CONCURRENCY AND COMPUTATION: PRACTICE AND EXPERIENCE
Concurrency Computat.: Pract. Exper. 2017; **29**:e4348
Published online 5 October 2017 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/cpe.4348

EDITORIAL

Computer architecture and high performance computing

This special issue of *Concurrency and Computation Practice and Experience* gathers eleven selected research articles that were previously presented at the Brazilian "XVII Simpósio em Sistemas Computacionais de Alto Desempenho," WSCAD 2016, held in conjunction with 28th International Symposium on Computer Architecture and High Performance Computing, *SBAC-PAD 2015*, *Florianópolis*, SC, Brazil, from the 19th to the 21st October 2015. Since 2000, this workshop has presented important and interesting research in the fields of computer architectures, high performance computing, and distributed systems.

The scope of the current special issue is broad and representative of the multidisciplinary nature of high performance and distributed computing, covering a wide range of subjects such as architecture issues, compiler optimization, analysis of HPC applications, job scheduling, and energy efficiency.

The title of the first paper is "An efficient virtual system clock for the wireless Raspberry Pi computer platform," by Diego L. C. Dutra, Edilson C. Corrêa, and Claudio L. Amorim [1]. In this paper, the authors present the design and experimental evaluation of an implementation of the RVEC virtual system clock in the Linux kernel for the EE (Energy-Efficient) Wireless Raspberry Pi (RasPi) platform. In the RasPi platform, the use of DVFS (Dynamic Voltage and Frequency) for reducing the energy consumption hinders the direct use of the cycle count of the ARM11 processor core for building an efficient system clock. Therefore, a distinct feature of RVEC is to obviate this obstacle, such that it can make use of the cycle count circuit for precise and accurate time measurements, concurrently with the use of DVFS by the operating system of the ARM11 processor core.

In the second contribution, entitled "Portability with efficiency of the advection of BRAMS between multi-core and many-core architectures," the authors, Manoel Baptista Silva Junior, Jairo Panetta, and Stephan Stephany [2], show the feasibility of writing a single portable code embedding both interfaces (the OpenMP programming interface and OpenACC). It presents acceptable efficiency when executed on nodes with multi-core or many-core architecture. The code chosen as a case study is the advection of scalars, a part of the dynamics of the regional atmospheric model Brazilian Regional Atmospheric Modeling System (BRAMS). The dynamics of this model is hard to parallelize due to data dependencies between adjacent grid points. Single-node executions of the advections of scalars for different grid sizes using OpenMP or OpenACC yielded similar speed-ups, showing the feasibility of the proposed approach.

In the third contribution, entitled "SMT-based context-bounded model checking for CUDA programs," the authors (Phillipe Pereira, Higo Albuquerque, Isabela da Silva, Hendrio Marques, Felipe Monteiro, Ricardo Ferreira, and Lucas Cordeiro) [3] present the ESBMC-GPU tool, an extension to the Efficient SMT-Based Context-Bounded Model Checker (ESBMC), which is aimed at verifying Graphics Processing Unit (GPU) programs written for the Compute Unified Device Architecture (CUDA) platform. ESBMC-GPU uses an operational model, that is, an abstract representation of the standard CUDA libraries, which conservatively approximates their semantics, in order to verify CUDA-based programs. It then explicitly explores the possible interleavings (up to the given context bound), while treats each interleaving itself symbolically. Additionally, ESBMC-GPU employs the monotonic partial order reduction and the two-thread analysis to prune the state space exploration.

The fourth contribution, entitled "Contextual Spaces Re-Ranking: accelerating the Re-sort Ranked Lists step on heterogeneous systems," by Flávia Pisani, Daniel C. G. Pedronette, Ricardo da S. Torres,

and Edson Borin [4], shows an improvement on the efficiency of one of Re-ranking algorithms, called Contextual Spaces Re-Ranking (CSRR). One of the approaches consists in parallelizing the algorithm with OpenCL to use the central and graphics processing units of an accelerated processing unit. The other is to modify the algorithm to a version that, when compared with the original CSRR, not only reduces the total running time of their implementations by a median of 1.6 times but also increases the accuracy score in most of their test cases. Different implementations for CSRR's Re-sort Ranked Lists step were explored as well, providing insights into graphics processing unit sorting, the performance impact of image descriptors, and the trade-offs between effectiveness and efficiency.

In the fifth contribution, entitled "Design methodology for workload-aware loop scheduling strategies based on genetic algorithm and simulation," Pedro H. Penna, Márcio Castro, Henrique C. Freitas, François Broquedis, and Jean-François Méhaut [5] introduce a design methodology to propose, study, and assess the performance of workload-aware loop scheduling strategies. In this methodology, a genetic algorithm is employed to explore the state space solution of the problem itself and to guide the design of new loop scheduling strategies. A simulator is used to evaluate their performance. As a proof of concept, they show how their methodology was used to propose and study a new workload-aware loop scheduling strategy named smart round-robin (SRR).

In the sixth paper, entitled "Exploiting performance, dynamic power and energy scaling in full-system simulators," Liana Duenha, Guilherme Madalozzo, Fernando Gehm Moraes, and Rodolfo Azevedo [6] present an extension of a framework for Multiprocessor Systems on Chip (MPSoCs) designs to support DVFS (Dynamic Voltage and Frequency Scaling) in MPSoCs simulators. They also evaluate three DVFS mechanisms. The experiments show that applying DVFS in the system can save power and energy consumption, with a negligible loss on performance.

In the seventh paper, entitled "Autotuning CUDA Compiler Parameters for Heterogeneous Applications using the OpenTuner Framework," Pedro Bruel, Marcos Amaris, and Alfredo Goldman [7] implement an autotuner for the CUDA compiler's parameters using the OpenTuner framework. The autotuner searches for a set of compilation parameters that optimizes the time to solve a problem. The authors analyze the performance speedups, in comparison with high-level compiler optimizations, achieved in three different GPU devices, for 17 heterogeneous GPU applications, 12 of which are from the Rodinia Benchmark Suite. The autotuner often beats the compiler's high-level optimizations, but is also underperformed for some problems.

In the eighth paper, which is entitled "Quantum computing simulation through reduction and decomposition optimizations with a case study of Shor's algorithm," Anderson Avila, Renata Reiser, and Mauricio Pilla [8] present reduction and decomposition optimizations for the Distributed Geometric Machine environment in the context of multidimensional quantum applications simulation, which is very relevant for the development and testing of new quantum algorithms. In order to overcome the increase in simulation complexity, the authors explore properties as the sparsity of the Identity operator and partiality of dense unitary transformations, better storage and distribution of quantum information are achieved. The main improvements are reached by decreasing replication and void elements inherited from quantum operators. In the evaluation of this proposal, Hadamard transformations from 21 to 28 qubits and Quantum Fourier Transforms from 26 to 28 qubits were simulated over CPU, sequentially and in parallel, and in graphics processing unit, showing reduced temporal complexity and, consequently, shorter simulation time.

The paper "Accelerating Pre-stack Kirchhoff Time Migration by Manual Vectorization" authored by Maicon Alves, Reynam Pestana, Rodrigo A.P. da Silva, and Lucia Drummond [9] study the Pre-stack Kirchhoff Time Migration (PKTM) which is a central process in petroleum exploration. As PKTM is computationally intensive, many papers have proposed the use of accelerators like GPU and FPGA to improve its execution time. On the other hand, although many off-the-shelf processors are endowed with a set of SIMD vector instructions, few papers tackle the problem considering vectorization and all of them consider that compilers can successfully vectorize the code. In this paper, the authors show that programming PKTM by using SIMD vector instructions manually is more efficient than the automatically and semi-automatically vectorized codes, provided by a hardware specific compiler and library.

The paper "Handling IoT platform heterogeneity with COISA, a compact OpenISA virtual platform," by Rafael Auler, Carlos Millani, Alexandre Brisighello, Alisson Linhares, and Edson

DOI: 10.1002/cpe

Borin [10], presents COISA, a compact virtual platform that relies on OpenISA, an instruction set architecture (ISA) that strives for easy emulation, to allow a single program to be deployed on many platforms, including tiny microcontrollers. By exploring the benefits of using a concrete ISA as a Virtual Machines language, experimental results indicate that COISA is easily portable and is capable of running unmodified guest applications in highly heterogeneous host platforms, including one with only 2 kB of RAM. For time-critical IoT applications on constrained platforms where extracting performance is of paramount importance, the authors propose the use of cloud-assisted translations, which employ static binary translation to deliver a binary fully converted to the native ISA used in the IoT device.

The last paper in this special issue, "Performance and Energy Efficiency Analysis of HPC Physics Simulation Applications in a Cluster of ARM Processors" authored by Jean Luca Bez, Eliezer E. Bernart, Fernando F. Dos Santos, Lucas Mello Schnorr, and Philippe Navaux [11], analyzes the feasibility and energy efficiency of using an unconventional cluster of low-power Advanced RISC Machines processors to execute two scientific parallel applications. For this purpose, they have selected two applications that present high computational and communication cost: Ondes3D that simulates geophysical events, and all-pairs N-Body that simulates astrophysical events. The authors compare and discuss the impact of different compilation directives and processor frequency and how they interfere in *Time-to-Solution* and *Energy-to-Solution*. The results demonstrate that by correctly tuning the application at compile time, for the Advanced RISC Machines architecture, it is possible to reduce both the execution time and the energy spent by computing simulations.

To conclude, we sincerely hope that this special issue stimulates your interest in the many issues related to the area of high performance computing. The topics covered in the papers are timely and important, and the authors have done an excellent job on presenting their different approaches. Regarding the reviewing process, our reviewers made a great effort to evaluate the papers. We would like to acknowledge their effort in providing us with the excellent feedback at the right time. Therefore, we wish to thank all the authors and reviewers. Finally, we would also like to express our gratitude to the Editor-in-Chief of CCPE, for his advice, vision, and support.

REFERENCES

1. Dutra DLC, Corrêa EC, Amorim CL. An efficient virtual system clock for the wireless raspberry pi computer platform. *Concurrency and Computation: Practice and Experience* 2017; **29**:e3960. https://doi.org/10.1002/cpe.3960
2. Silva Junior MB, Panetta J, Stephany S. Portability with efficiency of the advection of BRAMS between multi-core and many-core architectures. *Concurrency and Computation: Practice and Experience* 2017; **29**:e3959. https://doi.org/10.1002/cpe.3959
3. Pereira P, Albuquerque H, da Silva I, Marques H, Monteiro F, Ferreira R, Cordeiro L. SMT-based context-bounded model checking for CUDA programs. *Concurrency and Computation: Practice and Experience* 2017; **29**:e3934. https://doi.org/10.1002/cpe.3934
4. Pisani F, Pedronette DCG, Torres RS, Borin E. Contextual Spaces Re-Ranking: accelerating the Re-sort Ranked Lists step on heterogeneous systems. *Concurrency and Computation: Practice and Experience* 2017; **29**:e3962. https://doi.org/10.1002/cpe.3962
5. Penna PH, Castro M, Freitas HC, Broquedis F, Méhaut J-F. Design methodology for workload-aware loop scheduling strategies based on genetic algorithm and simulation. *Concurrency and Computation: Practice and Experience* 2017; **29**:e3933. https://doi.org/10.1002/cpe.3933
6. Duenha L, Madalozzo G, Moraes FG, Azevedo R. Exploiting performance, dynamic power and energy scaling in full-system simulators. *Concurrency and Computation: Practice and Experience* 2017; **29**:e4034. https://doi.org/10.1002/cpe.4034
7. Bruel P, Amaris M, Goldman A. Autotuning CUDA compiler parameters for heterogeneous applications using the OpenTuner framework. *Concurrency and Computation: Practice and Experience* 2017; **29**:e3973. https://doi.org/10.1002/cpe.3973
8. de Avila AB, Reiser RHS, Pilla ML. Quantum computing simulation through reduction and decomposition optimizations with a case study of Shor's algorithm. *Concurrency and Computation: Practice and Experience* 2017; **29**:e3961. https://doi.org/10.1002/cpe.3961
9. Melo Alves M, da Cruz Pestana R, Alves Prado da Silva R, Drummond LMA. Accelerating Pre-stack Kirchhoff Time Migration by Manual Vectorization. *Concurrency and Computation: Practice and Experience* 2017; **29**:e3935. https://doi.org/10.1002/cpe.3935

10. Auler R, Millani CE, Brisighello A, Linhares A, Borin E. Handling IoT platform heterogeneity with COISA, a compact OpenISA virtual platform. *Concurrency and Computation: Practice and Experience* 2017; **29**:e3932. https://doi.org/10.1002/cpe.3932
11. Bez JL, Bernart EE, dos Santos FF, Schnorr LM, Navaux POA. Performance and energy efficiency analysis of HPC physics simulation applications in a cluster of ARM processors. *Concurrency and Computation: Practice and Experience* 2017; **29**:e4041. https://doi.org/10.1002/cpe.4041

Alfredo Goldman
Computer Science, University of São Paulo, Rua do Matão, 1010
São Paulo, SP 05508090
Brazil
E-mail: gold@ime.usp.br

Luciana Arantes
UPMC, LIP6/REGAL/INRIA, 5, Place Jussieu
Paris 75005
France

Edward Moreno
Dept. of Computer and Information Science, Federal University of Sergipe
São Cristovão 49100-000
Brazil

DOI: 10.1002/cpe