

HiPEAC

info 55

OCTOBER 2018

Computing
Systems Week
Heraklion

'It's the memory, stupid': A conversation with Onur Mutlu

How HiPEAC is reinventing data centres, storage and networking

Database benchmarking breakthrough



Welcome to Heraklion



A conversation with Onur Mutlu



Data centres, storage and networking

<p>3 Welcome <i>Koen De Bosschere</i></p> <p>4 Policy corner Submitting a proposal for ICT-01-2019? What you should know <i>Sandro D'Elia</i></p> <p>6 News</p> <p>13 HiPEAC voices 'It's the memory, stupid': A conversation with Onur Mutlu</p> <p>16 Data centres, storage and networking special Delivering the data <i>Andrea Reale, Georgios Goumas, Manolis Katevenis, Angelos Bilas, Toni Cortés, Ramon Nou, Alberto Miranda, Dimitrios Soudris and Christoforos Kachris</i></p> <p>23 HiPEAC heads east Embedded intelligence at the heart of the Balkans <i>Radovan Stojanović and Rainer Leupers</i></p> <p>25 Innovation Europe Scaling up for HPC with AllScale's task-based parallelism <i>Charles Gillian</i></p> <p>26 Innovation Europe Putting transprecision computing into practice with OPRECOMP <i>Georgios Karakonstantis and Cristiano Malossi</i></p>	<p>27 Innovation Europe EXA2PRO to enhance productivity and ensure consistent performance <i>Lazaros Papadopoulos</i></p> <p>28 Innovation Europe AXIOM delivers scalable cyber-physical architectures for smart electronics <i>Laura Gales</i></p> <p>29 Technology transfer Creating Kaleao: A startup story <i>John Goodacre</i></p> <p>30 SME snapshot Turbocharging cloud applications with InAccel <i>Christoforos Kachris</i></p> <p>31 Technology opinion Discriminative performance evaluation: separate the bold and the beautiful <i>Martin Kersten and Ying Zhang</i></p> <p>34 HiPEAC futures HiPEAC Jobs: Four myths on choosing your career path – busted Career talk: Cristina Silvano, Politecnico di Milano Arm School Program: working to close the STEM skills gap Career talk: Leonidas Kosmidis, Barcelona Supercomputing Center Three-minute thesis: Tasks + dataflow = max performance on many cores</p>
--	--



Embedded intelligence at the heart of the Balkans



Better database benchmarks



Career talk: Cristina Silvano



HiPEAC is the European network on high performance and embedded architecture and compilation.



hipecac.net



@hipecac



hipecac.net/linkedin



HiPEAC has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 779656.

Cover photo: Sashkinw on Dreamstime.com

Back cover photo: Cristina Gottardi on Unsplash

Design: www.magelaan.be

Editor: Madeleine Gray

Email: communication@hipecac.net

Every other year, HiPEAC works on its Vision document, and 2018 is such a year. The summer break was a good opportunity for me to read about global societal challenges and how computing can help tackle them. I learned some good and some bad news. One piece of good news is that the global fertility rate has dropped almost everywhere on earth. Two children per family will soon be the average for most countries. Thanks to improved living conditions, billions of young people will have a chance to grow old. Population growth in the 20th century was caused by high birth rates; in the 21st century, it will be caused by increasing life expectancy.

Another piece of good news is that 88% of all people worldwide now live in 'middle income' countries, with access to basic healthcare, education, fresh water, and so on, and that in a couple of decades the 12% of people still living in extreme poverty today are expected to join this group. The bad news is that the global population today already consumes 70% more resources annually than the planet produces in that same year. Our pollution rate is faster than the earth's natural clean-up rate. Plastics in the ocean and carbon emissions are very visible examples of this problem.

Experts call it overshoot. This year's overshoot day was 1 August. If we want to save the planet for future generations, we must push the overshoot day back to 1 January of the next year, as it was in 1970 – that is, 50 years ago. Bringing global consumption back to the 1970 level while another three billion people join the world's population, all of whom are working hard to build a middle-class existence, is the mother of all societal challenges. It will only be possible by drastically reducing the resource consumption of modern lifestyles; for this transition, we will need all the knowledge and wisdom of the world. It is difficult to predict the future, but it is not difficult to predict that computers and artificial intelligence will be crucial in creating a sustainable living environment for 10 billion people by 2075. Since there is no time to lose, we'd better start working on it immediately.

You'll be able to read about the Vision in the next issue of the magazine. In the meantime, I hope you enjoy this special issue on data centres, storage and networking. In it, you'll find out how the HiPEAC community is playing its part to solve the sustainability challenge by reducing energy consumption and improving efficiency in computing systems.

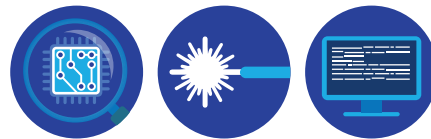
Koen De Bosschere, HiPEAC coordinator



A new European Commission call titled 'Computing technologies and engineering methods for cyber-physical systems of systems' opened on 16 October 2018, with a deadline of 28 March 2019. It addresses two areas that are at the core of the HiPEAC community's knowledge base. In this article, Sandro D'Elia, Programme Officer in the Digitising Industry Unit within the European Commission's DG CONNECT, gives some invaluable advice to anyone thinking of submitting a proposal.

Submitting a proposal for ICT-0

The European Commission is investing heavily in what we call the 'key digital technologies' relevant for industry in the 21st century: microelectronics, photonics and software. The ICT-01-2019 call within the Horizon 2020 programme is an important element of this strategy.



There are several significant differences between this call and previous ones on similar subjects. The first one is that the two different research areas of **computing and systems-of-systems have been merged**. The reason is simple: in both areas, the focus is on intelligence and computing capacity for cyber-physical systems (CPS) at the edge of the networks, and the target applications are industrial and professional markets, e.g. transport, manufacturing, energy, security, health-care, etc. It simply makes sense to view them together.

The second one is the new **focus on long-term, innovative research**. The target Technology Readiness Level (TRL) is 2 to 5. OK, so we all know that the TRL is not the best metric when dealing with software-intensive systems; nevertheless, it is clear that the objective is to identify new research directions rather than performing incremental development.

For computing, the call text mentions explicitly '**radically new solutions** overcoming the intrinsic limitations of today's computing **system architectures and software design** practices'. Does this mean neural networks writing code? Maybe. A compiler for neuromorphic architectures? Maybe. Evolutionary algorithms for reliable GPS? Maybe. A software stack for quantum computers? Maybe. Simulating a new computing architecture in field-programmable gate array (FPGA)? Maybe. Most probably, we will get something else, or a combination of different techniques that will help to build the foundations of software and system engineering for the 21st century.

*"Neural networks writing code?
A compiler for neuromorphic architectures? Evolutionary algorithms for reliable CPS? A software stack for quantum computers? Simulating a new computing architecture in FPGA?"*



1-2019? What you should know

Be aware that even though we are aiming for long-term objectives, this does not mean that you should limit your proposal to theoretical research alone. The call text clearly explains that projects will have to **'deliver a working prototype'** with at least two use cases relevant for **'industrial and professional domains'**. Don't forget that the overall objective of any EU-funded research in this area is industrial competitiveness.

For the systems-of-systems, the call targets **models, methods and tools** for the complete life-cycle of complex systems. We want to go beyond current model-based tools, targeting a new generation of **fully fledged solutions** that can support design, setup, operations, maintenance, reconfiguration, including by exploiting **'emerging technologies** such as augmented reality and artificial intelligence' and focusing on **'autonomic solutions** capable of guaranteeing the overall reliability and security' even when things go wrong.

You can see this as an extension of the concept of 'digital twin' beyond the sectors where it is already widely used (mainly optimization of operations and maintenance). The challenge is to develop methods and technologies that not only look great in a demo, but are usable and economically sustainable in real-world scenarios.

At both levels (computing and systems-of-systems), the targets are the critical cyber-physical systems of tomorrow, which will show cognitive behaviour and a high degree of autonomy compared to the systems we see today. Does this mean that this is an artificial intelligence (AI) call? Well, not really, because improvements in AI methods are not a direct objective, but we would expect AI techniques to be considered in many proposals.

To summarize, you could consider this call a taste of what will be coming in the future 'Horizon Europe' and 'Digital Europe' programmes, which, starting in 2021, will support European science, innovation and

industrial competitiveness in the era of autonomous systems. I am looking forward to seeing excellent proposals submitted by the HiPEAC community.

bit.ly/ICT-01-2019

"The overall objective of any EU-funded research in this area is industrial competitiveness"



Καλώς ήλθατε στο Ηράκλειο!

From the Minoans to Silicon Island: Crete is a truly majestic setting for our data-themed Computing Systems Week. From the Institute of Computer Science at the Foundation for Research and Technology – Hellas (FORTH), Manolis Katevenis and Angelos Bilas give us a whistle-stop tour.



Photo credit: Stavros Markopoulos

Kalós ílthate sto Irákleio! – Welcome to Heraklion!



Tell us a bit about FORTH.

Established 35 years ago, FORTH is a state-owned research foundation. It comprises six institutes, four of which are in Heraklion, where the main campus is located. Large and diverse, FORTH has research teams with interests spanning from basic and applied science, technology development, and industrial applications, to social and cultural studies. The Heraklion campus is part of a larger ecosystem that includes the University of Crete, Crete University Press, the Skinakas Observatory, the STEP-C startup incubator, the PRAXI network and more. It's a great environment for doing research.

The main focus of the Institute of Computer Science is clearly information and communi-

cation technology (ICT), but being part of this larger ecosystem creates interesting synergies all the time. FORTH has trained a few generations of engineers and scientists successfully in this environment; these people support industry and research in Greece, as well as in many other places in Europe and around the world.

What are Crete's main contributions to the high-tech sector?

Since the early years, the Institute of Computer Science has had close relationships with industry. FORTH pioneered the internet in Greece, founded one of the major telecom providers (FORTHNET), and has contributed technology used in several other products and services. Over the last few years we have observed a significant increase in industrial activity in Crete, in both service industries and technology-intensive products. Today, there are startups in areas such as datacentre management, enterprise storage systems, server design, augmented virtual reality, analytics for mobile,

travel services, business analytics, energy storage, and so on. Although best known as a holiday destination, Crete aspires to create a critical mass of industry around technology-intensive products and services. We hope that the current trends will continue.

Tell us three things about Crete that we might not know.

First, Heraklion has been transformed over the last 10 to 15 years. Traditionally, it had a reputation as a 'city with many problems'. Today, there is a vibrant, regenerated city centre and seafront area, improved city road network, and good sea

Mini phrase book

English	Greek
Where can I store my exabyte?	<i>Pou na apothikefso to exabyte mou?</i>
That's a lot of gigabits per second!	<i>Eine pola gigabits ana defterolepto!</i>
Anyone for a raki?	<i>Pame gia mia raki?</i>

and air transport connections. Thanks to these improvements, according to Euromonitor statistics, Heraklion was the 66th most popular destination in the world for travellers in 2016, with the highest rise in number of arrivals among all European cities.

Second, the climate in Heraklion is characterized by mild winters and pleasant summers, with the average maximum temperature during the summer hovering around 28-30°C. As an example, the average maximum temperature in July is 28.6°C, as opposed to a sweltering 34.5°C in Athens and 32.9°C in Thessaloniki. It is also noteworthy that the daily minimum temperature during the year deviates on average less than 10°C, whereas in winter, in particular, the deviation usually does not exceed 5°C.

At the same time, although the microclimates in Heraklion and many other places in Crete



Photo credit: Y Skolas

have small internal variations, there are big differences between microclimates across the island. The southeast part of the island is the sunniest, driest place in Europe; several mountainous areas have a climate which is categorized as 'alpine'; the west of the island is temperate with conifer trees; while the south has a 'North African' climate, and so on. Crete is also one of the most biodiverse places in Europe, with 160 plants which are unique to the island and around 1,700 other species. This means there are around 100 times more species of vegetation per square kilometre than the average on the European continent.

Lastly, a mini fact: Crete has been constantly inhabited for the last 8,000 years.

What are the must-dos for visitors to the city?

1. Walk around the city centre and follow the marked trails to various historical buildings. Visit the archaeological museum, where you can see the disc of Phaistos – the logo of FORTH.
2. Try some of the local wines and beers. There has been a lot of activity in the local industry over the last few years, with some great results.
3. If you have kids, check out the natural history museum in the city and the aquarium, which is about 30km from the city.

For a glimpse at some of the highlights of Crete, have a look at this promotional video by the Prefecture of Crete: bit.ly/Crete_video

It's all Greek to me

From microchips to megabytes, the influence of Greek is clear throughout the English language, particularly in scientific fields. Here is a quick round-up of some of HiPEAC's favourite Hellenicisms:

technology – from τέχνη ē (tékhnhē, 'art') and λόγος (logos, 'word/reason/account or explanation')

parallel – from παρά (para) ἄλληλος (allellos, along each other)

physics – from φύσις phýsis, nature

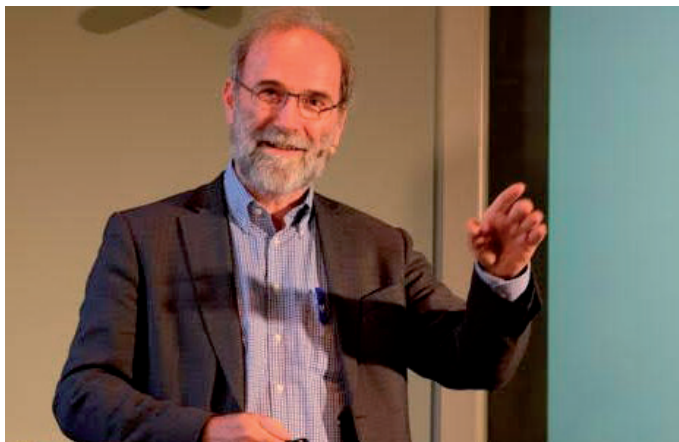
prototype – from πρῶτος (prōtos, 'first; earliest') + τύπος (týpos, 'blow, pressing; sort, type')

The trade in terminology goes both ways though: Angelos notes that there has been quite a bit of effort to translate terms such as 'bit' into Greek. 'Today, the English terms are used everywhere in Greece, but you can find some interesting translations and the debate on what is the proper translation can turn into a heated discussion!'



Photo credit: Y Skolas

HiPEAC Coordinator Koen De Bosschere wins ACM SIGARCH Alan D. Berenbaum Distinguished Service Award



In June, HiPEAC Coordinator Koen De Bosschere's exceptional dedication to HiPEAC was recognized with the ACM SIGARCH Alan D. Berenbaum Award, presented during ISCA 2018, the 45th International Symposium on Computer Architecture in Los Angeles. The first European to win the award, Professor De Bosschere was nominated for structuring and connecting the European research community in computer architecture and compilation.

In response to the news of this award, Professor De Bosschere commented: 'Leading the HiPEAC community, contributing to the HiPEAC Vision and observing its impact has always been a very rewarding experience for me. I feel very honoured to receive the Alan D. Berenbaum Distinguished Service Award and I would like to dedicate it to all our members and the staff, without whom there would be no HiPEAC.'

"Leading the HiPEAC community, contributing to the HiPEAC Vision and observing its impact has always been a very rewarding experience"

'I would also like to thank the European Commission for the uninterrupted support of our community during more than a decade,' he added.

Professor De Bosschere was nominated by his Ghent University colleague Professor Lieven Eeckhout, who comments: 'Koen really brought the systems community together in Europe, creating critical mass and putting systems research on the map in Europe. This has had massive impact at the political level with significantly increased funding for systems research in Europe. Koen's very generous and effective service work has had tremendous impact on many students and researchers in Europe and beyond.'

Professor De Bosschere strongly believes in the power of networking and collaboration and in the importance of investing in young talent. Under his leadership, HiPEAC has become the biggest network of its kind in Europe, engaging around 2,000 computing experts and organizing major annual events like the HiPEAC conference and the ACACES summer school. Its biennial roadmap, the HiPEAC Vision, has also become a strategic document influencing European research agendas in computing systems.



Koen De Bosschere speaking at MECO 2018 (see p.23)

Silexica raises \$18m in Series B funding

SILEXICA 

This summer saw HiPEAC member company Silexica complete a \$18 million Series B round of financing, led by EQT Ventures. The investment will be used to advance the company's software development solutions for autonomous driving.

In the face of the increased complexity of computing systems, since 2014 Silexica has been developing its SLX programming tools to analyse how software runs on heterogeneous multicore processors and provide deep system-level understanding for developers. Now with offices in Germany, the United States and Japan, Silexica's clients include Tier 1 automotive suppliers as well as leading original equipment manufacturers (OEMs) in 5G, aerospace and defence.

Maximilian Odendahl, chief executive officer at Silexica, said: 'This funding round will enable us to strengthen existing and upcoming SLX tools and solutions. In addition, it will see us develop an industry-first, vendor-neutral, multi-application Simulation Platform to help OEMs and all members of the supply chain optimize and integrate software provided for complex systems such as autonomous vehicles.'

For our interview with Maximum Odendahl on his experience of setting up Silexica, see HiPEACinfo 54 p.37



Photo credit: Eneko Illarramendi

Technology meets business at ACACES18



For this year's ACACES, the International Summer School on Advanced Computer Architecture and Compilation for High-Performance and Embedded Systems, HiPEAC teamed up with the European Union-funded Innovation Action TETRAMAX and the Coordination and Support Action Eurolab4HPC. Taking place once again in picturesque Fiuggi, Italy, the summer school, which year after year is attended by returning attendees, attracted a record 228 participants, selected from a total of 315 applicants. Many of these were able to attend thanks to the grants offered every year covering fees and accommodation.

Thanks to the participation of TETRAMAX and Eurolab4HPC, the summer school offered courses on high-performance computing and technology transfer, in addition to topics ranging from deep neural networks to the internet of things. Keeping up ACACES' tradition of excellence, the courses were delivered by experts in their fields from around the world, and were highly rated by participants.

The summer school was kicked off by Partha Ranganathan, Google, whose keynote speech titled 'End of Moore's Law: Or, a computer architect's mid-life crisis?' took a humorous look at the disruptions in the field of computer architecture. Covering the business aspect, Intel's Monia Chiavacci's keynote talk the following night set out the extraordinary entrepreneurial story of Italian company Yogitech, which was bought by Intel in 2016.

To help students prepare for the poster session, HiPEAC organized a 'pitch training session', where participants were challenged to present their work to audiences ranging from the oldest to the youngest, from potential investors or European Commission officials to fellow scientists. Meanwhile, in the 'Inspiring

Futures!' session, students had a chance to find out about different career paths available, with advice from startup founders and university professors. You can find out more about HiPEAC's career mentoring activities on p. 33.

We would like to thank everyone who contributed to making this year's summer school such a success. Details about next year's event will be available soon; check the HiPEAC website for further information.

Turn to p. 13 for our interview with ACACES18 tutor Onur Mutlu

Videos from the summer school are available on the HiPEAC YouTube channel:

bit.ly/ACACES_videos

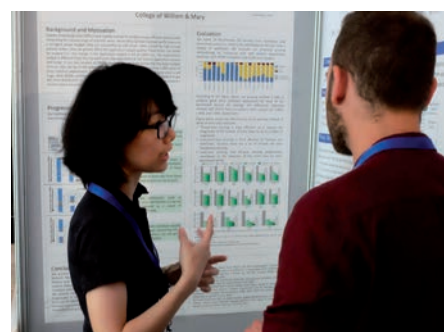


Photo credit: Eneko Illarramendi

PUMPS gets smarter thanks to AI component



The ever-popular Programming and tUning Massively Parallel Systems (PUMPS) summer school returned this year with a full programme on graphics processing units (GPUs). Dubbed PUMPS+AI, this year's edition had a specific focus on artificial intelligence (AI) and the key role GPUs are playing in this area.

Co-organized by Barcelona Supercomputing Center (BSC), the University of Illinois at Urbana-Champaign, the Universitat Politècnica de Catalunya-Barcelona Tech (UPC) and HiPEAC, the summer school took place on 16-20 July. Once again, it was extremely popular: of 108 applicants, 80 were accepted from 19 countries.

'The impact of AI is enormous, and the technologies under consideration at PUMPS will be crucial to its continued development,' commented Professor Mateo Valero (BSC-UPC), one of the PUMPS directors. 'We are delighted to have hosted the ninth edition of this hugely successful summer school.'

As always, PUMPS featured lectures from world-renowned experts such as Professor Wen-mei Hwu (University of Illinois at Urbana-Champaign) and David Kirk (NVIDIA). The programme this year covered topics including deep learning, high-level programming models, multi-GPU execution and CUDA algorithmic optimization strategies. HiPEAC co-founder and first coordinator Mateo Valero and Wen-mei Hwu are the summer school's directors, while local organization is provided by HiPEAC member Antonio Peña (BSC).

This year, HiPEAC hosted two workshops, both of which were well received. As at ACACES18, Xavier Salazar (BSC / HiPEAC) organized a session on pitching posters, to prepare students for the poster session, along with an 'Inspiring Futures!' careers session which explored different research career paths and funding options.

Next year's event will be held on 24-29 June 2019 at UPC, Barcelona; further information will be available on the PUMPS website in due course.

Further information: pumps.bsc.es

Workshop explores high-speed interfaces for satellites

On 24 April, leading organizations in the aerospace sector gathered in Pisa for a workshop looking at the future of high-speed data interfaces for on-board data systems in satellites. Organized by University of Pisa spinoff IngeniArs, the aim of the event was to highlight the role of SpaceFibre – the next-generation interconnect for on-board spacecraft communications – as the successor to the existing standard, SpaceWire.

Representing a significant step forward for the strategic aerospace domain, SpaceFibre is expected to succeed SpaceWire as a European technology which becomes a global standard. IngeniArs' commitment to SpaceFibre is demonstrated by the Horizon 2020 SME Instrument Phase 2 project SIMPLE, or Spacefibre IMPLementation design and test Equipment. With the aim of designing and implementing a system which is compliant with the new standard, along with test equipment capable of validating such a system, the project will deliver three products by the end of January 2019. These are: the SpaceFibre intellectual property (IP) core, the SpaceFibre Analyser PXI and SpaceART, the SpaceFibre Analyser Real Time.



Some of the major players in the aerospace sector, including Thales Alenia Space-Italy, National Instruments, Leonardo and Kayeser, as well as the European Space Agency and the Italian Space Agency, attended the workshop. Franca Malerba, the first Italian astronaut in space and mentor to IngeniArs in the SIMPLE project, chaired the event. Participants were enthusiastic about the IngeniArs products developed within SIMPLE, opening up the possibility of future commercial development.

For more information on IngeniArs, read our 'SME snapshot' profile in HiPEACinfo 53

bit.ly/HiPEACInfo53_hi-res

New CK framework with over 500 reusable AI components released

Grigori Fursin, cTuning Foundation



The open-source Collective Knowledge framework (CK) allows you to convert your code and data into unified CK components with a common Python API and JSON meta description and share them with others, allowing the community to adapt them to new research scenarios.

The latest release includes more than 540 CK packages shared by our community to automatically install various artificial intelligence (AI) / machine learning (ML) frameworks and libraries, models and datasets on Linux, MacOS, Windows and Android.

You can also reuse over 340 customizable CK programs from traditional systems benchmarks to emerging AI applications. This includes all workflows from the first ACM ReQuEST tournament to collaboratively benchmark and co-design the efficient software / hardware stack for deep learning inference from the cloud to the edge.

For all this and more, visit the revamped cKnowledge.org website, or access the Collective Knowledge framework on GitHub.

cKnowledge.org/request

github.com/ctuning/ck

FPGA-powered supercomputer

Dr Tassadaq Hussain, Assistant Professor, Riphah International University

Researchers from Barcelona Supercomputing Center, Riphah International University and UCERD Islamabad have joined forces to design a supercomputing system powered by field-programmable gate arrays (FPGAs).

Combining hardware accelerators, or co-processors, with multicore processors has become a popular way of achieving greater computing performance within a restricted power budget. These accelerators improve the performance of compute-intensive applications by executing a specific task.

FPGA architectures have established themselves as an attractive choice for high-performance computing systems architects, thanks to the considerable advancements in performance and power efficiency that they offer. By achieving higher performance per watt, FPGAs have proved themselves capable of competing with superscalar and graphics processing unit (GPU) accelerators, especially for high-performance computing applications.

This new FPGA-powered supercomputer system executes compute-intensive applications on FPGA. The design system uses message passing interface libraries for communication between compute nodes, while the communication between the Arm processors and FPGAs uses AXI4-Stream interfaces to execute a compute-intensive portion of an application. The Finite Impulse Response (FIR) filter has been implemented as a test application on the system to evaluate the computational capability of the Arm processors with and without hardware acceleration.

The results of this research were presented at the International Conference on Computing and Mathematical Sciences 2018 and have received patent number 311/2018 from the Government of Pakistan's Intellectual Property Organization.



PRiME Framework for runtime management

Ghaithaa Manla, University of Southampton



In June this year, PRiME, a multimillion euro UK research programme funded by the Engineering and Physical Sciences Research Council, released the PRiME Framework. Bringing university and industry partners together, PRiME is investigating future high-performance embedded systems utilizing many-core processors.

The PRiME Framework is for application- and platform-agnostic runtime management, enabling the portability of runtime management approaches. It allows users developing runtime managers for new systems to directly compare different approaches, as well as simplifying the development effort, enabling users to quickly prototype and test an idea in a few lines of code.

'The PRiME Framework is a step change in cross-layer system optimization of future many-core and heterogeneous systems. It benefits runtime management researchers and developers by reducing development overhead and enabling the direct comparison of approaches,' said Professor Bashir AL-Hashimi, University of Southampton, who leads the project.

prime-project.org

Book: Harnessing Performance Variability in Embedded and High-performance Many / Multi-core Platforms

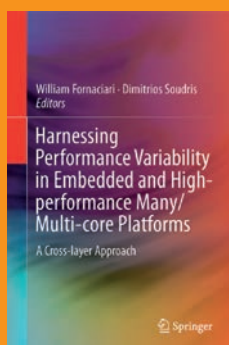
Editors: William Fornaciari, Dimitrios Soudris

In 2017, the HARPA project, led by HiPEAC member William Fornaciari (Politecnico di Milano – DEIB), was successfully completed. In addition to several patents being filed, the project resulted in the publication of this book, which describes the state of the art in research into the architectural design of heterogeneous multicore / manycore processors.

The authors present methods and tools to enable readers to develop heterogeneous multicore / manycore architectures that provide dependable performance – that is, correct functionality and timing guarantees for the expected lifetime of a platform under thermal, power and energy constraints. Various aspects of the reliability problem are discussed, at both circuit and architecture level, along with the intelligent selection of knobs and monitors in multicore platforms, and systematic design methodologies. The book also demonstrates how new techniques have been applied in case studies from different application domains.

Further information:

- bit.ly/HARPA_book
- harpa-project.eu



Book: Hardware Accelerators in Data Centers

Editors: Christoforos Kachris, Babak Falsafi, Dimitrios Soudris

Emerging cloud applications like machine learning, artificial intelligence, deep neural networks, and big data analytics have created the need for data centres that can process huge amounts of data without consuming excessive amounts of power. To meet these requirements, datacentre operators are adopting novel architectures incorporating hardware accelerators, leading to the era of heterogeneous distributed computing systems. However, using heterogeneous architectures brings its own challenges in the form of efficient resource utilization, efficient scheduling, and resource management.

Edited by three HiPEAC members, this book brings together the most promising recent research in this emerging domain. It sets out 13 system-level architectures showing how to efficiently use hardware accelerator in data centres, tackling challenges including energy efficiency, resource utilization and management, and programmability. The aim is for the book to serve as a reference text for researchers and engineers working in this exciting new field.

Further information:

- bit.ly/HW_accel_dc

Dates for your diary

PLDI 2019: Conference on Programming Language Design and Implementation

22-28 June 2019, Phoenix, Arizona, USA

Paper submissions: 16 November 2018

HiPEAC Paper Award conference

pldi19.sigplan.org

EFES 2018: European Forum for Electronic Components and Systems

20-22 November 2018, Lisbon, Portugal

efecs.eu

DAC19: Design Automation Conference

2-6 June 2019, Las Vegas, Nevada

Abstract submissions: 20 November 2018

HiPEAC Paper Award conference

dac.com/call-for-contributions

ISCA 2019: International Symposium on Computer Architecture

22-26 June 2019, Phoenix, Arizona, USA

Abstract submissions: 3 December 2018

HiPEAC Paper Award conference

iscaconf.org/isca2019

FCCM 2019: 27th IEEE International Symposium on Field-Programmable Custom Computing Machines

28-30 April 2019, San Diego, California, USA

Abstract submissions: 7 January 2019

HiPEAC Paper Award conference

fccm.org/2019

ICT 2018: Imagine Digital - Connect Europe

4-6 December 2018, Vienna, Austria

bit.ly/ICT18_Vienna

HiPEAC19: 14th International Conference on High Performance and Embedded Architecture and Compilers

21-23 January 2019, Valencia, Spain

hipeac.net/2019/valencia

DATE19: Design, Automation and Test in Europe

19-23 March 2019, Florence, Italy

date-conference.com

'It's the memory, stupid': A conversation with Onur Mutlu



'We're beyond computation; we know how to do computation really well, we can optimize it, we can build all sorts of accelerators ... but the memory – how to feed the data, how to get the data into the accelerators – is a huge problem.'

This was how ETH Zürich and Carnegie Mellon Professor Onur Mutlu opened his course on memory systems and memory-centric computing systems at HiPEAC's summer school, ACACES18. A prolific publisher – he recently bagged the top spot on the International Symposium on Computer Architecture (ISCA) hall of fame – Onur is passionate about computation and communication that are efficient and secure by design. For our special issue on data centres, storage, and networking, HiPEAC picked his brains on all things data-based.

How important is storing and moving data to the correct functioning of a computer system?

Critically important: all systems need to have reliable and efficient mechanisms to store and move data.

Fundamentally, a computing system consists of three components:

- computation units that operate on data
- data storage units – that is, memories – that keep the data intact
- data communication units that enable the movement of data to where it is needed

Most of the real estate of the processors we build is dedicated to data storage and communication units. Computation units actually occupy a very small area, thanks to Moore's Law and over 70 years dedicated to optimizing them. This arose from a design mindset in which all data has to be processed in the computation units.

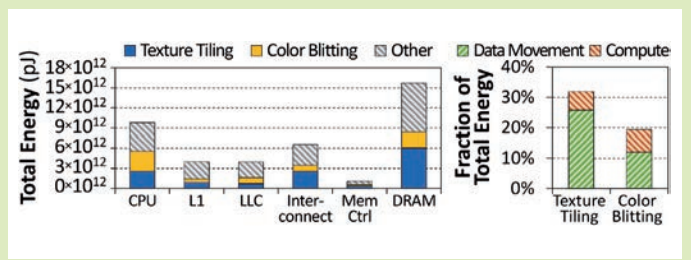
"We're beyond computation, we know how to do computation really well, but the memory is a huge problem"

This mindset has led to significant performance improvements in computation units. However, the same mindset came at the cost of data and movement units, which have essentially remained 'slaves' to the computation units. There are a few major downsides to the computation-unit-centric design approach.

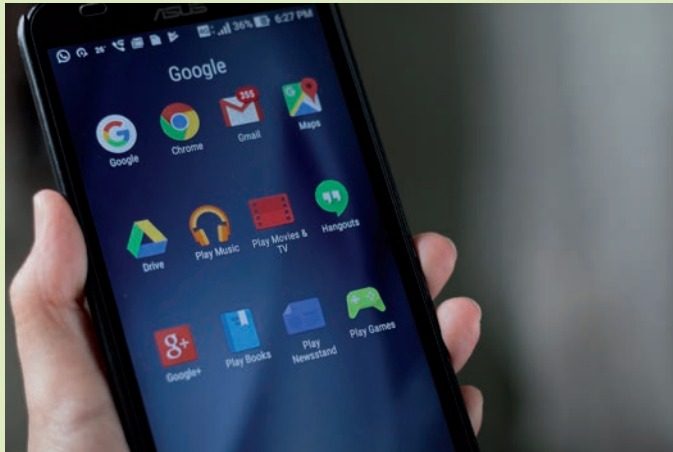
First, in many cases, more than 80% of the chip area is dedicated to caches, memories, memory controllers, interconnects and so on, whose sole purpose is to buffer data or control the buffering of data. Add main memory and storage, and we soon see 95-99% of the entire real estate of a system being dedicated to units that simply store and move data, or which control the storage and movement of data. If we were to design systems where, say, storage units could do computations on the data they store, we would likely have a much more balanced system.

Second, to get data quickly from storage units to the processor, we've created workarounds which are making systems ever more complex. These include sophisticated out-of-order and speculative execution engines, many levels of cache hierarchy, complex prefetching mechanisms and large amounts of multithreading. As well as making design harder, these workarounds have an adverse impact on reliability and energy efficiency.

Third, since processors are the only place where data can be processed, data moves around the system a lot, even if all we want to do is copy data from one place to another inside memory. Our results on four major mobile system workloads developed by Google show that, in a mobile system, more than 62% of the entire measured system energy is spent on moving data between memory and the computation units. We've also shown that, by processing data close to memory, using either very simple cores



Energy breakdown when scrolling through a Google Docs web page



or specialized accelerators in the logic layer of 3D-stacked memory, we can halve the energy used while at the same time doubling system performance on a mobile device.

What kind of applications could be enabled or improved thanks to enhanced data storage and movement?

The amount of data we store and manipulate is increasing at an unprecedented pace. Take bioinformatics, for example. Today we can sequence genomes much faster than our computational capacity allows us to analyse them. A major bottleneck is data movement between the memory and processor / field-programmable gate array (FPGA).

One of the goals of our bioinformatics research is to enable an embedded device that can perform comprehensive genome analysis in real time. To build such a device, you want to maximize energy efficiency, minimize latency and maximize security. All of these require minimizing data movement. As data movement is naturally energy hungry – a single memory access costs two to three orders of magnitude more energy than a complex arithmetic operation! – memory access also leads to performance loss, and increases security vulnerabilities by exposing the data to the outside world for longer.

Two directions we're pursuing to minimize data movement are:

New algorithms for genome analysis: we've shown that by designing algorithms (e.g., FastHASH) that take advantage of the

structure of the genome being analysed, one can reduce communication (and computation) by 10 to 100 times.

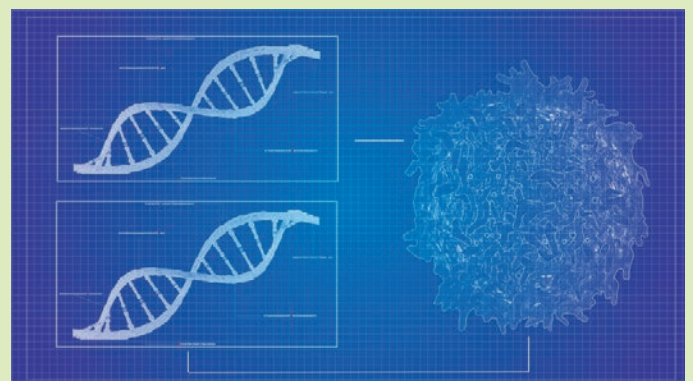
The processing near / inside memory paradigm: we have been examining massively parallel processing of the genome inside main memory using very simple bitwise operations and leveraging 3D-stacked memory + logic techniques. This can speed up genome analysis by almost four times, according to our initial work called GRIM-Filter.

Data movement greatly bottlenecks other key workloads we use daily, such as graph processing, which is foundational to numerous applications, including machine learning, social networks, information networks such as Wikipedia and bio-informatics. As these require frequent random memory accesses with little computation per memory access, the scalability of the workloads is limited to how fast we can move the data between the memory and the processor.

Here, again, processing with minimal data movement – that is, close to where the data resides using near-memory computation units – can greatly alleviate the data movement problem and accelerate graph processing. To this end, we've devised a graph-processing accelerator called Tesseract, where graph computation is performed in simple in-order cores at the logic layer of a 3D-stacked memory + logic chip. Tesseract delivers, on average, almost 14 times better performance and eight times lower energy consumption across five key graph processing workloads, including the PageRank system underlying search engines such as Google.

I could go on and on about the importance of reducing data movement, but I'll stick to one more example: databases. Databases are everywhere in our lives, storing massive amounts of data and responding to many types of query. The latency and throughput of these queries, which directly affect user experience and the types of problem we can solve in a given time, are largely dictated by how fast we can move data from the memory to the processor and how fast we can perform operations in parallel. Hence, processing inside the memory or storage units is extremely promising for database workloads.

“A single memory access costs two to three orders of magnitude more energy than a complex arithmetic operation”





We've come up with a system, *Ambit*, which minimally modifies memory (DRAM) chips such that they can perform bulk-bitwise operations in parallel. We've shown that in-memory processing using this system can deliver 12 times faster query execution. Since bulk bitwise operations are in workloads from genome analysis, to cryptography, to web searches, the impact could be significant.

What about storing data? Could this be improved?

Absolutely: another way to reduce energy wastage and performance loss is to make sure that the vast amounts of data we store are actually stored efficiently. There are two approaches.

The first is to design new technologies that are fundamentally more energy efficient, as they do not need to be refreshed as often as the memory technologies we have today (dynamic random access memory, or DRAM). Promising contenders include phase-change memory, spin torque transfer magnetic memory (STT-MRAM) and resistive random access memory (RRAM). We have shown that we can more than halve main memory energy consumption by carefully replacing DRAM with STT-MRAM.

The second involves changing the design mindset to maximize energy efficiency in current technologies like DRAM (used in main memory) and flash memory (used in solid state drives). Our solutions (called RAIDR and AVATAR) show that one can reduce the number of refreshes necessary to keep data intact in modern DRAM by four times through intelligent design of the memory controller. We can also reduce performance loss by two to four times with a combination of techniques that make the memory controller more intelligent in handling latency.

“In-memory processing can deliver 12 times faster query execution”

Can we just incrementally improve existing technologies to overcome data bottlenecks? Or do we need revolutionary change?

I think that both are needed. Improving existing technologies is easier in the short term, but their benefit is limited. Radically new approaches, like processing in memory, require a paradigm shift and a lot of cross-layer effort, but can provide huge benefits that may not be possible with incremental improvements.

That said, we cannot ignore the world if we want to change it; we should make it easy to gradually move from the existing paradigm to any radically new paradigm. This is why we've developed mechanisms which take advantage of processing in memory with minimal changes to the existing system, such as our work on 'PIM-enabled instructions'. Even this very simple approach improves performance by 40-50% and reduces energy consumption by about 25% on average across key data-intensive workloads.

How else can we make computing systems more efficient?

As we face increasingly harder problems in technology scaling at the device and circuit levels, I believe we have an exciting opportunity to reinvent how we architect computing systems, with a focus on energy efficiency. We've talked about processing in memory or storage, and more efficient new technologies for memory and storage. There's a third paradigm which is already taking place, but which needs to gather more speed: specialized architectures for key applications, such as genome sequence analysis. We need research across the hardware / software interface, intersecting all layers of the computing stack, to enable new paradigms that are fundamentally more efficient and sustainable.

A longer version of this interview and all research papers referenced are available at: bit.ly/Onur_Mutlu_interview

Videos of Onur Mutlu's ACACES18 course are available on the HiPEAC YouTube channel: bit.ly/ACACES_playlist

Delivering the data

How HiPEAC is revolutionizing data centres, storage and networking

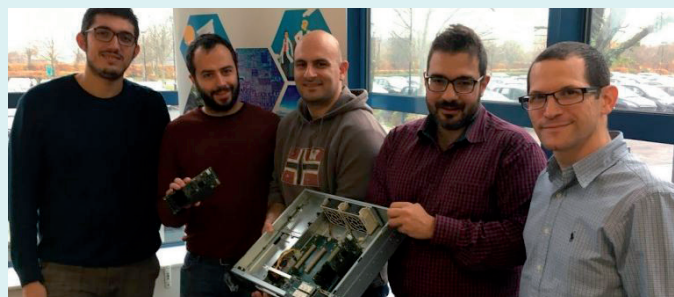
With data underpinning so much of the technology affecting our daily lives, from web searches to medical diagnoses, computing systems have to be able to move, store and process data efficiently. In this special feature, we look at how the HiPEAC community is reimagining the data centre, turbocharging interconnects and revitalizing supercomputing storage to remove bottlenecks and enhance performance.

DISAGGREGATING THE DATA CENTRE

Few would argue that the majority of data centres, sprawling over large areas and with the energy needs of a small town, are often bulky and inefficient. ‘From a 100ft perspective, today’s data centres can be seen as a collection of high-end servers interconnected via high-speed networks,’ says Andrea Reale, research engineer at IBM Research Ireland and co-ordinator of the dReDBox project. ‘Each of these servers is a monolithic unit with one or more processors, a varying number of memory banks, and, often, fast-storage and peripherals, all attached to the server’s motherboard.’

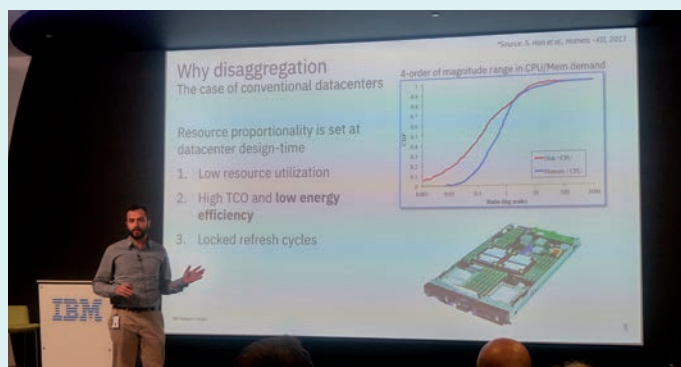
The problem, argues Andrea, is that the amount and proportion of resources of each server – such as the memory-to-central processing unit (CPU) ratio – is fixed during the data centre design and provisioning phase. Hence it is unable to respond to the extremely diverse requirements in terms of the type and amount of computing resources that modern cloud applications have. ‘This causes the resources of individual servers to be depleted asymmetrically, resulting in unused capacity due to fragmentation and, ultimately, in a large energy footprint and higher overall costs,’ he adds.

What if you could break up the server monolith, allocating sufficient hardware resources to meet individual computing



needs? That’s the idea behind the European Union (EU)-funded dReDBox – or Disaggregated Recursive Datacenter in a Box – project, explains Andrea. ‘We aim to disaggregate servers into composable, physical “resource units”, or bricks. Each brick provides a certain type of resource, mostly CPUs, memory and accelerators, and different types of bricks can be mixed and matched within resource trays.’ These custom resource trays can be used to build full racks, which in turn could form the basis of a full data centre.

To communicate between one another, ‘the bricks are connected by a circuit-switched, high-speed, low-latency, opto-electrical network’, says Andrea. ‘The system features an intelligent orchestration platform that can build on-demand computing platforms that precisely match the resource requirements of any application by connecting and allocating bricks on the fly.’

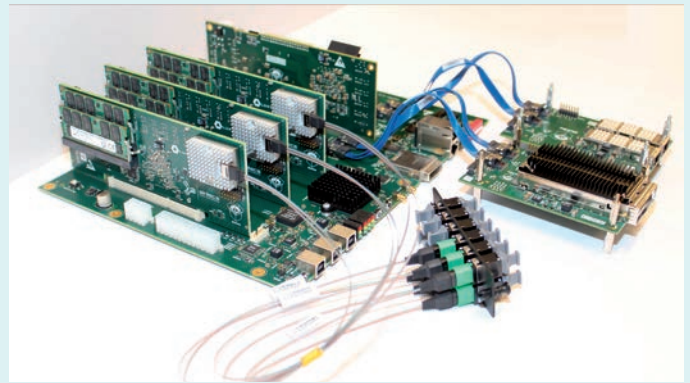


Efficiency, energy and cost gains

Pooling resources in this way offers significant efficiency gains, Andrea notes. 'Fragmentation is drastically reduced and unused bricks are automatically powered off, with an immediate impact on utilization, energy consumption and the total cost of ownership. Our initial projections forecast a reduction in resource fragmentation of up to four times for a mixed "cloud-style" workload in a medium-sized data centre, with eight times and 27 times the number of switch-off opportunities for CPUs and memory, respectively.'

In contrast to other approaches along similar lines, Andrea says that there is no need for developers to change their code or even recompile it to take advantage of dReDBox disaggregation. 'In addition, those applications which are aware of disaggregation can take further advantage of the system by using features such as a scale-up application programming interface (API) allowing application code to request additional hardware resources at runtime, a shared-memory API enabling many-to-many communication, and real zero-copy virtual machine migration.'

So far, the project team has built 'a small-scale but fully fledged hardware and software prototype of the system', says Andrea, and they are running end-to-end application benchmarks to evaluate its potential while working to increase the scale. 'We've chosen three use cases to test the system's capabilities: a high-performance computing (HPC) application for online 100G network traffic analysis exploiting disaggregated accelerators; a memory-intensive network function virtualization application exploiting the system's elasticity; and video surveillance and analytics applications using the system interconnect across nodes. The initial results are very promising,' he adds.



Creating such a complex system drew on the skills of a team with diverse backgrounds and skills, says Andrea, with expertise in fields such as computer architecture, optical networking, hardware design and integration, operating systems virtualization and distributed computing. 'Our target is to make the system a living proof of concept to study the benefits of disaggregation, in the hope that it could be a blueprint for the data centre of the future,' he concludes.

dReDBox is funded by the European Commission's Horizon 2020 programme under grant agreement number 687632. IBM Research – Ireland coordinates the project consortium which is made up of other 10 academic and industrial partners, including Barcelona Supercomputing Centre, the Foundation for Research and Technology – Hellas (FORTH), HUBER+SUHNER Polatis, Kinesense, Naudit, Sintecs, Telefonica, University College London, University of Thessaly, and Virtual Open Systems.

FURTHER INFORMATION:

dReDBox website: dredbox.eu

dReDBox video: bit.ly/dReDBox_video

A FITTER, MORE AGILE CLOUD

In the early days of the cloud, decommissioning privately owned data centres and migrating their applications to large-scale public or private clouds offered unprecedented savings in infrastructure,



operational and administrative costs, notes ACTiCLOUD scientific manager Georgios Goumas (Institute of Communication and Computer Systems, National Technical University of Athens).

However, now the cloud has come of age, it's looking a little... bloated, says Georgios. 'A single cloud data centre often hosts hardware equipment worth several million euros, may consume power in the order of tens of megawatts and services thousands of users bringing in highly diverse applications.' Yet, as Georgios points out, within this computational paradigm, there is paradoxically both resource waste and resource shortage.

The entire cloud stack is to blame, he says: 'Users tend to overestimate the resource demands of their applications and therefore overprovision their virtual machines with more

Data centres, storage and networking

computational capacity and memory space than actually needed. Meanwhile, resource managers leave resources unassigned to help mitigate the effects of peak traffic or performance penalties due to interference from other applications. Finally, the “PC architecture” which forms the dominant compute element in cloud offerings creates both resource fragmentation, with resources being scattered around compute nodes and data centres, and resource unavailability, since some resource-hungry applications place demands that significantly exceed what a typical cloud server can offer at a reasonable cost.’

Hence ACTiCLOUD takes a holistic approach to tackling resource waste and resource shortage. ‘We’re working on a deep stack involving cloud management, hypervisor technologies, language runtime systems and hardware architecture,’ explains Georgios. ‘Rather than placing expectations on programmers or application users in terms of more judicious resource requests, code reengineering or information processing, we only utilize information regarding expected and actual quality of service (QoS), where available. ACTiManager – our innovative resource manager accompanying OpenStack – monitors the health of running applications and steps in to cope with problems such as peak traffic, overload, underload and compromised QoS due to interference.’

In terms of hardware, ACTiCLOUD is working with two European hardware technologies to push capabilities beyond the state of the art: ‘Numascale systems, which scale up resources for big data, and KMAX systems, which scale out in densely integrated



racks suitable for low-power, form-factor data centres and edge computing. Both are virtualized with OnApp’s Microserver, a lightweight, type 1 hypervisor that reduces the virtualization overhead and manages flexible pools of resources at the rack scale,’ says Georgios. As a result, ACTiCLOUD is able to multiply computational and memory resources by at least an order of magnitude compared to typical cloud installations.

ACTiCLOUD combines these technologies with two European database frameworks: MonetDB, who pioneered the column-store approach, and Neo4j, world leaders in graph databases, to respond to the needs of resource-hungry applications, explains Georgios. ‘As both of these technologies use in-memory processing, there is no limit to the memory volume they can request. They form the critical substrate for applications requiring vast data processing, including business intelligence tools, streaming data analytics, fraud detection, genome annotation, intelligent transport, analysis of the Panama Papers and even a dating website!’

The goal is to transform these two database technologies into ‘database-as-a-service’ products. ‘Our goal is to evolve Neo4j from a framework that “analyses large graphs” to one which “analyses huge graphs in the cloud”. Neo4j’s version 3.4 already shows a strong ACTiCLOUD influence. Meanwhile, we are focusing on enhancing Monet DB’s elasticity capabilities to use



“These technologies form the critical substrate for applications including business intelligence tools, analysis of the Panama Papers and even a dating website!”

the exact amount of resources required for each database query.’ For more on Monet DB’s analysis of query performance, see p.31.

‘What makes ACTiCLOUD unique,’ says Georgios, ‘is that it bridges the three worlds of computer systems, cloud computing and big data analytics, while building on the work of multiple EU-funded projects.’ He points out that both KMAX systems, Kaleao’s micro-server architecture, and OnApp’s Microvisor were born within the EUROSERVER project, but that the connections don’t stop there.

‘ACTiCLOUD builds on partners’ collaborations within PRACE, on Umeå University’s long tradition of cloud computing projects such as CACTUS and ORBIT, and on the University of Manchester’s participation in projects such as TERAFLUX and AXEL.’

acticloud.eu

ACTiCLOUD has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement no 732366

BREAKING THROUGH THE BOTTLENECK: OPTIMIZING FOR I/O

If we want truly efficient computing systems, it’s time to start treating input / output (I/O) with respect, argues Manolis Katevenis, professor at the University of Crete and deputy director of the Institute of Computer Science (ICS) at the Foundation for Research and Technology – Hellas (FORTH).

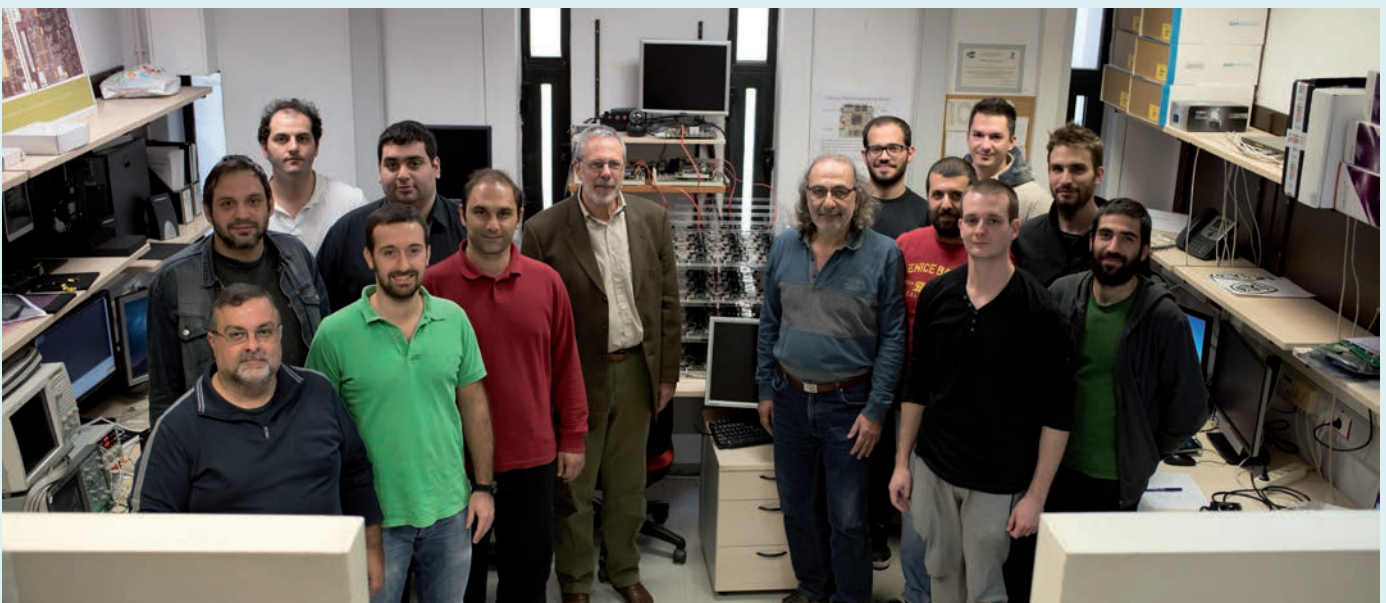
‘Broadly speaking, I/O is the ability to bring data in memory from storage or the network for processing by the central processing unit (CPU),’ he says. ‘Even in a simplified scenario where applications could create data in memory, I/O would still be important for synchronization purposes. As it is, today most applications and services need to process large amounts of data that have been accumulated or created externally to the application. Therefore, I/O has become a central aspect of modern data processing systems and computing infrastructure.’

In today’s increasingly connected world, where applications more than ever need to process data that has been produced by other applications, the need to optimize I/O is increasingly pressing, adds Angelos Bilas, also a professor and department chair at the University of Crete and senior researcher at FORTH. ‘Back-end

applications that serve mobile devices, users and services, such as maps, social media or recommendation-based systems, are typical cases where I/O is important,’ he notes, ‘and such applications are at the heart of the modern economy.’ He points out that in these cases ‘there is a constant need to move data between memory and storage or the network’. ‘Systems with slow I/O are not able to operate at full capacity and essentially waste computational resources and increase costs,’ he says.

The I/O bottleneck may be a longstanding issue, but the reasons for it have changed, according to Angelos, who points out that storage device and interconnect technologies have experienced dramatic change. ‘In the past, the bottleneck was indeed the storage devices – disks – themselves,’ he says. ‘Today, new storage technologies – mostly flash-based and emerging non-volatile memory devices – have resulted in the bottleneck shifting from the device to the host.’

This has led to a situation where it is easy to equip a server with fast and (relatively) low-cost devices, says Manolis, but the host CPU and the traditional software stack is unable to serve them all. ‘The



Data centres, storage and networking

I/O problem today is to a large extent a software-stack processing problem,' he says, adding that 'there is also a need to relax requirements on when systems synchronize during processing'.

'An interesting observation is that, as applications evolve, they allow us to design I/O – both data transfers and synchronization – in different ways, compared to what was possible in the past,' adds Angelos.

I/O for the exascale era

The issue of I/O optimization is only going to grow in importance with the advent of exascale computing systems, says Manolis. 'To scale up current systems, interconnects and the associated communication protocols need not only to transfer a lot of data, but also to do so efficiently – that is, consuming as few resources as possible, operating in a reliable manner, and offering predictable performance, especially at high loads,' he says. 'In addition, future systems will be heterogeneous, comprised of different types of processor and accelerator, and the interconnect will need to serve all these types of processing unit. These problems will apply to most applications, in both high-performance computing and data centres.'

Europe is tackling these issues with a number of European Union-funded research projects. ExaNoDe, ExaNeSt, EuroEXA and

VINEYARD are looking into 'building efficient interconnects that are tightly coupled with low-power processors, along with the required software stack', says Manolis. 'We are also examining how applications can run in the presence of failures, not to mention how future interconnects and the software stack can support heterogeneous processing units.' A further challenge, explains Angelos, is how to offer deterministic performance – the guarantee that a process will deliver a certain level of performance – although he notes that there is less work currently taking place in this direction.

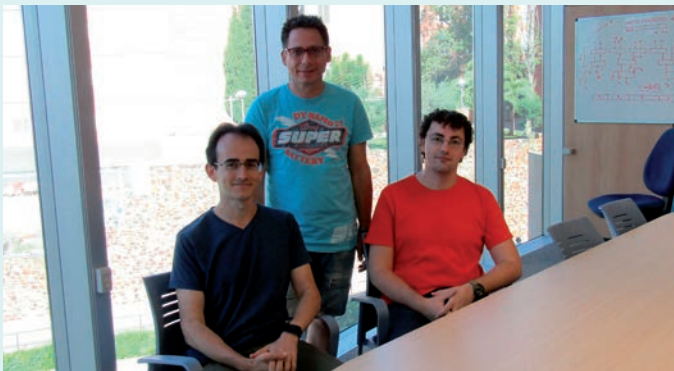
'Future computing systems may well take different forms and operate differently to what we know today,' concludes Manolis. 'Our goal is to ensure that they will be significantly more efficient to keep up with changing application needs.'

A video of Manolis Katevenis' talk 'I/O is remote load-store, and is not slower than compute, any more' is available to view on the HIPEAC YouTube channel: bit.ly/PER18_MK

FURTHER INFORMATION:

exanode.eu exanest.eu euroexa.eu vineyard-h2020.eu

REDUCING THE LATENCY GAP: A NEW STORAGE PARADIGM FOR SUPERCOMPUTING



A computer may have all the computing nodes in the world, but it is only as super as its ability to retrieve data. 'Accessing persistent data has traditionally been the bottleneck of supercomputing applications, with negative implications for applications needing to process large amounts of data,' explains Toni Cortés, leader of the storage systems group at Barcelona Supercomputing Center (BSC).

The problem, says Toni, is becoming more pronounced in the era of big data, with results produced by scientific applications needing in turn to be analysed by big-data applications. 'What's more, these newer applications bring different access patterns to

traditional supercomputing applications, thus putting additional pressure on the cluster's file system,' he adds. 'To complicate things further, many independent applications are often run concurrently on a supercomputer, meaning that data accesses from one application may interfere with another, thus slowing down both applications significantly unless they are properly coordinated.'

Bringing storage to compute nodes

To help overcome these issues, Toni's team has been investigating how to take advantage of new storage architectures. 'Fuelled by developments in non-volatile memories, there's an emerging trend to include high-density storage devices as local storage into compute nodes,' says Ramon Nou, a researcher in the storage systems group. 'In essence, this adds a high-capacity storage layer which is faster than NAND-based solid-state drives (SSDs), where supercomputing applications will be able to store and retrieve data efficiently. This brings storage closer to the computing resources and reduces the contention of the parallel file system.'

New paradigms and policies will be necessary to exploit this new storage layer, explains Ramon, along with new software to

enforce them. ‘On the other hand, since applications will now be able to run in isolation from a storage point of view, it allows you to design application-specific file systems optimized for certain application semantics,’ says Toni.

Optimizing for specific applications

The advantages of this new storage model are manifold, according to Toni. ‘First, restricting applications to node-local storage resources is a very effective way to isolate them and prevent them from arbitrarily causing contention to the parallel system. The global file system of the supercomputer would only need to be accessed before running an application (stage-in) and after an application stops running (stage-out). As an added bonus, using this new kind of device can typically be faster than accessing data in the parallel file system.’

Another advantage is that this approach can be used to deploy application-specific file systems designed to work optimally for specific workflows, avoiding the trade-offs necessary to optimize general input / output (I/O) patterns, explains Alberto Miranda, high-performance I/O researcher in the storage systems group at BSC. ‘For example, an application-specific file system could ignore the I/O synchronization of the POSIX standard, or could be tuned to access a huge number of small files, if this is how the application generates its output.’

This approach would also facilitate the planning of the stage-in and stage-out phases, says Alberto. ‘Since accesses to the parallel file systems would now be restricted to the start and end phases of a batch job, the supercomputer’s scheduler could effectively schedule them to get the best performance out of the parallel file system.’

‘We evaluated this scenario as part of our work for the EU-funded NEXTGenIO project, and we demonstrated that even with the extra overhead of the stage-in / stage-out phases, application I/O time is reduced significantly due to the benefits of a well-tuned file system and the reduced contention of the parallel file system,’ adds Alberto.



```

30 /////////////////////////////////////////////////// buffers
31 // specializations for data copy tasks
32 ///////////////////////////////////////////////////
33 template< void task_t(task_type::copy::operator()) {
34     std::error_code ec;
35     const auto tid = m_task_info->tid();
36     const auto type = m_task_info->type();
37     const auto src_backend = m_task_info->src_backend();
38     const auto src_rinfo = m_task_info->src_rinfo();
39     const auto dst_backend = m_task_info->dst_backend();
40     const auto dst_rinfo = m_task_info->dst_rinfo();
41     const auto auth = m_task_info->auth();
42
43     // helper lambda for error reporting
44     const auto log_error = [&] (const std::string& msg) {
45         m_task_info->update_status(task_status::finished_with_error,
46                                 urd_error::system_error, ec);
47
48         std::string r_msg = "[{}]" + msg + " : {}";
49
50         LOGGER_ERROR(r_msg.c_str(), tid, ec.message());
51         LOGGER_WARN("[{}]" I/O task completed with error", tid);
52     };
53
54     LOGGER_WARN("[{}]" Starting I/O task", tid);
55     m_task_info->update_status(task_status::running);
56
57     auto src = src_backend->get_resource(src_rinfo, ec);
58
59     if(ec) {
60         log_error("Could not access input data " + src_rinfo->to_string());
61         return;
62     }
63
64     auto dst = dst_backend->new_resource(dst_rinfo, src->ls_collection(), ec);
65
66     if(ec) {
67         log_error("Could not create output data " + dst_rinfo->to_string());
68         return;
69     }
70
71     ec = m_transferor->transfer(auth, m_task_info, src, dst);
72
73     if(ec) {
74         log_error("Transfer failed");
75         return;
76     }
77
78     LOGGER_WARN("[{}]" I/O task completed successfully", tid);
79     m_task_info->update_status(task_status::finished, urd_error::success,
80                             std::make_error_code(static_cast<std::error_code>::value{}));
81 }
82
83 NORMAL src/io/task.cpp operator {}() cpp 0 utf-8 17% 42/238 ln : 1 (89)tra-
    
```

Open challenges for research

Before this kind of application-specific file system can be put into practice, a number of open research questions need to be answered, however. ‘The first challenge consists of building a sufficient number of application-specific file systems to prevent application developers from having to create a new file system for each application,’ explains Ramon. ‘Fortunately, the access patterns and needs of applications are finite, and many applications could use different instances of the same file system.’

A further challenge, according to Toni, is that coordination between the job scheduler, the application-specific file system and the data scheduler is vital in order to make full use of available resources and to avoid interferences. This will require research into scheduling algorithms that take into account the new node-local storage resources and are able to optimize their usage. ‘These algorithms will need to rely on monitoring information to capture the current state of the supercomputer and the parallel file system – and given the number of nodes in contemporary supercomputers, a lot of problems need to be solved to make this system efficient and scalable.’

The benefits, though, would be felt across many application areas. ‘From personalized medicine to financial services, any application fed by large amounts of data could benefit from this kind of system,’ says Toni.

FURTHER INFORMATION:

bit.ly/BSC_storage_systems

GREEN LIGHT FOR FPGAS IN DATA CENTRES



In recent years, it's been hard to escape talk of a global data deluge. With huge amounts of data now powering developments in artificial intelligence and changing the way research is conducted, the floodgates are unlikely to close soon. This poses significant challenges for data centres, explains Dimitrios Soudris, associate professor at the Institute of Communication and Computation Systems (ICCS) and National Technical University of Athens (NTUA). 'Typical general-purpose accelerators can't cope with the increased traffic in data centres, so data centre operators are looking for novel systems that can provide much higher performance,' he says.

As a result, cloud computing and data centre operators have looked to heterogeneous computing architectures comprising different processor types to deliver high-performance and energy-efficient systems. 'We've seen general-purpose graphics processing units (GPGPUs) used in hyperscale data centres to provide cloud users with high-performance systems,' says Christoforos Kachris, senior researcher at ICCS/NTUA. 'Over the last few years, there's also been a growing trend to deploy field-programmable gate arrays (FPGAs) in data centres. These can be configured to tackle a specific, computationally intensive task and thus provide much higher performance at lower costs, compared to contemporary processors.'

However, anyone who's worked with FPGAs will know that they pose one major problem: programmability. 'Originally FPGAs had to be programmed – or configured – using hardware programming languages such as VHDL and Verilog, which describe circuits that are used to perform a specific task,' explains Christoforos. 'Hardware programming languages are much more difficult than higher level languages like C/C++ and Java, so FPGAs were mainly used for embedded systems and for modules based on a specific standard (compression or encryption). Alternatively, they were used for applications which needed high performance but where the volume of systems was low, making the development of application-specific integrated circuits prohibitively expensive.'

FPGAs for all

Fortunately, the larger FPGA vendors – principally Altera, which was acquired by Intel, and Xilinx – have started supporting the programming of FPGAs using OpenCL, says Christoforos. 'With data-driven applications, OpenCL can provide similar performance to handwritten modules developed in VHDL or Verilog.' This in turn has created a new market 'where cloud users can program FPGAs without grappling with the complexity of hardware description languages', he says.



While supporting programming with OpenCL has been beneficial, if FPGAs are to become more widespread in cloud computing developers with no knowledge of OpenCL should also be able to use them, argues Christoforos. 'Most cloud users prefer more widely used languages such as Python, Scala and Java,' he says. 'A marketplace offering third-party modules allows developers to speed up their applications without having to understand how FPGAs work. As this marketplace has developed, we've seen the emergence of a new ecosystem and market: acceleration as a service.'

The EU-funded VINEYARD project, in which Dimitrios and Christoforos take part, has contributed to this ecosystem by providing a high-level programming framework which allows end users to take advantage of hardware accelerators using common data-centre programming frameworks like Spark. The project has released an open-source repository of ready-to-use intellectual property (IP) cores for use in cloud FPGAs. 'The overall aim is for cloud users to use accelerators – both FPGAs and dataflow engines – in heterogeneous data centres in the same way as software packages and with the same flexibility as any other cloud service,' says Dimitrios.

Currently, VINEYARD IP cores focus on applications in neurocomputing, machine learning and financial applications. With their VINEYARD colleagues, Dimitrios and Christoforos have been investigating how accelerators could provide even more impressive speed-up rates through tailor-made architectures for specific applications. 'Through three different use cases, we've demonstrated that specialized architectures can run instructions up to 35 times faster while using 20% less energy,' says Dimitrios.

Read about Christoforos' acceleration-as-a-service start-up InAccel on p.30

FURTHER INFORMATION:

VINEYARD website [🔗 vineyard-h2020.eu](https://vineyard-h2020.eu)

IP repository [🔗 accel-store.com](https://accel-store.com)

VINEYARD has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 687628

Embedded intelligence at the heart of the Balkans



In June, HiPEAC coordinator Koen De Bosschere (Ghent University) and steering committee member Rainer Leupers (RWTH Aachen) presented HiPEAC at the Mediterranean Conference on Embedded Computing (MECO) in Budva, Montenegro. We caught up with MECO organizers Professor Radovan Stojanović and Professor Lech Józwiak to find out more about the embedded systems community in the Balkan Region.



What is the embedded systems ecosystem like in the Balkan region?

To a large extent, countries such as the United States, Western European countries, Japan, South Korea and China have a monopoly on embedded systems. Science and technology should be free from the barriers we often see in the political world, yet the technology gap between countries remains huge.

Lower income countries in Europe are often worryingly behind their higher income counterparts in technologies based on high-performance processors, systems-on-chip, integrated miniaturized sensors, and so on. Few hardware components for embedded systems are produced in those countries. However, the researchers and engineers in these countries have significant capacity in

developing theory, algorithms, software and smart systems applications, as well as hardware and software integration.

How does MECO contribute towards building this ecosystem?

When we established MECO, our main idea was to help reduce the gap between lower income countries in Europe and industry leaders. We also wanted to create a bridge between scientists, engineers and students from east and west, north and south.

Seven years on, we can proudly say that we've succeeded in these objectives. The conference is attended by researchers, developers and students from over 40 countries. While the majority of participants come from Southeast Europe, the Mediterranean belt and the former Soviet Union, there are also many participants from Western Europe, the United States and other advanced economies.

MECO covers the theory, technology, design and applications of systems supported by smart computers, such as embedded / cyber-physical systems and the internet of things. With an excellent citation score and strong keynote programme featuring outstanding, multidisciplinary speakers, including two Nobel laureates, the conference has become a major event in the computing systems calendar.

As the only international conference in this field in the Balkan region, MECO plays an important role in facilitating technology transfer. Most companies working on embedded technologies in this region are small or medium enterprises (SMEs), who have very limited in-house research and innovation capacity. Since most research is carried out at universities, innovation mostly

HiPEAC heads east

comes from collaboration between SMEs and university researchers, as well as the young engineers emerging from the universities.

MECO helps transfer technology directly to university professors, industry engineers and students participating in the conference; indirectly, it is transferred through the professors to students, and thereby to the region's industrial sector. Participants report that the conference has substantially improved the situation of researchers, developers and students throughout the Balkan region. More than 50% of attendees are young researchers or developers, and the already low-cost conference offers special rates for students.

What were some of the highlights from this year's edition?

This year's conference featured 12 keynote talks over three thematic days and showcased research from world-class universities, with particular emphasis on the symbiosis of modern computing, Nobel-prize-winning medicine, and a socially responsible economy. A number of talks addressed the theme of this issue of *HiPEACinfo*, including ones by Professor Veljko Milutinovic (University of Bloomington and University of Belgrade), Onur Mutlu (Carnegie Mellon and ETH Zürich) and Agnis Stibe (Paris ESLSA Business School). You can read all the presentations in the MECO library (see 'Further reading', below).

What are the main challenges in the embedded technologies sector over the next few years?

Embedded technologies should contribute towards improving society, the economy, the environment, transportation, health-care, energy, defence systems, manufacturing and so on. Smart technologies should also help improve education, expand the availability of information and communication to everyone, and help reduce poverty.

At the implementation level, challenges relate to safety, security, reliability, autonomous work, flexibility, performance, energy efficiency, correctness, testability, maintainability, manufacturability, non-recurring engineering and unit costs, etc. With regard to the ecosystem, the main challenge in the Balkan area is the development of adequate support for research, enterprise and innovation.

How can HiPEAC support researchers and engineers in the Balkan region?

MECO and MECO Net (Mediterranean Excellence in Computing and Ontology) are too small to draft a comprehensive vision and roadmap in embedded computing, but with the help of HiPEAC's 2,000-strong community we can achieve this. We aim to extend HiPEAC's activities within MECO, as well as disseminating HiPEAC's knowledge among MECO-related countries.

Participating in HiPEAC activities will not only enable knowledge transfer, but also support the establishment of international contacts and collaboration among scientists from academia and industry across Europe.

FURTHER READING:

MECO Net library

meconet.me/MECO Net-Library

Mediterranean Embedded Computing Resources

embeddedcomputing.me



Postcards from Montenegro

Over the last few years, RWTH Aachen's Rainer Leupers has been taking HiPEAC to regions which are currently underrepresented in the HiPEAC community.

This year's MECO, hosted in Budva, Montenegro – one of the oldest settlements on the Adriatic and a lively tourist destination – provided the perfect opportunity to present HiPEAC in the Balkan area.

'I was truly impressed to see that the MECO organizers were able to compile a fascinating, highly international conference programme, covering a wide spectrum of embedded ICT,' says Rainer. 'In addition to the latest results from local communities, we saw many inspiring keynotes from international speakers.' Opportunities for growth in the region, he suggests, include the region's strength 'in algorithm design for various embedded applications; in particular, medical electronics appeared to be an important field of growth'.

MECO attendees responded positively to the presentations by Rainer and HiPEAC coordinator Koen De Bosschere: 'I think they appreciated seeing how the combination of HiPEAC and the European Union (EU)-funded Innovation Action TETRAMAX provide opportunities in terms of networking infrastructure and technology transfer funding. We hope to attract many new members from the Western Balkans who will be active participants in our activities and open calls.'

Overall, the aim is to overcome the bridge between Eastern and Western European scientists, says Rainer. 'Over the past few years, Koen De Bosschere and I have visited nearly all the new EU member states, and now we are approaching countries who are on the verge of entering the EU. We've seen that many research groups and companies in Eastern European countries produce outstanding technologies and products, yet they are sometimes quite separate from Western European communities.'

'Our mission is to better connect these "hidden champions" to our well-established network structures, and thus to enable new collaborations such as project consortia beyond the "traditional" EU set-up.'

In this issue's feature on groundbreaking European research, we find out about task-based parallelism for exascale computing, transprecision computing, portability and cyber-physical systems.

Innovation Europe

SCALING UP FOR HPC WITH ALLSCALE'S TASK-BASED PARALLELISM



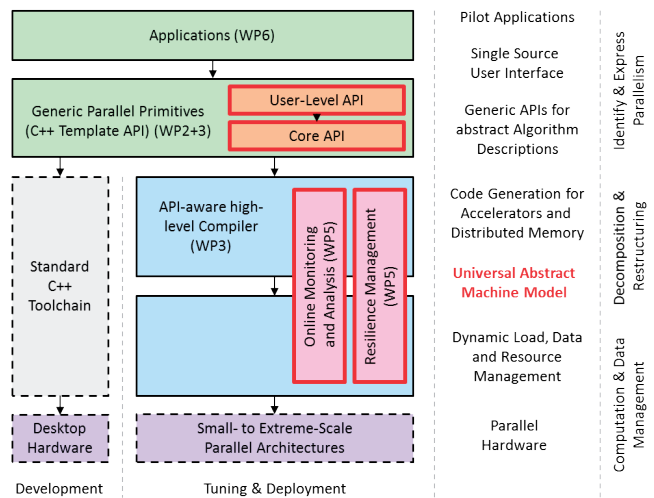
The tremendous challenge of developing applications efficiently utilizing the hardware provided by contemporary parallel systems of all scales is among the most limiting factors for the continued growth of high-performance computing (HPC). The AllScale project address this problem by providing a toolchain for the effective development and execution of such applications.

The design of the AllScale toolchain is based on three key principles:

1. Enabling the separation of responsibilities in the development of HPC applications.
2. Utilizing industry standard programming languages, such as C++, and preserving compatibility with existing development and debugging tools.
3. Employing advanced programming language, compilation and runtime system technology, such as the Insieme compiler and the HPX runtime system, to transparently integrate sophisticated services into parallel applications.

The AllScale toolchain boosts the productivity of parallel applications, as well as their portability and runtime efficiency. It will reduce energy needs, thus improving the resource-efficient utilization of small-scale to extreme-scale parallel systems.

In the diagram above right, you can see a view of the AllScale environment. On the left, you can see that the system runs on the desktop, suitable for development, while the right-hand side shows the full system running on a large computer.



The AllScale stack

The overall stack comprises the AllScale API, compiler, and runtime system. The user-facing API provides an extensible set of parallel primitives and data structures, which are converted by the compiler to fit the requirements of the runtime system. As part of this conversion, the compiler generates boiler-plate code covering, for example, data dependencies obtained through static program analysis. The runtime system is based on the tuple space paradigm, an approach that is better suited for emerging exascale systems than a combination of OpenMP and MPI.

Resiliency is a key part of the AllScale system, and is also enabled by the tuple paradigm. This uses a scheduler-dependent approach.

An advantage of using the AllScale system is that applications automatically benefit from its scheduling, resiliency and monitoring functionality. The project has three pilot applications which are being implemented with the toolchain. These cover a diverse range of application domains:

AMDADOS: An oceanography application modelling, in particular, oil dispersion from oil rig accidents. It is being actively developed by IBM R+D Dublin.

FINE™/Open developed by Numeca International to model high bypass turbofan jet engines.

iPIC3D: an application to model the interaction of solar wind with the earth's magnetosphere. It is maintained and developed by KTH.

International; Sweden: KTH Royal Institute of Technology;
Germany: Friedrich-Alexander Universität Erlangen-Nürnberg
BUDGET: €3.37 million
WEBSITE: allscale.eu

AllScale has received funding from the European Union's Horizon 2020 Programme under grant agreement no. 67160.

Interested in finding out more? Download AllScale publications via the project website allscale.eu.

NAME: AllScale: An Exascale Programming, Multi-objective Optimisation and Resilience Management Environment Based on Nested Recursive Parallelism

START/END DATE: 1 September 2015 – 30 September 2018

KEYWORDS: exascale computing, resiliency, task based parallelism

COORDINATOR: Professor Thomas Fahringer, University of Innsbruck, Austria

PARTNERS: Austria: University of Innsbruck; UK: Queen's University of Belfast; Ireland: IBM Research Ireland; Belgium: Numeca



AMDADOS models oil dispersion in the marine environment

PUTTING TRANSPRECISION COMPUTING INTO PRACTICE WITH OPRECOMP



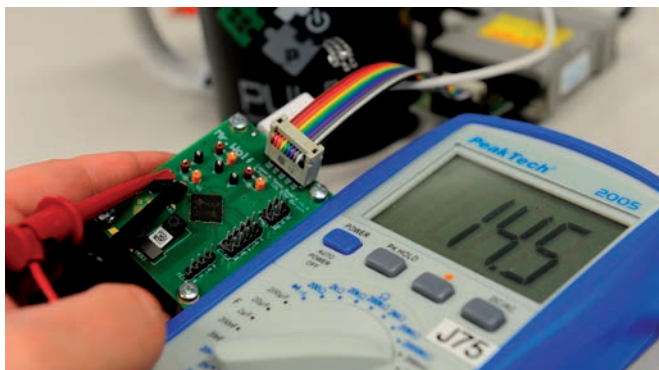
Open Transprecision Computing

Funded under the European Commission's Horizon 2020 programme, as part of the Future and Emerging Technologies Proactive initiative, OPRECOMP is a four-year research project which aims to build an innovative, reliable foundation for computing based on transprecision analytics.

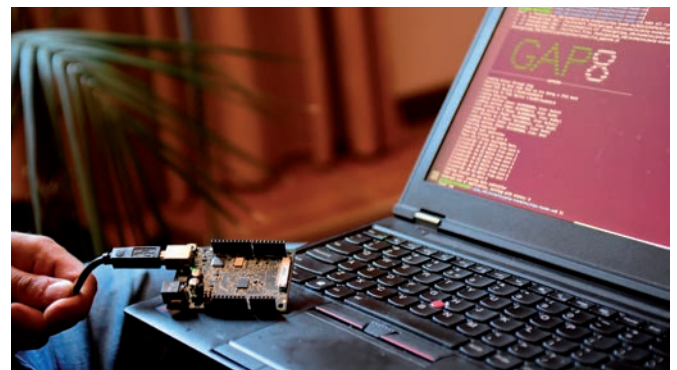
Guaranteed numerical precision of each elementary step in a complex computation has been the mainstay of traditional computing systems for many years. This era, fuelled by Moore's

law and the constant exponential improvement in computing efficiency, is at its twilight: from tiny nodes of the internet of things (IoT), to large high-performance computing (HPC) centres, sub-picojoule/operation energy efficiency is essential for practical operations. To overcome the 'power wall', a shift from traditional computing paradigms is now mandatory.

OPRECOMP aims at demolishing the ultra-conservative 'precise' computing abstraction and replacing it with a more flexible and efficient one, namely transprecision computing. The project is investigating the theoretical and practical understanding of the



Measuring the current consumption (14.5mA) of Mr. Wolf, a multicore PULP system with transprecision capabilities, while in operation



GAPUINO board running at the NIPS summer school on energy-aware transprecision computing in Perugia

energy efficiency boost obtainable when accuracy requirements on data being processed, stored and communicated can be lifted for intermediate calculations.

While approximate computing approaches have been used before, in OPRECOMP, for the first time ever, a complete framework for transprecision computing, covering devices, circuits, software tools and algorithms, along with the mathematical theory and physical foundations of the ideas will be developed. This will not only provide error bounds with respect to full precision results, but also will enable major energy efficiency improvements even when there is no freedom to relax end-to-end application quality-of-results.

The mission of OPRECOMP is to demonstrate using physical demonstrators that this idea holds in a huge range of application scenarios in the domains of IoT, big data analytics, deep

learning, and HPC simulations: from the sub-milliwatt to the megawatt range, spanning nine orders of magnitude. With a view to enabling industrial exploitation, we will prove quality and reliability and demonstrate that transprecision computing is the way to think about future systems.

NAME: Open TransPRECision COMPuting (OPRECOMP)

START/END DATE: 1 January 2017 – 31 December 2020

KEY THEMES: approximate and transprecision computing; architecture and circuit design; innovative technology; low-power, low-energy computing

PARTNERS: Switzerland: IBM Research GmbH, ETH Zürich;

France: CEA, Greenwaves Technologies; Italy: University of Perugia, University of Bologna, CINECA; Spain: Universitat Jaume I;

UK: Queen’s University Belfast; Germany: University of Kaiserslautern.

BUDGET: €5.9 million

WEBSITE: oprecomp.eu

NEW PROJECT EXA2PRO TO ENHANCE PRODUCTIVITY AND ENSURE CONSISTENT PERFORMANCE ACROSS EXASCALE COMPUTING PLATFORMS



Advancing scientific progress in domains such as physics, energy and materials science by developing and deploying applications in computing clusters of hundreds of thousand central processing unit (CPU) cores is exciting, yet very challenging. The European Commission-funded project EXA2PRO – Enhancing Programmability and boosting Performance Portability for Exascale Computing Systems – will design and develop a framework to enable the efficient deployment of applications on supercomputing systems.

This project is developing a set of tools that addresses challenges imposed by the increasing heterogeneity of supercomputing centres. The EXA2PRO framework will include:

- high-level software abstractions that provide performance portability – that is, consistent performance on different computing architectures
- memory management tools that improve memory utilization
- a runtime system that provides multi-criteria scheduling and load balancing in heterogeneous computing architectures

Additionally, the EXA2PRO framework will integrate novel software engineering tools that improve the design quality of the source code of applications designed for supercomputing centres, as well as allowing users to quantify the economic impact of source code quality improvements.



Four applications from four different domains – high-energy physics, materials science, chemical processes for cleaner environment and energy storage – will be deployed at Jülich supercomputing centre using the EXA2PRO framework. By exploiting the accelerators of heterogeneous systems, the simulation of more complex problems will be possible and the results that will be obtained are expected to have significant impact in the scientific community and industry.

To find out more, contact Dimitrios Soudris (project manager): dsoudris@microlab.ntua.gr

NAME: EXA2PRO: Enhancing Programmability and boosting Performance Portability for Exascale Computing Systems

START/END DATE: 1 May 2018 – 30 April 2021

KEYWORDS: high-performance computing, performance portability, heterogeneous computing systems.

PARTNERS: Greece: Institute of Communications and Computer Systems (ICCS); Centre for Research and Technology Hellas (CERTH); France: Institut National de Recherche en Informatique et en Automatique (INRIA); Centre National de la Recherche Scientifique (CNRS); Sweden: Linköping University; Germany: Forschungszentrum

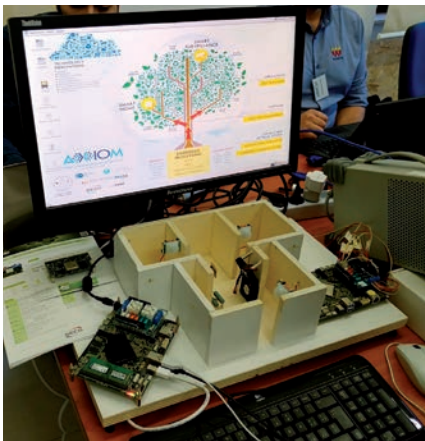
Jülich; UK: Maxeler Technologies Limited.

BUDGET: € 3.4 million

WEBSITE: exa2pro.eu

The EXA2PRO project has received funding from the European Union's Horizon 2020 Programme under grant agreement no. 801015.

AXIOM DELIVERS SCALABLE CYBER-PHYSICAL ARCHITECTURES FOR SMART ELECTRONICS



Regular readers of HiPEACinfo may remember our interview with AXIOM project partners in issue 48 and our peek under the hood of the AXIOM board in issue 51. Here, UDOO Community and Media Manager Laura Gales (AIDLIAB) explains how this innovative project created a flexible, energy-efficient board and the software to scale it up.

Last March saw the final review of the AXIOM project, where the multidisciplinary consortium – comprising three research institutions and four companies – presented the project's numerous accomplishments to stakeholders from the European Commission's Horizon 2020 programme.

The project, which started in 2015, was coordinated by Professor Roberto Giorgi of the University of Siena, who oversaw the collaborative effort between the European partners to create new software / hardware architectures for smart cyber-physical systems (CPS), to meet the growing needs of distributed and autonomous systems.

'The AXIOM board is unique the world in combining the Arm 64bit, the Arduino socket and reconfigurable logic,' explains Roberto. 'The board was designed with various applications in mind, including smart homes and living, and smart video surveillance.'

The project fulfilled all six principal objectives. Thanks to consortium member SECO, a company involved in the embedded systems market since 1979, it managed to surpass its first objective, producing a small board that is flexible, energy efficient and modularly scalable. In addition to managing the hardware development and production of board prototypes, SECO also devised a SMARC Rel. 2.0 compliant module for the industrial market. To produce a truly scalable system, consortium member the Foundation for Research and Technology-Hellas (FORTH) managed the interconnection development. The result was the 'AXIOM-link': a fast, low-latency, inexpensive and easy-to-operate way to interconnect multiple AXIOM Boards using widely available USB-C cables.

The second objective, easy programmability of multicore, multi-board FPGAs, was achieved by Barcelona Supercomputing Center (BSC) who contributed their extensive expertise in software toolchain development and high-performance computing. Early on, they presented the project with the OmpSs programming model they had developed previously, which was then further refined to fit the requirements of the AXIOM project.

BSC also supported the third objective, leveraging open-source software to manage the board, which was led by the engineers at Evidence, an enterprise focusing on open-source embedded systems.

The fourth objective, easy interfacing with the cyber-physical worlds, drew extensively upon the real-world experience of the enterprise members SECO, Vimar and Herta, all of whom brought expertise from their respective sectors – embedded systems, smart homes, and biometric security. Indeed, all partners contributed collectively towards the goals of the project, with each member bringing their field-specific expertise and knowledge. This approach was instrumental in successfully meeting the last two objectives: real-time guarantees for execution and contribution to standards.

One of the most interesting features of the platform is the possibility to interconnect a number of boards (a theoretical maximum of 255 boards) via USB-C cables. This allows considerable performance gains using the OmpSs programming model. Further research has been carried out by the University of Siena to demonstrate the possibility of running programs in a different execution model based on dataflow-threads.

You can see a video interview with Roberto Giorgi on the HiPEAC YouTube channel:

[🔗 youtu.be/nhGXAHY1Rgk](https://youtu.be/nhGXAHY1Rgk)

AXIOM received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 645496.



Now three years old, Cambridge start-up Kaleao, an innovation-driven server manufacturing company, is the third launched by Professor John Goodacre (University of Manchester / Arm). Here John, who is Kaleao's Chief Scientific Officer, gives an insight into how the company came about.

Creating Kaleao: a startup story



The way the Kaleao concept was born is one that will probably be familiar to many in this community: collaboration within a joint project. In Kaleao's case, it was EUROSERVER, a European project into energy-efficient servers, during which I started working with Giampietro Tecchioli, then chief technology officer of EuroTech, Italy. Giampietro headed up EuroTech's high-performance computing business and was looking at extending their product range to include general-purpose servers. At the time, my role at Arm involved helping manage university collaborations. Given the clear space between my contributions to Arm and the intellectual property and technology target for Kaleao, Arm provided me with an approved outside interest to work with Kaleao, while Giampietro agreed with EuroTech that he could make a start on the Kaleao business plan.

There are many ways to fund the start of a company, but the options for a manufacturing company are limited since there are significant capital costs associated with assembling the bill of materials and contract manufacturing services before anything can be sold. As the level of investment is beyond the typical European Union or national 'start-up' project, a comprehensive business plan was needed to find investors. Having developed this, we completed the founding team with a sales and marketing contact in the United States – the initial target market – and a

chief finance officer, while a team of young engineers in Crete, who had also worked on the EUROSERVER project, built up the engineering side.

The challenges of trying to create a company that wants to build something tangible, rather than license an idea or sell a few 0s and 1s to run on an existing platform, are no doubt pretty typical. Thanks to the team's previous experience of manufacturing, we knew it was important to focus as much effort on component selection and suppliers as on innovating and differentiating what we planned to build. Knowing that buying components in low volumes can cost orders of magnitude more than buying in volume, we needed to plan for success.

What we've designed is a server platform designed from the outset for scale-out applications to web-scale. We've come up with many new concepts that we've patented as fast as we could. The now

demonstrable sales pitch is: 'We deliver your scale-out workloads with a 10-fold performance density, using four times less energy and at a third of the price.' It opens lots of doors; who wouldn't be interested in that? We also offer a special price for academic research.

So where are we now? We've manufactured a number of early access machines, many already provided to customers, many looking for significant processing at the edge rather than being in the core data-centre segment. Having overcome unexpected delays during the design and development phase of the company, which have eaten into the funds required to ramp up mass production, we're now well on our way and we look forward to truly disrupting server markets with Kaleao KMAX.

[kaleao.com](https://www.kaleao.com)

Read more about Kaleao's novel system architecture in HiPEACinfo 48, p.18



Kaleao KMAX, 960Gb/s network, 192 8-core server instances

Turbocharging cloud machine learning applications with InAccel

COMPANY: InAccel

MAIN BUSINESS: acceleration-as-a-service for cloud applications

LOCATION: Delaware, United States and Athens, Greece

WEBSITE: inaccel.com



With the announcement of a successful €500,000 seed investment round led by Marathon Venture Capital in July, Athens-based InAccel has established itself on the technology start-up scene. Co-founded by HiPEAC member Dr Christoforos Kachris, the company provides field-programmable gate array (FPGA) accelerators for cloud computing.

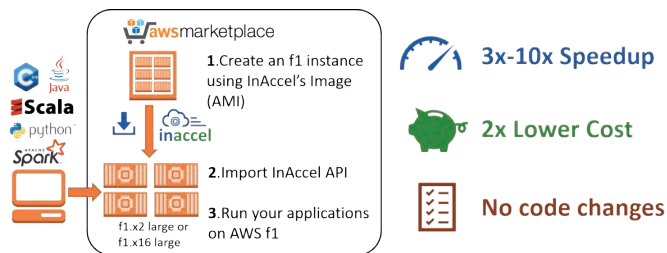
To respond to the needs of computationally intensive emerging applications like data analytics, machine learning and genomics, which require a large amount of processing power, warehouse data centres and cloud providers have turned to hardware accelerators such as FPGAs. Major technology companies such as Amazon, IBM, Intel, Baidu and Alibaba have started deploying FPGAs in their data centres, thanks to their ability to speed up application processing and reduce energy consumption.

‘Big data analytics and machine learning are crucial to the success of a company. However, associated costs are skyrocketing and computationally intensive algorithms are too slow in typical processors,’ said Christoforos, who is InAccel’s chief executive officer. ‘Our mission is to boost the performance of mission-critical algorithms with the use of FPGAs in the public cloud. The



InAccel co-founders Christoforos Kachris, Ioannis Stamelos and Elias Koromilas

use of cloud accelerators can help companies speed up their applications to gain a competitive advantage, and also to significantly reduce related operational expenses.’



InAccel provides ready-to-use intellectual property (IP) cores for FPGA platforms that can be deployed seamlessly in widely used frameworks such as Spark. As an example of the kind of speed-up that the company’s services can deliver, the company’s accelerated machine learning suite for Spark provides up to 12 times faster performance compared to 32-core contemporary processors. The service is also designed to be user friendly, requiring simply the replacement of the relevant library.

“Our mission is to boost the performance of mission-critical algorithms with the use of FPGAs in the public cloud”

Christoforos is well versed in FPGAs, having worked in the field for 17 years. He started at FPGA-pioneer Xilinx and continued his research in Athens Information Technology (AIT), FORTH (CARV) and the Microlab of the National Technical University of Athens (NTUA). InAccel’s founding team also consists of NTUA graduates and researchers Elias Koromilas and Ioannis Stamelos.

Having already deployed its first accelerators via the AWS Marketplace, InAccel will now use the seed investment to accelerate product development and distribution of its flagship products across public cloud vendors.

Find out more about InAccel in this talk by Christoforos Kachris at Spark Summit 2018

bit.ly/InAccel_Spark



Can we do better with database benchmarks? Professor Martin Kersten, CWI research fellow and chief executive of MonetDB Solutions, and Dr Ying Zhang, chief technology officer at MonetDB Solutions, think so. Here, they present the problem and their proposed solution, SQLscalpel.

Discriminative performance evaluation: Separate the bold and the beautiful

Standard benchmarks in the database arena have long been a focal point in business to aid customers to make an informed decision about a product's expected performance. The Transaction Processing Council (TPC) currently supports several benchmarks ranging from online transaction processing (OLTP) to the internet of things (IoT). Surprisingly, the number of publicly accessible results remains extremely low. Only a handful of vendors have undergone the rigorous process to obtain results for publication.

In the database community, the benchmarks TPC-C and TPC-H are commonly used to illustrate technical innovations, partly because their descriptions are easy to follow and the data generators easy to deploy. However, in many cases, only a fraction of the benchmarks are used to highlight scientific progress. Among other reasons, this is because implementing all the functionality requires an effort way beyond the resources of a single PhD candidate, which is all that most organizations are willing to dedicate. Moreover, setting up a comparative benchmark for illustrative purposes with the rigour required for a business decision is time consuming.

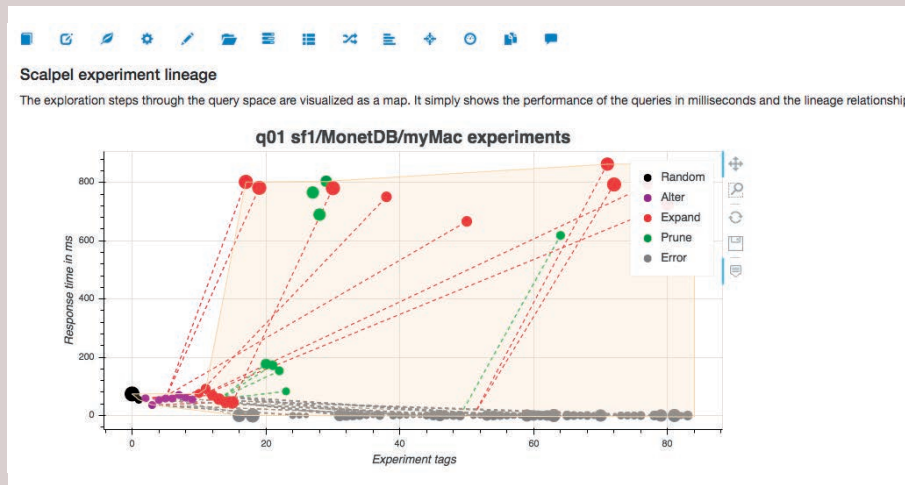
In a product development setting, benchmarks provide a yardstick for regression tests. Each new version of a database management system or deployment on a new hardware platform should ideally demonstrate better performance. Open-source systems are in that respect not different from commercial products. Performance stability and quality assurance over releases are priority number one.

Resource provisioning is one of the key challenges in modern high-performance computing (HPC) and big data applications, and has been the focus of many European research projects such as ACTiCLOUD, as described on p.17. Supercomputers are increasingly built around huge computer clusters, where distributed processing under the control of powerful software is needed. Database systems have entered this space by opening up their core for the inclusion of application-specific libraries. Within such a space, traditional benchmarks are ill suited.

After almost four decades of database performance assessments the state of affairs can be summarized as follows: 1) standard benchmarks are good for underpinning technical innovations, 2) standard benchmarks are hardly representative for real-life workloads, and 3) most benchmarking results are kept private unless they provide a success story for marketing purposes.

Better benchmarking

In a nutshell, we believe that database performance evaluation should step away from a predefined query set. Consider two systems A and B which may be different altogether or merely two versions of the same system. System B may be considered a better system overall, beating system A on all benchmarked TPC-H queries. This does not mean that no queries can be handled more efficiently in A. These queries might simply not be part of the benchmark suite. Moreover, the performance gain of B over A might have been obtained only in the restricted cases covered by the benchmark.



Therefore, the key questions to consider are ‘what queries perform relatively better on A?’ and ‘what queries run relatively better on B?’ Such queries can provide clues about the side-effects of new features or identify performance cliffs. We coined the term discriminative benchmark queries for these and the goal is to find them for any pair of systems quickly.

The second hurdle of database benchmarking is the lack of an easy way to report and share knowledge on experiments conducted – something akin to the purpose GitHub serves for software code. This hurdle stems from the so-called ‘deWitt’ clause in most end-user licence agreements, which seemingly prohibits publishing benchmark results for study and research, commentary, and professional advice, and which is certainly not intended by copyright laws.

Wouldn't we save an awful lot of energy if experimental data was shared more widely to ground technical innovations? Wouldn't we reduce the cost of running proof-of-concept projects by controlled sharing of experiences within larger teams or the database community at large?

To resolve these issues, we are in the process of realizing a software-as-a-service (SaaS) solution, called SQLscalpel. It steps away from fixed benchmark sets into

queries taken from the application and turning them into a grammar as a description of a much larger query space. SQLscalpel then explores this space using a guided random walk to find the discriminative queries. It leads to the following results:

- We extend the state-of-the-art in grammar-based database performance evaluation.
- We provide a fully fledged database performance repository to share information easily.
- We bootstrap the platform with a sizable number of online analytical processing (OLAP) cases and products.

SQLscalpel takes comparative performance assessment to a new level. It is inspired by a tradition in grammar-based testing of software. It assumes that the systems being compared understand more or less the same SQL dialect and that a large collection of queries can be conveniently described by a grammar. Minor differences in syntax are easily accommodated using dialect sections for the lexical tokens in the grammar.

A complete software stack is needed to manage benchmark projects and to deal with sharing potentially confidential information. We have taken inspiration from GitHub in the way they provide such a service. This includes the overall organization of projects and their legal structure to make the performance platform valuable to both industry and the research community.

In conclusion, HPC and big data applications need tools and software platforms to aid in the separation of the bold and the beautiful. Traditional approaches of benchmarks cast in concrete won't bring us much further. SQLscalpel will save an enormous amount of energy for (research) users when confronted with best-of-breed product selection, or best-of-breed functionality offerings, or simply when measuring progress in science by standing on the shoulders of giants.

FURTHER READING:

Adrian Vogelsgesang, Michael Haubenschild, Jan Finis, Alfons Kemper, Viktor Leis, Tobias Mühlbauer, Thomas Neumann, and Manuel Then. Get real: How benchmarks fail to represent the real world. In *Proceedings of the 7th International Workshop on Testing Database Systems, DBTest@SIGMOD 2018*, pages 1:1–1:6, Houston, TX, USA, June 15, 2018. ACM, 2018.

Martin L. Kersten, Panagiotis Koutsourakis, and Ying Zhang. Finding the pitfalls in query performance. In *Proceedings of the 7th International Workshop on Testing Database Systems, DBTest@SIGMOD 2018*, pages 3:1–3:6, Houston, TX, USA, June 15, 2018. ACM, 2018.



Over the last couple of years, HiPEAC Jobs has been organizing career-focused sessions. Named ‘Inspiring Futures!’, these bring together a range of representatives from academia and industry to share their thoughts on career growth. Here, Xavier Salazar (Barcelona Supercomputing Center) addresses some of the most common preconceptions voiced by students attending these events.

Four myths on choosing your career path – busted

‘Should I do a PhD or start working in industry?’

First of all, it’s important to understand the difference between the academic and industrial worlds. Whereas in academia the main goal is to provide a sound result on a given topic, in industry the emphasis is on providing solutions. Philippe Manet (Embedded Computing Specialists) explains this distinction using the example of firefighting: ‘Academics would have a deep understanding of the nature of fire and its chemistry. They would be able to design the best fire extinguisher. People in industry, on the other hand, would be able to set up the most efficient process to put out the fire in time with the resources allocated.’

For academics, an interesting exercise is to step back and think about the real-world problems which can be solved with the technologies being developed – as in HiPEAC’s Student Challenge. Industry people view this from the opposite perspective: they need to stay abreast of the latest technologies so that they provide better or faster solutions for problems.

‘I need to be an “A” student to do a PhD.’

Of course academic ability is important, but the grades you get become less important as you progress in academia. A good CV is built on publications in journals and at conferences – and you can start early here with special tracks aimed at students – poster sessions, research visits at other institutions, internships at companies and participation in challenges and hackathons. HiPEAC can support you with all of these activities, from displaying a poster at a HiPEAC event or participating in the HiPEAC Student Challenge to carrying out an industrial internship in a different country. In general, the most important requirement for a PhD is passion for the topic.

‘A PhD isn’t “real” work – it’s three more years of study.’

During their PhDs, students gain profound technical understanding of their research topic and related domains. However,

many other skills are also developed during this time. PhD students learn to be results oriented, capable of delivering work to deadlines, from presentations and posters to their final thesis. They gain experience of working in international, multicultural environments, such as the consortia of European Union-funded projects, and have to present results in public. Supervisors expect their students to be autonomous and take responsibility for their own work. Finally, they have to commit to achieving important, long-term goals, as anyone who has successfully defended a PhD thesis will tell you.

‘Once I’ve chosen a career path, there’s no going back.’

Industry recruiters are not only looking for people with a thorough understanding of their technology area – skills developed in academic research can be put to valuable use in industry, and vice versa. The main question anyone setting out on their career path should ask themselves is, ‘What do I want to do?’ If you need inspiration, check out the hundreds of job and internship opportunities on the HiPEAC Jobs portal.

Number of jobs posted in the last quarter by institution type	Jobs from selected institutions, sponsors, contributors and projects
58 jobs @ research centres	
61 jobs @ industry	
24 jobs @ SMEs	
78 jobs @ universities	
15 jobs @ H2020 Projects	
16 internship positions	

HiPEAC Jobs portal trends over the last quarter

FURTHER INFORMATION:

hipeac.net/jobs



‘Inspiring Futures!’ at ACACES 2018

Career talk: Cristina Silvano, Politecnico di Milano

How did you get into computer architecture?

I always liked maths and physics, but I applied to study electrical and computing engineering at the Politecnico di Milano because it seemed an innovative and fascinating field. After my MSc degree, I started my career as a design engineer working on microchips at Groupe Bull's research and development (R+D) lab in Italy. As part of a large, multinational organization, I got to travel abroad in France, Germany and the United States. My time at Bull gave me some really interesting experience: I was on the design team of several chipsets for Bull server systems and on the Bull-IBM team for the design of the first multiprocessor systems based on the PowerPC architecture.

You started out in industry and have ended up in academia. What made you decide to switch?

Nine years after joining Bull, I decided to leave my full-time position in industry to start a PhD in computer engineering so that I could follow my interests in scientific research and fulfil my wish to teach. The decision to switch was not easy: I was conscious of the difficulties and risks of starting a PhD at a later stage of my career, but I was determined to go where my aspirations led. I worked hard during that period and began investigating low-power design for embedded architectures. After my PhD, I decided to invest more time in academia as a post-doc. This turned out to be the right choice: a year later I was offered a permanent position as assistant professor at the University of Milan.

What are the main differences between industry and academia?

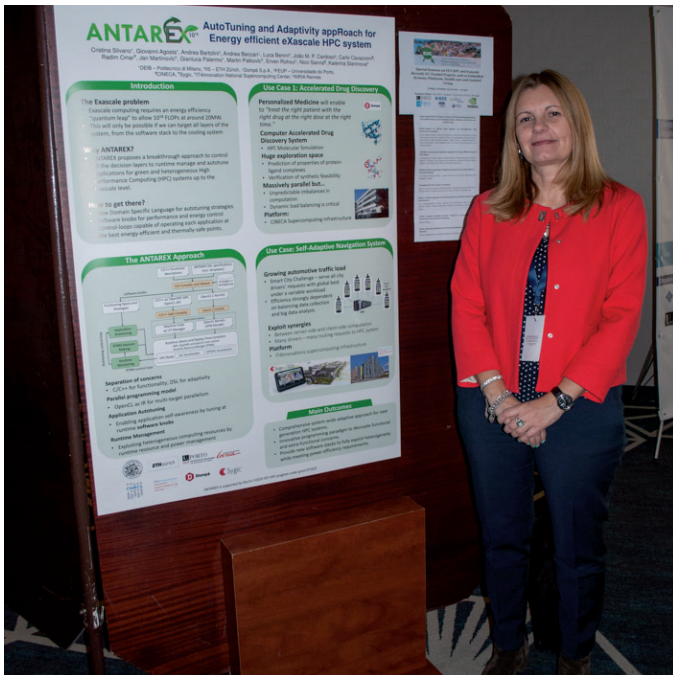
The advantage for me was that I could transfer what I had learnt during my industrial career to my academic career. While working in industry, in addition to valuable technical experience, I gained experience of working in a team and developed management skills in the context of a large international organization. I had the opportunity to see how ideas become solutions for commercial products.

“In industry, I saw how ideas become solutions for commercial products. In academia, you have more freedom to define your research paths”



However, in industry the research paths have sometimes already been set out by management and you're just part of a big team creating new products in the most efficient way. In academia, you have much more freedom to define your own research paths. The downside is that in order to follow them you need to spend a lot of your time finding research funding, which is an arduous and competitive process.

You're well known for your work on energy-efficient computer architectures. What are some of the highlights of your career? Early on in my PhD I started investigating low-power optimization and estimation techniques for embedded parallel processors. At that time, research on low-power bus encoding and instruction-level power analysis had only just started, so my papers on these



“A major challenge will be hardware acceleration of deep learning workloads, particularly maximizing performance per watt vs. absolute performance”

What are the main research challenges you see ahead?

The experience I gained researching energy-efficient computing for manycore architectures led me to a research path on application autotuning and runtime energy management for high-performance computing. This represents the core concept I proposed as scientific coordinator of the ongoing Horizon 2020 ANTAREX research project. I believe that a major challenge over the next few years will be hardware acceleration of deep learning workloads, particularly with regard to maximizing performance per watt versus absolute performance.

topics had significant research impact. These power-aware concepts were applied to the Lx/ST200 VLIW processors, designed in partnership between HP Labs and STMicroelectronics and used in a variety of embedded media processing products.

What advice would you give students about to start out on their careers, particularly women entering the field of computer architecture?

My main piece of advice to students is to follow your main interests in your career. Be curious, open-minded and keep training yourself. Working in high tech, you'll need to work continuously to make sure you keep up to date with the rapid evolution of the technologies. Keep up the passion and motivation throughout your career. Don't be worried if there are ups and downs; that's life. But never give up: life will teach you how to become resilient to events. Your knowledge and skills are your certainty and your capital to spend in the future.

After this, I started researching design space exploration techniques for application-specific manycore architectures to find power / performance tradeoffs. My research then moved to runtime resource management techniques for heterogeneous manycores. The societal and industrial impact of my research was demonstrated by the success of two European projects (MULTI-CUBE and 2PARMA) carried out under my scientific coordination.

My specific advice for young women is to believe in themselves and not let anybody tell them that they can't make it. Computer architecture is like any other scientific discipline: you need to work hard and to be passionate about the topic, but it is a really challenging and rapidly evolving field where we can do amazing things!

In 2017, I was also named IEEE Fellow for my contributions to energy-efficient computer architectures, which was a real honour. Most of these achievements have been thanks to the many great people I have met in my career, from the bosses and mentors who believed in me and encouraged me, to the colleagues and students with whom I have shared my passion for research.



Cristina with her PhD students



Arm School Program: working with partners to close the STEM skills gap



Nicholas Sample, Arm School Program Manager, explains how the new Arm School Program is addressing the STEM skills gap.

The skills gap

The gap between the STEM (science, technology, engineering and maths) skills required by industry and those acquired by learners on their journey through education is a key concern for any business which derives value from technological innovation. Across Europe, 40% of employers struggle to find people with the right skills, and the STEM skills gap shows no sign of narrowing.

Contributing factors include socio-economic drivers, such as the experiences of key groups of pupils underrepresented in STEM education and employment, and the impact of government policy.

While schools, colleges and universities are clearly central to closing the STEM skills gap, technology businesses also have a vital role to play.

Education systems

Understanding the STEM skills gap begins by analysing the experiences offered by schools, from pre-school through to university. The extent to which those experiences prepare students for a lifetime of engagement in STEM and create opportunities to work in STEM sectors is a key driver for individual empowerment and economic growth.

Consider also that education systems must prepare students for jobs of the future that might not exist yet, and the extent of the challenge becomes clear: it is not only about updating skillsets



and knowledge, but also about helping young people develop attributes that are often highly valued by STEM employers but difficult to assess, such as problem-solving, resilience and teamwork.

Technology businesses are accustomed to rapid change. In contrast, education systems change slowly. They are labour intensive, with an infrastructure that is expensive to maintain. Of course, every shift in the STEM sector cannot be mirrored by a change in school curricula. But even where education systems have identified the need to change – such as many countries' introduction of computer programming to their core curriculum – the process of change itself is often challenging at every level, from large-scale teacher training projects to the affordability of hardware.

The Arm School Program

The vision behind the new Arm School Program is to develop partnerships and work with our partners to close the STEM skills gap, through a real understanding of the complexities and challenges of STEM education. Through collaboration with schools, education companies, edtech organisations such as The Micro:bit Foundation, and educational research departments, we are identifying initiatives where Arm can make a real difference.

Our programme focuses on what really happens in the classroom, recognizing that children's school experiences are a powerful driver for change in closing the skills gap. We respect teachers' passion and expertise, and it's our role to support them as best we can. Since there is no one-size-fits-all solution, the training materials and content we're developing have adaptability and flexibility at their core.

The Arm School Program taps into Arm's own ecosystem, drawing on expertise developed by other Arm Education programmes. These include the Arm University Program, which provides college and university faculty with innovative teaching materials and technologies for computer engineering and related courses, and Arm Education Media, a subscription-based digital content hub which enables students and aspiring or practising engineers to keep up with the latest industry developments through online courses and textbooks.

FURTHER INFORMATION:

arm.com/resources/education

education@arm.com

Where will your HiPEAC career take you? We caught up with Dr Leonidas Kosmidis, Senior Researcher at Barcelona Supercomputing Center (BSC), to find out about embedded systems, HiPEAC internships and what catching a plane has to do with probabilistic timing analysis.

Probability in time and space

A journey from HiPEAC student to full member



Like many in the HiPEAC community, Leonidas' first passion at school was maths, before he was introduced to programming in his early teens and went on to study computer science. Perhaps the strongest attraction of the field, he says, was the ability to solve problems. 'You don't need much kit: with just a computer you can help advance science, unlike fields which require expensive equipment, such as chemistry and medicine.'

It was at the University of Crete and ICS-FORTH that Leonidas first became interested in computer architecture, and he went on to complete a PhD in computer architecture with applications in real time at the Universitat Politècnica de Catalunya-Barcelona Tech / BSC. 'With real-time computing systems – systems whose correct timing is as important as their functionality – you have to predict the worst-case execution time (WCET) of their tasks,' he explains. 'This typically involves using mathematics to provide strong guarantees regarding bounds. My PhD introduced probability into this area.'

Getting to the airport on time

Leonidas' PhD thesis focused on probabilistic timing analysis (PTA), a fundamental change from the deterministic approach used previously. 'Instead of providing a single value for WCET, it provides a distribution that attaches a probability to WCET values. Imagine that you need to know the maximum time it will take you to get to the airport so you don't miss your flight. Probabilistic timing analysis is the equivalent of saying with 95% probability it will take at most 20 minutes, with 99% probability we can get there in 40 minutes, and so on.'

Although PTA had been around for a long time, the method had not previously been applicable in practice, says Leonidas, because it relied on assumptions that are not satisfied in real systems, such as random timing behaviour. 'My research helped enable the use of PTA in existing systems: I showed how we can design hardware compatible with PTA, or how we can use existing

hardware with software modifications so we can introduce those properties,' he comments.

While introducing probability into the timing of safety-critical systems, such as those operating planes or cars, might seem dangerous, Leonidas suggests that this is not the case. 'Safety standards already define a maximum probability of failures which can occur in an hour of operation. It's important to bear in mind that even equipment such as a plane's engine has a probability of failure, as does computer hardware; there's no such thing as a sub-system which can never fail, so standards establish acceptable levels of failure.' Moreover, these systems are designed with multiple redundant levels of safety, to guarantee that a single failure doesn't compromise the entire system.

A solution for complex systems

With the hardware in many real-time systems reaching its sell-by date and lacking performance, upgrading to new equipment is increasingly necessary. However, the trend towards greater complexity – from out-of-order execution and branch prediction in single cores, to multicores, to heterogeneous architectures – means that the traditional approach of giving one, very pessimistic, bound, is no longer appropriate, says Leonidas.

'In the airport example above, whereas before there may have been a single straight road to the destination, now it's as if we're travelling through a city with many variables, such as traffic





Leonidas' PhD received an honourable mention in the EuroSys Roger Needham PhD Award 2018, while he won the best poster award at PUMPS 2018

lights, which make it harder to know how long it will take you to get there. The aim of my thesis was to show how we could modify the city infrastructure to make this easier.' Leonidas participated in the Spanish national three-minute thesis competition, and says that the experience was extremely valuable in helping him articulate the value of his research.

Exposure to industry

During his PhD, and thanks to a HiPEAC internship, Leonidas spent five months in Arm's Media Processing Department (MPD) in Cambridge, UK. 'I had always been fascinated by graphics processing unit (GPUs), so I chose this internship even though it was on a completely different topic to my PhD,' he says.

As it happened, this internship laid the foundations for Leonidas' future research, combining real-time systems and GPUs. Over the next four years he continued to work on embedded GPUs, starting a new research line within his research group independently from his advisors. In early 2018, in partnership with Airbus, he was awarded a European Space Agency project, GPU4Space, which considers the suitability of low-power GPUs for space applications. 'The hardware used for space payload

applications needs updating, and GPUs seem promising, but the power restrictions of satellites, which are powered by solar panels and batteries and have no fans, mean that energy-hungry GPUs aren't a possibility,' explains Leonidas. Fortunately, he emphasizes, the bulk of intellectual property (IP) in low-power, embedded GPU technology stems from Europe, with companies such as Arm, Imagination Technologies and Think Silicon featuring prominently.

Leonidas first started working with Airbus and Rapita Systems as part of the European Union-funded PROARTIS project, which explored the applicability of PTA to industry, and continued during the project's follow-up, PROXIMA, as well as in PROARTIS for Space and ESA-EFL, which transferred PTA technology to industry. Thanks to these projects, he also got to work with Cobham Gaisler, Airbus Defence and Space and SYSGO. 'This helped me adapt my PhD research so that it was relevant for industry,' he comments.

In July, Leonidas was granted full membership of HiPEAC. 'It's nice to have experienced everything that HiPEAC has to offer since I joined as a graduate student member: attending the summer school, several Computing Systems Weeks and other events. Participating in the community has allowed me to expand my network; attending the conference in particular has brought me into contact with embedded GPU companies,' he says. Participating in conferences in general is good for career development, he advises, both in terms of publishing and networking. Above all, he would recommend following your areas of interest, even when they are apparently diverse – you never know where it might lead.

"Attending the HiPEAC conference has brought me into contact with embedded GPU companies"

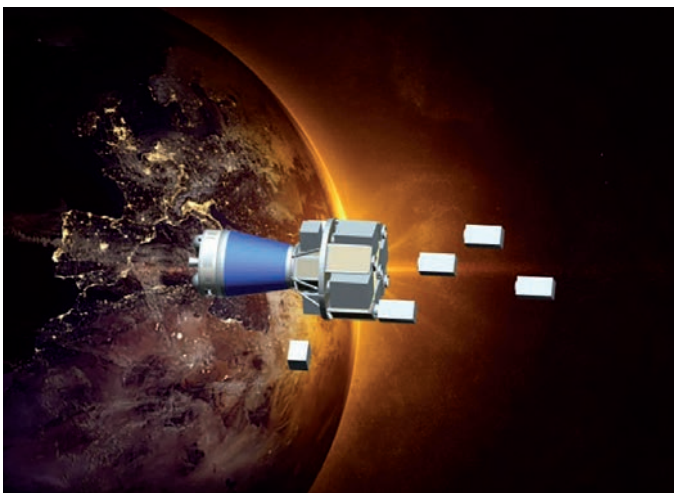


Photo © European Space Agency

HiPEAC's network includes almost 900 PhD students, producing an impressive body of research between them. This issue's featured thesis combines task-based parallelism with dataflow programming in order to resolve a host of different issues for greater performance scaling.

Three-minute thesis



NAME: Andreas Diavastos
RESEARCH CENTER: University of Cyprus
SUPERVISOR: Associate Professor Pedro Trancoso
DATE DEFENDED: November 2017
THESIS TITLE: Task Data-flow Execution on Many-core Systems

Featured research: Tasks + dataflow = max performance on many cores

In an effort to get ever-increasing performance levels out of processors, there is a growing trend of packing them with more cores. However, to get the most out of these hardware resources, the software must be able to effectively exploit parallel processing. In fact, this is one of a series of factors affecting performance scalability, including:

- the degree of parallelism
- programmability
- low-overhead runtime systems
- locality-aware execution
- efficient use of the available resources
- scalable architecture designs

What if we were to consider these different issues as aspects of the same problem? In my PhD, we set out to do just that. To provide greater parallelism, the task model – which describes work as asynchronous operations (tasks) – is used as an implementation of the dataflow paradigm, which models programs as a series of producer-consumer connections. Dataflow is the most appropriate model for exploiting large amounts of software parallelism, while a task-based implementation reduces runtime overheads and adapts easily to different architectures.

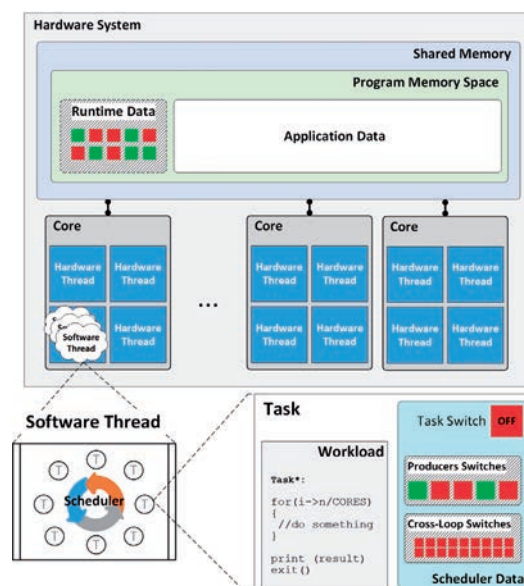
We combined the two to create the first data-driven multithreading implementation for a shared-memory many-core processor without requiring a hardware cache-coherence mechanism. This implementation comes in the form of a complete programming and execution framework equipped with a lightweight runtime system called SWITCHES. Avoiding the need for costly and complex hardware mechanisms, SWITCHES allows for simpler architecture designs that can be scaled more easily. This framework also incorporates machine-learning techniques to automatically produce scheduling policies that efficiently exploit the underlying hardware.

To address the programmability issue, we provide a source-to-source tool that automatically produces thread-based code from applications implemented using well-known programming language constructs (OpenMP). The code produced can then be compiled by any off-the-shelf C / C++ compiler, applying all existing optimizations. The framework extends the latest application programming interface (API) standard from OpenMP to support explicit task resource allocation mechanisms and variable loop task granularity as a way to increase data locality even for loop tasks with inter-dependencies. The framework along with the SWITCHES runtime system has been made publicly available on GitHub.

Since completing my PhD, I have moved to the National University of Singapore as a postdoc Research Fellow, where I am working on high-performance energy-efficient processor designs.

Andreas' supervisor Pedro Trancoso, now senior lecturer in the Department of Computer Science and Engineering at Chalmers University, commented: 'Andreas' thesis represents a breakthrough in exploiting ever-more complex hardware while minimizing system complexity for programmers. This approach gives us another way to scale up system performance as we push towards exascale computing in the next few years.'

Download SWITCHES here: github.com/diavastos/SWITCHES



SWITCHES runtime design

SPONSOR **HiPEAC** 2019

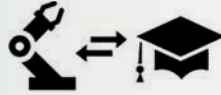


Reach scientific and technological experts

Largest computing systems research event in Europe

Extensive range of associated events

Be the first to find out about new technologies



Build industry-academia relations

20% industry participants

Attendees from around the world

Demonstrations and product announcements



Excellent recruitment opportunity

Around 25% of attendees are graduate students

Advertise vacancies on hipeac.net/jobs

HiPEAC Jobs booth and support arranging interviews

HiPEAC Conference

21-23 January 2019, Valencia

Sponsor the HiPEAC Conference and gain valuable exposure to the European computing systems community



Tailored sponsorship plans

	Gold from €4,000	Silver from €2,000	Bronze from €1,000
Free conference passes	3	2	2
Acknowledgement on HiPEAC website and conference communications	✓	✓	✓
Presentation slot during industry exhibition	20 min	20 min	10 min upon availability
Free mini-booth in industry exhibition	✓	✓	✓
Privileged booth location and customized booth	✓	✗	✗
Eligibility to be co-sponsor of student activities: travel grants, prizes, student poster session, best student presentation, etc.	✓	✗	✗

Interested in a tailored package? Contact us today

hipeac.net/2019/valencia/sponsorship

sponsorship@hipeac.net

@hipeac @hipeacjobs

hipeac.net/linkedin

