

HiPEAC *info* 45

COMPILATION **ARCHITECTURE**

APPEARS QUARTERLY
JANUARY 2016

**NETWORK OF EXCELLENCE ON
HIGH PERFORMANCE AND EMBEDDED
ARCHITECTURE AND COMPILE**

**WELCOME TO
THE HIPEAC 2016
CONFERENCE,
18-20 JANUARY 2016,
PRAGUE,
CZECH REPUBLIC**



**HiPEAC4 kicks off
in Prague!**



ACACES 2016 SUMMER SCHOOL: 10-16 JULY 2016, FIUGGI, ITALY

MESSAGE FROM THE HIPEAC COORDINATOR



First of all, I would like to wish you a healthy and prosperous 2016, personally as well as professionally.

For me, 2015 was the year in which European solidarity was stress tested. We had the acute financial problems in Greece, the massive influx of refugees in Europe, the threat of terrorism in several European countries, and the challenge of climate change. Rationally, all these challenges call for solutions at the European level, but the opposite seems to be happening: countries are told to solve their own financial problems, some countries are closing their borders for refugees, and many countries are focusing on local plans to fight global problems like terrorism and climate change. It shows that Europe is politically less united than we would hope. I nevertheless wish that our European political leaders will find effective solutions for these important challenges.

On the positive side, 2015 was also the year in which my new year's resolution of last year became reality: we successfully transited HiPEAC from a project in the Seventh Framework Programme into a project under Horizon 2020. This was a challenge because Horizon 2020 no longer calls for networks. Therefore, we had to morph the network into a coordination and support action. HiPEAC4 will offer most of the services we know from HiPEAC3, but on top of it, it will increasingly focus on the production of policy documents, on communication and dissemination and on job mobility in Europe. Therefore, we have extended the core HiPEAC team with three colleagues: Madeleine Gray from Barcelona will develop HiPEAC communication services, and Maureen Simpson and Catherine Inglis from Edinburgh will develop the HiPEAC recruitment services. I warmly welcome them in the HiPEAC community.

Together with Vicky and Eneko in Ghent, we have a very strong team to run HiPEAC4. Many of you will read this newsletter at the HiPEAC conference in Prague. The HiPEAC conference is the flagship event for the HiPEAC community. This year it will mark the end of HiPEAC3, and the start of HiPEAC4. I am thankful for the many volunteers who work very hard to make this event a successful and well-attended conference, and a sign of a thriving community in computing systems in Europe.

Take care,
Koen De Bosschere

CONTENT

HiPEAC ACTIVITY

- 3 REPORT ON THE THEMATIC SESSIONS IN THE MILANO COMPUTING SYSTEMS WEEK (SEP. 21-23, 2015)
- 5 ACACES 2016: 10TH – 16TH JULY, 2016, FIUGGI, ITALY

HiPEAC ANNOUNCE

- 6 BOOK: NEAR THRESHOLD COMPUTING
- 6 PARASUITE: PARALLEL BENCHMARKS FOR MULTI-CORE CPUS, CLUSTERS AND ACCELERATORS
- 7 ESPRESSO SOLVER

HiPEAC START-UP NEWS

- 8 HPC STARTUP APPENTRA AT SC15 IN AUSTIN, TEXAS

HiPEAC NEWS

- 9 MATEO VALERO BECOMES THE FIRST EUROPEAN TO RECEIVE THE SEYMOUR CRAY AWARD
- 9 HIPEAC ASSOCIATED MEMBER RECEIVES THE BENJAMIN FRANKLIN MEDAL
- 10 ALEXANDRA FERRERON AWARDED THE BRONZE MEDAL AT ACM STUDENT RESEARCH COMPETITION
- 10 ICCD BEST PAPER AWARD FOR TUDELFT PAPER
- 10 LEONEL SOUSA AND WALID A. NAJJAR SELECTED ACM DISTINGUISHED SCIENTISTS
- 11 MIGUEL ANGEL AGUILAR AWARDED GOLD MEDAL AT ACM SRC

- 11 SHAOTENG LIU AND ZHONGHAI LU RECEIVE BEST PAPER AWARD AT NOCS 2015
- 12 HIPEAC MEMBERS RECEIVE RECOGNITION FROM INTEL FOR THEIR RESEARCH ON TRANSACTIONAL MEMORY
- 12 INTEL INVESTS \$50M IN TU DELFT TO BUILD AN EFFECTIVE QUANTUM COMPUTER
- 13 IS 3D-STACKED MEMORY THE SOLUTION FOR HPC?
- 14 1ST INTERNATIONAL SYMPOSIUM ON MEMORY SYSTEMS (MEMSYS) 2015
- 14 1ST PANDORA SUMMER SCHOOL

HiPEAC IN THE SPOTLIGHT

- 15 THE ANTAREX PROJECT
- 16 NEW H2020 PROJECT: VINEYARD
- 17 THE CTUNING FOUNDATION WORKS WITH ARM TO ACCELERATE COMPUTER ENGINEERING
- 18 THE HOLISTIC PROJECT IN DEPENDABLE COMPUTER ARCHITECTURE

HiPEAC STUDENTS

- 18 INTERNSHIP REPORT: PAULO MARTINS
- 19 INTERNSHIP REPORT: MARCO BALBONI
- 19 COLLABORATION GRANT: JAVIER JALLE

PHD NEWS

UPCOMING EVENTS

Cover photography: Davis Staedtler under license CC BY-SA 2.0

MESSAGE FROM THE EDITORS

For the duration of HiPEAC₃, the HiPEAC newsletter has been edited by a team at Chalmers University of Technology. It has been an honour to serve our community during this time. We would like to thank all HiPEAC members for their contributions and support. We extend a special thanks to Ruben Titos who was the newsletter's

assistant editor until May 2014 and Paul Carpenter who has been the proofreader since July 2012. Madeleine Gray from the Barcelona Supercomputing Center will be the communications officer for the HiPEAC₄ CSA, starting next January. We welcome Maddy to the team and wish her good luck in her new task! We are confident

that under her direction the HiPEAC magazine will become an even more successful instrument for our community.

Miquel Pericàs and Per Stenström

REPORT ON THE THEMATIC SESSIONS IN THE MILANO COMPUTING SYSTEMS WEEK (SEP. 21-23, 2015)

The HiPEAC Computing Systems Week (CSW) was held in Milano in Sep. 21-23, 2015, and included a new set of six Thematic Sessions. They were all successful and well attended. We present a brief summary of the presentations and discussions.



OPTICAL INTERCONNECTION GOING TOWARDS SMALL SCALE TO ENABLE EFFICIENT LARGE SCALE COMPUTING SYSTEMS: LIGHT IS FILTERING DEEP INTO SILICON (MONDAY, SEP. 21ST, 2015)

Organized by Sandro Bartolini (UNISI), Jose M. Garcia (Univ. of Murcia), and Davide Bertozzi (Univ. of Ferrara)

This thematic session focused on the perspectives of optical technologies in the computing domain. It included six presentations:

- "Silicon Photonics: Designing for Complexity" (Wim Bogaerts, UGent, IMEC)
- "Present status of the industrial Silicon Photonics process at ST and next RD" (Guido Chiaretti, STMicroelectronics)
- "Photonic integrated technologies for large bandwidth density and low consumption intraboard interconnections" (Vito Soriano, CNIT)
- "Design Methods for High Performance Silicon Photonics Interconnects" (Sébastien Le Beux, Lyon Inst. of Nanotechnology)
- "Highly Scalable Silicon Photonics Interconnects for Large Scale Computing Systems" (Antonio La Porta, IBM Research Zürich)
- "Scalable and Agile Optical Interconnection for Future Computing Systems by Array Waveguide Grating Routers" (Roberto Proietti, Univ. of California at Davis)

CHALLENGES AND OPPORTUNITIES IN NEXT-GENERATION HPC SYSTEMS FOR REAL-TIME APPLICATIONS (MONDAY, SEP. 21ST, 2015)

Organized by José Flich (Univ. Politècnica de València)

This thematic session dealt with the integration of HPC and real-time QoS for new systems. It included the following presentations:

- "MANGO project: Exploring new heterogeneous architectures for HPC systems" (José Flich, Univ. Politècnica de València)
- "Mont-Blanc, an HPC supercomputer prototype based on commodity embedded technology: lesson learned and projection for future HPC architectures" (Filippo Mantovani, BSC)
- "The DEEP project: the Cluster-Booster approach to heterogeneous computing" (Damian Álvarez, JSC)
- "ANTAREX: AutoTuning and Adaptivity approach for Energy efficient eXascale HPC systems" (Cristina Silvano, POLIMI)
- "Potentials for multimedia processing mapped to heterogeneous HPC" (Mario Kovač, Univ. of Zagreb)
- "Multi-Scale Thermal Modeling and Simulation for Energy-Efficient HPC Systems" (David Atienza, EPFL)
- "Reconfigurability in HPC: opportunities and challenges" (Alessandro Cilardo, CeRICT/Univ. of Naples Federico II)
- "Unified Network Architecture for HPC Real-Time Systems" (José Flich, Univ. Politècnica de València)
- "Multi-layer run-time resource management for multi-many-core heterogeneous architectures" (William Fornaciari, POLIMI)

ADVANCED IMAGE PROCESSING IMPLEMENTATIONS – A MULTIDISCIPLINARY APPROACH (TUESDAY, SEP. 22ND, 2015)

Organized by Calliope-Louisa Sotiropoulou (Univ. of Pisa and INFN Pisa)

This thematic session dealt with the increasing needs for image processing in different domains: medicine, security, high-energy physics... It included the following presentations:

- “Advanced Image Processing Implementations: A multidisciplinary approach – an introduction” (Calliope-Louisa Sotiropoulou, Univ. of Pisa and INFN Pisa)
- “The Associative Memories Chips: the Past and the Future” (Alberto Stabile, UniMI and INFN Milano)
- “Advanced Embedded Systems for Real Time Pattern Matching” (Pierluigi Luciano, Univ. of Cassino and Southern Lazio, and INFN Pisa)
- “High Performance Computing for Medical Image Processing?” (Piergiorgio Cerello, INFN Torino)
- “Quantitative MRI of the brain using magnetic resonance fingerprinting” (Guido Buonincontri, INFN Pisa)

TACLE: TIMING AND ENERGY ANALYSIS TECHNIQUES FOR MULTICORE/MANYCORE SYSTEMS (TUESDAY, SEP. 22ND, 2015)

Organized by Kevin Hammond (Univ. of St. Andrews)

This thematic session focused on issues related to timing and energy analysis for manycore systems. It included the following presentations:

- “Using Parallel Patterns to drive Time and Energy Analysis for Multicore/Manycore Systems” (Kevin Hammond, Univ. of St Andrews)
- “Power and Energy Estimation at System-Level” (Santhosh Kumar Rethinagiri, BSC)
- “Timing Correctness for Hard Real-Time Multicore Applications” (Kai Lampka, Uppsala Univ.)

RISING VIRTUES OF HETEROGENEOUS SYSTEMS: DEBUGGABILITY (WEDNESDAY, SEP. 23ND, 2015)

Organized by Chris Fensch (Heriot-Watt Univ), Marisa Gil (UPC), Georgios Goumas (NTUA)

This thematic session focused on how systems are becoming increasingly complex, and the needs for good debugging techniques for parallel applications. It included the following presentations:

- “GPUVerify: Static Verification for GPU Programming” (Alastair F. Donaldson, Imperial College London)
- “A Scalable Tool for Automated Bug Localization” (Maksim Jenihhin, Tallinn Univ. of Technology)

The session concluded with a panel discussion on how to improve the debuggability and correctness of applications on heterogeneous systems.

SYSTEM-LEVEL INTERCONNECTS FOR EXASCALE-CLASS DATACENTERS AND HPC (WEDNESDAY, SEP. 23RD, 2015)

Organized by Fabien Chaix (FORTH), and Nikolaos Chrysos (FORTH)

This session dealt with the development of effective communication infrastructures including hardware and software designs, and tools that can be used to improve them, based on simulations and prototyping. The session included the following presentations:

- “Challenges when Bridging Networks for HPC” (José Flich, Univ. Politècnica de València)
- “Star-Replaced Networks: A Generalised Class of Dual-Port Server-Centric Data Centre Networks” (Javier Navaridas, Univ. of Manchester)
- “Simulation of HPC applications with SimGrid” (Frederic Suter, IN2P3 Computing Center / CNRS / Inria Avalon)
- “Unexplored energy aspects of scalable heterogeneous computing systems” (Holger Fröning, Univ. of Heidelberg)
- “Network on Chips for Modern Mixed Criticality Systems” (Marcello Coppola, STMicroelectronics)
- “Rate-based vs Delay-based Control for Dynamic and Voltage Frequency Scaling (DVFS) in Network-on-Chip (NoC)” (Mario Casu, Politecnico di Torino)

These Thematic Sessions in Milano were the last such meetings in the HiPEAC₃ Network of Excellence. Over the course of HiPEAC₃ the Thematic Sessions have been a successful mechanism to foster collaborations between research groups. We will therefore continue organizing similar meetings during the HiPEAC₄ collaborative project that starts in February 2016.

I would like to take the opportunity to thank all organizers, contributors and participants in the sets of Thematic Sessions organized over the past years: 68 Thematic Sessions in eight Computing Systems Weeks. Without their contribution and efforts, it would not have been possible to reach such a remarkable number.

These sessions have covered all topics from hardware and software, including applications, resiliency, debuggability, HPC and Real-Time support, and new approaches that surely will be growing in importance in the near future, such as silicon photonics and quantum computing.

As a final summary, the following are the Computing Systems Weeks with Thematic Sessions, during HiPEAC-3:

• Göteborg	April 24-25th, 2012	8 thematic sessions
• Ghent	October 15-17th, 2012	12 thematic sessions
• Paris	May 2-3rd, 2013	7 thematic sessions
• Tallinn	October 7-9th, 2013	7 thematic sessions
• Barcelona	May 12-16, 2014	7 thematic sessions
• Athens	October 8-10th, 2014	11 thematic sessions
• Oslo	May 4-8th, 2015	10 thematic sessions
• Milano	September 21-23rd, 2015	6 thematic sessions

Thanks again to all participants!

Xavier Martorell, Barcelona Supercomputing Center and Universitat Politècnica de Catalunya

ACACES 2016: 10TH - 16TH JULY, 2016, FIUGGI, ITALY

12th international Summer School on Advanced Computer Architecture and Compilation for high-performance and Embedded Systems

We are proud to announce the twelfth HiPEAC Summer School, which will take place in downtown Fiuggi, a historical hill town near Rome, during the third week of July. We will start on the Sunday evening with an opening keynote. The courses will start on Monday, and will be spread over two morning and two afternoon slots. There will be three parallel courses per slot, from

which the participants will be able to take one course. These courses have been allocated to slots in such a way that it will be possible to create a summer school program that matches your research interests.

The following world-class experts will present topics of this year's Summer School.

Instructor	Course
Uday Bondhugula, Indian Institute of Science	Building High-Level Compiler Optimizers and Code Generators for the Multicore Era
Nikil Dutt, University of California, Irvine	Variation-aware embedded system design and cross-layer memory optimizations
José Flich, Universitat Politècnica de València	Design of Efficient On-Chip Interconnects for Future Manycore Systems
Geoff Gregson, Northern Alberta Institute of Technology	New Venture and Entrepreneurship
Lizy John, The University of Texas at Austin	Performance Evaluation and Benchmarking
Rakesh Kumar, University of Illinois	Memory System Resilience
Timothy Roscoe, ETH Zürich	Operating Systems Directions for Future Hardware
Simha Sethumadhavan, Columbia University	Hardware Security
Steve Wilton, University of British Columbia	Reconfigurable Computing: Technology and Practice
Huiyang Zhou, North Carolina State University	Efficient GPU Computing

On Wednesday afternoon, participants will be given the opportunity to present their own work to the other participants during a huge poster session; and finally, on Friday evening there will be a farewell dinner and party.

The accommodation will be provided by a consortium of hotels in Fiuggi, all located closely together. There will be abundant Italian food, and the town of Fiuggi will provide plenty of opportunities to socialize in the evenings. At the end of the event, all participants will receive a certificate of attendance detailing the courses that

they took. If you are a student member of HiPEAC, you can apply for a grant that covers the registration fee.

You can find more information about the summer school at <http://www.hipeac.net/summerschool>. We look forward to seeing you there!

Koen De Bosschere, Summer school organizer





BOOK: NEAR THRESHOLD COMPUTING

Technology, Methods and Applications.

This book explores near-threshold computing (NTC), a design-space using techniques to run digital chips (processors) at close to the lowest possible voltage. Readers will be enabled with specific techniques to design chips that are extremely robust; tolerating variability and resilient against errors. Variability-aware voltage and frequency allocation schemes will be presented that will provide performance guarantees, when moving toward near-threshold manycore chips.

- Provides an introduction to near-threshold computing, enabling the reader with a variety of tools to face the challenges of the power/utilization wall;
- Demonstrates how to design efficient voltage regulation, so that each region of the chip can operate at the most efficient voltage and frequency point;
- Investigates how performance guarantees can be ensured when moving towards NTC manycores through variability-aware voltage and frequency allocation schemes.

Editors: Michael Hübner, Cristina Silvano.

URL: <http://link.springer.com/book/10.1007/978-3-319-23389-5>

Cristina Silvano, Politecnico di Milano

PARASUITE: PARALLEL BENCHMARKS FOR MULTI-CORE CPUS, CLUSTERS AND ACCELERATORS



Despite the ubiquity of parallel architectures in all computing segments, the research community often lacks benchmarks representative of parallel applications. The Inria Parallel Benchmark Suite (Parasuite) seeks to address this need by providing a set of representative parallel benchmarks for the architecture, compiler and system research communities. Parasuite targets the main contemporary parallel programming technologies: shared-memory multi-thread parallelism for multi-core, message-passing parallelism for clusters and fine-grained data-level parallelism for GPU architectures and SIMD extensions.

All benchmarks come with input datasets of various sizes, to accommodate use cases ranging from microarchitecture simulation to large-scale performance evaluation. Correctness checks on the computed results enable automated regression testing. In order to support computer arithmetic optimization and approximate computing research scenarios, the correctness checks favour accuracy metrics evaluating domain-specific relevance rather than bit-exact comparisons against an arbitrary reference output.

We encourage members of the HiPEAC community to experiment with Parasuite and use it in their own work. Your feedback is also welcome.

URL: <http://parasuite.inria.fr/>

Sylvain Collange, Inria Rennes

ESPRESSO SOLVER

ESPRESSO is an **ExaScale PaRallel FETI SOLver** developed at IT4Innovations. The main focus of the development team is to create a highly-efficient parallel solver that contains several Finite Element Tearing and Interconnect (FETI) based algorithms including a new Hybrid FETI method. This novel method is designed to run on massively parallel machines with thousands of compute nodes and hundreds of thousands of CPU cores. The algorithm can be seen as a multilevel FETI method designed to overcome the main bottleneck of standard FETI methods, a large coarse problem. This problem arises when solving large problems decomposed into a large number of subdomains. ESPRESSO is also being developed to support modern many-core accelerators.

We are currently developing four major versions of the solver:

- **ESPRESSO CPU** – is a CPU version that uses a sparse representation of system matrices.
- **ESPRESSO MIC** – is an Intel Xeon Phi accelerated version that works with dense representation of system matrices in the form of the Schur complement. This version is developed under the IPCC project awarded to IT4Innovations by Intel.
- **ESPRESSO GPU** – is a GPU-accelerated version that works with dense structures. The support for sparse structures using cuSolver is under development.
- **ESPRESSO GREEN** – is a power-efficient version developed under the H2020 READEX project. This version is in an early development stage.

SCALABILITY OF THE HYBRID FETI LINEAR SOLVER

The ESPRESSO solver is based on a highly-efficient communication layer on top of MPI 3.0 combined with shared memory parallelization inside the node. The communication layer was developed specifically for FETI solvers and uses several state-of-the-art communication hiding and avoiding techniques. These techniques provide better scalability. This layer allows good scaling on massively parallel machines as shown in Figure 1. Based on the observed

scalability of the solver, it is expected to be able to scale to several thousand nodes using hybrid parallelization.

INTEL XEON PHI AND NVIDIA GPU SUPPORT IN ESPRESSO

The ESPRESSO solver can take advantage of manycore accelerators to speedup the solver runtime. To achieve this, it uses a dense representation of sparse system matrices in the form of Schur complements (SC). Computing the SC is time consuming. Recently, new techniques have emerged that significantly reduce the processing time, and they have been implemented in PARDISO sparse direct solver, giving a speedup of up to ten times, depending on the matrix sizes and structure. Even though the SC computation is done only once during the preprocessing stage, it remains the main bottleneck. This approach can still be five to ten times slower than standard Hybrid FETI preprocessing, including “only” the factorization of a system matrix.

The main advantage of the SC approach in FETI solvers is the reduction of the iteration time. Instead of calling a solve routine of the sparse direct solver in every iteration, the solver can use the dense matrix-vector multiplication (GEMV) routine. This solve routine is, however, by nature a sequential operation. The GEMV offers the parallelism required by many-core accelerators and it delivers up to 4x speedup, depending on the hardware configuration.

Due to the additional work in the preprocessing stage, the solver needs to perform a large number of iterations to eliminate this initial penalty. The approach is most suitable for transient or ill-conditioned problems that usually perform hundreds or thousands of iterations.

Lubomír Říha, Tomáš Brzobohatý, IT4Innovations national supercomputing center

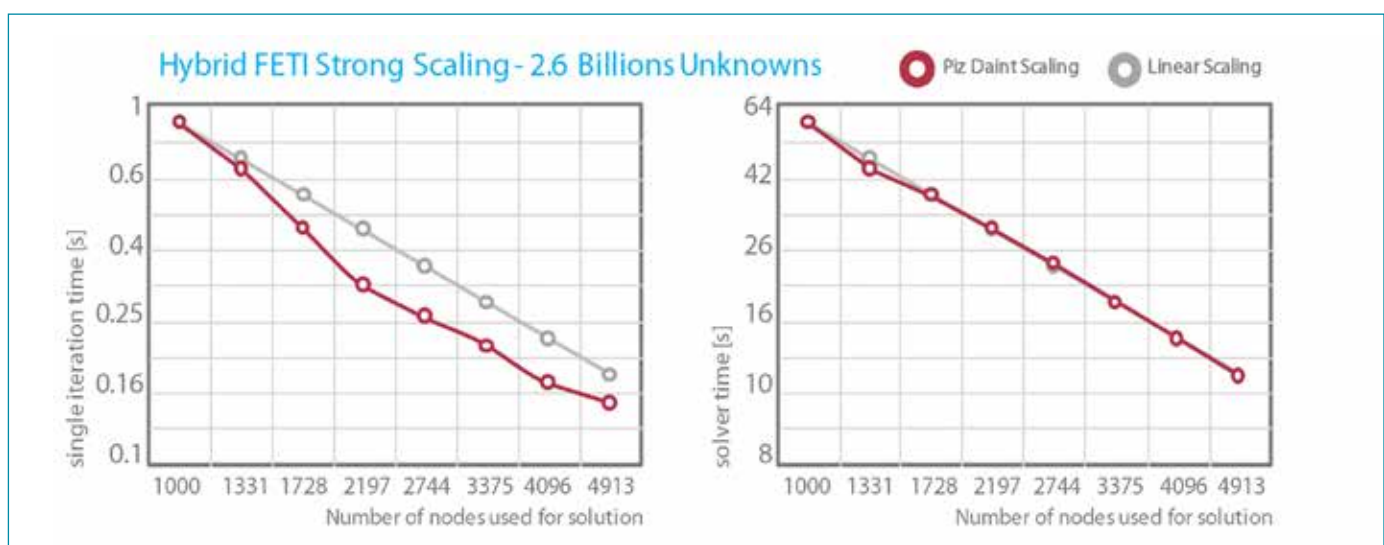


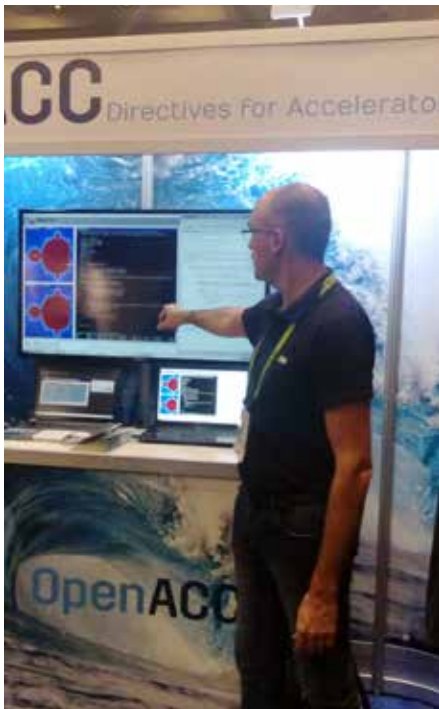
Figure 1. Strong scalability of Hybrid FETI in ESPRESSO running with 1000 to 4913 nodes on the CSCS Piz Daint supercomputer solving 2.6 billion unknowns. The figure shows scalability of both single iteration time (left) and entire solver runtime (right).

HPC STARTUP APPENTRA AT SC15 IN AUSTIN, TEXAS

Appentra is working to give value to the new Parallware tools as a powerful and modern solution to assist developers of parallel scientific programs.

Appentra Solutions is a member of ETP4HPC.

Parallware tools are the result of more than ten years of R+D in advanced compilation techniques for automatic detection of parallelism and automatic generation of parallel-equivalent code for multi-core and many-core computing systems.



SC15 in Austin is the world's largest gathering of HPC professionals. HPC is transforming the world, and we agree with this claim. Appentra wants to be part of this ecosystem and is developing the Parallware tools to help making parallel programming easier.

What were our activities at SC15?

- WACCPD 2015: Presentation of the paper "Experiences in extending Parallware to support OpenACC", a collaboration between Appentra and Oak Ridge National Laboratory (ORNL).
- OpenACC booth talk "Parallware and OpenACC", where Manuel Arenaz (CEO of Appentra) performed a live demo of the Parallware tools and answered technical and non-technical questions.
- OpenMP booth talk "Parallware and OpenMP", where Manuel Arenaz did a live demo of the Parallware tools and answered technical and non-technical questions. The recording of the talk and the slides are available at the OpenMP website: https://www.youtube.com/watch?v=_1kA6DMiTzw
- StartupHPC Conference: StartupHPC is a grass roots community for STEM- and HPC-inspired entrepreneurs, corporations, venture capital firms, academia, government agencies, and support organizations.

It was our first visit to the supercomputing conference, and we can only say that it was a great show. During these three 12-hour days, we managed to see just a quarter of the exhibit booths. Thanks to all of those who gave us the opportunity to present the work that we've been doing all this time. In July the project will be four years old, and SC15 has been an early gift for our team. We will not forget this experience easily: the city, the food, the friendly people, and the perfect organization of the event, especially seeing that we have a clear value proposition for the HPC market. We have confirmed and updated the roadmap to put in value the Parallware technology. We have a lot of work back in the office, but we are full of energy. It is impressive to see so much technology together, scientists and companies around the world trying to solve the great challenges of the society.

URL: <http://www.appentra.com/>

Manuel Arenaz, University of A Coruña & CEO of Appentra, Spain

MATEO VALERO BECOMES THE FIRST EUROPEAN TO RECEIVE THE SEYMOUR CRAY AWARD

Mateo Valero, UPC-Barcelona Tech professor and director of Barcelona Supercomputing Center (BSC), has become the first European researcher to receive the Seymour Cray Supercomputing Award, which is presented by the IEEE Computer Society. The award represents the highest international recognition within the field of high-performance computing.

The IEEE Computer Society presents this award 'in recognition of innovative contributions to high-performance computing systems that best exemplify the creative spirit demonstrated by Seymour Cray'. Mateo Valero was selected 'in recognition of seminal contributions to vector, out-of-order, multithreaded, and VLIW [Very Long Instruction Word] architectures', the society confirmed.

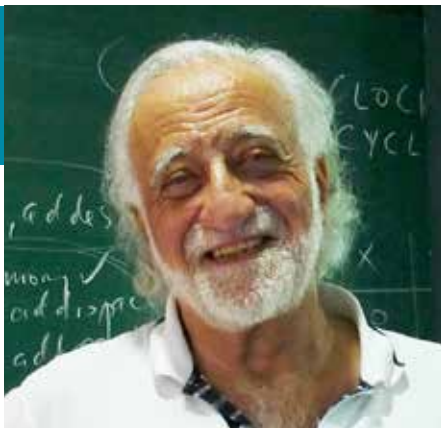
The award was presented during the SC15 (Supercomputing 15) conference in Austin (Texas, United States), the main international conference for high-performance computing, networks, storage and analysis (<http://www.sc15.supercomputing.org>).

Valero's research has focused principally on computer architecture. His work in this field has won him a number of awards, including the Eckert-Mauchly Award, which recognises contributions to digital systems and computer architecture, two national research awards and a European Research Council Advanced Grant for RoMoL, the research project Valero is currently leading at BSC, which centres on the design of multicore chips for the processors of the future.



Mateo Valero is one of the founding members of our HiPEAC Network of Excellence, a contribution that was highlighted during the award ceremony at SC15. The HiPEAC community congratulates Mateo for this great achievement!

Award Ceremony: <https://www.youtube.com/watch?v=E261V-CHetQ>



HIPEAC ASSOCIATED MEMBER RECEIVES THE BENJAMIN FRANKLIN MEDAL

Professor Yale Patt of University of Texas at Austin is the 2016 recipient of the Benjamin Franklin Medal in Computer and Cognitive Science for his pioneering contributions to the design of modern microprocessors that achieve higher performance by automatically identifying computer instructions that can be executed simultaneously.

Yale is, and has been, contributing in a major way to the HiPEAC network. He has given highly appreciated courses and keynote speeches in the successful HiPEAC ACACES summer school series and acted as general chair for the annual conference. He is also annually spending significant time with several prominent research groups in the HiPEAC network including Barcelona Supercomputing Center and Chalmers University of Technology.

The HiPEAC network congratulates Yale for this major recognition.

ALEXANDRA FERRERON AWARDED THE BRONZE MEDAL AT ACM STUDENT RESEARCH COMPETITION

Alexandra Ferreron (Universidad de Zaragoza) has been awarded the bronze medal at the ACM Student Research Competition



Alexandra Ferreron¹, PhD student from the Computer Architecture Group (gaZ) from Universidad de Zaragoza (Spain), and student member of HiPEAC, has obtained the bronze medal at the ACM Student Research Competition² at the Grace Hopper Celebration 2015³ held in Houston, Texas (USA).

Alexandra presented a summary of her PhD work. Her main research line is focused on content-management techniques for on-chip SRAM caches operating at ultra-low or near-threshold

voltages. Operating at very-low voltages has the promise of great energy savings, but it also affects the reliability of SRAM cells, as their functionality margins are reduced, which eventually degrades the system performance. Ferreron proposes to exploit the natural redundancy of the data to mitigate the impact of these SRAM failures with fault-aware content management techniques, relying on the underlying coherence protocol and the replacement policy to minimize performance degradation with very low overhead.

The Grace Hopper Celebration is the world's largest gathering of women technologists, and this year broke a record with over 12,000 attendees from institutions of all around the world, and over 200 technology companies. The ACM SRC of the Grace Hopper Celebration received 118 poster submissions, from which only 28 were accepted into the competition.

Alexandra was also selected as a Grace Hopper Scholar. Scholars receive a grant that covers all travel expenses related to the conference. This year, only 26% of the scholarship applications were accepted.

Jesús Alastruey and Alexandra Ferrerón, Universidad de Zaragoza

¹ <http://webdiis.unizar.es/~ferreron>

² <http://src.acm.org/>

³ <http://ghc.anitaborg.org/>

ICCD BEST PAPER AWARD FOR TUDELFT PAPER

The TUDelft Computer Engineering paper on memristor circuit design wins the Best Paper Award at the 33rd ICCD (International Conference on Computer Design), October 2015, New York City, USA. Lei Xie, Hoang Anh Du Nguyen, Mottaqiallah Taouil, Said

Hamdioui and Koen Bertels demonstrated in their paper, titled Fast Boolean Logic Mapped on Memristor Crossbar, that you can build and execute any logic function on a crossbar of memristors within a constant number of steps independent of

its functionality. The work outperforms the state-of-the-art in the fact that it increases the execution speed (up to 500 times) and reduces the energy (up to 3.7 times).

Koen Bertels, Delft University of Technology

LEONEL SOUSA AND WALID A. NAJJAR SELECTED ACM DISTINGUISHED SCIENTISTS

Two HiPEAC members have been selected as ACM Distinguished Scientists. This distinction awarded by ACM recognizes significant accomplishments or impact within the computing field. The HiPEAC network congratulates Leonel and Walid for this achievement.

<http://www.acm.org/press-room/news-releases/2015/distinguished-2015>



MIGUEL ANGEL AGUILAR AWARDED GOLD MEDAL AT ACM SRC

Miguel Angel Aguilar, PhD student at the RWTH Aachen University, was awarded with the gold medal in the ACM Student Research Competition (SRC) at PACT 2015

The ACM Student Research Competition (SRC) is a unique forum for undergraduate and graduate (Master and PhD) students to present their research at well-known ACM sponsored conferences in front of a panel of judges and attendees. The ACM SRC consists of three phases: i) extended abstract submission, ii) poster presentation, and iii) oral presentation. During the extended abstract phase students submit a written document describing their research to a conference hosting the ACM SRC. Based on the abstracts, a panel of judges invite the most promising works to participate in a poster session. During this poster session, students have the opportunity to present their work to the judges, who select three finalists to advance to the next round. The last round consists of an oral presentation at an important session of the conference to compete for the final award.

Miguel Angel Aguilar, research assistant at the Institute for Communication Technologies and Embedded Systems (ICE) of the RWTH Aachen University, was the winner of the ACM Student Research Competition at the conference PACT 2015. He was awarded with the gold medal for his work on parallelization of sequential embedded software.

The work of Miguel Angel focuses on developing technologies for optimized mapping of embedded applications to Multiprocessor System-on-Chips (MPSoCs), by automatically extracting multiple forms of parallelism hidden in applications. For this purpose Miguel Angel has developed a parallelization framework that takes as inputs a sequential application, and a model of the target MPSoC. Then a model of the application is built by combining static and dynamic analyses. While static analysis is based on

compile-time information, dynamic analysis is based on run-time information. This model is then analyzed by algorithms that expose multiple forms of parallelism, while considering the characteristics of the target MPSoC. The resultant information is used to automatically generate a parallel version of the application, and to provide source-level hints to the developers for an easy understanding of the identified parallelism. This framework has been successfully applied to commercial environments, such as Android-based mobile devices.

The research of Miguel Angel is in close connection with industry, as he is actively collaborating with the ICE spin-off Silexica Software Solutions GmbH, the leading provider of multicore programming tools for the embedded market.

Miguel Aguilar, RWTH Aachen University

SHAOTENG LIU AND ZHONGHAI LU RECEIVE BEST PAPER AWARD AT NOCS 2015

The paper "Highway in TDM NoCs" won the best paper award at NOCS 2015

The paper "Highway in TDM NoCs" authored by Shaoteng Liu, Zhonghai Lu (HiPEAC member) from KTH Royal Institute of Technology, Sweden and Axel Jantsch from Vienna University of Technology, Austria, has received The Best Paper Award at the 9th IEEE/ACM International Symposium on Networks-on-Chip (NOCS) held in Vancouver, Canada from September 28 to 30, 2015. NOCS is the premier forum for researchers to present their latest findings in the area of Networks-on-Chip. It is the premier event dedicated to interdisciplinary research on on-chip, chip-scale, and multichip package-scale communication technology, architecture, design methods, applications and systems.

The paper proposes a technique called highway, which can effectively enhance the bandwidth utilization of TDM (Time Division Multiplexing) connections in NoCs. TDM is a well-known technique to provide QoS guarantees in NoCs. However, it often suffers from low bandwidth utilization due to unallocated and idle (allocated but not used) time slots. A TDM highway is an express TDM connection composed of buffer queues, called highway channels, which can dynamically exploit those unallocated and idle slots. It can enhance throughput and reduce data transfer delay of the TDM connection while keeping TDM's QoS guarantees on minimum bandwidth and in-order packet delivery. The proposed technique has been efficiently implemented in hardware.

Zhonghai Lu, Royal Institute of Technology (KTH)

HIPEAC MEMBERS RECEIVE RECOGNITION FROM INTEL FOR THEIR RESEARCH ON TRANSACTIONAL MEMORY

Researchers from the Computing Systems Laboratory (CSLab) of the National Technical University of Athens (NTUA) have received an honorary recognition from Intel for delivering insightful feedback to the company on Intel's Transactional Synchronization Extensions (TSX).

Intel TSX provides a set of ISA extensions that allows programmers to specify regions of code for transactional synchronization and effectively achieve the performance of fine-grain locking with the programming simplicity and elegance of coarse-grain synchronization. CSLab researchers worked closely with Intel, applying TSX to a large number of application scenarios and stressing the system with a variety of concurrency levels, working sets and data structures. Their insight improved TSX functionality leading to a specification update in December 2014¹.

CSLab has a research tradition spanning more than three decades in the implementation and optimization of large-scale systems focusing on several aspects like system architecture, programming models, run-time systems and applications. CSLab researchers have been working on Transactional Memory and its application to real-life code during the past decade and together with Intel aspire to further strengthen this collaboration.

<http://research.cslab.ece.ntua.gr>

Nectarios Koziris, National Technical University of Athens (NTUA)

1 "Intel Core M Processor Family. Specification Update. December 2014. Revision 003. 330836-003". Intel. December 2014.



Award recipients: Prof. Nectarios Koziris, Dr. Georgios Goumas, Dr. Konstantinos Nikas, Dr. Nikos Anastopoulos, George Mappouras

INTEL INVESTS \$50M IN TU DELFT TO BUILD AN EFFECTIVE QUANTUM COMPUTER

On September 3, 2015, Intel and TU Delft signed a ten-year collaboration agreement. Intel will invest \$50 million in the Fault-tolerant quantum computing roadmap of the Qutech research institute. Intel will also provide technical support as well as facilities to speed up advancements in quantum computing.

Qutech is the quantum research institute of TU Delft and TNO founded two years ago in which physicists and engineers work together towards the goal of building the world's first working quantum computer. It currently has three research and technology roadmaps: Fault-tolerant quantum computing, Topological quantum computing and Quantum Internet. The collaboration with Intel focuses on research into fault-tolerant quantum computing in which professor Koen Bertels, head of the Computer Engineering Laboratory of TU Delft, is involved. The objective of this roadmap is to demonstrate a scalable architecture for simultaneously processing and protecting quantum information. The Computer Engineering Laboratory is responsible for defining the overall system architecture of the quantum computer that consists of the quantum processor controlled by a high performance, classical computer. The latter is necessary to perform quantum error correction and also controls the execution of the quantum circuits in the qubit plane. The architectural work will consist of both

theoretical work as well as building the control electronics for the 7-qubit superconducting quantum processor that is being developed by Prof. L. DiCarlo.

The collaboration with Intel is a long-term one with the final aim of building the first quantum computer in ten to twelve years from now and focuses on the fabrication of qubits, the development of interconnection and low-temperature control electronics and the overall system architecture.

URL: <http://qutech.nl/fault-tolerant-quantum-computing/>

Koen Bertels, Delft University of Technology



IS 3D-STACKED MEMORY THE SOLUTION FOR HPC?

Emerging 3D-stacking technology enables DRAM devices that support much higher bandwidths than traditional DIMMs. The first commercial products Hybrid Memory Cube and High Bandwidth Memory will soon hit the market, and some of the publicity surrounding these emerging memory devices suggests that they will bring significant performance improvements. Barcelona Supercomputing Center (BSC) researchers have analysed how 3D-stacked DRAMs will affect performance of high-performance computing (HPC) applications, and they concluded that simple replacement of conventional DIMMs with 3D-stacked devices may not lead to announced performance improvements. In order to properly exploit the benefits of the novel high-bandwidth memory solutions, BSC computer architects suggest rethinking about the design of the overall computer systems, mainly processors and memory controllers.

This analysis was presented during the MEMSYS 2015 conference as a result of the paper titled Another Trip to the Wall: How Much Will Stacked DRAM Benefit HPC?, written in collaboration with experts from Chalmers University and Lawrence Livermore National Laboratory. This paper

is a third chapter of the memory wall trilogy following Hitting the Memory Wall: Implications of the Obvious (1995) and Reflections on the memory wall (2004). Memory wall refers to the fact that memory latency is so large that most of the time processors are waiting for data from memory. “Technological evolutions and revolutions notwithstanding, the memory wall has imposed a fundamental limitation to system performance for over 20 years”, says Petar Radojkovic, Memory systems team lead at BSC.

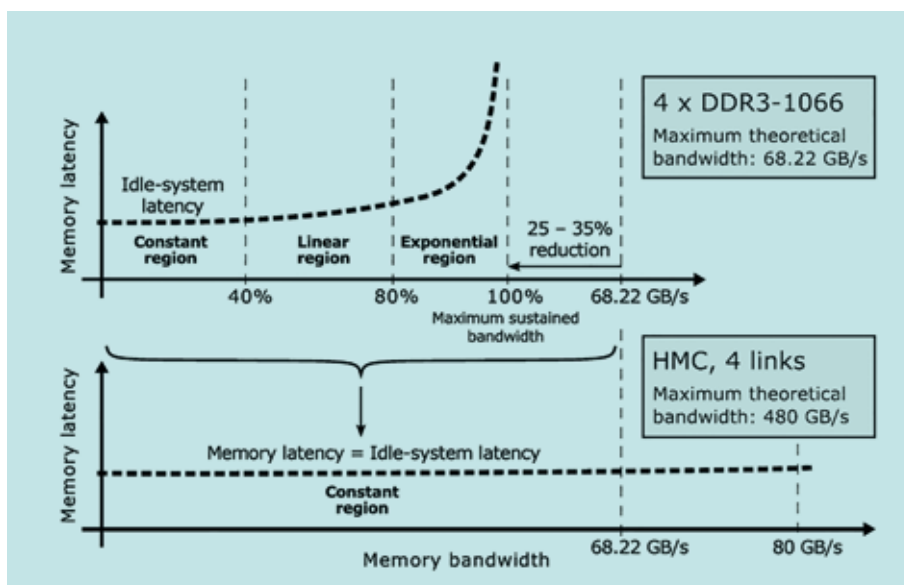
In the paper, BSC experts have recalled that the memory wall has always been defined in terms of main memory latency, not bandwidth. Higher bandwidth may lower memory latency, provided that the selected applications offer sufficient memory-level parallelism (MLP) and that processors can exploit it. But higher bandwidth cannot guarantee better performance because 3D-stacked DRAMs will not reduce idle-system memory latency. Therefore they will not improve the performance of applications with limited MLP. How well the available bandwidth we can be exploited, ultimately depends on the inherent MLP in our targeted workloads.



“Like the initial memory wall paper, this study points out something that most of computer architects “knew” without really understanding. And in order to fully exploit the potential of 3D-stacked DRAMs, we have to really understand what we can and what we cannot expect from these devices”, says Sally A. McKee, Professor at Computer Science Engineering Department of Chalmers University of Technology.

“Also, Hybrid Memory Cube (HMC) and High Bandwidth Memory (HBM) are much more than high-bandwidth memory devices. The logic layers in the HMC and HBM offer possibilities for in-memory processing and sophisticated memory controller functionality. Finding a way to use this innovation to build high-performance systems, however, will take time”, concludes Petar Radojkovic.

This paper is the first outcome of the collaboration between BSC and Samsung Electronics Co., Ltd. that started in 2013 in the context of memory technologies which are in line with Samsung’s high density memory solution including 3D-TSV technology for HPC systems. On one side, the collaboration focuses on analyzing how production HPC applications exercise the current DRAM memory system and evaluating the frequency and locality of memory errors in exiting DDR3 technologies. On the other side, the collaboration pursues the proposal of new architectures and management algorithms to exploit the upcoming non-volatile STT-MRAM memory technologies in HPC systems.



Memory latency depends on the used memory bandwidth, and the memory latency-bandwidth curve has three regions – constant, linear and exponential. Moving from conventional DDRx (upper figure) to high-bandwidth memory solutions (lower figure) will significantly reduce memory latency only for workloads located in the exponential region of the DDRx latency-bandwidth curve.

1st INTERNATIONAL SYMPOSIUM ON MEMORY SYSTEMS (MEMSYS'15)

In today's computing systems, everybody is concerned about the memory system. Thousands of researchers are applying clever ideas to tear down the memory and power walls, but many of these problems are by nature cross-disciplinary. Hence cross-layer solutions, from applications down to the circuit level, are mandatory. Following the strong increase in research in this area, Prof. Bruce Jacob (University of Maryland, USA) has therefore launched a new conference: The International Symposium on Memory Systems (MEMSYS). The goal of this event is to bring together people from industry and academia to discuss memory systems and their future. The symposium was held in the beautiful atmosphere of the old city Alexandria near Washington DC from 5th to 8th October 2015. The conference was organized as a single track of 20-minute paper presentations (regardless of whether the presentation concerned a short or a full research paper), with excellent panel

discussions and a very interesting keynote from J. Thomas Pawlowski (Micron).

Following the sessions it was possible to come into contact with famous researchers and industry representatives to exchange ideas and establish collaborations in a unique environment. In short: MEMSYS was an outstanding event that brought experts from applications, operating systems, system architecture, interconnect and circuit level together.

I hope I will get another opportunity to present a paper at MEMSYS next year.

URL: <http://www.memsys.io>

Matthias Jung, TU Kaiserslautern, <http://ems.eit.uni-kl.de>



Photo supported by Prof. Bruce Jacob

1ST PANDORA SUMMER SCHOOL

The PANDORA summer school on Progression and Diversity of Reconfigurable Architectures and Tools was part of an ongoing collaboration project named TEACHER (Teach AdvanCED Reconfigurable architectures and tools) between Karlsruhe Institute of Technology (KIT) and National Technical University of Athens (NTUA) for developing educational material related to advanced reconfigurable architectures and CAD algorithms. The project is funded through the program "Hochschulpartnerschaften mit Griechenland" from the German Academic Exchange Service, DAAD (<http://www.daad.gr/gr/26028/index.html>). The event was held at the School of ECE, National Technical University of Athens on September 14-18, 2015.

The aim of the event was to provide breakthrough knowledge to students and young researchers related to reconfigurable computing and CAD tools. More specifically, a binational selection of 20

German and Greek participants from the partner universities KIT and NTUA attended lectures, talks and hands-on labs given by experts in the fields of reconfigurable architectures and respective CAD algorithms. The topics within the summer school went beyond the usual lectures in these fields and cover not only established and commercial technologies, but also recent research results and trends. Moreover the focus was on providing the participants with the knowledge, methodologies and tools for developing and exploring their own reconfigurable architectures. We therefore employed the unique TEACHER framework with its virtual laboratory (<http://proteas.microlab.ntua.gr>), which is a perfect playground for quickly developing and evaluating custom FPGAs without the effort and expense of actual physical implementation.

The summer school's program included a design contest among all the participated students. The winners were five students from KIT and five students from NTUA, who were granted with a smart watch sponsored by Texas Instruments. Additionally, the organizers would like to thank the Friedrich-Ebert-Stiftung (FES) for sponsoring the social activities that took place during the PANDORA summer school.

More info: http://proteas.microlab.ntua.gr/teacher/PANDORA2015_SummerSchool.pdf

Kostas Siozios, National Technical University of Athens (NTUA)



THE ANTAREX PROJECT

Autotuning and Adaptivity Approach for Energy Efficient Exascale HPC Systems



The ANTAREX research project, coordinated by prof. Cristina Silvano from Politecnico di Milano, has been granted in the H2020 Future and Emerging Technologies programme on High Performance Computing. The project involves CINECA, the Italian Tier-0 Supercomputing Centre and IT4Innovations, the Czech Tier-1 Supercomputing Center. The Consortium also includes three top-ranked academic partners (ETH Zurich, University of Porto, and INRIA). Industrial partners include one of the Italian leading biopharmaceutical companies (Dompé) and the top European navigation software company (Sygic). Being one of the nineteen research projects in FET-HPC-2014, ANTAREX brings its partners to the forefront of European research in High Performance Computing. The project just started on September 1st, 2015.

The main goal of the ANTAREX project is to provide a breakthrough approach to mapping, runtime management and autotuning of applications for green and heterogeneous High Performance Computing systems up to the Exascale level. The approach will be based on the toolflow shown in Fig. 1. One key innovation of the proposed approach consists of introducing



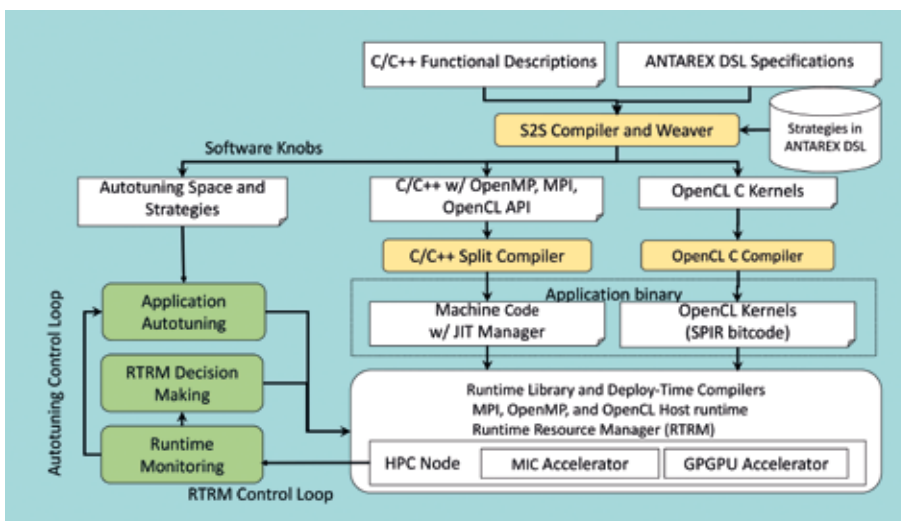
The ANTAREX Group Photo at the Kick-off Meeting held in September 2015 at CINECA (Italy).

a separation of concerns (where self-adaptivity and energy efficient strategies are specified aside from application functionalities) promoted by the definition of a Domain Specific Language (DSL) inspired by aspect-oriented programming concepts for heterogeneous systems. The new DSL will be introduced for expressing adaptivity, energy, and/or performance strategies and to enforce at runtime application autotuning and resource and power management. The goal is to support the parallelism, scalability and adaptability of a dynamic workload by exploiting the full system capabilities (including energy management) for emerging large-scale and extreme-scale systems, while reducing the Total Cost of Ownership (TCO) for companies and public organizations.

The ANTAREX project is driven by two use cases chosen to address the self-adaptivity and scalability characteristics of two highly relevant HPC application scenarios. The two use cases are: (1) a biopharmaceutical HPC application for accelerating drug discovery deployed on the 1.21 PetaFlops heterogeneous NeXTScale Intel-based IBM system at CINECA; and (2) a self-adaptive navigation system to be used in smart cities deployed on the server-side on a heterogeneous Intel-based 1.46 PetaFlops class system provided by IT4Innovations Supercomputing Center. All the key ANTAREX software innovations will be designed and engineered from the beginning to be scaled up to the Exascale level. Performance metrics extracted from the two use cases will be modelled to extrapolate the results towards Exascale systems. These use cases have been selected due to their significance in emerging application trends and thus by their direct economic exploitability and relevant social impact.

URL: <http://www.antarex-project.eu/>

Cristina Silvano, Politecnico di Milano



The ANTAREX tool-flow

NEW H2020 PROJECT: VINEYARD

Versatile Integrated Accelerator-based Heterogeneous Data Centres

H2020 VINEYARD Project

Project name:

VINEYARD: Versatile Integrated Accelerator-based Heterogeneous Data Centres

Project Coordinator: Prof. Dimitrios Soudris, ICCS, GR

Technical Project Management:

Dr. Christoforos Kachris, ICCS, GR

Partners: Institute of Communication and Computer Systems (ICCS), Maxeler Technologies, Bull SAS, Queen's University of Belfast (QUB), Foundation for Research and Technology-Hellas (FORTH), Science and Technology Facilities (STFC), Neurasmus BV, Neurocom Luxembourg, Athens Exchange (ATHEX), Leanxscale SL, Globaz SA

Start date: February 1st, 2016

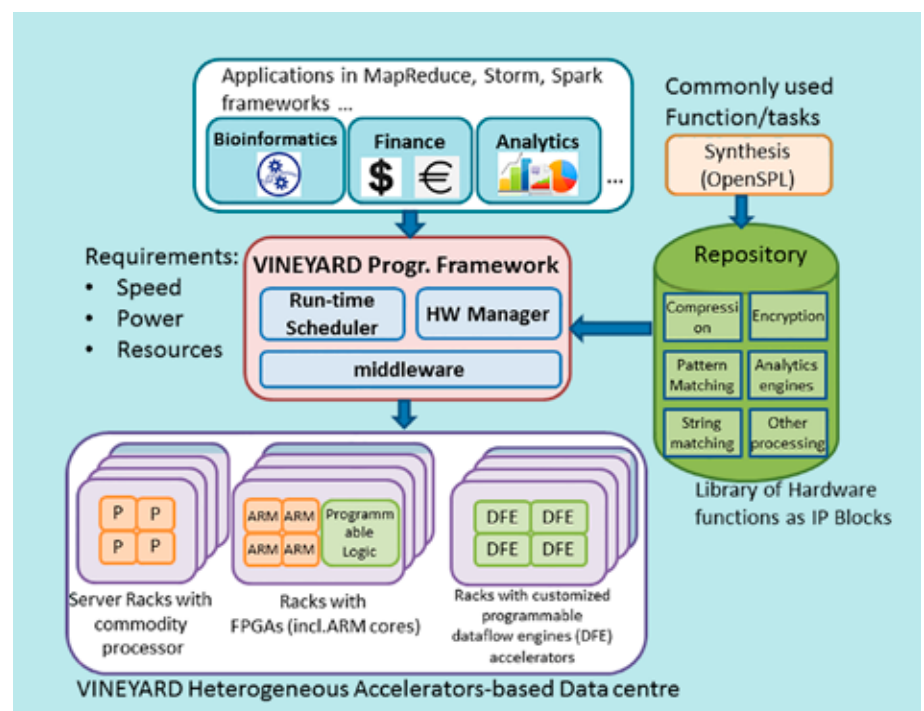
Duration: 36 months

Website: www.vineyard-h2020.eu

A consortium of HiPEAC (ICCS, Maxeler, QUB, FORTH and Neurasmus) and other partners have been granted a new H2020 project on customized and low-power data centers, called VINEYARD. It is worth noting that the consortium was formed largely during the HiPEAC CSWs. VINEYARD will develop an integrated platform for energy-efficient data centres based on new servers with novel, coarse-grain and fine-grain, programmable hardware accelerators. It will also build a high-level programming framework for allowing end-users to seamlessly utilize these accelerators in heterogeneous computing systems using typical data-centre programming frameworks (e.g. MapReduce, Storm, Spark, etc.). VINEYARD will develop two types of energy-efficient servers integrating two novel hardware accelerator types: coarse-grain programmable dataflow engines and fine-grain all-programmable FPGAs that accommodate multiple embedded cores. The former will be suitable for data centre applications that can be represented in dataflow graphs while the latter will be used for accelerating applications that need tight communication between the processor and the hardware accelerators.

Both types of programmable accelerators will be customized based on application requirements, resulting in higher performance and significantly reduced energy budgets. VINEYARD will additionally develop a new programming framework and the required system software to hide the programming complexity of the resulting heterogeneous system based on the hardware accelerators. This programming framework will also allow hardware accelerators to be swapped in and out of the heterogeneous infrastructure so as to offer efficient energy use. VINEYARD will foster the expansion of the soft-IP cores industry, currently limited to embedded systems, to include the data centre market. VINEYARD plans to demonstrate the advantages of its approach on three real use-cases a) a bioinformatics application for high-performance brain simulations, b) two critical financial applications, and c) a big-data analysis application.

*Christoforos Kachris,
National Technical University of Athens
(NTUA)*



THE CTUNING FOUNDATION WORKS WITH ARM TO ACCELERATE COMPUTER ENGINEERING

The non-profit cTuning Foundation has completed a technology transfer to ARM of Collective Knowledge (CK), an extensible framework for systematic and collaborative R&D combined with predictive analytics. Using CK, ARM was able to obtain valuable insights into performance of its products across a wide range of realistic scenarios in a fraction of the time required by conventional analysis.



Supported by the EU TETRACOM Coordination Action, the cTuning Foundation implemented Collective Knowledge and released it at <http://github.com/ctuning/ck> under a permissive open source license. Further development of Collective Knowledge is being coordinated by the cTuning Foundation and sponsored through activities of a start-up company called dividiti.

"Designing next-generation, high-performance, energy efficient computer systems requires a deep understanding of current and emerging real world workloads," said Ed Plowman, director of performance analysis strategy, ARM. "Performance data from systematic analysis of workloads is essential, but it does not by itself produce insights. Collective Knowledge applies leading edge statistical analysis and machine learning to deliver real world performance insights that has the potential to enable ARM to take computer engineering to a whole new level."

Dr Grigori Fursin, Chief Scientist of cTuning Foundation and CTO of dividiti, said: "We are passionate about systematic, collaborative and reproducible R&D. An open source, portable and easy to learn framework, Collective Knowledge dramatically simplifies creating, sharing and reusing knowledge between hardware vendors, software developers, tool providers, and end users. Collective Knowledge, co-developed with our growing interdisciplinary community similarly to Wikipedia, is our contribution to help solve grand challenges including performance and energy modeling, multi-objective optimization, and hardware/software co-design."

"We view Collective Knowledge as a catalyst for stimulating the flow of insights across industry and academia that will lead to breakthroughs in energy efficiency,

performance and reliability of computer systems", said Dr Anton Lokhmotov, CEO, dividiti. "Effective knowledge sharing and open innovation will enable new exciting applications in consumer electronics, robotics, automotive and healthcare at better quality, lower cost and faster time to market."

ABOUT CTUNING FOUNDATION

The cTuning Foundation (<http://ctuning.org>), established in 2008, is a non-profit organization crusading for reproducible and community-driven R&D in computer systems. The Foundation's current activities include coordinating further development of its open source Collective Knowledge framework, building a public repository of representative benchmarks and data sets, developing practical machine learning based techniques for software and hardware optimization, and crowdsourcing multi-objective autotuning. The cTuning Foundation works closely with ACM, IEEE and the broad R&D community to support and improve evaluation of research artifacts at leading conferences including PPOPP and CGO (<http://ctuning.org/ae>).

ABOUT DIVIDITI

dividiti (<http://dividiti.com>) is a UK based start-up founded in 2015 to pursue a vision of efficient and reliable computing everywhere. The key to this vision lies in accelerating computer systems' R&D by combining contributions from hardware vendors, software developers, tool providers and end users, across industry and academia. Collective Knowledge is the flagship effort of dividiti, supported by a growing community of its industry customers and academic collaborators.

Anton Lokhmotov, dividiti

More information: G. Fursin, A. Lokhmotov, E. Plowman. "Collective Knowledge: towards R&D sustainability." To be presented at the "Design, Automation and Test in Europe" conference in March 2016.

THE HOLISTIC PROJECT IN DEPENDABLE COMPUTER ARCHITECTURE

Five Greek university research groups joined forces on Hardware and Software Techniques for Multicore Processor Architectures Reliability Enhancement



The HOLISTIC research project (Hardware and Software Techniques for Multicore Processor Architectures Reliability Enhancement) funded by the European Union and the Greek Ministry of Education in the framework of the “Thales” research program, successfully concluded in November 2015.

Five Greek universities joined forces in the broader area of Dependable Computer Architecture and delivered novel hardware-based and software-based methods for the detection, diagnosis, recovery and repair of modern microprocessor architectures. The project research focused on protection techniques against (transient and permanent) hardware faults and design bugs in major subsystems of a computing

system: CPU cores, cache memories, main memory. In addition to protection techniques for individual hardware structures, methods have been developed to assess the efficiency of each solution, as well as scheduling techniques for the effective coordination of the methods at the multicore architecture level.

The project was coordinated by the University of Athens group of HiPEAC member Prof. Dimitris Gizopoulos; partner groups were led by HiPEAC member Prof. Dimitris Nikolos of the University of Patras, HiPEAC founder Prof. Manolis Katevenis of the Foundation for Research and Technology Hellas, as well as the groups of Prof. Kiamal Pekmestzi of the Technical University of Athens and Prof. Mihalis Psarakis of the University of Piraeus. The four-year project supported fully or in part the research activities of PhD students and post-doc researchers in the institutions including equipment and travel to international conferences.

The results of the HOLISTIC project were published in more than thirty papers in top-class journals, such as the IEEE Transactions on Computer-Aided Design (TCAD), the IEEE Transactions on VLSI Systems (TVLSI), the IEEE Transactions of Dependable and Secure Computing (TDSC) and top-tier computer architecture and design automation conferences: the ACM/IEEE International Symposium on Computer Architecture (ISCA), the ACM/IEEE Symposium on Microarchitecture (MICRO), the ACM/IEEE Design, Automation, and Test in Europe Conference (DATE), the ACM/IEEE Asia and South Pacific Design Automation Conference (ASP-DAC), the ACM FPGA Conference (FPGA), the IEEE VLSI Test Symposium (VTS), the IEEE International Test Conference (ITC), and the IEEE International Conference on Computer Design (ICCD).

*Dimitris Gizopoulos, University of Athens,
HOLISTIC project coordinator*



INTERNSHIP REPORT: PAULO MARTINS

Host Institution: Samsung Research and Development United Kingdom (SRUK)

Title: Key Management for Secure Multi-Media Communication

I am a PhD student at the University of Lisbon (UL) working on the implementation of public-key cryptographic algorithms for embedded systems under the supervision of Professor Leonel Sousa. Thanks to a HiPEAC Industrial PhD Internship, I had the chance to work with the team managed by Mr. Parashuram Chawan at the Samsung Research and Development United Kingdom (SRUK). During my stay there I was able to work firsthand with multiple security technologies, such as Trusted Execution Environment and SE Android; which has enabled me to get a wider view of the world of informatics security, while getting a perspective of how research is conducted in industry. Also, a new bond was built between UL and SRUK. I wish to thank HiPEAC for giving me the opportunity to carry out this collaboration. I would also like to thank Mr. Chawan’s team and the other colleagues at SRUK for their support and for making my stay such a pleasant experience.

Paulo Martins, University of Lisbon



INTERNSHIP REPORT: MARCO BALBONI

Host Institution: ARM Ltd, UK

Title: Transaction-level modelling of next-generation cache-coherent interconnects

My name is Marco Balboni, and I am a PhD student in Computer Architectures/ Science at MPSoC Group of Engineering Department of University of Ferrara (Italy), headed by prof. Davide Bertozzi, and a researcher at MicrelLab of University of Bologna (Italy), headed by prof. Luca Benini.

My research activity is focused on embedded systems and on virtualized, heterogeneous and many-core platforms and in particular on Networks-on-Chip, tackling the challenges of fault tolerance, routing, runtime reconfiguration, resource sharing and partitioning, but also taking into account new technologies such as photonics or wireless, applied to the interconnect fabric.

In order to broaden my knowledge, from October 2014 to January 2015, I was an intern in the HPC-correlation group headed by Andreas Hansson. The internship had

the goal of creating a Transaction-Level Model (TLM) of next-generation cache-coherent interconnects in the gem5 full-system simulator, mainly developed by ARM. As part of the project, the model would be correlated and evaluated using a wide range of mobile and server workloads. More precisely, first of all I executed several benchmarks on the Juno Versatile Express board, to let me compare the results from the gem5 simulator. I used public benchmarks, part of the LMBench set, in particular the ones that test performance in terms of bandwidth and latency. Then I studied and modified the models of caches and the coherent/non-coherent crossbar used as interconnection between the IP blocks of the target system. I managed the latencies, clarifying their usage inside the model, and added some new delays and features to the crossbar, to let the behaviour of the simulator be as close as possible to the real hardware reference platform.

During the internship period I improved my programming skills and also started working on the gem5 simulator, which I am now also using for my research. Furthermore, I enriched my knowledge of cache-coherence and transactions involving memories in a real system. Finally, I created some patches to optimize the simulator and improve its correlation with the real hardware.

I would like to thank HiPEAC for providing me the opportunity to visit ARM during the internship. I am also thankful to all the colleagues from ARM, for the support, assistance and friendship during my period there. In particular my manager Radhika, my boss Andreas H., and all the great persons of my group (Rekai, Renè, Roxana, Sascha, Andreas S., Omar, Stephan and all the others I am missing): it was a fantastic experience!

Marco Balboni, University of Ferrara



COLLABORATION GRANT REPORT: JAVIER JALLE

Host Institution: ESTEC (European Space Agency)

Title: Architectural solutions for the timing predictability of multicore processors in real-time systems.

Barcelona Supercomputing Center (BSC) and the European Space Agency (ESA) have been recently involved in a collaborative internship, focusing on the use of multicore processors for safety-critical systems. I spent three months at the Microelectronics section at ESTEC (ESA), working under the direction of Luca Fossati. During my HiPEAC internship I obtained knowledge to understand ESA software requirements and make use of the available technology at ESA. This allowed me to validate simulation tools and perform real experiments leading to two new paper submissions. Collaboration between both institutions will continue.

Javier Jalle, Barcelona Supercomputing Center

SUPPORTING GENERAL DATA STRUCTURES AND EXECUTION MODELS IN RUNTIME ENVIRONMENTS



Javier Fresno
Institution: Universidad de Valladolid
Advisor: Dr. Arturo González-Escribano
Graduation date: September 2015

The goal of this Ph.D. thesis is to create a runtime system for a generic parallel programming framework. For this, we address two common problems in parallel computing: unified support for dense and sparse data, and integration of data-mapping and dataflow parallelism. We propose a solution that decouples data representation, partitioning, and layout from the programmer's parallel strategy

and algorithmic decisions. Moreover, we introduce a new programming model based on the dataflow paradigm, where different activities can be arbitrarily linked, forming generic but structured networks that represent the overall computation.

ONLINE AUTO-TUNING FOR PERFORMANCE AND ENERGY THROUGH MICRO-ARCHITECTURE DEPENDENT CODE GENERATION



Fernando Endo
Institution: CEA
Advisors: Henri-Pierre Charles and Damien Couroussé
Graduation date: September 2015

Energy consumption is limiting the performance growth experienced in the last decades. New architectural and micro-architectural designs improve the energy efficiency of hardware, thanks to hardware specialization and core heterogeneity. Hence, software needs to cope with the lack of performance portability.

compilation and explore auto-tuning possibilities that are data-dependent. We demonstrated that our run-time auto-tuning of SIMD instructions to in-order cores can outperform a reference vectorized code run in similar out-of-order cores.

This thesis proposes a run-time auto-tuning framework for embedded systems. The proposed framework can both adapt code to a micro-architecture unknown prior to

The thesis is available at:
<http://bit.ly/iND5pqx>

SCALABLE AND BANDWIDTH-EFFICIENT MEMORY SUBSYSTEM DESIGN FOR REAL-TIME SYSTEMS



Manil Dev Gomony
Institution: Eindhoven University of Technology
Advisor: Prof. Kees Goossens and Dr. Benny Akesson
Graduation date: September 2015

Dynamic Random Access Memory (DRAM) is shared between the processing cores in heterogeneous multi-processor platforms that run applications with mixed time-criticality. To support the ever increasing number of applications with dynamic and diverse real-time requirements, the memory subsystem must be configurable and scalable, while keeping the area usage, power consumption and latency to the minimum. On the other hand, while designing memory subsystems the system designer has to make design choices of several system-level parameters such that

the memory is efficiently utilized. This thesis proposes a scalable real-time memory subsystem architecture and an automated design-flow for bandwidth-efficient design of the memory subsystem. The proposed architecture can be configured with multiple arbitration policies and it allows efficient use of multi-channel memories.

CROSS-LAYER RAPID PROTOTYPING AND SYNTHESIS OF APPLICATION-SPECIFIC AND RECONFIGURABLE MANY-ACCELERATOR PLATFORMS



Dionysios Diamantopoulos
Institution: National Technical
 University of Athens
Advisor: Prof. Dimitrios Soudris
Graduation date: September 2015

The future reserves the connection and interaction of IT/communications systems to the natural world, delimiting the transition to natural cyber systems. Such a transition will be supported by computing platforms that incorporate an increased systemic complexity.

This dissertation addresses emerging design challenges by developing methodologies and hardware/software co-design tools that enable the rapid synthesis of efficient architectures, i.e. a) virtualization

methodologies that accelerate the design flow for FPGAs and ASICs, b) many-accelerator heterogeneous architectural templates for energy-efficient computing and c) multi-objective synthesis techniques, both at a high abstraction level of programming (HLS) and at the physical silicon level.

PERFORMANCE PREDICTION: ANALYSIS OF THE SCALABILITY OF PARALLEL APPLICATIONS



Javier Panadero
Institution: Universitat Autònoma
 de Barcelona
Advisor: Dr. Emilio Luque
Graduation date: September, 2015.

Due to the complex interaction between the MPI applications and the HPC system, many applications may suffer performance inefficiencies when they scale to a large number of processes.

As the main contribution, we propose the methodology P3S (Prediction of Parallel Program Scalability), which allows us to analyze and predict strong scalability behavior for message-passing applications on a given system. The methodology strives to use a bounded analysis time, and a reduced set of computing resources, to

predict the application behavior at the large scale.

The output of the P3S methodology will be the predicted curve of application speedup. Using this information, the users can select the most appropriate resources to execute their applications on the target system, in order to use the system resources efficiently.

PARALLEL APPROACHES TO SHORTEST-PATH PROBLEMS FOR MULTILEVEL HETEROGENEOUS COMPUTING



Hector Ortega-Arranz
Institution: University of
 Valladolid
Advisor: Dr. Diego R. Llanos
 Ferraris and Dr. Arturo
 Gonzalez-Escribano
Graduation date: October 2015

There are graph algorithms that give solutions to the problem of finding shortest paths. These problems are key within the combinatorial optimization context due to their multiple real-world applications. The scientific community is increasingly interested in such graph algorithms, not only due to their wide-applicability, but also because they can be efficiently implemented using current parallel computing.

The emergence of new parallel programming models, together with modern GPUs, has enriched the performance of existing parallel algorithms, and has promoted the creation of new algorithms that

are even more efficient. The joint use of both GPUs and CPUs provides the perfect tool to face the most costly problems of shortest-path computing.

My Ph.D. thesis addresses both mentioned fields, through: (a) development of new GPU-based approaches to shortest-path problems, along with studies of GPU optimal configurations; and (b) design of solutions combining parallel and sequential algorithms for heterogeneous environments.

STATISTICAL COMPRESSION CACHE DESIGNS

Angelos Arelakis
Institution: Chalmers University
of Technology
Advisor: Prof. Per Stenström
Graduation date: October 2015

This thesis proposes architectural support for Huffman-based statistical compression caches. Statistics acquisition is handled by a hardware sampling mechanism, as value locality is stable over long time-periods. The cache is extended to adopt a novel floating-point specific compression method that compresses the semantic bit-fields in isolation. As cache data are of diverse types, multiple type-specific compression methods

are combined into a hybrid one that selects the best method through heuristics. We evaluate our designs and show 4x compression ratio. This offers speedups of 25% in multicore systems for cache-intensive workloads, with 30% lower energy overheads than a 4x larger cache



AUTOMATED DESIGN OF DOMAIN-SPECIFIC CUSTOM INSTRUCTIONS

Cecilia González-Álvarez
Institution: Ghent University
Advisors: Prof. Lieven Eeckhout,
 Prof. Daniel Jiménez-González
 and Prof. Carlos Álvarez.
Graduation date: November 2015

To accelerate an application, we can extend a processor with custom instructions (CIs). However, if application-specific CIs are scarcely used, their benefits will not compensate for the design cost. Domain-specific CIs target multiple applications, increasing their applicability but adding complexity. In this thesis, we explore automated methods to design CIs that improve the performance and energy efficiency of a domain. With a canonical representation at the basic block level, as well as domain-specific heuristics, we identify equivalent CIs that can accelerate different programs.

We observe that, with a limited specialization chip area, a mix of application and domain-specific CIs results in the best speedup. To improve the reusability of CIs, we create a canonical representation across basic blocks, and we introduce clustering-based partial matching to identify partially-similar CIs. At small chip areas, our CIs improve the energy-efficiency and performance of a domain.



INSTRUCTION-SET ARCHITECTURE SYNTHESIS FOR VLIW PROCESSORS

Roel Jordans
Institution: Eindhoven University
of Technology
Advisor: Prof. Dr. H. Corporaal and
 Dr. L. Jozwiak
Graduation date: December 2015

This dissertation presents my work on automatic data-path synthesis for very-long instruction-word (VLIW) application-specific instruction-set processors (ASIPs). First the required issue-width is estimated and an initial instruction-set of the VLIW processor is proposed based on the target application. This initial VLIW based processor architecture can then be further refined using one of several presented exploration strategies. The refinement process is organized in such a way that it allows for a very time efficient estimation of the energy consumption and temporal performance of the proposed architectures

through the use of our BuildMaster framework for intermediate exploration result caching. Through this method we manage to avoid over 90% of the traditionally required simulation time by reusing profile information from previously evaluated candidate architectures. This approach greatly increases the number of candidate architectures that can be evaluated within reasonable time.



ARCHITECTURAL AND RUNTIME ENHANCEMENTS FOR DYNAMICALLY CONTROLLED MULTI-LEVEL CONCURRENCY ON GPUS



Yash Ukidave
Institution: Northeastern
University
Advisor: Prof. David R. Kaeli
Graduation date: December 2015

GPUs have gained tremendous popularity as accelerators for a broad class of applications belonging to a number of important computing domains such as mobile SoCs and cloud engines. There is a growing need to provide improved utilization of compute resources and increased application throughput. We present new techniques that address design challenges for supporting concurrent execution of multiple application contexts on the GPU. We design a hardware/software-based mechanism to

enable multi-context execution, with QoS-aware resource allocation, while preserving adaptive multi-kernel execution. We also provide a machine learning based, interference-aware task scheduling mechanism on GPU clusters. We observe an overall improvement in system throughput of 37% when compared to state-of-the-art multi-tasking schemes on GPUs.

MPI LAYER TECHNIQUES TO IMPROVE NETWORK ENERGY EFFICIENCY



Branimir Dickov
Institution: Universitat Politècnica
de Catalunya
Advisor: Prof. Eduard Ayguadé,
Dr. Paul Carpenter and Dr. Miquel
Pericàs
Graduation date: December 2015

This thesis contemplates two directions for power savings in an HPC interconnection network. The first approach uses MPI data compression to save energy during communication phases. When compression of MPI data is possible, the link bandwidth is reduced in accordance with the compression rate, reducing link energy without incurring a performance penalty. The second approach shifts the links into low-power mode during computation phases, while

they are unused, thereby saving link energy. Here link wake-up latencies need to be considered to avoid a loss in performance. We propose a mechanism that accurately predicts when links are idle, allowing them to be switched to more power efficient mode.

UPCOMING EVENTS

22ND IEEE INTERNATIONAL SYMPOSIUM ON HIGH PERFORMANCE COMPUTER ARCHITECTURE (HPCA-22)

12-16 March 2016, Barcelona, Spain

<http://hpca22.site.ac.upc.edu/>

21st ACM SIGPLAN SYMPOSIUM ON PRINCIPLES AND PRACTICE OF PARALLEL PROGRAMMING (PPoPP 2016)

12-16 March 2016, Barcelona, Spain

<http://conf.researchr.org/home/PPoPP-2016>

2016 INTERNATIONAL SYMPOSIUM ON CODE GENERATION AND OPTIMIZATION (CGO 2016)

12-16 March 2016, Barcelona, Spain

<http://cgo.org/cgo2016/>

25TH INTERNATIONAL CONFERENCE ON COMPILER CONSTRUCTION (CC 2016)

12-16 March 2016, Barcelona, Spain

<http://cc2016.eew.technion.ac.il/>

DESIGN, AUTOMATION AND TEST IN EUROPE (DATE'16)

14-18 March, 2016, Dresden, Germany

<http://www.date-conference.com/>

ARCS 2016 - ARCHITECTURE OF COMPUTING SYSTEMS

4-7 April 2016, Nuremberg, Germany

<https://www3.cs.fau.de/arcs2016>

22ND IEEE REAL-TIME EMBEDDED TECHNOLOGY & APPLICATIONS SYMPOSIUM

11-14 April 2016, Vienna, Austria

<http://2016.rtas.org/>

23RD RECONFIGURABLE ARCHITECTURES WORKSHOP

23-24 May 2016, Chicago, Illinois USA

<http://raw.necst.it/>

INTERNATIONAL CONFERENCE ON EMBEDDED COMPUTER SYSTEMS: ARCHITECTURES, MODELING AND SIMULATION (SAMOS XVI)

18-21 July 2016, Samos Island, Greece

<http://samos-conference.com/>

19TH EUROMICRO CONFERENCE ON DIGITAL SYSTEM DESIGN

31 August - 2 September 2016, Limassol, Cyprus

<http://dsd2016.cs.ucy.ac.cy>

contributions If you are a HiPEAC member and would like to contribute to future HiPEAC newsletters, please visit <https://www.hipeac.net/publications/newsletter/>



HiPEAC INFO IS A QUARTERLY NEWSLETTER PUBLISHED BY THE HiPEAC NETWORK OF EXCELLENCE, FUNDED BY THE 7TH EUROPEAN FRAMEWORK PROGRAMME (FP7) UNDER CONTRACT NO. FP7/ICT 287759
WEBSITE: [HTTPS://WWW.HIPEAC.NET/](https://www.hipeac.net/)
SUBSCRIPTIONS: [HTTPS://WWW.HIPEAC.NET/PUBLICATIONS/NEWSLETTER/](https://www.hipeac.net/publications/newsletter/)