Published on DATE 2017 (https://past.date-conference.com/date17)

Home > Printer-friendly PDF > Printer-friendly PDF

Printer-friendly PDF

Goto Session:

- 1.1 Opening Session: Plenary, Awards Ceremony & Keynote Addresses
- UB01 Session 1
  - 2.1 Executive Panel: The Electronics Innovation Landscape: Opportunities, Challenges and Strategies
  - 2.2 Stochastic, Approximate and Neural Computing
  - 2.3 Cache memory management for performance and reliability
  - 2.4 Performance and Power Analysis
  - 2.5 Reliability and Energy-Efficiency: Two Pillars of NoC Design
  - 2.6 Advancing Test for Mixed-Signal and Microfluidic Circuits and Systems
- 2.7 EU Project Special Session: from Secure Clouds to reliable and variable HPC
- 2.8a Smart Medical Devices
  - 2.8b Smart Medical Devices, Part 2
- UB02 Session 2
- 3.0 LUNCH TIME KEYNOTE SESSION: Precision Medicine: Where Engineering and Life Science meet
- 3.1 IT&A Session: Parallel Ultra-Low-Power Computing for the IoT: Applications, Platforms, Circuits
- 3.2 IT&A Session: New Benchmarking Vectors for Emerging Devices, Circuits, and Architectures: Energy, Delay, and ... Accuracy
- 3.3 Hardware Trojans and Fault Attacks
- 3.4 Guardbanding and Approximation
- 3.5 Low-power brain inspired computing for embedded systems
- 3.6 Mechanisms for hardware fault testing, recovery and metastability management
- 3.7 Scheduling and Optimization
- 3.8 Addressing Challenges in Today's Datacenter Systems' Design
- 3.9 A tribute to Ralph Otten
- UB03 Session 3
- 4.1 IT&A Session: The Emergence of Silicon Photonics: From High Performance Computing to Data Centers and Quantum Computing
- 4.2 Logic, Interconnects, Neurons: New Realizations
- 4.3 Efficient memory design
- 4.4 From functional validation to functional qualification
- 4.5 Hot Topic Session: On How to Design and Manage Exascale Computing System Technologies
- 4.6 Fault modeling, test generation and diagnosis
- 4.7 Process variation management for today's and tomorrow's computing
- 4.8 CV Fair DATE 2017
- UB04 Session 4
- 5.1 IoT Day: IoT Perspectives
- 5.2 Emerging Computer Paradigms
- 5.3 Hot Topic Session: I'm Gonna Make an Approximation IoT Can't Refuse - Approximate Computing for Improving Power Efficiency of IoT and HPC
- 5.4 Solutions for efficient simulation and validation
- 5.5 Hot Topic Session: Spintronics-based Computing
- 5.6 Reuse and Integration of Test, Debug, and Reliability Infrastructure
- 5.7 Schedulability Analysis
- UB05 Session 5
- 6.1 IoT Day Hot Topic Session: IoT Enabling Technologies
- 6.2 IT&A Session: Panel: Ultra-Low-Power (ULP) Autonomously Powered Systems
- 6.3 Security Primitives
- 6.4 High-performance Reconfigurable Computing
- 6.5 Hot Topic Session: Memristor for Computing: Myth or Reality?
- 6.6 Industrial Experiences & EU Projects
- 6.7 Model-Based Design and Verification of Real-Time Systems
- 6.8 HiPEAC: European Network on High Performance and Embedded Architecture and Compilation
- UB06 Session 6
- 7.0 LUNCH TIME KEYNOTE SESSION
- UB07 Session 7
- 7.1 IoT Day Hot Topic Session: IoT Deployment
- 7.2 In-memory Computing and Security for Non-volatile Memory Technologies
- 7.3 Optimizing performance, energy and predictability via hardware/software codesign
- 7.4 Advances in Logic Synthesis
- 7.5 Hot Topic Session: The Engineering Challenges for Quantum Computing
- 7.6 Memory Reliability: Modeling and Mitigation
- 7.7 Resource management and analysis for embedded architectures
- 7.8 Smart Energy and Self-Powered Devices
- UB08 Interactive Presentations
<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:30</td>
<td>1.1.1</td>
<td>WELCOME ADDRESSES</td>
<td>David Atienza¹ and Giorgio Di Natale²</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>DATE 2017 General Chair, EPFL, CH; DATE 2017 Programme Chair, LIRMM, FR</td>
</tr>
<tr>
<td>08:45</td>
<td>1.1.2</td>
<td>PRESENTATION OF DISTINGUISHED AWARDS</td>
<td></td>
</tr>
</tbody>
</table>
09:15 1.1.3 KEYNOTE: DESIGN AUTOMATION IN THE ERA OF AI AND IOT: CHALLENGES AND PITFALLS
Speaker: Arvind Krishna, IBM Research, US
Abstract
The AI and IoT revolutions are twin phenomena that are reshaping business models, industries, and society. If we are to maximize their potential, we must overcome significant technical challenges with the help of the Design Automation and Test Community.

First, new computer architectures are required to accelerate solutions driven by cognitive computing, the term given to a comprehensive set of AI capabilities that includes not just machine learning but also data ingestion, data privacy, learning, reasoning, natural language, and conversation. These architectures must support each of these new technologies and manage extreme, cognitive workloads marked by unprecedented volumes of structured and unstructured data. This challenge poses important questions for the Design Automation and Test community about what new approaches can be taken.

A similar challenge is inherent in the rapid development of IoT, where the span of computing architecture varies from extremely low power constraints, limited bandwidth, and sporadic access at the “edge” of the network to the nearly infinite power and compute of data centers. This raises the question of how to maximize the design and placement of IoT systems, which will have to function for extended periods of time (up to ten years or more, like a pacemaker). Unlike smartphones, these systems can’t simply be disposed of, which raises significant security concerns.

In his talk exploring these challenges, Dr. Krishna will emphasize that solutions can only come from an integrated hardware-software co-design approach. He will also highlight some of the leading-edge technologies IBM Research is developing to drive further innovation in the computing stack as the era governed by Moore’s law comes to a close.

1.1.4 KEYNOTE: A NEW ERA OF HARDWARE MICROSERVICES IN THE CLOUD
Speaker: Doug Burger, Microsoft Research, US
Abstract
The Cloud is causing a major shift in both the business ecosystem and system infrastructures. The major hyperscale providers are building out highly-interconnected, worldwide computers at a scale that allows them to make significant first-party investments. This verticalization allows them to make cross-layer architectural changes more rapidly than would the old horizontal model. A second trend is the emergence of ultra-low latency requirements in the Cloud, moving storage, networking, and services from the millisecond to the microsecond regime. In this talk, I will describe how these architectural shifts are enabling the emergence of specialized hardware in datacenters, that enable services to be operated in the microsecond regime. On FPGAs, GPUs, and ASICs, these services can run with no CPU intervention, allowing much lower latencies and better cost structures than previously possible for key services such as deep learning. Over time this transition will enable a much broader collection of hardware IP to run at scale in the Cloud.

10:30 End of session
Coffee Break in Exhibition Area

On all conference days (Tuesday to Thursday), coffee and tea will be served during the coffee breaks at the below-mentioned times in the exhibition area.

Tuesday, March 28, 2017
- Coffee Break 10:30 - 11:30
- Coffee Break 16:00 - 17:00

Wednesday, March 29, 2017
- Coffee Break 10:00 - 11:00
- Coffee Break 16:00 - 17:00

Thursday, March 30, 2017
- Coffee Break 10:00 - 11:00
- Coffee Break 15:30 - 16:00

ÜB01 Session 1
Date: Tuesday 28 March 2017
Time: 10:30 - 12:30
Location / Room: Booth 1, Exhibition Area

<table>
<thead>
<tr>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>ÜB01.1 NOXIM-XT: A BIT-ACCURATE POWER ESTIMATION SIMULATOR FOR NOCS</td>
<td>Pierre Bomel, Université de Bretagne Sud, FR</td>
<td>André Rossi¹, Johann Laurent² and Erwan Moreac²</td>
</tr>
</tbody>
</table>
 ¹LERIA, Université d'Angers, Angers, France, FR; ²Lab-STICC, Université de Bretagne Sud, Lorient, FR |
Abstract
We have developed an enhanced version of Noxim (Noxim-XT) to estimate the energy consumption of a NoC in a SOC. Noxim-XT is used in a two-step methodology. First, applications are mapped on a SoC and their traffics are extracted by simulation with MPSOcBench. Second, Noxim-XT tests various hardware configurations of the NoC, and for each configuration, the application’s traffic is re-injected and replayed, an accurate performance and power breakdown is provided, and the user can choose different data coding strategies. With the help of Noxim XT, each configuration is bit-accurately estimated in terms of energy consumption. After simulation, a spatial mapping of the energy consumption is provided and highlights the hot-spots. Moreover, the new coding strategies allows significant energy saving. Noxim XT simulations and a FPGA-based prototype of a new coding strategy will be demonstrated at the U-booth to illustrate these works.

More information ...
Abstract

The unceasing shrinking process of CMOS technology is leading to its physical limits, impacting several aspects, such as performances, power consumption and many others. Alternative solutions are under investigation in order to overcome CMOS limitations. Among them, the memristor is one of promising technologies. Several works have been proposed so far, describing how to synthesize boolean logic functions on memristor-based crossbar architecture. However, depending on the synthesis parameters, different architectures can be obtained. In this demo, we show a Design Space Exploration (DSE) that we use to select the best crossbar configuration on the FPGA, which then interacts with the Just-In-Time compiler executing the program. The overall process requires hundreds of microseconds, and can be easily reverted should the outcome be unsatisfactory.

More information ...
UB01.8 MTA: MANCHESTER THERMAL ANALYZER
Presenter: Scott Ladeneim, University of Manchester, GB
Authors: Yi-Chung Chen, Vasili Pavlidis and Milan Mihajlovic, University of Manchester, GB
Abstract: The Manchester Thermal Analyzer (MTA) is a fast thermal analysis tool to compute temperature profiles of integrated circuits (ICs) in 3-D. The thermal simulations use the finite element method to discretize the heat equation in space coupled to an implicit time-integration method and are implemented with the open-source C++ library deal.ii. The MTA supports higher-order elements, several time-integration methods, and fully adaptive spatiotemporal refinement. State-of-the-art preconditioned iterative methods solve the linear systems arising from the discretized equations as efficiently as possible. Using shared memory parallelization, the MTA solves systems on the order of tens of millions enabling modeling ICs at the cell-level. We present a thermal simulation of an Intel Xeon processor within a FCLGA package with heatsink to show the diverse structures of modern ICs the MTA simulates. The MTA also models other 3-D structures such as bonded tiers, TSVs, heatsinks, and heat spreaders.
More information ...

UB01.9 SEFILE: A SECURE FILESYSTEM IN USERSPACE VIA SECUBE™
Presenter: Giuseppe Arofanulla, CINI, IT
Authors: Paolo Prinetto1 and Antonio Varriale2
1CINI & Politecnico di Torino, IT; 2Blu5 Labs Ltd., IT
Abstract: The SECube™ Open Source platform is a combination of three main cores in a single-chip design. Low-power ARM Cortex-M4 processor, a flexible and fast Field-Programmable-Gate-Array (FPGA), and an EAL5+ certified Security Controller (SmartCard) are embedded in an extremely compact package. This makes it a unique Open Source security environment where each function can be optimized, executed, and verified on its proper hardware device. In this demo, we present a Windows wrapper for a Filesystem in Userspace (FUSE) with an HDD firewall resorting to the hardware built-in capabilities, and the software libraries, of the SECube™.
More information ...

UB01.10 PER: METHOD AND TOOL FOR ANALYZING THE INTERPLAY BETWEEN PERFORMANCE, ENERGY AND SCALING IN MULTI- AND MANY-CORE PLATFORMS
Presenter: Fei Xia, Newcastle University, GB
Authors: Ashur Rafiev, Alexander Romanovsky and Alex Yakovlev, Newcastle University, GB
Abstract: Parallelization has been used to maintain a reasonable balance between energy consumption and performance in computing systems. However, the effectiveness of parallelization scaling is different for different hardware platforms. This is because the reliable operation region (ROR), a region defined in the voltage-throughput space for any hardware platform, is platform-dependent and its shape determines how effective parallelization scaling is in improving throughput and/or reducing power consumption. Although many of the interlinked issues are known, a unifying analysis method has just now been proposed to study the interplay between performance, energy, reliability and parallelization scaling. The method of bi-normalization of the ROR is designed to help achieve a meaningful cross-platform analysis of this interplay. The PER tool brings all these issues together and helps designers reason about hardware parallelization, DVFS and software parallelizability.
More information ...

2.1 Executive Panel: The Electronics Innovation Landscape: Opportunities, Challenges and Strategies
Date: Tuesday 28 March 2017
Time: 11:30 - 13:00
Location / Room: Auditorium A
Chair: Alberto Sangiovanni-Vincentelli, UCB, US

From autonomous driving to big data, from machine learning to cyber-physical systems, from robotics to the internet of everything, from brain-machine interfaces to the human intranet, innovation is moving at a pace that has never been seen before. To face the large investments and increasing global competition, mergers and acquisitions have sped up in all areas including the semiconductor industry that has been possibly the most decisive enabling factors of these disruptive technologies. The panel will address what are the structural factors to sustain innovations and what are the strategies that some of important actors in the industrial and research sector are embracing. The panel will also add the opportunities and difficulties of the different regions of the world in the changing social and economic landscape. The panel will begin with an introductory presentation about the state of technology and innovations in the areas outlined above. Then executives from IBM, ST Microelectronics and Leti will address the problems to face and the strategies to embrace in a challenging competitive landscape.

Panelists:
- Arvind Krishna, Sr. VP, Head of Research, IBM, US
- Marie-Noëlle Semeria, CEO, CEA/Leti, FR
- Benedetto Vigna, EVP & GM, Analog & MEMS Group, STMicroelectronics, IT

13:00 End of session
Lunch Break in Garden Foyer
Keynote Lecture session 3.0 in "Garden Foyer" 1350 - 1420

2.2 Stochastic, Approximate and Neural Computing
Date: Tuesday 28 March 2017
Time: 11:30 - 13:00
Location / Room: 4BC

The method of bi-normalization of the ROR is designed to help achieve a meaningful cross-platform analysis of this interplay. The PER tool brings all these issues together and helps designers reason about hardware parallelization, DVFS and software parallelizability.

End of session
Lunch Break in the Garden Foyer
Keynote Lecture session 3.0 in "Garden Foyer" 1350 - 1420

Lunch Break in the Garden Foyer
On all conference days (Tuesday to Thursday), a buffet lunch will be offered in the Garden Foyer, in front of the session rooms. Kindly note that this is restricted to conference delegates possessing a lunch voucher only. When entering the lunch break area, delegates will be asked to present the corresponding lunch voucher of the day. Once the lunch area is being left, re-entrance is not allowed for the respective lunch.
Stochastic and approximate computing is an approach developed to improve energy efficiency of computer hardware. First paper presents a framework for quantifying and managing accuracy in stochastic circuits design. Second paper deals with a new approximate multiplier design. Energy efficient hybrid stochastic-binary neural-networks are proposed in the third paper. The last paper addresses a new retraining method improving fault tolerance in RRAM crossbars.

**Presentation Title**: FRAMEWORK FOR QUANTIFYING AND MANAGING ACCURACY IN STOCHASTIC CIRCUIT DESIGN

**Speaker**: Florian Neugebauer, University of Passau, DE

**Authors**: Florian Neugebauer, Ilija Polan and John Hayes

**Abstract**: Stochastic circuits (SCs) offer tremendous area and power-consumption benefits at the expense of computational inaccuracies. Managing accuracy is a central problem in SC design and has no counterpart in conventional circuit synthesis. It raises a basic question: how to build a systematic design flow for stochastic circuits? We present, for the first time, a systematic design approach to control the accuracy of SCs and balance it against other design parameters. We express the (in)accuracy of a circuit processing n-bit stochastic numbers by the numerical deviation of the computed value from the expected result, in conjunction with a confidence level. Using the theory of Monte Carlo simulation, we derive expressions for the stochastic number length required for a desired level of accuracy, or vice versa. We discuss the integration of the theory into a design framework that is applicable to both combinational and sequential SCs. We show that, for combinational SCs, accuracy is independent of the circuit's size or complexity, a surprising result. We also show how the analysis can identify subtle errors in both combinational and sequential designs.

**Download Paper (PDF)**: Only available from the DATE venue WiFi

**Presentation Title**: ENERGY-EFFICIENT APPROXIMATE MULTIPLIER DESIGN USING BIT SIGNIFICANCE-DRIVEN LOGIC COMPRESSION

**Speaker**: Issa Qiqieh, School of Electrical and Electronic Engineering, Newcastle University, GB

**Authors**: Issa Qiqieh, Rishad Shafik, Ghaith Taranweh, Danil Sokolov and Alex Yakovlev, Newcastle University, GB

**Abstract**: Approximate arithmetic has recently emerged as a promising paradigm for many impression-tolerant applications. It can offer substantial reductions in circuit complexity, delay and energy consumption by relaxing accuracy requirements. In this paper, we propose a novel energy-efficient approximate multiplier design using a significance-driven logic compression (SDLK) approach. Fundamental to this approach is an algorithmic and configurable lossy compression of the partial product rows based on their progressive bit significance. This is followed by the commutative remapping of the resulting product terms to reduce the number of product rows. As such, the complexity of the multiplier in terms of logic cell counts and lengths of critical paths is drastically reduced. A number of multipliers with different bit-widths (4-bit to 128-bit) are designed in SystemVerilog and synthesized using Synopsys Design Compiler. Post-synthesis experiments showed that up to an order of magnitude energy savings, and reductions of 65% in critical delay and almost 45% in silicon area can be achieved for a 128-bit multiplier compared to an accurate equivalent. These gains are achieved with low accuracy losses estimated at less than 0.00071 mean relative error. Additionally, we demonstrate the energy-accuracy trade-offs for different degrees of compression, achieved through configurable logic clustering. In evaluating the effectiveness of our approach, a case study image processing application showed up to 68.3% energy reduction with negligible losses in image quality expressed as peak signal-to-noise ratio (PSNR).

**Download Paper (PDF)**: Only available from the DATE venue WiFi

**Presentation Title**: ENERGY-EFFICIENT HYBRID STOCHASTIC-BINARY NEURAL NETWORKS FOR NEAR-SENSOR COMPUTING

**Speaker**: Vincent Lee, University of Washington, US

**Authors**: Vincent Lee, Armin Alaghi, John Hayes, Visvesh Sathe and Luis Ceze

**Abstract**: Recent advances in neural networks (NNs) exhibit unprecedented success at transforming large, unstructured data streams into compact higher-level semantic information for tasks such as handwriting recognition, image classification, and speech recognition. Ideally, systems would employ near-sensor computation to execute these tasks at sensor endpoints to maximize data reduction and minimize data movement. However, near-sensor computing presents its own set of challenges such as operating power constraints, energy budgets, and communication bandwidth capacities. In this paper, we propose a stochastic-binary hybrid design which splits the computation between the stochastic and binary domains for near-sensor NN applications. In addition, our design uses a new stochastic adder and multiplier that are significantly more accurate than existing adders and multipliers. We also show that retraining the binary portion of the NN computation can compensate for precision losses introduced by shorter stochastic bit-streams, allowing faster run times at minimal accuracy losses. Our evaluation shows that our hybrid stochastic-binary design can achieve 9.8× energy efficiency savings, and application-level accuracies within 0.05% compared to conventional all-binary designs.

**Download Paper (PDF)**: Only available from the DATE venue WiFi

**Presentation Title**: ACCELERATOR-FRIENDLY NEURAL-NETWORK TRAINING: LEARNING VARIATIONS AND DEFECTS IN RRAM CROSSBAR

**Speaker**: Li Jiang, Shanghai Jiao Tong University, CN

**Authors**: Lerong Chen, Jiewen Li, Yiran Chen, Qiuping Deng, Jiyuan Shen, Xiaoyao Liang and Li Jiang

**Abstract**: RRAM crossbar consisting of memristor devices can naturally carry out the matrix-vector multiplication; it thereby has gained a great momentum as a highly energy-efficient accelerator for neuro-morphic computing. The resistance variations and stuck-at faults in the memristor devices, however, dramatically degrade not only the chip yield, but also the classification accuracy of the neural-networks running on the RRAM crossbar. Existing hardware-based solutions cause enormous overhead and power consumption, while software-based solutions are less efficient in tolerating stuck-at faults and large variations. In this paper, we propose an accelerator-friendly neural-network training method, by leveraging the inherent self-healing capability of the neural-network, to prevent the large-weight synapses from being mapped to the abnormal memristors based on the fault/variation distribution in the RRAM crossbar. Experimental results show the proposed method can pull the classification accuracy (10%-45% loss in previous works) up close to ideal level with ≤ 1% loss.
2.3 Cache memory management for performance and reliability

Date: Tuesday 28 March 2017
Time: 11:30 - 13:00
Location / Room: B2C
Chair: Dionisios Pnevmatikatos, Technical University of Crete, GR
Co-Chair: Cristina Silvano, Politecnico di Milano, IT

Cache memory design optimizations and management can have a significant effect on cost, performance, and reliability. The first paper proposes an asymmetric cache management policy for GPGPUs with hybrid main memories that significantly improve performance for memory intensive workloads. The second paper targets the optimization of the bank placement in GPUs’ last level cache, with the goal of maximizing the performance of the GPU’s on-chip network. The third paper proposes a methodology for jointly analyzing all the cache level configurations to determine and minimize the susceptibility of the caches to soft errors.
2.4 Performance and Power Analysis

Tuesday 28 March 2017

11:30  2.3.1 BEST PAPER AWARD CANDIDATE

Presentation Title

Speaker:
Lei Ju, Shandong University, CN

Authors:
Gianluca Palermo, Politecnico di Milano, IT
Lei Ju, Shandong University, CN
Guan Wang, Xiaojun Cai, Lei Ju, Chuanqi Zang, Mengying Zhao and Zhiping Jia, Shandong University, CN

Abstract

Memory intensive workloads become increasingly popular on general purpose graphics processing units (GPGPUs), and impose great challenges on the GPGPU memory subsystem design. On the other hand, with the recent development of non-volatile memory (NVM) technologies, hybrid memory combining both DRAM and NVM achieves high performance, low power and high density simultaneously, which provides a promising main memory design for GPGPUs. In this work, we explore the shared last-level cache management for GPGPUs with consideration of the underlying hybrid main memory. In order to improve the overall memory subsystem performance, we exploit the characteristics of both the asymmetric read/write latency of the hybrid main memory architecture, as well as the memory coalescing feature of GPGPU. In particular, to reduce the average cost of L2 cache misses, we prioritize cache blocks from DRAM or NVM based on observation that operations to NVM part of main memory have large impact on the system performance. Furthermore, the cache management scheme also integrates the GPU memory coalescing and cache bypassing techniques to improve the overall cache hit ratio. Experimental results show that in the context of a hybrid main memory system, our proposed L2 cache management policy improves performance against the traditional LRU policy and a state-of-the-art GPU cache strategy EABP (20) by up to 27.76% and 14%, respectively.

Download Paper (PDF; Only available from the DATE venue WiFi)

12:00  2.3.2 EFFECTIVE CACHE BANK PLACEMENT FOR GPUs

Presentation Title

Speaker:
Mohammad Sadossadati, Sharif University of Technology, IR

Authors:
Mohammad Sadossadati1, Amirhossein Mirhosseini2, Shahin Roozkhosh1, Nazhir Bakshi1 and Hamid Sarbazi-Azad1

1Sharif University of Technology, IR; 2University of Michigan, US

Abstract

The placement of the Last Level Cache (LLC) banks in the GPU on-chip network can significantly affect the performance of memory-intensive workloads. In this paper, we attempt to offer a placement methodology for the LLC banks to maximize the performance of the on-chip network connecting the LLC banks to the streaming multiprocessors in GPUs. We argue that an efficient placement needs to be derived based on a novel metric that considers the latency hiding capability of the CPUs through thread level parallelism. To this end, we propose a throughput aware metric, called Effective Latency Impact (ELI). Moreover, we define an optimization problem to formulate our placement approach based on the ELI metric mathematically. To solve this optimization problem, we deploy a heuristic solution as this optimization problem is NP-hard. Experimental results show that our placement approach improves the performance by up to 15.7% compared to the state-of-the-art ELI algorithm.

Download Paper (PDF; Only available from the DATE venue WiFi)

12:30  2.3.3 SOFT ERROR-AWARE ARCHITECTURAL EXPLORATION FOR DESIGNING RELIABILITY ADAPTIVE CACHE HIERARCHIES IN MULTI-CORES

Presentation Title

Speaker:
Semeen Rehman, Technische Universitat Dresden, DE

Authors:
Arun Subramaniam1, Semeen Rehman2, Muhammad Shaflaque3, Akash Kumar4 and Joerg Henkel5

1EECS, University of Michigan-Ann Arbor, US; 2Technische Universität Dresden, DE; 3Vienna University of Technology (TU Wien), AT; 4Technische Universität Dresden, DE; 5Karlsruhe Institute of Technology, DE

Abstract

Mainstream multi-core processors employ large multi-level on-chip caches making them highly susceptible to soft errors. We demonstrate that designing a reliable cache hierarchy requires understanding the vulnerability interdependencies across different cache levels. This involves vulnerability analyses depending upon the parameters of different cache levels (partition size, line size, etc.) and the corresponding cache access patterns for different applications. This paper presents a novel soft error-aware cache architectural space exploration methodology and vulnerability analysis of multi-level caches considering their vulnerability interdependencies. Our technique significantly reduces exploration time while providing reliability-efficient cache configurations. We also show applicability/benefits for ECC-protected caches under multi-bit fault scenarios.

Download Paper (PDF; Only available from the DATE venue WiFi)

13:00  2.4.1 BEST PAPER AWARD CANDIDATE

Presentation Title

Speaker:
Swaroop Ghosh, Pennsylvania State University, US

Authors:
Radha Krishna Aluru

1EECS, University of Michigan-Ann Arbor, US; 2University of South Florida, US; 3Pennsylvania State University, US

Abstract

Spin-Transfer Torque magnetic Random Access Memory (STT-RAM) is one of the emerging technologies in the Domain of Non-volatile dense memories especially preferred for the last level cache (LLC). The amount of current needed to reorient the magnetization at present (~100μA per bit) is too high, enabling the write operation. When we perform a full cache line (512-bit) Write, this extremely high current compared to MRAM will result in a Voltage droop in the conventional cache architecture. Due to this droop, the write operation will fail half way through when we attempt to write in the farthest Bank of the cache from the supply. In this paper, we will be proposing a new cache architecture to mitigate this problem of droop and make the write operation successful.

Download Paper (PDF; Only available from the DATE venue WiFi)

13:00  2.4.2 BEST PAPER AWARD CANDIDATE

Presentation Title

Speaker:
Lei Ju, Shandong University, CN

Authors:
Guan Wang, Xiaojun Cai, Lei Ju, Chuanqi Zang, Mengying Zhao and Zhiping Jia, Shandong University, CN

Abstract

In this paper, we attempt to offer a placement methodology for the LLC banks to maximize the performance of the on-chip network connecting the LLC banks to the streaming multiprocessors in GPUs. We argue that an efficient placement needs to be derived based on a novel metric that considers the latency hiding capability of the CPUs through thread level parallelism. To this end, we propose a throughput aware metric, called Effective Latency Impact (ELI). Moreover, we define an optimization problem to formulate our placement approach based on the ELI metric mathematically. To solve this optimization problem, we deploy a heuristic solution as this optimization problem is NP-hard. Experimental results show that our placement approach improves the performance by up to 15.7% compared to the state-of-the-art ELI algorithm.

Download Paper (PDF; Only available from the DATE venue WiFi)

2.4 Performance and Power Analysis

Date: Tuesday 28 March 2017
Time: 11:30 - 13:00
Location / Room: 3A
Chair: Gianluca Palermo, Politecnico di Milano, IT

Lunch Break in Garden Foyer

Keynote Lecture session 3.0 in "Garden Foyer" 1350 - 1420

Lunch Break in the Garden Foyer

On all conference days (Tuesday to Thursday), a buffet lunch will be offered in the Garden Foyer, in front of the session rooms. Kindly note that this is restricted to conference delegates possessing a lunch voucher only. When entering the lunch break area, delegates will be asked to present the corresponding lunch voucher of the day. Once the lunch area is being left, re-entrance is not allowed for the respective lunch.
In this work, we propose an icache performance estimation approach that focuses on a component necessary to handle the instruction parallelism in a very long instruction word (VLIW) processor: the instruction buffer (IB). Our annotation approach is founded on an intermediate level native-simulation framework. It is evaluated with reference to a cycle accurate instruction set simulator leading to an average cycle count error of 9.3% and an average speedup of 10.

### SIMULATION

CPU stall time (due to cache misses) is lower than 10%, while the evaluation time can be sped up by 35 times relative to the Gem5 full simulations.

### MEAS: A FAST ANALYTIC POWER MODEL FOR DRAM MEMORIES

**Speaker:** Sandeep Poddar, IBM Research, The Netherlands, NL

**Authors:**

Sandeep Poddar1, Rik Jongerius1, Leandro Florin1, Giovanni Mariani1, Gero Dittmann1, Andreea Anghel2 and Henk Corporaal2

1IBM Research, NL; 2IBM Research, CH; 3TU/e (Eindhoven University of Technology), NL

**Abstract:**

The design of an energy-efficient memory subsystem is one of the key issues that system architects face today. To achieve this goal, architects usually rely on system simulators and trace-based DRAM power models. However, their long execution makes the approach infeasible for the design-space exploration of next-generation exascale computing systems. Analytic models, in contrast, are orders of magnitude faster. In this paper, we propose a new analytic memory scheduler-agnostic power model (MeSAP) for DRAM. Our model achieves an average error of 20% for DDR3 and DDR4 memory systems, similar to a state-of-the-art trace-based approach but our analytic model is an order of magnitude faster. Furthermore, we integrate MeSAP into an analytic performance model of general-purpose processors and show its applicability to the design of a computing system targeting scientific image processing applications.

Download Paper (PDF; Only available from the DATE venue WiFi)
This session addresses challenges related to energy efficiency and reliability of NoCs. The first paper proposes an analytical approach to evaluate the reliability of adaptive routing algorithms. The second paper, an online monitoring and routing approach is proposed to address the aging-related degradation in electrical NoC. Finally, the third paper shows how to use network traffic-aware spatial parallelism to improve the energy efficiency of the Epiphany SoC.

### EBSP: MANAGING NOC TRAFFIC FOR BSP WORKLOADS ON THE 16-CORE ADAPTEVA EPIPHANY-III PROCESSOR

**Authors:**
Siddhartha ¹ and Nachiket Kapre ²

**Kailany Technological University, SG; ²University of Waterloo, CA**

**Abstract:**
We can deliver high performance and energy efficient operation on the multi-core NoC-based Adapteva Epiphany-III SoC for bulk-synchronous workloads using our proposed EBSP communication API. We characterize and automate per-furnace tuning of spatial parallelism for supporting (1) random access load-store style traffic suitable for irregular sparse computations, as well as (2) variable, data-dependent traffic patterns in neural networks or PageRank-style workloads in a manner tailored for the Epiphany NoC. We aggressively optimize traffic by exposing spatial communication structure to the fabric through offline pre-computation of destination addresses, un-rolling of message-passing loops, selective squelching of messages, and careful ordering of communication and compute. Using our approach, across a range of applications and datasets such as Sparse Matrix-Vector multiplication (Matix Market datasets), PageRank (BerkStan SNAP dataset), and Izhikevich spiking neural evaluation, we deliver speedups of 6.5-10× while lowering power use by 2× compared to XY routing, respectively. The impact of our proposed methodology on network latency, Energy-Delay-Product (EDP) and link utilization is negligible.

Download Paper (PDF; Only available from the DATE venue WiFi)

### ONLINE MONITORING AND ADAPTIVE ROUTING FOR AGING MITIGATION IN NOCS

**Authors:**
Zana Ghaderi, Ayed Alqahthani and Nader Bagherzadeh

**University of California, Irvine, US**

**Abstract:**
Scalability of Network-on-Chip (NoC) as a promising solution for many-core systems can be jeopardized due to reliability challenges such as aging in advanced silicon technology. Previous mitigation techniques to protect NoC are either offline, while aging is strictly influenced by runtime operating conditions, or impose significant overheads to the system. This paper presents an online monitoring method through a Centralized Aging Table (CAT) for routers in NoCs. Router's capacity in flits, which are the main stimuli in routers, is predictable and limited for a given period of time. Consequently, stress rate and temperature, which are the major sources of aging mechanisms such as Bias Temperature Instability (BTI) and Hot Carrier Injection (HCI), will be in the predictable ranges, as well. Hence, our methodology uses CAT which is populated by values that represent aging degradation for each different pairs of stress and temperature ranges during a given period of time. Furthermore, utilizing CAT, we propose an online adaptive aging-aware routing algorithm in order to avoid highly aged routers which eventually leads to age balancing between routers. Additionally, our proposed routing algorithm reduces maximum age of routers by changing the shortest paths between source-destination pairs adaptively, considering routers' ages across them in each given period of time. Extensive experimental analysis using gem5 simulator demonstrates that our online routing algorithm and monitoring methodology, CAT, improves delay degradation of maximum aged router and aging imbalance on average by 39% and 52% compared to XY routing, respectively. The impact of our proposed methodology on network latency, Energy-Delay-Product (EDP) and link utilization is negligible.

Download Paper (PDF; Only available from the DATE venue WiFi)

### RELIABILITY ASSESSMENT OF FAULT TOLERANT ROUTING ALGORITHMS IN NETWORKS-ON-CHIP: AN ANALYTIC APPROACH

**Authors:**
Sadia Moriam and Gerhard Fettweis

**Technische Universität Dresden, DE**

**Abstract:**
Recent scaling of transistor gate sizes has significantly increased the density of on-chip integrations and paved the way for many-core systems-on-chip with highly improved performances. The design of the interconnection network of these complex systems is a critical one and the network-on-chip is now the accepted efficient interconnect for such large core arrays. An unfortunate adverse effect of technology scaling is the increased susceptibility to failures resulting in failing links and routers in the network-on-chip. To keep the network connected, efficient fault adaptive routing algorithms are necessary to route around faults. To design and evaluate the fault resiliency of such adaptive routing algorithms, fast, accurate and flexible analytic models are required, especially in large networks for which simulations are extremely time costly. In this paper, we present an analytic approach to evaluate the reliability of adaptive routing algorithms based on algebraic manipulations of the channel dependency matrix. It allows also to evaluate the number of alternate paths between source-destination pairs, in the presence of any number of permanent faults in the network. The analytic model is general and can be adapted to evaluate network reliability for any network topology and with any adaptive routing algorithm based on the turn model. We present cycle-accurate simulations to compare the accuracy of the model for the 2-D mesh and the hexagonal networks. The model is able to estimate the network fault resilience with an accuracy of about 1% and more than 70 times faster than the cycle accurate simulation.

Download Paper (PDF; Only available from the DATE venue WiFi)
Abstract
Testing analog, mixed-signal and RF circuits represents the main cost component for testing complex SoCs. A promising solution to alleviate this cost is the machine learning-based test strategy. These test techniques are an indirect test approach that replaces costly specification measurements by simpler signatures. Machine learning algorithms are used to map these signatures to the performance parameters. Although this approach has a number of undoubtable advantages, it also opens new issues that have to be addressed before it can be widely adopted by the industry. In this paper we present a machine learning-based test for a complex mixed-signal system -i.e. a state-of-the-art pipeline ADC- that includes digital calibration. This paper shows how the introduction of digital calibration for the ADC has a serious impact in the proposed test as calibration completely decorrelates signatures from the target specification in the presence of local mismatch.

Download Paper (PDF; Only available from the DATE venue WiFi)
2.7 EU Project Special Session: from Secure Clouds to reliable and variable HPC

Date: Tuesday 28 March 2017
Time: 11:30 - 13:00
Location / Room: 3B
Chair: Lorena Anghel, TIMA Laboratory, FR

Coversing the major topics presented in DATE, the European Projects presented in this session show lessons learned, best practices, scientific methods and evaluation platforms, successful strategies and roadmaps solving research and industry concerns in Europe.

**Time** | **Label** | **Presentation Title** | **Authors**
--- | --- | --- | ---
11:30 | 2.7.1 | HARPA: TACKLING PHYSICALLY INDUCED PERFORMANCE VARIABILITY | Dimitrios Soudris, ICCS, GR
Speaker: Nikolaos Zompakis and Dimitrios Soudris
Authors: ICCS/NTUA, GR; NTUA, GR
Abstract: Continuously increasing application demands on both High Performance Computing (HPC) and Embedded Systems (ES) are driving the IC manufacturing industry on an ever-lasting scaling of devices in silicon. Nevertheless, integration and miniaturization of transistors comes with an important and non-negligible trade-off: time-zero and time-dependent performance variability. Increasing guard-bands to battle variability is not scalable, since worst-case design margins are prohibitive for downscaled technology nodes. This paper discusses the FP7-612069-HARPA project of the European Commission which aims to enable next-generation embedded and high-performance heterogeneous many-cores to cost-effectively confront variations by providing Dependable-Performance: correct functionality and timing guarantees throughout the expected lifetime of a platform under thermal, power, and energy constraints. The HARPA novelty is in seeking synergies in techniques that have been considered virtually exclusively in the ES or HPC domains (worst-case guaranteed partly proactive techniques in embedded, and dynamic best-effort reactive techniques in high-performance).
Download Paper (PDF; Only available from the DATE venue WiFi)

12:00 | 2.7.2 | DYNAMIC SOFTWARE RANDOMISATION: LESSONS LEARNED FROM AN AEROSPACE CASE STUDY | Leonidas Kosmidis, Barcelona Supercomputing Center and Universitat Politècnica de Catalunya, ES
Speaker: Leonidas Kosmidis, Jaume Abella and Francisco Cazorla
Authors: Barcelona Supercomputing Center and Universitat Politècnica de Catalunya, ES; Barcelona Supercomputing Center (BSC-CNS), ES; Barcelona Supercomputing Center and I3A-CSIC, ES
Abstract: Timing Validation and Verification (V&V) is an important step in real-time system design, in which a system’s timing behaviour is assessed via Worst Case Execution Time (WCET) estimation and scheduling analysis. For WCET estimation, measurement-based timing analysis (MBTA) techniques are widely-used and well-established in industrial environments. However, the advent of complex processors makes it more difficult for the user to provide evidence that the software is tested under stress conditions representative of those at system operation. Measurement-Based Probabilistic Timing Analysis (MBPTA) is a variant of MBTA followed by the PROXIMA European Project that facilitates formulating this representativeness argument. MBPTA requires certain properties to be applicable, which can be obtained by selectively injecting randomisation in platform’s timing behaviour via hardware or software means. In this paper, we assess the effectiveness of the PROXIMA’s dynamic software randomisation (DSR) with a space industrial case study executed on a real unmodified hardware platform and an industrial operating system. We present the challenges faced in its development, in order to achieve MBPTA compliance and the lessons learned from this process. Our results, obtained using a commercial timing analysis tool, indicate that DSR does not impact the average performance of the application, while it enables the use of MBPTA. This results in tighter pWCET estimates compared to current industrial practice.
Download Paper (PDF; Only available from the DATE venue WiFi)

12:15 | 2.7.3 | READEX: LINKING TWO ENDS OF THE COMPUTING CONTINUUM TO IMPROVE ENERGY-EFFICIENCY IN DYNAMIC APPLICATIONS | Per Gunnar Kjeldsberg, Norwegian University of Science and Technology, NO
Speaker: Per Gunnar Kjeldsberg, Andreas Gocht, Michael Gerndt, Riha Lubomir, Joseph Schuchart and Umbreen Sabir Mian
Authors: Norwegian University of Science and Technology, NO; Technische Universität Dresden, DE; Technische Universität München, DE; IT4Innovations, Ostrava, CZ; Universität Stuttgart, DE
Abstract: In both the embedded systems and High Performance Computing domains, energy-efficiency has become one of the main design criteria. Efficiently utilizing the resources provided in computing systems ranging from embedded systems to current petascale and future Exascale HPC systems will be a challenging task. Suboptimal designs can potentially cause large amounts of underutilized resources and wasted energy. In both domains, a promising potential for improving efficiency of scalable applications stems from the significant degree of dynamic behaviour, e.g., runtime alternation in application resource requirements and workloads. Manually detecting and leveraging this dynamism to improve performance and energy-efficiency is a tedious task that is commonly neglected by developers. However, using an automatic optimization approach, application dynamism can be analysed at design time and used to optimize system configurations at runtime. The European Union Horizon 2020 READEX (Runtime Exploitation of Application Dynamism for Energy-efficient eXascale computing) project will develop a tools-aided auto-tuning methodology inspired by the system scenario methodology used in embedded systems. Dynamic behaviour of HPC applications will be exploited to achieve improved energy-efficiency and performance. Driven by a consortium of European experts from academia, HPC resource providers, and industry, the READEX project aims at developing the first of its kind generic framework to split design time and runtime automatic tuning while targeting heterogeneous system at the Exascale level. This paper describes plans for the project as well as early results achieved during its first year. Furthermore, it is shown how project results will be brought back into the embedded systems domain.
Download Paper (PDF; Only available from the DATE venue WiFi)
**Abstract**

This is an overview paper that motivates and describes performed work done in the European Commission funded research project BASTION, which focuses on two critical problems of modern electronics: the No-Fault-Found (NFF) and CMOS ageing. New defect classes contributing to NFF have been identified, including timing related faults (TRF) at board level and interconnect resistive faults (IRF) at IC level. BASTION has addressed the mechanisms of ageing and developed several techniques to improve the longevity of electronic products. Embedded Instrumentation, monitors, and IEEE 1687 standard for reconfigurable scan networks (RSN) are seen as an important leverage that helped mitigating the impact of the above listed problems by facilitating a low-latency, scalable online system health monitoring and error localization infrastructure as well as integration of all heterogeneous technologies into a homogeneous demonstration platform. This paper helps the reader to get a general overview of the work performed and provides a collection of references to publications where the respective research results are described in detail.

Download Paper (PDF; Only available from the DATE venue WiFi)
EXPLORING THE UNKNOWN THROUGH SUCCESSIVE GENERATIONS OF LOW POWER AND LOW RESOURCE VERSATILE AGENTS

Speaker:
Martin Andraud, Eindhoven University of Technology, NL

Authors:
Martin Andraud\textsuperscript{1} and Marian Verhelst\textsuperscript{2}
\textsuperscript{1}Eindhoven University of Technology, NL; \textsuperscript{2}Katholieke Universiteit Leuven, BE

Abstract
The Phoenix project aims to develop a new approach to explore unknown environments, based on multiple measurement campaigns carried out by extremely tiny devices, called agents, that gather data from multiple sensors. These low power and low resource agents are configured specifically for each measurement campaign to achieve the exploration goal in the smallest number of iterations. Thus, the main design challenge is to build agents as much reconfigurable as possible. This paper introduces the Phoenix project in more details and presents first developments in the agent design.

Download Paper (PDF; Only available from the DATE venue WiFi)

2.8a Smart Medical Devices

Date: Tuesday 28 March 2017
Time: 11:30 - 12:30
Location / Room: Exhibition Theatre
Organiser:
Patrick Mayor, EPFL, CH

The goal of this session is to present concrete examples of smart medical devices, such as a novel surgical robot for hearing implant surgery, a measurement module for the identification of cancer cells through elastic properties, as well as a sensing pad for non-invasive wound monitoring.

Time Label Presentation Title Authors
11:30 2.8a.1 HEARRESTORE Speaker: Juan Ansó, UniBE, CH
11:50 2.8a.2 PATLISCI II Speaker: Hans Peter Lang, UniBAS, CH
12:10 2.8a.3 FLUSITEX Speaker: Daniel Ahmed, ETHZ, CH
12:30 End of session
13:00 Lunch Break in Garden Foyer
Keynote Lecture session 3.0 in "Garden Foyer" 1350 - 1420

Lunch Break in the Garden Foyer
On all conference days (Tuesday to Thursday), a buffet lunch will be offered in the Garden Foyer, in front of the session rooms. Kindly note that this is restricted to conference delegates possessing a lunch voucher only. When entering the lunch break area, delegates will be asked to present the corresponding lunch voucher of the day. Once the lunch area is being left, re-entrance is not allowed for the respective lunch.

2.8b Smart Medical Devices, Part 2

Date: Tuesday 28 March 2017
Time: 12:30 - 13:00
Location / Room: Exhibition Theatre
Organiser:
John Zhao, MathWorks, US

Time Label Presentation Title Authors
**Abstract**

Smart connected devices and Internet of Things (IoT) are emerging technologies that are impacting diverse industries, including automotive, energy, healthcare, retail, smart manufacturing, smart buildings and homes, smart transportation, etc. Combining internet-connected devices with cloud computing, machine learning, and other data analytics approaches is enabling products and solutions that are transforming the way we live and work. For example, Smart Medical Devices are key components of new products and solutions that may help healthcare professionals to improve health outcomes from anywhere, leading to increased value for the patient.

However, a system developer working on such products and services faces challenges in capturing, storing, and analyzing the Big Data generated from a multitude of devices. Also, integrating Smart Devices, IoT and Big Data raises specific challenges for data acquisition, reduction, and transmission, using increasingly sophisticated technologies such as RFID tags, Wireless Sensor Nodes and mobile devices.

Using the development of a Smart Medical Device based healthcare application as an example, this presentation will discuss how engineers and scientists creating smart devices and IoT systems can use MATLAB and Simulink to access and analyze huge data sets from devices, sensors, and databases; apply deep learning and other machine-learning techniques to develop predictive models; and design and test smart devices that wirelessly interact with cloud services like ThingSpeak™, an analytic IoT platform that can run MATLAB code on demand in the cloud.

---

**Authors**

Julien Lallet

**Abstract**

Current augmented reality (AR) eyewear solutions require large form factors, weight, cost and energy that reduce usability. In fact, connectivity, image processing, localization, and direction evaluation lead to high processing and power requirements. A multi-antenna system, patented by the industrial partner, enables a new generation of smart eye-wear that elegantly requires less hardware, connectivity, and power to provide AR functionalities. They will allow users to directionally locate nearby radio emitting sources that highlight objects of interest (e.g., people or retail items) by using existing standards like Bluetooth Low Energy, Apple’s iBeacon and Google’s Eddystone. This booth will report the current level of research addressed by the Computer Science Department of University of Verona and the company Wagoo LLC. In the presented demo, different objects emit an “I am here” signal and a prototype of the smart glasses shows the information related to the observed object.

**Authors**

Gianluca Benedetti and Walter Vendraminetto

1Wagoo LLC, IT; 2EDALab srl, IT

**Abstract**

A large number of models that are employed in the field of concurrent systems’ design, such as Petri nets, gate-level circuits, dataflow structures have an underlying static graph structure. Their semantics, however, is defined using additional entities, e.g. tokens or node/arc states, which collectively form the overall state of the system. We jointly refer to such formalisms as interpreted graph models. This demo will show the use of an open-source cross-platform Workcraft framework for capturing, simulation, synthesis, and verification of such models. The focus of our case study will be on synthesis from technology-independent formal specifications to verifiable circuit implementations.

**Authors**

Davide Quaglia, University of Verona, IT

**Presenters**

1Imran Latif, 2Edward Patel, 3Ludovic Apvrille, 4Renato Pacalet, 5Adrien Canuel

1Télécom ParisTech, FR

**Abstract**

Future 5G networks are expected to provide an increase of 10x in data rates. To meet these requirements, the equipment of baseband stations will be designed using mixed architectures, i.e., DSPs, FPGAs. However, efficiently programming these architectures is not trivial due to the drastic increase in complexity of their design space. To overcome this issue, we need to have unified tools capable of rapidly exploring, partitioning and prototyping the mixed architecture designs of 5G systems. At DATE 2017 University Booth, we demonstrate such a unified tool and show our latest achievements in the automatic code generation engine of TTool/DIPLODOCUS, a UML/SysML framework for the hardware/software co-design of data-flow systems, to support mixed architectures. Our demonstration will show the full design and evaluation of a 5G data-link layer receiver for both a DSP-based and an IP-based designs. We will validate the effectiveness of our solution by comparing automated vs manual designs.

**Authors**

Andrea Emrini, Nokia Bell Labs, France

1Nokia Bell Labs, France; 2Télécom ParisTech, FR

**Abstract**

Smart connected devices and Internet of Things (IoT) are emerging technologies that are impacting diverse industries, including automotive, energy, healthcare, retail, smart manufacturing, smart buildings and homes, smart transportation, etc. Combining internet-connected devices with cloud computing, machine learning, and other data analytics approaches is enabling products and solutions that are transforming the way we live and work. For example, Smart Medical Devices are key components of new products and solutions that may help healthcare professionals to improve health outcomes from anywhere, leading to increased value for the patient.

MATLAB AND SIMULINK IN THE SMART DEVICES AND BIG DATA ERA

Stefano Olivieri, MathWorks Academia Group, US

**Abstract**

Smart connected devices and Internet of Things (IoT) are emerging technologies that are impacting diverse industries, including automotive, energy, healthcare, retail, smart manufacturing, smart buildings and homes, smart transportation, etc. Combining internet-connected devices with cloud computing, machine learning, and other data analytics approaches is enabling products and solutions that are transforming the way we live and work. For example, Smart Medical Devices are key components of new products and solutions that may help healthcare professionals to improve health outcomes from anywhere, leading to increased value for the patient.

However, a system developer working on such products and services faces challenges in capturing, storing, and analyzing the Big Data generated from a multitude of devices. Also, integrating Smart Devices, IoT and Big Data raises specific challenges for data acquisition, reduction, and transmission, using increasingly sophisticated technologies such as RFID tags, Wireless Sensor Nodes and mobile devices.

Using the development of a Smart Medical Device based healthcare application as an example, this presentation will discuss how engineers and scientists creating smart devices and IoT systems can use MATLAB and Simulink to access and analyze huge data sets from devices, sensors, and databases; apply deep learning and other machine-learning techniques to develop predictive models; and design and test smart devices that wirelessly interact with cloud services like ThingSpeak™, an analytic IoT platform that can run MATLAB code on demand in the cloud.

**Authors**

Julien Lallet
More information ...
**LABSmiling:** A framework composed of a remotely accessible testbed and related SW tools, for analysis and design of low data-rate wireless personal area networks based on IEEE 802.15.4

**Presenter:**
Marco Santic, University of L’Aquila, IT

**Authors:**
Luigi Pomante, Walter Tiberti, Carlo Centofanti and Lorenzo Di Giuseppe, DEWS - Università di L’Aquila, IT

**Abstract**
Low data-rate wireless personal area networks (LR-WPANs) are even more present in the fields of IoT, wearable devices and health monitoring. The development, deployment and test of such systems, based on IEEE 802.15.4 standard (and its derivations, e.g. 15.4e), require the exploitation of a testbed when the network is not trivial and grows in complexity. This demo shows the framework of LabSmiling: a testbed and related SW tools that connect a meaningful (but still scalable) number of physical devices (sensor nodes) located in a real environment. It offers the following services: program, reset, switch on/off single devices; connect to devices up/down links to inject or receive commands/msgs/packets in/from the network; set devices as low level packet sniffers, allowing to test/debug protocol compliances or extensions. Advanced services are: possibility of design test scenarios for the evaluation of network metrics (throughput, latencies, etc.) and custom application verification.

**More information ...**

**15:00**  
End of session

**16:00**  
Coffee Break in Exhibition Area

On all conference days (Tuesday to Thursday), coffee and tea will be served during the coffee breaks at the below-mentioned times in the exhibition area.

Tuesday, March 28, 2017  
- Coffee Break 10:30 - 11:30
- Coffee Break 16:00 - 17:00

Wednesday, March 29, 2017  
- Coffee Break 10:00 - 11:00
- Coffee Break 16:00 - 17:00

Thursday, March 30, 2017  
- Coffee Break 10:00 - 11:00
- Coffee Break 15:30 - 16:00

---

### 3.0 LUNCH TIME KEYNOTE SESSION: Precision Medicine: Where Engineering and Life Science meet

**Date:** Tuesday 28 March 2017  
**Time:** 13:50 - 14:20  
**Location / Room:** Garden Foyer  
**Chair:**  
David Atienza, EPFL, CH

As we witness the relentless growth of computing power, storage capacity and communication bandwidth, we also see a major trend in bio-medical sciences to become more quantitative and amenable to benefit from the support of electronic systems. Moreover, societal and economic needs push us to develop and adopt health-management approaches that are more effective, less expensive and flexible enough to be personalized to individual and community needs. Within this frame, precision medicine promises to better society by applying engineering technology to personalized health, with devices that are in/on the body and ubiquitously connected. Examples from the Swiss-wide Nano-Tera.ch program will show various techniques related remote patient monitoring, emergency care as well as routine care. These examples show the advantages that stem from organized and optimized means to quantify clinical data, handle large data sets as well as controlling and personalizing therapy and drug administration.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
</table>

**Abstract**
As we witness the relentless growth of computing power, storage capacity and communication bandwidth, we also see a major trend in bio-medical sciences to become more quantitative and amenable to benefit from the support of electronic systems. Moreover, societal and economic needs push us to develop and adopt health-management approaches that are more effective, less expensive and flexible enough to be personalized to individual and community needs. Within this frame, precision medicine promises to better society by applying engineering technology to personalized health, with devices that are in/on the body and ubiquitously connected. Examples from the Swiss-wide Nano-Tera.ch program will show various techniques related remote patient monitoring, emergency care as well as routine care. These examples show the advantages that stem from organized and optimized means to quantify clinical data, handle large data sets as well as controlling and personalizing therapy and drug administration. Electronic design automation is a key technology to realize systems for precision medicine. Examples of specific EDA tools and methods encompass physical design of integrated sensors and their coupling to electronics, simulation of complex systems with bio-chemical stimuli, synthesis of decision making circuitry based on plurality of inexact inputs, policies design for therapies exploiting online data acquisition, and verification of life-critical applications under broadly-varying and unpredictable input conditions. Overall, precision medicine represents an important and large market opportunity. EDA is a necessary underlying technology to realize the promises of better and less expensive care for everyone.

**14:20**  
End of session
Coffee Break in Exhibition Area

On all conference days (Tuesday to Thursday), coffee and tea will be served during the coffee breaks at the below-mentioned times in the exhibition area.

Tuesday, March 28, 2017
- Coffee Break 10:00 - 11:30
- Coffee Break 16:00 - 17:00

Wednesday, March 29, 2017
- Coffee Break 10:00 - 11:00
- Coffee Break 16:00 - 17:00

Thursday, March 30, 2017
- Coffee Break 10:00 - 11:00
- Coffee Break 15:30 - 16:00

3.1 IT&A Session: Parallel Ultra-Low-Power Computing for the IoT: Applications, Platforms, Circuits

Date: Tuesday 28 March 2017
Time: 14:30 - 16:00
Location / Room: SBC

Organisers:
Luca Benini, ETHZ, CH
Davide Rossi, Università di Bologna, IT

Chair:
Luca Benini, ETHZ, CH

Co-Chair:
Davide Rossi, Università di Bologna, IT

This special session will give a deep dive into Ultra-low power computing for Internet-of-Things applications, starting from leading-edge MCU-based commercial solutions, moving to next generation highly-parallel ULP architectures based on open-source hardware & software, fast-forwarding to advanced research solutions based new models of computations.

14:30 3.1.1 BETTER THAN WORST CASE SIGNOFF STRATEGIES FOR LOW POWER IOT DEVICES
Speaker: Jose Pineda de Gyvez, NXP Semiconductors, NL
Authors: Jose Pineda and Hamed Fatemi, NXP Semiconductors, NL

Abstract
Portable consumer electronic devices are nowadays ubiquitous. Digital ubiquity, along with a lift in semiconductor utilization for consumer electronics, power autonomy, and device miniaturization are key challenges to attain digital convergence for seamless operability. Most of the state-of-the-art computing architectures are based on power-performance trade-offs. In fact, it is unconceivable to think that without power management any kind of competitive compute solution can be marketed in the entire application field. The relative slow innovation progress on battery technologies demands radical innovations for energy-efficient operation. The inability of battery technologies to keep pace with long operating times required by modern multi-purpose devices necessitates alternative (design) solutions that extend battery lifetime. In this presentation we will focus on signoff techniques aimed to yield designs with smaller area and lower power next to reducing signoff complexity because of sever process variability. More specifically, we make use of standard cell libraries characterized for a lower power next to reducing signoff complexity because of sever process variability. We evaluate the proposed techniques in a Cortex-M3 testchip designed in 40nm CMOS process. We will show measurement results that demonstrate the effectiveness of using better than worst case signoffs.

15:00 3.1.2 GAP: AN OPEN-SOURCE PULP-RISCV PLATFORM FOR NEAR-SENSOR ANALYTICS
Author: Eric Flamand, GreenWaves Technologies, FR

15:30 3.1.3 ENERGY-QUALITY SCALABLE ADAPTIVE VLSI CIRCUITS AND SYSTEMS BEYOND APPROXIMATE COMPUTING
Speaker and Author: Massimo Alioto, National University of Singapore, SG

Abstract
In this paper, the concept of energy-quality (EQ) scalable systems is introduced and explored, as novel design dimension to scale down energy in integrated systems for the Internet of Things (IoT). EQ-scalable systems explicitly trade off energy and quality at different levels of abstraction (“vertically”), and sub-systems (“horizontally”), creating new opportunities to improve energy efficiency for a given task and expected “quality”. The concept of quality slack, a taxonomy of techniques to trade off energy and quality and a general EQ-scalable architecture are presented. The generality of the EQ-scaling concept is shown through several examples, ranging from logic to analog circuits, to memories and Analog-Digital Converters. Challenges, opportunities and expected energy gains are discussed to gain an understanding of the potential of the EQ-scalable integrated circuits and systems. As a result, EQ scalable systems are expected to substantially improve the energy efficiency of systems for IoT, compensating the limited energy gains that will be offered by technology and voltage scaling.

Download Paper (PDF; Only available from the DATE venue WIFI)
### 3.2 Hot Topic Session: New Benchmarking Vectors for Emerging Devices, Circuits, and Architectures: Energy, Delay, and ... Accuracy

**Date:** Tuesday 28 March 2017  
**Time:** 14:30 - 16:00  
**Location / Room:** 4BC

**Organisers:**  
Xiaobo Sharon Hu, University of Notre Dame, US  
Michael Niemier, University of Notre Dame, US

**Chair:**  
Xiaobo Sharon Hu, University of Notre Dame, US

**Co-Chair:**  
Pierre-Emmanuel Gaillardon, The University of Utah at Salt Lake City, US

There is ever-growing interest in alternative computational models (e.g., neural networks, etc.), as well as how emerging technologies can best be exploited to address application-level needs. This hot topic session addresses the above issues from the perspective of benchmarking. It considers the impact of emerging devices, circuits, and architectures at the application level in the context of new metrics and benchmarking methodologies being developed via the Semiconductor Research Corporation (SRC). Subsequent presentations highlight benchmarking and design space exploration efforts that consider application-level energy and performance in the context of computational accuracy. They also highlight infrastructure that can be used to compare different devices, circuits, and architectures that ultimately address the same information processing task.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>14:30</td>
<td>3.2.1</td>
<td><strong>BEYOND-CMOS NON-BOOLEAN LOGIC BENCHMARKING: INSIGHTS AND FUTURE DIRECTIONS</strong></td>
<td>Azad Naeemi, Georgia Institute of Technology, US</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Authors:</strong> Chenyun Pan and Azad Naeemi, Georgia Institute of Technology, US</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong> Emerging technologies are facing significant challenges to compete with CMOS with respect to Boolean logic. There is an increasing need for using non-traditional circuits to realize the full potential of beyond-CMOS devices. This paper presents a uniform benchmarking methodology for non-Boolean computation based on the cellular neural network (CNN) for a variety of beyond-CMOS device technologies, including charge- based and spintronic devices. Three types of CNN implementations are benchmarked for a given input noise and recall accuracy target using analog, digital, and spintronic circuits. Results demonstrate that spintronic devices are promising candidates to implement CNNs, where up to 3× EDP improvement is predicted in domain wall devices compared to its conventional CMOS counterpart. This shows that alternative non-Boolean computing platforms are crucial for developing future emerging technologies.</td>
<td>Download Paper (PDF; Only available from the DATE venue WiFi)</td>
</tr>
<tr>
<td>15:00</td>
<td>3.2.2</td>
<td><strong>UNDERSTANDING THE DESIGN OF IBM NEUROSYNAPTIC SYSTEM AND ITS TRADEOFFS: A USER PERSPECTIVE</strong></td>
<td>Yiran Chen, Duke University, US</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Authors:</strong> Hsin-Pai Cheng, Wei Wen, Chunpeng Wu, Sicheng Li, Hai (Helen) Li and Yiran Chen, University of Pittsburgh, US</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong> As a large-scale commercial spiking-based neuromorphic computing platform, IBM TrueNorth processor consumed tremendous attentions in society. However, one of the known issues in TrueNorth design is the limited precision of synaptic weights. The current workaround is running multiple neural network copies in which the average value of each synaptic weight is close to that in the original network. We theoretically analyze the impacts of low data precision in the TrueNorth chip on inference accuracy, core occupation, and performance, and present a probability-biased learning method to enhance the inference accuracy through reducing the random variance of each computation copy. Our experimental results proved that the proposed techniques considerably improve the computation accuracy of TrueNorth platform and reduce the incurred hard-ware and performance overheads. Among all the tested methods, LITEA regularization achieved the best result, say, up to 2.74% accuracy enhancement when deploying MNIST application onto TrueNorth platform. In May 2016, IBM TrueNorth team implement convolutional neural networks (CNN) on TrueNorth processor and coincidently use a similar method, say, trinary weights, {−1, 0, 1}. It achieves near state-of-the-art accuracy on 8 standard datasets. In addition, to further evaluate TrueNorth performance on CNN, we test similar deep convolutional networks on True North, GPU and FPGA. Among all, GPU has the highest throughput. But if we consider energy consumption, TrueNorth processor is the most energy efficient one, say, &gt; 6000 frames/sec/Watt.</td>
<td>Download Paper (PDF; Only available from the DATE venue WiFi)</td>
</tr>
</tbody>
</table>
3.2.3 CELLULAR NEURAL NETWORK FRIENDLY CONVOLUTIONAL NEURAL NETWORKS - CNNS WITH CNNS

Speaker: Michael Niemier, University of Notre Dame, US
Authors: András Horváth, Michael Hillmer, Qiuwen Lou, X. Sharon Hu and Michael Niemier

Abstract
This paper will discuss the development and evaluation of a cellular neural network (CeNN)-friendly deep learning network that addresses the MNIST digit recognition problem. Prior work has shown that CeNNs leveraging emerging technologies such as tunnel transistors can improve energy or EDP of CeNNs, while simultaneously offering richer/more complex functionality. Important questions to address are what applications can benefit from CeNNs, and whether CeNNs can eventually outperform other alternatives at the application-level in terms of energy, performance, and accuracy. This paper begins to address these questions by using the MNIST problem as a case study.

Download Paper (PDF; Only available from the DATE venue WiFi)

3.3 Hardware Trojans and Fault Attacks

Date: Tuesday 28 March 2017
Time: 14:30 - 16:00
Location / Room: 2BC

Chair: Ilia Polian, University of Passau, DE
Co-Chair: Matthias Sauer, University of Freiburg, DE

This section focuses on two types of active attacks on system hardware modules: hardware Trojans (malicious modifications) and fault-injections into cryptographic modules. The papers cover Trojans that target coherence protocols in memory caches; Trojan detection based on measurement of path delays; detection of malware using machine learning; and fault attacks on the cryptographic hash function SHA-3.

3.3.1 ALGEBRAIC FAULT ANALYSIS OF SHA-3

Speaker: Pei Luo, Northeastern University, US
Authors: Pei Luo, Konstantinos Athanasiou, Yunsi Fei and Thomas Wahl, Northeastern University, US

Abstract
This paper presents an efficient algebraic fault analysis on all four modes of SHA-3 under relaxed fault models. This is the first work to apply algebraic techniques on fault analysis of SHA-3. Results show that algebraic fault analysis on SHA-3 is very efficient and effective due to the clear algebraic properties of Keccak operations. Comparing with previous work on differential fault analysis of SHA-3, algebraic fault analysis can identify the injected faults with much higher rates, and recover an entire internal state of the penultimate round with much fewer fault injections.

Download Paper (PDF; Only available from the DATE venue WiFi)

3.3.2 EVALUATING COHERENCE-EXPLOITING HARDWARE TROJAN

Speaker: Minsu Kim, Korea University, KR
Authors: Minsu Kim, Sunhee Kong, Boeui Hong, Lei Xu, Weidong Shi and Taeweon Suh

Abstract
Increasing complexity of integrated circuits and IP-based hardware designs have created the risk of hardware Trojans. This paper introduces a new type of threat, a coherence-exploiting hardware Trojan. This Trojan can be maliciously implanted in master components in a system, and continuously injects memory transactions onto the main interconnect. The injected traffic forces the eviction of cache lines, taking advantage of cache coherence protocols. This type of Trojans insidiously slows down the system performance, incurring Denial-of-Service (DoS) attack. We used a Xilinx Zynq-7000 device to implement the Trojan and evaluate its severity. Experiments revealed that the system performance can be severely degraded as much as 258% with the Trojan. A countermeasure to annihilate the Trojan attack is proposed in detail. We also found that AXI version 3.0 supports a seemingly irrelevant invalidation protocol through ACP, opening a door for the potential Trojan attack.

Download Paper (PDF; Only available from the DATE venue WiFi)
### 3.4 Guardbanding and Approximation

**Abstract**

This session starts with a guardbanding-based approach that uses cell libraries designed and classified for different temperature ranges for improving circuit timing as well as detecting hardware Trojans (HT) insertion. In this technique, the power profile is obtained by extracting the power parameters based techniques have been developed to detect these hardware Trojans that require golden circuit as a reference model, but due to the widespread usage of IPs, most of the system-on-chip (SoC) do not have a golden reference. Hardware Trojans in intellectual property (IP)-based SoC designs are considered as major concern for future integrated circuits. Most of the state-of-the-art runtime hardware Trojan detection techniques presume that Trojans will lead to anomaly in the SoC integration units. In this paper, we argue that an intelligent intruder may intrude the IP-based SoC without disturbing the normal SoC operation or violating any protocols. To overcome this limitation, we propose a methodology to extract the power profile of the microcontrollers instruction sets, which is in turn used to train a machine learning algorithm. In this technique, the power profile is obtained by extracting the power behavior of the micro-controllers for different assembly language instructions. This trained model is then embedded into the integrated circuits at the SoC integration level, which classifies the power profile during runtime to detect the intrusions. We applied our proposed technique on MC8051 microcontroller in VHDL, obtained the power profile of its instruction set and then applied deep learning, k-NW, decision tree and naive Bayesian based machine learning tools to train the models. The cross validation comparison of these learning algorithms, when applied to MC8051 Trojan benchmarks, shows that we can achieve 99.0% detection rate of 99.0% with less than 5% false positives and outperforms previous proposals for hardware-assisted malware detection.

**Authors**

Fatma Nur Esirci and Alp Arslan Bayrakci, Gebze Technical University, TR

**Presentation Title**

HARDWARE TROJAN DETECTION BASED ON CORRELATED PATH DELAYS IN DEFIANCE OF VARIATIONS WITH SPATIAL CORRELATIONS

**Speaker**

Fatma Nur Esirci and Alp Arslan Bayrakci, Gebze Technical University, TR

**Location / Room**

3A

**Time**

15:30

### 3.3.4 MALWARE DETECTION USING MACHINE LEARNING BASED ANALYSIS OF VIRTUAL MEMORY ACCESS PATTERNS

**Abstract**

Malicious software, referred as malware, continues to grow in sophistication. Past proposals for malware detection have primarily focused on software-based detectors which are vulnerable to being compromised. Thus, recent work has proposed hardware-assisted malware detection. In this paper, we introduce a new framework for hardware-assisted malware detection based on monitoring and classifying memory access patterns using machine learning. This provides for increased automation and coverage through reducing user input on specific malware signatures. The key insight underlying our work is that malware must change control flow and/or data structures, which leaves fingerprints on program memory accesses. Building on this, we propose an online framework for detecting malware that uses machine learning to classify malicious behavior based on virtual memory access patterns. Novel aspects of the framework include techniques for collecting and summarizing per-function/system-call memory access patterns, and a two-level classification architecture. Our experimental evaluation focuses on two important classes of malware: (i) kernel rootkits and (ii) memory corruption attacks on user programs. The framework has a detection rate of 99.0% with less than 5% false positives and outperforms previous proposals for hardware-assisted malware detection.

**Authors**

Zhixing Xu, Princeton University, US

**Presentation Title**

POWER PROFILING OF MICROCONTROLLER'S INSTRUCTION SET FOR RUNTIME HARDWARE TROJANS DETECTION WITHOUT GOLDEN CIRCUIT MODELS

**Speaker**

Zhixing Xu, Princeton University, US

**Location / Room**

3A

**Time**

15:45

**Download Paper (PDF; Only available from the DATE venue WiFi)**
Approximate computing is gaining more and more attention as potential solution to the problem of increasing energy demand in computing. Several recent developments constructively. Results show that when considering voltage scaling, up to 24.5% higher energy savings can be achieved compared to approaches that only consider switching activity. Our heuristic solver is able to find solutions within 0.1% of average energy savings compared to an exhaustive search, all while being up to 1,400x faster than simulation-based methods. Lately, several hardware approximate operators were invented, seeking the same performance benefits. However, a fair comparison between the usage of this new class of operators and classical fixed-point arithmetic with careful truncation or rounding, has never been performed. In this paper, we first compare approximate and fixed-point arithmetic operators in terms of power, area and delay, as well as in terms of induced error, using many state-of-the-art relevant metrics. In this paper, we show that considering a large set of parameters, existing approximate adders and multipliers tend to be dominated by truncated or rounded fixed-point ones. For a given accuracy level and when considering the whole computation data-path, fixed-point operators are several orders of magnitude more accurate while spending less energy to execute the application. A conclusion of this study is that the entropy of careful sizing is always lower than approximate operators, since it require significantly less bits to be processed in the data-path and stored. Approximated data therefore always contain on average a greater amount of costly erroneous, useless information. The second paper in the session compares the use of approximate arithmetic components (adders, multipliers) with truncation and rounding techniques to diminish the bit-width of the units. The third work proposes a source-to-source transformation to optimize the energy-accuracy tradeoff. This session concludes with two IP presentations on approximate errors symbolic system synthesis.
GAUSSIAN MIXTURE ERROR ESTIMATION FOR APPROXIMATE CIRCUITS

Speaker:
Amin Ghasemazar, The University of British Columbia, CA

Authors:
Amin Ghasemazar and Mieszko Lis, University of British Columbia, CA

Abstract
In application domains where perceived quality is limited by human senses, where data are inherently noisy, or where models are naturally inexact, approximate computing offers an attractive tradeoff between accuracy and energy or performance. While several approximate functional units have been proposed to date, the question of how these techniques can be systematically integrated into a design flow remains open. Ideally, units like adders or multipliers could be automatically replaced with their approximate counterparts as part of the design flow. This, however, requires accurately modelling approximation errors to avoid compromising output quality. Prior proposals have either focused on describing errors per-bit or significantly limited estimation accuracy to reduce otherwise exponential storage requirements. When multiple approximate modules are chained, these limitations become critical, and propagated error estimates can be orders of magnitude off. In this paper, we propose an approach where both input distributions and approximation errors are modelled as Gaussian mixtures. This naturally represents the multiple sources of error that arise in many approximate circuits while maintaining reasonable memory requirements. Estimation accuracy is significantly better than prior art (up to 7.2× lower Hellinger distance) and errors can be accurately propagated through a cascade of approximate operations; estimates of quality metrics like MSE and MED are within a few percent of simulation-derived values.

Download Paper (PDF; Only available from the DATE venue WiFi)

ENHANCING SYMBOLIC SYSTEM SYNTHESIS THROUGH ASPMT WITH PARTIAL ASSIGNMENT EVALUATION

Speaker:
Kai Neubauer, University of Rostock, DE

Authors:
Kai Neubauer1, Philipp Wanko2, Torsten Schaub2 and Christian Haubelt1
1University of Rostock, DE; 2University of Potsdam, DE

Abstract
The design of embedded systems is becoming continuously more complex such that efficient design methods are becoming crucial for competitive results regarding design time and performance. Recently, combined Answer Set Programming (ASP) and Quantifier Free Integer Difference Logic (QF-IDL) solvers has been shown to be a promising approach in system synthesis. However, this approach still has several restrictions limiting its applicability. In the paper at hand, we propose a novel ASP modulo theories (ASPmT) system synthesis approach, which (i) supports more sophisticated system models, (ii) tightly integrates the QF-IDL solver into the ASP solving, and (iii) makes use of partial assignment checking. As a result, more realistic systems are considered and an early exclusion of infeasible solutions improves the entire system synthesis.

Download Paper (PDF; Only available from the DATE venue WiFi)

3.5 Low-power brain inspired computing for embedded systems

Date: Tuesday 28 March 2017
Time: 14:30 - 16:00
Location / Room: 3C

Chair:
Johanna Sepulveda, TU Munich, DE

Co-Chair:
Andrea Bartolini, Universita' di Bologna - ETH Zurich, IT

Neural Networks are promising techniques for bringing brain inspired computing into embedded platforms. Energy efficiency is a primary concern in these computing domain. This track combines low power design techniques such as approximate computing and compression with state-of-the-art hardware architectures.
Approximate Computing for Spiking Neural Networks

Speaker: Qiuwen Chen and Qinru Qiu, Syracuse University, US
Authors: Qiuwen Chen and Qinru Qiu, Syracuse University, US
Abstract: Spiking Neural Networks (SNNs) are widely regarded as the third generation of artificial neural networks, and are expected to drive new classes of recognition, data analytics and computer vision applications. However, large-scale SNNs (e.g., of the scale of the human visual cortex) are highly compute and data intensive, requiring new approaches to improve their efficiency. Complementary to prior efforts that focus on parallel software and the design of specialized hardware, we propose AxSNN, the first effort to apply approximate computing to improve the computational efficiency of evaluating SNNs. In SNNs, the inputs and outputs of neurons are encoded as a time series of spikes. A spike at a neuron's output triggers updates to the potentials (internal states) of neurons to which it is connected. AxSNN determines spike-triggered neuron updates that can be skipped with little or no impact on output quality and selectively skips them to improve both compute and memory energy. Neurons that can be approximated are identified by utilizing various static and dynamic parameters such as the average spiking rates and current potentials of neurons, and the weights of synaptic connections. Such a neuron is placed into one of many approximation modes, wherein the neuron is sensitive only to a subset of its inputs and sends spikes only to a subset of its outputs. A controller periodically updates the approximation modes of neurons in the network to achieve energy savings with minimal loss in quality. We apply AxSNN to both hardware and software implementations of SNNs. For hardware evaluation, we designed SNNAP, a Spiking Neural Network Approximate Processor that embodies the proposed approximation strategy, and synthesized it to 45nm technology. The software implementation of AxSNN was evaluated on a 2.7 GHz Intel Xeon server with 128 GB memory. Across a suite of 6 image recognition benchmarks, AxSNN achieves 1.4-5.5X reduction in scalar operations for network evaluation, which translates to 1.2-3.62X and 1.26-3.9X improvement in hardware and software energies respectively, for no loss in application quality. Progressively higher energy savings are achieved with modest reductions in output quality.

Download Paper (PDF; Only available from the DATE venue WiFi)

Adaptive Weight Compression for Memory-Efficient Neural Networks

Speaker: Jong Hwan Ko, Georgia Institute of Technology, US
Authors: Jong Hwan Ko, Duckhwan Kim, Taesik Na, Jaeha Kung and Saibal Mukhopadhyay, Georgia Institute of Technology, US
Abstract: Neural networks generally require significant memory capacity/bandwidth to store/access a large number of synaptic weights. This paper presents an application of JPEG image encoding to compress the weights by exploiting the spatial locality and smoothness of the weight matrix. To minimize the loss of accuracy due to JPEG encoding, we propose to adaptively control the quantization factor of the JPEG algorithm depending on the error-sensitivity (gradient) of each weight. With the adaptive compression technique, the weight blocks with higher sensitivity are compressed less for higher accuracy. The adaptive compression reduces memory requirement, which in turn results in higher performance and lower energy of neural network hardware. The simulation for inference hardware for multilayer perceptron with the MNIST dataset shows up to 42X compression with less than 1% loss of recognition accuracy, resulting in 3X higher effective memory bandwidth and ~19X lower system energy.

Download Paper (PDF; Only available from the DATE venue WiFi)

Real-time Anomaly Detection for Streaming Data using Burst Code on a Neurosynaptic Processor

Speaker: Qinru Qiu, Syracuse University, US
Authors: Qinru Qiu, Syracuse University, US
Abstract: Real-time anomaly detection for streaming data is a desirable feature for mobile devices or unmanned systems. The key challenge is how to deliver required performance under the stringent power constraint. To address the paradox between performance and power consumption, brain-inspired hardware, such as the IBM Neurosynaptic System, has been developed to enable low power implementation of large-scale neural models. Meanwhile, inspired by the operation and the massive parallel structure of human brain, carefully structured inference model has been demonstrated to give superior detection quality than many traditional models while facilitates neuromorphic implementation. Implementing inference based anomaly detection on the neurosynaptic processor is not straightforward due to hardware limitations. This work presents a design flow and component library that flexibly maps learned detection network to the TrueNorth architecture. Instead of traditional rate code, burst code is adopted in the design, which represents numerical value using the phase of a burst of spike trains. This does not only reduce the hardware complexity, but also increases the results accuracy. A Corelet library, NeoInfer-TN, is developed for basic operations in burst code and two-phase pipelines are constructed based on the library components. The design can be configured for different tradeoffs between detection accuracy and throughput/energy. We evaluate the system using intrusion detection data streams. The results show higher detection rate than some conventional approaches and real-time performance, with only 50mW power consumption. Overall, it achieves 10^4-8 operations per watt-second.

Download Paper (PDF; Only available from the DATE venue WiFi)

Fast, Low Power Evaluation of Elementary Functions using Radial Basis Function Networks

Speaker: Parni Wijesinghe, Purdue University, US
Authors: Parni Wijesinghe, Chandika Liyanagepura and Kaushik Roy, Purdue University, US
Abstract: Fast and efficient implementation of elementary functions such as sin(), cos(), and log() are of ample importance in a large class of applications. The state of the art methods for function evaluation involves either expensive calculations such as multiplications, large number of iterations, or large Lookup-Tables (LUTs). Higher number of iterations leads to higher latency whereas large LUTs contribute to delay, higher area requirement and higher power consumption owing to data fetching and leakage. We propose a hardware architecture for evaluating mathematical functions, consisting a small LUT and a simple Radial Basis Function Network (RRBN), a type of an Artificial Neural Network (ANN). Our proposed method evaluates trigonometric, hyperbolic, exponential, logarithmic, and square root functions. This technique finds utility in applications where the highest priority is on performance and power consumption. In contrast to traditional ANNs, our approach does not involve multiplication when determining the post synaptic states of the network. Owing to the simplicity of the approach, we were able to attain more than 2.5x power benefits and more than 1.4x performance benefits when compared with traditional approaches, under the same accuracy conditions.

Download Paper (PDF; Only available from the DATE venue WiFi)
On all conference days (Tuesday to Thursday), coffee and tea will be served during the coffee breaks at the below-mentioned times in the exhibition area.

Tuesday, March 28, 2017
- Coffee Break 10:30 - 11:30
- Coffee Break 16:00 - 17:00

Wednesday, March 29, 2017
- Coffee Break 10:00 - 11:00
- Coffee Break 16:00 - 17:00

Thursday, March 30, 2017
- Coffee Break 10:00 - 11:00
- Coffee Break 15:30 - 16:00

3.6 Mechanisms for hardware fault testing, recovery and metastability management

Date: Tuesday 28 March 2017
Time: 14:30 - 16:00
Location / Room: 5A

Chair:
Jaume Abella, Barcelona Supercomputing Center (BSC), ES

Co-Chair:
Maria K. Michael, University of Cyprus, CY

Papers in this session provide new solutions for dealing with hardware faults and metastability issues, including testing and diagnosing mechanisms for NoCs, fault recovery approaches for 3D ICs, and containment solutions for metastability in sorting networks.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>14:30</td>
<td>3.6.1</td>
<td><strong>CHARKA: A RELIABILITY-AWARE TEST SCHEME FOR DIAGNOSIS OF CHANNEL SHORTS BEYOND MESH NOCS</strong>&lt;br&gt;<strong>Speaker:</strong>&lt;br&gt;Santosh Biswas, IIT Guwahati, IN</td>
<td><strong>Authors:</strong>&lt;br&gt;Biswajit Bhowmik¹, Jatindra Kumar Deka² and Santosh Biswas²&lt;br&gt;¹IIT Guwahati, IN; ²IIT GUWAHATI, IN&lt;br&gt;<strong>Abstract</strong>&lt;br&gt;This paper presents a fast and low cost on-line scheme named &quot;Charka&quot; that analyzes short faults in channels of octagon NoCs. Experimental results demonstrate that the proposed scheme achieves 100% coverage metrics and its on-line evaluation reveals compelling effect of these faults on system performance. We observe that the proposed scheme is up to 9X faster while packet latency is improved by 13.79-21.17% and energy consumption is reduced by 17.57-24.97%. Further, the test area overhead is reduced by 13-26% that shows 52-57.77% improvement. Download Paper (PDF; Only available from the DATE venue WiFi)</td>
</tr>
<tr>
<td>15:00</td>
<td>3.6.2</td>
<td><strong>RECOVERY-AWARE PROACTIVE TSV REPAIR FOR ELECTROMIGRATION IN 3D ICs</strong>&lt;br&gt;<strong>Speaker:</strong>&lt;br&gt;Shengcheng Wang, Chair of Dependable Nano Computing (CDNC), Karlsruhe Institute of Technology (KIT), DE</td>
<td><strong>Authors:</strong>&lt;br&gt;Shengcheng Wang¹, Hengyang Zhao², Sheldon Tan³ and Mehdi Tahoori¹&lt;br&gt;¹Karlsruhe Institute of Technology, DE; ²University of California, Riverside, US; ³University of California at Riverside, US&lt;br&gt;<strong>Abstract</strong>&lt;br&gt;Electromigration (EM) becomes a major reliability concern in three-dimensional integrated-circuits (3D ICs). To mitigate this problem, a typical solution is to use TSV redundancy in a reactive manner, maintaining the operability of a 3D chip in the presence of EM failures by detecting and replacing faulty TSVs with spares. In this work, we explore an alternative, more preferred approach to enhance the EM-related lifetime reliability of TSV grid, in which redundancy is used proactively to allow non-faulty TSVs to be temporarily deactivated. In this way, EM wear-out can be reversed by exploiting its recovery property. Applied to 3D benchmark designs, the recovery-aware proactive repair approach increases EM-related lifetime reliability (measured in mean-time-to-failure) of the entire TSV grid by up to 12X relative to the conventional reactive method, with less area overhead. Download Paper (PDF; Only available from the DATE venue WiFi)</td>
</tr>
<tr>
<td>15:30</td>
<td>3.6.3</td>
<td><strong>NEAR-OPTIMAL METASTABILITY-CONTAINING SORTING NETWORKS</strong>&lt;br&gt;<strong>Speaker:</strong>&lt;br&gt;Johannes Bund, Saarland University, DE</td>
<td><strong>Authors:</strong>&lt;br&gt;Johannes Bund¹, Christoph Lenzen² and Moti Medina²&lt;br&gt;¹Saarland University, DE; ²MPI-INF, DE&lt;br&gt;<strong>Abstract</strong>&lt;br&gt;Metastability in digital circuits is a spurious mode of operation induced by violation of setup/hold times of stateful components. It cannot be avoided deterministically when transitioning from continuously-valued to (discrete) binary signals. However, in prior work (Lenzen &amp; Medina ASYNC 2016) it has been shown that it is possible to fully and deterministically contain the effect of metastability in sorting networks. More specifically, the sorting operation incurs no loss of precision, i.e., any inaccuracy of the output originates from mapping the continuous input range to a finite domain. The downside of this prior result is inefficiency: for B-bit inputs, the circuit for a single comparison contains Theta(B^2) gates and has depth Theta(B). In this work, we present an improved solution with near-optimal Theta(Blog B) gates and asymptotically optimal Theta(log B) depth. On the practical side, our sorting networks improves over prior work for all input lengths $B &gt; 2$, e.g., for 16-bit inputs we present an improvement of more than 70% w.r.t. the depth of the sorting network and more than 60% improvement w.r.t. the cost of the sorting network. Download Paper (PDF; Only available from the DATE venue WiFi)</td>
</tr>
<tr>
<td>Time</td>
<td>Label</td>
<td>Presentation Title</td>
<td>Authors</td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
<td>------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>16:00</td>
<td>IP1-16 267</td>
<td><strong>3DFAR: A THREE-DIMENSIONAL FABRIC FOR RELIABLE MULTICORE PROCESSORS</strong></td>
<td>Speaker: Valeria Bertacco, University of Michigan, US Authors: Javad Bagherzadeh and Valeria Bertacco, University of Michigan, US</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
<td>In the past decade, silicon technology trends into the nanometer regime have led to significantly higher transistor failure rates. Moreover, these trends are expected to exacerbate with future devices. To enhance reliability, several approaches leverage the inherent core-level and processor-level redundancy present in large chip multiprocessors. However, all of these methods incur high overheads, making them impractical. In this paper, we propose 3DFAR, a novel architecture leveraging 3-dimensional fabrics layouts to efficiently enhance reliability in the presence of faults. Our key idea is based on a fine-grained reconfigurable pipeline for multicore processors, which minimizes routing delay among spare units of the same type by using physical layout locality and efficient interconnect switches, distributed over multiple vertical layers. Our evaluation shows that 3DFAR outperforms state-of-the-art reliable 2D solutions, at a minimal area cost of only 7% over an unprotected design.</td>
</tr>
<tr>
<td>16:01</td>
<td>IP1-17 933</td>
<td><strong>EVALUATING IMPACT OF HUMAN ERRORS ON THE AVAILABILITY OF DATA STORAGE SYSTEMS</strong></td>
<td>Speaker: Hossein Asadi, Sharif University of Technology, IR Authors: Mostafa Kishani, Reza Eftekhari and Hossein Asadi, Sharif University of Technology, IR</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
<td>In this paper, we investigate the effect of incorrect disk replacement service on the availability of data storage systems. To this end, we first conduct Monte Carlo simulations to evaluate the availability of disk subsystem by considering disk failures and incorrect disk replacement service. We also propose a Markov model that corroborates the Monte Carlo simulation results. We further extend the proposed model to consider the effect of automatic disk fail-over policy. The results obtained by the proposed model show that overlooking the impact of incorrect disk replacement can result up to three orders of magnitude unavailability underestimation. Moreover, this study suggests that by considering the effect of human errors, the conventional believes about the dependability of different RAID mechanisms should be revised. The results show that in the presence of human errors, RAID1 can result in lower availability compared to RAID5.</td>
</tr>
<tr>
<td>16:00</td>
<td></td>
<td>End of session</td>
<td>Coffee Break in Exhibition Area On all conference days (Tuesday to Thursday), coffee and tea will be served during the coffee breaks at the below-mentioned times in the exhibition area. Tuesday, March 28, 2017 ☐ Coffee Break 10:30 - 11:30 ☐ Coffee Break 16:00 - 17:00 Wednesday, March 29, 2017 ☐ Coffee Break 10:00 - 11:00 ☐ Coffee Break 16:00 - 17:00 Thursday, March 30, 2017 ☐ Coffee Break 10:00 - 11:00 ☐ Coffee Break 15:30 - 16:00</td>
</tr>
</tbody>
</table>

### 3.7 Scheduling and Optimization

**Date:** Tuesday 28 March 2017  
**Time:** 14:30 - 16:00  
**Location / Room:** 3B  
**Chair:** Rolf Ernst, TU Braunschweig, DE  
**Co-Chair:** Kai Lampka, Uppsala University, SE  

This session focuses on methods to optimize the design of real-time embedded systems. The first two presentations cover priority assignment and task partitioning for scheduling on multi-core systems. The last long presentation and interactive presentations focus on architectural and OS considerations.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>14:30</td>
<td>3.7.1</td>
<td><strong>THE CONCEPT OF UNSCHEDULABILITY CORE FOR OPTIMIZING PRIORITY ASSIGNMENT IN REAL-TIME SYSTEMS</strong></td>
<td>Speaker: Yecheng Zhao, Virginia Polytechnic Institute and State University, US  Authors: Yecheng Zhao and Haibo Zeng, Virginia Tech, US</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
<td>In the design optimization of real-time systems, the schedulability analysis is used to define the feasibility region within which tasks meet their deadlines, so that optimization algorithms can find the best solution within the region. However, the complexity of current schedulability analysis techniques often makes it difficult to leverage existing optimization frameworks and scale to large designs. In this paper, we consider the design optimization problems for real-time systems scheduled with fixed priority, where task priority assignment is part of the decision variables. We propose the concept of unschedulability core, a compact representation of the schedulability conditions, and develop efficient algorithms for its calculation. We present a new optimization procedure based on lazy constraint paradigm that leverages such a concept. Experimental results on two case studies show that the new optimization procedure provides optimal solutions, but is a few magnitudes faster than other exact algorithms (Branch-and-Bound, Integer Linear Programming).</td>
</tr>
</tbody>
</table>

Download Paper (PDF; Only available from the DATE venue WiFi)
## Scheduling Challenges in Today's Datacenter Systems Design

**Speaker:** Saravanan Ramanathan, Nanyang Technological University, SG  
**Authors:** Saravanan Ramanathan and Arvind Easwaran, Nanyang Technological University, SG  
**Abstract**  
Mixed-Criticality (MC) systems consolidate multiple functionalities with different criticalities onto a single hardware platform. Such systems improve the overall resource utilization while guaranteeing resources to critical tasks. In this paper, we focus on the problem of partitioned multiprocessor MC scheduling, in particular the problem of designing efficient partitioning strategies. We develop two new partitioning strategies based on the principle of evenly distributing the difference between total high-critical utilization and total low-critical utilization for the critical tasks among all processors. By balancing this difference, we are able to reduce the pessimism in uniprocessor MC schedulability tests that are applied on each processor, thus improving overall schedulability. To evaluate the schedulability performance of the proposed strategies, we compare them against existing partitioned algorithms using extensive experiments. We show that the proposed strategies are effective with both dynamic-priority Earliest Deadline First with Virtual Deadlines (EDF-VD) and fixed-priority Adaptive Mixed-Criticality (AMC) algorithms. Specifically, our results show that the proposed strategies improve schedulability by as much as 28.1% and 36.2% for implicit and constrained-deadline task systems respectively.  
[Download Paper (PDF; Only available from the DATE venue WiFi)]

### Conference Schedule

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>15:00</td>
<td>3.7.2</td>
<td>UTILIZATION DIFFERENCE BASED PARTITIONED SCHEDULING OF MIXED-CRITICALITY SYSTEMS</td>
<td>Saravanan Ramanathan, Nanyang Technological University, SG; Saravanan Ramanathan and Arvind Easwaran, Nanyang Technological University, SG</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.7.3</td>
<td>SCHEDULABILITY USING NATIVE NON-PREEMPTIVE GROUPS ON AN AUTOSAR/OSEK PLATFORM WITH CACHES</td>
<td>Leo Hatvani, Technische Universität Eindhoven, NL; Reinder J. Bril and Sebastian Altmeyer</td>
</tr>
<tr>
<td>16:00</td>
<td>IP1-18</td>
<td>GPUGUARD: TOWARDS SUPPORTING A PREDICTABLE EXECUTION MODEL FOR HETEROGENEOUS SOC</td>
<td>Björn Forsberg, ETH Zürich, CH; Andrea Marongiu and Luca Benini</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>16:01</td>
<td>A NON-INTRUSIVE, OPERATING SYSTEM INDEPENDENT SPINLOCK PROFILER FOR EMBEDDED MULTICORE SYSTEMS</td>
<td>Lin Li, Infineon Technologies, DE; Thomas Wild and Andreas Herkersdorf</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16:00</td>
<td></td>
<td>End of session</td>
<td></td>
</tr>
</tbody>
</table>

**Coffee Break** in Exhibition Area  
On all conference days (Tuesday to Thursday), coffee and tea will be served during the coffee breaks at the below-mentioned times in the exhibition area.

- **Tuesday, March 28, 2017**  
  - Coffee Break 10:30 - 11:30  
  - Coffee Break 16:00 - 17:00

- **Wednesday, March 29, 2017**  
  - Coffee Break 10:00 - 11:00  
  - Coffee Break 16:00 - 17:00

- **Thursday, March 30, 2017**  
  - Coffee Break 10:00 - 11:00  
  - Coffee Break 15:30 - 16:00
3.8.1 SERVER BENCHMARKING AND DESIGN WITH CLOUDSUITE 3.0
Speaker: Javier Picorel, EPFL, CH
Abstract
Since its inception, CloudSuite (cloudsuite.ch) has emerged as a popular suite of benchmarks both in industry and among academics for the performance evaluation of cloud services. The EuroCloud Server project blueprinted key optimizations in server SoCs based on the salient features of CloudSuite benchmarks that lead to an order of magnitude improvement in efficiency while preserving QoS. ARM-based server products (e.g., Cavium ThunderX) have now emerged following these guidelines and showcasing the improved efficiency. CloudSuite 3.0 is a major enhancement over prior releases both in benchmarks and infrastructure. It includes benchmarks that represent massive data manipulation with tight latency constraints such as in-memory data analytics using Apache Spark, a new real-time video streaming benchmark following today’s most popular video-sharing website setups, and a new web serving benchmark mirroring today’s multi-tier web server software stacks. To ease the deployment of CloudSuite into private and public cloud systems, the benchmarks are integrated into the Docker software container system and Google’s PerfKit Benchmark. Docker wraps each benchmark into a self-contained software package, guaranteeing the same execution regardless of the environment, while PerfKit automates the process of benchmarking cloud server systems with CloudSuite. CloudSuite 3.0 is supported to run both on real hardware and on our QEMU-based computer architecture simulation framework.

3.8.2 PROTECTING DATA IN FARM AND RDMA NETWORKS WITH CATAPULT
Speaker: Greg O’Shea, Microsoft, US
Abstract
FaRM is an in-memory, transactional database that runs distributed across a cluster of Windows Servers that are connected by a high-speed Remote Direct Memory Access (RDMA) network. Data in FaRM are stored in DRAM and exposed directly to the L2 network by the server’s RDMA network adapters, so that other members of the FaRM cluster can access the data with great efficiency. RDMA enables a network adapter to directly access the memory of another server in the same Ethernet network bypassing the operating system in both servers. This enables low-latency and high-bandwidth data access across the entire cluster. However, RDMA provides no security: the data are also accessible to every other server attached to the same Ethernet network, and message transfers are vulnerable to replay and modification. We present our work to protect data in FaRM using a bump-in-the-wire firewall for RDMA. Based upon the FPGA cards widely deployed in Windows Servers within Microsoft, the firewall exists as a barrier between a FaRM server’s RDMA adapter and the local Ethernet switch. It prevents packets from outside the FaRM cluster from ever reaching the server’s RDMA adapter, and it protects RDMA packets between members of the FaRM cluster by encapsulating them in DTLS tunnels. We show that implementing a similar level of protection in software can be prohibitively expensive.

16:00 End of session
Coffee Break in Exhibition Area
On all conference days (Tuesday to Thursday), coffee and tea will be served during the coffee breaks at the below-mentioned times in the exhibition area.

Tuesday, March 28, 2017
- Coffee Break 10:30 - 11:30
- Coffee Break 16:00 - 17:00

Wednesday, March 29, 2017
- Coffee Break 10:00 - 11:00
- Coffee Break 16:00 - 17:00

Thursday, March 30, 2017
- Coffee Break 10:00 - 11:00
- Coffee Break 15:30 - 16:00

3.9 A tribute to Ralph Otten
Date: Tuesday 28 March 2017
Time: 14:30 - 16:00
Location / Room: Auditorium A
Organiser: Giovanni De Micheli, EPFL, CH
Chair: Michael Burstein, CEO Billy.com, CA
Co-Chair: Giovanni De Micheli, EPFL, CH
Ralph Otten 
World renowned leaders in Physical Design will talk about accomplishments in this field over the last four decades, as a tribute to Ralph Otten, pioneer of this field and prematurely died in an accident.

14:30 3.9.1 CHIP DESIGN - PHYSICAL AND PHILOSOPHICAL
Author: Dave Liu, NTHU, TW

14:45 3.9.2 AUTOMATIC FLOORPLAN DESIGN
Author: Martin Wong, University of Illinois at Urbana Champaign, US
THE EVOLUTION OF FLOORPLANNING
Author: Antun Domic, Synopsys, US

FROM SILICON COMPILER TO PHYSICAL SYNTHESIS: RALPH OTTEN’S CONTRIBUTIONS TO EDA
Author: Patrick Groeneweld, Synopsys, US

DEALING WITH EXPLODING DESIGN RULE NUMBERS AND COMPLEXITY
Author: Raul Camposano, Sage Design Automation, US

IN MEMORIAM OF RALPH OTTEN: BREAKING DOWN THE COMPLEXITY OF LAYOUT DESIGN UNDER MOORE’S LAW
Author: Jochen Jess, Eindhoven University of Technology, NL

End of session
Coffee Break in Exhibition Area
On all conference days (Tuesday to Thursday), coffee and tea will be served during the coffee breaks at the below-mentioned times in the exhibition area.

Tuesday, March 28, 2017
- Coffee Break 10:30 - 11:30
- Coffee Break 16:00 - 17:00

Wednesday, March 29, 2017
- Coffee Break 10:00 - 11:00
- Coffee Break 16:00 - 17:00

Thursday, March 30, 2017
- Coffee Break 10:00 - 11:00
- Coffee Break 15:30 - 16:00
UB03.4  MATISSE: A TARGET-AWARE COMPILER TO TRANSLATE MATLAB INTO C AND OPENCL

Presenter:
Luís Reis, University of Porto, PT

Authors:
João Bispo and João Cardoso, University of Porto / INESC-TEC, PT

Abstract
Many engineering, scientific and finance algorithms are prototyped and validated in array languages, such as MATLAB, before being converted to other languages such as C for use in production. As such, there has been substantial effort to develop compilers to perform this translation automatically. Alternative types of computation devices, such as GPGPUs and FPGAs, are becoming increasingly more popular, so it becomes critical to develop compilers that target these architectures. We have adapted MATISSE, our MATLAB-compatible compiler framework, to generate C and OpenCL code for these platforms. In this demonstration, we will show how our compiler works and what its capabilities are. We will also describe the main challenges of efficient code generation from MATLAB and how to overcome them.

More information ...

UB03.5  A VOLTAGE-SCALABLE FULLY DIGITAL ON-CHIP MEMORY FOR ULTRA-LOW-POWER IOT PROCESSORS

Presenter:
Jun Shiomi, Kyoto University, JP

Authors:
Toshiu Ishihara and Hitotoshi Onodera, Kyoto University, JP

Abstract
A voltage-scalable RISC processor integrating standard-cell based memory (SCM) is demonstrated. Unlike conventional processors, the processor has Standard-Cell based Memories (SCMs) as an alternative to conventional SRAM macros, enabling it to operate at a 0.4 V single-supply voltage. The processor is implemented with the fully automated cell-based design, which leads to low design costs. By scaling the supply voltage and applying the back-gate biasing techniques, the power dissipation of the SCMs is less than 20 uW, enabling the SCMs to operate with ambient energy source only. In this demonstration, the SCMs of the processor operates with a lemon battery as the ambient energy source.

More information ...

UB03.6  RUNNING CONVOLUTIONAL LAYERS OF ALEXNET IN NEUROMORPHIC COMPUTING SYSTEM

Presenter:
Yongshin Kang, Incheon National University, KR

Authors:
Seban Kim, Taehwan Shin and Jaeyong Chung, Incheon National University, KR

Abstract
Neuromorphic hardware has drawn attention as an approach to deal with the issues of today’s computing platforms based on Von Neumann architecture when running deep learning models, but large-scale deep neural networks such as AlexNet have not been demonstrated yet in any neuromorphic systems. Since 2014, we have been developing a non-Von-Neumann computing system called INSight based on data flow architecture that aims at running large-scale deep neural networks in the neuromorphic fashion. We have now reached a major milestone and will demonstrate INSight running the convolutional layers of AlexNet. The proposed system is implemented with Xilinx Virtex 7 FPGA and performs the processing using 100K synapses mapped on LUTs without any array-type memories. It processes 1552 images per second and consumes 7.2W, resulting in the state-of-the-art energy efficiency.

More information ...

UB03.7  ACCELERATORS: RECONFIGURABLE SELF-TIMED DATAFLOW ACCELERATOR & FAST NETWORK ANALYSIS IN SILICON

Presenter:
Alessandro de Gennaro, Newcastle University, GB

Authors:
Danil Sokolov and Andrey Mokhov, Newcastle University, GB

Abstract
Many real-life applications require dynamically reconfigurable pipelines to handle incoming data items differently depending on their values or current operating mode. A demo will show the benefits of an asynchronous accelerator for ordinal pattern encoding with reconfigurable pipeline depth. This was designed, simulated and verified using dataflow structure formalism in Workcraft toolset. The self-timed chip, fabricated in TSMC 90nm, shows high resilience to voltage variation and configurable accuracy of the results. Applications with underlying graph models foster the importance of a fast and flexible approach to graph analysis. To support medicine discovery biological systems are modelled by graphs, and drugs can disconnect some of the connections. A demo will show how graphs can be automatically converted into VHDL designs, which are synthesised into a FPGA for the analysis: thousand times faster than in software. Single stand will be used for both case studies.

More information ...

UB03.8  NETWORKED LABS-ON-CHIPS

Presenter:
Andreas Grimmer, Johannes Kepler University Linz, AT

Authors:
Werner Haselmayr, Andreas Springer and Robert Wille, Johannes Kepler University Linz, AT

Abstract
Applications with underlying graph models foster the importance of a fast and flexible approach to graph analysis. To support medicine discovery biological systems are modelled by graphs, and drugs can disconnect some of the connections. A demo will show how graphs can be automatically converted into VHDL designs, which are synthesised into a FPGA for the analysis: thousand times faster than in software. Single stand will be used for both case studies.

More information ...

UB03.9  STACKADROP: A MODULAR DIGITAL MICROFLUIDIC BIOCHIP RESEARCH PLATFORM

Presenter:
Oliver Keszöcze, University of Bremen, DE

Authors:
Maximilian Luernert and Rolf Drechsler, University of Bremen & DFKI GmbH, DE

Abstract
Advances in microfluidic technologies have led to the emergence of Digital Microfluidic Biochips (DMFBs), which are capable of automating laboratory procedures. These DMFBs raised significant attention in industry and academia creating a demand for devices. Commercial products are available but come at a high price. So far, there are two open hardware DMFBs available: the DropBot from WheelerLabs and the OpenDrop from GaudiLabs. The aim of the StackADrop was to create a DMFB with many directly addressable cells while still being very compact. The StackADrop strives to provide means to experiment with different hardware setups. It’s main feature are the exchangeable top plates, supporting 256 high-voltage pins. It features SPI, UART and I2C connectors for attaching sensors/actuators and can be connected to a computer using USB for interactive sessions using a control software. The modularity allows to easily test different cell shapes, such as squares, hexagons and triangles.

More information ...
**PULP: A ULTRA-LOW POWER PLATFORM FOR THE INTERNET-OF-THINGS**

Presenter:
Francesco Conti, ETH Zurich, CH

Authors:
Stefan Mach1, Florian Zaruba1, Antonio Pullini2, Daniele Palossi1, Giovanni Rovere1, Florian Glaser1, Germain Haugou1, Schekeb Fatehi1 and Luca Benini2

1ETH Zurich, CH; 2ETH Zurich, CH and University of Bologna, IT

Abstract
The PULP (Parallel Ultra-Low Power) platform strives to provide high performance for IoT nodes and endpoints within a very small power envelope. The PULP platform is based on a tightly-coupled multi-core cluster and on a modular architecture, which can support complex configurations with autonomous I/O without SW intervention, HW-accelerated execution of hot computation kernels, fine-grain event-based computation - but can also be deployed in very simple configuration, such as the open source PULPino microcontroller. In this demonstration booth, we will showcase several prototypes using PULP chips in various configuration. Our prototypes perform demos such as real-time deep-learning based visual recognition from a low-power camera, and online biosignal acquisition and reconstruction on the same chip. Application scenarios for our technology include healthcare wearables, autonomous nano-UAVs, smart networked environmental sensors.

**IP1 Interactive Presentations**

**Date:** Tuesday 28 March 2017

**Time:** 16:00 - 16:30

**Location / Room:** IP sessions (in front of rooms 4A and 5A)

Interactive Presentations run simultaneously during a 30-minute slot. A poster associated to the IP paper is on display throughout the afternoon. Additionally, each IP paper is briefly introduced in a one-minute presentation in a corresponding session, prior to the actual Interactive Presentation. At the end of each afternoon Interactive Presentations session the award ‘Best IP of the Day’ is given.

**Table: Presenters and Details**

<table>
<thead>
<tr>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>IP1-1</td>
<td>STRUCTURAL DESIGN OPTIMIZATION FOR DEEP CONVOLUTIONAL NEURAL NETWORKS USING STOCHASTIC COMPUTING</td>
<td>Yanzhi Wang, Syracruse University, US</td>
</tr>
<tr>
<td>Authors</td>
<td>Zhe Li1, Ji Li2, Jinru Qiu1, Bo Yuan1, Jeffrey Draper2 and Yanzhi Wang1</td>
<td></td>
</tr>
<tr>
<td>1Syracuse University, US; 2University of Southern California, US;</td>
<td>3City University of New York, New York, City College, US</td>
<td></td>
</tr>
</tbody>
</table>

**Abstract**
Deep Convolutional Neural Networks (DCNNs) have been demonstrated as effective models for understanding image content. The computation behind DCNNs highly relies on the capability of hardware resources due to the deep structure. DCNNs have been implemented on different large-scale computing platforms. However, there is a trend that DCNNs have been embedded into light-weight local systems, which requires low power/energy consumptions and small hardware footprints. Stochastic Computing (SC) radically simplifies the hardware implementation of arithmetic units and has the potential to satisfy the small low-power needs of DCNNs. Local connectivities and down-sampling operations have made DCNNs more complex to be implemented using SC. In this paper, eight feature extraction designs for DCNNs using SC in two groups are explored and optimized in detail from the perspective of calculation precision, where we permute two SC implementations for inner-product calculation, two down-sampling schemes, and two structures of DCNN neurons. We evaluate the network in aspects of network accuracy and hardware performance for each DCNN using one feature extraction design out of eight. Through exploration and optimization, the accuracies of SC-based DCNNs are guaranteed compared with software implementations on CPU/GPU/binary-based ASIC synthesis, while area, power, and energy are significantly reduced by up to 776X.

**IP1-2**

**APPROXQA: A UNIFIED QUALITY ASSURANCE FRAMEWORK FOR APPROXIMATE COMPUTING**

Speaker:
Ting Wang, The Chinese University of Hong Kong, HK

Authors:
Ting Wang, Qian Zhang and Qiang Xu, The Chinese University of Hong Kong, HK

Abstract
Approximate computing, being able to trade off computation quality and computational effort (e.g., energy) by exploiting the inherent error-resilience of emerging applications (e.g., recognition and mining), has garnered significant attention recently. No doubt to say, quality assurance is indispensable for satisfactory user experience with approximate computing, but this issue has remained largely unexplored in the literature. In this work, we propose a novel framework namely ApproxQA to tackle this problem, in which approximation mode tuning and rollback recovery are considered in a unified manner when quality violation occurs. To be specific, ApproxQA resorts to a two-level controller, in which the high-level approximation controller tunes approximation modes at a coarse-grained scale based on Q-learning while the low-level rollback controller judiciously determines whether to perform rollback recovery at a fine-grained scale based on the target quality requirement. ApproxQA can provide statistical quality assurance even when the underlying quality checkers are not reliable. Experimental results on various benchmark applications demonstrate that it significantly outperforms existing solutions in terms of energy efficiency with quality assurance.

Download Paper (PDF; Only available from the DATE venue WiFi)

**IP1-3**

**EVOAPPROX8B: LIBRARY OF APPROXIMATE ADDERS AND MULTIPLIERS FOR CIRCUIT DESIGN AND BENCHMARKING OF APPROXIMATION METHODS**

Speaker:
Yanzhi Wang, Syracuse University, US

Authors:
Yanzhi Wang, Ji Li, Jinru Qiu, Bo Yuan, Jeffrey Draper, and Yanzhi Wang

Abstract
Approximate circuits and approximate circuit design methodologies attracted a significant attention of researchers as well as industry in recent years. In order to accelerate the approximate circuit and system design process and to support a fair benchmarking of circuit approximation methods, we propose a library of approximate adders and multipliers called EVOAPPROX8B. This library contains 430 non-dominated 8-bit approximate adders created from 13 conventional adders and 471 non-dominated 8-bit approximate multipliers created from 6 conventional multipliers. These implementations were evolved by a multi-objective Cartesian genetic programming. The EVOAPPROX8B library provides Verilog, Matlab and C models of all approximate circuits. In addition to standard circuit parameters, the error is given for seven different error metrics. The EVOAPPROX8B library is available at: www.fit.vutbr.cz/research/groups/etn2/approxlib

Download Paper (PDF; Only available from the DATE venue WiFi)
packages of the project are presented in this paper. These packages are on logic synthesis that aims to implement Boolean functions with nano-crossbar arrays with performance optimization methodology for switching nano-crossbar arrays that leads to the design and construction of an emerging nanocomputer. First two work integrate well developed conventional circuit design techniques into nano-crossbar arrays. Motivated by this, our project aims to develop a complete synthesis and such that each crosspoint can be used as a conventional electronic component such as a diode, a FET, or a switch. This is a unique opportunity that allows us to.

Nano-crossbar arrays have emerged as a strong candidate technology to replace CMOS in near future. They are regular and dense structures, and can be fabricated with reference to a cycle accurate instruction set simulator leading to an average cycle count error of 9.3% and an average speedup of 10.

The amount of current needed to reorient the magnetization at present (~100μA per bit) is too high, especially for the Write operation. When we perform a full cache line (512-bit) Write, this extremely high current compared to MRAM will result in a Voltage drop in the conventional cache architecture. Due to this droop, the write operation will fail halfway through when we attempt to write in the farthest Bank of the cache from the supply. In this paper, we will be proposing a new cache architecture to mitigate this problem of droop and make the write operation successful. Instead of continuously writing the entire Cache line (512-bit) in a single bank, our architecture will be writing these 512-bits in multiple different locations across the cache in parts of 8 (64-bit each).

In this work, we propose an icache performance estimation approach that focuses on a component necessary to handle the instruction parallelism in a very long instruction word (VLIW) processor: the instruction buffer (IB). Our annotation approach is founded on an intermediate level native- simulation framework. It is evaluated with reference to a cycle accurate instruction set simulator leading to an average cycle count error of 9.3% and an average speedup of 10.

High levels of integration in SoCs and SoPs is making pre as well as post-silicon validation of mixed-signal systems increasingly difficult due to: (a) lack of automated pre and post-silicon design checking algorithms and (b) lack of controllability and observability of internal circuit nodes in post-silicon. While digital scan chains provide models from the circuit to the functional level, 3) builds an efficient fault simulation framework. The processes of fault injection, faulty model abstraction and pre and post-silicon design checking algorithms and (b) lack of controllability and observability of internal circuit nodes in post-silicon. While digital scan chains provide models from the circuit to the functional level, 3) builds an efficient fault simulation framework. The processes of fault injection, faulty model abstraction and framework generation are reported in details, as well as how simulation is carried out. This abstraction process, which preserves the faulty behaviors, allows to reach a speed-up of some orders of magnitude and thus, making feasible an extensive analog faults campaign.

In this work, we propose an icache performance estimation approach that focuses on a component necessary to handle the instruction parallelism in a very long instruction word (VLIW) processor: the instruction buffer (IB). Our annotation approach is founded on an intermediate level native- simulation framework. It is evaluated with reference to a cycle accurate instruction set simulator leading to an average cycle count error of 9.3% and an average speedup of 10.

In this work, we propose an icache performance estimation approach that focuses on a component necessary to handle the instruction parallelism in a very long instruction word (VLIW) processor: the instruction buffer (IB). Our annotation approach is founded on an intermediate level native- simulation framework. It is evaluated with reference to a cycle accurate instruction set simulator leading to an average cycle count error of 9.3% and an average speedup of 10.

In this work, we propose an icache performance estimation approach that focuses on a component necessary to handle the instruction parallelism in a very long instruction word (VLIW) processor: the instruction buffer (IB). Our annotation approach is founded on an intermediate level native- simulation framework. It is evaluated with reference to a cycle accurate instruction set simulator leading to an average cycle count error of 9.3% and an average speedup of 10.

In this work, we propose an icache performance estimation approach that focuses on a component necessary to handle the instruction parallelism in a very long instruction word (VLIW) processor: the instruction buffer (IB). Our annotation approach is founded on an intermediate level native- simulation framework. It is evaluated with reference to a cycle accurate instruction set simulator leading to an average cycle count error of 9.3% and an average speedup of 10.

In this work, we propose an icache performance estimation approach that focuses on a component necessary to handle the instruction parallelism in a very long instruction word (VLIW) processor: the instruction buffer (IB). Our annotation approach is founded on an intermediate level native- simulation framework. It is evaluated with reference to a cycle accurate instruction set simulator leading to an average cycle count error of 9.3% and an average speedup of 10.

In this work, we propose an icache performance estimation approach that focuses on a component necessary to handle the instruction parallelism in a very long instruction word (VLIW) processor: the instruction buffer (IB). Our annotation approach is founded on an intermediate level native- simulation framework. It is evaluated with reference to a cycle accurate instruction set simulator leading to an average cycle count error of 9.3% and an average speedup of 10.

In this work, we propose an icache performance estimation approach that focuses on a component necessary to handle the instruction parallelism in a very long instruction word (VLIW) processor: the instruction buffer (IB). Our annotation approach is founded on an intermediate level native- simulation framework. It is evaluated with reference to a cycle accurate instruction set simulator leading to an average cycle count error of 9.3% and an average speedup of 10.

In this work, we propose an icache performance estimation approach that focuses on a component necessary to handle the instruction parallelism in a very long instruction word (VLIW) processor: the instruction buffer (IB). Our annotation approach is founded on an intermediate level native- simulation framework. It is evaluated with reference to a cycle accurate instruction set simulator leading to an average cycle count error of 9.3% and an average speedup of 10.

In this work, we propose an icache performance estimation approach that focuses on a component necessary to handle the instruction parallelism in a very long instruction word (VLIW) processor: the instruction buffer (IB). Our annotation approach is founded on an intermediate level native- simulation framework. It is evaluated with reference to a cycle accurate instruction set simulator leading to an average cycle count error of 9.3% and an average speedup of 10.

In this work, we propose an icache performance estimation approach that focuses on a component necessary to handle the instruction parallelism in a very long instruction word (VLIW) processor: the instruction buffer (IB). Our annotation approach is founded on an intermediate level native- simulation framework. It is evaluated with reference to a cycle accurate instruction set simulator leading to an average cycle count error of 9.3% and an average speedup of 10.

In this work, we propose an icache performance estimation approach that focuses on a component necessary to handle the instruction parallelism in a very long instruction word (VLIW) processor: the instruction buffer (IB). Our annotation approach is founded on an intermediate level native- simulation framework. It is evaluated with reference to a cycle accurate instruction set simulator leading to an average cycle count error of 9.3% and an average speedup of 10.
SECURECLOUD: SECURE BIG DATA PROCESSING IN UNTRUSTED CLOUDS
Speaker: Rafael Pires, University of Neuchâtel, CH

Abstract
We present the SecureCloud EU Horizon 2020 project, whose goal is to enable new big data applications that use sensitive data in the cloud without compromising data security and privacy. For this, SecureCloud designs and develops a layered architecture that allows for (i) the secure creation and deployment of secure microservices; (ii) the secure integration of individual microservices to full-fledged big data applications; and (iii) the secure execution of these applications within untrusted cloud environments. To provide security guarantees, SecureCloud leverages novel security mechanisms present in recent commodity CPUs, in particular, Intel’s Software Guard Extensions (SGX). SecureCloud applies this architecture to big data applications in the context of smart grids. We describe the SecureCloud approach, initial results, and considered use cases.

Download Paper (PDF; Only available from the DATE venue WiFi)

WCET-AWARE PARALLELIZATION OF MODEL-BASED APPLICATIONS FOR MULTI-CORES: THE ARGO APPROACH
Speaker: Steven Derrien, Université de Rennes 1, FR

Authors:
Steven Derrien1, Isabelle Puata2, Panayiotis Aliefagias3, Marcus Bednara4, Harald Bucher5, Clément David6, Yann Debray7, Umut Durağ7, Imen Fassi2, Christian Ferdinand8, Damien Hardy2, Angeliki Kritikakou2, Gerard Rauwerda5, Simon Reder5, Martin Sicks5, Timo Striffl3, Kim Sunesen5, Timon ter Braak9, Nikolaos Voros3 and Jürgen Becker5

1IRISA, FR; 2University of Rennes 1 / IRISA, FR; 3TGW, GR; 4IIS/Franhofer, DE; 5Karlsruhe Institute of Technology, DE; 6Scilab, FR; 7DLR, DE; 8Absint, FR; 9Recore, systems, FR

Abstract
Parallel architectures are nowadays not only confined to the domain of high performance computing, they are also increasingly used in embedded time-critical systems. The ARGO H2020 project provides a programming paradigm and associated tool flow to exploit the full potential of architectures in terms of development productivity, time-to-market, exploitation of the platform computing power and guaranteed real-time performance. In this paper we give an overview of the objectives of ARGO and explore the challenges introduced by our approach.

Download Paper (PDF; Only available from the DATE venue WiFi)

EXPLORING THE UNKNOWN THROUGH SUCCESSIVE GENERATIONS OF LOW POWER AND LOW RESOURCE VERSATILE AGENTS
Speaker: Martin Andraud, Eindhoven University of Technology, NL

Authors:
Martin Andraud1 and Marian Verhelst2
1Eindhoven University of Technology, NL; 2Katholieke Universiteit Leuven, BE

Abstract
The Phoenix project aims to develop a new approach to explore unknown environments, based on multiple measurement campaigns carried out by extremely tiny devices, called agents, that gather data from multiple sensors. These low power and low resource agents are configured specifically for each measurement campaign to achieve the exploration goal in the smallest number of iterations. Thus, the main design challenge is to build agents as much reconfigurable as possible. This paper introduces the Phoenix project in more details and presents first developments in the agent design.

Download Paper (PDF; Only available from the DATE venue WiFi)

POWER PROFILING OF MICROCONTROLLER'S INSTRUCTION SET FOR RUNTIME HARDWARE TROJANS DETECTION WITHOUT GOLDEN CIRCUIT MODELS
Speaker: Falah Awwad, College of Engineering / Department of Electrical Engineering, UAE University, AE

Authors:
Falah Awwad1, Syed Rafay Hasan2, Osman Hasan2 and Falah Awwad1
1School of Electrical Engineering and Computer Science National University of Sciences and Technologies (NUST), PK; 2Department of Electrical and Computer Engineering, Tennessee Technological University, US; 3College of Engineering, United Arab Emirates University, AE

Abstract
Globalization trends in integrated circuit (IC) design are leading to increased vulnerability of ICs against hardware Trojans (HT). Recently, several side channel parameters based techniques have been developed to detect these hardware Trojans that require golden circuit as a reference model, but due to the widespread usage of IP and most of the system-on-chip (SoC) do not have a golden reference system to achieve HT detection, HTs can be easily introduced to ICs. In this paper, we present a power profile based HT detection technique designed for the embedded system, which can detect the HTs without the need for the golden circuit.

Download Paper (PDF; Only available from the DATE venue WiFi)

ACCOUNTING FOR SYSTEMATIC ERRORS IN APPROXIMATE COMPUTING
Speaker: Martin Bruestel, Technical University Dresden, DE

Authors:
Martin Bruestel1 and Akash Kumar2
1Technical University Dresden, DE; 2Technische Universität Dresden, DE

Abstract
Approximate computing is gaining more and more attention as potential solution to the problem of increasing energy demand in computing. Several recent works focus on the application of deterministic approximate computing to arithmetic computations. Circuits for addition and multiplication are simplified, trading exactness for energy and/or speed. Recent approximation techniques for adders focus on modifications of individual full adders’ truth tables or shortening carry chains. While the resulting error is usually characterized with statistical measures over the range of possible input/output combinations, the actual adder is a static nonlinear system regarding arithmetic operations and signal processing. The resulting unexpected effects present a challenge for adopting approximate computing as a widespread and standard application-level optimization technique. This paper focuses on the deterministic effects of approximate multi-bit adders, which are especially evident for certain input data in an otherwise well specified systems, showing the necessity to look beyond purely statistical measures. We show which fundamental principles are violated depending on the chosen approximation scheme, and how this choice affects practical applications. This can serve as a basis for designers to make informed decisions about the use of approximate adders at the application level.

Download Paper (PDF; Only available from the DATE venue WiFi)
**GPUGUARD: TOWARDS SUPPORTING A PREDICTABLE EXECUTION MODEL FOR HETEROGENEOUS SOC**

**Speaker:** Björn Forsberg, ETH Zürich, CH

**Authors:**
- Björn Forsberg, ETH Zürich, CH
- Andrea Marongiu, University of Potsdam, DE
- Luca Benini, University of Bologna, IT

**Abstract**

The deployment of real-time workloads on commercial off-the-shelf (COTS) hardware is attractive, as it reduces the cost and time-to-market of new products. Most modern high-end embedded SoCs rely on a heterogeneous design, coupling a general-purpose multi-core CPU to a massively parallel accelerator, typically a programmable GPU, sharing a single global DRAM. However, because of non-predictable hardware arbiters designed to maximize average or peak performance, it is very difficult to provide timing guarantees on such systems. In this work we present our ongoing work on GPUguard, a software technique that predictably arbitrates main memory usage in heterogeneous SoCs. A prototype implementation for the NVIDIA Tegra TX1 SoC shows that GPUguard is able to reduce the adverse effects of memory sharing, while retaining a high throughput on both the CPU and the accelerator.

**Download Paper (PDF; Only available from the DATE venue WiFi)**
4.1 IT&A Session: The Emergence of Silicon Photonics: From High Performance Computing to Data Centers and Quantum Computing

Date: Tuesday 28 March 2017
Time: 17:00 - 18:30
Location / Room: SBC
Organiser: Luca Carloni, Columbia University, US
Chair: Luca Carloni, Columbia University, US

Recent years have seen major progress in the design and manufacturing of silicon photonics devices. This session provides an overview of the potential that this emerging technology offers for different types of systems and discusses the most important challenges that remain to be addressed. The first talk shows how silicon photonics components can be used to realize energy-efficient high-bandwidth optical interconnection networks. The second talk presents which further advances in manufacturing, packages and testing are needed in order to realize silicon photonics based products for data centers. Finally, the last talk explains how the generation of optical quantum states on an integrated platform can enable future practical implementations of quantum information processing systems.

18:30 End of session

Exhibition Reception in Exhibition Area
The Exhibition Reception will take place on Tuesday in the exhibition area, where free drinks for all conference delegates and exhibition visitors will be offered. All exhibitors are welcome to also provide drinks and snacks for the attendees.

4.2 Logic, Interconnects, Neurons: New Realizations

Date: Tuesday 28 March 2017
Time: 17:00 - 18:30
Location / Room: 48C
This session covers papers showing new approaches to realize optimized logic circuit using silicon nanowire reconfigurable transistors; intra- and inter-core optoelectronic interconnects for energy efficient communications; and magnetic skyrmions as novel nanoelectronic device for non-linear neuron networks.

**COMPACT MODELING AND CIRCUIT-LEVEL SIMULATION OF SILICON NANOPHOTONIC INTERCONNECTS**

**Authors:**
Rui Wu, Yuyang Wang, Zeyu Zhang, Chong Zhang, Clint Schow, John Bowers and Kwang-Ting Cheng, UC Santa Barbara, US

**Abstract**
Nanophotonic interconnects have been playing an increasingly important role in the datacom regime. Greater integration of silicon photonics demands photonic devices with ultra-small size (sub-10nm), high speed (>100m/s), with ultra-low depinning current density (MA/cm^2) and high defect tolerance compared to conventional magnetic domain wall motion devices. In this work, to the best of our knowledge, we are the first to propose a next-generation information carrier owing to ultra-small size (sub-10nm), high speed (>100m/s) with ultra-low depinning current density (MA/cm^2) and high defect tolerance compared to conventional magnetic domain wall motion devices. In this work, we develop compact models for a number of key photonic devices, which are extensively validated by the measurement data of a fabricated optical network-on-chip (ONoC). Implemented in SPICE-modeling and simulation support for design validation, optimization and design space exploration. In this work, we develop compact models for a number of key photonic devices, which are extensively validated by the measurement data of a fabricated optical network-on-chip (ONoC). Implemented in SPICE-

**Download Paper (PDF; Only available from the DATE venue WiFi)**

---

**AUTOMATIC PLACE-AND-ROUTE OF EMERGING LED-DRIVEN WIRES WITHIN A MONOLITHICALLY-INTEGRATED CMOS+III-V PROCESS**

**Authors:**
Tushar Krishna, Georgia Institute of Technology, US

**Abstract**
We leverage a recently demonstrated CMOS compatible III-V and Si monolithic integrated process to design photonic links comprising LEDs and photodiodes, as direct replacements for on-chip electrical wires. To enable VLSI-scale design of chips with such LED links, we create a library of opto-electronic standard cells, and model waveguides as traditional metal layers. This lets us integrate LED links into a commercial place-and-route tool, which treats them as electrical cells and wires for the most part, reducing design effort. We also add support for automated replacement of electrical nets with LED links. We find that LED-interconnect based designs substantially lower energy consumption vs. electrical copper wires (~39% reduction in the Network-on-Chip, ~27% reduction within a processor core) while achieving the same latency and bandwidth, demonstrating the promise of LED on-chip interconnects.

**Download Paper (PDF; Only available from the DATE venue WiFi)**

---

**A TUNABLE MAGNETIC SKYRMION NEURON CLUSTER FOR ENERGY EFFICIENT ARTIFICIAL NEURAL NETWORK**

**Authors:**
Zhezhi He, University of Central Florida, US

**Abstract**
Artificial neuron is one of the fundamental computing unit in brain-inspired artificial neural network. The standard CMOS based artificial neuron designs to implement non-linear neuron activation function typically consist of large number of transistors, which inevitably causes large area and power consumption. There is a need for novel nanoelectronic device that can intrinsically and efficiently implement such complex non-linear neuron activation function. Magnetic skyrmions are topologically stable chiral spin textures due to Dzyaloshinskii-Moriya interaction in bulk magnets or magnetic thin films. They are promising next-generation information carrier owing to ultra-small size (sub-10nm), high speed (>100m/s) with ultra-low depinning current density (MA/cm^2) and high defect tolerance compared to conventional magnetic domain wall motion devices. In this work, to the best of our knowledge, we are the first to propose a threshold-tunable artificial neuron based on magnetic skyrmion. Meanwhile, we propose a Skyrmion Neuron Cluster (SNC) to approximate non-linear soft-limiting neuron activation functions, such as the most popular sigmoid function. The device to system simulation indicates that our proposed SNC leads to 98.74% recognition accuracy in deep learning Convolutional Neural Network (CNN) with MNIST handwritten digits dataset. Moreover, the energy consumption of our proposed SNC is only 3.1 fJ/step, which is more than two orders lower than that of CMOS counterpart.

**Download Paper (PDF; Only available from the DATE venue WiFi)**

---

**COMPACT MODELING AND CIRCUIT-LEVEL SIMULATION OF SILICON NANOPHOTONIC INTERCONNECTS**

**Authors:**
Yuyang Wang, UC Santa Barbara, US

**Abstract**
Nanophotonic interconnects have been playing an increasingly important role in the datacom regime. Greater integration of silicon photonics demands modeling and simulation support for design validation, optimization and design space exploration. In this work, we develop compact models for a number of key photonic devices, which are extensively validated by the measurement data of a fabricated optical network-on-chip (ONoC). Implemented in SPICE-compatible Verilog-A, the models are used in circuit-level simulations of full optical links. The simulation results match well with the measurement data. Our model library and simulation approach enable the electro-optical (EO) co-simulation, allowing designers to include photonic devices in the whole system design space, and to co-optimize the transmitter, interconnect, and receiver jointly.

**Download Paper (PDF; Only available from the DATE venue WiFi)**
4.3 Efficient memory design

Date: Tuesday 28 March 2017
Time: 17:00 - 18:30
Location / Room: 2BC
Chair: Francisco Cazorla, CSIC and BSC, ES
Co-Chair: Cristina Silvano, Politecnico di Milano, IT

This session presents four papers on novel memory designs and efficient mapping in flash storage. The first two papers improve energy efficiency, with approximate caches on emerging technologies and with a novel DRAM tag-cache architecture. The third paper presents an energy-efficient memory hierarchy through software managed memories. At last, the session concludes with an adaptive page re-mapping architecture in flash-based storage with improved response time.
**Abstract**

STT-MRAM has attracted great interest for use as on-chip memory due to its high density, near-zero leakage and high endurance. However, its overall energy efficiency is limited by the energy requirements of spin-transfer torque switching during writes and reliable single-ended sensing during reads. Leveraging the ability of many applications to produce acceptable outputs under approximations to computations and data, we propose the use of approximate storage to improve the energy efficiency of STT-MRAM based caches. Towards this end, we explore a combination of different approximation techniques at the circuit and architecture levels that yield significant energy benefits for small probabilities of errors in reads, writes, and retention. A key challenge arises when introducing approximate storage into a cache - data that can tolerate different levels of approximation (or not at all) may be dynamically loaded into a cache line at different times. In addition, it is necessary to manage the approximations so as to obtain a desirable energy-quality tradeoff at the application level. We propose STAxCache (Spintronic Approximate Cache), an STT-MRAM based approximate L2 cache architecture that retains the full flexibility of a conventional cache, while allowing for different levels of approximation to different parts of a program’s memory address space. We introduce a simple interface that allows the programmer to specify the quality requirements for different data structures, and instructions in the ISA to expose this information to STAxCache. We utilize a device-to-architecture simulation framework to evaluate STAxCache and achieve 1.4x improvement in L2 cache energy for negligible (<0.5%) loss in application-level quality across a suite of 8 benchmarks.

Download Paper (PDF; Only available from the DATE venue WiFi)
From functional validation to functional qualification

4.4 From functional validation to functional qualification

Date: Tuesday 28 March 2017
Time: 17:00 - 18:30
Location / Room: 3A
Chair: Graziano Pravadelli, University of Verona, IT
Co-Chair: Elena Ioana Vatajelu, TIMA, FR

The section presents techniques and tools to generate testcases for functional validation and to define coverage metrics for functional qualification.
The growing race towards exascale computing is pushing the adoption of ever more heterogeneous systems into mainstream. The resources available on a chip, the level of

José L. Ayala, Universidad Complutense de Madrid, ES

Co-Chair:

Donatella Sciuto, Politecnico di Milano, IT

Chair:

Donatella Sciuto, Politecnico di Milano, IT

Organiser:

Time:

Date:

4.5 Hot Topic Session: On How to Design and Manage Exascale Computing System Technologies

Time

Label

Presentation Title

Authors

17:30

4.4.2

MINIME-VALIDATOR: VALIDATING HARDWARE WITH SYNTHETIC PARALLEL TESTCASES

Speaker:

Alper Sen, Bogazici University, TR

Authors:

Alper Sen1, Etem Deniz2 and Brian Kahne3

1Bogazici University, TR; 2TUBITAK, TR; 3NXP, US

Abstract

Programming of multicore architectures with large number of cores is a huge burden on the programmer. Parallel patterns ease this burden by presenting the developer with a set of predefined programming patterns that implement best practices in parallel programming. Since the behavior of patterns is well-known and understood they can also lower the burden for verification. In this work, we present a toolset, MINIME-Validator, for generating synthetic parallel testcases from a newly defined Parallel Pattern Markup Language (PPML) that uses the concept of parallel patterns. Our testcases mimic the behavior of real customer applications while being much smaller and can be used to generate traffic and validate e.g. inter-processor communication architectures.

Experiments show that synthetic testcases can be used for finding representative hardware communication problems. To the best of our knowledge, this is the first time synthetic testcases using parallel programming patterns are used for hardware validation.

Download Paper (PDF; Only available from the DATE venue WiFi)

18:00

4.4.3

COST-EFFECTIVE ANALYSIS OF POST-SILICON FUNCTIONAL COVERAGE EVENTS

Speaker:

Avi Ziv, IBM Research - Haifa, IL

Authors:

Farimah Farahmand1, Ronny Morad2, Avi Ziv2, Ziv Nevo2 and Prabhat Mishra1

1University of Florida, US; 2IBM Research - Haifa, IL

Abstract

Post-silicon validation is a major challenge due to the combined effects of debug complexity and observability constraints. Assertions as well as a wide variety of checkers are used in pre-silicon stage to monitor certain functional scenarios. Pre-silicon checkers can be synthesized to coverage monitors in order to capture the coverage of certain events and improve the observability during post-silicon debug. Synthesizing thousands of coverage monitors can introduce unacceptable area and energy overhead. On the other hand, absence of coverage monitors would negatively impact post-silicon coverage analysis. In this paper, we propose a framework for cost-effective post-silicon coverage analysis by identifying hard-to-detect events coupled with trace-based coverage analysis. This paper makes three major contributions. We propose a method to utilize existing debug infrastructure to enable coverage analysis in the absence of synthesized coverage monitors. This analysis enables us to identify a small percentage of coverage monitors that need to be synthesized in order to provide a trade-off between observability versus design overhead. To improve the observability further, we also present an observability-aware trace selection algorithm that gives priority to signals associated with important coverage monitors with negligible effect on debug observability. Our experimental results demonstrate that an effective combination of coverage monitor selection and trace analysis can drastically reduce (up to 10 times) the required coverage monitors without sacrificing observability.

Download Paper (PDF; Only available from the DATE venue WiFi)

18:30

IP2-9, 922

HEAD-MOUNTED SENSORS AND WEARABLE COMPUTING FOR AUTOMATIC TUNNEL VISION ASSESSMENT

Speaker:

Josue Ortiz, Complutense University of Madrid, ES

Authors:

Yuchao Ma and Hassan Ghasemzadeh, Washington State University, US

Abstract

As the second leading cause of blindness worldwide, glaucoma impacts a large population of individuals over 40. Although visual acuity often remains unaffected in early stages of the disease, visual field loss, expressed by tunnel vision condition, gradually increases. Glaucoma often remains undetected until it has moved into advanced stages. In this paper, we introduce a wearable system for automatic tunnel vision detection using head-mounted sensors and machine learning techniques. We develop several tasks, including reading and observation, and estimate visual field loss by analyzing user’s head movements while performing the tasks. An integrated computational module takes sensor signals as input, passes the data through several automatic data processing phases, and returns a final result by merging task-level predictions. For validation purposes, a series of experiments is conducted with 10 participants using tunnel vision simulators. Our results demonstrate that the proposed system can detect mild and moderate tunnel vision with an accuracy of 93.3% using a leave-one-subject-out analysis.

Download Paper (PDF; Only available from the DATE venue WiFi)

18:31

IP2-9, 273

AUTOMATIC EQUIVALENCE CHECKING FOR SYSTEMC-TLM 2.0 MODELS AGAINST THEIR FORMAL SPECIFICATIONS

Speaker:

Mehran Goli, University of Bremen, DE

Authors:

Mehran Goli, Janiss Stoppe and Rolf Drechsler, University of Bremen, DE

Abstract

The necessity to handle the increasing complexity of digital circuits has led to the usage of more and more abstract design paradigms. In particular, the Electronic System Level (ESL) has become an area of active research and industrial application, especially via SystemC and its Transaction Level Modeling (TLM) framework. Additionally, the usage of formal specification languages such as the Unified Modeling Language (UML) prior to the implementation (even at higher abstraction levels) is now a broadly accepted workflow. Utilizing this layer approach leaves the translation from the specification to the implementation to the designer, leaving the question unanswered how the equivalence of these should be verified. This paper proposes a novel, non-intrusive and broadly applicable approach to automatically validate the equivalence of the structural and behavioral information of a SystemC-TLM 2.0 model and its formal specification.

Download Paper (PDF; Only available from the DATE venue WiFi)

4.5 Hot Topic Session: On How to Design and Manage Exascale Computing System Technologies

Date: Tuesday 28 March 2017

Time: 17:00 - 18:30

Location / Room: 3C

Organiser:

Donatella Sciuto, Politecnico di Milano, IT

Chair:

Donatella Sciuto, Politecnico di Milano, IT

Co-Chair:

José L. Ayala, Universidad Complutense de Madrid, ES

The growing race towards exascale computing is pushing the adoption of ever more heterogeneous systems into mainstream. The resources available on a chip, the level of
integration and the speed of components have increased dramatically over the years. Moreover, To handle the stringent performance requirements of future exascale-class applications, High Performance Computing (HPC) systems need ultra-efficient heterogeneous compute nodes. However, we keep on adopting superseded approaches to the exploitation of these resources. In this session, the speakers will focus on this requirements providing insight on how to enable the definition and the efficient deployment of such a technology.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>17:00</td>
<td>4.5.1</td>
<td>TOWARDS EXASCALE COMPUTING WITH HETEROGENEOUS ARCHITECTURES</td>
<td>Kenneth O’Brien, Xilinx Inc., IE</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker:</td>
<td>Authors:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Kenneth O’Brien, Lorenzo Di Tucci, Gianluca Durelli, and Michaela Blott</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1Xilinx, IE; 2Politecnico di Milano, IT</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abstract</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>The goal of reaching exascale computing is made especially challenging by the highly heterogeneous nature of modern platforms and the energy they consume. As compute nodes typically utilize multiple multi-core CPU and are increasingly equipped with PCIe based accelerators, both are contributing to an ever more dynamic power consumption. In our study we evaluate our target application on a variety of heterogeneous platforms, including high end FPGA, GPU, and Xeon Phi accelerators, with respect to energy efficiency at a node and cluster level. We compare multiple implementations of our application, each built with a different modern parallel programming framework, with respect to execution performance, code complexity and energy efficiency. Later we extrapolate based on our findings, the implications of scaling this application towards exascale, with projections of computation achievable within the exascale power budget for our three architectures. Download Paper (PDF; Only available from the DATE venue WiFi)</td>
<td></td>
</tr>
<tr>
<td>17:18</td>
<td>4.5.2</td>
<td>FROM EXAFLOP TO EXAFLOW</td>
<td>Tobias Becker, Maxeler Technologies, GB</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker:</td>
<td>Authors:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tobias Becker, Pavel Burovsky, Anna Maria Nestorov, Kristina Palkikavka, Enrico Reggiani, and Georgi Gaydadjiev</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1Maxeler Technologies, GB; 2Maxeler Technologies Ltd, GB; 3Politecnico di Milano, IT; 4Maxeler / Imperial College, GB</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abstract</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Exascale computing is facing a gap between the ever increasing demand for application performance and the underlying chip technology that does no longer deliver the expected exponential increase in CPU performance. The industry is now progressively moving towards dedicated accelerators to deliver high performance and better energy efficiency. However, the question of programmability still remains. To address this challenge we propose a dedicated high-level accelerator programming and execution model where performance and efficiency are primary targets. Our model splits the computation into a conventional CPU-oriented part and a highly efficient fully programmable data flow part. We present a number of systematic transformations and optimisations targeting Maxeler dataflow systems that typically yield one to two orders of magnitude improvements in terms of both performance and energy efficiency. These significant gains are enabled by addressing fundamental algorithmic properties and on-demand numerical requirements. This approach is demonstrated by a case study from computational finance. Download Paper (PDF; Only available from the DATE venue WiFi)</td>
<td></td>
</tr>
<tr>
<td>17:36</td>
<td>4.5.3</td>
<td>HETEROGENEOUS EXASCALE SUPERCOMPUTING: THE ROLE OF CAD IN THE EXAFPGA PROJECT</td>
<td>Marco Santambrogio, Politecnico di Milano, IT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker:</td>
<td>Authors:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Marco Rabozzi, Giuseppe Natale, Emanuele Del Sozzo, Alberto Scolari, Marco D. Santambrogio and Luca Stornaiuolo, Politecnico di Milano, IT</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abstract</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Since the end of Moore’s law is limiting the growth of general purpose processors, High Performance Processing (HPC) systems are considering FPGA-based accelerators as a promising solution for several application fields. However, their employment poses challenges the research is still tackling, and existing tools and workflows do not naturally adapt to the scale and complexity of HPC domains. To help researchers and practitioners, this paper proposes CAOS, a platform that implements an FPGA development workflow tailored to HPC systems while being open to external contributions. Indeed, researchers and developers can plug into CAOS to experiment and compare their solutions at each step of the design flow. This paper describes the CAOS workflow and validates it against several case studies to assess its generality and highlight possible research contributions. Download Paper (PDF; Only available from the DATE venue WiFi)</td>
<td></td>
</tr>
<tr>
<td>17:54</td>
<td>4.5.4</td>
<td>AN OPEN RECONFIGURABLE RESEARCH PLATFORM AS STEPPING STONE TO EXASCALE HIGH-PERFORMANCE COMPUTING</td>
<td>Dirk Stroobandt, Ghent University, BE</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker:</td>
<td>Authors:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Dirk Stroobandt, Catalin Bogdan Ciobanu, Marco D. Santambrogio, Jose Gabriel Coutinho, Andreas Brokalojas, Dionisos Pnevmatikatos, Michael Huebner, Tobias Becker, and Alex J. W. Thom</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1Ghent University, BE; 2RU/VA, NL; 3Politecnico di Milano, IT; 4Imperial College London, GB; 5Hyivexis, GR; 6ECE Department, Technical University of Crete &amp; FORTH-ICS, GR; 7Ruhr-University Bochum, DE; 8Maxeler Technologies, GB; 9University of Cambridge, GB</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abstract</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>To handle the stringent performance and power requirements of future exascale-class applications, High Performance Computing (HPC) systems need ultra-efficient heterogeneous compute nodes and hardware accelerators with a high degree of specialization. Ideally, dynamic reconfiguration will be an intrinsic feature, so that specific HPC application features can be optimally accelerated, even if they regularly change over time. We create a new and flexible exploration platform for developing reconfigurable architectures, design tools and HPC applications with run-time reconfiguration built-in as a core fundamental feature instead of an add-on. Our project proposes an open research platform that covers the entire stack from architecture up to the application, focusing on the fundamental building blocks for run-time reconfigurable exascale HPC systems: new chip architectures with very low reconfiguration overhead, new tools that truly take reconfiguration as a central design concept, and applications that are tuned to maximally benefit from the proposed run-time reconfiguration techniques. Ultimately, this open platform will enable groundbreaking research towards new exascale computing platforms. Download Paper (PDF; Only available from the DATE venue WiFi)</td>
<td></td>
</tr>
<tr>
<td>18:12</td>
<td>4.5.5</td>
<td>GEOPM: A VEHICLE FOR EXASCALE COMMUNITY COLLABORATION TOWARD CO-DESIGNED ENERGY MANAGEMENT SOLUTIONS</td>
<td>Matthias Maiterth, Intel, US</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker:</td>
<td>Author:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abstract</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>The power scaling challenge associated with Exascale systems is a well-known issue. In this invited talk, we provide an overview of the Global Extensible Open Power Manager (GEOPM). GEOPM is an open source power management runtime framework which is being contributed to the HPC community to foster collaboration on new power management runtime techniques to address Exascale power challenges or enhance performance and power efficiency on today’s systems as well. Through GEOPM’s plug-in extensible architecture, it enables rapid prototyping of new runtime algorithms. This talk will cover GEOPM’s architecture, interfaces, and project status. For additional information, please visit: <a href="https://geopm.github.io/geopm/">https://geopm.github.io/geopm/</a></td>
<td></td>
</tr>
<tr>
<td>18:30</td>
<td></td>
<td>End of session</td>
<td></td>
</tr>
</tbody>
</table>

Exhibition Reception in Exhibition Area

The Exhibition Reception will take place on Tuesday in the exhibition area, where free drinks for all conference delegates and exhibition visitors will be offered. All exhibitors are welcome to also provide drinks and snacks for the attendees.
4.6 Fault modeling, test generation and diagnosis

Date: Tuesday 28 March 2017
Time: 17:00 - 18:30
Location / Room: SA

Chair: Stephan Eggersgluss, University of Bremen, DE
Co-Chair: Martin Keim, Mentor, DE

This session includes a presentation about new SAT-based ATPG techniques for robust initialization of transistor stuck-open faults. Further, a diagnosis method for arbiter physical unclonable functions to identify systematic manufacturing issues is presented. The last paper analyzes failure modes of Flash memories and proposes suitable fault models.

17:00  4.6.1  (Best Paper Award Candidate)
FAST AND WAVEFORM-ACCURATE HAZARD-AWARE SAT-BASED TSOF ATPG
Speaker: Jan Burchard, University of Freiburg, DE
Authors: Jan Burchard 1, Dominik Erb 1, Adit D. Singh 1, Sudhakar M. Reddy 2 and Bernd Becker 1
1University of Freiburg, DE; 2Auburn University, US; 3University of Iowa, US

Abstract
Opens are known to be one of the predominant defects in nanoscale technologies. Especially with an increasing number of complex cells in today’s VLSI designs intra-gate opens are becoming a major problem. The generation of tests for these faults is hard, as the timing of the circuit needs to be considered accurately to prevent the invalidation of the generated tests through hazards. Current test generation methods, including new cell aware tests that explicitly target open defects, ignore the possibility of hazard caused test invalidation. Such tests can fail to detect a significant fraction of the targeted opens. In this work we present a waveform-accurate hazard-aware test generation approach to target intra-gate opens. Our methodology is based on a SAT-based encoding and allows the generation of tests guaranteed to be robust against hazards. Experimental results for large benchmarks mapped to the state-of-the-art NanGate 45nm cell library including complex cells show the test generation efficiency of the proposed method. Large circuits were efficiently handled – even without the use of fault simulation. Our experiments show that on average, about 90.2% of conventional hazard-unaware tests will fail to detect the targeted opens because of test invalidation – these are reliably detected by our new test generation methodology. Importantly, our approach can also be applied to improve the effectiveness of commercial cell aware tests.

Download Paper (PDF; Only available from the DATE venue WiFi)

17:30  4.6.2  FAULT DIAGNOSIS OF ARBITER PHYSICAL UNCLONABLE FUNCTION
Speaker: Yu Hu, Institute of Computing Technology, Chinese Academy of Sciences, CN
Authors: Jing Ye 1, Qingli Guo 1, Yu Hui 1 and Xiaowei Li 1
1State Key Laboratory of Computer Architecture, Institute of Computing Technology, Chinese Academy of Sciences, CN; 2State Key Laboratory of Computer Architecture, Institute of Computing Technology, Chinese Academy of Sciences; University of Chinese Academy of Sciences, CN

Abstract
Physical Unclonable Function (PUF) has broad application prospects in the field of hardware security. If faults happen in PUF during manufacturing, the security of whole chip will be threatened. Fault diagnosis plays an important role in the yield learning process. However, since different manufactured PUFs with the same design have different Challenge-Response Pairs (CRPs), which cannot be predicted, the traditional fault diagnosis method based on comparing the fault-free responses of a design and the failing responses of chips is no longer suitable for diagnosing PUF. Therefore, this paper proposes a fault diagnosis method toward classic arbiter PUF. The stuck-at faults and the delay faults are considered. Based on the expected uniformity of arbiter PUF, a diagnostic challenge generation method and a corresponding CRP analysis method are proposed to distinguish faults within the arbiter PUF. Experimental results show that the diagnostic accuracy achieves 100.0% with good diagnostic resolution.

Download Paper (PDF; Only available from the DATE venue WiFi)

18:00  4.6.3  FPGA-BASED FAILURE MODE TESTING AND ANALYSIS FOR MLC NAND FLASH MEMORY
Speaker: Fei Wu, Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology, CN
Authors: Meng Zhang 1, Fei Wu 1, Qian Xia 1, He Huang 1, Jian Zhou 1 and Changsheng Xie 1
1Huazhong University of Science and Technology, CN; 2University of Central Florida, US

Abstract
With the improvement of flash memory storage density, data reliability and flash lifetime are decreased. Error correction codes (ECC) and error management schemes can boost both reliability and lifetime. However, in order to develop effective fault tolerance algorithms and management solutions, it is very necessary to have a more profound understanding of failure modes of flash memory. To enable such understanding, we design an experimental platform and scheme to clearly investigate flash failure modes. This paper examines various failure modes occurring at 2x-nm MLC NAND flash technologies, such as page allocation scheme-based program interference (PASBPI) errors (i.e., different page allocation schemes mean data can be programmed into flash pages in different ways, which can lead to different program interference errors), write errors of the least significant bit (LSB) and the most significant bit (MSB) and different data pattern-based read interference errors (i.e., different data values programmed into flash pages can cause differential read interference errors). We analyze these observed failure modes and explain why they exist. We hope it is helpful to understand these discovered failure modes to propose effective fault tolerance and error management algorithms.

Download Paper (PDF; Only available from the DATE venue WiFi)

18:30  4.6.4  (Best Paper Award Candidate)
RETRDMR: TROUBLESHOOTING NON-DETERMINISTIC FAULTS WITH RETROSPECTIVE DMR
Speaker: Ting Wang, The Chinese University of Hong Kong, HK
Authors: Ting Wang 1, Yannan Liu 1, Qiang Xu 1, Zhasbo Zhang 2, Zhiyuan Wang 2 and Xinli Gu 2
1The Chinese University of Hong Kong, HK; 2Huawei Technologies, Inc., US

Abstract
The most notorious faults for diagnosis in post-silicon validation are those that manifest themselves in a non-deterministic manner with system-level functional tests, where errors randomly appear from time to time even when applying the same workloads. In this work, we propose a novel diagnostic framework that resorts to dual-modular redundancy (DMR) for troubleshooting non-deterministic faults, namely RetroDMR. To be specific, we log the essential events (e.g., the sequence of thread migration) in the faulty run to record the mapping relationship between threads and their corresponding execution units. Then in the following diagnosis runs, we apply redundant multithreading (RMT) technique to reduce error detection latency, while at the same time we try to follow the thread migration sequence of the original run whenever possible. By doing so, RetroDMR significantly improves the reproduction rate and diagnosis resolution for non-deterministic faults, as demonstrated in our experimental results.

Download Paper (PDF; Only available from the DATE venue WiFi)
The session covers variable-aware solutions at the system and circuit level. Firstly, neuromorphic circuits are addressed and its relation with process variation. After that, variability

**4.7 Process variation management for today's and tomorrow's computing**

**Date:** Tuesday 28 March 2017  
**Time:** 17:00 - 18:30  
**Location / Room:** 3B  
**Chair:** Muhammad Shaﬁque, TU Wien, AT  

The session covers variable-aware solutions at the system and circuit level. Firstly, neuromorphic circuits are addressed and its relation with process variation. After that, variability is again addressed but, this time, for entire computing systems.

**Time** | **Label** | **Presentation Title** | **Authors**  
---|---|---|---  
18:31 | IP2-11 | CRITICAL PATH - ORIENTED THERMAL AWARE X-FILLING FOR HIGH UN-MODELED DEFECT COVERAGE | Fotios Varziotis, Computer Engineering, T.E.I. of Epirus, Greece, GR  
**Authors:** Fotios Varziotis*, Computer Engineering, T.E.I. of Epirus, Greece, GR; 2Department of Computer Science and Engineering, University of Ioannina, GR  
**Abstract**  
The thermal activity during testing can be considerably reduced by applying power-oriented filling of the unspecified bits of test vectors. However, traditional power-oriented X-fill methods do not correlate the thermal activity with delay failures, and they consume all the unspecified bits to reduce the power dissipation at every region of the core. Therefore, they adversely affect the un-modeled defect coverage of the generated test vectors. The proposed method identifies the unspecified bits that are more critical for delay failures, and it fills them in such a way as to create a thermal safe neighborhood around the most critical regions of the core. For the rest of the unspecified bits a probabilistic model based on output deviations is adopted to increase the un-modeled defect coverage of the test vectors. Experimental results show that the thermal activity and the inter-connection delays of critical regions of the core are comparable to those of the power-oriented X-fill methods, while the un-modeled defect coverage is as high as that of the random-fill method.

**Download Paper (PDF; Only available from the DATE venue WiFi)**

18:32 | IP2-12 | A COMPREHENSIVE METHODOLOGY FOR STRESS PROCEDURES EVALUATION AND COMPARISON FOR BURN-IN OF AUTOMOTIVE SOC | Paolo Bernardi, Politecnico di Torino, IT  
**Authors:** Paolo Bernardi*, Davide Appello*, Giampaolo Giacopelli1, Alessandro Motta1, Alberto Pagan1, Giorgio Pollicia1, Christian Rabbi1, Marco Restivo1, Prit Ruben, Ernesto Sanchez1, Claudio Maria Villa2 and Federico Venni1  
1Politecnico di Torino, IT; 2STMicroelectronics, IT; 3STMicroelectronics, IT; 4Falling University of Technology, EE  
**Abstract**  
Environmental and electrical stress phases are commonly applied to automotive devices during manufacturing test. The combination of thermal and electrical stress is used to give rise to early life latent failures that can be naturally found in a population of devices by accelerating aging processes through Burn-In test phases. This paper provides a methodology to evaluate and compare the stress procedures to be run during Burn-In; the proposed method takes into account several factors such as circuit activity, chip surface temperature and current consumption required by the stress procedure, and also considers Burn-In flow and tester limitations. A specific metric called Stress Coverage is suggested summing up all the stress contributions. Experimental results are gathered on an automotive device, showing the comparison between scan-based and functional stress run by a massively parallelized test equipment; reported figures and tables quantify the differences between the two approaches in terms of stress.

**Download Paper (PDF; Only available from the DATE venue WiFi)**

18:30 | End of session  

Exhibition Reception in Exhibition Area  
The Exhibition Reception will take place on Tuesday in the exhibition area, where free drinks for all conference delegates and exhibition visitors will be offered. All exhibitors are welcome to also provide drinks and snacks for the attendees.

**4.7.1 ROBUST NEUROMORPHIC COMPUTING IN THE PRESENCE OF PROCESS VARIATION**  
**Speaker:** Mehdi Kamal, University of Tehran, IR  
**Authors:** Mehdi Kamal, University of Tehran, IR; 2Mohammad Ali Maleki, Mehdi Kamal, Ali Afzali-Kusha and Massoud Pedram  
1University of Tehran, IR; 2University of Southern California, US  
**Abstract**  
In this paper, an approach for increasing the sustainability of inverter-based memristive neuromorphic circuits in the presence of process variation is presented. The approach works based on extracting the impact of process variations on the neurons characteristics during the test phase through a proposed algorithm. In this method, first, some combinations of inputs and weights (based on the neuromorphic circuit structure) are injected into the circuit and the features of the neurons are determined. Next these features which are back-annotated, are utilized in an efficient ex-situ training approach to determine the proper weights of the neurons. The approach provides a considerable improvement in the output accuracy. To evaluate the effectiveness of the proposed approach, some approximate applications are studied using 90nm technology. The results of the study reveal that using this framework provide, on average, 17X higher output accuracy compared to the cases that the impact of the process variation is not considered at all.

**Download Paper (PDF; Only available from the DATE venue WiFi)**

**4.7.2 AN ON-LINE FRAMEWORK FOR IMPROVING RELIABILITY OF REAL-TIME SYSTEMS ON "BIG-LITTLE" TYPE MPSoCs**  
**Speaker:** Yue Ma, University of Notre Dame, US  
**Authors:** Yue Ma1, Thidapat Chantem2, Robert Dick3, Shige Wang4 and X, Sharon Hu1  
1University of Notre Dame, US; 2Virginia Polytechnic Institute and State University, US; 3University of Michigan and Stryd, US; 4General Motors R&D, US  
**Abstract**  
Heterogeneous MPSoCs consisting of cores with different performance/power behaviors are widely used in many power-constrained real-time systems. Both soft-error reliability and lifetime reliability are key concerns in such systems. Although existing work have investigated related problems, they either focus on one of the two reliability concerns or propose complicated scheduling algorithms that cannot adequately address run-time workload and environment variations. This paper introduces an on-line heuristic to maximize soft-error reliability while satisfying a lifetime reliability constraint for soft real-time systems executed on MPSoCs composed of high-performance cores and low-power cores. Based on the run-time cores’ frequencies and utilizations, the heuristic performs workload migration between the high-performance cores and low-power cores to achieve improved soft-error reliability. Experimental results from both a hardware platform and a simulator show that the proposed algorithm reduces the probability of faults by at least 30% compared to a number of representative existing approaches while satisfying the same lifetime reliability constraints.

**Download Paper (PDF; Only available from the DATE venue WiFi)**
APPLICATION PERFORMANCE IMPROVEMENT BY EXPLOITING PROCESS VARIABILITY ON FPGA DEVICES

Speaker:
Konstantinos Maragos, National Technical University of Athens, GR

Authors:
Konstantinos Maragos¹, George Lentaris¹, Kostas Siozios¹, Dimitrios Soudris¹ and Vasilis Pavlidis²
¹National Technical University of Athens, GR; ²The University of Manchester, GR

Abstract
Process variability is known to be increasing with technology scaling in IC fabrication, thereby degrading the overall performance of the manufactured devices. The current paper focuses on the variability effect in FPGAs and the possibility to boost the performance of each device at run-time, after fabrication, based on the individual characteristics of this device. First, we develop a sensing infrastructure involving a wide network of customized ring oscillators to measure intra-chip and inter-chip variability in 28nm FPGAs, i.e., in eight Xilinx Zynq XC7Z020T-1CSG324 devices. Second, we develop a closed-loop framework based on dynamic reconfiguration of clock tiles, I/O data sniffing, HW/SW communication, and verification with test vectors, to dynamically increase the operating frequency in Zynq while preserving its correctness. Our results show intra-chip variability in the area of 5.2% to 7.7% and inter-chip variability up to 17%. Our framework improves the performance of example FIR designs by up to 90.3% compared to the SW tool reports and shows speed difference among devices by up to 12.4%.

Download Paper (PDF; Only available from the DATE venue WiFi)

18:30 End of session

Exhibition Reception in Exhibition Area

The Exhibition Reception will take place on Tuesday in the exhibition area, where free drinks for all conference delegates and exhibition visitors will be offered. All exhibitors are welcome to also provide drinks and snacks for the attendees.

4.8 CV Fair DATE 2017

Date: Tuesday 28 March 2017
Time: 17:00 - 18:30
Location / Room: Exhibition Theatre

Organiser:
Marisa Lopez-Vallejo, UPM, ES

Moderator:
Marisa Lopez-Vallejo, UPM, ES

The Curriculum Vitae (also known as a vita or CV) is the first point of contact between employee and employer. It must provide a concise overview of academic background and achievements. Furthermore, it usually should catch the attention of the readers, get them to take a closer look at you and ultimately invite you for an interview. Philippe Ory, Head of the EPFL Career Center, will open this CV Fair with a talk on the key issues that must be addressed when writing a CV.

Afterwards, organizations participating in the CV Fair will give a brief presentation with basic information about the company, potential positions or internships, what types of students are being sought, etc. The CV fair is designed to allow for students to engage in individual conversations with the company or organization team and ask specific questions that may have arisen during the presentation.

17:00  4.8.1 OPENING
Speaker: Philippe Ory, Head of the EPFL Career Center, CH

17:30  4.8.2 CADENCE PRESENTATION
Speaker: Anton Klotz, Cadence Design Systems, DE

17:40  4.8.3 HIPEAC PRESENTATION
Speaker: Xavier Salazar, Hipeac, ES

17:50  4.8.4 SMARTCARDIA PRESENTATION
Speaker: Srinivasan Murali, SmartCardia, CH

18:00  4.8.5 NESPRESSO PRESENTATION
Speaker: Martino Ruggiero, Nespresso, CH

18:10  4.8.6 NESTLé PRESENTATION
Speaker: Gian Paolo Perrucci, Mobility and Apps Solution Manager at Nestlé, CH

18:20  4.8.7 GAIT UP PRESENTATION
Speaker: Karim Kanoun, Mobile and Embedded Development Manager at Gait Up S.A., CH

18:30 End of session

Exhibition Reception in Exhibition Area
The Exhibition Reception will take place on Tuesday in the exhibition area, where free drinks for all conference delegates and exhibition visitors will be offered. All exhibitors are welcome to also provide drinks and snacks for the attendees.
NOXIM-XT: A BIT-ACCURATE POWER ESTIMATION SIMULATOR FOR NOCS

Presenter:
Pierre Bornel, Université de Bretagne Sud, FR

Authors:
André Rossi¹, Johann Laurenţ² and Erwan Moreac²
¹:LERIA, Université d'Angers, Angers, France, FR; ²:Lab-STICC, Université de Bretagne Sud, Lorient, FR

Abstract
We have developed an enhanced version of Noxim (Noxim-XT) to estimate the energy consumption of a NoC in a SOC. Noxim-XT is used in a two-step methodology. First, applications are mapped on a SoC and their traffics are extracted by simulation with MPSOcBench. Second, Noxim-XT tests various hardware configurations of the NoC, and for each configuration, the application’s traffic is re-injected and replayed, an accurate performance and power breakdown is provided, and the user can choose different data coding strategies. With the help of Noxim XT, each configuration is bit-accurately estimated in terms of energy consumption. After simulation, a spatial mapping of the energy consumption is provided and highlights the hot spots. Moreover, the new coding strategies allow significant energy saving. Noxim XT simulations and a FPGA-based prototype of a new coding strategy will be demonstrated at the U-booth to illustrate these works.

More information ...

RIMEDIO: WHEELCHAIR MOUNTED ROBOTIC ARM DEMONSTRATOR FOR PEOPLE WITH MOTOR SKILLS IMPAIRMENTS

Presenter:
Alessandro Palla, University of Pisa, IT

Authors:
Gabriele Meoni and Luca Fanucci, University of Pisa, IT

Abstract
People with reduced mobility experiment many issues in the interaction with the indoor and outdoor environment because of their disability. For those users even the simplest action might be a hard/impossible task to perform without the assistance of an external aid. We propose a simple and lightweight wheelchair mounted robotic arm with the focus on the human-machine interface that has to be simple and accessible for users with different kind of disabilities. The robotic arm is equipped with a 5 MP camera, force and proximity sensors and a 6 axis Inertial Measurement Unit on the end-effector that can be controlled using an app running on a tablet. When the user selects the object to reach (for instance a button) on the tablet screen, the arm autonomously carries out the task, using the camera image and the sensors measurements for autonomous navigation. The demonstrator consists in the robotic arm prototype, the Android tablet and a personal computer for arm setup and configuration.

More information ...

OPENCTMOD: AN OPEN SOURCE COLLABORATIVE MATLAB TOOLBOX FOR THE DESIGN AND SIMULATION OF CONTINUOUS-TIME SIGMA DELTA MODULATORS

Presenter:
Dang-Kien Germain Pham, LTCI, Télécom ParisTech, Université Paris-Saclay, 75013, Paris, France, FR

Author:
Chadi Jabbour, LTCI, Télécom ParisTech, Université Paris-Saclay, FR

Abstract
Simulating Continuous Time (CT) Sigma Delta Modulators (SDM) is commonly done using block level systems such as Simulink which is a highly time consuming task even at system level. Therefore, the existing design tools for SDM are either discrete time oriented (Schréier toolbox) or proprietary (Ulm toolbox). In this work, we propose a new Matlab/C toolbox for the design of CT SDM. Simulation is based on state space representation thereby allowing to support most of the existing SDM architectures. Moreover, the main non-idealities of the main blocks are modeled (opamp DC gain, finite GBW, DACs mismatch, ISI and quantizer offset). Besides, thanks to the modular and open source approach for this toolbox, every user can easily implement additional features and include it. During the forum, designs and simulations for various architectures of CT SDM will be performed to demonstrate the accuracy and efficiency of the proposed toolbox. The collaborative aspect will be also shown.

More information ...

MATISSE: A TARGET-AWARE COMPILER TO TRANSLATE MATLAB INTO C AND OPENCL

Presenter:
Luís Reis, University of Porto, PT

Authors:
João Bispo and João Cardoso, University of Porto / INESC-TEC, PT

Abstract
Many engineering, scientific and finance algorithms are prototyped and validated in array languages, such as MATLAB, before being converted to other languages such as C for use in production. As such, there has been substantial effort to develop compilers to perform this translation automatically. Alternative types of computation devices, such as GPGPUs and FPGAs, are becoming increasingly more popular, so it becomes critical to develop compilers that target these architectures. We have adapted MATISSE, our MATLAB-compatible compiler framework, to generate C and OpenCL code for these platforms. In this demonstration, we will show how our compiler works and what its capabilities are. We will also describe the main challenges of efficient code generation from MATLAB and how to overcome them.

More information ...

A VOLTAGE-SCALABLE FULLY DIGITAL ON-CHIP MEMORY FOR ULTRA-LOW-POWER IOT PROCESSORS

Presenter:
Jun Shiomi, Kyoto University, JP

Authors:
Tohru Ishihara and Hidetoshi Onodera, Kyoto University, JP

Abstract
A voltage-scalable RISC processor integrating standard-cell based memory (SCM) is demonstrated. Unlike conventional processors, the processor has Standard-Cell based Memories (SCMs) as an alternative to conventional SRAM macros, enabling it to operate at a 0.4 V single-supply voltage. The processor is implemented with the fully automated cell-based design, which leads to low design costs. By scaling the supply voltage and applying the back-gate biasing techniques, the power dissipation of the SCMs is less than 20 µW, enabling the SCMs to operate with ambient energy source only. In this demonstration, the SCMs of the processor operates with 0.4V single-supply voltage, showing the capability of the proposed architecture.

More information ...

GNOSTICS: AN ULTRA-FAST, HIGHLY EXTENSIBLE, CYCLE-ACCURATE GPU-BASED PARALLEL NETWORK-ON-CHIP SIMULATOR

Presenter:
Amir CHARIF, TIMA, FR

Authors:
Nacer-Eddine Zergainoh and Michael Nicolaidis, TIMA, FR

Abstract
With the continuous decrease in feature sizes and the recent emergence of 3D stacking, chips comprising thousands of nodes are becoming increasingly relevant, and state-of-the-art NoC simulators are unable to simulate such a high number of nodes in reasonable times. In this demo, we showcase GNOSTICS, the first detailed, modular and scalable parallel NoC simulator running fully on GPU (Graphics Processing Unit). Based on a unique design specifically tailored for GPU parallelism, GNOSTICS is able to achieve unprecedented speedups with no loss of accuracy. To enable quick and easy validation of novel ideas, the programming model was designed with high extensibility in mind. Currently, GNOSTICS accurately models a VC-based microarchitecture. It supports 2D and 3D mesh topologies with full or partial vertical connections. A variety of routing algorithms and synthetic traffic patterns, as well as dependency-driven trace-based simulation (Netrace), are implemented and will be demonstrated.

More information ...
ACCELERATORS: RECONFIGURABLE SELF-TIMED DATAFLOW ACCELERATOR & FAST NETWORK ANALYSIS IN SILICON

Presenter: Alessandro de Gennaro, Newcastle University, GB
Authors: Danil Sokolov and Andrey Mokhov, Newcastle University, GB

Abstract: Many real-life applications require dynamically reconfigurable pipelines to handle incoming data items differently depending on their values or current operating mode. A demo will show the benefits of an asynchronous accelerator for ordinal pattern encoding with reconfigurable pipeline depth. This was designed, simulated and verified using dataflow structure formalism in Workcraft toolset. The self-timed chip, fabricated in TSMC 90nm, shows high resilience to voltage variation and configurable accuracy of the results. Applications with underlying graph models foster the importance of a fast and flexible approach to graph analysis. To support medicine discovery biological connections are modelled by graphs, and drugs can disconnect some of the connections. A demo will show how graphs can be automatically converted into VHDL designs, which are synthesised into a FPGA for the analysis: thousand times faster than in software. Single stand will be used for both case studies.

SELINK: SECURING HTTP AND HTTPS-BASED COMMUNICATION VIA SECUBE™

Presenter: Airofaraully Giuseppe, CINI & Politecnico di Torino, IT
Authors: Paolo Prinetto¹ and Antonio Varriale²
1 Politecnico di Torino, IT; ²Blu5 Labs Ltd., IT

Abstract: The SEcube™ Open Source platform is a combination of three main cores in a single-chip design. Low-power ARM Cortex-M4 processor, a flexible and fast Field-Programmable-Gate-Array (FPGA), and an EAL5+ certified Security Controller (SmartCard) are embedded in an extremely compact package. This makes it a unique Open Source security environment where each function can be optimized, executed, and verified on its proper hardware device. In this demo, we present a client-server HTTP and HTTPS-based application, for which the traffic is encrypted resorting to the hardware built-in capabilities, and the software libraries, of the SEcube™. By doing so, we show how communication can be secured from an attacker capable of inspecting, and tampering, the regular communication.

GREENOPENHEVC: LOW POWER HEVC DECODER

Presenter: Menard Daniel, INSA Rennes, FR
Authors: Julien Heulot¹, Erwan Nogues¹, Maxime Pelcat² and Wassim Hamidouche¹
1 INSA Rennes, IETR, UBL, FR; ²Institut Pascal, Université Clermont-Ferrand, FR

Abstract: Video on mobile devices is a must-have feature with the prominence of new services and applications using video like streaming or conferencing. The new video standard HEVC is an appealing technology for service providers. Besides, with the recent progress of SoC, software video decoders are now a reality. The challenge is to provide power efficient design to fit with the compelling demand for long battery. We present here a practical set-up demonstrating that the new HEVC standard can be implemented in software on an embedded GPP multicore platform. Different techniques have been integrated to optimize the energy: data-level and thread level parallelisms, video aware Dynamic Voltage and Frequency Scaling. To push back the limits, algorithm level approximate computing is carried-out on the in-loop filtering. The subjective tests have demonstrated that the quality degradation is almost imperceptible. A mean power of less than 1 Watt is reported for a HD 1080p/24fps video decoding.

19:30 End of session
DESIGN FOR IOT
Author:
Lothar Thiele, Swiss Federal Institute of Technology Zurich, CH
Abstract
If visions and forecasts of industry come true then we will be soon surrounded by billions of interconnected embedded devices. We will interact with them in a cyber-human symbiosis, they will not only observe us but also our environment, and they will be part of many visible and ubiquitous objects around us. We have the legitimate expectation that the individual devices as well as the overall system behaves in a reliable and predictable manner. This is an indispensable requirement as it is infeasible to constantly maintain such a large set of devices. In addition, there are many application domains where we rely on a correct and fault-free system behavior. We expect trustworthy results from sensing, computation, communication and actuation due to economic importance or even catastrophic consequences if the overall system is not working correctly, e.g., in industrial automation, distributed control of energy systems, surveillance, medical applications, or early warning scenarios in the context of building safety or environmental catastrophes. Finally, trustworthiness and reliability are mandatory for the societal acceptance of human-cyber interaction and cooperation. It will be argued that we need novel architectural concepts, an associated design process and validations strategies to satisfy the strongly conflicting requirements and associated design challenges of platforms for CPS: Handle at the same time limited available resources, adaptive run-time behavior, and predictability. These challenges concern all components of an IoT system, e.g., computation, storage, wireless communication, energy management, harvesting, sensing and sensor interfaces, and actuation. The talk will be driven by examples from various application domains such as smart watches, zero-power systems, environmental sensing, and air pollution sensing.

THE INTERNET OF THINGS IN THE COGNITIVE ERA
Author:
Alesandro Curioni, IBM Zurich Research, CH
Abstract
Over next few years, the Internet of Things will become the biggest source of data on the planet. That’s where IBM’s Watson cognitive computing system comes in. Watson uses machine learning and other techniques to understand this data and turn it into insight, which can help automate tasks, enable manufacturers to design better products, innovate new services and enhance our overall quality of life. And with cognitive technologies, interactions with 'things' through natural language and voice commands will dramatically improve. This presentation will focus on how innovators in the design automation and embedded systems space can benefit from this trend and get access IBM Watson in the cloud.

5.2 Emerging Computer Paradigms
Date: Wednesday 29 March 2017
Time: 08:30 - 10:00
Location / Room: 4BC
Chair:
Jim Harkin, Ulster University, GB
This session presents recent advances in emerging computing strategies including Reversible Computing and Stochastic Computing with improvements in energy efficiency and reductions in computational complexity. An acceleration platform for the design exploration of Quantum Computers is also presented.

MAKE IT REVERSIBLE: EFFICIENT EMBEDDING OF NON-REVERSIBLE FUNCTIONS
Speaker:
Alwin Zulehner, Johannes Kepler University, Linz, AT
Authors:
Alwin Zulehner¹ and Robert Wille²
¹Johannes Kepler University, AT; ²Johannes Kepler University Linz, AT
Abstract
Reversible computation became established as a promising concept due to its application in various areas like quantum computation, energy-aware circuits, and further areas. Unfortunately, most functions of interest are non-reversible. Therefore, a process called embedding has to be conducted to transform a non-reversible function into a reversible one - a coNP-hard problem. Existing solutions suffer from the resulting exponential complexity and, hence, are limited to rather small functions only. In this work, an approach is presented which tackles the problem in an entirely new fashion. We divide the embedding process into matrix operations, which can be conducted efficiently on a certain kind of decision diagram. Experiments show that improvements of several orders of magnitudes can be achieved using the proposed method. Moreover, for many benchmarks exact results can be obtained for the first time ever.
Download Paper (PDF; Only available from the DATE venue WiFi)
5.3 Hot Topic Session: I’m Gonna Make an Approximation IoT Can’t Refuse - Approximate Computing for Improving Power Efficiency of IoT and HPC

Date: Wednesday 29 March 2017
Time: 08:30 - 10:00
Location / Room: 2BC
Organiser: Vincent Camus, EPFL, CH
Power efficiency is the primary concern of IoT-related applications, both at the sensor node and on its cloud-computing counterpart. Unfortunately, achieving high efficiency and robustness requires complex and conflicting design constraints. Fortunately, the inherent error resiliency of many IoT applications allows the use of Approximate Computing techniques at both hardware and software levels, leading to great benefits on power efficiency while having a minimal impact on the applications.

**Abstract**

Voltage scaling has been used as a prominent technique to improve energy efficiency in digital systems, as reduction in the supply voltage effects in quadratic reduction in energy consumption of the system. The energy efficiency is achieved at the cost of timing errors in the system, that are corrected through additional error detection and correction circuits. In this paper we are proposing voltage over-scaling based approximate operators for applications that can tolerate errors. We characterize the basic arithmetic operators using different operating tribes (combination of supply voltage, back biasing scheme and clock frequency) to generate models for approximate operators. Error-resilient applications can be mapped with the generated approximate operator models to achieve optimum trade-off between energy and error margin. Based on the dynamic speculation technique, best possible operating tribe is chosen at runtime based on the user definable error tolerance margin of the application. In our experiments in 28nm FDSOI, we achieve maximum energy efficiency of 89% for basic operators like 8-bit and 16-bit adders at the cost of 20% Bit Error Rate (ratio of faulty bits over total bits) by operating them in near-threshold regime.

**Download Paper (PDF; Only available from the DATE venue WiFi)**

---

**5.3.2 PUSHER THE LIMITS OF VOLTAGE OVER-SCALING FOR ERROR-RESILIENT APPLICATIONS**

**Speaker:** Olivier Sentieys, INRIA, FR

**Authors:** Rengerajan Ragavan1, Benjamin Barroso1, Cedric Killian1 and Olivier Sentieys1

1INRIA, FR; 2University of Rennes - INRIA, FR

**Abstract**

Voltage scaling has been used as a prominent technique to improve energy efficiency in digital systems, as reduction in the supply voltage effects in quadratic reduction in energy consumption of the system. The energy efficiency is achieved at the cost of timing errors in the system, that are corrected through additional error detection and correction circuits. In this paper we are proposing voltage over-scaling based approximate operators for applications that can tolerate errors. We characterize the basic arithmetic operators using different operating tribes (combination of supply voltage, back biasing scheme and clock frequency) to generate models for approximate operators. Error-resilient applications can be mapped with the generated approximate operator models to achieve optimum trade-off between energy and error margin. Based on the dynamic speculation technique, best possible operating tribe is chosen at runtime based on the user definable error tolerance margin of the application. In our experiments in 28nm FDSOI, we achieve maximum energy efficiency of 89% for basic operators like 8-bit and 16-bit adders at the cost of 20% Bit Error Rate (ratio of faulty bits over total bits) by operating them in near-threshold regime.

**Download Paper (PDF; Only available from the DATE venue WiFi)**

---

**5.3.3 COMBINING STRUCTURAL AND TIMING ERRORS IN OVERCLOCKED INEXACT SPECULATIVE ADDERS**

**Speaker:** Vincent Camus, EPFL, CH

**Authors:** Xun Jiao1, Vincent Camus2, Mattia Cacciotti3, Yu Jiang3, Christian Enz2 and Rajesh Gupta1

1UC San Diego, US; 2EPFL, CH; 3Tsinghua University, CN

**Abstract**

Worst-case design is used in IoT devices and high performance data centers to ensure reliability by adding extra safety margin, leading to a power efficiency loss. Recently, approximate computing has been proposed to trade off accuracy for efficiency. In this paper, we use an inexact speculative adder, which redesigns the adder architecture by shortening the critical path to save power consumption. Its overdesign introduces structural errors due to carry speculation. On the other hand, overclocking is used to reduce conservative timing guardbands but could introduce timing errors. In this paper, we apply a supervised learning model to overclocked inexact speculative adders to predict timing errors at bit level. We analyze these two types of errors and examine the joint effects of them.

**Download Paper (PDF; Only available from the DATE venue WiFi)**

---

**5.3.4 DVAFS: TRADING COMPUTATIONAL ACCURACY FOR ENERGY THROUGH DYNAMIC-VOLTAGE-ACCURACY-FREQUENCY-SCALING**

**Speaker:** Bert Moons, Katholieke Universiteit Leuven, BE

**Authors:** Bert Moons, Roel Uytterhoeven, Wim Dehaene and Marian Verhelst, Katholieke Universiteit Leuven, BE

**Abstract**

Several applications in machine learning and machine-to-human interactions tolerate small deviations in their computations. Digital systems can exploit this fault-tolerance to increase their energy-efficiency, which is crucial in embedded applications. Hence, this paper introduces a new means of Approximate Computing: Dynamic-Voltage-Accuracy-Frequency-Scaling (DVAFS), a circuit-level technique enabling a dynamic trade-off of energy versus computational accuracy that outperforms other Approximate Computing techniques. The usage and applicability of DVAFS is illustrated in the context of Deep Neural Networks, the current state-of-the-art in advanced recognition. These networks are typically executed on CPU’s or GPU’s due to their high computational complexity, making their deployment on battery-constrained platforms only possible through wireless connections with the cloud. This work shows how deep learning can be brought to IoT devices by running every layer of the network at its optimal computational accuracy. Finally, we demonstrate a DVAFS processor for Convolutional Neural Networks, achieving efficiencies of multiple TOPS/W.

**Download Paper (PDF; Only available from the DATE venue WiFi)**

---

**5.3.5 EXPLOITING COMPUTATION SKIP TO REDUCE ENERGY CONSUMPTION BY APPROXIMATE COMPUTING, AN HEVC ENCODER CASE STUDY**

**Speaker:** Daniel Menard, INSA Rennes, FR

**Authors:** Alexandre Mercat1, Justine Bonnot1, Maxime Pelcat2, Wassim Hamdouche1 and Daniel Menard1

1INSA Rennes, FR; 2IETR-INSIA, FR

**Abstract**

Approximate computing paradigm provides methods to optimize algorithms with considering both computational accuracy and complexity. This paradigm can be exploited at different levels of abstraction, from technological to application levels. Approximate computing at algorithm level aims at reducing computational complexity by approximating or skipping blocks function of the computation. Numerous applications in the signal and image processing domain integrate algorithms based on discrete optimization techniques. These techniques minimize a cost function by exploring the search space. In this paper, a new approach is proposed to exploit the computation-skipping approximate computing concept by using the SSSR technique. SSSR enables early selection of the best candidate configurations to reduce the search space. An efficient SSSR technique adjusts configuration selectivity to reduce execution complexity while selecting the functions most suitable to skip. The HEVC encoder in AI profile is used as a case study to illustrate the benefits of SSSR. In this application, two functions use discrete optimization to explore different solutions and select the one leading to the minimal cost in terms of bitrate/quality and computational energy: coding-tree partitioning and intra-mode prediction. By applying SSSR to this use case, energy reductions from 20% to 70% are explored through Pareto in Rate-Energy space.

**Download Paper (PDF; Only available from the DATE venue WiFi)**
5.4 Solutions for efficient simulation and validation

**Date:** Wednesday 29 March 2017

**Time:** 08:30 - 10:00

**Location / Room:** 3A

**Chair:**
Daniel Grosse, University of Bremen, DE

**Co-Chair:**
Alper Sen, Bogazici University, TR

The section introduces system-level frameworks for addressing memory tracing, timing estimation, real-time verification, and reliability degradation.

---

**LOCATION DETECTION FOR NAVIGATION USING IMUS WITH A MAP THROUGH COARSE-GRAINED MACHINE LEARNING**

**Speaker:**
Chen Luo, Rice University, US

**Authors:**
J. Jose Gonzales E., Chen Luo, Anshumali Shrivastava, Krishna Palem, Moon Yongshik, Soonhyun Noh, Dae Dong Park and Seongsoo Hong

**Abstract**
Location detection or localization supporting navigation has assumed significant importance in the recent past. In particular, techniques that exploit cheap inertial measurement units (IMU), the gyroscope and the accelerometer, have garnered attention, especially in an embedded computing context. However, these sensors measurements are quite unreliable, and it is widely believed that these sensors by themselves are too noisy for localization with acceptable accuracy. Consequently, several lines of work embody other costly alternatives to lower the impact of accumulated errors associated with IMU based approaches, invariably leading to very high energy costs resulting in lowered battery life. In this paper, we show that IMUs are sufficient by themselves if one augments them with known structural or geographical information about the physical area being explored by the user. By using the (em map) of the region being explored and the fact that humans typically walk in a structured manner, our approach sidesteps the challenges created by noise and concomitant accumulation of error. Specifically, we show that a simple coarse-grained machine learning approach mitigates the effect of the noisy perturbations in the information from our IMUs, provided we have accurate maps. Throughout, we rely on the principle of inexactness in an overarching manner and relax the need for absolute accuracy in return for significant lowering of resource (energy) costs. Namely, our approach is completely independent of any external guidance from sources including GPS, Bluetooth or WiFi support, and is this privacy preserving. Specifically, we show through experimental results that by relying on gyroscope and accelerometer data alone, we can correctly identify the path-segment where the user is walking/running on a known map, from sources including GPS, Bluetooth or WiFi support, and is this privacy preserving. Specifically, we show through experimental results that by relying on gyroscope and accelerometer data alone, we can correctly identify the path-segment where the user is walking/running on a known map, with an accuracy of 4.3 meters on the average using 0.44 Joules. This is a factor of 27X cheaper in energy lower than the "gold standard" that one could consider based on GPS support which, surprisingly, has an associated error of 8.7 meters on the average.

Download Paper (PDF; Only available from the DATE venue WiFi)
A NOVEL WAY TO EFFICIENTLY SIMULATE COMPLEX FULL SYSTEMS INCORPORATING HARDWARE ACCELERATORS

Speaker:
Nikolaos Tampouratzis, Technical University of Crete, GR

Authors:
Nikolaos Tampouratzis, Konstantinos Georgopoulos and Ioannis Papaefstathiou
1Technical University of Crete, GR; 2Telecommunication Systems Institute, Technical University of Crete, GR; 3Technical university of Crete, GR

Abstract
The breakdown of Dennard scaling coupled with the persistently growing transistor counts severely increased the importance of application-specific hardware acceleration; such an approach offers significant performance and energy benefits compared to general-purpose solutions. In order to thoroughly evaluate such architectures, the designer should perform a quite extensive design space exploration so as to evaluate the tradeoffs across the entire system. The design, until recently, has been predominantly done using Register Transfer Level (RTL) languages such as Verilog and VHDl, which, however, lead to a prohibitively long and costly design effort. In order to reduce the design time a wide range of both commercial and academic High-Level Synthesis (HLS) tools have emerged; most of those tools, handle hardware accelerators that are described in synthesizable SystemC. The problem today, however, is that most simulators used for evaluating the complete user applications (i.e. full-system CPU/Mem/Peripheral simulators) lack any type of SystemC accelerator support. Within this context this paper presents a novel simulation environment comprised of a generic SystemC accelerator and probably the most widely known full-system simulator (i.e. GEMS). The proposed system is the only solution supporting the very important feature of global synchronization across the integrated simulation; furthermore it has been evaluated based on two different computational/Intensive use cases and the final results demonstrate that the presented approach is orders of magnitude faster than the existing ones.

Download Paper (PDF; Only available from the DATE venue WiFi)
Weisheng Zhao, Beihang University, CN

Lionel Torres, LIRMM, CNRS/University of Montpellier, FR

Organisers:
Lionel Torres, LIRMM, CNRS/University of Montpellier, FR

Weisheng Zhao, Beihang University, CN

Chair:
Lionel Torres, LIRMM, CNRS/University of Montpellier, FR

Co-Chair:
Weisheng Zhao, Beihang University, CN

Numerous reports or industrial and academic works on emerging research devices identified magnetic tunnel junction (MTJ) (one of applications of Spintronics) as one of the most promising technologies to be part of the future of integrated systems. They provide non-volatility data, fast data access and low power operations. Indeed, MRAM or Magnetic memory based on the hybrid integration of MTJ have been commercialized since 2006 and used in a number of high-reliable applications. The aim of this session is to bring together the worldwide leading experts (from respectively USA, France, China, Japan and Germany) related to this hot topic to share the most recent results and discuss the future challenges. Different computing paradigms will be involved in this special session benefiting from interesting nature of spintronics devices. The invited speakers will talk about devices, design and compact modeling aspects, and applications, permitting a full development platform from devices to circuit & systems based on spintronics.

5.5 Hot Topic Session: Spintronics-based Computing

Date: Wednesday 29 March 2017
Time: 08:30 - 10:00
Location / Room: 3C

Organisers:
Lionel Torres, LIRMM, CNRS/University of Montpellier, FR
Weisheng Zhao, Beihang University, CN

Chair:
Lionel Torres, LIRMM, CNRS/University of Montpellier, FR

Co-Chair:
Weisheng Zhao, Beihang University, CN

Coffee Break in Exhibition Area

On all conference days (Tuesday to Thursday), coffee and tea will be served during the coffee breaks at the below-mentioned times in the exhibition area.

Tuesday, March 28, 2017
- Coffee Break 10:30 - 11:30
- Coffee Break 16:00 - 17:00

Wednesday, March 29, 2017
- Coffee Break 10:00 - 11:00
- Coffee Break 16:00 - 17:00

Thursday, March 30, 2017
- Coffee Break 10:00 - 11:00
- Coffee Break 15:30 - 16:00

5.5.1 MAGNETIC TUNNEL JUNCTION ENABLED ALL-SPIN STOCHASTIC SPIKING NEURAL NETWORK

Gopalakrishnan Srinivasan, Abhronil Sengupta and Kaushik Roy, Purdue University, US

Authors:
Kaushik Roy, Purdue University, US

Abstract
Biologically-inspired spiking neural networks (SNNs) have attracted significant research interest due to their inherent computational efficiency in performing classification and recognition tasks. The conventional CMOS-based implementations of large-scale SNNs are power intensive. This is a consequence of the fundamental mismatch between the technology used to realize the neurons and synapses, and the neuroscience mechanisms governing their operation, leading to area-expensive circuit designs. In this work, we present a three-terminal spintronic device, namely, the magnetic tunnel junction (MTJ)-heavy metal (HM) heterostructure that is inherently capable of emulating the neuronal and synaptic dynamics. We exploit the stochastic switching behavior of the MTJ in the presence of thermal noise to mimic the probabilistic spiking of cortical neurons, and the conditional change in the state of a binary synapse based on the pre- and post-synaptic spiking activity required for plasticity. We demonstrate the efficacy of a crossbar organization of our MTJ-HM based stochastic SNN in digit recognition using a comprehensive device-circuit-system simulation framework. The energy efficiency of the proposed system stems from the ultra-low switching energy of the MTJ-HM device, and the in-memory computation rendered possible by the localized arrangement of the computational units (neurons) and non-volatile synaptic memory in such crossbar architectures.

Download Paper (PDF; Only available from the DATE venue WiFi)
5.5.4 THREE-TERMINAL MTJ-BASED NONVOLATILE LOGIC CIRCUITS WITH SELF-TERMINATED WRITING MECHANISM FOR ULTRA-LOW-POWER VLSI PROCESSOR

Abstract
Magnetic random access memory (MRAM) has been widely studied for future nonvolatile working memory candidate. However, the mainstream current (spin transfer torque, STT or spin Hall effect, SHE) driven MRAMs (STT-MRAM or SHE-MRAM) face intrinsic problems in terms of high write power and long latency, significantly limiting the applications for low-power and high-speed working memories. The recently-developed new-generation MRAM, named VCMA-MRAM, which exploits the voltage-controlled magnetic anisotropy (VCMA) effect to write (or assist to write) data information into magnetic tunnel junctions (MTJs), holds the promise to efficiently overcome these problems. Despite the impressive possibility of improving write power and speed, this technology, however, is currently under intensive research and development (R&D), and some challenges still await answers. In this paper, we investigate the perspectives and challenges of VCMA-MRAM for working memories from a cross-layer (device/circuit/architecture) design point of view. We demonstrate that VCMA-MRAM outperforms STT-MRAM and SHE-MRAM in terms of area, speed, energy consumption and instruction-per-cycle (IPC) performance, benefiting from the low-power and high-speed VCMA-driven data writing mechanism. On the other hand, challenges in terms of device fabrication and circuit design should be efficiently addressed before practical applications.

Download Paper (PDF; Only available from the DATE venue WiFi)

11:00 End of session

Coffee Break in Exhibition Area
On all conference days (Tuesday to Thursday), coffee and tea will be served during the coffee breaks at the below-mentioned times in the exhibition area.

Tuesday, March 28, 2017
- Coffee Break 10:30 - 11:30
- Coffee Break 16:00 - 17:00

Wednesday, March 29, 2017
- Coffee Break 10:00 - 11:00
- Coffee Break 16:00 - 17:00

Thursday, March 30, 2017
- Coffee Break 10:00 - 11:00
- Coffee Break 15:30 - 16:00
This session deals with 3D reliability and repair, integration of compression into standard test infrastructure, and reusing silicon debug infrastructure to enhance functional performance.

**Time**  | **Label**  | **Presentation Title**  | **Authors**
--- | --- | --- | ---
08:30 | 5.6.1 | **FAULT CLUSTERING TECHNIQUE FOR 3D MEMORY BISR**  | Tianjian Li, Shanghai Jiao Tong University, CN  
Authors:  
Tianjian Li, Yan Han, Xiaoyao Liang, Hsien-Hsin S. Lee and Li Jiang  
1Shanghai Jiao Tong University, CN; 2TSMC / Georgia Tech, TW; 3Department of Computer Science and Engineering, Shanghai Jiao Tong University, CN  
**Abstract**  
Three Dimensional (3D) memory has gained a great momentum because of its large storage capacity, bandwidth and etc. A critical challenge for 3D memory is the significant yield loss due to the disruptive integration process: any memory die that cannot be successfully repaired leads to the failure of the whole stack. The repair ratio of each die must be as high as possible to guarantee the overall yield. Existing memory repair methods, however, follow the traditional way of using redundancies: a redundant row/column replaces a row/column containing few or even one faulty cell. We propose a novel technique specifically in 3D memory that can overcome this limitation. It can cluster faulty cells across layers to the same row/column in the same memory array so that each redundant row/column can repair more "faults". Moreover, it can be applied to the existing repair algorithms. We design the BIST and BISR modules to implement the proposed repair technique. Experimental results show more than 71% enhancement of the repair ratio over the global 3D GESP solution and 80% redundancy-cost reduction, respectively.

**Download Paper (PDF; Only available from the DATE venue WiFi)**

09:00 | 5.6.2 | **ARCHITECTURAL EVALUATIONS ON TSV REDUNDANCY FOR RELIABILITY ENHANCEMENT**  | Yen-Hao Chen, National Tsing Hua University, Taiwan, TW  
Authors:  
Yen-Hao Chen, Chien-Pang Chiu, Russell Barnes and TingTing Hwang  
1National Tsing Hua University, R.O.C, TW; 2University of California at Santa Barbara, US  
**Abstract**  
Three-dimensional Integrated Circuits (3D-ICs) is a next-generation technology that could be a solution to overcome the scaling problem. It stacks dies with Through-Silicon Vias (TSVs) so that signals can be transmitted through dies vertically. However, researchers have noticed that the aging effect due to the electromigration (EN) may result in faulty TSVs and affect the chip lifetime [1]. Several redundant TSV architectures have been proposed to address this issue. By replacing the faulty TSV with redundant TSVs which are added at design time, chips can achieve better reliability and longer lifetime. In this paper, we will study the tradeoff of various redundant TSV architectures in terms of effectiveness and cost. To allow the measurement of reliability more realistically, we propose a new standard, repair rate, to appraise the redundant TSV architectures. Moreover, to design a more flexible and efficient structure, we enhance the ring-based design [2] that can adjust the size of the TSV block and TSV redundancy.

**Download Paper (PDF; Only available from the DATE venue WiFi)**

09:30 | 5.6.3 | **REUSING TRACE BUFFERS TO ENHANCE CACHE PERFORMANCE**  | Neetu Jindal, PhD, IN  
Authors:  
Neetu Jindal, Preeti Ranjan Panda and Smruti R. Sarangi, Indian Institute of Technology Delhi, IN  
**Abstract**  
With the increasing complexity of modern Systems-on-Chip, the possibility of functional errors escaping design verification is growing. Post-silicon validation targets the discovery of these errors in early hardware prototypes. Due to limited visibility and observability, dedicated design-for-debug (DFD) hardware such as trace buffers are inserted to aid post-silicon validation. In spite of its benefit, such hardware incurs area overheads, which impose size limitations. However, the overhead could be overcome if the area dedicated to DFD could be reused in-field. In this work, we present a novel method for reusing an existing trace buffer as a victim cache of a processor to enhance performance. The trace buffer storage space is reused for the victim cache, with a small additional controller logic. Experimental results on several benchmarks and trace buffer sizes show that the proposed approach can enhance the average performance by up to 8.3% over a baseline architecture. We also propose a strategy for dynamic power management of the structure, to enable saving energy with negligible impact on performance.

**Download Paper (PDF; Only available from the DATE venue WiFi)**

09:45 | 5.6.4 | **OPTIMIZATION OF RETARGETING FOR IEEE 1149.1 TAP CONTROLLERS WITH EMBEDDED COMPRESSION**  | Sebastian Huhn, University of Bremen, DE  
Authors:  
Sebastian Huhn, Stephan Eggersglüß, Krishnendu Chakrabarty and Rolf Drechsler  
1University of Bremen, DE; 2Duke University, US; 3University of Bremen/DFKI GmbH, DE  
**Abstract**  
We present a formal optimization technique that enables retargeting for codeword-based IEEE 1149.1-compliant TAP controllers. The proposed method addresses the problem of high test data volume and Test Application Time (TAT) for a system-on-chip design during board or in-field testing, as well as during debugging. This procedure determines an optimal set of codewords with respect to given hardware constraints, e.g., embedded dictionary size and the interface to the Test Data Register in the IEEE 1149.1 Std. A complete traversal of the spanned search space is possible through the use of formal methods. An optimal set of codewords can be determined, which is directly utilized for retargeting. The proposed method is evaluated using test data with high-entropy, which is known to be the least amenable to compression, as well as input data for debugging and Functional Verification (FV) test data. Our results show a compression ratio improvement of more than 30% and a reduction in TAT up to 20% compared to previous techniques.

**Download Paper (PDF; Only available from the DATE venue WiFi)**
NOVEL MAGNETIC BURN-IN FOR RETENTION TESTING OF STTRAM

Speaker:
Swaroop Ghosh, Pennsylvania State University, US

Authors:
Mohammad Nasim Imtiaz Khan, Anirudh Iyengar and Swaroop Ghosh, Pennsylvania State University, US

Abstract
Spin-Transfer Torque RAM (STTRAM) is an emerging Non-Volatile Memory (NVM) technology that has drawn significant attention due to complete elimination of bitcell leakage. However, it brings new challenges in characterizing the retention time of the array during test. Significant shift of retention time under static (process variation (PV)) and dynamic (voltage, temperature fluctuation) variability furthers this issue. In this paper, we propose a novel magnetic burn-in (MBI) test which can be implemented with minimal changes in the existing test flow to enable STTRAM retention testing at short test time. The magnetic burn-in is also combined with thermal burn-in (MBI+BI) for further compression of retention and test time. Simulation results indicate MBI with 220Oe (at 25C) can improve the test time by 3.71x1013 X while MBI+BI with 2200e at 125C can improve the test time by 1.97x1014X.

Download Paper (PDF; Only available from the DATE venue WiFi)

10:00 End of session
Coffee Break in Exhibition Area

On all conference days (Tuesday to Thursday), coffee and tea will be served during the coffee breaks at the below-mentioned times in the exhibition area.

Tuesday, March 28, 2017
- Coffee Break 10:30 - 11:30
- Coffee Break 16:00 - 17:00

Wednesday, March 29, 2017
- Coffee Break 10:00 - 11:00
- Coffee Break 16:00 - 17:00

Thursday, March 30, 2017
- Coffee Break 10:00 - 11:00
- Coffee Break 15:30 - 16:00

5.7 Schedulability Analysis

Date: Wednesday 29 March 2017
Time: 08:30 - 10:00
Location / Room: 3B

Chair:
Petru Eles, Linköpings universitet, SE

Co-Chair:
Andreas Naderlinger, University of Salzburg, AT

The papers in this session introduce new schedulability analyses for real-time systems, including systems with precedence constraints, real-time networks-on-chip, and mixed-critical systems.

08:30 5.7.1 BOUNDING DEADLINE MISSES IN WEAKLY-HARD REAL-TIME SYSTEMS WITH TASK DEPENDENCIES

Speaker:
Zain A. H. Hammadeh, TU Braunschweig, DE

Authors:
Zain A. H. Hammadeh1, Sophie Quinton2, Rolf Ernst1, Rafik Henia3 and Laurent Rioux3
1TU Braunschweig, DE; 2Inria, FR; 3Thales Research & Technology, FR

Abstract
Real-time systems with functional dependencies between tasks often require end-to-end (as opposed to task-level) guarantees. For many of these systems, it is even possible to accept the possibility of longer end-to-end delays if one can bound their frequency. Such systems are called weakly-hard. In this paper we provide end-to-end deadline miss models for systems with task chains using Typical Worst-Case Analysis (TWCA). This bounds the number of potential deadline misses in a given sequence of activations of a task chain. To achieve this we exploit task chain properties which arise from the priority assignment of tasks in static-priority preemptive systems. This work is motivated by and validated on a realistic case study inspired by industrial practice and synthetic test cases.

Download Paper (PDF; Only available from the DATE venue WiFi)

09:00 5.7.2 REAL-TIME COMMUNICATION ANALYSIS FOR NETWORKS-ON-CHIP WITH BACKPRESSURE

Speaker:
Sebastian Tobuschat, TU Braunschweig, DE

Authors:
Sebastian Tobuschat and Rolf Ernst, TU Braunschweig, DE

Abstract
Networks-on-Chip (NoCs) for safety-critical domains require formal guarantees for the worst-case behavior of all real-time senders. The majority of existing analysis approaches is capable of providing such guarantees only under the assumption that the queues in the routers never overflow, i.e., that no backpressure occurs. This leads to overly pessimistic guarantees or unfulfilled design requirements in many setups using commercially available NoCs where buffer space is limited. Therefore, we propose an alternative analysis methodology providing formal timing guarantees for packet latencies also in a NoC where backpressure occurs. The analysis allows exploiting the behavior of individual traffic streams to determine safe upper bounds on the latency of individual packets. The correctness of the analysis is evaluated experimentally through comparison with simulation results.

Download Paper (PDF; Only available from the DATE venue WiFi)
### PROBABILISTIC SCHEDULABILITY ANALYSIS FOR FIXED PRIORITY MIXED CRITICALITY REAL-TIME SYSTEMS

**Speaker:**
Yasmina Abdeddaïm, Université Paris-East, LIGM, ESIEE Paris, FR

**Authors:**
Yasmina Abdeddaim\(^1\) and Dorin Maxim\(^2\)

1. Université Paris-East, LIGM, ESIEE-Paris, FR; 2. University of Lorraine - Loria - Inria Nancy Grand Est, FR

**Abstract**
In this paper we present a probabilistic response time analysis for mixed criticality real-time systems running on a single processor according to a fixed priority pre-emptive scheduling policy. The analysis extends the existing state of the art probabilistic analysis to the case of mixed criticalities, taking into account both the level of assurance at which each task needs to be certified, as well as the possible criticalities at which the system may execute. The proposed analysis is formally presented as well as explained with the aid of an illustrative example.

Download Paper (PDF; Only available from the DATE venue WiFi)

---

### IP2 Interactive Presentations

**Date:** Wednesday 29 March 2017

**Time:** 10:00 - 10:30

**Location / Room:** IP sessions (in front of rooms 4A and 5A)

Interactive Presentations run simultaneously during a 30-minute slot. A poster associated to the IP paper is on display throughout the morning. Additionally, each IP paper is briefly introduced in a one-minute presentation in a corresponding regular session, prior to the actual Interactive Presentation. At the end of each afternoon Interactive Presentations session the award 'Best IP of the Day' is given.

<table>
<thead>
<tr>
<th>Label</th>
<th>Presentation Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>IP2-1</td>
<td>COMPACT MODELING AND CIRCUIT-LEVEL SIMULATION OF SILICON NANOPHOTONIC INTERCONNECTS</td>
</tr>
<tr>
<td></td>
<td>Speaker: Yuyang Wang, UC Santa Barbara, US</td>
</tr>
<tr>
<td></td>
<td>Authors: Rui Wu, Yuyang Wang, Zeyu Zhang, Chong Zhang, Clint Schow, John Bowers and Kwang-Ting Cheng, UC Santa Barbara, US</td>
</tr>
<tr>
<td></td>
<td>Abstract Nanophotonic interconnects have been playing an increasingly important role in the datacom regime. Greater integration of silicon photonics demands modeling and simulation support for design validation, optimization and design space exploration. In this work, we develop compact models for a number of key photonic devices, which are extensively validated by the measurement data of a fabricated optical network-on-chip (ONoC). Implemented in SPICE-compatible Verilog-A, the models are used in circuit-level simulations of full optical links. The simulation results match well with the measurement data. Our model library and simulation approach enable the electro-optical (EO) co-simulation, allowing designers to include photonic devices in the whole system design space, and to co-optimize the transmitter, interconnect, and receiver jointly.</td>
</tr>
<tr>
<td></td>
<td>Download Paper (PDF; Only available from the DATE venue WiFi)</td>
</tr>
</tbody>
</table>

| IP2-2 | A TRUE RANDOM NUMBER GENERATOR BASED ON PARALLEL STT-MTJS |
|       | Speaker: Yuanzhuo Qu, University of Alberta, CA |
|       | Authors: Yuanzhuo Qu\(^1\), Jie Han\(^1\), Bruce Cockburn\(^1\), Yue Zhang\(^2\), Weisheng Zhao\(^2\) and Witold Pedrycz\(^1\) |
|       | 1. University of Alberta, CA; 2. Beihang University, CN |
|       | Abstract Random number generators are an essential part of cryptographic systems. For the highest level of security, true random number generators (TRNG) are needed instead of pseudo-random number generators. In this paper, the stochastic behavior of the spin transfer torque magnetic tunnel junction (STT-MTJ) is utilized to produce a TRNG design. A parallel structure with multiple MTJs is proposed that minimizes device variation effects. The design is validated in a 28-nm CMOS process with Monte Carlo simulation using a compact model of the MTJ. The National Institute of Standards and Technology (NIST) statistical test suite is used to verify the randomness quality when generating encryption keys for the Transport Layer Security or Secure Sockets Layer (TLS/SSL) cryptographic protocol. This design has a generation speed of 177.8 Mbit/s, and an energy of 0.64 pJ is consumed to set up the state in one MTJ. |
|       | Download Paper (PDF; Only available from the DATE venue WiFi) |
IP2-3 ENABLING AREA EFFICIENT RF ICS THROUGH MONOLITHIC 3D INTEGRATION

Speaker: Panagiotis Chaourani, KTH, Royal Institute of Technology, Stockholm, SE
Authors: Panagiotis Chaourani, Per-Erik Hellström, Saul Rodriguez, Raúl Onet and Ana Rusu, KTH, Royal Institute of Technology, SE
Abstract
The Monolithic 3D (M3D) integration technology has emerged as a promising alternative to dimensional scaling thanks to the unprecedented integration density capabilities and the low interconnect parasitics that it offers. In order to support technological investigations and enable future M3D circuits, M3D design methodologies, flows and tools are essential. Prospective M3D digital applications have attracted a lot of scientific interest. This paper identifies the potential of M3D RF/analog circuits and presents the first attempt to demonstrate such circuits. Towards this, a M3D custom design platform, which is fully compatible with commercial design tools, is proposed and validated. The design platform includes process characteristics, device models, LVS and DRC rules and a parasitic extraction flow. The envisioned M3D structure is built on a commercial CMOS process that serves as the bottom tier, whereas a SOI process is used as top tier. To validate the proposed design flow and to investigate the potential of M3D RF/analog circuits, a RF front-end design for Zig-Bee WPAN applications is used as case-study. The M3D RF front-end circuit achieves 35.5 % area reduction, while showing similar performance with the original 2D circuit.
Download Paper (PDF; Only available from the DATE venue WiFi)

IP2-4 RECONFIGURABLE THRESHOLD LOGIC GATES USING OPTOELECTRONIC CAPACITORS

Speaker: Baris Taskin, Drexel University, US
Authors: Ragh Kuttappa, Lunar Khun, Bahram Nabet and Baris Taskin, Drexel University, US
Abstract
This paper investigates the integration of optoelectronic devices with CMOS threshold logic gates to design reconfigurable Boolean functions. The weight of the optoelectronic cell can be altered by changing the optical power which is used to reconfigure the threshold logic (TL) gate. The proposed optoelectronic capacitor based TL (OECTL) gates are designed for i) simplistic AND/NAND gates and OR/NOR gates with large fan-in and ii) linearly separable Boolean functions that can be reconfigured to other linearly separable Boolean functions, constrained in reconfiguration by the specifics of TL operation. SPICE simulations in 65nm bulk CMOS technology with a Verilog-A model for the optoelectronic capacitor demonstrate i) AND/NAND gates and OR/NOR gates are 2X faster as fan0in increases and consumes low power ii) Boolean function can be reconfigured with 0.58X smaller delay and 0.46X lesser power of standard CMOS.
Download Paper (PDF; Only available from the DATE venue WiFi)

IP2-5 I-BEP: A NON-REDUNDANT AND HIGH-CONCURRENCY MEMORY PERSISTENCY MODEL

Speaker: Yuanchao Xu, Capital Normal University, CN
Authors: Yuanchao Xu, Zeyi Hou, Junfeng Yan, Lu Yang and Hu Wan, Capital Normal University, CN
Abstract
Byte-addressable, non-volatile memory (NVM) technologies enable fast persistent updates but incur potential data inconsistency upon a failure. Recent proposals present several persistency models to guarantee data consistency. However, they fail to express the minimal persist ordering as a result of inducing unnecessary ordering constraints. In this paper, we propose i-BEP, a non-redundant high concurrency memory persistency model, which expresses epoch dependency via persist directed acyclic graph instead of program order. Additionally, we propose two techniques, background persist and deferred eviction, to enhance the performance of i-BEP. We demonstrate that i-BEP can improve the performance by 15% for typical data structures on average over buffered epoch persistency (BEP) model.
Download Paper (PDF; Only available from the DATE venue WiFi)

IP2-6 SPMS: STRAND BASED PERSISTENT MEMORY SYSTEM

Speaker: Shuo Li, National University of Defense Technology, CN
Authors: Shuo Li1, Peng Wang2, Nong Xiao1, Guangyu Sun2 and Fang Liu1
1National University of Defense Technology, CN; 2Peking University, CN
Abstract
Byte-addressable, non-volatile memory (NVM) technologies enable fast persistent updates but incur potential data inconsistency upon a failure. Recent proposals present several persistency models to guarantee data consistency. However, they fail to express the minimal persist ordering as a result of inducing unnecessary ordering constraints. In this paper, we propose i-BEP, a non-redundant high concurrency memory persistency model, which expresses epoch dependency via persist directed acyclic graph instead of program order. Additionally, we propose two techniques, background persist and deferred eviction, to enhance the performance of i-BEP. We demonstrate that i-BEP can improve the performance by 15% for typical data structures on average over buffered epoch persistency (BEP) model.
Download Paper (PDF; Only available from the DATE venue WiFi)

IP2-7 ARCHITECTING HIGH-SPEED COMMAND SCHEDULERS FOR OPEN-ROW REAL-TIME SDRAM CONTROLLERS

Speaker: Leonardo Ecco, TU Braunschweig, DE
Authors: Leonardo Ecco1 and Rolf Ernst2
1Institute of Computer and Network Engineering, TU Braunschweig, DE; 2TU Braunschweig, DE
Abstract
As SDRAM modules get faster and their data buses wider, researchers proposed the use of the open-row policy in command schedulers for real-time SDRAM controllers. While the real-time properties of such schedulers have been thoroughly investigated, their hardware implementation was not. Hence, in this paper, we propose a highly-parallel and multi-stage architecture that implements a state-of-the-art open-row real-time command scheduler. Moreover, we evaluate such architecture from the hardware overhead and performance perspectives.
Download Paper (PDF; Only available from the DATE venue WiFi)

IP2-8 AUTOMATIC EQUIVALENCE CHECKING FOR SYSTEMC-TLM 2.0 MODELS AGAINST THEIR FORMAL SPECIFICATIONS

Speaker: Mehran Goli, University of Bremen, DE
Authors: Mehran Goli, Jannis Stopp and Rolf Drechsler, University of Bremen, DE
Abstract
The necessity to handle the increasing complexity of digital circuits has led to the usage of more and more abstract design paradigms. In particular, the Electronic System Level (ESL) has become an area of active research and industrial application, especially via SystemC and its Transaction Level Modeling (TLM) framework. Additionally, the usage of formal specification languages such as the Unified Modeling Language (UML) prior to the implementation (even at higher abstraction levels) is now a broadly accepted workflow. Utilizing this layered approach leaves the translation from the specification to the implementation to the designer, leaving the question unanswered how the equivalence of these should be verified. This paper proposes a novel, non-intrusive and broadly applicable approach to automatically validate the equivalence of the structural and behavioral information of a SystemC-TLM 2.0 model and its formal specification.
Download Paper (PDF; Only available from the DATE venue WiFi)
The use of Sobol sequences can lead to a higher energy efficiency in an SC circuit than using Halton sequences. A type of low discrepancy (LD) sequences, the Sobol sequence, is considered for energy-efficient implementations of SC circuits. The use of Sobol sequences presents a significant challenge for stochastic computing (SC) due to the long random binary bit streams required for accurate computation. In this paper, a type of low discrepancy (LD) sequences, the Sobol sequence, is considered for energy-efficient implementations of SC circuits. The use of Sobol sequences improves the output accuracy of a stochastic circuit with a reduced sequence length compared to the use of another type of LD sequences, the Halton sequence, and conventional LFSR-generated pseudorandom sequences. The use of Sobol sequences leads to a higher accuracy than using Halton sequences for basic arithmetic operations. Sobol sequence generators cost less energy than the Halton counterparts when multiple random sequences are required in a circuit, thus the use of Sobol sequences can lead to a higher energy efficiency in an SC circuit than using Halton sequences.
IP2-14  LOGIC ANALYSIS AND VERIFICATION OF N-INPUT GENETIC LOGIC CIRCUITS
Speaker: Hasan Baig, Technical University of Denmark, DK
Authors: Hasan Baig and Jan Madsen, Technical University of Denmark, DK
Abstract: Nature is using genetic logic circuits to regulate the fundamental processes of life. These genetic logic circuits are triggered by a combination of external signals, such as chemicals, proteins, light and temperature, to emit signals to control other gene expressions or metabolic pathways accordingly. As compared to electronic circuits, genetic circuits exhibit stochastic behavior and do not always behave as intended. Therefore, there is a growing interest in being able to analyze and verify the logical behavior of a genetic circuit model, prior to its physical implementation in a laboratory. In this paper, we present an approach to analyze and verify the Boolean logic of a genetic circuit from the data obtained through stochastic analog circuit simulations. The usefulness of this analysis is demonstrated through different case studies illustrating how our approach can be used to verify the expected behavior of an n-input genetic logic circuit.
Download Paper (PDF; Only available from the DATE venue WiFi)

IP2-15  A NOVEL WAY TO EFFICIENTLY SIMULATE COMPLEX FULL SYSTEMS INCORPORATING HARDWARE ACCELERATORS
Speaker: Nikolaos Tampouratzis, Technical University of Crete, GR
Authors: Nikolaos Tampouratzis, Konstantinos Georgopoulos and Ioannis Papaefstathiou
1Technical University of Crete, GR; 2Telecommunication Systems Institute, Technical University of Crete, GR; 3Technical university of Crete, GR
Abstract: The breakdown of Dennard scaling coupled with the persistently growing transistor counts severely increased the importance of application-specific hardware acceleration; such an approach offers significant performance and energy benefits compared to general-purpose solutions. In order to thoroughly evaluate such architectures, the designer should perform a quite extensive design space exploration so as to evaluate the tradeoffs across the entire system. The design, until recently, has been predominantly done using Register Transfer Level (RTL) languages such as Verilog and VHDL, which, however, lead to a prohibitively long and costly design effort. In order to reduce the design time a wide range of both commercial and academic High-Level Synthesis (HLS) tools have emerged; most of those tools, handle hardware accelerators that are described in synthesizable SystemC. The problem today, however, is that most simulators used for evaluating the complete user applications (i.e. full-system CPU/Mem/Peripheral simulators) lack any type of SystemC accelerator support. Within this context this paper presents a novel simulation environment comprised of a generic SystemC accelerator and probably the most widely known fullsystem simulator (i.e. GenS). The proposed system is the only solution supporting the very important feature of global synchronization across the integrated simulation; furthermore it has been evaluated based on two different computationallyintensive use cases and the final results demonstrate that the presented approach is orders of magnitude faster than the existing ones.
Download Paper (PDF; Only available from the DATE venue WiFi)

IP2-16  AUTOMATIC ABSTRACTION OF MULTI-DISCIPLINE ANALOG MODELS FOR EFFICIENT FUNCTIONAL SIMULATION
Speaker: Franco Fummi, Università degli Studi di Verona, IT
Authors: Enrico Fraccaroli1, Michele Lora1 and Franco Fummi2
1University of Verona, IT; 2Università di Verona, IT
Abstract: Multi-discipline components introduce problems when inserted within virtual platforms of Smart Systems for functional validation. This paper lists the most common emerging problems and it proposes a new solution of a set of solutions to them. It presents a set of techniques, unified in an automatic abstraction methodology, useful to achieve fast analog mixed-signal simulation even when different physical disciplines and modeling styles are combined into a single analog model. The paper makes use of a complex case study. It deals with multiple-discipline descriptions, non-electrical conservative models, non-linear equation systems, and mixed time/frequency domain models. The original component behavior has been modeled in Verilog-AMS by using electrical, mechanical and kinematic equations. Then, it has been abstracted and integrated within a virtual platform of a mixed-signal smart system for efficient functional simulation.
Download Paper (PDF; Only available from the DATE venue WiFi)

IP2-17  NOVEL MAGNETIC BURN-IN FOR RETENTION TESTING OF STRAM
Speaker: Swaroop Ghosh, Pennsylvania State University, US
Authors: Mohammad Nasim Imtiaz Khan, Anirudh Iyengar and Swaroop Ghosh, Pennsylvania State University, US
Abstract: Spin-Transfer Torque RAM (STTRAM) is an emerging Non-Volatile Memory (NVM) technology that has drawn significant attention due to complete elimination of bitcell leakage. However, it brings new challenges in characterizing the retention time of the array during test. Significant shift of retention time under static (process variation (PV)) and dynamic (voltage, temperature fluctuation) variability furthers this issue. In this paper, we propose a novel mag-netic burn-in (MBI) test which can be implemented with minimal changes in the existing test flow to enable STTRAM retention testing at short test time. The magnetic burn-in is also combined with thermal burn-in (MBI+BI) for further compression of retention and test time. Simula-tion results indicate MBI with 220Oe (at 25C) can improve the test time by 3.71x1013 X while MBI+BI with 220Oe at 125C can improve the test time by 1.97x1014X. MBI+BI with 220Oe at 125C can be implemented with minimal changes in the existing test flow to enable STTRAM retention testing at short test time. The magnetic burn-in is also combined with thermal burn-in (MBI+BI) for further compression of retention and test time. Simula-tion results indicate MBI with 220Oe at 125C can improve the test time by 3.71x1013 X while MBI+BI with 220Oe at 125C can improve the test time by 1.97x1014X.
Download Paper (PDF; Only available from the DATE venue WiFi)

IP2-19  AUTOMATIC CONSTRUCTION OF MODELS FOR ANALYTIC SYSTEM-LEVEL DESIGN SPACE EXPLORATION PROBLEMS
Speaker: Seyed-Hosein Attarzadeh-Niaiki, Shahid Beheshti University (SBU), IR
Authors: Seyed-Hosein Attarzadeh-Niaiki and Ingo Sander 1
1Shahid Beheshti University (SBU), IR; 2KTH Royal Institute of Technology, SE
Abstract: Due to the variety of application models and also the target platforms used in embedded electronic system design, it is challenging to formulate a generic and extensible analytic design-space exploration (DSE) framework. Current approaches support a restricted class of application and platform models and are difficult to extend. This paper proposes a framework for automatic construction of system-level DSE problem models based on a coherent, constraint-based representation of system functionality, flexible target platforms, and binding policies. Heterogeneous semantics is captured using constraints on logical clocks. The applicability of this method is demonstrated by constructing DSE problem models from different combinations of application and platforms models. Time-triggered and untimed models of the system functionality and heterogeneous target platforms are used for this purpose. Another potential advantage of this approach is that constructed models can be solved using a variety of standard and ad-hoc solvers and search heuristics.
Download Paper (PDF; Only available from the DATE venue WiFi)
NOXIM-XT: A BIT-ACCURATE POWER ESTIMATION SIMULATOR FOR NOCS
Presenter: Pierre Bonel, Université de Bretagne Sud, FR
Authors: André Ross¹, Johann Laurent² and Erwan Morec²
¹LERIA, Université d’Angers, Angers, France; ²Lab-STICC, Université de Bretagne Sud, Lorient, FR
Abstract
We have developed an enhanced version of Noxim (Noxim-XT) to estimate the energy consumption of a NoC in a SOC. Noxim-XT is used in a two-step methodology. First, applications are mapped on a SoC and their traffics are extracted by simulation with MPsOCBench. Second, Noxim-XT tests various hardware configurations of the NoC, and for each configuration, the application’s traffic is re-injected and replayed, an accurate performance and power breakdown is provided, and the user can choose different data coding strategies. With the help of Noxim XT, each configuration is bit-accurately estimated in terms of energy consumption. After simulation, a spatial mapping of the energy consumption is provided and highlights the hot-spots. Moreover, the new coding strategies allow significant energy saving. Noxim XT simulations and a FPGA-based prototype of a new coding strategy will be demonstrated at the U-booth to illustrate these works.
More information ...

RIMEDIO: WHEELCHAIR MOUNTED ROBOTIC ARM DEMONSTRATOR FOR PEOPLE WITH MOTOR SKILLS IMPAIRMENTS
Presenter: Alessandro Palla, University of Pisa, IT
Authors: Gabriele Meoni and Luca Fanucci, University of Pisa, IT
Abstract
People with reduced mobility experiment many issues in the interaction with the indoor and outdoor environment because of their disability. For those users even the simplest action might be a hard/impossible task to perform without the assistance of an external aid. We propose a simple and lightweight wheelchair mounted robotic arm with the focus on the human-machine interface that has to be simple and accessible for users with different kind of disabilities. The robotic arm is equipped with a 5 MP camera, force and proximity sensors and a 6 axis Inertial Measurement Unit on the end-effector that can be controlled using an app running on a tablet. When the user selects the object to reach (for instance a button) on the tablet screen, the arm autonomously carries out the task, using the camera image and the sensors measurements for autonomous navigation. The demonstrator consists in the robotic arm prototype, the Android tablet and a personal computer for arm setup and configuration.
More information ...

NDNNN: NEURAL NETWORKS DESIGNING NEURAL NETWORKS
Presenter: Brett Meyer, McGill University, US
Authors: Warren Gross, Sean Smithson, Ossama Ahmed and Guang Yang, McGill University, US
Abstract
Modern artificial neural networks currently achieve state-of-the-art results in various difficult problems, including image classification and speech recognition. However, both the performance and computational complexity of such models are heavily dependent on the design of characteristic hyper-parameters (e.g., numbers of hidden layers or nodes per layer) which are often manually optimized. With neural networks penetrating low-power mobile and embedded areas, the need now arises to optimize not only for performance, but also for implementation cost. In our work, we present a multi-objective design space exploration method leveraging machine learning based response surface modelling to reduce the number of solutions trained and evaluated. Experimental results are presented for several image recognition datasets, demonstrating the evolution of the approximated Pareto-optimal hyper-parameters and corresponding GPU code; all while exploring only a small fraction of the design space.
More information ...

MATISSE: A TARGET-AWARE COMPILER TO TRANSLATE MATLAB INTO C AND OPENCL
Presenter: Luis Reis, University of Porto, PT
Authors: João Bispo and João Cardoso, University of Porto / INESC-TEC, PT
Abstract
Many engineering, scientific and finance algorithms are prototyped and validated in array languages, such as MATLAB, before being converted to other languages such as C for use in production. As such, there has been substantial effort to develop compilers to perform this translation automatically. Alternative types of compilation devices, such as GPGPUs and FPGAs, are becoming increasingly more popular, so it becomes critical to develop compilers that target these architectures. We have adapted MATISSE, our MATLAB-compatible compiler framework, to generate C and OpenCL code for these platforms. In this demonstration, we will show how our compiler works and what its capabilities are. We will also describe the main challenges of efficient code generation from MATLAB and how to overcome them.
More information ...

SCCHARTS: SYNCHRONOUS STATECHARTS FOR SAFETY-CRITICAL APPLICATIONS
Presenter: Reinhard von Hanxleden, Kiel University, DE
Authors: Michael Menderl¹, Christian Motika², Christoph Daniel Schulze² and Steven Smyth ²
¹Bamberg University, DE; ²Kiel University, DE
Abstract
We present a visual language, SCCharts, designed for specifying safety-critical reactive systems. SCCharts use a statechart notation and provide determinate concurrency based on a synchronous model of computation (MoC), without restrictions common to previous synchronous MoCs. Specifically, we lift earlier limitations on sequential accesses to shared variables, by leveraging the sequentially constructive MoC. For further details, see [von Hanxleden et al., PLDI’14] and http://www.sccharts.com. The SCCharts demonstrator is an Eclipse Rich Client and part of KIELER (http://www.rtbsys.informatik.uni-kiel.de/en/research/kieler). The demonstration shows how to write an SCChart model using a textual notation, from which a visual model is generated on the fly using the Eclipse Layout Kernel (ELK). We also present a compilation chain that allows efficient synthesis of software and hardware.
More information ...

MULTI-CORE VERIFICATION: COMBINING MICROTESK AND SPIN FOR VERIFICATION OF MULTI-CORE MICROPROCESSORS
Presenter: Mikhail Chupilko, ISPRAS, RU
Authors: Alexander Kamkin, Mikhail Lebedev and Andrei Tatarnikov, ISPRAS, RU
Abstract
The complexity of modern cache coherence protocols (CCP) in multi-core microprocessors prevents from complete verification of shared memory subsystems by means of random test-program generators (TPG). The following steps are suggested to target the problem. The first step is to separately specify CCP features and generate CCP-specific events to be used in TPG when generating a test program (TP). The protocol is specified in Promela, with Spin making a test template (TT). Spin also produces UVR (or C+TESK) testbench to make the execution of the resulting TPs to be controllable and deterministic. The second step is to let TPG produce the memory access instructions causing desired CCP-specific behavior. As a TPG we use MicroTESK. Its Ruby-based TTs abstractly describe future TPs. MicroTESK processes that TP making TP with CCP-specific events. The resulting TP is executed together with the testbench to exactly reproduce the situation Spin had found to be important for such a protocol.
More information ...
The introduction and broad scale rollout of IoT applications puts pressing demands on semiconductor base technologies for computation, communication and sensing in terms of lowest cost, power dissipation, dependability, security and the ability to integrate heterogeneous devices and technologies. This session presents three research-oriented perspectives on the challenging aspects of IoT enabling technologies.
11:00  6.1.1  ULTRA-LOW-POWER CIRCUITS FOR IOT APPLICATIONS
Author:  Georges Gielen, Katholieke Universiteit Leuven, BE
Abstract:  IoT applications require ultra-low-power hardware solutions that communicate wirelessly. Challenges and some solutions in designing these will be highlighted.

11:30  6.1.2  STRUCTURAL HEALTH MONITORING FOR SMART CITIES: A HW/SW CODESIGN PERSPECTIVE
Author:  Jiang Xu, Hong Kong University of Science and Technology, HK
Abstract:  The structural integrity of civil structures is vital to economic prosperity and public safety. In developed countries and regions, a large number of transportation and residential infrastructures are aging rapidly. There is an urgent need and rapidly increasing demand for the ability to monitor the health conditions of civil structures in a real-time and distributed manner. This talk will share our experiences on developing large scale structural health monitoring systems from a HW/SW codesign perspective.

12:00  6.1.3  SECURITY IN THE INTERNET OF THINGS: A CHALLENGE OF SCALE
Speaker and Author:  Patrick Schaumont, Virginia Tech, US
Abstract:  Technological scaling has offered a windfall of benefits to electronics design. Increased transistor density has offered an exponential increase in computing capabilities over time, but without a corresponding increase in system cost. Information security has its own success story with scaling. Cryptographic algorithms become exponentially harder to break through a mere linear increase in encryption complexity or in key-length. In the Internet of Things, scaling is as much a security liability as it is an advantage. These security liabilities are new, poorly understood and poorly regulated. Some examples include the following: privacy of IoT data in the cloud; the safety consequences of poor information security in cyber-physical systems; the liabilities of long-lifetime devices that use outdated or poorly tested information security; the performance-limited information security in devices that run on the outskirts of the IoT using nothing but harvested energy. In this contribution we consider the security landscape for IoT. We consider the technological consequences of securely extending the Internet into the physical world of things. We identify current limitations, ongoing research efforts, and open challenges for the design community.

Download Paper (PDF; Only available from the DATE venue WiFi)

12:30  End of session
Lunch Break in Garden Foyer
Keynote Lecture session 7.0 in “Garden Foyer” 1350 - 1420

Lunch Break in the Garden Foyer
On all conference days (Tuesday to Thursday), a buffet lunch will be offered in the Garden Foyer, in front of the session rooms. Kindly note that this is restricted to conference delegates possessing a lunch voucher only. When entering the lunch break area, delegates will be asked to present the corresponding lunch voucher of the day. Once the lunch area is being left, re-entrance is not allowed for the respective lunch.

Date: Wednesday 29 March 2017
Time: 11:00 - 12:30
Location / Room: 4BC
Chair:  Jamil Kawa, Synopsys, US

In this executive session, we will discuss the prominent features and requirements of today’s autonomously powered systems and deliberate over various visions of what needs to happen next to take autonomously powered systems from their embryonic state to an advanced efficient state that is well thought through and efficiently architectured.

Moderator:
- Jamil Kawa, Synopsys, US

Panelists:
- Mario Konijnenburg, IMEC, BE
- Christoph Heer, Intel, DE
- Yankin Tanurhan, Synopsys, US
- Ali Keshavarzi, Cypress Semiconductor, US

12:30  End of session
Lunch Break in Garden Foyer
Keynote Lecture session 7.0 in “Garden Foyer” 1350 - 1420

Lunch Break in the Garden Foyer
On all conference days (Tuesday to Thursday), a buffet lunch will be offered in the Garden Foyer, in front of the session rooms. Kindly note that this is restricted to conference delegates possessing a lunch voucher only. When entering the lunch break area, delegates will be asked to present the corresponding lunch voucher of the day. Once the lunch area is being left, re-entrance is not allowed for the respective lunch.

6.3 Security Primitives
Date: Wednesday 29 March 2017
Time: 11:00 - 12:30
Location / Room: 2BC
Chair:  Berndt Gammel, Infineon Technologies, DE
Co-Chair:  Tim Güneysu, University of Bremen & DFKI, DE

This session discusses the implementation of basic primitives that are necessary building blocks for the secure systems: Physical unclonable functions (PUFs) are used for creating secret values which then are used as keys in cryptographic algorithms. Logical and physical security of these systems fundamentally relies on the presence of high quality random numbers.
11:00  6.3.1 SENSITIZED PATH PUF: A LIGHTWEIGHT EMBEDDED PHYSICAL UNCLONABLE FUNCTION
Speaker: Matthias Sauer, University of Freiburg, DE
Authors: Matthias Sauer1, Pascal Rialoli1, Linus Feiten1, Bernd Becker1, Ulrich Rührmair2 and Ilia Polian3
1University of Freiburg, DE; 2TU München, DE; 3University of Passau, DE

Abstract
Physical unclonable functions (PUFs) can be used for a number of security applications, including secure on-chip generation of secret keys. We introduce an embedded PUF concept called sensitized path PUF (SP-PUF) that is based on extracting entropy out of inherent timing variability of modules already present in the circuit. The new PUF sensitizes paths of nearly identical lengths and generates response bits by racing transitions through different paths against each other. SP-PUF has lower area overhead and higher speed than earlier embedded PUFs and requires no helper data stored in non-volatile memory beyond standard error-correction information for fuzzy extraction. Compared with standalone PUFs, the new design intrinsically and inseparably intertwines PUF behavior with functional circuitry, thus complicating invasive attacks or simplifying their detection. Moreover, SP-PUF can naturally define the contribution of a digital block to a system-wide “fusion PUF”. We present a systematic design flow to turn an arbitrary (sufficiently complex) circuit into an SP-PUF. The flow leverages state-of-the-art sensitization algorithms, formal filtering based on statistical analysis, and MAXSAT-based optimization of SP-PUF’s area overhead. Experiments show that SP-PUF extracts 256-bit keys with perfect reliability and nearly perfect uniqueness after fuzzy extraction for the majority of standard benchmarks circuits.

Download Paper (PDF; Only available from the DATE venue WiFi)

11:30  6.3.2 TEMPERATURE AWARE PHASE/FREQUENCY DETECTOR-BASED RO-PUFS EXPLOITING BULK-CONTROLLED OSCILLATORS
Speaker: Sha Tao, Royal Institute of Technology (KTH), SE
Authors: Sha Tao and Elena Dubrova, Royal Institute of Technology (KTH), SE

Abstract
Physical unclonable functions (PUFs) are promising hardware security primitives suitable for low-cost cryptographic applications. Ring oscillator (RO) PUF is a well-received silicon PUF solution due to its ease of implementation and entropy evaluation. However, the responses of RO-PUFs are susceptible to environmental changes, in particular, to temperature variations. Additionally, a conventional RO-PUF implementation is usually more power-hungry than other PUF alternatives. This paper explores circuit-level techniques to design low-power RO-PUFs with enhanced thermal stability. We introduce a power-efficient approach based on a phase/frequency detector (PFD) to perform pairwise comparisons of ROs. We also propose a temperature compensated bulk-controlled oscillator (BCO) and investigate its feasibility and usage in PFD-based RO-PUFs. Evaluation results demonstrate that the proposed techniques can effectively reduce the thermally induced errors in PUF responses while imposing a low power overhead. The PFD-based BCO-PUF is one of the best among existing RO-PUFs in terms of power efficiency.

Download Paper (PDF; Only available from the DATE venue WiFi)

12:00  6.3.3 CHACHA20-POLY1305 AUTHENTICATED ENCRYPTION FOR HIGH-SPEED EMBEDDED IOT APPLICATIONS
Speaker: Fabrizio De Santis, Technische Universität München, DE
Authors: Fabrizio De Santis, Andreas Schauer and Georg Sigl, Technische Universität München, DE

Abstract
The ChaCha20 stream cipher and the Poly1305 authenticator are cryptographic algorithms designed by Daniel J. Bernstein with the aim of ensuring high-security margins, while achieving high performance on a broad range of software platforms. % In response to the concerns raised about the reliability of the existing IETF/TLS cipher suite, its performance on software platforms, and the ease to realize secure implementations thereof, the IETF has recently published the RFC7905 and RFC7539 to promote the use and standardization of the ChaCha20 stream cipher and Poly1305 authenticator in the TLS protocol. % Most interestingly, the RFC7539 specifies how to combine together the ChaCha20 stream cipher and Poly1305 authenticator to construct a ChaCha20/Poly1305 authenticated encryption scheme with Channel Authentication (AEAD) scheme to provide confidentiality, integrity, and authenticity of data. % In this work, we present compact, constant-time, and fast implementations of the ChaCha20 stream cipher, Poly1305-Chacha20 authenticator, and ChaCha20-Poly1305 AEAD scheme for ARM Cortex-M4 processors, aimed at evaluating the suitability of such algorithms for high-speed and lightweight IoT applications, e.g. to deploy fast and secure TLS connections between IoT nodes and remote cloud servers, when AES hardware acceleration capabilities are not available.

Download Paper (PDF; Only available from the DATE venue WiFi)

12:15  6.3.4 TOWARDS POST-QUANTUM SECURITY FOR IOT ENDPOINTS WITH NTRU
Speaker: Johanna Sepulveda, TU Munich, DE
Authors: Oscar M. Guillen1, Thomas Poppelmann2, Jose M. Bermudo Mera1, Elena Fuentes Bongenaar3, Georg Sigl1 and Johanna Sepulveda1
1TU München, DE; 2Infineon Technologies, DE; 3Radboud University, NL

Abstract
The NTRU cryptosystem is one of the main alternatives for practical implementations of post-quantum, public-key cryptography. In this work, we analyze the feasibility of employing the NTRU encryption scheme, NTRUencrypt, in resource constrained devices such as those used for Internet-of-Things endpoints. We present an analysis of NTRUencrypt’s advantages over other cryptosystems for use in such devices. We describe four different NTRUencrypt implementations on an ARM Cortex M0-based microcontroller, compare their results, and show that NTRUencrypt is suitable for use in battery-operated devices. We present performance and memory footprint figures for different security parameters, as well as energy consumption in a resource constrained microcontroller to backup these claims. Furthermore, to the best of our knowledge, in this work we present the first time-independent implementation of NTRUencrypt.

Download Paper (PDF; Only available from the DATE venue WiFi)

12:30  6.3.5 LEVERAGING AGING EFFECT TO IMPROVE SRAM-BASED TRUE RANDOM NUMBER GENERATORS
Speaker: Mohammed Saber Golanbari, Karlsruhe Institute of Technology (KIT), DE
Authors: Saman Kiamehr1, Mohammad Saber Golanbari2 and Mehdi Tahoori 2
1Karlsruhe Institute of Technology (KIT), DE; 2Karlsruhe Institute of Technology, DE

Abstract
The start-up value of SRAM cells can be used as the random number vector or a seed for the generation of a pseudo random number. However, the randomness of the generated number is pretty low since many of the cells have a skewed start-up and their response biased toward zero or one. In this paper, we propose an approach to increase the randomness of SRAM-based True Random Number Generators (TRNGs) by leveraging transistor aging impact. The idea is to iteratively power-up the SRAM cells and put them under accelerated aging to make the cells less skewed and hence obtaining a more random vector. The simulation results show that the min-entropy of SRAM-based TRNG increases by 10X using this approach.

Download Paper (PDF; Only available from the DATE venue WiFi)
Time | Label | Presentation Title | Authors
--- | --- | --- | ---
12:31 | IP3-2, 718 | DESIGN AUTOMATION FOR OBfuscated CIRCUITS WITH MULTIPLE VIABLE FUNCTIONS | Shahrzad Keshavarz, University of Massachusetts Amherst, US; Christof Paar, EPFL, CH; Daniel Holcomb, University of Massachusetts Amherst, US; Horst Görtz Institut für IT-Sicherheit, Ruhr-Universität Bochum, DE

**Abstract**
Gate camouflaging is a technique for obfuscating the function of a circuit against reverse engineering attacks. However, if an adversary has pre-existing knowledge about the set of functions that are viable for an application, random camouflaging of gates will not obfuscate the function well. In this case, the adversary can target their search, and only needs to decide whether each of the viable functions could be implemented by the circuit. In this work, we propose a method for using camouflaged cells to obfuscate a design that has a known set of viable functions. The circuit produced by this method ensures that an adversary will not be able to rule out any viable functions unless she is able to uncover the gate functions of the camouflaged cells. Our method comprises iterated synthesis within an overall optimization loop to combine the viable functions, followed by technology mapping to deploy camouflaged cells while maintaining the plausibility of all viable functions. We evaluate our technique on cryptographic S-box functions and show that, relative to a baseline approach, it achieves up to 38% area reduction in PRESENT-style S-Boxes and 48% in DES S-boxes.

**Download Paper (PDF; Only available from the DATE venue WiFi)**

12:30 |  | Lunch Break in Garden Foyer |  
**Keynote Lecture** session 7.0 in "Garden Foyer" 1350 - 1420

**Lunch Break in the Garden Foyer**
On all conference days (Tuesday to Thursday), a buffet lunch will be offered in the Garden Foyer, in front of the session rooms. Kindly note that this is restricted to conference delegates possessing a lunch voucher only. When entering the lunch break area, delegates will be asked to present the corresponding lunch voucher of the day. Once the lunch area is being left, re-entrance is not allowed for the respective lunch.

### 6.4 High-performance Reconfigurable Computing

**Date:** Wednesday 29 March 2017  
**Time:** 11:00 - 12:30  
**Location / Room:** 3A

**Chair:** Philip Brisk, University of California, Riverside, US

**Co-Chair:** Mirjana Stojilovic, EPFL, CH

Reconfigurable architectures are seeing increased usage in high performance and scientific applications. This session addresses challenges in this space, which include optimizes arithmetic data paths, developing an in-memory architecture for optimizing database processing, and a case study on developing a high throughput Smith-Waterman accelerator.

**Time | Label | Presentation Title | Authors**
--- | --- | --- | ---
11:00 | 6.4.1 | AUTOMATING THE PIPELINE OF ARITHMETIC DATAPATHS | Florent de Dinechin, INSA-Lyon, FR; Mateo Is Ioan-1 and Florent de Dinechin-2  

**Abstract**
This article presents the new framework for semi-automatic circuit pipelining that will be used in future releases of the FloPoCo generator. From a single description of an operator or datapath, optimized implementations are obtained automatically for a wide range of FPGA targets and a wide range of frequency/latency trade-offs. Compared to previous versions of FloPoCo, the level of abstraction has been raised, enabling easier development, shorter generator code, and better pipeline optimization. The proposed approach is also more flexible than fully automatic pipelining approaches based on retiming: In the proposed technique, the incremental construction of the pipeline along with the circuit graph enables architectural design decisions that depend on the pipeline.

**Download Paper (PDF; Only available from the DATE venue WiFi)**

11:30 | 6.4.2 | OPERAND SIZE RECONFIGURATION FOR BIG DATA PROCESSING IN MEMORY | Paulo Cesar Santos, \(^1\) Geraldo Francisco de Oliveira Junior, \(^2\) Diego Gomes Tomé, \(^3\) Marco Antonio Zanata Alves, \(^3\) Eduarido Cunha de Almeida, \(^3\) and Luigi Carro, \(^4\)  

**Abstract**
Nowadays, applications that predominantly perform lookups over large databases are becoming more popular with column-stores as the database system architecture of choice. For these applications, Hybrid Memory Cubes (HMCs) can provide bandwidth of up to 320 GB/s and represents the best choice to keep the throughput for these ever increasing databases. However, even with the high available memory bandwidth and processing power, in order to achieve the peak performance, data movements through the memory hierarchy consumes an unnecessary amount of time and energy. In order to accelerate database operations, and reduce the energy consumption of the system, this paper presents the Reconfigurable Vector Unit (RVU) that enables massive and adaptive in-memory processing, extending the native HMC instructions and also increasing its effectiveness. RVU enables the programmer to reconfigure it to perform as a large vector unit or multiple small vector units to better adjust for the application needs during different computation phases. Due to its adaptability, RVU is capable of achieving performance increase of 27x on average and reduce the DRAM energy consumption in 29% when compared to an x86 processor with 16 cores. Compared with the state-of-the-art mechanism capable of performing large vector operations with fixed size, inside the HMC, RVU performed up to 12% better in terms of performance and improve in 53% the energy consumption.

**Download Paper (PDF; Only available from the DATE venue WiFi)**
6.5 Hot Topic Session: Memristor for Computing: Myth or Reality?

Date: Wednesday 29 March 2017
Time: 11:00 - 12:30
Location / Room: 3C

Organisers:
Koen Bertels, Delft University of Technology, NL
Said Hamdioui, Delft University of Technology, NL

Chair:
Akash Kumar, Technische Universität Dresden, DE

Co-Chair:
Koen Bertels, Delft University of Technology, NL

Both today’s technology and computer architectures are facing serious challenges/walls making them incapable to deliver the right computing power at pre-defined constraints for emerging applications such as big-data. However, a solution may be at your fingertips. This session discusses the emerging memristor device in enabling new memory technologies and new logic design styles, as well as its potential in enabling new computing paradigms such as memory intensive architectures and neuromorphic computing, due to its unique properties like the tight integration with CMOS and the ability to learn and adapt.

Time | Label | Presentation Title | Authors
--- | --- | --- | ---
11:00 | 6.5.1 | MEMRISTOR: WHAT IS IT ABOUT AND WHAT IS ITS POTENTIAL? | Said Hamdioui, Delft University of Technology, NL
11:30 | 6.5.2 | MEMRISTOR FOR MEMORY-INTENSIVE ARCHITECTURES | Shahar Kvatinsky, Technion/Israel Institute of Technology, IL
6.6 Industrial Experiences & EU Projects

Date: Wednesday 29 March 2017
Location / Room: SA

Chair:
Eugenio Villar, University of Cantabria, ES

This session addresses industrial research and practice on architecture, design, timing analysis techniques and analogue circuit sizing. The session will be rounded off by presentations of two European projects about to start, addressing cross-layer design of reconfigurable CPS and IoT for smart wearable applications.

Time Label Presentation Title Authors

12:00 6.6.3 MEMRISTOR FOR NEUROMORPHIC COMPUTING
Author: Gert Cauwenberghs, UC San Diego, US

12:30 End of session
Lunch Break in Garden Foyer
Keynote Lecture session 7.0 in "Garden Foyer" 1350 - 1420

Lunch Break in the Garden Foyer
On all conference days (Tuesday to Thursday), a buffet lunch will be offered in the Garden Foyer, in front of the session rooms. Kindly note that this is restricted to conference delegates possessing a lunch voucher only. When entering the lunch break area, delegates will be asked to present the corresponding lunch voucher of the day. Once the lunch area is being left, re-entry is not allowed for the respective lunch.

11:00 6.6.1 AN ASYNCHRONOUS NOC ROUTER IN A 14NM FINFET LIBRARY: COMPARISON TO AN INDUSTRIAL SYNCHRONOUS COUNTERPART
Speaker: Wayne Burleson, Advanced Micro Devices, Inc., US
Authors: Weiwei Jiang1, Davide Bertozzi1, Gabriele Miorandi2, Steven M. Nowick1, Wayne Burleson2 and Greg Sadowski2
1Columbia University, US; 2University of Ferrara, IT; 3Advanced Micro Devices, US

Abstract
An asynchronous high-performance low-power 5-port network-on-chip (NoC) router is introduced. The proposed router integrates low-latency input buffers using a circular FIFO design, and a novel end-to-end credit-based virtual channel (VC) flow control for a replicated switch architecture. This asynchronous router is then compared to an AMD synchronous router, in a realistic advanced 14nm FinFET library. This is the first such comparison, to the best of our knowledge, using a real synchronous router baseline already fabricated in several commercial products. Initial post-synthesis pre-layout experiments show dominating results for the asynchronous router, when compared to the synchronous router. In particular, 55% less area and 28% latency improvement are observed for the asynchronous implementation. Also, 88% and 58% savings in idle and active power, respectively, are obtained.

Download Paper (PDF; Only available from the DATE venue WiFi)

11:15 6.6.2 AN ADVANCED EMBEDDED ARCHITECTURE FOR CONNECTED COMPONENT ANALYSIS IN INDUSTRIAL APPLICATIONS
Speaker: Menbere Tekleyohannes, University of Kaiserslautern, DE
Authors: Menbere Tekleyohannes1, Mohammadsadegh Sadri1, Martin Klein2, Michael Siegrist2, Christian Weis1 and Norbert Wehn1
1University of Kaiserslautern, DE; 2Wipecot GmbH, DE

Abstract
In recent years, connected component analysis (CCA) has become one of the vital image/video processing algorithms due to its wide-range applicability in the field of computer vision. Numerous applications such as pattern recognition, object detection and image segmentation involve connected component analysis. In the context of camera-based inspection systems, CCA plays an important role for quality assurance. State-of-the-art hardware architectures offer high performance implementations of CCA using field programmable gate arrays (FPGAs). However, due to their high memory-demand, most of these implementations inhibit a large resource utilization. In this paper, we propose a hybrid software-hardware architecture of CCA for an industrial application using Xilinx Zynq-7000 All Programmable System on Chip (SoC). By offloading the most resource consuming part of the algorithm to the embedded CPU, we achieved high performance, while reducing the required resources on the FPGA. Our proposed architecture saves more than 30% of on-chip memory (Block RAMs) compared to state-of-the-art hardware architectures without affecting the throughput. Furthermore, due to the embedded CPU, our system provides a versatile and highly flexible feature extraction at run-time without the necessity to reconfigure the FPGA.

Download Paper (PDF; Only available from the DATE venue WiFi)

11:30 6.6.3 WORKLOAD DEPENDENT RELIABILITY TIMING ANALYSIS FLOW
Speaker: Ajith Sivadasan, TIMA Labs, FR
Authors: Ajith Sivadasan1, Armelle Notin2, Vincent Huard3, Etienne Maurin2, Sidi Ahmed Benhassain2 and Lorena Anghel4
1TIMA Labs, FR; 2STMicroelectronics, FR; 3TIMA, FR; 4Grenoble-Alpes University, FR

Abstract
Silicon measurements indicate the fact that the frequency limiting paths change as per aging and as a function of workload. This paper proposes a simulation flow that leads to the identification of such paths. Gate-level models provide an accurate estimate of aging of the critical paths by taking into consideration the stress experienced by corresponding standard cells for a given workload on the digital circuit and thereby providing a more accurate estimate of circuit aging.

Download Paper (PDF; Only available from the DATE venue WiFi)
### 6.7 Model-Based Design and Verification of Real-Time Systems

**Time:** Wednesday 29 March 2017 11:00 - 12:30

**Location / Room:** 3B

**Chair:**
Alain Girault, INRIA, FR

**Co-Chair:**
Amir Aminifar, IPFL Lausanne, CH

This session provides an overview of recent advances in model based design of embedded real-time systems. The first paper proposes an optimal deployment for data-flow applications on many-core chips. The second paper addresses the issue of simulation-based verification of embedded systems. It considers aspects of model based design of control systems in the context of event based real-time simulation. Last, but not least, the third paper discusses the workload monitoring of real-time systems by relying on a run-time feedback instead of offline assumptions.
6.8 HIPEAC: European Network on High Performance and Embedded Architecture and Compilation

Date: Wednesday 29 March 2017
Time: 11:00 - 12:30
Location / Room: Exhibition Theatre
Organiser: Catherine Roderick, Barcelona Supercomputing Center, ES
Moderator: Luca Fanucci, University of Pisa, IT

This session will showcase the activities of this network of research expertise. HIPEAC members come from both industry and academia and, together, form a community of expertise in Europe which reinforces and strengthens R&D activities. We offer funding for industrial PhD internships and short-term collaborations between early-career researchers and other research centres, as well as annual Tech Transfer Awards and communications and recruitment services. Annual HIPEAC activities include a high-profile conference, a

**Time** | **Label** | **Presentation Title** | **Authors**
---|---|---|---
11:00 | 6.7.1 | NEAR-OPTIMAL DEPLOYMENT OF DATAFLOW APPLICATIONS ON MANY-CORE PLATFORMS WITH REAL-TIME GUARANTEES | Stefanos Skalitsis, École Polytechnique Fédérale de Lausanne (EPFL), GR
**Speaker:** Stefanos Skalitsis and Alena Simalatar, EPFL, CH
**Abstract**
Safe and optimal deployment of data-streaming applications on many-core platforms requires the realistic estimation of task Worst-Case Execution Time (WCET). On the other hand, task WCET depends on the deployment solution, due to the varying number of interferences on shared resources, thus introducing a cyclic dependency. Moreover, WCET is still an over-approximation of the Actual Execution Time (AET), thus leaving room for run-time optimisation. In this paper we introduce an offline/online optimisation approach. In the offline phase, we first break the cyclic dependency and acquire safe and near-optimal solutions for tasks partitioning/placement, mapping, scheduling and buffer allocation. Then, we tighten the WCETs and update the scheduling function accordingly. In the online phase we introduce a safe distributed readjustment of the offline schedule, based on the AET. Experiments on a Kalray MPPA-256 platform show a tightening of the guaranteed latency up to 46% in the offline phase, and 41% latency reduction in the online phase. In total, we achieve more than 50% of latency reduction.
**Download Paper (PDF; Only available from the DATE venue WiFi)**

11:30 | 6.7.2 | SIMULATING PREEMPTIVE SCHEDULING WITH TIMING-AWARE BLOCKS IN SIMULINK | Andreas Naderlinger, University of Salzburg, AT
**Speaker and Author:**
**Abstract**
This paper introduces an extension of the modeling and simulation environment MATLAB/Simulink. It enables control and system engineers to consider software execution times, as well as the effects of scheduling and preemption inside software-in-the-loop (SIL) simulations. To this end, we present the concept of a Simulink block whose execution lasts for a finite amount of simulation time. During this time, the simulation engine continues to update the plant or other blocks with outputs that have already been calculated by the block. Execution time information is assumed to be known (or based on some random distribution). Source-level annotating the control software with target specific timing information enables a fine-grained and even a control-flow dependent simulation of the block. We outline the required synchronization with the simulation engine of Simulink. This timing-aware block consumes simulation time in the same sense as a task consumes CPU time on a target. We describe a mechanism to execute a set of such blocks with (potentially cyclic) data dependencies with a static priority scheduler inside Simulink, including support for preemption. The presented approach permits a development process, where a typical time invariant and platform agnostic model is incrementally transformed into a platform-specific one that makes the simulation more realistic.
**Download Paper (PDF; Only available from the DATE venue WiFi)**

12:00 | 6.7.3 | ONLINE WORKLOAD MONITORING WITH THE FEEDBACK OF ACTUAL EXECUTION TIME FOR REAL-TIME SYSTEMS | Biao Hu, Tech. Univ. Muenchen TUM, DE
**Speaker:** Biao Hu, Kai Huang, Gang Chen, Long Cheng and Alois Knoll
**Authors:**
1Tech. Univ. Muenchen TUM, DE; 2Sun Yat-Sen University, CN
**Abstract**
Guaranteeing the system workload within design bounds is a basic requirement for a real-time system. Design-time bounds are usually based on worst-case activation patterns and worst-case execution time. While using the worst-case assumptions for online monitoring can guarantee the system safety, it also introduces unexplored slacks due to tasks consuming less than their worst-case execution times. In this paper, we introduce a monitoring scheme with the feedback of actual execution time for real-time systems. By using this runtime feedback instead of offline assumptions, this monitoring scheme can accept events that are considered as violations offline, and thereby improve the system utilization. In the experiments of both MATLAB simulation and MicroCos-II running in a software processor implemented on an FPGA, different probability distributions of actual execution time are used in analyzing how much the benefit can be gained from the feedback scheme.
**Download Paper (PDF; Only available from the DATE venue WiFi)**

12:30 | IP3-6, 607 | LATENCY ANALYSIS OF HOMOGENEOUS SYNCHRONOUS DATAFLOW GRAPHS USING TIMED AUTOMATA | Guus Kuiper, University of Twente, NL
**Speaker:** Guus Kuiper and Marco Bekooij
**Authors:**
1University of Twente, NL; 2University of Twente + NXP semiconductors, NL
**Abstract**
There are several analysis models and corresponding temporal analysis techniques for checking whether applications executed on multiprocessor systems meet their real-time constraints. However, currently there does not exist an exact end-to-end latency analysis technique for Homogeneous Synchronous Dataflow (HSDF) with Auto-concurrency (HSDFa) models that takes the correlation between the firing durations of different firings into account. In this paper we present a transformation of strongly connected (HSDFa) models into timed automata models. This enables an exact end-to-end latency analysis because the correlation between the firing durations of different firings is taken into account. In a case study we compare the latency obtained using timed automata and a Linear Program (LP) based analysis technique that relies on a deterministic abstraction and compare their run-times as well. Exact end-to-end latency analysis results are obtained using timed automata, whereas this is not possible using deterministic timed-dataflow models.
**Download Paper (PDF; Only available from the DATE venue WiFi)**

End of session

Lunch Break in Garden Foyer
Keynote Lecture session 7.0 in *Garden Foyer* 1350 - 1420

Lunch Break in the Garden Foyer
On all conference days (Tuesday to Thursday), a buffet lunch will be offered in the Garden Foyer, in front of the session rooms. Kindly note that this is restricted to conference delegates possessing a lunch voucher only. When entering the lunch break area, delegates will be asked to present the corresponding lunch voucher of the day. Once the lunch area is being left, re-entrance is not allowed for the respective lunch.
research and development activities. In this session, after a brief introduction to HiPEAC, we highlight some of our members' innovative and groundbreaking research and development activities.

**HiPEAC Vision**, which recommends future actions and priorities for the European computing systems community and is a key source of reference for the European Commission. In addition, an heterogeneous platform for HPC architecture exploration will be deployed.

The **MANGO** project builds on these considerations and will set inherent architecture-level support for application-based customization as one of its underlying pillars. In addition, an heterogeneous platform for HPC architecture exploration will be deployed.
**GREENOPENHEVC: LOW POWER HEVC DECODER**

**Presenter:**
Menard Daniel, INSA Rennes, FR

**Authors:**
Julien Heulot1, Erwan Nogues1, Maxime Pelcat2 and Wassim Hamdouch1

1INS A Rennes, IETR, UBL, FR; 2Institut Pascal, Université Clermont-Ferrand, FR

**Abstract**

Video on mobile devices is a must-have feature with the prominence of new services and applications using video like streaming or conferencing. The new video standard HEVC is an appealing technology for service providers. Besides, with the recent progress of SoC, software video decoders are now a reality. The challenge is to provide power efficient design to fit with the compelling demand for long battery. We present here a practical set-up demonstrating that the new HEVC standard can be implemented in software on an embedded GPP multicore platform. Different techniques have been integrated to optimize the energy: data-level and thread level parallelisms, video aware Dynamic Voltage and Frequency Scaling. To push back the limits, algorithm level approximate computing is carried-out on the in-loop filtering. The subjective tests have demonstrated that the quality degradation is almost imperceptible. A mean power of less than 1 Watt is reported for a HD 1080p/24fps video decoding.

**More information ...**
**UB06.3 TTOOL5G: MODEL-BASED DESIGN OF A 5G UPLINK DATA-LINK LAYER RECEIVER FROM UML/SYSML DIAGRAMS**

**Presenter:** Andrea Emirici, Nokia Bell Labs France, FR  
**Authors:** Julien Lalitet, Imran Latif, Ludovic Aprvilli, Renaud Pacalet and Adrien Canuel  
**Abstract**  
Future 5G networks are expected to provide an increase of 10x in data rates. To meet these requirements, the equipment of baseband stations will be designed using mixed architectures, i.e., DSPs, FPGAs. However, efficiently programming these architectures is not trivial due to the drastic increase in complexity of their design space. To overcome this issue, we need to use detailed tools capable of rapidly exploring, partitioning and prototyping the mixed architecture designs of 5G systems. At DATE 2017 University Booth, we demonstrate such a unified tool and show our latest achievements in the automatic code generation engine of TTool/DIPLODOCUS, a UML/SysML framework for the hardware/software co-design of data-flow systems, to support mixed architectures. Our demonstration will show the full design and evaluation of a 5G data-link layer receiver for both a DSP-based and an IP-based designs. We will validate the effectiveness of our solution by comparing automated vs manual designs.

*More information...*

---

**UB06.4 WE DARE: WEARABLE ELECTRONICS DIRECTIONAL AUGMENTED REALITY**

**Presenter:** Davide Quaglia, University of Verona, IT  
**Authors:** Gianluca Benedetti and Walter Vendraminetto  
**Abstract**  
Current augmented reality (AR) eyewear solutions require large form factors, weight, cost and energy that reduce usability. In fact, connectivity, image processing, localization, and direction evaluation lead to high processing and power requirements. A multi-antenna system, patented by the industrial partner, enables a new generation of smart eye-wear that elegantly requires less hardware, connectivity, and power to provide AR functionalities. They will allow users to directionally locate nearby radio emitting sources that highlight objects of interest (e.g., people or retail items) by using existing standards like Bluetooth Low Energy, Apple’s iBeacon and Google’s Eddystone. This booth will report the current level of research addressed by the Computer Science Department of University of Verona and the company Wagoo LLC. In the presented demo, different objects emit an “I am here” signal and a prototype of the smart glasses shows the information related to the observed object.

*More information...*

---

**UB06.5 ITMD: RUN-TIME MANAGEMENT OF CONCURRENT MULTITHREADED APPLICATIONS ON HETEROGENEOUS MULTI-CORES**

**Presenter:** Karunakar Reddy Basireddy, University of Southampton, GB  
**Authors:** Amit Singh, Bashir M. Al-Hashimi and Geoff V. Merrett, University of Southampton, GB  
**Abstract**  
Heterogeneous multi-cores often need to deal with multiple applications having different performance requirements concurrently, which generate varying and mixed workloads. Runtime management is required for adapting to such performance requirements and workload variabilities, and to achieve energy efficiency. It is challenging to efficiently exploit different types of cores simultaneously and DVFS potential of cores. We present a run-time management approach that first selects thread-to-core mapping based on the performance requirements and resource availability. Then, it applies online adaptation by adjusting the voltage-frequency (V-f) levels to achieve energy optimization. We demonstrate the proposed run-time management approach on the Odroid-XU3, with various combinations of multi-threaded applications from PARSEC and SPLASH benchmarks. Results show an average improvement in energy efficiency up to 33% compared to existing approaches.

*More information...*

---

**UB06.6 BRAIN TO COMPUTER CONNECTIONS: A FAST TIME-DOMAIN APPROACH FOR BCI TRAINING**

**Presenter:** Vito Leonardo Gallo, Politecnico di Bari, Italy, IT  
**Authors:** Valerio Francesco Annese and Daniela De Venuto, Politecnico di Bari, IT  
**Abstract**  
We present a P300-based Brain Computer Interface (BCI) approach for the brain control of external devices through an innovative approach. The herein proposed HW/SW system acquires the signal from 6 EEG channels and synchronizes them with ad-hoc designed visual stimuli that evokes the P300 signal. The BCI signal processing comprises: (i) a Machine Learning stage, which is based on an algorithm (t-RIDE), which calibrates the system in ~190s (ii) a smart approach for the time-domain features extraction greatly reduces the computational effort, speeding up the classification and finally (iii) the on-line classification, which is entrusted to a linear classifier. Noteworthy results obtained in experimental setup are: (i) P300 spatio-temporal characterization in 1.95s, (ii) classification accuracy of 80.5±4.1% on single-trial, (iii) real time classification in 22ms (WC). As a PoC, supporting videos will show how the BCI outcomes can pilot a prototype car.

*More information...*

---

**UB06.7 PER: METHOD AND TOOL FOR ANALYZING THE INTERPLAY BETWEEN PERFORMANCE, ENERGY AND SCALING IN MULTI- AND MANY-CORE PLATFORMS**

**Presenter:** Fei Xia, Newcastle University, GB  
**Authors:** Ashur Rafiev, Alexander Romanovsky and Alex Yakovlev, Newcastle University, GB  
**Abstract**  
Parallelization has been used to maintain a reasonable balance between energy consumption and performance in computing systems. However, the effectiveness of parallelization scaling is different for different hardware platforms. This is because the reliable operation region (ROR), a region defined in the voltage-throughput space for any hardware platform, is platform-dependent and its shape determines how effective parallelization scaling is in improving throughput and/or reducing power consumption. Although many of the interlinked issues are known, a unifying analysis method has just now been proposed to study the interplay between performance, energy, reliability and parallelization scaling. The method of bi-normalization of the ROR is designed to help achieve a meaningful cross-platform analysis of this interplay. The PER tool brings all these issues together and helps designers reason about hardware parallelization, DVFS and software parallelizability.

*More information...*

---

**UB06.8 TIDES: NON-LINEAR WAVEFORMS FOR QUICK TRACE NAVIGATION**

**Presenter:** Janina Stoppe, University of Bremen, DE  
**Author:** Ralf Drechsler, University of Bremen / DFKI, DE  
**Abstract**  
System trace analysis is mostly done using waveform viewers -- tools that relate signals and their assignments at certain times. While generic hardware design is subject to some innovative visualisation ideas and software visualisation has been a research topic for much longer, these classic tools have been part of the design process since the earlier days of hardware design -- and have not changed much over the decades. Instead, the currently available programs have evolved to look practically the same, all following a familiar pattern that has not changed since their initial appearance. We argue that there is still room for innovation beyond the very classic waveform display though. We implemented a proof-of-concept waveform viewer (codenamed Tides) that has several unique features that go beyond the standard set of features for waveform viewers.

*More information...*
**SEFILE: A SECURE FILESYSTEM IN USERSPACE VIA SECUBE™**

**Presenter:** Giuseppe Airofarulla, CINI, IT  
**Authors:** Paolo Prinetto and Antonio Varriale  
1CINI & Politecnico di Torino, IT; 2BuS Labs Ltd., IT  

**Abstract**  
The SEcube™ Open Source platform is a combination of three main cores in a single-chip design. Low-power ARM Cortex-M4 processor, a flexible and fast Field-Programmable-Gate-Array (FPGA), and an EAL5+ certified Security Controller (SmartCard) are embedded in an extremely compact package. This makes it a unique Open Source security environment where each function can be optimized, executed, and verified on its proper hardware device. In this demo, we present a Windows wrapper for a Filesystem in Userspace (FUSE) with an HDD firewall resorting to the hardware built-in capabilities, and the software libraries, of the SEcube™.

**More information ...**

**LABSMILING: A FRAMEWORK, COMPOSED OF A REMOTELY ACCESSIBLE TESTBED AND RELATED SW TOOLS, FOR ANALYSIS AND DESIGN OF LOW DATA-RATE WIRELESS PERSONAL AREA NETWORKS BASED ON IEEE 802.15.4**

**Presenter:** Marco Santic, University of L’Aquila, IT  
**Authors:** Luigi Pomante, Walter Tiberti, Carlo Centofanti and Lorenzo Di Giuseppe, DEWS - Università di L’Aquila, IT  

**Abstract**  
Low data-rate wireless personal area networks (LR-WPANs) are even more present in the fields of IoT, wearable devices and health monitoring. The development, deployment and test of such systems, based on IEEE 802.15.4 standard (and its derivations, e.g. 15.4e), require the exploitation of a testbed when the network is not trivial and grows in complexity. This demo shows the framework of LabSmiling: a testbed and related SW tools that connect a meaningful (but still scalable) number of physical devices (sensor nodes) located in a real environment. It offers the following services: program, reset, switch on/off single devices; connect to devices up/down links to inject or receive commands/msgs/packets in/from the network; set devices as low level packet sniffers, allowing to test/debug protocol compliances or extensions. Advanced services are: possibility of design test scenarios for the evaluation of network metrics (throughput, latencies, etc.) and custom application verification.

**More information ...**

14:00 End of session

16:00 **Coffee Break** in Exhibition Area

On all conference days (Tuesday to Thursday), coffee and tea will be served during the coffee breaks at the below-mentioned times in the exhibition area.

- Tuesday, March 28, 2017  
  - Coffee Break 10:30 - 11:30  
  - Coffee Break 16:00 - 17:00

- Wednesday, March 29, 2017  
  - Coffee Break 10:00 - 11:00  
  - Coffee Break 16:00 - 17:00

- Thursday, March 30, 2017  
  - Coffee Break 10:00 - 11:00  
  - Coffee Break 15:30 - 16:00

**7.0 LUNCH TIME KEYNOTE SESSION**

**Date:** Wednesday 29 March 2017  
**Time:** 13:50 - 14:20  
**Location / Room:** Garden Foyer  
**Chair:** David Atienza, EPFL, CH

**Time** | **Label** | **Presentation Title** | **Authors**
---|---|---|---
13:50 | 7.0.1 | **INTERNET OF EVERYTHING IS OUR OPPORTUNITY** | Keith Willett, Director of Software Engineering for Merck Serono, CH

**Abstract**  
Merck Serono is working to revolutionize patient care and doctor assist through utilization of technology that is built of the Internet of Everything. Using global resources to consolidate medical devices under a single platform that will store, analyze and recommend patient care to physicians, Merck is leveraging the Internet of Everything to improve patient care. The IoE is not limited to medical devices, as everything from automobiles to light bulbs are looking for ways to connect to the Internet. These devices gather, store and analyze data to improve the user experience and create value for people and businesses that have yet to be recognized. However, connecting so many products will cause an increased strain on the network infrastructures, and most importantly expose personal information to potential threats; if not managed correctly. All companies connecting devices are having similar problems and are working to solve these issues. As the Internet of Everything continues to evolve, critical strategies will need to be in place for all companies to be successful. This presentation will discuss the strategies companies need to play in this space and how collaboration and cooperation will become more common in IoE.

14:20 End of session
Abstract

Werner Haselmayr, Andreas Springer and Robert Wille, Johannes Kepler University Linz, AT

Authors

Andreas Grimmer, Johannes Kepler University Linz, AT

Presenter

NETWORKED LABS-ON-CHIPS

More information ...

Abstract

Wednesday, March 29, 2017

Coffee Break 10:00 - 11:00

Coffee Break 16:00 - 17:00

Thursday, March 30, 2017

Coffee Break 10:00 - 11:00

Coffee Break 15:30 - 16:00

More information ...

R2MEDIQ: WHEELCHAIR MOUNTED ROBOTIC ARM DEMONSTRATOR FOR PEOPLE WITH MOTOR SKILLS IMPAIRMENTS

Presenter:

Alessandro Palla, University of Pisa, IT

Authors:

Gabriele Meoni and Luca Fanucci, University of Pisa, IT

Abstract

Object detection plays a crucial role in realizing intelligent indoor localization and mapping techniques. With the advantages of these techniques comes the complexity of computing hardware and the mobility. While the availability of open source computer vision algorithms and High-Level-Synthesis framework accelerates the development, the hybrid processing architecture of an All Programmable System on Chip (APSoC) enables efficient hardware-software partitioning. Using these tools, a generic platform was designed for evaluating the computer vision algorithms. Open source components such as Linux kernel and OpenCV libraries were integrated for evaluation of the algorithms on the software while Vivado HLS framework was used to synthesize the hardware counter parts. Algorithms such as Sobel filtering and Hough Line transformation were implemented and analyzed. The capabilities of this platform were used to realize a mobile object detection system for enhancing the localization techniques.

More information ...

FLEXPORT: FLEXIBLE PLATFORM FOR OBJECT RECOGNITION & TRACKING TO ENHANCE INDOOR LOCALIZATION AND MAPPING

Presenter:

Marko Rößler, Technische Universität Chemnitz, DE

Authors:

Christian Schott, Murali Padmanabha and Ulrich Heinkel, TU Chemnitz, DE

Abstract

Object detection plays a crucial role in realizing intelligent indoor localization and mapping techniques. With the advantages of these techniques comes the complexity of computing hardware and the mobility. While the availability of open source computer vision algorithms and High-Level-Synthesis framework accelerates the development, the hybrid processing architecture of an All Programmable System on Chip (APSoC) enables efficient hardware-software partitioning. Using these tools, a generic platform was designed for evaluating the computer vision algorithms. Open source components such as Linux kernel and OpenCV libraries were integrated for evaluation of the algorithms on the software while Vivado HLS framework was used to synthesize the hardware counter parts. Algorithms such as Sobel filtering and Hough Line transformation were implemented and analyzed. The capabilities of this platform were used to realize a mobile object detection system for enhancing the localization techniques.

More information ...

NETWORKED LABS-ON-CHIPS

Presenter:

Andreas Grimmer, Johannes Kepler University Linz, AT

Authors:

Werner Haselmayr, Andreas Springer and Robert Wille, Johannes Kepler University Linz, AT

Abstract

Labs-on-Chip (LoC) allow for the miniaturization, integration, and automation of medical and bio-chemical procedures. In recent years, different technologies have been considered. However, all of them have their drawbacks, e.g. electrowetting-based LoCs suffer from the evaporation of liquids, the fast degradation of the surface coatings, and the inferior biocompatibility, while flow-based LoCs require a complex and costly multilayer fabrication process. Hence, an alternative has recently been proposed in terms of Networked Labs-on-Chips. We present and demonstrate the NLoC technology where so-called droplets flow inside channels of micrometer-size. Networking functionalities enable the designer to dynamically select the operations to be conducted. These networking functionalities exploit hydrodynamic forces acting on droplets. Moreover, NLoC devices can be produced at low cost (e.g. using 3D printers). By this, drawbacks of established LoC-technologies are addressed.

More information ...
More information ...
IoT technologies have the potential to be a disruptive game changer for existing applications and services as well as an enabler for new businesses. This session provides viewpoints from industry as well as a startup company on the deployment and evolution of IoT-oriented services and products.

Andreas Herkersdorf, TU Muenchen, DE

Marilyn Wolf, Georgia Tech, US

Organisers:
Marilyn Wolf, Georgia Tech, US
Andreas Herkersdorf, TU Muenchen, DE

Chair:
Marilyn Wolf, Georgia Tech, US
Andreas Herkersdorf, TU Muenchen, DE

IoT technologies have the potential to be a disruptive game changer for existing applications and services as well as an enabler for new businesses. This session provides viewpoints from industry as well as a startup company on the deployment and evolution of IoT-oriented services and products.

Time Label Presentation Title Authors
14:30 IP7.1.1, 7107 A LOW-POWER IOT PROCESSOR INTEGRATING VOLTAGE-SCALABLE FULLY DIGITAL MEMORIES
Author:
Hidetoshi Ondotera, Kyoto University, JP

14:37 IP7.1.2, 7108 A SIMPLE, STATELESS, COST EFFECTIVE SYMMETRIC CRYPTOGRAPHY STRATEGY FOR ENERGY-HARVESTING IOT DEVICES
Author:
Jan Madsen, Technical University of Denmark, DK

14:44 IP7.1.3, 7109 RECONFIGURABLE MICROCONTROLLER FOR END NODES IN INTERNET OF THINGS
Author:
Wai-Chung Matthew Tang, Queen Mary University of London, GB

14:51 IP7.1.4, 7110 FURTHER SIMPLIFICATION OF APPROXIMATE ADDERS USING INPUT DATA RANGES IN IOT
Author:
Jeong-A Lee, Chosun University, KR

15:00 7.1.2 HOW ASIC DEVELOPMENT WILL CHANGE FOR FUTURE IOT MEMS SENSORS
Author:
Dirk Droste, Robert Bosch GmbH, DE

Abstract
The global ASIC community faces a strong trend towards new IoT applications - but, what is the concrete behind all fuzzy discussions for the ASIC design community? This talk will give an overview about the perspective of Bosch Sensortec: ASIC development to adapt to upcoming challenges in ASIC design for future IoT MEMS sensors with their broad span of new applications and features and their challenging requirements for low power, high performance and complex integration.

15:30 7.1.3 DISTRIBUTED WAYSIDE ARCHITECTURE - IOT FOR RAILWAY INFRASTRUCTURE
Speaker:
Olivier Kaiser, Siemens, CH

Abstract
Railway infrastructure is characterized by very long life cycles, e.g. 25 years or even more, and very harsh environmental conditions. The requirements for availability and safety are nonetheless very demanding to assure an efficient and safe operation. In addition, the fulfillment of these requirements has to be shown formally in so-called safety cases. These cases have to be confirmed by independent safety assessors and eventually government agencies. Under these circumstances, the adoption of new technologies in the railway industry can be a challenge. Over the last decades, the architecture of railway control systems has been more or less stable. The trackside equipment, i.e. points, signals, track vacancy detection etc., is connected via star-shaped cabling to an interlocking. This interlocking distributes the energy and assures the safety by controlling the trackside equipment accordingly. The star-shaped cabling limits the control range of every interlocking, thus there is a need for an interlocking in every station. Both, this cabling concept as well as the large number of interlocking installations lead to high costs. To bring the overall costs down, new concepts have to be implemented. The field elements have to be connected via bus systems, ideally based on the Internet Protocol. This reduces cabling and increases the distance over which the elements can be controlled. Thus, the number of cabinets and installations can be distinctly reduced. Furthermore, off-the-shelf communication equipment can be used to connect the field elements. In the long run, a centralized operation of the control equipment in data centers can be envisioned. However, installing an internet of things along the track, where all signals, points and level crossings are subscribers, is demanding for the following reasons. • The functional safety has to be provided in a way that it can be formally proven. • A very high availability is necessary to assure steady operation. If an element or the connection to an element breaks down, no or only reduced operation is possible. • Security problems could affect passenger safety; hence, the communication system has to fulfill highest standards. • Legacy interfaces (e.g. the four wire interface for point machines) have to be supported further. • The field elements have to be provided with power. If a data bus is introduced, an adequate power bus is needed too in order to achieve substantial cost savings. For several years, Siemens has been working on innovating the IoT in the railway infrastructure. We named the concept Distributed Wayside Architecture. First installations at DB in Germany and SBB in Switzerland showed that the challenges mentioned above can be overcome. Current work focuses on the power bus as well as on the scalability of the concepts to larger installations.
Non-volatile memories (NVMs) are playing an increasingly dominant role in the construction of energy-efficient systems thanks to reduced static power consumption. NVMs raise new opportunities and challenges in terms of enhancing computational efficiency and ensuring security, respectively. This session explores in-memory computing applied to emerging NVM technologies and goes on to investigate security and encryption strategies.

**7.2 In-memory Computing and Security for Non-volatile Memory Technologies**

**Date:** Wednesday 29 March 2017  
**Time:** 14:30 - 16:00  
**Location / Room:** 4BC  
**Chair:** Luca Amaru, Synopsys, US  
**Co-Chair:** Pierre-Emmanuel Gaillardon, University of Utah, US

Non-volatile memories (NVMs) are playing an increasingly dominant role in the construction of energy-efficient systems thanks to reduced static power consumption. NVMs raise new opportunities and challenges in terms of enhancing computational efficiency and ensuring security, respectively. This session explores in-memory computing applied to emerging NVM technologies and goes on to investigate security and encryption strategies.

**7.2.1 AUTOMATED SYNTHESIS OF COMPACT CROSSBARS FOR SNEAK-PATH BASED IN-MEMORY COMPUTING**

**Speaker:** Sumit Kumar Jha, University of Central Florida, US  
**Authors:** Dwaipayan Chakraborty and Sumit Kumar Jha, University of Central Florida, US  
**Abstract**  
The rise of data-intensive computational loads has exposed the processor-memory bottleneck in Von Neumann architectures and has intensified the need for in-memory computing. Existing literature on computing Boolean formula using sneak-paths in nanoscale memristor crossbars has only focussed on short Boolean formula, such as 1-bit addition. There are two open questions: (i) Can one synthesize sneak-path based crossbars for computing large Boolean formula? (ii) What is the size of a memristor crossbar that can compute a given Boolean formula using sneak paths? In this paper, we make progress on both these open problems. First, we show that the number of rows and columns required to compute a Boolean formula is at most linear in the size of the Reduced Ordered Binary Decision Diagram representing the Boolean function. Second, we demonstrate how Boolean Decision Diagrams can be used to synthesize nanoscale crossbars that can compute a given Boolean formula using naturally occurring sneak paths. In particular, we synthesize large logical circuits such as 128-bit adders for the first-time using sneak-path based crossbar computing.

**Download Paper (PDF; Only available from the DATE venue WiFi)**

**7.2.2 HYBRID SPIKING-BASED MULTI-LAYERED SELF-LEARNING NEUROMORPHIC SYSTEM BASED ON MEMRISTOR CROSSBAR ARRAYS**

**Speaker:** Yiran Chen, Professor, US  
**Authors:** Amir Hassan, Chaofei Yang, Chenchen Liu, Hai (Helen) Li and Yiran Chen, University of Pittsburgh, US  
**Abstract**  
Neuromorphic computing systems are under heavy investigation as a potential substitute for the traditional von Neumann systems in high-speed low-power applications. Recently, memristor crossbar arrays were utilized in realizing spiking-based neuromorphic system, where memristor conductance values correspond to synaptic weights. Most of these systems are composed of a single crossbar layer, in which system training is done off-chip, using computer-based simulations, then the trained weights are pre-programmed to the memristor crossbar array. However, multi-layered, on-chip trained systems become crucial for handling massive amount of data and to overcome the resistance shift that occurs to memristors overtime. In this work, we propose a spiking-based multi-layered neuromorphic computing system capable of online training. The system performance is evaluated using three different datasets showing improved results versus previous work. In addition, studying the system accuracy versus memristor resistance shift shows promising results.

**Download Paper (PDF; Only available from the DATE venue WiFi)**

**7.2.3 REVAMP : RERAM BASED VLIW ARCHITECTURE FOR IN-MEMORY COMPUTING**

**Speaker:** Anupam Chattopadhyay, School of Computer Science and Engineering, Nanyang Technological University, SG  
**Authors:** Debjyoti Bhattacharjee, Rajeswari Devadoss and Anupam Chattopadhyay, Nanyang Technological University, SG  
**Abstract**  
With diverse types of emerging devices offering simultaneous capability of storage and logic operations, researchers have proposed novel platforms that promise gains in energy-efficiency. Such platforms can be classified into two domains—application-specific and general-purpose. The application-specific in-memory computing platforms include machine learning accelerators, arithmetic units, and Content Addressable Memory (CAM)-based structures. On the other hand, the general-purpose computing platforms stem from the idea that several in-memory computing logic devices do support a universal set of Boolean logic operation and therefore, can be used for mapping arbitrary Boolean functions efficiently. In this direction, so far, researchers have concentrated on challenges in logic synthesis (e.g. depth optimization), and technology mapping (e.g. device count reduction). The important problem of efficient technology mapping of arbitrary logic network onto a crossbar array structure has been overlooked so far. In this paper, we propose, ReVAMP, a general-purpose computing platform based on Resistive RAM crossbar array, which exploits the parallelism in computing multiple logic operations in the same word. Further, we study the problem of instruction generation and scheduling for such a platform. We benchmark the performance of ReVAMP with respect to the state of the art architecture.

**Download Paper (PDF; Only available from the DATE venue WiFi)**
### 7.3 Optimizing performance, energy and predictability via hardware/software codesign

**Date:** Wednesday 29 March 2017  
**Time:** 14:30 - 16:00  
**Location / Room:** 2BC

**Chair:**  
Sasan Avesta, George Mason University, US  
**Co-Chair:**  
Stefano Di Carlo, Politecnico di Torino, IT

This session presents a variety of architectural solutions to improve performance/energy/predictability covering several hardware blocks: processor pipeline, caches, memory and on-chip I/O. The first paper proposes a hardware/software mechanism to classify accesses as private or shared. The second paper introduces a low-power asynchronous microprocessor design. The third paper proposes a coordinated approach to improve performance by partitioning multilevel caches. And the last paper proposes a hardware approach to increase the timing accuracy of I/O operations.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>14:30</td>
<td></td>
<td>TUNNEL FET BASED REFRESH-FREE-DRAM</td>
<td>Navneet Gupta, ISEP-Paris, FR</td>
</tr>
<tr>
<td>14:45</td>
<td></td>
<td>A WEAR-LEVELING-AWARE COUNTER MODE FOR DATA ENCRYPTION IN NON-VOLATILE MEMORIES</td>
<td>Fangting Huang, Huazhong University of Science and Technology, CN</td>
</tr>
<tr>
<td>15:30</td>
<td></td>
<td>COVERT: COUNTER OVERFLOW REDUCTION FOR EFFICIENT ENCRYPTION OF NON-VOLATILE MEMORIES</td>
<td>Navneet Gupta, ISEP-Paris, FR</td>
</tr>
<tr>
<td>16:00</td>
<td></td>
<td>TUNNEL FET BASED REFRESH-FREE-DRAM</td>
<td>Navneet Gupta, ISEP-Paris, FR</td>
</tr>
<tr>
<td>16:00</td>
<td></td>
<td>A WEAR-LEVELING-AWARE COUNTER MODE FOR DATA ENCRYPTION IN NON-VOLATILE MEMORIES</td>
<td>Fangting Huang, Huazhong University of Science and Technology, CN</td>
</tr>
<tr>
<td>16:00</td>
<td></td>
<td>COVERT: COUNTER OVERFLOW REDUCTION FOR EFFICIENT ENCRYPTION OF NON-VOLATILE MEMORIES</td>
<td>Navneet Gupta, ISEP-Paris, FR</td>
</tr>
</tbody>
</table>

**Coffee Break** in Exhibition Area  
On all conference days (Tuesday to Thursday), coffee and tea will be served during the coffee breaks at the below-mentioned times in the exhibition area.

Tuesday, March 28, 2017  
- Coffee Break 10:30 - 11:30  
- Coffee Break 16:00 - 17:00

Wednesday, March 29, 2017  
- Coffee Break 10:00 - 11:00  
- Coffee Break 16:00 - 17:00

Thursday, March 30, 2017  
- Coffee Break 10:00 - 11:00  
- Coffee Break 15:30 - 16:00
A HARDWARE IMPLEMENTATION OF THE MCAS SYNCHRONIZATION PRIMITIVE

Speaker: Smitru Sarangi, IIT Delhi, IN

Authors: Smitru Sarangi, IIT Delhi, IN

Abstract

Lock-based parallel programs are easy to write. However, they are inherently slow as the synchronization is blocking in nature. Non-blocking lock-free programs, which use atomic instructions such as compare-and-set (CAS), are significantly faster. However, lock-free programs are notoriously difficult to design and debug. This can be greatly eased if the primitives work on multiple memory locations instead of one. We propose MCAS, a hardware implementation of a multi-word compare-and-set primitive. Ease of programming aside, MCAS-based programs are 13.8X and 4X faster on an average than programs, which use atomic instructions such as compare-and-set (CAS), are significantly faster. However, lock-free programs are notoriously difficult to design and debug. This can be greatly eased if the primitives work on multiple memory locations instead of one. We propose MCAS, a hardware implementation of a multi-word compare-and-set primitive. Ease of programming aside, MCAS-based programs are 13.8X and 4X faster on an average than lock-based and traditional lock-free programs respectively. The area overhead, in a 32-core 400mm2 chip, is a mere 0.046%.

Download Paper (PDF; Only available from the DATE venue WiFi)
**BANDITS: DYNAMIC TIMING SPECULATION USING MULTI-ARMED BANDIT BASED OPTIMIZATION**

**Speaker:** Jeff Zhang, New York University, US

**Authors:** Jeff Zhang and Siddharth Garg, New York University, US

**Abstract**

Timing speculation has recently been proposed as a method for increasing performance beyond that achievable by conventional worst-case design techniques. Starting with the observation of fast temporal variations in timing error probabilities, we propose a run-time technique to dynamically determine the optimal degree of timing speculation (i.e., how aggressively the processor is over-clocked) based on a novel formulation of the dynamic timing speculation problem as a multi-armed bandit problem. By conducting detailed post-synthesis timing simulations on a 5-stage MIPS processor running a variety of workloads, the proposed adaptive mechanism improves processor's performance significantly comparing with a competing approach (about 8.3% improvement); on the other hand, it shows only about 2.8% performance loss on average, compared with the oracle results.

Download Paper (PDF; Only available from the DATE venue WiFi)

**DESIGN AND IMPLEMENTATION OF A FAIR CREDIT-BASED BANDWIDTH SHARING SCHEME FOR BUSES**

**Speaker:** Carles Hernandez, Barcelona Supercomputing Center (BSC), ES

**Authors:** Mladen Slijepcevic¹, Carles Hernandez², Jaume Abella³ and Francisco Cazorla⁴

¹Barcelona Supercomputing Center and Universitat Politecnica de Catalunya, ES; ²Barcelona Supercomputing Center, ES; ³Barcelona Supercomputing Center (BSC-CNS), ES; ⁴Barcelona Supercomputing Center and IIIA-CSIC, ES

**Abstract**

Fair arbitration in the access to hardware shared resources is fundamental to obtain low worst-case execution time (WCET) estimates in the context of critical real-time systems, for which performance guarantees are essential. Several hardware mechanisms exist for managing arbitration in those resources (buses, memory controllers, etc.). They typically attain fairness in terms of the number of slots each contender (e.g., core) gets granted access to the shared resource. However, those policies may lead to unfair bandwidth allocations for workloads with contenders issuing short requests and contenders issuing long requests. We propose a Credit-Based Arbitration (CBA) mechanism that achieves fairness in the cycles each core is granted access to the resource rather than in the number of granted slots. Furthermore, we implement CBA as part of a LEON3 4-core processor for the Space domain in an FPGA proving the feasibility and good performance characteristics of the design by comparing it against other arbitration schemes.

Download Paper (PDF; Only available from the DATE venue WiFi)

**AN ALGORITHM TO FIND OPTIMUM SUPPORT-REDUCING DECOMPOSITIONS FOR INDEX GENERATION FUNCTIONS.**

**Speaker:** Tsutomu Sasao, Meiji University, JP

**Authors:** Tsutomu Sasao, Meiji University, JP

**Abstract**

Index generation functions are useful for pattern matching, and routers in the internet, etc.. This paper presents an algorithm to find support-reducing decompositions for index generation functions. Let n be the number of the input variables, and let s be the number of bound variables. Then, the exhaustive search for finding an optimum support-reducing decomposition requires to check $\binom{n}{s}$ combinations. We found a special property of index generation functions that drastically reduces this search space. With this property, we developed a fast algorithm to find an exact optimum solution. For a given number of bound variables, it finds a decomposition with the fewest rails. Experimental results up to $n=60$ and $s=33$ are shown.

Download Paper (PDF; Only available from the DATE venue WiFi)
**Abstract**

Encoders play an important role in many areas such as memory addressing, data demultiplexing, or for interconnect solutions. However, design solutions for the automatic synthesis of corresponding circuits suffer from various drawbacks, e.g. they are often not scalable, do not exploit the full degree of freedom, or are applicable to realize certain codes only. All these problems are caused by the fact that existing design solutions have to explicitly guarantee a one-to-one mapping. In this work, we propose an alternative design approach which relies on dedicated description means for both, the specification of an encoder as well as its circuit. Based on that, synthesis can be conducted without the need to explicitly take care of guaranteeing one-to-one mappings. Experiments show that this indeed overcomes the drawbacks of current design solutions and leads to an improvement in the resulting number of gates by up to 92%.

Download Paper (PDF; Only available from the DATE venue WiFi)
7.5 Hot Topic Session: The Engineering Challenges for Quantum Computing

Date: Wednesday 29 March 2017
Time: 14:30 - 16:00
Location / Room: 3C

Organisers:
Koen Bertels, QuTech & Computer Engineering Lab, NL
Carmen G. Almudéver, QuTech & Computer Engineering Lab, NL

Chair:
Edoardo Charbon, Delft University of Technology, NL

Co-Chair:
Said Hamdioui, Delft University of Technology, NL

Quantum computers may revolutionize the field of computation by solving some complex problems that are intractable even for the most powerful current supercomputers. This session will explain the basic concepts of quantum computing and describe what the required layers are for building a quantum system. The different speakers in the session will then address the engineering challenges when building a quantum computer ranging from the core qubit technology, the control electronics, to the microarchitecture for the execution of quantum circuits and efficient quantum error correction and what compiler and system tools are needed in that context.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>14:30</td>
<td>7.5.1</td>
<td>WHAT IS QUANTUM COMPUTING ALL ABOUT?</td>
<td>Speaker: Carmen G. Almudever, Delft University of Technology, NL</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Authors: Carmen G. Almudever and Koen Bertels, Delft University of Technology, NL</td>
</tr>
<tr>
<td>15:00</td>
<td>7.5.2</td>
<td>QUANTUM PROCESSOR</td>
<td>Author: Andreas Wallraff, ETH Zurich, CH</td>
</tr>
<tr>
<td>15:30</td>
<td>7.5.3</td>
<td>CONTROL ELECTRONICS FOR QUANTUM COMPUTER</td>
<td>Author: Hendrik Bluhm, RWTH Aachen, DE</td>
</tr>
<tr>
<td>16:00</td>
<td></td>
<td>End of session</td>
<td>Coffee Break in Exhibition Area</td>
</tr>
</tbody>
</table>

On all conference days (Tuesday to Thursday), coffee and tea will be served during the coffee breaks at the below-mentioned times in the exhibition area.

Tuesday, March 28, 2017
- Coffee Break 10:30 - 11:30
- Coffee Break 16:00 - 17:00

Wednesday, March 29, 2017
- Coffee Break 10:00 - 11:00
- Coffee Break 16:00 - 17:00

Thursday, March 30, 2017
- Coffee Break 10:00 - 11:00
- Coffee Break 15:30 - 16:00

7.6 Memory Reliability: Modeling and Mitigation

Date: Wednesday 29 March 2017
Time: 14:30 - 16:00
Location / Room: 5A

Chair:
Jose Pineda De Gyvez, NXP, NL

Co-Chair:
Vikas Chandra, ARM, US

This session discusses new trends and solutions to model and mitigate resiliency challenges for advanced memory technologies. The first paper discusses unequal protection for...
more efficient memory resiliency. The second paper analyzes the aging impact on different memory components. Finally, the third paper proposes mitigation schemes for memory peripheral circuitry.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>14:30</td>
<td>7.6.1</td>
<td>(Best Paper Award Candidate) MVP ECC : MANUFACTURING PROCESS VARIATION AWARE UNEQUAL PROTECTION ECC FOR MEMORY RELIABILITY</td>
<td>Joon-Sung Yang, Sungkyunkwan University, KR</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Speaker:</strong></td>
<td>Joon-Sung Yang, Sungkyunkwan University, KR</td>
</tr>
<tr>
<td></td>
<td>Authors:</td>
<td>Josef Kinseher, Intel Deutschland, DE</td>
<td>Josef Kinseher¹, Leonhard Heiß¹ and Ilia Polian ²</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Authors:</strong></td>
<td>Josef Kinseher¹, Leonhard Heiß¹ and Ilia Polian ²</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Modern System-on-Chips rely heavily on the performance of their embedded memories which are also most susceptible to the increasing reliability challenges of today's nanoscale technology nodes. However, in contrast to memory care-cells, the effects of transistor aging inside the peripheral logic of SRAM architectures have received little attention. This study works out how BTI and HCI induced wear-out of the peripheral SRAM circuitry impacts various performance metrics of an industrially used memory library. We show that the degradation of the peripheral logic is the dominant driver for access speed loss while it tends to slightly lower memory read margin and lead to minor improvements of write margin. We furthermore show that in terms of access margin the degradation of SRAM control circuitry counteracts aging effects inside core-cells and sense amplifiers. Surprisingly, wear-out of peripheral circuitry can even improve access margin in case when the relative magnitude of BDTI is much lower compared with NBTI. Based on the example of an embedded memory library, this study further underlines the importance to analyze aging mechanisms at system level rather than for its individual interacting sub-circuits. Download Paper (PDF; Only available from the DATE venue WiFi)</td>
<td>Josef Kinseher¹, Leonhard Heiß¹ and Ilia Polian ²</td>
</tr>
<tr>
<td>16:00</td>
<td>7.6.2</td>
<td>ANALYZING THE EFFECTS OF PERIPHERAL CIRCUIT AGING OF EMBEDDED SRAM ARCHITECTURES</td>
<td>Daniel Kraak, Delft University of Technology, NL</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Speaker:</strong></td>
<td>Daniel Kraak, Delft University of Technology, NL</td>
</tr>
<tr>
<td></td>
<td>Authors:</td>
<td>Daniel Kraak, Innocent Agbo¹, Mottaqiallah Taouil¹, Said Hamdioui¹, Pieter Weckx², Stefan Cosemans², Francky Catthoor² and Wim Dehaene³</td>
<td>Daniel Kraak, Innocent Agbo¹, Mottaqiallah Taouil¹, Said Hamdioui¹, Pieter Weckx², Stefan Cosemans², Francky Catthoor² and Wim Dehaene³</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Modern System-on-Chips rely heavily on the performance of their embedded memories which are also most susceptible to the increasing reliability challenges of today's nanoscale technology nodes. However, in contrast to memory care-cells, the effects of transistor aging inside the peripheral logic of SRAM architectures have received little attention. This study works out how BTI and HCI induced wear-out of the peripheral SRAM circuitry impacts various performance metrics of an industrially used memory library. We show that the degradation of the peripheral logic is the dominant driver for access speed loss while it tends to slightly lower memory read margin and lead to minor improvements of write margin. We furthermore show that in terms of access margin the degradation of SRAM control circuitry counteracts aging effects inside core-cells and sense amplifiers. Surprisingly, wear-out of peripheral circuitry can even improve access margin in case when the relative magnitude of BDTI is much lower compared with NBTI. Based on the example of an embedded memory library, this study further underlines the importance to analyze aging mechanisms at system level rather than for its individual interacting sub-circuits. Download Paper (PDF; Only available from the DATE venue WiFi)</td>
<td>Daniel Kraak, Innocent Agbo¹, Mottaqiallah Taouil¹, Said Hamdioui¹, Pieter Weckx², Stefan Cosemans², Francky Catthoor² and Wim Dehaene³</td>
</tr>
<tr>
<td>15:00</td>
<td>7.6.3</td>
<td>MITIGATION OF SENSE AMPLIFIER DEGRADATION USING INPUT SWITCHING</td>
<td>Seungyeob Lee and Joon-Sung Yang, Sungkyunkwan University, KR</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Speaker:</strong></td>
<td>Seungyeob Lee and Joon-Sung Yang, Sungkyunkwan University, KR</td>
</tr>
<tr>
<td></td>
<td>Authors:</td>
<td>Daniel Kraak, Delft University of Technology, NL; imec, BE; KU Leuven, ESAT, BE</td>
<td>Daniel Kraak, Innocent Agbo¹, Mottaqiallah Taouil¹, Said Hamdioui¹, Pieter Weckx², Stefan Cosemans², Francky Catthoor² and Wim Dehaene³</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MITIGATION OF SENSE AMPLIFIER DEGRADATION USING INPUT SWITCHING</td>
<td>Daniel Kraak, Innocent Agbo¹, Mottaqiallah Taouil¹, Said Hamdioui¹, Pieter Weckx², Stefan Cosemans², Francky Catthoor² and Wim Dehaene³</td>
</tr>
<tr>
<td></td>
<td></td>
<td>To compensate for time-zero (due to process variation) and time-dependent (due to e.g. Bias Temperature Instability) variability, designers usually add design margins. Due to technology scaling, these variabilities become worse, leading to the need for bigger design margins. Typically, only worst-case scenarios are considered, which will not present the actual workload of the targeted application. Alternatively, mitigation schemes can be used to counteract the variability. This paper presents a run-time design-for-reliability scheme for memory Sense Amplifiers (SAs); SAs are an integral part of any memory system and are very critical for high performance. The proposed scheme mitigates the impact of time-dependent variability due to aging by using an on-line control circuit to create a balanced workload. The simulation results show that the proposed scheme can reduce the most critical figures-of-merit, namely the offset voltage shift and the sensing delay of the SA with up to ~40% and ~10%, respectively, depending on the stress conditions (temperature, voltage, workload). Download Paper (PDF; Only available from the DATE venue WiFi)</td>
<td>Daniel Kraak, Innocent Agbo¹, Mottaqiallah Taouil¹, Said Hamdioui¹, Pieter Weckx², Stefan Cosemans², Francky Catthoor² and Wim Dehaene³</td>
</tr>
<tr>
<td>16:00</td>
<td>IP3-15, 16</td>
<td>A BRIDGING FAULT MODEL FOR LINE COVERAGE IN THE PRESENCE OF UNDETECTED TRANSITION FAULTS</td>
<td>Josef Pomeranz, Purdue University, US</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Speaker and Author:</strong></td>
<td>Josef Pomeranz, Purdue University, US</td>
</tr>
<tr>
<td></td>
<td></td>
<td>A variety of fault models have been defined to capture the behaviors of commonly occurring defects and ensure a high quality of testing. When several fault models are used for test generation, it is advantageous if the existence of an undetectable fault in one model does not imply that a fault in the same component but from a different model is also undetectable. This allows a test set to cover the circuit more thoroughly when additional fault models are used. This paper studies the possibility of defining such fault models by considering transition faults as the first fault model, and bridging faults as the second fault model. The bridging faults are defined to cover lines for which transition faults are not detected. A test compaction procedure is developed to demonstrate the bridging fault coverage that can be achieved, and the effect on the number of tests. Download Paper (PDF; Only available from the DATE venue WiFi)</td>
<td>Josef Pomeranz, Purdue University, US</td>
</tr>
<tr>
<td>16:00</td>
<td></td>
<td>End of session</td>
<td>End of session</td>
</tr>
</tbody>
</table>

Coffee Break in Exhibition Area

On all conference days (Tuesday to Thursday), coffee and tea will be served during the coffee breaks at the below-mentioned times in the exhibition area.

Tuesday, March 28, 2017
- Coffee Break 10:30 - 11:30
- Coffee Break 16:00 - 17:00

Wednesday, March 29, 2017
- Coffee Break 10:00 - 11:00
- Coffee Break 16:00 - 17:00

Thursday, March 30, 2017
- Coffee Break 10:00 - 11:00
- Coffee Break 15:30 - 16:00
Embedded architectures have to often provide application performance guarantees despite stringent resource constraints. The talks in this session provide solutions to managing the limited resources of such platforms and analysing the impact of resource allocation - both from the power and performance perspective.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>14:30</td>
<td>7.7.1</td>
<td>(Best Paper Award Candidate) SCALABLE PROBABILISTIC POWER BUDGETING FOR MANY-CORES</td>
<td>Anuj Pathania, Karsruhe Institute of Technology, IN</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker:</td>
<td>Anuj Pathania¹, Heba Khr², Muhammad Shafique³, Tulika Mitra⁴ and Joerg Henkel⁵</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Authors:</td>
<td>¹Karsruhe Institute of Technology, DE; ²Karsruhe Institute of Technology (KIT), DE; ³Vienna University of Technology (TU Wien), AT; ⁴National University of Singapore, SG</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abstract:</td>
<td>Many-core processors exhibit hundreds to thousands of cores, which can execute lots of multi-threaded tasks in parallel. Restrictive power dissipation capacity of a many-core prevents all its executing tasks from operating at their peak performance together. Furthermore, the ability of a task to exploit part of the power budget allocated to it depends upon its current execution phase. This mandates careful rationing of the power budget amongst the tasks for full exploitation of the many-core. Past research proposed power budgeting techniques that redistribute power budget amongst tasks based on up-to-date information about their current phases. This phase information needs to be constantly propagated throughout the system and processed, inhibiting scalability. In this work, we propose a novel probabilistic technique for power budgeting which requires no exchange of phase information yet provides guarantees on judicial use of the power budget. The proposed probabilistic technique reduces the power budgeting overheads by 97.13% in comparison to a non-probabilistic approach, while providing almost equal performance on a simulated thousand-core system.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Download Paper (PDF; Only available from the DATE venue WiFi)</td>
<td></td>
</tr>
<tr>
<td>15:00</td>
<td>7.7.2</td>
<td>EXPLOITING SPORADIC SERVERS TO PROVIDE BUDGET SCHEDULING FOR ARINC653 BASED REAL-TIME VIRTUALIZATION ENVIRONMENTS</td>
<td>Matthias Beckert, Institute of Computer and Network Engineering, TU Braunschweig, DE</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker:</td>
<td>Matthias Beckert¹, Kai Björn Gemlau² and Rolf Ernst³</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Authors:</td>
<td>¹Institut für Datentechnik und Kommunikationsnetze - TU Braunschweig, DE; ²TU Braunschweig, DE</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abstract:</td>
<td>Virtualization techniques for embedded real-time systems typically employ TDMA schedule to achieve temporal isolation among different virtualized partitions. Due to the fixed TDMA schedule, worst case response times for IRQs and tasks are significantly increased. Recent publications introduced slack based IRQ shaping to mitigate this problem. While providing better response times for IRQs, those mechanisms neither improve task timings nor provide a work conserving scheduling. In order to provide such capabilities while still providing temporal isolation, we introduce a method based on the well known sporadic server model. In combination with a proposed budget scheduler the system is able to schedule a TDMA based configuration while providing better response times and the same amount of temporal isolation. We show correctness of the approach and evaluate it in a hypervisor implementation.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Download Paper (PDF; Only available from the DATE venue WiFi)</td>
<td></td>
</tr>
<tr>
<td>15:30</td>
<td>7.7.3</td>
<td>PROGRAMMING AND ANALYSING SCENARIO-AWARE DATAFLOW ON A MULTI-PROCESSOR PLATFORM</td>
<td>Reinier van Kampenhout, Eindhoven University of Technology, NL</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker:</td>
<td>Reinier van Kampenhout, Sander Stuijk and Kees Gossens, Eindhoven University of Technology, NL</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Authors:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abstract:</td>
<td>The FSM-SADF model of computation is especially suitable for analysing real-time applications with input-dependent behaviour such as different modes, variable execution times and scalable parallelism. Although FSM-SADF specifies which scenario transitions are possible, it does not specify how and when they are decided at runtime. Multiple actors of a scenario, e.g. video stream header parsing, may have to fire before it is known which scenario the application is in. We solve this causality dilemma with a concept for executing a sequence of scenarios, and demonstrate an implementation on multiple processors with rolling static-order scheduling. We furthermore present a platform-aware analysis model that covers concept and implementation, and integrate the contributions in a toolflow. A proof-of-concept confirms the low overhead of the implementation and the exact timing analysis of our model.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Download Paper (PDF; Only available from the DATE venue WiFi)</td>
<td></td>
</tr>
<tr>
<td>16:00</td>
<td>7.7.4</td>
<td>CHRT: A CRITICALITY- AND HETEROGENEITY-AWARE RUNTIME SYSTEM FOR TASK-PARALLEL APPLICATIONS</td>
<td>Myeonggyun Han, UNIST, KR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker:</td>
<td>Myeonggyun Han, Jinsu Park and Woongki Baek, UNIST, KR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Authors:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abstract:</td>
<td>Heterogeneous multiprocessing (HMP) is an emerging technology for high-performance and energy-efficient computing. While task parallelism is widely used in various computing domains from the embedded to machine-learning computing domains, relatively little work has been done to investigate the efficient runtime support that effectively utilizes the criticality of the tasks of the target application and the heterogeneity of the underlying HMP system with full resource management. To bridge this gap, we propose a criticality- and heterogeneity-aware runtime system for task-parallel applications (CHRT). CHRT dynamically estimates the performance and power consumption of the task-parallel application and robustly manages the full HMP system resources (i.e., core types, counts, and voltage/frequency levels) to maximize the overall efficiency. Our experimental results show that CHRT achieves significantly higher energy efficiency than the baseline runtime system that employs the breadth-first scheduler and the state-of-the-art criticality-aware runtime system.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Download Paper (PDF; Only available from the DATE venue WiFi)</td>
<td></td>
</tr>
</tbody>
</table>
### MobiXen: Porting Xen on Android Devices for Mobile Virtualization

**Speaker:** Jianguo Yao, Shanghai Jiao Tong University, CN  
**Authors:** Yaozu Dong¹, Jianguo Yao¹, Haibing Guan², Ananth. Krishna R¹ and Yunhong Jiang¹  
¹Intel, US; ²Shanghai Jiao Tong University, CN  
**Abstract**  
The mobile virtualization technology provides a feasible way to improve the manageability and security for embedded systems. This paper presents an architecture named MobiXen to address these challenges. In the MobiXen, both Xen’s physical memory space and virtual address space are shrunk as much as possible and thus Android owns more memory resource; optimizations are developed to reduce the virtualization overhead when Android is accessing system resources; new policies are implemented to achieve low suspend/resume latency. With these work adopted, MobiXen is customized as a high efficient mobile hypervisor. Detailed implementations shows that, most of the performance degradation brought by MobiXen is less than 3%, which is imperceptible by end users.  
**Download Paper (PDF; Only available from the DATE venue WiFi)**

### Optimisation Opportunities and Evaluation for GPGPU Applications on Low-End Mobile GPUs

**Speaker:** Leonidas Kosmidis, Barcelona Supercomputing Center and Universitat Politècnica de Catalunya, ES  
**Authors:** Matina Maria Trompouki¹ and Leonidas Kosmidis²  
¹Universitat Politècnica de Catalunya, ES; ²Barcelona Supercomputing Center and Universitat Politècnica de Catalunya, ES  
**Abstract**  
Previous works in the literature have shown the feasibility of general purpose computations for non-visual applications on low-end mobile graphics processors using graphics APIs. These works focused only on the functional aspects of the software, ignoring the implementation details and therefore their performance implications due to their particular micro-architecture. Since various steps in such applications can be implemented in multiple ways, we identify optimisation opportunities, explore the different options and evaluate them. We show that the implementation details can significantly affect the obtained performance with discrepancies up to 3 orders of magnitude and we demonstrate the effectiveness of our proposal on two embedded platforms, obtaining more than 16x speedup over benchmarks designed following OpenGL ES 2 best practices.  
**Download Paper (PDF; Only available from the DATE venue WiFi)**

### Smart Energy and Self-Powered Devices

**Date:** Wednesday 29 March 2017  
**Time:** 14:30 - 15:30  
**Location / Room:** Exhibition Theatre  
**Organiser:** Patrick Mayor, EPFL, CH  

The goal of this session is to present concrete examples of novel designs for next-generation energy-efficient computing architectures and real-time monitoring and management of smart grids, as well as robust low-power networks of acoustic detectors for natural hazard warning systems.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>14:30</td>
<td>7.8.1</td>
<td>YINS</td>
<td>Eugene Van Rooyen, Eaton, CH</td>
</tr>
<tr>
<td>14:50</td>
<td>7.8.2</td>
<td>SMARTGRID</td>
<td>Marco Pignati and Sergio Barreto, EPFL, CH</td>
</tr>
<tr>
<td>15:10</td>
<td>7.8.3</td>
<td>X-SENSE II</td>
<td>Jan Boeuf, ETHZ, CH</td>
</tr>
<tr>
<td>15:30</td>
<td></td>
<td>End of session</td>
<td></td>
</tr>
</tbody>
</table>
On all conference days (Tuesday to Thursday), coffee and tea will be served during the coffee breaks at the below-mentioned times in the exhibition area.

Tuesday, March 28, 2017
- Coffee Break 10:30 - 11:30
- Coffee Break 16:00 - 17:00

Wednesday, March 29, 2017
- Coffee Break 10:00 - 11:00
- Coffee Break 16:00 - 17:00

Thursday, March 30, 2017
- Coffee Break 10:00 - 11:00
- Coffee Break 15:30 - 16:00

### I3 Interactive Presentations

**Date:** Wednesday 29 March 2017  
**Time:** 16:00 - 16:30
**Location / Room:** IP sessions (in front of rooms 4A and 5A)

Interactive Presentations run simultaneously during a 30-minute slot. A poster associated to the IP paper is on display throughout the afternoon. Additionally, each IP paper is briefly introduced in a one-minute presentation in a corresponding regular session, prior to the actual Interactive Presentation. At the end of each afternoon Interactive Presentations session the award 'Best IP of the Day' is given.

<table>
<thead>
<tr>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
</table>
| IP3-1 | LEVERAGING AGING EFFECT TO IMPROVE SRAM-BASED TRUE RANDOM NUMBER GENERATORS | Mohammad Saber Golanbari, Karlsruhe Institute of Technology (KIT), DE  
Authors:  
Saman Kiamhori1, Mohammad Saber Golanbari2 and Mehdi Tahoori 3  
1Karlsruhe Institute of Technology (KIT), DE; 2Karlsruhe Institute of Technology, DE |
| Abstract | The start-up value of SRAM cells can be used as the random number vector or a seed for the generation of a pseudo random number. However, the randomness of the generated number is pretty low since many of the cells are largely skewed due to process variation and their start-up value leans toward zero or one. In this paper, we propose an approach to increase the randomness of SRAM-based True Random Number Generators (TRNGs) by leveraging transistor aging impact. The idea is to iteratively power-up the SRAM cells and put them under accelerated aging to make the cells less skewed and hence obtaining a more random vector. The simulation results show that the min-entropy of SRAM-based TRNG increases by 10X using this approach. |
| Download Paper (PDF; Only available from the DATE venue WiFi) |

<table>
<thead>
<tr>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
</table>
| IP3-2 | DESIGN AUTOMATION FOR OBFUSCATED CIRCUITS WITH MULTIPLE VIABLE FUNCTIONS | Shahrzad Keshavarz, University of Massachusetts Amherst, US  
Authors:  
Shahrzad Keshavarz1, Christof Paar2 and Daniel Holcomb1  
1University of Massachusetts Amherst, US; 2Horst Gortz Institut for IT-Security, Ruhr-Universitat Bochum, DE |
| Abstract | Gate camouflaging is a technique for defusing the function of a circuit against reverse engineering attacks. However, if an adversary has pre-existing knowledge about the set of functions that are viable for an application, random camouflaging of gates will not obfuscate the function well. In this case, the adversary can target their search, and only needs to decide whether each of the viable functions could be implemented by the circuit. In this work, we propose a method for using camouflaged cells to obfuscate a design that has a known set of viable functions. The circuit produced by this method ensures that an adversary will not be able to rule out any viable functions unless she is able to uncover the gate functions of the camouflaged cells. Our method comprises iterated synthesis within an overall optimization loop to combine the viable functions, followed by technology mapping to deploy camouflaged cells while maintaining the plausibility of all viable functions. We evaluate our technique on cryptographic S-box functions and show that, relative to a baseline approach, it achieves up to 38% area reduction in PRESENT-style S-Boxes and 48% in DES S-Boxes. |
| Download Paper (PDF; Only available from the DATE venue WiFi) |

<table>
<thead>
<tr>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
</table>
| IP3-3 | DOUBLE MAC: DOUBLING THE PERFORMANCE OF CONVOLUTIONAL NEURAL NETWORKS ON MODERN FPGAS | Jongeun Lee, UNIST, KR  
Authors:  
Dong Nguyen1, Daewoo Kim1 and Jongeun Lee2  
1UNIST, KR; 2Pusan National Institute of Science and Technology (UNIST), KR |
| Abstract | This paper presents a novel method to double the computation rate of convolutional neural network (CNN) accelerators by packing two multiply-and-accumulate (MAC) operations into one DSP block of off-the-shelf FPGAs (called Double MAC). While a general SIMD MAC using a single DSP block seems impossible, our solution is tailored for the kind of MAC operations required for a convolution layer. Our preliminary evaluation shows that not only can our Double MAC approach increase the computation throughput of a CNN layer by twice with essentially the same resource utilization, the network level performance can also be improved by 14–84% over a highly optimized state-of-the-art accelerator solution depending on the CNN hyper-parameters. |
| Download Paper (PDF; Only available from the DATE venue WiFi) |
IP3-4 BITMAN: A TOOL AND API FOR FPGA BITSTREAM MANIPULATIONS
Speaker: Dirk Koch, University of Manchester, GB
Authors: Khoa Pham, Edson Horta and Dirk Koch, University of Manchester, GB
Abstract
To fully support the partial reconfiguration capabilities of FPGAs, this paper introduces the tool and API BitMan for generating and manipulating configuration bitstreams. BitMan supports recent Xilinx FPGAs that can be used by the ISE and Vivado tool suites of the FPGA vendor Xilinx, including latest Virtex-6, 7 Series, UltraScale and UltraScale+ series FPGAs. The functionality includes high-level commands such as cutting out regions of a bitstream and placing or relocating modules on an FPGA as well as low-level commands for modifying primitives and for routing clock networks or rerouting signal connections at run-time. All this is possible without the vendor CAD tools for allowing BitMan to be used even with embedded CPUs. The paper describes the capabilities, API and performance evaluation of BitMan.
Download Paper (PDF; Only available from the DATE venue WiFi)

IP3-5 A GENERIC TOPOLOGY SELECTION METHOD FOR ANALOG CIRCUITS WITH EMBEDDED CIRCUIT SIZING DEMONSTRATED ON THE OTA EXAMPLE
Speaker: Andreas Gerlach, Robert Bosch Centre for Power Electronics, DE
Authors: Andreas Gerlach\(^1\), Thoralf Rosahl\(^2\), Frank-Thomas Eitrich\(^3\) and Jürgen Scheible\(^1\)
\(^1\)Robert Bosch Centre for Power Electronics, DE; \(^2\)Robert Bosch GmbH, DE
Abstract
We present a methodology for automatic selection and sizing of analog circuits demonstrated on the OTA circuit class. The methodology consists of two steps: a generic topology selection method supported by a "part-sizing" process and subsequent final sizing. The circuit topologies provided by a reuse library are classified in a topology tree. The appropriate topology is selected by traversing the topology tree starting at the root node. The decision at each node is gained from the result of the part-sizing, which is in fact a node-specific set of simulations. The final sizing is a simulation-based optimization. We significantly reduce the overall simulation effort compared to a classical simulation-based optimization by combining the topology selection with the part-sizing process in the selection loop. The result is an interactive user friendly system, which eases the analog designer’s work significantly when compared to typical industrial practice in analog circuit design. The topology selection method with sizing is implemented as a tool into a typical analog design environment.
Download Paper (PDF; Only available from the DATE venue WiFi)

IP3-6 LATENCY ANALYSIS OF HOMOGENEOUS SYNCHRONOUS DATAFLOW GRAPHS USING TIMED AUTOMATA
Speaker: Guus Kuiper, University of Twente, NL
Authors: Guus Kuiper\(^1\) and Marco Bekooij\(^2\)
\(^1\)University of Twente, NL; \(^2\)University of Twente + NXP semiconductors, NL
Abstract
There are several analysis models and corresponding temporal analysis techniques for checking whether applications executed on multiprocessor systems meet their real-time constraints. However, currently there does not exist an exact end-to-end latency analysis technique for Homogeneous Synchronous Dataflow (HSDF) with Auto-concurrency (HSDFa) models that takes the correlation between the firing durations of different firings into account. In this paper we present a transformation of strongly connected (HSDFa) models into timed automata models. This enables an exact end-to-end latency analysis because the correlation between the firing durations of different firings is taken into account. In a case study we compare the latency obtained using timed automata and a Linear Program (LP) based analysis technique that relies on a deterministic abstraction and compare their run-times as well. Exact end-to-end latency analysis results are obtained using timed automata, whereas this is not possible using deterministic timed-dataflow models.
Download Paper (PDF; Only available from the DATE venue WiFi)

IP3-7 COVERT: COUNTER OVERFLOW REDUCTION FOR EFFICIENT ENCRYPTION OF NON-VOLATILE MEMORIES
Speaker: Kartik Mohanram, ECE Dept, University of Pittsburgh, US
Authors: Shivam Swami and Kartik Mohanram, University of Pittsburgh, US
Abstract
Security vulnerabilities arising from data persistence in emerging non-volatile memories (NVMs) necessitate memory encryption to ensure data security. Whereas counter mode encryption (CME) is a stop-gap practical approach to address this concern, it suffers from frequent memory re-encryption (system freeze) for small-sized counters and poor system performance for large-sized counters. CME thus imposes heavy overheads on memory, system performance, and system availability in practice. We propose Counter OverFlow Reduction (COVERT), a CME-based memory encryption solution that performs on-demand memory allocation to reduce the memory encryption frequency of fast growing counters, while also retaining the area/performance benefits of small-sized counters. Our full-system simulations of a phase change memory (PCM) architecture across SPEC CPU2006 benchmarks show that for equivalent overhead and no impact to performance, COVERT simultaneously reduces the full memory re-encryption frequency from 6 minutes to 25 hours and doubles memory lifetime in comparison to state-of-the-art CME techniques.
Download Paper (PDF; Only available from the DATE venue WiFi)

IP3-8 A WEAR-LEVELING-AWARE COUNTER MODE FOR DATA ENCRYPTION IN NON-VOLATILE MEMORIES
Speaker: Fangting Huang, Huazhong University of Science and Technology, CN
Authors: Fangting Huang\(^1\), Dan Feng\(^2\), Yu Hua\(^3\) and Wen Zhou\(^2\)
\(^1\)Huazhong University of Science and Technology, CN; \(^2\)Wuhan National Lab for Optoelectronics, School of Computer Science and Technology, Huazhong University of Science and Technology, China, CN
Abstract
Counter-mode encryption has been widely used to resist NVMs from malicious attacks, due to its proved security and high performance. However, this scheme suffers from the counter size versus re-encryption problem, where per-line counters must be relatively large to avoid counter overflow, or re-encryption of the entire memory is required to ensure security. In order to address this problem, we propose a novel wear-leveling-aware counter mode for data encryption, called Resetting Counter via Remapping (RCR). The basic idea behind RCR is to leverage wear-leveling remappings to reset the line counter. With carefully designed procedure, RCR avoids counter overflow with much smaller counter size. The salient features of RCR include low storage overhead of counters, high counter cache hit ratio, and no extra re-encryption overhead. Compared with state-of-the-art works, RCR obtains significant performance improvements, e.g., up to a 57% reduction in the IPC degradation, under the evaluation of 8 memory-intensive benchmarks from SPEC 2006.
Download Paper (PDF; Only available from the DATE venue WiFi)
A NEW METHOD TO IDENTIFY THRESHOLD LOGIC FUNCTIONS

Speaker:
Seyed Nima Mozaffari, Spyros Tragoudas and Themistoklis Haniotakis, Southern Illinois University, US

Authors:
Seyed Nima Mozaffari, Spyros Tragoudas, Themistoklis Haniotakis.

Abstract
An Integer Linear Programming based method to identify current mode threshold logic functions is presented. The approach minimizes the transistor count and benefits from a generalized definition of threshold logic functions. Process variations are taken into consideration. Experimental results show that many more functions can be implemented with predetermined hardware overhead, and the hardware requirement of a large percentage of existing threshold functions is reduced.

Download Paper (PDF; Only available from the DATE venue WiFi)
A BRIDGING FAULT MODEL FOR LINE COVERAGE IN THE PRESENCE OF UNDETECTED TRANSITION FAULTS

Speaker and Author:
Irith Pomeranz, Purdue University, US

Abstract
A variety of fault models have been defined to capture the behaviors of commonly occurring defects and ensure a high quality of testing. When several fault models are used for test generation, it is advantageous if the existence of an undetectable fault in one model does not imply that a fault in the same component but from a different model is also undetectable. This allows a test set to cover the circuit more thoroughly when additional fault models are used. This paper studies the possibility of defining such fault models by considering transition faults as the first fault model, and bridging faults as the second fault model. The bridging faults are defined to cover lines for which transition faults are not detected. A test compaction procedure is developed to demonstrate the bridging fault coverage that can be achieved, and the effect on the number of tests.

Download Paper (PDF; Only available from the DATE venue WiFi)

CHRT: A CRITICALITY- AND HETEROGENEITY-AWARE RUNTIME SYSTEM FOR TASK-PARALLEL APPLICATIONS

Speaker:
Myeonggyun Han, UNIST, KR

Authors:
Myeonggyun Han, Jinsu Park and Woongki Baek, UNIST, KR

Abstract
Heterogeneous multiprocessing (HMP) is an emerging technology for high-performance and energy-efficient computing. While task parallelism is widely used in various computing domains from the embedded to machine-learning computing domains, relatively little work has been done to investigate the efficient runtime support that effectively utilizes the criticality of the tasks of the target application and the heterogeneity of the underlying HMP system with full resource management. To bridge this gap, we propose a criticality- and heterogeneity-aware runtime system for task-parallel applications (CHRT). CHRT dynamically estimates the performance and power consumption of the target task-parallel application and robustly manages the full HMP system resources (i.e., core types, counts, and voltage/frequency levels) to maximize the overall efficiency. Our experimental results show that CHRT achieves significantly higher energy efficiency than the baseline runtime system that employs the breadth-first scheduler and the state-of-the-art criticality-aware runtime system.

Download Paper (PDF; Only available from the DATE venue WiFi)

MOBIXEN: PORTING XEN ON ANDROID DEVICES FOR MOBILE VIRTUALIZATION

Speaker:
Jianqiu Yao, Shanghai Jiao Tong University, CN

Authors:
Yaozu Dong, Jangiu Yao, Haibing Guan, Ananth. Krishna R l and Yunhong Jiang

Intel, US; Shanghai Jiao Tong University, CN

Abstract
The mobile virtualization technology provides a feasible way to improve the manageabley and security for embedded systems. This paper presents an architecture named MobiXen to address these challenges. In the MobiXen, both Xen’s physical memory space and virtual address space are shrunk as much as possible and thus Android owns more memory resource; optimizations are developed to reduce the virtualization overhead when Android is accessing system resources; new policies are implemented to achieve low suspend/resume latency. With these work adopted, MobiXen is customized as a high efficient mobile hypervisor. Detailed implementations shows that, most of the performance degradation brought by MobiXen is less than 2%, which is imperceptible by end users.

Download Paper (PDF; Only available from the DATE venue WiFi)

OPTIMISATION OPPORTUNITIES AND EVALUATION FOR GPGPU APPLICATIONS ON LOW-END MOBILE GPUS

Speaker:
Leonidas Kosmidis, Barcelona Supercomputing Center and Universitat Politècnica de Catalunya, ES

Authors:
Matina Maria Trompouki and Leonidas Kosmidis

Universitat Politècnica de Catalunya, ES; Barcelona Supercomputing Center and Universitat Politècnica de Catalunya, ES

Abstract
Previous works in the literature have shown the feasibility of general purpose computations for non-visual applications on low-end mobile graphics processors using graphics APIs. These works focused only on the functional aspects of the software, ignoring the implementation details and therefore their performance implications due to their particular micro-architecture. Since various steps in such applications can be implemented in multiple ways, we identify optimisation opportunities, explore the different options and evaluate them. We show that the implementation details can significantly affect the obtained performance with discrepancies up to 3 orders of magnitude and we demonstrate the effectiveness of our proposal on two embedded platforms, obtaining more than 16x speedup over benchmarks designed following OpenGL ES 2 best practices.

Download Paper (PDF; Only available from the DATE venue WiFi)
UB08.1 COSSIM: A NOVEL, COMPREHENSIBLE, ULTRA-FAST, SECURITY-AWARE CPS SIMULATOR

Presenter: Nikolaos Tampouratzis, Technical University of Crete, GR
Authors: Antonios Nikitakis and Andreas Brokalakis, Synelixis Solutions Ltd, GR
Abstract: One of the main problems Cyber Physical Systems (CPS) and Highly Parallel Systems (HPS) designers face is the lack of simulation tools and models for system design and analysis. This is mainly because the majority of the existing simulation tools can handle efficiently only parts of a system (e.g. only the processing or only the network) while none of them supports the notion of security. Moreover, most of the existing simulators need extreme amounts of processing resources while faster approaches cannot provide the necessary precision and accuracy. COSSIM is an open-source framework that seamlessly simulates, in an integrated way, the networking and the processing parts of the CPS and Highly Parallel Heterogeneous Systems. In addition, COSSIM supports accurate power estimations while it is the first such tool supporting security as a feature of the design process. The complete COSSIM framework together with its sophisticated GUI will be presented.

More information ...

UB08.2 NETFI-2: AN AUTOMATIC METHOD FOR FAULT INJECTION ON HDL-BASED DESIGNS

Presenter: Alexandre Coelho, Université Grenoble Alpe, FR
Authors: Miguel Solinas, Juan Fraire, Nacer-Eddine Zergainoh, Pablo Ferreyra and Raouf Velazco, TIMA, FR
Abstract: Fault injection tools, which include fault simulation and emulation, are a well-known technique to evaluate the susceptibility of integrated circuits to the effects of radiation. This work presents a methodology to emulate Single Event Upsets (SEUs) and Single Event Transients (SETs) in a Field Programmable Gate Array (FPGA). The method proposed combines the flexibility of FPGA with the controllability provided by the MicroBlaze, to emulate HDL circuit and control the fault injection campaign. This approach has been integrated into a fault-injection platform, named NETFI (NETlist Fault Injection), developed by our research group, and received the name of NETFI-2. To validate this methodology fault injection campaign have been performed in Leon3 and Stochastic Bayesian Machine. Results on an Artix-7 FPGA show that NETFI-2 provides accurate measurements while improving the execution time of the experiment by more than 300% compared with analogous simulation-based campaigns.

More information ...

UB08.5 ITMD: RUN-TIME MANAGEMENT OF CONCURRENT MULTI-THREADED APPLICATIONS ON HETEROGENEOUS MULTI-CORES

Presenter: Karunaraker Reddy Basireddy, University of Southampton, GB
Authors: Amit Singh, Bashir M. Al-Hashimi and Geoff V. Merrett, University of Southampton, GB
Abstract: Heterogeneous multi-cores often need to deal with multiple applications having different performance requirements concurrently, which generate varying and mixed workloads. Runtime management is required for adapting to such performance requirements and workload variabilities, and to achieve energy efficiency. It is challenging to efficiently exploit different types of cores simultaneously and DVFS potential of cores. We present a run-time management approach that first selects thread-to-core mapping based on the performance requirements and resource availability. Then, it applies online adaptation by adjusting the voltage-frequency (V-f) levels to achieve energy optimization. We demonstrate the proposed run-time management approach on the Odroid-XU3, with various combinations of multi-threaded applications from PARSEC and SPLASH benchmarks. Results show an average improvement in energy efficiency up to 33% compared to existing approaches.

More information ...

UB08.6 GNOCS: AN ULTRA-FAST, HIGHLY EXTENSIBLE, CYCLE-ACCURATE GPU-BASED PARALLEL NETWORK-ON-CHIP SIMULATOR

Presenter: Amir CHARIF, TIMA, FR
Authors: Nacer-Eddine Zergainoh and Michael Nicolaides, TIMA, FR
Abstract: With the continuous decrease in feature sizes and the recent emergence of 3D stacking, chips comprising thousands of nodes are becoming increasingly relevant, and state-of-the-art NoC simulators are unable to simulate such a high number of nodes in reasonable times. In this demo, we showcase GNoCS, the first detailed, modular and scalable parallel NoC simulator running fully on GPU (Graphics Processing Unit). Based on a unique design specifically tailored for GPU parallelism, GNoCS is able to achieve unprecedented speeds up with no loss of accuracy. To enable quick and easy validation of novel ideas, the programming model was designed with high extensibility in mind. Currently, GNoCS accurately models a VC-based microarchitecture. It supports 2D and 3D mesh topologies with full or partial vertical connections. A variety of routing algorithms and synthetic traffic patterns, as well as dependency-driven trace-based simulation (Netrace), are implemented and will be demonstrated.

More information ...

UB08.8 SELINK: SECURING HTTP AND HTTPS-BASED COMMUNICATION VIA SECUBE™

Presenter: Airofarulla Giuseppe, CINI & Politecnico di Torino, IT
Authors: Paolo Prinetto 1 and Antonio Varriale 2
1Politecnico di Torino, IT; 2Blu5 Labs Ltd., IT
Abstract: The SECube™ Open Source platform is a combination of three main cores in a single-chip design. Low-power ARM Cortex-M4 processor, a flexible and fast Field-Programmable-Gate-Array (FPGA), and an EAL5+ certified Security Controller (SmartCard) are embedded in an extremely compact package. This makes it a unique Open Source security environment where each function can be optimized, executed, and verified on its proper hardware device. In this demo, we present a client-server HTTP and HTTPS-based application, for which the traffic is encrypted resorting to the hardware-built-in capabilities, and the software libraries, of the SECube™. By doing so, we show how communication can be secured from an attacker capable of inspecting, and tampering, the regular communication.

More information ...

UB08.9 HEPSYCODE: A SYSTEM-LEVEL METHODOLOGY FOR HW/SW CO-DESIGN OF HETEROGENEOUS PARALLELDEDICATED SYSTEMS

Presenter: Luigi Pomante, University of L’Aquila, IT
Authors: Giacomo Valente 1, Vittoriano Muttillo 1, Daniele Di Pompeos 1, Emilio Incerto 2 and Daniele Ciambrene 1
1University of L’Aquila, IT; 2Gran Sasso Science Institute, IT
Abstract: Heterogeneous parallel systems have been recently exploited for a wide range of application domains, for both the dedicated (e.g. embedded) and the general purpose products. Such systems can include different processor cores, memories, dedicated ICs and a set of connections between them. They are so complex that the design methodology plays a major role in determining the success of the products. So, this demo addresses the problem of the electronic system-level hw/sw co-design of heterogeneous parallel dedicated systems. In particular, it shows an enhanced CSP/SystemC-based design space exploration step (and related ESL-EDA prototype tools), in the context of an existing hw/sw co-design flow that, given the system specification and related F/NF requirements, is able to (semi)automatically propose to the designer: - a custom heterogeneous parallel architecture; - an HW/SW partitioning of the application; - a mapping of the partitioned entities onto the proposed architecture.
8.2 Hot Topic Session: No Power? No Problem! Exploiting Non-Volatility in Energy Constrained Environments

Date: Wednesday 29 March 2017
Time: 17:00 - 18:30
Location / Room: SBC

Speakers:
- Sujit Kokka Chhetri, UC Irvine, US
- Mohammad Al Faruque, University of California Irvine, US

Abstract

This talk will present a methodology to collect physical information (e.g., energy flows in the form of acoustics, vibration, electro-magnetic, etc.) effectively and efficiently from a manufacturing system using IoT infrastructure. Through applying information-theoretic analysis, we will show how to create a digital twin of the manufacturing system that may be used for process control (i.e., better decision making at different time-scales) and security. We will focus on the plug and play capability provided by the IoT, which will allow us to create digital twin of legacy manufacturing systems as well. We will demonstrate our work with an application in additive manufacturing system (3D printers). We will also present how in our recent work we have demonstrated that we can breach the confidentiality of a 3D printer by reconstructing an original 3D model from the printer's acoustic emission analysis.
With the rapid growth of the Internet of Things (IoT), demands for battery-less systems are ever increasing. Systems that can be powered by ambient energy sources would offer new opportunities and capabilities for personal entertainment, self-powered, computational systems have obvious societal benefits when deployed for medical monitoring, environmental sensing, etc. This hot topic session considers the current landscape of energy harvesting computing systems and highlights the need for power neutral systems. Subsequent presentations showcase emerging non-volatile memory and logic technologies that could enable battery-less computing systems.

### 8.2.1 ENERGY-DRIVEN COMPUTING: RETHINKING THE DESIGN OF ENERGY HARVESTING SYSTEMS

**Speaker:** Geoff Merrett, University of Southampton, GB  
**Authors:** Geoff Merrett and Bashir Al-Hashimi, University of Southampton, GB  
**Abstract**  
Energy harvesting computing has been gaining increasing traction over the past decade, fueled by technological developments and rising demand for autonomous and battery-free systems. Energy harvesting introduces numerous challenges to embedded systems but, arguably the greatest, is the required transition from an energy source that typically provides virtually unlimited power for a reasonable period of time until it becomes exhausted, to a power source that is highly unpredictable and dynamic (both spatially and temporally, and with a range spanning many orders of magnitude). The typical approach to overcome this is the addition of intermediate energy storage/buffering to smooth out the temporal dynamics of both power supply and consumption. This has the advantage that, if correctly sized, the system ‘looks like’ a battery-powered system; however, it also adds volume, mass, cost and complexity and, if not sized correctly, unreliability. In this paper, we consider energy-driven computing, where systems are designed from the outset to operate from an energy harvesting source. Such systems typically contain little or no additional energy storage (instead relying on tiny parasitic and decoupling capacitance), alleviating the aforementioned issues. Examples of energy-driven computing include transient systems (which power down when the supply disappears and efficiently continue execution when it returns) and power-neutral systems (which operate directly from the instantaneous power harvested, gracefully modulating their consumption and performance to match the supply). In this paper, we introduce a taxonomy of energy-driven computing, articulating how power-neutral, transient, and energy-driven systems present a different class of computing to conventional approaches.

Download Paper (PDF; Only available from the DATE venue WiFi)

### 8.2.2 NONVOLATILE PROCESSORS: WHY IS IT TENDING?

**Speaker:** Vijaykrishnan Narayanan, Penn State University, US  
**Authors:** Fang Su1, Kaisheng Ma2, Xueqing Li2, Tongda Wu1, Yongpan Liu1 and Vijaykrishnan Narayanan2  
1Tsinghua University, CN; 2Penn State University, US  
**Abstract**  
Energy harvesting has become a promising solution to power up Internet-of-Things (IoT) devices. In this scenario, the constrained power budget and frequent absence of ambient energy cause severe reliability issues and performance degradation on conventional CMOS computing circuits. Fortunately, the advent of nonvolatile processor (NVP) opens the possibility to compute continuously using an intermittent power supply. It is considered as a key component of the next generation IoT edge devices. In this work, we provide insights to the evolution of the NVP and its application in real world scenarios. Efforts on improving the performance of NVP and future research prospects are also discussed in this paper.

Download Paper (PDF; Only available from the DATE venue WiFi)

### 8.2.3 ADVANCED SPINTRONIC MEMORY AND LOGIC FOR NON-VOLATILE PROCESSORS

**Speaker:** X. Sharon Hu, University of Notre Dame, US  
**Authors:** Robert Perricone1, Ibrahim Ahmed2, Zhaoxin Liang2, Meghna Mankalale2, X. Sharon Hu1, Chris H. Kim2, Michael Niemier1, Sachin Sapatnekar2 and Jian-Ping Wang2  
1University of Notre Dame, US; 2University of Minnesota, US; 3University Of Minnesota, US  
**Abstract**  
Many ultra-low power Internet of things (IoT) systems may be powered by energy harvested from ambient sources (e.g., solar radiation, thermal gradients, and WiFi). However, these energy sources can vary significantly in terms of their strengths and on/off patterns. For volatile systems, the intermittent nature of the energy sources necessitates the use of backup/recovery schemes to guarantee computational correctness and forward progress, which incur performance, area and energy overhead. Non-volatile (NV) processors based on spintronic devices, such as Spin-Transfer Torque (STT) memory and All-Spin-Logic (ASL), are more attractive alternatives. These NV devices are capable of achieving forward progress without relying on backup/recovery schemes. This work establishes a general framework for evaluating NV device-based processors for energy harvesting applications. Results demonstrate that NV spintronic processors can achieve significant energy savings (up to 83X) versus a hybrid CMOS (computation) and STT-RAM (backup) implementation.

Download Paper (PDF; Only available from the DATE venue WiFi)

---

### 8.3 Secure Processor Components

**Date:** Wednesday 29 March 2017  
**Time:** 17:00 - 18:30  
**Location / Room:** 2BC

**Chair:** Patrick Schaumont, Virginia Tech, US  
**Co-Chair:** Nele Mentens, Katholieke Universiteit Leuven, BE

Security concerns have put significant demands on hardware design of processors. In this session, papers will be presented that describe processor components designed to improve their performance, protect them more efficiently against side channel attacks and thereby improve the overall performance of processors used in secure applications.
This session focuses on embedded systems for human activity recognition and control. These systems combine flexible and dynamic hardware architectures with advanced novel signal processing techniques for activity recognition, myoelectric prosthesis control, motor intention decoding and brain computer interface. Finally, we will have two interactive presentations focused on embedded systems for diagnosis.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>17:00</td>
<td>8.4.1</td>
<td>ADAPTIVE COMPRESSED SENSING AT THE FINGERTIP OF INTERNET-OF-THINGS SENSORS: AN ULTRA-LOW POWER ACTIVITY RECOGNITION</td>
</tr>
</tbody>
</table>

**Authors**
Ramin Fallahzadeh, Josué Pagan Ortiz, UCM, ES
Josué Pagan Ortiz, UCM, ES; Complutense University of Madrid, ES; Washington State University, US

**Abstract**
With the proliferation of wearable devices in the Internet-of-Things applications, designing highly power-efficient solutions for continuous operation of these technologies in life-critical settings emerges. We propose a novel ultra-low power framework for adaptive compressed sensing in activity recognition. The proposed design uses a coarse-grained activity recognition module to adaptively tune the compressed sensing module for minimized sensing/transmission costs. We pose an optimization problem to minimize activity specific sensing rates and introduce a polynomial time approximation algorithm using a novel heuristic dynamic optimization tree. Our evaluations on real-world data shows that the proposed autonomous framework is capable of generating feed-back with >80% confidence and improves power reduction performance of the state-of-the-art approach by a factor of two.

Download Paper (PDF; Only available from the DATE venue WiFi)
The papers in this session discuss the use of learning as well as energy efficient circuit level implementation techniques for Neural Networks and for Green Computing in general.
18:00 8.5.3 CHAIN-NN: AN ENERGY-EFFICIENT 1D CHAIN ARCHITECTURE FOR ACCELERATING DEEP CONVOLUTIONAL NEURAL NETWORKS

Speaker: Shihao Wang, Waseda University, JP
Authors: Shihao Wang, Dajiang Zhou, Xushen Han and Yoshimura Takeshi, Waseda University, JP

Abstract: Deep convolutional neural networks (CNN) have shown their good performances in many computer vision tasks. However, the high computational complexity of CNN involves a huge amount of data that the computational processor core and memory hierarchy which occupies the majority of the power consumption. This paper presents Chain-NN, a novel energy-efficient 1D chain architecture for accelerating deep CNNs. Chain-NN consists of the dedicated dual-channel process engines (PE). In Chain-NN, convolutions are done by the 1D systolic primitives composed of a group of adjacent PEs. These systolic primitives, together with the proposed column-wise scan input pattern, can fully reuse input operand to reduce the memory bandwidth requirement for energy saving. Moreover, the 1D chain architecture allows the systolic primitives to be easily reconfigured according to specific CNN parameters with fewer design complexity. The synthesis and layout of Chain-NN is under TSMC 28nm process. It costs 3751k logic gates and 352KB on-chip memory. The results show a 576-PE Chain-NN can be scaled up to 700MHz. This achieves a peak throughput of 806.4GOPS with 567.5mW and is able to accelerate the five convolutional layers in AlexNet at a frame rate of 362.2fps. 1421.0GOPS/W power efficiency is at least 2.5x to 4.1x times better than the state-of-the-art works.

Download Paper (PDF; Only available from the DATE venue WiFi)
8.6 Hot Topic Session: Self-aware Systems: Concepts and Applications

**Date:** Wednesday 29 March 2017

**Time:** 17:00 - 18:30

**Location / Room:** 5A

**Organisers:**
Axel Jantsch, TU Wien, AT
Nikil Dutt, UC Irvine, US

**Chair:**
Nikil Dutt, UC Irvine, US

**Co-Chair:**
Amir Rahmani, TU Wien, AT

This special hot topic session addresses concepts and applications of self-awareness for engineered systems. Interest in self-awareness continues to grow with applications in diverse domains such as automotive, space, military, consumer electronics, industrial control, health care, etc. The first talk outlines the concepts of self-awareness in psychology, and its applicability in computing, as well as in the engineering of adaptive systems. The second talk reviews the role of self-awareness in autonomous driving systems and explains how system self-awareness has become an important foundation for reliable and flexible platform management of autonomous cars. The third talk presents a remote health monitoring and diagnostic system for holistic perception of a patient’s situation, and demonstrates how self-awareness is leveraged through the use of wearable sensors, contextual knowledge of the patient’s health situation, and automated reasoning of the patient’s health situation.

**8.6.1 SELF-AWARE COMPUTING SYSTEMS: FROM PSYCHOLOGY TO ENGINEERING**

**Speaker and Author:**
Peter Lewis, Aston University, GB

**Abstract**

At the current time, there are several fundamental changes in the way computing systems are being developed, deployed and used. They are becoming increasingly large, heterogeneous, uncertain, dynamic and decentralised. These complexities lead to behaviours during run time that are difficult to understand or predict. One vision for how to rise to this challenge is to endow computing systems with increased self-awareness, in order to enable advanced autonomous environments, while reducing the need for a priori domain modelling at design or deployment time.

Download Paper (PDF; Only available from the DATE venue WiFi)

**8.6.2 SELF-AWARENESS IN AUTONOMOUS SYSTEMS: SELF-DRIVING CARS**

**Speaker:**
Rolf Ernst, TU Braunschweig, DE

**Authors:**
Johannes Schlatow1, Mischa Höst1, Rolf Ernst2, Marcus Nolte2, Inga Jatzkowski2, Markus Maurer4, Christian Herber3 and Andreas Herkersdorf4
1TU Braunschweig, Institute of Computer and Network Engineering, DE; 2TU Braunschweig, DE; 3Technische Universität München, DE; 4TU München, DE

**Abstract**

This work presents a trace-driven approach to self-awareness in autonomous systems. It describes how these concepts have been translated to the computing domain, and provides examples of how their explicit consideration can lead to systems better able to manage trade-offs between conflicting goals at run time in the context of a complex environment, while reducing the need for a prior domain modelling at design or deployment time.

Download Paper (PDF; Only available from the DATE venue WiFi)

**8.6.3 SELF-AWARENESS IN REMOTE HEALTH MONITORING SYSTEMS THROUGH WEARABLE ELECTRONICS**

**Speaker:**
Axel Jantsch, TU Wien, AT

**Authors:**
Arman Anzanour1, Iman Azimi2, Maximilian Güttinger1, Amir M. Rahmani2, Nima Taherinejad3, Pasi Liljeberg1, Axel Jantsch2 and Nikil Dutt4
1University of Turku, FI; 2University of California Irvine & TU Wien, US; 3Vienna University of Technology, AT; 4UC Irvine, US

**Abstract**

In healthcare, effective monitoring of patients plays a key role in detecting health deterioration early enough. Many signs of deterioration exist as early as 24 hours prior to a serious impact on the health of a person. As hospitalization times have to be minimized, in-home or remote early warning systems can fill the gap by allowing in-home care while having the potentially problematic conditions and their signs under surveillance and control. This work presents a remote monitoring and diagnostic system that provides a holistic perspective of patients and their health conditions. We discuss how the concept of self-awareness can be used in various parts of the system such as information collection through wearable sensors, confidence assessment of the sensory data, the knowledge base of the patient's health situation, and automation of reasoning about the health situation. Our approach to self-awareness provides (i) situation awareness to consider the impact of variations such as sleeping, walking, running, and resting, (ii) system personalization by reflecting parameters such as age, body mass index, and gender, and (iii) the attention property of self-awareness to improve the energy efficiency and dependability of the system via adjusting the priorities of the sensory data collection. We evaluate the proposed method using a full system demonstration.

Download Paper (PDF; Only available from the DATE venue WiFi)
The first paper in this session presents a novel open-source hardware/software infrastructure for dynamic binary translation. The second paper presents a mechanism to improve the performance of OpenMP loops optimized with our approach by up to 93%, the performance achieved by the original OpenMP loops, where the scheduling policy is not explicitly specified. Then, the scheduling decision is left to the default runtime, which in most of the cases does not yield the best performance. In this paper, we present a schedule-aware optimization approach enabled by exploiting the parallel slack existing in loops parallelized with OpenMP. Results show that using our hardware accelerators leads to a speed-up of 8x and a cost in energy 18x lower, compared with an equivalent software approach.

Dynamic Binary Translation (DBT) is often used in hardware/software co-design to take advantage of an architecture model while using binaries from another one. The co-development of the DBT engine and of the execution architecture leads to architecture with special support to these mechanisms. In this work, we propose a hardware accelerated Dynamic Binary Translation where the first steps of the DBT process are fully accelerated in hardware. Results show that using our hardware accelerators leads to a speed-up of 8x and a cost in energy 18x lower, compared with an equivalent software approach.
**DATE-Party DATE Party | Networking Event**

**Date:** Wednesday 29 March 2017  
**Time:** 19:00 – 23:00  
**Location / Room:** The Olympic Museum

The highlight of the DATE week will again be the DATE Party, which offers the perfect occasion to meet friends and colleagues in a relaxed atmosphere while enjoying local amenities. Thus, it states one of the main networking opportunities during the DATE week.

The party is scheduled on March 29, 2017, from 1900 to 2300, and will take place in Lausanne’s most outstanding museum location: The Olympic Museum. It is beautifully located in the heart of the city, with magnificent views over the Lake Geneva and the Swiss-French Alps. Since its renovation at the end of 2013, it now hosts more than 3,000 sqm of exhibit space and a new scenography which perfectly reflects the idea and spirit behind and how rich and diverse Olympism is. Some of the themes highlighted include sports, history, culture, design, sociology, and technology.

During the evening, all delegates will have the chance to visit the different expositions for free.

**Please kindly note that it is not a seated dinner. Drinks and snacks (flying buffet) will be served in the TOM Café! All delegates, exhibitors and their guests are invited to attend the party. Please be aware that entrance is only possible with a valid party ticket. Each full conference registration includes a ticket for the DATE Party (which needs to be booked during the online registration process though). Additional tickets can be purchased on-site at the registration desk (subject to availability of tickets). Price for extra ticket: CHF 80.00 per person.**
### 9.1 Wearable and Smart Medical Devices Day: New tools and devices for chronic and acute care

**Date:** Thursday 30 March 2017  
**Time:** 08:30 - 10:00  
**Location / Room:** 5BC

**Organisers:**  
José L. Ayala, Universidad Complutense de Madrid, ES  
Chris Van Hoof, IMEC, BE

**Chair:**  
José L. Ayala, Universidad Complutense de Madrid, ES

**Co-Chair:**  
Mario Konijnenburg, IMEC, BE

This session will present the recent advances in medical devices for the clinical practice. We will attend how Industry and Academia work on designing novel wearable, ASICs and computational systems that help on promoting the novel healthcare paradigms in the treatment of chronic and acute diseases.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:30</td>
<td>9.1.1</td>
<td>WEARABLE ROBOTICS IN CLINICAL PRACTICE: PROSPECTS</td>
<td>José Luis Pons, CSIC, ES</td>
</tr>
<tr>
<td>09:00</td>
<td>9.1.2</td>
<td>OVERCOMING HEARING LOSS THROUGH NEW IMPLANT TECHNOLOGIES</td>
<td>Carl Van Himbeeck, Cochlear Technology Centre, BE</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Hearing loss is a big unmet medical need. There is a significant and growing group of people with significant hearing loss who could benefit from implant technologies. A broad range of implant and clinical solutions are developed to improve the access to the users and the professionals.</td>
<td></td>
</tr>
<tr>
<td>09:30</td>
<td>9.1.3</td>
<td>CIRCUITS AND SYSTEMS AS ENABLERS FOR NOVEL HEALTHCARE PARADIGMS</td>
<td>Mario Konijnenburg, imec, BE</td>
</tr>
</tbody>
</table>

10:00  
End of session  
Coffee Break in Exhibition Area  
On all conference days (Tuesday to Thursday), coffee and tea will be served during the coffee breaks at the below-mentioned times in the exhibition area.

- **Tuesday, March 28, 2017**  
  - Coffee Break 10:30 - 11:30  
  - Coffee Break 16:00 - 17:00

- **Wednesday, March 29, 2017**  
  - Coffee Break 10:00 - 11:00  
  - Coffee Break 16:00 - 17:00

- **Thursday, March 30, 2017**  
  - Coffee Break 10:00 - 11:00  
  - Coffee Break 15:30 - 16:00

### 9.2 Emerging Schemes for Memory Management

**Date:** Thursday 30 March 2017  
**Time:** 08:30 - 10:00  
**Location / Room:** 4BC

**Chair:**  
Arne Heittman, RWTH, DE

**Co-Chair:**  
Costin Anghel, ISEP, FR

This topic covers aspects of emerging memory architectures and functional blocks with respect to performance and endurance enhancement. In particular caches, FTL, logic-in-memory and error correction schemes covering strategies like error correction wear leveling and cache replacement are covered. NVMs like PCM, Flash and RRAMs are considered in this track.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>09:00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>09:30</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
A LOG-AWARE SYNERGIZED SCHEME FOR PAGE-LEVEL FTL DESIGN

Speaker:
Chu Li, Huazhong University of Science & Technology, CN

Authors:
Chu Li1, Sungjin Lee1, Yu Hua1, Fang Wang1, Chuntao Jiang2 and Wei Zhou1

1Huazhong University of Science and Technology, CN; 2Illinois Institute of Technology, US

Abstract
NAND flash-based Solid State Drives (SSDs) employ the Flash Translation Layer (FTL) to perform logical-to-physical address translation. Modern page-level FTLs selectively cache the address mappings in the limited SRAM while storing the mapping table in flash pages (called translation pages). However, many extra accesses to the translation pages are required for address translation, which decreases the performance and lifetime of an SSD. In this paper, we propose a Log-aware Synergized scheme for page-level FTL to reduce the extra overheads, called LSFTL. The contribution of LSFTL consists of two key elements: (i) By exploiting garbage collection overhead via reserving a small portion of each translation page as a logging area to hold multiple updates to the entries of that translation page. (ii) "Log-aware flash back" reduces the number of translation page updates by evicting multiple dirty cache line that share the same translation page in a single transaction. Extensive experimental results of trace-driven simulations show that LSFTL decreases the system response time by 39.40% on average, and up to 58.35%, and reduces the block erase count by 37.55% on average, and up to 39.99%, compared to the well-known DFTL.

Download Paper (PDF; Only available from the DATE venue WiFi)

MALRU: MISS-PENALTY AWARE LRU-BASED CACHE REPLACEMENT FOR HYBRID MEMORY SYSTEMS

Speaker:
Chen Di, Huazhong University of Science and Technology, CN

Authors:
Di Chen, Hai Jin, Xiaofei Liao, Haikun Liu, Rentsong Guo and Dong Liu, Huazhong University of Science and Technology, CN

Abstract
Current DRAM based memory systems face the scalability challenges in terms of storage density, power, and cost. Hybrid memory architecture composed of emerging Non-Volatile Memory (NVM) and DRAM is a promising approach to large-capacity and energy-efficient main memory. However, hybrid memory systems pose a new challenge to on-chip cache management due to the asymmetrical penalty of memory access to DRAM and NVM in case of cache misses. Cache hit rate is no longer an effective metric for evaluating memory access performance in hybrid memory systems. Current cache replacement policies that aim to improve cache hit rate are not efficient either. In this paper, we take into account the asymmetry of cache miss penalty on DRAM and NVM, and advocate a more general metric, Average Memory Access Time (AMAT), to evaluate the performance of hybrid memories. We propose a miss penalty-aware LRU-based (MALRU) cache replacement policy for hybrid memory systems. MALRU is aware of the source (DRAM or NVM) of missing blocks and prevents high-latency NVM blocks as well as low-latency DRAM blocks with good temporal locality from being evicted. Experimental results show that MALRU improves system performance against LRU and the state-of-the-art HAP policy by up to 20.4% and 11.7% (11.1% and 5.7% on average), respectively.

Download Paper (PDF; Only available from the DATE venue WiFi)

ENDURANCE MANAGEMENT FOR RESISTIVE LOGIC-IN-MEMORY COMPUTING ARCHITECTURES

Speaker:
Saeideh Shirinzadeh, University of Bremen, DE

Authors:
Saeideh Shirinzadeh1, Mathias Soeken2, Pierre-Emmanuel Gaillardon3, Giovanni De Micheli4 and Rolf Drechsler5
1Group of Computer Architecture, University of Bremen, Germany; 2EPFL, CH; 3University of Utah, US; 4Integrated Systems Laboratory, EPFL, Lausanne, Switzerland, CH; 5Group of Computer Architecture, University of Bremen, Germany

Abstract
Resistive Random Access Memory (RRAM) is a promising non-volatile memory technology which enables modern in-memory computing architectures. Although RRAMs are known to be superior to conventional memories in many aspects, they suffer from a low write endurance. In this paper, we focus on balancing memory write traffic as a solution to extend the lifetime of resistive crossbar architectures. As a case study, we monitor the write traffic in a Programmable Logic-in-Memory (PLM) architecture, and propose an endurance management scheme for it. The proposed endurance-aware compilation is capable of handling different trade-offs between write balance, latency, and area of the resulting PLM implementations. Experimental evaluations on a set of benchmarks including large arithmetic and control functions show that the standard deviation of writes can be reduced by 86.65% on average compared to a naive compiler, while the average number of instructions and RRAM devices also decreases by 36.45% and 13.67%, respectively.

Download Paper (PDF; Only available from the DATE venue WiFi)

LIVE TOGETHER OR DIE ALONE: BLOCK COOPERATION TO EXTEND LIFETIME OF RESISTIVE MEMORIES

Speaker:
David Kaeli, Northeastern University, US

Authors:
Mohammad Khvare Tavana, Amir Kayvan Ziaabari and David Kaeli, Northeastern University, US

Abstract
Block-level cooperation is an endurance management technique that operates on top of error correction mechanisms to extend memory lifetimes. Once an error recovery scheme fails to recover from faults in a data block, the entire physical page associated with that block is disabled and becomes unavailable to the physical address space. To reduce the page waste caused by early block failures, other blocks can support the failed block, working cooperatively to keep it alive and extend the page’s lifetime. We combine the proposed technique with different error recovery schemes, such as Error Correction Pointers (ECP) and Aegis, to increase memory lifetimes. Block cooperation is realized through metadata sharing in ECP, where one data block shares its unused metadata with another data block. When combined with Aegis, block cooperation is realized through reorganizing data layout, where blocks possessing few faults help failed blocks bring them back from the dead. Employing block cooperation at a single level (or multiple levels) on top of ECP and Aegis, we can increase memory lifetimes by 28% (37%), and 8% (14%) on average, respectively.

Download Paper (PDF; Only available from the DATE venue WiFi)

DAC: DEDUP-ASSISTED COMPRESSION SCHEME FOR IMPROVING LIFETIME OF NAND STORAGE SYSTEMS

Speaker:
Jisung Park, Seoul National University, KR

Authors:
Jisung Park1, Sungjin Lee2 and Jihong Kim1
1Seoul National University, KR; 2Inha University, KR

Abstract
Thanks to an aggressive scaling of semiconductor devices, the capacity of NAND flash-based solid-state drives (SSDs) has increased greatly. However, this benefit comes at the expense of a serious degradation of NAND device’s lifetime. In order to improve the lifetime of flash-based SSDs, various data reduction techniques, such as deduplication, lossless compression, and delta compression, are rapidly adopted to SSDs. Although each technique has been extensively studied, how to efficiently combine these techniques for maximizing their synergy effects is not investigated well. In this paper, we propose a novel deduplication-assisted compression (DAC) scheme that integrates existing data reduction techniques so that potential benefits of individual ones can be maximized while overcoming their inherent limitations. By doing so, DAC greatly reduces the amount of write traffic sent to SSDs. DAC also requires negligible hardware resources by utilizing existing hardware modules. Our evaluation results show that the proposed DAC decreases the amount of written data by up to 30% over a simple integration reduplication and lossless compression.

Download Paper (PDF; Only available from the DATE venue WiFi)

**Date:** Thursday 30 March 2017  
**Time:** 08:30 - 10:00

**Organisers:**  
Anupam Chattopadhyay, Nanyang Technological University, SG  
Muhammad Shafique, CARE-Tech, TU Wien, AT

**Chair:**  
Ahmad Sadeghi, TU Darmstadt, DE

**Co-Chair:**  
Muhammad Shafique, CARE-Tech, TU Wien, AT

The goal of this special session is to revisit the depth and breadth of CPS security, with focus on practical system and design automation aspects. In a practical system, the possible sources of security vulnerabilities and recent attacks are discussed, and it is argued that there are significant varieties of attacks that need to be accounted for in a holistic manner.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
</table>
| 10:01  | IP4-8 | LIFETIME ADAPTIVE ECC IN NAND FLASH PAGE MANAGEMENT | Shunzhuo Wang, Huazhong University of Science and Technology, CN  
Authors:  
Shunzhuo Wang1, Fei Wu1, Zhonghai Lu1, You Zhou1, Qin Xiong1, Meng Zhang1 and Changsheng Xie1  
1Huazhong University of Science and Technology, CN; 2KTH Royal Institute of Technology, SE

**Abstract**  
With increasing density, NAND flash memory has decreasing reliability. Furthermore, raw bit error rate (RBER) of flash memory grows at an exponential rate as program/erase (P/E) cycle increases. Thus, error correction codes (ECCs), usually stored in the out-of-band area (OOB) of flash pages, are widely employed to ensure the reliability. However, the worstcase oriented ECC is largely under-utilized in the early stage, i.e. when P/E cycles are small, and the required ECC redundancy may be too large to be stored in the OOB. In this paper, we propose LAE-FTL, which employs a lifetimeadaptive ECC scheme, to improve the performance and lifetime of NAND flash memory. In the early stage, weak ECCs can guarantee the reliability and the OOB is large enough to store the ECCs. Thus, LAE-FTL employs weak ECCs and adaptively uses small and incremental codewords as P/E cycle increases to improve data transfer and decoding parallelism. In the late stage with large P/E cycles, strong ECCs are needed and the ECC redundancies become too large to fit in the OOB. Thus, LAE-FTL stores the exceeding ECC redundancies in the data space of flash pages and stores user data in a cross-page fashion. Finally, our evaluation results of trace-driven simulations show that LAE-FTL improves the read performance by up to 63.42%, compared to the worst-case oriented ECC scheme in the early stage, and significantly improve reliability of flash memory at low data accessing overhead in the late stage.

Download Paper (PDF; Only available from the DATE venue WiFi)

---

**Time** | **Label** | **Presentation Title** | **Authors**
|--------|-------|--------------------|---------|
| 10:02  | IP4-9 | 3D-DPE: A 3D HIGH-BANDWIDTH DOT-PRODUCT ENGINE FOR HIGH-PERFORMANCE NEUROMORPHIC COMPUTING | Miguel Lastras-Montaño, University of California, Santa Barbara, US  
Authors:  
Miguel Angel Lastras-Montaño1, Bhaswar Chakrabarti1, Dmitri B. Strukov1 and Kwang-Ting Cheng2  
1UC Santa Barbara, US; 2HKUST, HK

**Abstract**  
We present and experimentally validate 3D-DPE, a general-purpose dot-product engine, which is ideal for accelerating artificial neural networks (ANNs). 3D-DPE is based on a monolithically integrated 3D CMOS-memristor hybrid circuit and performs a high-dimensional dot-product operation (a recurrent and computationally expensive operation in ANNs) within a single step, using analog current-based computing. 3D-DPE is made up of two subsystems, namely a CMOS subsystem serving as the memory controller and an analog memory subsystem consisting of multiple layers of high-density memristive crossbar arrays fabricated on top of the CMOS subsystem. Their integration is based on a high-density area-distributed interface, resulting in much higher connectivity between the two subsystems, compared to the traditional interface of a 2D system or a 3D system integrated using through silicon vias. As a result, 3D-DPE’s single-step dot-product operation is not limited by the memory bandwidth, and the input dimension of the operations scales well with the capacity of the 3D memristive arrays. To demonstrate the feasibility of 3D-DPE, we designed and fabricated a CMOS memory controller and monolithically integrated 2 layers of titanium-oxide memristive crossbars. Then we performed the analog dot-product operation under different input conditions in two scenarios: (1) with devices within the same crossbar layer and (2) with devices from different layers. In both cases, the devices exhibited low voltage operation and analog switching behavior with high tuning accuracy.

Download Paper (PDF; Only available from the DATE venue WiFi)

---

**Time** | **Label** | **Presentation Title** | **Authors**
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>10:00</td>
<td></td>
<td>End of session</td>
<td></td>
</tr>
</tbody>
</table>

Coffee Break in Exhibition Area

On all conference days (Tuesday to Thursday), coffee and tea will be served during the coffee breaks at the below-mentioned times in the exhibition area.

Tuesday, March 28, 2017
- Coffee Break 10:30 - 11:30
- Coffee Break 16:00 - 17:00

Wednesday, March 29, 2017
- Coffee Break 10:00 - 11:00
- Coffee Break 16:00 - 17:00

Thursday, March 30, 2017
- Coffee Break 10:00 - 11:00
- Coffee Break 15:30 - 16:00
**Abstract**

The stream cipher ChaCha20 and the MAC function Poly1305 have been published as IETF RFC 7539. Since then, the industry is starting to use it more often. For example, it has been implemented by Google in their Chrome browser for TLS and also support has been added to OpenSSL, as well as OpenSSH. It is often claimed, that the algorithms are designed to be resistant to side-channel attacks. However, this is only true, if the only observable side-channel is the timing behavior. In this paper, we show that ChaCha20 is susceptible to power and EM side-channel analysis, which also translates to an attack on Poly1305, if used together with ChaCha20 for key generation. As a first countermeasure, we analyze the effectiveness of randomly shuffling the operations of the ChaCha round function.

Download Paper (PDF; Only available from the DATE venue WiFi)
This session features methods that extract desired implementation options from the huge design space of digital systems. The first talk presents a method to pick valuable operating points from a Pareto optimal set of task mappings for an efficient online resource management. The second presentation presents a rapid estimation framework to evaluate performance/area metrics of various accelerator options for an application at an early design phase. A design space exploration for implementing convolutional layers of operating points from a Pareto optimal set of task mappings for an efficient online resource management. The fourth talk presents an HLS scheduling method that is optimized for incorporating Radix 8 Booth multipliers. The session concludes with two short introductions of interactive presentations.

### 9.4 Design Space Exploration

**Date:** Thursday 30 March 2017  
**Time:** 08:30 - 10:00  
**Location / Room:** 3A

**Chair:** Lars Bauer, KIT Karlsruhe, DE  
**Co-Chair:** Alberto Del Barrio, Universidad Computense de Madrid, ES

This session features methods that extract desired implementation options from the huge design space of digital systems. The first talk presents a method to pick valuable operating points from a Pareto optimal set of task mappings for an efficient online resource management. The second presentation presents a rapid estimation framework to evaluate performance/area metrics of various accelerator options for an application at an early design phase. A design space exploration for implementing convolutional layers of neural networks is presented in the third talk in order to maximize the performance. The fourth talk presents an HLS scheduling method that is optimized for incorporating Radix 8 Booth multipliers. The session concludes with two short introductions of interactive presentations.

**Time | Label | Presentation Title**
--- | --- | ---
08:30 | 9.4.1 | AUTOMATIC OPERATING POINT DISTILLATION FOR HYBRID MAPPING METHODOLOGIES  
**Speaker:** Behnaz Pourmohseni, Friedrich-Alexander-Universität Erlangen-Nürnberg, DE  
**Authors:** Behnaz Pourmohseni¹, Michael Glaß² and Jürgen Teich¹  
¹Friedrich-Alexander-Universität Erlangen-Nürnberg, DE; ²Ulm University, DE  
**Abstract**  
Efficient execution of applications on heterogeneous many-core platforms requires mapping solutions that address different aspects of run-time dynamism like resource availability, energy budgets, and timing requirements. Hybrid mapping methodologies employ a static design space exploration (DSE) to obtain a set of mapping alternatives termed operating points that trade off quality properties (compute performance, energy consumption, etc.) and resource requirements (number of allocated resources of each type, etc.) among which one is selected at run-time by a run-time resource manager (RRM). Given multiple quality properties and the presence of heterogeneous resources, the DSE typically delivers a substantially large set of operating points handling of which may impose an intolerable run-time overhead to the RRM. This paper investigates the problem of truncation of operating points termed operating point distillation, such that (a) an acceptable run-time overhead is achieved, (b) on-line quality requirements are met, and (c) dynamic resource constraints are satisfied, i.e., application embeddability is preserved. We propose an automatic design-time distillation methodology that employs a hyper grid-based approach to retain diverse trade-off options wrt. quality properties, while selecting representative operating points based on their resource requirements to achieve a high level of run-time embeddability. Experimental results for a variety of applications show that compared to existing truncation approaches, proposed methodology significantly enhances the run-time embeddability while achieving a competitive and often improved efficiency in the distilled quality properties.

**Download Paper (PDF; Only available from the DATE venue WiFi)**

09:00 | 9.4.2 | DESIGN SPACE EXPLORATION OF FPGA-BASED ACCELERATORS WITH MULTI-LEVEL PARALLELISM  
**Speaker:** Guanwen Zhong, National University of Singapore, SG  
**Authors:** Guanwen Zhong¹, Alok Prakash², Siqi Wang¹, Yun (Eric) Liang², Tulika Mitra¹ and Smail Niar³  
¹National University of Singapore, SG; ²Nanyang Technological University, SG; ³Peking University, CN; ⁴LAMIH-University of Valenciennes, FR  
**Abstract**  
Applications containing compute-intensive kernels with nested loops can effectively leverage FPGAs to exploit fine- and coarse-grained parallelism. HLS tools used to translate these kernels from high-level languages (e.g., C/C++), however, are inefficient in exploiting multiple levels of parallelism automatically, thereby producing sub-optimal accelerators. Moreover, the large design space resulting from the various combinations of fine- and coarse-grained parallelism options makes exhaustive design space exploration prohibitively time-consuming with HLS tools. Hence, we propose a rapid estimation framework, MPSeeker, to evaluate performance/area metrics of various accelerator options for an application at an early design phase. Experimental results show that MPSeeker can rapidly (in minutes) explore the complex design space and accurately estimate performance/area of various design points to identify the near-optimal (95.7%) performance of the optimal on average) combination of parallelism options.

**Download Paper (PDF; Only available from the DATE venue WiFi)**
9.5 Modeling and optimization of Internet-of-things (IoT) devices

Date: Thursday 30 March 2017
Time: 08:30 - 10:00
Location / Room: 3C

Chair: William Fornaciari, Politecnico di Milano, IT

Co-Chair: Shusuke Yoshimoto, Osaka University, JP

Modeling and optimization of Internet-of-things (IoT) devices from energy sources to computing components including battery, energy harvesting system, power converter, and microprocessor

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>10:00</td>
<td>IP4-10</td>
<td>A SCHEDULABILITY TEST FOR SOFTWARE MIGRATION ON MULTICORE SYSTEMS</td>
</tr>
</tbody>
</table>

Speaker: Jung-Eun Kim, Department of Computer Science at the University of Illinois at Urbana-Champaign, US

Authors: Jung-Eun Kim, Richard Bradford 2, Tarek Abdelzaher 3 and Lui Sha 3

1Department of Computer Science, University of Illinois at Urbana-Champaign, US; 2Rockwell Collins, Cedar Rapids, IA, US; 3University of Illinois, US

Abstract
This paper presents a new schedulability test for safety-critical software undergoing a transition from single-core to multicore systems - a challenge faced by multiple industries today. Our migration model consists of a schedulability test and execution model. Its properties enable us to obtain a utilization bound that places an allowable limit on total task execution times. Evaluation results demonstrate the advantages of our scheduling model over competing resource partitioning approaches, such as Periodic Server and TDMA.

Download Paper (PDF; Only available from the DATE venue WiFi)
MEASUREMENT AND VALIDATION OF ENERGY HARVESTING IOT DEVICES

Authors:
Luks Sigrist, ETH Zurich, CH

Abstract
With the appearance of wearable devices and the IoT, energy harvesting nodes are becoming more and more important. The design and evaluation of these small standalone sensors and actuators, which harvest limited amounts of energy, requires novel tools and methods. Fast and accurate measurement systems are required to capture the rapidly changing harvesting scenarios and characterize leakage currents and energy efficiencies. The need for real-world experiments creates a demand for compact and portable equipment to perform in-situ power measurements and environmental logging. This work presents the RocketLogger, a hand-held measurement device that combines both portability and accuracy. The custom analog front-end allows logging at sampling rates up to 64 kSPS. The fast range switching within 1 us guarantees continuous power measurements starting from 4 pW at 1 mV up to 2.75 W at 5.5 V. The software provides remote control and manages data acquisition of up to 13 Mb/sec in real-time. We extensively characterize the RocketLogger's performance, demonstrate the need for its properties in three use-cases at different stages of the system design flow, and show its advantages in measuring and validating new harvesting-driven devices for the IoT.

Download Paper (PDF; Only available from the DATE venue WiFi)

A METHODOLOGY FOR THE DESIGN OF DYNAMIC ACCURACY OPERATORS BY RUNTIME BACK BIAS

Authors:
Daniele Zahier Pagliari, Politecnico di Torino, IT

Abstract
Mobile and IoT applications must balance increasing processing demands with limited power and cost budgets. Approximate computing achieves this goal leveraging the error tolerance features common in many emerging applications to reduce power consumption. In particular, adequate (i.e., energy/quality-configurable) hardware operators are key components in an error tolerant system. Existing implementations of these operators require significant architectural modifications, hence they are often design-specific and tend to have large overheads compared to accurate units. In this paper, we propose a methodology to design adequate datapath operators in an automatic way, which uses threshold voltage scaling as a knob to dynamically control the power/accuracy tradeoff. The method overcomes the limitations of previous solutions based on supply voltage scaling, in that it introduces lower overheads and it allows fine-grain regulation of this tradeoff. We demonstrate our approach on a state-of-the-art 28nm FDSOI technology, exploiting the strong effect of back biasing on threshold voltage. Results show a power consumption reduction of as much as 39% compared to solutions based only on supply voltage scaling, at iso-accuracy.

Download Paper (PDF; Only available from the DATE venue WiFi)

A SCAN-CHAIN BASED STATE RETENTION METHODOLOGY FOR IOT PROCESSORS OPERATING ON INTERMITTENT ENERGY

Authors:
Pascal Alexander Hager, ETH Zurich, CH

Abstract
Future IoT systems are tightly constrained by cost and size and will often be operated from an energy harvester's output. Since these batteryless systems operate on intermittent energy they have to be able to retain their state during the power outages in order to guarantee computation progress. Due to the lack of large energy buffers the state needs to be saved quickly using residual energy only. In related work, the state is retained in-place by replacing all flip-flops with state retentive flip-flops (SRFF), which are powered by auxiliary supplies for retention or incorporate non-volatile memory cells. However, these SRFFs increase the power consumption during active operation impairing the overall systems efficiency. In this paper, we present a scan-chain based state retention approach, where the state is moved to memory using only 4.5pJ/b. Since our approach does not introduce any power overhead, this energy cost is negligible. SRFFs increase the power consumption during active operation impairing the overall systems efficiency. In this paper, we present a scan-chain based state retention approach, where the state is moved to memory using only 4.5pJ/b. Since our approach does not introduce any power overhead, this energy cost is negligible. We demonstrate our approach on a state-of-the-art 28nm FDSOI technology, exploiting the strong effect of back biasing on threshold voltage. Results show a power consumption reduction of as much as 39% compared to solutions based only on supply voltage scaling, at iso-accuracy.

Download Paper (PDF; Only available from the DATE venue WiFi)

A CIRCUIT-EQUIVALENT BATTERY MODEL ACCOUNTING FOR THE DEPENDENCY ON LOAD FREQUENCY

Authors:
Yukai Chen, Politecnico di Torino, IT

Abstract
Circuit-equivalent battery models are considered de-facto standard for modeling and simulation of digital systems due to many practical advantages. In spite of the many variants of models proposed in the literature, none of them accounts for one important feature of the battery dynamics, namely, the dependency on the frequency of current load profile. For a given average current value, current loads with different spectral distributions may have quite different impacts on the battery discharge. This is a very well-know issue in the design of hybrid energy storage systems, where different types of storage devices are used, each with different storage efficiency for different load frequency ranges. We propose a basic modification to a state-of-the-art model that incorporates this load frequency dependency, as well as a methodology to identify the frequency-sensitive parameters of the model from publicly available data (e.g., datasheets). Results show that frequency-agnostic models can significantly overestimate the battery state-of-charge, and that this effect is far from being negligible.

Download Paper (PDF; Only available from the DATE venue WiFi)

ADAPTIVE POWER DELIVERY SYSTEM MANAGEMENT FOR MANY-CORE PROCESSORS WITH ON/OFF-CHIP VOLTAGE REGULATORS

Authors:
Haoran Li, The Hong Kong University of Science and Technology, HK

Abstract
The power delivery system (PDS) plays a crucial role of guaranteeing the proper functionality of many-core processors. However, as PDS is usually optimized to provide power to the target chip at its best performance level, its energy efficiency can be seriously degraded under highly dynamic workloads, making it a major source of system power losses. On-chip voltage regulators (VR), which are able to achieve fast and fine-grained power control, have been proposed as a choice for PDS implementation and provided design opportunities for improving system energy efficiency. In this paper, we propose the adaptive Quantized Power Management (QPM) scheme to dynamically adjust the PDS with both on-chip and off-chip VRs based on run-time workloads. Experimental results on different applications show that QPM applied on a hybrid PDS with both on/off-chip voltage regulators (VR) achieves 74.1% average overall energy efficiency, 12.3% higher than the conventional PDS with single off-chip VR.

Download Paper (PDF; Only available from the DATE venue WiFi)
The last paper discusses methods based on machine learning for timing error detection. The first two papers discuss optimizations for yield and performances of analog circuits. The third paper proposes methods for flip-flop soft error protection in sequential circuits. 

Switched-capacitor voltage regulators (SCVRs) are widely used in on-chip power management, due to high step-down efficiency and feasibility of integration. In this work, we present theoretical analysis and optimization methodology for flying and decoupling capacitance values for area-constrained on-chip SCVRs to achieve the highest system-level power efficiency. The proposed models for efficiency and droop voltage are validated with on-chip 2:1 SCVR implementations in both 65nm and 32nm CMOS, which show high model accuracy. The maximum and average error of the predicted optimal ratio between flying and decoupling capacitance are 5% and 1.7%, respectively.

Download Paper (PDF; Only available from the DATE venue WiFi)

In this work, we present theoretical analysis and optimization methodology for flying and decoupling capacitance values for area-constrained on-chip SCVRs to achieve the highest system-level power efficiency. The proposed models for efficiency and droop voltage are validated with on-chip 2:1 SCVR implementations in both 65nm and 32nm CMOS, which show high model accuracy. The maximum and average error of the predicted optimal ratio between flying and decoupling capacitance are 5% and 1.7%, respectively.

Download Paper (PDF; Only available from the DATE venue WiFi)

Switched-capacitor voltage regulators (SCVRs) are widely used in on-chip power management, due to high step-down efficiency and feasibility of integration. In this work, we present theoretical analysis and optimization methodology for flying and decoupling capacitance values for area-constrained on-chip SCVRs to achieve the highest system-level power efficiency. The proposed models for efficiency and droop voltage are validated with on-chip 2:1 SCVR implementations in both 65nm and 32nm CMOS, which show high model accuracy. The maximum and average error of the predicted optimal ratio between flying and decoupling capacitance are 5% and 1.7%, respectively.

Download Paper (PDF; Only available from the DATE venue WiFi)

Dynamic timing errors (DTEs), that are caused by the timing violations of sensitized critical timing paths, have emerged as an important threat to the reliability of digital circuits. Existing approaches model the DTEs without considering the impact of input operands on dynamic path sensitization, resulting in loss of accuracy. The diversity of input operands leads to complex path sensitization behaviors, making it hard to represent in DTE modeling. In this paper, we propose SLoT, a supervised learning model to predict the output of functional units (FUs) to be one of two timing classes: {timing correct, timing erroneous}, as a function of input operands and clock period. We apply random forest classification (RFC) method to construct SLoT, by using input operands, computation history and circuit toggling as input features and outputs' timing classes as labels. The outputs timing classes are measured using gate-level simulation (GLS) of a post place-and-route design in TSMC 45nm process. For evaluation, we apply SLoT to several FUs and on average 95% predictions are consistent with GLS, which is 6.3X higher compared to the existing instruction-level model. SLoT-based reliability analysis of FUs under different benchmark datasets can achieve 0.7-4.8% average difference compared with GLS-based analysis, and execute more than 20X faster than GLS.

Download Paper (PDF; Only available from the DATE venue WiFi)

Switched-capacitor voltage regulators (SCVRs) are widely used in on-chip power management, due to high step-down efficiency and feasibility of integration. In this work, we present theoretical analysis and optimization methodology for flying and decoupling capacitance values for area-constrained on-chip SCVRs to achieve the highest system-level power efficiency. The proposed models for efficiency and droop voltage are validated with on-chip 2:1 SCVR implementations in both 65nm and 32nm CMOS, which show high model accuracy. The maximum and average error of the predicted optimal ratio between flying and decoupling capacitance are 5% and 1.7%, respectively.

Download Paper (PDF; Only available from the DATE venue WiFi)

Dynamic timing errors (DTEs), that are caused by the timing violations of sensitized critical timing paths, have emerged as an important threat to the reliability of digital circuits. Existing approaches model the DTEs without considering the impact of input operands on dynamic path sensitization, resulting in loss of accuracy. The diversity of input operands leads to complex path sensitization behaviors, making it hard to represent in DTE modeling. In this paper, we propose SLoT, a supervised learning model to predict the output of functional units (FUs) to be one of two timing classes: {timing correct, timing erroneous}, as a function of input operands and clock period. We apply random forest classification (RFC) method to construct SLoT, by using input operands, computation history and circuit toggling as input features and outputs' timing classes as labels. The outputs timing classes are measured using gate-level simulation (GLS) of a post place-and-route design in TSMC 45nm process. For evaluation, we apply SLoT to several FUs and on average 95% predictions are consistent with GLS, which is 6.3X higher compared to the existing instruction-level model. SLoT-based reliability analysis of FUs under different benchmark datasets can achieve 0.7-4.8% average difference compared with GLS-based analysis, and execute more than 20X faster than GLS.

Download Paper (PDF; Only available from the DATE venue WiFi)

Switched-capacitor voltage regulators (SCVRs) are widely used in on-chip power management, due to high step-down efficiency and feasibility of integration. In this work, we present theoretical analysis and optimization methodology for flying and decoupling capacitance values for area-constrained on-chip SCVRs to achieve the highest system-level power efficiency. The proposed models for efficiency and droop voltage are validated with on-chip 2:1 SCVR implementations in both 65nm and 32nm CMOS, which show high model accuracy. The maximum and average error of the predicted optimal ratio between flying and decoupling capacitance are 5% and 1.7%, respectively.

Download Paper (PDF; Only available from the DATE venue WiFi)
Abstract
Starting from a set of starting points, the multiple-starting-point optimization searches the local optima by gradient-guided local search. The global optimum is selected from these local optima. The region-hit property of the multiple-starting-point optimization makes the multiple-starting-point approach more likely to reach the global optimum. However, for non-smooth objective functions, e.g., worst-case optimization, the traditional gradient-based local search methods may get stuck at non-optimal points, even if the objective function is smooth. In this paper, we propose a subgradient-based multiple-starting-point algorithm for non-smooth optimization of analog circuits. Subgradients instead of traditional gradients are used to guide the local search of the non-smooth optimization. The Shor's R algorithm is used to accelerate the subgradient-based local search. A two-stage optimization strategy is proposed to deal with the constraints in analog circuit optimization. Our experiments on 2 circuits show that the proposed method is very efficient for worst-case optimization. The proposed approach can achieve much better solutions with less simulations, compared with the traditional gradient-based method, smoothing approximation method, smooth relaxation method and differential evolution algorithms.

Download Paper (PDF; Only available from the DATE venue WiFi)
NOISE-SENSITIVE FEEDBACK LOOP IDENTIFICATION IN LINEAR TIME-VARYING ANALOG CIRCUITS

Speaker:
Peng Li, Texas A&M University, US

Authors:
Ang Li, Peng Li, Tingwen Huang and Edgar Sánchez-Sinencio
1Texas A&M University, US; 2Texas A&M University at Qatar, QA

Abstract
The continuing scaling of VLSI technology and design complexity has rendered robustness of analog circuits a significant concern. Parasitic effects may introduce unexpected marginal instability within multiple noise-sensitive loops and hence jeopardize circuit operation and processing precision. The Loop Finder algorithm has been recently proposed to allow detection of noise-sensitive return loops for circuits that are described using a linear time-invariant (LTI) system model. However, many practical circuits such as switched-capacitor filters and mixers present time-varying behaviors which are intrinsically coupled with noise propagation and introduce new noise generation mechanisms. For the first time, we take an in-depth look into the marginal instability of linear periodically time-varying (LPTV) analog circuits and further develop an algorithm for efficient identification of noise-sensitive loops, unifying the solution to noise sensitivity analysis for both LTI and LPTV circuits.

Download Paper (PDF; Only available from the DATE venue WiFi)

10:00
End of session

Coffee Break in Exhibition Area

On all conference days (Tuesday to Thursday), coffee and tea will be served during the coffee breaks at the below-mentioned times in the exhibition area.

Tuesday, March 28, 2017
- Coffee Break 10:30 - 11:30
- Coffee Break 16:00 - 17:00

Wednesday, March 29, 2017
- Coffee Break 10:00 - 11:00
- Coffee Break 16:00 - 17:00

Thursday, March 30, 2017
- Coffee Break 10:00 - 11:00
- Coffee Break 15:30 - 16:00

9.7 Front-row seats for Temperature and Variability

Date: Thursday 30 March 2017
Time: 08:30 - 10:00
Location / Room: 3B

Chair:
Marina Zapater Sancho, EPFL, CH

Co-Chair:
Giovanni Ansaloni, USI, CH

This session sets highlights on the impact of temperature and variability sources at the overall system level. Firstly, an approach that incorporates leakage in thermal simulations is presented and a thermal simulation framework is devised. After that, temperature is not only estimated but also minimized in the context of global interconnects by means of an analytic methodology. Finally, timing variability plays its role in the session and its effects in variable-latency designs.
Coffee Break 10:00 - 11:00

10:30 | IP4-18, 57 | POWER PRE-CHARACTERIZED MESHING ALGORITHM FOR FINITE ELEMENT THERMAL ANALYSIS OF INTEGRATED CIRCUITS

Speaker: Shohdy Abdelkader, Software Developer, EG
Authors: Shohdy Abdelkader 1, Alaa ElRouby 2 and Mohamed Dessouky 1
1Mentor, EG; 2Electric and Electronic Department, Faculty of Engineering and Natural Science, Yildirim Beyazit University, TR

Abstract
In this paper we present an adaptive meshing technique suitable for steady state finite element (FE) based thermal analysis of integrated circuits (ICs). The algorithm presented is a non iterative one where the technology used is first pre-characterized. The characterization results are then used for scanning the layout to detect high power regions then fine meshing them. Finally, the analysis is done only once. This makes it faster than conventional iterative adaptive meshing methods. The algorithm results showed comparable accuracy and better performance when compared to the flux based (iterative) and the power aware (non iterative) algorithms.

Download Paper (PDF; Only available from the DATE venue WiFi)

10:45 | IP4-19, 58 | A THERMALLY-AWARE ENERGY MINIMIZATION METHODOLOGY FOR GLOBAL INTERCONNECTS

Speaker: Afzali Kusha, Tehran University, IR
Authors: Soheil Nazar Shahsavani 1, Alireza Shafaei Bejestan 1, Shahin Nazarian 1 and Massoud Pedram 2
1University of Southern California, US; 2JUSC, US

Abstract
As a result of the Temperature Effect Inversion (TEI) in FinFET-based designs, gate delays decrease with the increase of temperature. In contrast, the resistive characteristic and hence delay of global interconnects increase with the temperature. However, as shown in this paper, if buffers are judiciously inserted in global interconnects, the buffer delay decrease is more pronounced than the interconnect delay increase, resulting in an overall performance improvement at higher temperatures. More specifically, this work models the delay of buffer-inserted global interconnects vs. temperature in order to derive the optimal number and size of buffers for a given interconnect length and temperature. Furthermore, the paper addresses the problem of minimizing the buffered interconnect energy consumption by changing the supply voltage level or FinFET threshold voltage, and also presents a temperature-aware optimization policy for solving this problem. Simulation results show average interconnect energy savings of 16% with no performance penalty for five different benchmarks implemented on a 14nm FinFET technology.

Download Paper (PDF; Only available from the DATE venue WiFi)
Today everything from the door locks, a heating system or vehicle can be connected to internet opening the endless possibilities of future innovative technologies. As more low-power and internet-connected gadgets and sensors are integrated to our lives, an increase in demand for developing secure and trustworthy IoT-based systems is becoming the key element to make winning products.

Although, there has been a steady increase in improving the security, still proper authentication and encrypted communications are not common; making the overall Internet as a network of insecure things. This session proposes a journey through several speeches to show the advances in technologies that master the security aspects of IoT.

The session starts with an in-depth overview of security challenges and the trends in the IoT ecosystem against cyber-threats. Then, introduces the STM32 and the secure IoT platforms based on STM32 called SECube. Finally, the session provides some real use cases for smart vehicle, where IoT have a big impact on the type of applications and services that can be deployed using the association between vehicle and the homes of their owners. Last but not least, all the pre-registered attendees are eligible to get one of IoT platforms presented by the speakers via the www.secube.eu web site.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Speaker</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:30</td>
<td>9.8.1</td>
<td>CHALLENGES FOR SECURE IOT</td>
<td>Paolo Prinette, Politecnico di Torino, IT</td>
<td></td>
</tr>
<tr>
<td>08:45</td>
<td>9.8.2</td>
<td>MITIGATING THE RISKS IN IOT WITH AN EFFECTIVE SECURITY OFFER</td>
<td>Michele Scarlatella, STMicroelectronics, FR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>The IoT will change our lives, bringing huge benefits and making a positive impact on society and the economy, but it requires trusted systems with efficient security and privacy mechanisms from devices to the Cloud. For years digital security technologies have proven their efficiency in telecom, banking and ID applications. Technical solutions exist and can be reused as a toolbox to provide security and privacy for the IoT. In this session we will describe how STMicroelectronics’ scalable security offer based on STM32 microcontrollers and STSAFE secure microcontrollers make it possible to build secure IoT solutions with the right level of robustness. The STMicroelectronics scalable offer for IoT security can also be adapted to efficiently combat various threats. STMicroelectronics, a global semiconductor leader supplying the market with the most advanced technologies and solutions and a 20-year presence in security, is committed to contributing to a more secure connected world.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>09:00</td>
<td>9.8.3</td>
<td>UNIVERSITY EXPERIENCES USING A SECURE IOT PLATFORM BASED ON STM32</td>
<td>George Kornarors, Univ. of Applied Sciences of Crete, GR</td>
<td></td>
</tr>
<tr>
<td>09:15</td>
<td>9.8.4</td>
<td>SECURE™: THE SECURE COMMERCIAL IOT PLATFORM</td>
<td>Antonio Varriale, Blu5 Labs Ltd, MT</td>
<td></td>
</tr>
<tr>
<td>09:30</td>
<td>9.8.5</td>
<td>SECURE COMMUNICATION IN AUTOMOTIVE</td>
<td>Giovanni Gherardi, Energica Motor Company, IT</td>
<td></td>
</tr>
</tbody>
</table>
On all conference days (Tuesday to Thursday), coffee and tea will be served during the coffee breaks at the below-mentioned times in the exhibition area.

Tuesday, March 28, 2017
- Coffee Break 10:30 - 11:30
- Coffee Break 16:00 - 17:00

Wednesday, March 29, 2017
- Coffee Break 10:00 - 11:00
- Coffee Break 16:00 - 17:00

Thursday, March 30, 2017
- Coffee Break 10:00 - 11:00
- Coffee Break 15:30 - 16:00

### IP4 Interactive Presentations

**Date:** Thursday 30 March 2017  
**Time:** 10:00 - 10:30  
**Location / Room:** IP sessions (in front of rooms 4A and 5A)

Interactive Presentations run simultaneously during a 30-minute slot. A poster associated to the IP paper is on display throughout the morning. Additionally, each IP paper is briefly introduced in a one-minute presentation in a corresponding regular session, prior to the actual Interactive Presentation. At the end of each afternoon Interactive Presentations session the award 'Best IP of the Day' is given.

<table>
<thead>
<tr>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
</table>
| [IP4-1] | **1024-CHANNEL 3D ULTRASOUND DIGITAL BEAMFORMER IN A SINGLE 5W FPGA**               | Aya Ibrahim, EPFL, CH  
Authors: Federico Angiolini, Aya Ibrahim, William Simon, Ahmet Caner Yüzügüler, Marcel Arditi, Jean-Philippe Thiran and Giovanni De Micheli  
1EPFL, CH; 2École Polytechnique Fédérale de Lausanne (EPFL), CH  
**Abstract**  
3D ultrasound, an emerging medical imaging technology that is presently only used in hospitals, has the potential to enable breakthrough telemedicine applications; provided that its cost and power dissipation can be minimized. In this paper, we present an FPGA architecture suitable for a portable medical 3D ultrasound device. We show an optimized design for the digital part of the imager, including the delay calculation block, which is its most critical part. Our computationally efficient approach requires a single FPGA for 3D imaging, which is unprecedented. The design is scalable; a configuration supporting a 32×32 channel probe, which enables high-quality imaging, consumes only about 5W.  
Download Paper (PDF; Only available from the DATE venue WiFi) |
| [IP4-2] | **LAANT: A LIBRARY TO AUTOMATICALLY OPTIMIZE EDP FOR OPENMP APPLICATIONS**           | Arthur Lorenzon, Jeckson Delligostin Souza and Antonio Carlos Schneider Beck Filho, Universidade Federal do Rio Grande do Sul, BR  
**Abstract**  
Efficiently exploiting thread level parallelism from new multicore systems has been challenging for software developers. While blindly increasing the number of threads may result in disproportionate increase in energy consumption. For this reason, rightly choosing the number of threads is essential to reach the best compromise between both. However, such task is extremely difficult: besides the huge number of variables involved, many of them will change according to different aspects of the system at hand and are only possible to be defined at run-time. To address this complex scenario, we propose LAANT, a novel library to automatically find the optimal number of threads for OpenMP applications, by dynamically considering their particular characteristics, input set, and the processor architecture. By executing nine well-known benchmarks on three real multicore processors, LAANT improves the EDP (Energy-Delay Product) by up to 61%, compared to the standard OpenMP execution; and by 44%, when the dynamic adjustment of the number of threads of OpenMP is activated.  
Download Paper (PDF; Only available from the DATE venue WiFi) |
| [IP4-3] | **IMPROVING THE ACCURACY OF THE LEAKAGE POWER ESTIMATION OF EMBEDDED CPUs**          | Shiao-Li Tsao, National Chiao Tung University, TW  
**Abstract**  
Previous studies have used on-chip thermal sensors (diodes) to estimate the leakage power of a CPU. However, an embedded CPU equips only a few thermal sensors and may suffer from considerable spatial temperature variances across the CPU core, and leakage power estimation based on insufficient temperature information introduces errors. According to our experiments, the conventional leakage power models may have up to 22.9% estimation error for a 70-nm embedded CPU. In this study, we first evaluated the accuracy of leakage power estimates based on thermal sensors on different locations of a CPU and suggested locations that can reduce the error to 0.9%. Then, we proposed temperature-referred and counter-tracked estimation (TRACE) that relies on temperature sensors and hardware activity counters to estimate leakage power. The simulation results demonstrated that employing TRACE could reduce the error to 3.4%. Experiments were also conducted on a real platform to verify our findings.  
Download Paper (PDF; Only available from the DATE venue WiFi) |
LIFETIME ADAPTIVE ECC IN NAND FLASH PAGE MANAGEMENT

Speaker:
Shunhuo Wang, Huazhong University of Science and Technology, CN
Authors:
Shunhuo Wang1, Fei Wu1, Zhonghai Lu2, You Zhou1, Qin Xiong1, Meng Zhang1 and Changsheng Xie1
1Huazhong University of Science and Technology, CN; 2KTH Royal Institute of Technology, SE

Abstract
With increasing density, NAND flash memory has decreasing reliability. Furthermore, raw bit error rate (RBER) of flash memory grows at an exponential rate as program/erase (P/E) cycle increases. Thus, error correction codes (ECCs), usually stored in the out-of-band area (OOB) of flash pages, are widely employed to ensure the reliability. However, the worst case oriented ECC is largely under-utilized in the early stage, i.e. when P/E cycles are small, and the required ECC redundancy may be too large to be stored in the OOB. In this paper, we propose LAE-FTL, which employs a lifetime-adaptive ECC scheme, to improve the performance and lifetime of NAND flash memory. In the early stage, weak ECCs can guarantee the reliability and the OOB is large enough to store the ECCs. Thus, LAE-FTL employs weak ECCs as deduplication, lossless compression, and delta compression, are rapidly adopted to SSDs. Although each technique has been extensively studied, how to efficiently combine these techniques for maximizing their synergistic effects is not investigated well. In this paper, we propose a novel dedup-assisted compression (DAC) scheme that integrates existing data reduction techniques so that potential benefits of individual ones can be maximized while overcoming their inherent limitations. By doing so, DAC greatly reduces the amount of write traffic sent to SSDs. DAC also requires negligible hardware resources by utilizing existing hardware modules. Our evaluation results show that the proposed DAC decreases the amount of written data by up to 30% over a simple integration reduplication and lossless compression.

Download Paper (PDF; Only available from the DATE venue WiFi)
A NEW SAMPLING TECHNIQUE FOR MONTE CARLO-BASED STATISTICAL CIRCUIT ANALYSIS

Speaker:
Miguel Lastras-Montaño, University of California, Santa Barbara, US

Authors:
Miguel Angel Lastras-Montaño1, Bhawar Chakrabarti1, Dmitri B. Strukov1 and Kwang-Ting Cheng2

1UC Santa Barbara, US; 2HKUST, HK

Abstract
We present and experimentally validate 3D-DPE, a general-purpose dot-product engine, which is ideal for accelerating artificial neural networks (ANNs). 3D-DPE is based on a monolithically integrated 3D CMOS-memristor hybrid circuit and performs a high-dimensional dot-product operation (a recurrent and computationally expensive operation in ANNs) within a single step, using analog current-based computing. 3D-DPE is made up of two subsystems, namely a CMOS subsystem serving as the memory controller and an analog memory subsystem consisting of multiple layers of high-density memristive crossbar arrays fabricated on top of the CMOS subsystem. Their integration is based on a high-density area-distributed interface, resulting in much higher connectivity between the two subsystems, compared to the traditional interface of a 2D system or a 3D system integrated using through silicon vias. As a result, 3D-DPE’s single-step dot-product operation is not limited by the memory bandwidth, and the input dimension of the operations scales well with the capacity of the 3D memristive arrays. To demonstrate the feasibility of 3D-DPE, we designed and fabricated a CMOS memory controller and monolithically integrated 2 layers of titanium-oxide memristive crossbars. Then we performed the analog dot-product operation under different input conditions in two scenarios: (1) with devices within the same crossbar layer and (2) with devices from different layers. In both cases, the devices exhibited low voltage operation and analog switching behavior with high tuning accuracy.

Download Paper (PDF; Only available from the DATE venue WiFi)
This paper addresses the problem of automatic technology migration of analog IC designs. The proposed approach introduces a new level of abstraction, for EDA tools addressing analog IC design, allowing a systematic and effortless adaption of a design to a new technology. The new abstraction level is based on generic cell libraries, which includes topology and testbenches descriptions for specific circuit classes. The new approach is implemented and tested using a state-of-the-art multi-objective multi-constraint-circuit-level optimization tool, and is validated for the sizing and optimization of continuous-time comparators, including technology migration between two different design nodes, respectively, XFAB 350 nm technology (XH035) and ATNEL 150 nm SOI technology (AT77K).

Download Paper (PDF; Only available from the DATE venue WiFi)
Abstract

Genomics, molecular dynamics, and machine learning are just the most recent examples of fields where FPGAs could provide the means to achieve interesting breakthroughs. However, HDL programming requires considerable multi-disciplinary skills, experience, large budgets, time, and a bit of wizardry. Given that most implementations are short-lived, the investment simply does not pay off. In this demo we propose a multi-vendor LLVM-based automated framework that can transparently - without the user or developer being aware of it - offload computing-intensive code fragments to FPGAs. The system relies on a performance monitor to detect computing-intensive code sections and, if they are suitable for offloading, extracts the Data Flow Graph and uses it to program an overlay pre-programmed on the FPGA, which then interacts with the Just-In-Time compiler executing the program. The overall process requires hundreds of microseconds, and can be easily reverted should the outcome be unsatisfactory.

More information ...
UB09.8 TIDES: NON-LINEAR WAVEFORMS FOR QUICK TRACE NAVIGATION

Presenter:
Jannis Stoppe, University of Bremen, DE

Author:
Ralf Drechsler, University of Bremen / DFKI, DE

Abstract
System trace analysis is mostly done using waveform viewers -- tools that relate signals and their assignments at certain times. While generic hardware design is subject to some innovative visualisation ideas and software visualisation has been a research topic for much longer, these classic tools have been part of the design process since the earlier days of hardware design -- and have not changed much over the decades. Instead, the currently available programs have evolved to look practically the same, all following a familiar pattern that has not changed since their initial appearance. We argue that there is still room for innovation beyond the very classic waveform display though. We implemented a proof-of-concept waveform viewer (codenamed Tides) that has several unique features that go beyond the standard set of features for waveform viewers.

More information ...

UB09.9 HEPSYCODE: A SYSTEM-LEVEL METHODOLOGY FOR HW/SW CO-DESIGN OF HETEROGENEOUS PARALLEL DEDICATED SYSTEMS

Presenter:
Luigi Pomante, University of L'Aquila, IT

Authors:
Giacomo Valente 1, Vittoriano Muttillo 1, Daniele Di Pompeo 1, Emilio Incerto 2 and Daniele Ciambrotte 1
1University of L'Aquila, IT; 2Gran Sasso Science Institute, IT

Abstract
Heterogeneous parallel systems have been recently-exploited for a wide range of application domains, for both the dedicated (e.g. embedded) and the general purpose products. Such systems can include different processor cores, memories, dedicated ICS and a set of connections between them. They are so complex that the design methodology plays a major role in determining the success of the products. So, this demo addresses the problem of the electronic system-level hw/sw co-design of heterogeneous parallel dedicated systems. In particular, it shows an enhanced CSP/SystemC-based design space exploration step (and related ESL-EDA prototype tools), in the context of an existing hw/sw co-design flow that, given the system specification and related F/NF requirements, is able to (semi)automatically propose to the designer: - a custom heterogeneous parallel architecture; - an HW/SW partitioning of the application; - a mapping of the partitioned entities onto the proposed architecture.

More information ...

UB09.10 PULP: A ULTRA-LOW POWER PLATFORM FOR THE INTERNET-OF-THINGS

Presenter:
Francesco Conti, ETH Zurich, CH

Authors:
Stefan Mach 1, Florian Zaruba 1, Antonio Pullini 1, Daniele Palossi 1, Giovanni Rovere 1, Florian Glaser 1, Germain Haugou 1, Schekeb Fateh 1 and Luca Benini 2
1ETH Zurich, CH; 2ETH Zurich, CH and University of Bologna, IT

Abstract
The PULP (Parallel Ultra-Low Power) platform strives to provide high performance for IoT nodes and endpoints within a very small power envelope. The PULP platform is based on a tightly-coupled multi-core cluster and on a modular architecture, which can support complex configurations with autonomous I/O without SW intervention, HW-accelerated execution of hot computation kernels, fine-grain event-based computation - but can also be deployed in very simple configuration, such as the open source PULPino microcontroller. In this demonstration booth, we will showcase several prototypes using PULP chips in various configuration. Our prototypes perform demos such as real-time deep-learning based visual recognition from a low-power camera, and online biosignal acquisition and reconstruction on the same chip. Application scenarios for our technology include healthcare wearables, autonomous nano-UAVs, smart networked environmental sensors.

More information ...

10.1 Wearable and Smart Medical Devices Day: Diagnosis and prevention systems

Date: Thursday 30 March 2017
Time: 11:00 - 12:30
Location / Room: SBC

Organisers:
José L. Ayala, Universidad Complutense de Madrid, ES
Chris Van Hoof, IMEC, BE

Chair:
Olivier Romain, Université de Cergy-Pontoise, FR

Co-Chair:
Mario Konijnenburg, IMEC, BE

This session will present novel approaches, techniques and devices for the improvement of diagnosis and prevention systems. Improved bioanalytics-on-chip designs, wearables in the prevention of elderly, computational mechanisms for prevention of symptoms, and bioelectronics medicines will be covered.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>11:00</td>
<td>10.1.1</td>
<td>ENABLING TECHNOLOGIES FOR NEXT GENERATION BIOANALYTICS ON CHIP</td>
</tr>
</tbody>
</table>

Author:
Carlota Guiducci, EPFL, CH

Abstract
The adoption of lab-on-chip based solutions in clinical practice and in the framework of the most common bioanalytics protocols has long been sought for the possibility to fine control the movement of fluids and the flow of molecules and particles. Nevertheless, the existing solutions inherently limit both throughput and the possibility to sense and manipulate single particles. A few years ago, we undertook a major challenge in this context, starting from the consideration that the lack of solutions to localize electric fields in micro-regions and to control their distribution over the height of the chambers fundamentally limited the efficiency and the scalability of these systems. Our strategy, based on monolithic process, results in highly conductive and singularly addressable vertical microelectrodes, fully integrated in high aspect-ratio microfluidics. We have applied this novel process to develop a new generation of microfluidic flow cytometers that could successfully detect, for the first time, activated T lymphocytes in a cellular sample. In this talk we will describe as well our contribution to the integration of biosensors on IC layers and to solve the issues related to the specific surface treatments involved in the analytical protocol.
<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>10:12</td>
<td>10.1.2</td>
<td>CARBON NANOTUBES ENABLE MAJOR ENERGY EFFICIENCY BENEFITS FOR SUB-10NM DIGITAL SYSTEMS</td>
</tr>
<tr>
<td>10:13</td>
<td>10.1.3</td>
<td>AN OPTIMAL APPROACH FOR LOW-POWER MIGRAINE PREDICTION MODELS IN THE STATE-OF-THE-ART WIRELESS MONITORING DEVICES</td>
</tr>
<tr>
<td>10:14</td>
<td>10.1.4</td>
<td>WEARABLE ELECTRONICS - WHAT IS IT GOOD FOR - AND WHAT IS MISSING TO SUPPORT THE QUALITY OF LIFE OF ELDERLY PEOPLE?</td>
</tr>
</tbody>
</table>

**10.2 Hot Topic Session: EDA as an Emerging Technology Enabler**

**Date:** Thursday 30 March 2017  
**Time:** 11:00 - 12:30  
**Location / Room:** 4BC  
**Organisers:**  
Pierre-Emmanuel Gaillardon, The University of Utah at Salt Lake City, US  
Mathias Soeken, EPFL, CH  
**Chair:**  
Mathias Soeken, EPFL, CH  
**Co-Chair:**  
Ian O'Connor, Ecole Centrale de Lyon, FR

In this Hot Topic Session, we demonstrate how design automation enables emerging technologies. Four talks will be provided. The first talk will review how logic synthesis has and is still enabling today's technologies and will also outline the requirements of design automation for the technologies of tomorrow. The other three talks present several emerging technologies, such as carbon nanotubes, spin wave devices, quantum-dot cellular automata, nanomagnetic logic and quantum computing, and illustrate how design automation is still enabling today's technologies and will also outline the requirements of design automation for the technologies of tomorrow.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>11:00</td>
<td>10.2.1</td>
<td>LOGIC OPTIMIZATION AND SYNTHESIS: TRENDS AND DIRECTIONS IN INDUSTRY</td>
</tr>
<tr>
<td>11:22</td>
<td>10.2.2</td>
<td>CARBON NANOTUBES ENABLE MAJOR ENERGY EFFICIENCY BENEFITS FOR SUB-10NM DIGITAL SYSTEMS</td>
</tr>
</tbody>
</table>

**Authors**

1. Luca Amaru, Synopsys Inc., US  
2. Luca Amaru, Patrick Vuillod, Jiong Luo and Janet Olson, Synopsys, US  
3. Gage Hills, Stanford University, US  

**Abstract**

Logic synthesis is a key design step which optimizes abstract circuit representations and links them to technology. With CMOS technology moving into the deep nanometer regime, logic synthesis needs to be aware of physical informations early in the flow. With the rise of enhanced functionality nanodevices, research on technology needs the help of logic synthesis to capture advantageous design opportunities. This paper deals with the synergy between logic synthesis and technology, from an industrial perspective. First, we present new synthesis techniques which embed detailed physical informations at the core optimization engine. Experiments show improved Quality of Results (QoR) and better correlation between RTL synthesis and physical implementation. Second, we discuss the application of these new synthesis techniques in the early assessment of emerging nanodevices with enhanced functionality. Finally, we argue that new synthesis methods can push further the progress of electronics, as we have reached a multiforking point of technology where choices are tougher than ever.

**Authors**

1. Department of Electrical Engineering, Stanford University, US;  
2. IMEC, BE;  
3. IMEC, BE;  
4. imec, vzw, BE;  
5. IMEC, BE;  
6. imec, BE;  
7. NU Singapore, SG
### 10.2.3 WAVE PIPELINING FOR MAJORITY-BASED BEYOND-CMOS TECHNOLOGIES

**Speaker:** Odysseas Zografos, imec, BE  
**Authors:** Odysseas Zografos1, Anton De Meester1, Eleonora Testa2, Mathias Soeken2, Pierre-Emmanuel Gaillardon3, Giovanni De Micheli3, Luca Amaru3, Praveen Raghavani3, Francy Catthoor1 and Rudy Lauwereins1  
1IMEC, BE; 2EPFL, CH; 3University of Utah, US; 4École Polytechnique Fédérale de Lausanne (EPFL), CH; 5Synopsys, US; 6imec, BE  
**Abstract**  
The performance of some emerging nanotechnologies benefits from wave pipelining. The design of such circuits requires new models and algorithms. Thus we show how Majority-Inverter Graphs (MIG) can be used for this purpose and we extend the related optimization algorithms. The resulting designs have increased throughput, something that has traditionally been a weak point for the majority of non-charge-based technologies. We benchmark the algorithm on MIG netlists with three different technologies, Spin Wave Devices (SWD), Quantum-dot Cellular Automata (QCA), and NanoMagnetic Logic (NML). We find that the wave pipelined version of the netlists have an improvement in throughput over power of 23x, 13x, and 5x for SWD, QCA, and NML, respectively. In terms of throughput over area ratio, the improvement is 5x, 8x, and 3x, respectively.  
Download Paper (PDF; Only available from the DATE venue WiFi)

### 12:07 10.2.4 DESIGN AUTOMATION FOR QUANTUM ARCHITECTURES

**Speaker:** Martin Roetteler, Microsoft Research, US  
**Authors:** Martin Roetteler, Krysta M. Svore, Dave Wecker and Nathan Wiebe, Microsoft, US  
**Abstract**  
We survey recent strides made towards building a software framework that is capable of compiling quantum algorithms from a high-level description down to physical gates that can be implemented on a fault-tolerant quantum computer. We discuss why compilation and design automation tools such as the ones in our framework are key for tackling the grand challenge of building a scalable quantum computer. We then describe specialized libraries that have been developed using the LIQUi|> programming language. This includes reversible circuits for arithmetic as well as new, truly quantum approaches that rely on quantum computer architectures that allow the probabilistic execution of gates, a model that can reduce time and space overheads in some cases. We highlight why these libraries are useful for the implementation of many quantum algorithms. Finally, we survey the tool REV that facilitate resource efficient compilation of higher-level irreversible programs into lower-level reversible circuits while trying to optimize the memory footprint of the resulting reversible networks. This is motivated by the limited availability of qubits for the foreseeable future.  
Download Paper (PDF; Only available from the DATE venue WiFi)

### Lunch Break in the Garden Foyer

On all conference days (Tuesday to Thursday), a buffet lunch will be offered in the Garden Foyer, in front of the session rooms. Kindly note that this is restricted to conference delegates possessing a lunch voucher only. When entering the lunch break area, delegates will be asked to present the corresponding lunch voucher of the day. Once the lunch area is being left, re-entrance is not allowed for the respective lunch.

### 10.3 Side-Channel Attacks

**Date:** Thursday 30 March 2017  
**Time:** 11:00 - 12:30  
**Location / Room:** 2BC  
**Chair:** Oscar Reparaz, Katholieke Universiteit Leuven, BE  
**Co-Chair:** Wieland Fischer, Infineon Technologies, DE  
This session introduces new side-channel attacks techniques against cryptographic primitives, namely leakage resilient protocols and storage encryption based on AES. Also a power measurement setup specifically targeting static power consumption is presented and evaluated from the side-channel attack viewpoint.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
</table>
| 10:00 | 10.3.1 | SIDE-CHANNEL PLAINTEXT-RECOVERY ATTACKS ON LEAKAGE-RESILIENT ENCRYPTION | Thomas Unterluggauer, Graz University of Technology, AT  
Thomas Unterluggauer, Mario Werner and Stefan Mangard, Graz University of Technology, AT  
**Abstract**  
Differential power analysis (DPA) is a powerful tool to extract the key of a cryptographic implementation from observing its power consumption during the en-/decryption of many different inputs. Therefore, cryptographic schemes based on frequent re-keying such as leakage-resilient encryption aim to inherently prevent DPA on the secret key by limiting the amount of data being processed under one key. However, the original asset of encryption, namely the plaintext, is disregarded. This paper builds on this observation and shows that the re-keying countermeasure does not only protect the secret key, but also induces another DPA vulnerability that allows for plaintext recovery. Namely, the frequent re-keying in leakage-resilient streaming modes causes constant plaintexts to be attackable through first-order DPA. Similarly, constant plaintexts can be revealed from re-keyed block ciphers using templates in a second-order DPA. Such plaintext recovery is particularly critical whenever long-term key material is encrypted and thus leaked. Besides leakage-resilient encryption, the presented attacks are also relevant for a wide range of other applications in practice that implicitly use re-keying, such as multi-party communication and memory encryption with random initialization for the key. Practical evaluations on both an FPGA and a microcontroller support the feasibility of the attacks and thus suggest the use of cryptographic implementations protected by mechanisms like masking in scenarios that require data encryption with multiple keys.  
Download Paper (PDF; Only available from the DATE venue WiFi)
10.4 Emerging Architectures for Reconfigurable Computing

Date: Thursday 30 March 2017
Time: 11:00 - 12:30
Location / Room: 3A

Chair:
Alessandro Cilardo, University of Naples Federico II, IT

Co-Chair:
Florent de Dinechin, ENS-Lyon, FR

This session presents a view of future reconfigurable architectures. These include a field programmable transistor array, a programmable methodology for power gating FPGA routing network, and a dynamic instruction issue technique for coarse grain reconfigurable architectures.

Time | Label | Presentation Title | Authors
-----|-------|---------------------|----------
11:30 | 10.3.2 | (Best Paper Award Candidate) STATIC POWER SIDE-CHANNEL ANALYSIS OF A THRESHOLD IMPLEMENTATION PROTOTYPE CHIP | Thorben Moos, Horst Göritz Institute for IT-Security, Ruhr-Universität Bochum, DE

Authors:
Thorben Moos1, Amir Moradi2 and Bastian Richter1
1Ruhr-Universität Bochum, DE; 2Ruhr University Bochum, DE

Abstract

The static power consumption of modern CMOS devices has become a substantial concern in the context of the side-channel security of cryptographic hardware. The continuous growth of the leakage power dissipation in nanometer-scaled CMOS technologies is not only inconvenient for effective low power designs, but does also create a new target for power analysis adversaries. In this paper, we present the first experimental results of a static power side-channel analysis targeting an ASIC implementation of a provably first-order secure hardware masking scheme. The investigated 150 nm CMOS prototype chip realizes the PRESENT-80 lightweight block cipher as a threshold implementation and allows us to draw a comparison between the information leakage through its dynamic and static power consumption. By employing a sophisticated measurement setup dedicated to static power analysis, including a very low-noise DC amplifier as well as a climate chamber, we are able to recover the key of our target implementation with significantly less traces compared to the corresponding dynamic power analysis attack. In particular, for a successful third-order attack exploiting the static currents, less than 200 thousand traces are needed. Whereas for the same attack in the dynamic power domain around 5 million measurements are required. Furthermore, we are able to show that only-first-order resistant approaches like the investigated threshold implementation do not significantly increase the complexity of a static power analysis. Therefore, we firmly believe that this side channel can actually become the target of choice for real-world adversaries against masking countermeasures implemented in advanced CMOS technologies.

Download Paper (PDF; Only available from the DATE venue WiFi)

12:00 | 10.3.3 | SIDE-CHANNEL POWER ANALYSIS OF XTS-AES | Chao Luo, Northeastern University, CN

Authors:
Chao Luo, Yunsi Fei and A. Adam Ding, Northeastern University, US

Abstract

XTS-AES is an advanced mode of AES for data protection of sector-based devices. Compared to other AES modes, it features two secret keys instead of one, and an additional tweak for each data block. These characteristics make the mode resistant against cryptoanalysis attacks, and also make side-channel attacks on it more challenging. In this paper, we propose two attack methods on XTS-AES overcoming these challenges. In the first attack, we analyze side-channel leakage of the particular modular multiplication in XTS-AES mode. In the second one, we utilize the relationship between two consecutive block tweaks and propose a method to work around the masking of ciphertext by the tweak. These attacks are verified on an FPGA implementation of XTS-AES. The results show that XTS-AES is susceptible to side-channel power analysis attacks, and therefore dedicated protections are required for security of XTS-AES in storage devices.

Download Paper (PDF; Only available from the DATE venue WiFi)

12:30 | IPS-1, 702 | FORMAL MODEL FOR SYSTEM-LEVEL POWER MANAGEMENT DESIGN | Mirela Simonovic, AggiOS, RS

Authors:
Mirela Simonovic1, Vojin Zivovnjovic2 and Lazar Saranovac3
1University of Belgrade, RS; 2AggiOS Inc., US; 3University of Belgrade, School of Electrical Engineering, RS

Abstract

In this paper we present a new formal model, called p-FSM, for system-level power management design. The p-FSM is a modular, compositional, hierarchical, and unified model for hardware and software components. The model encapsulates power management control mechanisms, operating states and properties of a component that affect power, energy and thermal aspects of the system. Component dependencies are modeled through a component-based interface. By connecting multiple p-FSMs we gradually compose the model of the whole system which ensures correct-by-construction system-level control sequencing. The model can also be used to formally verify the functional correctness of the power management design.

Download Paper (PDF; Only available from the DATE venue WiFi)

12:30 | End of session | Lunch Break in Garden Foyer

Keynote Lecture session 11.1 in "Garden Foyer" 1320 - 1350

Lunch Break in the Garden Foyer

On all conference days (Tuesday to Thursday), a buffet lunch will be offered in the Garden Foyer, in front of the session rooms. Kindly note that this is restricted to conference delegates possessing a lunch voucher only. When entering the lunch break area, delegates will be asked to present the corresponding lunch voucher of the day. Once the lunch area is being left, re-entrance is not allowed for the respective lunch.

10.4 Emerging Architectures for Reconfigurable Computing

Date: Thursday 30 March 2017
Time: 11:00 - 12:30
Location / Room: 3A

Chair:
Alessandro Cilardo, University of Naples Federico II, IT

Co-Chair:
Florent de Dinechin, ENS-Lyon, FR

This session presents a view of future reconfigurable architectures. These include a field programmable transistor array, a programmable methodology for power gating FPGA routing network, and a dynamic instruction issue technique for coarse grain reconfigurable architectures.

Time | Label | Presentation Title | Authors
-----|-------|---------------------|----------
10.5 Emerging NoC Directions

Date: Thursday 30 March 2017
Time: 11:00 - 12:30
Location / Room: 3C

Chair:
Jiang Xu, Hong Kong University of Science and Technology, HK

Co-Chair:
Tushar Krishna, GeorgiaTech, US

This session presents papers on emerging directions in NoC design. The first paper uses machine learning for effective power management in NoCs. The next three papers use emerging technologies - wireless, 3D, and Optical - for efficient on-chip communications.
MACHINE LEARNING ENABLED POWER-AWARE NETWORK-ON-CHIP DESIGN

Speaker:
Avinash Kodı, Ohio University, US

Authors:
Dominic DiTomaso1; Ashif Sikder1; Avinash Kodı1 and Ahmed Louri2
1Ohio University, US; 2George Washington University, US

Abstract
Although Network-on-Chips (NoCs) are fast becoming pervasive as the interconnect fabric for multicore architectures and systems-on-chips, they still suffer from excessive static and dynamic power consumption. High dynamic power consumption results from switching and storing data within routers/links while excess static power is consumed when routers and links are not utilized for communication and yet have to be powered up. In this paper, we propose LESSON (Learning Enabled Sleepy Storage Links and Routers in NoCs) to reduce both static and dynamic power consumption by power-gating the links and routers at low network utilization and moving the data storage from within the routers to the links at high network utilization. As the network utilization increases from low-to-high, to accommodate more traffic, we design the same channels to flow traffic in either direction, thereby avoiding complex routing or look-ahead wake-up algorithms. Machine learning algorithms predict when to power-gate the channels and routers and when to increase the channel bandwidths such that power savings are maximized while performance penalty is minimized. Our results show that we can improve total network power consumption when compared to conventional NoC buffer designs by 85.6% and when compared with aggressive NoC buffer designs by 31.7%. Our predictor shows marginal performance penalties and by dynamically changing the direction of the links, we can improve packet latency by 14%.

Download Paper (PDF; Only available from the DATE venue WiFi)

PERFORMANCE EVALUATION AND DESIGN TRADE-OFFS FOR WIRELESS-ENABLED SMART NOC

Speaker:
Karthi Duraisamy, Washington State University, US

Authors:
Karthi Duraisamy and Partha Pande, Washington State University, US

Abstract
SMART (Single-Cycle Multi-hop Asynchronous Repeated Traversal) NoC architectures enable single cycle data transfers, even between the physically far apart nodes. However, enabling single cycle hops over long distance restricts the achievable clock frequency of the system. In other words, increasing the NoC clock frequency lowers the number of hops that can be traversed in a single-cycle in a conventional SMART NoC. In this work, we demonstrate that by integrating wireless links and a novel look-ahead request mechanism in the SMART NoC, it is possible to enable low-latency and energy efficient data transfers, even when the system is designed with high clock frequencies. For various applications considered in this work, our wireless-enabled SMART (WISSMART) NoC achieves on an average 33% reduction in message latency compared to the wireline SMART mesh NoC. This network level improvement translates into 16% savings in full system energy-delay-product.

Download Paper (PDF; Only available from the DATE venue WiFi)

ROBUST TSV-BASED 3D NOC DESIGN TO COUNTERACT ELECTROMIGRATION AND CROSSTALK NOISE

Speaker:
Partha Pande, Washington State University, US

Authors:
Sourav Das1; Janardhan Rao Doppa1; Partha Pande1 and Krishnendu Chakrabarty2
1Washington State University, US; 2Duke University, US

Abstract
A 3D network-on-chip (3D NoC) is an enabler for the design of high-performance and energy-efficient manycore chips. Most popular 3D NoCs utilize the Through-Silicon-Via (TSV)-based vertical links (VLS) as the communication pillars between the planar dies. However, the TSVs in a 3D NoC may fail due to both workload-induced stress and crosstalk capacitance. This failure negatively affects the overall achievable performance of the 3D NoC. In this work, we analyze the joint effects of workload-induced stress and crosstalk on the TSV mean-time-to-failure (MTTF) and hence the 3D NoC lifetime. We demonstrate that if we only consider the effects of electromigration on the TSVs due to workload-induced stress then the estimated MTTF and the subsequently lifetime of 3D NoC are too optimistic. Due to the combined effects of workload and crosstalk noise, the lifetime of 3D NoC reduces significantly. Subsequently, we demonstrate that a spare TSV allocation methodology considering the joint effects of workload and crosstalk noise enhances the lifetime of the 3D NoC by a factor of 4.6 compared to when only the workload is considered for a given spare budget of 5%.

Download Paper (PDF; Only available from the DATE venue WiFi)

PERFORMANCE AND ENERGY AWARE WAVELENGTH ALLOCATION ON RING-BASED WDM 3D OPTICAL NOC

Speaker:
Jiating Luo, INRIA/IRISA, FR

Authors:
Jiating Luo1; Ashraf Elantably1; Pham Van-Dung1; Cedric Killian1; Daniel Chillet1; Sébastien Le Beux2; Olivier Sentieys3 and Ian O’Connor 2
1INRIA/IRISA, FR; 2Lyon Institute of Nanotechnology, FR; 3INRIA, FR

Abstract
Optical Network-on-Chip (ONoC) is a promising communication medium for large-scale Multiprocessor System on Chip (MPSoC). ONoC outperforms classical electrical NoC in terms of throughput and latency. The medium can support multiple transactions at the same time on different wavelengths by using Wavelength Division Multiplexing (WDM). Moreover multiple wavelengths can be used as high-bandwidth channel to reduce transmission time. However, multiplexing signals simultaneously a waveguide can lead to inter-channel crosstalk noise. This problem impacts the Signal to Noise Ratio (SNR) of the optical signal, which leads to an increase in the Bit Error Rate (BER) at the receiver side. In this paper we first formulate the crosstalk noise and execution time models and then propose a Wavelength Allocation (WA) method in a ring-based WDM ONoC allowing to search for performance and energy trade-offs, based on the application constraints. As result, most promising WA solutions are highlighted for a defined application mapping onto 16-core WDM ONoC.

Download Paper (PDF; Only available from the DATE venue WiFi)

Lunch Break in Garden Foyer

Keynote Lecture session 11.0 in "Garden Foyer" 1320 - 1350

Lunch Break in the Garden Foyer

On all conference days (Tuesday to Thursday), a buffet lunch will be offered in the Garden Foyer, in front of the session rooms. Kindly note that this is restricted to conference delegates possessing a lunch voucher only. When entering the lunch break area, delegates will be asked to present the corresponding lunch voucher of the day. Once the lunch area is being left, re-entrance is not allowed for the respective lunch.

Approximate computing and neural networks for novel communication and multimedia systems

Date: Thursday 30 March 2017
Time: 11:00 - 12:30
Location / Room: SA

Chair:
Norbert Wehn, Technical University Kaiserslautern, DE

Co-Chair:
Gerogios Keramidas, Think Silicon, GR
In this session ideas related to approximate computing and neural networks are presented, which can be applied in novel communication and multimedia systems.

### 10.6.1 Exploiting Special-Purpose Function Approximation for Hardware-Efficient QR-Decomposition

**Speaker:** Jochen Rust, University of Bremen, DE  
**Authors:** Jochen Rust¹ and Steffen Paul²  
¹University of Bremen, DE; ²University Bremen, DE  
**Abstract**  
Efficient signal processing takes a key role in application-specific circuit design. For instance, future mobile communication standards, e.g., high-performance industrial mobile communication, require high data rates, low latency and/or high energy-efficiency. Hence, sophisticated algorithms and computing schemes must be explored to satisfy these challenging constraints. In this paper we leverage the paradigm of approximate computing to enable hardware-efficient QR-decomposition for channel precoding. For an efficient computation of the Givens-Rotation, bivariate, non-linear numeric functions are taken into account. An effective design method is introduced leading to highly adapted (special-purpose) functions. For evaluation, our work is tested with different configurations in a Tomlinson-Harashima precoding downlink environment. In addition, a corresponding HDL implementation is set up and logic and physical CMOS synthesis is performed. The comparison to actual reference proves to be a powerful approach for future mobile communication systems.

**Download Paper (PDF; Only available from the DATE venue WiFi)**

### 10.6.2 Embracing Approximate Computing for Energy-Efficient Motion Estimation in High Efficiency Video Coding

**Speaker:** Muhammad Shaflaque, Vienna University of Technology (TU Wien), AT  
**Authors:** Walaa El-Harouni¹, Semeen Rehman², Bharath Srinivas Prabakaran², Akash Kumar³, Rehan Hafiz⁴ and Muhammad Shaflaque⁵  
¹Private Researcher, DE; ²Technische Universität Dresden, DE; ³Technische Universität Dresden, DE; ⁴ETU, PK; ⁵Vienna University of Technology (TU Wien), AT  
**Abstract**  
Approximate Computing is an emerging paradigm for developing highly energy-efficient computing systems. It leverages the inherent resilience of applications to trade output quality with energy efficiency. In this paper, we present a novel approximate architecture for energy-efficient motion estimation (ME) in high efficiency video coding (HEVC). We synthesized our designs for both ASIC and FPGA design flows. ModelSim gate-level simulations are used for functional and timing verification. We comprehensively analyze the impact of heterogeneous approximation modes on the power/energy-quality tradeoffs for various video sequences. To facilitate reproducible results for comparisons and further research and development, the RTL and behavioral models of approximate SAD architectures and constituting approximate modules are made available at https://sourceforge.net/projects/pacclib/.

**Download Paper (PDF; Only available from the DATE venue WiFi)**

### 10.6.3 Hardware Architecture of Bidirectional Long Short-Term Memory Neural Network for Optical Character Recognition

**Speaker:** Vladimir Rybaklin, University of Kaiserslautern, DE  
**Authors:** Vladimir Rybaklin¹, Mohammad Reza Yousefi², Norbert Wehn³ and Didier Stricker⁴  
¹University of Kaiserslautern, DE; ²Augmented Vision Department, German Research Center for Artificial Intelligence (DFKI), DE; ³German Research Center for Artificial Intelligence (DFKI), DE  
**Abstract**  
Optical Character Recognition is conversion of printed or handwritten text images into machine-encoded text. It is a building block of many processes such as machine reading, text-to-speech conversion and text mining. Bidirectional Long Short-Term Memory Neural Networks have shown a superior performance in character recognition with respect to other types of neural networks. In this paper, to the best of our knowledge, we propose the first hardware architecture of Bidirectional Long Short-Term Memory Neural Network with Connectionist Temporal Classification for Optical Character Recognition. Based on the new architecture, we present an FPGA hardware accelerator that achieves 459 times higher throughput than state-of-the-art. Visual recognition is a typical task on mobile platforms that usually use two scenarios either the task runs locally on embedded processor or offloaded to a cloud to be run on high performance machine. We show that computationally intensive visual recognition task benefits from being migrated to our dedicated hardware accelerator and outperforms high-performance CPU in terms of runtime, while consuming less energy than low power systems with negligible loss of recognition accuracy.

**Download Paper (PDF; Only available from the DATE venue WiFi)**

### 10.7 Adaptive and Resilient Cyber-Physical Systems

**Speaker:** Tianchan Guan, Columbia University, US  
**Authors:** Tianchan Guan¹, Xiaoyang Zeng² and Mingoo Seok³  
¹Fudan University, CN; ²Columbia University, US  
**Abstract**  
Neural associative memory (AM) is one of the critical building blocks for cognitive workloads such as classification and recognition. It learns and retrieves memories as human brain does, i.e., changing the strengths of plastic synapses (weights) based on inputs and retrieving information by information itself. One of the key challenges in designing AM is to extend memory capacity (i.e., memories that a neural AM can learn) while minimizing power and hardware overhead. However, prior arts show that memory capacity scales slowly, often logarithmically or in a square root with the total bits of synaptic weights. This makes it prohibitive in hardware and power to achieve large memory capacity for practical applications. In this paper, we propose a synaptic model called recursive synaptic bit reuse, which enables near-linear scaling of memory capacity with total synaptic bits. Also, our model can handle input data that are correlated, more robustly than the conventional model. We experiment our proposed model in Hopfield Neural Networks (HNN) which contains the total synaptic bits of 5kB to 327kB and find that our model can increase the memory capacity as large as 30X over conventional models. We also study hardware cost via VLSI implementation of HNNs in a 65nm CMOS, confirming that our proposed model can achieve up to 10X area savings at the same capacity over conventional synaptic model.

**Download Paper (PDF; Only available from the DATE venue WiFi)**

### End of Session

**Lunch Break in Garden Foyer**  
**Keynote Lecture session 11.0 in "Garden Foyer" 1320 - 1350**

**Lunch Break in the Garden Foyer**  
On all conference days (Tuesday to Thursday), a buffet lunch will be offered in the Garden Foyer, in front of the session rooms. Kindly note that this is restricted to conference delegates possessing a lunch voucher only. When entering the lunch break area, delegates will be asked to present the corresponding lunch voucher of the day. Once the lunch area is being left, re-entry is not allowed for the respective lunch.
The session contains four regular papers and four IP papers addressing different aspects of adaptivity and resilience for Cyber-Physical Systems. The topic of the first paper is implemented in the micro-controllers of drones. Experiments with a quad-rotor drone are conducted to show the effectiveness of the proposed method. The fast growth of civil drones raises significant security challenges. A legitimate drone may be hijacked by GPS spoofing for illegal activities, such as terrorist attacks. The target of this paper is to develop techniques to let drones detect whether they have been hijacked using onboard motion sensors (accelerometers and memory usage. Two model partition schemes are also designed to minimize non-parallel data delivery time, including both wakeup time and transmission time. Experimental results show that when the number of worker nodes increases from 2 to 4, MoDNN can accelerate the DNN computation by 2.17-4.28×. Besides the paralleled execution, the performance speedup also partially comes from the significant reduction of the data delivery time, e.g., 30.02% w.r.t. conventional 2D-grids partition.

EFFICIENT DRONE HIJACKING DETECTION USING ONBOARD MOTION SENSORS

The fast growth of civil drones raises significant security challenges. A legitimate drone may be hijacked by GPS spoofing for illegal activities, such as terrorist attacks. The target of this paper is to develop techniques to let drones detect whether they have been hijacked using onboard motion sensors (accelerometers and gyroscopes). Ideally, the linear acceleration and angular velocity measured by motion sensors can be used to estimate the position of a drone, which can be compared with the position reported by GPS to detect whether the drone has been hijacked. However, the position estimation by motion sensors is very inaccurate due to the significant error accumulation over time. In this paper, we propose a novel method to detect hijacking based on motion sensors measurements and GPS, which overcomes the accumulative error problem. The computational complexity of our method is very low, and thus is suitable to be implemented in the micro-controllers of drones. Experiments with a quad-rotor drone are conducted to show the effectiveness of the proposed method.
12:30  IPS-3, 813  ANOMALIES IN SCHEDULING CONTROL APPLICATIONS AND DESIGN COMPLEXITY  
Speaker: Amir Aminifar, Swiss Federal Institute of Technology in Lausanne, CH  
Authors: Amir Aminifar¹ and Enrico Bini²  
¹Swiss Federal Institute of Technology in Lausanne (EPFL), CH; ²University of Turin, IT  
Abstract  
Today, many control applications in cyber-physical systems are implemented on shared platforms. Such resource sharing may lead to complex timing behaviors and, in turn, instability of control applications. This paper highlights a number of anomalies demonstrating complex timing behaviors caused as a result of resource sharing. Such anomalous scenarios, then, lead to a dramatic increase in design complexity, if not properly considered. Here, we demonstrate that these anomalies are, in fact, very improbable. Therefore, design methodologies for these systems should mainly be devised and tuned towards the majority of cases, as opposed to anomalies, but should also be able to handle such anomalous scenarios.  
Download Paper (PDF; Only available from the DATE venue WiFi)

12:31  IPS-4, 843  CONTRACT-BASED INTEGRATION OF AUTOMOTIVE CONTROL SOFTWARE  
Speaker: Tobias Sehnke, IAV GmbH, DE  
Authors: Tobias Sehnke¹, Matthias Schultalbers² and Rolf Ernst³  
Abstract  
The functionalities of automotive control are distributed over a large number of independently developed components that are interconnected by complex data dependencies. During integration it is critical to ensure the functional correctness of each component, due to the safety-critical nature of the automotive system. Thus existing integration processes ensure that interfaces are syntactically correct. Still in many cases communicated signals are semantically incompatible. This results in complicated errors that are hard to detect and fix. Moreover, existing component languages do not provide applicable means for the description and control of correspondency requirements. In this paper we present a novel methodology for an automated identification of integration errors in automotive control software. The key aspect of our approach are contracts, which are used to disclose domain level requirements. These contracts are then checked during integration supported by existing tools. A case study involving an existing engine control software shows the applicability of our approach by detecting a significant number of formerly unknown integration errors.  
Download Paper (PDF; Only available from the DATE venue WiFi)

12:32  IPS-5, 736  MODELING AND INTEGRATING PHYSICAL ENVIRONMENT ASSUMPTIONS IN MEDICAL CYBER-PHYSICAL SYSTEM DESIGN  
Speaker: Chunhui Guo, Illinois Institute of Technology, US  
Authors: Zhicheng Fu¹, Chunhui Guo¹, Shangqing Ren¹, Yu Jiang² and Lui Sha³  
¹Illinois Institute of Technology, US; ²Tsinghua University, CN; ³University of Illinois at Urbana-Champaign, US  
Abstract  
Implicit physical environment assumptions made by safety critical cyber-physical systems, such as medical cyber- physical systems (M-CPS), can lead to catastrophes. Several recent U.S. Food and Drug Administration (FDA) medical device recalls are due to implicit physical environment assumptions. In this paper, we develop a mathematical assumption model and composition rules that allow M-CPS engineers to explicitly and precisely specify assumptions about the physical environment in which the designed M-CPS operates. Algorithms are developed to integrate the mathematical assumption model with system model so that the safety of the system can be not only validated by both medical and engineering professionals but also formally verified by existing formal verification tools. We use an FDA recalled medical ventilator scenario as a case study to show how the mathematical assumption model and its integration in M-CPS design may improve the safety of the ventilator and M-CPS in general.  
Download Paper (PDF; Only available from the DATE venue WiFi)

12:33  IPS-6, 535  A UTILITY-DRIVEN DATA TRANSMISSION OPTIMIZATION STRATEGY IN LARGE SCALE CYBER-PHYSICAL SYSTEMS  
Speaker: Bei Yu, The Chinese University of Hong Kong, HK  
Authors: Soumi Chattopadhyay¹, Ansuman Banerjee¹ and Bei Yu²  
¹Indian Statistical Institute, IN; ²The Chinese University of Hong Kong, HK  
Abstract  
In this paper, we examine the problem of data dissemination and optimization in the context of a large scale distributed cyber-physical system (CPS), and propose a novel rule-based mechanism for effective observation collection and transmission. Our work rests on the idea that all observations on all parameters are not required at all times, and thereby, selective data transmission can reduce sensor workload significantly. Experiments show the efficacy of our proposal.  
Download Paper (PDF; Only available from the DATE venue WiFi)

12:30  End of session  
Lunch Break in Garden Foyer  
Keynote Lecture session 11.0 in "Garden Foyer" 1320 - 1350  
Lunch Break in the Garden Foyer  
On all conference days (Tuesday to Thursday), a buffet lunch will be offered in the Garden Foyer, in front of the session rooms. Kindly note that this is restricted to conference delegates possessing a lunch voucher only. When entering the lunch break area, delegates will be asked to present the corresponding lunch voucher of the day. Once the lunch area is being left, re-entrance is not allowed for the respective lunch.
UB10 Session 10
Date: Thursday 30 March 2017
Time: 12:00 - 14:30
Location / Room: Booth 1, Exhibition Area

<table>
<thead>
<tr>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>UB10.1</td>
<td>A FRAMEWORK FOR VARIATION-AWARE ANALOG CIRCUITS SIZING</td>
<td>Ons Lahiouel, Concordia University, CA; Mohamed H. Zaki and Sofiene Tahar, Concordia University, CA</td>
</tr>
<tr>
<td>UB10.2</td>
<td>TFA: TRANSPARENT CODE OFFLOADING ON FPGA</td>
<td>Roberto Rigamonti, HEIG-VD/HES-SO, CH; Anthony Convers, Baptiste Delporte, Xavier Ruppen and Alberto Dassatti, HEIG-VD/HES-SO, CH</td>
</tr>
</tbody>
</table>

More information ...

UB10 Session 10
Date: Thursday 30 March 2017
Time: 12:00 - 14:30
Location / Room: Booth 1, Exhibition Area

<table>
<thead>
<tr>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>UB10.1</td>
<td>A FRAMEWORK FOR VARIATION-AWARE ANALOG CIRCUITS SIZING</td>
<td>Ons Lahiouel, Concordia University, CA; Mohamed H. Zaki and Sofiene Tahar, Concordia University, CA</td>
</tr>
<tr>
<td>UB10.2</td>
<td>TFA: TRANSPARENT CODE OFFLOADING ON FPGA</td>
<td>Roberto Rigamonti, HEIG-VD/HES-SO, CH; Anthony Convers, Baptiste Delporte, Xavier Ruppen and Alberto Dassatti, HEIG-VD/HES-SO, CH</td>
</tr>
</tbody>
</table>

More information ...

UB10 Session 10
Date: Thursday 30 March 2017
Time: 12:00 - 14:30
Location / Room: Booth 1, Exhibition Area

<table>
<thead>
<tr>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>UB10.1</td>
<td>A FRAMEWORK FOR VARIATION-AWARE ANALOG CIRCUITS SIZING</td>
<td>Ons Lahiouel, Concordia University, CA; Mohamed H. Zaki and Sofiene Tahar, Concordia University, CA</td>
</tr>
<tr>
<td>UB10.2</td>
<td>TFA: TRANSPARENT CODE OFFLOADING ON FPGA</td>
<td>Roberto Rigamonti, HEIG-VD/HES-SO, CH; Anthony Convers, Baptiste Delporte, Xavier Ruppen and Alberto Dassatti, HEIG-VD/HES-SO, CH</td>
</tr>
</tbody>
</table>

More information ...

UB10 Session 10
Date: Thursday 30 March 2017
Time: 12:00 - 14:30
Location / Room: Booth 1, Exhibition Area

<table>
<thead>
<tr>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>UB10.1</td>
<td>A FRAMEWORK FOR VARIATION-AWARE ANALOG CIRCUITS SIZING</td>
<td>Ons Lahiouel, Concordia University, CA; Mohamed H. Zaki and Sofiene Tahar, Concordia University, CA</td>
</tr>
<tr>
<td>UB10.2</td>
<td>TFA: TRANSPARENT CODE OFFLOADING ON FPGA</td>
<td>Roberto Rigamonti, HEIG-VD/HES-SO, CH; Anthony Convers, Baptiste Delporte, Xavier Ruppen and Alberto Dassatti, HEIG-VD/HES-SO, CH</td>
</tr>
</tbody>
</table>

More information ...

UB10 Session 10
Date: Thursday 30 March 2017
Time: 12:00 - 14:30
Location / Room: Booth 1, Exhibition Area

<table>
<thead>
<tr>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>UB10.1</td>
<td>A FRAMEWORK FOR VARIATION-AWARE ANALOG CIRCUITS SIZING</td>
<td>Ons Lahiouel, Concordia University, CA; Mohamed H. Zaki and Sofiene Tahar, Concordia University, CA</td>
</tr>
<tr>
<td>UB10.2</td>
<td>TFA: TRANSPARENT CODE OFFLOADING ON FPGA</td>
<td>Roberto Rigamonti, HEIG-VD/HES-SO, CH; Anthony Convers, Baptiste Delporte, Xavier Ruppen and Alberto Dassatti, HEIG-VD/HES-SO, CH</td>
</tr>
</tbody>
</table>

More information ...

UB10 Session 10
Date: Thursday 30 March 2017
Time: 12:00 - 14:30
Location / Room: Booth 1, Exhibition Area

<table>
<thead>
<tr>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>UB10.1</td>
<td>A FRAMEWORK FOR VARIATION-AWARE ANALOG CIRCUITS SIZING</td>
<td>Ons Lahiouel, Concordia University, CA; Mohamed H. Zaki and Sofiene Tahar, Concordia University, CA</td>
</tr>
<tr>
<td>UB10.2</td>
<td>TFA: TRANSPARENT CODE OFFLOADING ON FPGA</td>
<td>Roberto Rigamonti, HEIG-VD/HES-SO, CH; Anthony Convers, Baptiste Delporte, Xavier Ruppen and Alberto Dassatti, HEIG-VD/HES-SO, CH</td>
</tr>
</tbody>
</table>

More information ...

UB10 Session 10
Date: Thursday 30 March 2017
Time: 12:00 - 14:30
Location / Room: Booth 1, Exhibition Area

<table>
<thead>
<tr>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>UB10.1</td>
<td>A FRAMEWORK FOR VARIATION-AWARE ANALOG CIRCUITS SIZING</td>
<td>Ons Lahiouel, Concordia University, CA; Mohamed H. Zaki and Sofiene Tahar, Concordia University, CA</td>
</tr>
<tr>
<td>UB10.2</td>
<td>TFA: TRANSPARENT CODE OFFLOADING ON FPGA</td>
<td>Roberto Rigamonti, HEIG-VD/HES-SO, CH; Anthony Convers, Baptiste Delporte, Xavier Ruppen and Alberto Dassatti, HEIG-VD/HES-SO, CH</td>
</tr>
</tbody>
</table>

More information ...

UB10 Session 10
Date: Thursday 30 March 2017
Time: 12:00 - 14:30
Location / Room: Booth 1, Exhibition Area

<table>
<thead>
<tr>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>UB10.1</td>
<td>A FRAMEWORK FOR VARIATION-AWARE ANALOG CIRCUITS SIZING</td>
<td>Ons Lahiouel, Concordia University, CA; Mohamed H. Zaki and Sofiene Tahar, Concordia University, CA</td>
</tr>
<tr>
<td>UB10.2</td>
<td>TFA: TRANSPARENT CODE OFFLOADING ON FPGA</td>
<td>Roberto Rigamonti, HEIG-VD/HES-SO, CH; Anthony Convers, Baptiste Delporte, Xavier Ruppen and Alberto Dassatti, HEIG-VD/HES-SO, CH</td>
</tr>
</tbody>
</table>

More information ...

UB10 Session 10
Date: Thursday 30 March 2017
Time: 12:00 - 14:30
Location / Room: Booth 1, Exhibition Area

<table>
<thead>
<tr>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>UB10.1</td>
<td>A FRAMEWORK FOR VARIATION-AWARE ANALOG CIRCUITS SIZING</td>
<td>Ons Lahiouel, Concordia University, CA; Mohamed H. Zaki and Sofiene Tahar, Concordia University, CA</td>
</tr>
<tr>
<td>UB10.2</td>
<td>TFA: TRANSPARENT CODE OFFLOADING ON FPGA</td>
<td>Roberto Rigamonti, HEIG-VD/HES-SO, CH; Anthony Convers, Baptiste Delporte, Xavier Ruppen and Alberto Dassatti, HEIG-VD/HES-SO, CH</td>
</tr>
</tbody>
</table>

More information ...
More information ...

Abstract
Future 5G networks are expected to provide an increase of 10x in data rates. To meet these requirements, baseband stations will be designed using mixed architectures, i.e., DSPs, FPGAs. However, efficiently programming these architectures is not trivial due to the drastic increase in complexity of their design space. To overcome this issue, we need to have unified tools capable of rapidly exploring, partitioning and prototyping the mixed architecture designs of 5G systems. At DATE 2017 University Booth, we demonstrate such a unified tool and show our latest achievements in the automatic code generation engine of TTool/DIPLODOCUS, a HLS/SMIL framework for the hardware/software co-design of data-flow systems, to support mixed architectures. Our demonstration will show the full design and evaluation of a 5G data-link layer receiver for both a DSP-based and an IP-based designs. We will validate the effectiveness of our solution by comparing automated vs manual designs.

More information ...

Abstract
Advances in microfluidic technologies have led to the emergence of Digital Microfluidic Biochips (DMFBs), which are capable of automating laboratory procedures. These DMFBs raised significant attention in industry and academia creating a demand for devices. Commercial products are available but come at a high price. So far, there are two open hardware DMFBs available: the DropBot from WheelerLabs and the OpenDrop from GaudiLabs. The aim of the StackADrop was to create a DMFB with many directly addressable cells while still being very compact. The StackADrop strives to provide means to experiment with different hardware setups. Its main feature are the exchangeable top plates, supporting 256 high-voltage pins. It features SPI, UART and 12C connectors for attaching sensors/actuators and can be connected to a computer using USB for interactive sessions using a control software. The modularity allows to easily test different cell shapes, such as squares, hexagons and triangles.

More information ...

More information ...

More information ...

More information ...

More information ...

More information ...

More information ...

More information ...

More information ...

More information ...

Abstract
We argue that there is still room for innovation beyond the very classic waveform display though. We implemented a proof-of-concept waveform viewer (codenamed Tides) that has several unique features that go beyond the standard set of features for waveform viewers.

More information ...
WE DARE: WEARABLE ELECTRONICS DIRECTIONAL AUGMENTED REALITY

Abstract
Current augmented reality (AR) eyewear solutions require large form factors, weight, cost and energy that reduce usability. In fact, connectivity, image processing, localization, and direction evaluation lead to high processing and power requirements. A multi-antenna system, patented by the industrial partner, enables a new generation of smart eye-wear that elegantly requires less hardware, connectivity, and power to provide AR functionalities. This will allow us to directionally locate nearby radio emitting sources that highlight objects of interest (e.g., people or retail items) by using existing standards like Bluetooth Low Energy, Apple’s iBeacon and Google’s Eddystone. This booth will report the current level of research addressed by the Computer Science Department of University of Verona and the company Wagoo LLC. In the presented demo, different objects emit an “I am here” signal and a prototype of the smart glasses shows the information related to the observed object.

More information ...
### 11.1 Wearable and Smart Medical Devices Day: HW and SW design constraints in medical devices

**Date:** Thursday 30 March 2017  
**Time:** 14:00 - 15:30  
**Location / Room:** SBC

**Organisers:**  
José L. Ayala, Universidad Complutense de Madrid, ES  
Chris Van Hoof, IMEC, BE

**Chair:**  
Maurizio Rossi, University of Trento, IT  
José L. Ayala, Universidad Complutense de Madrid, ES

**Co-Chair:**  
José L. Ayala, Universidad Complutense de Madrid, ES

This session will present the current efforts on making optimal electronic designs for biomedical devices. Therefore, the issues of low power consumption, reconfigurability and design challenges will be analysed in a broad range of medical applications.

#### 11.1.1 RECONFIGURABLE EMBEDDED SYSTEMS APPLICATIONS FOR VERSATILE BIOMEDICAL MEASUREMENTS

**Speaker:** Luca Cerina, Politecnico di Milano, IT  
**Authors:** Luca Cerina\(^1\) and Marco D. Santambrogio\(^2\)  
\(^1\)Politecnico di Milano, Italy; \(^2\)Politecnico di Milano, Italy

**Abstract**  
Nowadays, the majority of the monitoring devices used in clinical settings is limited to specific applications and powered by highly specialized microcontrollers and pre-programmed DSP systems. Moreover, these kind of devices are usually connected to a high capacity battery to operate in case of power blackout. Nevertheless, considering that all the measured bio-signals depends from an amperometric or potentiometric transducer, it should be viable to integrate them on a single device with multiple probes, reprogrammable sensor-fusion capabilities and on-board signal processing. Within this context, in this paper, we present a design concept for such a device. Exploiting FPGA reconfigurability, various analog front-ends can be connected to the device and configured to return the measured signal or the output of the desired signal processing to the user. Multiple case studies with different sensors and end-user applications are presented. This paper addresses these challenges by way of some recent examples of ASICs designed for ECG, EIT (Electrical Impedance Tomography) and PPG (Photoplethysmography) sensors as well as for non-invasive blood pressure monitoring.

**Download Paper (PDF; Only available from the DATE venue WiFi)**

#### 11.1.2 ULTRA LOW POWER MICROELECTRONICS FOR WEARABLE AND MEDICAL DEVICES

**Speaker:** Pierre-François Rüedi, CSEM, CH  
**Authors:** Pierre-François Rüedi, André Bischof, Marcin Kamil Augustyniak, Pascal Persechini, Jean-Luc Nagel, Marc Pons, Stephane Emery and Olivier Chételat, CSEM S.A., CH

**Abstract**  
The requirements for wearables and portable medical devices present a number of challenges in terms of integration, autonomy and connectivity, and demand a careful co-design of hardware and software to reach optimum performance. This paper addresses these challenges by way of some recent examples of ASICs designed for ECG, EIT (Electrical Impedance Tomography) and PPG (Photoplethysmography) sensors as well as for non-invasive blood pressure monitoring.

**Download Paper (PDF; Only available from the DATE venue WiFi)**

#### 11.1.3 DESIGN CHALLENGES FOR WEARABLE EMG APPLICATIONS

**Speaker:** Elisabetta Farella, Fondazione Bruno Kessler - ICT Center, IT  
**Authors:** Bojan Milošević\(^1\), Simone Benatti\(^2\) and Elisabetta Farella\(^1\)  
\(^1\)Fondazione Bruno Kessler (FBK), IT; \(^2\)Università di Bologna, IT

**Abstract**  
Wearable technologies are changing the way we deal with health and fitness in our daily life. Nevertheless, while MEMS-enabled inertial sensors have conquered the consumer market, physiological monitoring has still to face barriers due to the complexity and costs of physical interfaces (e.g. electrodes), the degree of intuitiveness of the interaction and the processing required to reach satisfying performance. These limitations are mitigated by the embedded systems’ growing integration of interfacing capabilities and efficient computing power. In this paper, we describe the main applications and the related technologies for the acquisition and processing of myoelectric (EMG) signals. Starting from well established active sensors and bench-top setups, we introduce a recent design based on the combination of an integrated Analog Front End (AFE) and embedded processing. This solution provides high quality signal acquisition and on-board digital processing capabilities with a contained power consumption. The system was tested within the prosthesis control application scenario, one of the most stringent EMG applications, achieving a 90% gesture recognition accuracy with real time on-board processing at a power consumption of 30mW. Such promising results highlight the current trend in shifting EMG applications from dedicated analog solutions towards integrated digital devices, favouring the development of advanced, modular and low-power wearable solutions.

**Download Paper (PDF; Only available from the DATE venue WiFi)**
On all conference days (Tuesday to Thursday), coffee and tea will be served during the coffee breaks at the below-mentioned times in the exhibition area.

**Tuesday, March 28, 2017**
- Coffee Break 10:30 - 11:30
- Coffee Break 16:00 - 17:00

**Wednesday, March 29, 2017**
- Coffee Break 10:00 - 11:00
- Coffee Break 16:00 - 17:00

**Thursday, March 30, 2017**
- Coffee Break 10:00 - 11:00
- Coffee Break 15:30 - 16:00

---

### 11.2 Emerging Technologies for Future Memory Design

**Date:** Thursday 30 March 2017  
**Time:** 14:00 - 15:30  
**Location / Room:** 4BC

**Chair:**  
Weisheng Zhao, Beihang University, CN

**Co-Chair:**  
Jean-Michel Portal, Aix-Marseille Université, FR

Memory design based on emerging technologies is critical for the future VLSI design targeting low power and high performance. This session involves novel design method and evaluation tool for emerging technologies (i.e. STT-MRAM, Racetrack memory, Phase Change Memory and Ferroelectric memory etc.) including variation aware design, novel architecture implementation and reliability concern.

### 11.2.1 HYBRID VC-MTJ/CMOS NON-VOLATILE STOCHASTIC LOGIC FOR EFFICIENT COMPUTING

**Speaker:** Shaodi Wang, University of California, Los Angeles, US

**Authors:** Shaodi Wang¹, Septadeep Pal¹, Tianmu Li², Andrew Pan², Cecile Grezes², Pedram Khalili-Amiri², Kang L. Wang² and Puneet Gupta²  
¹University of California, Los Angeles, US; ²UCLA, US

**Abstract**  
In this paper, we propose a non-volatile stochastic computing (SC) scheme using voltage-controlled magnetic tunnel junction (VC-MTJ) and negative differential resistance (NDR). The proposed design includes a VC-MTJ based true stochastic bit stream generator and VC-MTJ and NDR based stochastic adder, multiplier, register, which are experimentally demonstrated using 60nm VC-MTJ and CMOS NDR connected on die. These components are then used to realize FIR filter and AdaBoost (machine-learning algorithm). 3X - 37X energy advantage is shown for the proposed SC compared with CMOS binary arithmetic ASIC and SC designs.

[Download Paper (PDF; Only available from the DATE venue WiFi)]

### 11.2.2 DESIGN AND BENCHMARKING OF FERROELECTRIC FET BASED TCAM

**Speaker:** Xunzhao Yin, University of Notre Dame, US

**Authors:** Xunzhao Yin, Michael Niemier and X. Sharon Hu, University of Notre Dame, US

**Abstract**  
We consider how emerging transistor technologies, specifically ferroelectric field effect transistors (or FeFETs), can realize compact and energy efficient ternary content addressable memories (TCAMs). As Moore's Law-based performance scaling trends slow and many computational tasks of interest are now more data-centric than compute-centric, researchers are looking to improve performance/save energy by integrating efficient and compact logic-processing elements into various levels of the memory hierarchy. Potential benefits include reduced I/O traffic, energy/delay from data transfers, etc. A TCAM is an example of a logic-in-memory element that is ubiquitous in routers, caches, databases, and even neural networks. Not surprisingly, researchers continue to study how emerging technologies could lead to improved TCAMs. Recent work has considered how non-volatile (NV) memory technologies (e.g., resistive random access memory (ReRAM) or magnetic tunnel junctions (MTJs)) could best be used to construct low energy, NV TCAMs. However, acceptable Ron-Roff ratios and the two terminal nature of these devices introduce energy and area overheads. Due to hysteresis in a device's I-V curve, an FeFET-based NV TCAM, offers low area overhead, as well as search energies and search speeds that are superior to other TCAM designs (i.e., based on MTJ, ReRAM and CMOS in array- and architectural-level evaluations.)

[Download Paper (PDF; Only available from the DATE venue WiFi)]
15:00  11.2.3  LEVERAGING ACCESS PORT POSITIONS TO ACCELERATE PAGE TABLE WALK IN DWM MAIN MEMORY
Speaker: Chengmo Yang, University of Delaware, US
Authors: Hoda Aghaei Khouzani 1, Pouya Fotouhi 1, Chengmo Yang 1 and Guang R. Gao 2
1University of Delaware, US; 2Department of Electrical and Computer Engineering, University of Delaware, US
Abstract: Domain Wall Memory (DWM) with ultra-high density and comparable read/write latency to SRAM/DRAM is an attractive replacement for CMOS-based devices. Unlike SRAM/DRAM, DWM has non-uniform data access latency that is proportional to the number of shift operations. While previous works have demonstrated the feasibility of using DWM as main memory and have proposed different ways to alleviate the impact of shift operations, none of them have addressed the performance-critical metadata accesses, in particular page table accesses. To bridge this gap, this paper aims at accelerating page table walk in DWM main memory from two innovative aspects. First of all, we propose a new page table layout and leverage the positions of access ports in DWM to differentiate the state of page table entries. In addition, we propose a technique to pre-align the access ports to the positions to be accessed in the near future, thus hiding shift latency to the maximum extent. Since both address translation and context switching are affected by page table access latency, the proposed technique can effectively improve system performance and user experience.
Download Paper (PDF; Only available from the DATE venue WiFi)

15:15  11.2.4  VAET-STT: A VARIATION AWARE ESTIMATOR TOOL FOR STT-MRAM BASED MEMORIES
Speaker: Sarath Mohanachandran Nair 1, KIT, Germany, DE
Authors: Sarath Mohanachandran Nair 1, Rajendra Bishnoi 1, Mohammad Saber Golanbari 1, Fabian Oboril 1 and Mehdi Tahoori 1
1Karlsruhe Institute of Technology, DE; 2Karlsruhe Institute of Technology, DE
Abstract: Spin Transfer Torque Magnetic Random Access Memory (STT-MRAM) is a promising candidate to replace CMOS based on-chip memories due to its advantages such as non-volatility, high density and scalability. However, its stochastic switching and higher sensitivity to process variation compared to CMOS memories can significantly affect its performance, energy and reliability. Although a few works exist which analyze the impact of process variation at the bit-cell level, such analysis at the system level is missing. We have bridged this gap in our work. Specifically, we quantify the effect of stochasticity and process variations from the cell-level to the overall memory system and perform a variation-aware memory configuration optimization for energy or performance while meeting reliability constraints. Our system-level variation-aware framework has been built on top of the well-known NVSim engine. The results show that our framework can provide more realistic margins and the optimized variation-aware memory configuration could be significantly different from the conventional framework.
Download Paper (PDF; Only available from the DATE venue WiFi)

15:30  11.2.8  PROTECT NON-VOLATILE MEMORY FROM WEAR-OUT ATTACK BASED ON TIMING DIFFERENCE OF ROW BUFFER HIT/MISS
Speaker: Haiyu Mao, Tsinghua University, CN
Authors: Haiyu Mao 1, Xian Zhang 1, Guangyu Sun 2 and Jiwu Shu 1
1Tsinghua University, CN; 2Peking University, CN
Abstract: Non-volatile Memories (NVMs), such as PCM and ReRAM, have been widely proposed for future main memory design because of their low standby power, high storage density, fast access speed. However, these NVMs suffer from the write endurance problem. In order to prevent a malicious program from wearing out NVMs deliberately, researchers have proposed various wear-leveling methods, which remap logical addresses to physical addresses randomly and dynamically. However, we discover that side channel leakage based on NVM row buffer hit information can reveal details of address remappings. Consequently, it can be leveraged to side-step the wear-leveling. Our simulation shows that the proposed attack method in this paper can wear out a NVM within 137 seconds, even with the protection of state-of-the-art wear-leveling schemes. To counteract this attack, we further introduce an effective countermeasure named Intra-Row Swap (IRS) to hide the wear-leveling details. The basic idea is to enable an additional intra-row block swap when a new logical address is remapped to the memory row. Experiments demonstrate that IRS can secure NVMs with negligible timing/energy overhead, compared with previous works.
Download Paper (PDF; Only available from the DATE venue WiFi)

15:32  11.2.8  EFFECTS OF CELL SHAPES ON THE ROUTABILITY OF DIGITAL MICROFLUIDIC MICROCHIPS
Speaker: Oliver Kesztocze, University of Bremen, DE
Authors: Kevin Leonard Schneider 1, Oliver Kesztocze 1, Jannis Stoppe 1 and Rolf Drechsler 2
1University of Bremen, DE; 2University of Bremen/DFKI GmbH, DE
Abstract: Digital microfluidic biochips (DMFBs) are an emerging technology promising a high degree of automation in laboratory procedures by means of manipulating small discretized amounts of fluids. A crucial part in conducting experiments on biochips is the routing of discretized droplets. While doing so, droplets must not enter each others’ interference region to avoid unintended mixing. This leads to cells in the proximity of the droplet being impassable for others. For different cell shapes, the effect of these temporary blockages varies as the adjacency of cells changes with their shapes. Yet, no evaluation with respect to routability in relation to cell shapes has been conducted so far. This paper analyses and compares various tessellations for the field of cells. Routing benchmarks are mapped to these and the results are compared in order to determine if and how cell shapes affect the performance of DMFBs, showing that certain cell shapes are superior to others.
Download Paper (PDF; Only available from the DATE venue WiFi)
This session introduces new approaches for building reconfigurable accelerators and heterogeneous architectures integrating big-little cores, FPGA hardware, and coarse grain reconfigurable array architectures targeting emerging applications such as Hadoop map reduce framework and neural networks.

**Authors**
Soheil Hashemi, Nicholas Anthony, Hokchhay Tann, Iris Bahar and Sherief Reda, Brown University, US

**Abstract**

Understanding the impact of precision quantization on the accuracy and energy of neural networks

Speaker:
Sohel Hashemi, Brown University, US

Authors:
Sohel Hashemi, Nicholas Anthony, Hokchhay Tann, Iris Bahar and Sherief Reda, Brown University, US

Abstract

Deep neural networks are gaining in popularity as they are used to generate state-of-the-art results for a variety of computer vision and machine learning applications. At the same time, these networks have grown in depth and complexity in order to solve harder problems. Given the limitations in power budgets dedicated to these networks, the importance of low-power, low-memory solutions has been stressed in recent years. While a large number of dedicated hardware using different precisions has recently been proposed, there exists no comprehensive study of different bit precisions and arithmetic in both inputs and network parameters. In this work, we address this issue and perform a study of different bit-precisions in neural networks (from floating-point to fixed-point, powers of two, and binary). In our evaluation, we consider and analyze the effect of precision scaling on both network accuracy and hardware metrics including memory footprint, power and energy consumption, and design area. We also investigate training-time methodologies to compensate for the reduction in accuracy due to limited bit precision and demonstrate that in most cases, precision scaling can deliver significant benefits in design metrics at the cost of very modest decreases in network accuracy. In addition, we propose that a small portion of the benefits achieved when using lower precisions can be forfeited to increase the network size and therefore the accuracy. We evaluate our experiments, using three well-recognized networks and datasets to show its generality. We investigate the trade-offs and highlight the benefits of using lower precisions in terms of energy and memory footprint.

Download Paper (PDF; Only available from the DATE venue WiFi)
In this paper, we present a high speed yet energy efficient approximate divider where the division operation is performed by multiplying the dividend by the inverse of the divisor. In this structure, truncated value of the dividend is multiplied exactly (approximately) by the approximate inverse value of divisor. To assess the efficacy of the proposed divider, its design parameters are extracted and compared to those of a number of prior art dividers in a 45nm CMOS technology. Results reveal that this structure provides 66% and 52% improvements in the area and energy consumption, respectively, compared to the most advanced prior art approximate divider. In addition, delay and energy consumption of the division operation are reduced about 94.4% and 99.93%, respectively, compared to the most advanced prior art approximate divider. In addition, delay and energy consumption of the division operation are reduced about 94.4% and 99.93%, respectively, compared to the most advanced prior art approximate divider.

Abstract

In this paper, we present a high speed yet energy efficient approximate divider where the division operation is performed by multiplying the dividend by the inverse of the divisor. In this structure, truncated value of the dividend is multiplied exactly (approximately) by the approximate inverse value of divisor. To assess the efficacy of the proposed divider, its design parameters are extracted and compared to those of a number of prior art dividers in a 45nm CMOS technology. Results reveal that this structure provides 66% and 52% improvements in the area and energy consumption, respectively, compared to the most advanced prior art approximate divider. In addition, delay and energy consumption of the division operation are reduced about 94.4% and 99.93%, respectively, compared to those of an exact SRT radix-4 divider. Finally, the efficacy of the proposed divider in image processing application is studied.

Advanced prior art approximate divider. In addition, delay and energy consumption of the division operation are reduced about 94.4% and 99.93%, respectively, compared to those of an exact SRT radix-4 divider. Finally, the efficacy of the proposed divider in image processing application is studied.

In this paper, we present a high speed yet energy efficient approximate divider where the division operation is performed by multiplying the dividend by the inverse of the divisor. In this structure, truncated value of the dividend is multiplied exactly (approximately) by the approximate inverse value of divisor. To assess the efficacy of the proposed divider, its design parameters are extracted and compared to those of a number of prior art dividers in a 45nm CMOS technology. Results reveal that this structure provides 66% and 52% improvements in the area and energy consumption, respectively, compared to the most advanced prior art approximate divider. In addition, delay and energy consumption of the division operation are reduced about 94.4% and 99.93%, respectively, compared to those of an exact SRT radix-4 divider. Finally, the efficacy of the proposed divider in image processing application is studied.

In this paper, we present a high speed yet energy efficient approximate divider where the division operation is performed by multiplying the dividend by the inverse of the divisor. In this structure, truncated value of the dividend is multiplied exactly (approximately) by the approximate inverse value of divisor. To assess the efficacy of the proposed divider, its design parameters are extracted and compared to those of a number of prior art dividers in a 45nm CMOS technology. Results reveal that this structure provides 66% and 52% improvements in the area and energy consumption, respectively, compared to the most advanced prior art approximate divider. In addition, delay and energy consumption of the division operation are reduced about 94.4% and 99.93%, respectively, compared to those of an exact SRT radix-4 divider. Finally, the efficacy of the proposed divider in image processing application is studied.

In this paper, we present a high speed yet energy efficient approximate divider where the division operation is performed by multiplying the dividend by the inverse of the divisor. In this structure, truncated value of the dividend is multiplied exactly (approximately) by the approximate inverse value of divisor. To assess the efficacy of the proposed divider, its design parameters are extracted and compared to those of a number of prior art dividers in a 45nm CMOS technology. Results reveal that this structure provides 66% and 52% improvements in the area and energy consumption, respectively, compared to the most advanced prior art approximate divider. In addition, delay and energy consumption of the division operation are reduced about 94.4% and 99.93%, respectively, compared to those of an exact SRT radix-4 divider. Finally, the efficacy of the proposed divider in image processing application is studied.

In this paper, we present a high speed yet energy efficient approximate divider where the division operation is performed by multiplying the dividend by the inverse of the divisor. In this structure, truncated value of the dividend is multiplied exactly (approximately) by the approximate inverse value of divisor. To assess the efficacy of the proposed divider, its design parameters are extracted and compared to those of a number of prior art dividers in a 45nm CMOS technology. Results reveal that this structure provides 66% and 52% improvements in the area and energy consumption, respectively, compared to the most advanced prior art approximate divider. In addition, delay and energy consumption of the division operation are reduced about 94.4% and 99.93%, respectively, compared to those of an exact SRT radix-4 divider. Finally, the efficacy of the proposed divider in image processing application is studied.

In this paper, we present a high speed yet energy efficient approximate divider where the division operation is performed by multiplying the dividend by the inverse of the divisor. In this structure, truncated value of the dividend is multiplied exactly (approximately) by the approximate inverse value of divisor. To assess the efficacy of the proposed divider, its design parameters are extracted and compared to those of a number of prior art dividers in a 45nm CMOS technology. Results reveal that this structure provides 66% and 52% improvements in the area and energy consumption, respectively, compared to the most advanced prior art approximate divider. In addition, delay and energy consumption of the division operation are reduced about 94.4% and 99.93%, respectively, compared to those of an exact SRT radix-4 divider. Finally, the efficacy of the proposed divider in image processing application is studied.

In this paper, we present a high speed yet energy efficient approximate divider where the division operation is performed by multiplying the dividend by the inverse of the divisor. In this structure, truncated value of the dividend is multiplied exactly (approximately) by the approximate inverse value of divisor. To assess the efficacy of the proposed divider, its design parameters are extracted and compared to those of a number of prior art dividers in a 45nm CMOS technology. Results reveal that this structure provides 66% and 52% improvements in the area and energy consumption, respectively, compared to the most advanced prior art approximate divider. In addition, delay and energy consumption of the division operation are reduced about 94.4% and 99.93%, respectively, compared to those of an exact SRT radix-4 divider. Finally, the efficacy of the proposed divider in image processing application is studied.
This session focuses on issues related to timing and layout in the presence of manufacturing variability and photolithographic limitations. The first paper reduces pessimism in timing analysis by estimating path sensitization while accounting for delay variations. The second paper enables patterning with reduced wirelength and overlay violation through placement refinement. The third paper improves manufacturability with an optimization algorithm for cut locations in line-end process. The last paper discusses clock tree synthesis to reduce delay sensitivity mismatch with gate delay circuitry.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>14:00</td>
<td>11.4.1</td>
<td>QUANTIFYING ERROR: EXTENDING STATIC TIMING ANALYSIS WITH PROBABILITY TRANSITIONS</td>
<td>Kevin Murray, University of Toronto, CA&lt;br&gt;Kevin Murray, SA, Andrea Suardi, VA, Vaughn Betz, and George Constantinides&lt;br&gt;1University of Toronto, CA; 2Imperial College, GB</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
<td>Timing analysis is a cornerstone of the digital design process. Statistical Static Timing Analysis was introduced to reduce pessimism by modeling device delay variations. However, it ignores circuit logic, which may cause some timing paths to never or only rarely be sensitized. We introduce a general timing analysis approach and tool to calculate the probability that individual timing paths are sensitized, enabling the calculation of bounding delay distributions over all input combinations. We show the connection to the well-known 4SAT problem and present approaches to improve scalability, achieving average results 46 to 32% less pessimistic than Static Timing Analysis while running 14.6 to 44.0 times faster than Monte-Carlo timing simulation. Download Paper (PDF; Only available from the DATE venue WiFi)</td>
</tr>
<tr>
<td>14:30</td>
<td>11.4.2</td>
<td>ON REFINING STANDARD CELL PLACEMENT FOR SELF-ALIGNED DOUBLE PATTERNING</td>
<td>Ting-Chi Wang, National Tsing Hua University, TW&lt;br&gt;Ye-Hong Chen, Sheng-He Wang and Ting-Chi Wang, National Tsing Hua University, TW</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
<td>In this paper, we study the problem of refining a standard cell placement for self-aligned double patterning (SADP), which asks to simultaneously refine a detailed placement and find a valid SADP layout decomposition such that both overlay violation and wirelength are as small as possible. We present an algorithm that adopts the technique of white space insertion for an SADP-aware single-row cell placement problem. Based on the single-row algorithm, we then describe an approach to the addressed placement refinement problem. Finally, we report encouraging experimental results to support the efficacy of our approach. Download Paper (PDF; Only available from the DATE venue WiFi)</td>
</tr>
<tr>
<td>15:00</td>
<td>11.4.3</td>
<td>CUT MASK OPTIMIZATION FOR MULTI-PATTERNING DIRECTED SELF-ASSEMBLY LITHOGRAPH</td>
<td>Wachirawit Ponghiran, School of Electrical Engineering, KAIST, KR&lt;br&gt;Wachirawit Ponghiran, Seongbo Shim, and Youngsoo Shin&lt;br&gt;1School of Electrical Engineering, KAIST, KR; 2Dept. of Electrical Engineering, KAIST, KR; 3KAIST, KR</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
<td>Line-end cut process has been used to create very fine metal wires in sub-14nm technology. Cut patterns split regular line patterns into a number of wire segments with some segments being used as actual routing wires. In sub-7nm technology, cuts are smaller than optical resolution limit, and a directed self-assembly lithography with multiple patterning (MP-DSAL) is considered as a patterning solution. We address cut mask optimization problem for MP-DSAL, in which cut locations are determined in such a way that cuts are grouped into manufacturable clusters and assigned to one of masks without MP coloring conflicts; minimizing wire extensions is also pursued in the process. Only a restricted version of this problem has been addressed before while we do not assume any such restrictions. The problem is formulated as ILP first, and a fast heuristic algorithm is also proposed for application to larger circuits. Experimental results indicate that the ILP can remove all coloring conflicts, and reduce total wire extensions by 93% on average compared to those obtained by the restricted approach. Heuristic achieves a similar result with less than 1% of coloring conflicts and 91% reduction in total wire extensions. Download Paper (PDF; Only available from the DATE venue WiFi)</td>
</tr>
<tr>
<td>15:15</td>
<td>11.4.4</td>
<td>CLOCK DATA COMPENSATION AWARE CLOCK TREE SYNTHESIS IN DIGITAL CIRCUITS WITH ADAPTIVE CLOCK GENERATION</td>
<td>Saibal Mukhopadhyay, Georgia Institute of Technology, US&lt;br&gt;Taesik Na, Jong Hwan Ko and Saibal Mukhopadhyay, Georgia Institute of Technology, US</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
<td>Adaptive clock generation to track critical path delay enables lowering supply voltage with improved timing slack under supply noise. This paper presents how to synthesize clock tree in adaptive clocking to fully exploit the clock data compensation (CDC) effect in digital circuits. The paper first provides analytical proof of ideal CDC effect for ring oscillator based clock generation. Second, the paper analyzes non-ideal CDC effect in a gate dominated critical path and wire dominated clock tree design. The paper shows the delay sensitivity mismatch between clock tree and critical path can degrade CDC effect by analyzing timing slack under power supply noise (PSN). Finally, the paper proposes simple but efficient clock tree synthesis (CTS) technique to maximize timing slack under PSN in digital circuits with adaptive clock generation. Download Paper (PDF; Only available from the DATE venue WiFi)</td>
</tr>
<tr>
<td>15:30</td>
<td>IPS-11</td>
<td>TIMING-AWARE WIRE WIDTH OPTIMIZATION FOR SADP PROCESS</td>
<td>Youngsoo Song, KAIST, KR&lt;br&gt;Youngsoo Song, Sangmin Kim and Youngsoo Shin, School of Electrical Engineering, KAIST, KR</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
<td>With the scaling of the minimum feature size, RC delay of interconnect is relatively getting more critical in next node technology. SADP is one of the popular processes used in sub-7nm technology. For SADP process, we can increase wire width using patterns formed by block mask, which can reduce wire resistance of critical nets. We determine the direction and length of each wire widening, so that the resulting layout is conflict-free. We convert this as a maximum weight independent set problem and solve this by formulating an ILP. For various test circuits, the wire resistance of critical nets was reduced on average by 18.5%, which led to 9.9% reduction in clock period. The wire width optimization in SADP process can give an insight into timing optimization through the enhancement of fabrication process. Download Paper (PDF; Only available from the DATE venue WiFi)</td>
</tr>
</tbody>
</table>
End of session

Coffee Break in Exhibition Area

On all conference days (Tuesday to Thursday), coffee and tea will be served during the coffee breaks at the below-mentioned times in the exhibition area.

Tuesday, March 28, 2017
- Coffee Break 10:30 - 11:30
- Coffee Break 16:00 - 17:00

Wednesday, March 29, 2017
- Coffee Break 10:00 - 11:00
- Coffee Break 16:00 - 17:00

Thursday, March 30, 2017
- Coffee Break 10:00 - 11:00
- Coffee Break 15:30 - 16:00

11.5 Smart Energy and Automotive Systems

Date: Thursday 30 March 2017
Time: 14:00 - 15:30
Location / Room: 3C

Chair:
Geoff Merrett, University of Southampton, GB

Co-Chair:
Michele Magno, ETHZ, CH

This session presents the state of the art in efficient automotive software, smart battery systems and the latest strives toward energy neutral wireless communications systems.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
</table>
| 14:00  | 11.5.1             | (Best Paper Award Candidate) ON REDUCING BUSY WAITING IN AUTOSAR VIA TASK-RELEASE-DELTA-BASED RUNNABLE REORDERING | Robert Höttger1, Olaf Spinczyk2 and Burkhard Igel1  
1  
FH-Dortmund, DE;  
TU-Dortmund, DE |

Abstract

The increasing amount of innovative software technologies in the automotive domain comes with challenges regarding inevitable distributed multi-core and many-core methodologies. Approaches for general purpose solutions have been studied over decades but do not completely meet the specific constraints (e.g. timing, safety, reliability, affinity, etc.) for AUTOSAR compliant applications. AUTOSAR utilizes a spinlock mechanism in combination with the priority ceiling protocol in order to provide mutually exclusive access to shared resources. The essential disadvantages of spinlocks are unpredictable task response times on the one hand and wasted computation time caused by busy waiting periods on the other hand. In this paper, we propose a concept of task-release-delta-based runnable reordering for the purpose of sequentializing parallel accesses to shared resources, resulting in reduced task response times, improved timing predictability, and increased parallel efficiency respectively. To achieve this, runnables that represent smallest executable program parts in AUTOSAR are reordered based on precedence constraints. Our experiments among industrial use cases show that task response times can be reduced by up to 18.2%.

Download Paper (PDF; Only available from the DATE venue WiFi)

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>14:30</td>
<td>11.5.2</td>
<td>POWER NEUTRAL PERFORMANCE SCALING FOR ENERGY HARVESTING MP-SOCS</td>
<td>Benjamin Fletcher, Domenico Balsamo and Geoff Merrett, University of Southampton, GB</td>
</tr>
</tbody>
</table>

Abstract

Using energy ‘harvested’ from the environment to power autonomous embedded systems is an attractive ideal, alleviating the burden of periodic battery replacement. However, such energy sources are typically low-current and transient, with high temporal and spatial variability. To overcome this, large energy buffers such as supercapacitors or batteries are typically incorporated to achieve energy neutral operation, where the energy consumed over a certain period of time is equal to the energy harvested. Large energy buffers, however, pose environmental issues in addition to increasing the size and cost of systems. In this paper we propose a novel power neutral performance scaling approach for multiprocessor system-on-chips (MP-SoCs) powered by energy harvesting. Under power neutral operation, the system’s performance is dynamically scaled through DVFS and DPM such that the instantaneous power consumption is approximately equal to the instantaneous harvested power. Power neutrality means that large energy buffers are no longer required, while performance scaling ensures that available power is effectively utilised. The approach is experimentally validated using the Samsung Exynos5422 big.LITTLE SoC directly coupled to a monocrystalline photovoltaic array, with only 47mF of intermediate energy storage. Results show that the proposed approach is successful in tracking harvested power, stabilising the supply voltage to within 5% of the target value for over 93% of the test duration, resulting in the execution of 69% more instructions compared to existing static approaches.

Download Paper (PDF; Only available from the DATE venue WiFi)
Efficient Decentralized Active Balancing Strategy for Smart Battery Cells

Speaker:
Nitin Shivaraman, Nanyang Technological University, SG
Authors:
Nitin Shivaraman, Arvind Easwaran and Sebastian Steinhorst
1Nanyang Technological University, SG; 2Technical University of Munich, DE

Abstract
Among series-connected cells in large battery packs, such as those found in electric vehicles, a charge imbalance develops over time due to manufacturing and temperature variations. Therefore, active balancing strategies can be employed in Battery Management Systems (BMSs) to attain a charge balance among cells by transferring charge between them, maximizing the usable capacity of the battery pack. Recently, decentralized BMS architectures with smart battery cells have been developed, in which balancing strategies can operate by local cooperation between the cells without requiring global coordination. In this paper, we propose a decentralized active balancing strategy for smart cells where we identify boundary cells having special properties. These boundary cells enable to divide the global balancing problem into independent subproblems, where local decisions on charge transfers eventually converge to a globally balanced battery pack. The proposed strategy is implemented in a simulator framework and compared with two decentralized state-of-the-art strategies. Our results show significantly improved performance and scalability of the proposed strategy in terms of charge transfer losses and communication overhead between cells, while maintaining a comparable time to balance.

Download Paper (PDF; Only available from the DATE venue WiFi)

Wulora: An Energy Efficient Iot End-Node for Energy Harvesting and Heterogeneous Communication

Speaker:
Michele Magno, ETH Zurich, CH
Authors:
Michele Magno, Faryal Ait Aoudia, Matthieu Gautier, Olivier Berder and Luca Benini
1ETH Zurich, CH; 2Irisa - University of Rennes, FR; 3University of Rennes 1, IRISA, INRIA, FR; 4Irisa - University of Rennes, FR; 5Università di Bologna, IT

Abstract
Intelligent connected objects, which build the IoT, are electronic devices usually supplied by batteries that significantly limit their life-time. These devices are expected to be deployed in very large numbers, and manual replacement of their batteries will severely restrict their large-scale or widearea deployments. Therefore energy efficiency is of the utmost importance in the design of these devices. The wireless communication between the distributed sensor devices and the host stations can consume significant energy, even more when data needs to reach several kilometers of distance. In this paper, we present an energy-efficient multi-sensing platform that exploits energy harvesting, long-range communication and ultra-low-power shortrange wake-up radio to achieve self sustainability in a kilometer range network. The proposed platform is designed with power efficiency in mind and exploits the always-on wake-up radio as both receiver and a power management unit to significantly reduce the quiescent current even continuously listening the wireless channel. Moreover the platform allows the building of an heterogeneous long-short range network architecture to reduce the latency and reduce the power consumption in listening phase at only 4.6µW. Experimental results and simulations demonstrate the benefits of the proposed platform and heterogeneous network.

Download Paper (PDF; Only available from the DATE venue WiFi)

Ultra Low-Power Visual Odometry for Nano-Scale Unmanned Aerial Vehicles

Speaker:
Daniele Falossi, ETH Zurich, CH
Authors:
Daniele Falossi, Andrea Marongiu and Luca Benini
1ETH - Zurich, CH; 2Swiss Federal Institute of Technology in Zurich (ETHZ), CH; 3Università di Bologna, IT

Abstract
One of the fundamental functionalities for autonomous navigation of Unmanned Aerial Vehicles (UAVs) is the hovering capability. State-of-the-art techniques for implementing hovering on standard-size UAVs process camera stream to determine position and orientation (visual odometry). Similar techniques are considered unaffordable in the context of nano-scale UAVs (i.e. few centimeters of diameter), where the ultra-constrained power-energies of tiny rotor-crafts limit the on-board computational capabilities to those of low-power microcontrollers. In this work we study how the emerging ultra-low-power parallel computing paradigm could enable the execution of complex hovering algorithmic flows onto nano-scale UAVs. We provide insight on the software pipeline, the parallelization opportunities and the impact of several algorithmic enhancements. Results demonstrate that the proposed software flow and architecture can deliver unprecedented GOPS/W, achieving 117 frame-per-second within a power envelope of 10 mW.

Download Paper (PDF; Only available from the DATE venue WiFi)

Long Range Wireless Sensing Powered by Plant-Microbial Fuel Cell

Speaker:
Maurizio Rossi, University of Trento, IT
Authors:
Maurizio Rossi, Pietro Tosato, Luca Gemma, Luca Torquati, Cristian Catania, Sergio Camalò and Davide Brunelli, University of Trento, IT

Abstract
Going low power and having a low or neutral impact on the environment is key for embedded systems, as pervasive and wearable consumer electronics is growing. In this paper, we present a self-sustaining, ultra-low power device, supplied by a Plant-Microbial Fuel Cell (PMFC) and capable of smart sensing and long-range communication. The use of a PMFC as a power source is challenging but has many advantages like the only requirement of watering the plant. The system uses aggressive power management thanks to FRAM technology exploited to retain microcontroller status and to shutdown electronics without losing context information. Experimental results show that the proposed system paves the way to energy neutral sensors powered by biosystems available almost anywhere on Earth.

Download Paper (PDF; Only available from the DATE venue WiFi)
MULTI-ARMED BANDITS FOR EFFICIENT LIFETIME ESTIMATION IN MPSOC DESIGN

Authors:
Antonio Miele, Politecnico di Milano, IT

Abstract
The recent ability of some vehicles to handle autonomously the lane change maneuvers, and the progressive equipment of roads and vehicles with ITS-G5 units motivate this paper to consider the case of road narrowing that requires a lane change because one lane is occupied by road works for maintenance, incidents and so on. This paper extends the approaches of cooperative speed synchronization at intersections. Because of the complexity of the overall system, it considers each automatic lane change as a mobile (unfixed) intersection in which vehicles synchronize their velocities. The wireless communication allows each vehicle to increase its field of view to negotiate its merging with the other equipped vehicles. Hence, the proposed approach introduces a kind of automatic "courtesy" between equipped vehicles. The paper defines the intersection point between each pair of vehicles and the suited protocol to safely reach the new lane. The protocol can be handled by the new work item (NWJ) that has been created at ETSI to realize platooning and cooperative adaptive cruise control. Besides enhancing safety, the simulation results show that the main advantage of the approach is the energy saving by smoothing the traffic.

Download Paper (PDF; Only available from the DATE venue WiFi)
15:00  11.6.3 HARDWARE-BASED ON-LINE INTRUSION DETECTION VIA SYSTEM CALL ROUTINE FINGERPRINTING
Speaker: Yiorgos Makris, The University of Texas at Dallas, US
Authors: Liwei Zhou and Yiorgos Makris, The University of Texas at Dallas, US
Abstract
We introduce a hardware-based methodology for performing on-line intrusion detection in microprocessors. The proposed method extracts fingerprints from the basic blocks of the routine executed in response to a system call and examines their validity using a Bloom filter. Implementation in hardware renders spoofing attacks, to which operating system or hypervisor-level intrusion detection methods are vulnerable, ineffective. The proposed method is evaluated using kernel rootkits which covertly modify the system call service routines of a Linux operating system running on a 32-bit x86 architecture, implemented in the Simics simulation environment, while hardware overhead is evaluated using a predictive 45nm PDK.
Download Paper (PDF; Only available from the DATE venue WiFi)

15:30  IPS-16, 935 EVALUATING MATRIX REPRESENTATIONS FOR ERROR-TOLERANT COMPUTING
Speaker: Pareesa Golnari, Princeton University, US
Authors: Pareesa Ameen Golnari and Sharad Malik, Princeton University, US
Abstract
We propose a methodology to determine the suitability of different data representations in terms of their error-tolerance for a given application with accelerator-based computing. This methodology helps match the characteristics of a representation to the data access patterns in an application. For this, we first identify a benchmark of key kernels from linear algebra that can be used to construct applications of interest using any of several widely used data representations. This is then used in an experimental framework for studying the error tolerance of a specific data format for an application. As case studies, we evaluate the error-tolerance of seven data-formats on sparse matrix to vector multiplication, diagonal add, and two machine learning applications (i) principal component analysis (PCA), which is a statistical technique widely used in data analysis and (ii) movie recommendation system with Restricted Boltzmann Machine (RBM) as the core. We observe that the Dense format behaves well for complicated data accesses such as diagonal accessing but is poor in utilizing local memory. Sparse formats with simpler addressing methods and a careful selection of stored information, e.g., CRS and ELLPACK, demonstrate a better error-tolerance for most of our target applications.
Download Paper (PDF; Only available from the DATE venue WiFi)

15:31  IPS-17, 131 SIMULATION-BASED DESIGN PROCEDURE FOR SUB 1 V CMOS CURRENT REFERENCE
Speaker: Dmitry Osipov, University of Bremen, DE
Authors: Dmitry Osipov and Steffen Paul, University of Bremen, DE
Abstract
This paper presents a new compact current reference and a simulation-based design procedure to establish the circuit parameters quickly and efficiently. To verify the proposed design procedure, two sub 1-V example circuits for two different reference current values (80 nA and 800 nA) were designed and simulated using 0.35 µm CMOS technology. The circuits are robust against supply voltage variation without the need for external bandgap. A line sensitivity of approximately 1-2%/V over the supply voltage range from sub 1 V is achieved in both cases. The simulated temperature coefficient (TC) values are 93 ppm°C and 197 ppm°C in the temperature range from 0°C to 120°C for the 800 nA and 80 nA references, respectively.
Download Paper (PDF; Only available from the DATE venue WiFi)

15:30  End of session
Coffee Break in Exhibition Area
On all conference days (Tuesday to Thursday), coffee and tea will be served during the coffee breaks at the below-mentioned times in the exhibition area.

Tuesday, March 28, 2017
- Coffee Break 10:30 - 11:30
- Coffee Break 16:00 - 17:00

Wednesday, March 29, 2017
- Coffee Break 10:00 - 11:00
- Coffee Break 16:00 - 17:00

Thursday, March 30, 2017
- Coffee Break 10:00 - 11:00
- Coffee Break 15:30 - 16:00
state-of-the-art biologically-inspired techniques and devices that demonstrate the efficacy of such methods to designs focused on smart, low-power, and secure systems on chip. Recent years have inspired techniques such as evolutionary algorithms and artificial neural networks have been used in the mainstream circuit design community infrequently. While advanced well-tuned techniques are employed in current integrated circuits to increase the lifetime of cyber-physical, IoT and other systems, major concerns and important product differentiators such as power, security and variability continue to be major design factors. For many applications a sacrifice of performance or accuracy is acceptable in exchange for extremely low power consumption. However, even when this sacrifice is possible, other conflicting performance features must still be taken into account. Biologically-inspired techniques and devices that demonstrate the efficacy of such methods to designs focused on smart, low-power, and secure systems on chip.
15:12 11.8.3 SEMANTIC DRIVEN HIERARCHICAL LEARNING FOR ENERGY-EFFICIENT IMAGE CLASSIFICATION
Speaker: Priyadarshini Panda, Purdue University, US
Authors: Priyadarshini Panda and Kaushik Roy, Purdue University, US
Abstract: Machine-learning algorithms have shown outstanding image recognition performance for computer vision applications. While these algorithms are modeled to mimic brain-like cognitive abilities, they lack the remarkable energy-efficient processing capability of the brain. Recent studies in neuroscience reveal that the brain resolves the competition among multiple visual stimuli presented simultaneously with several mechanisms of visual attention that are key to the brain’s ability to perform cognition efficiently. One such mechanism known as saliency based selective attention simplifies complex visual tasks into characteristic features and then selectively activates particular areas of the brain based on the feature (or semantic) information in the input. Interestingly, we note that there is a significant similarity among underlying characteristic semantics (like color or texture) of images across multiple objects in real world applications. This presents us with an opportunity to decompose a large classification problem into simpler tasks based on semantic or feature similarity. In this paper, we propose semantic driven hierarchical learning to construct a tree-based classifier inspired by the biological visual attention mechanism for optimizing energy-efficiency of machine learning classifiers. We exploit the inherent feature similarity across images to identify the input variability and use recursive optimization procedure, to determine data partitioning at each tree node, whereby, learning the feature hierarchy. A set of binary classifiers is organized on top of the learnt hierarchy to minimize the overall test-time complexity. The feature based-learning allows selective activation of only those branches and nodes of the classification tree that are relevant to the input while keeping the remaining nodes idle. The proposed framework has been evaluated on Caltech-256 dataset and achieves 3.7x reduction in test complexity for 1.2% accuracy improvement over state-of-the-art one-vs-all tree-based method, and even higher improvements in test-time (of 5.5x) when some loss in output accuracy (up to 2.5%) is acceptable.
Download Paper (PDF; Only available from the DATE venue WiFi)

15:44 11.8.4 MACHINE LEARNING FOR RUN-TIME ENERGY OPTIMISATION IN MANY-CORE SYSTEMS
Speaker: Rishad Shafik, Newcastle University, GB
Authors: Dwipaayen Biswas, 1Vibshna Balogopal, 1Rishad Shafik, 1Bashir Al-Hashimi 1 and Geoff Merrett 1
1University of Southampton, GB; 2Newcastle University, GB
Abstract: In recent years, the focus of computing has moved away from performance-centric serial computation to energy-efficient parallel computation. This necessitates run-time optimisation techniques to address the dynamic resource requirements of different applications on many-core architectures. In this paper, we report on intelligent run-time algorithms which have been experimentally validated for managing energy and application performance in many-core embedded system. The algorithms are underpinned by a cross-layer system approach where the hardware, software and application layers work together to optimise the energy-performance trade-off. Algorithm development is motivated by the biological process of how a human brain (acting as an agent) interacts with the external environment (system) changing their respective states over time. This leads to a pay-off for the action taken, and the agent eventually learns to take the optimal/best decisions in future. In particular, our online approach uses a model-free reinforcement learning algorithm that suitably selects the appropriate voltage-frequency scaling based on workload prediction to meet the applications’ performance requirements and achieve energy savings of up to 16% in comparison to state-of-the-art techniques, when tested on four ARM A15 cores of an ODROID-XU3 platform.
Download Paper (PDF; Only available from the DATE venue WiFi)
A voltage-scalable RISC processor integrating standard-cell based memory (SCM) is demonstrated. Unlike conventional processors, the processor has Standard-Cell based Memories (SCMs) as an alternative to conventional SRAM macros, enabling it to operate at a 0.4 V single-supply voltage. The processor is implemented with the fully automated cell-based design, which leads to low design costs. By scaling the supply voltage and applying the back-gate biasing techniques, the power dissipation of the SCMs is less than 20 μW, enabling the SCMs to operate with ambient energy source only. In this demonstration, the SCMs of the processor operates with a lemon battery as the ambient energy source.
**UB11.6 GNOCs: AN ULTRA-FAST, HIGHLY EXTENSIBLE, CYCLE-ACCURATE GPU-BASED PARALLEL NETWORK-ON-CHIP SIMULATOR**

**Presenter:**
Amir CHARIF, TIMA, FR

**Authors:**
Nacer-Eddine Zergainoh and Michael Nicolaidis, TIMA, FR

**Abstract:**
With the continuous decrease in feature sizes and the recent emergence of 3D stacking, chips comprising thousands of nodes are becoming increasingly relevant, and state-of-the-art NoC simulators are unable to simulate such a high number of nodes in reasonable times. In this demo, we showcase GNOCs, the first detailed, modular and scalable parallel NoC simulator running fully on GPU (Graphics Processing Unit). Based on a unique design specifically tailored for GPU parallelism, GNOCs is able to achieve unprecedented speedups with no loss of accuracy. To enable quick and easy validation of novel ideas, the programming model was designed with high extensibility in mind. Currently, GNOCs accurately models a VC-based microarchitecture. It supports 2D and 3D mesh topologies with full or partial vertical connections. A variety of routing algorithms and synthetic traffic patterns, as well as dependency-driven trace-based simulation (Netrace), are implemented and will be demonstrated

**More information ...**

**UB11.7 EMU: RAPID FPGA PROTOTYPING OF NETWORK SERVICES IN C#**

**Presenter:**
Salvator Galea, University of Cambridge, GB

**Authors:**
Nik Sultana1, Pietro Bressana1, David Greaves1, Robert Soulé2, Andrew W Moore1 and Noa Zilberman1
1 University of Cambridge, GB; 2Università della Svizzera italiana, CH

**Abstract:**
General-purpose CPUs and OS abstractions impose overheads that make it challenging to implement network functions and services in software. On the other hand, programmable hardware such as FPGAs suffer from low-level programming models, which make the rapid development of network services cumbersome. We demonstrate Emu, a framework that makes use of an HLS tool (Kiwi) and enables the execution of high-level descriptions of network services, written in C#, on both x86 and Xilinx FPGA. Emu therefore opens up new opportunities for improved performance and power usage, and enables developers to more easily write network services and functions. We demonstrate C# implementations of network functions, such as Memcached and DNS Server, using Emu running on both x86 and NetFPGA-SUME platform and show that they are competitive to natively written hardware counterparts while providing a superior development and debug environment.

**More information ...**

**UB11.9 HEPSYCODE: A SYSTEM-LEVEL METHODOLOGY FOR HW/SW CO-DESIGN OF HETEROGENEOUS PARALLEL DEDICATED SYSTEMS**

**Presenter:**
Luigi Pomante, University of L’Aquila, IT

**Authors:**
Luigi Pomante, University of L’Aquila, IT; Luigi Pomante, University of L’Aquila, IT; Daniele Di Pompeo1, Emilio Incerto2 and Daniele Cimbrone1
1 University of L’Aquila, Italy; 2Gran Sasso Science Institute, Italy

**Abstract:**
Heterogeneous parallel systems have been recently exploited for a wide range of application domains, for both the dedicated (e.g. embedded) and the general purpose products. Such systems can include different processor cores, memories, dedicated ICS and a set of connections between them. They are so complex that the design methodology plays a major role in determining the success of the products. So, this demo addresses the problem of the electronic system-level hw/sw co-design of heterogeneous parallel dedicated systems. In particular, it shows an enhanced CSP/SystemC-based design space exploration step (and related ESL-EDA prototype tools), in the context of an existing hw/sw co-design flow that, given the system specification and related F/NF requirements, is able to (semi)automatically propose to the designer: - a custom heterogeneous parallel architecture; - an HW/SW partitioning of the application; - a mapping of the partitioned entities onto the proposed architecture.

**More information ...**

16:30 End of session

**IP5 Interactive Presentations**

**Date:** Thursday 30 March 2017
**Time:** 15:30 - 16:00
**Location / Room:** IP sessions (in front of rooms 4A and 5A)

Interactive Presentations run simultaneously during a 30-minute slot. A poster associated to the IP paper is on display throughout the morning. Additionally, each IP paper is briefly introduced in a one-minute presentation in a corresponding regular session, prior to the actual Interactive Presentation. At the end of each afternoon Interactive Presentations session the award ‘Best IP of the Day’ is given.

<table>
<thead>
<tr>
<th>Label</th>
<th>Presentation Title</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>IP5-1</strong></td>
<td>FORMAL MODEL FOR SYSTEM-LEVEL POWER MANAGEMENT DESIGN</td>
</tr>
</tbody>
</table>

**Speaker:**
Mirela Simonovic, Aggis, RS

**Authors:**
Mirela Simonovic1, Vojin Zivojnovic3 and Lazar Saranovac3
1 University of Belgrade, RS; 2 AGGIOS Inc., US; 3 University of Belgrade, School of Electrical Engineering, RS

**Abstract:**
In this paper we present a new formal model, called p-FSM, for system-level power management design. The p-FSM is a modular, compositional, hierarchical, and unified model for hardware and software components. The model encapsulates power management control mechanisms, operating states and properties of a component that affect power, energy and thermal aspects of the system. Inter-component dependencies are modeled through a component-based interface. By connecting multiple p-FSMs we gradually compose the model of the whole system which ensures correct-by-construction system-level control sequencing. The model can also be used to formally verify the functional correctness of the power management design.

**Download Paper (PDF; Only available from the DATE venue WiFi)**
IPS-2  EXTINGUISHING MEMORY CAPACITY OF NEURAL ASSOCIATIVE MEMORY BASED ON RECURSIVE SYNAPTIC BIT REUSE

Speaker: Tianchan Guan, Columbia University, US
Authors: Tianchan Guan1, Xiaoyang Zeng1 and Mingoo Seok2
1Indian Statistical Institute, IN; 2The Chinese University of Hong Kong, HK

Abstract: Neural associative memory (AM) is one of the critical building blocks for cognitive workloads such as classification and recognition. It learns and retrieves memories as human brain does, i.e., changing the strengths of plastic synapses (weights) based on inputs and retrieving information by itself. One of the key challenges in designing AM is to extend memory capacity (i.e., memories that a neural AM can learn) while minimizing power and hardware overhead. However, prior arts show that memory capacity scales slowly, often logarithmically or in square root with the total bits of synaptic weights. This makes it prohibitive in hardware and power to achieve large memory capacity for practical applications. In this paper, we propose a synaptic model called recursive synaptic bit reuse, which enables near-linear scaling of memory capacity with total synaptic bits. Also, our model can handle input data that are correlated, more robustly than the conventional model. We experiment our proposed model in Hopfield Neural Networks (HNN) which contains the total synaptic bits of 5kB to 327kB and find that our model can increase the memory capacity as large as 30X over conventional models. We also study hardware cost via VLSI implementation of HNNs in a 65nm CMOS, confirming that our proposed model can achieve up to 10X area savings at the same capacity over conventional synaptic model.

Download Paper (PDF; Only available from the DATE venue WiFi)

IPS-3  ANOMALIES IN SCHEDULING CONTROL APPLICATIONS AND DESIGN COMPLEXITY

Speaker: Amir Aminifar, Swiss Federal Institute of Technology in Lausanne, CH
Authors: Amir Aminifar1 and Enrico Bini2
1Swiss Federal Institute of Technology in Lausanne (EPFL), CH; 2University of Turin, IT

Abstract: Today, many control applications in cyber-physical systems are implemented on shared platforms. Such resource sharing may lead to complex timing behaviors and, in turn, instability of control applications. This paper highlights a number of anomalies demonstrating complex timing behaviors caused as a result of resource sharing. Such anomalous scenarios, then, lead to a dramatic increase in design complexity, if not properly considered. Here, we demonstrate that these anomalies are, in fact, very improbable. Therefore, design methodologies for these systems should mainly be devised and tuned towards the majority of cases, as opposed to anomalies, but should also be able to handle such anomalous scenarios.

Download Paper (PDF; Only available from the DATE venue WiFi)

IPS-4  CONTRACT-BASED INTEGRATION OF AUTOMOTIVE CONTROL SOFTWARE

Speaker: Tobias Sehnke, IAV GmbH, DE
Authors: Tobias Sehnke1, Matthias Schultalbers2 and Rolf Ernst3

Abstract: The functionalities of automotive control are distributed over a large number of independently developed components that are interconnected by complex data dependencies. During integration it is critical to ensure the functional correctness of each component, due to the safety-critical nature of the automotive system. Thus existing integration processes ensure that interfaces are syntactically correct. Still in many cases communicated signals are semantically incompatible. This results in complicated errors that are hard to detect and fix. Moreover, existing component languages do not provide applicable means for the description and control of correspondent requirements. In this paper we present a novel methodology for an automated identification of integration errors in automotive control software. The key aspect of our approach are contracts, which are used to disclose domain level requirements. These contracts are then checked during integration supported by existing tools. A case study involving an existing engine control software shows the applicability of our approach by detecting a significant number of formerly unknown integration errors.

Download Paper (PDF; Only available from the DATE venue WiFi)

IPS-5  MODELING AND INTEGRATING PHYSICAL ENVIRONMENT ASSUMPTIONS IN MEDICAL CYBER-PHYSICAL SYSTEM DESIGN

Speaker: Chunhui Guo, Illinois Institute of Technology, US
Authors: Zhicheng Fu1, Chunhui Guo1, Shangqing Ren1, Yu Jiang2 and Lui Sha3
1Illinois Institute of Technology, US; 2Tsinghua University, CN; 3University of Illinois at Urbana-Champaign, US

Abstract: Implicit physical environment assumptions made by safety critical cyber-physical systems, such as medical cyber-physical systems (M-CPS), can lead to catastrophes. Several recent U.S. Food and Drug Administration (FDA) medical device recalls are due to implicit physical environment assumptions. In this paper, we develop a mathematical assumption model and composition rules that allow M-CPS engineers to explicitly and precisely specify assumptions about the physical environment in which the designed M-CPS operates. Algorithms are developed to integrate the mathematical assumption model with system model so that the safety of the system can be not only validated by both medical and engineering professionals but also formally verified by existing formal verification tools. We use an FDA recalled medical ventilator scenario as a case study to show how the mathematical assumption model and its integration in M-CPS design may improve the safety of the ventilator and M-CPS in general.

Download Paper (PDF; Only available from the DATE venue WiFi)

IPS-6  A UTILITY-DRIVEN DATA TRANSMISSION OPTIMIZATION STRATEGY IN LARGE SCALE CYBER-PHYSICAL SYSTEMS

Speaker: Bei Yu, The Chinese University of Hong Kong, HK
Authors: Soumi Chattopadhyay1, Ansuman Banerjee1 and Bei Yu2
1Indian Statistical Institute, IN; 2The Chinese University of Hong Kong, HK

Abstract: In this paper, we examine the problem of data dissemination and optimization in the context of a large scale distributed cyber-physical system (CPS), and propose a novel rule-based mechanism for effective observation collection and transmission. Our work rests on the idea that all observations on all parameters are not required at all times, and thereby, selective data transmission can reduce sensor workload significantly. Experiments show the efficacy of our proposal.

Download Paper (PDF; Only available from the DATE venue WiFi)
**IP5-7**  
**PROTECT NON-VOLATILE MEMORY FROM WEAR-OUT ATTACK BASED ON TIMING DIFFERENCE OF ROW BUFFER HIT/MISS**  
**Speaker:** Haiyu Mao, Tsinghua University, CN  
**Authors:** Haiyu Mao, Xiaochun Zhang, Guangyu Sun, Jiwu Shu  
**Affiliations:** 1) Tsinghua University, CN; 2) Peking University, CN  
**Abstract:** Non-volatile Memories (NVMs), such as PCM and ReRAM, have been widely proposed for future main memory design because of their low standby power, high storage density, fast access speed. However, these NVMs suffer from the write endurance problem. In order to prevent a malicious program from wearing out NVMs deliberately, researchers have proposed various wear-leveling methods, which remap logical addresses to physical addresses randomly and dynamically. However, we discover that side channel leakage based on NVM row buffer hit information can reveal details of address remappings. Consequently, it can be leveraged to side-step the wear-leveling. Our simulation shows that the proposed attack method in this paper can wear out a NVM within 137 seconds, even with the protection of state-of-the-art wear-leveling schemes. To counteract this attack, we further introduce an effective countermeasure named Intra-Row Swap (IRS) to hide the wear-leveling details. The basic idea is to enable an additional intra-row block swap when a new logical address is remapped to the memory row. Experiments demonstrate that IRS can secure NVMs with negligible timing/energy overhead, compared with previous works.

*Download Paper (PDF; Only available from the DATE venue WiFi)*

**IP5-8**  
**EFFECTS OF CELL SHAPES ON THE ROUTABILITY OF DIGITAL MICROFLUIDIC BIOCHIPS**  
**Speaker:** Oliver Keszeöce, University of Bremen, DE  
**Authors:** Kevin Leonard Schneider, Oliver Keszeöce, Jannis Stoppe, Rolf Drechsler  
**Affiliations:** 1) University of Bremen, DE; 2) University of Bremen/DFKI GmbH, DE  
**Abstract:** Digital microfluidic bios (DMFBs) are an emerging technology promising a high degree of automation in laboratory procedures by means of manipulating small discretized amounts of fluids. A crucial part in conducting experiments on biochips is the routing of discretized droplets. While doing so, droplets must not enter each others’ interference region to avoid unintended mixing. This leads to cells in the proximity of the droplet being impassable for others. For different cell shapes, the effect of these temporary blockages varies as the adjacency of cells changes with their shapes. Yet, no evaluation with respect to routability in relation to cell shapes has been conducted so far. This paper analyses and compares various tessellations for the field of cells. Routing benchmarks are mapped to these and the results are compared in order to determine if and how cell shapes affect the performance of DMFBs, showing that certain cell shapes are superior to others.

*Download Paper (PDF; Only available from the DATE venue WiFi)*

**IP5-9**  
**LESS: BIG DATA SKETCHING AND ENCRYPTION ON LOW POWER PLATFORM**  
**Speaker:** Amey Kulkarni, University of Maryland Baltimore County, US  
**Authors:** Amey Kulkarni, Colin Shea, Houman Homayoun and Tinoosh Mohsenin  
**Affiliations:** 1) University of Maryland, Baltimore County, US; 2) University of Maryland Baltimore County, US; 3) George Mason University, US  
**Abstract:** Ever-growing IoT demands big data processing and cognitive computing on mobile and battery operated devices. However, big data processing on low power embedded cores is challenging due to their limited communication bandwidth and on-chip storage. Additionally, IoT and cloud-based computing demand low overhead security kernel to avoid data breaches. In this paper, we propose a Light-weight Encryption using Scalable Sketching (LESS) framework for big data sketching and encryption using One-Time Random Linear Projections (OTRPL). OTRPL encoded matrix makes the Known Plaintext Attacks (KPA) ineffective, and attackers cannot gain significant information from plaintext-ciphertext pair. LESS framework can reduce data up to 67% with 3.81~dB signal-to-reconstruction error rate (SNR). This framework has two important kernels “sketching” and “sketch-reconstruction”, the latter is computationally-intensive and costly. We propose to accelerate the sketch reconstruction using Orthogonal Matching Pursuit (OMP) on a domain specific many-core hardware named Power Efficient Nano Cluster (PENC) designed by authors. Detailed performance and power analysis suggests that PENC platform has 15x and 200x less energy consumption and 8x and 177x faster reconstruction time as compared to low power ARM CPU, and K1 GPU, respectively. To demonstrate efficiency of LESS framework, we integrate it with Hadoop MapReduce platform for objects and scenes identification application. The full hardware integration consists of tiny ARM cores which perform task scheduling and objects identification application, while PENC acts as an accelerator for sketch reconstruction. The full hardware integration results show that the LESS framework achieves 46% reduction in data transfers with very low execution overhead of 0.11% and negligible energy overhead of 0.001% when tested for 2.6GB streaming input data. The heterogeneous LESS framework requires 2x less transfer time and achieves 2.25x higher throughput per watt compared to MapReduce platform.

*Download Paper (PDF; Only available from the DATE venue WiFi)*

**IP5-10**  
**TRUNCAPP: A TRUNCATION-BASED APPROXIMATE DIVIDER FOR ENERGY EFFICIENT DSP APPLICATIONS**  
**Speaker:** Shaghayegh Vahdat, University of Tehran, IR  
**Authors:** Shaghayegh Vahdat, Mehdi Kamali, Ali Azfali-Kusha, Zainalabedin Navabi and Massoud Pedram  
**Affiliations:** 1) University of Tehran, IR; 2) University of Southern California, US  
**Abstract:** In this paper, we present a high speed yet energy efficient approximate divider where the division operation is performed by multiplying the divisor by the inverse of the divisor. In this structure, truncated value of the dividend is multiplied exactly (approximately) by the approximate inverse value of divisor. To assess the efficacy of the proposed divider, its design parameters are extracted and compared to those of a number of prior art dividers in a 45nm CMOS technology. Results reveal that this structure provides 66% and 52% improvements in the area and energy consumption, respectively, compared to the most advanced prior art approximate divider. In addition, delay and energy consumption of the division operation are reduced about 94.4% and 99.93%, respectively, compared to those of an exact SRT radix-4 divider. Finally, the efficacy of the proposed divider in image processing application is studied.

*Download Paper (PDF; Only available from the DATE venue WiFi)*

**IP5-11**  
**TIMING-AWARE WIRE WIDTH OPTIMIZATION FOR SADP PROCESS**  
**Speaker:** Youngsoo Song, KAIST, KR  
**Authors:** Youngsoo Song, Sangmin Kim and Youngsoo Shin, School of Electrical Engineering, KAIST, KR  
**Abstract:** With the scaling of the minimum feature size, RC delay of interconnect is relatively getting more critical in next node technology. SADP is one of the popular processes used in sub-7nm technology. For SADP process, we can increase wire width using patterns formed by block mask, which can reduce wire resistance of critical nets. We determine the direction and length of each wire widening, so that the resulting layout is conflict-free. We convert this as a maximum weight independent set problem and solve this by formulating an ILP. For various test circuits, the wire resistance of critical nets was reduced on average by 18.5%, which led to 9.9% reduction in clock period. The wire width optimization in SADP process can give an insight into timing optimization through the enhancement of fabrication process.

*Download Paper (PDF; Only available from the DATE venue WiFi)*
IPS-12  
**FORMAL TIMING ANALYSIS OF NON-SCHEDULED TRAFFIC IN AUTOMOTIVE SCHEDULED TSN NETWORKS**  
Speaker: Jürgen Teich, Friedrich-Alexander-Universität Erlangen-Nürnberg, DE  
Authors: Fedor Siminov1, Michael Gläf2, Felix Reimann3 and Jürgen Teich1  
1Friedrich-Alexander-Universität Erlangen-Nürnberg, DE; 2Uni University, DE; 3Audi Electronics Venture GmbH, DE  
**Abstract**  
To cope with requirements for low latency, the upcoming Ethernet standard Time-Sensitive Networking (TSN) provides enhancements for scheduled traffic, enabling mixcriticality networks where critical messages are sent according to a system-wide schedule. While these networks provide a completely predictable behavior of the scheduled traffic by construction, timing analysis of the critical non-scheduled traffic with hard deadlines remains an unsolved issue. State-of-the-art analysis approaches consider the interference that unscheduled messages impose on each other, but there is currently no approach to determine the worst-case interference that can be imposed by scheduled traffic, the so-called schedule interference (SI), without relying on restrictions of the shape of the schedule. Considering all possible interference scenarios during each calculation of the SI is impractical, as it results in an explosion of the computation time. As a remedy, this paper proposes a) an approach to integrate the analysis of the worst-case SI into state-of-the-art timing analysis approaches and b) preprocessing techniques that reduce the computation time of the SI calculation by several orders of magnitude without introducing any pessimism.

Download Paper (PDF; Only available from the DATE venue WiFi)

IPS-13  
**ULTRA LOW-POWER VISUAL ODOMETRY FOR NANO-SCALE UNMANNED AERIAL VEHICLES**  
Speaker: Daniele Palossi, ETH Zurich, CH  
Authors: Daniele Palossi1, Andrea Marongiu2 and Luca Benini3  
1ETH - Zurich, CH; 2Swiss Federal Institute of Technology in Zurich (ETHZ), CH; 3Università di Bologna, IT  
**Abstract**  
One of the fundamental functionalities for navigation of Unmanned Aerial Vehicles (UAVs) is the hovering capability. State-of-the-art techniques for implementing hovering on standard-size UAVs process camera stream to determine position and orientation (visual odometry). Similar techniques are considered unaffordable in the context of nano-scale UAVs (i.e. few centimeters of diameter), where the ultra-constrained power-envelopes of tiny rotor-crafts limit the on-board computational capabilities to those of low-power microcontrollers. In this work we study how the emerging ultra-low-power parallel computing paradigm could enable the execution of complex hovering algorithmic flows onto nano-scale UAVs. We provide insight on the software pipeline, the parallelization opportunities and the impact of several algorithmic enhancements. Results demonstrate that the proposed software flow and architecture can deliver unprecedented GOPS/W, achieving 117 frame-per-second within a power envelope of 10 mW.

Download Paper (PDF; Only available from the DATE venue WiFi)

IPS-14  
**LONG RANGE WIRELESS SENSING POWERED BY PLANT-MICROBIAL FUEL CELL**  
Speaker: Maurizio Rossi, University of Trento, IT  
Authors: Maurizio Rossi, Pietro Tosato, Luca Gemma, Luca Torquati, Cristian Catania, Sergio Camalò and Davide Brunelli, University of Trento, IT  
**Abstract**  
Going low power and having a low or neutral impact on the environment is key for embedded systems, as pervasive and wearable consumer electronics is growing. In this paper, we present a self-sustaining, ultra-low power device, supplied by a Plant-Microbial Fuel Cell (PMFC) and capable of smart sensing and long-range communication. The use of a PMFC as a power source is challenging but has many advantages like the only requirement of watering the plant. The system uses aggressive power management thanks to FRAM technology exploited to retain microcontroller status and to shutdown electronics without losing context information. Experimental results show that the proposed system paves the way to energy neutral sensors powered by biosystems available almost anywhere on Earth.

Download Paper (PDF; Only available from the DATE venue WiFi)

IPS-15  
**ON THE COOPERATIVE AUTOMATIC LANE CHANGE: SPEED SYNCHRONIZATION AND AUTOMATIC "COURTESY"**  
Speaker: Alexandre Lombard, UTBM, FR  
Authors: Alexandre Lombard1, Florent Perronet2, Abdeljalil Abbas-Turki3 and Abdellah El-Moudni1  
1UTBM, FR; 2Université de Technologie de Belfort-Montbéliard, FR  
**Abstract**  
The recent ability of some vehicles to handle autonomously the lane change maneuvers, and the progressive equipment of roads and vehicles with ITS-GS units motivate this paper to consider the case of road narrowing that requires a lane change because one lane is occupied by road works for maintenance, incidents and so on. This paper extends the approaches of cooperative speed synchronization at intersections. Because of the complexity of the overall system, it considers each automatic lane change as a mobile (unfixed) intersection in which vehicles synchronize their velocities. The wireless communication allows each vehicle to increase its field of view to negotiate its merging with the other equipped vehicles. Hence, the proposed approach introduces a kind of automatic "courtesy" between equipped vehicles. The paper defines the intersection point between each pair of vehicles and the suited protocol to safely reach the new lane. The protocol can be handled by the new work item (NWI) that has been created at ETSI to realize platooning and cooperative adaptive cruise control. Besides enhancing safety, the simulation results show that the main advantage of the approach is the energy saving by smoothing the traffic.

Download Paper (PDF; Only available from the DATE venue WiFi)

IPS-16  
**EVALUATING MATRIX REPRESENTATIONS FOR ERROR-TOLERANT COMPUTING**  
Speaker: Pareesa Golnari, Princeton University, US  
Authors: Pareesa Ameneh Golnari and Sharad Malik, Princeton University, US  
**Abstract**  
We propose a methodology to determine the suitability of different data representations in terms of their error-tolerance for a given application with accelerator-based computing. This methodology helps match the characteristics of a representation to the data access patterns in an application. For this, we first identify a benchmark of key kernels from linear algebra that can be used to construct applications of interest using any of several widely used data representations. This is then used in an experimental framework for studying the error tolerance of a specific data format for an application. As case studies, we evaluate the error-tolerance of seven data formats on sparse matrix to vector multiplication, diagonal add, and two machine learning applications i) principal component analysis (PCA), which is a statistical technique widely used in data analysis and ii) movie recommendation system with Restricted Boltzmann Machine (RBM) as the core. We observe that the Dense format behaves well for complicated data accesses such as diagonal accessing but is poor in utilizing local memory. Sparse formats with simpler addressing methods and a careful selection of stored information, e.g., CRS and ELLPACK, demonstrate a better error-tolerance for most of our target applications.

Download Paper (PDF; Only available from the DATE venue WiFi)
IP5-17  SIMULATION-BASED DESIGN PROCEDURE FOR SUB 1 V CMOS CURRENT REFERENCE

**Speaker:**
Dmitry Osipov, University of Bremen, DE

**Authors:**
Dmitry Osipov and Steffen Paul, University of Bremen, DE

**Abstract:**
This paper presents a new compact current reference and a simulation-based design procedure to establish the circuit parameters quickly and efficiently. To verify the proposed design procedure, two sub 1-V example circuits for two different reference current values (80 nA and 800 nA) were designed and simulated using 0.35 µm CMOS technology. The circuits are robust against supply voltage variation without the need for external bandgap. A line sensitivity of approximately 1-2%/V over the supply voltage range from sub 1 V is achieved in both cases. The simulated temperature coefficient (TC) values are 93 ppm/°C and 197 ppm/°C in the temperature range from 0°C to 120°C for the 800 nA and 80 nA references, respectively.

Download Paper (PDF; Only available from the DATE venue WiFi)

12.1 Wearable and Smart Medical Devices Day: Industry panel: Industrial challenges for tomorrow's medical devices and tools

**Date:** Thursday 30 March 2017
**Time:** 16:00 - 17:30
**Location / Room:** SBC

**Organisers:**
José L. Ayala, Universidad Complutense de Madrid, ES
Chris Van Hoof, IMEC, BE

**Chair:**
Nick Van Helleputte, IMEC, BE

**Co-Chair:**
José L. Ayala, Universidad Complutense de Madrid, ES

This panel will analyze the Industrial challenges for tomorrow's medical devices and tools. We expect the invited industries to provide their view in how technology, market and users will drive the evolution of medical devices.

**Panelists:**
- Adrian Ionescu, Xsensio, CH
- David Bailey, Sensimed, CH
- Kamiar Aminian, GaitUp, CH
- Carl Van Himbeeck, Cochlear, BE

17:30  End of session

12.2 Advances in Microfluidics and Neuromorphic Architectures

**Date:** Thursday 30 March 2017
**Time:** 16:00 - 17:30
**Location / Room:** 4BC

**Chair:**
Tsung-Yi Ho, National Tsing Hua University, TW

**Co-Chair:**
Li Jiang, Shanghai Jiao Tong University, CN

This session consists of four presentations from emerging applications in EDA such as microfluidics and neural networks. The first presentation proposes a progressive optimization procedure for the synthesis of fault-tolerant flow-based microfluidics. The second presentation presents a hybrid microfluidic platform that enables single-cell analysis on a heterogeneous cells. Next presentation discusses automatic verification on networked labs-on-chip architecture. The final presentation proposes synthesis method for parallel convolutional layers of convolutional neural network.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>16:00</td>
<td>12.2.1</td>
<td>FAST ARCHITECTURE-LEVEL SYNTHESIS OF FAULT-TOLERANT FLOW-BASED MICROFLUIDIC BIOCHIPS</td>
<td>Wei-Lun Huang¹, Ankur Gupta², Sudip Roy², Tsung-Yi Ho¹ and Paul Pop³</td>
</tr>
</tbody>
</table>

¹ National Tsing Hua University, TW; ² Indian Institute of Technology Roorkee, IN; ³ Technical University of Denmark, DK

**Abstract:**
Microfluidic-based lab-on-a-chips have emerged as a popular technology for implementation of different biochemical test protocols used in medical diagnostics. However, in the manufacturing process or during operation of such chips, some faults may occur that leads to damage of the chip, which in turn results in wastage of expensive reagent fluids. In order to make the chip fault-tolerant, the state-of-the-art technique adopts simulated annealing (SA) based approach to synthesize a fault-tolerant architecture. However, the SA method is time consuming and non-deterministic with over-simplified model that usually derive sub-optimal results. Thus, we propose a progressive optimization procedure for the synthesis of fault-tolerant flow-based microfluidic biochips. Simulation results demonstrate that our method is efficient compared to the state-of-the-art techniques and can provide near-optimal and effective solutions in 88% (on average) less CPU time compared to state-of-the-art technique over three benchmark bioproducts.

Download Paper (PDF; Only available from the DATE venue WiFi)
### 12.3 Security Tools

**Date:** Thursday 30 March 2017  
**Time:** 16:00 - 17:30  
**Location:** Room: 2BC

#### 12.3.2 REGISTER TRANSFER LEVEL INFORMATION FLOW TRACKING FOR PROVABLY SECURE HARDWARE DESIGN

**Speaker:** Ryan Kastner, University of California, San Diego, US  
**Authors:** Armati Ardeshiricham, Wei Hu, Joshua Marxen and Ryan Kastner  
**Abstract**  
Information Flow Tracking (IFT) provides a formal methodology for modeling and reasoning about security properties related to integrity, confidentiality, and logical side channel. Recently, IFT has been employed for secure hardware design and verification. However, existing hardware IFT techniques either require designers to rewrite their hardware specifications in a new language or do not scale to large designs due to a low level of abstraction. In this work, we propose Register Transfer Level IFT (RTLIFT), which enables verification of security properties in an early design phase, at a higher level of abstraction, and directly on RTL code. The proposed method enables a precise understanding of all logical flows through RTL design and allows various tradeoffs in IFT precision. We show that RTLIFT achieves over 5x speedup in verification performance as compared to gate level IFT while minimizing the required effort for the designer to verify security properties on RTL designs.

**Download Paper (PDF; Only available from the DATE venue WiFi)**
12.4 Formal and Predictive Models for System Design

**Date:** Thursday 30 March 2017  
**Time:** 16:00 - 17:30  
**Location / Room:** 3A  
**Chair:** Jürgen Teich, Friedrich-Alexander-Universität Erlangen-Nürnberg, DE  
**Co-Chair:** Michael Huebner, Ruhr-University Bochum, DE

The first paper presents a predictive approach to measure the impact of platform changes on the application performance. The second paper introduces a unifying approach for detecting several variable-time cryptographic implementations. We place a prototype implementation of dudect in the public domain.

### 12.4.1 SAMPLING-BASED BINARY-LEVEL CROSS-PLATFORM PERFORMANCE ESTIMATION

**Speaker:** Xinnian Zheng, University of Texas at Austin, US  
**Authors:** Xinnian Zheng, Haris Vikalo, Shuang Song, Lzy K. John and Andreas Gerstlauer, The University of Texas at Austin, US  
**Abstract:** Fast and accurate performance estimation is a key challenge in modern system design. Recently, machine learning-based approaches have emerged that allow predicting the performance of an application on a target platform from executions on a different host. However, existing approaches rely on expensive instrumentation that requires source code to be available. We propose a novel sampling-based, binary-level cross-platform prediction method that accurately predicts performance of a workload on a target by relying on various performance statistics sampled on a host using built-in hardware counters. Compounding the problem, in analog/mixed-signal circuits such information leakage channels can cross the analog/digital or digital/analog interface, making their detection even harder. To this end, in this paper we introduce a PCHIP-based methodology which enables systematic formal evaluation of information flow policies in analog/mixed-signal designs. As we demonstrate, by integrating information flow tracking across the digital and analog domain, our method is able to detect sensitive data leakage from the digital domain to the analog domain and vice versa, without requiring any modification of the current analog/mixed-signal circuit design flow.

**Download Paper (PDF; Only available from the DATE venue WiFi)**

### 12.4.2 A LAYERED FORMAL FRAMEWORK FOR MODELING OF CYBER-PHYSICAL SYSTEMS

**Speaker:** George Ungureanu, KTH Royal Institute of Technology, SE  
**Authors:** George Ungureanu and Ingo Sander, KTH Royal Institute of Technology, SE  
**Abstract:** Designing cyber-physical systems is highly challenging due to its manifold interdependent aspects such as composition, timing, synchronization and behavior. Several formal models exist for description and analysis of these aspects, but they focus mainly on a single or only a few system properties. We propose a formal composable framework which tackles these concerns in isolation, while capturing interaction between them as a single, layered model. This yields a holistic, fine-grained, hierarchical and structured view of a cyber-physical system. We demonstrate the various benefits for modeling, analysis and synthesis through a typical example.

**Download Paper (PDF; Only available from the DATE venue WiFi)**
12.5 Power Modeling, Estimation and Verification

Date: Thursday 30 March 2017
Time: 16:00 - 17:30
Location / Room: 3C
Chair: Pascal Vivet, CEA-Leti, FR
Co-Chair: Hiroshi Nakamura, University of Tokyo, JP

This session covers a wide scope on power modeling and estimation in circuit design. The first paper presents a new model for modeling electromigration in power grid network, taking into account transient effects. The second paper introduces a fast and accurate thermal simulator for 3D circuits, taking into account thermal leakage dependency. The third paper proposes a new identification technique of fine grain power sources for multi-core without the knowledge of the thermal model. The last paper presents rule based checking for quick verification at implementation level of the power intent defined in UPF.

### 12.5.1 PHYSICS-BASED ELECTROMIGRATION MODELING AND ASSESSMENT FOR MULTI-SEGMENT INTERCONNECTS IN POWER GRID NETWORKS

**Speaker:** Xiaoyi Wang, Beijing University of Technology, CN
**Authors:** Xiaoyi Wang¹, Hongyu Wang², Jian He², Sheldon X.-D. Tan³, Yici Cai⁴ and Shengqi Yang²

**Abstract**
Electromigration (EM) is considered to be one of the most important reliability issues for current and future ICs in 10nm technology and below. In this paper we focus on the EM stress evaluation for one-dimensional multi-segment interconnect wires in which all the segments have the same direction, which is a common routing structure for power grid networks. The proposed method, which is based on integral transform technique, could efficiently calculate the hydrostatic stress evolution for multi-segment metal wires stressed with different current densities. The new method can also naturally consider the pre-existing residual stresses coming from thermal or other stress sources. Based on this new transient EM assessment method, a full-chip assessment algorithm for power grid networks is then proposed. The new algorithm is also based on the IR-drop metrics for failure assessment of the power grid networks. However, it finds the precise location and time of EM-induced void nucleation by directly checking the time-changing hydrostatic stresses of all the wires. The resulting EM assessment method can ensure sufficient accuracy of the EM verification for large scale power grid networks without sacrificing the efficiency. The accuracy of the proposed transient analysis approach is validated against the numerical analysis. Also the resulting EM-aware full-chip power grid reliability analysis has been demonstrated and compared with existing methods.

**Download Paper (PDF; Only available from the DATE venue WiFi)**

### 12.5.2 A FAST LEAKAGE AWARE THERMAL SIMULATOR FOR 3D CHIPS

**Speaker:** Hameedah Sultan, IIT Delhi, IN
**Authors:** Hameedah Sultan and Smruti R. Sarangi, IIT Delhi, IN

**Abstract**
In this paper, we propose, 3DSim, which is an ultrafast thermal simulator for 3D chips. It simulates the effects of both dynamic and leakage power. Our technique captures the steady state as well as the transient response with a high speed and good accuracy. 3DSim uses an approach based on Green’s functions, where a Green’s function is defined as the impulse response of a unit power source. Our approach incorporates the effects of the leakage-temperature feedback loop, exploits the radial symmetry in the thermal profile, and uses Hankel transforms to yield a closed form solution for the leakage aware Green’s function. To further speed up the technique, we use fast numerical discrete Hankel transforms, and pre-compute and store certain functions in a lookup table. Our approach fundamentally converts a 3D problem to a set of 1D problems, thus leading to a 68X speedup as compared to competing simulators with an error limited to 1.5C.

**Download Paper (PDF; Only available from the DATE venue WiFi)**
12.6 Efficient design methodologies for high-performance analog circuits and systems.

Date: Thursday 30 March 2017
Time: 16:00 - 17:30
Location / Room: SA

Chair: Nuno Horta, Instituto de Telecomunicacoes, PT
Co-Chair: Deuk Heo, Washington State University, US

This session presents area- and energy-efficient design methodologies for high-performance analog circuits and systems. These papers include an energy-efficient asynchronous digital design method for digitally-assisted analog circuits, a design method of high-density energy storage components, and a robust communication link design for power conversion, and can be traded off for the cost of analog components. A4A flow, A2A interfaces, and Workcraft tools are used for development of power management via multiple power domains can effectively save power by dynamically turning off idle domains. To control domains of a design, introducing low power intent complicates the physical implementation and verification process. During the physical implementation stage, the optimization or manual ECO could be tedious, and error-prone on power/ground signal connections. Therefore, in this paper, we focus on low power rule checking at the physical implementation stage for multiple power domain design. Existing methods adopt an iterative approach, which identifies one error at a time, thus possibly requiring multiple iterations. Different from them, we propose a fast low power rule checking approach to detect all errors at one time. To do so, we separate all paths into inner-domain and cross-domain paths and extract cross-domain net topology before power rule verification. Based on the global topology, we can verify the correctness of connections and detect all errors at the same time. Experimental results show the effectiveness and efficiency of our approach, achieving 3.62X speedups to detect all errors compared with the iterative approach. Moreover, our approach can identify complicated bugs to facilitate subsequent bug fixing.

Download Paper (PDF; Only available from the DATE venue WiFi)

17:00 12.5.3 BENEFITS OF ASYNCHRONOUS CONTROL FOR ANALOG ELECTRONICS: MULTIPHASE BUCK CASE STUDY

Speaker: Danil Sokolov, Newcastle University, GB
Authors: Danil Sokolov1, Vladimir Dubikhin1, Victor Khomenko1, David Lloyd2, Andrey Mokhov1 and Alex Yakovlev1

Abstract: Analogue and mixed signal (AMS) electronics becomes increasingly complex and needs to be digitally enhanced by its own control circuitry. The RTL synthesis flow routed for digital logic is however optimized for synchronous data processing and produces inefficient control for AMS. In this paper we demonstrate the evident benefits of asynchronous circuits in the context of AMS systems, and propose an asynchronous design for analog electronics (A4A) flow for their specification, synthesis, and formal verification. A library of specialized analog-to-asynchronous (A2A) components is developed for interfacing analog signals to asynchronous control. A4A flow is automated in the Workcraft framework and evaluated using a multiphase buck converter case study. The simulation results show improved response time, voltage ripple, and peak current of the buck when controlled asynchronously. These benefits lead to the higher efficiency of power conversion, and can be traded off for the cost of analog components. A4A flow, A2A interfaces, and Workcraft tools are used for development of power converters at Dialog Semiconductor.

Download Paper (PDF; Only available from the DATE venue WiFi)

17:15 12.5.4 FAST LOW POWER RULE CHECKING FOR MULTIPLE POWER DOMAIN DESIGN

Speaker: Iris Hui-Ru Jiang, National Chiao Tung University, TW
Authors: Chien-Pang Lu1 and Iris Hui-Ru Jiang2

Abstract: Power management via multiple power domains can effectively save power by dynamically turning off idle domains. To control domains of a design, introducing low power intent complicates the physical implementation and verification process. During the physical implementation stage, the optimization or manual ECO could be tedious, and error-prone on power/ground signal connections. Therefore, in this paper, we focus on low power rule checking at the physical implementation stage for multiple power domain design. Existing methods adopt an iterative approach, which identifies one error at a time, thus possibly requiring multiple iterations. Different from them, we propose a fast low power rule checking approach to detect all errors at one time. To do so, we separate all paths into inner-domain and cross-domain paths and extract cross-domain net topology before power rule verification. Based on the global topology, we can verify the correctness of connections and detect all errors at the same time. Experimental results show the effectiveness and efficiency of our approach, achieving 3.62X speedups to detect all errors compared with the iterative approach. Moreover, our approach can identify complicated bugs to facilitate subsequent bug fixing.

Download Paper (PDF; Only available from the DATE venue WiFi)

17:30 End of session
The papers in this session propose optimization techniques to improve the lifetime and performance of emerging technologies like persistent memory and scalable many-cores.

Semeen Rehman, Technische Universität Dresden, DE

Co-Chair:

Amit Singh, University of Southampton, GB

Chair:

Location / Room:

Date:

12.7 Software optimization for emerging memory architectures and technologies

12.7.2

16:30

HIGH-DENSITY MOM CAPACITOR ARRAY WITH NOVEL MORTISE-TENON STRUCTURE FOR LOW-POWER SAR ADC

Speaker:
Pang-Yen Chou, Technical University of Munich, DE

Authors:

Nai-Chen Chen, Pang-Yen Chou, Helmut Graeb, and Mark Po-Hung Lin

National Chung Cheng University, TW; Technische Universität München, DE; TU Muenchen, DE

Abstract

The design of capacitor structures have great impact on capacitance density, parasitic capacitance, routability, and matching quality of capacitor network in a SAR ADC, which may affect power, performance, and area of the whole data converter. Most of the recent studies focused on common-centroid placement and routing optimization of the capacitor network. Only few of them investigated the structures of highly integrated capacitors. In this paper, a novel mortise-tenon metal-oxide-metal capacitor structure is proposed, which has the advantages of high capacitance density and small parasitic capacitance. Based on the proposed structure, an integer-linear-programming based capacitor sizing and routing parasitic matching method is further introduced. Experimental results show that the proposed structure and method can achieve the best capacitance density and matching quality of the capacitor network in a SAR ADC.

Download Paper (PDF; Only available from the DATE venue WiFi)

17:00

12.7.3

ADAPTIVE INTERFERENCE REJECTION IN HUMAN BODY COMMUNICATION USING VARIABLE DUTY CYCLE INTEGRATING DDR RECEIVER

Speaker:

Shreyas Sen, Purdue University, US

Authors:

Shovan Maity, Debayan Das and Shreyas Sen

Purdue University, US; ECE, Purdue University, US

Abstract

Connected smart wearable devices are becoming increasingly popular with the advent of cheap, miniaturized, ultra-low-power computing and communication. Human Body Communication (HBC) is emerging as an alternative to Wireless Body Area Network (WBAN) for communication among these devices, as it provides higher energy-efficiency and security. One of the biggest bottlenecks of HBC is the interference picked up due to the human body antenna effect, with Signal to Interference Ratio often worse than -20dB. An interference robust integrating dual data rate (DDR) receiver is introduced which can adapt itself to changing interference conditions and provide high interference rejection by Pulse Width Modulation of integration clock, thus dynamically changing its duty cycle. The theory, architecture of the receiver is developed along with the adaptation algorithm to train the receiver to find the optimum duty cycle of operation. System-level simulations show >20 dB of rejection even in presence of variable interference frequencies.

Download Paper (PDF; Only available from the DATE venue WiFi)

17:30

End of session
PEGASUS: EFFICIENT DATA TRANSFERS FOR PGAS LANGUAGES ON NON-CACHE-COHERENT MANY-CORES

Speaker: Manuel Mohr, Karlsruhe Institute of Technology, DE
Authors: Manuel Mohr and Carsten Tradowsky, Karlsruhe Institute of Technology, DE

Abstract
To improve scalability, some many-core architectures abandon global cache coherence, but still provide a shared address space. Partitioning the shared memory and communicating via messages is a safe way of programming such machines. However, accessing pointered data structures from a foreign memory partition is expensive due to the required serialization.

In this paper, we propose a novel data transfer technique that avoids serialization overhead for pointered data structures by managing cache coherence in software at object granularity. We show that for PGAS programming languages, the compiler and runtime system can completely handle the necessary cache management, thus requiring no changes to application code. Moreover, we explain how cache operations working on address ranges complement our data transfer technique. We propose a novel non-blocking implementation of range-based cache operations by offloading them to an enhanced cache controller. We evaluate our approach on a non-cache-coherent many-core architecture using a distributed-kernel benchmark suite and demonstrate a reduction of communication time of up to 39.8%.

Download Paper (PDF; Only available from the DATE venue WiFi)

Digital-Microfluidic Biochips for Quantitative Analysis: Bridging the Gap between Microfluidics and Microbiology

Speaker: Krishnendu Chakrabarty, Duke University, US
Authors: Mohamed Ibrahim and Krishnendu Chakrabarty, Duke University, US

Abstract
Digital-microfluidics technology has shown considerable promise for advancing sample preparation and point-of-care diagnostics; therefore, it has the potential to transform microbiology and biochemistry research. Over the past decade, a number of microfluidics design-automation techniques have been developed for on-chip droplet manipulation. However, these methods overlook the myriad complexities of biomolecular protocols and they have yet to make a significant impact in biochemistry/microbiology research. A paradigm shift in biochip design automation and a “phase transition” in research are clearly needed to bridge this gap between microfluidics and microbiology. In this paper, we explain how researchers from design-automation and embedded systems can play a key role in this transition. We present a new synthesis flow that uses realistic models of biomolecular protocols and cyberphysical adaptation to address real-world microbiology applications. We also present a list of metrics that can be used for the assessment of design-automation techniques for microbiology applications.

Download Paper (PDF; Only available from the DATE venue WiFi)
<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>16:30</td>
<td>12.8.2</td>
<td>THE CASE FOR SEMI-AUTOMATED DESIGN OF MVLSI BIOCHIPS</td>
<td>Speaker: Jeffrey McDaniel, University of California, Riverside, US Authors: Jeffrey McDaniel, William H. Grover and Philip Brisk, University of California, Riverside, US</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
<td>In recent years, significant interest has emerged in the problem of fully automating the design of microfluidic very large scale integration (mVLSI) chips, a popular class of Lab-on-a-Chip (LoC) devices that can automatically execute a wide variety of biological assays. To date, this work has been carried out with little to no input from LoC designers. We conducted interviews with approximately 100 LoC designers, biologists, and chemists from academia and industry; uniformly, they expressed frustration with existing design solutions, primarily commercially available software such as AutoCAD and Solidworks; however, they expressed limited interest and considerable skepticism about the potential for “push-button” end-to-end automation. In response, we have developed a semi-automated mVLSI drawing tool that is designed specifically to address the pain points elucidated by our interviewees. We have used this tool to rapidly reproduce several previously published LoC architectures and generate fabrication ready specifications.</td>
</tr>
<tr>
<td>17:00</td>
<td>12.8.3</td>
<td>SYNTHESIS OF ON-CHIP CONTROL CIRCUITS FOR MVLSI BIOCHIPS</td>
<td>Speaker: Seetal Potluri, Technical University of Denmark, DK Authors: Seetal Potluri1, Alexander Schneider2, Martin Herslev-Petersen2, Paul Pop2 and Jan Madsen2</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
<td>Microfluidic VLSI (mVLSI) biochips help perform biochemistry at miniaturized scales, thus enabling cost, performance and other benefits. Although biochips are expected to replace biochemical labs, including point-of-care devices, the off-chip pressure actuators and pumps are bulky, thereby limiting them to laboratory environments. To address this issue, researchers have proposed methods to reduce the number of off-chip pressure sources, through integration of on-chip pneumatic control logic circuits fabricated using three-layer monolithic membrane valve technology. Traditionally, mVLSI biochip physical design was performed assuming that all of the control logic is off-chip. However, the problem of mVLSI biochip physical design changes significantly, with introduction of on-chip control, since along with physical synthesis, we also need to (i) perform on/off-chip control partitioning, (ii) on-chip control circuit design and (iii) the integration of on-chip control in the placement and routing design tasks. In this paper we present a design methodology for logic synthesis and physical synthesis of mVLSI biochips that use on-chip control. We show how the proposed methodology can be successfully applied to generate biochip layouts with integrated on-chip pneumatic control.</td>
</tr>
<tr>
<td>17:15</td>
<td>12.8.4</td>
<td>SCHEDULING AND OPTIMIZATION OF GENETIC LOGIC CIRCUITS ON MICROFLUIDIC BIOCHIPS</td>
<td>Speaker: Tsung-Yi Ho, National Tsing Hua University, TW Authors: Yu-Jhih Chen1, Sumit Sharma2, Sudip Roy3 and Tsung-Yi Ho1</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
<td>Synthetic biologists design genetic logic circuit using living cells. A challenge in this task is the difficulty in constructing bigger logic circuits with several living cells due to the crosstalk effect among the biological cells. In order to remove the crosstalk effect, current practice is to use separate chambers on a flow-based microfluidic biochip to isolate each reaction zone. The state-of-the-art technique assumes different reaction times for each gates in a genetic logic circuit. This assumption is pessimistic as each gate has different reaction rate from others. Hence, it will cause unnecessary waiting time for faster gates and this may in turn increase the total experiment completion time significantly. In this paper, we propose a genetic logic circuit synthesis technique for flow-based microfluidic biochip considering different reaction time of each logic gate. Simulation results show that the proposed scheme reduces the total experiment completion time. We further minimize the number of control valves and optimize the routing of flow and control layers in the chip layout, which in turn reduces the design cost.</td>
</tr>
<tr>
<td>17:30</td>
<td></td>
<td>End of session</td>
<td></td>
</tr>
</tbody>
</table>

Source URL: https://past.date-conference.com/date17/booklet/proof_reading