Parallel and distributed algorithms

Guillaume Aupy, Xueyan Tang

Parallel and Distributed Algorithms

Guillaume Aupy[1] | Xueyan Tang[2]

[1]Inria Bordeaux Sud – Ouest & Univ. of Bordeaux, France

[2]School of Computer Science and Engineering, Nanyang Technological University, Singapore

Correspondence

Email: guillaume.aupy@inria.fr,

Email: asxytang@ntu.edu.sg

Abstract

We introduce the papers submitted to the special issue of *Computation, Concurrency: Practice and Experience* on parallel and distributed algorithms.

KEYWORDS:

Editorial

1 | INTRODUCTION

Welcome to this special issue! This issue compiles latest research results on parallel and distributed algorithms. Selected papers from the algorithm track of the 17th International Conference on Algorithms and Architectures for Parallel Processing (ICA3PP-2017) and from the 12th Scheduling for Large Scale Systems Workshop were invited for extension and submission to this issue. The general objectives of these two events is to target both theoretical and experimental algorithms for novel parallel architectures. We also received submissions from an open call.

2 | THEMES OF THIS SPECIAL ISSUE

Five papers were finally accepted to this special issue. These papers cover a wide range of topics coming from parallel computing: run time-systems to deal with heterogeneous architecture, linear algebra applications, but also energy-efficient issues linked to infrastructure such as mobile computing or cloud computing.

2.1 | Runtime systems to deal with heterogeneous architecture

Runtime systems have gained tremendous momentum with the increasing complexity of novel parallel architectures. From an application point of view, leaving the hard scheduling part to a highly optimized system (such as StarPU, Parsec or Pycomps) is very interesting. In this issue we welcome two very different paper, one offering solution to improve the performance of these runtime systems, and another where an application is using such implementations.

Fast Approximation Algorithms for Task-Based Runtime Systems [1], conducts a theoretical study of the HeteroPrio scheduling policy for complex computations on heterogeneous platforms. Heterogeneity is common in high performance computing with specialized accelerators like GPUs providing efficient computational power. It proves approximation bounds compared to the optimal schedule in both cases of independent tasks and general task graphs. The paper also provides an experimental evaluation of HeteroPrio on real task graphs from dense linear algebra computation.

Dataflow programming models are becoming popular as a means to deliver a good balance between performance and portability in the post-petascale era. Evaluation of Dataflow Programming Models for Electronic Structure Theory [2], evaluate two programming paradigms for expressing scientific applications in a dataflow form: "explicit dataflow" where the dataflow is specified explicitly by the developer, and "implicit dataflow" where a task scheduling runtime derives the dataflow using per-task data-access information embedded in a serial program. The paper compares these two models in terms of programmability, resource utilization, and scalability.

2.2 | Matrix factorization with MapReduce

The third paper, Distributed Geometric Nonnegative Matrix Factorization and Hierarchical Alternating Least Squares-based Nonnegative Tensor Factorization with the MapReduce Paradigm [3], proposes various parallel and distributed computation strategies for the latent factors in both factorization models based on partitioning the computational tasks according to the MapReduce paradigm. These factorization models are widely used in machine learning and data analysis for feature extraction and dimensionality reduction. Experiments using various large-scale data sets demonstrated that the proposed algorithms are efficient and robust to noisy data.

2.3 | Energy-efficiency

In energy-efficient computing, shutdown techniques have been developed to adapt the number of switched-on servers to the actual workload. The fourth paper, Quantifying the Impact of Shutdown Techniques for Energy-Efficient Data Centers [4], evaluates the potential gain of shutdown techniques on recent server architectures and future energy-aware architectures by taking into account shutdown and boot up costs in time and energy. The evaluation exploits real traces collected on production infrastructures with several shutdown policies, with and without workload prediction.

The low performance and reduced power-capacity of mobile devices severely limit the complexity and duration of mobile gaming. The fifth paper, A Mirroring Architecture for Sophisticated Mobile Games using Computation-Offloading [5], explores moving the compute-intensive parts of mobile games to execute on remote infrastructure. It proposes a system for computation-offloading that supports the demanding performance requirements of sophisticated mobile games. A real-world prototype is implemented and applied to the popular game OpenTTD, which demonstrates that the system can significantly improve performance and power consumption, while also delivering smooth gameplay.

ACKNOWLEDGEMENTS

References

1. Olivier Beaumont, Lionel Eyraud-Dubois, and Suraj Kumar, *Fast Approximation Algorithms for Task-Based Runtime Systems*. Concurrency and Computation: Practice and Experience [this issue].

2. Heike Jagode, Anthony Danalis, Reazul Hoque, Mathieu Faverge, and Jack Dongarra, *Evaluation of Dataflow Programming Models for Electronic Structure Theory*. Concurrency and Computation: Practice and Experience [this issue].

3. Rafal Zdunek and Krzysztof Fonal, *Distributed Geometric Nonnegative Matrix Factorization and Hierarchical Alternating Least Squares-based Nonnegative Tensor Factorization with the MapReduce Paradigm*. Concurrency and Computation: Practice and Experience [this issue].

4. Issam Rais, Anne-Cécile Orgerie, Martin Quinson, and Laurent Lefèvre. *Quantifying the Impact of Shutdown Techniques for Energy-Efficient Data Centers*. Concurrency and Computation: Practice and Experience [this issue].

5. Minghai Jiang, Otto Visser, Wishnu Prasetya, and Alexandru Iosup, *A Mirroring Architecture for Sophisticated Mobile Games using Computation-Offloading*. Concurrency and Computation: Practice and Experience [this issue].

AUTHOR BIOGRAPHIES

Guillaume Aupy received his PhD from ENS Lyon in 2014. He is currently a researcher at Inria Bordeaux Sud-Ouest. His research interests include data-aware scheduling at the different levels of the memory hierarchy, but also energy-efficient and reliable scheduling strategies in High-Performance Computing. He was the Technical Program vice-chair of the ACM/IEEE SC'17 conference.

Xueyan Tang received the B.Eng. degree in computer science and engineering from Shanghai Jiao Tong University in 1998, and the Ph.D. degree in computer science from The Hong Kong University of Science and Technology in 2003. He is currently an Associate Professor with the School of Computer Science and Engineering, Nanyang Technological University, Singapore. His research interests include distributed systems, cloud computing, mobile and pervasive computing, and wireless sensor networks. He has served as an Associate Editor of the IEEE Transactions on Parallel and Distributed Systems, and a Program Co-Chair of the IEEE ICPADS 2012 and the IEEE CloudCom 2014 conferences.