Argonne NATIONAL LABORATORY

MPICH: A High-Performance Open Source MPI Library for Leadership-class HPC Systems

Ken Raffenetti, Hui Zhou, Ethan Wong, Rajeev Thakur

Cass BoF Days 2026

February 12th

U.S. DEPARTMENT OF ENERGY

Schedule

- Welcome - Ken Raffenetti
- MPICH Update - Hui Zhou
- Partner updates
  - HPE (Kim McMahon)
  - OSU (Nat Shineman)
  - Partec (Simon Pickartz)
  - Intel (Gengbin Zheng)
- User presentations
  - Nathan Nichols (ANL/ALCF)
  - Junchao Zhang (ANL/PETSc)
- Q&A

MPICH: Status and Upcoming Releases
http://www.mpich.org

Ken Raffenetti, Hui Zhou, Ethan Wong, Rajeev Thakur

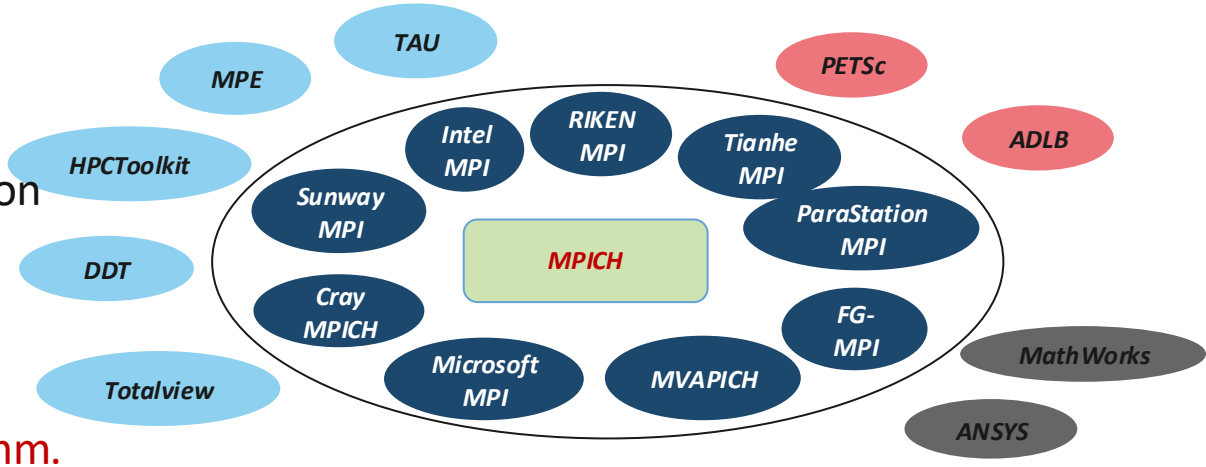
Argonne National Laboratory

MPICH turns 34 (since 1992)

U.S. DEPARTMENT OF ENERGY

The MPICH Project

- Funded by DOE for 34 years
- Has been a key influencer in the adoption of MPI
  - First/most comprehensive implementation of every MPI standard
  - Allows supercomputing centers to not compromise on what features they demand from vendors
- DOE R&D100 award in 2005 for MPICH
- DOE R&D100 award in 2019 for UCX (MPICH internal comm. layer)
- 2024 ACM Software System Award
- MPICH and its derivatives are the world's most widely used MPI implementations



MPICH is not just a software

It's an Ecosystem

MPICH in 6 of Top 10 Supercomputers

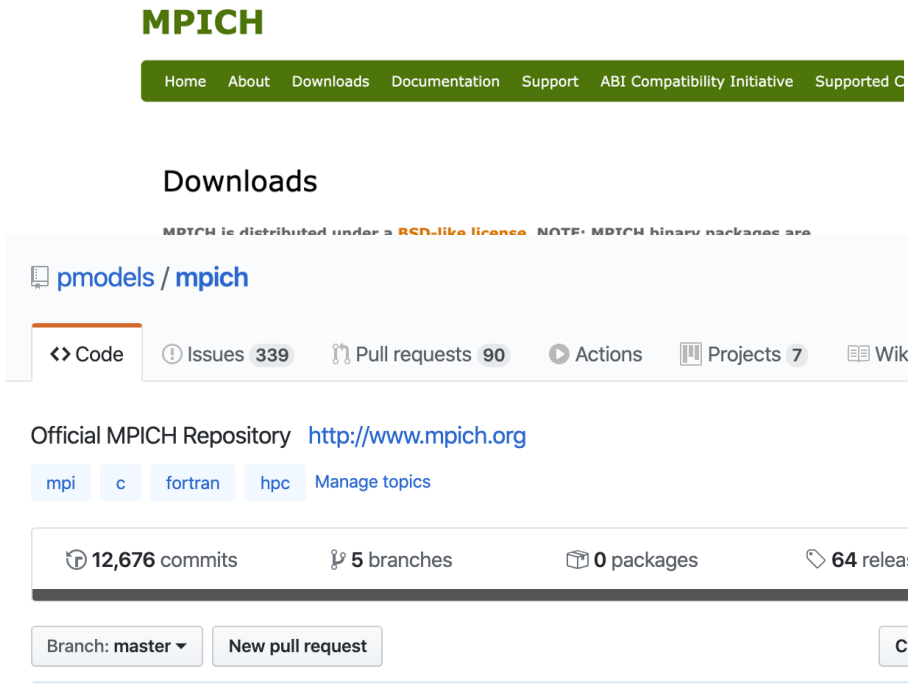
1. **El Capitan, LLNL, USA (Cray MPI)**
2. **Frontier, ORNL, USA (Cray MPI)**
3. **Aurora, ANL, USA (Intel MPI for Aurora)**
4. Jupiter Booster, EuroHPC/FZJ, Germany
5. Eagle, Microsoft, USA
6. **HPC 6, Eni S.p.A, Italy (Cray MPI)**
7. Fugaku, RIKEN, Japan
8. **Alps, CSCS, Switzerland (Cray MPI)**
9. **LUMI, EuroHPC/CSC, Finland (Cray MPI)**
10. Leonardo, EuroHPC/CINECA, Italy

ABI Compatibility

- MPICH have binary compatibility for MPI implementations
  - Started in 2013
  - Explicit goal of maintaining ABI compatibility between multiple MPICH derivatives
  - Collaborators:
    - MPICH (since v3.1, 2013)
    - Intel MPI Library (since v5.0, 2014)
    - Cray MPICH (starting v7.0, 2014)
    - MVAPICH2 (starting v2.0, 2017)
    - Parastation MPI (starting v5.1.7-1, 2017)
- MPI ABI supported starting MPICH v5.0.0
  - New standard moving forward
  - Opt-in with mpicc_abi

MPICH Distribution Model

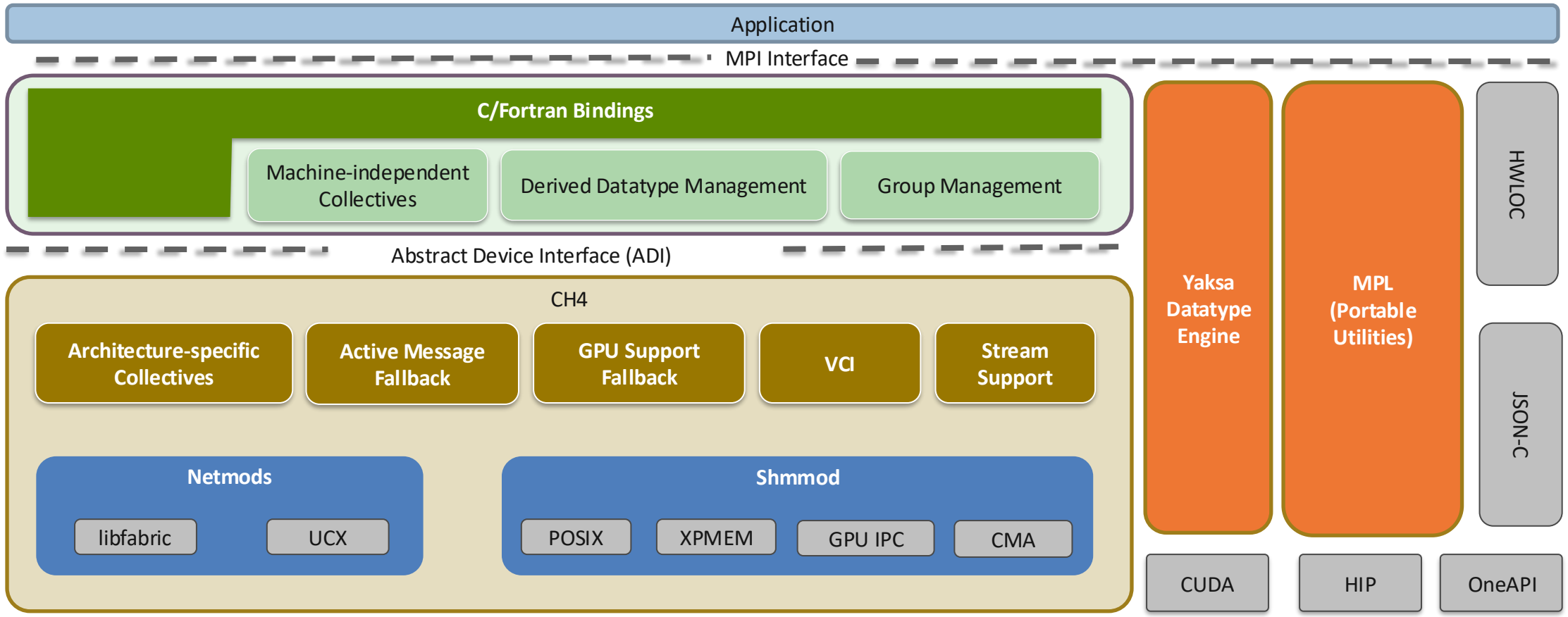
- Source Code Distribution
  - MPICH Website, Github
- Binary Distribution through OS Distros and Package Managers
  - Redhat, CentOS, Debian, Ubuntu, Homebrew (Mac)
- Distribution through HPC Package Managers
  - Spack, OpenHPC, E4S
- Distribution through Vendor Derivatives



MPICH Releases

- MPICH currently follow a 12-month cycle for major releases
  - Actively develop features in the main branch and backport bug fixes to the stable branch
  - Point stable releases based on accumulated bug fixes
  - Preview alpha releases based on major feature development
  - Beta release in November, feature frozen into new stable branch
  - GA release usually in January
- Current stable release is in the 5.0.x series
  - We just released mpich-5.0.0 on 2/4/2026
- Upcoming major release is in the 5.1.x series
  - New development has started
  - Planned beta release in Nov. 2026
  - Bug fixes will be back ported to 5.0.x

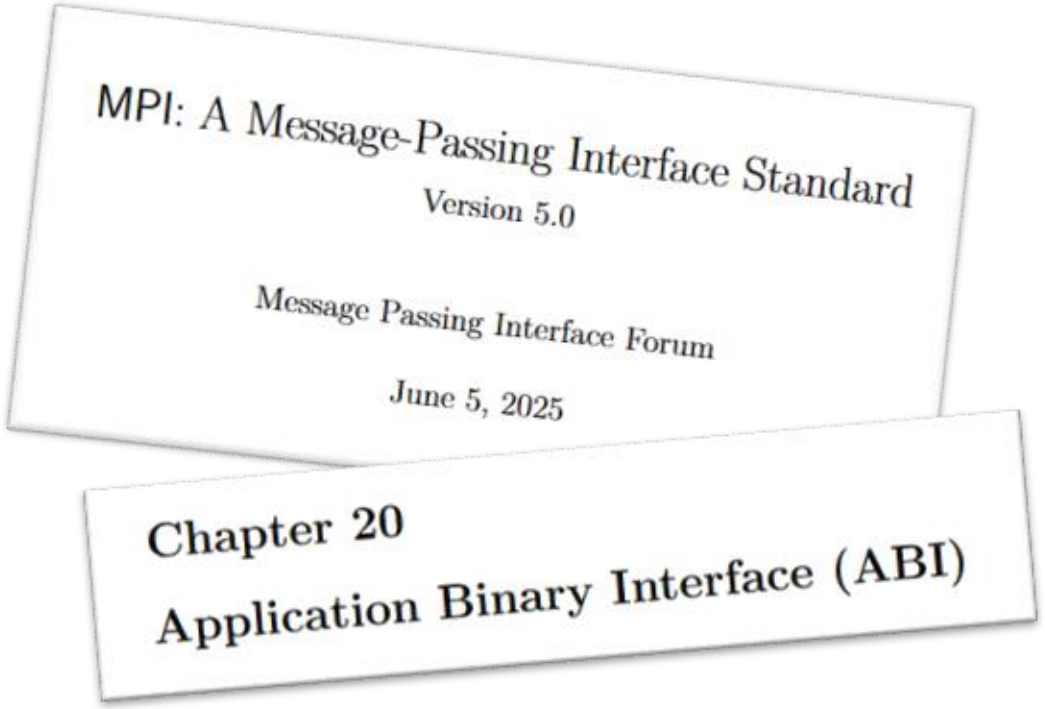
MPICH Layered Structure

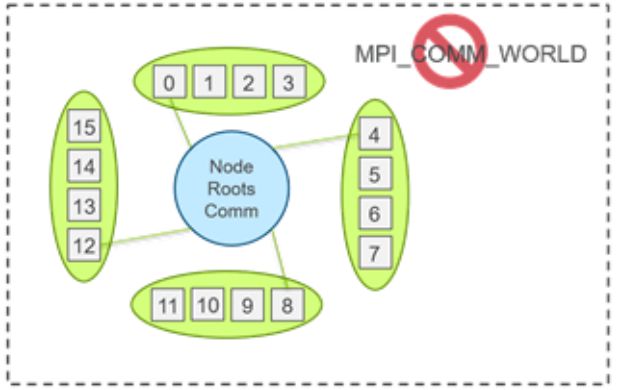


MPICH 5.0 Release

MPICH 5.0 Release

- Full support for MPI 5.0 standard including support for MPI ABI
- Internal datatypes
- True MPI Sessions implementation
- CH4 native netmod RNDV path





Full Support for MPI-ABI

- Build MPICH ABI by default
- Build both by –enable-mpi-abi
- MPI ABI may become the default in the next release

- Switched to internal datatypes to support ABI built-in datatypes

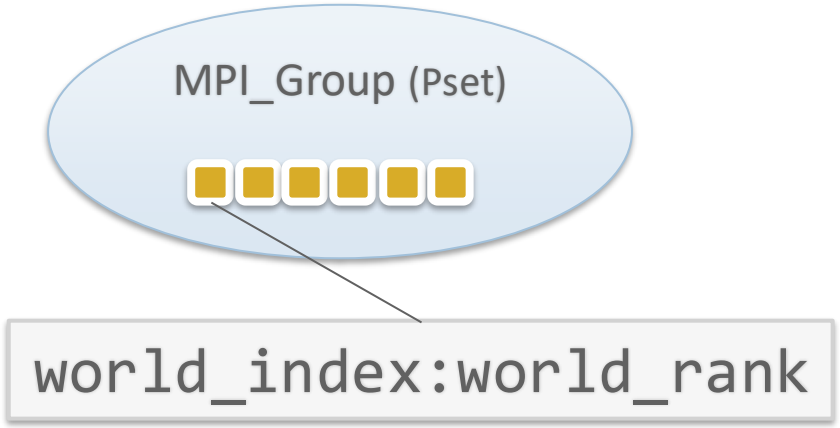
```
$ ./autogen.sh
$ ./configure --prefix=[path]
              --enable-mpi-abi
$ make install

$ mpicc_abi -o cpi examples/cpi.c
$ mpirun -n 2 ./cpi
$ ldd ./cpi |grep libmpi

libmpi_abi.so.0 =>
/opt/MPI/lib/libmpi_abi.so.0
(0x00007ff1883e9000)
```

Implementing True MPI Sessions

- Remove the dependency on MPI_COMM_WORLD
- Device independent MPI Groups
- Supporting dynamism



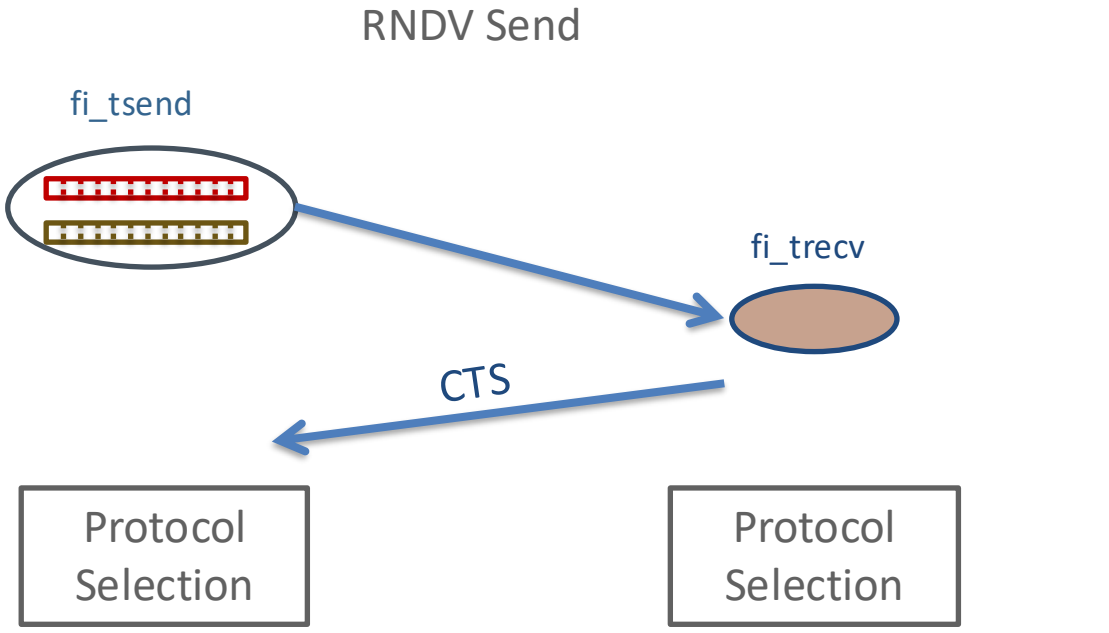
CH4:OFI RNDV path

- Supllement lower-layer network with Rendezvous protocols

Set:

MPIR_CVAR_CH4_OFI_EAGER_THRESHOLD

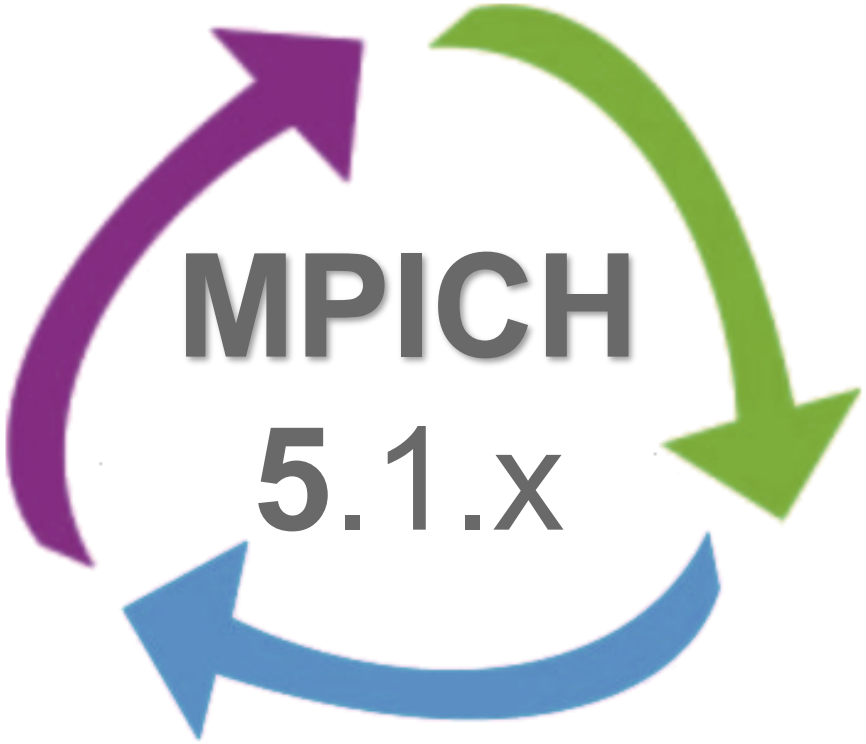
to enable the native RNDV path



Pipelining | RDMA read | RDMA write | direct

MPICH 5.1 Development Plan (RFC)

- Fortran bindings on top of MPI ABI
- Redesign Collective Selection Framework
- Pipelined collective algorithms
- Improve MPI dynamism via sessions and dynamic processes



MPICH Fortran Binding on top of MPI ABI

- `mpifort` will still link with MPICH ABI even with `-enable-mpi-abi`
- Fortran MPI binary won't work with MPI ABI yet
- We will migrate MPICH Fortran binding to an independent package
  - Can be configured and built independent of C MPI
  - Will depend on a few MPIX extensions
    - User-defined error handlers
    - User-defined Ops
    - Attributes
    - Stub code will be used to build against any C MPI

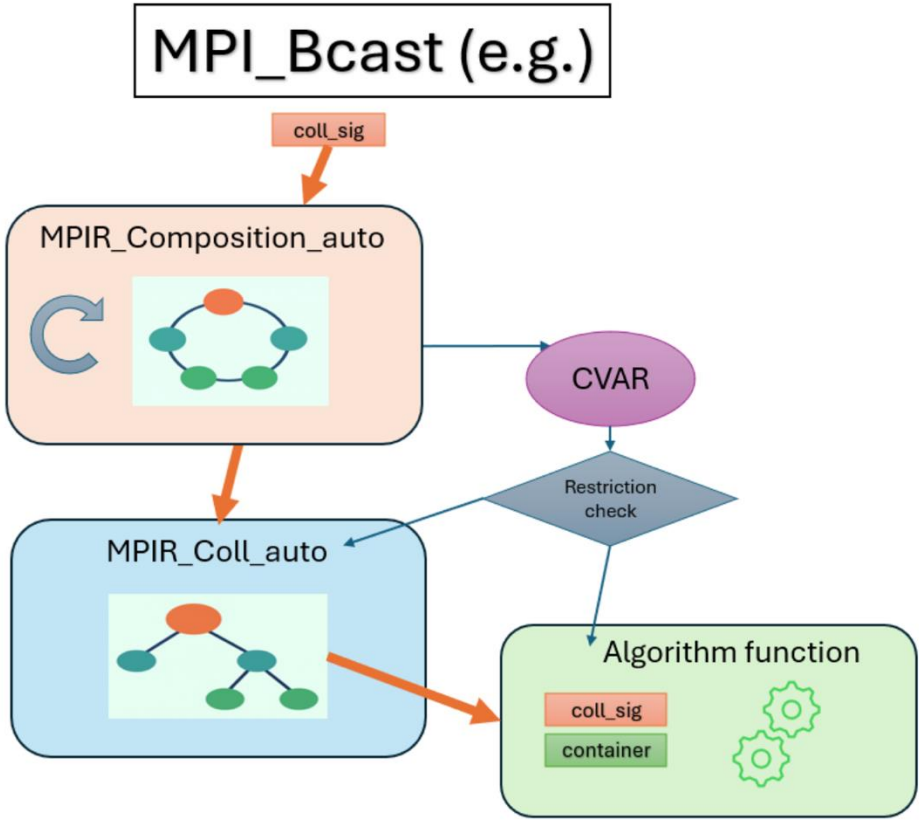
- Plan to make MPI ABI default in 5.1.x
  - Only build a single ABI based on configure option
  - `mpicc` and `mpifort` link with the consistent ABI
  - Build MPICH Fortran binding on top of Open MPI once ABI is supported

Fortran Binding + MPI ABI

Redesign Collective Selection Framework

- MPICH supports custom collective algorithm tuning via JSON configuration files
- However, old design doesn't work well with algorithms across ADI and device modules

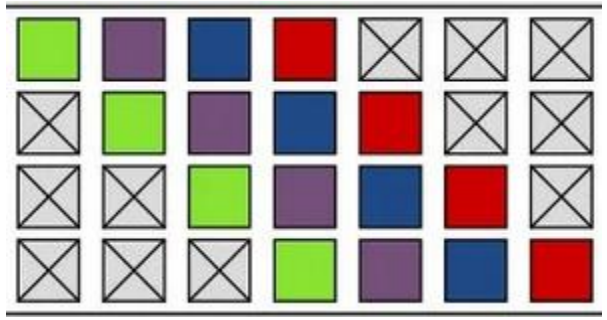
- The plan:
  - Single spec file for all algorithms and conditions
    - Support JSON subtrees
  - Support device-layer algorithms
  - Better algorithm usage debugging / reporting



On-going Focuses

- Pipelined Collectives
  - MPI should out-perform alternatives in collectives
  - MPI should remain an excellent abstraction: portable and performant

- Better support for dynamism
  - MPI is all about collective parallelism
  - However, MPI can use dynamism in
    - Getting in
    - Getting out
    - Recovering from fault
  - Dynamism is necessary for MPI to step out of the ivory tower
    - More robust
    - More performant
    - ULFM



Latency hiding w. pipelining

MPI Sessions

MPI Dynamic Processes

ULFM: User-Level Fault Tolerance

Argonne NATIONAL LABORATORY

Thank you!

- https://www.mpich.org
- Mailing list: discuss@mpich.org, devel@mpich.org, announce@mpich.org
- Issues and Pull requests: https://github.com/pmodels/mpich
- Weekly development call every Thursday at 9am (central).
    - Agenda and meeting links sent to devel@mpich.org ahead of the meeting

CASS

U.S. DEPARTMENT OF ENERGY