Activity Report 2018

Team CAIRN

Energy Efficient Computing Architectures

Joint team with Inria Rennes – Bretagne Atlantique

D3 – Architecture

Table of contents

Project-Team CAIRN

Creation of the Project-Team: 2009 January 01

CAIRN *is a common project with CNRS, University of Rennes 1, and ENS Rennes, located on two campuses: Rennes (Beaulieu) and Lannion (*ENSSAT*).*

Keywords:

Computer Science and Digital Science:

A1.1. - Architectures
A1.1.1. - Multicore, Manycore
A1.1.2. - Hardware accelerators (GPGPU, FPGA, etc.)
A1.1.8. - Security of architectures
A1.1.9. - Fault tolerant systems
A1.1.10. - Reconfigurable architectures
A1.1.12. - Non-conventional architectures
A1.2.5. - Internet of things
A1.2.6. - Sensor networks
A2.2. - Compilation
A2.2.1. - Static analysis
A2.2.4. - Parallel architectures
A2.2.6. - GPGPU, FPGA...
A2.2.7. - Adaptive compilation
A4.4. - Security of equipment and software
A8.10. - Computer arithmetic

Other Research Topics and Application Domains:

B4.5. - Energy consumption
B4.5.1. - Green computing
B4.5.2. - Embedded sensors consumption
B6.2.2. - Radio technology
B6.2.4. - Optic technology
B6.6. - Embedded systems
B8.1. - Smart building/home
B8.1.1. - Energy for smart buildings
B8.1.2. - Sensor networks for smart buildings

1. Team, visitors, external collaborators

Research Scientists
François Charot [Inria, Researcher]
Silviu Ioan Filip [Inria, Researcher, from Oct 2018]
Tomofumi Yuki [Inria, Researcher]

Faculty Members
Olivier Sentieys [Team leader, Professor, Univ. Rennes, Inria Chair, HDR]
Emmanuel Casseau [Professor, Univ. Rennes, ENSSAT, Lannion, HDR]

Daniel Chillet [Professor, Univ. Rennes, Enssat, Lannion, HDR]
Steven Derrien [Professor, Univ. Rennes, Istic, Rennes, HDR]
Cédric Killian [Associate Professor, Univ. Rennes, Iut, Lannion]
Angeliki Kritikakou [Associate Professor, Univ. Rennes, Istic, Rennes]
Patrice Quinton [Ecole Normale Supérieure de Rennes, Emeritus, Rennes]
Christophe Wolinski [Professor, Univ. Rennes, Director of Esir, Rennes, HDR]

Post-Doctoral Fellows
Silviu Ioan Filip [Inria, until Sep 2018]
Mansureh Shahraki Moghaddam [Inria]
Lei Mo [Inria]

PhD Students
Audrey Lucas [CNRS, granted by DGA-PEC, Lannion, from Jan. 2016]
Genevieve Ndour [Univ. Rennes, granted by CEA Leti, Grenoble, from May 2016]
Joel Ortiz Sosa [Inria, Lannion, from Oct. 2016]
Nicolas Roux [Inria, granted by Brittany Region/LTC, Lannion, from Oct. 2016]
Mael Gueguen [Univ. Rennes, MENRT grant, Rennes, from Nov. 2016]
Minh Thanh Cong [Univ. Rennes, granted by USTH, Rennes, from May 2017]
Thibaut Marty [Univ. Rennes, granted by H2020 ARGO and Brittany Region, Rennes, from Sep. 2017]
Petr Dobias [Univ. Rennes, MENRT grant, Lannion, from Oct. 2017]
Van Phu Ha [Inria, granted by ANR Artefact, Rennes, from Nov. 2017]
Romain Mercier [Inria, granted by DGA and Inria, Lannion, from Oct. 2018]
Minyu Cui [Univ. Rennes, granted by CSC-ENS, Rennes, from Sep. 2018]
Jaechul Lee [Univ. Rennes, granted by Brittany Region/LTC, Lannion, from Dec. 2018]
Davide Pala [Inria, granted by IPL ZEP, Rennes, from Jan. 2018]
Joseph Paturel [Inria, granted by RAPID FLODAM, Rennes, from Sep. 2018]
Aymen Gammoudi [Univ. Rennes, Lannion, until Jun. 2018]
Rafail Psiakis [Univ. Rennes, MENRT grant, Rennes, until Dec. 2018]
Simon Rokicki [Univ. Rennes, granted by ENS Rennes, Rennes, until Dec. 2018]
Gabriel Gallin [CNRS, granted by CominLabs, Lannion, until Nov. 2018]
Jiating Luo [Univ. Rennes, granted by China Gov., Lannion, until Jul. 2018]
Van Dung Pham [Inria, granted by CominLabs, Lannion, until Dec. 2018]

Technical staff
Arnaud Carer [Research Engineer (half time), Univ. Rennes, Lannion]
Mickael Dardaillon [Univ. Rennes, Rennes, until Aug. 2018]
Ali Hassan El Moussawi [Inria, Rennes, until Jun. 2018]
Pierre Guilloux [Univ. Rennes, Lannion, until Dec. 2018]
Pierre Halle [Inria, Lannion]
Simon Rokicki [Inria, Rennes, from Nov 2018]

Administrative Assistants
Nadia Derouault [Assistant, Inria, Rennes]
Emilie Carquin [Assistant, Univ. Rennes, Enssat, Lannion]

2. Overall Objectives

2.1. Overall Objectives

Abstract — The Cairn project-team researches new architectures, algorithms and design methods for flexible, secure, fault-tolerant, and energy-efficient domain-specific system-on-chip (SoC). As performance and energy-efficiency requirements of SoCs, especially in the context of multi-core architectures, are continuously increasing, it becomes difficult for computing architectures to rely only on programmable processors

solutions. To address this issue, we promote/advocate the use of reconfigurable hardware, i.e., hardware structures whose organization may change before or even during execution. Such reconfigurable chips offer high performance at a low energy cost, while preserving a high level of flexibility. The group studies these systems from three angles: (i) The invention and design of new reconfigurable architectures with an emphasis on flexible arithmetic operator design, dynamic reconfiguration management and low-power consumption. (ii) The development of their corresponding design flows (compilation and synthesis tools) to enable their automatic design from high-level specifications. (iii) The interaction between algorithms and architectures especially for our main application domains (wireless communications, wireless sensor networks and digital security).

Keywords — **Architectures:** Embedded Systems, System-on-Chip, Reconfigurable Architectures, Hardware Accelerators, Low-Power, Computer Arithmetic, Secure Hardware, Fault Tolerance. **Compilation and synthesis:** High-Level Synthesis, CAD Methods, Numerical Accuracy Analysis, Fixed-Point Arithmetic, Polyhedral Model, Constraint Programming, Source-to-Source Transformations, Domain-Specific Optimizing Compilers, Automatic Parallelization. **Applications:** Wireless (Body) Sensor Networks, High-Rate Optical Communications, Wireless Communications, Applied Cryptography.

The scientific goal of the CAIRN group is to research new hardware architectures for domain-specific SOCs, along with their associated design and compilation flows. We particularly focus on on-chip integration of specialized and reconfigurable accelerators. Reconfigurable architectures, whose hardware structure may be adjusted before or even during execution, originate from the possibilities opened up by Field Programmable Gate Arrays (FPGA) [60] and then by Coarse-Grain Reconfigurable Arrays (CGRA) [63], [75] [1]. Recent evolutions in technology and modern hardware systems confirm that reconfigurable systems are increasingly used in recent and future applications (see e.g. Intel/Altera or Xilinx/Zynq solutions). This architectural model has received a lot of attention in academia over the last two decades [66], and is now considered for industrial use in many application domains. One first reason is that the rapidly changing standards or applications require frequent device modifications. In many cases, software updates are not sufficient to keep devices on the market, while hardware redesigns remain too expensive. Second, the need to adapt the system to changing environments (e.g., wireless channel, harvested energy) is another incentive to use runtime dynamic reconfiguration. Moreover, with technologies at 28 nm and below, manufacturing problems strongly impact electrical parameters of transistors, and transient errors caused by particles or radiations also often appear during execution: error detection and correction mechanisms or autonomic self-control can benefit from reconfiguration capabilities.

As chip density increased, power or energy efficiency has become "the Grail" of all chip architects. With the end of Dennard scaling [70], multicore architectures are hitting the *utilisation wall* and the percentage of transistors in a chip that can switch at full frequency drops at a fast pace [64]. However, this unused portion of a chip also opens up new opportunities for computer architecture innovations. Building specialized processors or hardware accelerators can come with orders-of-magnitude gains in energy efficiency. Since from the beginning of CAIRN in 2009, we advocate the interest of heterogeneous multicores, in which general-purpose processors (GPPs) are integrated with specialized accelerators, especially when built on reconfigurable hardware, which provides the best trade-off between power, performance, cost and flexibility. During the period, it therefore turns out that the time has come for these heterogeneous manycore architectures.

> Standard multicore architectures enable flexible software on fixed hardware, whereas reconfigurable architectures make possible **flexible software on flexible hardware**.

However, designing reconfigurable systems poses several challenges: the definition of the architecture structure itself, along with its dynamic reconfiguration capabilities, and its corresponding compilation or synthesis tools. The scientific goal of CAIRN is therefore to leverage the background and past experience of its members to tackle these challenges. We propose to approach energy efficient reconfigurable architectures from three angles: (i) the invention and the design of new reconfigurable architectures or hardware accelerators, (ii) the development of their corresponding compilers and design methods, and (iii) the exploration of the interaction between applications and architectures.

3. Research Program

3.1. Panorama

The development of complex applications is traditionally split in three stages: a theoretical study of the algorithms, an analysis of the target architecture and the implementation. When facing new emerging applications such as high-performance, low-power and low-cost mobile communication systems or smart sensor-based systems, it is mandatory to strengthen the design flow by a joint study of both algorithmic and architectural issues.
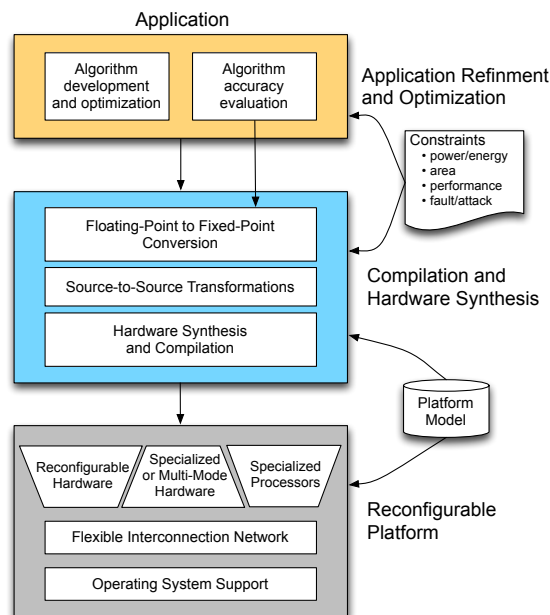
Figure 1. CAIRN*'s general design flow and related research themes*

Figure 1 shows the global design flow we propose to develop. This flow is organized in levels which refer to our three research themes: application optimization (new algorithms, fixed-point arithmetic, advanced representations of numbers), architecture optimization (reconfigurable and specialized hardware, application-specific processors, arithmetic operators and functions), and stepwise refinement and code generation (code transformations, hardware synthesis, compilation).

In the rest of this part, we briefly describe the challenges concerning **new reconfigurable platforms** in Section 3.2 and the issues on **compiler and synthesis tools** related to these platforms in Section 3.3.

3.2. Reconfigurable Architecture Design

Nowadays, FPGAs are not only suited for application specific algorithms, but also considered as fully-featured computing platforms, thanks to their ability to accelerate massively parallelizable algorithms much faster than their processor counterparts [78]. They also support to be dynamically reconfigured. At runtime, partially reconfigurable regions of the logic fabric can be reconfigured to implement a different task, which allows for a better resource usage and adaptation to the environment. Dynamically reconfigurable hardware can also cope with hardware errors by relocating some of its functionalities to another, sane, part of the logic fabric. It

could also provide support for a multi-tasked computation flow where hardware tasks are loaded on-demand at runtime. Nevertheless, current design flows of FPGA vendors are still limited by the use of one partial bitstream for each reconfigurable region and for each design. These regions are defined at design time and it is not possible to use only one bitstream for multiple reconfigurable regions nor multiple chips. The multiplicity of such bitstreams leads to a significant increase in memory. Recent research has been conducted in the domain of task relocation on a reconfigurable fabric. All of the related work was conducted on architectures from commercial vendors (e.g., Xilinx, Altera) which share the same limitations: the inner details of the bitstream are not publicly known, which limits applicability of the techniques. To circumvent this issue, most dynamic reconfiguration techniques are either generating multiple bitstreams for each location [62] or implementing an online filter to relocate the tasks [72]. Both of these techniques still suffer from memory footprint and from the online complexity of task relocation.

Increasing the level and grain of reconfiguration is a solution to counterbalance the FPGA penalties. Coarse-grained reconfigurable architectures (CGRA) provide operator-level configurable functional blocks and word-level datapaths [79], [67], [77]. Compared to FPGA, they benefit from a massive reduction in configuration memory and configuration delay, as well as for routing and placement complexity. This in turns results in an improvement in the computation volume over energy cost ratio, although with a loss of flexibility compared to bit-level operations. Such constraints have been taken into account in the design of DART[9], Adres [75] or polymorphous computing fabrics[11]. These works have led to commercial products such as the PACT/XPP [61] or Montium from Recore systems, without however a real commercial success yet. Emerging platforms like Xilinx/Zynq or Intel/Altera are about to change the game.

In the context of emerging heterogenous multicore architecture, CAIRN advocates for associating general-purpose processors (GPP), flexible network-on-chip and coarse-grain or fine-grain dynamically reconfigurable accelerators. We leverage our skills on microarchitecture, reconfigurable computing, arithmetic, and low-power design, to discover and design such architectures with a focus on: reduced energy per operation; improved application performance through acceleration; hardware flexibility and self-adaptive behavior; tolerance to faults, computing errors, and process variation; protections against side channel attacks; limited silicon area overhead.

3.3. Compilation and Synthesis for Reconfigurable Platforms

In spite of their advantages, reconfigurable architectures, and more generally hardware accelerators, lack efficient and standardized compilation and design tools. As of today, this still makes the technology impractical for large-scale industrial use. Generating and optimizing the mapping from high-level specifications to reconfigurable hardware platforms are therefore key research issues, which have received considerable interest over the last years [65], [80], [76], [74], [73]. In the meantime, the complexity (and heterogeneity) of these platforms has also been increasing quite significantly, with complex heterogeneous multi-cores architectures becoming a *de facto* standard. As a consequence, the focus of designers is now geared toward optimizing overall system-level performance and efficiency [71]. Here again, existing tools are not well suited, as they fail at providing a unified programming view of the programmable and/or reconfigurable components implemented on the platform.

In this context, we have been pursuing our efforts to propose tools whose design principles are based on a tight coupling between the compiler and the target hardware architectures. We build on the expertise of the team members in High Level Synthesis (HLS) [5], ASIP optimizing compilers [12] and automatic parallelization for massively parallel specialized circuits [2]. We first study how to increase the efficiency of standard programmable processors by extending their instruction set to speed-up compute intensive kernels. Our focus is on efficient and exact algorithms for the identification, selection and scheduling of such instructions [6]. We address compilation challenges by borrowing techniques from high-level synthesis, optimizing compilers and automatic parallelization, especially when dealing with nested loop kernels. In addition, and independently of the scientific challenges mentioned above, proposing such flows also poses significant software engineering issues. As a consequence, we also study how leading edge software engineering techniques (Model Driven

Engineering) can help the Computer Aided Design (CAD) and optimizing compiler communities prototyping new research ideas [4].

Efficient implementation of multimedia and signal processing applications (in software for DSP cores or as special-purpose hardware) often requires, for reasons related to cost, power consumption or silicon area constraints, the use of fixed-point arithmetic, whereas the algorithms are usually specified in floating-point arithmetic. Unfortunately, fixed-point conversion is very challenging and time-consuming, typically demanding up to 50% of the total design or implementation time. Thus, tools are required to automate this conversion. For hardware or software implementation, the aim is to optimize the fixed-point specification. The implementation cost is minimized under a numerical accuracy or an application performance constraint. For DSP-software implementation, methodologies have been proposed [7] to achieve fixed-point conversion. For hardware implementation, the best results are obtained when the word-length optimization process is coupled with the high-level synthesis [68]. Evaluating the effects of finite precision is one of the major and often the most time consuming step while performing fixed-point refinement. Indeed, in the word-length optimization process, the numerical accuracy is evaluated as soon as a new word-length is tested, thus, several times per iteration of the optimization process. Classical approaches are based on fixed-point simulations [69]. Leading to long evaluation times, they can hardly be used to explore the design space. Therefore, our aim is to propose closed-form expressions of errors due to fixed-point approximations that are used by a fast analytical framework for accuracy evaluation [10].

4. Application Domains

4.1. Panorama

keywords: Wireless (Body) Sensor Networks, High-Rate Optical Communications, Wireless Communications, Applied Cryptography, Machine Learning.

Our research is based on realistic applications, in order to both discover the main needs created by these applications and to invent realistic and interesting solutions.

Wireless Communication is our privileged application domain. Our research includes the prototyping of (subsets of) such applications on reconfigurable and programmable platforms. For this application domain, the high computational complexity of the 5G Wireless Communication Systems calls for the design of high-performance and energy-efficient architectures. In **Wireless Sensor Networks** (WSN), where each wireless node is expected to operate without battery replacement for significant periods of time, energy consumption is the most important constraint. Sensor networks are a very dynamic domain of research due, on the one hand, to the opportunity to develop innovative applications that are linked to a specific environment, and on the other hand to the challenge of designing totally autonomous communicating objects.

Other important fields are also considered: hardware cryptographic and security modules, high-rate optical communications, machine learning, and multimedia processing.

5. Highlights of the Year

5.1. Highlights of the Year

5.1.1. *Awards*

Petr Dobias received the A. Richard Newton Young Fellow Award at IEEE/ACM Design Automation Conference (DAC), San Francisco, 2018.

Davide Pala received the A. Richard Newton Young Fellow Award at IEEE/ACM Design Automation Conference (DAC), San Francisco, 2018.

6. New Software and Platforms

6.1. Gecos

Generic Compiler Suite

KEYWORDS: Source-to-source compiler - Model-driven software engineering - Retargetable compilation

SCIENTIFIC DESCRIPTION: The Gecos (Generic Compiler Suite) project is a source-to-source compiler infrastructure developed in the Cairn group since 2004. It was designed to enable fast prototyping of program analysis and transformation for hardware synthesis and retargetable compilation domains.

Gecos is Java based and takes advantage of modern model driven software engineering practices. It uses the Eclipse Modeling Framework (EMF) as an underlying infrastructure and takes benefits of its features to make it easily extensible. Gecos is open-source and is hosted on the Inria gforge.

The Gecos infrastructure is still under very active development, and serves as a backbone infrastructure to projects of the group. Part of the framework is jointly developed with Colorado State University and between 2012 and 2015 it was used in the context of the FP7 ALMA European project. The Gecos infrastructure is currently used by the EMMTRIX start-up, a spin-off from the ALMA project which aims at commercializing the results of the project, and in the context of the H2020 ARGO European project.

FUNCTIONAL DESCRIPTION: GeCoS provides a programme transformation toolbox facilitating parallelisation of applications for heterogeneous multiprocessor embedded platforms. In addition to targeting programmable processors, GeCoS can regenerate optimised code for High Level Synthesis tools.

- Participants: Tomofumi Yuki, Thomas Lefeuvre, Imèn Fassi, Mickael Dardaillon, Ali Hassan El Moussawi and Steven Derrien
- Partner: Université de Rennes 1
- Contact: Steven Derrien
- URL: http://gecos.gforge.inria.fr

6.2. ID-Fix

Infrastructure for the Design of Fixed-point systems

KEYWORDS: Energy efficiency - Dynamic range evaluation - Accuracy optimization - Fixed-point arithmetic - Analytic Evaluation - Embedded systems - Code optimisation

SCIENTIFIC DESCRIPTION: The different techniques proposed by the team for fixed-point conversion are implemented on the ID.Fix infrastructure. The application is described with a C code using floating-point data types and different pragmas, used to specify parameters (dynamic, input/output word-length, delay operations) for the fixed-point conversion. This tool determines and optimizes the fixed-point specification and then, generates a C code using fixed-point data types (ac_fixed) from Mentor Graphics. The infrastructure is made-up of two main modules corresponding to the fixed-point conversion (ID.Fix-Conv) and the accuracy evaluation (ID.Fix-Eval).

FUNCTIONAL DESCRIPTION: ID.Fix focuses on computational precision accuracy and can provide an optimised specification using fixed point arithmetic from a C source code with floating point data types. Fixed point arithmetic is very widely used in embedded systems as it provides better performance and is much more energy efficient. ID.Fix used an analytic programme model which means it can explore more solutions and thereby produce much more efficient code.

- Participant: Olivier Sentieys
- Partner: Université de Rennes 1
- Contact: Olivier Sentieys
- URL: http://idfix.gforge.inria.fr

6.3. Platforms

6.3.1. Zyggie: a Wireless Body Sensor Network Platform

KEYWORDS: Health - Biomechanics - Wireless body sensor networks - Low power - Gesture recognition - Hardware platform - Software platform - Localization

SCIENTIFIC DESCRIPTION: Zyggie is a hardware and software wireless body sensor network platform. Each sensor node, attached to different parts of the human body, contains inertial sensors (IMU) (accelerometer, gyrometer, compass and barometer), an embedded processor and a low-power radio module to communicate data to a coordinator node connected to a computer, tablet or smartphone. One of the system's key innovations is that it collects data from sensors as well as on distances estimated from the power of the radio signal received to make the 3D location of the nodes more precise and thus prevent IMU sensor drift and power consumption overhead. Zyggie can be used to determine posture or gestures and mainly has applications in sport, healthcare and the multimedia industry.

FUNCTIONAL DESCRIPTION: The Zyggie sensor platform was developed to create an autonomous Wireless Body Sensor Network (WBSN) with the capabilities of monitoring body movements. The Zyggie platform is part of the BoWI project funded by CominLabs. Zyggie is composed of a processor, a radio transceiver and different sensors including an Inertial Measurement Unit (IMU) with 3-axis accelerometer, gyrometer, and magnetometer. Zyggie is used for evaluating data fusion algorithms, low power computing algorithms, wireless protocols, and body channel characterization in the BoWI project.

The Zyggie V2 prototype (see Figure 2) includes the following features: a 32-bit micro-controller to manage a custom MAC layer and process quaternions based on IMU measures, and an UWB radio from DecaWave to measure distances between nodes with Time of Flight (ToF).

- Participants: Arnaud Carer and Olivier Sentieys
- Partners: Lab-STICC, Université de Rennes 1
- Contact: Olivier Sentieys
- URL: https://bowi.cominlabs.u-bretagneloire.fr/zyggie-wbsn-platform

Figure 2. CAIRN's Ziggie platform for WBSN

6.3.2. E-methodHW: an automatic tool for the evaluation of polynomial and rational function approximations

KEYWORDS: function approximation, FPGA hardware implementation generator

SCIENTIFIC DESCRIPTION: E-methodHW is an open source C/C++ prototype tool written to exemplify what kind of numerical function approximations can be developed using a digit recurrence evaluation scheme for polynomials and rational functions.

FUNCTIONAL DESCRIPTION: E-methodHW provides a complete design flow from choice of mathematical function operator up to optimised VHDL code that can be readily deployed on an FPGA. The use of the E-method allows the user great flexibility if targeting high throughput applications.

- Participants: Silviu-Ioan Filip, Matei Istoan
- Partners: Université de Rennes 1, Imperial College London
- Contact: Silviu-Ioan Filip
- URL: https://github.com/sfilip/emethod

6.3.3. *Hybrid-DBT*

KEYWORDS: Dynamic Binary Translation, hardware acceleration, VLIW processor, RISC-V

SCIENTIFIC DESCRIPTION: Hybrid-DBT is a hardware/software Dynamic Binary Translation (DBT) framework capable of translating RISC-V binaries into VLIW binaries. Since the DBT overhead has to be as small as possible, our implementation takes advantage of hardware acceleration for performance critical stages (binary translation, dependency analysis and instruction scheduling) of the flow. Thanks to hardware acceleration, our implementation is two orders of magnitude faster than a pure software implementation and enable an overall performance improvements by 23% on average, compared to a native RISC-V execution.

- Participants: Simon Rokicki, Steven Derrien
- Partners: Université de Rennes 1
- URL: https://github.com/srokicki/HybridDBT

6.3.4. *Comet*

KEYWORDS: Processor core, RISC-V instruction-set architecture

SCIENTIFIC DESCRIPTION: Comet is a RISC-V pipelined processor with data/instruction caches, fully developed using High-Level Synthesis. The behavior of the core is defined in a small C code which is then fed into a HLS tool to generate the RTL representation. Thanks to this design flow, the C description can be used as a fast and cycle-accurate simulator, which behaves exactly like the final hardware. Moreover, modifications on the core can be done easily at the C level. Figure 3 depicts the place and route of a Comet core in a 28nm FDSOI technology.

- Participants: Simon Rokicki, Steven Derrien, Olivier Sentieys, Davide Pala, Joseph Paturel
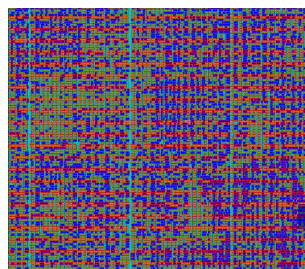- Partners: Université de Rennes 1
- URL: https://gitlab.inria.fr/srokicki/Comet

Figure 3. Floorplan of the Comet core (150μm×150μm in a 28nm FDSOI technology)

7. New Results

7.1. Reconfigurable Architecture Design

7.1.1. *Algorithmic Fault Tolerance for Timing Speculative Hardware*

Participants: Thibaut Marty, Tomofumi Yuki, Steven Derrien.

Timing speculation, also known as overclocking, is a well known approach to increase the computational throughput of processors and hardware accelerators. When used aggressively, timing speculation can lead to incorrect/corrupted results. As reported in the literature, timing errors can cause large numerical errors in the computation, and such occasional large errors can have devastating effect on the final output. The frequency of such errors depends on a number of factors, including the intensity of overclocking, operating temperature, voltage drops, variability within and across boards, input data, and so on. This makes it extremely difficult to determine a "safe" overclocking speed analytically or empirically. Several circuit-level error mitigating techniques have been proposed, but they are difficult to implement in modern FPGAs, and often involve significant area overhead. Instead of resorting to circuit level technique, we propose to rely on light-weight algorithm-level error detections techniques. This allows us to augment accelerators with low overhead mechanism to protect against timing errors, enabling aggressive timing speculation. We have demonstrated in [36] the validity of our approach for convolutional neural networks, where we use overclocking for the convolution stages. Our prototype on ZC706 demonstrated 68-77% computational throughput with negligible (<1%) area overhead.

7.1.2. *Adaptive Dynamic Compilation for Low-Power Embedded Systems*

Participants: Steven Derrien, Simon Rokicki.

Single ISA-Heterogeneous multi-cores such as the ARM big.LITTLE have proven to be an attractive solution to explore different energy/performance trade-offs. Such architectures combine Out of Order cores with smaller in-order ones to offer different power/energy profiles. They however do not really exploit the characteristics of workloads (compute-intensive vs. control dominated). In this work, we propose to enrich these architectures with VLIW cores, which are very efficient at compute-intensive kernels. To preserve the single ISA programming model, we resort to Dynamic Binary Translation as used in Transmeta Crusoe and NVidia Denver processors. Our proposed DBT framework targets the RISC-V ISA, for which both OoO and in-order implementations exist. Since DBT operates at runtime, its execution time is directly perceptible by the user, hence severely constrained. As a matter of fact, this overhead has often been reported to have a huge impact on actual performance, and is considered as being the main weakness of DBT based solutions. This is particularly true when targeting a VLIW processor: the quality of the generated code depends on efficient scheduling; unfortunately scheduling is known to be the most time-consuming component of a JIT compiler or DBT. Improving the responsiveness of such DBT systems is therefore a key research challenge. This is however made very difficult by the lack of open research tools or platform to experiment with such platforms. To address these issues, we have developed an open hardware/software platform supporting DBT. The platform was designed using HLS tools and validated on a FPGA board. The DBT uses RISC-V as host ISA, and can be retargeted to different VLIW configurations. Our platform uses custom hardware accelerators to improve the reactivity of our optimizing DBT flow. Our results [43], [27] show that, compared to a software implementation, our approach offers speed-up by $8\times$ while consuming $18\times$ less energy. We have also shown how our approach can be used to support runtime configurable VLIW cores. Such cores enable fine grain exploration of energy/performance trade-off by dynamically adjusting their number of execution slots, their register file size, etc. Our first experimental results have shown that this approach leads to best-case performance and energy efficiency when compared against static VLIW configurations [42].

7.1.3. *Hardware Accelerated Simulation of Heterogeneous Platforms*

Participants: Minh Thanh Cong, François Charot, Steven Derrien.

When considering designing heterogeneous multi-core platforms, the number of possible design combinations leads to a huge design space, with subtle trade-offs and design interactions. To reason about what design is best for a given target application requires detailed simulation of many different possible solutions. Simulation frameworks exist (such as gem5) and are commonly used to carry out these simulations. Unfortunately, these are purely software-based approaches and they do not allow a real exploration of the design space. Moreover, they do not really support highly heterogeneous multi-core architectures. These limitations motivate the study of the use of hardware to accelerate the simulation, and in particular of FPGA components. In this context, we are currently investigating the possibility of building hardware accelerated simulators of heterogeneous multicore architectures using the HAsim/LEAP infrastructure. Two aspects are currently under development. The first one concerns the deployment of simulator models on the hybrid Xeon CPU-Arria 10 FPGA Intel platforms. The second one concerns the definition of simulation models of hardware accelerators. The core processor brick is a RISCV core.

7.1.4. *Dynamic Fault-Tolerant Scheduling onto Multi-Core Systems*

Participants: Emmanuel Casseau, Petr Dobias.

Demand on multi-processor systems for high performance and low energy consumption still increases in order to satisfy our requirements to perform more and more complex computations. Moreover, the transistor size gets smaller and their operating voltage is lower, which goes hand in glove with higher susceptibility to system failure. In order to ensure system functionality, it is necessary to conceive fault-tolerant systems. Temporal and/or spatial redundancy is currently used to tackle this issue. Actually, multi-processor platforms can be less vulnerable when one processor is faulty because other processors can take over its scheduled tasks. In this context, we investigate how to dynamically map and schedule tasks onto homogeneous faulty processors. We developed several run-time algorithms based on the primary/backup approach which is commonly used for its minimal resources utilization and high reliability [31], [30]. The aim of our work was to reduce the complexity of the algorithm in order to target real-time embedded systems without sacrificing reliability. This work is done in collaboration with Oliver Sinnen, PARC Lab., the University of Auckland.

7.1.5. *Run-Time Management on Multicore Platforms*

Participant: Angeliki Kritikakou.

In real-time mixed-critical systems, Worst-Case Execution Time analysis (WCET) is required to guarantee that timing constraints are respected at least for high criticality tasks. However, the WCET is pessimistic compared to the real execution time, especially for multicore platforms. As WCET computation considers the worst-case scenario, it means that whenever a high criticality task accesses a shared resource in multi-core platforms, it is considered that all cores use the same resource concurrently. This pessimism in WCET computation leads to a dramatic under utilization of the platform resources, or even failing to meet the timing constraints. In order to increase resource utilization while guaranteeing real-time guarantees for high criticality tasks, previous works proposed a run-time control system to monitor and decide when the interferences from low criticality tasks cannot be further tolerated. However, in the initial approaches, the points where the controller is executed were statically predefined. We propose a dynamic run-time control in [21] which adapts its observations to on-line temporal properties, increasing further the dynamism of the approach, and mitigating the unnecessary overhead implied by existing static approaches. Our dynamic adaptive approach allows to control the ongoing execution of tasks based on run-time information, and increases further the gains in terms of resource utilization compared with static approaches.

7.1.6. *Energy Constrained and Real-Time Scheduling and Assignment on Multicores*

Participants: Olivier Sentieys, Angeliki Kritikakou, Lei Mo.

Multicore architectures have been used to enhance computing capabilities, but the energy consumption is still an important concern. Embedded application domains usually require less accurate, but always in-time, results. Imprecise Computation (IC) can be used to divide a task into a mandatory subtask providing a baseline Quality-of-Service (QoS) and an optional subtask that further increases the baseline QoS. Combining dynamic voltage and frequency scaling, task allocation and task adjustment, we can maximize the system

QoS under real-time and energy supply constraints. However, the nonlinear and combinatorial nature of this problem makes it difficult to solve. In [25], we formulate a Mixed-Integer Non-Linear Programming (MINLP) problem to concurrently carry out task-to-processor allocation, frequency-to-task assignment and optional task adjustment. We provide a Mixed-Integer Linear Programming (MILP) form of this formulation without performance degradation and we propose a novel decomposition algorithm to provide an optimal solution with reduced computation time compared to state-of-the-art optimal approaches (22.6% in average). We also propose a heuristic version that has negligible computation time. In [24], we focus on QoS maximizing for dependent IC-tasks under real-time and energy constraints. Compared with existing approaches, we consider the joint-design problem, where task-to-processor allocation, frequency-to-task assignment, task scheduling and task adjustment are optimized simultaneously. The joint-design problem is formulated as an NP-hard Mixed-Integer Non-Linear Programming and it is safely transformed to a Mixed-Integer Linear Programming (MILP) without performance degradation. Two methods (basic and accelerated version) are proposed to find the optimal solution to MILP problem. They are based on problem decomposition and provide a controllable way to trade-off the quality of the solution and the computational complexity. The optimality of the proposed methods is proved rigorously, and the experimental results show reduced computation time (23.7% in average) compared with existing optimal methods. Finally, in [50] we summarize the problem and the methods for imprecise computation task mapping on multicore Wireless Sensor Networks.

7.1.7. *Real-Time Energy-Constrained Scheduling in Wireless Sensor and Actuator Networks*
Participants: Angeliki Kritikakou, Lei Mo.

Wireless Sensor and Actuator Networks (WSANs) are emerging as a new generation of Wireless Sensor Networks (WSNs). Due to the coupling between the sensing areas of the sensors and the action areas of the actuators, the efficient coordination among the nodes is a great challenge. In our work in [23] we address the problem of distributed node coordination in WSANs aiming at meeting the user's requirements on the states of the Points of Interest (POIs) in a real-time and energy-efficient manner. The node coordination problem is formulated as a non-linear program. To solve it efficiently, the problem is divided into two correlated subproblems: the Sensor-Actuator (S-A) coordination and the Actuator-Actuator (A-A) coordination. In the S-A coordination, a distributed federated Kalman filter-based estimation approach is applied for the actuators to collaborate with their ambient sensors to estimate the states of the POIs. In the A-A coordination, a distributed Lagrange-based control method is designed for the actuators to optimally adjust their outputs, based on the estimated results from the S-A coordination. The convergence of the proposed method is proved rigorously. As the proposed node coordination scheme is distributed, we find the optimal solution while avoiding high computational complexity. The simulation results also show that the proposed distributed approach is an efficient and practically applicable method with reasonable complexity. In addition, the design of fast and effective coordination among sensors and actuators in Cyber-Physical Systems (CPS) is a fundamental, but challenging issue, especially when the system model is a priori unknown and multiple random events can simultaneously occur. In [37], we propose a novel collaborative state estimation and actuator scheduling algorithm with two phases. In the first phase, we propose a Gaussian Mixture Model (GMM)-based method using the random event physical field distribution to estimate the locations and the states of events. In the second phase, based on the number of identified events and the number of available actuators, we study two actuator scheduling scenarios and formulate them as Integer Linear Programming (ILP) problems with the objective to minimize the actuation delay. We validate and demonstrate the performance of the proposed scheme through both simulations and physical experiments for a home temperature control application.

7.1.8. *Real-Time Scheduling of Reconfigurable Battery-Powered Multi-Core Platforms*
Participants: Daniel Chillet, Aymen Gammoudi.

Reconfigurable real-time embedded systems are constantly increasingly used in applications like autonomous robots or sensor networks. Since they are powered by batteries, these systems have to be energy-aware, to adapt to their environment, and to satisfy real-time constraints. For energy-harvesting systems, regular recharges of battery can be estimated. By including this parameter in the operating system, it is then possible to develop some strategy able to ensure the best execution of the application until the next recharge. In this

context, operating system services must control the execution of tasks to meet the application constraints. Our objective concerns the proposition of a new real-time scheduling strategy that considers execution constraints such as the deadline of tasks and the energy for heterogeneous architectures. For such systems, we first addressed homogeneous architectures including P identical cores. We assumed that they can be reconfigured dynamically by authorizing the addition and/or removal of periodic tasks and each core schedules its local tasks by using the EDF algorithm [20]. This work is extended to address heterogeneous systems for which each task has different execution parameters. The objective of this extension work is to develop a new strategy for mapping N tasks to P heterogeneous cores of a given distributed system [32]. For these two architectures models, we formulated the problem as an Integer Linear Program (ILP) optimization problem. Assuming that the energy consumed by the communication is dependent on the distance between cores, we proposed a mapping strategy to minimize the total cost of communication between cores by placing the dependent tasks as close as possible to each other. The proposed strategy guarantees that, when a task is mapped into the system and accepted, it is then correctly executed prior to the task deadline. Finally, as on-line scheduling is targeted for this work, we proposed heuristics to solve these problems in efficient way. These heuristics are based on the previous packing strategy developed for the mono-core architecture case. Experimental results reveal the effectiveness of the proposed strategy by comparing the derived heuristics with the optimal ones, obtained by solving an ILP problem

7.1.9. *Improving the Reliability of Wireless NoC*
Participants: Olivier Sentieys, Joel Ortiz Sosa.

Wireless Network-on-Chip (WiNoC) is one of the most promising solutions to overcome multi-hop latency and high power consumption of modern many/multi core System-on- Chip (SoC). However, the design of efficient wireless links faces challenges to overcome multi-path propagation present in realistic WiNoC channels. In order to alleviate such channel effect, we propose a Time-Diversity Scheme (TDS) to enhance the reliability of on-chip wireless links using a semi-realistic channel model in [45]. First, we study the significant performance degradation of state-of-the-art wireless transceivers subject to different levels of multi-path propagation. Then we investigate the impact of using some channel correction techniques adopting standard performance metrics. Experimental results show that the proposed Time-Diversity Scheme significantly improves Bit Error Rate (BER) compared to other techniques. Moreover, our TDS allows for wireless communication links to be established in conditions where this would be impossible for standard transceiver architectures. Results on the proposed complete transceiver, designed using a 28-nm FDSOI technology, show a power consumption of 0.63mW at 1.0V and an area of 317 μm^2. Full channel correction is performed in one single clock cycle.

7.1.10. *Optical Network-on-Chip (ONoC) for 3D Multiprocessor Architectures*
Participants: Jiating Luo, Van Dung Pham, Cédric Killian, Daniel Chillet, Olivier Sentieys.

Photonics on silicon is now a technology that offers real opportunities in the context of multiprocessor interconnect. The optical medium can support multiple transactions at the same time on different wavelengths by using Wavelength Division Multiplexing (WDM). Moreover, multiple wavelengths can be gathered as high-bandwidth channel to reduce transmission latency. However, multiple signals sharing simultaneously a waveguide lead to inter-channel crosstalk noise. This problem impacts the Signal to Noise Ratio (SNR) of the optical signal, which increases the Bit Error Rate (BER) at the receiver side. We formulated the crosstalk noise and latency models and then proposed a Wavelength Allocation (WA) method in a ring-based WDM ONoC to reach performance and energy trade-offs based on application constraints. We show that for a 16-cluster ONoC architecture using 12 wavelengths, more than 10^5 allocation solutions exist and only 51 are on a Pareto front giving a tradeoff between latency and energy per bit derived from the BER. These optimized solutions reduce the execution time of the application by 37% and the energy from 7.6fJ/bit to 4.4fJ/bit. In [22], we define high-level mechanisms which can handle wavelength allocation protocol of the communication medium for each data transfer between tasks. Indeed, the optical wavelengths are a shared resource between all the electrical computing clusters and are allocated at run time according to application needs and quality of service. We produce the communication configurations which are defined by the number of wavelengths for each communication, the level of quality for the communications, and the laser power levels. In [35], we define

an Optical-Network-Interface (ONI) to connect a cluster of processors to the optical communication medium. This interface, constrained by the 10 Gb/s data-rate of the lasers, integrates Error Correcting Codes (ECC), laser drivers, and a communication manager. The ONI can select, at run-time, the communication mode to use depending on performance, latency or power constraints. The use of ECC is based on redundant bits which increases the transmission time, but saves power for a given Bit Error Rate (BER). Furthermore, the use of several wavelengths in parallel reduces latency and increases bandwidth, but also increases communication loss.

7.2. Compilation and Synthesis for Reconfigurable Platform

7.2.1. *Compile Time Simplification of Sparse Matrix Code Dependences*
Participant: Tomofumi Yuki.

Analyzing array-based computations to determine data dependences is useful for many applications including automatic parallelization, race detection, computation and communication overlap, verification, and shape analysis. For sparse matrix codes, array data dependence analysis is made more difficult by the use of index arrays that make it possible to store only the nonzero entries of the matrix (e.g., in $A[B[i]]$, B is an index array). Here, dependence analysis is often stymied by such indirect array accesses due to the values of the index array not being available at compile time. Consequently, many dependences cannot be proven unsatisfiable or determined until runtime. Nonetheless, index arrays in sparse matrix codes often have properties such as monotonicity of index array elements that can be exploited to reduce the amount of runtime analysis needed. In this work, we contribute a formulation of array data dependence analysis that includes encoding index array properties as universally quantified constraints. This makes it possible to leverage existing SMT solvers to determine whether such dependences are unsatisfiable and significantly reduces the number of dependences that require runtime analysis in a set of eight sparse matrix kernels. Another contribution is an algorithm for simplifying the remaining satisfiable data dependences by discovering equalities and/or subset relationships. These simplifications are essential to make a runtime-inspection-based approach feasible.

7.2.2. *Automatic Parallelization Techniques for Time-Critical Systems*
Participants: Steven Derrien, Mickael Dardaillon.

Real-time systems are ubiquitous, and many of them play an important role in our daily life. In hard real-time systems, computing the correct results is not the only requirement. In addition, the results must be produced within pre-determined timing constraints, typically deadlines. To obtain strong guarantees on the system temporal behavior, designers must compute upper bounds of the Worst-Case Execution Times (WCET) of the tasks composing the system. WCET analysis is confronted with two challenges: (i) extracting knowledge of the execution flow of an application from its machine code, and (ii) modeling the temporal behavior of the target platform. Multi-core platforms make the latter issue even more challenging, as interference caused by concurrent accesses to shared resources have also to be modeled. Accurate WCET analysis is facilitated by *predictable* hardware architectures. For example, platforms using ScratchPad Memories (SPMs) instead of caches are considered as more predictable. However SPM management is left to the programmer-managed, making them very difficult to use, especially when combined with complex loop transformations needed to enable task level parallelization. Many researches have studied how to combine automatic SPM management with loop parallelization at the compiler level. It has been shown that impressive average-case performance improvements could be obtained on compute intensive kernels, but their ability to reduce WCET estimates remains to be demonstrated, as the transformed code does not lend itself well to WCET analysis.

In the context of the ARGO project, and in collaboration with members of the PACAP team, we have studied how parallelizing compilers techniques should be revisited in order to help WCET analysis tools. More precisely, we have demonstrated the ability of polyhedral optimization techniques to reduce WCET estimates in the case of sequential codes, with a focus on locality improvement and array contraction. We have shown on representative real-time image processing use cases that they could bring significant improvements of WCET estimates (up to 40%) provided that the WCET analysis process is guided with automatically generated flow

annotations [34]. Our current research direction aims [41] at studying the impact of compiler optimization on WCET estimates, and develop specific WCET aware compiler optimization flows. More specifically, we explore the use of iterative compilation (WCET-directed program optimization to explore the optimization space), with the objective to (i) allow flow facts to be automatically found and (ii) select optimizations that result in the lowest WCET estimates. We also explore to which extent code outlining helps, by allowing the selection of different optimization options for different code snippets of the application.

7.2.3. Design of High Throughput Mathematical Function Evaluators
Participant: Silviu Ioan Filip.

The evaluation of mathematical functions is a core component in many computing applications and has been a core topic in computer arithmetic since the inception of the field. In [28], we proposed an automatic method for the evaluation of functions via polynomial or rational approximations and its hardware implementation, on FPGAs. These approximations are evaluated using Ercegovac's iterative E-method adapted for FPGA implementation. The polynomial and rational function coefficients are optimized such that they satisfy the constraints of the E-method. It allows for an effective way to perform design space exploration when targeting high throughput.

7.2.4. Robust Tools for Computing Rational Chebyshev Approximations
Participant: Silviu Ioan Filip.

Rational functions are useful in a plethora of applications, including digital signal processing and model order reduction. They are nevertheless known to be much harder to work with in a numerical context than other, potentially less expressive families of approximating functions, like polynomials. In [19] we have proposed the use of a numerically robust way of representing rational functions, the barycentric form (*i.e.,* a ratio of partial fractions sharing the same poles). We use this form to develop scalable iterative algorithms for computing rational approximations to functions which minimize the uniform norm error. Our results are shown to significantly outperform previous state of the art approaches.

8. Bilateral Contracts and Grants with Industry

8.1. Bilateral Contracts with Industry

Collaboration with Huawei Technologies, Sophia Antipolis: In the context of Image Signal Processing (ISP), the project aims at building a proof of concept of an environment able to automatically optimize the precision of every operator (fixed-point or floating-point arithmetic) in a complex, multi-kernel algorithm and find the best tradeoff between cost/power and image quality.

9. Partnerships and Cooperations

9.1. National Initiatives

9.1.1. Labex CominLabs - 3DCORE (2014-2018)
Participants: Olivier Sentieys, Daniel Chillet, Cédric Killian, Jiating Luo, Van Dung Pham.

3DCORE (3D Many-Core Architectures based on Optical Network on Chip) is a project investigating new solutions based on silicon photonics to enhance by 2 to 3 magnitude orders energy efficiency and data rate of on-chip interconnect in the context of a many-core architecture. Moreover, 3DCore will take advantage of 3D technologies to design a specific optical layer suitable for a flexible and energy efficient high-speed optical network on chip (ONoC). 3DCORE involves CAIRN, FOTON (Rennes, Lannion) and Institut des Nanotechnologies de Lyon. For more details see https://3d-opt-many-cores.cominlabs.u-bretagneloire.fr.

9.1.2. Labex CominLabs - RELIASIC (2014-2018)

Participant: Emmanuel Casseau.

RELIASIC (Reliable Asic) will address the issue of fault-tolerant computation with a bottom-up approach, starting from an existing application as a use case (a GPS receiver) and adding some redundant mechanisms to allow the GPS receiver to be tolerant to transient errors due to low voltage supply. RELIASIC involves CAIRN, Lab-STICC (Lorient) and IETR (Rennes, Nantes). In this project, CAIRN is in charge of the analysis and design of arithmetic operators for fault tolerance. We focus on the hardware implementations of conventional arithmetic operators such as adders, multipliers. We also propose a lightweight design and assessment framework for arithmetic operators with reduced-precision redundancy. For more details see https://reliasic. cominlabs.u-bretagneloire.fr

9.1.3. Labex CominLabs & Lebesgue - H-A-H (2014-2018)

Participants: Arnaud Tisserand, Gabriel Gallin, Audrey Lucas.

H-A-H for *Hardware and Arithmetic for Hyperelliptic Curves Cryptography* is a project on advanced arithmetic representation and algorithms for hyper-elliptic curve cryptography. It will provide novel implementations of HECC based cryptographic algorithms on custom hardware platforms. H-A-H involves CAIRN (Lannion) and IRMAR (Rennes). For more details see http://h-a-h.inria.fr/.

9.1.4. Labex CominLabs - BBC (2016-2020)

Participants: Olivier Sentieys, Cédric Killian, Joel Ortiz Sosa.

The aim of the BBC (on-chip wireless Broadcast-Based parallel Computing) project is to evaluate the use of wireless links between cores inside chips and to define new paradigms. Using wireless communications enables broadcast capabilities for Wireless Networks on Chip (WiNoC) and new management techniques for memory hierarchy and parallelism. The key objectives concern improvement of power consumption, estimation of achievable data rates, flexibility and reconfigurability, size reduction and memory hierarchy management. In this project, CAIRN will address new low-power MAC (media access control) technique based on CDMA access as well as broadcast-based fast cooperation protocol designed for resource sharing (bandwidth, distributed memory, cache coherency) and parallel programming. For more details see https://bbc. cominlabs.u-bretagneloire.fr

9.1.5. Labex CominLabs - SHERPAM (2014-2018)

Participant: Patrice Quinton.

Heart failure and peripheral artery disease patients require early detection of health problems in order to prevent major risk of morbidity and mortality. Evidence shows that people recover from illness or cope with a chronic condition better if they are in a familiar environment (i.e., at home) and if they are physically active (i.e., practice sports). The goal of the Sherpam project is to design, implement, and validate experimentally a monitoring system allowing biophysical data of mobile subjects to be gathered and exploited in a continuous flow. Transmission technologies available to mobile users have been improved a lot during the last two decades, and such technologies offer interesting prospects for monitoring the health of people anytime and anywhere. The originality of the Sherpam project is to rely simultaneously and in an agile way on several kinds of wireless networks in order to ensure the transmission of biometric data, while coping with network disruptions. Sherpam also develops new signal processing algorithms for activity quantification and recognition which represent now a major social and public health issue (monitoring of elderly patient, personalized quantification activity, etc.). Sherpam involves research teams from several scientific domains and from several laboratories of Brittany (IRISA/CASA, LTSI, M2S, CIC-IT 1414-CHU Rennes and LAUREPS). For more details see https://sherpam.cominlabs.u-bretagneloire.fr

9.1.6. DGA RAPID - FLODAM (2017–2021)

Participants: Olivier Sentieys, Angeliki Kritikakou.

FLODAM is an industrial research project for methodologies and tools dedicated to the hardening of embedded multi-core processor architectures. The goal is to: 1) evaluate the impact of the natural or artificial environments on the resistance of the system components to faults based on models that reflect the reality of the system environment, 2) the exploration of architecture solutions to make the multi-core architectures fault tolerant to transient or permanent faults, and 3) test and evaluate the proposed fault tolerant architecture solutions and compare the results under different scenarios provided by the fault models. For more details see https://flodam.fr

9.2. European Initiatives

9.2.1. H2020 ARGO

Participants: Steven Derrien, Olivier Sentieys, Mickael Dardaillon, Ali Hassan El Moussawi.

> Program: H2020-ICT-04-2015
> Project acronym: ARGO
> Project title: WCET-Aware Parallelization of Model-Based Applications for Heterogeneous Parallel Systems
> Duration: Feb. 2016 - Feb. 2019
> Coordinator: KIT
> Other partners: KIT (Germany), UR1/Inria/CAIRN, Recore Systems (Netherlands), TEI-WG (Greece), Scilab Ent. (France), Absint (Ger.), DLR (Ger.), Fraunhofer (Ger.)

Increasing performance and reducing cost, while maintaining safety levels and programmability are the key demands for embedded and cyber-physical systems, e.g. aerospace, automation, and automotive. For many applications, the necessary performance with low energy consumption can only be provided by customized computing platforms based on heterogeneous many-core architectures. However, their parallel programming with time-critical embedded applications suffers from a complex toolchain and programming process. ARGO will address this challenge with a holistic approach for programming heterogeneous multi- and many-core architectures using automatic parallelization of model-based real-time applications. ARGO will enhance WCET-aware automatic parallelization by a cross-layer programming approach combining automatic tool-based and user-guided parallelization to reduce the need for expertise in programming parallel heterogeneous architectures. The ARGO approach will be assessed and demonstrated by prototyping comprehensive time-critical applications from both aerospace and industrial automation domains on customized heterogeneous many-core platforms.

9.2.2. ANR International ARTEFaCT

Participants: Olivier Sentieys, Van Phu Ha, Tomofumi Yuki.

> Program: ANR International France-Switzerland
> Project acronym: ARTEFaCT
> Project title: AppRoximaTivE Flexible Circuits and Computing for IoT
> Duration: Feb. 2016 - Dec. 2019
> Coordinator: CEA
> Other partners: CEA-LETI, CAIRN, EPFL

The ARTEFaCT project aims to build on the preliminary results on inexact and exact near-threshold and sub-threshold circuit design to achieve major energy consumption reductions by enabling adaptive accuracy control of applications. ARTEFaCT proposes to address, in a consistent fashion, the entire design stack, from physical hardware design, up to software application analysis, compiler optimizations, and dynamic energy management. We do believe that combining sub-near-threshold with inexact circuits on the hardware side and, in addition, extending this with intelligent and adaptive power management on the software side will produce outstanding results in terms of energy reduction, i.e., at least one order of magnitude, in IoT applications. The project will contribute along three research directions: (1) approximate, ultra low-power circuit design, (2) modeling and analysis of variable levels of computation precision in applications, and (3) accuracy-energy trade- offs in software.

9.3. International Initiatives

9.3.1. Inria International Labs

EPFL-Inria
Associate Team involved in the International Lab:

9.3.1.1. IoTA

Title: Ultra-Low Power Computing Platform for IoT leveraging Controlled Approximation

International Partner (Institution - Laboratory - Researcher):

Ecole Polytechnique Fédérale de Lausanne (Switzerland) - Christian Enz

Start year: 2017

See also: https://team.inria.fr/cairn/IOTA

Energy issues are central to the evolution of the Internet of Things (IoT), and more generally to the ICT industry. Current low-power design techniques cannot support the estimated growth in number of IoT objects and at the same time keep the energy consumption within sustainable bounds, both on the IoT node side and on cloud/edge-cloud side. This project aims to build on the preliminary results on inexact and exact sub/near-threshold circuit design to achieve major energy consumption reductions by enabling adaptive accuracy control of applications. IoTA proposes to address, in a consistent fashion, the entire design stack, from hardware design, up to software application analysis, compiler optimizations, and dynamic energy management. The main scientific challenge is twofold: (1) to add adaptive accuracy to hardware blocks built in near/sub threshold technology and (2) to provide the tools and methods to program and make efficient use of these hardware blocks for applications in the IoT domain. This entails developing approximate computing units, on one side, and methods and tools, on the other side, to rigorously explore trade-offs between accuracy and energy consumption in IoT systems. The expertise of the members of the two teams is complementary and covers all required technical knowledge necessary to reach our objectives, i.e., ultra low power hardware design (EPFL), approximate operators and functions (Inria, EPFL), formal analysis of precision in algorithms (Inria), and static and dynamic energy management (Inria, EPFL). Finally, the proof of concept will consist of results on (1) an adaptive, inexact or exact, ultra-low power microprocessor in 28 nm process and (2) a real prototype implemented in an FPGA platform combining processors and hardware accelerators. Several software use-cases relevant for the IoT domain will be considered, e.g., embedded vision, IoT sensors data fusion, to practically demonstrate the benefits of our approach.

9.3.2. Inria International Partners

9.3.2.1. LRS

Title: Loop unRolling Stones: compiling in the polyhedral model

International Partner (Institution - Laboratory - Researcher):

Colorado State University (United States) - Department of Computer Science - Prof. Sanjay Rajopadhye

9.3.2.2. HARAMCOP

Title: Hardware accelerators modeling using constraint-based programming

International Partner (Institution - Laboratory - Researcher):

Lund University (Sweden) - Department of Computer Science - Prof. Krzysztof Kuchcinski

9.3.2.3. SPINACH

Title: Secure and low-Power sensor Networks Circuits for Healthcare embedded applications

International Partner (Institution - Laboratory - Researcher):

University College Cork (Ireland) - Department of Electrical and Electronic Engineering - Prof. Liam Marnane and Prof. Emanuel Popovici

Arithmetic operators for cryptography, side channel attacks for security evaluation, energy-harvesting sensor networks, and sensor networks for health monitoring.

9.3.2.4. DARE

Title: Design space exploration Approaches for Reliable Embedded systems

International Partner (Institution - Laboratory - Researcher):

IMEC (Belgium) - Francky Catthoor

Methodologies to design low cost and efficient techniques for safety-critical embedded systems, Design Space Exploration (DSE), run-time dynamic control mechanisms.

9.3.2.5. Informal International Partners

LSSI laboratory, Québec University in Trois-Rivières (Canada), Design of architectures for digital filters and mobile communications.

Department of Electrical and Computer Engineering, University of Patras (Greece), Wireless Sensor Networks, Worst-Case Execution Time, Priority Scheduling.

Karlsruhe Institute of Technology - KIT (Germany), Loop parallelization and compilation techniques for embedded multicores.

Ruhr - University of Bochum - RUB (Germany), Reconfigurable architectures.

University of Science and Technology of Hanoi (Vietnam), Participation of several CAIRN's members in the Master ICT / Embedded Systems.

9.4. International Research Visitors

9.4.1. Visits of International Scientists

Martin Kumm, University of Kassel, Germany, July 2018.

Son Tran Giang, Lecturer at ICTLab, Vietnam, December 2018.

9.4.2. Visits to International Teams

E. Casseau spent 3 weeks as a visiting researcher in the Parallel and Reconfigurable Lab. of the Electrical and Computer Engineering department, the University of Auckland, New Zealand, in December 2018.

P. Dobias (Phd student) spent 5 months in the Parallel and Reconfigurable Lab. of the Electrical and Computer Engineering department, the University of Auckland, New Zealand, from November 2018 until March 2019.

10. Dissemination

10.1. Promoting Scientific Activities

10.1.1. Scientific Events Organisation

10.1.1.1. General Chair, Scientific Chair

- E. Casseau was General Co-Chair of DASIP, Conference on Design and Architectures for Signal and Image Processing, in Porto, Portugal, October 10-12, 2018.
- D. Chillet was General Chair of 10th Workshop on Rapid Simulation and Performance Evaluation: Methods and Tools (RAPIDO), Manchester, United Kingdom, January 22-24, 2018.

10.1.1.2. Member of the Organizing Committees

- E. Casseau is a member of DASIP Steering Committee, Conference on Design and Architectures for Signal and Image Processing.

10.1.2. Scientific Events Selection

10.1.2.1. Chair of Conference Program Committees

- O. Sentieys was Track Chair at IEEE NEWCAS and Co-Chair of the D8 Track on Architectural and Microarchitectural Design at IEEE/ACM DATE.

10.1.2.2. Member of the Conference Program Committees

- D. Chillet was member of the technical program committee of HiPEAC RAPIDO, HiPEAC WRC, MCSoC, DCIS, ComPAS, DASIP, LP-EMS, ARC.
- S. Derrien was a member of technical program committee of IEEE FPL, IEEE FPT and ARC.
- A. Kritikakou was a member of technical program committee of IEEE RTAS, ECRTS, SAMOS.
- O. Sentieys was a member of technical program committee of IEEE/ACM DATE, IEEE FPL, ACM ENSSys, ACM SBCCI, IEEE ReConFig, CROWNCOM.
- T. Yuki was a member of technical program committee of CGO conference and of Impact workshop.

10.1.3. Member of the Editorial Boards

- D. Chillet is member of the Editor Board of Journal of Real-Time Image Processing (JRTIP).
- O. Sentieys is member of the editorial board of Journal of Low Power Electronics.

10.1.4. Invited Talks

- D.Chillet gave an invited talk at FETCH (École d'hiver Francophone sur les Technologies de Conception des Systèmes embarqués Hétérogènes), Saint Malo, France, January 2018 on "Gestion des fautes au niveau tâche pour architectures MPSoC et Reconfigurables - Aspects multiprocesseur et reconfiguration dynamique".
- C. Killian gave an invited talk at FETCH (École d'hiver Francophone sur les Technologies de Conception des Systèmes embarqués Hétérogènes), Saint Malo, France, January 2018 on "Energy-performance tradeoffs in optical Network-on-Chips".
- C. Killian gave an invited talk at OPTICS (4th International Workshop on Optical/Photonic Interconnects for Computing Systems), in conjunction with IEEE/ACM Design Automation and Test in Europe (DATE), Dresden, Germany, march 2018 on "Offline optimization of wavelength allocation and laser to deal with Energy-Performance tradeoffs in nanophotonic interconnects".
- C. Killian gave an invited talk at a thematic day *Photonique sur silicium pour les architectures de calcul* organized by GDR SoC^2, Lyon, France, November 2018 on "Digital architectures to enhance Optical NoCs efficiency".
- O. Sentieys gave an invited talk at FETCH (École d'hiver Francophone sur les Technologies de Conception des Systèmes embarqués Hétérogènes), Saint Malo, France, January 2018 on "Playing with number representations for energy efficiency: an introduction to approximate and stochastic computing".
- O. Sentieys gave a Keynote at the Third Workshop on Approximate Computing (AxC), in conjunction with IEEE European Test Symposium (ETS), Bremen, Germany, June 2018 on "Playing with number representations and operator-level approximations" [59].
- O. Sentieys gave a tutorial at the Embedded Systems Week (ESWEEK), September 2018 on "A Comprehensive Analysis of Approximate Computing Techniques: From Component- to Application-Level" [55].
- T. Yuki gave an invited talk at TAPAS Workshop, Freiburg im Breisgau, Germany, August 2018 on "Polyhedral Static Analysis for the X10 Language".

10.1.5. Leadership within the Scientific Community

- E. Casseau is a member of the French National University Council in Signal Processing and Electronics (CNU - Conseil National des Universites, 61ème section) since 2018.

- D. Chillet is member of the Board of Directors of Gretsi Association.
- D. Chillet is co-animator of the topics "Connected Objects" and "Near Sensor Computing" of GDR SoC[2].
- F. Charot and O. Sentieys are members of the steering committee of a CNRS Spring School for graduate students on embedded systems architectures and associated design tools (ARCHI).
- O. Sentieys is a member of the steering committee of a CNRS spring school for graduate students on low-power design (ECOFAC).
- O. Sentieys is a member of the steering committee of GDR SoC[2].

10.1.6. Scientific Expertise

- O. Sentieys served as a jury member in the EDAA Outstanding Dissertations Award (ODA).

10.2. Teaching - Supervision - Juries

10.2.1. Teaching Responsibilities

- C. Wolinski is the Director of ESIR.
- O. Sentieys is responsible of the "Embedded Systems" major of the SISEA Master by Research.
- D. Chillet is the responsible of the ICT Master of University of Science and Technology of Hanoi.
- C. Killian is the responsible of the second year of the Physical Measurement DUT at IUT, Lannion.

ENSSAT stands for *"École Nationale Supérieure des Sciences Appliquées et de Technologie"* and is an *"École d'Ingénieurs"* of the University of Rennes 1, located in Lannion. ISTIC is the Electrical Engineering and Computer Science Department of the University of Rennes 1. ESIR stands for *"École supérieure d'ingénieur de Rennes"* and is an *"École d'Ingénieurs"* of the University of Rennes 1, located in Rennes.

10.2.2. Teaching

E. Casseau: signal processing, 21h, ENSSAT (L3)

E. Casseau: low power design, 6h, ENSSAT (M1)

E. Casseau: real time design methodology, 57h, ENSSAT (M1)

E. Casseau: computer architecture, 24h, ENSSAT (M1)

E. Casseau: VHDL design, 42h, ENSSAT (M1)

E. Casseau: SoC and high-level synthesis, 33h, Master by Research (SISEA) and ENSSAT (M2)

S. Derrien, optimizing and parallelising compilers, 14h, Master of Computer Science, ISTIC(M2)

S. Derrien, advanced processor architectures, 8h, Master of Computer Science, ISTIC(M2)

S. Derrien, high level synthesis, 20h, Master of Computer Science, ISTIC(M2)

S. Derrien, computer science research projects, 10h, Master of Computer Science, ISTIC(M1)

S. Derrien: introduction to operating systems, 8h, ISTIC (M1)

S. Derrien, principles of digital design, 20h, Bachelor of EE/CS, ISTIC(L2)

S. Derrien, computer architecture, 48h, Bachelor of Computer Science, ISTIC(L3)

F. Charot: computer architectures, 16h, ESIR (L3)

D. Chillet: embedded processor architecture, 20h, ENSSAT (M1)

D. Chillet: multimedia processor architectures, 24h, ENSSAT (M2)

D. Chillet: low-power digital CMOS circuits, 6h, Telecom Bretagne (M2)

C. Killian: digital electronics, 62h, IUT Lannion (L1)

C. Killian: signal processing, 36h, IUT Lannion (L2)

C. Killian: automated measurements, 56h, IUT Lannion (L2)

C. Killian: measurement chain, 58h, IUT Lannion (L2)

C. Killian: embedded systems programming, 12h, IUT Lannion (L2)

C. Killian: automatic control, 18h, IUT Lannion (L2)

A. Kritikakou: computer architecture 1, 32h, ISTIC (L3)

A. Kritikakou: computer architecture 2, 44h, ISTIC (L3)

A. Kritikakou: C and unix programming languages, 102h, ISTIC (L3)

A. Kritikakou: operating systems, 96h, ISTIC (L3)

A. Kritikakou: multitasking operating systems, 20h, ISTIC (M1)

O. Sentieys: VLSI integrated circuit design, 24h, ENSSAT (M1)

O. Sentieys: VHDL and logic synthesis, 18h, ENSSAT (M1)

C. Wolinski: computer architectures, 92h, ESIR (L3)

C. Wolinski: design of embedded systems, 48h, ESIR (M1)

C. Wolinski: signal, image, architecture, 26h, ESIR (M1)

C. Wolinski: programmable architectures, 10h, ESIR (M1)

C. Wolinski: component and system synthesis, 10h, Master by Research (ISTIC) (M2)

10.2.3. Supervision

PhD: Gabriel Gallin, Hardware arithmetic units and cryptoprocessors for hyperelliptic curve cryptography, Nov. 2018, A. Tisserand.

PhD: Aymen Gammoudi, Scheduling and Mapping Strategies for Software Tasks on Energy-Constrained Reconfigurable Architectures, June 2018, D. Chillet, M.Khalgui.

PhD: Jiating Luo, Architectural and Protocol Exploration for 3D Optical Network-on-Chip, Jul. 2018, D. Chillet, C. Killian, S. Le-Beux.

PhD: Mai-Thanh Tran, Towards Hardware Synthesis of a Flexible Radio from a High-Level Language, Nov. 2018, E. Casseau, M. Gautier.

PhD: Van Dung Pham, Architectural Exploration of Network Interface for Energy Efficient 3D Optical Network-on-Chip, Dec. 2018, O. Sentieys, D. Chillet, C. Killian, S. Le-Beux.

PhD: Rafail Psiakis, Performance Optimization Mechanisms for Fault-Resilient VLIW Processors, Dec. 2018, A. Kritikakou, O. Sentieys.

PhD: Simon Rokicki, Hardware acceleration of Dynamic Binary Translation, Dec. 2018, S. Derrien, E. Rohou.

PhD in progress: Minh Thanh Cong, Hardware Accelerated Simulation of Heterogeneous Multicore Platforms, May 2017, F. Charot, S. Derrien.

PhD in progress: Minyu Cui, Energy-Quality-Time Fault Tolerant Task Mapping on Multicore Architectures, Oct. 2018, E. Casseau, A. Kritikakou.

PhD in progress: Petr Dobias, Energy-Quality-Time Fault Tolerant Task Mapping on Multicore Architectures, Oct. 2017, E. Casseau.

PhD in progress: Mael Gueguen, Improving the performance and energy efficiency of complex heterogeneous manycore architectures with on-chip data mining, Nov. 2016, O. Sentieys, A. Termier.

PhD in progress: Van-Phu Ha, Application-Level Tuning of Accuracy, Nov. 2017, T. Yuki, O. Sentieys.

PhD in progress: Jaechul Lee, Energy-Performance Trade-Off in Optical Network-on-Chip, Dec. 2018, D. Chillet, C. Killian.

PhD in progress: Audrey Lucas, Software support resistant to passive and active attacks for asymmetric cryptography on (very) small computation cores, Jan. 2016, A. Tisserand.

PhD in progress: Thibaut Marty, Compiler support for speculative custom hardware accelerators, Sep. 2017, T. Yuki, O. Sentieys.

PhD in progress: Romain Mercier, Fault Tolerant Network on Chip for Deep Learning Algorithms, Oct. 2018, D. Chillet, C. Killian, A. Kritikakou.

PhD in progress: Genevieve Ndour, Approximate Computing with High Energy Efficiency for Internet of Things Applications, Apr. 2016, A. Tisserand, A. Molnos (CEA LETI).

PhD in progress: Joel Ortiz Sosa, Study and design of a digital baseband transceiver for wireless network-on-chip architectures, Nov. 2016, O. Sentieys, C. Roland (Lab-STICC).

PhD in progress: Davide Pala, Non-Volatile Processors for Intermittently-Powered Computing Systems, Jan. 2018, O. Sentieys, I. Miro-Panades (CEA LETI).

PhD in progress: Joseph Paturel, Design-space exploration of fault-tolerant multicores, Sep. 2018, O. Sentieys, A. Kritikakou.

PhD in progress: Nicolas Roux, Sensor-aided Non-Intrusive Appliance Load Monitoring: Detecting Activity of Devices through Low-Cost Wireless Sensors, Oct. 2016, O. Sentieys, B. Vrigneau.

10.3. Popularization

10.3.1. Articles and contents

Article (in French) about the Embrace project in *Le Mag numérique*: http://www.lemag-numerique.com/2018/01/embrace-vers-radio-hf-numerique-10653

Article in Emergences on "durcir les multi-cœurs contre les rayonnements ionisants": http://emergences.inria.fr/2018/newsletter-n51/L51-FLODAM

11. Bibliography

Major publications by the team in recent years

[1] R. DAVID, S. PILLEMENT, O. SENTIEYS. *Energy-Efficient Reconfigurable Processsors*, in "Low Power Electronics Design", C. PIGUET (editor), Computer Engineering, Vol 1, CRC Press, August 2004, chap. 20

[2] S. DERRIEN, S. RAJOPADHYE, P. QUINTON, T. RISSET. *12*, in "High-Level Synthesis From Algorithm to Digital Circuit", P. COUSSY, A. MORAWIEC (editors), Springer Netherlands, 2008, pp. 215-230, http://dx.doi.org/10.1007/978-1-4020-8588-8

[3] C. HURIAUX, A. COURTAY, O. SENTIEYS. *Design Flow and Run-Time Management for Compressed FPGA Configurations*, in "IEEE/ACM Design, Automation and Test in Europe (DATE)", March 2015, https://hal.inria.fr/hal-01089319

[4] J.-M. JÉZÉQUEL, B. COMBEMALE, S. DERRIEN, C. GUY, S. RAJOPADHYE. *Bridging the Chasm Between MDE and the World of Compilation*, in "Journal of Software and Systems Modeling (SoSyM)", October 2012, vol. 11, n° 4, pp. 581-597 [*DOI :* 10.1007/s10270-012-0266-8], https://hal.inria.fr/hal-00717219

[5] B. LE GAL, E. CASSEAU, S. HUET. *Dynamic Memory Access Management for High-Performance DSP Applications Using High-Level Synthesis*, in "IEEE Transactions on VLSI Systems", 2008, vol. 16, n[o] 11, pp. 1454-1464

[6] K. MARTIN, C. WOLINSKI, K. KUCHCINSKI, A. FLOCH, F. CHAROT. *Constraint Programming Approach to Reconfigurable Processor Extension Generation and Application Compilation*, in "ACM transactions on Reconfigurable Technology and Systems (TRETS)", June 2012, vol. 5, n° 2, pp. 1-38, http://doi.acm.org/10.1145/2209285.2209289

[7] D. MENARD, D. CHILLET, F. CHAROT, O. SENTIEYS. *Automatic Floating-point to Fixed-point Conversion for DSP Code Generation*, in "Proc. ACM/IEEE CASES", October 2002

[8] D. MENARD, O. SENTIEYS. *Automatic Evaluation of the Accuracy of Fixed-point Algorithms*, in "IEEE/ACM Design, Automation and Test in Europe (DATE-02)", Paris, March 2002

[9] S. PILLEMENT, O. SENTIEYS, R. DAVID. *DART: A Functional-Level Reconfigurable Architecture for High Energy Efficiency*, in "EURASIP Journal on Embedded Systems (JES)", 2008, pp. 1-13

[10] R. ROCHER, D. MENARD, O. SENTIEYS, P. SCALART. *Analytical Approach for Numerical Accuracy Estimation of Fixed-Point Systems Based on Smooth Operations*, in "IEEE Transactions on Circuits and Systems. Part I, Regular Papers", October 2012, vol. 59, n° 10, pp. 2326 - 2339 [*DOI :* 10.1109/TCSI.2012.2188938], http://hal.inria.fr/hal-00741741

[11] C. WOLINSKI, M. GOKHALE, K. MCCABE. *A polymorphous computing fabric*, in "IEEE Micro", 2002, vol. 22, n° 5, pp. 56–68

[12] C. WOLINSKI, K. KUCHCINSKI, E. RAFFIN. *Automatic Design of Application-Specific Reconfigurable Processor Extensions with UPaK Synthesis Kernel*, in "ACM Trans. on Design Automation of Elect. Syst.", 2009, vol. 15, n° 1, pp. 1–36, http://doi.acm.org/10.1145/1640457.1640458

Publications of the year

Doctoral Dissertations and Habilitation Theses

[13] A. GAMMOUDI. *Tasks scheduling and placement for reconfigurable architectures under energy constraint*, Université de rennes 1 ; Université de Carthage (Tunisie), June 2018, https://hal.inria.fr/tel-01956241

[14] L. JIATING. *Architectural and Protocol Exploration for 3D Optical Network-on-Chip*, Université de Rennes 1 [UR1], July 2018, https://hal.inria.fr/tel-01956255

[15] V. D. PHAM. *Architectural Exploration of Network Interface for Energy Efficient 3D Optical Network-on-Chip*, Université de rennes 1, December 2018, https://hal.inria.fr/tel-01956229

[16] R. PSIAKIS. *Performance Optimization Mechanisms for Fault-Resilient VLIW Processors*, Université de Rennes 1, December 2018, https://hal.inria.fr/tel-01956233

[17] S. ROKICKI. *Hardware Accelerated Dynamic Binary Translation*, Université de Rennes 1 [UR1], December 2018, https://hal.archives-ouvertes.fr/tel-01959136

[18] M.-T. TRAN. *Towards Hardware Synthesis of a Flexible Radio from a High-Level Language*, Université de Rennes 1 [UR1], December 2018, https://hal.inria.fr/tel-01942187

Articles in International Peer-Reviewed Journals

[19] S.-I. FILIP, Y. NAKATSUKASA, L. TREFETHEN, B. BECKERMANN. *Rational Minimax Approximation via Adaptive Barycentric Representations*, in "SIAM Journal on Scientific Computing", August 2018, vol. 40, n°4, pp. A2427-A2455, Author files., forthcoming, https://hal.inria.fr/hal-01942974

[20] A. GAMMOUDI, A. BENZINA, M. KHALGUI, D. CHILLET. *Energy-Efficient Scheduling of Real-Time Tasks in Reconfigurable Homogeneous Multi-core Platforms*, in "IEEE Transactions on Systems, Man, and Cybernetics: Systems", July 2018, pp. 1 - 14 [*DOI : 10.1109/TSMC.2018.2865965*], https://hal.inria.fr/hal-01934955

[21] A. KRITIKAKOU, T. MARTY, M. ROY. *DYNASCORE: DYNAmic Software COntroller to increase REsource utilization in mixed-critical systems*, in "ACM Transactions on Design Automation of Electronic Systems", January 2018, vol. 23, n°2, art ID n°13 p. [*DOI : 10.1145/3110222*], https://hal.archives-ouvertes.fr/hal-01559696

[22] J. LUO, C. KILLIAN, S. LE BEUX, D. CHILLET, O. SENTIEYS, I. O'CONNOR. *Offline Optimization of Wavelength Allocation and Laser Power in Nanophotonic Interconnects*, in "ACM Journal on Emerging Technologies in Computing Systems", July 2018, vol. 14, n°2, pp. 1 - 19 [*DOI : 10.1145/3178453*], https://hal.inria.fr/hal-01934870

[23] L. MO, X. CAO, Y.-Q. SONG, A. KRITIKAKOU. *Distributed Node Coordination for Real-Time Energy-Constrained Control in Wireless Sensor and Actuator Networks*, in "IEEE internet of things journal", May 2018, pp. 1-12 [*DOI : 10.1109/JIOT.2018.2839030*], https://hal.inria.fr/hal-01825524

[24] L. MO, A. KRITIKAKOU, O. SENTIEYS. *Controllable QoS for Imprecise Computation Tasks on DVFS Multicores with Time and Energy Constraints*, in "IEEE Journal on Emerging and Selected Topics in Circuits and Systems", July 2018, vol. 8, n°4, pp. 708-721 [*DOI : 10.1109/JETCAS.2018.2852005*], https://hal.inria.fr/hal-01831297

[25] L. MO, A. KRITIKAKOU, O. SENTIEYS. *Energy-Quality-Time Optimized Task Mapping on DVFS-enabled Multicores*, in "IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems", July 2018, pp. 1 - 10 [*DOI : 10.1109/TCAD.2018.2857300*], https://hal.inria.fr/hal-01843918

[26] T. H. NGUYEN, P. SCALART, M. GAY, L. BRAMERIE, C. PEUCHERET, F. GOMEZ AGIS, O. SENTIEYS, J.-C. SIMON, M. JOINDOT. *New metric for IQ imbalance compensation in optical QPSK coherent systems*, in "Photonic Network Communications", December 2018, vol. 36, n°3, pp. 326-337 [*DOI : 10.1007/s11107-018-0783-7*], https://hal.inria.fr/hal-01941892

[27] S. ROKICKI, E. ROHOU, S. DERRIEN. *Hybrid-DBT: Hardware/Software Dynamic Binary Translation Targeting VLIW*, in "IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems", August 2018, pp. 1-14 [*DOI : 10.1109/TCAD.2018.2864288*], https://hal.archives-ouvertes.fr/hal-01856163

International Conferences with Proceedings

[28] N. BRISEBARRE, G. CONSTANTINIDES, M. ERCEGOVAC, S.-I. FILIP, M. ISTOAN, J.-M. MULLER. *A High Throughput Polynomial and Rational Function Approximations Evaluator*, in "ARITH 2018 - 25th IEEE Symposium on Computer Arithmetic", Amherst, MA, United States, IEEE, June 2018, pp. 99-106 [*DOI : 10.1109/ARITH.2018.8464778*], https://hal.inria.fr/hal-01774364

[29] A. COURTAY, M. LE GENTIL, O. BERDER, A. CARER, P. SCALART, O. SENTIEYS. *Zyggie: A Wireless Body Area Network platform for indoor positioning and motion tracking*, in "ISCAS 2018 - IEEE International Symposium on Circuits and Systems", Florence, Italy, IEEE, May 2018, pp. 1-5 [*DOI :* 10.1109/ISCAS.2018.8351278], https://hal.archives-ouvertes.fr/hal-01804927

[30] P. DOBIAS, E. CASSEAU, O. SINNEN. *Comparison of Different Methods Making Use of Backup Copies for Fault-Tolerant Scheduling on Embedded Multiprocessor Systems*, in "DASIP 2018 - Conference on Design and Architectures for Signal and Image Processing", Porto, Portugal, October 2018, pp. 1-7, https://hal.inria.fr/hal-01942186

[31] P. DOBIAS, E. CASSEAU, O. SINNEN. *Restricted Scheduling Windows for Dynamic Fault-Tolerant Primary/Backup Approach-Based Scheduling on Embedded Systems*, in "SCOPES '18 - 21th International Workshop on Software and Compilers for Embedded Systems", Sankt Goar, Germany, May 2018, pp. 27-30 [*DOI :* 10.1145/3207719.3207724], https://hal.inria.fr/hal-01942185

[32] A. GAMMOUDI, D. CHILLET, M. KHALGUI, A. BENZINA. *Mapping of Periodic Tasks in Reconfigurable Heterogeneous Multi-core Platforms*, in "ENASE 2018 - 13th International Conference on Evaluation of Novel Approaches to Software Engineering", Funchal, France, SCITEPRESS - Science and Technology Publications, March 2018, pp. 99-110 [*DOI :* 10.5220/0006698500990110], https://hal.inria.fr/hal-01936163

[33] M. GUEGUEN, O. SENTIEYS, A. TERMIER. *Accelerating Itemset Sampling using Satisfiability Constraints on FPGA*, in "DATE 2019 - IEEE/ACM Design, Automation and Test in Europe", Florence, Italy, March 2019, pp. 1-6, https://hal.inria.fr/hal-01941862

[34] T. LEFEUVRE, E. K. KASNAKLI, I. FASSI, I. PUAUT, C. CULLMANN, S. DERRIEN, G. GEBHARD. *Using Polyhedral Techniques to Tighten WCET Estimates of Optimized Code: A Case Study with Array Contraction*, in "DATE 2018 - Design Automation and Test Europe", Dresden, Germany, IEEE, March 2018, pp. 925-930 [*DOI :* 10.23919/DATE.2018.8342142], https://hal.inria.fr/hal-01815499

[35] J. LUO, V.-D. PHAM, C. KILLIAN, D. CHILLET, I. O'CONNOR, O. SENTIEYS, S. LE BEUX. *Run-Time management of energy-performance trade-off in Optical Network-on-Chip*, in "DCIS 2018 - XXXIII Conference on Design of Circuits and Integrated Systems", Lyon, France, November 2018, pp. 1-6, https://hal.archives-ouvertes.fr/hal-01937350

[36] T. MARTY, T. YUKI, S. DERRIEN. *Enabling Overclocking with HLS Tools through Algorithm-Level Error Detection*, in "FPT 2018 - International Conference on Field-Programmable Technology", Naha, Japan, December 2018, https://hal.inria.fr/hal-01942429

[37] L. MO, A. KRITIKAKOU, X. CAO. *Collaborative State Estimation and Actuator Scheduling for Cyber-Physical Systems under Random Multiple Events*, in "AdHoc-Now 2018 - 17th International Conference on Ad Hoc Networks and Wireless", Saint Malo, France, Springer, September 2018, pp. 267-279 [*DOI :* 10.1007/978-3-030-00247-3_24], https://hal.inria.fr/hal-01857496

[38] L. MO, A. KRITIKAKOU, O. SENTIEYS. *Energy-Quality-Time Optimized Task Mapping on DVFS-enabled Multicores*, in "ESWEEK 2018 - Embedded Systems Week", Torino, Italy, September 2018, pp. 1-11, https://hal.inria.fr/hal-01941764

[39] L. Mo, A. Kritikakou, O. Sentieys. *Approximation-aware Task Deployment on Asymmetric Multicore Processors*, in "DATE 2019 - 22nd IEEE/ACM Design, Automation and Test in Europe", Florence, Italy, March 2019, pp. 1-6, https://hal.inria.fr/hal-01940358

[40] R. Psiakis, A. Kritikakou, O. Sentieys. *Fine-Grained Hardware Mitigation for Multiple Long-Duration Transients on VLIW Function Units*, in "DATE 2019 IEEE/ACM Design, Automation and Test in Europe", Florence, Italy, March 2019, pp. 1-4, https://hal.inria.fr/hal-01941860

[41] I. Puaut, M. Dardaillon, C. Cullmann, G. Gebhard, S. Derrien. *Fine-Grain Iterative Compilation for WCET Estimation*, in "WCET 2018 - 18th International Workshop on Worst-Case Execution Time Analysis", Barcelona, Spain, July 2018, pp. 1-12 [*DOI :* 10.4230/OASIcs.WCET.2018.9], https://hal.inria.fr/hal-01889944

[42] S. Rokicki, E. Rohou, S. Derrien. *Supporting Runtime Reconfigurable VLIWs Cores Through Dynamic Binary Translation*, in "DATE 2018 - IEEE/ACM Design, Automation & Test in Europe Conference & Exhibition", Dresden, Germany, IEEE, March 2018, pp. 1009-1014 [*DOI :* 10.23919/DATE.2018.8342160], https://hal.archives-ouvertes.fr/hal-01653110

[43] S. Rokicki, E. Rohou, S. Derrien. *Aggressive Memory Speculation in HW/SW Co-Designed Machines*, in "DATE 2019 - IEEE/ACM Design, Automation and Test in Europe", Florence, Italy, March 2019, https://hal.archives-ouvertes.fr/hal-01941876

[44] M. Soltan Mohammadi, K. Cheshmi, M. Mehri Dehnavi, A. Venkat, T. Yuki, M. Mills Strout. *Extending Index-Array Properties for Data Dependence Analysis*, in "LCPC 2018 - 31st International Workshop on Languages and Compilers for Parallel Computing", Salt Lake City, United States, October 2018, https://hal.inria.fr/hal-01942363

[45] J. O. Sosa, O. Sentieys, C. Roland. *A Diversity Scheme to Enhance the Reliability of Wireless NoC in Multipath Channel Environment*, in "Twelfth IEEE/ACM International Symposium on Networks-on-Chip (NOCS)", Torino, Italy, October 2018, pp. 1-8 [*DOI :* 10.1109/NOCS.2018.8512165], https://hal.inria.fr/hal-01941761

Conferences without Proceedings

[46] A. Lucas. *A Simulator for Evaluating the Leakage in Arithmetic Circuits*, in "CryptArchi 2018 - International Workshop on Cryptographic architectures embedded in logic devices", Lorient, France, June 2018, pp. 1-24, https://hal.archives-ouvertes.fr/hal-01841048

[47] A. Lucas, A. Tisserand. *Microcontroller Implementation of Simultaneous Protections Against Observation and Perturbation Attacks for ECC*, in "SECRYPT: 15th International Conference on Security and Cryptography", Porto, Portugal, July 2018, pp. 1-8 [*DOI :* 10.5220/0006884605700577], https://hal.archives-ouvertes.fr/hal-01826303

Scientific Books (or Scientific Book chapters)

[48] D. Menard, G. Caffarena, J. A. Lopez, D. Novo, O. Sentieys. *Fixed-point refinement of digital signal processing systems*, in "Digitally Enhanced Mixed Signal Systems", 2018, forthcoming, https://hal.inria.fr/hal-01941898

[49] D. MENARD, G. CAFFARENA, J. A. LOPEZ, D. NOVO, O. SENTIEYS. *Analysis of Finite Word-Length Effects in Fixed-Point Systems*, in "Handbook of Signal Processing Systems", S. S. BHATTACHARYYA (editor), 2019, pp. 1063-1101 [*DOI :* 10.1007/978-3-319-91734-4_29], https://hal.inria.fr/hal-01941888

[50] L. MO, A. KRITIKAKOU, O. SENTIEYS. *Imprecise Computation Task Mapping on Multi-Core Wireless Sensor Networks*, in "Encyclopedia of Wireless Networks", October 2018, pp. 1 - 6, forthcoming [*DOI :* 10.1007/978-3-319-32903-1_261-1], https://hal.inria.fr/hal-01900174

Research Reports

[51] P. IANNETTA, L. GONNORD, L. MOREL, T. YUKI. *Semantic Array Dataflow Analysis*, Inria Grenoble Rhône-Alpes, December 2018, n⁰ RR-9232, https://hal.archives-ouvertes.fr/hal-01954396

[52] T. MARTY, T. YUKI, S. DERRIEN. *Algorithm Level Timing Speculation for Convolutional Neural Network Accelerators*, Univ Rennes, Inria, CNRS, IRISA, France, June 2018, n⁰ RT-0500, pp. 1-17, https://hal.inria.fr/hal-01811231

[53] M. SOLTAN MOHAMMADI, K. CHESHMI, G. GOPALAKRISHNAN, M. HALL, M. MEHRI DEHNAVI, A. VENKAT, T. YUKI, M. MILLS STROUT. *Sparse Matrix Code Dependence Analysis Simplification at Compile Time*, Arxiv, July 2018, https://hal.inria.fr/hal-01942381

Other Publications

[54] G. BERTHOU, A. CARER, H.-P. CHARLES, S. DERRIEN, K. MARQUET, I. MIRO-PANADES, D. PALA, I. PUAUT, F. RASTELLO, T. RISSET, E. ROHOU, G. SALAGNAC, O. SENTIEYS, B. YARAHMADI. *The Inria ZEP project: NVRAM and Harvesting for Zero Power Computations*, March 2018, 10th Annual Non-Volatile Memories Workshop (NVMW), Poster, https://hal.inria.fr/hal-01941766

[55] A. BOSIO, D. MENARD, O. SENTIEYS. *A Comprehensive Analysis of Approximate Computing Techniques: From Component- to Application-Level*, September 2018, pp. 1-2, ESWEEK 2018 - Embedded Systems Week, https://hal.inria.fr/hal-01941755

[56] A. BOSIO, D. MENARD, O. SENTIEYS. *A Comprehensive Analysis of Approximate Computing Techniques: From Component- to Application-Level*, March 2019, DATE 2019 - 22nd IEEE/ACM Design, Automation and Test in Europe, https://hal.inria.fr/hal-01941757

[57] S.-I. FILIP, A. JAVEED, L. N. TREFETHEN. *Smooth random functions, random ODEs, and Gaussian processes*, December 2018, To appear in SIAM Review., https://hal.inria.fr/hal-01944992

[58] N. ROUX, B. VRIGNEAU, O. SENTIEYS. *Estimating Power Loads from Partial Appliance States*, March 2018, NILM 2018 - 4th International Workshop on Non-Intrusive Load Monitoring, Poster, https://hal.inria.fr/hal-01941877

[59] O. SENTIEYS. *Playing with number representations and operator-level approximations*, June 2018, Keynote at the Third Workshop on Approximate Computing (AxC), in conjunction with IEEE European Test Symposium (ETS), https://hal.inria.fr/hal-01941868

References in notes

[60] S. HAUCK, A. DEHON (editors). *Reconfigurable Computing: The Theory and Practice of FPGA-Based Computation*, Morgan Kaufmann, 2008

[61] V. BAUMGARTE, G. EHLERS, F. MAY, A. NÜCKEL, M. VORBACH, M. WEINHARDT. *PACT XPP — A Self-Reconfigurable Data Processing Architecture*, in "The Journal of Supercomputing", 2003, vol. 26, n° 2, pp. 167–184

[62] C. BECKHOFF, D. KOCH, J. TORRESEN. *Portable module relocation and bitstream compression for Xilinx FPGAs*, in "24th Int. Conf. on Field Programmable Logic and Applications (FPL)", 2014, pp. 1–8

[63] C. BOBDA. *Introduction to Reconfigurable Comp.: Architectures Algorithms and Applications*, Springer, 2007

[64] S. BORKAR, A. A. CHIEN. *The Future of Microprocessors*, in "Commun. ACM", May 2011, vol. 54, n° 5, pp. 67–77, http://doi.acm.org/10.1145/1941487.1941507

[65] J. M. P. CARDOSO, P. C. DINIZ, M. WEINHARDT. *Compiling for reconfigurable computing: A survey*, in "ACM Comput. Surv.", June 2010, vol. 42, 13:1 p. , http://doi.acm.org/10.1145/1749603.1749604

[66] K. COMPTON, S. HAUCK. *Reconfigurable computing: a survey of systems and software*, in "ACM Comput. Surv.", 2002, vol. 34, n° 2, pp. 171–210, http://doi.acm.org/10.1145/508352.508353

[67] J. CONG, H. HUANG, C. MA, B. XIAO, P. ZHOU. *A Fully Pipelined and Dynamically Composable Architecture of CGRA*, in "IEEE Int. Symp. on Field-Program. Custom Comput. Machines (FCCM)", 2014, pp. 9–16, http://dx.doi.org/10.1109/FCCM.2014.12

[68] G. CONSTANTINIDES, P. CHEUNG, W. LUK. *Wordlength optimization for linear digital signal processing*, in "IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems", October 2003, vol. 22, n° 10, pp. 1432- 1442

[69] M. COORS, H. KEDING, O. LUTHJE, H. MEYR. *Fast Bit-True Simulation*, in "Proc. ACM/IEEE Design Automation Conference (DAC)", Las Vegas, june 2001, pp. 708-713

[70] R. H. DENNARD, F. H. GAENSSLEN, V. L. RIDEOUT, E. BASSOUS, A. R. LEBLANC. *Design of ion-implanted MOSFET's with very small physical dimensions*, in "IEEE Journal of Solid-State Circuits", 1974, vol. 9, n° 5, pp. 256–268

[71] A. HORMATI, M. KUDLUR, S. MAHLKE, D. BACON, R. RABBAH. *Optimus: efficient realization of streaming applications on FPGAs*, in "Proc. ACM/IEEE CASES", 2008, pp. 41–50

[72] H. KALTE, M. PORRMANN. *REPLICA2Pro: Task Relocation by Bitstream Manipulation in Virtex-II/Pro FPGAs*, in "3rd Conference on Computing Frontiers (CF)", 2006, pp. 403–412

[73] J.-E. LEE, K. CHOI, N. D. DUTT. *Compilation Approach for Coarse-Grained Reconfigurable Architectures*, in "IEEE Design and Test of Computers", 2003, vol. 20, n° 1, pp. 26-33, http://doi.ieeecomputersociety.org/10.1109/MDT.2003.1173050

[74] H. LEE, D. NGUYEN, J.-E. LEE. *Optimizing Stream Program Performance on CGRA-based Systems*, in "52nd IEEE/ACM Design Automation Conference", 2015, pp. 110:1–110:6, http://doi.acm.org/10.1145/2744769. 2744884

[75] B. MEI, S. VERNALDE, D. VERKEST, H. DE MAN, R. LAUWEREINS. *ADRES: An architecture with tightly coupled VLIW processor and coarse-grained reconfigurable matrix*, in "Proc. FPL", Springer, 2003, pp. 61–70

[76] N. R. MINISKAR, S. KOHLI, H. PARK, D. YOO. *Retargetable Automatic Generation of Compound Instructions for CGRA Based Reconfigurable Processor Applications*, in "Proc. ACM/IEEE CASES", 2014, pp. 4:1–4:9, http://doi.acm.org/10.1145/2656106.2656125

[77] Y. PARK, H. PARK, S. MAHLKE. *CGRA express: accelerating execution using dynamic operation fusion*, in "Proc. Int. Conf. on Compilers, Architecture, and Synthesis for Embedded Systems", New York, NY, USA, CASES'09, ACM, 2009, pp. 271–280, http://doi.acm.org/10.1145/1629395.1629433

[78] A. PUTNAM, A. CAULFIELD, E. CHUNG, D. CHIOU, K. CONSTANTINIDES, J. DEMME, H. ESMAEILZADEH, J. FOWERS, G. GOPAL, J. GRAY, M. HASELMAN, S. HAUCK, S. HEIL, A. HORMATI, J.-Y. KIM, S. LANKA, J. LARUS, E. PETERSON, S. POPE, A. SMITH, J. THONG, P. XIAO, D. BURGER. *A reconfigurable fabric for accelerating large-scale datacenter services*, in "ACM/IEEE 41st International Symposium on Computer Architecture (ISCA)", June 2014, pp. 13-24, http://dx.doi.org/10.1109/ISCA.2014. 6853195

[79] G. THEODORIDIS, D. SOUDRIS, S. VASSILIADIS. *2*, in "A survey of coarse-grain reconfigurable architectures and CAD tools", Springer Verlag, 2007

[80] G. VENKATARAMANI, W. NAJJAR, F. KURDAHI, N. BAGHERZADEH, W. BOHM, J. HAMMES. *Automatic compilation to a coarse-grained reconfigurable system-on-chip*, in "ACM Trans. on Emb. Comp. Syst.", 2003, vol. 2, n° 4, pp. 560–589, http://doi.acm.org/10.1145/950162.950167