Thesis Proposal

Scheduling of real-time and heterogeneous multiprocessor platform

Funding : JCJC fellowship from the French National Research Agency (ANR)
Place : Laboratoire d'Automatique et d'Informatique pour les Systèmes (LIAS), Poitiers
Start date : September 2022
Supervision : Pascal RICHARD and Antoine BERTOUT (Real-time team)
(The supervisor whose name is underlined is the referent supervisor, contact : antoine.bertout@univ-poitiers.fr)

Context

MultiProcessor Systems On Chips (MPSoCs) embedded in real-time systems are made of increasingly specialised computing units (CPUs, GPUs, NPU's, etc.). This heterogeneity offers a better use of the resources (processing units, power consumption, etc.) but systems may be harder to predict. Critical real-time systems must provide logical but also timing guarantees.

The software of these systems is represented as a set of tasks with timing constraints, such as a deadline by which the execution of a task must be completed. For the hardware part, these heterogeneous systems are often described in the literature as "unrelated" platforms [1]. In this classification, it is possible to assign a different execution speed to each task/processor pair. In particular, this generalises the so-called "homogeneous" category, where the processors can have different, but constant speeds for all tasks.

The temporal characteristics of such applications require special treatment by the operating system via its scheduler. The algorithm used by the latter must allow efficient management of resources (multi-core processors, memory, etc.) and respect of the timing constraints of the applications whatever the system's operation duration.

This PhD thesis is part of SHRIMP project, funded by a JCJC fellowship from the French National Research Agency (ANR)

Description of the subject

Goal The thesis aims at proposing a global (allowing task migration between processors) and dynamic (online) real-time scheduler for unrelated platforms. This scheduler must be able to manage tasks without a predefined arrival date (sporadic tasks) and to react online to events, while ensuring the respect of task deadlines. Existing solutions are static. They do not allow for a satisfactory use of resources. For example, they cannot take advantage of the early termination (before the end of its worst-case execution time) of a task. Moreover, the task models considered in this work are not adapted to the characteristics of modern applications (dependencies) and realistic (monolithic execution time worst-case for a task possibly running on different processors). The project aims to consider first a special case of "unrelated" platforms called "consistent" for which there is a comparison order between the processors but where the speeds of the processors are not necessarily constant (as for the homogeneous platforms). This category makes it possible to represent ARM big.LITTLE type architectures with slow and fast processors, of different architectures but with the same instruction set. Next, we will be critical of the classically used task model and propose a scheduling algorithm capable of scheduling dependent tasks. This latter model would allow for a more accurate representation of tasks with code sections whose execution time could vary depending on the processor used.

Validation The developed solutions will have to be formally validated and compared with seminal schedulers. Through simulations, attention will be paid to the performance of the scheduler, e.g. on the supported usage load or on the number of context changes (preemptions, migrations) which have a strong impact on the applicability of the results. The recruited researcher will be able to rely on an existing prototype of simulator developed in Python. The project also focuses on the practical evaluation of the solution. The scheduling algorithms will have to be implemented on a realistic testbed (e.g. real-time OS or patch). This last task will be specifically carried out by two interns funded by the project, intervening at different times during the thesis. The recruited researcher will be involved in their supervision.

International Collaboration Building on existing works [3, 2] on the subject, the thesis will be carried out in collaboration with Professor Joël Goossens from the Université Libre de Bruxelles.

Required expertise

The candidate should have a master's degree in computer science or an engineering degree and have knowledge of theoretical computer science, computer systems, mathematics and programming. A good level of written and spoken English is required. Skills in French are appreciated.

Documents to be provided

- CV
- Covering letter
- Master's (or equivalent) grades and rankings
- Last internship report
- Any other document deemed necessary by the applicant to enrich the application

References

[1] S. K. Baruah, « Feasibility Analysis of Preemptive Real-Time Systems upon Heterogeneous Multiprocessor Platforms », in Real-Time Systems Symposium, IEEE, 2004, p. 37-46.

[2] A. Bertout, J. Goossens, E. Grolleau, R. Jamil et X. Poczekajlo, « Workload assignment for global real-time scheduling on unrelated clustered platforms », Real-Time Systems, p. 1-32, 2021.

[3] A. Bertout, J. Goossens, E. Grolleau et X. Poczekajlo, « Template schedule construction for global real-time scheduling on unrelated multiprocessor platforms », in 2020 Design, Automation & Test in Europe Conference & Exhibition (DATE), IEEE, 2020, p. 216-221.

[4] E. Massa, G. Lima, B. Andersson et V. Petrucci, « Heterogeneous Quasi-Partitioned Scheduling », in 2021 IEEE Real-Time Systems Symposium (RTSS), 2021, p. 266-278. doi : 10.1109/RTSS52674.2021.00033.