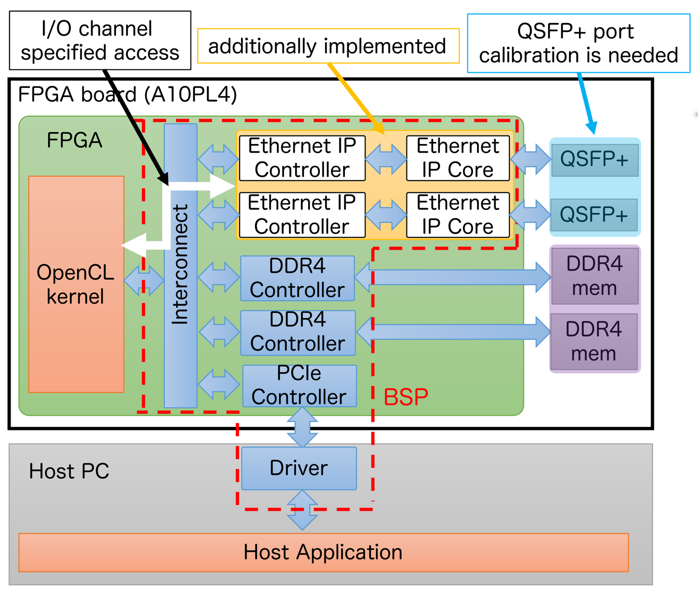
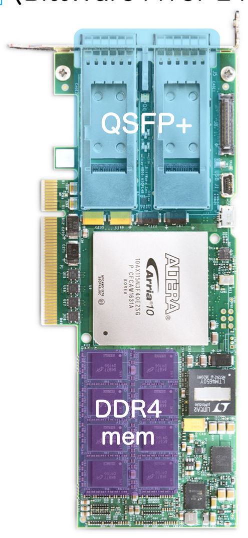
Numerical Computation

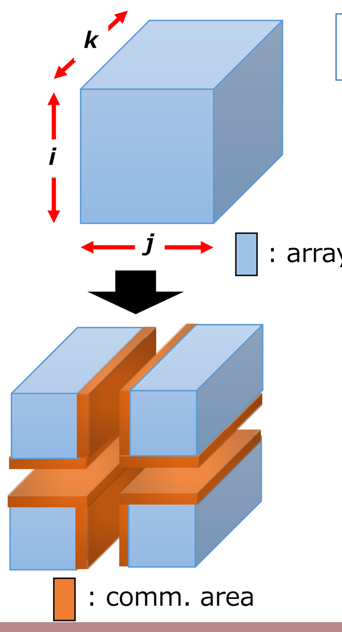
OpenCL-ready High Speed FPGA Networking [1]

- Dev. tool: Intel FPGA for OpneCL
- Board Support Package (BSP) is a hardware component to support multiple different boards
  - Which FPGA chip is used on the board
  - What kind of peripherals are support by the board
- Basically, only minimum interfaces are supported
  - To perform inter FPGA communication, implementing network controller and integrating it into the BSP are required

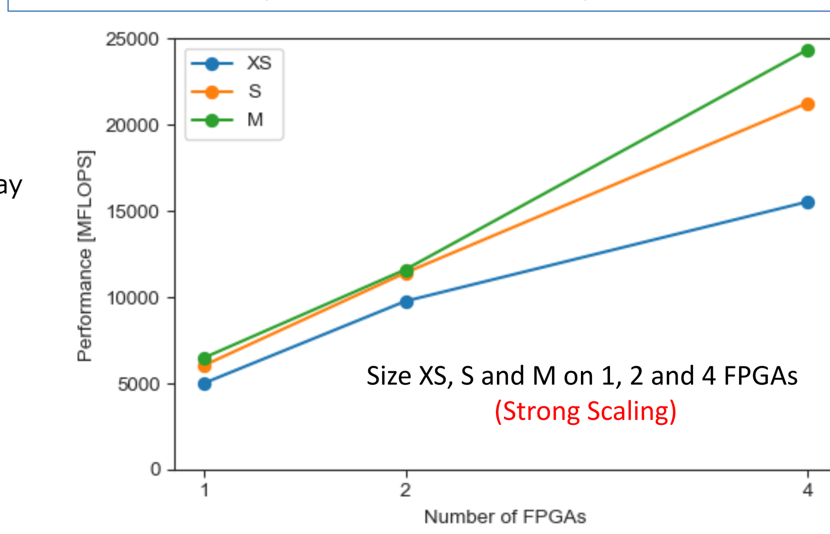


Our FPGA board (BittWare A10PL4)



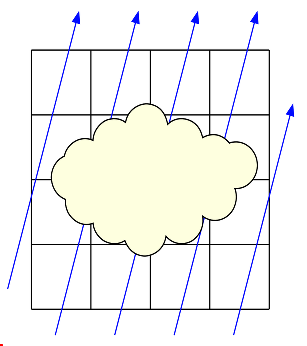


3D-Poisson Equation Solver, 19-point stencil computation

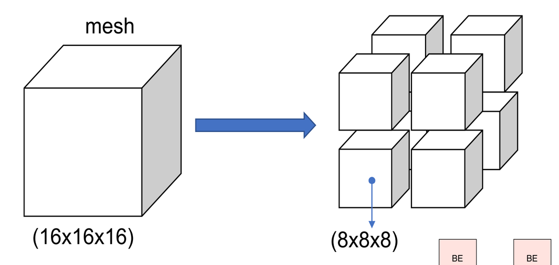


Authentic Radiation Transfer (ART) on FPGA [2]

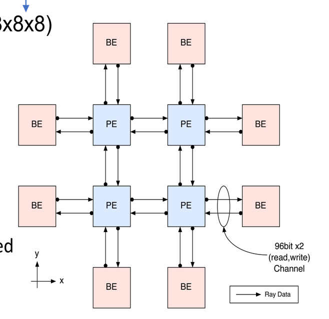
- Accelerated Radiative transfer on grids Oct-Tree (ARGOT) has been developer in Center for Computational Sciences, University of Tsukuba
  - ART is one of algorithms used in ARGOT and dominant part (90% or more of computation time) of ARGOT program
- ART is ray tracing based algorithm
  - problem space is divided into meshes and reactions are computed on each mesh
  - Memory access pattern depends on ray direction
- Not suitable for SIMD architecture

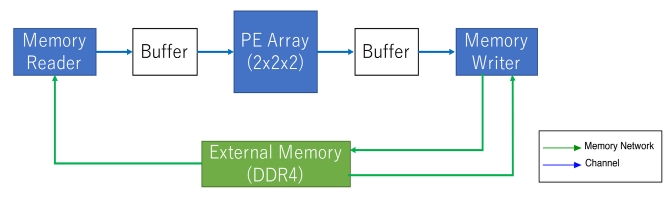


- Problem space is divided into small blocks
  - e.g. (16, 16, 16) → 8 × (8, 8, 8)
  - PE is assigned to each of small blocks

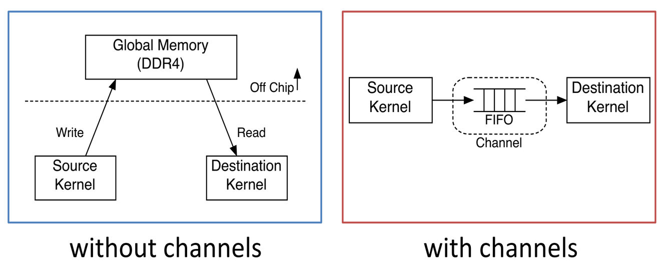


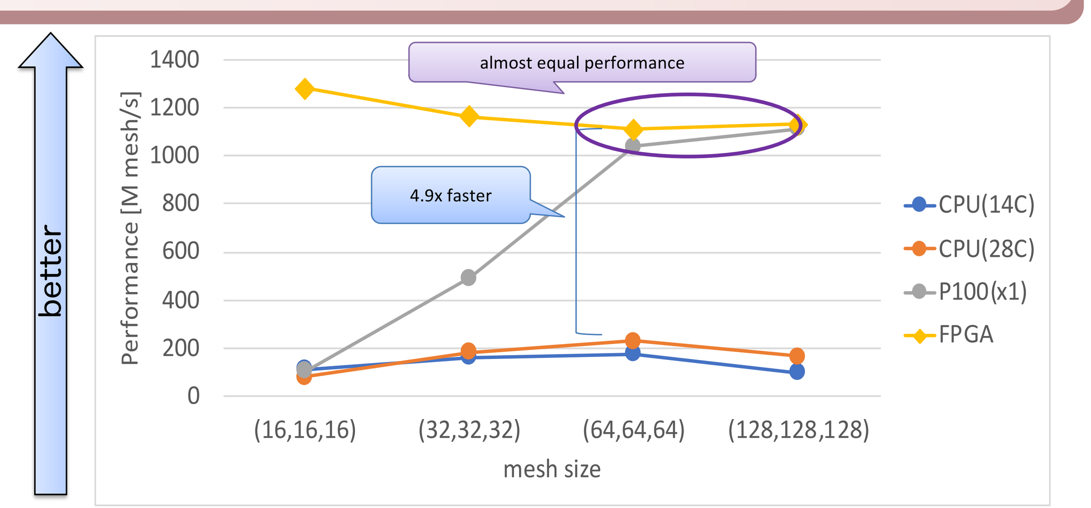
- PEs are connected by channels each other
  - PE: Processing Element
  - BE: Boundary Element
- Kernel of PEs and BEs are started automatically by autorun attribute
  - Lower control overhead and resource usage because of decreasing number of host controlled kernels





- Our implementation uses channel based approach
  - One of extensions to OpenCL for FPGAs by Intel
- It enables inter kernel communication much faster
  - No external memory (DDR) access is required
  - Lower resource utilization than DDR access





Size	CPU(14C)	CPU(28C)	P100	FPGA
(16,16,16)	112.4	77.2	105.3	1282.8
(32,32,32)	158.9	183.4	490.4	1165.2
(64,64,64)	175.0	227.2	1041.4	1111.0
(128,128,128)	95.4	165.0	1116.1	1133.5

Reference

[1] Ryohei Kobayashi, Yuma Oobata, Norihisa Fujita, Yoshiki Yamaguchi, and Taisuke Boku, OpenCL-ready High Speed FPGA Network for Reconfigurable High Performance Computing, HPC Asia 2018, pp.192-201, January 2018

[2] Norihisa Fujita, Ryohei Kobayashi, Yoshiki Yamaguchi, Yuuma Oobata, Taisuke Boku, Makito Abe, Kohji Yoshikawa, and Masayuki Umemura: Accelerating Space Radiate Transfer on FPGA using OpenCL (Accepted), International Symposium on Highly-Efficient Accelerators and Reconfigurable Technologies (HEART 2018)

Acknowledgment

This research is a part of the project titled "Development of Computing-Communication Unified Supercomputer in Next Generation" under the program of "Research and Development for Next-Generation Supercomputing Technology" by MEXT. We thank Intel University Program for providing us both of hardware and software.

Automatic Tuning of Parallel 1-D FFT on Cluster of Intel Xeon Phi Processors

Background

The fast Fourier transform (FFT) is widely used in science and engineering. Parallel FFTs on distributed-memory parallel computers require intensive all-to-all communication, which affects their performance. How to overlap the computation and the all-to-all communication is an issue that needs to be addressed for parallel FFTs.

Overview

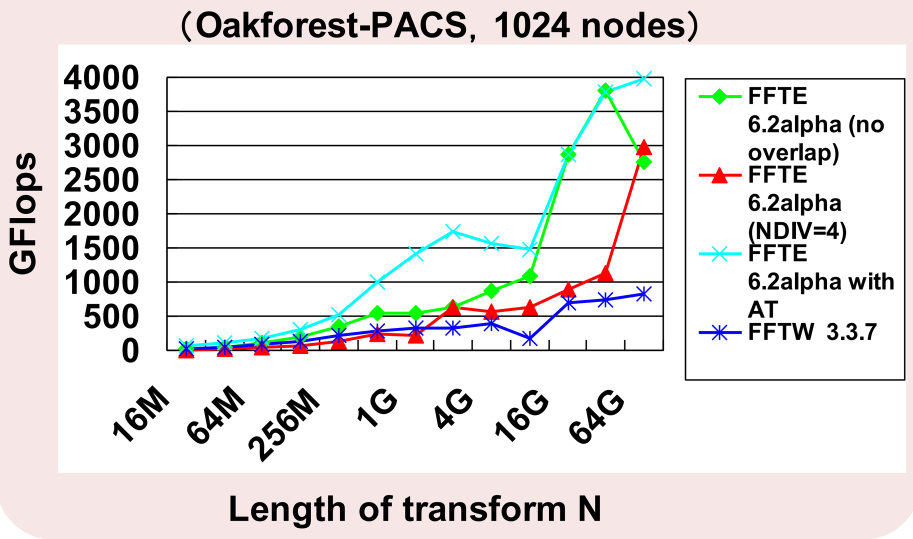
We proposed an automatic tuning (AT) of computation-communication overlap for parallel 1-D FFT. We used a computation-communication overlap method that introduces a communication thread with OpenMP. An automatic tuning facility for selecting the optimal parameters of the computation-communication overlap, the radices, and the block size was implemented.

Performance

To evaluate the parallel 1-D FFT with AT, we compared its performance against those of FFTE 6.2alpha (http://www.ffte.jp/) , FFTE 6.2alpha with AT, and FFTW 3.3.7.

The performance results demonstrate that the proposed implementation of a parallel 1-D FFT with automatic tuning is efficient for improving the performance on cluster of Intel Xeon Phi processors.

Performance of parallel 1-D FFTs (Oakforest-PACS, 1024 nodes)

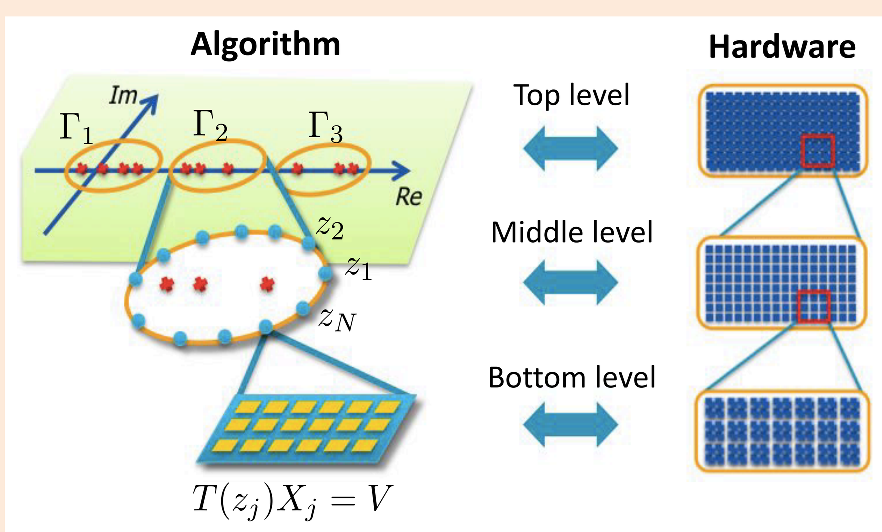


Development of Parallel Sparse Eigensolver Package: z - Pares

The aim of this research project is to develop numerical software for large- scale eigenvalue problems for post-petascale computing environment. An eigensolver based on contour integral (the SS method) has been proposed by Sakurai and Sugiura [3]. This method has a hierarchical structure and is suitable for massively parallel supercomputers [2]. Moreover, the SS method can be applicable for nonlinear eigenvalue problem [1]. Block Krylov method [4] improves the performance of the method. Based on these newly designed algorithms, we have developed a massively parallel software z-Pares freely-available from http://zpares.cs.tsukuba.ac.jp/. MATLAB version is also available in our webpage. We have also developed CISS eigensolver in SLEPc.

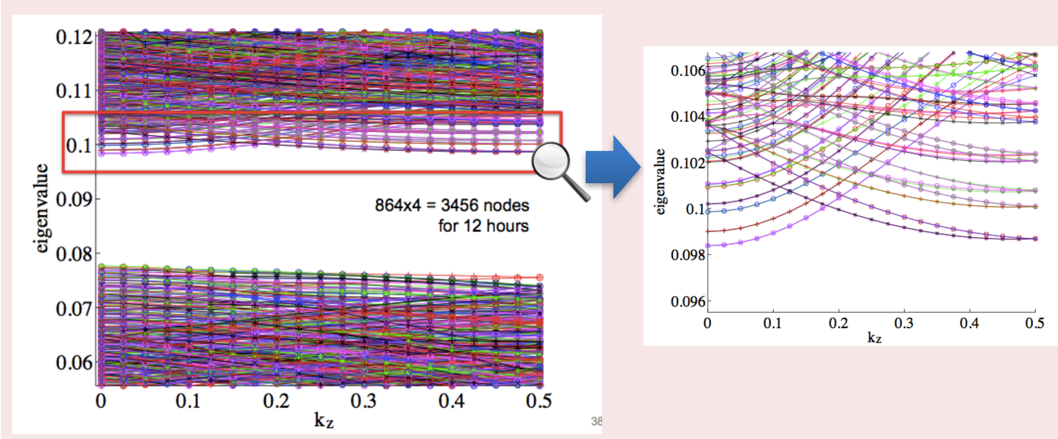
Hierarchical Parallel Structure

Hardware is grouped according to a hierarchical structure of the algorithm.



Numerical Example on the K Computer

Application for band calculation with real space density functional theory (RSDFT) [2].



Band structure of silicon nanowire of 9,924 atoms. (matrix size = 8,719,488, Number of cores = 6,144)

*The results are tentative since they are obtained by early access to the K computer.

Reference

[1] J. Asakura, T. Sakurai, H. Tadano, T. Ikegami and K. Kimura, A numerical method for nonlinear eigenvalue problems using contour integrals, JSIAM Letters, 1 (2009) 52-55.

[2] Y. Futamura, T. Sakurai, S. Furuya and J.-I. Iwata, Efficient algorithm for linear systems arising in solutions of eigenproblems and its application to electronic-structure calculations, Proc. 10th International Meeting on High-Performance Computing for Computational Science (VECPAR 2012), 7851 (2013), 226-235.

[3] T. Sakurai and H. Sugiura, A projection method for generalized eigenvalue problems, J. Comput. Appl. Math., 159 (2003) 119-128.

[4] H. Tadano, T. Sakurai and Y. Kuramashi, Block BiCGGR: A new block Krylov subspace method for computing high accuracy solutions, JSIAM Letters, 1 (2009) 44-47.

Acknowledgment

This research was partially supported by the Core Research for Evolutional Science and Technology (CREST) Project of JST and the Grant-in-Aid for Scientific Research of Ministry of Education, Culture, Sports, Science and Technology, Japan, Grant number: 23105702.

contact address: pr@ccs.tsukuba.ac.jp