Title

Application Performance on Intel Xeon Phi - Being Prepared for KNL and Beyond

ISBN

978-3-319-46078-9

Authors

Gerber, Richard A

Milfeld, Kent

Newburn, Chris J

et al.

Publication Date

2016

Peer reviewed

Application Performance on Intel Xeon Phi – Being Prepared for KNL and Beyond

Richard A. Gerber[1], Kent Milfeld[2], Chris J. Newburn[3], and Thomas Steinke[4]

[1] National Energy Research Scientific Computing Center
at Lawrence Berkeley National Lab. (NERSC)
ragerber@lbl.gov
[2] Texas Advanced Computing Center (TACC)
milfeld@tacc.utexas.edu
[3] Intel Corporation
chris.newburn@intel.com
[4] Zuse Institute Berlin (ZIB)
steinke@zib.de

The ISC16 workshop "Application Performance on Intel Xeon Phi – Being Prepared for KNL and Beyond" brought together about 100 members of a world-wide community of application developers and technology experts working to prepare scientific application codes to run at high performance on high performance computing systems powered by the Intel Xeon Phi processor. Intel released details of the Xeon Phi processor codenamed "Knight's Landing" or "KNL" at ISC16 and this workshop featured the first public KNL application performance results, delivered through a series of peer-reviewed presentations, lightening talks, and keynotes from Intel's Chief Architect for KNL, Avinash Sodani, and John McCaplin from the Texas Advanced Computing Center (TACC). The first KNL-based supercomputers are arriving this year at the National Energy Research Scientific Computing Center at Lawrence Berkeley National Laboratory, Argonne National Laboratory, TACC, and Los Alamos National Laboratory. Smaller KNL-based systems will be available this year at other places, e.g. a Cray KNL evaluation system at the Zuse Institute Berlin.

The Intel Xeon Phi Users Group (IXPUG) organized the workshop, which followed a similar event at ISC15 and built upon recent longer workshops in Berkeley, CA; Ostrava, Czech Republic, and St. Petersburg, Russia and Birds of a Feather sessions at ISC15 and SC15. The next IXPUG event, IXPUG 2016, is scheduled to be held at Argonne National Laboratory outside Chicago in September 2016. IXPUG is an independent organization working to build an international community to share challenges, experiences and best-practice methods for the optimization of HPC workloads on the Intel Xeon Phi. IXPUG workshops cover topics in application performance and scalability challenges at all levels - from single processor, to moderately-scaled cluster, up to large HPC configurations with many Xeon Phi devices. IXPUG also provides an effective conduit for application developers to interact directly with Intel engineers and other experts. Further information about IXPUG can be found at the IXPUG website (http://ixpug.org).

The Xeon Phi has a number of architectural features that provide opportunities for large gains in performance using the manycore, power efficient processing cores But taking advantage of these features can present a number of challenges for programmers used to coding for traditional Xeon-type processors. Many if not most existing codes have to be refactored to take advantage of the Xeon Phi's on-chip High Bandwidth Memory (HBM), longer 512-bit vector units, and up to 72 cores on a single socket. The papers in this workshop cover optimization and scalability topics in real-world HPC applications, e.g. data layouts and code restructuring for efficient SIMD operation, work distribution and thread management. Aspects related to KNL features have considerable weight in the studies. The versatility and value of tools for development, debugging and performance analysis is also covered. The keynotes presented recent information about the KNL processor, and trends in HPC system performance; and Lightning Talks provide late-breaking work and experiences on Intel Xeon Phi systems. Keynotes and Lightning Talks are available at https://www.ixpug.org.

Call for Papers

Papers presented at the workshop were selected from submissions solicited from the community through a call for submissions issued in March 2016. Submissions were reviewed by the program committee (listed below) and accepted papers are scheduled to be published by Springer in post-workshop ISC'16 Workshop Proceedings.

A summary of the call follows:

The IXPUG workshop is about sharing ideas, implementations, and experiences that will help users take advantage of new technologies such as AVX512 operations, high-bandwidth memory (HBM) and OmniPath. These architectural advances in Vectorization, Memory, and Communications on the Intel Xeon Phi platform will help boost adoption of many-core architecture in HPC as well as other computational spaces.

In the workshop you will experience an open forum with fellow application programmers, Intel Phi architecture designers, and compiler and tool experts. In addition to the technical paper presentations, the program will include a morning keynote on Intel microprocessors and an afternoon presentation on HPC performance trends. There will also be two Lightning Talks sessions and the workshop will conclude with a discussion.

IXPUG welcomes paper submissions on innovative work from KNC and KNL users in academia, industry and government labs, describing original discoveries and experiences that will promote and prescribe efficient use of many-core and multicore systems.

Topics of interest are (but not limited to):

- Vectorization: Data layout in cache for efficient SIMD operations, SIMD directives and operations, and 2-core tiling with 2D interconnected mesh

- Memory: Data layout in memory for efficient access (data preconditioning), access latency concerns (prefetch, streams, costs for HBM), partitioning of DDR and HBM for applications (memory policies)
- Communication: including early experiences with OmniPath
- Thread and Process Management: Process and thread affinity issues, SMT (simultaneous multi-threading, in core), balancing processes and threads
- Programming Models: OpenMP 4.x, hStreams, using MPI 3 on Xeon Phi, hybrid programming (MPI/OpenMP, others)
- Algorithms and Methods: including scalable and vectorizable algorithms
- Software Environments and Tools
- Benchmarking & Profiling Tools
- Visualization

Program Committee

Damian Alvarez-Mallon	Forschungszentrum Jülich
Ryan Coleman	Sandia National Lab
Douglas Doerfler	NERSC/Berkeley Lab
Richard Gerber	NERSC/Berkeley Lab
Antonio Gomez	TACC
Simon Hammond	Sandia National Lab
Rahul Hardikar	Indian Institute of Science
Helen He	NERSC/Berkeley Lab
Dave M Hiatt	big denominator
Michael Klemm	Intel Corp.
Lars Koesterke	TACC
Rakesh Krishnaiyer	Intel Corp.
Olli-Pekka Lehto	CSC - IT Center for Science Ltd.
John Linford	ParaTools, Inc.
Simon McIntosh-Smith	Bristol Univ.
John Michalakes	NREL
Kent Milfeld	TACC
Chris J Newburn	Intel Corp.
Dmitry Prohorov	Intel Corp.
Karthik Raman	Intel Corp.
Carlos Rosales	TACC
Hideki Saito	Intel Corp.
Abhinav Sarje	Berkeley Lab
Thomas Steinke	Zuse Institute Berlin
Estela Suarez	Forschungszentrum Jülich
Srinath Vadlamani	Paratools, Inc.
Jerome Vienne	TACC

Workshop Organizers

Richard A. Gerber	National Energy Research Scientific Computing Center at Lawrence Berkeley National Lab. (NERSC)
Kent Milfeld	Texas Advanced Computing Center (TACC)
Chris J. Newburn	Intel Corporation
Thomas Steinke	Zuse Institute Berlin (ZIB)