SC21 MPICH BoF

MPI@Intel

Maria Garzaran

Intel® MPI 2021.4 Update

- What's new:
  - 3rd Generation Intel® Xeon® Scalable Processors support and optimizations
  - Release and release_mt library paths unification
  - Migrated to libfabric 1.13.x
- Enhancements:
  - Intel GPU buffers support optimizations (basic pipelining/overlap support)
  - AWS/GCP performance tuning (OFI/efa and OFI/tcp)
  - Intel® Ethernet 800 Series Network Adapter support enhancements
    - OFI/psm3 provider integration and tuning
  - Startup time optimization

MPICH for Aurora Update

- Support for Intel GPUs
- Support for multiple NICs per node
- Improvements to Threading
- Contributions to Collectives
- Others

MPICH Supports Intel GPUs

- Support for Intel GPUs is feature complete (two-sided, one-sided, and collectives are supported)
  - Uses oneAPI Level Zero
  - Includes optimizations for intra-node communication using Inter Process Communication (IPC) primitives
  - Yaksa data type engine supports Intel GPUs
  - Additional optimizations are on the way

Multiple NICs per Node

- Balanced distribution of ranks to NICs based on rank-to-socket affinity
- Optimizations for a single rank to use multiple NICs
  - Stripe large messages across multiple NICs
  - Hash different messages through different NICs
- Environment variables allow user to choose options about NIC utilization
- Added communicator info hints to distribute traffic across NICs

Improved Threading Support

- Collaborated with Argonne to add efficient support for threading
- Add support for threading using OFI FI_THREAD_SAFE mode
- Added communicator info hint for VCI (context) selection
- Support for affinity for asynchronous progress thread
- Tests to validate additional threading support

Others

- Support for QMPI
- Infrastructure to support algorithm selection for each collective
- Coming soon optimized high radix algorithms for collectives (see our EuroMPI 2021 paper)


intel