

HIPEAC

INFO

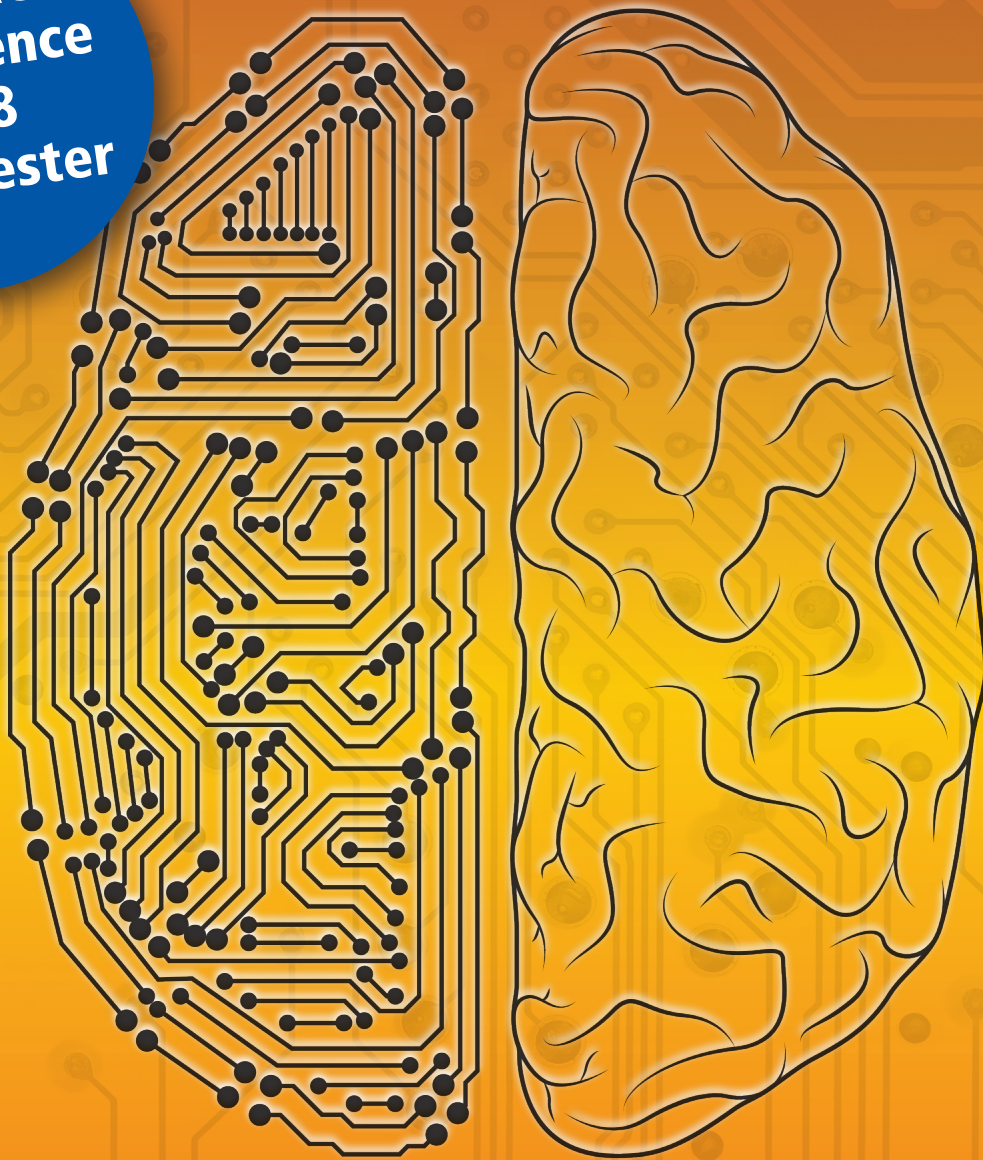
53

COMPILATION

ARCHITECTURE

APPEARS QUARTERLY | JANUARY 2018

HiPEAC
Conference
2018
Manchester



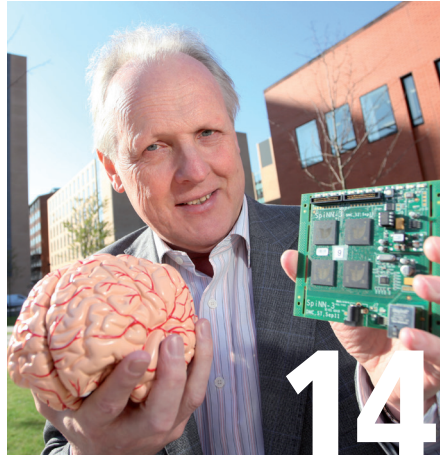
Fast learners: The unstoppable rise of machine learning

The European spin-offs taking over tech

Smarter systems, from customized architectures to workloads



Welcome to Manchester



Machine learning special feature



Innovation Europe

<p>3 Welcome <i>Koen De Bosschere</i></p> <p>4 Policy corner Artificial intelligence: the parsley in every soup <i>Sandro D'Elia</i></p> <p>6 News</p> <p>14 Machine learning special Fast learners: The unstoppable rise of machine learning <i>Steve Furber, José Manuel García Carrasco, Valentin Radu, Håkan Grahn and Oscar Deniz Suarez</i></p> <p>21 Innovation Europe LEGaTO: Plugging the software-support gap for low-energy computing <i>Osman Ünsal, Adrián Cristal and Anna Molinet</i></p> <p>22 Innovation Europe Silver lining: New ACTiCLOUD architecture for efficient management of cloud computing resources <i>Georgios Goumas</i></p> <p>23 Innovation Europe Exploiting heterogeneity through collaboration <i>Clara Pezuela</i></p> <p>25 Technology transfer Armed for success: Amanieu Systems <i>Mikel Luján and Amanieu d'Antras</i></p> <p>27 Technology transfer More ParaFormance for your money <i>Chris Brown</i></p> <p>28 Computing for innovation Cognitive discovery: Pushing the frontiers of R+D with AI <i>Costas Bekas</i></p>	<p>29 SME snapshot To the moon and back with IngeniArs <i>Camilla Giunti</i></p> <p>30 Peac performance SCRATCH: Automated generation of application-specific soft-GPGPU architectures <i>Pedro Tomás and Gabriel Falcão</i></p> <p>32 Peac performance Smarter worksharing on heterogeneous computing systems <i>Sabri Pllana</i></p> <p>33 Technology opinion On high-performance machine learning <i>Kemal A. Delic, David M. Penkler and Martin Walker</i></p> <p>35 HiPEAC futures Computing systems jobs: what's new? Career talk: Trevor Carlson, National University of Singapore Multiphysics made easier</p>
---	---

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.687698.

Cover photo: Evcrow, Dreamstime - Design: www.magelaan.be

Editor: Madeleine Gray - Email: communication@hipecac.net



Cognitive discovery



High-performance machine learning



Career talk, Singapore-style learning



First of all, I would like to wish you a healthy and prosperous 2018, personally as well as professionally.

I will remember 2017 as the year of artificial intelligence. One notable event was that the startup Vicarious found a way to break CAPTCHAs by means of machine learning. The fact that CAPTCHA stands for ‘Completely Automated Public Turing test to tell Computers and Humans Apart’ confronts us with the bare fact that computers have started outperforming us in cognitive tasks that were considered exclusively human until now.

A second notable event was that DeepMind proved that it was possible to train their technology to play strategy board games from scratch to world champion level in a matter of hours, just by giving it the rules of the game and letting it play against itself. Rather than basing their knowledge on human games, AlphaGo Zero and AlphaZero learned everything from their own successes and failures. The fact that a computer program can discover more about strategy board games in a few hours than a human player can from carefully studying the masters is quite humbling.

According to Gartner, machine learning is currently at the peak of inflated expectations. China and Russia are investing billions of euros to become the leader in artificial intelligence. All military organizations are evaluating the potential of artificial intelligence in weapon systems. In the future, the military capacity of a country may no longer be measured by the amount of firepower and the size of its army, but by the sophistication of its smart weapons and the quality of its cyber army. It could mark the start of a new arms race.

On a brighter note, 2018 also brings us HiPEAC5. In HiPEAC5 we will focus on stimulating collaboration between academic and industrial researchers, and on connecting them to the innovation community in Europe. You will hear more about it in future issues of this magazine.

Many of you will read this at the HiPEAC conference in Manchester. The HiPEAC conference is the flagship event for the HiPEAC community. I am thankful to the many volunteers who work very hard to make this event successful, and a sign of a thriving community in computing systems in Europe.

Koen De Bosschere, HiPEAC coordinator



Thanks to technological advances, artificial intelligence is poised to transform the world as we know it; but there are a myriad of obstacles to be dealt with before trustworthy, secure, reliable AI systems become common. Sandro D'Elia, Programme Officer in the Digitising Industry Unit at DG CONNECT, explains how the European Commission is laying the groundwork.

Artificial intelligence



“Just think of the difficulty of explaining a decision taken by a neural network in terms which are understandable by human beings (or, even worse, by lawyers)”

In the South of Italy, we have a saying: ‘parsley goes in every soup’. It’s easy to understand why: parsley is easy to find, cheap, blends with everything, and improves almost every recipe. It is just like artificial intelligence (AI), which is now becoming the parsley of high-tech: any application, from medical diagnosis to automatic translation or even robot bees, seems ready to benefit from some kind of AI technology. In European Union (EU) jargon, we would say that AI has been ‘mainstreamed’.

What happened, and why? The key concepts behind all those technologies which go under the name of ‘AI’, from deep learning to genetic algorithms, have not evolved dramatically over the last few years, but nevertheless we have seen some amazing developments. The best example is deep learning / neural networks: on one side, recent hardware has made neural networks usable in real-world domains; in parallel, the emergence of software libraries and open datasets for training has significantly reduced the cost of developing applications. As a result, interesting applications are popping up everywhere, from computer vision to business intelligence, and the trend seems to be growing.

A very interesting aspect of these applications is that often the software is not the critical factor. Many code libraries are open source and reusable across different domains; what is really important is the data used to train the software.

Does this mean that we should simply get used to a world where software is a commodity, and data is the most important asset? Maybe, but there is a lot more. There are significant problems to be solved to make AI in general a mature technology, for example in the areas of time criticality, energy cost and reliability. Moreover, hardware for AI has very specific requirements, and we can expect that computing architectures will have to evolve significantly to support AI requirements. However, probably the most important issue is the interaction between AI applications and humans, which can potentially change the way in which we interact with the physical world.

This is a huge problem; just think of the difficulty of explaining a decision taken by a neural network in terms which are understandable by human beings (or, even worse, by lawyers). Today ‘explainable AI’ or ‘accountable AI’ is an open research issue, but it could soon become a serious obstacle for the deployment of AI-based applications.

At the European Commission, we believe that AI has enormous potential to make our society better and our economy stronger, but this will not happen by itself. We need to ‘roll out’ AI, making it accessible for developers and innovators in all sectors, and making sure that AI skills are widely distributed.

ce: the parsley in every soup

This is why you will find a specific call (ICT-26-2018-2020) in the 2018-2020 Work Programme, aiming to build a ‘European AI-on-demand platform’. By ‘platform’, we mean an organization capable of bringing together researchers, companies and start-ups, becoming a model in AI technologies, developing what is needed by the market and boosting technology transfer, particularly towards small and medium enterprises (SMEs) and non-tech companies. In other words, we want to grow the AI community by putting together

research and industry, just as HiPEAC is doing for the computing community.

The resulting ecosystem will support the roll out of technologies across industry and academia, but more work is needed to develop basic AI technologies, to make them practically and securely usable in the industry, and to make sure that people can trust artificial intelligence. The European Commission is aware of this, and we are preparing a comprehensive initiative for 2018 which will address the various

aspects of AI: industrial capacity, the impact on the job market and society, legal issues and, of course, technology.

Today, European industry has a strong market position in ‘embodied’ AI applications, like robotics and embedded systems; we want to make sure that in the years to come Europe is a world leader in all the areas of artificial intelligence with high economic and social value. Stay tuned, because the future will bring us some very interesting news.

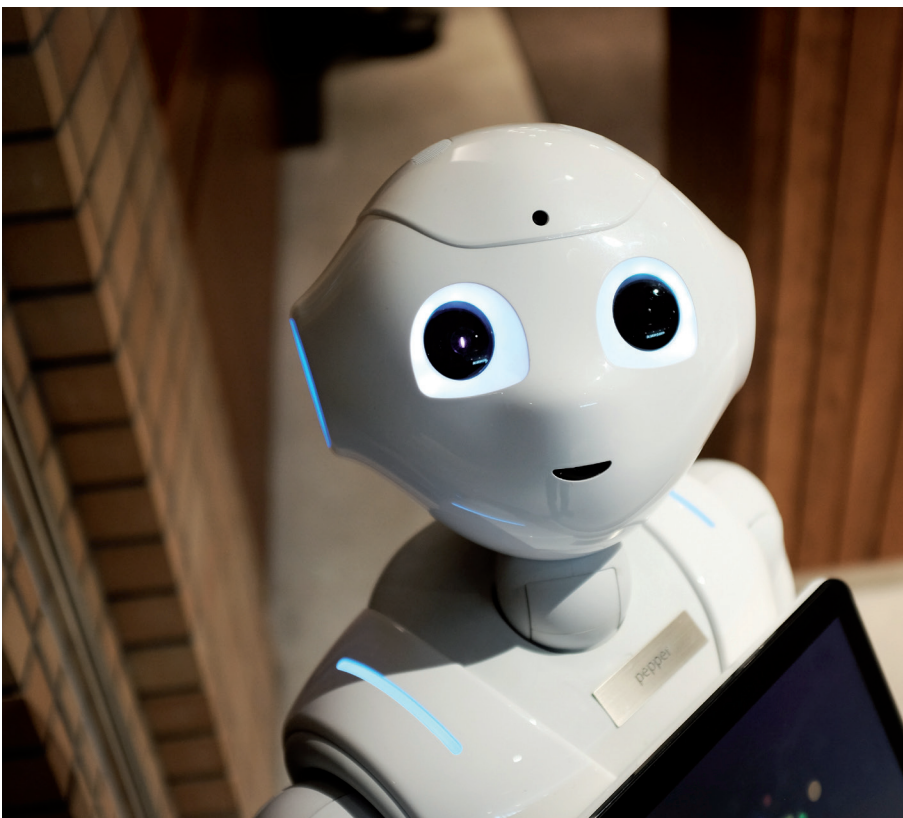


Photo credit: Alex Knight on Unsplash

“The European Commission is preparing a comprehensive initiative for 2018 which will address the various aspects of AI”

Whippin' Piccadilly

HiPEAC18 local hosts Mikel Luján and Antoniu Pop, University of Manchester, give us a flavour of what makes this year's conference location special.

1. Why is Manchester the ideal location for the HiPEAC conference?

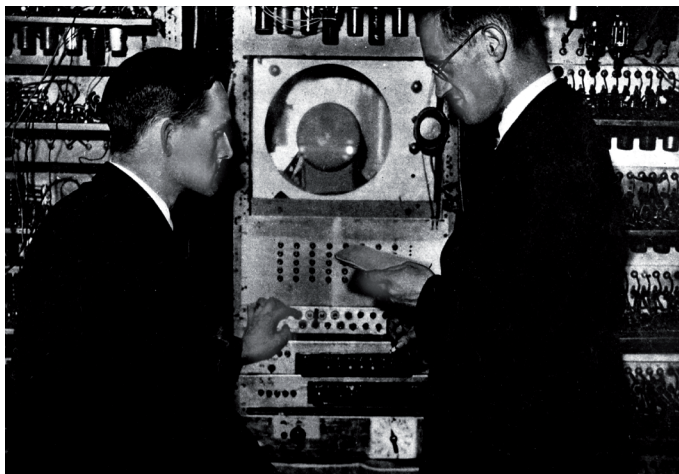
Manchester has a rich and distinguished history in computing. 2018 marks the 70th anniversary of the Manchester 'Baby', or Small Scale Experimental Machine. In other words, in June 2018, it will 70 years since the world's first stored-program computer successfully executed its first program. The 'Baby' was a testbed for the Williams-Kilburn tube, the first random-access digital storage device (i.e. an early form of computer memory). Bringing the HiPEAC conference to Manchester is the perfect homage to this historical milestone. Find out more in this short film: bit.ly/Manchester_Baby

Most people associate Manchester with great football, music and the Industrial Revolution. However, recently Manchester and the English North West have been undergoing a quiet reinvention. This is transforming Manchester's reputation into a tourism hotspot, with the Lonely Planet and the New York Times calling it a top travel destination.

Manchester welcomes the HiPEAC community to the UK at this time of exciting computing developments. It's impossible to ignore that Brexit has, unfortunately, created uncertainty and misunderstandings. However, the UK research community remains strongly committed to the ethos and critical role of the HiPEAC conference, and wishes to contribute to HiPEAC's continued success.

2. Tell us about some of the work at the University of Manchester.

The University of Manchester is the largest single-site university in the UK with more than 40,000 students and 10,000 staff. Thus, there are many exciting things happening across the campus, such as the Square Kilometre Array (SKA) telescope headquarters, the Graphene Flagship and the Human Brain Flagship.



Tom Kilburn and Frederic Williams with the Baby



Wander around the John Rylands Library

Thanks to our collaborations in the European Union-funded Horizon 2020 programme, we are advancing the EU high-performance computing (HPC) roadmap in EuroExa, ExaNoDe, ExaNest, ECOSCALE, and Eurolab-4-HPC. We are also actively participating in how to program heterogeneous systems and clouds in the E2DATA and ACTiCLOUD projects.

3. What machine learning technologies are you most excited about?

There are many exciting new technologies emerging, from both academia and industry – such as DeepMind, from whom we have a great keynote this year. The key theme that drives our collaboration with the Manchester machine learning (ML) group is the interaction between computational efficiency and statistical efficiency. Within HiPEAC, we're well aware of computational inefficiency, and much great work goes into ML-specific hardware, parallelizing, optimizing, and approximating ML algorithms.

Statistical efficiency refers to how a ML technique can make good use of smaller amounts of data more suitable for the internet of things/ edge computing and smartphones. Most deep learning systems are tremendously statistically inefficient, requiring full data centres of training data to build their networks/models. We have investigated methods for feature selection and extraction, as well as actively researching efficient modular learning methods, all of which contribute to statistically efficient re-use of learning systems.

4. What shouldn't we miss in Manchester?

Make sure you visit the fantastic gastro pubs in the city centre (the Ox noble, Mr Thomas's Chop House, The Wharf, Sinclair's Oyster Bar). Near the conference venue, stop by the John Rylands library and enjoy getting lost around the Town Hall and Spinningfields. If walking is your thing, head south towards the Whitworth Art Gallery or east towards the Northern Quarter. Finally, if you have more time, get out of the city and visit the Jodrell Bank Observatory.

Happy new HiPEAC!

A new phase in HiPEAC's evolution has begun: HiPEAC 5 officially started on 1 December. Over the next two years, we will be further consolidating links with industry and connecting the research and innovation communities in Europe, in support of the European Commission's Digitising Industry initiative.

To help HiPEAC reach out to industry contacts, we are joined by two new partners, ARTEMIS Industry Association and Innovalia. The four annual HiPEAC events will continue, and we will still provide financial support for research and industry placements. We've also started work on the next HiPEAC Vision, giving policy makers and industry representatives invaluable insights into the future of computing systems.

If you're working on European Union-funded research, HiPEAC is here to help you get greater visibility for your project, whether that be

through our roadshow events, media outreach or articles in the HiPEAC magazine. Meanwhile, if you're looking for top-quality staff, look no further than our recruitment services, which include the HiPEAC Jobs portal, travelling careers unit and mentoring sessions.

All of this would not be possible without the generous support of the European Commission and our industry sponsors, to whom we are grateful for their continued trust in the project.

The project is due to run until 29 February 2020.

Want to find out more about how HiPEAC 5 can help you meet your research or industry goals? Contact communication@hipeac.net.

For information on all of HiPEAC's activities, visit our website: hipeac.net

Motors for Europe: CSW Stuttgart

The autumn edition of Computing Systems Week, HiPEAC's biannual networking event, took place in Stuttgart on 25-27 October. Paying homage to the region's most famous export, the theme for this event was the automotive industry, or smart mobility more generally. Over the three days, 152 attendees from 22 countries attended sessions on trends in automotive engineering, architectures for autonomous driving, big data in mobility and transport and more.

Participants also learned about the European Commission's Smart Anything Everywhere

initiative, which provides funding and expert support for innovation through digital technologies (see p.8 for news on this programme). Other sessions presented distributed platforms for the industrial internet of things, low-power architectures for next-generation cloud and cyber-physical infrastructure, and simplifying/optimizing heterogeneity.

The HiPEAC Industry Partner Programme showcased industry innovations from the local area and beyond, while the Student Programming Challenge and 'Inspiring Futures!' session offered HiPEAC's students the

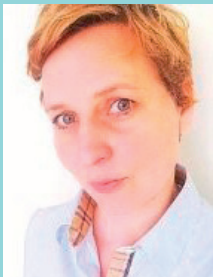
chance to prove their programming mettle and get advice on both business and research career paths. The HiPEAC Jobs wall also displayed the impressive range of open vacancies on the HiPEAC Jobs portal (see p.36 for more on this), while the poster session allowed researchers to share European project findings and industry representatives to scout for highly qualified new team members.

The next edition of Computing Systems Week will take place in spring 2018 – check hipeac.net for further information.

hipeac.net/csw/2017/stuttgart



First TETRAMAX call now open



Katrien Van Impe, Dissemination and Communication Officer, TETRAMAX

TETRAMAX focuses on customized low-energy computing for cyber-physical systems and the internet of things within the framework of the European Smart Anything Everywhere (SAE) initiative. Over

the course of the project (September 2017 – August 2021), there will be several open calls offering you the opportunity to contribute to TETRAMAX technology transfer experiments (TTX), with significant funding opportunities.

At the end of November, TETRAMAX announced the first call for bilateral TTX, which require one academic and one industry partner from two different EU countries or associated countries. In justified cases, both partners can be small/medium enterprises (SMEs).

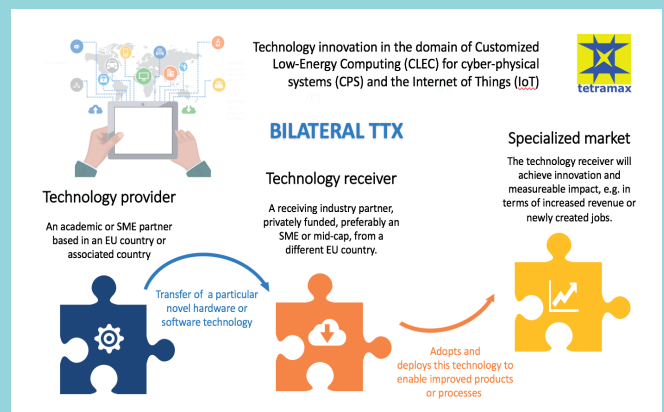
One academic or SME partner transfers a particular novel hardware or software technology in the domain of ‘Customized Low-Energy Computing for Cyber-Physical Systems or the Internet of Things’ to a receiving industry partner (privately funded, preferably an SME or mid-cap) from a different European Union country. The receiving partner deploys this technology to improve products or processes, for example in product cost or performance gains, or reduced power

requirements. Thereby, the technology receiver will achieve innovation and measurable impact, for example in terms of increased revenue or newly created jobs. Funding of up to €50,000 is available for these projects, which should last between six and 12 months.

The closing date is 28 February 2018. We look forward to receiving your applications. Please send any questions to opencalls@tetramax.eu.

Further information: tetramax.eu/ttx/calls

The TETRAMAX project has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement number 761349



€ 60,000 towards your next cyber-physical product



How about a cash injection and expert support to make the cyber-physical products you’ve always dreamed of a reality?

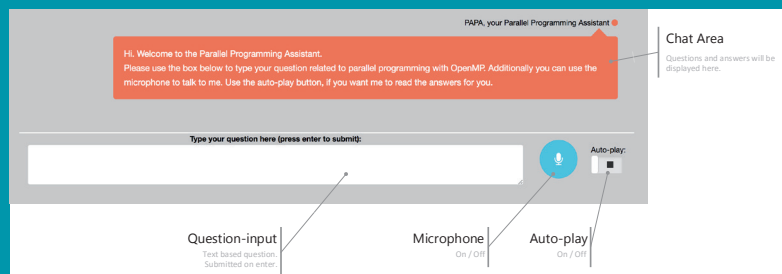
Led by CEA-Leti, FED4SAE – which stands for Federated CPS Digital Innovation Hubs for the Smart Anything Everywhere Initiative – is an acceleration programme available to any European company looking to develop new products and business models based on cyber-physical systems, and thereby lead the digitization of European industry. Co-funded by the European Commission, the programme is designed for European start-ups, SMEs and midcaps addressing exciting new markets, such as smart cities, smart agriculture, smart food, smart health and wellbeing, smart building, smart transport and others.

Beneficiaries will get access to advanced platforms (advanced technologies and testbeds) and industrial platforms, as well as appropriate technical, business and innovation management support to turn their ideas into commercial products. They will also receive up to €60,000 grant funding to support first development. This first-level investment is expected to be further completed beyond the acceleration programme through private and public funding. The first call for applications is open until 6 February 2018.

Find out more on the FED4SAE website: fed4sae.eu

FED4SAE has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement no. 761708

A Siri for parallel programming



Know someone who needs a helping hand with parallel programming? Researchers at Linnaeus University have developed a cognitive-based digital assistant to help developers' code get the most out of a given platform's resources.

'While several models for parallel programming have been developed, it's still easy for beginners to make mistakes that may lead to lower performance or unexpected program behaviour,' explains Sabri Pllana, leader of the High-Performance Computing Center and Associate Professor at Linnaeus University. 'In a similar way to Apple's Siri, our Parallel Programming Assistant (PaPA) can answer questions related to parallel programming. You can ask it questions and it will search its knowledge database for an appropriate answer, interacting in real time through text and speech.'

Students studying parallel programming at Linnaeus University have been evaluating PaPA, with preliminary results showing that the assistant gives helpful answers for novice programmers. In turn, the students have shown willingness to use the digital assistant as they develop their applications. Based on this, the researchers believe that PaPA could be used as an educational resource for introductory parallel programming courses.

Further information: bit.ly/PPA_Linnaeus

Samsung Galaxy GameDev website

Samsung has launched an online resource to support developers. The website includes tutorials, user guides and tech articles to help develop Vulkan graphics rendering and optimize gaming applications. The page also provides code samples and additional development tools.

Future additions to the page will include a tool suite by the EU-funded project LPGPU2 (Low-Power Parallel Computing on GPUs), which will help developers optimize software for low-power devices.

Further information: developer.samsung.com/game

ReQuEST: First multi-objective SW/HW co-design competition at ASPLOS'18

The first Reproducible Quality-Efficient Systems Tournament (ReQuEST) will debut at ASPLOS'18, the ACM International Conference on Architectural Support for Programming Languages and Operating Systems. ReQuEST aims to provide an open-source tournament framework, a common experimental methodology and an open repository of design knowledge. These will be used for continuous evaluation and multi-objective optimization of the quality vs. efficiency Pareto optimality of a wide range of real-world applications, models and libraries across the whole software/hardware stack. The tournament is organized by a consortium of leading universities (Washington, Cornell, Toronto, Cambridge, EPFL) and the cTuning foundation.

ReQuEST will use the established artefact evaluation methodology together with the Collective Knowledge framework validated at leading ACM/IEEE conferences to reproduce results, display them on a live dashboard and share artefacts with the community. Distinguished entries will be presented at the associated ASPLOS'18 workshop and published in the ACM Digital Library. To win, the results of an entry do not necessarily have to lie on the Pareto frontier; the originality, reproducibility, adaptability, scalability, portability, ease of use, etc., of entries will also be taken into consideration.

The first ReQuEST competition will focus on deep learning for image recognition with an ambitious long-term goal of building a public repository of customizable, reusable and optimized artificial intelligence (AI) artefacts across diverse datasets and platforms, from the internet of things to supercomputers. Future competitions will consider other emerging workloads, as suggested by our Industrial Advisory Board.

For any queries regarding the Industrial Advisory Board, including participation and sponsorship, please contact anton@dividiti.com.

Further information: cknowledge.org/request

The night is dark and full of data



Ahsan J. Awan, KTH Royal Institute of Technology and Universitat Politècnica de Catalunya – Barcelona Tech

Near-data processing (NDP) enables the data to be processed where it resides, whether that be in storage or main memory. It helps to avoid costly back-and-forth movement of data between the host CPU and storage devices for applications that are bound by the latency of frequent accesses to the main memory.

Among the challenges of NDP architecture design are the identification of specialized logic that matches the requirements of data-intensive workloads, cost-effective integration of logic and memory, unconventional programming models and the lack of interoperability with caches and virtual memory.

Project Night-King focuses on the provisioning of programmable accelerators in-memory and in-storage for data-intensive workloads that follow the map-reduce programming model, such as sql queries, graph analytics, statistical queries on streams and machine learning

workloads in Apache Spark, Apache Flink, etc. The NDP architecture comprises a template-based design to support generality.

Mappers and Reducers can be programmed on C/C++ and can be synthesized using Vivado High-Level Synthesis. The final bit stream is generated at compile time along with the vendor-provided IPs (memory and flash controllers) and loaded into NDP-augmented servers. A runtime system is needed to dynamically balance the load between the CPUs and near-data accelerators. Using the roofline model for the near-data accelerators' augmented scale-up servers, we estimate a speed-up of four times for Spark MLLib.

In the next phase, we'll be verifying our hypothesis by testing the prototypes on Intel HARP or IBM Power + CAPI-enabled servers, which we see as emulation systems for our research. We will also build the runtime system.

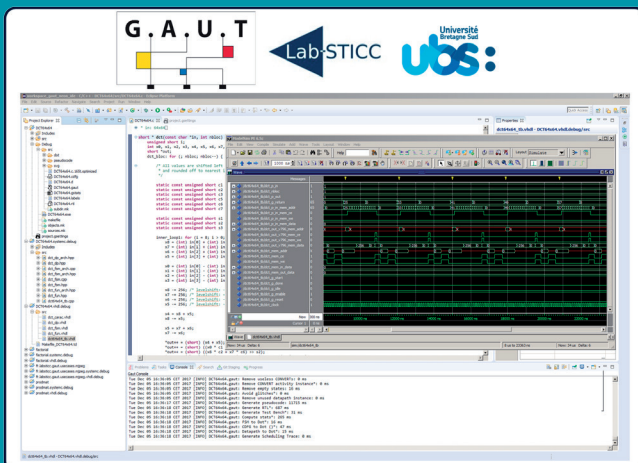
Further information:

Ahsan J. Awan. PhD thesis: 'Performance Characterization and Optimization of In-Memory Data Analytics on a Scale-Up Server'
bit.ly/AhsanJAWan_PhDthesis-2

New release of GAUT, the Eclipse plug-in

Philippe Coussy, Université de Bretagne-Sud

We are pleased to announce that a new version of our high-level synthesis tool, GAUT, is now available. GAUT 3.0 is a free, open-source Eclipse plug-in (CeCILL-B licence). You can access the tool's integrated development environment, lab examples, video tutorials, the full source code and programming environment.



The new version is a suitable platform for researchers and students. This first public release was developed for pedagogical purposes and does not include advanced optimization features, which will be addressed in future versions.

Starting from a C/C++ input description and a set of synthesis options, GAUT 3.0 automatically generates a hardware architecture composed of a controller and a datapath, as well as memory and communication interfaces.

GAUT generates IEEE P1076 compliant RTL level VHDL and SystemC projects. The VHDL files are an input for commercial, off-the-shelf, logical synthesis tools like Vivado from Xilinx or Design Compiler from Synopsys. Windows, Linux and MacOS platforms are supported (32- or 64-bit).

To download GAUT, visit our website: gaut.fr

An overview of GAUT 3.0 is available in this video: bit.ly/GAUT_3-0_video



Jesús Labarta receives Ken Kennedy award

The Association for Computing Machinery (ACM) and IEEE Computer Society (IEEE CS) have awarded HiPEAC member Professor Jesús Labarta, Computer Sciences director at Barcelona Supercomputing Center (BSC), the ACM-IEEE CS Ken Kennedy Award. Professor Labarta was presented with the award at SC17, the annual international supercomputing conference, which took place in Denver, USA, in November.

Selected for his seminal contributions to programming models and performance analysis tools for high-performance computing, Professor Labarta is the first non-American researcher to receive this award.

The Ken Kennedy Award was established in 2009 to recognize substantial contributions to programmability and productivity in computing and significant community service or mentoring contributions. Throughout his career, Professor Labarta has developed tools for scientists and engineers working in parallel programming.

Congratulations on winning this award!

For further information about Jesús Labarta's work, check out our interview with him in HiPEACinfo 52 bit.ly/HiPEACinfo52

Video interviews are also available on the Performance Optimisation and Productivity website and HiPEAC YouTube channel bit.ly/POP_video_JL

Mateo Valero awarded honorary doctorate by CINVESTAV, México

The Mexican Center for Research and Advanced Studies of the National Polytechnic Institute (Cinvestav, according to its initials in Spanish) has awarded an honorary doctorate to HiPEAC co-founder Professor Mateo Valero, Director of Barcelona Supercomputing Center.

The doctorate was awarded in recognition of Professor Valero's excellent work in all aspects of supercomputing development and research, in particular his collaboration in driving forward supercomputing in Mexico. It was presented by Dr. Pablo Rudomín Zevnovaty, Professor Emeritus of the Physiology, Biophysics and Neuroscience Department at Cinvestav, who had nominated Professor Valero for the academic award.

Cinvestav Director José Mustre de León was also at the event, which ended with Professor Valero's keynote speech, titled 'From Classical to Runtime Aware Computer Architectures'.

On behalf of the HiPEAC community, congratulations!



Miguel Angel Aguilar wins RWTH Aachen ICT Young Researcher Award

RWTH Aachen University has honoured the HIPEAC affiliated student Miguel Angel Aguilar with the ICT Young Researcher Award 2017 for his contributions to ICT research at the university. Miguel Angel is a PhD student at the Institute for Communication Technologies and Embedded Systems (ICE), under the supervision of HIPEAC steering committee member Professor Rainer Leupers. The award comes with €3,000 for research-related purposes, and it was presented to Miguel Angel by Professor Stefan Kowalewski, coordinator of the ICT area at RWTH Aachen.

Miguel Angel's research focuses on novel compiler technologies to automatically optimize legacy sequential software for efficient execution on modern heterogeneous multicore systems. He has been developing a parallelization framework that takes as inputs sequential applications and a model of the target embedded multicore system. The framework automatically generates parallel versions of applications, and provides source-level hints to the developers to help them understand the optimization opportunities identified. This framework has been successfully applied to commercial environments. In addition, some of these research results have been deployed in the industry through Silexica GmbH.

Congratulations to Miguel Angel on winning this award!

Dobrý deň, Košice!

Over the past few years, HiPEAC has been visiting different new European Union member states, with the aim of getting more people from these countries involved in the network. On 9 October, HiPEAC coordinator Koen De Bosschere and steering committee member Rainer Leupers visited the Technical University of Košice (TUKE), Slovakia, to present HiPEAC and learn about innovations in the region. We spoke to Prof. Ing. Stanislav Kmet, Rector of the TUKE, to find out more about technology and innovation in the region.

What is the advanced computing field in Slovakia like?

We've been carrying out some fascinating projects in this area. One example is the Aurel supercomputer; among the 500 most powerful in the world with a theoretical performance of 128 teraflops, this computer is available for use by the Slovak Academy of Sciences as well as universities, including the TUKE.

Another significant step was the development of the Slovak Academic Network (SANET). With over 500 members and up to 300,000 connected computers, SANET represents one of the largest data networks in Slovakia. SANET is also connected to the Czech Republic, Austria and Poland, as well as to the pan-European data network GÉANT. Starting in 2015, the TUKE has been establishing advanced cloud services consisting of over 50 servers and based on the 100Gbps optical backbone connecting all major universities as part of SANET.

Last but not least, the National Telepresence Infrastructure project aims to support research, development and technology transfer, connecting over 200 communication rooms at universities and research institutions. Further information can be found on the NTI website: nti.sk.

“At the TUKE, we support technology transfer through expert consultation and access to top-of-the-line research infrastructure”



The TUKE-HiPEAC workshop

How are you promoting innovation in Eastern Slovakia?

The Košice Self-government Region (KSR) is second only to the Bratislava Region in terms of national research potential, as witnessed by the number of entities conducting research, development and innovation activities. There are four universities taking in, on average, 19,000 students per year, which play a central role in further acceleration.

The Slovak Academy of Sciences, with seven research institutes and two internationally recognized research, development and innovation (R+D+I) clusters (the Košice IT Valley and the Cluster for Automation Technologies and Robotics), significantly complements the R+D potential of the KSR. These institutions are furthering innovation through two university science parks, as well as a centre for research in progressive materials and technologies for current and future applications in Košice.

In terms of private enterprise, some of the most dynamic R+D organizations are ZTS-VVÚ, CEIT Biomedical Engineering, Embraco Slovakia and GlobalLogic Slovakia. The R+D+I ecosystem in the region is significantly supported by 14 active industrial parks.

At the TUKE itself, we aim to support technology transfer through expert consultation and access to top-of-the-line research infrastructure. UVP TECHNICOM is our research and transfer centre for innovative applications with the support of knowledge technologies; we aim for this to become a hub at the centre of a regional innovation ecosystem. The centre's pre-incubation services contribute significantly to the creation of new spin-off or start-up companies.



HiPEAC coordinator Koen De Bosschere (left) and steering committee member Rainer Leupers (right) at the workshop

In 2014 the Startup Centre TUKE was formed, the first of its kind in the region, as part of the University Centre for Innovation, Technology Transfer and Intellectual Property Protection (UCITT). The ultimate goal is to help both students and the general population of Košice and Prešov implement their innovative ideas into a commercially usable product or service.

As part of UVP TECHNICOM, the TUKE incubator helps ensure accelerate the formation and development process for small/medium high-tech companies. This is particularly designed for the outputs of relevant research and innovation activities at the TUKE which have been through the pre-incubation process at the Startup Centre.

What is your vision of the future of computing at the TUKE?

The challenges at the TUKE centre around implementing the fast, dynamic, borderless, disruptive side of innovation through technology services to meet the needs of the product manufacturing and service sectors, through dynamic ICT methods and tools (including interactive demos, webinars, challenges, hackathons, etc.). These will complemented in a one-stop shop by:

- Business services (idea incubation, business acceleration, demand-offer matchmaking and brokerage, access to finance) to support start-ups and web entrepreneurs as well as corporates.
- Skill-building services (serious games, role play, participative lessons/webinars, virtual experiments in teaching factories, professional courses) to help users take full advantage of new technologies, providing an operational framework that will stimulate trust, confidence and investment.
- Sector-specific expertise dealing with issues relating to the competitive environment, commitment to research and development, cross-border cooperation, availability of resources, etc.

Tell us about some of your current technology projects.

We're currently working on some exciting projects in key areas. We've prepared a strategic development concept for Industry 4.0, formulated within the design of an extensive multidisciplinary project in close collaboration with major industrial initiatives. Through UCITT, the TUKE is also involved in the Horizon 2020 project MIDIH, or 'Manufacturing Industry Digital Innovation Hubs for Industry 4.0 implementation', an Innovation Action with 21 beneficiaries from 12 EU countries.

There are also several ongoing projects in the area of machine learning and cognitive computing more broadly at the TUKE. These include cloud-based human-robot interaction, cloud-based computational intelligence, intelligent rehabilitation with gaming, big data and intelligence, computer-aided design support for hepatic encephalopathy, intelligent robots – a collaboration with Japan through ERASMUS+, intelligence for ambient assisted living (ERASMUS) and a Microsoft Azure machine learning award for cloud-based ambient assisted living. We have a large number of collaboration activities with universities from Italy and Japan in this area.

How can HiPEAC help you achieve your goals?

In my opinion, it will be mainly through the active involvement of the TUKE researchers in HiPEAC activities, as well as through promoting HiPEAC in the region. We're already using HiPEAC outputs in the form of reports in the educational process at our university.

Further information:
tuke.sk

Machine learning special

Thanks to rapid advances in computing, machine learning has evolved from arresting idea to ubiquitous reality. Taking inspiration from biology to shape new computer architectures and algorithms, it is powering a whole host of innovative applications. We spoke to some of the HiPEAC experts working in this fascinating field, creating brain-inspired computers and dolls which can detect how you're feeling, enabling faster medical diagnosis and smart fitness applications, providing smart solutions for archiving historical documents and much more.

Fast learners

The unstoppable rise of machine learning

VIRTUAL BRAINS AND BRAINIER COMPUTERS



Trying to build a computer that mimics the working of the human brain sounds intriguing, but what can it teach us? 'We really have very little idea how information is represented, stored and recovered in the brain,' explains Professor Steve Furber, University of

Manchester. 'So trying to build machines based upon some of the things we do know about the low-level behaviours of neurons and synapses, how they interconnect, and so on, may help us make progress in understanding the brain while suggesting some new approaches to computer design.'

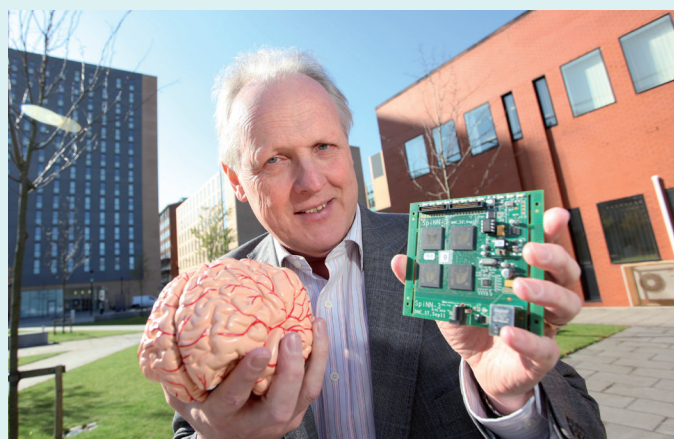
Steve notes that 'the brain, like the computer, is an information processing system. It receives inputs from eyes, ears, touch, etc., and uses these inputs in combination with its stored memories ("experience") to decide how to control its actuators (muscles) to deliver the outcomes that it seeks'. There are crucial differences, though. 'While the power of the computer is based upon its ability to process very simple things very fast, the power of the brain is based upon processing very complex things rather slowly,' says Steve.

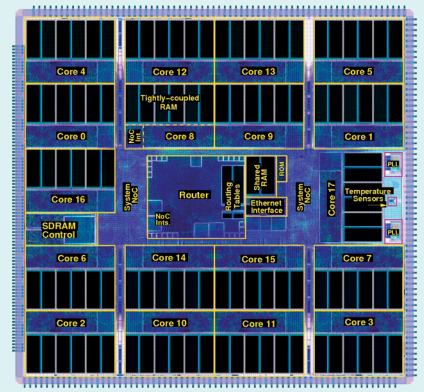
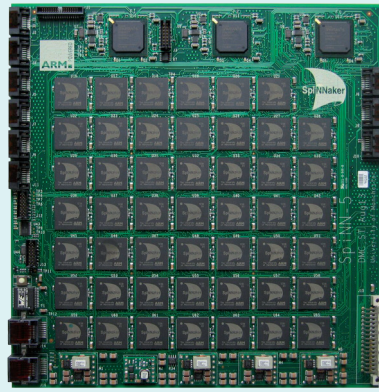
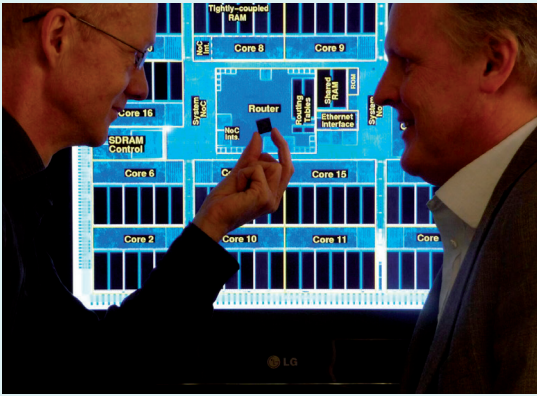
Using the powerful computers available today, we can 'build simulated models of brain regions to test hypotheses about brain function', says Steve. Conversely, 'we can use what we know about the brain to suggest ways to build better computers', such as:

- **increasing parallelism** – 'the brain is massively parallel!'
- **coping with variability and component failure** as technology shrinks: 'The brain is highly tolerant of component failure; how does it do this and what can we learn?'
- **energy efficiency**: 'all predictions as to how powerful a computer would have to be to model the human brain in real time point to exascale or beyond, and we are struggling to reach exascale within a 20MW power budget. Yet the brain uses just 20W!'

SpiNNaker: Spiking Neural Network Architecture

Steve's work in this field includes SpiNNaker, a computing platform which emulates the way the brain neurons fire signals in real time. SpiNNaker is Manchester's contribution to the flagship €1 billion European Human Brain Project, whose goal is to accelerate the fields of neuroscience, computing and brain-related medicine.





‘SpiNNaker has been designed from the silicon upwards to deliver unique brain-modelling capabilities. The most notable example is the way the neuron outputs – “spikes” – are routed around the machine as tiny packets on a packet-switched fabric to enable the machine to emulate the very high connectivity of the biological system, where neurons have thousands of inputs from other neurons,’ Steve explains.

Building such an ambitious machine ‘on academic research budgets’ has been challenging, says Steve: ‘SpiNNaker already has half a million Arm processor cores, and the team is now expanding towards the original goal of a million. Getting the hardware both cheap enough to build and reliable enough to use has been a lot of work.’ However, the challenges don’t stop with the hardware. ‘Because of the unique architecture, the software has had to be developed from the bare metal upwards, and many special algorithms are required to map a problem to the machine, configure all the packet routing resources, and so on,’ he explains.

The results have been worth it: the machine is up and running reliably, with software support in place to make it usable even without detailed knowledge of the machine itself. Plus, as well the big machine in Manchester, there are over 90 smaller SpiNNaker systems in use around the world.

SpiNNaker can support detailed brain models, such as a cortical microcolumn model, delivering the same results as those obtained running the same model on a supercomputer, according to Steve. ‘It has also run artificial networks for constraint satisfaction, where we have shown the ability of a stochastic spiking neural network to solve problems such as Sudoku and map colouring.’ The machine’s full potential is yet to be exploited, though, with all the jobs run so far using only about 1% of the big machine’s capacity.

Neural networks

Giving computers the ability to learn without being explicitly programmed requires a different approach to algorithmic design. ‘An artificial neural network,’ explains Steve, ‘is quite different

from a conventional sequential algorithm’, in that ‘the computation is broken up into many very small parts, and each part is assigned to an individual “neuron” – a small processing unit that receives a number of inputs and uses these to decide what its output should be. Its output is one of the many inputs to one or more neurons in the next layer of the computation’.

The neural network is able to ‘learn’ – or ‘adjust its behaviour in appropriate ways’ – by ‘changing the importance each neuron assigns to each of its many inputs’, says Steve. ‘Like a child learning to ride a bike or play the piano, there will be many mistakes to start off with, but gradually the neural network changes itself to reduce errors and improve the outputs until the final result is an Olympic cyclist or a concert pianist!’

He notes that neural networks have dominated machine learning, with products such as Amazon’s Alexa, Apple’s Siri, and so on, using deep neural networks to understand the user’s speech. Despite this, approaches haven’t changed dramatically since the 1980s. ‘We will have to keep returning to the biology for further inspiration,’ says Steve.

In addition, much of the significant progress achieved is still in niche areas, meaning that the idea that artificial intelligence will soon reach a point where it amplifies its own capabilities far beyond human intelligence is exaggerated, Steve believes. ‘There has been relatively little progress in artificial general intelligence of the sort anticipated by Turing, and no machine has convincingly passed his test. From the little we do know about natural intelligence, it seems to be me to be far more complex than a single parameter that can be amplified by such a process.’

“SpiNNaker has been developed from the silicon upwards to deliver unique brain-modelling capabilities”

HARNESSING HPC FOR MACHINE LEARNING



The idea of neural networks is not new, going back to work by Frank Rosenblatt in the 1950s and 1960s, as Professor José Manuel García Carrasco of the University of Murcia notes, but progress at the time was stymied by a lack of computing power. ‘It was the introduction of custom accelerators that broke the teraflops barrier in 2006, namely NVIDIA

graphics processing units (GPUs) for general-purpose computing that enabled researchers to revisit artificial intelligence and machine learning, as their algorithmic approach is inherently parallel.’

Thanks to this, machine learning based on hardware accelerators has now become a pervasive tool, according to José Manuel, with both industry and the academic community fully embracing machine learning as a major application domain, and numerous hardware solutions being explored by different companies (such as NVIDIA, Intel, Microsoft, IBM and Google) and academic groups.

José Manuel’s research group at the University of Murcia were interested in investigating the potential of high-performance computing for deep learning, or deep neural networks. ‘Architecturally, a deep neural network is modelled using layers of artificial neurons: computational units able to receive inputs, combine them and apply an activation function along with a threshold to determine if messages are passed along. Deep neural networks are characterized by adaptive weights along paths between neurons. These weights can be tuned by an algorithm that learns from observed data to improve the model.’

Platforms for deep learning

Platforms need to be able to meet the requirements of deep learning’s two main steps, explains José Manuel: the learning process and the inference process. ‘During the learning process, the target platform has to crunch a huge amount of data as fast as possible. To do that, the platform has to have as many cores as possible, as well as a high bandwidth memory.’

Traditionally this process relied upon graphics cards from NVIDIA, but other options are now available. ‘Intel entered the competition around 2013 with its Xeon Phi line, and last year Google introduced its Tensor Processing Unit, a hardware accelerator designed for running the TensorFlow framework.’ José Manuel notes that other platforms, such as application-specific integrated circuit (ASIC) or field-programmable gate array (FPGA) designs, could also be appropriate.

With regard to the inference process, the two most important constraints are power use and real time processing, notes José Manuel, as much of the inference process is carried out in embedded devices. ‘Some vendors offer a scaled-down version of the same architecture, but others, like Intel, have brought out new ones, like the Movidius neural stick. Again, ASICs and FPGAs could have an important role here.’

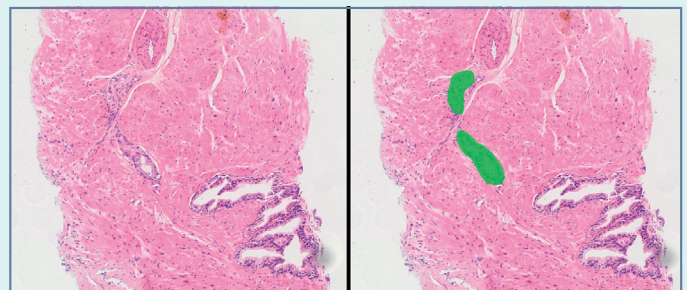
Taking advantage of the arrival of the Intel Xeon Phi, José Manuel’s group started coding a deep neural network from scratch using C++, with the aim of gaining a profound understanding of the main features of deep neural networks. ‘Through this we tackled the parallelization of deep neural networks for Intel manycore architecture, and learned a lot about vectorization, memory usage, scaling to use all system nodes, etc. With only slight changes in the code, we have tested our implementation for the two Phi generations (KNC and KNL) as well as Xeon line processors.’

Business and healthcare applications

José Manuel’s research group is currently testing several deep learning frameworks (such as TensorFlow, Caffe and Theano) for real problems, including:

Business: standardizing company inventory data, which is often stored in different formats and uses different nomenclature to identify the same thing, into a master inventory, doing so very quickly.

Healthcare: in collaboration with the Reina Sofía Hospital in Murcia, the group is applying deep learning to improve the objectivity and efficiency of histopathologic slide analysis. As a case study, they are testing prostate cancer identification in biopsy specimens.



Deep learning can be used to improve histopathologic slide analysis: a) Original image; b) after the inference process, the image is labelled as an image containing cancer, highlighting in green the likely tumorous areas

Their research methodology consists of adjusting the many parameters of a deep learning network, with the aim of obtaining the highest accuracy possible. ‘The higher the amount of data, the higher the accuracy,’ says José Manuel. ‘To keep learning times tractable, you need to figure out which parameters will be best for the specific problem.’ He explains that these parameters range from the type of architecture, activation and cost functions, the number of layers and number of ‘neurons’ in each layer, to other minor parameters that can have a major impact on the learning process, such as how to initialize the weights and biases, the learning rate, the size of batches, etc.

The group will continue to work on the optimization process, so that the deep neural network can solve the problem in hand with the highest performance. Other challenges for the future include how to use sparse multi-layer perceptron models, moving to low-precision arithmetic and using concepts from approximate computing, and finally scaling out to tens or hundreds of thousands of cores.

‘The “Trends in Machine Learning” workshop at ISCA, the International Symposium on Computer Architecture, offered a very interesting overview of this area,’ says José Manuel. ‘I hope to see a similar workshop at the HiPEAC conference in the next few years.’

MACHINE LEARNING GOES MOBILE



Valentin Radu, Research Associate at the University of Edinburgh, was an early adopter of personal sensing with mobile devices, which was once ‘limited to eccentric enthusiasts’: ‘I remember the mixed reactions when I told people that I logged WiFi access points on my smartphone to track

my journeys and accelerometer to monitor activity.’ Now, as he points out, personal sensing is commonplace, with commercial offerings like the Fitbit and Apple Health being decidedly mainstream. ‘These offer just a glimpse into the emerging opportunities for building smarter digital assistants and shifting the direction of healthcare from treatment to prevention, by continuously monitoring everyday activities and sensing contexts.’

Context detection and activity recognition on mobile devices pose specific problems, however. ‘Applications running on battery-powered devices are designed around a limited energy budget to supply sensing, compute and user interaction. The cost of network communication is not negligible, either.’ As these devices tap into personal sensor data, privacy is also an issue, says Valentin, as ‘uploading raw data to the cloud exposes the user to unnecessary risks avoidable only by performing computations partly or entirely on the mobile device’.

A further challenge is that ‘no two users are the same,’ explains Valentin, ‘so algorithms must be robust enough to handle various mobility patterns across users, making this extremely difficult to model with traditional signal processing methods’.

From HiPEAC-supported research to start-up

So how can machine learning help? ‘By building robust models directly from data, which can generalize beyond just observations at hand. Machine learning models can be trained on servers at scale and deployed to run detections on mobile devices, with minimal battery impact,’ says Valentin. Deep learning is particularly promising: ‘In our recent article “Multimodal Deep Learning for Activity and Context Recognition” in *Interactive, Mobile, Wearable and Ubiquitous Technologies*, we show that deep learning achieves consistently better performance across a multitude of detection tasks, while staying within a manageable energy budget for modern smartphones and wearable devices.’

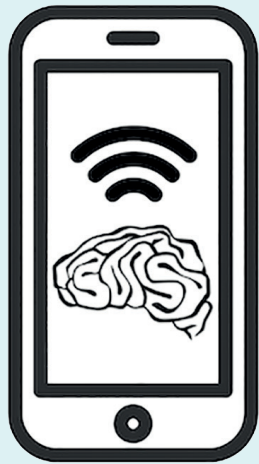
Valentin’s research in deep learning began during a research stay at the Mobile Systems Group at the University of Cambridge, during which he worked with partners at Bell Labs – a visit funded by a HiPEAC Collaboration Grant. The exceptional results he witnessed encouraged him to explore context detection for the home automation market further, eventually leading to the



Machine learning special

creation of a start-up, DeepContext. 'DeepContext delivers context understanding to smart-home devices and digital assistant technologies (Amazon Alexa and Google Home) to improve user interaction with home appliances and the relevance of information received by aligning with users' contexts and activities,' he explains.

DeepContext



As for the future, Valentin sees machine learning as being at the core of an emerging technological revolution. 'We will see more large-scale automation and optimization of processes impacting our everyday lives, and mobile computing is no exception. We will see better and more energy-efficient applications with algorithms constructed on or optimized using machine learning. Hardware and support libraries will also be affected by machine learning, with designs being improved and their execution time accelerated,' he says.

Valentin is also participating in the Bonseyes project, which aims to transform artificial intelligence (AI) development from a cloud-centric model to an edge device-centric model through a marketplace and an open AI platform. 'We're looking at how to accelerate the execution of deep neural networks on embedded and resource-constrained devices. There is some really exciting work coming out of this project, and I'm hopeful that these advances will impact not just mobile computing but high-performance computing (HPC) more generally.'

Further information:

deepcontext.tech

bonseyes.com

Bonseyes has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 732204 and the Swiss State Secretariat for Education, Research and Innovation (SERI) under contract number 16.0159

TAMING BIG DATA WITH MACHINE LEARNING



It's hard to ignore the revolution powered by big data, from providing potent research resources to merrily disrupting industry sectors with a wealth of new business models. Two examples, Professor Håkan Grahn of Blekinge Institute of Technology (BTH) notes, are recommender systems for online

purchasing and advanced data analytics for self-driving cars, a topic explored in depth in *HiPEACinfo* 52.

Such is the volume of data that 'there is no practical possibility of using it without computers', says Håkan. 'Machine learning, described in Tom Mitchell's 1997 book *Machine Learning* as "the study of computer algorithms that improve automatically through experience", is a powerful approach to extract knowledge from data, by building models to solve various classification and regression tasks,' he adds.

Håkan highlights three main challenges when creating machine learning systems to extract value from big data:

Scalability: how to design algorithms that scale well when we increase the data size as well as the number of nodes in the execution platform.

Limited execution resources, for example computational capabilities, memory constraints and power/energy consumption.

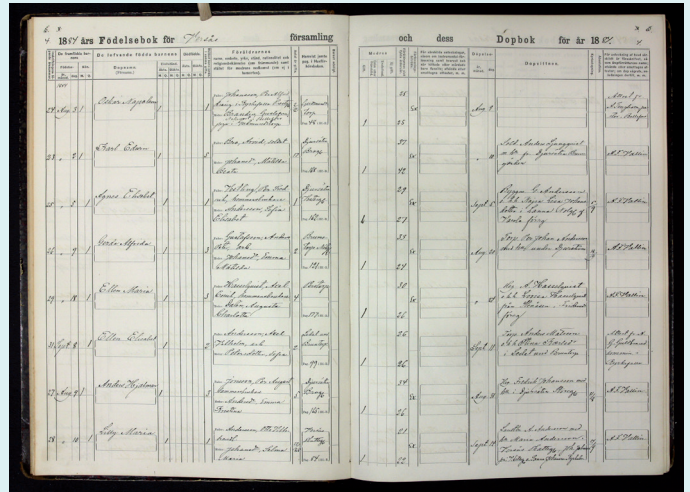
Data stream mining: in many applications, data arrives (or is generated) in real time as a stream, so the algorithm has only a limited time to make a decision and in most cases only has one chance to look at the data before it is gone.

Håkan's group at BTH researches the interaction between machine learning/big data analytics and computer system engineering. 'Our focus is on how to develop scalable, resource-efficient solutions, which is of particular interest for embedded, battery-powered devices.' He points out that, with the growth of the internet of things and the subsequent deployment of numerous devices, many of which collect data, there will be a requirement for a large amount of data analysis on the devices themselves. 'As a result, the algorithms must be very resource efficient. Our studies have shown that we can reduce the energy consumption in data stream mining applications by up to 90% in some cases, with only marginal effects on accuracy,' says Håkan.

To take this exciting area of research forward, Håkan is leading the BigData@BTH project, or ‘Scalable resource-efficient systems for big data analytics’, to give its official title, running from 2014-2020. Financed by the Knowledge Foundation in Sweden, the project includes nine industrial partners. ‘One case study that we’ve done in conjunction with a company partner is the development of an automatic system based on deep learning for classifying and sorting incoming customer mail. Another example is with Arkiv Digital, who have over 60 million historical documents, where we work on image quality enhancement and content analysis based on pattern recognition, for example.’

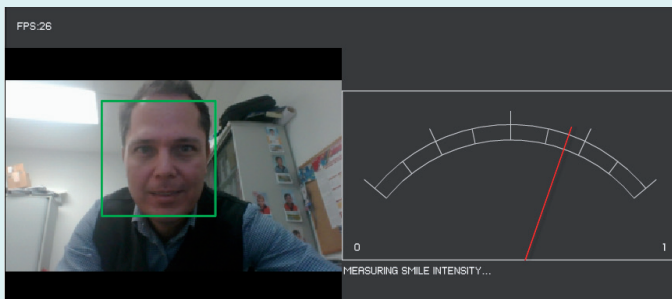
Further information:

bth.se/bigdata



Machine learning can be used to enhance image quality of historical documents

HERE'S LOOKING AT YOU, KID



Ever get the feeling you're being watched? In the future, many inanimate objects might be checking you out, according to Professor Oscar Deniz Suarez, University of Castilla-La Mancha. ‘Thanks to our brains, our eyes are arguably our richest sensor. Likewise, the internet of things paradigm could reach its full potential if we had “eyes everywhere”.’ Just a few of the things he foresees as having ‘eyes’ over the next 10 years are mini robots, headsets, cars, forests, lamps and streets. While surveillance is an obvious application, a few others he cites are ‘intelligent toys; drones equipped with vision which can detect and track people, cattle, objects, or measure crowds, for example; headsets that augment our vision; etc.’

While still a long way from human capabilities, Oscar points out that computer vision has progressed enormously over the last few years; previously confined to restricted conditions, such as quality control in manufacturing plants, it is now ‘an increasingly horizontal capability that can be used in many novel applications’.

‘The main challenge in this field is computing power,’ says Oscar. ‘Mobile and efficient high-performance computing (HPC) is what facilitates the deployment of vision on a larger scale.’ The

EU-funded project he coordinates, Eyes of Things (EoT), aims to overcome some of the roadblocks to ubiquitous machine vision, which also include power consumption and cost. As the project identifies, the only practical solution in many cases is cloud services, which pose problems with bandwidth (particularly for images and video) and privacy concerns, as the data involved (images) is sensitive.

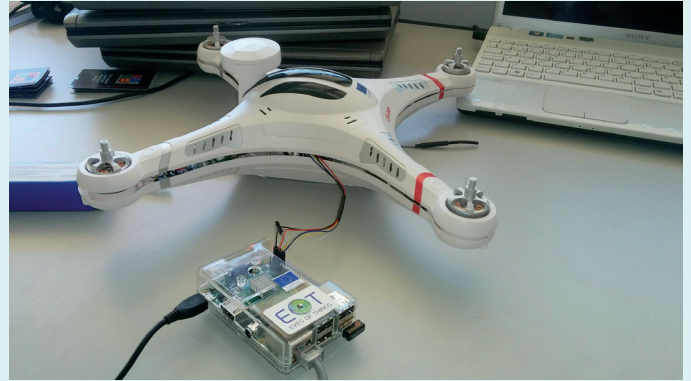
The project's solution is an optimized, embedded core vision platform, which allows the user to develop mobile artificial vision applications with minimal power use. Based on the Intel Movidius Myriad 2 system-on-chip (SoC), which was specifically designed for intensive computer vision operations, the platform enables deep learning while keeping power use low. In addition, explains Oscar, the software libraries and protocols implemented for the platform have also been carefully selected, ported and optimized for this sole purpose.

Two convolutional neural network frameworks have been implemented for the EoT platform:

1. `tiny_dnn`, a well-known open-source library in C++ which includes a deep learning inference engine optimized for limited computational resources.
2. the Fathom framework, a proprietary library developed by Intel Movidius to run convolutional neural networks targeting the Myriad 2 SoC hardware.

These have been tested using a digit-recognition network and an emotion recognition network provided by partner nVISO.

Machine learning special



Researchers have embedded the EoT board in a doll, which can be used to recognize a child's emotions from facial expressions and give feedback through a speaker incorporated in the doll; alternatively, the emotions can be registered to provide information for therapy. Oscar points out that local processing is a fundamental achievement here, meaning that the privacy issues associated with sending pictures via the internet can be avoided. 'In EoT, each captured image is stored in (volatile) memory, processed and deleted. The 12-layer emotion recognition network was trained on 6258 images and outputs one of seven facial expressions.' The energy efficiency is also impressive: results show the dolls can perform emotion recognition with tolerable latency (244ms) for up to 13 hours continuously on a 4000mAh battery.

So is there anything Oscar thinks shouldn't have eyes? 'Images of people must be protected, and all existing regulations relating to privacy and surveillance apply. The technological progress we are witnessing will only bring closer scrutiny and more work on the regulatory side. The only novel situation I can think of is that of wearable cameras, such as the recent Google Clips. But even in that case images are processed inside the device, and only metadata, if any, is streamed out.'

Eyes of Things has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 643924



This issue brings news of EU-funded research on software support for heterogeneous low-energy computing, efficient cloud resource management and how collaboration can help bring heterogeneity into the mainstream.

Innovation Europe

PLUGGING THE SOFTWARE-SUPPORT GAP FOR LOW-ENERGY COMPUTING



Due to fundamental limitations of scaling at the atomic scale, coupled with heat density problems of packing an ever-increasing number of transistors in a unit area, Moore's Law – the observation that the number of transistors in a dense integrated circuit doubles approximately every two years – has slowed down. Heterogeneity aims to solve the problems associated with the end of Moore's Law by incorporating more specialized compute units in the system hardware and by utilizing the most efficient compute unit for each computation. However, while software-stack support for heterogeneity is relatively well developed for performance, for power- and energy-efficient computing it is severely lacking.

This is where the European Union-funded project LEGaTO – or Low Energy Toolset for Heterogeneous Computing – comes in. According to LEGaTO coordinators and HiPEAC members Osman Ünsal and Adrián Cristal (Barcelona Supercomputing Center), 'in the LEGaTO project we will leverage task-based programming models to provide a software ecosystem for made-in-Europe heterogeneous hardware composed of central and graphics processing units (CPUs and GPUs), field-programmable gate arrays (FPGAs) and dataflow engines. Our aim is one order of magnitude energy savings from the edge to the converged cloud/high-performance computing'.

LEGaTO aims to deliver the following results:

- one order of magnitude improvement in energy-efficiency for heterogeneous hardware through the use of the energy-optimized programming model and runtime
- reduction in size of the trusted computing base by at least an order of magnitude
- fivefold decrease in mean time to failure through energy-efficient software-based fault tolerance.

- fivefold increase in FPGA designer productivity through the design of novel features for hardware design using dataflow languages

The toolset will be put to the test in three use cases:

- **Healthcare:** as well as demonstrating a decrease in energy consumption in the healthcare sector, LEGaTO will also show that the toolset will increase healthcare application resilience and security – both critical requirements in this area.
- **Internet of things (IoT), smart homes and smart cities:** this application will demonstrate the ease of programming and energy saving possible thanks to the LEGaTO toolset. Sensor information and actuator instructions will be received and sent via the secure IoT gateway to be developed.
- **Machine learning:** here, the project will demonstrate how to improve energy efficiency by employing accelerators and tuning the accuracy of computations at runtime. The use case will explore object detection using convolutional neural networks (CNNs) for automated driving systems and CNN- and long short-term memory (LSTM)-based methods for realistic rendering of graphics for gaming and multi-camera systems.

NAME: LEGaTO: Low Energy Toolset for Heterogeneous Computing

START/END DATE: 01/12/2017 – 30/11/2020

KEY THEMES: Heterogeneous computing, low-energy, software toolset, OmpSs

START/END DATE: Spain: Barcelona Supercomputing Center; Germany: Universität Bielefeld, Technische Universität Dresden, Christmann Informationstechnik + Medien GmbH & Co. KG, Helmholtz-Zentrum für Infektionsforschung GmbH; Switzerland: Université de Neuchâtel; Sweden: Chalmers Tekniska Högskola AB, Data Intelligence Sweden AB; Israel: TECHNION - Israel Institute of Technology; UK: Maxeler Technologies Limited.

BUDGET: €5.51M

WEBSITE: legato-project.eu

SILVER LINING NEW ARCHITECTURE FOR EFFICIENT MANAGEMENT OF CLOUD COMPUTING RESOURCES

Georgios Goumas, *Institute of Communication and Computer Systems*



Despite their proliferation as a dominant computing paradigm, cloud computing systems lack effective mechanisms to manage their vast resources efficiently. Resources are stranded and fragmented, ultimately limiting cloud applicability only to classes of applications that pose moderate resource demands.

Enter ACTiCLOUD, a three-year Horizon 2020 project creating a novel cloud architecture that breaks existing scale-up and share-nothing barriers and enables the holistic management of physical resources, at both the local and distributed cloud site levels.

ACTiCLOUD responds to four typical scenarios of resource inefficiency in state-of-the-art cloud offerings, as shown in Figure 1 below.

Scenario 1 (Figure 1a): The standard practice of cloud service providers is to plan conservatively and reserve system resources for the infrequent cases of peak traffic. This strategy clearly leaves large amounts of resources unused.

ACTiCLOUD solution: to improve resource efficiency and utilization through effective consolidation.

Scenario 2 (Figure 1b): Current server architectures are unable to serve resource requests that exceed the levels provided by single servers. This is a critical shortcoming of state-of-the-art cloud offerings, prohibiting resource-hungry applications from enjoying cloud benefits.

ACTiCLOUD response: particular focus on applications that rely on large in-memory databases with non-conventional main memory demands.

Scenario 3 (Figure 1c): Despite resources being available, the fact that they are scattered around means cloud sites are unable to host a new service.

ACTiCLOUD solution: to identify resource fragmentation before devising and applying efficient migration and co-scheduling policies.

Scenario 4 (Figure 1d): Problems arise due to interference between applications that compete for shared resources, when these are misplaced within the cloud platform.

ACTiCLOUD solution: to identify and mitigate resource interference through appropriate migration and co-location actions.

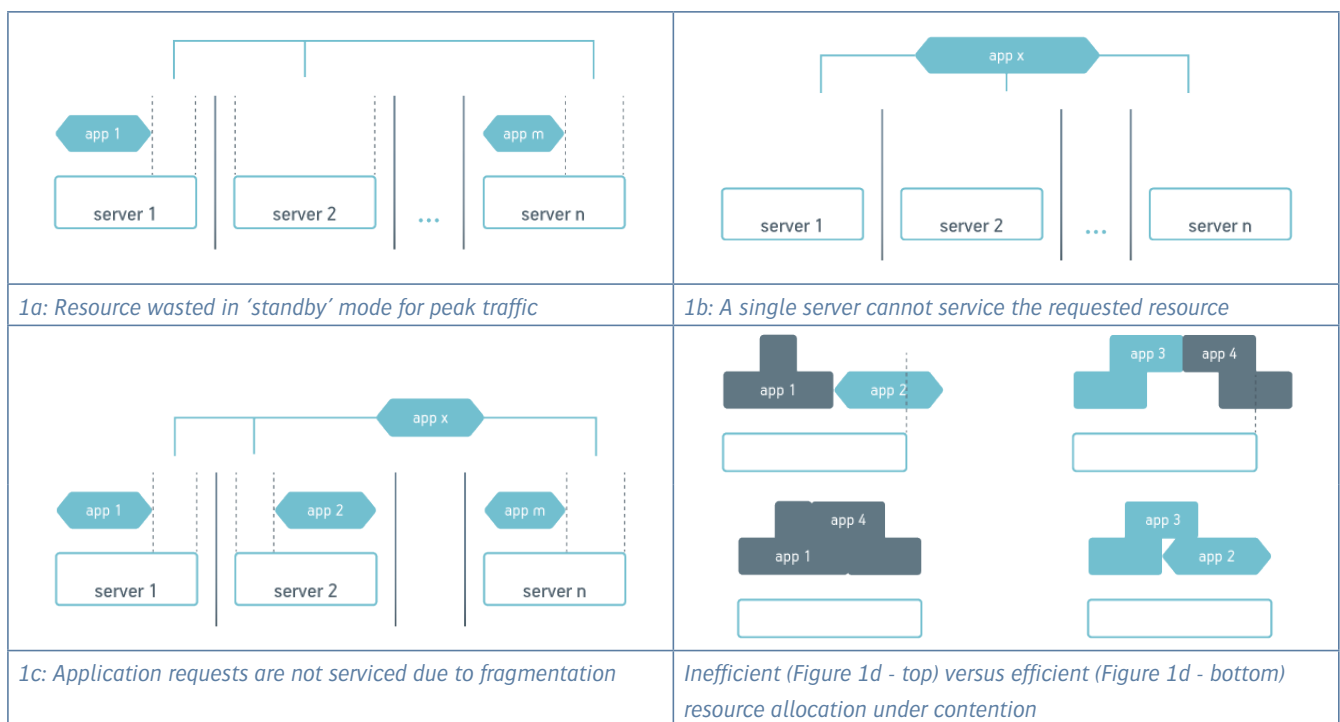


Figure 1: Scenarios to which ACTiCLOUD responds

To overcome these challenges, ACTiCLOUD innovates holistically across the cloud architecture, building on top of novel hardware support for true disaggregation and fluidity of resources. The project advances virtualization technology to support virtual machine execution with minimal overheads, effective pooling of cloud resources at the rack level, and advanced mechanisms for resource monitoring and management.

ACTiCLOUD utilizes this substrate and extends the mechanisms and policies of state-of-the-art cloud managers in order to break the two critical barriers that hinder fluidity of cloud resources today: the server barrier and the datacentre barrier. In this way, ACTiCLOUD-enabled systems allocate resources efficiently, avoid interference, and establish a close collaboration between geographically distributed cloud sites (see Figure 2, below). Finally, ACTiCLOUD extends system software and language runtimes to offer the abundance of cloud resources to applications that need them, such as business intelligence applications that rely heavily on fast, in-memory database support.

The project builds on cutting-edge European technologies for cloud servers brought into the project by Numascale and Kaleao, and extends OnApp's Microvisor, an innovative hypervisor for virtualizing resources at the rack-scale, developed during the EU-funded EUROSERVER project. In addition,

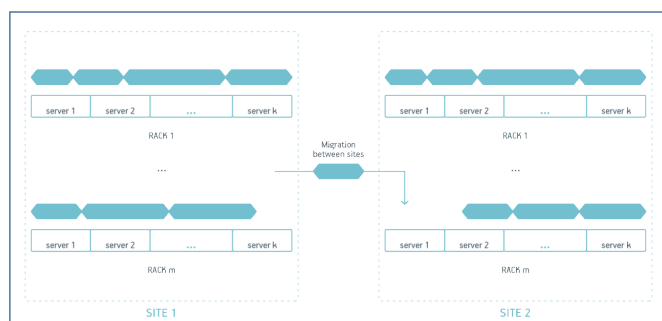


Figure 2

ACTiCLOUD brings together highly acclaimed academic institutions to address key OpenStack and JVM research challenges, and extend their capabilities. Finally, ACTiCLOUD enables the efficient execution of MonetDB, the column-store database pioneer, and Neo4j, the world-leader in graph databases, to provide novel ACTiCLOUD-enabled database-as-a-service (DBaaS) products, in addition to supporting traditional cloud applications through infrastructure-as-a-service (IaaS) offerings (see Figure 3, below).

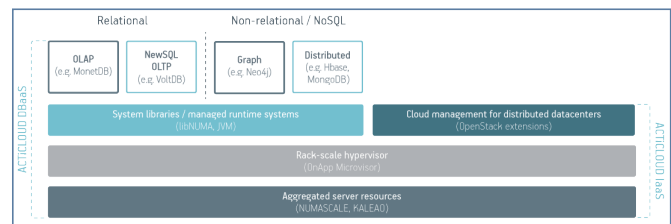


Figure 3: ACTiCLOUD architecture and services

PROJECT: ACTivating resource efficiency and large databases in the CLOUD (ACTiCLOUD)

START/END DATE: 01/01/2017 - 31/12/2019

KEY THEMES: cloud computing, resource efficiency

PARTNERS: Greece: Institute of Communication and Computer Systems – ICCS; Norway: Numascale; UK: Kaleao, OnApp, University of Manchester; Netherlands: MonetDB; Sweden: Neo4j, UMEA University.

BUDGET: €4.73M

WEBSITE: acticloud.eu

The ACTiCLOUD project has received funding from the European Union's Horizon 2020 programme under grant agreement no. 732366.

Find out more about ACTiCLOUD at EnESCE, the Workshop on Energy-efficient Servers for Cloud and Edge Computing, at the HiPEAC conference on 24 January 2018

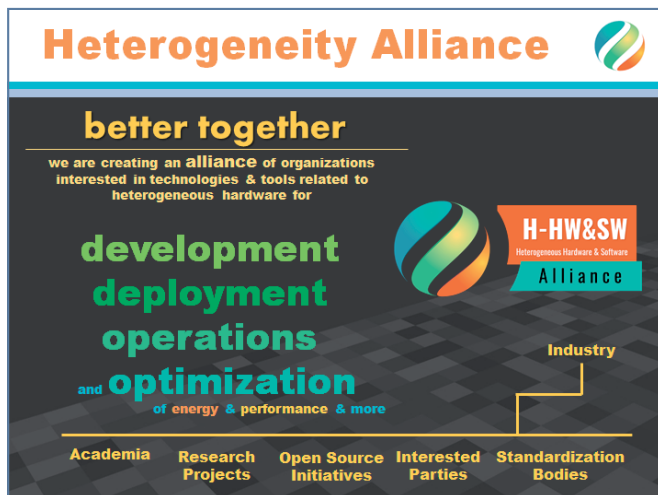
EXPLOITING HETEROGENEITY THROUGH COLLABORATION

Clara Pezuela, Head of IT Market, Research and Innovation Group, Atos

The Heterogeneity Alliance, launched by the team behind the Horizon 2020 TANGO (Transparent heterogeneous hardware Architecture deployment for eNergy Gain in Operation) project, represents a community united by a desire to fully exploit heterogeneous hardware. The alliance aims to support collaborative research, as well as integrating and promoting results produced by business and academic members.

As indicated in our report for HiPEACinfo 51, heterogeneous hardware is increasingly disrupting the IT landscape, bringing with it a need for appropriate software and programming methodologies. As in the past with traditional computing, the time has come for high-level heterogeneous programming abstractions. TANGO is providing a practical response to this challenge with its software toolbox; however, in order to avoid duplication of work and fragmented approaches, the project decided that collaboration was the way forward.





The Heterogeneity Alliance: better together

The Heterogeneity Alliance is formed of different organizations managed by a governance structure that pursues a common objective: to influence the heterogeneity market. It was initially launched as a formal association (non-profit and non-legal) by the TANGO project, but has been rapidly expanded with a series of related EU projects (RAPID, SHARCS, P-SOCRATES, ECOSCALE, HERCULES, VINEYARD) and several independent organizations. You can see the full list of members on our website: heterogeneityalliance.eu/alliance-members

One of the Alliance's main goals is to involve anyone interested in these technology areas. Bringing together a range of expertise, the objective is to found a common, open-source, extendable set of technologies and tools around the development of heterogeneous hardware and software. Based on technologies created by Alliance members, the aim is for these to influence the market and become attractive, easy to use and broader in scope and value, making them viable for mass adoption.

Reference architecture and online catalogue

In addition to our promotional activities, the Alliance is currently working on the creation of a reference architecture. We are also working on an online catalogue of tools and technologies, in line with our vision and with the reference architecture, to support the community developing for heterogeneous architectures.

The Alliance architecture focuses on all phases of the development lifecycle for heterogeneous hardware, from design time to enhanced execution, parallel programming and optimized runtime. We have also considered a number of factors, such as energy, performance, real time, data locality and security. This will enable new ways of developing and executing next-generation applications.

The reference architecture can be downloaded from the heterogeneity website (see below), while the catalogue is already available and being continuously populated.

Working with HiPEAC

The objectives of the Heterogeneity Alliance are closely in line with HiPEAC's mission to steer and increase European research in high-performance and embedded computing systems, while promoting collaboration between different stakeholders in this field. The Alliance provides a way to create marketable results from European research. Conversely, HiPEAC's networking events and communication channels allow the Alliance to reach institutions across disciplines, as well as helping to bridge the gap between academia and industry, and between European and non-European institutions.

If you are working at a research centre, an academic institution or a company that is involved in any part of the development lifecycle, being a part of the Alliance allows you to help influence the heterogeneity market. Its open innovation process also provides the opportunity to engage with other potential competitors, partners or customers. Meanwhile, if your organization provides solutions for key markets such as high-performance computing, parallel programming, the internet of things and big data, you could benefit from the state-of-the-art tools and technologies provided by the Alliance's online catalogue and reference architecture.

If you want to build something that matters, join the Alliance and start benefiting from heterogeneity now. Contact us: heterogeneityalliance.eu/contact

You can download the Heterogeneity Alliance reference architecture and navigate through online catalogue from the website: heterogeneityalliance.eu/resources

The Heterogeneity Alliance keynote speech and session 'Heterogeneity Alliance: Better Together' takes place at the HiPEAC conference in Manchester on 22 January.

TANGO is funded by the European Commission under the Horizon 2020 Framework Programme for Research and Innovation under grant agreement no. 68758.



When Amanieu d'Antras developed a technology which could provide an alternative to Arm 32-bit hardware support without significantly impacting performance, he and his PhD supervisor, Professor Mikel Luján of the University of Manchester, realized they were onto something good. This technology has since been licensed and a start-up launched off the back of it. Here Mikel and Amanieu explain how it they made it happen.

Armed for success: Amanieu Systems

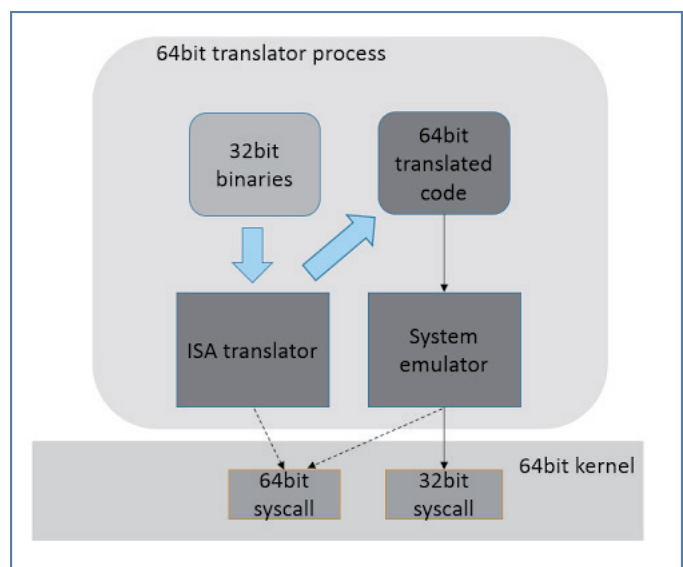
Remove AArch32 hardware support while maintaining performance

'Current computer architectures – Arm, MIPS, PowerPC, SPARC, x86 – have evolved from a 32-bit to a 64-bit architecture,' explains Professor Mikel Luján, University of Manchester. 'Computer architects often consider whether it would be possible to eliminate hardware support for a subset of the instruction set in order to reduce hardware complexity, which could improve performance, reduce power usage and accelerate processor development.'

'The latest Arm processors (Armv8) introduced a new 64-bit execution mode and instruction set, also known as AArch64,' says Mikel. 'Some Armv8 processors are capable of running existing 32-bit Arm applications directly in AArch32 mode, but maintaining this support comes at a significant cost in hardware complexity, power usage and development time.' Unsurprisingly, then, the trend is for the support to be withdrawn: Cavium does not include hardware support for AArch32 in their ThunderX processors, nor does Qualcomm for their Centriq processors.

Finding a solution which would avoid the need for 32-bit hardware support for Armv8 architecture was therefore a priority. Over the course of his PhD studies at the University of Manchester, Mikel's student Amanieu d'Antras undertook research which resulted in a demonstration that the performance offered by a dynamic binary translation was similar to still having the hardware support for running AArch32. 'In other words,' says Amanieu, 'the research developed the main techniques needed so that users of future Arm processors with hardware support for AArch64 alone would not notice a difference in performance, even when executing applications for AArch32.'

This research was highly successful, with one paper scooping up a Distinguished Paper Award at PLDI 2017, the ACM SIGPLAN Conference on Programming Language Design and Implementation. The impact went beyond the academic world, however: the product, the Tango Binary Translator, has been licensed by



fabless semiconductor company Spreadtrum Communications and is now being commercialized by a new start-up, Amanieu Systems.

Bringing the technology to market

'When the first research paper was published in ACM TACO (*Transactions on Architecture and Code Optimization*), we were approached by a few companies who wanted to know if the technology was real or if it was just able to run a few benchmarks,' says Amanieu. 'Seeing such interest, and with one licensee of the technology already in place, we decided it was the right time to establish Amanieu Systems and make the first product available for evaluation.'

The market for this technology is the Arm ecosystem: the companies designing and producing Arm processors. 'We are seeing great interest in our product both from Arm systems focusing on smartphones running Android and Arm systems targeting data centres,' adds Mikel.



Technology transfer

What advice would they give other researchers thinking about creating a start-up based on their technology? ‘Every market is different, so it’s difficult to give specific advice,’ says Mikel. ‘One thing I would say is that once you have an impressive, novel technology, it’s really important to talk with potential clients and find out more about their interests. Learn from these conversations to get a better idea of how to present your technology and how it could evolve to improve aspects you hadn’t originally thought of.’

Networks and professional development in this area are of great help, too. In addition to receiving training on commercialization from the UK’s Royal Society (royalsociety.org) offered to Royal Society University Fellows, Mikel mentions interactions in the HiPEAC network and his participation in the EuroLab-4-HPC project (eurolab4hpc.eu) as being particularly useful.

Start-ups made in Europe

If the start-up scene in Europe isn’t as dynamic as in other parts of the world, it’s more to do with a shortage of private finance than a lack of talent, argues Mikel. ‘Europe is not short of bright and capable people, nor great ideas. What does need to change is for the venture capital available to increase and, with it, the amount of risk venture capitalists are willing to take. When you invest and provide long-term trust in people with good ideas without too much interference or bureaucracy, start-ups can grow and thrive.’

He sees the process of launching a start-up as valuable in and of itself. ‘Even when they don’t succeed in their first incarnation, a whole team of people would have acquired very important skills, and a set of second- and third-generation start-ups tend to pop up, as long as we don’t stigmatize those who have made unsuccessful attempts,’ he adds.

As for HiPEAC’s role in facilitating technology transfer, Mikel notes that it is ‘already a magnificent forum for researchers in

Europe, and we are seeing increasing participation by companies. Being able to share commercial success stories and celebrating them is a great beginning’. He sees the HiPEAC summer school, ACACES – ‘an amazing platform for bringing together early career and experienced researchers for training’ – as being able to play a key part here: ‘HiPEAC could build on the ACACES platform to share first-person success stories of how research results have generated market-ready products and services.’

With the first-ever technology transfer track being planned for next year’s ACACES, this could be the ideal opportunity to do just that. Watch this space for further information.

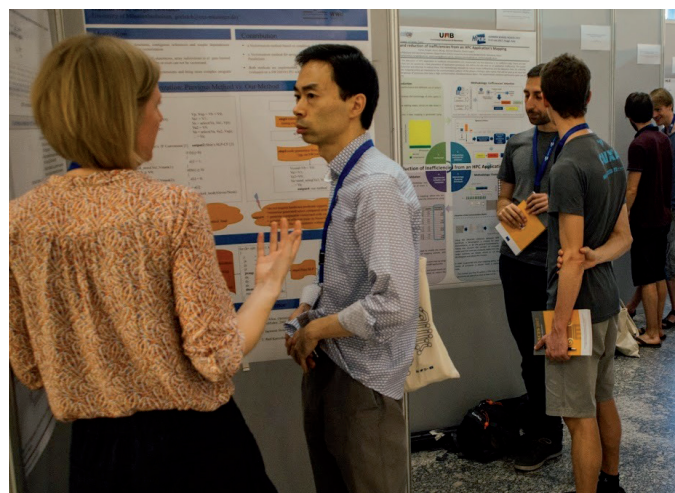
amanieusystems.com

Further reading:

Amanieu d’Antras, Cosmin Gorgovan, Jim Garside, and Mikel Luján. 2016. ‘Optimizing Indirect Branches in Dynamic Binary Translators.’ *ACM Transactions on Architecture and Code Optimization (ACM TACO)* 13, 1, Article 7 (April 2016)
[doi.org/ 10.1145/2866573](https://doi.org/10.1145/2866573)

Amanieu D’Antras, Cosmin Gorgovan, Jim Garside, Mikel Luján. 2017. ‘Low overhead dynamic binary translation on ARM.’ *Proceedings of the ACM SIGPLAN Conference on Programming Language Design and Implementation 2017 (PLDI 17)*: 333-346
doi.org/10.1145/3062341.3062371. Distinguished Paper Award.

Amanieu d’Antras, Cosmin Gorgovan, Jim Garside, John Goodacre, and Mikel Luján. 2017. ‘HyperMAMBO-X64: Using Virtualization to Support High-Performance Transparent Binary Translation.’ *Proceedings of the 13th ACM SIGPLAN/SIGOPS International Conference on Virtual Execution Environments (VEE ’17)*, 228-241.
doi.org/10.1145/3050748.3050756



The HiPEAC ACACES summer school could be an ideal seedbed for tech transfer

In *HiPEACinfo 50*, we reported on how the University of St Andrews was poised to commercialize results of European Commission-funded projects through ParaFormance. Here, Chief Technology Officer Chris Brown sets out how the new company has been steadily building up its portfolio.



More ParaFormance for your money

ParaFormance Technologies Ltd. was formed in October 2017 as a spinout from the University of St Andrews using €537,000 of Scottish Enterprise innovation funding to exploit results from successful European Union (EU) projects. Benefiting from constructive feedback from the HiPEAC community, the ParaFormance tools help developers of all skill levels build safer, faster code for multicore systems. ParaFormance developers produce multicore software quicker, enabling them to meet customer needs sooner, reducing bugs, and improving company profitability.

ParaFormance is now open for business, offering a free 30-day full evaluation licence. The tools are available to download from our website or from the Visual Studio and Eclipse marketplaces.

Opportunities in high performance computing (HPC)

ParaFormance sees the HPC space as an exciting opportunity for its products. As an example, we're involved in a new collaboration with Slovenian HPC centre, Arctur. The Arctur high-performance computing (HPC) Challenge is a joint initiative where applicants can win up to €350,000 of HPC resources, including subsidized access to state-of-the-art HPC and cloud infrastructures. ParaFormance is a perfect fit for the HPC challenge, allowing software developers to quickly and easily scale up their applications to exploit HPC resources. Find out more at paraformance.com/arctur-hpc-challenge.html.

Tech transfer tip #1: Network, network, network

Go to big events like the annual Supercomputing conference in the USA. SC17 brought us into contact with reputable HPC vendors and technologists, including Oak Ridge Labs, ICHEC, DST and the Slovenian HPC centre, Arctur.

New ParaFormance Features

We've been working hard to integrate some exciting new features to showcase at the HiPEAC conference in January 2018.

OpenMP Support

Our new OpenMP support enables developers to use ParaFormance to parallelize their applications quickly and easily using our sophisticated and advanced refactoring technology. Our safety-checking tool ensures that parallelization won't introduce any bugs into your application.

Performance prediction

Have you ever wondered what the performance increase will be for your application before you've begun to parallelize it? Our unique new feature predicts speedups of the application before they are parallelized, using advanced performance models of the parallel algorithm and the hardware it is running on.

Visual Studio Support

ParaFormance now supports Microsoft Visual Studio as well as Eclipse, and is available on multiple platforms, including Windows, Mac OS X and Linux. Download a free trial by going to bit.ly/ParaFormance_VS. The Eclipse version can be downloaded at marketplace.eclipse.org/content/paraformance.

Tech transfer tip #2: Assemble a great team

Make sure your budding spinout has the human resources to succeed. While we already had solid technological expertise thanks to our research team, one of our first steps was to find a commercial champion.

ParaFormance at HiPEAC18

If you're at the HiPEAC 2018 Conference in Manchester, don't miss our ParaFormance tutorial at 14:00 on 25 January, where you can follow along with our demonstrations. You can also visit our company stand, where you can talk to one of our team members and request a personal demonstration, and we will also be presenting technical results at the HLPGPU workshop on 23 January.

FURTHER INFORMATION:

Email: enquiries@paraformance.com
paraformance.com

Is this the beginning of a paradigm change of how we do business and undertake research? In this article, Dr Costas Bekas, Distinguished Research Staff Member, Manager – Foundations of Cognitive Solutions, IBM Research – Zurich, explains how cognitive computing will allow us to extract the full value from huge datasets, whether for scientific discovery or new business models.

Cognitive discovery: Pushing



“Modern machine learning methods are starting to make the massive extraction of technical knowledge from highly unstructured sources possible”

Cognitive computing is the new frontier of the information age. Computers have evolved into indispensable tools of our societies, having modernized numerous aspects of our everyday lives. From the very first electronic general-purpose machines of the 1940s, they have facilitated the acquisition, storage and access of huge amounts of data. Since then, we have learned how to program computers to enable the use of tools such as the internet, social networks and simulations of the natural world that go well beyond the wildest imaginations of the computer pioneers of the 50s and 60s.

Cognitive computing turns our trusted programmable machines into cognitive companions. Systems are not programmed to simply achieve a task, but rather are developed to analyse in ways that are natural and complementary to us. They have the ability to debate and test our ideas in natural language as they decipher incredible volumes of data and give us insights that ultimately free us and allow us the space to tap into the deepest of human capabilities: intuition and intelligence.

Cognitive systems mimic the way we humans reason and how we express ourselves in unstructured ways. For example, speech and vision can be simulated in order to achieve feats in a small fraction of the time previously required. Pharmaceuticals and materials discovery, cancer treatment research, understanding both complex natural ecosystems and manmade ecosystems

such as the economy and technology are just a few ways in which cognitive systems can help humanity advance.

Today, technical research and development is facing a series of disruptions:

- The volume of public as well as proprietary technical knowledge, made available in the form of publications and technical documents/reports, is simply exploding. For instance, close to half a million papers in the field of materials science were published in 2016 alone. Moreover, these documents hold highly complex information, also known as ‘dark’ information: technical plots and diagrams, tables and formulas are just a few examples.

The situation can be even more challenging with internal company documents, as they may contain handwritten information that is nonetheless crucially informative. Hence, systematic extraction and organization of this vast ocean of bits and bytes into a knowledge base that allows deep search and inference is imperative.

- In the wake of the Industry 4.0 revolution, an unprecedented wave of information has the potential to flood corporate and public data holding systems.

Internet of things (IoT) systems also generate an enormous wealth of data that needs to be transformed into

the frontiers of R+D with AI

valuable knowledge and actionable items. Thus, the systematic ingestion of these streams into the knowledge base is a must.

- Finally, it is also becoming clear that supercomputing systems do not benefit from Moore’s law as they did in the past. As a result, our computing infrastructure needs to become ever more complex at all architectural and software levels and components in order to keep up top-line performance. However, this comes at the cost of significant loss of productivity and sometimes performance. The simple reason: parallelizing algorithms on highly complex systems is by no means straightforward.

methods, which in turn are powered by advances in computing, are starting to make the massive extraction of technical knowledge from highly unstructured sources possible.

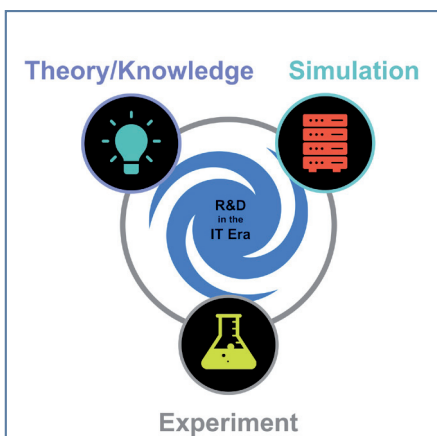
Knowledge graph technologies, as well as powerful graph inference and analytics methods allow unprecedented fidelity in knowledge representation. Early versions of these algorithms, such as Google’s PageRank, made the internet search revolution possible. Advanced methods such as spectral centralities, graph simplification and comparison allow for advanced knowledge analysis and hypothesis generation. Powerful inference methodologies as well as new causality methods can give great insights and deep reasoning based on data.

Artificial intelligence (AI) offers great promise with regard to overcoming these barriers. For instance, huge leaps in natural language processing and computer vision, powered by deep learning and other modern machine learning (ML)

Last but not least, ML-based surrogate models for physical systems provide great insights for simulations. This allows users to focus only on the models that have the best chances of advancing our knowledge and thus provide value.

“Powerful inference methodologies as well as new causality methods can give great insights and deep reasoning based on data”

AI is creating inroads for a whole new series of tools that expedite scientific and engineering progress. In the hands of experts, we expect significant improvement for timely innovation impact and perhaps groundbreaking science. This is the great promise cognitive discovery brings to the world. We have merely scratched the surface; the pace of innovation in this area is simply staggering. This is the dawning of a new era of an immense increase in research and development productivity.



Registered as an innovative start-up, University of Pisa spin-off IngeNiArs provides cutting-edge technology for the aerospace, healthcare and automotive sectors. Here, Marketing Manager Camilla Giunti explains what marks IngeNiArs out as a rising star in the technology world.

To the moon and back



COMPANY: IngeNiArs S.r.l.

MAIN BUSINESS: Design, development and commercialization of electronics systems, informatics systems and innovative services

in the aerospace, healthcare and automotive sectors

LOCATION: Pisa, Italy

WEBSITE: ingeniars.com



Innovative Italian start-up IngeNiArs was born in 2014 out of its joint founders' extensive experience in the areas of electronics systems, very-large scale integration design and advanced computer science engineering research. As a spin-off of the University of Pisa, it continuously promotes technology transfer from research outcomes to the market.

The name IngeNiArs, a fusion of the Latin words *ingenium* and *ars*, conveys a strong correlation between creative art and engineering skill. The key to IngeNiArs' success is the ability to combine these skills to create outstanding products and services. The company responds to the ever-increasing demand for innovation in the strategic aerospace, healthcare and automotive sectors, offering highly advanced hardware/software solutions and managing the full lifecycle of electronics, microelectronics and embedded systems.

In the aerospace field, IngeNiArs offers hardware description language intellectual property (IP) cores and hardware for high-speed, highly reliable links for telemetry and science data, as well as efficient solutions for state-of-the-art communication technologies such as SpaceWire, SpaceFibre and WizardLink, for both flight hardware and ground testing equipment.

For the healthcare market, IngeNiArs offers a family of innovative, interactive and advanced gateways, developed in cooperation with expert doctors, which are ideal for remotely monitoring medical parameters and the lifestyle of patients affected by chronic disease, such as chronic heart failure, chronic obstructive pulmonary disease, diabetes, hypertension, etc.

In the automotive sector, IngeNiArs offers design services to leading companies in the automotive electronics market. Drawing upon its experience in the development of digital systems for processing, communication, networking and security, the company supports the development of products to be used in ground vehicles.

Despite its relative youth, IngeNiArs already has several major achievements, such as obtaining a prime contract with the European Space Agency and winning the European Commission's Horizon 2020 SME Instrument Phase 1 and 2 projects. The second phase of the SME Instrument, in particular, is extremely competitive, with only around 3% of applicants being successful.

IngeNiArs' success was thanks to its SIMPLE (Spacefibre IMPLementation design and test Equipment) innovation project proposal. SIMPLE will produce three main outcomes:

- an innovative solution for the development of a SpaceFibre IP-core
- equipment for testing and validating designs related to SpaceFibre technology
- a module for the National Instruments test equipment platform with PXI interface



Aimed particularly at companies, agencies and research and development centres, these products will help speed up the development and testing of aerospace systems with high-speed communication requirements. But beyond this, the development of IngeNiArs' SpaceFibre technology will help advance the strategic aerospace field, helping SpaceFibre follow SpaceWire in becoming an internationally accepted technology originating from Europe.

Sabri Pllana, Linnaeus University, explains how his team is using machine learning for optimal worksharing on heterogeneous computing systems.

Smarter worksharing



Sabri Pllana, Linnaeus University

A node of heterogeneous computing systems typically comprises one or more host CPUs and several accelerators, also known as devices, such as the NVIDIA graphics processing unit (GPU) or the Intel Xeon Phi.

For instance, the Tianhe-2 supercomputer (rank 2 in TOP500 list, top500.org) combines two Intel Xeon E5 CPUs with three Intel Xeon Phi co-processors at each computing node. Compute nodes of Piz Daint (rank 3 in TOP500 list) combine Intel Xeon E5 CPUs with NVIDIA Tesla P100 GPUs. While heterogeneous computing systems provide high performance and energy efficiency, sharing the work between host CPUs and accelerators such that the overall program execution time is minimized is challenging.

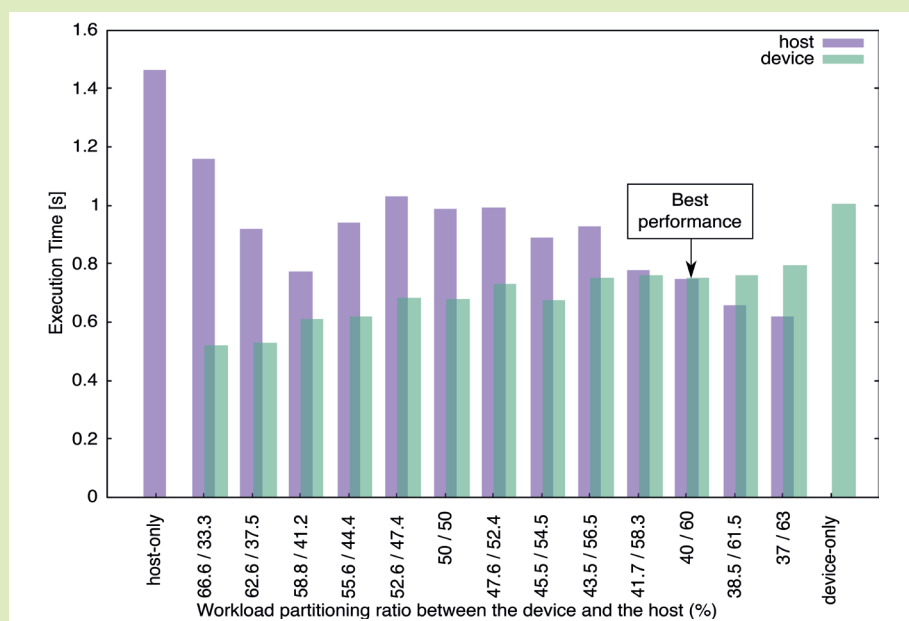
host (the bar on the far left), the execution time is higher than if 60% of the DNA sequence is processed by the host and the remaining 40% by the device. We propose using machine learning to determine the optimal workload sharing for a given DNA sequence and available host and device cores.

We use machine learning to split the DNA sequence between the host and the device based on the performance prediction, such that the load is balanced between the host and the device and the overall execution time is reduced. We developed the performance prediction model using the Boosted Decision Tree Regression, a supervised learning algorithm. The model training is performed using a set of 11 DNA sequences of different organisms (alpaca, armadillo, chimpanzee, coelacanth, duck, ferret, guinea pig, molly, elephant, turtle and zebra fish). The trained model enables us to satisfactorily share the load between the host and device.

Figure 1 illustrates the challenge of worksharing for DNA sequence analysis on a heterogeneous computing node that comprises two host CPUs (12 cores each) and one Intel Xeon Phi device (61 cores). In this example, we use all available cores of the host CPUs and all available cores of the Xeon Phi device. If 100% of the DNA sequence is analysed by the device (the bar on the far right) or 100% of the DNA sequence is analysed by the

Further reading:

Suejb Memeti and Sabri Pllana. 'A machine learning approach for accelerating DNA sequence analysis'. International Journal of High Performance Computing Applications, published online on June 26, 2016



“The trained performance prediction model enables us to satisfactorily share the load between the host and device”

Figure 1. An example of a DNA sequence analysis application running on the host and device



Pedro Tomás, INESC-ID of the Instituto Superior Técnico at the University of Lisbon, and Gabriel Falcão of the Instituto de Telecomunicações at the University of Coimbra, explain how their SCRATCH framework delivers twice the performance on half the energy, by tailoring the processing to the application.

SCRATCH: Automated generation of appli

Massively parallel processors, including graphics processing units (GPUs), have gradually occupied a prominent place in high-performance computing systems, with over 65% of the top 50 systems on the latest Green500 list being equipped with such devices. However, as the amount of data generated each year continues to rise, there is still enormous pressure to deliver the required processing performances within reasonable power and energy budgets in the years to come. On the other hand, while there is a permanent quest for more energy efficient computing systems, it is also important for new solutions and products to remain compatible with legacy code.

The SCRATCH framework, recently presented at MICRO-50 (the 50th International Symposium on Microarchitecture), aims at addressing this problem. SCRATCH

represents an open-source end-to-end solution (from OpenCL software to register transfer language (RTL) and field-programmable gate array (FPGA) implementation) for the development of application-specific soft-GPU architectures, operating under the AMD Southern Islands instruction set architecture (ISA).

The framework allows the architecture to be easily customized on a per-application basis to pursue higher performance and energy efficiency levels (see Figure 1). Application specificity is obtained by employing a special-purpose architecture-trimming tool that, by analysing the application source code, is able to free valuable resources, which can then be re-used to improve processing parallelism, e.g., by introducing more compute cores or more vectorized functional units.

Starting from the MIAOW soft-GPU architecture, developed at the University of Wisconsin-Madison to comply with the AMD Southern Islands ISA, we extended the set of supported instructions and validated their correct execution through a comprehensive set of benchmarks and tests. The revised architecture supports a total of 156 instructions, which allows a wider range of applications than in the original design to be supported. Additionally, we introduced a fast prefetch memory buffer capable of minimizing the slow access (and latency) to external global memory, and a dual clock mechanism to allow parts of the computing subsystem to operate four times faster than the clock frequency of the original MIAOW architecture (where the critical path resides).

To support the generation of application-specific soft-GPGPU architectures, we further developed an architecture-pruning tool (see Figure 1). By analysing the application's source code, the tool is able to remove all logic and hardware associated with the decoding and execution of unused instructions, generating optimized (application-specific) soft-GPU architectures with reduced area requirements.

Although the technology developed supports all kind of programs and applications, we focused on emerging applications and areas, namely related to computer vision and artificial intelligence, with a particular focus on image classification problems, which take advantage of convolutional neural networks (CNNs)

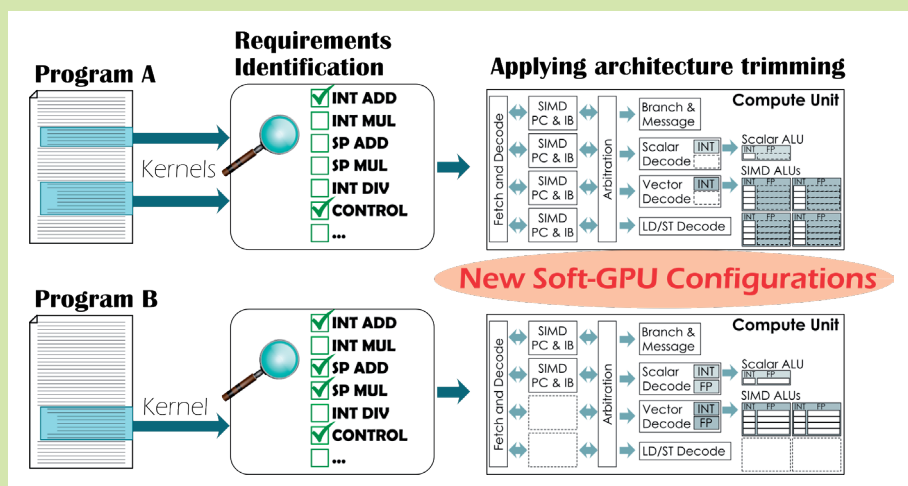


Figure 1: During compile time, the instructions present in each kernel indicate which functional units shall be instantiated on the reconfigurable fabric. This information is used by the architecture-trimming tool to automatically generate application-specific soft-GPU architectures.

Application-specific soft-GPGPU architectures

or other deep learning approaches – see for example Figure 2.

Compared with the original MIAOW architecture, we improved processing performance and energy-efficiency levels by two orders of magnitude, using a Xilinx Virtex 7 FPGA. Additionally, by allocating the freed resources to instantiate additional (and useful) computing elements, we attained 2.4x speedup and 2.1x energy-efficiency gains when comparing the application-specific (optimized) architecture against the generic unspecific one. This is achieved through a combination of a

significant reduction in power needs and with the employment of area savings to increase processing parallelism by up to four times.

Naturally, as FPGA technology advances and is increasingly adopted by a larger community of system developers, the technology proposed in this paper will become more and more widely applicable. The developed tool is user friendly and attractive for application developers, who often do not have the skills for programming the FPGA using hardware description language (HDL). Furthermore,

SCRATCH allows the development and testing of additional architectural optimizations that could provide new performance or energy-efficiency gains. For example, one can adjust the bit width of the datapath to provide additional gains in terms of area and power, especially since in many applications (e.g. CNNs) it is perfectly acceptable to reduce numerical precision.

The SCRATCH framework proposed in our paper (see ‘Further reading’, below) is therefore a full end-to-end solution, providing users a way to compile an OpenCL program, trim the design to satisfy the application-specific requirements (optional step), synthesize, implement and run the application on Xilinx FPGAs.

The MIAOW2.0 architecture and SCRATCH tool are publicly available on GitHub, under repositories MIAOW2 and Trimming-Tool, respectively, for the community to try out. They represent ongoing work and are therefore subject to continual updates with new ideas and solutions being gradually developed and released.

github.com/scratch-gpu

Further reading:

P. Duarte, P. Tomas, and G. Falcao. ‘SCRATCH: An End-to-End Application-Aware Soft-GPGPU Architecture and Trimming Tool’, Proceedings of IEEE/ACM International Symposium on Microarchitecture (MICRO), Boston, MA, United States, October 2017.

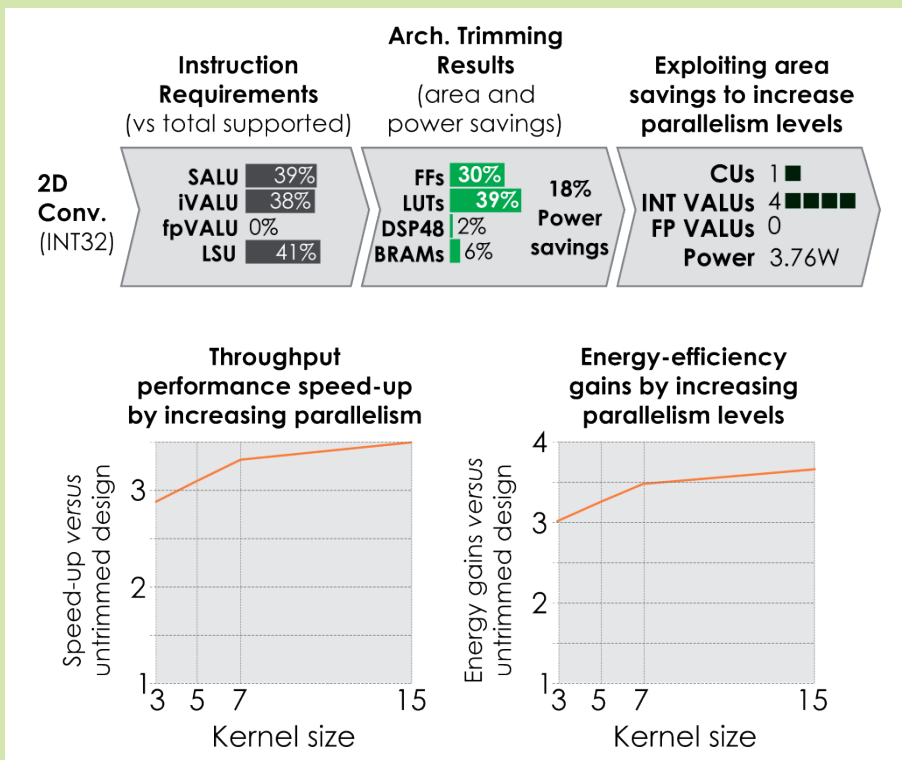


Figure 2: By applying architecture trimming, important area and power savings are made. The exposed FPGA resources are then exploited to increase parallelism levels and improve throughput performance and energy-efficiency levels.

The internet of things is coming, and, as we are all aware, it will bring with it a deluge of data. Here, Kemal A. Delic, David M. Penkler (Hewlett Packard Enterprise) and independent technology specialist Martin Walker argue that, properly executed, high-performance machine learning could be the contemporary equivalent of the microscope or telescope in furthering scientific progress.

On high-performance



At a basic level, machine learning is about presenting a computer program with enough training samples representing the measured attributes or successive states of a system to achieve a satisfactory rate of recognition of new, unseen samples, or prediction of future values. Recognition or prediction here is to be understood as producing correct results. A measure of correctness on unseen samples or known future values is necessary in order to check that the topology and training samples adequately capture the system, although in some domains, such as language translation, it is difficult to define a sharp metric to measure ‘correctness’. While the principles of machine learning were set out a long time ago, technology, methods and large datasets have only recently made it practical for large-scale, industrial deployments on a wide variety of problems.

“Problems will have different levels of complexity, and will require optimal choices of infrastructure, data volume and type of algorithm”

High-performance machine learning aims to achieve the shortest possible training time and execute inference or recognition in the most efficient way, while minimizing energy consumption. Neural networks – better called multilayer weighted networks (MWNs) – are currently the most frequently used mechanism to capture training sessions within a compact model used for inference or recognition. Inference and recognition are two distinct acts for which one must find optimal solutions (datasets, algorithms, infrastructure) for efficient and effective problem solving. Problems will have different levels of complexity, and will require optimal choices of infrastructure, data volume and type of algorithm. Thus, for example, one can think of a space (see Figure 1) in which problem complexity determines the resources required, expressed as storage requirements (xbytes) and necessary computing power (xFLOPS).

The universal approximation theorem for MWNs states that any function of compact support (or large class of bounded functions) can be approximated arbitrarily accurately by an appropriately weighted multilayer network. This theorem is the basis for the belief that multilayer weighted networks can be trained to reproduce observations of those natural phenomena that can be described by numerical simulation or modelling – that is, those for which the governing mathematical equations, typically systems of partial differential equations, are known.

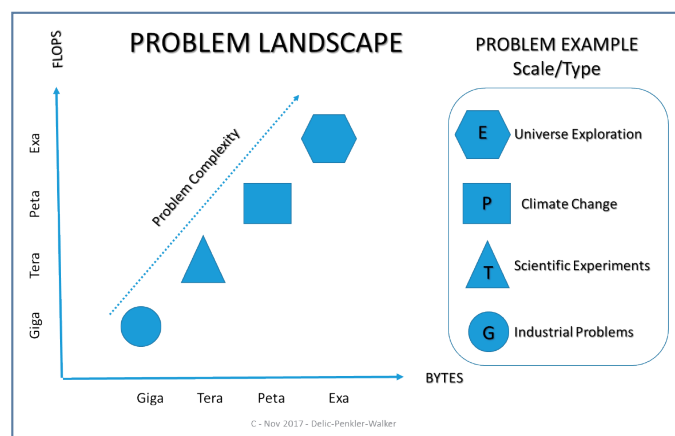


Figure 1 – Problem complexity versus storage/computation

Machine learning

The application programs underlying numerical simulations can be used to train appropriate MWNs. The resulting trained MWNs could then be run (perform inferences) anywhere, without needing to port the underlying application programs to different machines. In this way, MWNs provide a bridge between artificial intelligence and traditional high-performance computing.

Approximation to solutions of partial differential equations with MWNs of course requires determination of the size and topology of the networks needed, in addition to determining the weights through training. Attention needs to be paid to the impact of network size, topology, and weight determination on the accuracy of the resulting approximations.

In future, machine learning will need to respond to extreme requirements for the field of exascale computing, which will potentially resolve grand challenges or so-called ‘moonshot’ projects in different domains of scientific inquiry or industrial development. The forthcoming roll out of the internet of things will create huge data repositories, called data lakes, with a vast volumes of a wide variety of data reaching exabyte sizes (10^{18}). To deal expeditiously with such volumes and variety of data, we will need exaflops (10^{18}) of computing power.

Performance will be about shortening training duration by several orders of magnitude and radically improving the inferencing process. We believe that a hybrid infrastructure – such as central processing unit (CPU) plus graphics processing unit (GPU) – will be best for training purposes, while specialized chips – such as tensor processing unit (TPU) or Tofino – will be necessary for efficient inferencing execution. Overall, this will ensure latency-critical problems are addressed properly.

Clearly, Big Science will require large and novel infrastructure. At the same time, with judicious choices in algorithms, data-lake content feed, and adequate infrastructure, machine learning may enable scientific advances similar to those enabled by the invention of the microscope and telescope a couple of centuries ago.

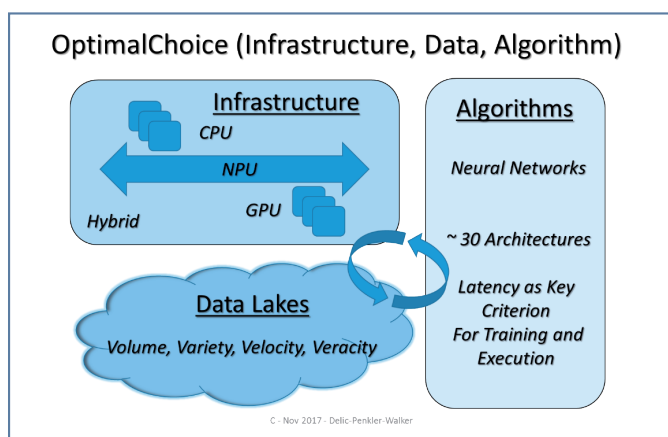


Figure 2 – Architecture of hyperscale, high-performance ML system

Further reading:

Christopher M. Bishop. Pattern Recognition and Machine Learning Springer-Verlag New York, Inc.: Secaucus, NJ, USA, 2006

Karlijn Willems. ‘How Machines Learn: A Practical Guide’ bit.ly/How_machines_learn

Kemal Delic. ‘Big Science Will Require a Big and Different Infrastructure’ bit.ly/BigScience_Delic_BVEX

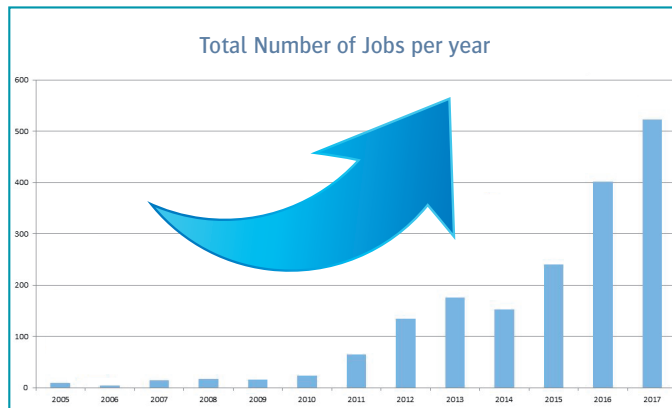
“Machine learning may enable scientific advances similar to those enabled by the invention of the microscope and telescope”

Computing systems jobs: what's new?

Smarter searching on the HiPEAC Jobs portal

According to the 2017 HiPEAC Vision, we are entering the artificial intelligence era, with all that entails both for how we interact with machines and how we instruct machines what to do. New computing systems and technologies need to be developed to address this new paradigm. That's why you'll now find machine learning as a HiPEAC core skill on the HiPEAC Jobs portal, allowing you to upload and filter vacancies covering all aspects of this field.

If your institution is developing the high-end computing systems that power neural networks or optimize systems training, or if you're developing machine-learning applications that need advanced heterogeneous computational platforms, this will help you find the right people from HiPEAC's pool of specialist personnel. Meanwhile, if you're looking for a new opportunity in this exciting area, it will now be easier to find the perfect match. So far, the portal has featured 67 machine learning job opportunities, 27 of which were in the last quarter.



This growth is due in large part to our focus on HiPEAC Jobs activities over the last year, including the travelling careers centre at major conferences and careers sessions at HiPEAC events. You can find a full list of recruitment support services on the careers centre webpage:

hipeac.net/jobs/career-center

Use your network to spread the word

All this growth is only worthwhile if you find the portal useful and it keeps on sourcing the right candidates for the vacancies. We need to reach all those from your universities interested in doing a PhD or engineering career at a HiPEAC institution, whether final-year PhD students looking for a post-doc position or senior researchers who want to advance their careers.

Number of Jobs posted in the last quarter by institution	Jobs from relevant institutions, sponsors, contributors & projects
37 Jobs @ Research Centres	
62 Jobs @ Industry	
23 Jobs @ SMEs	
54 Jobs @ Universities	
36 Jobs @ H2020 Projects	
25 Machine Learning Jobs	

More than 500 jobs posted in 2017

In early December we reached the milestone of 500 job vacancies posted on the HiPEAC Jobs portal in 2017, and we are on course to reaching more than 1,000 job vacancies since the beginning of HiPEAC 4. The portal's user numbers have continued to increase month on month, and visitors to the portal almost doubled in 2017, compared to previous years. The number of open positions on the website has also been consistently breaking records, while the total number of new opportunities in the last quarter reached a new high of 176.

You don't have to be an HR professional to use the portal; if you need new team members, it takes less than 10 minutes to upload a vacancy and get it out to the HiPEAC community. What recruiters value most are the specialist profiles they get via the portal.

How you can help:

- Forward the HiPEAC monthly job opportunities email to your students, colleagues and other university departments. If you're not currently receiving this, let us know by emailing recruitment@hipeac.net.
- If your university is preparing a careers event, let us know – or just put in contact with your institution's careers centre. We can provide material and publicize the event.
- Promote your summer school, Master's or PhD programmes and show future students the great career opportunities they can get as part of HiPEAC.
- Got any more ideas on how we can reach the next level? Contact us at recruitment@hipeac.net

Looking for your next opportunity or have a post you need to fill? Visit the HiPEAC Jobs to check out the numerous opportunities and upload your vacancies: hipeac.net/jobs

Career talk: Trevor Carlson, National University of Singapore



What are you currently working on?

Right now, my focus is on bringing efficient and flexible computation to the internet of things (IoT) hierarchy. While accelerators are one modern means to efficiency, they remain application specific and are optimal for a set of specific, pre-defined tasks. Unfortunately, the precise needs of the future compute infrastructure are not known, as applications, especially those that run in today's data centres, change frequently. I'd like to get closer to answering the question: how can we build both efficient and flexible solutions for future needs?

As an example, here in Singapore the Smart Nation initiative aims to improve living conditions by leveraging IoT and distributed sensing, which requires a large number of distributed devices. Deploying hundreds of thousands of IoT devices needs to be affordable, energy efficient and flexible for yet-to-be-designed algorithms. Replacing thousands of devices because they are no longer efficient for new applications is not sustainable. I'm looking to develop processors that are efficient, high performance and configurable to meet future needs.

What trends are you keeping an eye on in high-efficiency microarchitectures and performance analysis?

I've really enjoyed two recent trends in architecture research. The first was the direct result of work by McFarlin, et al. which describes how high-performance processors receive significant performance benefits primarily from speculation. The programs they evaluated do not exhibit significant dynamic schedule variability, which means that they are inefficiently re-learning the same instruction schedule

over and over again. Several papers over the past few years have taken advantage of this knowledge to improve performance and efficiency by using the speculative knowledge from a larger core into a smaller, more efficient one.

Second, I am very excited to see that the open-source hardware movement, specifically with RISC-V and the many projects that build upon it, has allowed for a great deal of experimentation in computer architecture. Researchers can now jump in, design and evaluate new ideas, and can work to evaluate the efficiency of the processor directly down to the silicon. In addition, the work on the new RISC-V vector instruction set shows how these platforms can serve to bring back interesting ideas for efficiency and performance.

As for performance analysis, this really makes up the foundation needed for most architecture research. The biggest advance we've seen recently has been Intel's TopDown methodology to use real hardware counters with a single run to determine performance bottlenecks. While this works well for current platforms, future platform development still requires stimulation. Recent work on field-programmable gate array (FPGA)-based simulation for performance analysis and energy efficiency show how FPGA-based platforms might one day be commonplace and accelerate our research.

You've worked in both industry and academia – what are the main differences?

I have been truly lucky to work in industry, at an innovation hub and in academia. While working in industry, I was able to build solutions for products that were

“Real-world requirements – and the limits they place on research – often produce much more innovative and impactful work”

about to hit the market, and would touch the lives of many people. I was able to run my own team and pursue my own directions, while developing new, patentable ideas and helping co-workers on the other side of the world. It was often fast-paced, demanding and rewarding.

Academia, on the other hand, can give you the time to reflect. As researchers, we need to know when to dig deep and when the idea isn't worth it. In industry, someone has already determined that the idea is good; we just need to find the most efficient way of getting there.

For my own research, I have found that having too much of an open-ended mandate interferes with good ideas. Although it seems counter intuitive, real-world requirements – and the limits they place on research – often produce much more innovative and impactful work. As an example, when we were collaborating with Intel at the Exascience Lab in Leuven during my PhD studies, we wanted to simulate next-generation high-performance computing (HPC) platforms. We couldn't find a simulator that met our needs, so we built a new one, along with new sampling and simulation methodologies to speed things up.

In addition, as researchers we face many failures along the way. The biggest adjustment I had to make was to see these as lessons and learn how to fail faster, thereby learning more quickly from my mistakes.

How does work culture differ between the USA, Europe and Asia?

In the USA, our team had pre-defined goals and often worked long hours to meet them. There was an implicit expectation that we would work to get the job done, even if that meant working late and at weekends.

In my first job in Europe, at imec in Belgium, we worked hard during the working day, and were rewarded with ample vacation time. In Belgium, every office must have a window, and the corner office on my floor was occupied by PhD students – a big change from a culture where the corner window offices were reserved for senior management. I was also surprised to see wine and beer in the work cafeteria, seeing this as an appreciation for food and life, instead of a taboo against drinking at work.

Sweden's work culture is defined by 'fika': officially, this translates as 'coffee break', but in reality it's a block of time where people come together to have (strong) coffee and share ideas. This has a huge impact on the group's culture: everyone

knows one another informally, information is shared and communication is strong.

My first impressions of the work culture in Singapore is that people are highly driven and willing to share insights and suggestions. The country and the university value high impact and high-quality research, and provide the resources required to complete that work.

Each culture brings its own rewards; my personal challenge has been to jump into each with an open mind, and to try to integrate my favourite aspects to create a truly global workplace.

How does being in Singapore influence your perspective on computing research?

Singapore as a nation has prioritized the development of a knowledge-based economy, recently investing S\$19 billion to continue its development as a research and development hub. One aspect is the Smart Nation Sensing Platform, which aims to improve the ability of the country to monitor and react to the environment. I expect that my future research directions will be shaped by the need for a more efficient and flexible sensing and analysis platform. I feel that Singapore aims to foster research which has an impact on the community, and I hope to make a meaningful contribution.



Singapore aims to become Smart Nation. Photos: Mike Enerio and Duy Nguyen on Unsplash



In 2017, Jan Zapletal (VŠB - Technical University of Ostrava) won first prize in the Joseph Fourier award for computational sciences. The award, a joint initiative by Atos France and the French Embassy in the Czech Republic, recognizes outstanding doctoral work in computer sciences, and attracts competitors from across the Czech Republic. Here Jan, whose thesis was supervised by Jiří Bouchala, tells us more about his research.

Multiphysics made easier

The boundary element method

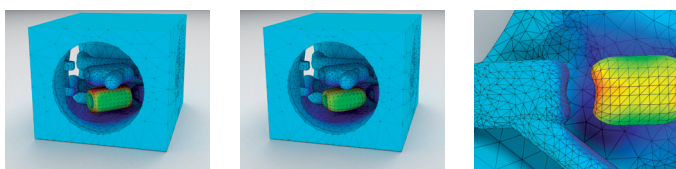
The boundary element method (BEM) is a numerical approach for solving partial differential equations. Its key advantage over volume-based methods is dimension reduction, since only the boundary of the domain has to be discretized. In addition to simplifying mesh generation and storage, this aspect leads to much smaller systems of equations. Over the course of my studies, I tackled computational problems in the areas of heat conduction, electrostatics, wave scattering and shape optimization, and gained experience in implementing efficient solvers based on BEM.

Shape optimization with BEM

While undertaking an internship at TU Graz, I participated in the FP7 Marie Curie project (Controlled Component and Assembly-Level Optimization of Industrial Devices). Its aim was to provide a tool to optimize the shape of high-voltage electronic devices to prevent electric failures. The tasks within the project included mathematical modelling of electrical fields using a BEM solver (TU Graz) and the implementation of a multi-resolution optimization algorithm (University of Cambridge). The computational cost measured to assess completion of the objective decreased by almost 18%.

Development of an HPC-optimized BEM library

The key output of this thesis is the high-performance computing (HPC)-oriented C++ boundary element library. This is being developed at the IT4Innovations National Supercomputing Center, and can be deployed not only on high-performance computers but also on modern workstations. Currently, the library is able to solve 3D problems in heat transfer, electrostatics, time-harmonic sound, electromagnetic wave scattering and linear elasticity.



To take full advantage of modern processors, BEM4I leverages several layers of parallelism. To fully utilize the potential of modern HPC systems, intra-node operation of the solver is crucial. This is especially pronounced on clusters with manycore systems and wide single instruction, multiple data (SIMD) registers, represented by the Xeon Phi (co-)processors. Failure to implement efficient threading or SIMD approaches on such systems leads to a waste of computational power accompanied by inefficient use of energy. Experiments performed on various HPC systems have shown that vectorization is becoming a crucial part of the scientific code design process.

To deploy BEM4I to massively parallel architectures, the library can be linked to the domain decomposition ESPRESO library, also being developed at IT4Innovations. This combination leads to a method utilizing all available parallelization layers, including Message Passing Interface (MPI) over distributed memory, threading within a single node and SIMD vectorization.

Applications of BEM4I

The parallelism layers in BEM4I allow us to tackle problems with up to millions of surface degrees of freedom, corresponding to up to millions of volume unknowns. The library can be used to solve problems in the areas of:

- noise prediction and shape optimization of sound barriers
- distribution of electromagnetic signals
- exterior problems for heat conduction or wave scattering
- non-linear large-scale contact problems in linear elasticity

MORE INFORMATION:

Jan Zapletal. PhD thesis: 'The Boundary Element Method for Shape Optimization in 3D'

bit.ly/BEM_shape_optimization_3D

BEM4I library

bem4i.it4i.cz

ESPRESO library

espresso.it4i.cz

Thanks to our sponsors for making #HiPEAC18 a great success!

Atos

BARCO

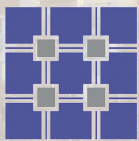
$\frac{d\vec{v}}{dt}$

arm

DeepMind

Google

AIXIOM



Embedded
Computing
Specialists

polly labs

SAMSUNG

Springer

SUNDANCE

SYSGO
EMBEDDING INNOVATIONS

THALES

Think Silicon
ultra-low power | vivid graphics

Sponsors correct at the time of going to print. For the full list, see hipeac.net/2018/manchester

Join the community

@hipeac

hipeac.net/linkedin



hipec.net

This project has received funding from the European Union's Horizon2020 research and innovation programme under grant agreement no. 687698

