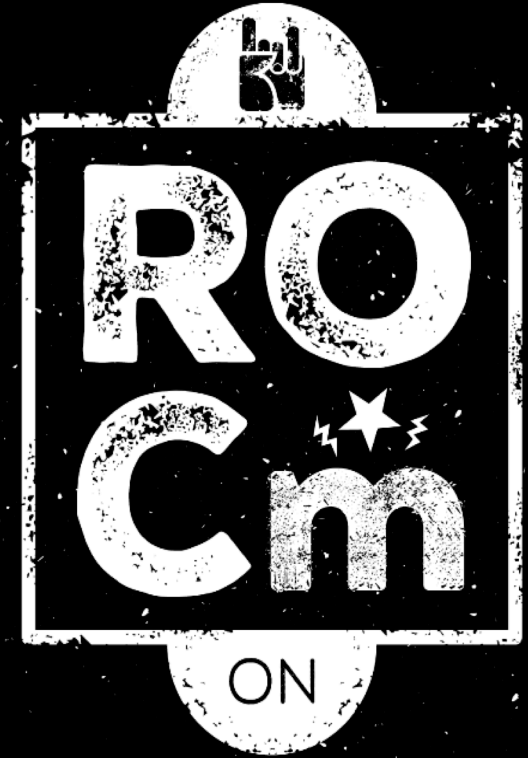


UCX-ROCm: ROCm Integration into UCX

{Khaled Hamidouche, Brad Benton}@AMD Research



ROCm: An open platform for GPU computing exploration

ROCm Software Platform

An Open Source foundation for Hyper Scale and HPC-class GPU computing

Graphics core next headless Linux® 64-bit driver

- Large memory single allocation
- Peer-to-Peer Multi-GPU
- Peer-to-Peer with RDMA
- Systems management API and tools



HSA drives rich capabilities into the ROCm hardware and software

- User mode queues
- Architected queuing language
- Flat memory addressing
- Atomic memory transactions
- Process concurrency & preemption



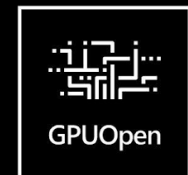
Rich compiler foundation for HPC developer

- LLVM native GCN ISA code generation
- Offline compilation support
- Standardized loader and code object format
- GCN ISA assembler and disassembler
- Full documentation to GCN ISA



“Open Source” tools and libraries

- Rich Set of “Open Source” math libraries
- Tuned “Deep Learning” frameworks
- Optimized parallel programming frameworks
- CodeXL profiler and GDB debugging



ROCm

Leverages OpenUCX For Scale-up and Scale-out Distributed Programming Models

- Next generation open source HPC communication framework
- Built off the foundation of MXM, UCCS, PAMI
- Broad Industry support including IBM, ARM, Mellanox, Nvidia, and AMD
- Rich platform for supporting MPI, OpenSHMEM, PGAS



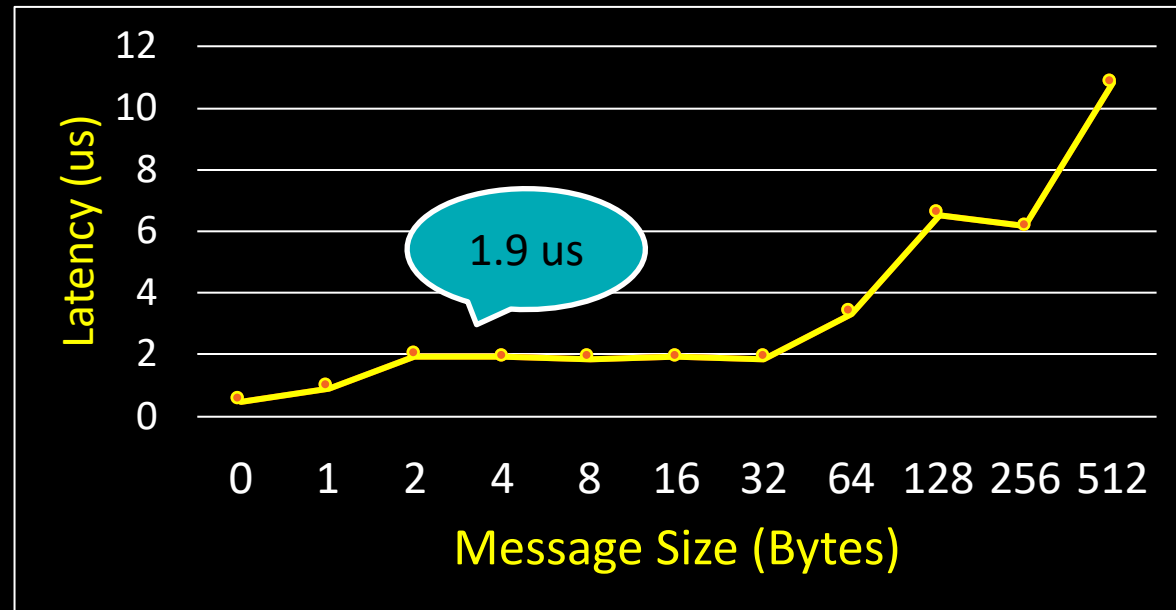
ROCm for Distributed Systems

- ▲ CPU can directly accesses GPU memory
 - Expose entire GPU frame buffer as addressable memory through PCIe BAR (LargeBar feature)
 - Map GPU pages to CPU pages
 - Allow CPU to directly load/store from/to GPU memory

- ▲ HCA to directly access GPU memory : ROCnRDMA feature
 - Leverages Mellanox's PeerDirect feature
 - Allows IB HCA to directly read/write data from/to GPU memory
 - Available and enabled by default in ROCm

UCX over ROCm: Intra-node support

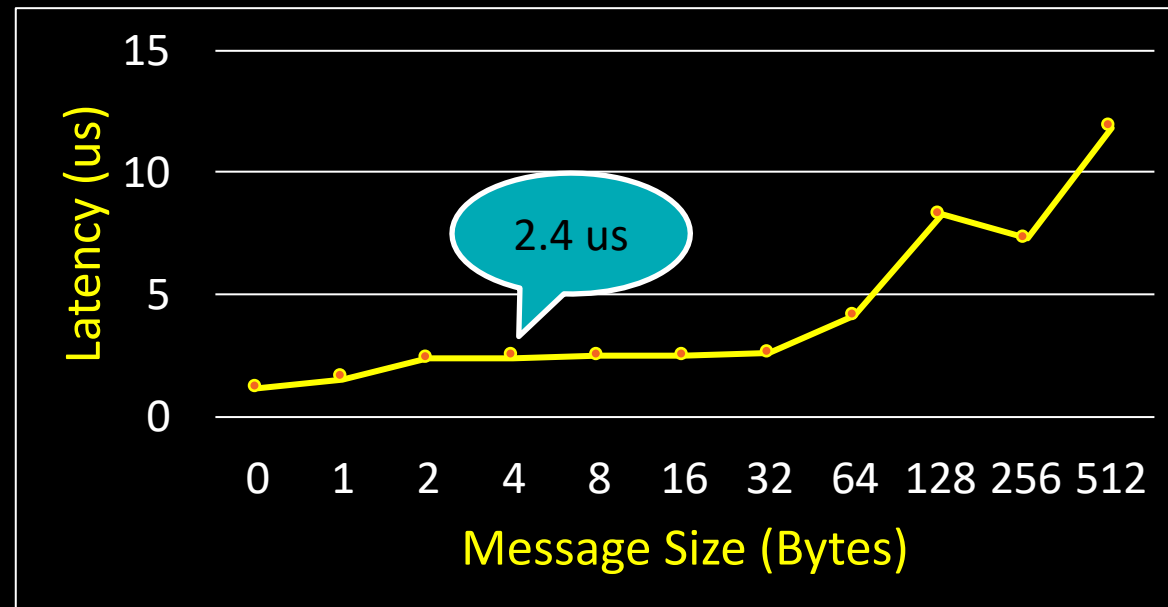
- ▲ Zero-copy based design
 - `uct_rocm_cma_ep_put_zcopy`
 - `uct_rocm_cma_ep_get_zcopy`
- ▲ Zero-copy based implementation
 - Similar to the CMA UCT code in UCX
 - ROCm provides similar functions to the original CMA for GPU memories
 - `hsaKmtProcessVMWrite`
 - `hsaKmtProcessVMRead`
- ▲ IPC for intra-node communication
 - Working on providing ROCm-IPC support in UCX
- ▲ Test-bed:
 - AMD FIJI GPUs, Intel CPU, Mellanox Connect-IB
 - OMB latency benchmark



- ▶ ROCM-CMA provides efficient support for large messages
- ▶ **1.9 us** for 4 Bytes transfer for intra-node D-D
- ▶ **43 us** for 512KBytes transfer for intra-node

UCX over ROCm: Inter-node Support

- ▲ Takes advantage of LargeBar capability to support eager protocols
 - Eager protocols can run directly from GPU buffers
- ▲ Take advantage of ROCnRDMA to design rendezvous (RNDV) protocols
- ▲ Optimization and tuning work in progress
 - Enhanced and optimized GPU-Aware protocols Pipeline, ...etc.



- ▶ LargeBar feature provides efficient support for eager protocol
- ▶ **2.4 us** for 4 Bytes transfer for inter-nodes

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

© 2018 Advanced Micro Devices, Inc. All rights reserved. AMD, the AMD Arrow logo, FirePro and combinations thereof are trademarks of Advanced Micro Devices, Inc. in the United States and/or other jurisdictions. ARM is a registered trademark of ARM Limited in the UK and other countries. PCIe is a registered trademarks of PCI-SIG Corporation. OpenCL and the OpenCL logo are trademarks of Apple, Inc. and used by permission of Khronos. OpenVX is a trademark of Khronos Group, Inc. Other names are for informational purposes only and may be trademarks of their respective owners. Use of third party marks / names is for informational purposes only and no endorsement of or by AMD is intended or implied.