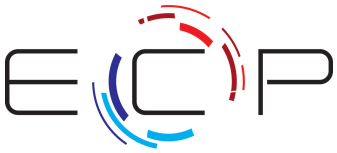
ECP Community BOF

MPI@Intel

Nusrat Islam
nusrat.islam@intel.com



Contributions to Open Source MPICH - I

Supporting Intel GPU in Yaksa using oneAPI Level Zero (in open source Yaksa)

- Support for packing/unpacking of non-contiguous data
- Support for reduction operations
  - Worked with Argonne to add support for host-based reduction

Supporting GPU in MPI Communication

- Infrastructure to support Intel GPUs (in MPICH 4.0a1 release)
  - Support for fallback path for pt2pt and collectives
- Support for reduction and one-sided compute operations in GPU (coming soon)
  - Leverages the reduction support in Yaksa
  - No need to move data to the host before performing reduction
- Optimizations for in-node Inter-Process Communication (IPC) with Intel GPUs (coming soon)

Support for Intel GPUs is now feature complete!!!

Contributions to Open Source MPICH - II

Support for Multiple NICs per Node (upstreaming in progress)

- Map ranks to NICs based on Numa node affinity of the ranks in a balanced manner
- Optimizations for a single rank to use multiple NICs
  - Stripe large messages across multiple NICs
  - Multiplex different messages through different NICs
- Configuration options for application programmers to use multiple NICs

Contributions to Open Source MPICH - III

Support for Lightweight Profiling (in MPICH main)

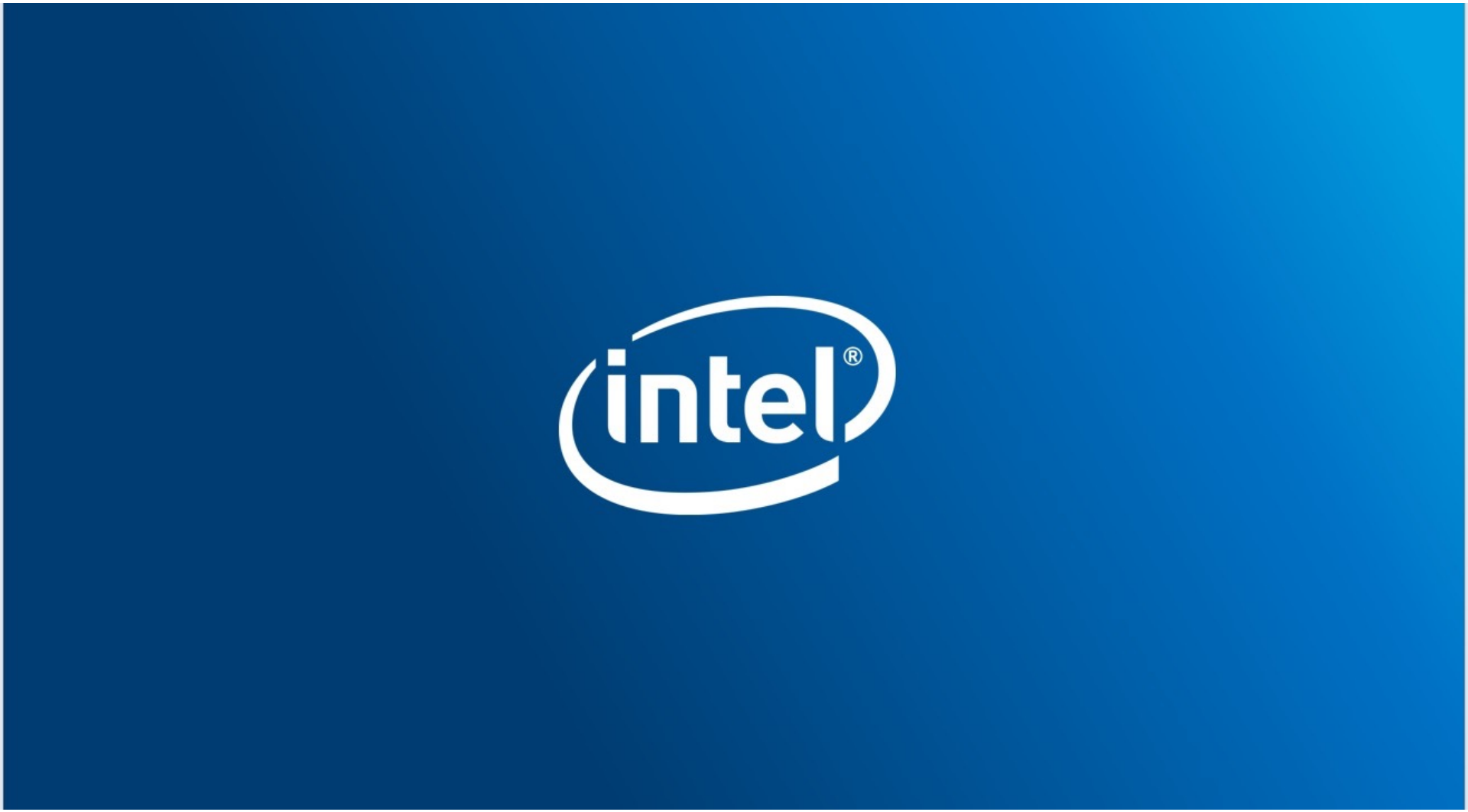
- Implement "QMPI" support in MPICH
  - Leading MPI Forum's work to standardize QMPI
  - Supports multiple tools simultaneously
  - Continues support for "legacy" PMPI tools
  - Minimal performance impact
- Implement profiling information to expose multi-NIC usage

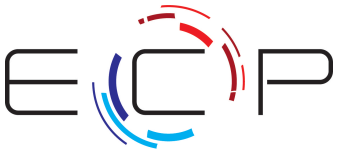
Others (in MPICH 4.0a1 release)

- Support for MPI 4.0
  - New info hints
  - New persistent API for collectives
- Improved intra-node pt2pt communication with a per-process shared queue
- Infrastructure to support algorithm selection for collectives

Intel® MPI Library 2021 Features

- Amazon* AWS/EFA, Google* GCP support enhancements
  - Enable support and specific tuning for AWS EFA H/W (using their OFI provider)
  - Enable support for Google GCP NIC H/W (using OFI/TCP)
  - Support for Microsoft Azure and Oracle cloud
- Intel GPU pinning and GPU buffers support
  - Optimal placement of ranks and efficient data transfer to/from GPUs
- Optimizations for Intel® Xeon® Platinum 9282/9242/9222/9221 family
  - Platform recognition and specific tuning for HW parameters
- Mellanox* ConnectX*-3/4/5/6 (FDR/EDR/HDR) support enhancements
  - Evaluate HW specific features of MLNX solutions
- Distributed Asynchronous Object Storage (DAOS) file system support
  - Optimized stack for integration with DAOS
- mpitune_fast functionality improvements
  - Special tool to optimize MPI library tuning time for specific application and/or cluster topology/scale


intel


EXASCALE COMPUTING PROJECT