International Journal of Business, Human and Social Sciences
ISSN: 2517-9411
Vol:7, No:10, 2013

A Consideration of the Achievement of Productive Level Parallel Programming Skills

Tadayoshi Horita, Masakazu Akiba, Mina Terauchi, Tsuneo Kanno

***Abstract*—**This paper gives a consideration of the achievement of productive level parallel programming skills, based on the data of the graduation studies in the Polytechnic University of Japan. The data show that most students can achieve only parallel programming skills during the graduation study (about 600 to 700 hours), if the programming environment is limited to GPGPUs. However, the data also show that it is a very high level task that a student achieves productive level parallel programming skills during only the graduation study. In addition, it shows that the parallel programming environments for GPGPU, such as CUDA and OpenCL, may be more suitable for parallel computing education than other environments such as MPI on a cluster system and Cell.B.E. These results must be useful for the areas of not only software developments, but also hardware product developments using computer technologies.

***Keywords*—**Parallel computing, programming education, GPU, GPGPU, CUDA, OpenCL, MPI, Cell.B.E.

I. Introduction

PROGRAMMING skills of engineers are very important in the areas of not only software developments, but also hardware product developments using computer technologies such as embedded systems, car, ship or airplane form designs using flow simulations. Furthermore, a recent trend in computing hardware areas is the increasing of the number of computing cores in a chip, rather than the higher performance inside a single computing core. Owing to such trend, we can use non-expensive parallel computing environments such as GPU devices with their software development kits for GPGPUs, PC cluster systems which consist of non-expensive multicore CPUs, and so on. These environments provide their higher computing performances for the lower cost. However, to utilize their high performances sufficiently, programmers and/or engineers must achieve certain level programming skills of not only sequential processing on a single core but also parallel computing in these environments. In addition, the performance of the program is strongly related to such skills of the programmer. Therefore, the achievement of parallel programming skills in these areas is very important.

The authors consider that the education of parallel programming is very suitable for universities and/or even graduate schools, because such education concerns not only software programming but also parallel computing architectures, that is, its difficulty is more than high school level in common cases. The recent works dealing with the education of parallel programming in universities are, e.g., [1] and [2]. They introduce useful tools and/or curriculum to teach parallel programming to students. However, they give few data concerning the achievement of parallel programming skills, including the comparison between parallel programming environments such as CUDA and MPI in terms of it. In addition, the author can find no work which gives such data.

On the other hand, for some years, the authors have trained third and fourth grade students to achieve parallel programming skills as parts of their graduation studies in the Polytechnic University of Japan ("SYOKUGYO DAI" for short) [3]. Therefore in this paper, we give many data and a consideration concerning the achievement of productive level parallel programming skills, including the comparison between parallel programming environments, through such training experiences.

II. Overview of the Polytechnic University of Japan

The Polytechnic University of Japan, called "SYOKUGYO DAI" for short of Japan, was established by the Japanese government and is under the management of Japan Organization for Employment of the Elderly, Persons with Disabilities and Job Seekers. The university undertakes the tasks under the articles of the Human Resources Development Promotion Law. One of the tasks is the education and training of vocational instructors[1]. For this task, the university has the undergraduate course called "Chouki-katei". This is a four-year course for high school graduates or those having the equivalent education to become vocational training instructors or technical leaders. Most new students are 17 or 18 years old.

Upon graduation, students are conferred not only the bachelor's degree of engineering but also the license of vocational training instructor, which is essential license to work as the instructor of Japan. The course features many hours of experiments and practice, which are essential for vocational training instructors, and a requirement for subjects related to vocational capabilities development.

One of the departments of this course was the department of information systems engineering, which has already dismissed since Mar. 2012. In addition, the acceptance of new students of this course has already been stopped since Apr. 2011[2], however

Tadayoshi Horita, Masakazu Akiba, Mina Terauchi, and Tsuneo Kanno are with the Polytechnic University of Japan (SYOKUGYO DAI), 2-32-1, Ogawanishimachi, Kodaira-shi, Tokyo 187-0035 Japan (phone: +81-42-346-7681; fax: +81-42-344-5609; e-mail: horita@uitec.ac.jp).

[1] Other tasks of SYOKUGYO DAI are listed as follows: 1) Research and training to improve the quality of vocational instructors. 2) Advanced vocational training that provides standards for vocational training nationwide. 3) Surveys and research on the development and improvement of vocational capabilities and provision of various information. 4) Counseling and assistance on development and improvement of vocational capabilities. Note that SYOKUGYO DAI is the only one university of Japan which undertakes these tasks.

[2] Instead of the Chouki-katei, another four-year course called

in the following; this department of the course is detailed in this paper because this paper is strongly related to it.

The subjects of the normal type entrance examination of this course for Japanese new students are English I and II, mathematics I, II, III, A, B and C, and a selection from {physics I and II } and {chemistry I and II }[3].

III. Related Subjects and Environments Concerning the Graduation Study

The graduation study and the programming subjects in this department, which are strongly related to the graduation study, are listed in Table I. Note that except the subjects in Table I, students must take some another exercise type subjects which require programming skills.

The corresponding duration (hours) per unit of each subject in Table I is listed in Table II.

Concerning the type of subject, in exercise type subject, a modern PC or workstation is assigned to each student, and basically, each student must do programmings by oneself. Concerning the subjects of the fourth grade in Table I including the graduation study, not only a modern assigned PC and a desk but also electric files, including source code files which are denoted in Section IV, as the results of past years' graduation studies on the network file server, are available by each student.

Other factors which influence the quality of results of graduation study is denoted as follows: The timing that the student belonged to the authors' laboratory[4]. The case which the student changes the theme of the graduation study on the way[5]. The timing that a student achieves the unofficial job offer, which is the employment after the graduation[6]. The case which the student is governmental-sponsored foreign student[7].

TABLE I
Related Subjects

Subject name	Grade	Unit	Type
Data structure and algorithm	3rd	2	classroom
Computer programming I(C grammar)	1st	1.5	exercise
Computer programming II(Java grammar)	1st	1.5	exercise
Computer programming III(signal processing)	2nd	1.5	exercise
Computer programming IV(advanced programming technique)	2nd	1.5	exercise
Assembler programming	2nd	2	exercise
Practice on computer system design	3rd	2	exercise
Practice on system implementation I	4th	2	exercise
Practice on system implementation II	4th	2	exercise
Graduation study	4th	8	exercise

"Sougou-katei" was started on Apr. 2012.

[3] These symbols of subjects such as I, II, A, and B are categories of their subjects of Japanese high schools.

[4] Usually, the timing which each student is assigned to a laboratory is the beginning of the latter period of the third grade. However, the assignment of some students may change at the timing of the beginning of the fourth grade because of the special reasons, such as professors' personnel changes and so on.

[5] Evidently, the theme needs to be changed if the laboratory assignment of the student is changed.

[6] A student doesn't want to attend the university during having no job offer because of the anxiety. In addition, especially recent years, job hunting of Japan is very heavy task for students because the job offer situation is bad.

[7] According to the governmental policy, SYOKUGYO DAI does not permit governmental-sponsored foreign students to hunt job.

TABLE II
Corresponding Duration (Hours) per Unit

	type	duration(hour)
before Apr. 2008	classroom	10.5
before Apr. 2008	exercise	21.0
after Apr. 2008	classroom	12.5
after Apr. 2008	exercise	25.0

IV. Technical Terms and Parallel Computing Environments

The technical terms and parallel computing environments are listed in the following.

- PC cluster system: This system consists of multiple PCs which are connected with each other via high speed LANs. In this paper, in each PC, the Linux is installed as its OS, and the OpenMPI [4] is installed as its parallel computing library.
- KURO-BOX [5]: This is a NAS (network attached storage) barebone, which is produced by Kurouto-shikou Inc. A KURO-BOX has a PowerPC CPU, and a Gbit LAN NIC (Network interface controller).
- KURO-BOX cluster system: This system consists of multiple KURO-BOXes which are connected with each other via the Gbit LAN. The Linux and the OpenMPI library are installed in each KURO-BOX.
- Cell.B.E.[8]: This is the heterogeneous multicore CPU, which consists of the main processor called the Power Processing Element(PPE) and eight fully functional co-processors called the Synergistic Processing Elements(SPEs). In this paper, Sony PlayStation3, which has the Cell.B.E. is used, to which Linux and C programming development software [6] are installed.
- GPU[9]: This is originally developed for image processing, and is implemented on devices such as graphic cards.
- GPGPU[10]: The technique to utilize computing abilities of GPUs for general purpose computing except image processing.
- CUDA[11] toolkit [7]: This is the software development kit, based on C language, to execute GPGPU on NVIDIA inc. GPU devices[12].
- OpenCL[13][8]: This is the cross-platform framework for parallel computing on heterogeneous computing environment, such as a mixed one of multicore-CPUs, GPUs, the Cell.B.E, and so on.
- BP algorithm[14][9]: This is one of learning methods to train multilayer perceptrons. This algorithm consists of two parts. One of them, which is called the forward calculation, calculates each output value of neuron from the input to output layers. On the other hand, the other, which called the backward calculation, calculates the modification of

[8] Cell.B.E. stands for Cell Broadband Engine.

[9] GPU stands for Graphics Processing Unit.

[10] GPGPU stands for General-purpose computing on GPUs.

[11] CUDA stands for Compute Unified Device Architecture.

[12] Windows, Linux, and Mac versions are available on the NVIDIA web site.

[13] OpenCL stands for Open Computing Language.

[14] BP stands for Back-Propagation.

each weight value. The fourth grade students who engage their graduation studies of authors' laboratory can use source code files of BP simulation programs on the file server. Concerning the source codes, C-language codes for single CPU and CUDA codes for parallel computing on GPUs are available since Apr. 2006 and Aug. 2010, respectively.

V. Results of Graduation Studies in Each Year and Considerations

Tables III and IV show the result of graduation study per student in each year. In these tables, the value of "hours" indicates the total time (hours) which the student ake concerning his/her graduation study. Before 2010, these values are approximate ones calculated based on mainly the data of Tables I and II in Section III. Since 2011, these values are measured by checking the picture images of the laboratory scene which are captured by the network camera per ten minutes.

TABLE III
Results of Graduation Studies (1)

student	hours	results
		2007
A	564	A KURO-BOX cluster system (consisting of 9 KURO-BOXes) is constructed. A PC cluster system (consisting of 5 PCs) is constructed. Benchmark programs are executed in these clusters.
B	508	The low level MPI programming of the BP simulation is done in the PC cluster constructed by the student A.
		2008
C	564	The Cell.B.E. programming environment on the Sony PlayStation3 is constructed. The programming of the BP simulation using a PPE and a SPE of the Cell.B.E is done.
D	564	The construction of the KURO-BOX cluster system (consisting of 16 KURO-BOXes) is constructed by extending the KURO-BOX cluster of the last year. Benchmark programs are executed in this cluster. The presentation of these results are given in the IEICE CPSY study group.
		2009
E	667	The Cell.B.E. programming modification of the BP simulation is done based on the program and the programming environment of the last year, using the vector-type data structure.
F	667	Another PC cluster system (consisting of 4 PCs) is constructed. Benchmark programs are executed in this cluster.
G	550	The CUDA programming environment (consisting of the toolkit and a GPU device) is installed to a Windows PC. The low level CUDA programming of the BP simulation, which uses only one CUDA core, is done.

Here, concerning productive level parallel programming skills, if someone can generate a sufficiently parallelized programming code of an algorithm which leads to a product creation, we consider that this person has a productive level parallel programming skill. In other words, such person can understand the algorithm and its concept which has some difficulties to be understood and for the parallelized coding, and also can understand the architecture of the parallel computing hardware and the parallel computing language grammars, and then can apply these grammars to the algorithm considering the architecture.

In addition, the BP algorithm is treated as one which may lead to a product creation, because the BP algorithm has some difficulties to be understood and for the parallelized coding, even though it is now already old and not practical. On the other hand, the algorithms such as a matrix multiplication and a moving average calculation are not treated as one which may lead to a product creation because of their easiness.

TABLE IV
Results of Graduation Studies (2)

student	hours	results
		2010
H	600	The OpenCL programming environment (consisting of the toolkit and a GPU device) is installed to a Windows PC. The sufficiently parallelized OpenCL programming of the BP simulation, including the use of the vector-type data structure, is done.
I	667	The sufficiently parallelized MPI programming of the BP simulation is done using the PC cluster of the last year.
J	667	The high level CUDA programming of the forward calculation part of the BP simulation is done, by referring the author's CUDA code of the same calculation.
K	667	The construction of the Cell.B.E programming environment which enables the use of two or more SPEs is tried, but that is failed.
		2011
L	734	The teaching materials of CUDA programmings are developed. The topics in these materials are, a matrix multiplication, an image converter to binary colors, and a template matching.
M	587	The teaching materials of OpenCL programmings are developed. The topics in these materials are, a moving average calculation, and parallel calculations of addition, subtraction, multiplication and division between couples of scalar values.
N	746	The OpenCL programming environment (consisting of the toolkit and a GPU device) is installed to a Windows PC. The "Hello World" sample program is executed on this environment.

From the above and Tables III and IV, we can see the followings.

- The students which have sufficiently achieved the parallel programming skills are only the four students of H and I in 2010 and L and M in 2011.
  - Especially, the students which have not only achieved the parallel programming skills but also understood the concept of BP algorithm are only the two students of H and I in 2010.
  - At least, this reason includes that these two students are governmental-sponsored foreign students, who are forbidden to hunt job, then they could sufficiently pay attention to their graduation study works.
- From the above, it is a very high level challenge that a student achieves productive level parallel programming skills during only the graduation study.
- On the other hand, most students can achieve only parallel programming skills during the graduation study (about 600 to 700 hours), if the case of the student N in 2011 can be treated as an exception.
- Concerning parallel programming environments, we consider that the environments for GPGPU, which are CUDA and OpenCL in this paper, are more suitable for parallel computing education than other environments

such as the MPI on cluster systems and the Cell.B.E., at least for students in SYOKUGYO DAI, because among these four students, three students chose CUDA or OpenCL and the characteristics of CUDA and OpenCL for GPGPU are similar to each other. Concerning this topic, we consider that CUDA and OpenCL are easier than the other environments in terms of installation and programming. For example, in CUDA and OpenCL, the coding for accessing a global memory from each computing core is easier than in others, because the global variables can be defined and accessed like the coding for global matrix variables in C programming. On the other hand, for example, the coding of the DMA (Direct Memory Access) transfer between the PPE and the SPE in the Cell.B.E. is very difficult for programmers, especially for beginners including the students of SYOKUGYO DAI.

VI. Conclusions

A consideration of the achievement of productive level parallel programming skills, based on the data of graduation studies in the Polytechnic University of Japan, is given. The data show that most students can achieve only parallel programming skills during the graduation study, but also show that it is a very high level challenge that a student achieves productive level parallel programming skills during only the graduation study. In addition, it can be seen that the environments for GPGPU, which are CUDA and OpenCL in this paper, are more suitable for parallel computing education, at least for students in SYOKUGYO DAI, than other environments such as MPI on a cluster system and Cell.B.E.

As future works, we will study an educational method to accelerate the achievement of parallel programming skills based on the result given in this paper.

Acknowledgment

The authors thank SYOKUGYO DAI students for their diligent works in graduation studies.

References

[1] C. Ivica, J.T. Riley, and C. Shubert. StarHPC – Teaching parallel programming within elastic compute cloud. *Proc. Int'l Conf ITI, pp.353-356*, 2009.

[2] W.B. Gardner. Third-year parallel programming for CS undergraduates. *Proc. Int'l Conf FECS, pp.8–13, CSREA*, 2011.

[3] (2013) SYOKUGYO DAI homepage. [Online]. Available: http://www.uitec.jeed.or.jp /overview/index.html.

[4] (2012) Open MPI website. [Online]. Available: http://www.open-mpi.org/.

[5] (2012) KURO-BOX/HG (in Japanese) homepage. [Online]. Available: http://www.kuroutoshikou.com/modules/display/index.php?cid=218&mode=1.

[6] (2012) PLAYSTATION3 Linux Information Site. [Online]. Available: http://cell.fixstars.com/ps3linux/.

[7] (2012) CUDA Downloads homepage. [Online]. Available: http://developer.nvidia.com/cuda-downloads.

[8] (2012) OpenCL homepage. [Online]. Available: http://www.khronos.org/ opencl/.

[9] J.E. Dayhoff. *"Neural Network Architectures: An Introduction"*. Van Nostrand Reinhold, 1989.