

Introduction of Open MPI+UCX+UCC

Shinji Sumimoto

The University of Tokyo

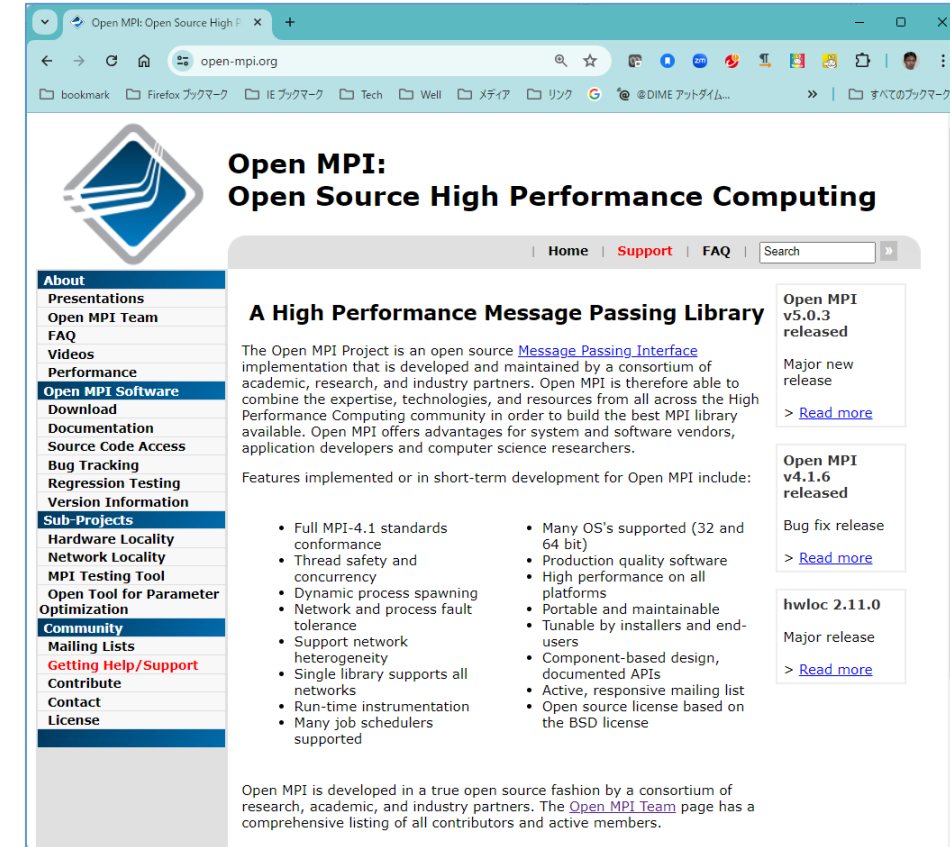
2024/6/28

Presentation Overview

- Introduction of
 - Open MPI
 - UCX
 - UCC

What is Open MPI

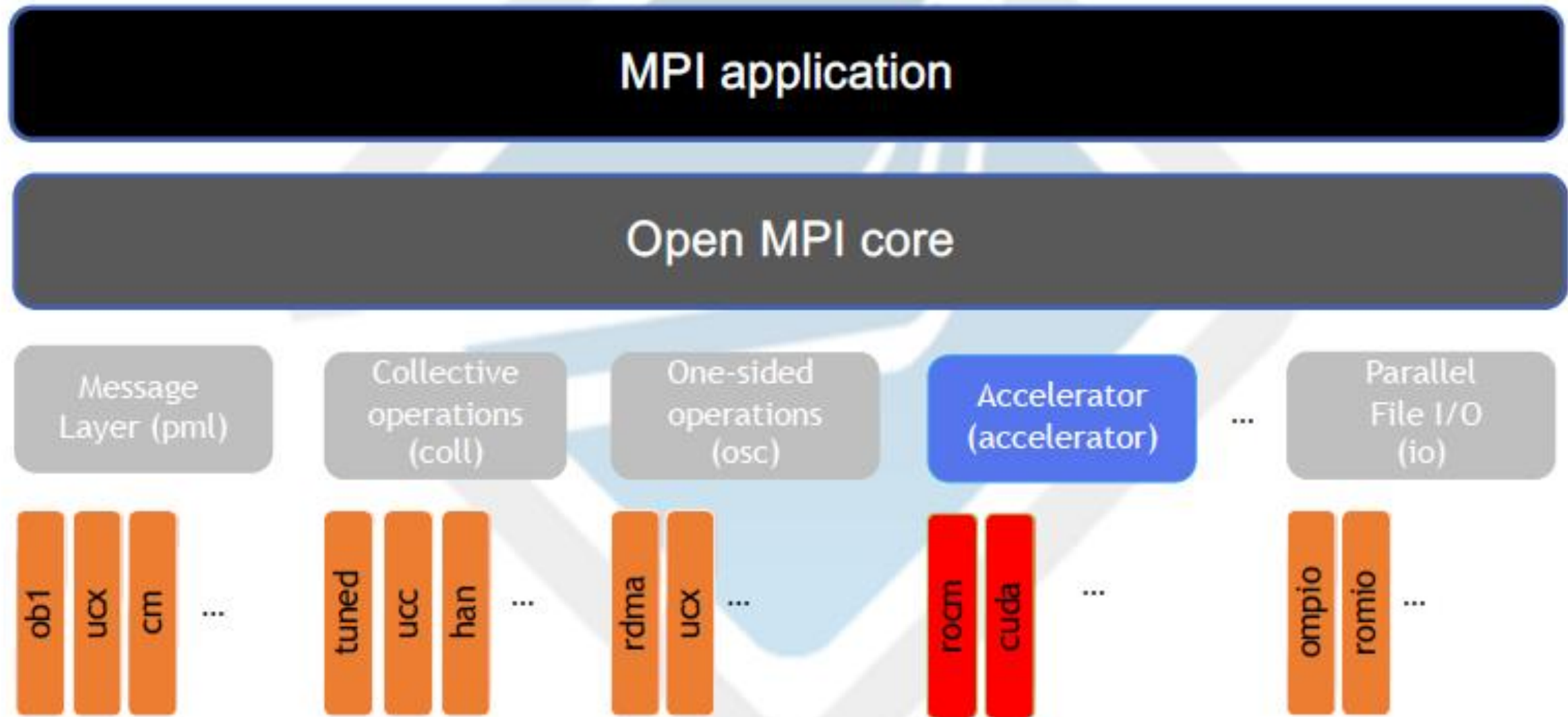
- [Open MPI](https://www.open-mpi.org/) is a community-based open-source implementations of the [MPI standard](https://www.open-mpi.org/)
- Short list of features
 - Full MPI-4.1 standards conformance
 - Thread safety and concurrency
 - Dynamic process spawning
 - Network and process fault tolerance
 - Run-time instrumentation
 - Many job schedulers supported
 - Accelerators ready (CUDA/ROCM/ZE)
- Composed of dynamic components that can be configured in/out depending on the setup and user needs



<https://www.open-mpi.org/>

From George Bosilca

Open MPI Software Stacks for GPU

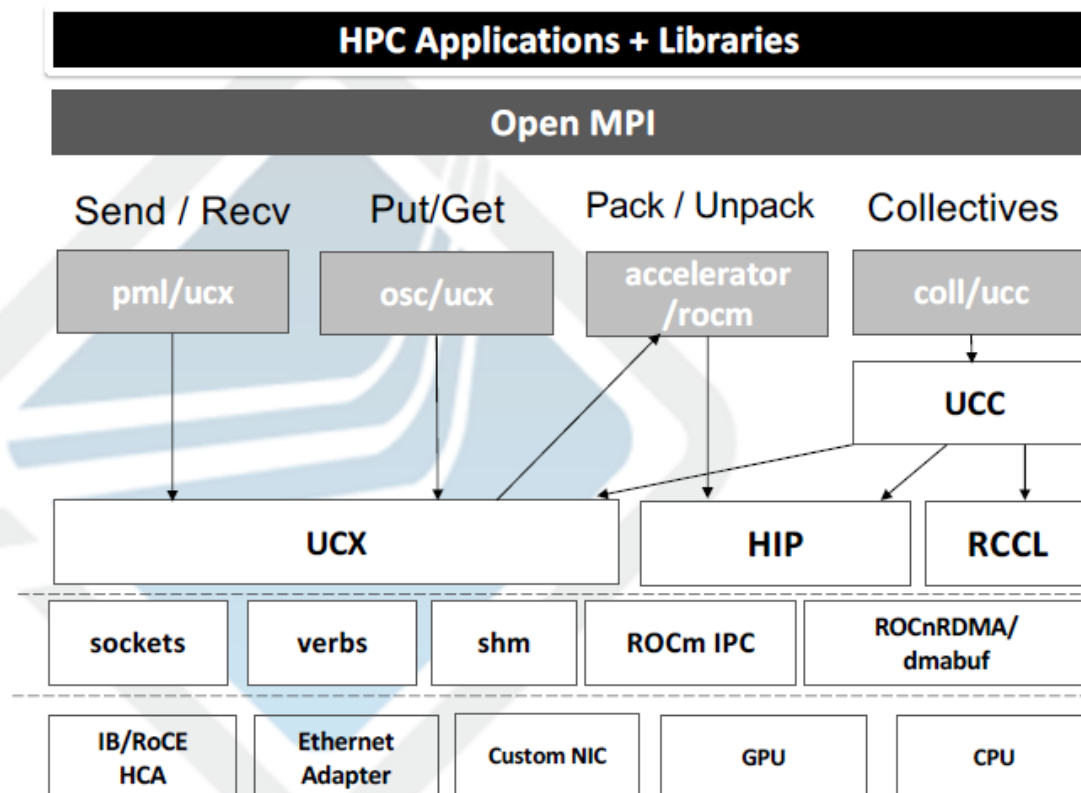


Open MPI Software Stacks

<https://www-lb.open-mpi.org/papers/sc-2023/Open-MPI-SC23-BOF.pdf>

ROCm Aware Open MPI Software Stack with UCX and UCC

- Recommended software stack for InfiniBand and RoCE networks
- Most stable and best tested configuration



UCX and UCC by UCF Consortium

Unified Communication Framework (UCF) Consortium



MISSION: Collaboration between industry, laboratories, and academia to create production grade communication frameworks and open standards for data centric, ML/AI, and high-performance applications

Projects & Working Groups

- **UCX – Unified Communication X** – www.openucx.org
- **UCC – Collective Library**
- **OpenSNAPI – Smart network Project**
- SparkUCX – www.sparkucx.org
- UCD – Advanced Datatype Engine
- HPCA Benchmark – Benchmarking Effort

Board members

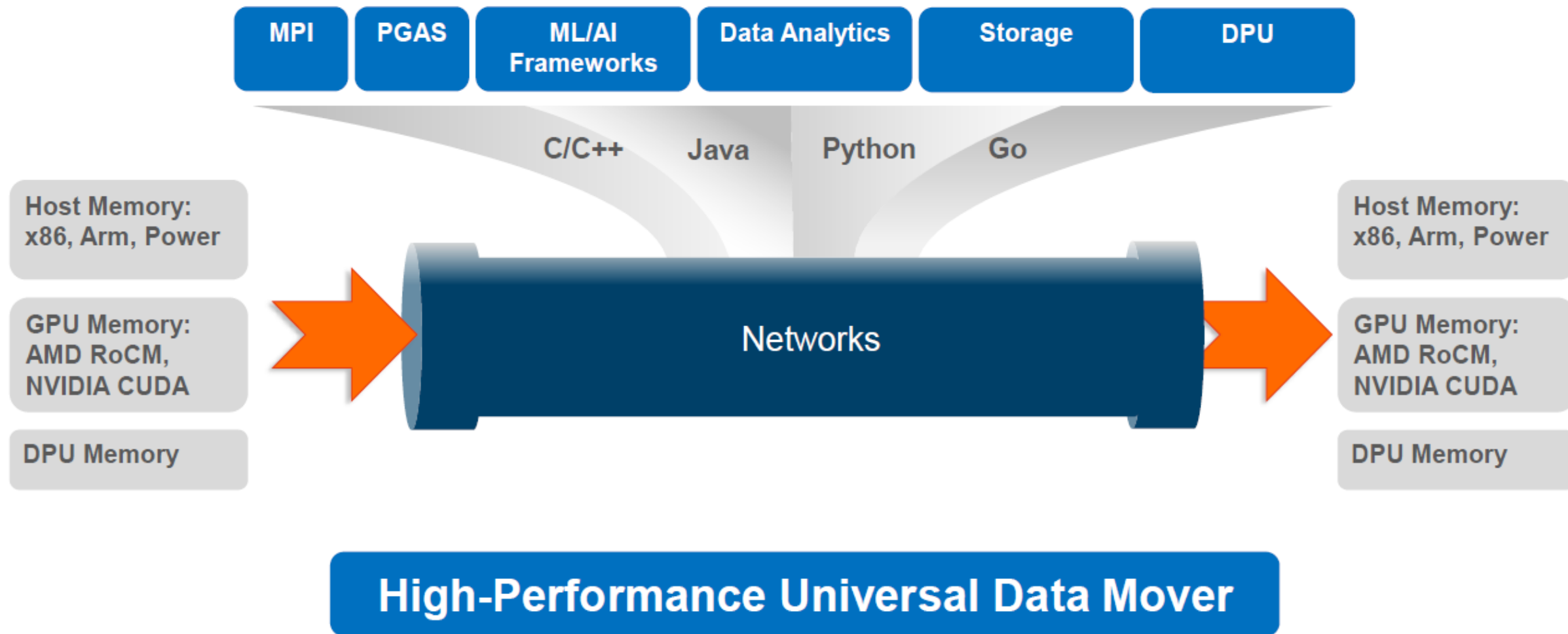
- **Jeff Kuehn**, UCF Chairman (Los Alamos National Laboratory)
- **Gilad Shainer**, UCF President (Nvidia)
- **Pavel Shamis**, UCF Treasurer (Arm)
- **Brad Benton**, Board Member (AMD)
- **Yanfei Guo**, Board Member (Argonne National Laboratory)
- **Perry Schmidt**, Board Member (IBM)
- **Dhableswar K. (DK) Panda**, Board Member (Ohio State University)
- **Steve Poole**, Board Member (Open Source Software Solutions)



Join

<https://www.ucfconsortium.org> or info@ucfconsortium.org

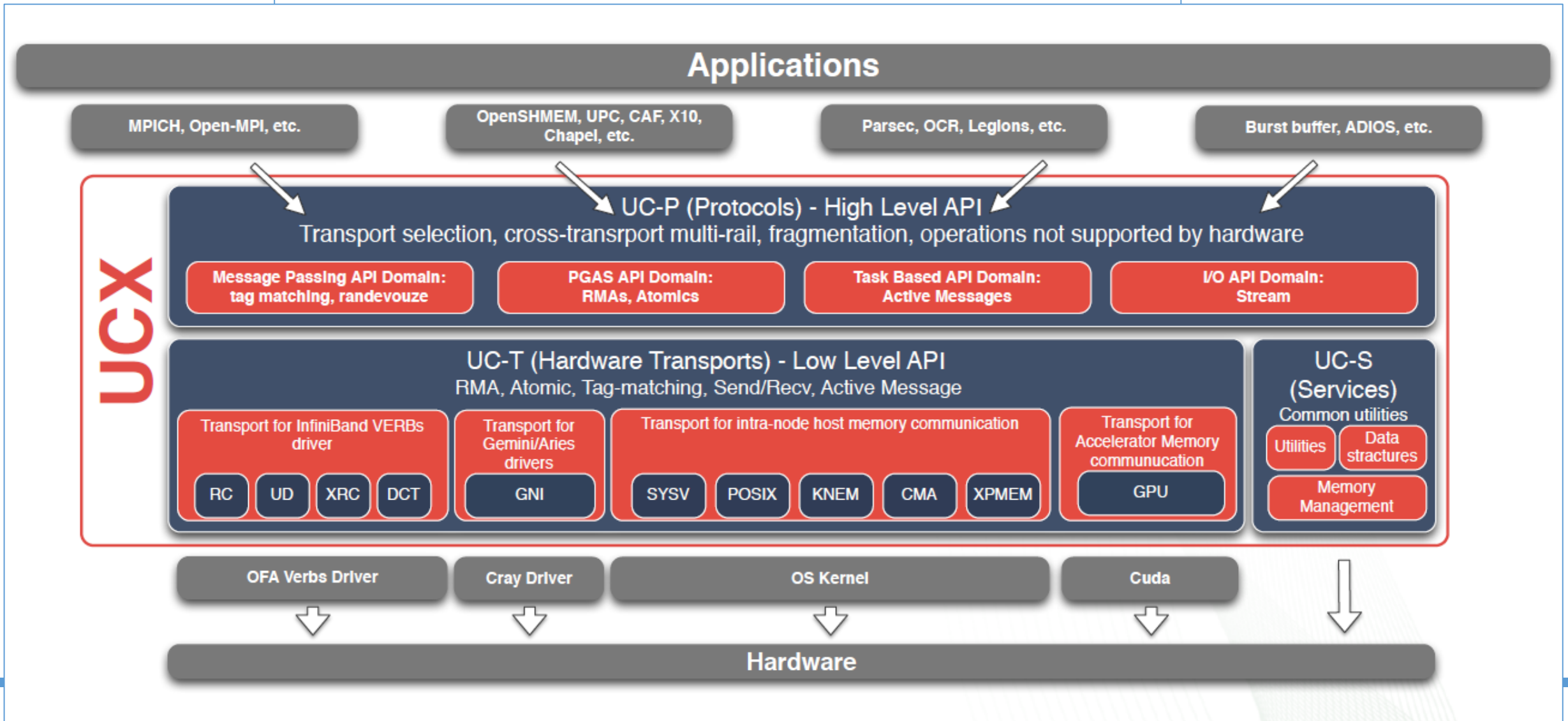
Why UCX ?



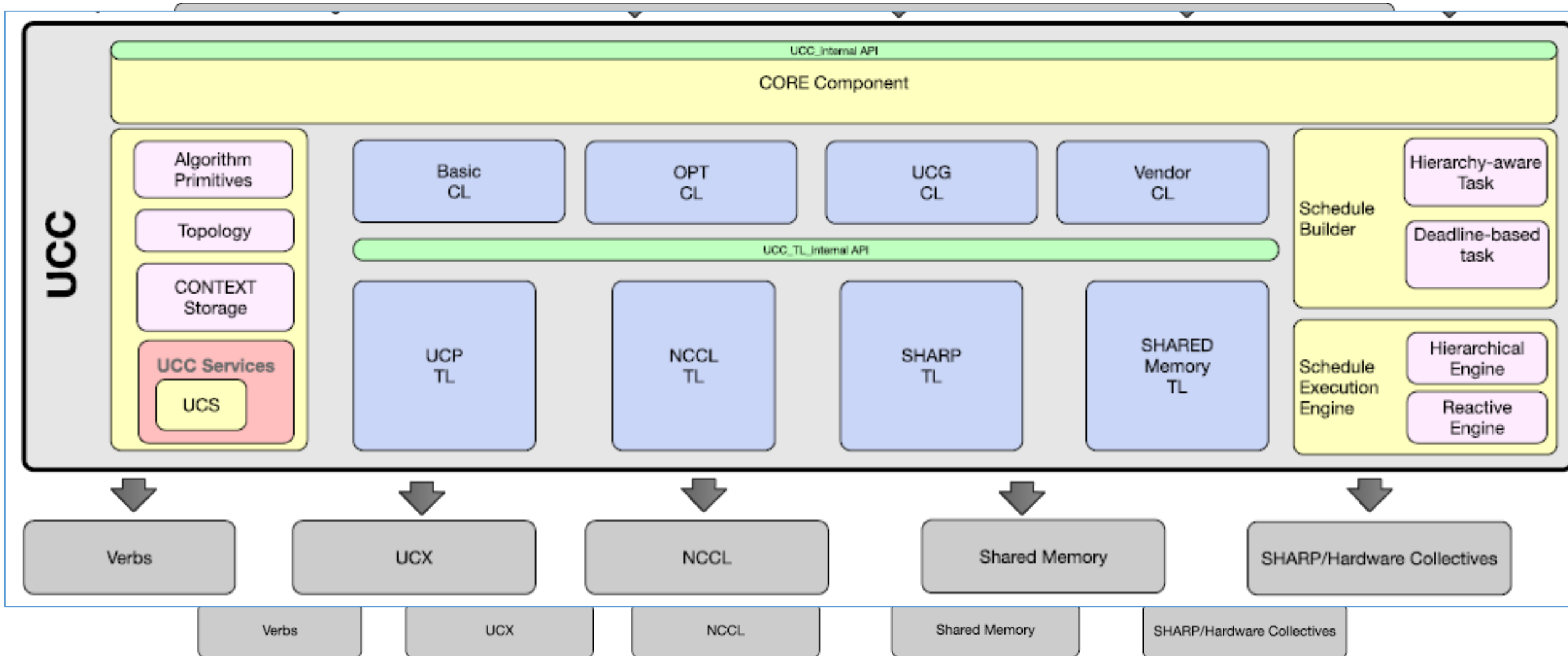
<https://openucx.org/wp-content/uploads/2021/07/UCF-UCX-BOF-ISC-2021.pdf>

UCX: High-level Overview

<https://openucx.org/wp-content/uploads/2015/11/UCX-booth-presentation.pdf>



UCC Reference Implementation: Component Diagram



<https://openucx.org/wp-content/uploads/2021/07/UCF-UCX-BOF-ISC-2021.pdf>

- Unified collective stack for HPC and DL/ML workloads
 - Need to support a wide variety of semantics
 - Need to optimize for different performance sensitives - latency, bandwidth, throughput
 - Need for flexible resource scheduling and ordering model
- Unified collective stack for software and hardware transports
 - Need for complex resource management - scheduling, sharing, and exhaustion
 - Need to support multiple semantic differences – reliability, completion
- Unify parallelism and concurrency
 - Concurrency – progress of a collective and the computation
 - Parallelism – progress of many independent collectives
- Unify execution models for CPU, GPU, and DPU collectives
 - Two-way execution model – control operations are tightly integrated
 - Do active progress, returns values, errors, and callbacks with less overhead
 - One-way execution model – control operations are loosely integrated
 - passive progress, and handle return values (GPU/DPUs)

<https://openucx.org/wp-content/uploads/2021/07/UCF-UCX-BOF-ISC-2021.pdf>

Questions?