Hardware Architecture Modeling for Massively Parallel Real-time System Scheduling

PhD proposal (under CIFRE contract)
Zhen Zhang, Arnaud Lallouet, Chong Li
Huawei Technologies – France Research Center
2012 Labs / Central Software Institute / DAL
20 quai du Point du Jour, 92100 Boulogne-Billancourt

Dumitru Potop-Butucaru
INRIA, team AOSTE
2 Rue Simone IFF, 75012 Paris, France

Context
Huawei is a leading global ICT solutions provider. Through our dedication to customer-centric innovation and strong partnerships, we have established end-to-end capabilities and strengths across the carrier networks, enterprise, consumer, and cloud computing fields. We are committed to creating maximum value for telecom carriers, enterprises and consumers by providing competitive ICT solutions and services. Our products and solutions ranging from processors, servers to mobile phones have been deployed in over 170 countries, serving more than one third of the world’s population. Huawei has launched in 2014 the French Research Center (FRC), focusing on mathematics and algorithmic science, with more than 80 researchers. The Distributed Algorithms Lab (DAL) develops algorithms and programming tools to support massively parallel big-data applications, high performance machine learning, computer vision and real-time embedded technologies.

Established in 1967, INRIA is the only French public research body fully dedicated to computational sciences. Combining computer sciences with mathematics, INRIAs 3,500 researchers strive to invent the digital technologies of the future. Educated at leading international universities, they creatively integrate basic research with applied research and dedicate themselves to solving real problems, collaborating with the main players in public and private research in France and abroad and transferring the fruits of their work to innovative companies. The researchers at Inria published over 4,450 articles in 2012. They are behind over 250 active patents and 112 startups. The 180 project teams are distributed in eight research centers located throughout France. The AOSTE research team (http://www.inria.fr/en/teams/aoste) promotes the use of synchronous formalisms for the high-level modeling, the full formal design, and the distributed real-time implementation of embedded software. The team builds upon prior work by its members on the SyncCharts, Esterel, and SynDEx formalisms, which included extensive algorithmic studies on dedicated modeling, compilation, analysis, and optimization techniques. Our main expertise is in the fields of formal semantics of synchronous reactive systems, and optimized mapping (i.e. distribution and scheduling) between application algorithms and physical architectures descriptions.

Project
Facing forthcoming Artificial Intelligence needs, future embedded real-time systems will become massively parallel both on the hardware and the software part. Some new large-scale intelligent applications will have hard real-time constraints on components such as control, image processing,
object recognition, machine learning, deep learning, multiple-sensor information fusion and on-line AI decision. It will be very hard to implement such applications on a massively parallel hardware, for example a many-core system, while meeting all critical constraints. Therefore early stage scheduling of functional tasks, communications relying on non-functional constraints on an abstract hardware architecture model will be an efficient solution at all steps of system design, implementation, verification, validation and test.

In this project, we will study concurrent functional specification models, such as dataflow synchronous [5], bridging parallel models such as BSP [7], Multi-BSP [8, 6] or PRAM, and state-of-the-art scheduling algorithms (based on both constraint solving engines or dedicated heuristics [2]) on top of such models. Our goals are to:

- Define a general model for massively parallel execution platforms
- Define efficient scheduling algorithms.

Constraint Programming has already been used, by numerous research groups, to model and solve scheduling problems under increasingly complex non-functional requirements (real-time, allocation, etc.). However, recent work by our teams [3, 4] and others [1] has shown that existing modeling approaches and solvers do not scale beyond certain limits, and put into evidence factors that influence these limits – the problem size, its complexity, or the system load. In particular, our work has shown that the allocation and scheduling of parallelized real-time applications onto massively parallel hardware under high system load is untractable in practice.

This thesis proposal aims at better understanding the limitations of constraint solving techniques and the factors influencing the empirical complexity of scheduling problems. Based on this better understanding, our objective is to extend the reach of such approaches by means of improved encoding of the scheduling problems and by improvements to constraint solving techniques/engines. We also aim at developing efficient dedicated heuristics for solving the considered problems.

Student

We are seeking for a Master’s degree graduate in Computer Science from a top-level university/engineer school. The candidate shall have a strong background in more than one of the following fields: real-time systems, artificial intelligence, computer vision, parallel algorithms, machine learning, deep learning, graph algorithms, constraint programming, image processing, etc. In addition, the right person should love the practice as well as the theory: having a good coding skill is essential to transfer beautiful algorithms to perfect programs.

Please send by email to the contacts hereafter an archive containing a CV, a motivation letter, at least two recommendation letters, all undergraduate and graduate marks and a detailed description of the courses followed. Reports of scholar or personal projects, as well as any achievement, prize or distinction will be appreciated.

Work environment

The work will take place in the offices of Huawei FRC (located in Boulogne-Billancourt) and Inria Paris (located in Paris). Supervision will be done by Dr. Zhen Zhang, Prof. Arnaud Lallouet and Dr.-Ing. Chong Li from the Huawei part and by Dr. Dumitru Potop-Butucaru from the INRIA part. Huawei FRC provides a challenging scientific environment, cutting-edge parallel hardware and software, digital library, travel funds, company restaurant and a salary competitive with similar positions in French industry. The successful candidate will receive a job offer in 2016, and the actual doctoral work will begin as soon as possible thereafter.

Contact

Zhen Zhang, Huawei Technologies, zhen.zhang.fr@huawei.com
Arnaud Lallouet, Huawei Technologies, arnaud.lallouet@huawei.com
References


