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 Hot Topic 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
  - IP1 Interactive Presentations
    - 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

- UB04 Session 4
  - Exhibition-Reception Exhibition Reception
  - 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
  - IP2 Interactive Presentations

- 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
  - IP3 Interactive Presentations
1. Opening Session: Plenary, Awards Ceremony & Keynote Addresses

**Date:** Tuesday 28 March 2017  
**Time:** 08:30 - 10:30  
**Location / Room:** Auditorium A

**Chair:** David Atienza, EPFL, CH  
**Co-Chair:** Giorgio Di Natale, LIRMM, FR

<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><strong>WELCOME ADDRESSES</strong></td>
<td>Speakers: David Atienza and Giorgio Di Natale</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Speakers:</strong></td>
<td>1 DATE 2017 General Chair, EPFL, CH; 2 DATE 2017 Programme Chair, LIRMM, FR</td>
</tr>
<tr>
<td>08:45</td>
<td>1.1.2</td>
<td><strong>PRESENTATION OF DISTINGUISHED AWARDS</strong></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>
</tr>
</thead>
<tbody>
<tr>
<td>UB01.1</td>
<td>NOXIM-XT: A BIT-ACCURATE POWER ESTIMATION SIMULATOR FOR NOCS</td>
</tr>
</tbody>
</table>

Presenter: Pierre Bomel, Université de Bretagne Sud, FR
Authors: André Rossi1, Johann Laurent2 and Erwan Moreac2
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 ...
**UB01.2** TFA: TRANSPARENT CODE OFFLOADING ON FPGA

**Presenter:**
Roberto Rigamonti, HEIG-VD/HES-SO, CH

**Authors:**
Anthony Convers, Baptiste Delporte, Xavier Ruppen and Alberto Dassati, HEIG-VD/HES-SO, CH

**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 ...**

**UB01.3** DEMONSTRATION OF HW/SW CO-PROCESSING WITH FPGA FOR FAST VISUAL NAVIGATION OF ROVERS

**Presenter:**
Konstantinos Maragos, National Technical University of Athens, GR

**Authors:**
George Lentaris and Dimitrios Soudris, National Technical University of Athens, GR

**Abstract:**
Autonomy, speed and accuracy constitute vital factors for the successful rover-exploration missions. However, the extremely low performance of the on-board space-grade CPUs in conjunction with the increased complexity of the sophisticated computer vision algorithms become a serious bottleneck for fast rover navigation. In this work, we present a HW/SW co-design solution based on FPGA to accelerate visual odometry algorithms tailored to the needs of future Mars exploration missions being scheduled by European Space Agency. For demonstration purposes, we use a Xilinx Kintex-7 FPGA to process images and perform feature detection, description, and matching. The FPGA communicates via ethernet port with the host CPU, which performs filtering and egomotion estimation with absolute orientation. We present the navigation path of a hypothetical moving rover which processes successively stores images acquired by a hypothetical Martian surface while live-recording the CPU-FPGA co-processing.

**More information ...**

**UB01.4** 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 Tatarinov, 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 UVM (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 TT 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 ...**

**UB01.5** 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 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 ...**

**UB01.6** MARGOT: APPLICATION ADAPTATION THROUGH RUNTIME AUTOTUNING

**Presenter:**
Gianluca Palermo, Politecnico di Milano, IT

**Authors:**
Davide Gadioli, Emanuele Vitali and Cristina Silvano, Politecnico di Milano, IT

**Abstract:**
Several classes of applications expose parameters that influence their extra-functional properties, such as the quality of the result or the performance. This leads the application designer to tune these parameters to find the configuration that produces the desired outcome. Given that the application requirements and the resources assigned to each application might vary at runtime, finding a one-fit-all configuration is not a trivial task. For this reason, we implemented the mARGOt framework that enhances an application with an adaptation layer in order to continuously tune the parameters according to the evolving situation. More in detail, mARGOt is composed of a monitoring infrastructure, an application-level adaptation engine and an extra-functional configuration framework based on the separation of concerns paradigm between functional and extra-functional aspects. At the booth, we plan to demonstrate the effectiveness of the proposed infrastructure on three real-life applications.

**More information ...**

**UB01.7** XBARGEN: A TOOL FOR DESIGN SPACE EXPLORATION OF MEMRISTOR BASED CROSSBAR ARCHITECTURES

**Presenter:**
Marcello Traiola, LIRMM, FR

**Authors:**
Mario Barbarescu and Alberto Bosio

1University of Naples Federico II, IT; 2University of Montpellier - LIRMM laboratories, FR

**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 memristors-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 basis of workload dependent and independent parameters, such as area, time and power consumption. The main advantage is that it does not require any simulation and thus it avoid any runtime overheads. The demo aims to show the tool prototype on a selected set of benchmarks which will be synthesized on a memristor-based crossbar circuit.
2.2 Stochastic, Approximate and Neural Computing

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 ...

SEFILE: A SECURE FILESYSTEM VIA SECUBE™

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 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 ...

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

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, the 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 ...

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 address 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

End of session

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.

2.2 Stochastic, Approximate and Neural Computing

Date: Tuesday 28 March 2017
Time: 11:30 - 13:00
Location / Room: 4BC
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.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
<th>Speaker</th>
<th>Abstract</th>
</tr>
</thead>
<tbody>
<tr>
<td>11:30</td>
<td>2.2.1</td>
<td>FRAMEWORK FOR QUANTIFYING AND MANAGING ACCURACY IN STOCHASTIC CIRCUIT DESIGN</td>
<td>Florian Neugebauer, University of Passau, DE</td>
<td>Florian Neugebauer 1, Ilia Polian 1 and John Hayes 2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1University of Passau, DE; 2University of Michigan, US</td>
<td></td>
<td>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.</td>
</tr>
<tr>
<td>12:00</td>
<td>2.2.2</td>
<td>ENERGY-EFFICIENT APPROXIMATE MULTIPLIER DESIGN USING BIT SIGNIFICANCE-DRIVEN LOGIC COMPRESSION</td>
<td>Issa Qiieh, School of Electrical and Electronic Engineering, Newcastle University, GB</td>
<td>Issa Qiieh, Rishad Shafik, Ghaith Tariawneh, Danil Sokolov and Alex Yakovlev, Newcastle University, GB</td>
<td>Approximate arithmetic has recently emerged as a promising paradigm for many imprecision-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 (SDLG) 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).</td>
</tr>
<tr>
<td>12:30</td>
<td>2.2.3</td>
<td>ENERGY-EFFICIENT HYBRID STOCHASTIC-BINARY NEURAL NETWORKS FOR NEAR-SENSOR COMPUTING</td>
<td>Vincent Lee, University of Washington, US</td>
<td>Vincent Lee 1, Armin Alaghi 1, John Hayes 2, Visvesh Sathe 1 and Luis Ceze 1</td>
<td>University of Washington, US; 2University of Michigan, US</td>
</tr>
<tr>
<td>12:45</td>
<td>2.2.4</td>
<td>ACCELERATOR-FRIENDLY NEURAL-NETWORK TRAINING: LEARNING VARIATIONS AND DEFECTS IN RRAM CROSSBAR</td>
<td>Li Jiang, Shanghai Jiao Tong University, CN</td>
<td>Lerong Chen 1, Jiawen Li 1, Yiran Chen 2, Qiuping Deng 3, Jiuyuan Shen 1, Xiaoyao Liang 1 and Li Jiang 4</td>
<td>Shanghai Jiao Tong University, CN; 2University of Pittsburgh, US; 3Lynmax Research, CN; 4Department of Computer Science and Engineering, Shanghai Jiao Tong University, CN</td>
</tr>
</tbody>
</table>
2.3 Cache memory management for performance and reliability

Date: Tuesday 28 March 2017
Time: 11:30 - 13:00
Location / Room: 2BC
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.
<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
<th>Abstract</th>
</tr>
</thead>
</table>
| 11:30   | 2.3.1 | **SHARE LAST-LEVEL CACHE MANAGEMENT FOR GPGPUS WITH HYBRID MAIN MEMORY** | Le 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 GPGPUS** | Mohammad Sadrosadati, Sharif University of Technology, IR; Mohammad Sadrosadati, Amirhossein Mirhosseini, Shahin Roodkhosh, Hazhir Bakhishi and Hamid Sarbazi-Azad, University 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 GPUs 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 approach. 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** | Semeen Rehman, Technische Universitat Dresden, DE; Arun Subramaniyan, Semeen Rehman, Muhammad Shafique, Akash Kumar and Joerg Henkel, 1EECS, University of Michigan-Ann Arbor, US; 2Technische Universitat Dresden, DE; 3Vienna University of Technology (TU Wien), AT; 4Technische Universitaet 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 vulnera-bility 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   | 1 | **DROOP MITIGATING LAST LEVEL CACHE ARCHITECTURE FOR STTRAM** | Swaroop Ghosh, Pennsylvania State University, US; Radha Krishna Aluru and Swaroop Ghosh, University of South Florida, 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, essential 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 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. 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). The various simulation results obtained (both circuit and micro-architectural) comparing our proposed architecture against the conventional are presented in detail. Download Paper (PDF; Only available from the DATE venue WiFi) |
| 13:00   | 2 | **END OF SESSION** | **END OF SESSION** | **END OF SESSION** |

**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
2.5 Reliability and Energy-Efficiency: Two Pillars of NoC Design

**Time** | **Label** | **Presentation Title** | **Authors**
--- | --- | --- | ---
11:30 | 2.4.1 | **GATSIM: ABSTRACT TIMING SIMULATION OF GPUs** | Andreas Gerstlauer, The University of Texas at Austin, US
Authors: 
Kishore Punniyamurthy, Behzad Boroujerdian and Andreas Gerstlauer, The University of Texas at Austin, US
**Abstract**
General-Purpose Graphic Processing Units (GPUs) have become an integral part of heterogeneous system architectures. Ever increasing complexities have made rapid, early performance evaluation of GPU-based architectures and applications a primary design concern. Traditional cycle-accurate GPU simulators are too slow, while existing analytical or source-level estimation approaches are often inaccurate. This paper proposes a novel abstract GPU performance simulation approach that is based on flexible separation of functional and timing models, combining a fast functional execution either on existing simulators or native GPU hardware with a light, fast and abstract timing model. Micro-architecture timing of individual GPU cores is abstracted through static, one-time pre-characterization of code, and only the dynamic scheduling effects are simulated. Using a native GPU for functional execution and excluding pre-characterization, our GPU simulation achieves a throughput of more than 80 MIPS. This is on average 400x faster with 4% error compared to a cycle-accurate GPU simulator for standard GPU benchmarks. Moreover, our simple timing model provides flexibility to target different GPU configurations with little or no extra effort.

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

12:00 | 2.4.2 | **MESAP: A FAST ANALYTIC POWER MODEL FOR DRAM MEMORIES** | Sandeep Poddar, IBM Research, The Netherlands, NL
Authors: 
Sandeep Poddar, Rik Jongerius, Leandro Fiorin, Giovanni Mariani, Gero Dittmann, Andreea Anghel and Henk Corporaal
**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)

12:30 | 2.4.3 | **AFEC: AN ANALYTICAL FRAMEWORK FOR EVALUATING CACHE PERFORMANCE IN OUT-OF-ORDER PROCESSORS** | Kecheng Ji, Southeast University, CN
Authors: 
Kecheng Ji, Ming Ling, Qin Wang, Longxing Shi and Jiaping Pan
**Abstract**
Evaluating cache performance is becoming critically important to predict the overall performance of out-of-order processors. Non-blocking caches, which are very common in out-of-order CPUs, can reduce the average cache miss penalty by overlapping multiple outstanding memory requests and merging different cache misses with the same cacheline address into one memory request. Normally, memory-level-parallelism (MLP) has been used as a metric to describe the concurrency of memory access. Unfortunately, due to the extremely dynamic dependences among the program memory references, it is very difficult to quantify MLP without time-consuming simulations. Moreover, the merging of multiple cache misses, which makes the average cache miss service time less than the physical DDR access latency, is seldom considered in the existing researches. In this paper, we propose a cache performance evaluation framework based on program trace analysis and analytical models to fast estimate MLP and the effective cache miss service time without simulations. Comparing with the results by Gem5 simulations of MobyBench 2.0, Mibench 1.0 and Medabench II, the average accuracy of the modeled MLP and the average cache miss service time is higher than 91% and 92%, respectively. Combined with cache misses calculated by the stack distance theory, the average absolute error of 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.

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

13:00 | 1P1-5, 88 | **MODELING INSTRUCTION CACHE AND INSTRUCTION BUFFER FOR PERFORMANCE ESTIMATION OF VLIW ARCHITECTURES USING NATIVE SIMULATION** | Omaya Matoussi, Grenoble INP, TIMA laboratory, FR
Authors: 
Omaya Matoussi, Frédéric Pétrot, TIMA Laboratory at Grenoble, FR; TIMA Laboratory, Grenoble Institute of Technology, FR
**Abstract**
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.

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

2.5 Reliability and Energy-Efficiency: Two Pillars of NoC Design

**Date:** Tuesday 28 March 2017
**Time:** 11:30 - 13:00
**Location / Room:** 3C
**Chair:** Sebastien Le Beux, Ecole Central du Lyon, FR
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. In the second paper, an online monitoring and routing approach is proposed to address the aging-related degradation in electrical NoCs. Finally, the third paper shows how to use network traffic-aware spatial parallelism to improve the energy efficiency of the Epiphany SoC.

**Abstract**

Siddhartha and Nachiket Kapre

**Authors**

Siddhartha and Nachiket Kapre

**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-fine-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 (Matrix Market datasets), PageRank (BerkStan SNAP dataset), and Izhikevich spiking neural evaluation, we deliver speedups of 6.5-10× while lowering power use by 2× over optimized ARM-based compute. Using our approach, across a range of applications and datasets such as Sparse Matrix-Vector multiplication (Matrix Market datasets), PageRank (BerkStan SNAP dataset), and Izhikevich spiking neural evaluation, we deliver speedups of 6.5-10× while lowering power use by 2× over optimized ARM-based compute. Using our approach, across a range of applications and datasets such as Sparse Matrix-Vector multiplication (Matrix Market datasets), PageRank (BerkStan SNAP dataset), and Izhikevich spiking neural evaluation, we deliver speedups of 6.5-10× while lowering power use by 2× over optimized ARM-based compute.

**Download Paper (PDF): Only available from the DATE venue WiFi**
11:30:30 2.6.1 (Best Paper Award Candidate) ON THE LIMITS OF MACHINE LEARNING-BASED TEST: A CALIBRATED MIXED-SIGNAL SYSTEM CASE STUDY

Speaker: Gildas Leger, Instituto de Microelectronica de Sevilla, IMSE-CNM, (CSIC - Universidad de Sevilla), ES

Authors: Manuel Barragán 1, Gildas Leger 2, Antonio Gines 3, Eduardo Peralas 4 and Adoracion Rueda 3

1TIMA Laboratory, FR; 2Instituto de Microelectronica de Sevilla, IMSE-CNM, (CSIC - Universidad de Sevilla), ES; 3Instituto de Microelectronica de Sevilla, IMSE-CNM, (CSIC-Universidad de Sevilla), ES; 4Institute of Microelectronics of Sevilla, IMSE-CNM, (CSIC-Universidad de Sevilla), ES

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 undoubtably 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)

12:00:00 2.6.2 AN EXTENSION OF CROHN’S SENSITIVITY THEOREM TO MISMATCH ANALYSIS OF 1-PORT RESISTOR NETWORKS

Speaker and Author: Sebastien Cluquinnois, STMicroelectronics, FR

Abstract

An analytical expression of statistical mismatch properties of 1-port resistor networks and associated figure-of-merit is proposed, and related to Cohr’s sensitivity theorem. This expression is then used to demonstrate matching properties of R-ladders. Experimental verification of this formula is done by comparing theoretical results to Monte-Carlo simulations of random R-networks up to 10 resistors, which are generated by a new graph-based algorithm. Further analysis is performed on this figure-of-merit for all generated networks, leading to more insights into matching properties of R-networks.

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

12:30:30 2.6.3 TESTING MICROFLUIDIC FULLY PROGRAMMABLE VALVE ARRAYS (FPVAs)

Speaker: Chunfeng Liu, Technical University of Munich, DE

Authors: Chunfeng Liu 1, Bing Li 2, Bhargab B. Bhattacharya 3, Krishnendu Chakrabarty 4, Tsung-Yi Ho 5 and Ulf Schlüchtermann 6

1Technical University of Munich (TUM), DE; 2TU München (TUM), DE; 3Indian Statistical Institute, IN; 4Duke University, US; 5National Tsing Hua University, TW; 6TU München, DE

Abstract

Fully Programmable Valve Array (FPVA) has emerged as a new architecture for the next-generation flow-based microfluidic biochips. This 2D-array consists of regularly-arranged valves, which can be dynamically configured by users to realize microfluidic devices of different shapes and sizes as well as interconnections. Additionally, the regularity of the underlying structure renders FPVAs easier to integrate on a tiny chip. However, these arrays may suffer from various manufacturing defects such as blockage and leakage in control and flow channels. Unfortunately, no efficient method is yet known for testing such a general-purpose architecture. In this paper, we present a novel formulation using the concept of flow paths and cut-sets, and describe an ILP-based hierarchical strategy for generating compact test sets that can detect multiple faults in FPVAs. Simulation results demonstrate the efficacy of the proposed method in detecting manufacturing faults with only a small number of test vectors.

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

13:00:30 IP1–6, 228 ANALOG FAULT TESTING THROUGH ABSTRACTION

Speaker: Enrico Fraccaroli, Università degli Studi di Verona, IT

Authors: Enrico Fraccaroli and Franco Fummi, Università degli Studi di Verona, IT

Abstract

Despite analog SPICE-like simulators have reached their maturity, most of them were not originally conceived for simulating faulty circuits. With the advent of smart systems, fault testing has to deal with models encompassing both analog and digital blocks. Due to their complexity, the industry is still lacking of effective testing approaches for these analog and mixed-signal (AMS) models. The current problem is the computational time required for implementing an analog fault simulation campaign. To this end, the work presented in this paper is an automatic procedure which: 1) injects faults in an analog circuit, 2) abstracts both faulty and fault-free 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.

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

13:01:00 IP1–7, 65 BISCC: EFFICIENT PRE THROUGH POST SILICON VALIDATION OF MIXED-SIGNAL/RF SYSTEMS USING BUILT IN STATE CONSISTENCY CHECKING

Speaker: Abhijit Chatterjee, Georgia Institute of Technology, US

Authors: Sabyasachi Deyati 1, Barry Muldrey 2 and Abhijit Chatterjee 2

1Georgia Institute of Technology, US; 2Georgia Tech, US

Abstract

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 observability of internal digital circuit states, analog scan chains suffer from signal integrity, bandwidth and circuit loading issues. In this paper, we propose a novel technique based on built-in state consistency checking that allows both pre as well as post-silicon validation of mixed-signal/RF systems without the need to rely on manually generated checks. The method is supported by a design-for-validation (DFV) methodology which systematically injects a minimum amount of circuitry into mixed-signal systems for design bug detection and diagnosis purposes. The core idea is to apply two spectrally diverse stimuli to the circuit under test (CUT) in such a way that they result in the same circuit state (observed voltage/current values at internal or external circuit nodes). By comparing the resulting state values, design bugs are detected efficiently without the need for manually generated checks. No assumption is made about the nature of the detected bugs; the stimulus applied is steered towards those that are the most likely to detect design bugs. Test cases for both pre and post-silicon design bug detection and diagnosis prove the viability of the proposed BISCC approach.

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

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>13:00</td>
<td>End of session</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lunch Break in Garden Foyer</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Keynote Lecture session 3.0 in &quot;Garden Foyer&quot; session 1350 - 1420</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lunch Break in the Garden Foyer</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>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.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**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

Covering 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.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
</table>
| 11:30  | 2.7.1 HARPA: TACKLING PHYSICALLY INDUCED PERFORMANCE VARIABILITY | Speaker: Dimitrios Soudris, ICCS, GR  
Authors: Nikolaos Zompakis 1 and Dimitrios Soudris 2  
1ICCS/NTUA, GR; 2NTUA, 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 | Speaker: Leonidas Kosmidis, Barcelona Supercomputing Center and Universitat Politècnica de Catalunya, ES  
Authors: Leonidas Kosmidis 1, Jaume Abella 2 and Francisco Cazorla 3  
1Barcelona Supercomputing Center and Universitat Politècnica de Catalunya, ES; 2Barcelona Supercomputing Center (BSC-CNS), ES; 3Barcelona 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 randomization (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 | Speaker: Per Gunnar Kjeldsberg, Norwegian University of Science and Technology, NO  
Authors: Per Gunnar Kjeldsberg 1, Andreas Gocht 2, Michael Gerdt 3, Riha Lubomir 4, Joseph Schuchart 5 and Umbreen Sabir Mian 6  
1Norwegian University of Science and Technology, NO; 2Technische Universität Dresden, DE; 3Technische Universität München, DE; 4IT4Innovations, Ostrava, CZ; 5Universität 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 alternative 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) |
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 CMS ageing. New defect classes contributing to NFF have been identified, including timing related faults (TRF) at board level and intermittent 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 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)

13:00 End of session
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.

2.8a 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.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>11:30</td>
<td>2.8a.1</td>
<td>HEARRESTORE</td>
<td>Juan Ansó, UniBE, CH</td>
</tr>
<tr>
<td>11:50</td>
<td>2.8a.2</td>
<td>PATLISCI II</td>
<td>Hans Peter Lang, UniBAS, CH</td>
</tr>
<tr>
<td>12:10</td>
<td>2.8a.3</td>
<td>FLUSITEX</td>
<td>Daniel Ahmed, ETHZ, CH</td>
</tr>
<tr>
<td>12:30</td>
<td></td>
<td>End of session</td>
<td></td>
</tr>
<tr>
<td>13:00</td>
<td></td>
<td>Lunch Break in Garden Foyer</td>
<td></td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.8a.1</td>
<td>HEARRESTORE</td>
<td>Juan Ansó, UniBE, CH</td>
</tr>
<tr>
<td>12:10</td>
<td>2.8a.3</td>
<td>FLUSITEX</td>
<td>Daniel Ahmed, ETHZ, CH</td>
</tr>
<tr>
<td>13:00</td>
<td></td>
<td>End of session</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lunch Break in Garden Foyer</td>
<td></td>
</tr>
</tbody>
</table>

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.
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 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.

More information ...

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 increasingly sophisticated technologies such as RFID tags, Wireless Sensor Nodes and mobile devices.

Using the development of a Smart Medical 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.

More information ...

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.

More information ...

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 ...

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.

More information ...

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.

More information ...

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 ...

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.

More information ...

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.

More information ...
UB02.4 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 GPUs 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 ...

UB02.5 A VOLTAGE-SCALABLE FULLY DIGITAL ON-CHIP MEMORY FOR ULTRA-LOW-POWER 10T PROCESSORS
Presenter:
Jun Shiomi, Kyoto University, JP
Authors:
Tohru Ishihara and Hitodeshi 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 that 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 lithium battery as the ambient energy source.
More information ...

UB02.6 MARGOT: APPLICATION ADAPTATION THROUGH RUNTIME AUTOTUNING
Presenter:
Gianluca Palermo, Politecnico di Milano, IT
Authors:
Davide Gadioli, Emanuele Vitali and Cristina Silvano, Politecnico di Milano, IT
Abstract:
Several classes of applications expose parameters that influence their extra-functional properties, such as the quality of the result or the performance. This leads the application designer to tune these parameters to find the configuration that produces the desired outcome. Given that the application requirements and the resources assigned to each application might vary at runtime, finding a one-fit-all configuration is not a trivial task. For this reason, we implemented the mARGOt framework that enhances an application with an adaptation layer in order to continuously tune the parameters according to the evolving situation. More in detail, mARGOt is composed of a monitoring infrastructure, an application-level adaptation engine and an extra-functional configuration framework based on the separation of concerns paradigm between functional and extra-functional aspects. At the booth, we plan to demonstrate the effectiveness of the proposed infrastructure on three real-life applications.
More information ...

UB02.7 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 ...

UB02.8 TIDES: NON-LINEAR WAVEFORMS FOR QUICK TRACE NAVIGATION
Presenter:
Jannis Stoppe, University of Bremen, DE
Author:
Rolf 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 ...

UB02.9 SEFILE: A SECURE FILESYSTEM IN USERSPACE VIA SECUCE™
Presenter:
Giuseppe Airofarulla, CINI, IT
Authors:
Paolo Prinetto and Antonio Varialle
1CINI & Politecnico di Torino, IT; 2Bu5 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 ...

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 commandsmsgs/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
**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.

---

**Time** | **Label** | **Presentation Title** | **Authors**
--- | --- | --- | ---
14:30 | 3.1.1 | BETTER THAN WORST CASE SIGNOFF STRATEGIES FOR LOW POWER IOT DEVICES | 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 process spread (e.g. -1σ corner), tighter voltage margin (e.g. Vdd-5%) and typical operating temperature instead of targeting the worst-case PVT corner (e.g. -3σ corner, Vdd-10%, 125oC). 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 | Eric Flamand, GreenWaves Technologies, FR

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

---

**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>
</table>
| 14:30 | 3.2.1 | **BEYOND-CMOS NON-BOOLEAN LOGIC BENCHMARKING: INSIGHTS AND FUTURE DIRECTIONS**  
Speaker: Azad Naeemi, Georgia Institute of Technology, US  
Authors: Chenyun Pan and Azad Naeemi, Georgia Institute of Technology, US  
Abstract: 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.  
Download Paper (PDF; Only available from the DATE venue WiFi) |
| 15:00 | 3.2.2 | **UNDERSTANDING THE DESIGN OF IBM NEUROSYNAPTIC SYSTEM AND ITS TRADEOFFS: A USER PERSPECTIVE**  
Speaker: Yiran Chen, Duke University, US  
Authors: Hsin-Pai Cheng, Wei Wen, Chunpeng Wu, Sicheng Li, Hai (Helen) Li and Yiran Chen, University of Pittsburgh, US  
Abstract: As a large-scale commercial spiking-based neuromorphic computing platform, IBM TrueNorth processor received 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 regulation 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 pro-cessor 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 through-put. But if we consider energy consumption, TrueNorth processor is the most energy efficient one, say, > 6000 frames/sec/Watt.  
Download Paper (PDF; Only available from the DATE venue WiFi) |
3.2.3 CELLULAR NEURAL NETWORK FRIENDLY CONVOLUTIONAL NEURAL NETWORKS - CNNS WITH CNNS

Speaker: Michael Niemer, University of Notre Dame, US
Authors: András Horváth\textsuperscript{1}, Michael Hillmer\textsuperscript{2}, Qiuwen Lou\textsuperscript{2}, X, Sharon Hu\textsuperscript{2} and Michael Niemer\textsuperscript{2}

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)

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\textsuperscript{1}, Sunhee Kong\textsuperscript{1}, Boeui Hong\textsuperscript{1}, Lei Xu\textsuperscript{2}, Weidong Shi\textsuperscript{2} and Taeweon Suh\textsuperscript{1}

Abstract
Increasing complexity of integrated circuits and IP-based hardware designs has 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)
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. To overcome the limitations of traditional hardware Trojan detection techniques, which are often based on golden circuit reference models, we introduce a novel methodology that relies on power profiling of microcontroller instruction sets.

**Abstract**

Malicious software, referred to 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

Faiq Khalid Lodhi, College of Engineering / Department of Electrical Engineering, UAE University, AE

Fatma Nur Esirci, Gebze Technical University, TR

Princeton University, US

Intel corp, US

**Presentation Title**

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

**Download Paper (PDF; Only available from the DATE venue WiFi)**
lifetime. This is followed by several approximate computation techniques that optimize the energy consumption of the circuits. 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.

### 3.4.1 OPTIMIZING TEMPERATURE GUARDBANDS

**Speaker:** Hussam Amrouch, Karlsruhe Institute of Technology (KIT), DE

**Authors:** Hussam Amrouch¹, Behnam Khaleghi² and Joerg Henkel¹

1Karlsruhe Institute of Technology, DE; 2Sharif University of Technology, IR

**Abstract**
We introduce the first temperature guardbands optimization based on thermal-aware logic synthesis and thermalaware timing analysis. The optimized guardbands are obtained solely due to using our so-called thermal-aware cell libraries together with existing tool flows and not due to sacrificing timing constraints (i.e. no trade-offs). We demonstrate that temperature guardbands can be optimized at design time through thermalaware logic synthesis in which more resilient circuits against worst-case temperatures are obtained. Our static guardband optimization leads to 18% smaller guardbands on average. We also demonstrate that thermal-aware timing analysis enables designers to accurately estimate the required guardbands for a wide range of temperatures without over/under-estimations. Therefore, temperature guardbands can be optimized at operation time through employing the small, yet sufficient guardband that corresponds to the current temperature rather than employing throughout a conservative guardband that corresponds to the worst-case temperature. Our daptive guardband optimization results, on average, in a 22% higher performance along with 9.2% less energy. Neither thermal-aware logic synthesis nor thermal-aware timing analysis would be possible without our thermal-aware cell libraries. They are compatible with use of existing commercial tools. Hence, they allow designers, for the first time, to automatically consider thermal concerns within their design tool flows even if they were not designed for that purpose.

Download Software: This work is publicly available at http://ces.tec.kit.edu/dependable-hardware.php

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

### IP1-3.4.2 THE HIDDEN COST OF FUNCTIONAL APPROXIMATION AGAINST CAREFUL DATA SIZING: A CASE STUDY

**Speaker:** Benjamin Barrois, University of Rennes 1 / IRISA, FR

**Authors:** Benjamin Barrois¹, Olivier Senteley² and Daniel Menard³

¹University of Rennes - INRIA, FR; ²INRIA, FR; ³INSA Rennes, FR

**Abstract**
Many applications are error-resilient, allowing for the introduction of approximations in the calculations, as long as a certain accuracy target is met. Traditionally, fixed-point arithmetic is used to relax accuracy, by optimizing the bit-width. This arithmetic leads to important benefits in terms of delay, power and area. 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 metrics and by emphasizing the issue of data sizing. To perform this analysis, we developed a design exploration framework, APXPERF, which guarantees that all operators are compared using the same operating conditions. Moreover, operators are compared in several classical real-life applications leveraging 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.

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

### 3.4.3 HIGH-LEVEL SYNTHESIS OF APPROXIMATE HARDWARE UNDER JOINT PRECISION AND VOLTAGE SCALING

**Speaker:** Seogoo Lee, The University of Texas at Austin, US

**Authors:** Seogoo Lee¹, Lizy John² and Andreas Gerstlauer¹

¹The University of Texas at Austin, US; ²UT Austin, US

**Abstract**
In recent years, approximate computing has emerged as a promising approach to trade off quality of computed outputs for energy savings. In this paper, we present an approximate high-level synthesis (A HLS) approach that outputs a quality-energy optimized register-transfer-level implementation from an accurate high-level C description. Existing AHLS work only considers switching activity for energy savings under hardware approximations. By contrast, we aim to provide a general AHLS solution that also considers voltage scaling given a reduced processing time. To maximize voltage and associated energy reductions, we include both operation-level approximations by bit rounding and more aggressive operation eliminations as approximation techniques. Optimally exploiting scaling opportunities under such approximations requires tight interaction with scheduling tasks. We address this problem by combining an optimization pass that estimates the scheduling impact of approximations with fast yet accurate quality-energy models and an efficient optimization solver to find near-optimal solutions 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.

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

### 16:00 ACCOUNTING FOR SYSTEMATIC ERRORS IN APPROXIMATE COMPUTING

**Speaker:** Martin Brustel, Technical University Dresden, DE

**Authors:** Martin Brustel¹ and Akash Kumar²

¹Technical University Dresden, DE; ²Technische Universitaet 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)
16:01 IP1-14, 467 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)

16:02 IP1-15, 215 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) solving 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 solving 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)

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

3.5 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: Sanchari Sen, Purdue University, US
Authors: Sanchari Sen, Swagath Venkataramani and Anand Raghunathan, Purdue 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 efficiencies 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 Saiabal 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: Qiwen Chen and 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 detection 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 10^8 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: Parami Wijesinghe, Purdue University, US
Authors: Parami Wijesinghe, Chamika Liyanagedera 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 (RBFN), 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)
Coffee Break

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>
</tr>
</thead>
<tbody>
<tr>
<td>14:30</td>
<td>3.6.1</td>
<td>CHARKA: A RELIABILITY-AWARE TEST SCHEME FOR DIAGNOSIS OF CHANNEL SHORTS BEYOND MESH NOCS</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker: Santosh Biswas, IIT Guwahati, IN</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Authors: Biswajit Bhattacharyya, Jatindra Kumar Deka and Santosh Biswas</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1IIT Guwahati, IN; 2IIT Guwahati, IN</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abstract: 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.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Download Paper (PDF; Only available from the DATE venue WiFi)</td>
</tr>
</tbody>
</table>

| 15:00 | 3.6.2 | RECOVERY-AWARE PROACTIVE TSV REPAIR FOR ELECTROMIGRATION IN 3D ICs |
|       |       | Speaker: Shengcheng Wang, Chair of Dependable Nano Computing (CDNC), Karlsruhe Institute of Technology (KIT), DE |
|       |       | Authors: Shengcheng Wang1, Hengyang Zhao1, Sheldon Tan2 and Mehdi Tahoori 1 |
|       |       | 1Karlsruhe Institute of Technology, DE; 2University of California, Riverside, US |
|       |       | Abstract: 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) |

| 15:30 | 3.6.3 | NEAR-OPTIMAL METASTABILITY-CONTAINING SORTING NETWORKS |
|       |       | Speaker: Johannes Bund, Saarland University, DE |
|       |       | Authors: Johannes Bund1, Christoph Lenzen2 and Moti Medina2 |
|       |       | 1Saarland University, DE; 2MPI-INF, DE |
|       |       | Abstract: 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 & 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(log 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 > 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) |
3DFAR: A THREE-DIMENSIONAL FABRIC FOR RELIABLE MULTICORE PROCESSORS

Valeria Bertacco, University of Michigan, US
Javad Bagherzadeh and Valeria Bertacco, University of Michigan, US

Abstract
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.

EVALUATING IMPACT OF HUMAN ERRORS ON THE AVAILABILITY OF DATA STORAGE SYSTEMS

Hossein Asadi, Sharif University of Technology, IR
Mostafa Kishani, Reza Eftekhari and Hossein Asadi, Sharif University of Technology, IR

Abstract
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.

THE CONCEPT OF UNSCHEDULABILITY CORE FOR OPTIMIZING PRIORITY ASSIGNMENT IN REAL-TIME SYSTEMS

Yecheng Zhao, Virginia Polytechnic Institute and State University, US
Yecheng Zhao and Haibo Zeng, Virginia Tech, US

Abstract
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 a 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).
3.8 Addressing Challenges in Today's Datacenter Systems' Design
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.

15:15 3.8.2 PROTECTING DATA IN FA RM 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.
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
More information ...

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

Authors:
Tohru Ishihara and Hitoshito 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 ...

RUNNING CONVOLUTIONAL LAYERS OF ALEXNET IN NEUROMORPHIC COMPUTING SYSTEM

Authors:
Yongshin Kang, 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 ...

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

Authors:
Alessandro de Gennaro, Newcastle University, GB

Abstract
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 an FPGA for the analysis: thousand times faster than in software. Single stand will be used for both case studies.

More information ...

NETWORKED LABS-ON-CHIPS

Authors:
Andreas Grimmer, Johannes Kepler University Linz, AT

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 an FPGA for the analysis: thousand times faster than in software. Single stand will be used for both case studies.

More information ...

STACKADROP: A MODULAR DIGITAL MICROFLUIDIC BIOCHIP RESEARCH PLATFORM

Authors:
Oliver Keszööze, University of Bremen, 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 WheelElabs 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 ...

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

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 ...

RUNNING CONVOLUTIONAL LAYERS OF ALEXNET IN NEUROMORPHIC COMPUTING SYSTEM

Authors:
Yongshin Kang, 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 ...

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

Authors:
Alessandro de Gennaro, Newcastle University, GB

Abstract
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 an FPGA for the analysis: thousand times faster than in software. Single stand will be used for both case studies.

More information ...

NETWORKED LABS-ON-CHIPS

Authors:
Andreas Grimmer, Johannes Kepler University Linz, AT

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 an FPGA for the analysis: thousand times faster than in software. Single stand will be used for both case studies.

More information ...

STACKADROP: A MODULAR DIGITAL MICROFLUIDIC BIOCHIP RESEARCH PLATFORM

Authors:
Oliver Keszööze, University of Bremen, 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 WheelElabs 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 ...

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

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

Francesco Conti, ETH Zurich, CH

Authors:
Stefan Mach1, Florian Zaruba1, Antonio Pullini1, Daniele Palossi1, Giovanni Rovere1, Florian Glaser1, Germain Haugou1, Schiekke Fateh1 and Luca Benni2

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 ...
**IP1-4**

**DROOP MITIGATING LAST LEVEL CACHE ARCHITECTURE FOR STTRAM**

**Speaker:**
Swaroop Ghosh, Pennsylvania State University, US

**Authors:**
Radha Krishna Aluru and Swaroop Ghosh

*1*University of South Florida, US; *2*Pennsylvania 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, 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 drop, 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). The various simulation results obtained (both circuit and micro-architectural) comparing our proposed architecture against the conventional are presented in detail.

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

**IP1-5**

**MODELING INSTRUCTION CACHE AND INSTRUCTION BUFFER FOR PERFORMANCE ESTIMATION OF VLIW ARCHITECTURES USING NATIVE SIMULATION**

**Speaker:**
Omayma Matoussi, Grenoble INP, TIMA laboratory, FR

**Authors:**
Omayma Matoussi*1* and Frédéric Pétrot*2*

*1*Tima Laboratory at Grenoble, FR; *2*TIMA Laboratory, Grenoble Institute of Technology, FR

**Abstract**

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.

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

**IP1-6**

**ANALOG FAULT TESTING THROUGH ABSTRACTION**

**Speaker:**
Enrico Fraccaroli, Università degli Studi di Verona, IT

**Authors:**
Enrico Fraccaroli and Franco Fummi, Università degli Studi di Verona, IT

**Abstract**

Despite analog SPICE-like simulators have reached their maturity, most of them were not originally conceived for simulating faulty circuits. With the advent of smart systems, fault testing has to deal with models encompassing both analog and digital blocks. Due to their complexity, the industry is still lacking of effective testing approaches for these analog and mixed-signal (AMS) models. The current problem is the computational time required for implementing an analog fault simulation campaign. To this end, the work presented in this paper is an automatic procedure which: 1) injects faults in an analog circuit, 2) abstracts both faulty and fault-free models from the circuit to the functional level, 3) builds an efficient fault simulation framework. The processes of fault injection, faulty model abstraction and verification 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.

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

**IP1-7**

**BISCC: EFFICIENT PRE THROUGH POST SILICON VALIDATION OF MIXED-SIGNAL/RF SYSTEMS USING BUILT IN STATE CONSISTENCY CHECKING**

**Speaker:**
Abhijit Chatterjee, Georgia Institute of Technology, US

**Authors:**
Sabyasachi Deyati, Barry Muldrey 1 and Abhijit Chatterjee

*1*Georgia Institute of Technology, US; *2*Georgia Tech, US

**Abstract**

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 observability of internal digital circuit states, analog scan chains suffer from signal integrity, bandwidth and circuit loading issues. In this paper, we propose a novel technique based on built-in state consistency checking that allows both pre as well as post-silicon validation of mixed-signal/RF systems without the need to rely on manually generated checks. The method is supported by a design-for-validation (DFV) methodology which systematically inserts a minimum amount of circuitry into mixed-signal systems for design bug detection and diagnosis purposes. The core idea is to apply two spectrally diverse stimuli to the circuit under test (CUT) in such a way that they result in the same circuit state (observed voltage/current values at internal or external circuit nodes). By comparing the resulting state values, design bugs are detected efficiently without the need for manually generated checks. No assumption is made about the nature of the detected bugs; the stimulus applied is steered towards those that are the most likely to detect design bugs. Test cases for both pre and post-silicon design bug detection and diagnosis prove the viability of the proposed BISCC approach.

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

**IP1-8**

**COMPUTING WITH NANO-CROSSBAR ARRAYS: LOGIC SYNTHESIS AND FAULT TOLERANCE**

**Speaker:**
Mustafa Altun, Istanbul Technical University, TR

**Authors:**
Mustafa Altun, Valentina Ciriani 1 and Mehdi Tahoori

*1*Istanbul Technical University, TR; *2*University of Milan, IT; *3*Karlsruhe Institute of Technology, DE

**Abstract**

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 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 integrate well developed conventional circuit design techniques into nano-crossbar arrays. Motivated by this, our project aims to develop a complete synthesis and performance optimization methodology for switching nano-crossbar arrays that leads to the design and construction of an emerging nanocomputer. First two work 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 area optimization, and fault tolerance that aims to provide a full methodology in the presence of high fault densities and extreme parametric variations in nano-crossbar architectures.

**Download Paper (PDF); Only available from the DATE venue WiFi**
IP1-9 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)

IP1-10 WCET-AWARE PARALLELIZATION OF MODEL-BASED APPLICATIONS FOR MULTI-CORES: THE ARGO APPROACH
Speaker: Steven Derrien, Universite de Rennes, FR
Authors:
Steven Derrien1, Isabelle Puaut2, Panayiotis Aliefargsi3, Marcus Bednara4, Harald Bucher5, Clément David6, Yann Debray5, Umut Durak7, Imen Fassi2, Christian Ferdinand8, Damien Hardy2, Angeliki Kritikakou2, Gerard Rauwerda9, Simon Reder5, Martin Sick5, Timo Stripf5, Kim Sunesen5, Timon ter Braak9, Nikolaos Voros2 and Jürgen Becker5
1IRISA, FR; 2University of Rennes 1 / IRISA, FR; 3TWG, GR; 4iSIS/Franshofer, DE; 5Karlsruhe Institute of Technology, DE; 6Scilab, FR; 7DLR, DE; 8Absint, FR; 9Recure 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)

IP1-11 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 Marien 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)

IP1-12 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 Awwad2
1School of Electrical Engineering and Computer Science National University of Sciences and Technology (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 IoT, most of the system-on-chip (SoC) do not have a golden reference. The SoC reconfigurable systems 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 micro-controllers instruction set, 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 MCB8051 micro-controller in VHDL, obtained the power profile of its instruction set and then applied deep learning, k-NN, decision tree and naive Bayesian based machine learning tools to train the models. The cross validation comparison of these learning algorithm, when applied to MCB8051 Trojan benchmarks, shows that we can achieve 87% to 99% accuracy. To the best of our knowledge, this is the first work in which the power profile of a microprocessor's instruction set is used in conjunction with machine learning for runtime HT detection.
Download Paper (PDF; Only available from the DATE venue WiFi)

IP1-13 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 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 Forsberg¹, Andrea Marongiu² and Luca Benini³
¹ETH Zürich, CH; ²Swiss Federal Institute of Technology in Zurich (ETHZ), CH; ³Università di 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 three different types of system 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.

Time | Label | Presentation Title
--- | --- | ---
17:00 | 4.1.1 | ENERGY-PERFORMANCE OPTIMIZED DESIGN OF SILICON PHOTONIC INTERCONNECT NETWORKS FOR HIGH-PERFORMANCE COMPUTING
Speaker and Author: Keren Bergman, Columbia University, US
Authors: Meisam Bahadori1, Sebastien Rumley1, Robert Polster1, Alexander Gazman1, Matt Traverso2, Mark Webster2, Kaushik Patel1 and Keren Bergman1
Abstract: We present detailed electrical and optical models of the elements that comprise a WDM silicon photonic link. The electronics is assumed to be based on 65 nm CMOS node and the optical modulators and demultiplexers are based on microring resonators. The goal of this study is to analyze the energy consumption and scalability of the link by finding the right combination of (number of channels X data rate per channel) that fully covers the available optical power budget. Based on the set of empirical and analytical models presented in this work, a maximum capacity of 0.75 Tbps can be envisioned for a point-to-point link with an energy consumption of 1.9 pJ/bit. Sub-pJ/bit energy consumption is also predicted for aggregated bitrates up to 0.35 Tbps.

17:30 | 4.1.2 | RAPID GROWTH OF IP TRAFFIC IS DRIVING ADOPTION OF SILICON PHOTONICS IN DATA CENTERS
Speaker and Author: Kaushik Patel, Cisco Systems, US
Abstract: With the dramatic growth in consumers using Mobile plus Video data and the corresponding increase in IP traffic, more Data Centers are required together with a need to scale the capacity within the Data Centers. Moore’s law continues to push advances in CMOS technology enabling the design of larger capacity ASICs used to build Switches and Routers in the Data Centers. The cost, power dissipation and face plate optical density challenges are being solved by Silicon Photonics deployed in smaller form factor pluggable optics with a longer term transition to embedded optics. This march towards higher data rates, lower cost and lower power dissipation requires major advances in the cost, volume wafer manufacturing, optical packaging and test for Silicon Photonics based products. The focus of this talk will be on how Cisco is addressing these multiple development and manufacturing challenges as Silicon Photonics based products are released in the market.

18:00 | 4.1.3 | GENERATION OF COMPLEX QUANTUM STATES VIA INTEGRATED FREQUENCY COMBS
Speaker: Roberto Morandotti, INRS-EMT, CA
Authors: Christian Reimer1, Michael Kues2, Piotr Roztocki2, Benjamin Wetzel2, Brent E. Little4, Sai T. Chu5, Luca Caspani5, David J. Moss5 and Roberto Morandotti1
1INRS-EMT, CA; 2INRS-EMT & University of Glasgow, CA; 3INRS-EMT & University of Sussex, CA; 4Xian Institute of Optics and Precision Mechanics, CN; 5City University of Hong Kong, CN; 6University of Strathclyde, GB; 7Swinburne University of Technology, AU
Abstract: The generation of optical quantum states on an integrated platform will enable low cost and accessible advances for quantum technologies such as secure communications and quantum computation. We demonstrate that integrated quantum frequency combs (based on high-Q microring resonators made from a CMOS-compatible, high refractive-index glass platform) can enable, among others, the generation of heralded single photons, cross-polarized photon pairs, as well as bi- and multi-photon entangled qubit states over a broad frequency comb covering the S, C, L telecommunications band, constituting an important cornerstone for future practical implementations of photonic quantum information processing.

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: 4BC
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.

**Time** | **Label** | **Presentation Title**
--- | --- | ---
17:00 | 4.2.1 | **EXPLOITING TRANSISTOR-LEVEL RECONFIGURATION TO OPTIMIZE COMBINATIONAL CIRCUITS**
**Speaker:** Michael Raizt a, Technische Universität Dresden, DE
**Authors:** Michael Raizta1, Jens Trommer2, Akash Kumar3, Marcus Völpe4, Dennis Walter5, Walter Weber6 and Thomas Mikolajick7
1Technische Universität Dresden and CFAED, DE; 2Namlab gGmbH, DE; 3Technische Universität Dresden, DE; 4SNT University of Luxembourg, LU; 5Technische Universität Dresden, DE; 6NamLab gGmbH and CFAED, DE; 7NamLab GmbH / TU Dresden, DE
**Abstract**
Silicon nanowire reconfigurable field effect transistors (SiNW RFETs) abolish the physical separation of n-type and p-type transistors by taking up both roles in a configurable way within a doping-free technology. However, the potential of transistor-level reconfigurability has not been demonstrated in larger circuits, so far. In this paper, we present first steps to a new compact and efficient design of combinational circuits by employing transistor-level reconfiguration. We contribute new basic gates realized with silicon nanowires, such as 2/3-XOR and MUX gates. Exemplifying our approach with 4-bit, 8-bit and 16-bit conditional carry adders, we were able to reduce the number of transistors to almost one half. With our current case study we show that SiNW technology can reduce the required chip area by 16%, despite larger size of the individual transistor, and improve circuit speed by 26%.

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

17:30 | 4.2.2 | **AUTOMATIC PLACE-AND-ROUTE OF EMERGING LED-DRIVEN WIRES WITHIN A MONOLITHICALLY-INTEGRATED CMOS+III-V PROCESS**
**Speaker:** Tushar Krishna, Georgia Institute of Technology, US
**Authors:** Tushar Krishna1, Arya Balachandran2, Siau Ben Chiah3, Li Zhang1, Bing Wang4, Cong Wang5, Kenneth Lee Eng Kiong6, Jurgen Michel7 and Li-Shiuan Peh8
1Georgia Institute of Technology, US; 2NTU, SG; 3SMART, SG; 4MIT, US; 5Professor, National University of Singapore, SG
**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)

18:00 | 4.2.3 | **A TUNABLE MAGNETIC SKYRMION NEURON CLUSTER FOR ENERGY EFFICIENT ARTIFICIAL NEURAL NETWORK**
**Speaker:** Deliang Fan, University of Central Florida, US
**Authors:** Zhezhi He1 and Deliang Fan2
1Department of ECE, University of Central Florida, US; 2University 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 new neuron information carrier owing to ultra-small size (sub-10nm), high speed (>100m/s) with ultra-low depinning current density(MA/cm²-2) 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 3.11 J/bit, which is more than two orders lower than that of CMOS counterpart.

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

18:30 | IP2-1 | **COMPACT MODELING AND CIRCUIT-LEVEL SIMULATION OF SILICON NANOPHOTONIC INTERCONNECTS**
**Speaker:** Yuyang Wang, UC Santa Barbara, US
**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 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)
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. The last paper concludes with an adaptive page re-mapping architecture in flash-based storage with improved response time.
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 1.44x improvement in L2 cache energy for negligible (< 0.5%) loss in application-level quality across a suite of 8 benchmarks.

Abstract
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 1.44x 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)
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.

### Time, Label, Presentation Title, Authors

**18:30**  
**IP2-5, 328**  
**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 WN, 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)**

**18:31**  
**IP2-6, 880**  
**SPMS: STRAND BASED PERSISTENT MEMORY SYSTEM**  
**Speaker:** Shuo Li, National University of Defense Technology, CN  
**Authors:** Shuo Li\(^1\), Peng Wang\(^2\), Nong Xiao\(^1\), Guangyu Sun\(^2\) and Fang Liu\(^1\)  
\(^{1}\)National University of Defense Technology, CN; \(^{2}\)Peking University, CN  
**Abstract**  
Emerging non-volatile memories enable persistent memory, which offers the opportunity to directly access persistent data structures residing in main memory. In order to keep persistent data consistent in case of system failures, most prior work relies on persist ordering constraints which incur significant overheads. Strand persistency minimizes persist ordering constraints. However, there is still no proposed persistent memory design based on strand persistency due to its implementation complexity. In this work, we propose a novel persistent memory system based on strand persistency, called SPMS. SPMS consists of cacheline-based strand group tracking components, a volatile strand buffer and ultra-capacitors incorporated in persistent memory modules. SPMS can track each strand and guarantee its atomicity. In case of system failures, committed strands buffered in the strand buffer can be flushed back to persistent memory within the residual energy window provided by the ultra-capacitors. Our evaluations show that SPMS outperforms the state-of-the-art persistent memory system by 6.6% and has slightly better performance than the baseline without any consistency guarantee. What’s more, SPMS reduces the persistent memory write traffic by 30%, with the help of the strand buffer.

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

**18:32**  
**IP2-7, 72**  
**ARCHITECTING HIGH-SPEED COMMAND SCHEDULERS FOR OPEN-ROW REAL-TIME SDRAM CONTROLLERS**  
**Speaker:** Leonardo Ecco, TU Braunschweig, DE  
**Authors:** Leonardo Ecco\(^1\) and Rolf Ernst\(^2\)  
\(^{1}\)Institute of Computer and Network Engineering, TU Braunschweig, DE; \(^{2}\)TU 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 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)**

**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.

---

### Time, Label, Presentation Title, Authors

**17:00**  
**4.4.1**  
**DATA FLOW TESTING FOR VIRTUAL PROTOTYPES**  
**Speaker:** Muhammad Hassan, University of Bremen, DE  
**Authors:** Muhammad Hassan\(^1\), Vladimir Herdt\(^1\), Hoang M. Le\(^1\), Mingsong Chen\(^1\), Daniel Grosse\(^3\) and Rolf Drechsler\(^3\)  
\(^{1}\)University of Bremen, DE; \(^{2}\)East China Normal University, CN; \(^{3}\)University of Bremen/DFKI GmbH, DE  
**Abstract**  
Data flow testing (DFT) has been shown to be an effective testing strategy. DFT features a high fault detection rate while avoiding the intense scalability problems to achieve full path coverage. In this paper we propose to apply data flow testing for SystemC virtual prototypes (VPs). Our contribution is twofold: First, we develop a set of SystemC specific coverage criteria for data flow testing. This requires to consider the SystemC semantics of using non-preemptive thread scheduling with shared memory communication and event-based synchronization. Second, we explain how to automatically compute the data flow coverage result for a given VP using a combination of static and dynamic analysis techniques. The coverage result provides clear suggestions for the testing engineer to add new testcases in order to improve the coverage result. Our experimental results on real-world VPs demonstrate the applicability and efficacy of our analysis approach and the SystemC specific coverage criteria to improve the testsuite.

**Download Paper (PDF; Only available from the DATE venue WiFi)**
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; Lorenzo Di Tucci, Gianluca Durelli, and Michaela Blott</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abstract</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>
</tr>
<tr>
<td>17:18</td>
<td>4.5.2</td>
<td>FROM EXAFLOP TO EXAFLOW</td>
<td>Tobias Becker, Maxeler Technologies, GB; Pavel Burovsky, Anna Maria Nestorov, Kristina Paliakareva, Enrico Reggiani, Georgiy Gaydadjiev, Maxeler Technologies, GB; Maxeler Technologies Ltd, GB; Politecnico di Milano, IT; Maxeler / Imperial College, GB</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abstract</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>
</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; Marco Rabozzi, Giuseppe Natale, Emanuele Del Sozzo, Alberto Scolari, Marco D. Santambrogio and Luca Stornaiuolo, Politecnico di Milano, Italy</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abstract</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>
</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; Catalin Bogdan Ciobanu, Marco D. Santambrogio, Jose Gabriel Coutinho, Andreas Brokalski, Dionisos Pnevmatikatos, Michael Huebner, Tobias Becker, and Jakob E. T. Thom, Ghent University, BE; UVA, NL; Politecnico di Milano, IT; Imperial College London, GB; Synelixis, GR; ECE Department, Technical University of Crete &amp; FORTH-ICS, GR; Ruhr-University Bochum, DE; Maxeler Technologies, GB; University of Cambridge, GB</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abstract</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>
</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 Walterth, Intel, US; Jonathan Eastep, Intel, US</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abstract</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>
</tr>
<tr>
<td>18:30</td>
<td></td>
<td>End of session</td>
<td>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.</td>
</tr>
</tbody>
</table>
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 by identifying systematic manufacturing issues is presented. The last paper analyzes failure modes of Flash memories and proposes suitable fault models.

<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>4.6.1</td>
<td>FAST AND WAVEFORM-ACCURATE HAZARD-AWARE SAT-BASED TSOF ATPG</td>
<td>Jan Burchard, University of Freiburg, DE; Dominik Erb, Adit D. Singh, Sudhakar M. Reddy, Bernd Becker</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Authors:</strong></td>
<td>Jan Burchard 1, Dominik Erb 1, Adit D. Singh 2, Sudhakar M. Reddy 3 and Bernd Becker 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
<td>Open repairs 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 open causing 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-unknown 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)</td>
</tr>
<tr>
<td>17:30</td>
<td>4.6.2</td>
<td>FAULT DIAGNOSIS OF ARBITER PHYSICAL UNCLONABLE FUNCTION</td>
<td>Yu Hu, Institute of Computing Technology, Chinese Academy of Sciences, CN</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Authors:</strong></td>
<td>Jing Ye 1, Qingli Guo 1, Yu Hu 1 and Xiaowei Li 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
<td>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 novel diagnostic method toward classic arbiter PUF. The stuck-at faults and 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)</td>
</tr>
<tr>
<td>18:00</td>
<td>4.6.3</td>
<td>FPGA-BASED FAILURE MODE TESTING AND ANALYSIS FOR MLC NAND FLASH MEMORY</td>
<td>Fei Wu, Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology, CN</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Authors:</strong></td>
<td>Meng Zhang 1, Fei Wu 1, Qian Xia 1, He Huang 1, Jian Zhou 1 and Changsheng Xie 1</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
<td>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 (PASSPI) 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), read 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)</td>
</tr>
<tr>
<td>18:30</td>
<td>4.6.3</td>
<td>RETRODMR: TROUBLESHOOTING NON-DETERMINISTIC FAULTS WITH RETROSPECTIVE DMR</td>
<td>Ting Wang, The Chinese University of Hong Kong, HK</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Authors:</strong></td>
<td>Ting Wang 1, Yannan Liu 1, Qian Xu 1, Zhaobo Zhang 2, Zhiyuan Wang 2 and Xinli Gu 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
<td>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)</td>
</tr>
</tbody>
</table>
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 heterogeneous MPSoCs consisting of cores with different performance/power behaviors. 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. In this paper, an on-line heuristic is introduced to maximize soft-error reliability while satisfying a lifetime reliability constraint for soft real-time systems on heterogeneous MPSoCs. The approach works based on extracting the impact of process variations on the neurons characteristics during the test phase through a proposed algorithm. The approach then determines the features which are back-annotated, and utilizes 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 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)
**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)

---

**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.

---

**4.8 Openings 2017**

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

**Time** | **Label** | **Presentation Title** | **Authors**
--- | --- | --- | ---
17:00 | 4.8.1 | OPENING | Philippe Ory, Head of the EPFL Career Center, CH
17:30 | 4.8.2 | CADENCE PRESENTATION | Anton Klotz, Cadence Design Systems, DE
17:40 | 4.8.3 | HIPEAC PRESENTATION | Xavier Salazar, Hipec, ES
17:50 | 4.8.4 | SMARTCARDIA PRESENTATION | Srinivasan Murali, SmartCardia, CH
18:00 | 4.8.5 | NESPRESSO PRESENTATION | Martino Ruggiero, Nespresso, CH
18:10 | 4.8.6 | NESTLé PRESENTATION | Gian Paolo Perrucci, Mobility and Apps Solution Manager at Nestlé, CH
18:20 | 4.8.7 | GAIT UP PRESENTATION | 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.

---

**UB04 Session 4**

**Date:** Tuesday 28 March 2017  
**Time:** 17:30 - 19:30  
**Location / Room:** Booth 1, Exhibition Area

**Label** | **Presentation Title** | **Authors**
--- | --- | ---
A variety of routing algorithms and synthetic traffic patterns, as well as dependency-driven trace-based simulation (Netrace), are implemented and will be demonstrated in mind. Currently, GNoCS accurately models a VC-based microarchitecture. It supports 2D and 3D mesh topologies with full or partial vertical connections. 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.

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 in the SIs 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.

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 in the SIs 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.

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 MPSDbench. 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.

An open source collaborative MATLAB toolbox for the design and simulation of Continuous-Time Sigma Delta Modulators is proposed. 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 (Schreier 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, 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.

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 in the SCSs 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.

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 in the SCSs 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.

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 in the SCSs 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.

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 in the SCSs 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.

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 in the SCSs 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.
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.
More information ...

SELINK: SECURING HTTP AND HTTPS-BASED COMMUNICATION VIA SECUBE™
Presenter:
Airofarulla Giuseppe, CINI & Politecnico di Torino, IT
Authors:
Paolo Prinetto1 and Antonio Varriale2
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 ...

GREENOPENHEVC: LOW POWER HEVC DECODER
Presenter:
Menard Daniel, INSA Rennes, FR
Authors:
Julien Heulot1, Erwan Nogues1, Maxime Pelcat2 and Wassim Hamidouche1
1INSA 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 ...

19:30 End of session

Exhibition-Reception Exhibition Reception
Date: Tuesday 28 March 2017
Time: 18:30 - 19:30
Location / Room: 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.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>19:30</td>
<td>End of session</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5.1 IoT Day: IoT Perspectives
Date: Wednesday 29 March 2017
Time: 08:30 - 10:00
Location / Room: 5BC

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

Chair:
Marilyn Wolf, Georgia Tech, US

Co-Chair:
Andreas Herkersdorf, TU Muenchen, DE

The DATE 2017 Special Day on IoT will be kicked-off by perspective talks from academia and industry sharing their views and experience from backgrounds of large distributed sensor networks and cognitive computing. The entire spectrum of IoT devices and computing, storage and communication infrastructure, from smallest form factor sensors to Cloud backbone systems will be considered.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>19:30</td>
<td>End of session</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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.

MAKE IT REVERSIBLE: EFFICIENT EMBEDDING OF NON-REVERSIBLE FUNCTIONS

Author: Alwin Zulehner, Johannes 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)
Energy Efficient Stochastic Computing with Sobol Sequences

Abstract
Energy efficiency 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 similar or 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.

Design Automation and Design Space Exploration for Quantum Computers

Abstract
A major hurdle to the deployment of quantum linear systems algorithms and recent quantum simulation algorithms lies in the difficulty to find inexpensive reversible circuits for arithmetic using existing hand coded methods. Motivated by recent advances in reversible logic synthesis, we synthesize arithmetic circuits using classical design automation flows and tools. The combination of classical and reversible logic synthesis enables the automatic design of large components in reversible logic starting from well-known hardware description languages such as Verilog. As a prototype example for our approach we automatically generate high quality networks for the reciprocal $1/x$, which is necessary for quantum linear systems algorithms.

Quantum Computing is RAPIDLY evolving especially after the discovery of several efficient quantum algorithms solving intractable classical problems such as Shor's factoring algorithm. However the realization of a large-scale physical quantum computer is very challenging and the number of qubits that are currently under development is still very low, namely less than 15. In the absence of large size platforms, quantum computer simulation is critical for developing and testing quantum algorithms and investigating the different challenges facing the design of quantum computer hardware. What makes quantum computer simulation on classical computers particularly challenging are the memory and computational resource requirements. In this paper, we introduce a universal quantum computer simulator, called QX, that takes as input a specially designed quantum assembly language, called QASM, and provides, through aggressive optimisations, high simulation speeds and large number of qubits. QX allows the simulation of up to 34 fully entangled qubits on a single node using less than 270 GB of memory. Our experiments using different quantum algorithms show that QX achieves significant simulation speedup over similar state-of-the-art simulation environment.

Design Automation and Design Space Exploration for Quantum Computers

Abstract
Motivated by recent advances in reversible logic synthesis, we synthesize arithmetic circuits using classical design automation flows and tools. The combination of classical and reversible logic synthesis enables the automatic design of large components in reversible logic starting from well-known hardware description languages such as Verilog. As a prototype example for our approach we automatically generate high quality networks for the reciprocal $1/x$, which is necessary for quantum linear systems algorithms.

A major hurdle to the deployment of quantum linear systems algorithms and recent quantum simulation algorithms lies in the difficulty to find inexpensive reversible circuits for arithmetic using existing hand coded methods. Motivated by recent advances in reversible logic synthesis, we synthesize arithmetic circuits using classical design automation flows and tools. The combination of classical and reversible logic synthesis enables the automatic design of large components in reversible logic starting from well-known hardware description languages such as Verilog. As a prototype example for our approach we automatically generate high quality networks for the reciprocal $1/x$, which is necessary for quantum linear systems algorithms.

Quantum Computing is RAPIDLY evolving especially after the discovery of several efficient quantum algorithms solving intractable classical problems such as Shor's factoring algorithm. However the realization of a large-scale physical quantum computer is very challenging and the number of qubits that are currently under development is still very low, namely less than 15. In the absence of large size platforms, quantum computer simulation is critical for developing and testing quantum algorithms and investigating the different challenges facing the design of quantum computer hardware. What makes quantum computer simulation on classical computers particularly challenging are the memory and computational resource requirements. In this paper, we introduce a universal quantum computer simulator, called QX, that takes as input a specially designed quantum assembly language, called QASM, and provides, through aggressive optimisations, high simulation speeds and large number of qubits. QX allows the simulation of up to 34 fully entangled qubits on a single node using less than 270 GB of memory. Our experiments using different quantum algorithms show that QX achieves significant simulation speedup over similar state-of-the-art simulation environment.

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

Design Automation and Design Space Exploration for Quantum Computers

Abstract
A major hurdle to the deployment of quantum linear systems algorithms and recent quantum simulation algorithms lies in the difficulty to find inexpensive reversible circuits for arithmetic using existing hand coded methods. Motivated by recent advances in reversible logic synthesis, we synthesize arithmetic circuits using classical design automation flows and tools. The combination of classical and reversible logic synthesis enables the automatic design of large components in reversible logic starting from well-known hardware description languages such as Verilog. As a prototype example for our approach we automatically generate high quality networks for the reciprocal $1/x$, which is necessary for quantum linear systems algorithms.

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

Design Automation and Design Space Exploration for Quantum Computers

Abstract
A major hurdle to the deployment of quantum linear systems algorithms and recent quantum simulation algorithms lies in the difficulty to find inexpensive reversible circuits for arithmetic using existing hand coded methods. Motivated by recent advances in reversible logic synthesis, we synthesize arithmetic circuits using classical design automation flows and tools. The combination of classical and reversible logic synthesis enables the automatic design of large components in reversible logic starting from well-known hardware description languages such as Verilog. As a prototype example for our approach we automatically generate high quality networks for the reciprocal $1/x$, which is necessary for quantum linear systems algorithms.

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

Design Automation and Design Space Exploration for Quantum Computers

Abstract
A major hurdle to the deployment of quantum linear systems algorithms and recent quantum simulation algorithms lies in the difficulty to find inexpensive reversible circuits for arithmetic using existing hand coded methods. Motivated by recent advances in reversible logic synthesis, we synthesize arithmetic circuits using classical design automation flows and tools. The combination of classical and reversible logic synthesis enables the automatic design of large components in reversible logic starting from well-known hardware description languages such as Verilog. As a prototype example for our approach we automatically generate high quality networks for the reciprocal $1/x$, which is necessary for quantum linear systems algorithms.

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

Design Automation and Design Space Exploration for Quantum Computers

Abstract
A major hurdle to the deployment of quantum linear systems algorithms and recent quantum simulation algorithms lies in the difficulty to find inexpensive reversible circuits for arithmetic using existing hand coded methods. Motivated by recent advances in reversible logic synthesis, we synthesize arithmetic circuits using classical design automation flows and tools. The combination of classical and reversible logic synthesis enables the automatic design of large components in reversible logic starting from well-known hardware description languages such as Verilog. As a prototype example for our approach we automatically generate high quality networks for the reciprocal $1/x$, which is necessary for quantum linear systems algorithms.

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

Design Automation and Design Space Exploration for Quantum Computers

Abstract
A major hurdle to the deployment of quantum linear systems algorithms and recent quantum simulation algorithms lies in the difficulty to find inexpensive reversible circuits for arithmetic using existing hand coded methods. Motivated by recent advances in reversible logic synthesis, we synthesize arithmetic circuits using classical design automation flows and tools. The combination of classical and reversible logic synthesis enables the automatic design of large components in reversible logic starting from well-known hardware description languages such as Verilog. As a prototype example for our approach we automatically generate high quality networks for the reciprocal $1/x$, which is necessary for quantum linear systems algorithms.
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 resilience 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.

**Chair:**
Christian Enz, EPFL, CH

**Co-Chair:**
Anca Molinco, CEA Leti, FR
5.4 Solutions for efficient simulation and validation

Date: Wednesday 29 March 2017
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.

5.4.1 PERFORMANCE IMPACTS AND LIMITATIONS OF HARDWARE MEMORY-ACCESS TRACE COLLECTION

Speaker: Graham Holland, Simon Fraser University, CA

Authors: Nicholas C. Doyle¹, Eric Matthews¹, Graham Holland¹, Alexandra Fedorova² and Lesley Shannon¹

1Simon Fraser University, CA; ²University of British Columbia, CA

Abstract

In today’s multicore architectures, complex interactions between applications in the memory system can have a significant, and highly variable, impact on application execution time. System designers typically use hardware counters to profile execution behaviours and diagnose performance problems. However, hardware counters are not always sufficient and some problems are best identified with full memory access traces. Collecting these traces in software is very expensive. Our work explores using dedicated hardware for memory-access trace collection. We focus on analyzing the limitations of hardware data collection and its impacts on application performance. The key feature of our study is that it is performed on actual hardware using two very different CPU platforms: 1) the PolyBlaze multicore soft processor and 2) the ARM Cortex-A9. In both cases, the data collection is implemented on an FPGA. Using micro-benchmarks designed to test the bounds of memory access behaviour, we illustrate the operational regions of data collection and the impact on system performance. By examining the bandwidth bottlenecks that limit the rate of data collection, as well as hardware architecture choices that can aggravate the impact on application performance, we provide guidelines that can be used to extrapolate our analysis to other systems and processor architectures.

Download Paper (PDF; Only available from the DATE venue WiFi)
CONTEXT-SENSITIVE TIMING AUTOMATA FOR FAST SOURCE LEVEL SIMULATION

Speaker:
Sebastian Ottlik, FZI Research Center for Information Technology, DE

Authors:
Sebastian Ottlik1, Christoph Gerum2, Alexander Vieh1, Wolfgang Rosenstiel1 and Oliver Bringmann1

1FZI Research Center for Information Technology, DE; 2University of Tuebingen, DE; 3FZI Forschungszentrum Informatik, DE; 4University of Tuebingen / FZI, DE

Abstract
We present a novel technique for efficient source level timing simulation of embedded software execution on a target platform. In contrast to existing approaches, the proposed technique can accurately approximate time without requiring a dynamic cache model. Thereby the dramatic reduction in simulation performance inherent to dynamic cache modeling is avoided. Consequently, our approach enables an exploitation of the performance potential of source level simulation for complex microarchitectures that include caches. Our approach is based on recent advances in context-sensitive binary level timing simulation. However, a direct application of the binary level approach to source level simulation reduces simulation performance similarly to dynamic cache modeling. To overcome this performance limitation, we contribute a novel pushdown automaton based simulation technique. The proposed context-sensitive timing automata enable an efficient evaluation of complex simulation logic with little overhead. Experimental results show that the proposed technique provides a speedup of an order of magnitude compared to existing context selection techniques. Simulation performance is similar to a state of the art accelerated cache simulation. The accelerated simulation is only applicable in specific circumstances, whereas the proposed approach does not suffer this limitation.

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

MARS: A FLEXIBLE REAL-TIME STREAMING PLATFORM FOR TESTING AUTOMATION SYSTEMS

Speaker:
Alexandru Moga, ABB Research, CH

Authors:
Raphael Eidenbenz, Alexandru Moga, Thanikasav Senthanvi and Carsten Franke, ABB Corporate Research, CH

Abstract
The paper makes use of a complex case study. It deals with multiple-discipline descriptions, non-electrical conservative models, non-linear equation systems, common emerging problems and it proposes a set of solutions to them. It presents a set of techniques, unified in an automatic abstraction methodology, to accommodate active as well as quiescent reliability decay rates of underlying mechanical components. It uses path based reliability modeling to estimate the reliability degradation of component based systems that are controlled by software logic. Its efficacy is further demonstrated using a thorough case study.

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

SERD: A SIMULATION FRAMEWORK FOR ESTIMATION OF SYSTEM LEVEL RELIABILITY DEGRADATION

Speaker:
Saurav Ghosh, IIT Kharagpur, IN

Authors:
Saurav Kumar Ghosh1 and Dey Soumyajit2

1Dept. of CSE, IIT Kharagpur, IN; 2IIT Kharagpur, IN

Abstract
Development of highly reliable embedded controlsystems is typically performed following the model driven engineering paradigm. Such systems involve software control interaction of mechanical subsystems. The aging of the overall system depends on the physical aging or reliability decay of the underlying mechanical components. The reliability of such components degrade according to their rate of usage, which again is governed by the software control logic and input environment. Such dependencies of component reliabilities make the problem of deriving system level reliability degradation using exact methods combinatorially intractable. Given the fact that model driven system design advocates the usage of initial high level system models, methods for early stage lifetime reliability and reliability degradation estimation based on such initial models should definitely aid in robust high assurance engineering of such software controlled physical systems. The present work proposes SERD, a lightweight, scalable simulation framework for embedded control systems. It can accommodate active as well as quiescent reliability decay rates of underlying mechanical components. It uses path based reliability modeling to estimate the reliability degradation of component based systems that are controlled by software logic. Its efficacy is further demonstrated using a thorough case study.

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

MARS: A FLEXIBLE REAL-TIME STREAMING PLATFORM FOR TESTING AUTOMATION SYSTEMS

Speaker:
Alexandru Moga, ABB Research, CH

Authors:
Raphael Eidenbenz, Alexandru Moga, Thanikasav Senthanvi and Carsten Franke, ABB Corporate Research, CH

Abstract
The paper makes use of a complex case study. It deals with multiple-discipline descriptions, non-electrical conservative models, non-linear equation systems, common emerging problems and it proposes a set of solutions to them. It presents a set of techniques, unified in an automatic abstraction methodology, to accommodate active as well as quiescent reliability decay rates of underlying mechanical components. It uses path based reliability modeling to estimate the reliability degradation of component based systems that are controlled by software logic. Its efficacy is further demonstrated using a thorough case study.

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 Tampouratzis1, Konstantinos Georgopoulos2 and Ioannis Papaefstathiou3

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/Mems/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 computationally 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)

AUTOMATIC ABSTRACTION OF MULTI-DISCIPLINE ANALOG MODELS FOR EFFICIENT FUNCTIONAL SIMULATION

Speaker:
Franco Funmi, Università degli Studi di Verona, IT

Authors:
Enrico Fraccaroli1, Michele Lora1 and Franco Funmi2

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 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)
### 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

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.

### Automatic Construction of Models for Analytic System-level Design Space Exploration Problems

**Speaker:**  
Seyed-Hosein Attarzadeh-Niaki, Shahid Beheshti University (SBU), IR  
Shahid Beheshti University (SBU), IR; KTH Royal Institute of Technology, SE

**Authors:**  
Seyed-Hosein Attarzadeh-Niaki and Ingo Sander

**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 platform 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)
OPPORTUNISTIC WRITE FOR FAST AND RELIABLE STT-MRAM

Speaker:
Mehdi Tahoori, Karlsruhe Institute of Technology, DE
Authors:
Nour Sayed1, Mojtaba Ebrahimi2, Rajendra Bishnoi1 and Mehdi Tahoori1
1Karlsruhe Institute of Technology, DE; 2Karlsruhe Institute of Technology, DE

Abstract
Due to the stochastic switching behavior of the bitcell in Spin Transfer Torque Magnetic Random Access Memory (STT-MRAM), an excessive write margin is required to guarantee an acceptable level of reliability and yield. This prevents the usage of STT-MRAM in fast memories such as L1 or L2 caches. The excessive write margin of STT-MRAM can be reduced to a large extent by an opportunistic write (i.e., terminating the write process before all bit switchings are completed) and by reducing thermal stability factor. The bits with unfinished writes have to be processed by robust Error Correction Codes (ECCs). However, such coding schemes have relatively large decoding latencies, which increases the overall read latency significantly. Moreover, thermally induced retention failures can limit the applicability of such schemes. In this paper, we exploit the fact that error detection is much faster than correction. Therefore, the errors can be detected quickly and all erroneous data can be reverted before they arrive critical parts of the system (e.g., commit stage or memory ports). We also provide an adaptive approach to manage temperature-dependent retention failures at runtime. Hence, our proposed approach enables the use of STT-MRAM technology for fast cache applications.

Download Paper (PDF; Only available from the DATE venue WiFi)
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 08:30 - 10:00

**Location / Room:** SA

**Chair:** Paolo Bernardi, Politecnico di Torino, IT

**Co-Chair:** Alberto Bosio, LIRMM, FR

**Time** | **Label** | **Presentation Title**
--- | --- | ---
08:30 | 5.6.1 | **FAULT CLUSTERING TECHNIQUE FOR 3D MEMORY BISR**
**Speaker:** Tianjian Li, Shanghai Jiao Tong University, CN
**Authors:**
1. Tianjian Li, Yan Han 1, Xiaoyao Liang 1, Hsien-Hsin S. Lee 2 and Li Jiang 3
2. Shanghai Jiao Tong University, CN; 3. TSMC / Georgia Tech, TW; 2. Department 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)

---

**Time** | **Label** | **Presentation Title**
--- | --- | ---
09:00 | 5.6.2 | **ARCHITECTURAL EVALUATIONS ON TSV REDUNDANCY FOR RELIABILITY ENHANCEMENT**
**Speaker:** Yen-Hao Chen, National Tsing Hua University, Taiwan, TW
**Authors:**
1. Yen-Hao Chen 1, Chien-Pang Chiu 1, Russell Barnes 2 and TingTing Hwang 1
2. National Tsing Hua University, R.O.C, TW; 2. University 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 (EM) 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)

---

**Time** | **Label** | **Presentation Title**
--- | --- | ---
09:30 | 5.6.3 | **REUSING TRACE BUFFERS TO ENHANCE CACHE PERFORMANCE**
**Speaker:** Neetu Jindal, PhD, IN
**Authors:**
1. Neetu Jindal 1, Preeti Ranjan Panda and Smriti 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)

---

**Time** | **Label** | **Presentation Title**
--- | --- | ---
09:45 | 5.6.4 | **OPTIMIZATION OF RETARGETING FOR IEEE 1149.1 TAP CONTROLLERS WITH EMBEDDED COMPRESSION**
**Speaker:** Sebastian Huhn, University of Bremen, DE
**Authors:**
1. Sebastian Huhn 1, Stephan Eggersglüß 1, Krishnendu Chakrabarty 2 and Rolf Drechsler 3

**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 25°C) can improve the test time by 3.71x10^13 X while MBI+BI with 2200e at 125°C can improve the test time by 1.97x10^14X.

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

Coffee Break

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)
5.7.3 PROBABILISTIC SCHEDULABILITY ANALYSIS FOR FIXED PRIORITY MIXED CRITICALITY REAL-TIME SYSTEMS

Speaker:
Yasmina Abdeddaïm, Université Paris-Est, LIGM, ESIEE Paris, FR

Authors:
Yasmina Abdeddaim\textsuperscript{1} and Dorin Maxim\textsuperscript{2}
\textsuperscript{1}University Paris-Est, LIGM, ESIEE-Paris, FR; \textsuperscript{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)

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
\begin{itemize}
  \item Coffee Break 10:30 - 11:30
  \item Coffee Break 16:00 - 17:00
\end{itemize}

Wednesday, March 29, 2017
\begin{itemize}
  \item Coffee Break 10:00 - 11:00
  \item Coffee Break 16:00 - 17:00
\end{itemize}

Thursday, March 30, 2017
\begin{itemize}
  \item Coffee Break 10:00 - 11:00
  \item Coffee Break 15:30 - 16:00
\end{itemize}

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.

Label | Presentation Title | Authors
\hline
IP2-1 | COMPACT MODELING AND CIRCUIT-LEVEL SIMULATION OF SILICON NANOPHOTONIC INTERCONNECTS | Speaker: Yuyang Wang, UC Santa Barbara, US 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 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)

IP2-2 | A TRUE RANDOM NUMBER GENERATOR BASED ON PARALLEL STT-MTJS | Speaker: Yuanzhuo Qu, University of Alberta, CA Authors: Yuanzhuo Qu\textsuperscript{1}, Jie Han\textsuperscript{1}, Bruce Cockburn\textsuperscript{1}, Yue Zhang\textsuperscript{2}, Weisheng Zhao\textsuperscript{2} and Witold Pedrycz\textsuperscript{1}
\textsuperscript{1}University of Alberta, CA; \textsuperscript{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 stochasic 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 AUTOMATIC EQUIVALENCE CHECKING FOR SYSTEMC-TLM 2.0 MODELS AGAINST THEIR FORMAL SPECIFICATIONS

Authors:
Mehran Goli, Jannis 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. 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)
**IP2-9 (Best Paper Award Candidate)**

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

**Speaker:** Josue Ortiz, Complutense University of Madrid, ES

**Authors:** Yuchao Ma and Hassan Ghasezmzadeh, 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 visions with an accuracy of 93.3% using a leave-one-subject-out analysis.

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

**IP2-10 (Best Paper Award Candidate)**

**RETRORDMR: TROUBLESHOOTING NON-DETERMINISTIC FAULTS WITH RETROSPECTIVE DMR**

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

**Authors:** Ting Wang¹, Yannan Liu¹, Qiang Xu¹, Zhaobo Zhang², Zhiyuan Wang² and Xinli Gu²

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-modal 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)

**IP2-11**

**CRITICAL PATH - ORIENTED THERMAL AWARE X-FILLING FOR HIGH UN-MODELED DEFECT COVERAGE**

**Speaker:** Fotios Vartziotis, Computer Engineering, T.E.I. of Epirus, Greece, GR

**Authors:** FOTIOS VARTZIOTIS¹ and Chrysovalantis Kavoussianos²

¹T.E.I of Epirus, University of Ioannina, 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)

**IP2-12**

**A COMPREHENSIVE METHODOLOGY FOR STRESS PROCEDURES EVALUATION AND COMPARISON FOR BURN-IN OF AUTOMOTIVE SOC**

**Speaker:** Paolo Bernardi, Politecnico di Torino, IT

**Authors:** Paolo Bernardi¹, Davide Appello², Giampiero Giacopelli³, Alessandro Motta², Alberto Pagani², Giorgio Pollaccia³, Christian Rabbi², Marco Restivo¹, Prit Ruberg⁴, Ernesti Sanchez¹, Claudio Maria Villa² and Federico Venini³

¹Politecnico di Torino, IT; ²STMicroelectronics, IT; ³STMicroelectronics, IT; ⁴Tallinn 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 procedure 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)

**IP2-13**

**ENERGY EFFICIENT STOCHASTIC COMPUTING WITH SOBOL SEQUENCES**

**Speaker:** Siting Liu, University of Alberta, CA

**Authors:** Siting Liu and Jie Han, University of Alberta, CA

**Abstract**

Energy efficiency 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.

Download Paper (PDF; Only available from the DATE venue WiFi)
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 Tampouratzis1, Konstantinos Georgopoulos2 and Ioannis Papaefstathiou1
1Technical University of Crete, GR; 2Telecommunication Systems Institute, Technical University of Crete, GR; 2Technical 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. GEN5). 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; 2Universita` 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 approach to solve 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 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 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 3.71x10^13 X while MBI+BI with 220Oe at 125C can improve the test time by 1.97x10^14 X. While MBI+BI with 220Oe at 125C can improve the test time by 1.97x10^14 X while MBI+BI with 220Oe at 125C can improve the test time by 1.97x10^14 X while MBI+BI with 220Oe at 125C can improve the test time by 1.97x10^14 X while MBI+BI with 220Oe at 125C can improve the test time by 1.97x10^14 X. Moreover, it provides a better test capability of 220Oe at 25C compared to 220Oe at 125C.
Download Paper (PDF; Only available from the DATE venue WiFi)

IP2-18
AUTOMATIC CONSTRUCTION OF MODELS FOR ANALYTIC SYSTEM-LEVEL DESIGN SPACE EXPLORATION PROBLEMS
Speaker:
Seyed-Hosein Attarzadeh-Niaki, Shahid Beheshti University (SBU), IR
Authors:
Seyed-Hosein Attarzadeh-Niaki1 and Ingo Sander2
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 extendable 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)

UB05 Session 5
Date: Wednesday 29 March 2017
Time: 10:00 - 12:00
Location / Room: Booth 1, Exhibition Area
Label | Presentation Title | Authors
**NOXIM-XT: A BIT-ACCURATE POWER ESTIMATION SIMULATOR FOR NOCS**

**Presenter:** Pierre Bornel, Université de Bretagne Sud, FR

**Authors:** André Rossi1, Johann Laurent2 and Erwan Moresc2

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 ...**

**RIMEDIA: 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 ...**

**NNDN: NEURAL NETWORKS DESIGNING NEURAL NETWORKS**

**Presenter:** Brett Meyer, McGill University, US

**Authors:** Warren Gross, Sean Smithson, Oussama 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 characteristics 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**
Modern 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 GPUs 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 Mendler2, 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.rtsys.informatik.uni-kiel.de/en/research/kieler). The demonstration shows how a textual notation, from which a visual model is generated, can be used to illustrate 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 unceasing shrinking process of CMOS technology is a physical limit, 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 memristors-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 basis of workloaddependent and independent parameters, such as area, time and power consumption. The main advantage is that it does not require any simulation and thus it avoid any runtime overheads. The demo aims to show the tool prototype on a selected set of benchmarks which will be synthesized on a memristor-based crossbar circuit.

More information ...

MTA: MANCHESTER THERMAL ANALYZER
Presenter: Scott Laderenheim, University of Manchester, GB
Authors: Yi-Chung Chen, Vasillis Pavlidis and Milan Mihajlović, 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, heat sinks, and heat spreaders.

More information ...

SEFILE: A SECURE FILESYSTEM IN USERSPACE VIA SECUBE™
Presenter: Giuseppe Airofarulla, CINI, IT
Authors: Paolo Prinetto and Antonio Varriale
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.

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 ...

6.1 IoT Day Hot Topic Session: IoT Enabling Technologies
Date: Wednesday 29 March 2017
Time: 11:00 - 12:30
Location / Room: SBC

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

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

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.
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.

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.

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)

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

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 1, Pascal Raaij1, Linus Feiten 2, Bernd Becker 1, Ulrich Rührmair 2 and Ilia Polian 3
Authors: Matthias Sauer 1, Pascal Raaij 1, Linus Feiten 2, Bernd Becker 1, Ulrich Rührmair 2 and Ilia Polian 3

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 concept 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 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 implementing 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 PUF-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 an Authenticated Encryption with Associated Data (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. Guillen 1, Thomas Pöppelmann 2, Jose M. Bermudo Mera 3, Georg Sigl 1 and Johanna Sepulveda 1

Abstract

NTRU is a lattice-based public-key cryptosystem that is being considered for post-quantum security. This work presents a systematic design flow to turn an arbitrary (sufficiently complex) circuit into an NTRU encrypt, in resource constrained devices such as those used for Internet-of-Things endpoints. We present an analysis of NTRUcrypt’s advantages over other cryptosystems for use in such devices. We describe four different NTRUcrypt implementations on an ARM Cortex M0-based microcontroller, compare their results, and show that NTRUcrypt 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 NTRUcrypt.

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

12:30 IP3-1, 206 LEVERAGING AGING EFFECT TO IMPROVE SRAM-BASED TRUE RANDOM NUMBER GENERATORS

Speaker: Mohammad Saber Golanbar, Karlsruhe Institute of Technology (KIT), DE
Authors: Saman Kiamehr 1, Mohammad Saber Golanbar 2 and Mehdi Tahoori 2

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’ response skew 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)
### OPERAND SIZE RECONFIGURATION FOR BIG DATA PROCESSING IN MEMORY

**Speaker:**
Luigi Carro, UFRGS, BR

**Authors:**
Paulo Cesar Santos 1, Geraldo Francisco de Oliveira Junior 2, Diego Gomes Tomé 3, Marco Antonio Zanata Alves 3, Eduardo Cunha de Almeida 3 and Luigi Carro 4

1UFRGS - Universidade Federal do Rio Grande do Sul, BR; 2Universidade Federal do Rio Grande do Sul, BR; 3UFPR, BR; 4UFRGS, BR

**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 vectors 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)
to its unique properties like the tight integration with CMOS and the ability to learn and adapt 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 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.

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.
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.

12:00 6.5.3 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-entrance 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 Miorandi1, Steven M. Nowick1, Wayne Burleson2 and Greg Sadowski3
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 Sadr1, Martin Klein2, Michael Siegrist2, Christian Weis1 and Norbert Wehn1
1University of Kaiserslautern, DE; 2WiPotec 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 Zynq7000 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 Notin1, Vincent Huard2, Etienne Maurin2, Sidi Ahmed Benhassain3 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)
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 control systems in the context of event based real-time simulation. 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.

### 6.6.4 Probabilistic Timing Analysis on Time-Randomized Platforms for the Space Domain

**Speaker:** Francisco J. Cazorla, Barcelona Supercomputing Center and Spanish National Research Council (III-CSIC), ES

**Authors:**
- Mikel Fernandez 1, David Morales 2, Leonidas Kosmidis 3, Alen Bardizbanyan 4, Ian Broster 5, Carles Hernandez 6, Eduard Quinones 7, Jaume Abella 8

1Barcelona Supercomputing Center, ES; 2BSC, ES; 3Barcelona Supercomputing Center and Universität Politècnica de Catalunya, ES; 4Chalmers University of Technology, SE; 5Rapita Systems LTD, GB; 6Barcelona Supercomputing Center (BSC-CNS), ES; 7Barcelona Supercomputing Center and IIIA-CSIC, ES; 8RESA, IT

**Abstract**

Timing Verification is a fundamental step in real-time embedded systems, with measurement-based timing analysis (MBTA) being the most common approach used to that end. We present a Space case study on a real platform that has been modified to support a probabilistic variant of MBTA called MBPTA. Our platform provides the properties required by MBPTA with the predicted WCET estimates with MBPTA being competitive to those with current MBTA practice while providing more solid evidence on their correctness for certification.

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

### 6.6.5 Cross-Layer Design of Reconfigurable Cyber-Physical Systems

**Speaker and Author:** Michael Masin, IBM Research, IL

**Abstract**

In the last few years, besides the concepts of embedded and interconnected systems, also the notion of Cyber-Physical Systems (CPS) has emerged: embedded computational collaborating devices, capable of sensing and controlling physical elements and, often, responding to humans. The continuous interaction between the physical and the computing layers makes their design and maintenance extremely complex. Uncertainty management and runtime reconfigurability, to mention the most relevant ones, are rarely tackled by available commercial and academic toolchains. In this context, the Cross-layer modell-based framework for multi-objective design of Reconfigurable systems in uncerRainy hybrid envirOnments (CERBERO) EU project aims at developing a design environment for CPS based of two pillars: 1) a cross-layer model-based approach to describe, optimize, and analyze the system and all its different views concurrently and 2) an advanced adaptivity support based on a multi-layer autonomous engine. In this work, we describe the necessary components and the required developments for seamless design of reusable and reconfigurable CPS and System of Systems in uncertain hybrid environments.

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

### 6.6.6 InspeX: Design and Integration of a Portable/Wearable Smart Spatial Exploration System

**Speaker and Author:** Suzanne Lesecq, CEA, LETI, Minatc Campus, FR

**Abstract**

The INSPEX H2020 project main objective is to integrate automotive-equivalent spatial exploration and obstacle detection functionalities into a portable/wearable multi-sensor, miniaturised, low power device. The INSPEX system will detect and localise in real-time static and mobile obstacles under various environmental conditions in 3D. Potential applications range from safer human navigation in reduced visibility, small robot/drone obstacle avoidance systems to navigation for the visually/mobility impaired, latter being the primary use-case considered in the project.

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

### 7.0 Cross-Layer Design of Reconfigurable Cyber-Physical Systems

**Speaker:** Michael Masin, IBM Research, IL

**Abstract**

In the last few years, besides the concepts of embedded and interconnected systems, also the notion of Cyber-Physical Systems (CPS) has emerged: embedded computational collaborating devices, capable of sensing and controlling physical elements and, often, responding to humans. The continuous interaction between the physical and the computing layers makes their design and maintenance extremely complex. Uncertainty management and runtime reconfigurability, to mention the most relevant ones, are rarely tackled by available commercial and academic toolchains. In this context, the Cross-layer modell-based framework for multi-objective design of Reconfigurable systems in uncertain hybrid environments (CERBERO) EU project aims at developing a design environment for CPS based of two pillars: 1) a cross-layer model-based approach to describe, optimize, and analyze the system and all its different views concurrently and 2) an advanced adaptivity support based on a multi-layer autonomous engine. In this work, we describe the necessary components and the required developments for seamless design of reusable and reconfigurable CPS and System of Systems in uncertain hybrid environments.

Download Paper (PDF; Only available from the DATE venue WiFi)
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
researchers, summer school and two Computing Systems Weeks, which are networking and knowledge-exchange gatherings. We also produce a biennial technology roadmap, the 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 this session, after a brief introduction to HiPEAC, we highlight some of our members’ innovative and groundbreaking research and development activities.

**Time** | **Presentation Title** | **Authors**
---|---|---
11:00 | **ACCELERATED DATA CENTERS FOR CLOUD COMPUTING: THE VINEYARD PLATFORM** | Dimitrios Soudris, National Technical Univ. of Athens and ICCS, GR
**Abstract**
VINEYARD aims to develop the technology and the ecosystem that will enable the efficient integration of the hardware acceleration in the data centres, seamlessly. The deployment of energy-efficient hardware accelerators will be used to improve significantly the performance of cloud computing applications and reduce the energy consumption in data centres.

VINEYARD is developing an integrated framework for energy-efficient data centres based on programmable hardware accelerators. It is working towards a high-level programming framework that allows end-users to seamlessly utilize these accelerators in heterogeneous computing systems by using typical data-centre cluster frameworks (i.e. Spark). VINEYARD is also developing two types of novel energy-efficient servers integrating two kinds of hardware accelerators: programmable dataflow-based accelerators and FPGA-based accelerators. The servers coupled with dataflow-based accelerators are suitable for cloud computing applications that can be represented in dataflow graphs while the latter will be used for accelerating applications that need tight communication between the processor and the hardware accelerators.

VINEYARD also fosters the establishment of an ecosystem that will empower open innovation based on hardware accelerators as data-centre plugins, thereby facilitating innovative enterprises (large industries, SMEs, and creative start-ups) to develop novel solutions using VINEYARD’s leading edge developments.

11:15 | **HIGH-PERFORMANCE PARALLELISATION OF REAL-TIME APPLICATIONS WITH THE UPSCALE SDK** | Luis Miguel Pinho, Polytechnic of Porto, PT
**Abstract**
Nowadays, the prevalence of computing systems in our lives is so ubiquitous that it would not be far-fetched to state that we live in a cyber-physical world dominated by computer systems. These systems demand for more and more computational performance to process large amounts of data from multiple data sources, some of them with guaranteed processing response times. In other words, systems are required to deliver their results within pre-defined (and sometimes extremely short) time bounds. Examples can be found for instance in intelligent transportation systems for fuel consumption reduction in cities or railway, or autonomous driving of vehicles.

To cope with such performance requirements, chip designers produced chips with dozens or hundreds of cores, interconnected with complex networks on chip. Unfortunately, the parallelization of the computing activities brings many challenges, among which how to provide timing guarantees, as the timing behaviour of the system running within a many-core processor depends on interactions on shared resources that are most of the time not known by the system designer.

P-SOCRATES (Parallel Software Framework for Time-Critical Many-core Systems) is an FP7 European project, which developed a novel methodology to facilitate the deployment of standardized parallel architectures for real-time applications. This methodology was implemented (based on existing models and components) to provide an integrated software development kit, the UpScale SDK, to fully exploit the huge performance opportunities brought by the most advanced many-core processors, whilst ensuring a predictable performance and maintaining (or even reducing) development costs of applications. The presentation will provide an overview of the UpScale SDK, its underlying methodology, and the results of its application on relevant industrial use-cases.

11:30 | **POWER-AWARE SOFTWARE MAPPING OF PARALLEL APPLICATIONS ONTO HETEROGENEOUS MPSOCs** | Gereon Onnebrink, RWTH Aachen University, DE
**Abstract**
With the ever-increasing need of computational power, heterogeneous multi- and many-processor SoCs provide the best trade-off between performance, cost, and power. However, one of the biggest hurdles to exploit multicore architectures from the SW side is how to efficiently develop performance and power co-optimised parallel applications. Making the right decisions in the vast SW design space can hardly be done by the programmer in a reasonable time frame, especially, when performing a manual design process. Considering an application that has been properly partitioned into multiple concurrent tasks, and programmed in a parallel language, the process of mapping those tasks onto the processors with the optimal voltage and frequency setting is a huge challenge for a certain design goal. An automatic approach is needed that determines the optimal decision, given an optimisation constraint. A great amount of research has been conducted at ICE aiming to optimise the performance of a parallelised application. The Silexica GmbH, a VC-backed spin-off from ICE, continues on this track of producing novel compiler technology and tools for programming embedded multicore platforms, and offers the tools and knowledge to the industry.

In order to co-optimise for power, accurate power modelling has to be integrated into the existing performance driven framework. ICE’s electronic system-level power estimation methodology is a more than consequent starting point. The methodology takes the available power information from a reference power trace and back-annotates it to determine the coefficients of a linear power model. Several case studies have shown power estimation errors with less than 5%.

Based on the power modelling capability, a novel power-aware SW mapping heuristic has been implemented. This algorithm is verified in several case studies and used to identify the gain of sophisticated power management techniques by providing the power-performance trade-off.

11:45 | **OVERVIEW OF MANGO: EXPLORING MANYCORE ARCHITECTURES FOR NEXT-GENERATION HPC SYSTEMS** | José Flich, Technical University of Valencia, ES
**Abstract**
The performance/power efficiency wall poses the major challenge faced nowadays by HPC. Looking straight at the heart of the problem, the hurdle to the full exploitation of today computing technologies ultimately lies in the gap between the applications’ demand and the underlying computing architecture: the closer the computing system matches the structure of the application, the most efficiently the available computing power is exploited. Consequently, enabling a deeper customization of architectures to applications is the main pathway towards computation power efficiency. In addition to mere performance and power-efficiency, it is of paramount importance to meet new nonfunctional requirements posed by emerging classes of applications. In particular, a growing number of HPC applications demand some form of time-predictability, or more generally Quality-of-Service (QoS), particularly in those scenarios where correctness depends on both performance and timing requirements and the failure to meet either of them is critical.

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

Presentation Title: GREENOPENHEVC: LOW POWER HEVC DECODER

Authors: Menard Daniel, INSA Rennes, FR

Abstract

We have developed a low-power HEVC decoder that is designed for embedded systems. The decoder is implemented on a multi-core platform and uses an efficient memory management technique. The decoder is optimized for low power consumption and high performance. The decoder has been tested on a set of video sequences and has achieved a significant reduction in power consumption compared to existing decoders.

More information ...

NOXIM-XT: A BIT-ACCURATE POWER ESTIMATION SIMULATOR FOR NOCS

Presentation Title: NOXIM-XT: A BIT-ACCURATE POWER ESTIMATION SIMULATOR FOR NOCS

Authors: André Rossi1, Johann Laurent2 and Erwan Moreac2

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 traffic is 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 ...

END OF SESSION

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

Presenter:
Andrea Emrici, Nokia Bell Labs France, FR

Authors:
Julien Laliet1, Imran Latif1, Ludovic Aprville2, Renaud Pacatel2 and Adrien Canuel2
1Nokia Bell Labs France, FR; 2 Télécom 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.

More information ...

WE Dare: Wearable Electronics Directional Augmented Reality

Presenter:
Davide Quaglia, University of Verona, IT

Authors:
Gianluca Benedetti1 and Walter Vendraminetto2
1Wagoo LLC, IT; 2 EDA Lab srl

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 ...

ITMD: Run-Time Management of Concurrent Multi-threaded Applications on Heterogeneous Multi-cores

Presenter:
Karunakar Reddy Basireddy, University of Southampton, GB

Authors:
Amit Singh, Bashar 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 ...

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 ...

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 Rafiee, 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 ...

TIDES: Non-linear Waveforms for Quick Trace Navigation

Presenter:
Jannis Stopp, 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 Prinetto1 and Antonio Varriale2
1CINI & Politecnico di Torino, IT; 2BiUS 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™.

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.

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 |
--- | --- | --- |
13:50 | 7.0.1 | INTERNET OF EVERYTHING IS OUR OPPORTUNITY |

Author:
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 coopertition will become more common in IoE.
Networking functionalities enable the designer to dynamically select the operations to be conducted. These networking functionalities exploit hydrodynamic forces 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 coatings, and the inferior biocompatibility, while flow-based LoCs require a complex and costly multilayer fabrication process. Hence, an alternative has recently 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 and the sensors measurements for autonomous navigation. The demonstrator consists in the robotic arm prototype, the Android tablet and a personal computer for robotic arm setup and configuration.

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.

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.

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

Presenters:
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 ...

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

Presenters:
Andreas Grimmer, Johannes Kepler University Linz, AT
Authors:
Andreas Grimmer, Johannes Kepler University Linz, AT
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.

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

Presenters:
Antonios Nikitakis and Andreas Brokalakis, Synelixis Solutions Ltd, GR
Authors:
Nikolaos Tampouratzis, Technical University of Crete, GR
Presenter:
Anastasios Nikolakis and Andreas Brokalakis, Synelixis Solutions Ltd, GR
Abstract
The robotic arm testing of the RIMEDIO project was presented. It is a 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.
SCCHARTS: SYNCHRONOUS STATECHARTS FOR SAFETY-CRITICAL APPLICATIONS

Presenter: Reinhard von Hanxleden, Kiel University, DE
Authors: Michael Mendler 1, Christian Motika 2, Christoph Daniel Schulze 2 and Steven Smyth 2
1 Bamberg University, DE; 2 Kiel 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.scccharts.com. The SCCharts demonstrator is an Eclipse Rich Client and part of KIELER (http://www.rtsys.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 ...

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 ...

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 Raliev, 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 ...

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

Presenter: Ainofarulla Giuseppe, CINI & Politecnico di Torino, IT
Authors: Paolo Prinetto 1 and Antonio Varrile 2
1 Politecnico di Torino, IT; 2 Blu5 Labs Ltd., IT

Abstract
The SÉcube™ 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 SÉcube™. By doing so, we show how communication can be secured from an attacker capable of inspecting, and tampering, the regular communication.

More information ...

STACKADROP: A MODULAR DIGITAL MICROFLUIDIC BIOCHIP RESEARCH PLATFORM

Presenter: Oliver Keszöce, University of Bremen, DE
Authors: Maximilian Lunert 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 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 Benni 2
1 ETH Zurich, CH; 2 ETH 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 ...
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

Co-Chair:

Chair:

Andreas Herkersdorf, TU Muenchen, DE

Marilyn Wolf, Georgia Tech, US

Organisers:

Location / Room: SBC

Time: 14:30 - 16:00

7.1 IoT Day Hot Topic Session: IoT Deployment

Date: Wednesday 29 March 2017

Time: 14:30 - 16:00

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

Chair:
Marilyn Wolf, Georgia Tech, US

Co-Chair:
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.

<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>IP7.1.1,</td>
<td>A LOW-POWER IOT PROCESSOR INTEGRATING VOLTAGE-SCALABLE FULLY DIGITAL MEMORIES</td>
<td>Hidetoshi Ondotera, Kyoto University, JP</td>
</tr>
<tr>
<td>14:37</td>
<td>IP7.1.2,</td>
<td>A SIMPLE, STATELESS, COST EFFECTIVE SYMMETRIC ENCRYPTION STRATEGY FOR ENERGY-HARVESTING IOT DEVICES</td>
<td>Jan Madsen, Technical University of Denmark, DK</td>
</tr>
<tr>
<td>14:44</td>
<td>IP7.1.3,</td>
<td>RECONFIGURABLE MICROCONTROLLER FOR END NODES IN INTERNET OF THINGS</td>
<td>Wai-Chung Matthew Tang, Queen Mary University of London, GB</td>
</tr>
<tr>
<td>14:51</td>
<td>IP7.1.4,</td>
<td>FURTHER SIMPLIFICATION OF APPROXIMATE ADDERS USING INPUT DATA RANGES IN IOT</td>
<td>Jeong-A Lee, Chosun University, KR</td>
</tr>
<tr>
<td>15:00</td>
<td></td>
<td>HOW ASIC DEVELOPMENT WILL CHANGE FOR FUTURE IOT MEMS SENSORS</td>
<td>Dirk Droste, Robert Bosch GmbH, DE</td>
</tr>
<tr>
<td>15:30</td>
<td>7.1.3</td>
<td>DISTRIBUTED WAYSIDE ARCHITECTURE - IOT FOR RAILWAY INFRASTRUCTURE</td>
<td>Peter Hefti, Siemens, CH</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Olivier Kaiser, Siemens, CH</td>
</tr>
</tbody>
</table>

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.
- The availability and safety are nonetheless very demanding to assure an efficient and safe operation.
- 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.
- 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.
emerging NVM technologies and goes on to investigate security and encryption strategies. New opportunities and challenges in terms of enhancing computational efficiency and ensuring security, respectively. This session explores in-memory computing applied to 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.

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.
Co-Chair: 
Sasan Avesta, George Mason University, US

Chair: 
Stefano Di Carlo, Politecnico di Torino, IT

Location / Room: 
2BC

Time: 
16:00

Label: 
IP3-7

Presentation Title: 
COVERT: COUNTER OVERFLOW REDUCTION FOR EFFICIENT ENCRYPTION OF NON-VOLATILE MEMORIES

Authors: 
Navneet Gupta, ISEP-Paris, FR

Speaker: 
Navneet Gupta, ISEP-Paris, FR

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)

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>
<th>Abstract</th>
</tr>
</thead>
<tbody>
<tr>
<td>14:30</td>
<td>7.3.1</td>
<td>ACCURATE PRIVATE/SHARED CLASSIFICATION OF MEMORY ACCESSES: A RUN-TIME ANALYSIS SYSTEM FOR THE LEON3 MULTI-CORE PROCESSOR</td>
<td>Nam Ho, Department of Computer Science, University of Paderborn, DE</td>
<td>Related work has presented simulation-based experiments to classify data accesses in a shared memory multi-core into private and shared. This information can be used to selectively turn on/off cache coherency mechanisms for data blocks, which can save memory bus bandwidth, minimize energy consumption, and reduce application runtimes. In this paper, we present an implementation of a private/shared classification mechanism on a LEON3 SPARC multi-core processor running the Linux 2.6 kernel. Our mechanism is paged-based and allows for classifying and counting data accesses at run-time. Compared to previous work, our system provides more accurate, i.e., realistic, data as it includes a real multi-core architecture and an OS. Additionally, our prototype allows us to quantitatively evaluate the overhead for the classification mechanism. We test our system with sequential and parallel benchmarks from the MiBench, ParaMiBench, PARSEC, and SPLASH2 application suites. The results show that parallel benchmarks are promising targets for selectively controlling coherency mechanisms and that the run-time overheads induced by our mechanism are rather small. Download Paper (PDF; Only available from the DATE venue WiFi)</td>
</tr>
<tr>
<td>15:00</td>
<td>7.3.2</td>
<td>DESIGN OF A LOW POWER, RELATIVE TIMING BASED ASYNCHRONOUS MSP430 MICROPROCESSOR</td>
<td>Dipanjan Bhadra, University of Utah, US</td>
<td>Power dissipation is one of the primary design constraints in modern digital circuits. From a magnitude of hand-held portable devices to big data analytics using high-performance computing, low energy dissipation is a key requirement for most modern devices. This paper showcases an elegant low power circuit design methodology based on Relative Timing driven asynchronous techniques. A low power MSP430 microprocessor design based on a novel asynchronous finite state machine implementation is presented. The design showcases the power benefits of the proposed asynchronous implementation over the synchronous counterpart and avoids major architectural modification which would directly influence the performance or power consumption. The implemented asynchronous MSP430 exhibits a minimum of 8X power benefit over the synchronous design for an almost identical pipeline structure and comparable throughput. The paper further elaborates on the novel asynchronous state machine implementation used for the design and presents an efficient method to design communicating asynchronous finite state machines in clock-less systems. Download Paper (PDF; Only available from the DATE venue WiFi)</td>
</tr>
<tr>
<td>15:30</td>
<td>7.3.3</td>
<td>A COORDINATED MULTI-AGENT REINFORCEMENT LEARNING APPROACH TO MULTI-LEVEL CACHE CO-PARTITIONING</td>
<td>Dipanjan Bhadra and Kenneth Stevens, University of Utah, US</td>
<td>Power dissipation is one of the primary design constraints in modern digital circuits. From a magnitude of hand-held portable devices to big data analytics using high-performance computing, low energy dissipation is a key requirement for most modern devices. This paper showcases an elegant low power circuit design methodology based on Relative Timing driven asynchronous techniques. A low power MSP430 microprocessor design based on a novel asynchronous finite state machine implementation is presented. The design showcases the power benefits of the proposed asynchronous implementation over the synchronous counterpart and avoids major architectural modification which would directly influence the performance or power consumption. The implemented asynchronous MSP430 exhibits a minimum of 8X power benefit over the synchronous design for an almost identical pipeline structure and comparable throughput. The paper further elaborates on the novel asynchronous state machine implementation used for the design and presents an efficient method to design communicating asynchronous finite state machines in clock-less systems. Download Paper (PDF; Only available from the DATE venue WiFi)</td>
</tr>
<tr>
<td>15:45</td>
<td>7.3.4</td>
<td>GPIOCP: TIMING-ACCURATE GENERAL PURPOSE I/O CONTROLLER FOR MANY-CORE REAL-TIME SYSTEMS</td>
<td>Preeti Ranjan Panda, Indian Institute of Technology Delhi, IN</td>
<td>Modern SoC / NOC chips often provide General-Purpose I/O (GPIO) pins for connecting devices that are not directly integrated within the chip. Timing accurate control of devices connected to GPIO is often required within embedded real-time systems -- i.e. I/O operations should occur at exact times, with minimal error, neither being significantly early or late. This is difficult to achieve due to the latencies and contentions present in architecture, between CPU instigating the I/O operation, and the device connected to the GPIO -- software drivers, RTOSs, buses and bus contentions all introduce significant variable latencies before the command reaches the device. This is compounded in NOC devices utilising a mesh interconnect between CPUs and I/O devices. The contribution of this paper is a resource efficient programmable I/O controller, termed the GPIO Command Processor (GPIOCP), that permits applications to instigate complex sequences of I/O operations at an exact time, so achieving timing-accuracy at a single clock cycle level. Also, I/O operations can be programmed to occur at some point in the future, periodically, or reactively. The GPIOCP is a parallel I/O controller, supporting cycle level timing accuracy across several devices connected to GPIO simultaneously. The GPIOCP exploits the tradeoff between placing using a full sequential CPU to control each GPIO connected device, which achieves some timing accuracy at high resource cost; and poor timing-accuracy achieved where the application CPU controls the device remotely. The GPIOCP has efficient hardware cost compared to CPU approaches, with the additional benefits of total timing accuracy (CPU solutions do not provide this in general) and parallel control of many I/O devices. Download Paper (PDF; Only available from the DATE venue WiFi)</td>
</tr>
<tr>
<td>16:00</td>
<td>IP3-10</td>
<td>A HARDWARE IMPLEMENTATION OF THE MCAS SYNCHRONIZATION PRIMITIVE</td>
<td>Smruti Sarangi, IIT Delhi, IN</td>
<td>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 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)</td>
</tr>
</tbody>
</table>
**7.4 Advances in Logic Synthesis**

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

**Chair:** Paolo Ienne, EPFL, CH  
**Co-Chair:** Tsutomu Sasao, Meiji University, JP

This session focuses on new results in logic synthesis. The first two papers present specialized synthesis algorithms for index generating functions and encoder circuits. The last two papers discuss efficient encoding with SAT of short-circuit detection and combinational delay optimization.

### Time | Label | Presentation Title | Authors
--- | --- | --- | ---
14:30 | 7.4.1 | AN ALGORITHM TO FIND OPTIMUM SUPPORT-REDUCING DECOMPOSITIONS FOR INDEX GENERATION FUNCTIONS. | Tsutomu Sasao, Meiji University, JP  
**Authors:** Tsutomu Sasao, Kyu Matsuura and Yukihito Iguchi, 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 \( 2^n \) 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)
15:00  7.4.2 TAKING ONE-TO-ONE MAPPINGS FOR GRANTED: ADVANCED LOGIC DESIGN OF ENCODER CIRCUITS
Speaker:   
Robert Wille, Johannes Kepler University, Linz, AT
Authors:   
Alwin Zulehner 1 and Robert Wille 2
1Johannes Kepler University, AT; 2Johannes Kepler University Linz, AT
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)

15:30  7.4.3 ANALYSIS OF SHORT-CIRCUIT CONDITIONS IN LOGIC CIRCUITS
Speaker:   
João Afonso, INESC-ID, PT
Authors:   
João Pedro 1 and Jose Monteiro 2
1INESC-ID, PT; 2INESC-ID, IST, U Lisboa, PT
Abstract
This paper offers a novel approach for the analysis of input conditions that cause a short-circuit in a logic circuit, that is, that create a direct path from the power supply to ground. We model the logic circuit as a graph where edges represent transistors which are either open or closed, function of the input conditions. From this graph we derive a Quantified Boolean Formula (QBF) problem whose solution identifies the existence of a valid input combination that creates a path in the graph between the pair of nodes that represent the power source and ground, without ever enumerating all input combinations. We build the QBF problem incrementally, minimising the number of active nodes and hence of possible states. In the end, we obtain a relatively simple CNF expression, function only of the circuit inputs, that is handled by a generic SAT solver. We present results that demonstrate the practical applicability of our method on circuit instances that are intractable by alternative methods.
Download Paper (PDF; Only available from the DATE venue WiFi)

15:45  7.4.4 BUSY MAN’S SYNTHESIS: COMBINATIONAL DELAY OPTIMIZATION WITH SAT
Speaker:   
Mathias Soeken, EPFL, CH
Authors:   
Mathias Soeken 1, Giovanni De Micheli 1 and Alan Mishchenko 2
1EPFL, CH; 2UC Berkeley, US
Abstract
Boolean SAT solving can be used to find a minimum-size logic network for a given small Boolean function. This paper extends the SAT formulation to find a minimum-size network under delay constraints. Delay constraints are given in terms of input arrival times and the maximum depth. After integration into a depth-optimizing mapping algorithm, the proposed SAT formulation can be used to perform logic rewriting to reduce the logic depth of a network. It is shown that to be effective the logic rewriting algorithm requires (i) a fast SAT formulation and (ii) heuristics to quickly determine whether the given delay constraints are feasible for a given function. The proposed algorithm is more versatile than previous algorithms, which is confirmed by the experimental results.
Download Paper (PDF; Only available from the DATE venue WiFi)

16:00  IP3-13, 799 TECHNOLOGY MAPPING WITH ALL SPIN LOGIC
Speaker:   
Azadeh Davoodi, University of Wisconsin - Madison, US
Authors:   
Boyu Zhang 1 and Azadeh Davoodi 2
1University of Wisconsin-Madison, US; 2University of Wisconsin - Madison, US
Abstract
This work is the first to propose a technology mapping algorithm for All Spin Logic (ASL) device. The ASL device is the most actively-pursued one among spintronics devices which themselves fall under emerging post-CMOS nano-technologies. We identify the shortcomings of directly applying the classical technology mapping with ASL devices, and propose techniques to extend the classical procedure to handle these shortcomings. Our results show that our ASL-aware technology mapping algorithm can achieve on-average 9.15% and up to 27.27% improvement in delay (when optimizing delay) with slight improvement in area, compared to the solution generated by classical technology mapping. In a broader sense, our results show that we need for developing circuit-level CAD tools that are aware of and optimized for emerging technologies in order to better assess their promise as we move to the post-CMOS era.
Download Paper (PDF; Only available from the DATE venue WiFi)

16:01  IP3-14, 734 A NEW METHOD TO IDENTIFY THRESHOLD LOGIC FUNCTIONS
Speaker:   
Spyros Tragoudas, Southern Illinois University Carbondale, US
Authors:   
Seyed Nima Mozaffari, Spyros Tragoudas and Themistoklis Haniotakis, Southern Illinois University, US
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)
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

### 7.5 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.

### 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.

### Time  |  Label  | Presentation Title
--- | --- | ---
14:30  | 7.6.1  | (Best Paper Award Candidate) MVP ECC : MANUFACTURING PROCESS VARIATION AWARE UNEQUAL PROTECTION ECC FOR MEMORY RELIABILITY  
Speaker: Joon-Sung Yang, Sungkyunkwan University, KR  
Authors: Seungyoo Lee and Joon-Sung Yang, Sungkyunkwan University, KR  
**Abstract** With a development of process technology, a memory density has been increased. However, a smaller feature size makes the memory susceptible to soft errors. For reliability enhancement, ECC with single bit error correction and double bit error detection is widely used. As multiple bit cell upset become dominant, there is a need for stronger ECC. ECC such as RS or BCH code requires significantly large overhead and longer latency. To overcome the problem, this paper introduces an unequal protection ECC assigning stronger level of protection to weak memory cells and normal level to normal cells. Information from manufacturing characterization test is utilized to identify weak memory cells with low design margins. Instead of equally treating all memory cells, the proposed ECC focuses more on the weak cells since they are more susceptible to soft errors. Compared to conventional ECCs, experimental results show that the proposed ECC considerably enhances memory reliability with the same code length.  
Download Paper (PDF; Only available from the DATE venue WiFi)

15:00  | 7.6.2  | ANALYZING THE EFFECTS OF PERIPHERAL CIRCUIT AGING OF EMBEDDED SRAM ARCHITECTURES  
Speaker: Josef Kinseher, Intel Deutschland, DE  
Authors: Josef Kinseher 1, Leonhard Heiß 1 and Ilia Polian 2  
1Intel Deutschland, DE; 2University of Passau, DE  
**Abstract** 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 core-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 PBTI 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)

15:30  | 7.6.3  | MITIGATION OF SENSE AMPLIFIER DEGRADATION USING INPUT SWITCHING  
Speaker: Daniel Kraak, Delft University of Technology, NL  
Authors: Daniel Kraak 1, Innocent Agbo 2, Mottaqiallah Taouil 1, Said Hamdioui 3, Stefan Cosemans 1, Francky Catthoor 1 and Wim Dehaene 2  
1Delft University of Technology, NL; 2imec, BE; 3KU Leuven, ESAT, BE  
**Abstract** 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 variability becomes more severe, 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)

16:00  | IP3-15, 16  | A BRIDGING FAULT MODEL FOR LINE COVERAGE IN THE PRESENCE OF UNDETECTATED 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)

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

---

7.7 Resource management and analysis for embedded architectures  
**Date:** Wednesday 29 March 2017
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>SCALABLE PROBABILISTIC POWER BUDGETING FOR MANY-CORES</td>
<td>Anuj Pathania, Karlsruhe 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 Henkel¹</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Authors:</td>
<td>¹Karlsruhe Institute of Technology, DE; ²Karlsruhe Institute of Technology (KIT), DE; ³Vienna University of Technology (TU Wien), AT; ⁴National University of Singapore, SG</td>
</tr>
<tr>
<td>14:50</td>
<td>7.7.2</td>
<td>EXPLOITING SPORADIC SERVERS TO PROVIDE BUDGET SCHEDULING</td>
<td>Matthias Beckert, Institut für Computer und Network Engineering, TU Braunschweig, DE</td>
</tr>
<tr>
<td></td>
<td></td>
<td>FOR ARINC653 BASED REAL-TIME VIRTUALIZATION ENVIRONMENTS Speaker:</td>
<td>Matthias Beckert¹, Kai Björn Gemlau² and Rolf Ernst²</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>15:00</td>
<td>7.7.3</td>
<td>PROGRAMMING AND ANALYSING SCENARIO-AWARE DATAFLOW ON A</td>
<td>Reinier van Kampenhout, Sander Stuijk and Kees Gossens, Eindhoven University of Technology, NL</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MULTI-PROCESSOR PLATFORM</td>
<td>Authors:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speaker:</td>
<td>Reinier van Kampenhout¹, Kai Björn Gemlau² and Rolf Ernst²</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Authors:</td>
<td>¹Institut für Datentechnik und Kommunikationsnetze - TU Braunschweig, DE; ²Eindhoven University of Technology, NL</td>
</tr>
<tr>
<td>15:30</td>
<td>7.7.4</td>
<td>CHRT: A CRITICALITY- AND HETEROGENEITY-AWARE RUNTIME SYSTEM FOR TASK-PARALLEL APPLICATIONS Speaker:</td>
<td>Myeonggyun Han, UNIST, KR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Authors:</td>
<td>Myeonggyun Han, Jinsu Park and Woongki Baek, UNIST, KR</td>
</tr>
<tr>
<td>16:00</td>
<td>IP3-16</td>
<td>CREATING A CRITICALITY- AND HETEROGENEITY-AWARE RUNTIME SYSTEM FOR TASK-PARALLEL APPLICATIONS Speaker:</td>
<td>Myeonggyun Han, Jinsu Park and Woongki Baek, UNIST, KR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Authors:</td>
<td>Myeonggyun Han, Jinsu Park and Woongki Baek, UNIST, KR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Download Paper (PDF; Only available from the DATE venue WiFi)</td>
<td>Download Paper (PDF; Only available from the DATE venue WiFi)</td>
</tr>
</tbody>
</table>

Abstract:

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 the critical 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.

Download Paper (PDF; Only available from the DATE venue WiFi)
**MOBIKEN: PORTING XEN ON ANDROID DEVICES FOR MOBILE VIRTUALIZATION**

**Speaker:** Jianguo Yao, Shanghai Jiao Tong University, CN  
**Authors:**  
Yaozu Dong¹, Jianguo Yao², 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.

**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.

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.

**7.8 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  

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**  
  - 10:30 - 11:30  
  - 16:00 - 17:00  

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

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

---

**7.8.1 YINS**  
**Speaker:** Eugene Van Roojen, Eaton, CH  

**7.8.2 SMARTGRID**  
**Speakers:** Marco Pignati and Sergio Barreto, EPFL, CH  

**7.8.3 X-SENSE II**  
**Speaker:** Jan Boutel, ETHZ, CH  

End of session
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.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>16:00</td>
<td>Coffee Break</td>
<td>in Exhibition Area</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tuesday, March 28, 2017</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Coffee Break 10:30 - 11:30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Coffee Break 16:00 - 17:00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Wednesday, March 29, 2017</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Coffee Break 10:00 - 11:00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Coffee Break 16:00 - 17:00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Thursday, March 30, 2017</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Coffee Break 10:00 - 11:00</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Coffee Break 15:30 - 16:00</td>
</tr>
</tbody>
</table>

### 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 Kiamehr 1, Mohammad Saber Golanbari 2 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)

| IP3-2 | DESIGN AUTOMATION FOR OBFUSCATED CIRCUITS WITH MULTIPLE VIABLE FUNCTIONS           | Shahrzad Keshavarz, University of Massachusetts Amherst, US  
Authors:  
Shahrzad Keshavarz 1, Christof Paar 2 and Daniel Holcomb 3  
1University of Massachusetts Amherst, US;  
2Horst Gortz Institut for IT-Security, Ruhr-Universitat Bochum, DE  
Abstract  
Gate camouflaging is a technique for defuscating 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)

| IP3-3 | DOUBLE MAC: DOUBLING THE PERFORMANCE OF CONVOLUTIONAL NEURAL NETWORKS ON MODERN FPGAS | Jongeun Lee, UNIST, KR  
Authors:  
Dong Nguyen 1, Daewoo Kim 1 and Jongeun Lee 2  
1UNIST, KR;  
2Ulsan 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: Khoo 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 Gerlach1, Thoralf Rosahl2, Frank-Thomas Eitrich2 and Jürgen Scheible1
1Robert Bosch Centre for Power Electronics, DE; 2Robert 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 Kuiper1 and Marco Bekooij2
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)

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 Huang1, Dan Feng2, Yu Hua3 and Wen Zhou 2
1Huazhong University of Science and Technology, CN; 2Wuhan 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)
**IP3-9**

**TUNNEL FET BASED REFRESH-FREE-DRAM**

**Speaker:**
Navneet Gupta, ISEP-Paris, FR

**Authors:**
Navneet Gupta1, Adam Makosiej1, Andrei Vladimirescu1, Amara Amara1 and Costin Anghel2
1: Institut supérieur d’électronique de Paris, France; LETI, Commissariat à l’Energie Atomique et aux Energies Alternatives (CEA-Leti) France; 2: LETI, Commissariat à l’Energie Atomique et aux Energies Alternatives (CEA-Leti), FR; 3: Institut Supérieur d’Electronique de Paris (ISEP), FR

**Abstract**

A refresh free and scalable ultimate DRAM (uDRAM) bitcell and architecture is proposed for embedded application. uDRAM 1T1C bitcell is designed using access Tunnel FETs. Proposed design is able to store the data stably during retaining elimination needing the data refresh. This is achieved using negative differential resistivity property of TFETs and storage capacitor leakage. uDRAM allows scaling of storage capacitor by 87% and 80% in comparison to DDR and eDRAMs, respectively. Implemented design have sub-array read/write access times of < 4ns. Bitcell area of 0.0275μm2 is achieved in 28nm FDSOI-CMOS and is scalable further with technology shrink. Estimated throughput gain is 3.8% to 18% in comparison to CMOS DRAMs by refresh removal.

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

**IP3-10**

**A HARDWARE IMPLEMENTATION OF THE MCAS SYNCHRONIZATION PRIMITIVE**

**Speaker:**
Smruti Sarangi, IIT Delhi, IN

**Authors:**
Srshy Patel, Rajshekar Kalayappan, Ishani Mahajan and Smruti R. 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 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)

**IP3-11**

**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)

**IP3-12**

**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
1: Barcelona Supercomputing Center and Universitat Politècnica de Catalunya, ES; 2: Barcelona Supercomputing Center, ES; 3: Barcelona Supercomputing Center (BSC-CNFS), ES; 4: Barcelona 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)

**IP3-13**

**TECHNOLOGY MAPPING WITH ALL SPIN LOGIC**

**Speaker:**
Azadeh Davoodi, University of Wisconsin - Madison, US

**Authors:**
Boyu Zhang1 and Azadeh Davoodi2

**Abstract**

This work is the first to propose a technology mapping algorithm for All Spin Logic (ASL) device. The ASL device is the most actively-pursued one among spintronics devices which themselves fall under emerging post-CMOS nano-technologies. We identify the shortcomings of directly applying the classical technology mapping with ASL devices, and propose techniques to extend the classical procedure to handle these shortcomings. Our results show that our ASL-aware technology mapping algorithm can achieve on-average 9.15% and up to 27.27% improvement in delay (when optimizing delay) with slight improvement in area, compared to the solution generated by classical technology mapping. In a broader sense, our results show the need for developing circuit-level CAD tools that are aware of and optimized for emerging technologies in order to better assess their promise as we move to the post-CMOS era.

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

**IP3-14**

**A NEW METHOD TO IDENTIFY THRESHOLD LOGIC FUNCTIONS**

**Speaker:**
Spyros Tragoudas, Southern Illinois University Carbondale, US

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

**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)
IP3-15
A BRIDGING FAULT MODEL FOR LINE COVERAGE IN THE PRESENCE OF UNDETECTED TRANSITION FAULTS
Speaker and Author:
Irith Pomerantz, 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)

IP3-16
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)

IP3-17
MOBIXEN: PORTING XEN ON ANDROID DEVICES FOR MOBILE VIRTUALIZATION
Speaker:
Jianpu Yao, Shanghai Jiao Tong University, CN
Authors:
Yaozu Dong1, Jianguo Yao2, Haibing Guan2, Ananth. Krishna R1 and Yunhong Jiang 1
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)

IP3-18
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 Kosmidis 2
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)

IP7 Ten Cent Chip Challenge - Interactive Presentations
Date: Wednesday 29 March 2017
Time: 16:00 - 16:30
Location / Room: IP session (in front of room SBC)

<table>
<thead>
<tr>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>IP7.1.1</td>
<td>A LOW-POWER IOT PROCESSOR INTEGRATING VOLTAGE-SCALABLE FULLY DIGITAL MEMORIES</td>
<td>Hidetoshi Ondotera, Kyoto University, JP</td>
</tr>
<tr>
<td>IP7.1.2</td>
<td>A SIMPLE, STATELESS, COST EFFECTIVE SYMMETRIC CRYPTOGRAPHY STRATEGY FOR ENERGY-HARVESTING IOT DEVICES</td>
<td>Jan Madsen, Technical University of Denmark, DK</td>
</tr>
<tr>
<td>IP7.1.3</td>
<td>RECONFIGURABLE MICROCONTROLLER FOR END NODES IN INTERNET OF THINGS</td>
<td>Wai-Chung Matthew Tang, Queen Mary University of London, GB</td>
</tr>
<tr>
<td>IP7.1.4</td>
<td>FURTHER SIMPLIFICATION OF APPROXIMATE ADDERS USING INPUT DATA RANGES IN IOT</td>
<td>Jeong-A Lee, Chosun University, KR</td>
</tr>
</tbody>
</table>

UB08 Session 8
Date: Wednesday 29 March 2017
Time: 16:00 - 18:00
Location / Room: Booth 1, Exhibition Area

<table>
<thead>
<tr>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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 of 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 integrates, in a simulated 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 tool supporting security as a feature of the design process. The complete COSSIM framework together with its sophisticated GUI will be presented.

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 HDI 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 to analogous simulation-based campaigns.

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.

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 huge 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 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.

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

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.

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 Muttillio 1, Daniele Di Pompeos1, Emilio Incerto 2 and Daniele Ciambrone 1

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.
8.1 IoT Day Hot Topic Session: Challenges and Potentials for IoT Rollout

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

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

Chair:
Andreas Herkersdorf, TU Muenchen, DE

Co-Chair:
Marilyn Wolf, Georgia Tech, US

Realizing the potential of IoT will require coordinated advances in multiple markets: applications, software systems, and VLSI. Understanding the requirements on IoT devices requires understanding the stack in which they operate. This session pulls together several points of view on the big picture of IoT rollout and their implications for device and system design.

<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>8.1.1</td>
<td>ULTRA-LOW POWER AND DEPENDABILITY FOR IOT DEVICES</td>
<td>Santiago Pagani, KIT Karlsruhe, DE; Joerg Henkel1, Santiago Pagani1, Hussam Amrouch1, Lars Bauer1 and Farzad Samie1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abstract</td>
<td>Recent advances in technologies have allowed the design of small-size low-power and low-cost devices that can be connected to the Internet, enabling the emerging paradigm of Internet-of-things (IoT). IoT covers an ever-increasing range of applications, e.g., health-care monitoring, smart homes and buildings, etc. In this invited paper, we discuss and summarize the IoT paradigm with a special focus on energy consumption and methodologies for its minimization. Furthermore, we also discuss about reliability in the context of IoT devices. In all, this paper attempts to be a starting point for readers interested in developing energy-efficient IoT devices.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Download Paper (PDF); Only available from the DATE venue WiFi</td>
<td></td>
</tr>
<tr>
<td>17:30</td>
<td>8.1.2</td>
<td>SMARTER SPACES THROUGH LOCAL(I2ED) OBJECT INTERACTIONS</td>
<td>Jean-Marie Bonnin, Telecom Bretagne, FR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abstract</td>
<td>The technologies necessary for the development of pervasive applications are now widely available and accessible for many uses: short/long-range and low energy communications, a broad variety of visible (smart objects) or invisible (sensors and actuators) objects, as well as the democratization of the Internet of Things (IoT). Large areas of our living spaces are now instrumented. The concept of Smart Spaces is about to emerge, based on both massive and apposite interactions between individuals and their everyday working and living environments. The potential applications are boundless. However, many scenarios are often designed in an ad-hoc manner depending on the target area of application. Resources (sensors / actuators, connected objects etc.) are used in silos which prevents using them for implementing several pervasive computing scenarios. They can only be used in the environment they had especially been developed for (for example &quot;classical&quot; home automation tasks: comfort, entertainment, surveillance). They are difficult to adapt to increasingly complex situations, even though the environments in which they evolve are more open, or change over time (new sensors added, failures, mobility etc.) As fine decisions can be made close to the objects producing and acting on the data. Local data characterization and local processing de-emphasize the computing and storage resources of the cloud. Therefore, developing a comprehensive set of new interactions models between objects in the field could help pervasive application designers in the development phase with the side effect to ease the life cycle management, and make objects more useful and more durable.</td>
</tr>
<tr>
<td>18:00</td>
<td>8.1.3</td>
<td>DEPLOYING IOT FOR INSTRUMENTATION AND ANALYSIS OF MANUFACTURING SYSTEMS</td>
<td>Sujit Roska Chhetri, UC Irvine, US</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Author</td>
<td>Mohammad Al Faruque, University of California Irvine, US</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abstract</td>
<td>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.</td>
</tr>
</tbody>
</table>

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
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.

### Time  Label Presentation Title Authors

17:00 8.2.1 ENERGY-DRIVEN COMPUTING: RETHINKING THE DESIGN OF ENERGY HARVESTING SYSTEMS  Geoff Merrett, 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)**

17:30 8.2.2 NONVOLATILE PROCESSORS: WHY IS IT TRENDING?  Vijaykrishnan Narayanan, Penn State University, US
**Authors:**
Fang Su, Kaisheng Ma, Xueqing Li, Tongda Wu, Yongpan Liu, and Vijaykrishnan Narayanan
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)**

18:00 8.2.3 ADVANCED SPINTRONIC MEMORY AND LOGIC FOR NON-VOLATILE PROCESSORS  X. Sharon Hu, University of Notre Dame, US
**Authors:**
Robert Perricone, Ibrahim Ahmed, Zhaoxin Liang, Meghna Mankalale, X. Sharon Hu, Chris H. Kim, Michael Niemier, Sachin Sapatnekar, and Jian-Ping Wang
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)**

18:30 End of session

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.

**8.4 Advanced systems for healthcare and assistive technologies**

**Date:** Wednesday 29 March 2017  
**Time:** 17:00 - 18:30  
**Location / Room:** 3A  
**Chair:** Ruben Braojos, EPFL, CH  
**Co-Chair:** Luca Fanucci, University of Pisa, IT

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.

**Time** | **Label** | **Presentation Title** | **Authors**
--- | --- | --- | ---
17:00 | 8.4.1 | **ADAPTIVE COMPRESSED SENSING AT THE FINGERTIP OF INTERNET-OF-THINGS SENSORS: AN ULTRA-LOW POWER ACTIVITY RECOGNITION**  
**Speaker:** Josué Pagán Ortiz, UCM, ES  
**Authors:**  
Josué Pagán Ortiz, UCM, ES;  
Hassan Ghahmemzadeh

*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 feedback 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.

8.4.2 A ZYNQ-BASED DYNAMICALLY RECONFIGURABLE HIGH DENSITY MYOELECTRIC PROSTHESIS CONTROLLER

Speaker: Linus Witschen, Paderborn University, DE
Authors: Alexander Boschmann1, Georg Thombsen1, Linus Witschen1, Alex Wiens1 and Marco Platzner2
1Paderborn University, DE; 2University of Paderborn, DE

Abstract
The combination of high-density electromyographic (HD EMG) sensor technology and modern machine learning algorithms allows for intuitive and robust prosthesis control of multiple degrees of freedom. However, HD EMG real-time processing poses a challenge for common microprocessors in an embedded system. With the goal set on an autonomous prosthesis capable of performing training and classification of an amputee’s HD EMGs signals, the focus of this paper lies in the acceleration of the computationally expensive parts of the embedded signal processing chain: the feature extraction and classification. Using the Xilinx Zynq as a low-cost off-the-shelf system, we present a solution capable of processing 192 HD EMG channels with controller delays below 120 milliseconds, suitable for highly responsive real-world prosthesis control, achieving speed-ups up to 2.8 as compared to a software-only solution. Using dynamic FPGA reconfiguration, the system is able to trade-off increased controller delay against improved classification accuracy when signal quality is decreased due to noisy channels. Offloading feature extraction and classification to the FPGA also reduced the system’s power consumption, making it more suitable to be used in a battery-powered setup. The system was validated using real-time experiments with online HD EMG data from an amputee to control a state-of-the-art prosthesis.

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

18:00 8.4.3 MICROWATT END-TO-END DIGITAL NEURAL SIGNAL PROCESSING SYSTEMS FOR MOTOR INTENTION DECODING

Speaker: Zhewei Jiang, Columbia University, US
Authors: Zhewei Jiang1, Chisung Bae2, Joongsong Kang2, Sang Joon Kim2 and Mingoo Seok1
1Columbia University, US; 2Samsung Electronics, KR

Abstract
This paper presents microwatt end-to-end digital signal processing (DSP) systems for deployment-stage real-time upper-limb movement intent prediction. This brain computer interface (BCI) DSP systems feature intercellular spike detection, sorting, and decoding operations for a 96-channel prosthetic implant. We design the algorithms for those operations to achieve minimal computation complexity while matching or advancing the accuracy of state-of-art BCI sorting and movement decoding. Based on those algorithms, we architect the DSP hardware with the focus on hardware reuse and event-driven operation. The VLSI implementation of the proposed architecture in a 65-nm high-VTH shows that it can achieve 7.7μW at the supply voltage of 300mV in the post-layout simulation. The area is 0.16 mm2.

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

18:15 8.4.4 AN EMBEDDED SYSTEM REMOTELY DRIVING MECHANICAL DEVICES BY P300 BRAIN ACTIVITY

Speaker: Daniela De Venuto, Politecnico di Bari, IT
Authors: Valerio F. Annese1, Giovanni Mezzina2 and Daniela De Venuto2
1Politecnico di Bari, IT; 2Dept. of Electrical and Information Engineering, Politecnico di Bari, IT

Abstract
In this paper we present a P300-based Brain Computer Interface (BCI) for the remote control of a mechatronic actuator, such as a wheelchair, or even a car, driven by EEG signals to be used by tetraplegic and paralytic users or just for safe drive in case of car. The P300 signal, an Evoked Related Potential (ERP) devoted to the cognitive brain activity, is induced for purpose by visual stimulation. The ERP data are collected by 6 smart wireless electrodes from the parietal-cortex area and online classified by a linear threshold classifier, based on a suitable stage of Machine Learning (ML). The ML is implemented on a µPC dedicated to the system and where the data acquisition and processing is performed. The main improvement in remote driving car by EEG, regards the approach used for the intentions recognition. In this work, the classification is based on the P300 and not just on the average of more not well identify potentials. This approach reduces the number of electrodes on the EEG helmet. The ML stage is based on a custom algorithm (t-RIDE) which tunes the following classification stage on the user’s “cognitive chronometry”. The ML algorithm starts with a fast calibration phase (just ~190s for the first learning). Furthermore, the BCI presents a functional approach for time-domain features extraction, which reduces the amount of data to be analyzed, and then the system response times. In this paper, a proof of concept of the proposed BCI is shown using a prototype car, tested on 5 subjects (aged 26 ± 3). The experimental results show that the novel ML approach allows a complete P300 spatio-temporal characterisation in 1.95s using 38 target brain visual stimuli (for each direction of the car path). In free-drive mode, the BCI classification reaches 80.5 ± 4.1% on single-trial detection accuracy while the worst-case computational time is 19.65ms ± 10.1. The BCI system here described can be also used on different mechatronic actuators, such as robots.

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

18:31 8.4.1 1024-CHANNEL 3D ULTRASOUND DIGITAL BEAMFORMER IN A SINGLE 5W FPGA

Speaker: Aya Ibrahim, EPFL, CH
Authors: Federico Angiolini1, Aya Ibrahim1, William Simon1, Ahmet Caner Yüzügüler1, Marcel Arditi1, Jean-Philippe Thiran1 and Giovanni De Micheli2
1EPFL, CH; 2École Polytechnique Fédérale de Lausanne (EPFL), CH

Abstract
3D ultrasound, an emerging medical imaging tech-nique 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 targets 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. The channel probe, which enables high-quality imaging, consumes only about 5W.

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

18:30 End of session
Efficiently exploiting thread level parallelism from new multicore systems has been challenging for software developers. While blindly increasing the number of threads may lead to performance gains, it can also 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, the number of threads of OpenMP is activated.

Efficiently exploiting thread level parallelism from new multicore systems has been challenging for software developers. While blindly increasing the number of threads may lead to performance gains, it can also 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, the number of threads of OpenMP is activated.

In this paper, we illustrate "choke points" as a vital consequence of process variation in the Near Threshold Computing (NTC) domain. Choke points are sensitized logic gates with increased delay deviation, due to process variation.

In this paper, we illustrate "choke points" as a vital consequence of process variation in the Near Threshold Computing (NTC) domain. Choke points are sensitized logic gates with increased delay deviation, due to process variation.

In this paper, we illustrate "choke points" as a vital consequence of process variation in the Near Threshold Computing (NTC) domain. Choke points are sensitized logic gates with increased delay deviation, due to process variation.

In this paper, we illustrate "choke points" as a vital consequence of process variation in the Near Threshold Computing (NTC) domain. Choke points are sensitized logic gates with increased delay deviation, due to process variation.

In this paper, we illustrate "choke points" as a vital consequence of process variation in the Near Threshold Computing (NTC) domain. Choke points are sensitized logic gates with increased delay deviation, due to process variation.

In this paper, we illustrate "choke points" as a vital consequence of process variation in the Near Threshold Computing (NTC) domain. Choke points are sensitized logic gates with increased delay deviation, due to process variation.

In this paper, we illustrate "choke points" as a vital consequence of process variation in the Near Threshold Computing (NTC) domain. Choke points are sensitized logic gates with increased delay deviation, due to process variation.

In this paper, we illustrate "choke points" as a vital consequence of process variation in the Near Threshold Computing (NTC) domain. Choke points are sensitized logic gates with increased delay deviation, due to process variation.

In this paper, we illustrate "choke points" as a vital consequence of process variation in the Near Threshold Computing (NTC) domain. Choke points are sensitized logic gates with increased delay deviation, due to process variation.

In this paper, we illustrate "choke points" as a vital consequence of process variation in the Near Threshold Computing (NTC) domain. Choke points are sensitized logic gates with increased delay deviation, due to process variation.

In this paper, we illustrate "choke points" as a vital consequence of process variation in the Near Threshold Computing (NTC) domain. Choke points are sensitized logic gates with increased delay deviation, due to process variation.

In this paper, we illustrate "choke points" as a vital consequence of process variation in the Near Threshold Computing (NTC) domain. Choke points are sensitized logic gates with increased delay deviation, due to process variation.

In this paper, we illustrate "choke points" as a vital consequence of process variation in the Near Threshold Computing (NTC) domain. Choke points are sensitized logic gates with increased delay deviation, due to process variation.

In this paper, we illustrate "choke points" as a vital consequence of process variation in the Near Threshold Computing (NTC) domain. Choke points are sensitized logic gates with increased delay deviation, due to process variation.

In this paper, we illustrate "choke points" as a vital consequence of process variation in the Near Threshold Computing (NTC) domain. Choke points are sensitized logic gates with increased delay deviation, due to process variation.

In this paper, we illustrate "choke points" as a vital consequence of process variation in the Near Threshold Computing (NTC) domain. Choke points are sensitized logic gates with increased delay deviation, due to process variation.

In this paper, we illustrate "choke points" as a vital consequence of process variation in the Near Threshold Computing (NTC) domain. Choke points are sensitized logic gates with increased delay deviation, due to process variation.

In this paper, we illustrate "choke points" as a vital consequence of process variation in the Near Threshold Computing (NTC) domain. Choke points are sensitized logic gates with increased delay deviation, due to process variation.

In this paper, we illustrate "choke points" as a vital consequence of process variation in the Near Threshold Computing (NTC) domain. Choke points are sensitized logic gates with increased delay deviation, due to process variation.

In this paper, we illustrate "choke points" as a vital consequence of process variation in the Near Threshold Computing (NTC) domain. Choke points are sensitized logic gates with increased delay deviation, due to process variation.

In this paper, we illustrate "choke points" as a vital consequence of process variation in the Near Threshold Computing (NTC) domain. Choke points are sensitized logic gates with increased delay deviation, due to process variation.

In this paper, we illustrate "choke points" as a vital consequence of process variation in the Near Threshold Computing (NTC) domain. Choke points are sensitized logic gates with increased delay deviation, due to process variation.

In this paper, we illustrate "choke points" as a vital consequence of process variation in the Near Threshold Computing (NTC) domain. Choke points are sensitized logic gates with increased delay deviation, due to process variation.

In this paper, we illustrate "choke points" as a vital consequence of process variation in the Near Threshold Computing (NTC) domain. Choke points are sensitized logic gates with increased delay deviation, due to process variation.

In this paper, we illustrate "choke points" as a vital consequence of process variation in the Near Threshold Computing (NTC) domain. Choke points are sensitized logic gates with increased delay deviation, due to process variation.

In this paper, we illustrate "choke points" as a vital consequence of process variation in the Near Threshold Computing (NTC) domain. Choke points are sensitized logic gates with increased delay deviation, due to process variation.

In this paper, we illustrate "choke points" as a vital consequence of process variation in the Near Threshold Computing (NTC) domain. Choke points are sensitized logic gates with increased delay deviation, due to process variation.

In this paper, we illustrate "choke points" as a vital consequence of process variation in the Near Threshold Computing (NTC) domain. Choke points are sensitized logic gates with increased delay deviation, due to process variation.
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 diverse domains such as automotive, space, military, consumer electronics, industrial control, health care, etc. 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.

**Time** | **Label** | **Presentation Title** | **Authors**
---|---|---|---
18:31 | IP4-3, 68 | **IMPROVING THE ACCURACY OF THE LEAKAGE POWER ESTIMATION OF EMBEDDED CPUS** | Shiao-Li Tsao, National Chiao Tung University, TW
**Speaker:** Shiao-Li Tsao, National Chiao Tung University, TW  
**Authors:** Ting-Wu Chin, Shiao-Li Tsao, Kuo-Wei Hung and Pei-Shu Huang, 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)**

18:30 | **End of session** | | |

---

**Time** | **Label** | **Presentation Title** | **Authors**
---|---|---|---
18:00 | 8.6.3 | **SELF-AWARENESS IN REMOTE HEALTH MONITORING SYSTEMS THROUGH WEARABLE ELECTRONICS** | Axel Jantsch, TU Wien, AT
**Speaker:** Axel Jantsch, TU Wien, AT  
**Authors:** Arman Anzanour1, Iman Azimi1, Maximilian Göttinger1, Amir M. Rahmani2, Nima Taherinejad3, Pasi Liljeberg4, Axel Jantsch2 and Nikil Dutt3  
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 having 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, 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)**

18:00 | 8.6.2 | **SELF-AWARENESS IN AUTONOMOUS SYSTEMS: SELF-DRIVING CARS** | Rolf Ernst, TU Braunschweig, DE
**Speaker:** Rolf Ernst, TU Braunschweig, DE  
**Authors:** Johannes Schlatow1, Mischa Höst2, Rolf Ernst2, Marcus Nolte3, Inga Jatzkowskij4, Markus Maurer2, 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**  
One vision for how to rise to this challenge is to endow computing systems with increased self-awareness, in order to enable advanced autonomous adaptive behaviour. A desire for self-awareness has arisen in a variety of areas of computer science and engineering over the last two decades, and more recently a more fundamental understanding of what self-awareness concepts might mean for the design and operation of computing systems has been developed. This draws on self-awareness theories from psychology and other related fields, and has led to a number of contributions in terms of definitions, architectures, algorithms and case studies. This paper introduces some of the main aspects of self-awareness from psychology, that have been used in developing associated notions in computing. It then 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 priori domain modelling at design or deployment time.
**Download Paper (PDF; Only available from the DATE venue WiFi)**

18:30 | **End of session** | | |
### 8.7 Instruction-level and thread-level parallelism in embedded systems

**Date:** Wednesday 29 March 2017  
**Time:** 17:00 - 18:30  
**Location / Room:** 3B

**Chair:**  
Oliver Bringmann, Universität Tübingen, DE

**Co-Chair:**  
Jürgen Teich, Friedrich-Alexander-Universität Erlangen-Nürnberg, DE

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 floating point to fixed point conversion by exploiting word-level parallelism. The third paper presents a WCET analysis for multiple tasks on single-core systems.

### Time | Label | Presentation Title  
--- | --- | ---  
17:00 | 8.7.1 | HARDWARE-ACCELERATED DYNAMIC BINARY TRANSLATION  
**Speaker:** Simon Rokicki, Université de Rennes 1 / IRISA, FR  
**Authors:** Simon Rokicki¹, Erven Rohou² and Steven Derrien¹  
¹IRISA, FR; ²IRISA, FR  
**Abstract**  
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 shows 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.  
Download Paper (PDF; Only available from the DATE venue WiFi)

17:30 | 8.7.2 | SUPERWORD LEVEL PARALLELISM AWARE WORD LENGTH OPTIMIZATION  
**Speaker:** Ali Hassan El Moussawi, IRISA, FR  
**Authors:** Ali Hassan El Moussawi¹ and Steven Derrien²  
¹IRISA, FR; ²IRISA, FR  
**Abstract**  
Many embedded processors do not support floating-point arithmetic in order to comply with strict cost and power consumption constraints. But, they generally provide support for SIMD as a mean to improve performance for little cost overhead. Achieving good performance when targeting such processors requires the use of fixed-point arithmetic and efficient exploitation of SIMD data-path. To reduce time-to-market, automatic SIMDization -- such as superword level parallelism (SLP) extraction -- and floating-point to fixed-point conversion methodologies have been proposed. In this paper we show that applying these transformations independently is not efficient. We propose a SLP-aware word length optimization algorithm to jointly perform float-to-fixed-point conversion and SLP extraction. We implement the proposed approach in a source-to-source compiler framework and evaluate it on several embedded processors. Experimental results illustrate the validity of our approach.  
Download Paper (PDF; Only available from the DATE venue WiFi)

18:00 | 8.7.3 | SCHEDULABILITY-AWARE SPM ALLOCATION FOR PREEMPTIVE HARD REAL-TIME SYSTEMS WITH ARBITRARY ACTIVATION PATTERNS  
**Speaker:** Arno Luppold, Hamburg University of Technology, DE  
**Authors:** Arno Luppold¹ and Heiko Falk²  
¹Hamburg University of Technology, DE; ²Hamburg University of Technology (TUHH), DE  
**Abstract**  
In hard real-time multi-tasking systems each task has to meet its deadline under any circumstances. If one or several tasks violate their timing constraints, compiler optimizations can be used to optimize the Worst-Case Execution Time (WCET) of each task with a focus on the system’s schedulability. Existing approaches are limited to single-tasking or strictly periodic multi-tasking systems. We propose a compiler optimization to perform a schedulability-aware static instruction Scratchpad Allocation for arbitrary activation patterns and deadlines. The approach is based on Integer-Linear Programming and is evaluated for the Infineon TriCore TCI796 microcontroller.  
Download Paper (PDF; Only available from the DATE venue WiFi)

18:30 | IP4-4, 636 | SCHEDULE-AWARE LOOP PARALLELIZATION FOR EMBEDDED MPSoCs BY EXPLOITING PARALLEL SLACK  
**Speaker:** Miguel Angel Aguilar, RWTH Aachen University, DE  
**Authors:** Miguel Angel Aguilar¹, Rainer Leupers¹, Gerd Ascheid¹, Nikolaos Kavvadias² and Liam Fitzpatrick²  
¹RWTH Aachen University, DE; ²Silexica Software Solutions GmbH, DE  
**Abstract**  
MPSoC programming is still a challenging task, where several aspects have to be taken into account to achieve a profitable parallel execution. Selecting a proper scheduling policy is an aspect that has a major impact on the performance. OpenMP is an example of a programming paradigm that allows to specify the scheduling policy on a per loop basis. However, choosing the best scheduling policy and the corresponding parameters is not a trivial task. In fact, there is already a large amount of software parallelized with OpenMP, 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 on an embedded multicore device, show that the performance achieved by OpenMP loops optimized with our approach outperform by up to 93%, the performance achieved by the original OpenMP loops, where the scheduling policy is not specified.  
Download Paper (PDF; Only available from the DATE venue WiFi)
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.

### Time | Label | Presentation Title | Authors
--- | --- | --- | ---
08:30 | 9.1.1 | WEARABLE ROBOTICS IN CLINICAL PRACTICE: PROSPECTS | José Luis Pons, CSIC, ES
09:00 | 9.1.2 | OVERCOMING HEARING LOSS THROUGH NEW IMPLANT TECHNOLOGIES | Carl Van Himbeeck, Cochlear Technology Centre, BE
     | Abstract | 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.
10:00 | 9.1.3 | CIRCUITS AND SYSTEMS AS ENABLERS FOR NOVEL HEALTHCARE PARADIGMS | Mario Konijnenburg, imec, BE

### Time | Label | Authors
--- | --- | ---
08:30 - 11:30 | Coffee Break | in Exhibition Area
16:00 - 17:00 | Coffee Break | in Exhibition Area

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.

### Time | Label | Presentation Title | Authors
--- | --- | --- | ---

### Paper 1: A LOG-AWARE SYNERGIZED SCHEME FOR PAGE-LEVEL FTL DESIGN

**Authors:**
Chu Li, Huazhong University of Science & Technology, CN

**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 lines 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)**

### Paper 2: MALRU: MISS-PENALTY AWARE LRU-BASED CACHE REPLACEMENT FOR HYBRID MEMORY SYSTEMS

**Authors:**
Chen Di, 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)**

### Paper 3: ENDURANCE MANAGEMENT FOR RESISTIVE LOGIC-IN-MEMORY COMPUTING ARCHITECTURES

**Authors:**
Saeideh Shirinzadeh, University of Bremen, DE

**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 (PLiM) 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 PLiM 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 86.69% 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)**

### Paper 4: LIVE TOGETHER OR DIE ALONE: BLOCK COOPERATION TO EXTEND LIFETIME OF RESISTIVE MEMORIES

**Authors:**
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 write 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% (17%), and 8% (14%) on average, respectively.

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

### Paper 5: DAC: DEDUP-ASSISTED COMPRESSION SCHEME FOR IMPROVING LIFETIME OF NAND STORAGE SYSTEMS

**Authors:**
Jisung Park, Seoul National 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 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)**
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)
09:45 9.3.1  SECURE CYBER-PHYSICAL SYSTEMS: CURRENT TRENDS, TOOLS AND OPEN RESEARCH PROBLEMS  
Speaker: Anupam Chattopadhyay, Nanyang Technological University, SG  
Authors: Anupam Chattopadhyay¹, Alok Prakash¹ and Muhammad Shafique²  
¹Nanyang Technological University, SG; ²Vienna University of Technology (TU Wien), AT  
Download Paper (PDF; Only available from the DATE venue WiFi)

08:45 9.3.2  DON'T FALL INTO A TRAP: PHYSICAL SIDE-CHANNEL ANALYSIS OF CHACHA20-POLY1305  
Speaker: Bernhard Jungk, Temasek Laboratories @ Nanyang Technological University, SG  
Authors: Bernhard Jungk¹ and Shivam Bhasin²  
¹Temasek Laboratories @ Nanyang Technological University, SG; ²TL@NTU, SG  
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)

09:00 9.3.3  THE ROWHAMMER PROBLEM AND OTHER ISSUES WE MAY FACE AS MEMORY BECOMES DENSER  
Speaker and Author: Onur Mutlu, ETH Zurich, CH  
Abstract  
As memory scales down to smaller technology nodes, new failure mechanisms emerge that threaten their correct operation. If such failure mechanisms are not anticipated and corrected, they can not only degrade system reliability and availability but also, perhaps even more importantly, open up security vulnerabilities: a malicious attacker can exploit the exposed failure mechanism to take over the entire system. As such, new failure mechanisms in memory can become practical and significant threats to system security. In this work, we discuss the RowHammer problem in DRAM, which is a prime (and perhaps the first) example of how a circuit-level failure mechanism in DRAM can cause a practical and widespread system security vulnerability. RowHammer, as it is popularly referred to, is the phenomenon that repeatedly accessing a row in a modern DRAM chip causes bit flips in physically-adjacent rows at consistently predictable bit locations. It is caused by a hardware failure mechanism called DRAM disturbance errors, which is a manifestation of circuit-level cell-to-cell interference in a scaled memory technology. Researchers from Google Project Zero recently demonstrated that this hardware failure mechanism can be effectively exploited by user-level programs to gain kernel privileges on real systems. Several other recent works demonstrated other practical attacks exploiting RowHammer. These include remote takeover of a server vulnerable to RowHammer, takeover of a victim virtual machine by another virtual machine running on the same system, and takeover of a mobile device by a malicious user-level application that requires no permissions. We analyze the root causes of the RowHammer problem and examine various solutions. We also discuss what other vulnerabilities may be lurking in DRAM and other types of memories, e.g., NAND flash memory or Phase Change Memory, that can potentially threaten the foundations of secure systems, as the memory technologies scale to higher densities. We conclude by describing and advocating a principled approach to memory reliability and security research that can enable us to better anticipate and prevent such vulnerabilities.  
Download Paper (PDF; Only available from the DATE venue WiFi)

09:15 9.3.4  COMPROMISING FPGA SOCS USING MALICIOUS HARDWARE BLOCKS  
Speaker: Nisha Jacob, Fraunhofer AISEC, DE  
Authors: Nisha Jacob¹, Carsten Rolfes¹, Andreas Zankl¹, Johann Heynszl¹ and Georg Sigl²  
¹Fraunhofer Institute for Applied and Integrated Security (AISEC), DE; ²Technische Universität München, DE  
Abstract  
Modern FPGA System-on-Chips (SoCs) combine high performance application processors with reconfigurable hardware. This allows to enhance complex software systems with reconfigurable hardware accelerators. Unfortunately, even when state-of-the-art software security mechanisms are implemented, this combination creates new security threats. Attacks on the software are now possible through the reconfigurable hardware as these cores share resources with the processor and may contain unwanted functionality. In this paper, we discuss software protection mechanisms offered in conventional SoCs and how they can be circumvented by malicious hardware blocks. As a concrete example, we show how the malicious functionality within an IP core accesses and replaces critical memory sections. We refer to this type of attacks as hardware-assisted attacks against running software systems. We carry-out a proof-of-concept on the Xilinx Zynq device which runs a Linux OS and a software application that verifies system updates. The malicious IP core replaces the public key used to verify system updates, thus, allowing an attacker to maliciously update the FPGA SoC. Additionally, we propose a countermeasure that can be applied against such threats in the form of a security wrapper for hardware modules.  
Download Paper (PDF; Only available from the DATE venue WiFi)

09:30 9.3.5  INSPIRING TRUST IN OUTSOURCED INTEGRATED CIRCUIT FABRICATION  
Speaker and Author: Siddharth Garg, New York University, US  
Abstract  
Outsourcing of integrated circuits (ICs) is typically outsourced to an external semiconductor foundry to reduce cost. However, this can come at the expense of trust. How can a designer ensure the integrity of the ICs fabricated by an external foundry? The talk will discuss a new approach for insuring trust in outsourced IC fabrication by complementing the untrusted (outsourced) with an IC fabricated at a low-end but trusted foundry. This approach is referred to as split fabrication. We present two different ways in which split fabrication can be used to enhance security: logic obfuscation and verifiable ASICs.  
Download Paper (PDF; Only available from the DATE venue WiFi)

09:45 9.3.6  ANALYZING SECURITY BREACHES OF COUNTERMEASURES THROUGHOUT THE REFINEMENT PROCESS IN HARDWARE DESIGN FLOW  
Speaker: Jean-Luc Danger, Secure-IC, FR  
Authors: Sylvain Guille, Jean-Luc Danger, Philippe Nguyen, Robert Nguyen and Youssef Souissi, Secure-IC S.A.S., FR  
Abstract  
Side-channel and fault injection attacks are two threats on devices carrying sensitive information. Protections are thus implemented at design time. However, CAD (Computer Aided Design) tools can compromise them, in ways we detail pedagogically in this paper. Then, we explain how a simulation-based methodology allows to check for non-regression, and find problems in case some are introduced while refining the design description from RTL (Register Transfer Level) source code to GDS (Graphic Display System) stream format.  
Download Paper (PDF; Only available from the DATE venue WiFi)
### 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.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
</table>
| 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 |
| 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) Liang³, Tulika Mitra¹ and Small Niar⁴  
¹National University of Singapore, SG; ²Nanyang Technological University, SG; ³Peking University, CN; ⁴LAMIH-University of Valenciennes, FR |

---

**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>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>09:30</td>
<td>9.4.3</td>
<td>DESIGN SPACE EXPLORATION OF FPGA ACCELERATORS FOR CONVOLUTIONAL NEURAL NETWORKS</td>
<td>Jongeun Lee, UNIST, KR</td>
</tr>
<tr>
<td>Authors:</td>
<td></td>
<td></td>
<td>Atul Rahman1, Sangyun Oh2, Jongeun Lee3 and Kiyoung Choi4</td>
</tr>
<tr>
<td>1Samsung Electronics, KR; 2UNIST, KR; 3Ulsan National Institute of Science and Technology (UNIST), KR; 4Seoul National University, KR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abstract</td>
<td></td>
<td></td>
<td>The increasing use of machine learning algorithms, such as Convolutional Neural Networks (CNNs), makes the hardware accelerator approach very compelling. However, the question of how to best design an accelerator for a given CNN has not been answered yet, even on a very fundamental level. This paper addresses that challenge, by providing a novel framework that can universally and accurately evaluate and explore various architectural choices for CNN accelerators on FPGAs. Our exploration framework is more extensive than that of any previous work in terms of the design space, and takes into account various FPGA resources to maximize performance including DSP resources, on-chip memory, and off-chip memory bandwidth. Our experimental results using some of the largest CNN models including one that has 16 convolutional layers demonstrate the efficacy of our framework, as well as the need for such a high-level architecture exploration approach to find the best architecture for a CNN model.</td>
</tr>
<tr>
<td>Download Paper (PDF; Only available from the DATE venue WiFi)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>09:45</td>
<td>9.4.4</td>
<td>A SLACK-BASED APPROACH TO EFFICIENTLY DEPLOY RADIX 8 BOOTH MULTIPLIERS</td>
<td>Alberto Antonio Del Barrio Garcia, Universidad Complutense de Madrid, ES</td>
</tr>
<tr>
<td>Authors:</td>
<td></td>
<td></td>
<td>Alberto Antonio Del Barrio Garcia and Hermida Roman, Complutense University of Madrid, ES</td>
</tr>
<tr>
<td>Abstract</td>
<td></td>
<td></td>
<td>In 1951 A. Booth published his algorithm to efficiently multiply signed numbers. Since the appearance of such algorithm, it has been widely accepted that radix 4-based Booth multipliers are the most efficient. They allow the height of the multiplier to be halved, at the expense of a simple recoding that consists of just shifts and negations. Theoretically, higher radix should produce even larger reductions, especially in terms of area and power, but the recoding process is much more complex. Notably, in the case of radix 8 it is necessary to compute 3X, X being the multiplicand. In order to avoid the penalty due to this calculation, we propose decoupling it from the product and considering 3X as an extra operation within the application's Dataflow Graph (DFG). Experiments show that typically there is enough slack in the DFGs to do this without degrading the performance of the circuit, which permits the efficient deployment of radix 8 multipliers that do not calculate the 3X multiple. Results show that our approach is 10% and 17% faster than radix 4 and radix 8 Booth based implementations, respectively, and 12% and 10% more energy efficient in terms of Energy Delay Product.</td>
</tr>
<tr>
<td>Download Paper (PDF; Only available from the DATE venue WiFi)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
08:30 9.5.1 MEASUREMENT AND VALIDATION OF ENERGY HARVESTING IOT DEVICES
Speaker: Lukas Sigrist, ETH Zurich, CH
Authors: Lukas Sigrist1, Andres Gomez1, Roman Lim2, Stefan Lippuner1, Matthias Leubin1 and Lothar Thiele2
1ETH Zurich, CH; 2Swiss Federal Institute of Technology 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 properties: portability and data acquisition and analysis. The RocketLogger measures current, voltage and temperature with a resolution of 1 mV, 1 mμA and 0.01°C, respectively. The range switching is performed from 4 pW 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)

09:00 9.5.2 A METHODOLOGY FOR THE DESIGN OF DYNAMIC ACCURACY OPERATORS BY RUNTIME BACK BIAS
Speaker: Daniele Jahier Pagliari, Politecnico di Torino, IT
Authors: Daniele Jahier Pagliari1, Yves Durand2, David Coriat2, Anca Molnos2, Edith Beigne2, Enrico Macii1 and Massimo Poncino1
1Politecnico di Torino, IT; 2CEA-Leti, FR
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 modulation 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)

09:30 9.5.3 A SCAN-CHAIN BASED STATE RETENTION METHODOLOGY FOR IOT PROCESSORS OPERATING ON INTERRMITENT ENERGY
Speaker: Pascal Alexander Hager, ETH Zurich, CH
Authors: Pascal Alexander Hager1, Hamed Fatemi2, Jose Pineda2 and Luca Benini1
1ETH Zurich, CH; 2NXP Semiconductors, NL; 3Università di Bologna, IT
Abstract
Future IoT systems are tightly constraint 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 supported by auxiliary supplies for retention or incorporate non-volatile memory cells. However, these SRFFs increase the power consumption during active operation impeding 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 pays off after an on-time of just 100us compared to state-of-the-art in-place solutions. Moreover, compared to a software mechanism, our approach requires 6.6x less energy to move the state and is 5.8x faster.
Download Paper (PDF; Only available from the DATE venue WiFi)

09:45 9.5.4 A CIRCUIT-EQUIVALENT BATTERY MODEL ACCOUNTING FOR THE DEPENDENCY ON LOAD FREQUENCY
Speaker: Yukai Chen, Politecnico di Torino, IT
Authors: Yukai Chen, Enrico Macii and Massimo Poncino, 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 the 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-known 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)

10:00 11.4 ADAPTIVE POWER DELIVERY SYSTEM MANAGEMENT FOR MANY-CORE PROCESSORS WITH ON/OFF-CHIP VOLTAGE REGULATORS
Speaker: Haoran Li, The Hong Kong University of Science and Technology, HK
Authors: Haoran Li, Jiang Xu, Zhe Wang, Peng Yang, Rafael Kioji Vivas Maeda and Zhongyuan Tian, 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 for 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 choices 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)
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)
This paper presents the study and implementation of a new efficient yield optimization technique for multi-objective optimization-based automatic analog integrated circuit sizing. The approach uses a commercial electrical simulator and standard process design kit (PDK) models to perform, during the optimization process, the same Monte Carlo (MC) simulations that designers use. The proposed yield estimation technique reduces the number of required MC simulations by using the k-means algorithm, with a variable number of clusters, to select only a handful potential solutions where the MC simulations are performed. Due to the use of a commercial simulator tool and foundry supplied PDK models the developed methodology provides the most accurate and reliable results, and also, the variable k-means algorithm is able to achieve 91% reduction in the total number of the MC simulations required for an optimization, when considering MC simulations for all solutions. Moreover, this new approach presents a 50% increase in speed performance when comparing to a previous yield optimization technique also using k-means and MC simulations.

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 10:00 - 11:00

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.

AN EFFICIENT LEAKAGE-AWARE THERMAL SIMULATION APPROACH FOR 3D-ICS USING CORRECTED LINEARIZED MODEL AND ALGEBRAIC MULTIGRID

Speaker: Chao Yan, Microelectronics Dept., Fudan University, CN
Authors: Chao Yan1, Hengliang Zhu1, Dian Zhou2, and Xuan Zeng1
1Fudan University, CN; 2University of Texas at Dallas, US

Abstract
Thermal control has become a great challenge for 3D-ICs due to the ever increasing power density and 3D integration. Among techniques to address the problem, fast thermal simulation approach is basically required to accurately characterize the runtime temperature variations of 3D-ICs. In this paper, we propose an accurate and fast leakage-aware thermal simulation approach for 3D-ICs with consideration of both heatsink cooling and microfluidic cooling. First, the proposed approach is based on a corrected linearized model for leakage power approximation, which is proved to be equivalent to the Newton-Chord method for solving nonlinear algebra equations. A convergence comparison is presented in this paper to show that such approach is more efficient than other methods for leakage-aware thermal simulation. Second, an aggregation-based algebraic multigrid (AMG) preconditioned iterative linear solver is adopted that greatly reduces the computation time for solving the linear equations during calculation, which makes the proposed approach even more efficient. Numerical experiments show that the proposed approach can achieve 8x-139x speedup in comparison with the state-of-the-art methods, and with almost negligible average temperature error no more than 0.025K and maximum temperature error no more than 0.095K.

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

CANDY-TM: COMPARATIVE ANALYSIS OF DYNAMIC THERMAL MANAGEMENT IN MANY-CORES USING MODEL CHECKING

Abstract
Dynamic thermal management (DTM) techniques based on task migration provide a promising solution to mitigate thermal emergencies and thereby ensuring safe operation and reliability of Many-Core systems. These techniques can be classified as central or distributed based on the basis of a central DTM controller for the whole system or individual DTM controllers for each core or set of cores in the system, respectively. However, having a trustworthy comparison between central (c-) and distributed (d-) DTM techniques to find out the most suitable one for a given system is quite challenging. This is primarily due to the systemic difference between cDTM and dDTM controllers, and the inherent non-exhaustiveness of simulation and emulation methods conventionally used for DTM analysis. In this paper, we present a novel methodology called CANDy-TM (stands for Comparative Analysis of Dynamic Thermal Management) that employs Model Checking to perform formal comparative analysis for cDTM and dDTM techniques. We identify a set of generic functional and performance properties to provide a common ground for their comparison. We demonstrate the usability and benefits of our methodology by comparing state-of-the-art cDTM and dDTM techniques, and illustrate which technique is good w.r.t. thermal stability and other task migration parameters. Such an analysis helps in selecting the most appropriate DTM for a given chip.

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

POWER PRE-CHARACTERIZED MESHING ALGORITHM FOR FINITE ELEMENT THERMAL ANALYSIS OF INTEGRATED CIRCUITS

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)

A THERMALLY-AWARE ENERGY MINIMIZATION METHODOLOGY FOR GLOBAL INTERCONNECTS

Authors
Shohdy Abdelkader, Software Developer, EG
Syed Ali Asadullah Bukhari1, Faiq Khalid Lodhi2, Osman Hasan3, Muhammad Shafique4 and Joerg Henkel4
1National University of Sciences and Technology - School of Electrical Engineering and Computer Science, PK; 2School of Electrical Engineering and Computer Science National University of Sciences and Technology (NUST), PK; 3Vienna University of Technology (TU Wien), AT; 4Karlsruhe Institute of Technology, DE

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.
9.8 The Internet of INSECURE Things

Date: Thursday 30 March 2017
Time: 08:30 - 10:00
Location / Room: Exhibition Theatre

Organiser: Marcello Coppola, STMicroelectronics, FR

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 networks of insecure things. This session proposes a journey through several speeches to show the advances in technologies that master the security aspects of IoT.

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 key element to make winning products.

The growth and diffusion of high technology consumer communication devices and the following tech skills in average user are pushing industry to put connectivity/network functions in devices. Automotive industry is riding as well this wave. Vehicles are nowadays implementing new "Cyber Physical Features" by collecting information from the physical system and processing it via interconnected cyber systems, creating thus new challenges for safety and security. In addition, an increasing number of vehicles are nowadays connected to the Web, and the capillarity of interconnected IoT devices are drawing the future for the customer expectations in term of innovative services. Historically, security was first of all achieved with isolation of subsystems and, nowadays, with the growing number of interconnected systems that are indirectly interconnected with IoT services highlight how component level countermeasures are important but not enough to enforce protection in a modern vehicle. A multi-level, coordinated, system wide approach is necessary such as isolation of safety critical components, secure gateways, virtualization, trusted software injection and execution, but not only. It requires also a re-design of vehicle data transport infrastructure with new communication standards with the adoption of secure protocols like sCAN.

In this session practical design methods and experiences are presented centered on STM32 devices. Gateways and connected IoT devices networks need to be secured as well as the devices themselves. Suitable safeguards must be integrated to prevent network interfaces and embedded firmware updates from becoming security holes themselves; these safeguards refer to securing the data stored by the device, secure communication and protecting the device from cyber-attacks. Software and hardware development approaches are outlined along with practical experiences that meets the appropriate security level of modern IoT platforms.

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.

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.

In this session we will introduce 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.
### 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**<br>**Speaker:** Aya Ibrahim, EPFL, CH  
**Authors:** Federico Angiolini¹, Aya Ibrahim¹, William Simon¹, Ahmet Caner Yüzügüler¹, Marcel Arditi², Jean-Philippe Thiran¹ and Giovanni De Micheli²<br>¹EPFL, CH; ²École Polytechnique Fédérale de Lausanne (EPFL), CH |  
**Abstract**<br>3D ultrasound, an emerging medical imaging technique 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. |
| IP4-2 | **LAANT: A LIBRARY TO AUTOMATICALLY OPTIMIZE EDP FOR OPENMP APPLICATIONS**<br>**Speaker:** Arthur Francisco Lorenzon, Federal University of Rio Grande do Sul, BR  
**Authors:** Arthur Lorenzon, Jeckson Dellagostin Souza and Antonio Carlos Schneider Beck Filho, Universidade Federal do Rio Grande do Sul, BR |  
**Abstract**<br>Efficiently exploiting thread level parallelism from new multicore systems has been challenging for software developers. While blindly increasing the number of threads may also 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. |
| IP4-3 | **IMPROVING THE ACCURACY OF THE LEAKAGE POWER ESTIMATION OF EMBEDDED CPUs**<br>**Speaker:** Shiao-Li Tsao, National Chiao Tung University, TW  
**Authors:** Ting-Wu Chin, Shiao-Li Tsao, Kuo-Wei Hung and Pei-Shu Huang, National Chiao Tung University, TW |  
**Abstract**<br>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. |
LIFETIME ADAPTIVE ECC IN NAND FLASH PAGE MANAGEMENT

Author:
Shunhuo Wang, Huazhong University of Science and Technology, CN
Shunhuo Wang1, Fei Wu1, Zonghai Lu1, You Zhou1, Qin Xiong1, Meng Zhang1 and Changsheng Xie1
Huazhong 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 a lifetime-adequate 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.

Download Paper (PDF; Only available from the DATE venue WiFi)
### A NEW SAMPLING TECHNIQUE FOR MONTE CARLO-BASED STATISTICAL CIRCUIT ANALYSIS

**Authors:** Hiwa Mahmoudi and Horst Zimmermann, Vienna University of Technology, AT

**Abstract:**

A new sampling technique for improving automated analog yield optimization using a two step exploration strategy. First, a global optimization phase relies on a modified Lipschitzian optimization to sample the potential optimal sub-regions of the feasible design space. The search locates a design point near the optimal solution that is used as a starting point by a local optimization phase. The local search constructs linear interpolating surrogate models of the yield to explore the basin on a modified Lipschitzian optimization to sample the potential optimal sub-regions of the feasible design space. The search locates a design point near the optimal solution that is used as a starting point by a local optimization phase. The local search constructs linear interpolating surrogate models of the yield to explore the basin of convergence and to rapidly reach the global optimum. Experimental results show that our approach locates higher quality design points in terms of yield rate within 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)**

### FLYING AND DECOUPLING CAPACITANCE OPTIMIZATION FOR AREA-CONSTRAINED ON-CHIP SWITCHED-CAPACITOR VOLTAGE REGULATORS

**Authors:** Xiaoyang Mi, Arizona State University, US

**Abstract:**

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 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)**

### ENHANCING ANALOG YIELD OPTIMIZATION FOR VARIATION-AWARE CIRCUITS SIZING

**Authors:** Ors Lahisuel, Concordia University, CA

**Abstract:**

This paper presents a novel approach for improving automated analog yield optimization using a two step exploration strategy. First, a global optimization phase relies on a modified Lipschitzian optimization to sample the potential optimal sub-regions of the feasible design space. The search locates a design point near the optimal solution that is used as a starting point by a local optimization phase. The local search constructs linear interpolating surrogate models of the yield to explore the basin of convergence and to rapidly reach the global optimum. Experimental results show that our approach locates higher quality design points in terms of yield rate within less run time and without affecting the accuracy.

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

### A NEW SAMPLING TECHNIQUE FOR MONTE CARLO-BASED STATISTICAL CIRCUIT ANALYSIS

**Authors:** Hiwa Mahmoudi, Vienna University of Technology, AT

**Abstract:**

Variability is a fundamental issue which gets exponentially worse as CMOS technology shrinks. Therefore, characterization of statistical variations has become an important part of the design phase. Monte Carlo-based simulation method is a standard technique for statistical analysis and modeling of integrated circuits. However, crude Monte Carlo sampling based on pseudorandom selection of parameter variations suffers from low convergence rates and thus, providing high accuracy is computationally expensive. In this work, we present an extensive study on the performance of two widely used techniques, Latin Hypercube and Low Discrepancy sampling methods, and compare their speed-up and accuracy performance properties. It is shown that these methods can exhibit a better efficiency as compared to the pseudorandom sampling but only in limited applications. Therefore, we propose a new sampling scheme that exploits the benefits of both methods by combining them. Through representative circuit examples, it is shown that the proposed sampling technique provides a major improvement in terms of computational effort and offers better properties as compared to each solely.

**Download Paper (PDF; Only available from the DATE venue WiFi)**
algorithm results showed comparable accuracy and better performance when compared to the flux based (iterative) and the power aware (non iterative) algorithms. Power regions then fine meshing them. Finally, the analysis is done only once. This makes it faster than conventional iterative adaptive meshing methods. The 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

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 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.

Dynamic thermal management (DTM) techniques based on task migration provide a promising solution to mitigate thermal emergencies and thereby ensuring safe operation and reliability of Many-Core systems. These techniques can be classified as central or distributed on the basis of a central DTM controller for the whole system or individual DTM controllers for each core or set of cores in the system, respectively. However, having a trustworthy comparison between central (c-) and distributed (d-) DTM techniques to find out the most suitable one for a given system is quite challenging. This is primarily due to the systemic difference between cDTM and dDTM controllers, and the inherent non-exhaustiveness of simulation and emulation methods conventionally used for DTM analysis. In this paper, we present a novel methodology called CAnDy-TM (stands for Comparative Analysis of Dynamic Thermal Management) that employs Model Checking to perform formal comparative analysis for cDTM and dDTM controllers. We identify a set of generic functional and performance properties to provide a common ground for their comparison. We demonstrate the usability and benefits of our methodology by comparing state-of-the-art cDTM and dDTM techniques, and illustrate which technique is good w.r.t. thermal stability and other task migration parameters. Such an analysis helps in selecting the most appropriate DTM for a given chip.

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 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.

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 ATNML 150 nm SOI technology (AT77K).

More information ...
TFA: TRANSPARENT CODE OFFLOADING ON FPGA

Presenter:
Roberto Rigamonti, HEIG-VD/HE5-SO, CH

Authors:
Anthony Convers, Baptiste Delporte, Xavier Ruppen and Alberto Dassatti, HEIG-VD/HE5-SO, CH

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 ...

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

Presenter:
Mario Rüüller, 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 and 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 ...

AF3-MC: DEVELOPMENT OF MIXED CRITICALITY SYSTEMS USING MBSE

Presenter:
Thomas Boehm, fortiss, DE

Authors:
Johannes Eder and Sebastian Voss, fortiss, DE

Abstract
AutoFOCUS3 (https://af3.fortiss.org/) is an open-source model-based development tool, including a number of different analysis- and verification tools as well as design space exploration functionality, task scheduling dependent on a number of system requirements (timing, resource, energy, etc.), and code generators targeting C-based VHW. The presented demonstrator illustrates both a SW tool demonstrator and a corresponding HW demonstrator setup to show how a seamless model-based system approach could look like, w.r.t. to mixed-critical applications integrated on a (COTS) MC-platform. A floating ball can be controlled by an person by moving his hand over an US sensor, providing input to the control loop implemented in the high criticality part of the system. The low criticality part of the system which is running on the same CPU consists of the computation of the digits of PI and of the Fibonacci sequence, providing computationally intensive neighbors to the control loop.

More information ...

XBARGEN: A TOOL FOR DESIGN SPACE EXPLORATION OF MEMRISTOR BASED CROSSBAR ARCHITECTURES.

Presenter:
Marcello Traiola, LIRMM, FR

Authors:
Mario Barbareschi 1 and Alberto Bosio 2
1University of Naples Federico II, IT; 2University of Montpellier - LIRMM laboratories, FR

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 memristors-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 basis of workload dependent and independent parameters, such as area, time and power consumption. The main advantage is that it does not require any simulation and thus it avoid any runtime overheads. The demo aims to show the tool prototype on a selected set of benchmarks which will be synthesized on a memristor-based crossbar circuit.

More information ...

EMU: RAPID FPGA PROTOTYPING OF NETWORK SERVICES IN C#

Presenter:
Salvator Galea, University of Cambridge, GB

Authors:
Nik Sultana 1, Pietro Bressana 2, David Greaves 2, Robert Soulié 3, Andrew W Moore 1 and Noa Zilberman 2
1University 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-x86 platform and show that they are competitive to natively written hardware counterparts while providing a superior development and debug environment.

More information ...
**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 ...**

**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, Emilico Incerto 2 and Daniele Ciabrinone 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/NI 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 ...**

**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 Fatehi 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 Koninjenburg, 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><strong>ENABLING TECHNOLOGIES FOR NEXT GENERATION BIOANALYTICS ON CHIP</strong></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 throughout 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 open 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.
plays a central role in their development.

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. The other three talks present several emerging

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

Co-Chair:
Mathias Soeken, EPFL, CH

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 plays a central role in their development.

Time | Label | Presentation Title
--- | --- | ---
11:20 | 10.1.2 | BIOELECTRONICS MEDICINES - BRIDGING BIOLOGY WITH TECHNOLOGY
 **Author:**
Firat Yazicioglu, GSK, BE

11:45 | 10.1.3 | AN OPTIMAL APPROACH FOR LOW-POWER MIGRAINE PREDICTION MODELS IN THE STATE-OF-THE-ART WIRELESS MONITORING DEVICES
 **Speaker:**
Josué Pagán, Universidad Complutense de Madrid, ES

 **Authors:**
Josué Pagán¹, Ramin Fallahzadeh², Hassan Ghasemzadeh³, Jose Manuel Moya⁴, José Luis Risco Martín ¹ and Jose L. Ayala ¹
¹Complutense University of Madrid, ES; ²School of Electrical Engineering and Computer Science, Washington State University, US; ³Washington State University, US; ⁴University Politécnica de Madrid, ES

 **Abstract**
Wearable monitoring devices for ubiquitous health care are becoming a reality that has to deal with limited battery autonomy. Several researchers focus their efforts in reducing the energy consumption of these motes: from efficient micro-architectures, to on-node data processing techniques. In this paper we focus in the optimization of the energy consumption of monitoring devices for the prediction of symptomatic events in chronic diseases in real time. To do this, we have developed an optimization methodology that incorporates information of several sources of energy consumption: the running code for prediction, and the sensors for data acquisition. As a result of our methodology, we are able to improve the energy consumption of the computing process up to 90% with a minimal impact on accuracy. The proposed optimization methodology can be applied to any prediction modeling scheme to introduce the concept of energy efficiency. In this work we test the framework using Grammatical Evolutionary algorithms in the prediction of chronic migraines.

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

12:05 | 10.1.4 | WEARABLE ELECTRONICS - WHAT IS IT GOOD FOR - AND WHAT IS MISSING TO SUPPORT THE QUALITY OF LIFE OF ELDERLY PEOPLE?
 **Author:**
Ralf Brederlow, Kilby Labs at Texas Instruments, DE

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.

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 plays a central role in their development.

Time | Label | Presentation Title
--- | --- | ---
11:00 | 10.2.1 | LOGIC OPTIMIZATION AND SYNTHESIS: TRENDS AND DIRECTIONS IN INDUSTRY
 **Speaker:**
Luca Amaru, Synopsys Inc., US

 **Authors:**
Luca Amaru, Patrick Vuillod, Jiong Luo and Janet Olson, Synopsys, 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 syn- ergy 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 implementa- tion. 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.

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

11:22 | 10.2.2 | CARBON NANOTUBES ENABLE MAJOR ENERGY EFFICIENCY BENEFITS FOR SUB-10NM DIGITAL SYSTEMS
 **Speaker:**
Gage Hills, Stanford University, US

 **Authors:**
Gage Hills¹, Max Shulaker², Chi-Shuen Lee³, Peter Debacker⁴, Marie Garcia Bardon ³, Dmitry Yakimets ⁵, Romain Ritzenthaler ⁵, Iuliana Radu ⁵, Francky Cathoor⁶, Praveen Raghavan⁷, Aaron Thean⁷, H.-S. Philip Wong ³ and Subhashis Mitra ³
¹Department of Electrical Engineering, Stanford University, US; ²MIT, US; ³Stanford University, US; ⁴imec vzw, BE; ⁵IMEC, BE; ⁶imec, BE; ⁷NUSingapore, SG
11:45 10.3.1 SIDE-CHANNEL PLAINTEXT-RECOVERY ATTACKS ON LEAKAGE-RESILIENT ENCRYPTION

**Speaker:** Thomas Unterluggauer, Graz University of Technology, AT

**Authors:** 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 encryption/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.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
</table>
| 10:40 | 10.4.1 | **Almost-Always Coherence:** Myths and Reality | Thorben Moos, Horst Göritz Institute for IT-Security, Ruhr-Universität Bochum, DE  
Authors: Thorben Moos, Horst Göritz Institute for IT-Security, Ruhr-Universität Bochum, DE |
| 11:00 | 10.4.2 | **SIDE-CHANNEL POWER ANALYSIS OF XTS-AES** | Chao Luo, Yunsi Fei and A. Adam Ding, Northeastern University, US  
Authors: Chao Luo, Yunsi Fei and A. Adam Ding, Northeastern University, US |
| 12:30 | 10.4.3 | **FORMAL MODEL FOR SYSTEM-LEVEL POWER MANAGEMENT DESIGN** | Mirela Simonovic, Aggius, RS  
Authors: Mirela Simonovic, Aggius, RS |
| 13:30 | 10.4.4 | **SYNTHESIS OF HARDWARE DEPENDENT PATTERNS** | Mirela Simonovic, Vojia Zivojnovic and Lazar Saranoc  
Authors: Mirela Simonovic, Vojia Zivojnovic and Lazar Saranoc |

---

**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 Emerging Technologies for Reconfigurable Architectures

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
</table>
| 10:30 | 10.3.1 | **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 Moos, Horst Göritz Institute for IT-Security, Ruhr-Universität Bochum, DE |
| 11:30 | 10.3.2 | **SIDE-CHANNEL POWER ANALYSIS OF XTS-AES** | Chao Luo, Yunsi Fei and A. Adam Ding, Northeastern University, US  
Authors: Chao Luo, Yunsi Fei and A. Adam Ding, Northeastern University, US |
| 12:30 | 10.3.3 | **FORMAL MODEL FOR SYSTEM-LEVEL POWER MANAGEMENT DESIGN** | Mirela Simonovic, Aggius, RS  
Authors: Mirela Simonovic, Aggius, RS |
| 13:30 | 10.3.4 | **SYNTHESIS OF HARDWARE DEPENDENT PATTERNS** | Mirela Simonovic, Vojia Zivojnovic and Lazar Saranoc  
Authors: Mirela Simonovic, Vojia Zivojnovic and Lazar Saranoc |

---

**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.

---

**Download Paper (PDF; Only available from the DATE venue WiFi)**
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.
10.6 Approximate computing and neural networks for novel communication and multimedia systems

Date: Thursday 30 March 2017
Time: 10:00 - 12:30
Location: 5A

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.

11:00 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 references prove our work to be a powerful approach for future mobile communication systems.

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

11:30 10.6.2 **EMBRACING APPROXIMATE COMPUTING FOR ENERGY-EFFICIENT MOTION ESTIMATION IN HIGH EFFICIENCY VIDEO CODING**

**Speaker:** Muhammad Shafique, Vienna University of Technology (TU Wien), AT  
**Authors:** Walaa El-Harouni², Semeen Rehman¹, Bharath Srinivas Prabakaran², Akash Kumar³, Rehan Hafiz² and Muhammad Shafique³  
¹Private Researcher, DE; ²Technische Universität Dresden, DE; ³Technische Universitäet Dresden, DE; ⁴TU, 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 comparisons 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/paclib/.

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

12:00 10.6.3 **HARDWARE ARCHITECTURE OF BIDIRECTIONAL LONG SHORT-TERM MEMORY NEURAL NETWORK FOR OPTICAL CHARACTER RECOGNITION**

**Speaker:** Vladimir Rybalkin, University of Kaiserslautern, DE  
**Authors:** Vladimir Rybalkin¹, Mohammad Reza Yousefi², Norbert Weinl¹ and Didier Stricker³  
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 learning, 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)

12:30 436 **EXTENDING MEMORY CAPACITY OF NEURAL ASSOCIATIVE MEMORY BASED ON RECURSIVE SYNAPTIC BIT REUSE**

**Speaker:** Tianchan Guan, Columbia University, US  
**Authors:** Tianchan Guan¹, Xiaoyang Zeng² and Mingoo Seok²  
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 humans 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 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 5k 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.

**Date:** Thursday 30 March 2017
**Time:** 11:00 - 12:30
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.

**Abstract**

Although Deep Neural Networks (DNN) are ubiquitously utilized in many applications, it is generally difficult to deploy DNNs on resource-constrained devices, e.g., mobile platforms. Some existing attempts mainly focus on client-server computing paradigm or DNN model compression, which require either infrastructure support or special training phases, respectively. In this work, we propose MoDNN - a local distributed mobile computing system for DNN applications. MoDNN can partition DNN models onto several mobile devices to accelerate DNN computations by alleviating device-level computing cost 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 parallelized 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.

Download Paper (PDF; Only available from the DATE venue WiFi)
10.8a Smart and Wearable Sensors for Health

Date: Thursday 30 March 2017
Time: 11.00 - 12.00
Location / Room: Exhibition Theatre
Organiser: Patrick Mayor, EPFL, CH
Moderator: Martin Rajman, EPFL, CH

The goal of this session is to present three concrete examples of innovative wearable devices: a contactless monitoring system using dedicated imaging to accurately measure heart and respiratory rates of neonates, wearable devices integrated in smart textiles for the long-term monitoring of obese patients, as well as a prototype of next-generation, high-quality, mobile ultrasound imaging device.

Time | Label | Presentation Title | Authors
--- | --- | --- | ---
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 corresponded 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 is 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, Shangping Ren2, Yu Jiang2 and Liu 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 cyberphysical 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 models 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.

10.8a Smart and Wearable Sensors for Health

Date: Thursday 30 March 2017
Time: 11.00 - 12.00
Location / Room: Exhibition Theatre
Organiser: Patrick Mayor, EPFL, CH
Moderator: Martin Rajman, EPFL, CH

The goal of this session is to present three concrete examples of innovative wearable devices: a contactless monitoring system using dedicated imaging to accurately measure heart and respiratory rates of neonates, wearable devices integrated in smart textiles for the long-term monitoring of obese patients, as well as a prototype of next-generation, high-quality, mobile ultrasound imaging device.
11:00  10.8a.1   NEWBORNCARE
Speaker: Martin Wolf, USZ, CH

11:20  10.8a.2   OBESENSE
Speaker: Jean-Philippe Thiran, EPFL, CH

11:40  10.8a.3   ULTRASOUNDTOGO
Speaker: Federico Angiolini, EPFL, CH

12:00  End of session
12:30  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.

10.8b IoT Edge Devices

Date: Thursday 30 March 2017
Time: 12:00 - 12:30
Location / Room: Exhibition Theatre

12:00  10.8b.1   MENTOR'S CUSTOM/ANALOG SOLUTIONS FOR IOT EDGE DEVICES
Speaker: Nicolas Williams, Mentor, US

12:30  End of session
12:30  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

UB10.1 A FRAMEWORK FOR VARIATION-AWARE ANALOG CIRCUITS SIZING
Presenter: Ons Lahiouel, Concordia University, CA
Authors: Mohamed H. Zaki and Sofiene Tahar, Concordia University, CA
Abstract
Today's analog design faces significant challenges due to circuit complexity and short time-to-market windows. The proposed demonstration presents new techniques for enhancing variation-aware circuits sizing. The sizing problem is encoded using nonlinear constraints. A new algorithm using Satisfiability Modulo Theory (SMT) solving techniques exhaustively explores the design space and computes a continuous set of feasible sizing solutions. Two methods for the computation of parametric yield are implemented. The first method combines the advantages of sparse regression and SMT solving techniques for reliable and accelerated yield estimation. The second approach employs a statistical classifier to reduce the number of simulations. An optimization process using a two-step exploration strategy is also integrated to find the feasible design point with the highest yield. Experimental results show that our approach locates higher quality of design point within less run time.

More information ...

UB10.2 TFA: TRANSPARENT CODE OFFLOADING ON FPGA
Presenter: Roberto Rigamonti, HEIG-VD/HES-SO, CH
Authors: Anthony Convers, Baptiste Delporte, Xavier Ruppen and Alberto Dassatti, HEIG-VD/HES-SO, CH
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 ...
TOOLS: MODEL-BASED DESIGN OF A 5G UPLINK DATA-LINK LAYER RECEIVER FROM UML/SYML DIAGRAMS

Presenter: Andrea Enrici, Nokia Bell Labs, FR
Authors: Julien Lallet, Imran Latif, Ludovic Appvrire, Renaud Paccard and Adrien Canuel

Abstract
Future 5G networks are expected to provide an increase of 10x in data rates. To meet these requirements, the hardware design 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 system 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/SyML 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 ...

AF3-MC: DEVELOPMENT OF MIXED CRITICALITY SYSTEMS USING MSBE

Presenter: Thomas Boehm, fortiss, DE
Authors: Johannes Eder and Sebastian Voss, fortiss, DE

Abstract
AF3MC (https://af3.fortiss.org/) is an open-source model-based development tool, including a number of different analysis- and verification tools as well as design space exploration functionality, task scheduling dependent on a number of system requirements (timing, resource, energy, etc.), and code generators targeting C-code or VHDL. The presented demonstrator illustrates both a SW tool demonstrator and a corresponding HW demonstrator setup to show how a seamless model-based system approach could look like, w.r.t. to mixed-critical applications integrated on a (COTS) MC-platform. A floating ball can be controlled by a person by moving his hand over an US sensor, providing input to the control loop implemented in the high criticality part of the system. The low criticality part of the system which is running on the same CPU consists of the computation of the digits of Pi and of the Fibonacci sequence, providing computationally intensive neighbors to the control loop.

More information ...

STACKADROP: A MODULAR DIGITAL MICROFLUIDIC BIOCHIP RESEARCH PLATFORM

Presenter: Oliver Keszlbece, University of Bremen, DE
Authors: Maximilian Luennert 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 Wheelerc Laboratories and the OpenDrop from Gaudi Labs. 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 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 ...

MARGOT: APPLICATION ADAPTATION THROUGH RUNTIME AUTOTUNING

Presenter: Gianluca Palermo, Politecnico di Milano, IT
Authors: Davide Gadioli, Emanuele Vitali and Cristina Silvano, Politecnico di Milano, IT

Abstract
Several classes of applications expose parameters that influence their extra-functional properties, such as the quality of the result or the performance. This leads the application designer to tune these parameters to find the configuration that produces the desired outcome. Given that the application requirements and the resources assigned to each application might vary at runtime, finding a one-fit-all configuration is not a trivial task. For this reason, we implemented the mARGOt framework that enhances an application with an adaptation layer in order to continuously tune the parameters according to the evolving situation. More in detail, mARGOt is composed of a monitoring infrastructure, an application-level adaptation engine and an extra-functional configuration framework based on the separation of concerns paradigm between functional and extra-functional aspects. At the booth, we plan to demonstrate the effectiveness of the proposed infrastructure on three real-life applications.

More information ...

EMU: RAPID FPGA PROTOTYPING OF NETWORK SERVICES IN C#

Presenter: Salvatore Galea, University of Cambridge, GB
Authors: Nik Sultanii, Pietro Bressana, David Greaves, Robert Soulé, Andrew W Moore and Noa Zilberman
2University of Cambridge, GB; 3Università 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 ...

TIDES: NON-LINEAR WAVEFORMS FOR QUICK TRACE NAVIGATION

Presenter: Janne Stoppel, University of Bremen, DE
Author: Rolf 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 ...
As technology slows down, matures, and the industry consolidates, we are presented with opportunities for applying our talents for the analysis, modeling, optimization and solution of difficult large scale problems in adjacent fields. This talk is about one such opportunity in the area of radiation therapy, where Medical Physicists work hand-in-hand with Oncologists to provide life-saving treatments for Cancer. Making the transition from EDA to Medicine required some significant sacrifices and humility—but the end result is a commercial and scientific success and a far greater level of relevance to people’s lives.

**More information ...**
DESIGN CHALLENGES FOR WEARABLE EMG APPLICATIONS

Speaker: Josué A. Ayala, Universidad Complutense de Madrid, ES

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 proliferation, the availability of new low-power sensors, and the development of on-chip signal processing capabilities with a contained power consumption. 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)
11.2 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 (Best Paper Award Candidate)
HYBRID VC-MTJ/CMOS NON-VOLATILE STOCHASTIC LOGIC FOR EFFICIENT COMPUTING

Speaker:
Shaodi Wang, University of California, Los Angeles, US

Authors:
Shaodi Wang, Saptadeep Pal, Tianmu Li, Andrew Pan, Cecile Grezes, Pedram Khalili-Amiri, Kang L. Wang and Puneet Gupta

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)
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)

VAET-STT: A VARIATION AWARE ESTIMATOR TOOL FOR STT-MRAM BASED MEMORIES

Speaker: Sarath Mohanachandran Nair, KIT, Germany, DE
Authors: Sarath Mohanachandran Nair 1, Rajendra Bishnoi 2, Mohammad Saber Golanbari 1, Fabian Oborn 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)

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 2, Guanyu 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)

EFFECTS OF CELL SHAPES ON THE ROUTABILITY OF DIGITAL MICROFLUIDIC BIOCHIPS

Speaker: Oliver Kes佐zce, University of Bremen, DE
Authors: Kevin Leonard Schneider 1, Oliver Kes佐zce 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 other's 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)
reconfigurable array architectures targeting emerging applications such as Hadoop map reduce framework and neural networks. This session introduces new approaches for building reconfigurable accelerators and heterogeneous architectures integrating big-little cores, FPGA hardware, and coarse grain.

Houman Homayoun, George Mason University, US

Co-Chair:

Georgios Keramidas, Think Silicon S.A./Technological Educational Institute of Western Greece, GR

Chair:

Location / Room: 3BC

Time: 11:00 - 12:30

11.3 Exploiting Heterogeneity for Big Data Computing

Date: Thursday 30 March 2017

Time: 14:00 - 15:30

Location / Room: 3BC

Chair:

Georgios Keramidas, Think Silicon S.A./Technological Educational Institute of Western Greece, GR

Co-Chair:

Houman Homayoun, George Mason University, US

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:

Sherief Reda, Brown University, US

Speaker:

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

11.3.1 A NOVEL ZERO WEIGHT/ACTIVATION-AWARE HARDWARE ARCHITECTURE OF CONVOLUTIONAL NEURAL NETWORK

Speaker:

Dongyoung Kim, Seoul National University, KR

Authors:

Dongyoung Kim, Junwhan Ahn and Sungjoo Yoo, Seoul National University, KR

Abstract

It is imperative to accelerate convolutional neural networks (CNNs) due to their ever-widening application areas from server, mobile to IoT devices. Based on the fact that CNNs can be characterized by significant amount of zero values in both kernel weights (under quality-preserving pruning) and activations (when rectified linear units are applied), we propose a novel architecture of hardware accelerator for CNNs which exploits zero values in both weights and activations. We also report a zero-induced load imbalance problem encountered in the zero-aware parallel architecture and present a zero-aware kernel allocation. In our experiments, we designed a cycle-accurate model, RTL and layout designs of the proposed architecture. In our evaluations with two real deep CNNs, pruned AlexNet and VGG, our proposed architecture offers 4×/1.8× times (AlexNet [1]) and 5.2×/2.1× times (VGG-16 [2]) speedup compared with state-of-the-art zero-agnostic/zero activation-aware architectures.

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

11.3.2 A MECHANISM FOR ENERGY-EFFICIENT REUSE OF DECODING AND SCHEDULING OF X86 INSTRUCTION STREAMS

Speaker:

Antonio Carlos S. Beck, Universidade Federal do Rio Grande do Sul, BR

Authors:

Marcelo Brandalero and Antonio Carlos Schneider Beck, Universidade Federal do Rio Grande do Sul, BR

Abstract

Current superscalar x86 processors decompose each CISC instruction (variable-length and with multiple addressing modes) into multiple RISC-like µops at runtime so they can be pipelined and scheduled for concurrent execution. This challenging and power-hungry process, however, is usually repeated several times on the same instruction sequence, inefficiently producing the very same decoded and scheduled µops. Therefore, we propose a transparent mechanism to save the decoding and scheduling transformation for later reuse, so that next time the same instruction sequence is found it can automatically bypass the costly pipeline stages involved. We use a coarse-grained reconfigurable array as a means to save this transformation, since its structure enables the recovery of µops already allocated in time and space, and also larger ILP exploitation than superscalar processors. The technique can reduce the energy consumption of a powerful 8-issue superscalar by 31.4% at low area costs, while also improving performance by 32.6%.

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

11.3.3 UNDERSTANDING THE IMPACT OF PRECISION QUANTIZATION ON THE ACCURACY AND ENERGY OF NEURAL NETWORKS

Speaker:

Soheil Hashemi, Nicholas Anthony, Hokchhay Tann, Iris Bahar and Sherief Reda, Brown University, US

Authors:

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)
15:30 IPS-9, 763
LESS: BIG DATA SKETCHING AND ENCRYPTION ON LOW POWER PLATFORM
Speaker:
Amey Kulkarni, University of Maryland Baltimore County, US
Authors:
Amey Kulkarni1, Colin Shea2, Houman Homayoun3 and Tinoosh Mohsenin2
1University of Maryland, Baltimore County, US; 2University of Maryland Baltimore County, US; 3George 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 (OTRLP). OTRLP 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 (SRER). 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)

15:31 IPS-10, 656
TRUNCAPP: A TRUNCATION-BASED APPROXIMATE DIVIDER FOR ENERGY EFFICIENT DSP APPLICATIONS
Speaker:
Shaghayegh Vahdat, University of Tehran, IR
Authors:
Shaghayegh Vahdat1, Mehdi Kamal1, Ali Afzali-Kusha1, Zainalabedin Navabi1 and Massoud Pedram2
1University of Tehran, IR; 2University 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 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.
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
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>
<th>Speaker</th>
</tr>
</thead>
<tbody>
<tr>
<td>14:00</td>
<td>11.4.1</td>
<td>QUANTIFYING ERROR: EXTENDING STATIC TIMING ANALYSIS WITH PROBABILISTIC TRANSITIONS</td>
<td>Kevin Murray, University of Toronto, CA</td>
<td>Kevin Murray1, Andrea Suardi2, Vaughn Betz1 and George Constantinides2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1University of Toronto, CA; 2Imperial College, GB</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abstract</td>
<td>Timing analysis is a cornerstone of the digital design process. Statistical Static Timing Analysis was introduced to reduce pessimism by modelling 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 #SAT 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.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Download Paper (PDF; Only available from the DATE venue WiFi)</td>
<td></td>
<td></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</td>
<td>Ye-Hong Chen, Sheng-He Wang and Ting-Chi Wang, National Tsing Hua University, TW</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abstract</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 first 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.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Download Paper (PDF; Only available from the DATE venue WiFi)</td>
<td></td>
<td></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</td>
<td>Wachirawit Ponghiran1, Seongbo Shim2 and Youngsoo Shin3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Authors</td>
<td>1School of Electrical Engineering, KAIST, KR; 2Dept. of Electrical Engineering, KAIST, KR; 3KAIST, KR</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abstract</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.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Download Paper (PDF; Only available from the DATE venue WiFi)</td>
<td></td>
<td></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</td>
<td>Taesik Na, Jong Hwan Ko and Saibal Mukhopadhyay, Georgia Institute of Technology, US</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Authors</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abstract</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.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Download Paper (PDF; Only available from the DATE venue WiFi)</td>
<td></td>
<td></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</td>
<td>Youngsoo Song, Sangmin Kim and Youngsoo Shin, School of Electrical Engineering, KAIST, KR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Authors</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abstract</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.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Download Paper (PDF; Only available from the DATE venue WiFi)</td>
<td></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

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 Höttger, Dortmund University of Applied Sciences and Arts, DE
Authors: Robert Höttger¹, Olaf Spinczyk² and Burkhard Igel¹
¹FH-Dortmund, DE; ²TU-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>
</table>
| 14:30 | 11.5.2 | POWER NEUTRAL PERFORMANCE SCALING FOR ENERGY HARVESTING MP-SOCS | Benjamin Fletcher, University of Southampton, GB
Authors: Benjamin Fletcher, Domenico Balsamo and Geoff Merrett, University of Southampton, GB
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)
15:00 11.5.3 EFFICIENT DECENTRALIZED ACTIVE BALANCING STRATEGY FOR SMART BATTERY CELLS
Speaker: Nitin Shivaraman, Nanyang Technological University, SG
Authors: Nitin Shivaraman, Avind 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)

15:00 11.5.4 WULORA: AN ENERGY EFFICIENT IOT END-NODE FOR ENERGY HARVESTING AND HETEROGENEOUS COMMUNICATION
Speaker: Michele Magno, ETH Zurich, CH
Authors: Michele Magno, Faysal Al 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 wideareas deployments. Therefore energy efficiency is of 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 short-range 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)

15:30 IPS-12, 84 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 Glaß2, Felix Reimann1 and Jürgen Teich1
1Friedrich-Alexander-Universität Erlangen-Nürnberg, DE; 2TUM 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 mixedcriticality 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-calulation by several orders of magnitude without introducing any pessimism.
Download Paper (PDF; Only available from the DATE venue WiFi)

15:30 IPS-13, 368 ULTRA LOW-POWER VISUAL ODOMETRY FOR NANO-SCALE UNMANNED AERIAL VEHICLES
Speaker: Daniele Falossi, ETH Zurich, CH
Authors: Daniele Falossi, 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 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-capabilities 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)

15:32 IPS-14, 598 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)
The recent ability of some vehicles to handle autonomously 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.

11.6 Dependable microprocessors and systems

**Date:** Thursday, March 30, 2017
**Time:** 14:00 - 15:30

**Chair:**
Maksim Jenihhin, Tallinn University of Technology, EE

**Co-Chair:**
Antonio Miele, Politecnico di Milano, IT

The section presents two papers investigating the effects of soft errors on critical registers and hardware methods to detect intrusion attacks in microprocessors. A third paper provides a solution for estimating multiprocessor expected lifetime.

### 11.6.1 CHARACTERIZATION OF STACK BEHAVIOR UNDER SOFT ERRORS

**Authors:**
Junchi Ma, School of Computer Science and Engineering, Southeast University, CN

**Abstract**
As process technology scales, electronic devices become more susceptible to soft error induced by radiation. The stack in the memory implements procedure calls and its behavior under soft error has not been studied yet. To analyze the effects of soft error on the stack behavior, we conduct a series of fault injection experiment in the IA-32 instruction set architecture. The injection targets are the ESP register (used as the stack pointer) and the EBP register (used as the stack-frame base pointer). We obtain a few important observations from the fault injection experiment. Results show that injections on ESP lead to silent data corruption (SDC) or benign only if the flipped ESP points to another return address when executing the RET instruction, otherwise most of the injections cause crash. The injected bits of these SDC and benign cases are distributed in the particular bits (4-7) and the reason for the distribution is given. Moreover, flipped EBP may cause a series of infinite return operations, which is defined as return cycle. We describe the basic mechanism of return cycle and the essential condition for its occurrence.

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

### 11.6.2 MULTI-ARMED BANDITS FOR EFFICIENT LIFETIME ESTIMATION IN MPSOC DESIGN

**Authors:**
Calvin Ma, Aditya Mahajan and Brett Meyer, McGill University, CA

**Abstract**
Reliability in integrated circuits is becoming a critical issue with the miniaturization of electronics. Smaller process technologies have led to higher power densities, resulting in higher temperatures and earlier device wear-out. One way to mitigate failure is by over-provisioning resources and remapping tasks from failed components to components with spare capacity, or slack. Since the slack allocation design space is large, finding the optimal is difficult, as brute-force approaches are impractical. During design space exploration, device lifetimes are typically evaluated using Monte-Carlo Simulation (MCS) by sampling each design equally; this method is inefficient since poor designs are evaluated as accurately as good designs. A better method will focus sampling time on the designs that are difficult to distinguish, reducing the time required to evaluate a set of designs; this can be accomplished using Multi-armed Bandit (MAB) Algorithms. This work demonstrates that MAB achieve the same level of accuracy as MCS in 1.45 to 5.26 times fewer samples.

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  IP5-16 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)

15:31  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)

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

11.7 Formal Methods and Verification: Core Technologies and Applications
Date: Thursday 30 March 2017
Time: 14:00 - 15:30
Location / Room: 3B
Chair: Barbara Jobstmann, EPFL / Cadence, CH
Co-Chair: Christoph Scholl, University of Freiburg, DE

The session consists of three papers on formal verification and its applications. The first paper presents the use of grammar-based techniques for the analysis of high-end processor designs at the netlist level. The second paper considers a computer algebra-based technique to reverse engineer the irreducible polynomial used in the implementation of multipliers in finite fields. The third paper applies probabilistic model checking in a case study analyzing the dependability of optical communication networks with double-ring topologies (which have been proposed for multicast traffic in metropolitan areas).
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 witnessed a significant development and progress in these fields. The goal of this Special Session is to present latest research results from worldwide leading experts addressing inspired techniques such as evolutionary algorithms and artificial neural networks have been used in the mainstream circuit design community infrequently. Recent years have exchange for extremely low power consumption. However, even when this sacrifice is possible, other conflicting performance features must still be taken into account. Biologically-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

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 such as evolutionary algorithms and artificial neural networks have been used in the mainstream circuit design community infrequently. Recent years have witnessed a significant development and progress in these fields. The goal of this Special Session is to present latest research results from worldwide leading experts addressing 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.
AN EVOLUTIONARY APPROACH TO RUNTIME VARIABILITY MAPPING AND MITIGATION ON A MULTI-RECONFIGURABLE ARCHITECTURE
Speaker: Simon Bale, University of York, GB
Authors: Simon Bale, Pedro Campos, Martin Albrecht Trefzer, James Walker and Andy Tyrell, University of York, GB
Abstract
Intrinsic device variability has become a significant problem in deep sub-micron technology nodes. The stochastic variations in device performance, which are a result of structural irregularities at the atomic scale, can impact both the yield and reliability of a circuit design. In this paper we describe a novel multi-reconfigurable FPGA architecture, the programmable analogue and digital array (PANDA), which can tackle this problem by allowing post-fabrication reconfiguration of the effective transistor gate widths in a circuit. We demonstrate the advantages of this architecture by creating a frequency variability map of the array using ring oscillators in order to ascertain the location of any frequency outliers. We then show that it is possible, using an evolutionary algorithm, to select alternative transistor configurations which minimise the difference in frequency between one of these outliers and the chips median frequency of operation. Such methods can be used to increase system performance and reliability by presenting an array with more uniform performance characteristics.
Download Paper (PDF; Only available from the DATE venue WiFi)

TOWARDS LOW POWER APPROXIMATE DCT ARCHITECTURE FOR HEVC STANDARD
Speaker: Zdenek Vasiczek, Brno University of Technology, CZ
Authors: Zdenek Vasiczek, Vojtech Mrazek and Lukas Sekanina, Brno University of Technology, CZ
Abstract
Video processing performed directly on IoT nodes is one of the most performance as well as energy demanding applications for current IoT technology. In order to support real-time high-definition video, energy-reduction optimizations have to be introduced at all levels of the video processing chain. This paper deals with an efficient implementation of Discrete Cosine Transform (DCT) blocks employed in video compression based on the High Efficiency Video Coding (HEVC) standard. The proposed multiplierless 4-input DCT implementations contain approximate adders and subtrators that were obtained using genetic programming. In order to manage the complexity of evolutionary approximation and provide formal guarantees in terms of errors of key circuit components, the worst and average errors were determined exactly by means of Binary decision diagrams. Under conditions of our experiments, approximate 4-input DCTs show better quality/power trade-offs than relevant implementations available in the literature. For example, 25% power reduction for the same error was obtained in comparison with a recent highly optimized implementation.
Download Paper (PDF; Only available from the DATE venue WiFi)

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, thereby, 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)

MACHINE LEARNING FOR RUN-TIME ENERGY OPTIMISATION IN MANY-CORE SYSTEMS
Speaker: Rishad Shaftik, Newcastle University, GB
Authors: Dwapayan Biswas, Vibsarna Balagopal, Rishad Shaftik, Bashir Al-Hashimi and Geoff Merrett
University of Southampton, GB; Newcastle 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)

AN EVOLUTIONARY APPROACH TO HARDWARE ENCRYPTION AND TROJAN-HORSE MITIGATION
Speaker: Ernesto Sanchez, Politecnico di Torino, IT
Authors: Andrea Marcelli, Marco Restifo, Ernesto Sanchez and Giovanni Squillero, Politecnico di Torino, IT
Abstract
New threats, grouped under the name of hardware attacks, became a serious concern in recent years. In a global market, untrusted parties in the supply chain may jeopardize the production of integrated circuits with intellectual-property piracy, illegal overproduction and hardware Trojan-horses (HT) injection. While one way to protect from overproduction is to encrypt the design by inserting logic gates that prevents the circuit from generating the correct outputs unless the right key is used, reducing the number of poorly-controllable signals is known to minimize the chances for an attacker to successfully hide the trigger for some malicious payload. Several approaches successfully tackled independently these two issues. This paper proposes a novel technique based on a multi-objective evolutionary algorithm able to increase hardware security by explicitly targeting both the minimization of rare signals and the maximization of the efficacy of logic encryption. Experimental results demonstrate the proposed method is effective in creating a secure encryption schema for all the circuits under test and in reducing the number rare signals on six circuits over nine, outperforming the current state of the art.
Download Paper (PDF; Only available from the DATE venue WiFi)
## Abstract

Tohru Ishihara and Hidetoshi Onodera, Kyoto University, JP

### UB11.5

**A VOLTAGE-SCALEABLE 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

### UB11.4

**AF3-MC: DEVELOPMENT OF MIXED CRITICALITY SYSTEMS USING MBSE**

**Presenter:** Thomas Boehm, fortiss, DE

**Authors:** Johannes Eder and Sebastian Voss, fortiss, DE

**Abstract**

AutoFOCUS3 ([https://af3.fortiss.org/](https://af3.fortiss.org/)) is an open-source model-based development tool, including a number of different analysis- and verification tools as well as design space exploration functionality, task scheduling dependent on a number of system requirements (timing, resource, energy, etc.), and code generators targeting C-code or VHDL. The presented demonstrator illustrates both a SW tool demonstrator and a corresponding HW demonstrator setup to show how a seamless model-based system approach could look like, w.r.t. to mixed-critical applications integrated on a (COTS) MC-platform. A floating ball can be controlled by an person by moving his hand over an US sensor, providing input to the control loop implemented in the high criticality part of the system. The low criticality part of the system can be controlled by a computer running on the same CPU consists of the computation of the digits of Pi and of the Fibonacci sequence, providing computationally intensive neighbors to the control loop.

### UB11.2

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

**Presenter:** Alexandre Coelho, Université Grenoble Alpes, FR

**Authors:** Miguel Solinas, Juan Fraire, Nacer Eddine Zergainoh and Raoul Velazco, Univ. Grenoble Alpes, CNRS, Grenoble INP, FR

**Abstract**

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.

### UB11.1

**TGV: TESTER GENERIC AND VERSATILE FOR RADIATION EFFECTS ON ADVANCED VLSI CIRCUITS**

**Presenter:** Miguel Solinas, TIMA, FR

**Authors:** Alexandre Coelho Coelho, Juan Fraire, Nacer Eddine Zergainoh, Pablo Ferreyra and Raoul Velazco, TIMA, FR

**Abstract**

The purpose of this work is to describe a novel tester for radiation effects experiments, called TGV (Tester Generic and Versatile) based on a commercial development board ZEDBOARD. The main idea is to implement the whole DUT (Device Under Test) board architecture controlled by an FPGA, whose configuration is obtained from compiling the description of key features of the DUT in a high-level language such as C. This tester constitutes a powerful tool with generic capabilities for the functional validation and test under radiation of any digital circuit, with a particular focus on processor-like circuits. In this way, there is only a minor hardware development, limited to wiring the DUT pins to the ones of the tester connector. During the demonstration will be shown details of TGV platform, its use being illustrated be means of fault injection experiments which reproduces in a realistic way the random occurrence in time and location of SEUs in sensitive targets of the considered circuit.

---

### Table: UB11 Session 11

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>15:30</td>
<td>End of session</td>
<td>Coffee Break in Exhibition Area</td>
</tr>
<tr>
<td>15:30</td>
<td>Coffee Break</td>
<td>10:00 - 11:30</td>
</tr>
<tr>
<td>16:00</td>
<td>Coffee Break</td>
<td>10:00 - 11:00</td>
</tr>
<tr>
<td>17:00</td>
<td>Coffee Break</td>
<td>10:00 - 11:00</td>
</tr>
</tbody>
</table>

---

**Authors**

**Presenters**

**Location / Room:** Booth 1, Exhibition Area

**Date:** Thursday, March 30, 2017

**Time:** 14:30 - 16:30

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.

When: March 28, 2017

- Coffee Break 10:00 - 11:30
- Coffee Break 16:00 - 17:00

When: March 29, 2017

- Coffee Break 10:00 - 11:00
- Coffee Break 16:00 - 17:00

When: March 30, 2017

- Coffee Break 10:00 - 11:00
- Coffee Break 15:30 - 16:00

---

**Coffee Break**

15:30 - 16:00

**End of session**

---

**More information ...**

---

**Coffee Break**

10:00 - 11:00

**End of session**

---

**Coffee Break**

10:30 - 11:30

**End of session**

---

**Coffee Break**

16:00 - 17:00
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.

More information ...

HEPSYCODE: A SYSTEM-LEVEL METHODOLOGY FOR HW/SW CO-DESIGN OF HETEROGENEOUS PARALLEL DEDICATED SYSTEMS

Giacomo Valente, University of L'Aquila, IT
Luigi Pomante, University of L'Aquila, IT
Presenter:
Pietro Bressana, David Greaves, Robert Soulé, Andrew W Moore, Noa Zilberman
Authors:

Abstract

HEPSYCODE is a system-level methodology for the co-design of heterogeneous parallel dedicated systems. It supports modularity, scalability, and abstraction, and is able to automatically generate hardware/ software co-optimized and correctly-by-construction designs. With HEPSYCODE, architects and designers are enabled to rapidly explore and evaluate new ideas, and to quantify the design space before the actual implementation.

More information ...

EMU: RAPID FPGA PROTOTYPING OF NETWORK SERVICES IN C#

Salvator Galea, University of Cambridge, GB
Authors:

Abstract

EMU is a framework that 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 ...

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

Amir CHARIF, TIMA, FR
Mirela Simonovic, Aggios, RS
Authors:

Abstract

GNOCS is a cycle-accurate, high-speed, and highly extensible NoC simulator that is able to simulate thousands of nodes in reasonable times. In this paper, 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 speedsups with no loss of accuracy. To enable quick and easy validation of new 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 ...

FORMAL MODEL FOR SYSTEM-LEVEL POWER MANAGEMENT DESIGN

Mirela Simonovic, Vojin Zivojnovic, Lazar Saranovac
Authors:

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)
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.

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 engines 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 system.

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, Xiyan Zhou, Guangyu Sun, and Jiwu Shu

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, and fast access speed. However, these NVMs suffer from the write endurance problem. In order to prevent a malicious program from erasing 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 Keszöcze, University of Bremen, DE
Authors: Kevin Leonard Schneider, Oliver Keszöcze, Jannis Stopp, and Rolf Drechsler

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)

IP5-9

LESS: BIG DATA SKETCHING AND ENCRYPTION ON LOW POWER PLATFORM

Speaker: Ameyulkarni, University of Maryland Baltimore County, US
Authors: Ameyulkarni, Colin Shea, Houman Homayoun and Tinoosh Mohsenin

Abstract
Every-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 3.81-4dB 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 17x 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 sketching. The full hardware integration results show that the LESS framework achieves 45% 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 Afzali-Kusha, Zainalabedin Navabi and Massoud Pedram

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, 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)
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 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 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 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 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 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 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 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 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 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 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 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 forms with simpler addressing methods and a tech...
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)
### 12.2.2 INFORMATION FLOW TRACKING ON RTL DESIGNS

**Title:** COSYN: EFFICIENT SINGLE-CELL ANALYSIS USING A HYBRID MICROFLUIDIC PLATFORM

**Authors:**
- Mohamed Ibrahim, Duke University, US
- Ryan Kastner, University of California, San Diego, US
- Armaiti Ardeshiricham, UCSD, US
- Ulf Schlichtmann, UCSD, US

**Abstract:**
Single-cell genomics is used to advance our understanding of diseases such as cancer. Microfluidic solutions have recently been developed to classify cell types or perform single-cell biochemical analysis on pre-isolated types of cells. However, new techniques are needed to efficiently classify cells and conduct biochemical experiments on multiple cell types concurrently. System integration and design automation are major challenges in this context. To overcome these challenges, we present a hybrid microfluidic platform that enables complete single-cell analysis on a heterogeneous pool of cells. We combine this architecture with an associated design-automation and optimization framework, referred to as Co-Synthesis (CoSyn). The proposed framework employs real-time resource allocation to coordinate the progression of concurrent cell analysis. Simulation results show that CoSyn efficiently utilizes platform resources and outperforms baseline techniques.

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

### 12.2.3 NETWORKED LAB-ON-CHIP ARCHITECTURES

**Title:** VERIFICATION OF NETWORKED LAB-ON-CHIP ARCHITECTURES

**Authors:**
- Andreas Grimmer, Johannes Kepler University of Linz, AT
- Andreas Springer, Johannes Kepler University University Linz, AT
- Robert Wille, Johannes Kepler University Linz, AT

**Abstract:**
Labs-on-Chips (LoCs) revolutionize conventional biochemical processes and may even replace laboratories by integrating and minimizing their functionalities on a single chip. In a promising and emerging realization of LoCs, small volumes of reagents, so-called droplets, transport the biological sample and flow in closed channels of sub-millimeter diameters. This realization is called Networked Labs-on-Chips (NuLoCs). The architecture of an NuLoC defines different paths through which the droplets can flow. These paths are realized by splitting channels into multiple successor channels - so-called bifurcations. However, whether the architecture indeed allows to route droplets along the desired paths and, hence, correctly executes the intended experiment is not guaranteed. In this work, we present the first automatic solution for verifying whether an NuLoC architecture allows to correctly route the droplets. Our evaluations demonstrate the applicability and importance of the proposed solution on a set of NuLoC architectures.

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

### 12.2.4 ACTIVATION-PARALLEL CONVOLUTIONAL STRUCTURES

**Title:** SYNTHESIS OF ACTIVATION-PARALLEL CONVOLUTION STRUCTURES FOR NEUROMORPHIC ARCHITECTURES

**Authors:**
- Seban Kim, Incheon National University, KR

**Abstract:**
Convolutional neural networks have demonstrated continued success in various visual recognition challenges. The convolutional layers are implemented in the activation-serial or fully parallel manner on neuromorphic computing systems. This paper presents an unrolling method that generates parallel structures for the convolutional layers depending on a required level of parallel processing. We analyze the resource requirements for the unrolling of the two-dimensional filters, and propose methods to deal with practical considerations such as stride, borders, and alignment. We apply the propose methods to practical convolutional neural networks including AlexNet and the generated structures are mapped onto a recent neuromorphic computing system. This demonstrates that the proposed methods can improve the performance or reduce the power consumption significantly even without area penalty.

**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

**Chair:**
Francesco Regazzoni, AlaRI/USI, CH

**Co-Chair:**
Georg Sigl, TU Munich, DE

Security tools provide support to build secure systems. Such techniques have made great progress in past years with improvements in SAT solvers, theorem provers and available computing power. This session includes papers that perform information flow checks on hardware designs, to check for information leaks either directly through analysis of the design or indirectly through timing channels.

### 12.3.1 INFORMATION FLOW TRACKING ON RTL DESIGNS

**Title:** REGISTER TRANSFER LEVEL INFORMATION FLOW TRACKING FOR PROVABLY SECURE HARDWARE DESIGN

**Authors:**
- Ryan Kastner, University of California, San Diego, US
- Joshua Marxen, UCSD, US
- Ryan Kastner, UCSD, US

**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)**
expressing multiple models of computations. The final paper compares two alternative implementations to realize faithfully the logical execution time model of computation.

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

12.3.2 DUDE, IS MY CODE CONSTANT TIME?
Speaker: Oscar Reparaz, KU Leuven/COSIC, BE
Authors: Oscar Reparaz, Josep Balasch and Ingrid Verbauwhede
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.

17:30 End of session
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 networks, stresses coming from thermal or other stress sources.

**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.

**In this paper, we propose, 3DSim, which is an ultrafast 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.

**A FAST LEAKAGE AWARE THERMAL SIMULATOR FOR 3D CHIPS**

**In this paper, 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.**

**The second paper introduces a fast and accurate thermal simulator for 3D circuits, taking into account thermal leakage dependency.** To further speed up our 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.
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 the measurements of the thermal sensors and the total power consumption measurement. Our identification technique is blind as it does not require design knowledge of the thermal model to identify the power sources. Furthermore, our technique makes no use of the performance counters, which reduces its overhead, and works seamlessly with dynamic voltage and frequency scaling. We implement our technique on a real multi-core CPU-GPU processor-based system, and we show the ability to identify the runtime power consumption of the individual cores using just the total power measurement and the measurements of the thermal sensors under different workloads. We also verify the superior accuracy of our approach using results from a controlled simulation environment.

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

17:00 12.5.3 BLIND IDENTIFICATION OF POWER SOURCES IN PROCESSORS
Speaker: Sherief Reda, Brown University, US
Authors: Sherief Reda1 and Adel Belouchrani2
1Brown University, US; 2ENP, Algeria, DZ

Abstract
The ability to measure power consumption is at the heart of power and thermal management techniques. Modern processors are equipped with hardware monitoring mechanisms that can measure total power. However, this lumped measurement is not sufficient if there is a need to execute fine-grain thermal and power management techniques. This paper proposes a new direction for identifying the fine-grain sources of power consumption in many-core processors. For the first time, we show that it is possible to simultaneously identify both the power consumption of different cores and the thermal model of the chip from just the measurements of the thermal sensors and the total power consumption measurement. Our identification technique is blind as it does not require design knowledge of the thermal model to identify the power sources. Furthermore, our technique makes no use of the performance counters, which reduces its overhead, and works seamlessly with dynamic voltage and frequency scaling. We implement our technique on a real multi-core CPU-GPU processor-based system, and we show the ability to identify the runtime power consumption of the individual cores using just the total power measurement and the measurements of the thermal sensors under different workloads. We also verify the superior accuracy of our approach using results from a controlled simulation environment.

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
1MediaTek, TW; 2National Chiao Tung University, TW

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
Architectural optimizations are also presented to improve energy and performance of applications executing on GPU-based platforms.

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

12.7 Software optimization for emerging memory architectures and technologies
Date: Thursday 30 March 2017
Time: 16:00 - 17:30
Location / Room: 3B
Chair:
Amit Singh, University of Southampton, GB
Co-Chair:
Semeen Rehman, Technische Universität Dresden, DE

The papers in this session propose optimization techniques to improve the lifetime and performance of emerging technologies like persistent memory and scalable many-cores. Architectural optimizations are also presented to improve energy and performance of applications executing on GPU-based platforms.

12.7.2 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 Chen¹, Pang-Yen Chou¹, Helmut Graeb² and Mark Po-Hung Lin ¹
¹National Chung Cheng University, TW; ²Technische Universität München, DE; ³TU 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)

12.7.3 ADAPTIVE INTERFERENCE REJECTION IN HUMAN BODY COMMUNICATION USING VARIABLE DUTY CYCLE INTEGRATING DDR RECEIVER
Speaker:
Shreys Sen, Purdue University, US
Authors:
Shovan Maity¹, Debayan Das ¹ and Shreys Sen²
¹Purdue University, US; ²ECE, 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 in 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 adaptive 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)

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)

12.8 Hot Topic Session: Cyberphysical Microfluidic Biochips: EDA Challenges and Opportunities to Bridge the Gap between Microfluidics and Microbiology

Date: Thursday 30 March 2017
Time: 16:00 - 17:30
Location / Room: Exhibition Theatre

Organisers:
Paul Pop, Technical University of Denmark, DK
Seetal Potluri, Technical University of Denmark, DK

Chair:
Jan Madsen, Technical University of Denmark, DK

Co-Chair:
Seetal Potluri, Technical University of Denmark, DK

Microfluidic biochips (also called lab-on-a-chip) are replacing the conventional biochemical analyzers by integrating all the necessary functions for biochemical analysis using microfluidics. The current trend is towards cyberphysical biochip platforms that integrate novel sensors and actuators, as well as on-chip control circuits. Motivated by the similarity to microelectronics, researchers have started to propose EDA tools for the synthesis of microfluidic biochips. However, we advocate for a paradigm shift, to bridge the formidable barrier that separates engineering (or chip design) from practical biochemistry and microbiology. The special session will serve as a “call to arms” for more focused and relevant research to increase the adoption of microfluidics in translational research.

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>Jeffrey McDaniel, University of California, Riverside, US</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Speaker:</strong></td>
<td>Jeffrey McDaniel, William H. Grover and Philip Brisk, University of California, Riverside, US</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Authors:</strong></td>
<td>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 &quot;push-button&quot; 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>Seetal Potluri, Technical University of Denmark, DK</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Speaker:</strong></td>
<td>Seetal Potluri, Alexander Schneider, Martin Herslev-Petersen, Paul Pop, and Jan Madsen</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Authors:</strong></td>
<td>Authors: Seetal Potluri, Alexander Schneider, Martin Herslev-Petersen, Paul Pop, and Jan Madsen</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>Tsung-Yi Ho, National Tsing Hua University, TW</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Speaker:</strong></td>
<td>Authors: Yu-Jhih Chen, Sumit Sharma, Sudip Roy and Tsung-Yi Ho</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Authors:</strong></td>
<td>Authors: Yu-Jhih Chen, Sumit Sharma, Sudip Roy and Tsung-Yi Ho</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>
</tbody>
</table>

Source URL: https://past.date-conference.com/date17/booklet/proof_reading