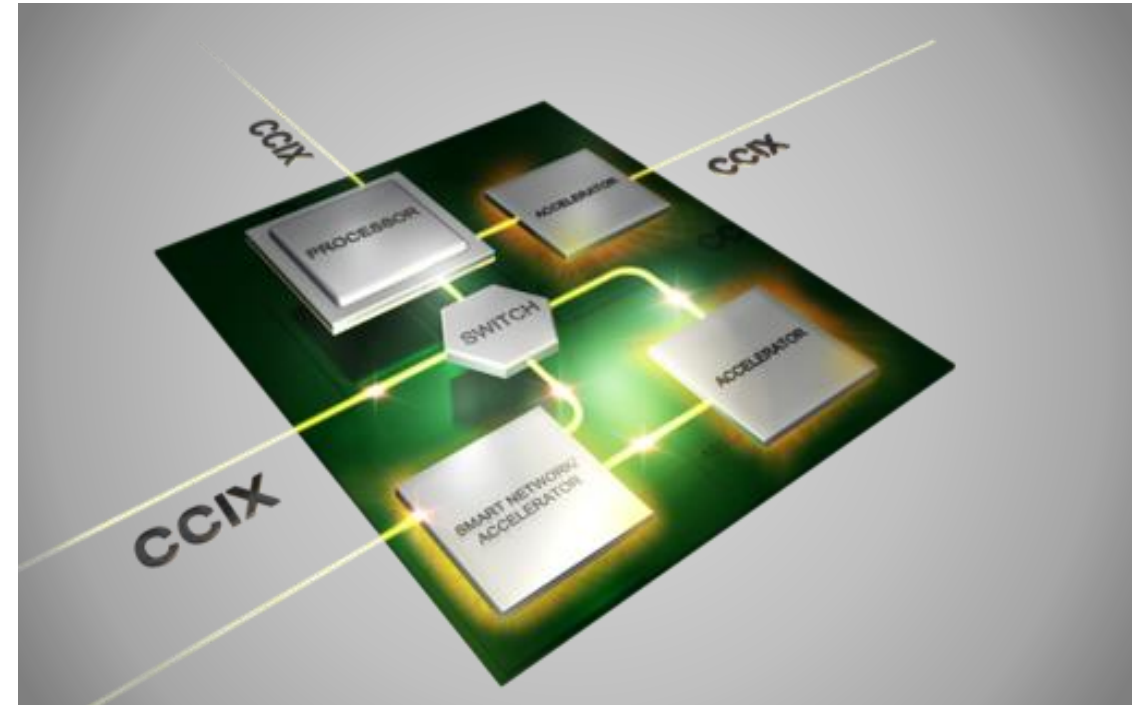
COMPARISONS

Standard	Physical Layer	Topology	Unidirectional Bandwidth	Mechanicals	Coherence
CCIX	PCIe PHY	p2p and switched	32-50GB/s x16	PCIe	Full cache coherency between processors and accelerators
GenZ	IEEE 802.3 Short and Long Haul PHY	p2p and switched	Signaling Rates: 16, 25, 28, 56 GT/s Multiple link widths: 1 to 256 lanes	Supports existing PCIe mechanicals/form factors Will develop new, Gen-Z specific mechanicals/form factors	Does not specify cache coherent agent operations, but does specify protocols that support cache coherent agents
OpenCAPI 3.0	BlueLink 25Gbs PHY Used for OpenCAPI & NVLINK	p2p	25GB/s x8	In definition, see Zaius design for a possible approach	Coherent access to memory Cache coherence not supported until v4.0

IMPLICATIONS FOR OFA

CCIX

- Straightforward extension of PCIe
- In box/node system interconnect
- Preserves existing mechanicals, connectors, etc.
- OFA visibility limited
- Communication with CCIX-attached devices managed by vendor-specific drivers/libraries
- Address translation services via ATS/PRI

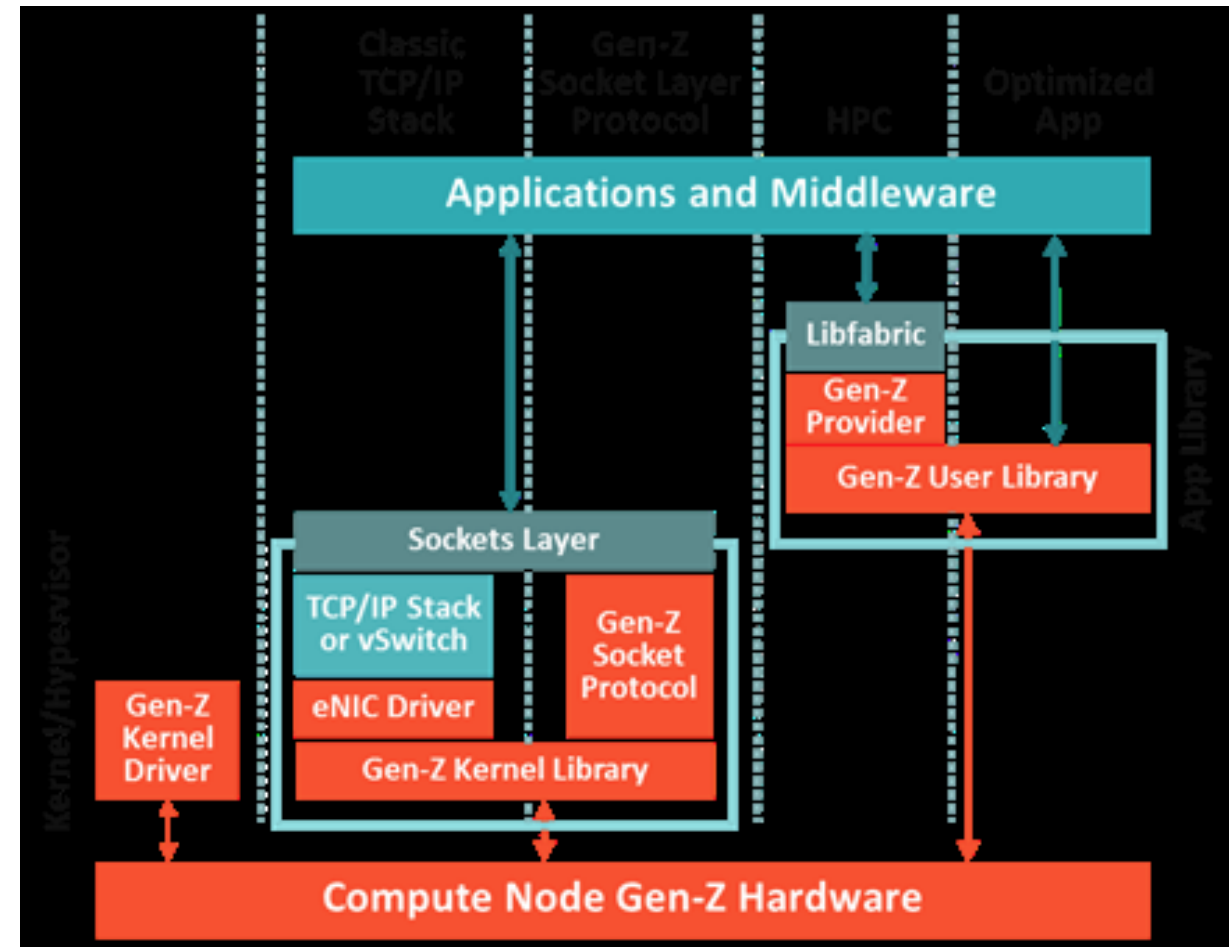


<http://www.ccixconsortium.com/>

IMPLICATIONS FOR OFA

Gen-Z

- Defines new mechanicals/connectors
- But will also integrate with existing mechanical form factors, connectors, and cables
- Exposure to OFA:
 - Messaging interfaces
 - Block I/O interfaces
 - Persistent memory
- Gen-Z working on libfabric integration

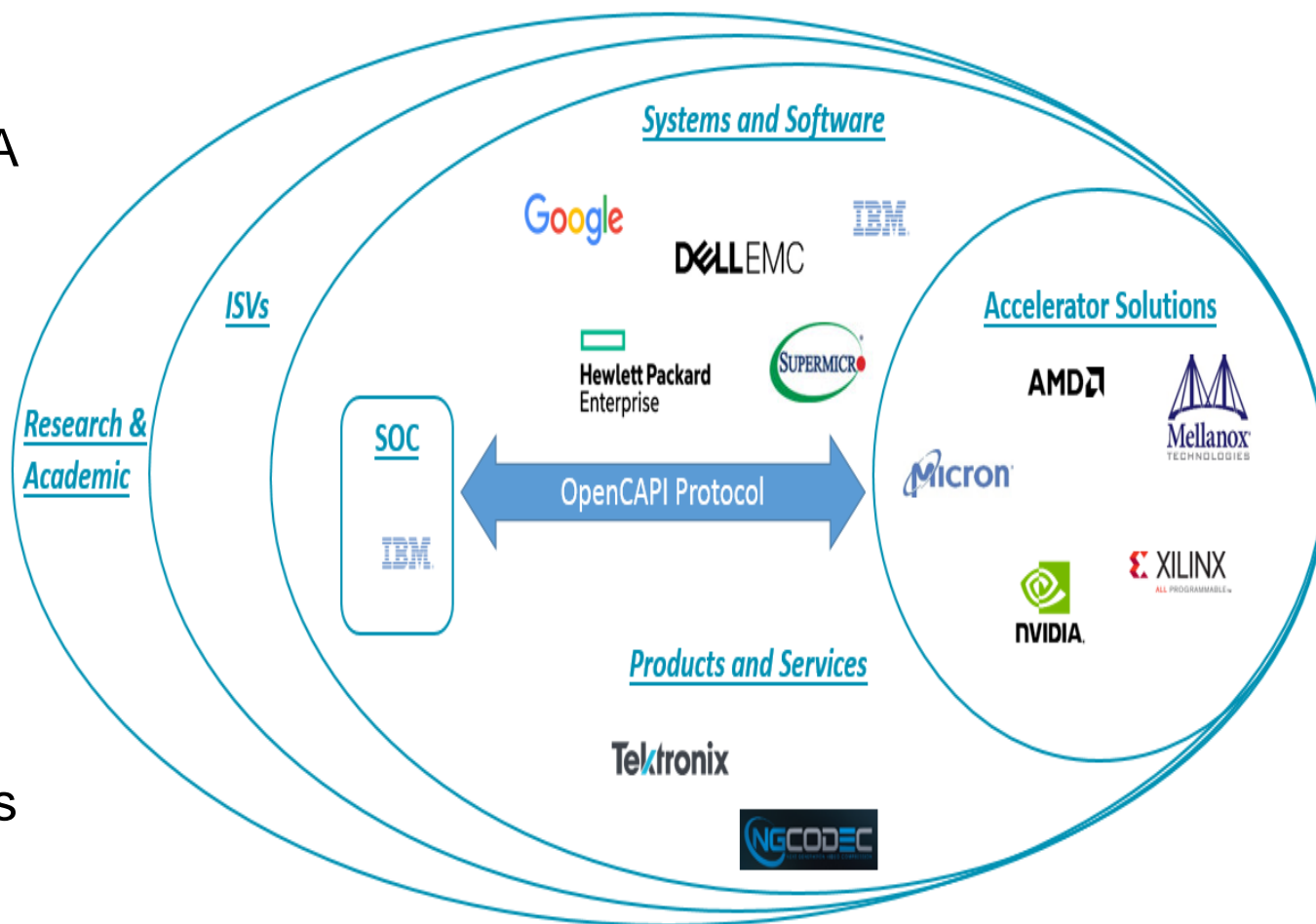


<http://www.genzconsortium.org/>

IMPLICATIONS FOR OFA

OpenCAPI

- Integrated into POWER9 (e.g., Zaius)
- Supports features of interest to NIC/FPGA vendors
 - Virtual address translation services
 - Aggregation of accelerator & system memory
- OFA visibility limited
- Communication with OpenCAPI-attached devices managed by vendor-specific drivers/libraries
- Accelerator holds virtual address; address translation managed by the host



<http://opencapi.org/wp-content/uploads/2016/11/OpenCAPI-Overview-SC16-vf.pptx>

MOVING FORWARD WITH 3 STANDARDS?

- **Are there paths to convergence?**

- Vendors, integrators anxious for clarity

- **CCIX & Gen-Z**

From: <http://www.ccixconsortium.com/about-us.html>

GenZ is a new data access technology that enables memory operations to directly attach and disaggregate memory and storage. CCIX extends the processor's coherency domain to heterogeneous components. These heterogeneous "nodes" would then get access to the large and disaggregated storage and memory through the GenZ fabric.

- **OpenCAPI**

- Supported in POWER9 architecture via BlueLink (25Gbs on-chip)
- OpenCAPI architecture implementable with non-POWER processors
- As with CCIX, potential to bridge OpenCAPI/Gen-Z

- **In the near term, expect FPGA-based bridging for interconnect transitions**

Questions

Thoughts

Discussion



OPENFABRICS
ALLIANCE

13th ANNUAL WORKSHOP 2017

THANK YOU

Brad Benton

Advanced Micro Devices



DISCLAIMER & ATTRIBUTION

The information presented in this document is for informational purposes only and may contain technical inaccuracies, omissions and typographical errors.

The information contained herein is subject to change and may be rendered inaccurate for many reasons, including but not limited to product and roadmap changes, component and motherboard version changes, new model and/or product releases, product differences between differing manufacturers, software changes, BIOS flashes, firmware upgrades, or the like. AMD assumes no obligation to update or otherwise correct or revise this information. However, AMD reserves the right to revise this information and to make changes from time to time to the content hereof without obligation of AMD to notify any person of such revisions or changes.

AMD MAKES NO REPRESENTATIONS OR WARRANTIES WITH RESPECT TO THE CONTENTS HEREOF AND ASSUMES NO RESPONSIBILITY FOR ANY INACCURACIES, ERRORS OR OMISSIONS THAT MAY APPEAR IN THIS INFORMATION.

AMD SPECIFICALLY DISCLAIMS ANY IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR ANY PARTICULAR PURPOSE. IN NO EVENT WILL AMD BE LIABLE TO ANY PERSON FOR ANY DIRECT, INDIRECT, SPECIAL OR OTHER CONSEQUENTIAL DAMAGES ARISING FROM THE USE OF ANY INFORMATION CONTAINED HEREIN, EVEN IF AMD IS EXPRESSLY ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

ATTRIBUTION

© 2017 Advanced Micro Devices, Inc. All rights reserved. AMD, the AMD Arrow logo and combinations thereof are trademarks of Advanced Micro Devices, Inc. in the United States and/or other jurisdictions. Other names are for informational purposes only and may be trademarks of their respective owners.