HiPEAC info 37

COMPILATION ARCHITECTURE

APPEARS QUARTERLY
JANUARY 2014

NETWORK OF EXCELLENCE ON HIGH PERFORMANCE AND EMBEDDED ARCHITECTURE AND COMPILATION

WELCOME TO THE HIPEAC 2014 CONFERENCE, JANUARY 20-22, 2014, VIENNA, AUSTRIA

SEVENTH FRAMEWORK PROGRAMME

WWW.HIPEAC.NET

SPRING COMPUTING SYSTEMS WEEK, MAY 13-15, 2014, BARCELONA, SPAIN

MESSAGE FROM THE HIPEAC COORDINATOR

First of all, I would like to wish you a healthy and prosperous 2014, personally as well as professionally. The start of a new year is always a good occasion to make an evaluation of the previous year, and to make resolutions for the coming year.

I will remember 2013 as the year of the NSA leaks. For me, this proves once more that the internet democratizes truth by empowering individual people. Even the most powerful country in the world was not able to prevent the leaks, or to contain them.

Many internet users were shocked when they learned that major corporations like Google and Facebook were asked to disclose information about them. Of course, these companies disclosed the information, because they had to; they might even have been paid for the extra work it took. What surprised me most was the indignation of many politicians that they were being spied on, when everybody knows that this is what secret services do, including their own. The scary thing however is that if secret services can hack communication devices, we must assume that everybody can. Hopefully, people now realize that digital communication gives a false sense of security and cannot be trusted.

HiPEAC has been warning of this danger for half a decade in its bi-annual roadmaps. We have always stressed that making systems secure is a major challenge, and even that it is impossible to make systems safe without first making them secure. In the 2013 roadmap, dependability is one of the three core challenges, along with energy efficiency and system complexity.

Many of you will read this newsletter at the HiPEAC conference in Vienna. This conference is the flagship networking event of HiPEAC. This year, we used the innovative journal-first publication model for the third time, leading to 37 thoroughly reviewed and revised high quality papers that will be presented at the conference.

2014 will be a special year for HiPEAC because HiPEAC celebrates its 10th anniversary. Started as a bold initiative by Mateo Valero to organize the computing systems community in Europe, and with the help of many volunteers, HiPEAC succeeded in building a vibrant European research network. I believe that HiPEAC has created a voice for the European computing systems community, and that this voice is being heard.

The 10th year of HiPEAC is a good opportunity to look back on what we have accomplished so far, to be thankful for the generous support we received from the European Commission, and to make plans for the next 10 years. There are many ideas on how to further improve the impact of our community. If you want to learn more about it, you definitely have to join us for the celebration of 10 years of HiPEAC in Barcelona on May 14, 2014.

Take care!

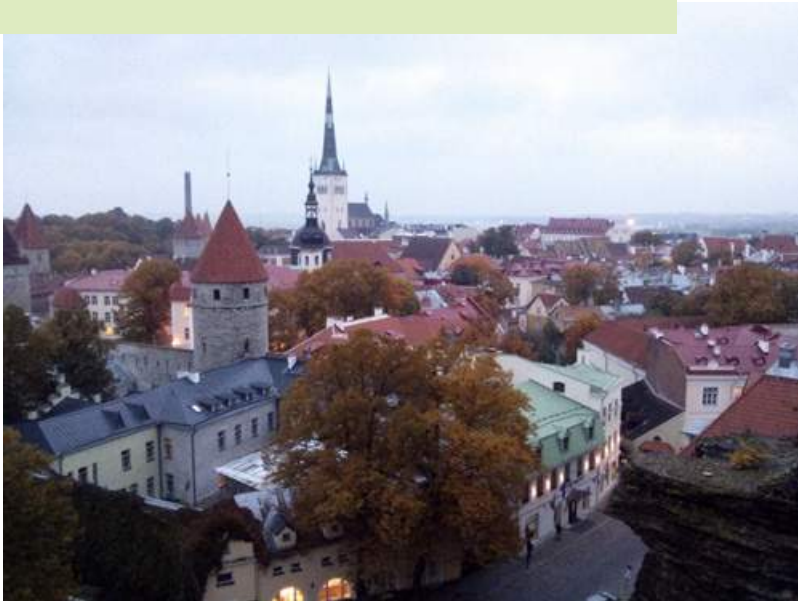
Koen De Bosschere

CONTENT

HIPEAC AUTUMN COMPUTING SYSTEMS WEEK, TALLINN

This CSW featured a program with seven thematic sessions as well as the second edition of the HiPEAC Industry Partner Program (HIPP) event

The program for the last Computing Systems Week in Tallinn (October 7-9, 2013) was designed around seven thematic sessions, which resulted from a call launched the first week of June. Before presenting a short summary of each of the sessions and their main conclusions, I just want to thank the organizers for the enthusiasm in preparing them and for their contributions to this summary for the HiPEAC Newsletter. Additional details about the presentations and slides for most of them are available through the HiPEAC website.



EVALUATION OF HETEROGENEOUS SYSTEMS AND THEIR APPLICATIONS

This thematic session explored how to evaluate heterogeneous and hybrid systems with respect to performance, energy consumption, synchronization efficiency, and/or code quality. The session was the 4th installment in a series of sessions on heterogeneous and hybrid systems. Previous sessions covered topics in the areas of programming models, compilers, runtime systems, and architectural support.

The session followed a twofold scheme, with a first slot devoted to invited talks and a second slot offering an opportunity for HiPEAC members to share their experiences and evaluations in the research area of heterogeneous and hybrid platforms. The first invited speaker, Julia Fedorova from Intel Russia, introduced new VTune capabilities in CPU/GPU activities correlation; audience questions and participation showed a real interest in the tool and the possibilities for research and performance improvement of applications, hoping to have a Linux version available as soon as possible. After this, Per Gunnar Kjeldsberg from NTNU presented a design of system scenario methodology to evaluate performance and energy consumption in heterogeneous multi-processor systems; and Alessio Sclocco from Vrije Universiteit in Amsterdam and the Netherlands eScience Center, shared his results on using modern many-core architectures to accelerate radio astronomy processing.

Keynote about The Graphene Flagship, by Prof. Fogelström (Chalmers)

The second slot included presentations from the following speakers: Stefan Wallentowitz, (PhD student at TU München, Germany), Magnus Jahre (NTNU, Norway) and Erik Tomusk (PhD student at UoE, UK).

SILICON PHOTONICS FOR NEXT GENERATION COMPUTING SYSTEMS

This session aimed at shading light on the current maturity of silicon photonic technology, with the expected perspective of becoming part of the evolution of future computing systems.
The session hosted a variety of perspectives, from close-to-silicon points of view up to network and system level ones. Guido Chiaretti (ST Microelectronics) highlighted STM's view and expected roadmap, along with key milestones, from board level up to completely on-chip adoption of silicon photonics, through the mandatory silicon interposer step. Takahiro Nakamura (Photonics Electronics Technology Research Association) described the results of the Japanese PETRA association in the study and, specifically, in the actual implementation of a real, fully-functional, photonically-integrated active silicon interposer prototype. Sergei F. Mingaleev (VPI Photonics) emphasized the extreme need for design and simulation tools for complex photonic circuits for on-chip photonics to become a viable design resource for computer systems. The talk highlighted criticalities and advanced methods to overcome current limitations and issues. Yvain Thonnart (LETI) described the research performed in LETI about the implications of physical and applicative constraints in the design choices of a photonic on-chip interconnection. Some physical and network-level issues have a major role in ranking the relative importance of alternative design solutions. Kostas Katrinis (IBM Research) described IBM's current efforts in adopting photonics in commercial solutions, specifically for HPC systems to significantly reduce latency and communication delay. Davide

Bertozzi (University of Ferrara) reported about a cross-benchmarking framework between an optical interconnection network and its aggressive electrical baseline in 40nm technology. The study pointed out the critical role of the network interface architecture, which accounts for most of the complexity in an optical network-on-chip, but which is unfortunately often overlooked. Finally, Sandro Bartolini (University of Siena) discussed the application of on-chip silicon photonics in cache-coherent chip multi-processors and highlighted the critical interactions between CMP architectural issues and far lower-level silicon interconnection design choices in the pursuit of low-power and performance.

TEACHING IN THE FACE OF MOOCS

In this thematic session the organizer, David Black-Schaffer, discussed how massive online open courses (MOOCs) are likely to bring competition to teaching, much as we've seen competition in research for many years. To compete with internationally produced online courses, we need to focus on providing real value to the students in their physical in-class meetings. One approach to this is to "flip the classroom", wherein lectures are put online and students participate in active problem solving during the class time.

During the session there was a wide-ranging discussion of how this method could be adapted to a range of classes and how to address the difficulty of providing teachers with the time and training needed to make the most of a flipped-classroom teaching approach. The discussion fit in well with Dr. Khaled Benkrid's presentation of ARM's work in providing educational resources for teachers to use. After our discussions, several of the attendees were actively interested in pursuing this form of teaching in their own classrooms.

Dr. David Black-Schaffer (Uppsala) during the MOOCs session.

STORAGE, PERSISTENT I/O, AND THE EVOLVING MEMORY HIERARCHY

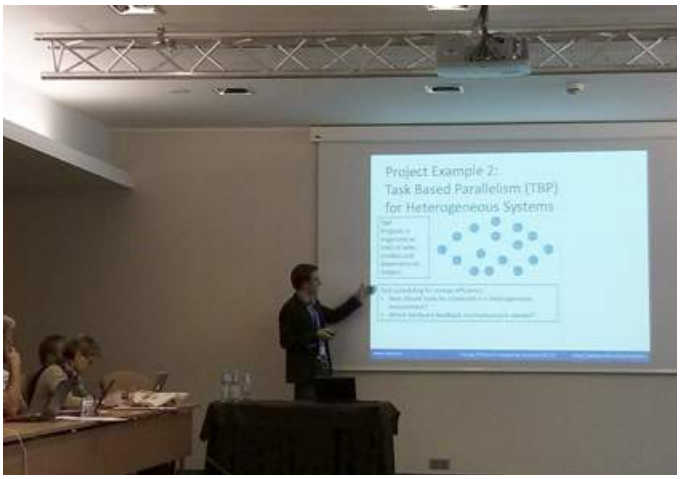
Storage and persistent I/O is deemed an important part of current and future servers and systems, due to the changing nature of computation from memory- to data-centric. This trend towards data-centric computation means that storage is often an important performance bottleneck and it is therefore a significant part of the total cost of ownership, for both servers and HPC systems.

The thematic session was organized around two talks by Dr. Simone Lavizzari (R&D – Technology Development, Micron) and Dr. Anna Queralt (Barcelona Supercomputing Center). The talk by Simone Lavizzari was on trends and evolution of non-volatile memories, covering NAND/DRAM challenges and new memory concepts. Memories are becoming increasing important because they are increasingly fundamental to the definition of an electronic system. At present, the industry standard technologies are still DRAM and Flash, which have been able to guarantee cost sustainability thanks to continuous scaling. NAND/DRAM miniaturization is, however, becoming increasingly difficult, and new applications are requiring higher memory density and better performance. In consequence, there are good opportunities for the alternative memory technologies to enter into the market and replace or displace the standard ones.

The talk by Anna Queralt was on the potential, limitations, and challenges in cloud data management. The speaker focussed on current trends in data storage and management in the cloud, from cloud storage solutions, which provide access to the data through a coarse-grained interface, to data services, which use a fine-grained interface. She also provided an overview of existing approaches in this area and the kind of problems they can solve. However, when it comes to sharing data in an efficient way, current solutions are too inflexible, which opens the door to new approaches that enable external users to personalize the way they consume and process data.

TIMING ANALYSIS FOR MULTICORE/MANYCORE ARCHITECTURES

Timing analysis for multicore/manycore architectures presents several new technical challenges. Solving these challenges is essential to the use of these architectures in mixed criticality, embedded and other settings. This thematic session introduced some of the technical challenges and outlined possible solutions. The session was organized around four presentations. The first one, by Bjørn Lisper (Mälardalen University, Sweden) presented the TACLe EU COST Action on Code-Level Timing Analysis. The Action gathers all the prominent European groups in the area, as well as groups in neighboring areas. The Action is motivated particularly by the challenges to code-level timing analysis brought by the rapid transition to multi-core technology, and anticipated future evolution to many-core architectures. These challenges require a coordinated effort to be met.



The second talk, by Peter Puschner (Universitaet Wien, Austria Technische), was entitled Achieving Time Predictability in the MPSoC Age., and presented some key principles that must be followed when constructing the hardware and software architecture for embedded multi-core systems, if those systems are to be used in safety-critical or mixed-criticality applications. In the third talk, Raimund Kirner (University of Hertfordshire) presented the concept of mixed time-criticality, which allows describing different types of real-time system requirements, including service criticality, and described what types of performance specifications should be supported by the programming models. Finally, Kevin Hammond (University of St Andrews, UK) presented Predictable Timing Behavior on x86 Multicores using High-Level Patterns of Parallelism, showing that it is possible to accurately predict timing behaviour for real x86 multicores.

TECHNOLOGY TRANSFER IN COMPUTING SYSTEMS: THE TETRACOM APPROACH

This session introduced the new TETRACOM FP7 Project. The mission of the TETRACOM Coordination Action is to boost European academia-to-industry technology transfer (TT) in all domains of Computing Systems. While many other European and national initiatives focus on training of entrepreneurs and support for start-up companies, the key differentiator of TETRACOM is a novel instrument called the Technology Transfer Project (TTP). The TTP activities will be complemented by Technology Transfer Infrastructures (TTIs), which provide training, service, and dissemination actions. Altogether, TETRACOM is conceived as the major pilot project of its kind in the area of Computing Systems, acting as a TT catalyst for the mutual benefit of academia and industry. The project's primary success metrics are the number and value of coordinated TTPs as well as the amount of newly introduced European TT actors. The session featured three examples of successful TTPs: "An ultra-fast just-in-time-simulator" by Mike O´Boyle (University of Edinburgh), "From the FP6 Hartes project to BlueBee" by Koen Bertels (TU Delft) and "CoScale: from PhD to Product", by F. Ryckbosch (CoScale). The session finished with a short talk by Laurent Julliard (Kalray) presenting the industry viewpoint with a talk entitled "MPPA: what it takes to create and market a manycore processor".

DEPENDABILITY CHALLENGES

This thematic session aimed to increase awareness of dependability-related activity in EU funded projects, identify key challenges related to dependability, offer a forum for networking and exchange that fosters collaboration, and facilitate the formation of consortia for future research proposals. The session included invited presentations from seven projects related to dependability, each of which reported on the challenges they address, their interesting findings, and their current status. Two of the projects have recently completed, DIAMOND and EUROCLOUD, two are in progress, DESYRE and PARADIME, whereas RoMoL, BASTION and HARPA are either about to start or have recently kicked off. The projects covered a wide spectrum of problems and solutions spanning all layers of the compute stack. These include tools for error detection and diagnosis (DIAMOND), fault-tolerance for cloud servers (EUROCLOUD), hardware and software adaptation to battle permanent and transient faults (DESYRE), exploring energy-reliability trade-offs (PARADIME), leveraging task-based programming and a dataflow-based runtime for enhanced resiliency (RoMoL), battling aging and no failure found field returns (BASTION), and a cross layer approach to harnessing performance variability due to time-dependent variations (HARPA). Finally, we were pleased that two of the presentations were given by Estonian colleagues, helping to highlight the strong expertise of their institutions in the area of dependability.

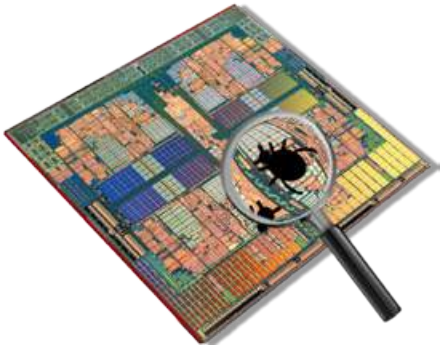
CSW social event in Tallinn

As a final note for this article, I just want to stress the importance of this HiPEAC-3 instrument to promote your research areas in the community, to share your own research results, and to build a network of researchers from which you can form a consortium for a future project proposal. Volunteering to organize a thematic session is an opportunity to contribute to the HiPEAC community, helping the HiPEAC network to work on the challenges identified in the **HiPEAC roadmap**. Being innovative in the organization of these sessions, escaping from the "traditional" format, will be very welcomed. Thanks in advance for submitting a proposal in the forthcoming calls for thematic sessions.

Eduard Ayguadé,
Barcelona Supercomputing Center (BSC) and Universitat Politècnica de Catalunya (UPC), Spain

DO YOU TRUST IN MULTICORE HW/SW SYSTEMS?

The 1st International Workshop on Multicore Application Debugging (MAD-2013) was held at TU München on November 14-15, 2013



The trend towards multicore architectures has a fundamental impact on the way embedded systems are designed. A lot of current research is directed towards novel multicore architectures, programming languages and compilers. However, new software validation and debugging technologies still seem to be an afterthought. Classical debug approaches do not scale in the multicore era, as they do not cover new types of software and system problems related to parallel execution, such as data races and deadlocks. Correct software execution may become architecture configuration dependent. Moreover, system developers face limited observability within SoC platforms, platform heterogeneity, and skyrocketing complexity of upcoming manycore systems with hundreds of integrated processing elements. These challenges demand radically new debug approaches.

Discussion of these topics among a group of experts was the major purpose of the MAD-2013 workshop, organized by Rainer Leupers (RWTH Aachen University) and Andreas Herkersdorf and Samarjit Chakraborty (TU München) during Nov 14-15, 2013 at TU München´s new Garching Campus. Around 45 attendees accepted the personal invitation to present and analyse the state of the art and future approaches in the important domain of multicore applications debugging. Thanks to a good academia-industry balance, the workshop program including entirely new ideas as well as contemporary best practices. Participating companies included tool providers (Synopsys, ACE, Lauterbach), hardware/IP providers (Freescale, Infineon, Intel, National Instruments, TOPS), and system houses (Bosch, BMW, Cassidian, Samsung). The academic views and ideas, amongst others, were represented by TU Vienna, ETH Zurich, University of Erlangen, BSC, and the National University of Singapore. The nine workshop sessions ranged from industrial needs, through HW tracing and debug support, novel SW tools and concepts, and observability issues, to compiler support for debugging and runtime verification. Besides a Bavarian-style social event, the workshop program also included a significant number of discussion slots.

According to on-site survey results, the informal workshop format and its mix of contents were well received. It became clear that multicore debugging is a major challenge in the light of future system complexity levels. The topic is still being approached from various angles, and better coordination between the stakeholders seems necessary. For instance, HW tracing of complex SoCs currently tends to generate an enormous amount of data, which needs to be mined for useful and comprehensible information. On the SW tools side, higher "system-wide" debug abstraction levels are required to avoid dozens of simultaneous debugger windows on a SW developer's screen. In contrast to e.g. SW compiler tools, the area of multicore system debugging is heavily underrepresented in academic research, and there is a high potential for tighter collaborations between academia and industry. The MAD workshop community will therefore pursue several follow-up activities, including special sessions at major conferences as well as joint R&D projects. Moreover, a second MAD workshop is already on the horizon.

Anyone interested to work on this new hot topic is invited to visit the workshop website (https://lis.ei.tum.de/mad2013), which contains the relevant materials and presentation slides.

MAD-2013 participants.

Rainer Leupers
RWTH Aachen University, Germany

2nd COMPILER, ARCHITECTURE AND TOOLS CONFERENCE, HAIFA

170 attendees gathered from around the world for the two-day HiPEAC sponsored event

Following the successful event last year, researchers and industry members from around the world gathered again in Haifa to attend the second Compiler, Architecture and Tools conference. This year the event spanned two full days, November 18th and 19th, featuring 19 talks by speakers from Spain, Poland, Sweden, UK, Switzerland, France, Mexico, Brazil, US, Russia and Israel. The program consisted of six sessions on Optimizing GPGPUs, Synthesis and SIMD, New Architectures, Programming Languages, Runtime and Performance, and Hardware-Software Co-Optimization, with lively breaks in between for refreshments and lunch. A new Posters and Ice-cream session was introduced this year, held at the top floor terrace of the Intel Development Center, closing the first day of the conference. Among the main highlights of the event were four keynote talks delivered by four distinguished guests from Europe and the US: Prof. Antonio Gonzalez from Intel Barcelona Research Center and UPC (Spain) presented his vision on the next revolution in computing, one that is likely to come from the bottom and in which computing will expand to be present everywhere. Prof. Uri Weiser from the Technion (Israel) talked about the changes in the computing trends that are likely to drive on-die heterogenous designs and the introduction of on-die memristors that will change the way we architect our computing elements. Prof. Kathryn McKinley from Microsoft Research (USA) shared her insights on the software aspect of facing hardware heterogeneity and the necessity of new programming and system abstractions for establishing a parallel heterogeneous ecosystem in the post-Dennard era. And Prof. Michael O'Boyle from the University of Edinburgh (UK) explored in his talk why compiler-based code optimizations have often failed to deliver in the past and how we should let go of certainty, focus on data-driven approaches including machine-learning, and embrace the challenges of the heterogenous many-core era. A total of 170 guests attended the event, coming from more than a dozen countries. More than 80% of the attendees were from industry and nearly 20% from academia. Very positive feedback was received on the technical content, participation, and organization. The event was organized jointly by Ayal Zaks, Ohad Shacham, Michael Behar and Gadi Haber from Intel Haifa, Dorit Nuzman from IBM Haifa, and Erez Petrank from the Technion. It took place at the "Green" IDC9 building (LEED® Gold certified), was hosted by Intel's Software and Services and Architecture Groups (SSGi, IDGz), and endorsed by HiPEAC.

Prof. Antonio Gonzalez giving the opening keynote

Conference organizers

Please visit http://software.intel.com/compilerconf2013 for more information.

Ayal Zaks, Ohad Shacham*, Dorit Nuzman^, *Intel / ^IBM, Israel*

SEMINAR ON PARALLEL PROCESSING FOR ENERGY EFFICIENCY (PP4EE)

An international seminar on Parallel Processing for Energy Efficiency (PP4EE) took place at the Norwegian University of Science and Technology (NTNU) on October 3, 2013.

PP4EE was a one-day seminar that gathered a selection of excellent researchers to discuss trends and needs in energy-efficient processor design. The program covered the main abstraction levels, from applications down to parallel programming models and multicore architectures. The event was attended by more than 80 participants, forming an active audience comprising students, faculty members and representatives from industry. The seminar was organized by the Computer Architecture and Design research group (CARD) at NTNU together with NTNU HPC-section, ARM Norway and the EMECS Erasmus Mundus. The main presenters included Hiroshi Okuda (University of Tokyo), Georgios Goumas (National Techn. University of Athens), Magnus Jahre (NTNU), Ana Varbanescu (University of Amsterdam), Javed Absar (ARM), Juan M. Cebrián (NTNU), Guillermo Miranda (Universitat Politècnica de Catalunya and Barcelona Supercomputing Centre) and Jan Chr. Meyer (NTNU). The main goal of the seminar was to stimulate collaborative research involving partners from several institutions, bridging a variety of research topics. In addition, the participants had the chance to get an overall idea of the challenges in future heterogeneous multicore designs. Most of the presentations are available at the seminar's webpage and on high-quality video stream (see link below). The presentations are also part of a distinguished lecture series given to students at the very successful EU-funded Erasmus Mundus International Master Program in Embedded Computing Systems (EMECS). This program is a cooperation between NTNU and the universities in Kaiserslautern and Southampton.

For more information, see

- https://research.idi.ntnu.no/multicore/seminar http://video.adm.ntnu.no/openVideo/serier/52526a41cfbad
- http://www.ntnu.edu/ime/eecs

Juan Manuel Cebrian, Per Gunnar Kjeldsberg and Lasse Natvig, NTNU, Norway

NTNU at night. Photo by Igor Barros Barbosa

ACACES 2014: JULY 13TH - 19TH , 2014, FIUGGI, ITALY

10th International Summer School on Advanced Computer Architecture and Compilation for High-Performance and Embedded Systems

We are proud to announce the tenth HiPEAC Summer School, which will take place in downtown Fiuggi, a historical hill town near Rome, during the third week of July. We start on Sunday evening with an opening keynote. The twelve courses start on Monday, and are spread over two morning and two afternoon slots. There are three parallel courses per slot, from which the participants can select one course. The courses have been allocated to slots in such a way that it will be possible to create a summer school program that matches your research interests.
For this tenth edition, we have exceptionally re-invited experts from the previous editions. The following world-class experts will present the topics of this year's Summer School.

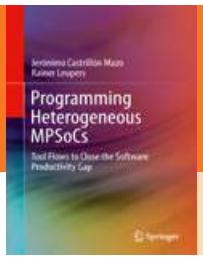
Instructor	Course
Pradip Bose - IBM T.J. Watson, USA	Energy-Efficient Resilience in Next Generation Systems
Derek Chiou - Microsoft/University of Texas at Austin, USA	Fast and Accurate Simulators of Computer Systems
Jan Coppens - iMinds - Incubation & Entrepreneurship, Belgium	Opportunity Recognition & Business Modelling
Erik Hagersten - Uppsala University, Sweden	High-Performance and Low-Energy Design, Modeling and Use of the Memory Hierarchy
Mike Hind - IBM T.J. Watson, USA	Changing the Foundation: How the Multicore Era Has Impacted Software and What the Future Holds
Christos Kozyrakis - Stanford University, USA	Resource Efficient Cloud Computing
Avi Mendelson - Technion, Israel	From Virtualization to Data Centers
Walid Najjar - University of California Riverside, USA	Beyond Streaming - New Horizons for Reconfigurable Computing
Dan Sorin - Duke University, USA	Fault Tolerant Computer Architecture
Theodore Ts'o- Google, USA	File Systems and Storage Technologies
David Whalley - Florida State University, USA	Architectural and Compilation Techniques for Addressing Embedded Design Constraints
Ayal Zaks - Intel/Technion, Israel	Heterogenous Compilation using OpenCL

On Wednesday afternoon, participants are given the opportunity to present their own work to other participants during a huge poster session. Finally, on Friday evening there will be a farewell dinner and party.
Accommodation will be provided by a consortium of hotels in Fiuggi, all located closely together. There will be abundant Italian food, and the town of Fiuggi will provide plenty of opportunities to socialize in the evenings. At the end of the event, all participants will receive a certificate of attendance detailing the courses they took. If you are a student member of HiPEAC, you can apply for a grant that covers the registration fee. In this newsletter, you will also find a summer school poster. Please post it at some visible place in your department.

You can find more information about the summer school at http://www.hipeac.net/summerschool.

We look forward to seeing you there!

Koen De Bosschere
Summer school organizer



NEW BOOK: "PROGRAMMING HETEROGENEOUS MPSOCS: TOOL FLOWS TO CLOSE THE SOFTWARE PRODUCTIVITY GAP"

Jeronimo Castrillon and Rainer Leupers

The book introduces and provides a detailed description of the problems and challenges of programming heterogeneous Multi-Processor Systems on Chip (MPSoCs). It then presents four tool flows that automate time-consuming programming tasks in order to close the gap between ever increasing software requirements and programmer productivity. The tool flows are demonstrated on case studies with applications from the multimedia and baseband processing domains and on representative heterogeneous hardware platforms.

The first tool flow presented in the book addresses the problems of extracting parallelism from sequential applications written using the C programming language. This includes the means for detecting different kinds of parallelism (e.g., task, data and pipeline level parallelism) and selecting the most effective parallelization strategy. The second flow deals with analyzing and mapping parallel applications written as Kahn Process Networks (KPNs) while considering various constraints (e.g., real time and resource constraints). An extension of this flow is then described that allows the use of dedicated hardware and optimized software routines when mapping applications from the Software Defined Radio (SDR) domain. Finally, the fourth flow serves at analyzing and producing mapping configurations for multiple applications at design time.

The tool flows, methodologies and algorithms presented in the book are thoroughly described and motivated through relevant examples. As target systems, mostly Virtual Platforms are used, which allows the consideration of platform features of future MPSoCs (e.g., tiled systems and MPSoCs with a hardware-supported runtime system). For validation purposes a commercial platform is also used. This book is the result of over six years of research on programming methodologies and tools at the chair for Software for Systems on Silicon at RWTH Aachen University. The techniques exposed in the book are integrated in the so-called MPSoC Application Programming Studio (MAPS) of the same institution.

Jeronimo Castrillon, Rainer Leupers
RWTH Aachen University, Germany

NEW BOOK: "DESIGNING 2D AND 3D NETWORK-ON-CHIP ARCHITECTURES"

K. Tatas, K. Siozios, D. Soudris and A. Jantsch

Moore's law continues unabated and new design challenges lead to new design methodologies and even paradigm shifts. One such recent development is the introduction of three-dimensional (3D) integration technology. Efficiently utilizing novel technologies poses new design challenges and therefore requires novel design methodologies and EDA tools. Training engineers in these methodologies and design techniques is mostly done at the graduate level and once these technologies become the established paradigm, at the undergraduate level.

Network-on-Chip technology has been a popular research topic for a while now, and is the current design paradigm for multi- and many-core architectures. It is also a natural complement to 3D integration technology. Its multi-faceted and multidisciplinary nature imposes a number of challenges, in both industrial and academic environments.

While at the graduate level it is common or even preferable to use papers, case studies and assignments as the main teaching tools, at the undergraduate level a suitable textbook is indispensable. Since there is an increased need to include an introduction to Networks-on-Chip in undergraduate curricula, such a textbook is required, and has been missing from the literature for too long. At the same time, the large body of research work in the field must also be made available in an organized way for graduate students, researchers and professionals to use as reference.

Therefore, the purpose of this book, apart from being used as reference for researchers in the field on Network-on-Chips, it is also used as an undergraduate or graduate level textbook as an introduction to Network-on-Chip technology, providing students and practicing engineers with the fundamentals, as well as many details in the multiple facets of Network-on-Chip design, including recent work on 3D NoCs. For this purpose, each chapter contains a number of questions, problems and design assignments that can be used in the class or laboratory. Additionally, the book can be used as an introductory course in NoC design either at the undergraduate or graduate level. Also, the book presents four case studies from academia and industry, which can be used in a second (graduate-level) course in NoC design. Finally, the second part of the book contains proposed design projects covering all aspects of NoC design, starting with design space exploration and high-level simulation, down to the VLSI design of NoC components.

K. Tatas, K. Siozios, D. Soudris and A. Jantsch
National Technical University of Athens (NTUA), Greece

HIPEAC MINI-SABBATICAL - VALENTÍN PUENTE

Report for mini-sabbatical at the IBM T.J. Watson research center in New York, USA

From October 2012 until September 2013, I have been working as visiting researcher at the IBM T.J. Watson research center in Yorktown Heights, New York. I was collaborating with researchers in the MYRA Department, and invited by its manager, Dr. Jaime Moreno. This center's long history of past successes, with many developments of key components in today's computers, percolates through to its current working philosophy. Openness and easy access to knowledge are paramount, making it more enjoyable to work and learn here. In some ways, the center is closer to a regular University than a corporate environment.

The focus of interest of this stay was the study of the impact of emerging implementation technologies on the micro architecture of next-generation computers. The main emphasis was on 3D stacking and non-volatile memories. I was involved in the team's work in exploring how to take advantage of 3D stacking in the company's future processors. Interaction with this team, which includes both architects and technologists, provided a comprehensive and up-to-date vision of the potential and limitations of the technology. I was also working with memory strategists and technologists analyzing and understanding the true potential of upcoming non-volatile memories. There is a broad consensus about their key relevance in the medium term. Although conventional memories are still pervasive, long-term scalability limitations will oblige the adoption of this type of memories. Finally, I had the opportunity to interact with key persons in charge of next-generation processors, which provided a golden opportunity to understand and experience the development and research processes behind this kind of systems. After this enriching experience, we have agreed to keep the collaboration active and we hope that this will continue to generate value for both parties.

Overall, from both the personal and professional standpoints, it was a very beneficial stay, which provided great opportunities to share ideas with researchers. Finally, I would like to thank HiPEAC for providing two months' support during my year-long stay at IBM's T.J. Watson Research Center. I would also like to thank Dr. Jaime Moreno and all the members of the MYRA Department for their hospitality.

Valentín Puente,
University of Cantabria, Spain

HIPEAC MEMBER ELEVATED TO IEEE FELLOW

Paolo Faraboschi has recently been elevated to IEEE Fellow "for contributions to embedded processor architecture and system-on-chip technology"

Paolo Faraboschi is a Distinguished Technologist at HP Labs, where he currently works on energy-efficient servers. From 2004 to 2010 he led the HP Labs group in Barcelona on system-level modeling and simulation. From 1995 to 2003 he worked at HP Labs Cambridge as the principal architect of the Lx/ST200 family of VLIW embedded cores, widely used in consumer SoCs for video and imaging products and HP's printers. Paolo is an active member of the computer architecture community: guest co-editor of IEEE Micro TopPicks 2012, Program Chair for Computing Frontiers 2012, HiPEAC10 (2010), MICRO41 (2008) and MICRO34 (2001). He holds 24 patents and co-authored the book "Embedded Computing: a VLIW approach to architecture, compiler and tools". Before joining HP in 1994, he received a Ph.D. in EECS (1993) from the University of Genoa, Italy.

HIPEAC CO-FOUNDER RECEIVES AN ERC ADVANCED GRANT

Per Stenström of Chalmers University of Technology has received 2.3 M Euros over five years from the European Research Council to undertake a project called MECCA: Meeting Challenges in Computer Architecture. The project will commence in 2014.

For several decades, computer technology has doubled computational performance biennially. This performance growth rate has enabled the dramatic innovation in information technology that now embraces our society. Before 2004, application developers could exploit this performance growth rate with no effort. However, since 2004, power consumption of single-core computer chips exceeded the allowable limits, so from that point and onwards, parallel computer architectures became the norm. Parallelism is currently completely exposed to application developers and managing it is difficult and time-consuming. This seriously cuts down software productivity, which can stall progress in information technology.

Technology forecasts predict that by 2020 there will be hundreds of processors on a computer chip. Apart from the challenge of managing parallelism on such architectures, keeping power consumption within allowable limits will remain a key roadblock for maintaining historical performance growth rates. To avoid limiting this rate of growth, power efficiency must increase by at least an order of magnitude in the next ten years. Finally, as computer chips are universal components, they are used not only as number-crunching devices in supercomputers and data centers, but increasingly as key components in embedded controllers, in e.g. automotive applications where stringent timing responses are mandatory. Delivering predictable and tight response times using parallel architectures is a challenging and unsolved problem.

Prof. Per Stenström

MECCA takes a novel, interdisciplinary and unconventional approach to address three important challenges facing computer architecture – the three Ps: Parallelism, Power, and Predictability in a unified framework. Unlike earlier, predominantly single disciplinary approaches, MECCA bridges layers in computing systems from the programming language/model, to the compiler, to the run-time/OS, down to the architecture layer. This opens up the exchanging of information across layers and its use in novel ways to manage parallelism and architectural resources automatically "under the hood" to meet challenging performance, power, and predictability requirements for future computers systems.

The key approach will be to leverage emerging task-based programming models. The hypothesis is that by conveying domain-specific information provided through annotations (e.g., quality of service requirements and data structure usage) it is possible to enable a new class of optimizations at the compiler, run-time and at the architecture level. These optimizations can be used to offer a substantial improvement in efficiency, performance and power. This approach is also expected to open up for approaches to guarantee predictable parallel program performance, a prerequisite for using heterogeneous multi-core systems in safety-critical systems. In this holistic framework, MECCA will address critical issues for managing parallelism, architecture resources and enabling deterministic execution of parallel programs with the goal of making heterogeneous multi-core systems broadly applicable across computing segments from HPC to embedded systems.

Per Stenström is professor at Chalmers University of Technology. His research interests are in parallel computer architecture. He has authored or co-authored three textbooks and more than 130 publications in this area. He co-founded the EC funded HiPEAC Network on Excellence. He has been program chairman of the IEEE/ACM Symposium on Computer Architecture, the IEEE High-Performance Computer Architecture Symposium, and the IEEE Parallel and Distributed Processing Symposium and acts as Senior Associate Editor of ACM TACO and Associate Editor-in-Chief of JPDC. He is a Fellow of the ACM and the IEEE and member of Academia Europaea, the Royal Swedish Academy of Engineering Sciences and the Royal Spanish Academy of Engineering.

HIPEAC WELCOMES VILNIUS UNIVERSITY

A HiPEAC delegation visited Vilnius to welcome new members from Lithuania

In September we welcomed the first HiPEAC member from Lithuania, Julius Žilinskas from Vilnius University. To attract new members from Lithuania, a meeting of HiPEAC delegation and Lithuanian computing systems researchers has been organized by J. Žilinskas on 7th of November at Vilnius University Institute of Mathematics and Informatics. The members of the HiPEAC delegation were the coordinator Koen De Bosschere (Ghent University), Marc Duranton (CEA), and

Emre Özer (ARM). The presentation of the HiPEAC network of excellence was given by Koen De Bosschere, who also introduced his research programs at the Electronics and Information Systems (ELIS) department at Ghent University. Marc Duranton presented the HiPEAC Roadmap, while Emre Özer described the initiatives of ARM. Jesus Carretero presented the new COST action IC1305 Network for Sustainable Ultrascale Computing (NESUS) as well as studies and research at Computer Science and Engineering Department at University Carlos III of Madrid.

Representatives of several Lithuanian institutions described their research and studies programs. Several talks were given by representatives of Vilnius University: Director of the Institute of Mathematics and Informatics Gintautas Dzemyda reviewed activities of the institute, vice dean of the Faculty of Physics Juozas Šulskus described the applied problems solved in Lithuanian Saulėtekis Technological Valley Supercomputing Center, vice director of Digital Science and Computing Center Arūnas Stašionis presented the applications and services of this center, Kristina Lapin and Vytautas Čyras reviewed research performed by Computing Group at the Faculty of Mathematics and Informatics. Arnas Kačeniauskas reviewed infrastructures, software, and applications of the Laboratory of Parallel Computing at Vilnius Gediminas Technical University, and vice dean of the Faculty of Informatics at Kaunas University of Technology Robertas Damaševičius described research and studies of embedded computing at this university.

The talks were followed by the discussion about the opportunities that HiPEAC can offer to new members and the application procedure to join the network. The meeting resulted in several new HiPEAC members from Lithuania.

Koen De Bosschere

FIRST OFFICIAL RELEASE OF COMPSS PROGRAMMING ENVIRONMENT

BSC team releases a set of tools that helps developers to program and execute their applications efficiently on distributed computational infrastructures

The Grid Computing and Clusters team at Barcelona Supercomputing Center is proud to announce a new release, and the first official release, of the programming environment COMPSs. COMPSs is a set of tools that helps developers to program and execute their applications efficiently on distributed computational infrastructures such as clusters, grids and clouds.

COMPSs has been available for the last years to MareNostrum users and to the Spanish Supercomputing Network, and has already been adopted in several research projects, including OPTIMIS, VENUS-C, transPLANT, EUBrazilOpenBio and EGI. In these projects COMPSs has been applied to implement use cases provided by various communities across diverse disciplines as biomedicine, engineering, biodiversity, chemistry, astrophysics and earth sciences. Last year, COMPSs was downloaded more than 500 times, and it is being used by about 20 groups in real applications. COMPSs has recently attracted interest from areas such as genomics and biodiversity, where specific courses and dissemination actions have been performed.

In recent years, the team's efforts have been focussed on emerging virtualization technologies, adopted by cloud environments. In such systems, COMPSs provides scalability and elasticity features by dynamically adapting the number of resources to the actual workload.

The current release is interoperable with both public and private cloud providers, such as Amazon EC2, OpenNebula, BSC EMOTIVE Cloud, and with OCCI-compliant offerings.

The packages and the complete list of features are available in the Downloads page. A virtual appliance is also available to test the functionalities of COMPSs through a step-by-step tutorial guiding the user to develop and execute a set of example applications.

Also available are a user guide and papers published in relevant conferences and journals.

For more information on COMPSs please visit our webpage: http://compss.bsc.es

DELFT UNIVERSITY OF TECHNOLOGY AND IMEC INTRODUCE 3D-COSTAR TO OPTIMIZE TEST FLOWS OF 3D STACKED INTEGRATED CIRCUITS

Last October, Delft University of Technology (TU Delft) and nanoelectronics research center IMEC presented 3D-COSTAR, a new test flow cost modeling tool for 2.5/3D stacked integrated circuits (ICs). 3D-COSTAR aims to optimize the test flow of 3D stacked ICs (SICs), taking into account the yields and costs of design, manufacturing, packaging, test, and logistics.

Due to its many high-precision steps, semiconductor manufacturing is prone to defects. Consequently, every IC needs to undergo electrical tests to weed out defective parts and guarantee outgoing product quality to the customer. For TSV-based 2.5D- and 3D-SICs, which typically contain complex die designs in advanced technology nodes, testing is even more critical. In addition, there are many possible test moments in their manufacturing flow: pre-bond (before stacking), mid-bond (on a partial stack), post-bond (on a completed stack), and final testing (on a packaged device).

Although testing is expensive, filtering out the bad components in an early stage is critical to save costs later on in the production process. "There is not a 'one-size-fits-all' test flow that covers all stacked-die products. The test flow needs to be optimized based on yield and cost parameters of an individual product and that is a complex optimization problem," stated Dr. Said Hamdioui, Associate Professor at TU Delft. "And different test flows, executed after manufacturing, may require different design-for-test features, which need to be incorporated in the various dies during their early design stages."

3D-COSTAR uses input parameters that cover the entire 2.5D-/3D-SIC production flow: 1) design; 2) manufacturing; 3) test; 4) packaging; and 5) logistics. It is aware of the stack build-up (2.5D versus 3D, multiple towers; face-to-face or face-to-back) and stacking process (die-to-die, die-to-wafer, or wafer-to-wafer). The tool produces three key analysis parameters: 1) product quality, expressed as defect level (test escape rate) in DPPM (defective parts per million); 2) overall stack cost; and 3) breakdown per cost type.

"3D-COSTAR has proven to be a crucial tool to analyze the many complex trade-offs in 3D test flows, in terms of both cost and DPPM," said Erik Jan Marinissen, Principal Scientist at IMEC. "Among others, we have used 3D-COSTAR to determine when pre-bond testing of the interposer in 2.5D-SICs pays off and what its maximum-allowed test cost can be. In some cases, the overall stack cost reduction amounts to 40%, showcasing that upstream testing can help avoid downstream costs. The tool also demonstrated under which circumstances mid-bond testing (on partially-completed stacks) can be avoided without compromising a high stack yield." "Together with IMEC, Cascade Microtech has recently demonstrated the feasibility of direct probing of large-array fine-pitch microbumps to avoid the usage of dedicated pre-bond pads," stated Ken Smith, Principal Engineer, Cascade Microtech, Inc. (of Beaverton, Oregon, USA). "Analysis with 3D-COSTAR clearly showed up to 50% overall cost benefit of doing microbump probing using an advanced probe cell such as was demonstrated with Pyramid Probe® RBI technology on our CM300 probe station."

Mottaqiallah Taouil, Said Hamdioui*, Erik Jan Marinissen^, *Delft University of Technology (the Netherlands), ^IMEC (Belgium)*

TALLINN - HIPEAC COMPUTING SYSTEMS WEEK

Last HiPEAC Computing Systems Week of this year was held in Tallinn. It started with several talks about dependability, reliability and fault tolerance activity in on-going FP7 projects in Europe. Later there was a slot about timing analysis for multicore/manycore architectures. We got a brief insight of the new EU funding programme for research and innovation – Horizon 2020 together with ways it continues the FP7 project. Last day I attended the meetings related to heterogeneous systems and their applications and predictable and real-time embedded systems.

During this workshop I had an opportunity to talk directly to other researchers about things that I had problems with and also to hear what others are doing. Presentations from other speakers gave me an insight what other research groups are working on and gave me some new ideas how to cooperate together to resolve the problems that we are facing.
One presentation specially intrigued me, it was about novel ways of learning using "flipped classroom" effect with interactive lectures that are listened at home and teaching part of the lectures where students are working in pairs and solving problems. It showed good results in Sweden and hopefully it will be used in the rest of the world.

Evening social event in the Lennusadam Seaplane Harbour Museum was organized very nicely. Watching pictures from Estonian nature in an informal atmosphere and beautiful ambient of the museum surrounded by ships, submarine and airplanes was a perfect relaxation after the workshop.
This way I would like to thank HiPEAC for helping me as a member of new member states to come to this workshop and I encourage other researcher from NMS to use this opportunity and join us on upcoming events.

Daniel Hofman,
University of Zagreb

BSC RESEARCHERS AWARDED SC13 BEST STUDENT PAPER

Supercomputing with Commodity CPUs: Are Mobile SoCs Ready for HPC?

The Heterogeneous Architectures team of the Computer Sciences Department of the Barcelona Supercomputing Center (BSC) has been awarded the SC13 best student paper, announced in the awards ceremony at the International Conference for High Performance Computing, Networking, Storage and Analysis (SC13), held in Denver from 17-22nd November. Each year, the SC technical papers committee identifies one paper as the best technical article written primarily by a student and presented in the conference's technical program. Papers must be identified as student papers when submitted to be eligible for this award.

On Tuesday, Nikola Rajovic, PhD student from the Heterogeneous Architectures team of the BSC Computer Sciences Department presented the scientific paper titled "Supercomputing with Commodity CPUs: Are Mobile SoCs Ready for HPC?" The paper shows that mobile processors, used in smartphones and tablets, have promising qualities to replace current technologies, and become mainstream in High Performance Computing in the near future. Provided they implement a reduced set of extra features not needed by mobile devices, they could lead to faster, cheaper, and more energy-efficient supercomputers. The awarded paper is a research result of the European funded Mont-Blanc project, coordinated by BSC, whose aim is to design a new type of computer architecture capable of setting future global HPC standards, built from today's energy efficient solutions used in embedded and mobile devices. This award values the impressive dedication and effort of the BSC Computer Sciences researchers, as well as confirming the European research quality in leading research projects that can influence the future of the HPC market.

INTERNSHIP REPORT - JOSÉ M. CECILIA

Implementation of the Ant Colony Optimisation algorithm on emergent massively parallel processors

Thanks to the HiPEAC collaboration grants program, last summer I worked with Prof. Wen-Mei W. Hwu and his research group (the IMPACT group) in the Department of Electrical and Computer Engineering (ECE) at University of Illinois (USA). The objective of IMPACT is to provide critical research, architecture innovation, and algorithm and compiler prototypes for heterogeneous parallel architectures. The group achieves portable performance and energy efficiency for emerging real-world applications by developing novel hardware, compiler, and algorithmic solutions.

The work performed during our collaboration was focused on the efficient implementation of the ACO algorithm on emergent massively parallel processors. Ant Colony Optimisation (ACO) is a population-based meta-heuristic that has been successfully applied to many NP-complete problem. As a population-based algorithm, its computation is intrinsically massively parallel, so it is theoretically well suited to be implemented on massively parallel architectures. However, the parallel designs for the ACO algorithm proposed so far on these architectures are either task-based approaches, which compromise GPU parallelism, or highly computationally demanding to avoid serialization.

In this research project, we rethink the parallelization strategies for the ACO algorithm for solving the Travelling Salesman Problem (TSP) on massively parallel architectures. We focus our efforts on the tour construction stage, which is the most challenging part of this algorithm. A vectorization-based design developed using basic parallel primitives, such as prefix scan, stencil and reduction, which are better suited to the GPU parallelism model, is used to enhance performance and decrease the arithmetic intensity. The state of the work is under minor corrections and will be ready to be submitted to a conference before the end of January. We are also working on comparing the most recent generations of accelerators from Nvidia (Kepler K20x and K40) and Intel Xeon Phi for this kind of novel computational pattern.

Personally, I am grateful to Prof. Wen-Mei Hwu and his group for a very pleasant internship and the new collaboration opportunities they offer, especially to my new friends Nacho Navarro and Javier Cabezas for introducing me to the American culture and food; Li-Wen Chang who spent some time with me sharing clever ideas about the implementation and Izzat El Hajj, Hee-Seok Kim and Christopher Rodrigues for sharing their lab, ideas and tea with me. I am also grateful to my advisor José M. García for helping me in taking this decision. Finally, I am grateful to HiPEAC for providing the opportunity to participate in this internship and also for supporting the collaboration.

FP7 SAVE PROJECT

Software/Hardware technologies for implementing and exploiting heterogeneous system architectures using self-adaptiveness and virtualization for energy/performance optimization

Project name: SAVE (Self-Adaptive Virtualisation-Aware High-Performance/Low-Energy Heterogeneous System Architectures)
Coordinator: Politecnico di Milano, Italy
Partners: Politecnico di Milano, Italy
STMicroelectronics, France
Technological Educational Institute of Crete, Greece
Virtual Open Systems, France
Maxeler Technologies, UK
ARM Ltd., UK
University of Paderborn, Germany
Start date: September 2013
Duration: 36 months
Website: http://www.fp7-save.eu

The global ICT footprint is expected to more than triple by 2020, caused by data centres that provide information at our fingertips and the mobile devices we use to access it. Nowadays, high performance systems are designed to serve rather static workloads with high-performance requirements, although we are moving towards a highly flexible, on-demand computing scenario characterized by varying workloads, constituted by diverse applications with different performance requirements and criticality. This mismatch between demand and supply of computing power causes high energy dissipation.

A promising approach to address the challenges posed by this scenario is to better exploit specialised computing resources integrated in a heterogeneous system architecture by taking advantage of their individual characteristics to optimise the whole system performance/energy trade-off.

Heterogeneity, however, comes at the cost of greater complexity. System architects need, for example, to consider the efficiency of the various computational resources as well as the application workload, which often leads to inefficiency in resource exploitation and a suboptimal performance/energy trade-off.

SAVE addresses these limitations by exploiting self-adaptivity and hardware-assisted virtualisation to allow the system to autonomously decide which specialised computing resources are exploited to achieve the most efficient execution, based on user-defined optimisation goals, such as performance, energy, reliability, as well as seeking to ease the exploitation of resources heterogeneity.

SAVE defines crosscutting SW/HW technologies for implementing self-adaptive systems exploiting GPUs and FPGA-based dataflow engines (DFEs) that enhance heterogeneous architectures to cope with the increased variety and dynamics of high-performance and embedded computing workloads.

Virtualisation and self-adaptation are jointly exploited to obtain a new self-adaptive virtualisation-aware Heterogeneous System Architecture.

To this end, SAVE will develop:

ADVANCED RUN-TIME SELF-ADAPTIVENESS OS SUPPORT LAYER
implementing a "smart" orchestrator as part of the host OS able to dynamically and seamlessly partition and distribute the various tasks on the available resources, based on changeable workloads and/or optimisation goals (optimise performance for a given energy budget, or minimise energy consumption without violating the expected QoS).

HARDWARE-ASSISTED VIRTUALIZATION SUPPORT FOR GPUs/DFEs
overcoming the limitations of today's solutions (e.g., performance overhead, limited control) enabling hardware virtualisation for GPUs and DFEs, expected to become more widely available as required by HPC and ES use cases.

HYPERVISOR EXTENSIONS
exposing a virtual computation interface to applications and allowing the hypervisor to handle GPUs and DFEs as first-class schedulable entities.

Partners at the kick-off meeting of SAVE.

FP7 STREP HARPA PROJECT

Project name: HARPA (Harnessing Performance Variability)
Coordinator: Prof. William Fornaciari, Politecnico di Milano
Partners: Politecnico di Milano, Italy
IMEC, Belgium)
ICCS/NTUA, Greece
UCY, Cyprus
IT4I, Czech Republic
THALES, France
HENESIS, Italy
Start date: September 2013
Duration: 36 months
Website: www.harpa-project.eu

GOAL AND CHALLENGES

The overall goal of the HARPA project is to provide architectures (both Embedded Systems (ES) and High Performance Computing (HPC)-oriented) with efficient mechanisms to offer performance dependability guarantees in the presence of unreliable time-dependent variations and aging throughout the lifetime of the system. This goal will be achieve by utilizing both proactive techniques (in the absence of hard failures) and reactive techniques (in the presence of hard failures).

The term "performance dependability guarantee" refers to time-criticality in ES (i.e., meeting deadlines), and, in the case of HPC, a predefined bound on the performance deviation from the nominal specifications. The promise is to achieve this reliability guarantee in both domains within a reasonable energy overhead (e.g. less than 10% average). A significant improvement is hence achieved compared to the SotA, which currently provides guarantees but with at least 50% overhead. In addition, we will provide better flexibility in the platform design while still achieving power savings of at least 20%. To the best of our knowledge, this is the first project to attempt a holistic approach for providing dependable performance guarantees on both ES and HPC systems. This is done while taking into account various non-functional factors, such as timing, reliability, power, and ageing effects. The HARPA project aims to address several scientific challenges in this direction:

(i) Shaving margins. Similar to the circuit technique Razor, but with different techniques at the microarchitecture and middleware, our aim is to introduce margin shaving concepts into aspects of a system that are typically over-provisioned for the worst case.

(ii) A more predictable system with real-time guarantees, where needed. The different monitors, knobs, and the HARPA engine will make the target system more predictable and proactively act on performance variability prior to hard failures.

(iii) Implementation of effective platform monitors and knobs. HARPA will select the appropriate monitors and knobs and their correct implementation to reduce efficiency and performance overheads.

TECHNICAL APPROACH: HARPA ENGINE OVERVIEW

Figure 1(a) below shows the main concepts of the HARPA architecture and the main components of an architecture that can provide performance-dependability guarantees. Note that this generic framework applies to both embedded systems and high-performance general-purpose systems. The main elements that distinguish a HARPA-enabled system are: (i) monitors and knobs, (ii) user requirements and (iii) the HARPA Engine. Conceptually, the HARPA Engine mainly consists of a feedback loop (Figure 1(b)), where the different metrics (performance, timing, power, temperature, errors and manifestations of time-dependent variations etc.) of the system are continuously monitored. The HARPA engine actuates the knobs to bias the execution flow as desired, based on the state of the system and the performance requirements (timing/throughput) of the application.

The concepts that are to be developed within the HARPA context address equally both the HPC and ES domains. More specifically, from the HPC domain we will use Disaster and Flood Management Simulation, while from the ES domain we will use applications for Radio frequency spectrum sensing, Face Detection, Object Recognition, and Real-time data monitoring for wearable human motion acquisition system.

Project Coordinator, Prof. William Fornaciari, Politecnico di Milano

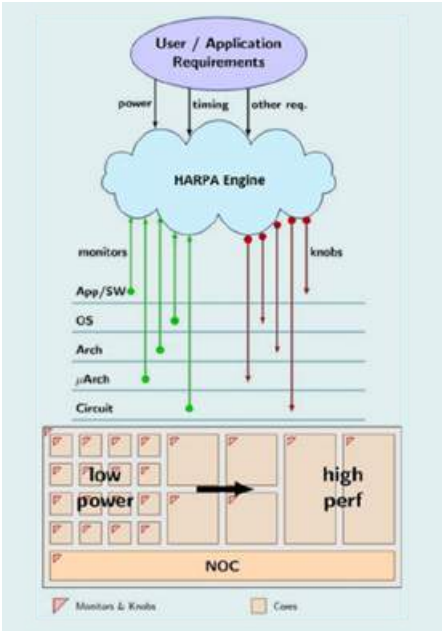


Figure 1a: HARPA High-Level View

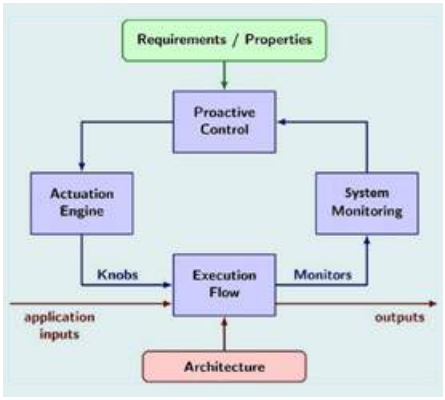
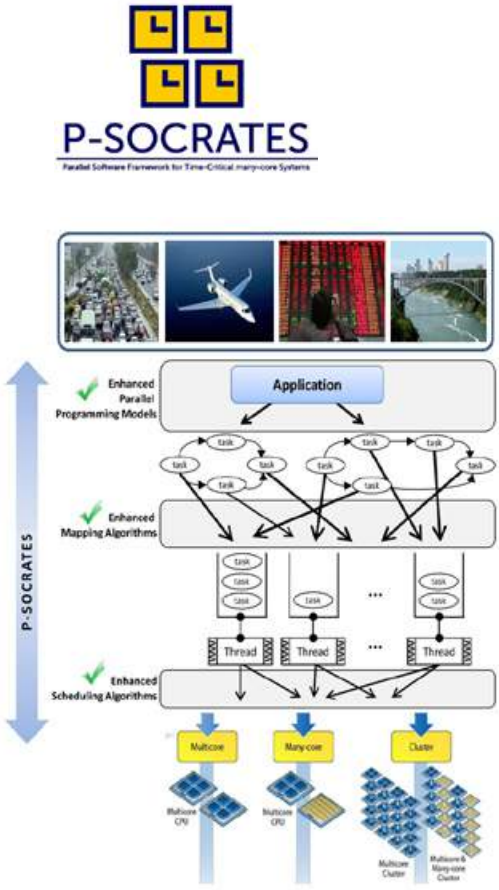


Figure 1b: HARPA Engine

FP7 P-SOCRATES PROJECT

This EU project will develop a novel software infrastructure to allow time-critical systems to exploit the performance opportunities of the newest processor architectures

Project name: P-SOCRATES (Parallel Software Framework for Time-Critical Many-core Systems)
Coordinator: Dr. Luis Miguel Pinho, Instituto Superior de Engenharia do Porto
Partners: Instituto Superior de Engenharia do Porto, Portugal
Barcelona Supercomputing Center, Spain
Universitá di Modena e Reggio Emilia, Italy
ETH Zurich, Switzerland
Evidence, Italy
Active Technologies, Italy
Atos, Spain
Start date: October 2013
Duration: 36 months
Website: http://www.p-socrates.eu



P-SOCRATES (Parallel Software Framework for Time-Critical Many-core Systems) will allow current and future applications with high-performance and real-time requirements to fully exploit the huge performance opportunities brought by the most advanced many-core processors, whilst ensuring a predictable performance and maintaining (or even reducing) development costs of applications. The technology developed within P-SOCRATES will be evaluated within different system scenarios provided by project advisory board members.

"There is a real necessity for new software development tools for parallel systems", said Dr. Claudio Scordino, of Evidence, the SME responsible for the operating system development. "P-SOCRATES will provide an entirely new generic design framework, from the conceptual design of the system functionality to its physical implementation, to facilitate the deployment of standardized parallel architectures in all kinds of systems", he concluded.

"The computing technology developed in the project will allow a deeper understanding of many-core off-the-shelf processors, enabling new kinds of applications to be developed on top of such platforms", added Dr. Michele Ramponi, of Active Technologies.

Nowadays, the prevalence of electronic and computing systems with time-critical requirements is rapidly increasing. Example applications include cars and airplanes, smart grids and traffic management, business monitoring, etc.

All these systems demand more and more computational performance to process large amounts of data from multiple data sources with guaranteed processing response times.

As a result, the computer electronic devices on which these systems depend are constantly required to become more and more powerful and reliable, while remaining affordable. Many-core processor architectures allow these performance requirements to be achieved, by integrating dozens or hundreds of cores, interconnected with complex networks on chip. This radical shift in chip design paved the way for parallel computing.

Unfortunately, parallelization brings many challenges, by drastically affecting the system's timing behavior. Therefore, although multi-core processors are promising candidates for improving the responsiveness of these systems, providing timing guarantees becomes harder, because the timing behavior of the system running on a multi-core processor depends on interactions that are usually not known by the system designer. This causes system analysts to struggle to provide timing guarantees for such platforms.

"P-SOCRATES brings together world-class researchers, with the complementary expertise required to tackle these complex challenges. P-SOCRATES will represent a reference point for the implementation of workload-intensive applications with time-criticality requirements", said Dr. Luis Miguel Pinho, the Project Coordinator.

The P-SOCRATES consortium exploit synergies and strengths between different computing segments to successfully exploit the performance opportunities brought by parallel programming models used in the high-performance domain (e.g. OpenMP, OpenCL) and the newest many-core processors currently available in the market, while providing timing guarantees. To that end, the project brings together teams from the high-performance (Barcelona Supercomputer Center and Atos), embedded (Swiss Federal Institute of Technology Zurich and Active Technologies) and real-time (Instituto Superior de Engenharia do Porto, University of Modena and Evidence) computing segments, in a stack that ranges from the application provider to the hardware manufacturer.

FP7 PARADIME PROJECT

Parallel Distributed Infrastructure for Minimization of Energy

Project name: paraDIME (Parallel Distributed Infrastructure for Minimization of Energy)
Coordinator: Dr. Adrián Cristal, Barcelona Supercomputing Center
Partners: Barcelona Supercomputing Center, Spain
IMEC, Belgium
TU Dresden, Germany
Université de Neuchâtel, Switzerland
AoTerra GmbH, Germany
Start date: October 2012
Duration: 36 months
Website: www.paradime-project.eu

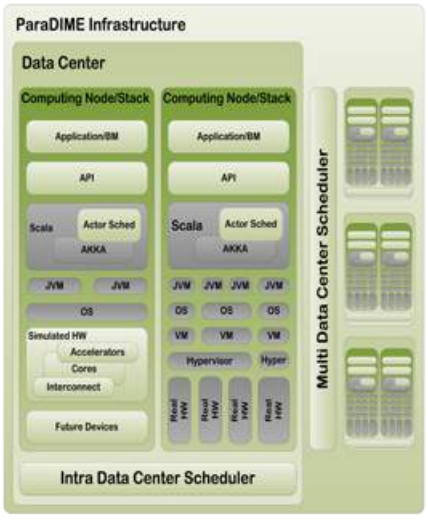
The ParaDIME Project is about using the most energy-efficient computing techniques and taking the best advantage of renewable energy sources in order to radically reduce overall energy consumption, beginning with the chip in your phone and working all the way up to the data center that is streaming your favorite television program to your home computer. Ultimately, however, it is about saving money. The increasing power and energy consumption of modern computing devices is perhaps the largest threat to technology minimization and the associated gains in performance and productivity. On the one hand, we expect technology scaling to finally come face-to-face with the problem of "dark silicon" (only segments of a chip can function concurrently due to power restrictions), which will push us to use devices with completely new characteristics. On the other hand, as core counts increase, the shared memory model based on cache coherence will severely limit code scalability and increase energy consumption. Therefore, to overcome these problems, we need new computing paradigms that are radically more energy efficient. These problems of energy consumption scale beyond the computing node to the data center, where consumption is independent of the computing load of the system. The nodes keep state in memory and on local disk, which means that they cannot be turned off even if the load is low. Moreover, making individual servers energy-proportional requires architectural changes.

In ParaDIME, we will explore scheduling between the computing nodes by employing a mechanism for maintaining the state that allows nodes to be switched off nodes when their load is low. At the same time, Data Centers have high power consumption even when they are idle because CPU loads are often kept between 10% and 30% so that they can react to sudden peaks in load. In ParaDIME, we will employ a mechanism that raises the CPU load to 90% while at the same time adhering to Service Level Agreements.

Taken a step further, there is little scholarly work that analyzes energy consumption and the distribution of computation across multiple data centers. In particular, scheduling in cloud computing does not usually take carbon footprint into consideration. In ParaDIME, we schedule tasks across multiple data centers that have the least ecological impact per VMh, taking into account the current power mix (of renewable vs. non-renewable energy sources) as well as their current power usage efficiency. Our work will build upon recent advancements in replicating data across distant sites.

The main objectives are: Objective 1: To develop an energy-aware programming model driving an associated ecosystem, the ParaDIME Computing Node / Stack (applications, runtime and architecture) that combines energy efficient SW programming and HW design methodologies and utilizes new emerging devices at the limit of CMOS scaling to radically decrease energy consumption; to quantify the energy savings from employing these methodologies by running several real-world power-hungry applications to stress test the computing node / stack. Objective 2: To build a reference Data Center (Infrastructure as a Service or IaaS) platform that incorporates new energy conscious workload scheduling techniques utilizing information from the runtime to radically decrease energy consumption; to quantify the energy savings from employing these techniques by running several real-world power-hungry applications to stress test this Data Center platform.

The outcomes of the project will be: a roadmap that indicates the most promising individual or combination of energy efficient hardware and software methodologies by quantifying the energy cost / benefit for the Computing Node / Stack and (possibly) the Data Center. A reference Individual and Multiple Data Center Infrastructure that will help European companies develop "beyond green" technology product offerings based on the most promising energy efficient computing methodologies.



FP7 PROJECT CLERECO: CROSS-LAYER EARLY RELIABILITY EVALUATION FOR THE COMPUTING CONTINUUM

Early Reliability Evaluation for Embedded and High-Performance Computing Systems in the Computing Continuum

Project name: Clereco (Cross-Layer Early Reliability Evaluation for the Computing cOntinuum)
Coordinator: Politecnico di Torino
Partners: Politecnico di Torino, Italy
University of Athens, Greece
Centre National de la Recherche Scientifique – LIRMM, France
Intel Labs Barcelona, Spain
Yogitech SpA, Italy
Thales, France
ABB, Norway
Start date: October 2013
Duration: 36 months
Website: http://www.clereco.eu/

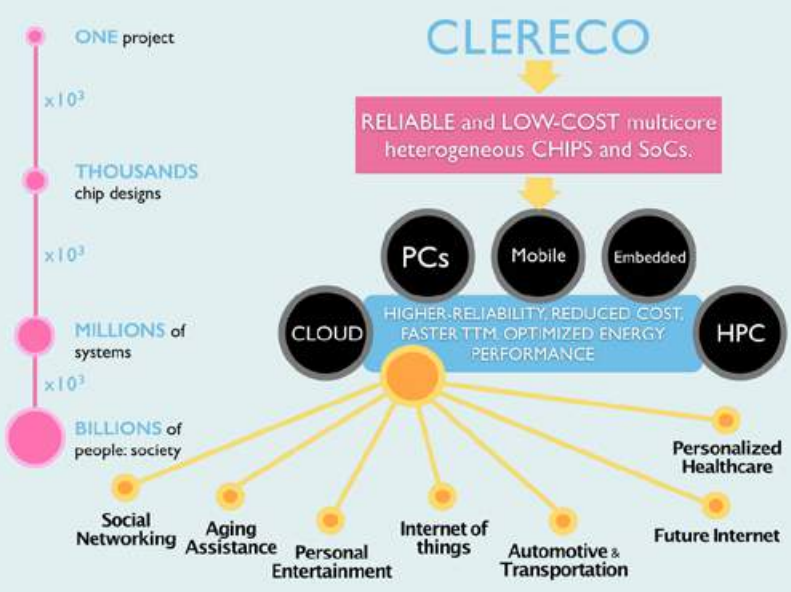
FP7 Collaboration Project CLERECO aims to investigate new design methods for early reliability evaluation of digital systems in the forthcoming computing continuum. Advanced multifunctional computing systems based on future manufacturing technologies hold the promise of a significant increase in the computational capability that will offer end-users ever-improving services and functionalities (next generation mobile devices, cloud services, etc.).Reliability of electronic systems will become an ever-increasing challenge for information and communication technologies and must be guaranteed without penalizing or slowing down the characteristics of the final products.

The CLERECO research project recognizes the importance of accurately evaluating the reliability of systems early in the process to be one of the most important and challenging tasks toward this goal. Being able to precisely evaluate the reliability of a system means being able to carefully plan for specific countermeasures rather than resorting to worst-case approaches.

The CLERECO project will be fundamental in the development of scaled systems for the next decade. The CLERECO framework for efficient reliability evaluation and therefore efficient exploitation of reliability oriented design approaches starting with the earliest phases of the design process will enable circuit integration to continue at exponential rates. It will enable the design and manufacture of future systems for the computing continuum at a minimum cost, contrary to existing worst-case design solutions for reliability.

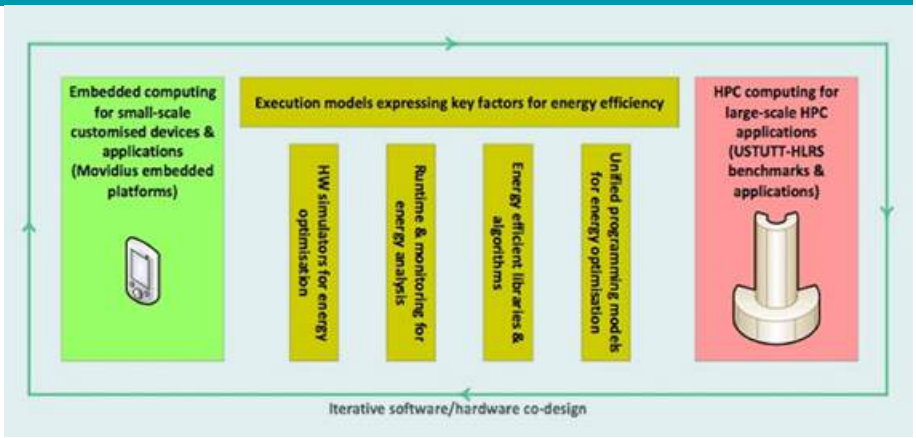
Applications of such chips will play a major role in our society and can be seen through the prism of future computing systems ranging from avionics, automobile, smartphones, mobile and embedded systems, PCs, and future servers utilized in the settings of Data Centers, Grid Computing, Cloud Computing and other types of HPC systems.

The CLERECO project consortium includes: (1) Politecnico di Torino from Italy (the TestGroup of the Department of Computer and Control Engineering) which acts as the project coordinator, (2) University of Athens from Greece (the Computer Architecture Lab of the Department of Informatics & Telecommunications), (3) Centre National de la Recherche Scientifique - Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier from France, (4) Intel Labs Barcelona from Spain (Architecture and Design Research group), (5) Thales SA from France, (6) Yogitech S.P.A. from Italy and (7) ABB AS from Norway (Corporate Research).



FP7 EXCESS PROJECT

Project name: EXCESS (Execution Models for Energy-Efficient Computing Systems)
Coordinator: Philippas Tsigas, Chalmers University of Technology
Partners: Chalmers University of Technology, Sweden
Linköping University, Sweden
Movidius Ltd., Ireland
Tromsø University, Norway
Stuttgart University, Germany
Start date: September 2013
Duration: 36 months
Website: http://www.excess-project.eu



Computing technology is currently at the beginning of a disruptive transition from petascale to exascale (2010 – 2020), posing a great challenge on energy efficiency. Future computing systems will be constrained by power and energy, and will feature massive parallelism, heterogeneity, and heavy exploitation of data locality.
Existing approaches to system design tend to focus on individual layers of the system stack, neglecting potential improvements in performance and energy that could have been leveraged by coordinated system-wide optimisation. The three-year European FP7 project EXCESS, which started in October 2013, aims to provide radically new energy execution models, forming foundations for energy-efficient computing paradigms that will enable two orders of magnitude improvements in energy efficiency for computing systems.
EXCESS plans to develop a clear understanding of where energy-performance is wasted and a continuous process to reduce this energy waste, leading to significant improvements in energy efficiency for computing systems. There are currently no power complexity models of computation that are relevant to computer engineers, and the lack of holistic integrated approaches limits the exploitation of existing solutions and their reduction in energy consumption.
For instance, although some chips implement mechanisms to control energy consumption, most of them are ignored in practice because it is hard for the programmer to use them. As a result, software power optimisation is often applied only as a secondary step, once the application is written, thereby preventing the software engineer from making power-aware architectural decisions early in the development flow, where the majority of power savings can be achieved.
To reach the next level of energy efficiency, the interaction of hardware and software needs to be optimised through an iterative software/hardware co-design process. EXCESS is addressing the challenge of developing power complexity models of computation that are relevant to computer engineers via a holistic approach that will examine energy losses also at software and hardware layer matching.
In order to synthesize energy optimisation in embedded computing and performance optimisation in High Performance Computing, EXCESS will bridge the gaps between those two domains through novel models that allow for holistic optimisation across all layers of the computer system. The model's "bridge" is supported by the four "pillars" shown in the picture.

Coordinated by the Chalmers University of Technology, EXCESS unites forces from both High Performance Computing and embedded computing: universities of Linköping and Tromsø, the High Performance Computing centre of HLRS at the University of Stuttgart, and a European embedded multi-core SME, Movidius.

METHODOLOGY TO EVALUATE PERFORMANCE OF THE I/O SYSTEMS ON HIGH PERFORMANCE COMPUTERS

Sandra Méndez
Universitat Autònoma de Barcelona, Spain
Advisor: Dra. Dolores Rexachs
July 2013

In this thesis, we have proposed a methodology to analyse performance of the Input/Output systems of High Performance Computers considering the behavior of parallel scientific applications. We represent the I/O characteristics of the application in an I/O model and we analyze the relationship with the I/O system configuration performance. We have extracted the application's I/O patterns and extensively evaluated different I/O systems against a wide variety of I/O characteristics to simulate the user I/O patterns. From this analysis, we can compare the behavior and give information to help select the most appropriate I/O system configuration.

HIGH-PERFORMANCE AND FAULT-TOLERANT TECHNIQUES FOR MASSIVE DATA DISTRIBUTION IN ONLINE COMMUNITIES

Daniel Higuero
Universidad Carlos III de Madrid, Spain
Advisors: Dr. Jesus Carretero and Dr. Florin Isaila
July 2013

In this work we provide a solution for efficient and reliable data distribution in a geographically distributed scenario. We focus on a solution that 1) optimizes resource utilization, 2) does not require infrastructure changes, and 3) provides fault-tolerant capabilities. Our solution is composed of three components. First, a community detection module groups users that request similar data. Second, a transfer scheduling module produces a transfer plan using a distributed linear approach to distribute the required data, improving resource utilization. Finally, a distribution controller manages the distribution process controlling a dynamic server infrastructure.

NEW HARDWARE SUPPORT FOR TRANSACTIONAL MEMORY AND PARALLEL DEBUGGING IN MULTICORE PROCESSORS

Lois Orosa
Universidade de Santiago de Compostela, Spain
Advisors: Prof. Elisardo Antelo and Prof. Javier D. Bruguera
September 2013

This thesis contributes to the area of hardware support for parallel programming by introducing new hardware elements in multicore processors, with the aim of improving performance, and optimizing new tools, abstractions and applications related to parallel programming, including transactional memory and data race detectors. Specifically, we develop a new hardware filter for reducing the use of signatures in a transactional memory system. We also develop a hardware asymmetric data race detector and tolerator. Finally, we propose a new module of hardware signatures that solves some of the problems that we found in the previous tools related to the lack of flexibility and adaptability.

HYBRID CACHES: DESIGN AND DATA MANAGEMENT

Alejandro Valero
Universitat Politècnica de València, Spain
Advisor: Prof. Julio Sahuquillo and Prof. Salvador Petit
September 2013

This thesis proposes hybrid eDRAM/SRAM cache architectures to leverage the advantages offered by each technology. Architectural mechanisms are devised to hold the MRU data in fast SRAM technology, while the remaining data are stored in energy-efficient eDRAM technology. Compared with typical SRAM designs, hybrid caches significantly reduce energy and area with a minimal impact on performance. The core research developed in this dissertation was recognized with the Intel Doctoral Student Award, received by Alejandro in 2012. His main research interests include energy-aware memory hierarchy design, reliability, and high-performance cache replacement algorithms. Alejandro is currently with Prof. David Kaeli at Northeastern University in Boston, where he is working on the Multi2Sim simulator and GPUs.

HARDWARE ACCELERATION FOR JUST-IN-TIME COMPILATION IN EMBEDDED SYSTEMS

Alexandre Carbon
CEA LIST, France
Advisors: Dr. Henri-Pierre Charles and Dr. Yves Lhuillier
October 2013

In this thesis, we explore hardware acceleration opportunities for Just-In-Time Compilation in embedded systems to reduce its performance impact on small and sparse resources. We highlight common optimization opportunities between the different technologies using dynamic compilation (virtual machines, binary translation, dynamically-typed languages, multistage, etc.), especially concerning associative array management and dynamic memory allocation. We consequently propose a new ARM core functional unit for JIT acceleration. Our experiments highlight significant performance gains compared to standard libraries and LLVM existing software optimizations.

HIGH PERFORMANCE JAVA FOR MULTI-CORE SYSTEMS

Sabela Ramos, Universidade da Coruña, Spain
Advisors: Prof. Guillermo López Taboada, Prof. Juan Touriño Domínguez
October 2013

This thesis is focused on enabling Java for High Performance Computing (HPC) in the many-core era. The interest in Java within the HPC community has been rising during the last years thanks to its noticeable performance improvements and its productivity features. In addition to an analysis of the potential of the use of hardware accelerators in Java, this Thesis presents a Java message-passing solution with optimized blocking and non-blocking collectives. This support intends to provide shared memory programming with the scalability of distributed memory paradigms, abstracting multithreading details from the user and outperforming MPI libraries like Open MPI or MPICH2.

SIMULATION OF LARGE CHIP MULTIPROCESSORS RUNNING MULTITHREADED APPLICATIONS

Alejandro Rico
Barcelona Supercomputing Center, Spain
Advisors: Assoc. Prof. Alex Ramirez, Prof. Mateo Valero
October 2013

In this thesis we introduce simulation techniques for chip multiprocessors with hundreds of cores running multithreaded applications. The simulation models for these architectures employ multiple levels of abstraction with different accuracy/speed trade-offs. These simulation models lie on top of a trace-driven engine for multithreaded applications that combines a trace-driven simulation front-end with a runtime system for the dynamic scheduling and synchronization of parallel tasks to simulated threads. We evaluate and validate our abstract simulation models against real machines and realize that they are accurate for different purposes and they are faster and more insightful than functional simulation in execution-driven simulators. We also explore the use of trace-driven memory simulation using filtered traces, and extend the state-of-the-art techniques to improve their accuracy for simulating multithreaded applications.

DYNAMICALLY RECONFIGURABLE ARCHITECTURES FOR VIDEO CODING AND HYPERSPECTRAL IMAGING SYSTEMS

Teresa G. Cervero García
Universidad de Las Palmas de Gran Canaria, Spain
Advisors: Prof. Dr. Roberto Sarmiento and Assoc. Prof. Dr. Sebastián López
November 2013

FPGA-based embedded systems are gaining relevance for implementing a wide range of applications. Part of this success is due to their balanced compromise between performance and flexibility, but also because of their capability for exploiting the dynamic reconfigurability. Going further to the traditional dynamic reconfigurability strategies, it is possible to take advantage of this feature for adjusting the performance of a system by scaling the number of processing elements involved in the computation, according to the demands brought up at run-time. In this sense, this Thesis is focused on developing scalable designs, in which the hardware scalability might be adjusted and managed at run time. As a whole, these contributions (the scalable designs and the reconfigurability mechanism) are focused on filling the gap in the development of FPGA-based context-aware embedded systems.

UPCOMING EVENTS

20th International Symposium on High-Performance Computer Architecture (HPCA-20)
February 15-19, 2014, Orlando, Florida, USA http://www.hpcaconf.org

International Symposium on Code Generation and Optimization (CGO 2014)
February 15-19, 2014, Orlando, Florida, USA http://www.cgo.org/cgo2014/

The 27th International Conference on Architecture of Computing Systems (ARCS 2014)
February 25-28, 2014, Luebeck, Germany http://www.arcs2014.iti.uni-luebeck.de/

16th Design, Automation and Test in Europe Conference (DATE 2014)
March 24-28, 2014, Dresden, Germany http://www.date-conference.com

The International Conference on Compiler Construction (CC 2014)
April 5-13, 2014, Grenoble, France http://www.etaps.org

4th European LLVM Conference
April 7-8, 2014, Edinburgh, Scotland http://llvm.org/devmtg/2014-04

The European Conference on Computer Systems (EuroSys 2014)
April 13-16, 2014, Amsterdam, The Netherlands http://eurosys2014.vu.nl/

19th International Conference on Reliable Software Technologies (Ada-Europe 2014)
June 23-27, 2014, Paris, France http://www.ada-europe2014.org/

The 11th International Conference on Computing Frontiers (CF 2014)
May 20 - 22, 2014, Cagliari, Italy http://www.computingfrontiers.org/

The 28th International Conference on Supercomputing (ICS 2014)
June 10-13, 2014, Munich, Germany http://ics-conference.org/

The 9th IEEE Symposium on Industrial Embedded Systems (SIES 2014)
June 18-20, 2014, Pisa, Italy http://retis.sssup.it/sies2014/

International Conference on Embedded Computer Systems: Architectures, Modeling, and Simulation (SAMOS XIV)
July 14-17, 2014, Samos, Greece http://www.samos-conference.com/

contributions If you are a HiPEAC member and would like to contribute to future HiPEAC newsletters, please visit http://www.hipeac.net/hipeacinfo

HiPEAC info 37
COMPILATION ARCHITECTURE

HiPEAC info is a quarterly newsletter published by the HiPEAC network of excellence, funded by the 7th European Framework Programme (FP7) under contract no. FP7/ICT 287759
website: http://www.hipeac.net.
subscriptions: http://www.hipeac.net/newsletter

10TH INTERNATIONAL SUMMER SCHOOL ON ADVANCED COMPUTER ARCHITECTURE AND COMPILATION FOR HIGH-PERFORMANCE AND EMBEDDED SYSTEMS (ACACES 2014), JULY 13-19, 2014, FIUGGI, ITALY