



LEVERAGING INTEL MPI FOR BEST PERFORMANCE ON INTEL[®] XEON PHI[™] AND INTEL[®] OMNIPATH HARDWARE

IXPUG'16

Dmitry Durnov, Intel MPI team

Agenda

- Intel MPI 2017 Beta U1 product availability
- New features overview
- Competitive results
- Useful links
- Q/A

Intel MPI 2017 Beta U1 is available!

Key features:

- Topology aware SHM collectives
- Intel® Xeon® processor E5-2600 v4 product family + Intel® Omni-Path Fabric tuning
- Intel® Xeon Phi™ Processor codenamed Knights Landing (KNL) tuning (node level)
- Memory binding management features
- Asynchronous progress control
- Enhanced OpenFabrics Interfaces (OFI) support
- Process deployment enhancements
- Intel MPI benchmark improvements

Intel MPI 2017 Beta U1 is available!

Join Intel® Parallel Studio XE 2017 Beta program:



<https://software.intel.com/en-us/articles/intel-parallel-studio-xe-2017-beta>

The beta program officially ends **June 28th, 2016**.

The beta license provided will expire **October 7th, 2016**.

Topology aware SHM collectives

Allow to get a very low collective operation latency

Available for the following collective operations:

- MPI_Barrier
- MPI_Bcast
- MPI_Reduce
- MPI_Allreduce

Topology aware SHM collectives

Implemented as a set of new collective operations and available via

I_MPI_ADJUST family control:

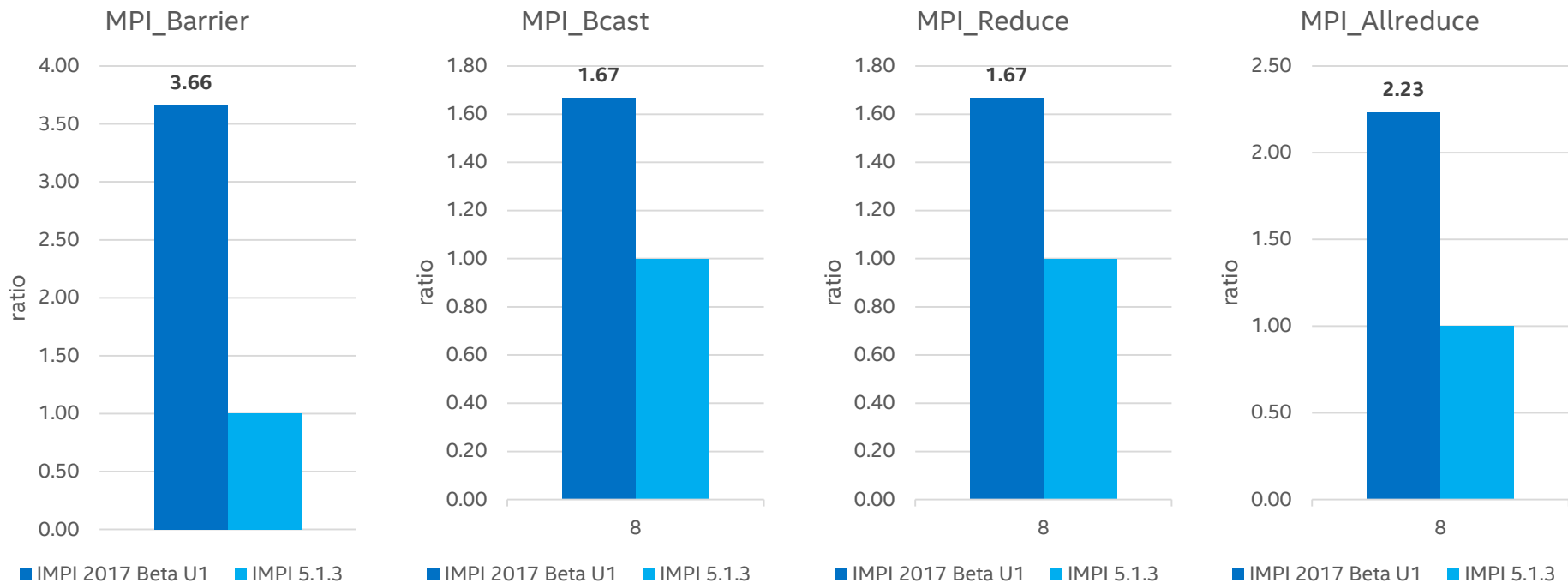
I_MPI_ADJUST_BARRIER=<7|8|9>

I_MPI_ADJUST_BCAST=<9|10|11>

I_MPI_ADJUST_REDUCE=<8|9|10>

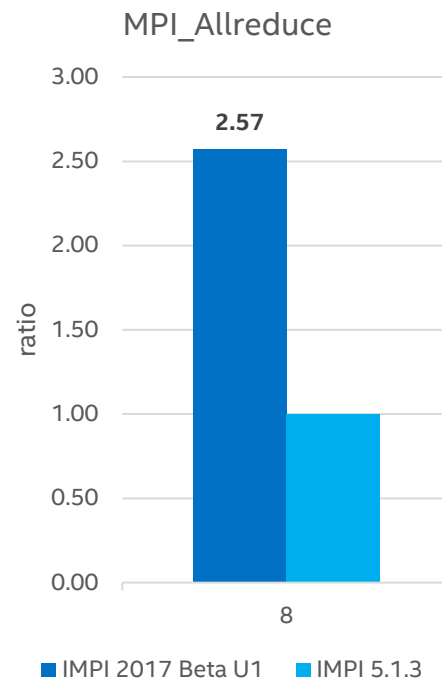
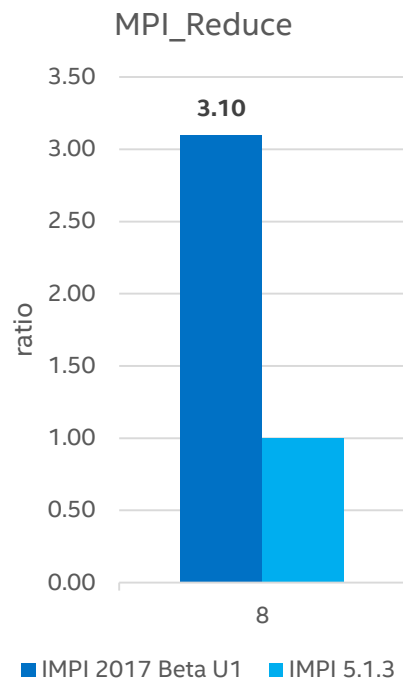
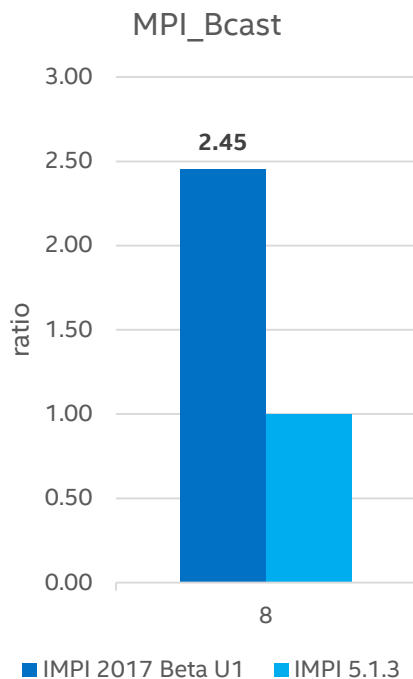
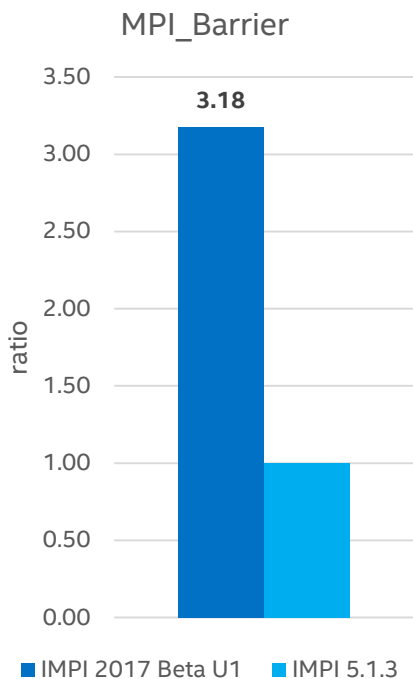
I_MPI_ADJUST_ALLREDUCE=<10|11|12>

Topology aware SHM collectives. Xeon. Intranode



Note: IMB-MPI1 4.1.1. N1P44. Intel® Xeon® E5-2699 v4 @ 2.20GHz. Higher is better

Topology aware SHM collectives. Xeon Phi. Intranode



Note: IMB-MPI1 4.1.1. N1P64. Intel® Xeon Phi (KNL). Higher is better
Results were obtained with pre-release HW. Final results may vary.

Memory binding management feature

- Provides user friendly interface for memory allocation control
 - General NUMA awareness
 - HBM/MCDRAM awareness (Xeon Phi specific)
- Available via the following env variables:
 - I_MPI_BIND_NUMA, I_MPI_BIND_ORDER
 - I_MPI_BIND_WIN_ALLOCATE
 - I_MPI_HBW_POLICY
- Fine grain control for MPI_Win_allocate_shared via MPI_Info mechanism

Memory binding management feature.

I_MPI_HBW_POLICY example.

There are 3 kinds of MPI process memory we can control:

- Application buffers
- Internal MPI buffers
- Application buffers allocated for MPI_Win_allocate_shared/MPI_Win_allocate

I_MPI_HBW_POLICY=<user buffers policy>[, [mpi buffers policy]][, win_allocate policy]]

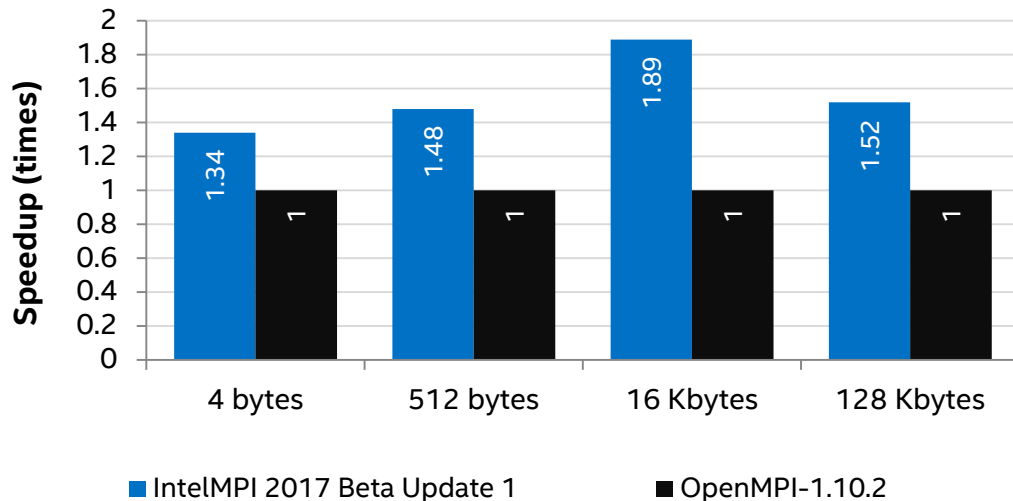
The following values are available:

Value	Note
hbw_preferred	Try to allocate MCDRAM first. If not available – allocate DDR.
hbw_bind	Try to allocate MCDRAM. If not available – fail.
hbw_interleave	MCDRAM/DDR interleaved allocation

Intel® Xeon® processor E5-2600 v4 product family + Intel® Omni-Path Fabric tuning

Superior Performance with Intel® MPI Library 2017 Beta U1

2304 Processes, 64 nodes (Omni-Path), Linux* 64
Relative (Geomean w/o vector ops) MPI Latency Benchmarks
(Higher is Better)



Configuration Info:

Hardware: CPU: Intel® Xeon E5-2697 v4 @ 2.30GHz; 128 GB RAM. Interconnect: Intel® Corporation Omni-Path HFI Silicon 100 Series [discrete] (rev 10)

Software: RHEL* 6.7; IFS 10.0.1.0.50; Libfabric 1.3.0; Intel® MPI Library 2017 Beta Update 1; Intel® MPI Benchmarks 4.1.1 (built with Intel® C++ Compiler XE 17.0.0 Beta for Linux*);

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. * Other brands and names are the property of their respective owners. Benchmark Source: Intel Corporation.

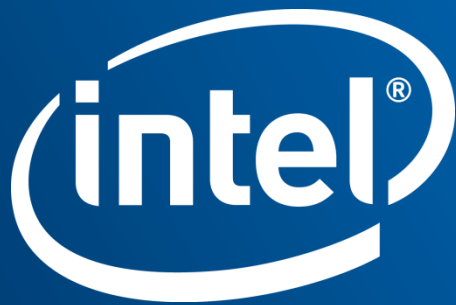
Optimization Notice: Intel's compilers may or may not optimize to the same degree for non-Intel microprocessors for optimizations that are not unique to Intel microprocessors. These optimizations include SSE2, SSE3, and SSSE3 instruction sets and other optimizations. Intel does not guarantee the availability, functionality, or effectiveness of any optimization on microprocessors not manufactured by Intel. Microprocessor-dependent optimizations in this product are intended for use with Intel microprocessors. Certain optimizations not specific to Intel microarchitecture are reserved for Intel microprocessors. Please refer to the applicable product User and Reference Guides for more information regarding the specific instruction sets covered by this notice. Notice revision #20110804

Links/Contacts

<https://software.intel.com/en-us/intel-mpi-library>

<https://software.intel.com/en-us/articles/intel-parallel-studio-xe-2017-beta>

mail: dmitry.durnov@intel.com



Legal Disclaimer & Optimization Notice

INFORMATION IN THIS DOCUMENT IS PROVIDED "AS IS". NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. INTEL ASSUMES NO LIABILITY WHATSOEVER AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO THIS INFORMATION INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT.

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.

Copyright © 2015, Intel Corporation. All rights reserved. Intel, Pentium, Xeon, Xeon Phi, Core, VTune, Cilk, and the Intel logo are trademarks of Intel Corporation in the U.S. and other countries.

Optimization Notice

Intel's compilers may or may not optimize to the same degree for non-Intel microprocessors for optimizations that are not unique to Intel microprocessors. These optimizations include SSE2, SSE3, and SSSE3 instruction sets and other optimizations. Intel does not guarantee the availability, functionality, or effectiveness of any optimization on microprocessors not manufactured by Intel. Microprocessor-dependent optimizations in this product are intended for use with Intel microprocessors. Certain optimizations not specific to Intel microarchitecture are reserved for Intel microprocessors. Please refer to the applicable product User and Reference Guides for more information regarding the specific instruction sets covered by this notice.

Notice revision #20110804