

The Message Passing Interface: Version 3.0 and What Comes Next?

November 20th, 2013, Denver, CO

Martin Schulz, LLNL

Abstract

Last year, the MPI forum released MPI 3.0 with large enhancements like nonblocking and neighborhood collectives, enhanced RMA, and a tools information interface. However, this milestone is not the end of the standardization process: the forum continues to discuss new additions to the standard, such as support for fault tolerance and improved support for hybrid models. We will use this BoF to discuss the changes in MPI 3.0 and to continue an active discussion with HPC community on priorities for MPI 3.1 / 4.0 in an effort to make the forum more approachable for a wider community.

Structure of the BoF Session

The session started with a quick overview of features on MPI 3.0 (presented by Martin Schulz, LLNL). This talk included an overview of how far the major MPI implementations have progressed in implementing MPI 3.0. This corresponding table is included below. Following this overview, representatives of the three major open source implementations presented more details on the progress in MPICH (Rajeev Thakur, ANL), MVAPICH (Raghu Raja Chandrasekar, OSU) and Open MPI (Nathan Hjelm, LANL) as well as shared some first success stories on using MPI 3.0 features. The session was completed by a presentation by Martin Schulz, LLNL covering the structure of the MPI forum, the standardization body for MPI, and how the community can participate in the standardization of MPI. This talk was then followed by an open Q&A session.

MPI 3.0 Features and Current Status of Implementations

The following table lists major features added in MPI 3.0 as well as the status of their implementation in most major MPI distributions. Release dates are estimates and are subject to change at any time. Empty cells indicate no *publicly announced* plan to implement/support that feature. Information as of 11/20/2013.

	MPICH	MVAPICH	Cray	Tianhe	Intel	IBM PE	IBM Platform	Open MPI	SGI	Fujitsu	MS
NB collectives	✓	✓	✓	✓	✓	Q1 '14	✓	✓	✓		
Neighborhood collectives	✓	✓	✓	✓	✓	Q1 '14	Q3 '14	Q4 '13 (in nightly snapshots)	Q1 '14		
RMA	✓	✓	✓	✓	✓	Q1 '14	Q3 '14	Q1 '14	Q4 '13		
Shared memory	✓	✓	✓	✓	✓	Q1 '14	Q2 '15	Q1 '14	Q4 '13		
Tools Interface	✓	Q1 '14	Q2 '14	Q1 '14	Q1 '14	Q1 '14	Q2 '15	✓	Q1 '14		
Non-collective comm. create	✓	✓	✓	✓	✓	Q1 '14	Q2 '15	Q4 '13 (in nightly snapshots)	✓		
Fo8 Bindings (MPI-3 + errata)	(Q1 '14)	(Q2 '14)	(Q2 '14)	(Q1 '14)	(Q1 '14)	(Q1 '14)	(Q3 '14)	(✓)	(Q1 '14)		
New Datatypes	✓	✓	✓	✓	✓	Q1 '14	Q2 '15	✓	✓		
Large Counts	✓	✓	✓	✓	✓	Q1 '14	Q2 '15	✓	Q1 '14		
Matched Probe	✓	✓	✓	✓	✓	Q1 '14	Q3 '14	✓	✓		

[Information collected by Pavan Balaji, ANL]

Updates on Open Source MPI Implementations

MPICH – presented by Rajeev Thakur, ANL

The latest release of MPICH, which was made available as MPICH 3.1 shortly after SC13, includes all major features of MPI 3.0, aside from the recently updated Fortran 2008 bindings. The latter will be updated in Q1/14 and available in the next general release, MPICH 3.2 for SC14 along with other improvements such as improved support for accelerators as well as enhanced MPI-PGAS interoperability.

MVAPICH – presented by Raghu Raja Chandrasekar, OSU

The latest MVAPICH releases are closely coupled with the MPICH releases and MVAPICH therefore also offers most of the MPI 3.0 features in its latest release. In his talk Raghu specially highlighted performance improvements that were possible through the enhanced RMA model in MPI 3.0 as well as the wide variety of control and performance variables available through the new implementation of the MPI Tool Information interface (MPI_T).

Open MPI – presented by Nathan Hjelm, LANL

As the two other two open source MPI implementations, also Open MPI has made large progress towards fully supporting MPI 3.0 and is already now making most features available either in its latest release or through nightly snapshots. Open MPI also already supports the reworked version of the Fortran 2008 bindings. The remaining features are expected to be available in Q1/14.

Invitation to Join the MPI Forum

After the ratification of MPI 3.0 in September 2012, the MPI Forum decided to continue to meet and work towards future versions of the MPI standard. This includes minor corrections in the form of errata items, minor additions of functionality, as well as the addition of major new concepts. For the latter, the forum currently focuses on concepts for fault tolerance and hybrid programming models.

The MPI Forum itself has an open membership and interested parties are welcome to join and to participate in the forum. This can reach from submitting comments to the respective mailing lists, to participation in the virtual meetings of the various working groups, and all the way to physical attendance of MPI forum meetings. The latter are held four times per year – typically three times in the US (alternating between San Jose, CA and Chicago, IL) and one time co-located with EuroMPI outside the US.

Additional Information / Slides

The slides for this BoF will be posted on the MPI 4.0 Section of the MPI Forum pages at:

<http://meetings.mpi-forum.org/>

These pages also include a list of all working groups and how to participate in them, as well as information on the voting rules and process for additions to MPI.