

# MPI Forum ABI WG Status and Discussion

Dublin meeting  
September, 2008

# Context

- “convenience” item
- Will allow dynamic run-time re-link (e.g. through `LD_LIBRARY_PATH`)
- Of most use to software providers that layer on top of MPI (ISVs, tools, parallel apps, etc.)
- Some push back from implementers

# Status

- Focus on C bindings
- Hope that fortran WG will eliminate fortran specific bindings in the long run
- First step: agree on a common mpi.h spec
- After that: other items as needed
  - Environmental items (e.g. mpiexec)
  - Calling/Linkage conventions
  - ...

# Participating organizations

Initiating organizations:

- MathWorks
- Intel Corp.
- Microsoft Corp.
- QLogic
- HP

# Members of working group

- [alexander.supalov at intel.com](mailto:alexander.supalov@intel.com)
- [davdav\\_2000 at yahoo.com](mailto:davdav_2000@yahoo.com)
- [david.solt at hp.com](mailto:david.solt@hp.com)
- [dries.kimpe at cs.kuleuven.be](mailto:dries.kimpe@cs.kuleuven.be)
- [edric.ellis at mathworks.co.uk](mailto:edric.ellis@mathworks.co.uk)
- [erezh at microsoft.com](mailto:erezh@microsoft.com)
- [gil at dev.mellanox.co.il](mailto:gil@dev.mellanox.co.il)
- [hayasaka at bx.jp.nec.com](mailto:hayasaka@bx.jp.nec.com)
- [howardp at cray.com](mailto:howardp@cray.com)
- [jeffb at lanl.gov](mailto:jeffb@lanl.gov)
- [jsquyres at cisco.com](mailto:jsquyres@cisco.com)
- [kannan.narasimhan at hp.com](mailto:kannan.narasimhan@hp.com)
- [keller at hirs.de](mailto:keller@hirs.de)
- [kmm at cray.com](mailto:kmm@cray.com)
- [lindahl at pbm.com](mailto:lindahl@pbm.com)
- [mercierg at mcs.anl.gov](mailto:mercierg@mcs.anl.gov)
- [ndebard at lanl.gov](mailto:ndebard@lanl.gov)
- [rabenseifner at hirs.de](mailto:rabenseifner@hirs.de)
- [sameer at cs.uoregon.edu](mailto:sameer@cs.uoregon.edu)
- [schliephake at hirs.de](mailto:schliephake@hirs.de)
- [shirley at eecs.utk.edu](mailto:shirley@eecs.utk.edu)
- [terry.dontje at sun.com](mailto:terry.dontje@sun.com)
- [treumann at us.ibm.com](mailto:treumann@us.ibm.com)
- [yann.kalemkarian at bull.net](mailto:yann.kalemkarian@bull.net)
- [jjolly at sandia.gov](mailto:jjolly@sandia.gov)
- [koziol at hdfgroup.org](mailto:koziol@hdfgroup.org)
- [srimks11 at gmail.com](mailto:srimks11@gmail.com)

# Status of mpi.h effort

- Spreadsheet in development that
  - Covers items in annex A
    - constants, type definitions, info values and keys
  - Classifies items, (Organized by Bill Gropp's paper on GMPI)
    - Compile-time values
      - Used in declarations (MPI\_MAX\_ERROR\_STRING, etc.)
      - Other (MPI\_ANY\_SOURCE, MPI\_ERR\_TRUNCATE, etc.)
    - Init-time constants (MPI\_INT, etc.)
    - Opaque objects (MPI\_Comm, MPI\_Datatype, etc.)
    - Defined objects (MPI\_Status)
    - Defined pointers (MPI\_BOTTOM, MPI\_STATUS\_NULL, etc.)
  - Exposes alternate implementations
    - OpenMPI, MPICH 1 and 2, HP, LAMPI, NEC
  - Beginnings of an ABI column (focus of WG session)

Does the forum want to pursue this?

# Path Forward

- MPI ABI standard, separate from MPI API standard, but under the umbrella of the MPI Forum
- Dynamic link is the main objective (binary compatibility)
- Startup (mpirun/mpiexec) is a secondary objective
- Include all major language bindings in the ABI standard (C, C++, Fortran) as we go
- Tied to MPI 2.1 and forward
- Implementers may choose to comply (or not)
- May be referenced in procurements
- May be implemented as a morph layer or native
- Initial reference implementation will be a morph layer (tbd who provides)
- May be different per platform (Linux, Windows, ...)
- Different voting rules separate from MPI API forum (?)
- Need participation from major implementers in order to make progress

# Focus groups

mpi.h: MPICH 2 (1.0.7) as base with mods as needed (e.g. max values for compile time values used in declarations)

Proceed with a reference morph layer implementation targeting (at least) MPICH 2.0, OpenMPI

Load a different module and it just works

Demonstrate with a ring program

Defer startup consideration

- Linux (Jeff, Alexander) on X86-64 cluster Red Hat Linux/SUSE
- Windows (Alexander, Erez) – on X86-64 cluster
- Bring back to the December forum meeting

# Technical tasks

- Common mpi.h (spreadsheet)