Special Issue Proposal for the Parallel Computing Journal: HeteroPar 2016 and HCW 2016 Workshops Editorial

Loris Marchal* Erik Saule† Oliver Sinnen‡

January 12, 2020

Heterogeneity in computing platforms has been a hot topic for some time now, as attested by the continuous existence of the Heterogeneity in Computing Workshop (HCW) for 27 years, and of the International Workshop on Algorithms, Models and Tools for Parallel Computing on Heterogeneous Platforms (HeteroPar) for 16 years. These workshops have seen a wide variety of heterogeneous architectures, ranging from heterogeneous networks of workstations to big.LITTLE processors. Nowadays, heterogeneous computing mainly takes the form of hybrid platforms made of both regular CPU cores and accelerators (such as GPUs) dedicated to specific computations. While these platforms are now commonplace (138 machines of the November 2018 Top500 ranking use such accelerators), their efficient use is still challenging: many classical HPC applications mainly make use of only a single type of the cores (CPUs or GPUs) at a time. Besides, energy optimization has emerged as an important objective in HPC, which again motivates the use of GPUs or other energy-efficient accelerators, at least for some part of the computations.

This special issue of the Parallel Computing journal hosts four articles, which are all significant extensions of a first publication in either the HCW 2016 workshop (held in conjunction with IEEE IPDPS and chaired by Erik Saule) or in the HeteroPar 2016 workshop (held in conjunction with EuroPar 2016 and chaired by Loris Marchal). Following these venues, we proposed to all authors of accepted papers in these workshops to submit an extended version of their work. Nine manuscripts were received, and underwent a rigorous journal peer-review process. Eventually, we selected the four best reviewed articles. They all study how to efficiently use hybrid machines, made of several types of computing units (CPUs, GPUs, Xeon Phis or FPGAs). Two of them are also concerned with energy issues.

The first paper is entitled "*Utility-Based Resource Management in an Oversubscribed Energy-Constrained Heterogeneous Environment Executing Parallel Applications*". It studies a computing platform made of different clusters. Each cluster contains homogeneous computing nodes, however different clusters embed different nodes, such as clusters with specialized hardware that are designed to process specific tasks (e.g., GPUs or other accelerators). The paper proposes and studies several resource allocation heuristics that increase the system utility.

The second paper studies linear algebra algorithms for hybrid platforms and is entitled "*Resource Aggregation for Task-Based Cholesky Factorization on Top of Modern Architectures*". It first acknowledges that each type of architecture (such as CPU or GPU) has an optimal task size (or block size in case of matrix operations). Optimal GPU block sizes are generally much larger than optimal CPU block sizes. This paper proposes to alleviate the platform heterogeneity by aggregating multiple CPUs so as to use larger block sizes for all resources. It presents experiments with several types of accelerators: GPUs or Intel Xeon Phi Knights Landing processors.

The third paper focuses on the energy efficiency of GPUs and is entitled "*DVFS-Aware Application Classification to improve GPGPUs Energy Efficiency*'". It studies the dynamic voltage and

*Univ Lyon, EnsL, UCBL, CNRS, Inria.
†UNC Charlotte
‡Univ. of Auckland

frequency scaling (DVFS) for GPUs and proposes a classification method for both applications and architectures. Thanks to this classification, it is possible to predict the impact of DVFS on a given application's execution time, power and energy consumption.

The fourth paper, entitled "*Empirical Modeling and Simulation of an Heterogeneous Cloud Computing Environment*". It studies a hybrid Cloud environment embedding FPGAs for better energy efficiency. It proposes a framework as well as allocation algorithms to effectively managing such a system. It also presents a simulation tool to study the impact on performance and energy consumption when taking advantage of FPGAs.

The guest editors thank all the authors who submitted their work for consideration. We are also much indebted to the many colleagues who kindly made their time available to assist in the reviewing process. We finally thank associate editor Michaela Taufer for her help in preparing this special issue, as well as past and current editors in chief Jeff Hollingsworth and Ümit V. Çatalyurek for offereing us an opportunity to publish it. We hope you will enjoy reading the selected articles and appreciate their contributions as we did.