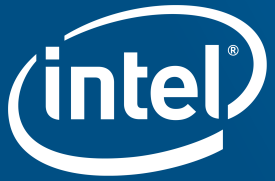


OpenSHMEM Birds of a Feather

November 15, 2017

Legal Notices and Disclaimers

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Learn more at intel.com, or from the OEM or retailer.

No computer system can be absolutely secure.

Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance. Consult other sources of information to evaluate performance as you consider your purchase. For more complete information about performance and benchmark results, visit http://www.intel.com/performance.

Intel, the Intel logo, Xeon and Xeon Phi and others are trademarks of Intel Corporation in the U.S. and/or other countries. *Other names and brands may be claimed as the property of others.

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. For more complete information visit www.intel.com/benchmarks.

© 2017 Intel Corporation.

OpenSHMEM Involvement

Actively engaged in OpenSHMEM specification committee

- Lead effort to specify new communication management API (contexts)

Engaged in open source development targeting the Open Fabrics Interface

- Sandia OpenSHMEM (SOS) v1.3.4 (released Nov. 7)
  - Full support for OpenSHMEM 1.3 specification on OFI libfabric and Portals 4
  - Bugfix release with performance and stability improvements
- Sandia OpenSHMEM (SOS) v1.4.0a1 (released Nov. 9)
  - Full support for OpenSHMEM 1.4 draft specification
    - Sockets provider fully supported
    - PSM2 provider supported, SHMEM_THREAD_MULTIPLE support is work-in-progress
  - Performance tuning and support for additional providers is work-in-progress
- Available at: https://github.com/Sandia-OpenSHMEM/SOS

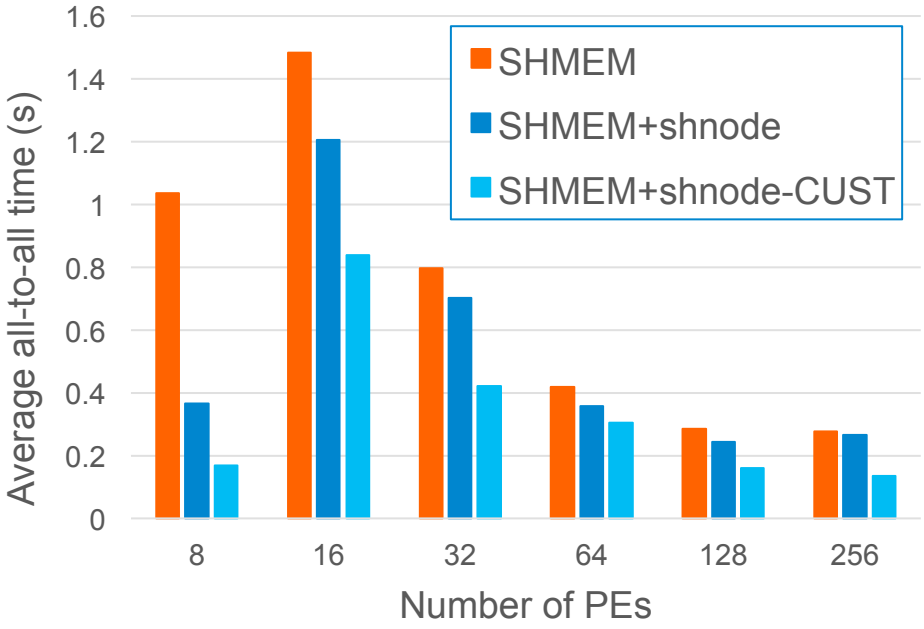
Exploring Locality/Topology Exposure in OpenSHMEM

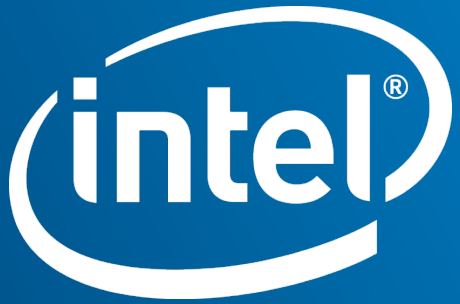
ISx strong scaling experiment

- Cori: Cray XC40, Cray SHMEM v7.5.5
- 4 nodes with 8 – 256 total PEs
- 1.5 billion data set

Application-level topology optimizations:

1. *SHMEM*: Baseline
2. *SHMEM+shnode*: On-node transfers use memcpy (communication avoidance)
3. *SHMEM+shnode-CUST*: Issue on-node transfers last to overlap with remote transfers (locality-aware overlap)




intel®

LIBFABRIC SCALABILITY

CRAY XC40

SHMEM

Evaluate libfabric SHMEM performance on high-performance interconnect

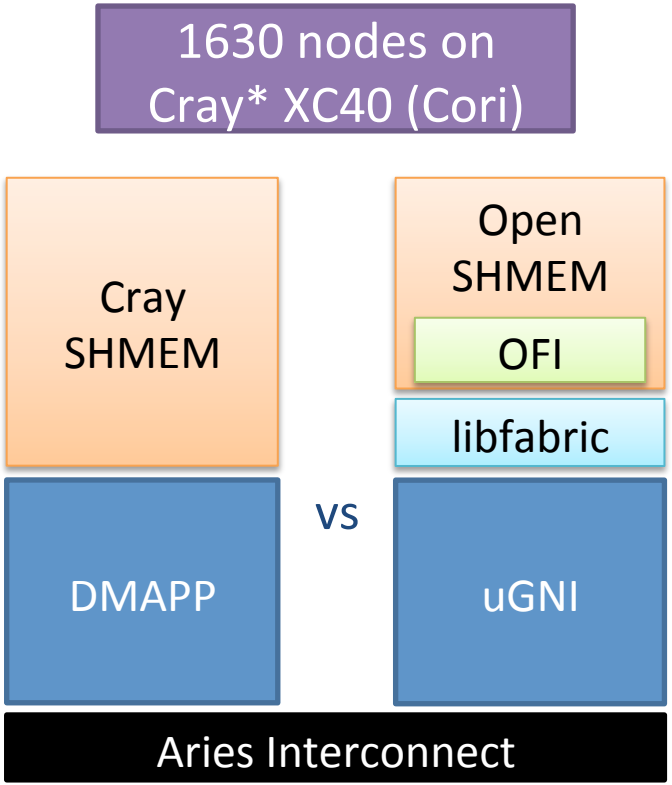
Provider implementation that uses the Cray* uGNI hardware and network interface for communication

* Other names and brands may be claimed as the property of others

LIBFABRIC SCALABILITY

CRAY XC40

SHMEM



- **Cray* SHMEM**
  - Cray* Aries, Dragonfly* topology
  - CLE (Cray* Linux*), SLURM*
  - DMAPP
    - Designed for PGAS
    - Optimized for small messages
- **Sandia* OpenSHMEM / libfabric**
  - uGNI
    - Designed for MPI and PGAS
    - Optimized for large messages
- https://www.nersc.gov/users/computational-systems/cori/configuration

* Other names and brands may be claimed as the property of others

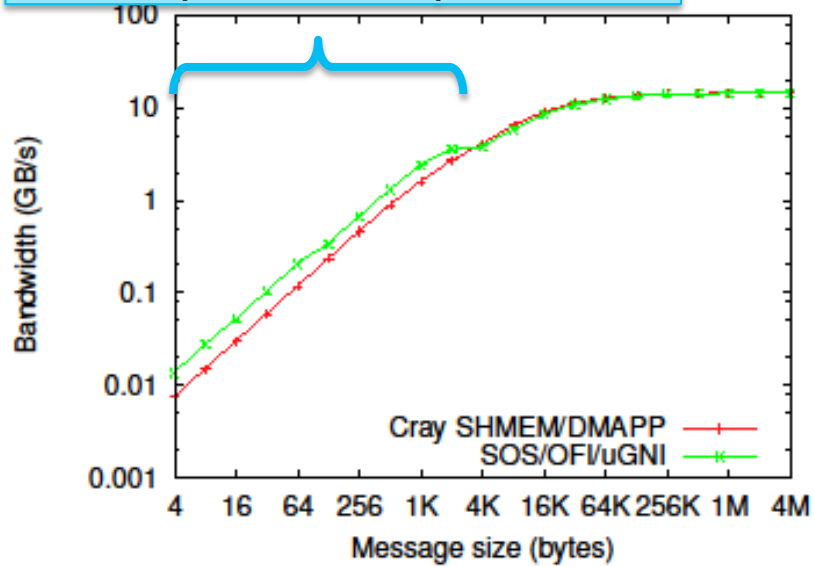
LIBFABRIC SCALABILITY

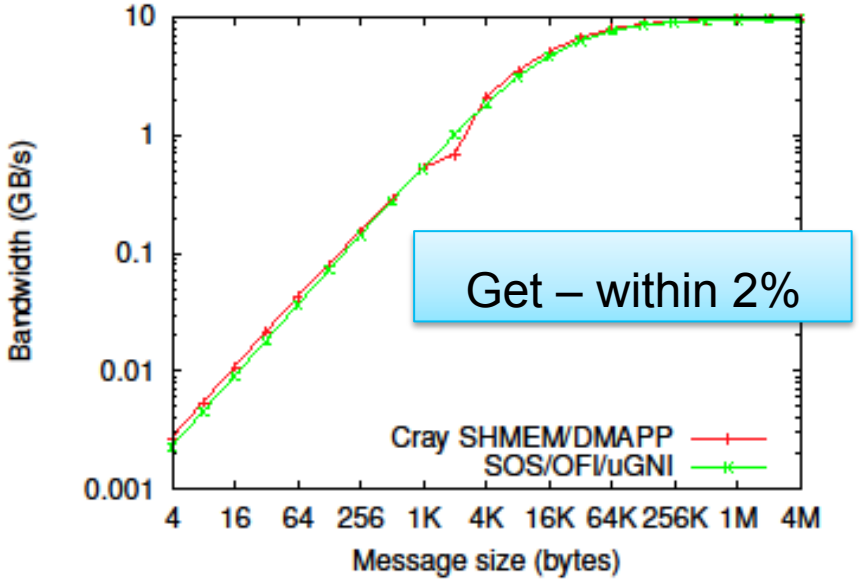
CRAY XC40

SHMEM

Blocking Get/Put B/W

Put – up to 61% improvement





Get – within 2%

Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance. Consult other sources of information to evaluate performance as you consider your purchase. For more complete information about performance and benchmark results, visit http://www.intel.com/performance.

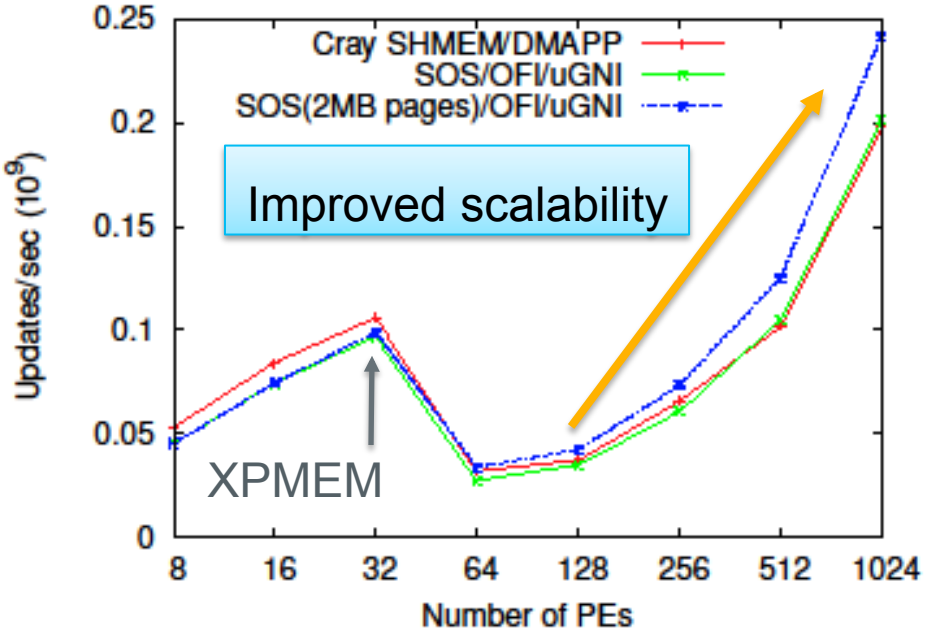
* Other names and brands may be claimed as the property of others

LIBFABRIC SCALABILITY

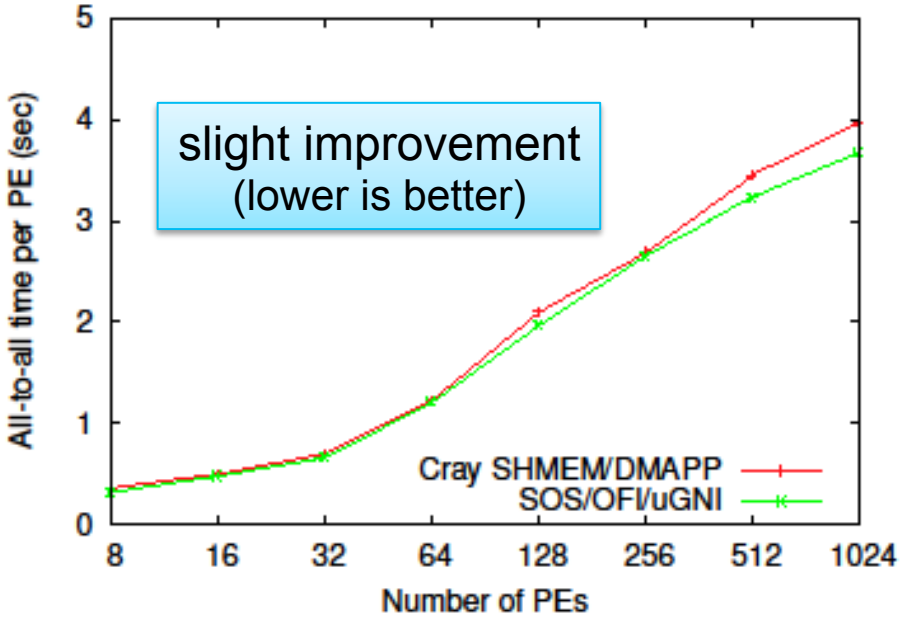
CRAY XC40

SHMEM

GUPS Scaling



NAS ISx (Integer Sort) weak scaling



Tests document performance of components on a particular test, in specific systems. Differences in hardware, software, or configuration will affect actual performance. Consult other sources of information to evaluate performance as you consider your purchase. For more complete information about performance and benchmark results, visit http://www.intel.com/performance.

* Other names and brands may be claimed as the property of others