Electron Dynamics Simulation with Time-Dependent Density Functional Theory on Large Scale Many-Core Systems

Yuta Hirokawa
Graduate School of Systems and Information Engineering, University of Tsukuba, Japan
hirokawa@hpcs.cs.tsukuba.ac.jp

1. POSTER SUMMARY

Many-core processors such as the Intel Xeon Phi (Knights Corner, KNC) and GPUs have been providing a new solution for HPC with high performance/power ratio. However, it is difficult to exploit the high sustained performance due to the different characteristics of matured traditional processors such as Intel Xeon CPU. In this study, I implement an electron dynamics simulation called ARTED [6] [1] [5] as real application code for material science to three types of many-core processors: KNC and KNL (Knights Landing) as Intel MIC, and Fujitsu SPARC64 XIfx for Fujitsu FX100. Fujitsu SPARC64 VIIIfx on K Computer is also evaluated for reference.

Three supercomputers are examined in this research. COMA at University of Tsukuba with KNC [2], Fujitsu FX100 at Nagoya University with SPARC64 XIfx processor [3] [7], and Oakforest-PACS at JCAHPC (Joint Center for Advanced HPC) with KNL [4]. The environment of these systems are shown in Table 3.

The target code ARTED currently focuses on two materials, Silicon and SiO_2, for the fundamental material simulation in quantum level. In the parallel computation shown in Table 1, the computation on Silicon implies large number of threads with small vector parallelism. Conversely, SiO_2 case implies small number of threads with large vector parallelism. The many-core processor strongly requires the large degree of parallelism with the thread- and SIMD vector-levels. I show the variance of performance with these different target of simulations on these processors.

While the main calculation contributes to the 3D stencil computation, we avoid to decompose the physical domain into parallel processes because the domain size is relatively small. Instead of it, a number of wave dimensions are parallelized by MPI. As a result, MPI_Allreduce for vector data is the dominant communication among processes. Fig. 3 shows the MPI_Allreduce latency with Silicon case on various systems and processors. KNC and KNL cause huge communication latency as compared with Xeon and SPARC64 CPUs. However, the communication overhead occupies just around 1–10 % of computation time and it can be neglected in Silicon case.

I optimized the computation for 25-points stencil with periodic boundary condition which dominates the total computation time in each single thread. Here, both the explicit and compiler vectorizations are applied. The explicit vectorization is coded with 256/512-bit SIMD instructions explicitly in C language with intrinsic operations. In the compiler vectorization, the indirect reference for neighboring points is avoided to optimize the memory reference sequence. For cyclic boundary condition, MOD operation is also replaced with table lookup since it takes long cycle. For the explicit vectorization, GATHER intrinsic under the unit-stride memory access causes large latency to be the performance bottleneck on KNC. The aligned LOAD and bit-shift instructions avoid this problem. In porting to the KNL, the compatibility of SIMD instructions between KNC and KNL is incomplete. However, the replacement can be automated with preprocessor at compile-time.

The number of OpenMP threads per KNC is set to 240 to use entire cores with SMT on 60 cores. KNL has 68 cores, however I use up to only 64 cores (256 threads with SMT) due to the configuration on a chip. As a result, the stencil computation performance achieves 591.2 GFLOPS on two KNC and 707.2 GFLOPS on a KNL for Silicon problem while 215.1 GFLOPS is achieved on a SPARC64 XIfx for SiO_2 as shown in Fig. 4. KNL achieves three times faster performance of the KNC for SiO_2 case, and gets 23.2 % of performance efficiency for Si case.

Fig. 5 shows the entire computation time for all systems. On KNC, the Symmetric mode executes the application with cooperative computation between the CPU and Xeon Phi using MPI to utilize entire resources in each node. This mode needs to balance the load among heterogeneous processors, and it is found that the best performance ratio between CPU and KNC is around 0.6–0.75:1.0. Consequently, the Symmetric mode achieves 1.97 times better performance compared with the CPU-only utilization at strong scaling. In SiO_2 case, FX100 CPU is faster than the KNC and Xeon, while the performance of Silicon case is comparable with Xeon. The performance of KNC are influenced with degree of thread parallelism especially. In conclusion, I confirmed the single node performance of KNL on this problem much exceeds that of KNC and SPARC processor on FX100.

Currently, the strong scaling performance is limited by the problem size. It is expected to improve the thread parallelism with employment of oblique lattice. This improvement increases the thread parallelism, while vector paral-

SC'16 November 13–18, 2016, Salt Lake City, Utah, USA

DOI: 10.1145/1235

lelism is decreased by physical symmetry of the material.

The future works imply:

1. Further optimization on KNC, KNL, and FX100 CPU to get close to their theoretical peak.
2. Implementation and performance comparison with GPU. (It is undergoing by collaboration with NVIDIA.)

2. REFERENCES

[1] ARTED. source code repository. https://github.com/ARTED/ARTED.

[2] Center for Computational Sciences, University of Tsukuba. http://www.ccs.tsukuba.ac.jp/eng/.

[3] Information Techonology Center, Nagoya University. http://www.icts.nagoya-u.ac.jp/en/center/.

[4] JCAHPC: Joint Center for Advanced HPC. http://jcahpc.jp/eng/index.html.

[5] M. Schultze, K. Ramasesha, C. D. Pemmaraju, S. A. Sato, D. Whitmore, A. Gandman, J. S. Prell, L. J. Borja, D. Prendergast, K. Yabana, D. M. Neumark, and S. R. Leone. Attosecond band-gap dynamics in silicon. *Science*, 346(6215):1348–1352, 2014.

[6] S. A. Sato and K. Yabana. Maxwell + TDDFT multi-scale simulation for laser-matter interactions. *J. Adv. Simulat. Sci. Eng.*, 1(1):98–110, 2014.

[7] T. Yoshida and *et al.* Sparc64 XIfx: Fujitsu's Next-Generation Processor for High-Performance Computing. In *HotChips 25*, 2014.