

# MPI Advance : Open-Source Message Passing Optimizations

Amanda Bienz, Derek Schafer, Anthony Skjellum (speaker)

[askjellum@tntech.edu](mailto:askjellum@tntech.edu)

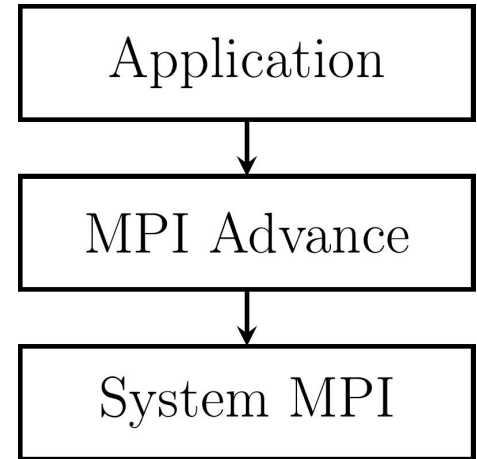
EuroMPI 2023

September 11<sup>th</sup>, 2023



# Motivation

- A collection of lightweight libraries that sit on top of MPI
- Four main motivations:
  - Early access to new MPI functionality before MPI implementations
  - Conceptual prototyping + sharing of pre-standard features
  - Run new on older MPIs (for some cases)
  - Optimizations of existing APIs
- Middle ground enables optimization opportunities from each end
  - Applications adapt code to new to APIs
  - Implementations get implementation knowledge
- Current Collections:
  - Locality Aware Collectives
  - MPIPCL



# Locality Aware MPI Collective Operations

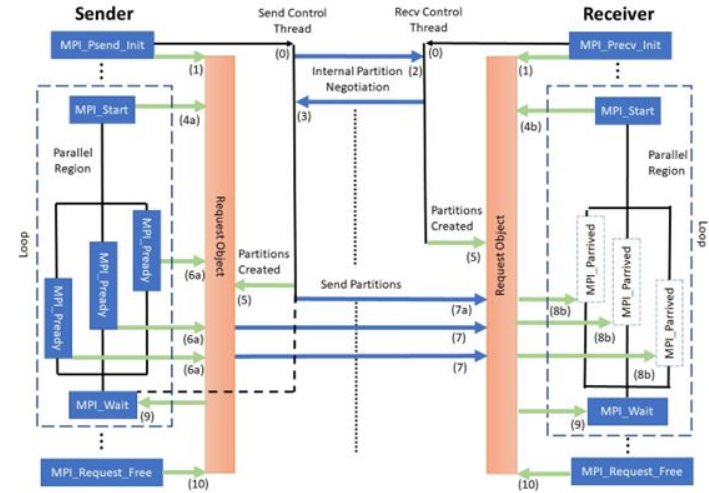
- Designed to fit codebases with minimal changes to existing code
- Locality-Aware Collectives: Allgather, Alltoall, Alltoallv
- Persistent Neighborhood Collectives:
  - Neighbor Alltoallv, Neighbor Alltoallw
  - Requires use of special topology communicator (`dist_graph_create`)
- Optimizations of the two above libraries w/ respect to GPUs
- Integrated into Trilinos

Bienz A, Gropp WD, Olson LN. Reducing communication in algebraic multigrid with multi-step node aware communication. *The International Journal of High Performance Computing Applications*. 2020;34(5):547-561.



# MPIPCL

- Implements all MPI 4.0 partitioned communication APIs
- Layered library on top of existing MPI implementations
- Technical details:
  - Uses MPI Persistent P2P APIs
  - Uses progress thread for partition negotiation
- Partitioned collectives under prototyping and experimentation



M. G.F. Dosanjh, A. Worley, D. Schafer, P. Soundararajan, S. Ghafoor, A. Skjellum, P. V. Bangalore, R. E. Grant, Implementation and evaluation of MPI 4.0 partitioned communication libraries, *Parallel Computing*, Volume 108, 2021, <https://doi.org/10.1016/j.parco.2021.102827>.

# Looking Forward

- Contributors across spectrum of MPI-4+ features
- Tighter integration of the sub-libraries
- Abstraction of common features
  - Both libraries require custom MPI Request API
  - Each has their own set, so a shared interface would allow better cohesion
  - Feedback mechanism from users
  - Bootstrap progress thread support for add-on feature sets
- <https://github.com/mpi-advance>

# Thank you!

Funding in part is acknowledged from NSF Grants CCF-2151022, CCF-1918987, CCF-1562306, CCF-1822191, CCF-1821431, OAC-1923980, OAC-1549812, OAC-1925603, OAC-2201497, and CCF-2151020, as well as and the U.S. Department of Energy's National Nuclear Security Administration (NNSA) under the Predictive Science Academic Alliance Program (PSAAP-III), Award DE-NA0003966.

