OPENFABRICS ALLIANCE

Panel Discussion:

The Future of I/O From a CPU Architecture Perspective

Brad Benton
AMD, Inc.

10TH ANNUAL INTERNATIONAL DEVELOPER WORKSHOP

#OFADevWorkshop

Future of I/O from CPU/Arch Perspective
Issues

- Move to Exascale involves more parallel processing across more processing elements
  - GPUs, FPGAs, Specialized accelerators, Processing In Memory (PIM), etc.
- These elements frequently do not live in the same coherent memory domain
- Results in much data movement to bring data to the needed processing element

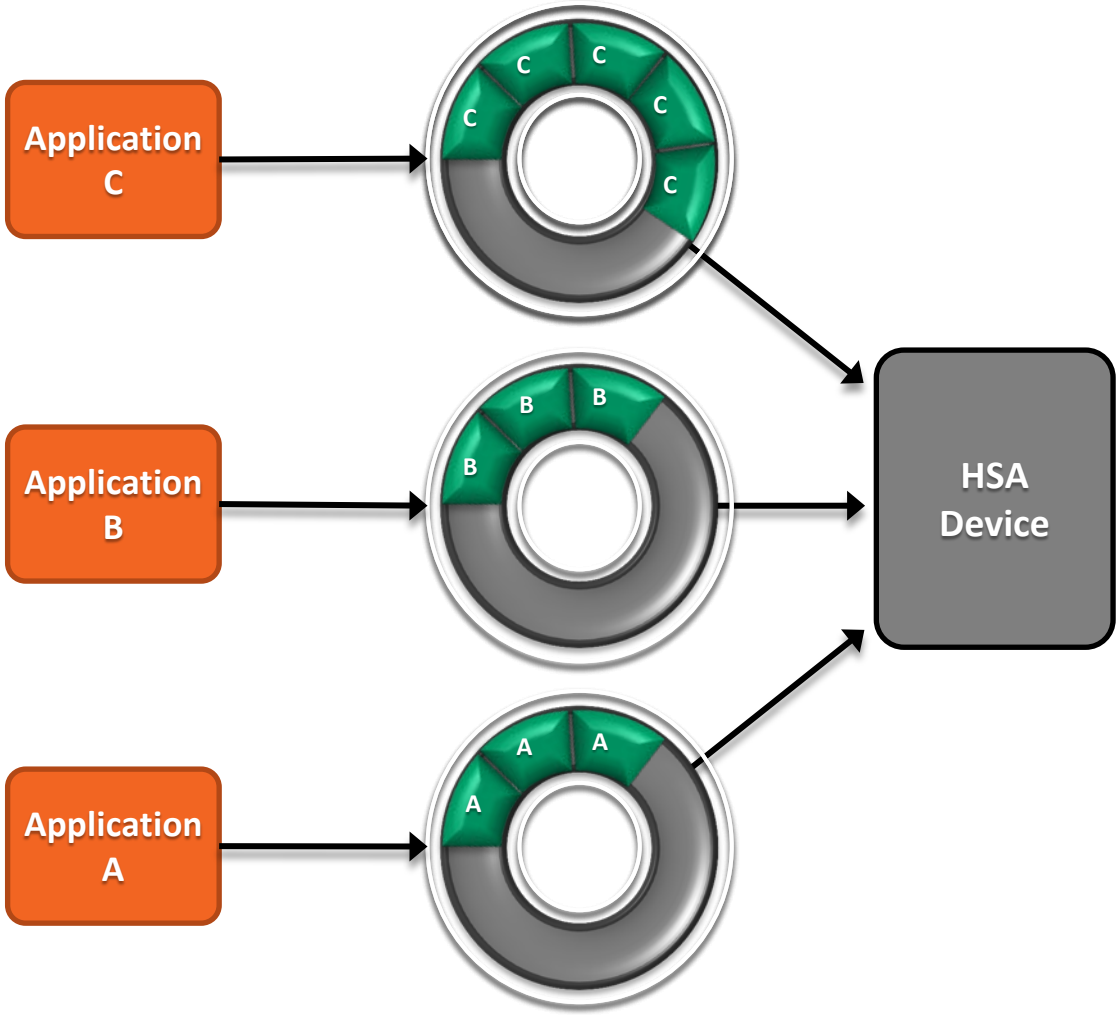
Future of I/O from CPU/Arch Perspective
Heterogeneous System Architecture (HSA)

- HSA Foundation
  - "Not-for-Profit Industry Consortium of SOC and SOC IP vendors, OEMs, academia, OSVs and ISVs defining a consistent heterogeneous platform architecture to make it dramatically easier to program heterogeneous parallel devices"
  - Develops open, royalty-free industry specifications and APIs for heterogeneous computing
  - http://www.hsafoundation.com/
- Goal: Bring Accelerators (and other devices) forward as first class citizens
  - Unified address space
  - Operate in pageable system memory
  - Full memory coherence
  - User mode dispatch/scheduling

Future of I/O from CPU/Arch Perspective
HSA Foundation Membership

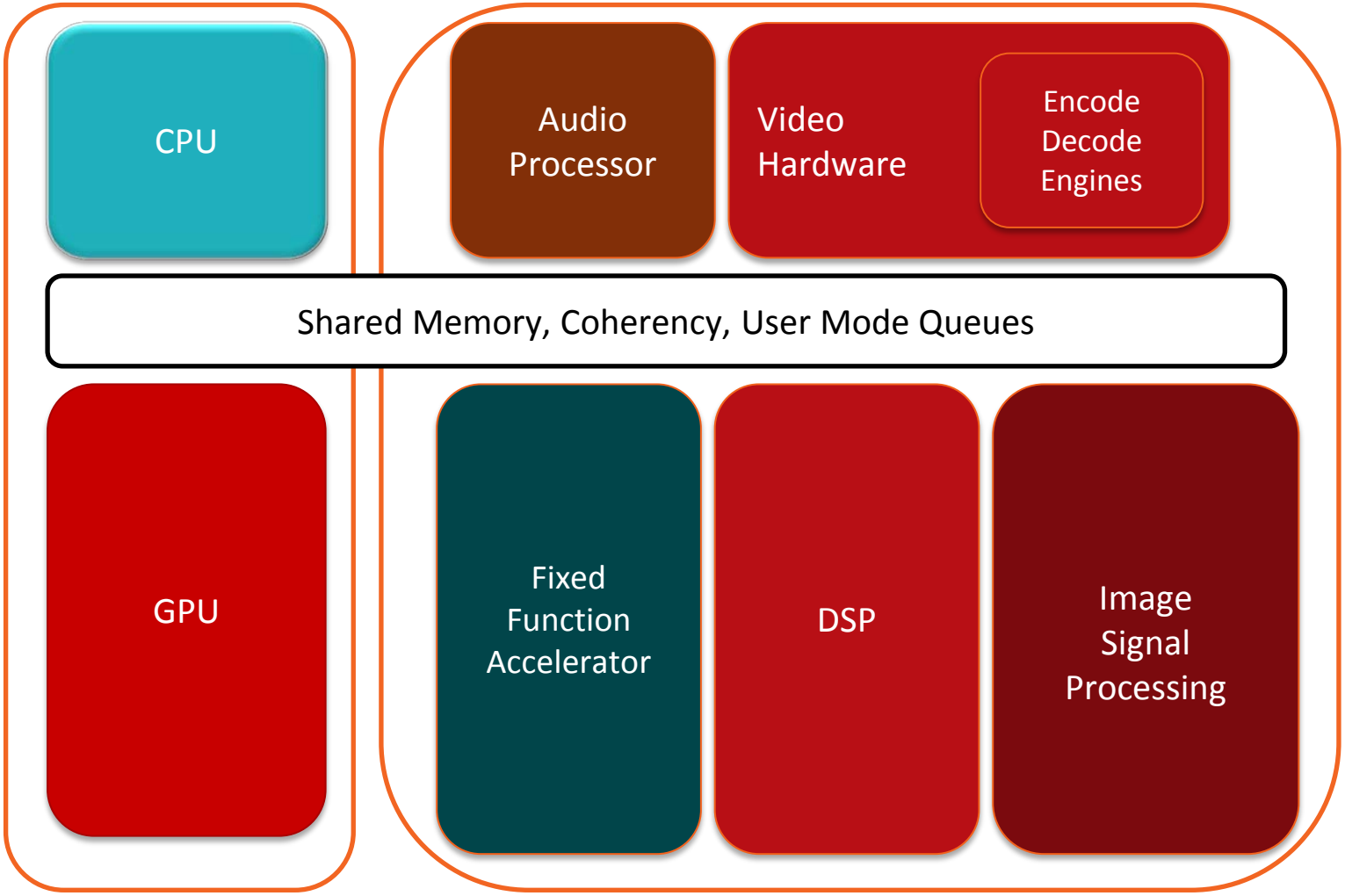
- 50+ Members
  - Multiple categories of membership
- Founding Members
  - AMD
  - ARM
  - Imagination
  - MediaTek
  - Qualcomm
  - Samsung
  - Texas Instruments

Future of I/O from CPU/Arch Perspective
Application-level access to HSA devices



- Application talks directly to hardware
  - No system call
  - No kernel driver involved
  - Hardware scheduling
  - Low dispatch latency
- Standardized binary packet format
  - Architected Queueing Language (AQL)

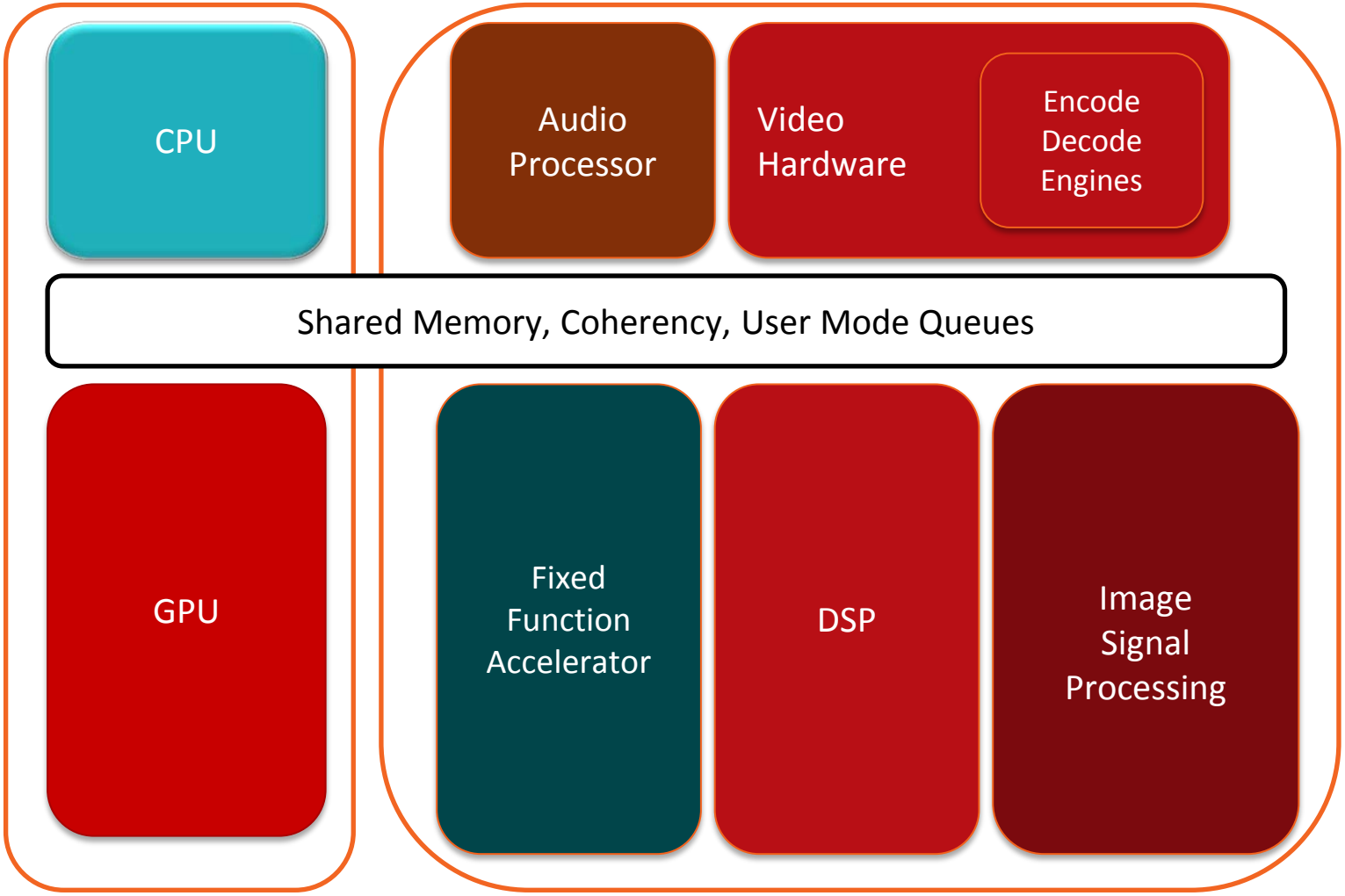
Future of I/O from CPU/Arch Perspective
HSA is Designed to go Beyond the GPU



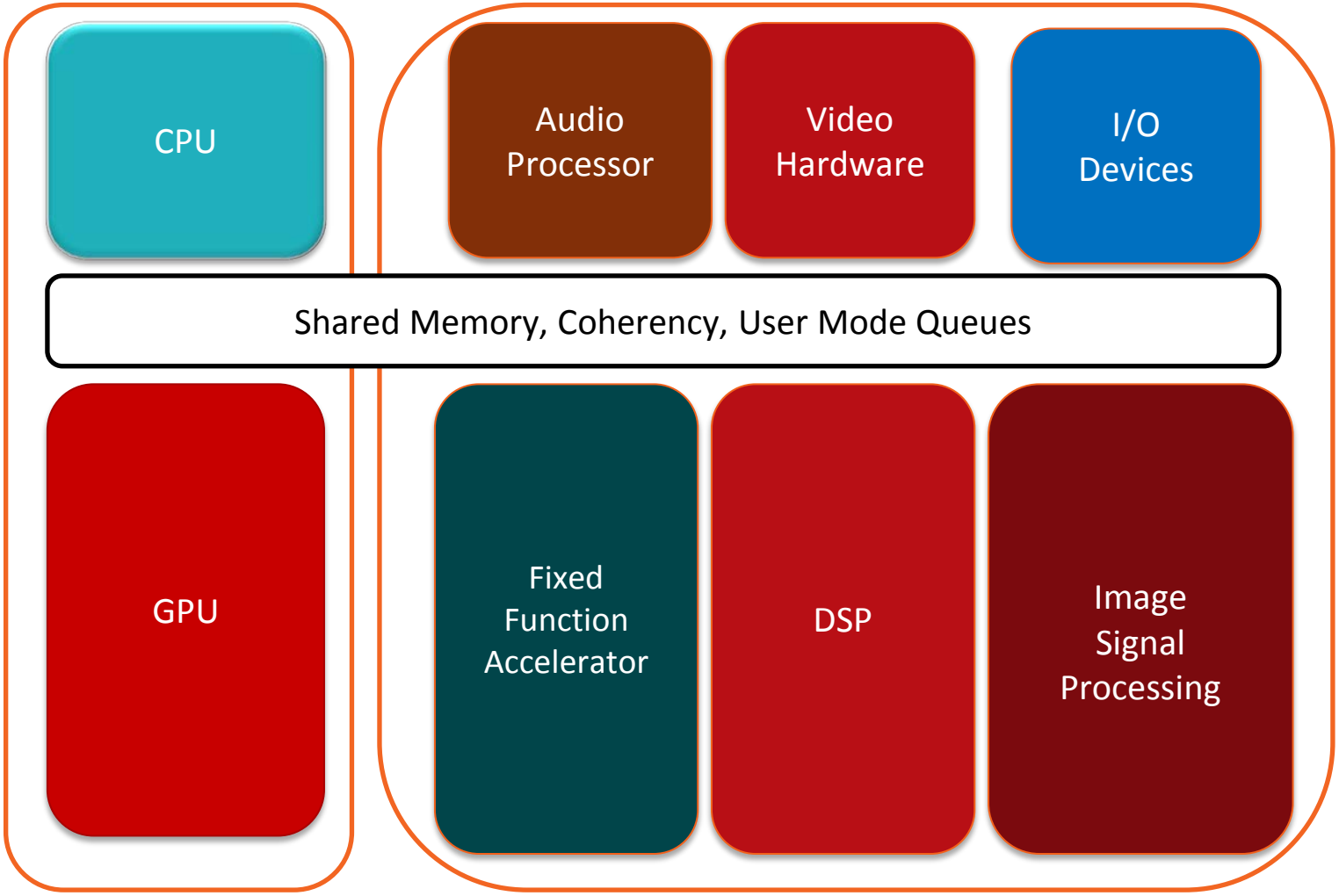
Future of I/O from CPU/Arch Perspective
HSA and I/O Devices

- Okay…HSA looks interesting from the viewpoint of accelerators, how is this related to I/O?
- I/O devices can participate, as well!
- The HSA and its programming model are designed to allow any I/O devices capable of page fault handling to participate in a shared virtual address space.

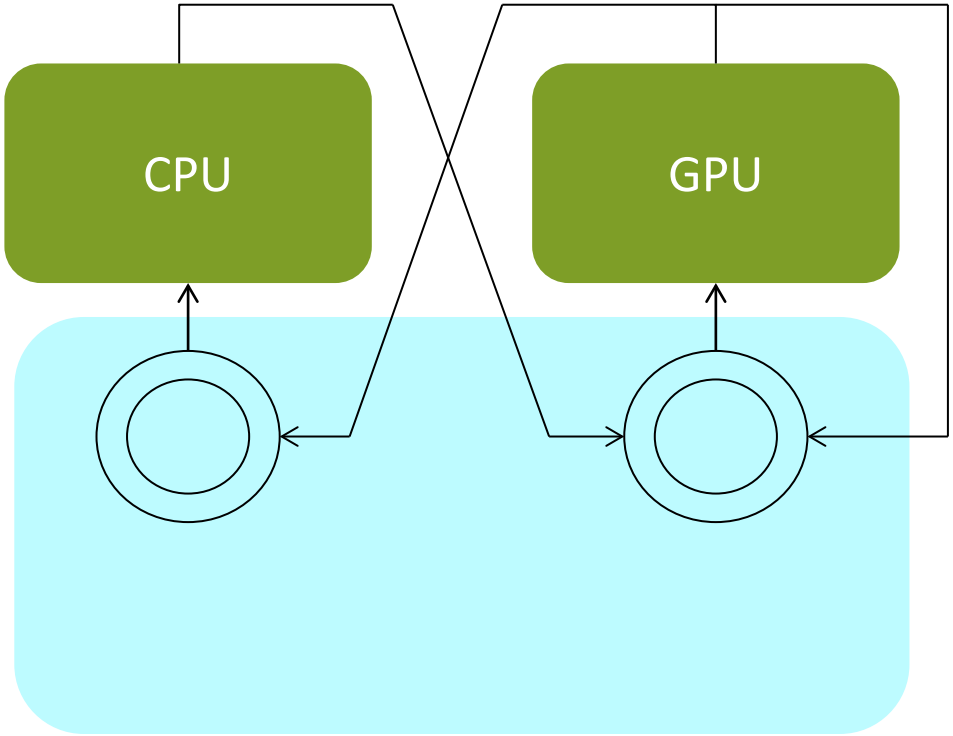
Future of I/O from CPU/Arch Perspective
HSA is Designed to go Beyond the GPU



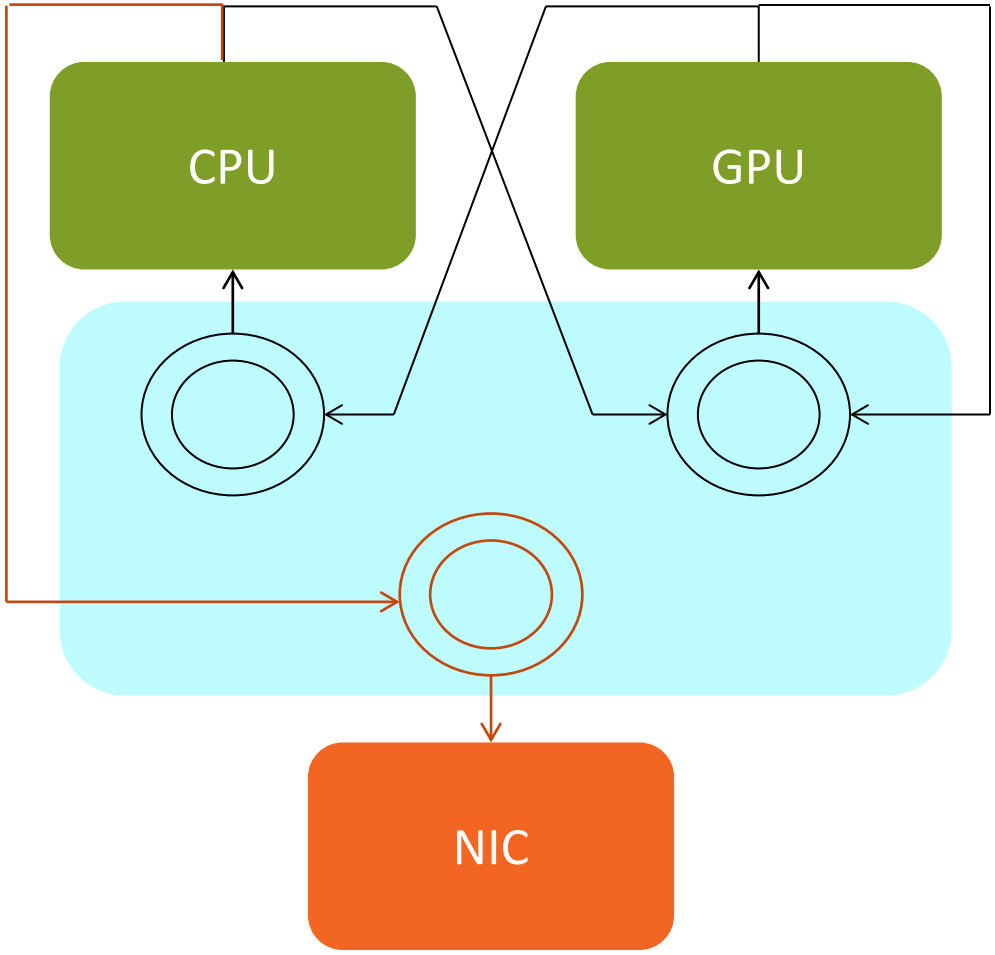
Future of I/O from CPU/Arch Perspective
HSA is Designed to go Beyond the GPU



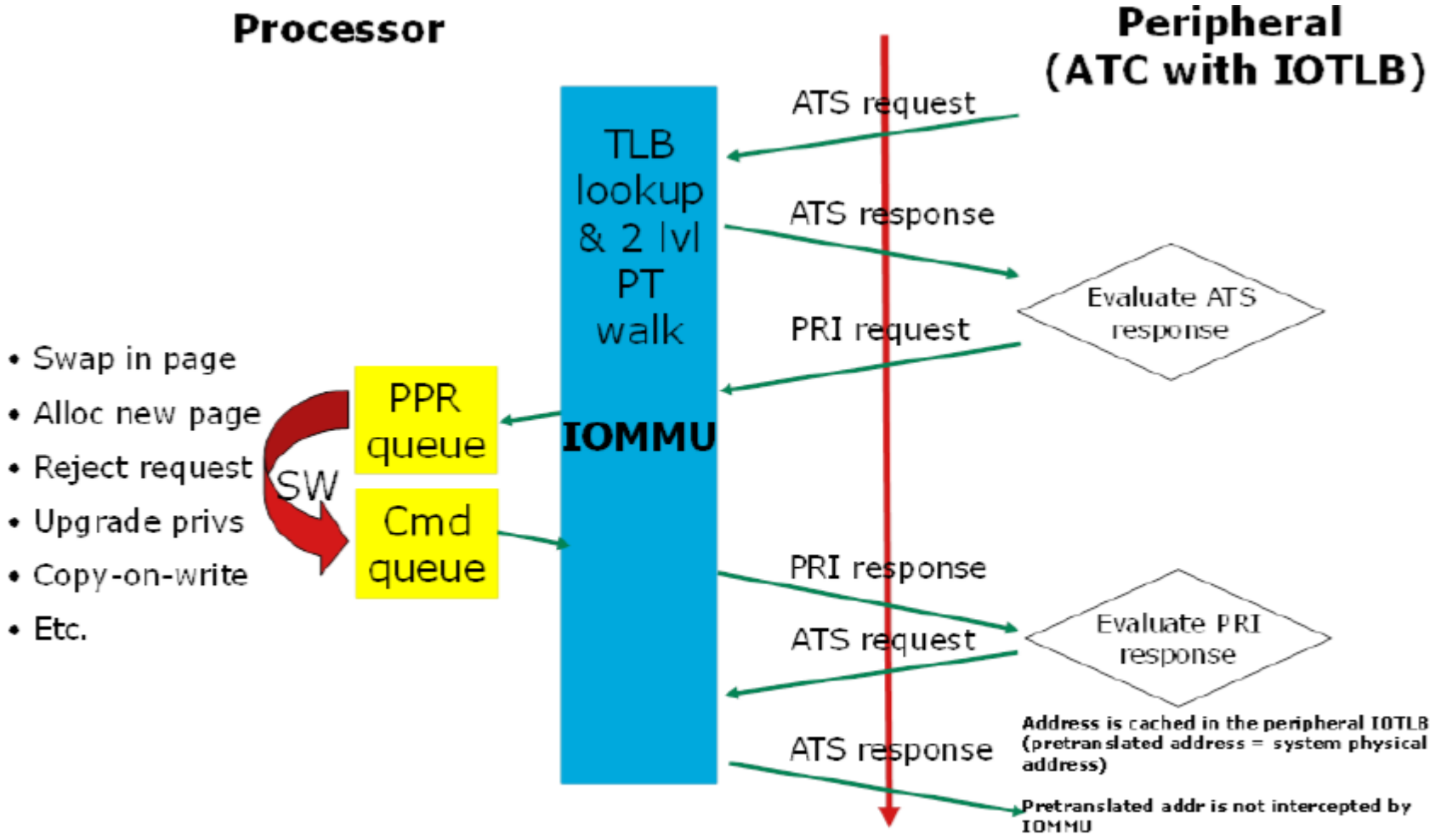
Future of I/O from CPU/Arch Perspective
HSA CPU/GPU Queueing



Future of I/O from CPU/Arch Perspective
HSA CPU/GPU Queueing + NIC Queueing



Future of I/O from CPU/Arch Perspective
HSA IOMMU/Device Interactions



Future of I/O from CPU/Arch Perspective
HSA and I/O Devices

- Implications for OFA Software stack
  - Unified memory addressing via system-wide accessible page tables
  - Support for no pinning, on-demand paging options
  - Aligns with requirements from MPI and PGAS input to OFIWG to simplify memory registration

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

Thank You

OPENFABRICS ALLIANCE

10TH ANNUAL INTERNATIONAL DEVELOPER WORKSHOP