Goto Session:

- 1.1 Opening Session
- UB01 Session 1
  - 2.1 EXECUTIVE SESSION: How to Handle Today's Design Complexity
  - 2.2 Panel: Emerging vs. Established Technologies: a Two Sphinxes' Riddle at the Crossroads?
  - 2.3 Making automotive systems safer and more energy efficient
  - 2.4 Modern Challenges in Analog and Mixed-Signal Design
  - 2.5 Low-Power and Efficient Architectures
  - 2.6 Real-Time memory hierarchies
  - 2.7 Yield and Reliability for Robust Systems
  - 2.8 Hot Topic: Technology Transfer towards Horizon 2020
- UB02 Session 2
  - 3.1 EXECUTIVE SESSION: Advanced Technology Challenges & Opportunities
  - 3.2 Panel: The World Is Going... Analog & Mixed-Signal! What about EDA?
  - 3.3 Secure Hardware Primitives and Implementations
  - 3.4 Modeling and Optimization of Power Distribution Networks
  - 3.5 Robust Architectures
  - 3.6 Cyber Physical Systems: Security and Co-design
  - 3.7 On line Strategies for Reliability
  - 3.8 Hot Topic: Mission Profile Aware Design - The Solution for Successful Design of Tomorrows Automotive Electronics
- UB03 Session 3
  - 4.1 EXECUTIVE SESSION: Addressing Challenges of Reliable Chips
  - 4.2 Hot Topic: Multicore Systems in Safety Critical Electronic Control Units for Automotive and Avionics
  - 4.3 Secure Device Identification
  - 4.4 "Almost there" emerging technologies
  - 4.5 Memory System Architectures
  - 4.6 Code Generation and Optimization for Embedded Platforms
  - 4.7 Dependable System Design
  - 4.8 State-of-the-art in Verification: European Tertulia IC Design - Enabling AMS Structured Verification / Verification in FPGA & IP design flows
- UB04 Session 4
  - Exhibition-Reception Exhibition Reception
  - 5.1 SPECIAL DAY Hot Topic: Predictable Multi-Core Computing
  - 5.2 Hot Topic: Hacking and Protecting Hardware: Threats and Challenges
  - 5.3 Reliable Systems in the Age of Variability
  - 5.4 Prediction and optimization of timing variations
  - 5.5 Boosting the Scalability of Formal Verification Technologies
  - 5.6 Emerging logic technologies
  - 5.7 Test Generation and Optimization
  - 5.8 Hot Topic: System Integration - The Bridge between More than Moore and More Moore
- UB05 Session 5
  - IP2 Interactive Presentations
  - 6.1 SPECIAL DAY Hot Topic: The fight against Dark Silicon
  - 6.2 Embedded Tutorial: Emerging Transistor Technologies: From Devices to Architectures
  - 6.3 Management of Micro/Macro Renewable Energy Storage Systems
  - 6.4 Power delivery and distribution
  - 6.5 Beyond EDA: Extending the Application Domain of Formal Methods
  - 6.6 Model-Based Design and Hardware/Software Interfaces
  - 6.7 Hardening Approaches at Different Design Levels
  - 6.8 First Time Right in Analog Design Enabling New Business Cases
- UB06 Session 6
  - 7.0 Special Day Keynote
- UB07 Session 7
  - 7.1 SPECIAL DAY Panel: HW/SW Co-Development - The Industrial Workflow
  - 7.2 Embedded Tutorial: Cross Layer Resiliency in Real World
  - 7.3 Low-power methods and multicore architectures for mobile health applications
  - 7.4 Runtime memory optimization and GPU/multicore architectures
1.1 Opening Session

Date: Tuesday, March 25, 2014
Time: 08:30 - 10:30
Location / Room: Grosse Saal

Organiser:
Gerhard Fettweis, Technische Universität Dresden, DE, Contact Gerhard Fettweis
08:30 1.1.1 WELCOME ADDRESSES
Speakers:
Gerhard Fettweis1 and Luca Fanucci2
1Technische Universität Dresden, DE; 2University of Pisa, IT

08:50 1.1.2 PRESENTATION OF DISTINGUISHED AWARDS
Speaker:
DATE Executive Committee, ,
Abstract
DATE 2014 Best Paper Awards
EDAA Lifetime Achievement Award 2014 (Rolf Ernst, TU Braunschweig, DE)
EDAA Outstanding Dissertation Awards 2013
ACM SIGDA Distinguished Service Award (Peter Marwedel, TU Dortmund, DE)
DATE Fellow Award (Enrico Macii, Politecnico di Torino, IT)
IEEE/CEDA Outstanding Service Contribution Award 2013 (Enrico Macii, Politecnico di Torino, IT)
IEEE CS TTTC Outstanding Contribution Award (Enrico Macii, Politecnico di Torino, IT)
IEEE Fellow Award (Cecilia Metra, University of Bologna, IT)
Read More ...

09:10 1.1.3 KEYNOTE ADDRESS: SYSTEM DESIGN CHALLENGES FOR NEXT GENERATION WIRELESS AND EMBEDDED SYSTEMS
Speaker:
David Fuller, National Instruments, US
Abstract
Application demands in our embedded world are growing dramatically. Consumer expectations and the industry’s forward-looking technology roadmaps paint a picture of a connected world full of intelligent devices once thought to have fixed functionalities. Researchers exploring next generation wireless systems, Internet of Things (IoT), and even machine-to-machine (M2M) communications face many challenges in making this vision a reality. Where once a single, isolated design flow addressed the discrete application, heterogeneous multi-processing architectures must be considered and embraced along with the connections to other devices and systems, and real-world sensor data. As the systems grow in complexity, new design approaches must also be developed and employed to expedite the research, design, and development cycle. David Fuller will outline challenges system designers face in developing cyber-physical systems and explore a graphical system design approach that includes hardware abstraction and comprehends a heterogeneous multiprocessing environment while embracing different models of computation. Through this new approach, system designers can shorten design cycles and the time to prototype ultimately accelerating deployment.

09:50 1.1.4 KEYNOTE ADDRESS: THE GROWING IMPORTANCE OF MICROELECTRONICS FROM A FOUNDRY PERSPECTIVE
Speaker:
Gerd Teepe, GLOBALFOUNDRIES, DE
Abstract
Microelectronics is the dominant industrial technology of today. Its rate of innovation, spelled out by Moore’s Law, is exceptional by any commercial metric, especially, as it has been on this trajectory for almost 40 years. It is not surprising, that other industrial sectors are taking advantage of the innovation engine of the semiconductors for its own product innovation: Cars are safer and more economic, medical diagnostics are performing to a significantly higher level, and energy efficiency from the generation to the consumer is a lot more efficient. “The Internet” has become the basis for our communication, organization and planning in our economies with significant impact to our society. However, the Semiconductor industry is under a powerful transformation marked by the following trends: - Design Complexity is facing new challenges, as technological complexity is transferred to the design space at an accelerated pace - The SOC is dominating the design space - Intelligent Things are emerging with unprecedented cognitive and motion capabilities - The supply chain transformation is in full motion, with the foundry model at the forefront With these powerful trends in motion, we will have to rethink our approach towards semiconductors as part of the industrial system. It will not be sufficient any more to "enhance" traditional products like Cars, TVs, machines or phones with semiconductor content to make them perform at a higher level to increase its value to consumers. We need to rethink the connected world around us to truly assess the next generation of intelligent applications, which we are about to enter.

10:30 End of session
Coffee Break in Exhibition Area
On Tuesday-Thursday the coffee and lunch breaks will be located in the Exhibition Area (Terrace Level).

UB01 Session 1
Date: Tuesday, March 25, 2014
Time: 10:30 - 12:30
Location / Room: University Booth, Booth 3, Exhibition Area
Quantum circuits use quantum-mechanical properties of certain physical systems, such as superposition and entanglement, to perform massively parallel calculations. They provide polynomial algorithms for problems for which only inefficient algorithms with asymptotically-exponential running time are known in conventional models of computation. Building a scalable quantum computer that can process a large number of quantum bits (qubits) is one of the grand challenges of modern science. While first small quantum computers have been experimentally demonstrated and a number of implementation technologies have been suggested, all of them encounter difficulties when it comes to scaling. The central difficulty is the high susceptibility of such circuits to noise and decoherence, which necessitates the use of special quantum error correction. Topological quantum computing (TQC) is a paradigm that offers a path to scalability. It strikes a balance between systematic, intuitive methods to design large computations, and relatively loose requirements on the vulnerability of individual qubits to errors. The availability of a platform for implementing large quantum algorithms constitutes the need for methods to manage design complexity, including automatic synthesis, optimization, compilation, verification and visualization of TQC circuits. Topological quantum circuits are based on a three-dimensional cluster of qubits which supports highly efficient topological quantum error-correcting codes. In this way, the circuits can operate even though its individual qubits are subject to relatively high error rates. We will present the first environment for design of TQC circuits. The environment allows the user to graphically enter the structure of a circuit, add, delete and re-shape individual qubits, and perform optimization and compilation (both manually and by global replacement). The circuits are represented on an intermediate technology-independent level, where "logical qubits" that consist of a large number of physical qubits perform error-corrected operations. For example, the circuit in Fig. 1 shows an error-corrected CNOT gate implemented by four logical qubits represented by colored structures. The optimized representation can be translated into instruction sequences for a classical computer that operates the actual quantum hardware.

More information ...
Two sphinxes, the knight and the knave, the lady and the tiger, are just few instances of difficult puzzles that have kept logicians and mathematicians busy for the last 5,000 years. Today, you
crossroads have always been challenging: they require a decision; in Egyptian and Greek mythology they were often guarded by two sphinxes trying to cheat the traveler with their riddles. The
Giovanni De Micheli, EPFL, CH,
Chair:
Marco Casale-Rossi, Synopsys, Inc., US,
Organiser:
Location / Room:
Date:
2.2 Panel: Emerging vs. Established Technologies: a Two Sphinxes' Riddle at the Crossroads?

13:00 End of session
13:00 Lunch Break in Exhibition Area
Sandwich lunch

2.1 EXECUTIVE SESSION: How to Handle Today’s Design Complexity
Date: Tuesday, March 25, 2014
Time: 11:30 - 13:00
Location / Room: Saal 1
Organiser:
Yervant Zorian, Fellow & Chief Architect, Synopsys, US, Contact Yervant Zorian
Executives:
Sanjeev Agarwala, Fellow & Silicon Director, Texas Instruments, US
Paul Lo, Senior Vice President, Synopsys, US
Rainer Kress, Head Design Methodology, Infineon, DE
Wolfgang Maier, Director, IBM, DE

The widening gap between growing SOC complexity and designer productivity limits traditional chip design methods and flows. This results in several new approaches and innovative methods that work to elevate the limitations of different aspects of complex SOC design. Executives in this session will discuss the impact of complexity and the new opportunities it may bring in designing today’s SOC.

Time | Label | Presentation Title | Authors
---|---|---|---
13:00 | End of session | | 
13:00 | Lunch Break in Exhibition Area | | 

2.2 Panel: Emerging vs. Established Technologies: a Two Sphinxes’ Riddle at the Crossroads?
Date: Tuesday, March 25, 2014
Time: 11:30 - 13:00
Location / Room: Konferenz 6
Organiser:
Marco Casale-Rossi, Synopsys, Inc., US, Contact Marco Casale-Rossi
Chair:
Giovanni De Micheli, EPFL, CH, Contact Giovanni De Micheli

Crossroads have always been challenging: they require a decision; in Egyptian and Greek mythology they were often guarded by two sphinxes trying to cheat the traveler with their riddles. The two sphinxes, the knight and the knave, the lady and the tiger, are just few instances of difficult puzzles that have kept logicians and mathematicians busy for the last 5,000 years. Today, you
are walking down Moore's Law road when you come to a crossroads: one road brings you into the land of emerging technologies: 14, 10 and 7 nanometer, FDSOI, FinFET, 3D-IC,... beyond and below; the other road holds you into the land of established technologies: 28, 40, 65, and 90 nanometers, possibly even above, A&M/S, MEMS,... Choosing the right road is critical to lead your project and your company to success, but making the right decision is increasingly difficult, as it encompasses complex technical and economic considerations. However, unlike the mythological traveler, you won’t run into the sphinxes but, rather, into some of our industry best experts; unlike the sphinxes, they will strive to provide you with honest advice about the "road conditions", and you are allowed to ask multiple questions to them to figure out which road is the best for you.

Panelists:
- Rob Atkenu, ARM Ltd., US
- Antun Domici, Synopsys, Inc., US
- Manfred Horstmann, GLOBALFOUNDRIES, DE
- Robert Hum, Mentor Graphics Corp., US
- Philippe Magarshack, STMicroelectronics, FR

13:00 End of session
Lunch Break in Exhibition Area
Sandwich lunch

2.3 Making automotive systems safer and more energy efficient

Date: Tuesday, March 25, 2014
Time: 11:30 - 13:00
Location / Room: Konferenz 1

Chair:
Bart Vermeulen, NXP, NL, Contact Bart Vermeulen

Co-Chair:
Sebastian Steinhorst, TUM-CREATE, SG, Contact Sebastian Steinhorst

With the transition from hydraulic and mechanical automotive systems to electronic systems, the requirements on safety and energy efficiency are becoming increasingly important. The papers in this session address these issues by presenting robustness improvements at component and system level, advanced energy management at network level, and multi-variant design space exploration.

Time | Label | Presentation Title | Authors
--- | --- | --- | ---
11:30 | 2.3.1 | EMULATION-BASED ROBUSTNESS ASSESSMENT FOR AUTOMOTIVE SMART-POWER ICS | Manuel Harrant\(^1\), Thomas Nirmaier\(^1\), Jerome Kirsch\(^1\), Christoph Grimm\(^2\) and Georg Pelz\(^1\)
|  |  | Speakers: | Infineon Technologies AG, DE; TU Kaiserslautern, DE
|  |  | Abstract | In this paper we present a concept for assessing the robustness of automotive smart power ICs through lab measurements with respect to application variance and parameter spread. Classical compliance to the product specification, where only minimum and maximum values are defined, is not enough to assess device robustness since complex transients of application components influence single specification parameters. That is why application testing becomes a necessary task to reduce device failures, which may occur in the application. One solution would be to enhance traditional lab verification methods with a concept that considers application and parameter spread. This innovative concept is demonstrated on an electronic throttle control application. It has been emulated in real-time, including power amplification and application-relevant parameters. Within this application space, Monte Carlo experiments were carried out to evaluate the influence of parameter spread on selected system characteristics. Finally, an appropriate metric was used to quantify the robustness of the micro-electronic device within its application.

12:00 | 2.3.2 | STARTUP ERROR DETECTION AND CONTAINMENT TO IMPROVE THE ROBUSTNESS OF HYBRID FLEXRAY NETWORKS | Alexander Kordes\(^1\), Bart Vermeulen\(^2\), Abhijit Deb\(^1\) and Michael Wahl\(^1\)
|  |  | Speakers: | University of Siegen, DE; NXP Semiconductors, NL
|  |  | Abstract | The research and development on in-vehicle networks (IVNs) is driven by two main requirements: bandwidth and robustness. In this paper we address the robustness requirement. We focus on FlexRay IVNs that are used for safety-critical applications. We analyze and discuss faults that may affect the startup and operation of a FlexRay network. These failures may not only occur during the startup phase of the vehicle, but they may also happen due to a bus problem that requires the bus to be reinitialized during normal operation. Here any startup failure leads to a critical situation like a brake system failure. The fault scenarios we discuss in this paper are the resetting leading cold start node (RLCN), the dead cold start node (DLCN), and the babbling idiot (BI). These faults are described in literature, but neither the precise behavior of all involved nodes, nor a clear solution is provided to contain their impact. The idea of a bus guardian (BG) is given in a draft specification of the FlexRay consortium, but no details are given. In this paper, we extend these ideas by investigating and implementing a detailed (BG) concept, based on our fault analysis. We subsequently evaluate the successful containment of the three fault types in simulation. We also quantity the chip area cost of our solution.

12:30 | 2.3.3 | A SELF-PROPAGATING WAKEUP MECHANISM FOR POINT-TO-POINT NETWORKS WITH PARTIAL NETWORK SUPPORT | Jan Reinke Seyler, Thilo Streichert, Juri Warkentin, Matthias Spägel, Michael Gläß and Jürgen Teich\(^1\)
|  |  | Speakers: | Daimler AG, DE; TU Kaiserslautern, DE
|  |  | Abstract | As a result of the increased demand for bandwidth, current automotive networks are getting more heterogeneous. New technologies like Ethernet as a packet-switched point-to-point network are introduced. Nevertheless, the requirements on stand-by power consumption and short activation times are still the same as for existing field buses. Ethernet does not provide wakeup mechanisms that are sufficient for automotive systems. As a remedy, this paper introduces a novel physical-layer mechanism called Low Frequency Wakeup that is largely independent of the communication technology and topology used. It provides parallel and remote wake up for all nodes even in a point-to-point network as well as full support of partial networking. The overall wakeup detection time is smaller than 10 ms and every node can actively feed a wakeup signal asynchronously to all other nodes. In terms of latency, it is shown that Low Frequency Wakeup reaches a reduction of more than 50 % for a three-hop network and more than 50 % for a five-hop network in comparison to the current state-of-the-art technology for automotive point-to-point networks.

This paper proposes a novel design method for modern automotive electrical and electronic (E/E) architecture component platforms. The addressed challenge is to derive an optimized component platform termed Baukasten where components, i.e., different manifestations of Electronic Control Units (ECUs), are reused across different car configurations, models, or even OEM companies. The proposed approach derives an efficient graph-based exploration model from defined functional variants. From this, a novel symbolic formulation of multi-variant resource allocation, task binding, and message routing serves as input for a state-of-the-art hybrid optimization technique to derive the individual architecture for each functional variant and the resulting Baukasten at once. For the first time, this enables a concurrent analysis and optimization of individual variants and the Baukasten. Given each manifestation of a component in the Baukasten induces production, storage, and maintenance overhead, we particularly investigate the trade-off between the number of different hardware variants and other established design objectives like monetary cost. We apply the proposed technique to a real-world automotive use case, i.e., a subsystem within the safety domain, to illustrate the advantages of the multi-variant-based design space exploration approach.

Abstract

In this paper, we propose SAFE (Security Aware FlexRay scheduling Engine), to provide a problem definition and a design framework for FlexRay segment schedule to address the new challenge on security. From a high level specification of the application, the architecture and communication middleware are synthesized to satisfy security requirements, in addition to extensibility, costs, and end-to-end latencies. The proposed design process is applied to two industrial case studies consisting of a set of active safety functions and an X-by-wire system respectively.

Abstract

When a single bit is flipped as a result of a transient error in an electronic circuit, its effect can have a severe impact if the circuit is deployed in safety critical domains such as automotive, aeronautics, and industrial automation. In the design phase it is therefore essential to evaluate, and where necessary improve, the resilience of a circuit to all possible transient errors. In this paper, we present a method to analyze the transient error resiliency of a digital circuit. This method is based on an analytical model. It models a transient error as a random function and finds the vulnerable number of bits for each node. We perform a case study on a circuit implementation of a well-known adaptive filter algorithm. The results from the analytical and simulation models show that the analytical model is accurate enough to estimate the effects of transient errors on the performance of a digital circuit. Our analytical method also reduces the run time significantly in a design phase.

2.4 Modern Challenges in Analog and Mixed-Signal Design

Date: Tuesday, March 25, 2014
Time: 11:30 - 13:00
Location / Room: Konferenz Z
Chair: Georges Gielen, KU Leuven, BE, Contact Georges Gielen
Co-Chair: Günhan Dündar, Bogazici University, TR, Contact Günhan Dündar

The session addresses complex challenges in analogue and mixed-signal modeling and design. The regular papers present a novel, zonotope based approach to non-linear macromodeling and a new layout technique in analogue IC design that avoids failures due to IR-drop and electromigration. The two short papers discuss new mechanisms to select solutions from multi-dimensional spaces.

2.4.1 ELECTROMIGRATION-AWARE AND IR-DROP AVOIDANCE ROUTING IN ANALOG MULTIPORT TERMINAL STRUCTURES

Abstract

This paper describes an electromigration-aware and IR-Drop avoidance routing approach considering multiport multiterminal (MP/MT) signal nets of analog integrated circuits (IC). The effects of current densities and temperature in the interconnects may cause the malfunction/failure of a circuit due to IR-Drop or electromigration (EM). These become increasingly more relevant with the ongoing reduction in circuit sizes caused by the evolution of the nanoscale integration processes. Therefore, EM and IR-Drop effects must be taken into account in the design of both power networks and signal wires of analog and mixed-signal ICs, making their impact on the circuits’ reliability negligible. In previous EM and IR-Drop-aware analog IC routing approaches, `dot-models’ are assumed for the terminals, i.e., each terminal has only one port that need to be routed, however, in practice, analog standard cells usually contain multiple electrically-equivalent locations, often distributed over different fabrications layers, where legal connections can be made, i.e., MP terminals, which need to be properly explored. The design flow is detailed, and the applicability of the approach is demonstrated with experimental results, and also, by generating the routing of an analog circuit structure for the UMC 130nm design process.
12:00 2.4.2 ZONOTOPE-BASED NONLINEAR MODEL ORDER REDUCTION FOR FAST PERFORMANCE BOUND ANALYSIS OF ANALOG CIRCUITS WITH MULTIPLE-INTERVAL-VALUED PARAMETER VARIATIONS

Speakers: Yang Song, Sai Manoj Pd and Hao Yu, Nanyang Technological University, SG

Abstract

It is challenging to efficiently evaluate performance bound of high-precision analog circuits with multiple parameter variations at nano-scale. In this paper, a nonlinear model order reduction is proposed to deploy zonotope-based model for multiple-interval valued parameter variations. As such, one can have a zonotope-based reachability analysis to generate a set of trajectories with performance bound defined. By further constructing local parameterized subspaces to approximate a number of zonotopes along the set of trajectories, one can perform nonlinear model order reduction to generate the performance bound under parameter variations. As shown by numerical experiments, the zonotope-based nonlinear macromodeling by order of 19 achieves up to 500x speedup when compared to Monte Carlo simulations of the original model, and up to 50% smaller error when compared to previous parameterized nonlinear macromodeling under the same order.

12:30 2.4.3 IMPLEMENTATION ISSUES IN THE HIERARCHICAL COMPOSITION OF PERFORMANCE MODELS OF ANALOG CIRCUITS

Speakers: Manuel Velasco-Jiménez, Rafael Castro-López, Elsaendia Roca and Francisco Fernández, IMSE-CNMT-CSIC and Universidad de Sevilla, ES

Abstract

Emerging hierarchical design methodologies based on the use of Pareto-optimal fronts (PoFs) are promising candidates to reduce the bottleneck caused by the design of complex analog circuits. However, little work has been reported about how to transmit the information provided by the PoFs of low hierarchical level blocks through the hierarchy to compose the performance models of higher level blocks. This composition actually poses several problems such as the dependence of the PoF performances on the surrounding circuitry and the complexity of dealing with multi-dimensional PoFs in order to explore more efficiently the design space. To deal with these problems, this paper proposes new mechanisms to represent and select candidate solutions from multi-dimensional PoFs that are transformed to the changing operating conditions enforced by the surrounding circuitry. These mechanisms are demonstrated with the generation of the performance model of an active filter by composing previously generated PoFs of operational amplifiers.

12:45 2.4.4 MODELING OF AN ANALOG RECORDING SYSTEM DESIGN FOR ECG AND AP SIGNALS

Speakers: Nil Heidmann1, Neo Hellwege1, Tim Hoehlein1, Thomas Westphall1, Dagmar Peters-Drolshagen1 and Steffen Paul1
1University of Bremen, DE; 2University Bremen, DE

Abstract

The recording of neural activities has turned out to be a promising approach to understand the basic function of specific brain parts like the visual or motor cortex. However, the development and design of advanced neural recording systems is very challenging since the number of parallel measurement channels increases continuously. Beside the analog recording channels digital preprocessing becomes mandatory to handle the corresponding amount of data and to adapt this data to the available transmission bandwidth. In this paper we present the design as well as the behavioral modeling of an analog recording front-end. Simulation and measurement results demonstrate the performances of the system for recording neural signals. Since simulation of this analog front-end is very time consuming but essential for large fully-integrated designs, a mixed-signal model approach is introduced that enables a significant simulation acceleration of integrated and external analog front-ends. The simulation can be accelerated by a factor of up to 22.2 for a single front-end. The proposed system has been fabricated in a 0.35 µm CMOS technology and performances have been measured. This demonstrates that the behavioral model is compatible to the transistor level design. A neural spike detector shows the transient performance of the modeled design on real input stimuli.

13:00 411 MODELBASED HIERARCHICAL OPTIMIZATION STRATEGIES FOR ANALOG DESIGN AUTOMATION

Speakers: Engin Afacan1, Gunhan Dundar1, Falk Baskaya1, Sinme Ay1 and Francisco Fernandez2
1Bogazici University, TR; 2University de Sevilla, TR

Abstract

The design of complex analog circuits by using flat optimization-based approaches is inefficient, even impossible, due to the high number of design variables and the growth of the cost of performance evaluation with the circuit size. Over the past two decades, top-down hierarchical design approaches have been developed and applied. They are based on hierarchical circuit decomposition and specification transmission from top-level to lower level blocks. However, such specification transmission is usually performed with little knowledge on the feasibility of the specifications, leading, therefore, to costly redesign iterations. Even if the specification transmission is successful, there is no guarantee that it is optimal in terms of e.g., power consumption or area occupation. To palliate this problem, two novel model-based hierarchical synthesis methods are proposed in this paper: Model-Based Hierarchical Optimization (MBHO) and Improved Model-Based Hierarchical Optimization (IMBHO). They are based on the concurrent design at higher and lower hierarchical levels and appropriate communication between the different processes. Experimental results on a filter example comparing the new approaches and the conventional top-down design approach are provided.

13:01 925 A NOVEL LOW POWER 11-BIT HYBRID ADC USING FLASH AND DELAY LINE ARCHITECTURES

Speakers: Haun-Cheng Lee and Jacob Abraham, the University of Texas at Austin, US

Abstract

This paper presents a novel low power 11-bit hybrid ADC using flash and delay line architectures, where a 4-bit flash ADC is followed by a 7-bit delay-line ADC. This hybrid ADC inherits accuracy and power efficiency from flash ADCs and delay-line ADCs, respectively. Also, in order to reduce the power of the first stage flash ADC, a power-saving technique is adopted by biasing the DC tail current of the pre-amplifiers at 5µA instead of the operational current, 47µA in stand-by mode. The hybrid ADC was designed and simulated in a commercial 65nm process. With a 1.1 V supply and 100 MHz, the ADC achieves an SNDR of 60 dB and consumes 1.6 mW, which results in a figure of merit (FOM) of 19.4 fJ/conversion-step without any calibration technique. Also, Monte Carlo simulations are performed with a 3σ device mismatch for the SNDR estimation, and the SNDR is observed to be better than 58.5 dB.

13:02 752 SEMI-SYMBOLIC ANALYSIS OF MIXED-SIGNAL SYSTEMS INCLUDING DISCONTINUITIES

Speakers: Cama Radoljic, Christoph Grimm, Javier Moreno and Xiao Pan, TU Kaiserslautern, DE

Abstract

The paper describes an approach for semi-symbolic analysis of mixed-signal systems that contain discontinuous functions, e.g. due to modeling comparators. For modeling and semi-symbolic simulation, we use extended Affine Arithmetic. Affine Arithmetic is currently limited to accurate analysis of linear functions and mild non-linear functions, but not yet discontinuities. In this paper we extend the approach to also handle discontinuities. For demonstration, we symbolically analyze a ΣΔ-modulator.
This session presents three papers on energy efficiency in memory-intensive systems. The first paper aims at energy-efficient scheduling of cooperative-thread arrays on GPGPUs for memory intensive workloads through throttling of warps on different cores. The second paper leverages the application-specific knowledge of the next-generation parallelized high-efficiency video encoder to design a distributed scratchpad memory system with adaptive SPM data allocation and power management. The third paper explores the feasibility of non-volatile memories for intensive workloads through throttling of warps on different cores.

### 2.5 Low-Power and Efficient Architectures

**Date:** Tuesday, March 25, 2014  
**Time:** 11:30 - 13:00  
**Location / Room:** Konferenz 3

**Chair:**  
Cristina Silvano, Politecnico di Milano, IT, Contact Cristina Silvano

**Co-Chair:**  
Todd Austin, University of Michigan, US, Contact Todd Austin

This session presents three papers on energy efficiency in memory-intensive systems. The first paper aims at energy-efficient scheduling of cooperative-thread arrays on GPGPUs for memory intensive workloads through throttling of warps on different cores. The second paper leverages the application-specific knowledge of the next-generation parallelized high-efficiency video encoder to design a distributed scratchpad memory system with adaptive SPM data allocation and power management. The third paper explores the feasibility of non-volatile memories for intensive workloads through throttling of warps on different cores.

#### 13:00 End of session

**Lunch Break** In Exhibition Area  
Sandwich lunch

---

### 2.5.1 ENERGY-EFFICIENT SCHEDULING FOR MEMORY-INTENSIVE GPGPU WORKLOADS

**Speakers:**  
Seokwoo Song¹, Minseok Lee¹, John Kim¹, Woong Seo¹, Yeongon Cho² and Soojung Ryu²  
¹KAIST, KR; ²Samsung, KR

**Abstract**  
High performance for a GPGPU workload is obtained by maximizing parallelism and fully utilizing the available resources. However, this is not necessarily energy-efficient, especially for memory-intensive GPGPU workloads. In this work, we propose Throttle CTA (cooperative-thread array) Scheduling (TCS) where we leverage two type of throttling - throttling the number of active cores and throttling of warp execution in the cores - to improve energy-efficiency for memory-intensive GPGPU workloads. The algorithm requires the global CTA or thread block scheduler to reduce the number of cores with assigned thread blocks while leveraging the local warp scheduler to throttle memory requests for some of the cores to further reduce power consumption. The proposed TCS scheduling does not require off-line analysis but can be done dynamically during execution. Instead of relying on conventional metrics such as miss-per-kilo-instruction (MPKI), we leverage the memory access latency metric to determine the memory intensity of the workloads. Our evaluations show that TCS reduces energy by up to 48% (38% on average) across different memory-intensive workload while having very little impact on performance for compute-intensive workloads.

---

### 2.5.2 DSVM: ENERGY-EFFICIENT DISTRIBUTED SCRATCHPAD VIDEO MEMORY ARCHITECTURE FOR THE NEXT-GENERATION HIGH EFFICIENCY VIDEO CODING

**Speakers:**  
Felipe Sampaio¹, Mohammad Shafique², Bruno Zatt², Sergio Bampi² and Jörg Henkel²  
¹Federal University of Rio Grande do Sul, BR; ²Karlsruhe Institute of Technology (KIT), DE; ³Federal University of Pelotas, BR; ⁴Federal University of Rio Grande do Sul, BR

**Abstract**  
An energy-efficient distributed Scratchpad Video Memory Architecture (DSVM) for the next-generation parallel High Efficiency Video Coding is presented. Our DSVM combines private and overlapping (shared) Scratchpad Memories (SPMs) to support data reuse within and across different cores concurrently executing multiple parallel HEVC threads. We developed a statistical method to size and design the organization of the SPMs along with a supporting memory reading policy for energy efficiency. The key is to leverage the HEVC and video content knowledge. Furthermore, we integrate an adaptive power management policy for SPMs to manage the power states of different memory parts at run time depending upon the varying video content properties. Our experimental results illustrate that our dSVM architecture reduces the overall memory energy consumption by up to 51%-61% compared to parallelized state-of-the-art solutions [11]. The dSVM external memory energy savings increase with an increasing number of parallel HEVC threads and size of search window. Moreover, our SPM power management reacts to the current video properties and achieves up to 54% on-chip leakage energy savings.
2.6 Real-Time memory hierarchies

Time Label Presentation Title Authors
13:00 IP1-8, 766 EVX: VECTOR EXECUTION ON LOW POWER EDGE CORES Speakers: Milovan Dulic\(^1\), Oscar Pakmari, Aaron Smith\(^1\), Osman Unsall\(^1\), Adrian Cristal\(^1\), Mateo Valero\(^1\) and Doug Burger\(^2\)
\(^1\)Barcelona Supercomputing Center, ES; \(^2\)Microsoft Research, US

Abstract
In this paper, we present a vector execution model that provides the advantages of vector processors on low power, general purpose cores, with limited additional hardware. While accelerating data-level parallel (DLP) workloads, the vector model increases the efficiency and hardware resources utilization. We use a modest dual issue core based on an Exploit Data Graph Execution (EDGE) architecture to implement our approach, called EVX. Unlike most DLP accelerators which utilize additional hardware and increase the complexity of low power processors, EVX leverages the available resources of EDGE cores, and with minimal costs allows for specialization of the resources. EVX adds a control logic that increases the core area by 2.1%. We show that EVX yields an average speedup of 3x compared to a scalar baseline and outperforms multimedia SIMD extensions.

13:01 IP1-9, 730 PROGRAM AFFINITY PERFORMANCE MODELS FOR PERFORMANCE AND UTILIZATION Speakers: Ryan Moore and Bruce Childers, University of Pittsburgh, US

Abstract
Multithreaded applications have a wide variety of behavior, causing complex interactions with today's chip multiprocessor machines. Application threads may have large private working sets, and may compete for cache space and memory bandwidth. These threads benefit from large private caches. Other threads may share data or communicate, and thus, execute more quickly if using shared caches. Many applications fall somewhere in between, requiring careful thread-to-core assignments to maximize performance. Yet because of the large number of thread-to-core assignments on today's chip multiprocessors, it is time and energy prohibitive to exhaustively try and determine the best assignment. In this paper, we present and demonstrate application performance models that predict application performance given a proposed thread-to-core assignment. We show how these models can be quickly built and used to select thread-to-core assignments for multiple programs and to improve system utilization.

13:02 IP1-10, 791 ADVANCED SIMD: EXTENDING THE REACH OF CONTEMPORARY SIMD Architectures Speakers: Matthias Boettcher, Giacomo Gabrielli, Mbou Eyouze, Alastair Reid and Bashir M. Al-Hashimi
\(^1\)University of Southampton, GB; \(^2\)ARM Ltd., GB

Abstract
SIMD extensions have gained widespread acceptance in modern microprocessors as a way to exploit data-level parallelism in general-purpose cores. Popular SIMD architectures (e.g. Intel SSE/AVX) have evolved by adding support for wider registers and datatpaths, and advanced features like indexed memory accesses, per-lane predication and inter-lane instructions, at the cost of additional silicon area and design complexity. This paper evaluates the performance impact of such advanced features on a set of workloads considered hard to vectorize for traditional SIMD architectures. Their sensitivity to the most relevant design parameters (e.g. register/datapath width and L1 data cache configuration) is quantified and discussed. We developed an ARMv7 NEON based ISA extension (ARGON), augmented a cycle accurate simulation framework for it, and derived a set of benchmarks from the Berkeley dwarfs. Our analyses demonstrate how ARGON can, depending on the structure of an algorithm, achieve speedups of 1.5x to 16x.

13:03 IP1-11, 898 A TIGHTLY-COUPLED HARDWARE CONTROLLER TO IMPROVE SCALABILITY AND PROGRAMMABILITY OF SHARED-MEMORY HETEROGENEOUS CLUSTERS Speakers: Paolo Burgio\(^1\), Robin Danko\(^2\), Andrea Marongiu, Philippe Coussy and Luca Benini
\(^1\)University of Bologna, Université de Bretagne-Sud, IT; \(^2\)University of Bologna, IT; \(^3\)University of Bologna, IT

Abstract
Modern designs for embedded many-core systems increasingly include application-specific units to accelerate key computational kernels with orders-of-magnitude higher execution speed and energy efficiency compared to software counterparts. A promising architectural template is based on heterogeneous clusters, where simple RISC cores and specialized HW units (HWPU) communicate in a tightly-coupled manner via L1 shared memory. Efficiently integrating processors and a high number of HW Processing Units (HWPU)s in such an system poses two main challenges, namely, architectural scalability and programmability. In this paper we describe an optimized Data Pump (DP) which connects several accelerators to a restricted set of communication ports, and acts as a virtualization layer for programming, exposing FIFO queues to offload “HW tasks” to them through a set of lightweight APIs. In this work, we aim at optimizing both these mechanisms, for respectively reducing modules area and making programming sequence easier and lighter.
Robust design is tolerance of rare events. Understanding design robustness helps predict and enhance yield.

The papers in this session deal with analysis and management of memory hierarchies for complex real-time systems, both from the deterministic and the probabilistic point of view.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>12:00</td>
<td>2.6.2</td>
<td>WCET-CENTRIC DYNAMIC INSTRUCTION CACHE LOCKING</td>
<td>Huping Ding¹, Yun Liang² and Tulika Mitra¹</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cache locking is an effective technique to improve timing predictability in real-time systems. In static cache locking, the locked memory blocks remain unchanged throughout the program execution. Thus, static locking may not be effective for large programs where multiple memory blocks are competing for few cache lines available for locking. In comparison, dynamic cache locking overcomes cache space limitation through time-multiplexing of locked memory blocks. Prior dynamic locking techniques partition the program into regions and take independent locking decisions for each region. We propose a flexible loop-based dynamic cache locking approach. We do not only select the memory blocks to be locked but also the locking points (e.g., loop level). We judiciously allow memory blocks from the same loop to be locked at different program points for WCET improvement. We design a constraint-based approach that incorporates a global view to decide on the number of locking slots at each loop entry point and then select the memory blocks to be locked for each loop. Experimental evaluation shows that our dynamic cache locking approach achieves substantial improvement of WCET compared to prior techniques.</td>
<td></td>
</tr>
<tr>
<td>12:30</td>
<td>2.6.3</td>
<td>MINIMIZING STACK MEMORY FOR HARD REAL-TIME APPLICATIONS ON MULTICORE PLATFORMS</td>
<td>Chuansheng Dong and Haibo Zeng, McGill University, CA</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Multicore platforms are increasingly used in real-time embedded applications. In the development of such applications, an efficient use of RAM memory is as important as the effective scheduling of software tasks. Preemption Threshold Scheduling is a well-known technique for controlling the degree of preemption, possibly improving system schedulability, and allowing savings in stack space. In this paper, we target at the optimal mapping of tasks to cores and the assignment of the scheduling parameters for systems scheduled with preemption thresholds. We formulate the optimization problems using Mixed Integer Linear Programming framework, and propose an efficient heuristic as an alternative. We demonstrate the efficiency and quality of both approaches with extensive experiments using random systems as well as two industrial case studies.</td>
<td></td>
</tr>
<tr>
<td>12:45</td>
<td>2.6.4</td>
<td>TIME-PREDICTABLE EXECUTION OF MULTITHREADED APPLICATIONS ON MULTICORE SYSTEMS</td>
<td>Ahmed Alhammad and Rodolfo Pellizzoni, University of Waterloo, CA</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>In multicore systems, contention for access to main memory between application threads complicates timing analysis and may lead to pessimistic bounds on execution time. This is particularly problematic for real-time applications, which require provable bounds on worst-case performance. In this work, we employ a predictable execution model to schedule memory accesses performed by application threads without relying on unpredictable hardware arbiters. In addition, we statically schedule application’s threads with the objective to minimize the application’s makespan. Our experimental evaluation on 4-core system with NAS Parallel Benchmarks indicates that the proposed execution scheme yields an aggregated improvement of 21% over contention execution in which application’s threads uncontrollably accessing the main memory.</td>
<td></td>
</tr>
</tbody>
</table>

2.7 Yield and Reliability for Robust Systems

Date: Tuesday, March 25, 2014
Time: 11:30 - 13:00
Location / Room: Konferenz 5

Chair: Joan Figueras, UPC, ES, Contact Joan Figueras

Co-Chair: Jose Pineda de Gyvez, NXP, NL, Contact Jose Pineda de Gyvez

Robustness is increasingly a requirement for SoCs and memories, and effects such as wearout, BTI, and soft errors are important to consider as part of design. Another important component of robust design is tolerance of rare events. Understanding design robustness helps predict and enhance yield.
Bias Temperature Instability Analysis of FinFET Based SRAM Cells

Seyab Khan1, Innocent Agbo2, Said Hamloul, Halil Kukner, Ben Kaczer, Praveen Raghavan1 and Franky Catthoor4

1Technical University Delft, NL; 2TU Delft, NL; 3Delft University of Technology, NL; 4IMEC, BE; 5Imec, BE

Abstract
Bias Temperature Instability (BTI) poses a major reliability challenge for today’s and future nano-devices as it degrades their performance. This paper provides a comprehensive analysis of BTI impact, in terms of time-dependent degradation, on FinFET based SRAM cells. The evaluation metrics are read Static Noise Margin (SNM), hold SNM and Write Trip Point (WTP); while the aspects investigated consist dependence on the supply voltage, cell strength, and design styles (6 versus 8 Transistor cell). A comparison between FinFET and planar CMOS based SRAM cell degradation is also covered. The simulation results for FinFET based cells show that: (a) Read SNM of the cell degrades more by (6.72%) than the other metrics (6.82% in WTP and 14.19% in hold SNM) (b) 12% increment in the cells supply voltage enhances its read SNM by 9% (c) Strengthening only the pull-down transistors in the cell by 1.5 reduces BTI induced read SNM degradation by 26.61% (d) 8T SRAM cells has 1.43 higher WTP than 6T cell; however, the cells suffer from 31.13 higher read SNM and 8.05 higher hold SNM degradations than 6T SRAM cells and (e) FinFET based SRAM cells are more vulnerable to BTI degradation than planar CMOS based cells.

Simulation Reduction Technique for SRAM Yield Analysis

Manish Rana and Ramon Canal, Universitat Politecnica de Catalunya, ES

Abstract
Estimating extremely low SRAM failure probabilities by conventional Monte Carlo (MC) approach requires hundreds-of-thousands simulations making it an impractical problem. To alleviate this problem, failure-probability estimation methods with a smaller number of simulations have recently been proposed, most notably variants of consecutive mean-shift based Importance Sampling (IS). In this method, a large amount of time is spent simulating data points that will eventually be discarded in favor of other data points with minimum norm. This can potentially increase the simulation time by orders of magnitude. To solve this very important limitation, in this paper, we introduce SSFB-a novel SRAM failure-probability estimation method that has much better cognition of the data points compared to conventional approaches. The proposed method starts with radial simulation of a single point and reduces discarded simulations by: a) random sampling only when it reaches a failure boundary and after that continues again with radial simulation of a chosen point, and b) random sampling is performed only within a specific failure-range which decreases in each iteration. The proposed method is also scalable to higher dimensions (more input variables) as sampling is done on the surface of the hyper-sphere, rather than within-the-hypersphere as other techniques do. Our results show that using our method we can achieve an overall 40x reduction in simulations compared to consecutive mean-shift IS methods while remaining within the 0.01-Sigma accuracy.

An Integrated Framework for Early-Stage Memory Robustness Analysis

Shrikanth Ganapathy1, Ramon Canal2, Dan Alexandrescu1, Enrico Costanera3, Antonio Gonzalez2 and Antonio Rubio1

1Universitat Politecnica de Catalunya, ES; 2RoC Technologies, FR; 3intel and Universitat Politecnica de Catalunya, ES

Abstract
With the growing importance of parametric (process and environmental) variations in advanced technologies, it has become a serious challenge to design reliable, fast and low-power embedded memories. Adopting a variation-aware design paradigm requires a holistic perspective of memory-wide metrics such as yield, power and performance. However, accurate estimation of such metrics is largely dependent on circuit implementation styles, technology parameters and architecture-level specifics. In this paper, we propose a fully automated tool — INFORMER that helps high-level designers estimate memory reliability metrics rapidly and accurately. The tool relies on accurate circuit-level simulations of failure mechanisms such as ageing, soft-errands and parametric failures. The obtained statistics can then help couple low-level metrics with higher-level design choices. A new technique for rapid estimation of low-probability failure events is also proposed. We present three use-cases of our prototype tool to demonstrate its diverse capabilities in autonomously guiding large SRAM-based robust memory designs.
new concepts. Furthermore, it will exemplify several success stories from both academic and industrial perspectives. The community is currently investigating novel ways of stimulating additional academia-industry technology transfer. This special session contributes by discussing concrete transfer experiences and research results in computing technologies and EDA actually make it into industrial practice? In the context of the transition into the Horizon 2020 framework program, the European research projects produce many excellent results, and the quality of research papers at DATE and other major European conferences is often outstanding. But how many academic

Abstract

Counterfeit ICs have become an issue for semiconductor manufacturers due to impacts on their reputation and lost revenue. Counterfeit ICs are either products that are intentionally mislabeled or legitimate products that are extracted from electronic waste. The former is easier to detect whereas the latter is harder since they are identical to new devices but display degraded performance due to environmental and use stress conditions. Detecting counterfeit ICs that are extracted from electronic waste requires an approach that can approximate the age of manufactured devices based on their parameters. In this paper, we present a methodology that uses information on both fresh and aged ICs and tries to distinguish between the fresh and aged population based on an estimate of the age. Since analog devices age mainly due to their bias stress, input signals play less of a role. Hence, it is possible to use simulation models to approximate the aging process, which would give us access to a large population of aged devices. Using this information, we can construct a statistical model that approximates the age of a given circuit. We use a Low noise amplifier (LNA) and an NMOS LC oscillator to demonstrate that individual aged devices can be accurately classified using the proposed method.

Luca Fanucci, University of Pisa, IT

SUCCESSFUL TECHNOLOGY TRANSFER - SHARING OF EXPERIENCE

Speaker: Johannes Stahl, Synopsys, Inc., US

Abstract: We will highlight where we see the value of cooperation with universities. We will refer to what the researchers need to do and what industry has to do to make for a successful technology transfer. Our contribution will be based on many years of experience working with RWTH Aachen as our lead university partner.

Johannes Stahl, Synopsys, Inc., US

Abstract: We will present some technology transfer experiences from research to industry of innovative integrated circuit architectures for different application fields (automotive, multimedia and industrial). Starting from the analysis of the relevant scenarios we will discuss the adopted Research/Industry collaboration model based on know-how and human resources sharing. Our intent is to highlight main key points in order to have a successful research/industry technology transfer up to the market.

Luciano Fanucci, University of Pisa, IT

Abstract: Leveraging European Research to Create Value: Experiences from bringing advanced system-software R&D to global industrial use by a European high-tech SME. Using some of the European R&D projects and its results in our 38 years of history demonstrating both business and failure, we will explain some critical success factors in the different phases of technology transfer in our specific domain.

Speaker: Marco Roodzant, ACE Associated Compiler Experts bv, NL

Abstract: Experiences from bringing advanced system-software R&D to global industrial use by a European high-tech SME. Using some of the European R&D projects and its results in our 38 years of history demonstrating both business and failure, we will explain some critical success factors in the different phases of technology transfer in our specific domain.

Speaker: Marco Roodzant, ACE Associated Compiler Experts bv, NL

Abstract: Experiences from bringing advanced system-software R&D to global industrial use by a European high-tech SME. Using some of the European R&D projects and its results in our 38 years of history demonstrating both business and failure, we will explain some critical success factors in the different phases of technology transfer in our specific domain.

Speaker: Marco Roodzant, ACE Associated Compiler Experts bv, NL

Abstract: Experiences from bringing advanced system-software R&D to global industrial use by a European high-tech SME. Using some of the European R&D projects and its results in our 38 years of history demonstrating both business and failure, we will explain some critical success factors in the different phases of technology transfer in our specific domain.

Speaker: Marco Roodzant, ACE Associated Compiler Experts bv, NL

Abstract: Experiences from bringing advanced system-software R&D to global industrial use by a European high-tech SME. Using some of the European R&D projects and its results in our 38 years of history demonstrating both business and failure, we will explain some critical success factors in the different phases of technology transfer in our specific domain.

Speaker: Marco Roodzant, ACE Associated Compiler Experts bv, NL
<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>12:30</td>
<td>2.8.5</td>
<td>OPEN SOURCE TECHNOLOGY TRANSFER: SCENARIOS AND VALUE CREATION</td>
<td>Albert Cohen, INRIA, FR</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>When value creation and business cases are at stake, free and open-source software is perceived as an opportunity and also as a threat. We will go through selected examples of the successful transfer of research results into industrial use, based exclusively or in part on open source platforms. The talk will build on personal experience conducting research in production compilers, and collecting the experiences of fellow researchers at IRILL, a joint initiative of INRIA and two French Universities promoting research and innovation on free software.</td>
<td></td>
</tr>
<tr>
<td>12:45</td>
<td>2.8.6</td>
<td>SUPPORTING INTERNATIONAL TECHNOLOGY TRANSFER: OBJECTIVES AND OBSTACLES</td>
<td>Bernd Janson, consultant, DE</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Founded in 1984 ZENIT in Mülheim/Ruhr, North Rhine Westphalia, Germany, offers qualified support especially for SMEs who are engaged in innovation business like research and innovation funding programmes or international technology transfer. As part of the Enterprise Europe Network and as National Contact Point for SMEs ZENIT is focused on consultancy services for innovative companies and other players like universities and research centres. To get new insights on the impact of FP7 (Framework Programme 7 of the European Union) projects into the NRW-market ZENIT started a series of interviews with companies and universities funded in FP7. First results show that positive effects on science, basic research and further application-oriented research and development are dominating. Nevertheless there are some incidences for positive effects on the NRW-market through innovative products and processes but they are quite rare yet. One important barrier for FP7 project innovation is the fact that the contract negotiation with many players often offers only a suboptimal basis for technology transfer business.</td>
<td></td>
</tr>
</tbody>
</table>

13:00 End of session

Lunch Break in Exhibition Area

Sandwich lunch

**UB02 Session 2**

**Date:** Tuesday, March 25, 2014

**Time:** 12:30 - 15:00

**Location / Room:** University Booth, Booth 3, Exhibition Area

<table>
<thead>
<tr>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>UB02.01</td>
<td>QUANTUMEDA: A VISUALIZATION AND DESIGN ENVIRONMENT FOR TOPOLOGICAL QUANTUM CIRCUITS</td>
<td>Ila Pollan, Wolfgang Waltner and Alexandru Paler, University of Passau, DE</td>
</tr>
<tr>
<td></td>
<td><strong>Abstract</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Quantum circuits use quantum-mechanical properties of certain physical systems, such as superposition and entanglement, to perform massively parallel calculations. They provide polynomial algorithms for problems for which only inefficient algorithms with asymptotically-exponential running time are known in conventional models of computation. Building a scalable quantum computer that can process a large number of quantum bits (qubits) is one of the grand challenges of modern science. While first small quantum computers have been experimentally demonstrated and a number of implementation technologies have been suggested, all of them encounter difficulties when it comes to scaling. The central difficulty is the high susceptibility of such circuits to noise and decoherence, which necessitates the use of special quantum error correction. Topological quantum computing (TQC) is a paradigm that offers a path to scalability. It strikes a balance between systematic, intuitive methods to design large computations, and relatively loose requirements on the vulnerability of individual qubits to errors. The availability of a platform for implementing large quantum algorithms constitutes the need for methods to manage design complexity, including automatic synthesis, optimization, compaction, verification and visualization of TQC circuits. Topological quantum circuits are based on a three-dimensional cluster of qubits which supports highly efficient topological quantum error-correcting codes. In this way, the circuits can operate even though its individual qubits are subject to relatively high error rates. We will present the first environment for design of TQC circuits. The environment allows the user to graphically enter the structure of a circuit, add, delete and re-shape individual qubits, and perform optimization and compaction (both manually and by global replacement). The circuits are represented on an intermediate technology-independent level, where &quot;logical qubits&quot; that consist of a large number of physical qubits perform error-corrected operations. For example, the circuit in Fig. 1 shows an error-corrected CNOT gate implemented by four logical qubits represented by colored structures. The optimized representation can be translated into instruction sequences for a classical computer that operates the actual quantum hardware.</td>
<td></td>
</tr>
<tr>
<td>UB02.02</td>
<td>AN AUTOMATED DESIGN FLOW FOR FAST PROTOTYPING OF SIMULINK MODELS ONTO MPSOC</td>
<td>Francesco Robino and Johnny Öberg, Royal Institute of Technology, SE</td>
</tr>
<tr>
<td></td>
<td><strong>Abstract</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Simulink is a modelling environment suitable to model embedded systems at system-level. However there is no standard to rapidly prototype Simulink models onto modern multiprocessor systems-on-chip (MPSoC). In this demonstration we show how our NoC System Generator tool can be used as part of an automated platform-based design flow to synthesize a Simulink model to a network-on-chip based MPSoC implementation on FPGA. The performance of the generated prototype scales with the number of processors.</td>
<td></td>
</tr>
</tbody>
</table>

More information ...
Building a Prototyping Platform for Investigating the Impact of Attacks against Automotive Networks

Authors: Alexander Stühring, Günter Ehmen, and Sibylle Fritsche

University of Oldenburg, DE; zOFFIS, DE

Abstract
The University of Oldenburg is working on solutions to ensure a secure communication in the automotive domain. This is a key requirement for safe applications in the context of future CadX applications. In order to achieve this goal, we are using a self-developed prototyping platform to analyze and demonstrate the impact of attacks on in-vehicle buses and wireless networks. Moreover, the visitors are able to start attacks and observe the consequences in a simulated driving scenario.

More information ...

ID.FIX: An EDA Tool for Fixed-Point Refinement of Embedded Systems

Authors: Olivier Sentieys, Daniel Menard, and Nicolas Simon

INRIA, FR; ZINSA Rennes, FR; University of Rennes, FR

Abstract
Most of digital image and signal processing algorithms are implemented into architectures based on fixed-point arithmetic, to satisfy the cost and power consumption constraints of embedded systems. The fixed-point conversion process (or refinement) is crucial for reducing the time-to-market. Design tools to automate this phase and to explore the design space are thus required. The ID.FIX EDA tool based on the compiler infrastructure GECOS allows for the conversion of a floating-point C source code into a C code using fixed-point data types. The data word-lengths are optimized in order to minimize the implementation cost under accuracy constraint. To obtain low optimization time, an analytical approach is used to evaluate the fixed-point computation accuracy. This approach is valid for systems made-up of any (smooth) arithmetic operations.

More information ...

MicroTESK: Reconfigurable Open-Source Framework for Test Program Generation

Authors: Andrei Tatarnikov, Alexander Kamkin, and Artem Kotouchnyak

Institute for System Programming of the Russian Academy of Sciences (ISP RAS), RU

Abstract
Test program generation plays a major role in functional verification of microprocessors. Due to tremendous growth in complexity of modern designs and rigid constraints on time to market, it becomes an increasingly difficult task. In spite of powerful test program generation tools available in the market, development of functional tests is still known to be the bottleneck of the microprocessor design cycle. The common problem is that it takes a significant effort to reconfigure a test program generation environment for a new microprocessor design. The model-based approach applied in the state-of-the-art tools, like Genesis-PRO (IBM Research), still does not provide enough flexibility since creating a microprocessor model is difficult and requires special knowledge and skills. MicroTESK, the open-source test program generation framework being developed at ISP RAS, offers an approach to ease customization by using light-weight formal specifications to describe the target microprocessor architecture. The approach helps reduce the effort needed to create a microprocessor model and, consequently, minimize the time required to create functional tests. In addition to gaining flexibility, the use of formal specifications also allows automated extraction of knowledge about test situations that occur in a microprocessor (coverage model), thus, facilitating creating directed tests and improving test coverage. By the present moment, a demo prototype of MicroTESK has been implemented. It uses the SimNML architecture description language to specify the target microprocessor architecture and provides a convenient Ruby-based language for creating test templates that serve as an abstract description of test programs to be generated. The current version of the framework focuses primarily on RISX microprocessors including ARM, MIPS, and SPARK. Supported test generation methods include random, combinational, template-based, and model-based generation. Flexible architecture of the framework allows adding support for new test generation methods.
**FAULTIFY: PROBABILISTIC CIRCUIT FAULT EMULATION**

**Authors:**
David May and Walter Stechele, TUM, DE

**Abstract**
We want to demonstrate an FPGA-based probability-aware fault emulator and its corresponding algorithms in the context of a real-time H.264 decoder. The demo will show that reliability constraints can be relaxed inside the circuit without noticeable degradation of the image quality when carefully investigating where the constraints can be relaxed. We will show how this investigation can be done using our emulator and we will show the effect of a relaxed robustness of the circuit in real-time.

**More information...**

15:00 End of session
16:00 **Coffee Break** in Exhibition Area

On Tuesday-Thursday the coffee and lunch breaks will be located in the Exhibition Area (Terrace Level).

### 3.1 EXECUTIVE SESSION: Advanced Technology Challenges & Opportunities

**Date:** Tuesday, March 25, 2014  
**Time:** 14:30 - 16:00  
**Location / Room:** Saal 1

**Organiser:**  
Yervant Zorian, Fellow & Chief Architect, Synopsys, US, [Contact Yervant Zorian](mailto:)

**Executives:**  
Lorent Remont, Vice President, STMicroelectronics, FR  
Joachim Kunkel, Senior Vice President & GM, Synopsys, US  
Rudy Lauwereins, VP, IMEC, BE  
Wenchi Chang, Senior Manager, TSMC, NL  
Gerd Teepe, VP, Global Foundries, DE

The continuous technology scaling and their new applications are dramatically impacting the semiconductor industry. This may also significantly affect the dependency between eco-system players necessitating smooth interdependency between them. The executives in this session will discuss upcoming innovations in the semiconductor industry and their impact on the solutions offered by the eco system players.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>16:00</td>
<td>End of session</td>
<td><a href="#">Coffee Break</a> in Exhibition Area</td>
<td></td>
</tr>
</tbody>
</table>

On Tuesday-Thursday the coffee and lunch breaks will be located in the Exhibition Area (Terrace Level).

### 3.2 Panel: The World Is Going... Analog & Mixed-Signal! What about EDA?

**Date:** Tuesday, March 25, 2014  
**Time:** 14:30 - 16:00  
**Location / Room:** Konferenz 6

**Organiser:**  
Marco Casale-Rossi, Synopsys, Inc., US, [Contact Marco Casale-Rossi](mailto:)

**Chair:**  
Pietro Palella, STMicroelectronics, IT, [Contact Pietro Palella](mailto:)

Contrary to a common belief, the world is not going digital! Analog and mixed-signal electronics is more and more important and yet pervasive. This is due both to the increasing systems integration, by nature leading to heterogeneity, and to the complex, digital computing functions being complemented by scores of on-chip analog functions, interfacing/interacting with people, environment, and other systems. Specialty silicon foundries are now stable members of top ten revenue rankings. This technology trend demands for more design automation in both implementation and verification domains. Lossless interfaces between digital and analog design environments, multi-technology support, mixed-signal simulation engines - but also debugging aids - are no longer a nice to have. According to IBS, the cost of implementing and verifying the mixed-signal functions is generally over 50% of the design costs even though the mixed-signal transistors can be as low as 3% of the total! What are the critical requirement, moving forward, and what is EDA industry doing to serve the needs of this increasingly important semiconductor industry segment?

**Panelists:**
- Mario Anton, Micronas, DE
- Ori Gaizur, TowerJazz, IL
- Robert Hum, Mentor Graphics Corp., US
- Rainer Kress, Infineon Technologies, DE
- Paul Lo, Synopsys, Inc., US

16:00 End of session  
**Coffee Break in Exhibition Area**  
On Tuesday-Thursday the coffee and lunch breaks will be located in the Exhibition Area (Terrace Level).

### 3.3 Secure Hardware Primitives and Implementations

**Date:** Tuesday, March 25, 2014  
**Time:** 14:30 - 16:00  
**Location / Room:** Konferenz 1

**Chair:**  
Paolo Maistri, TIMA, FR, [Contact Paolo Maistri](mailto:)

Contrary to a common belief, the world is not going digital! Analog and mixed-signal electronics is more and more important and yet pervasive. This is due both to the increasing systems integration, by nature leading to heterogeneity, and to the complex, digital computing functions being complemented by scores of on-chip analog functions, interfacing/interacting with people, environment, and other systems. Specialty silicon foundries are now stable members of top ten revenue rankings. This technology trend demands for more design automation in both implementation and verification domains. Lossless interfaces between digital and analog design environments, multi-technology support, mixed-signal simulation engines - but also debugging aids - are no longer a nice to have. According to IBS, the cost of implementing and verifying the mixed-signal functions is generally over 50% of the design costs even though the mixed-signal transistors can be as low as 3% of the total! What are the critical requirement, moving forward, and what is EDA industry doing to serve the needs of this increasingly important semiconductor industry segment?

**Panelists:**
- Mario Anton, Micronas, DE
- Ori Gaizur, TowerJazz, IL
- Robert Hum, Mentor Graphics Corp., US
- Rainer Kress, Infineon Technologies, DE
- Paul Lo, Synopsys, Inc., US

16:00 End of session  
**Coffee Break in Exhibition Area**  
On Tuesday-Thursday the coffee and lunch breaks will be located in the Exhibition Area (Terrace Level).
optimize approaches for their efficient characterization and robust design, including order reduction, variability impact, via planning, decoupling capacitor selection, and thermal effects.

The performance and robustness of 3D power distribution networks is of critical importance for state-of-the-art electronic designs. The papers in this session discuss new modeling and optimization approaches for their efficient characterization and robust design, including order reduction, variability impact, via planning, decoupling capacitor selection, and thermal effects.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>14:30</td>
<td>3.3.1</td>
<td>LIGHTWEIGHT CODE-BASED CRYPTOGRAPHY: QC-MDPC MCELIECE ENCRYPTION ON RECONFIGURABLE DEVICES</td>
</tr>
</tbody>
</table>

**Speakers:**
- Ingo von Maurich, Tim Günsters, Ruhr-Universität Bochum, DE
- Patrick Haddad, Yannick Teglia, Florent Bernard and Viktor Fischer

**Abstract**
Security in true random number generation in cryptography is based on entropy per bit at the generator output. The entropy is evaluated using stochastic models. Several recent works propose stochastic models based on assumptions related to selected physical analog phenomena such as noisy signals and on the knowledge of the principle of randomness extraction from the obtained noisy analog signal. However, these assumptions simplify considerably the underlying analog processes, which include several noise sources. In this paper, we present a new comprehensive multilevel approach, which enables to build the stochastic model based on detailed analysis of noise sources starting at transistor level and on conversion of the noise to the clock jitter exploited at the generator level. Using this approach, we can estimate proportion of the jitter coming only from the thermal noise, which is included in the total clock jitter.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>15:00</td>
<td>3.3.2</td>
<td>ON THE ASSUMPTION OF MUTUAL INDEPENDENCE OF JITTER REALIZATIONS IN P-TRNG STOCHASTIC MODELS</td>
</tr>
</tbody>
</table>

**Speakers:**
- Ingo von Maurich, Tim Günsters, Ruhr-Universität Bochum, DE
- Patrick Haddad, Yannick Teglia, Florent Bernard and Viktor Fischer

**Abstract**
Security in true random number generation in cryptography is based on entropy per bit at the generator output. The entropy is evaluated using stochastic models. Several recent works propose stochastic models based on assumptions related to selected physical analog phenomena such as noisy signals and on the knowledge of the principle of randomness extraction from the obtained noisy analog signal. However, these assumptions simplify considerably the underlying analog processes, which include several noise sources. In this paper, we present a new comprehensive multilevel approach, which enables to build the stochastic model based on detailed analysis of noise sources starting at transistor level and on conversion of the noise to the clock jitter exploited at the generator level. Using this approach, we can estimate proportion of the jitter coming only from the thermal noise, which is included in the total clock jitter.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>15:30</td>
<td>3.3.3</td>
<td>CLOCK-MODULATION BASED WATERMARK FOR PROTECTION OF EMBEDDED PROCESSORS</td>
</tr>
</tbody>
</table>

**Speakers:**
- Jedrzej Kufel, Peter Wilson, Stephen Hill, Bashir Al-Hashimi, Paul N. Whatmough and James Myers

**Abstract**
This paper presents a novel watermark generation technique for the protection of embedded processors. In previous work, a load circuit is used to generate detectable watermark patterns in the ASIC power supply. This approach leads to hardware area overheads. We propose removing the dedicated load circuit entirely, instead to compensate the reduced power consumption the watermark power pattern is emulated by reusing existing clock gated sequential logic as a zero-overhead load circuit and modulating the clock-gating enable signal with the watermark sequence. The proposed technique has been validated through experiments using two ASICs in 65nm CMOS, one with an ARM Cortex-M0 microcontroller and one with a Cortex-A5 microprocessor. Silicon measurements results verify the viability of the technique for embedded processors. Furthermore, the proposed clock modulation technique demonstrates a significant area reduction, without compromising the detection performance. In our experiments an area overhead reduction of 98% was achieved. Through reuse of existing logic and reduction of watermark hardware implementation costs, the proposed clock modulation technique offers an improved robustness against removal attacks.

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

On Tuesday-Thursday the coffee and lunch breaks will be located in the Exhibition Area (Terrace Level).

### 3.4 Modelling and Optimization of Power Distribution Networks

**Date:** Tuesday, March 25, 2014

**Time:** 14:30 - 16:00

**Location / Room:** Konferenz 2

**Chair:**
Luca Daniel, MIT, US, Contact Luca Daniel

**Co-Chair:**
Stefano Grivet-Talocia, Politecnico di Torino, IT, Contact Stefano Grivet-Talocia

The performance and robustness of 3D power distribution networks is of critical importance for state-of-the-art electronic designs. The papers in this session discuss new modeling and optimization approaches for their efficient characterization and robust design, including order reduction, variability impact, via planning, decoupling capacitor selection, and thermal effects.
<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
</table>
| 14:30| 3.4.1 | SENSITIVITY-BASED WEIGHTING FOR PASSIVITY ENFORCEMENT OF LINEAR MACROMODELS IN POWER INTEGRITY APPLICATIONS | Speakers: Andrea Uboli\(^1\), Stefano Grivet-Talocia\(^1\), Michelangelo Bandinu\(^2\) and Alessandro Chinea\(^2\)

\(^1\)Politecnico di Torino, IT; \(^2\)IdemWorks s.r.l., IT

Abstract
The electrical performance of Power Distribution Networks (PDNs) is usually assessed by computing frequency responses through quasi-static or full-wave electromagnetic solvers. Such responses, often available in the scattering form, are then fed to suitable macromodeling algorithms for the extraction of compact reduced-order behavioral models that can be seamlessly simulated in the time domain by standard circuit solvers. Such algorithms perform a rational fitting of the raw scattering responses, followed by a passivity check and enforcement step. The resulting macromodel is typically very accurate when compared to the raw scattering responses. It may however happen that the responses of the PDN macromodel exhibit significant deviation from the true system responses under realistic loading conditions, which include appropriate models for active device blocks, decoupling capacitors, voltage regulators, etc. We highlight the source of this accuracy loss, and we propose a sensitivity-based weighting strategy that is able to optimize and tune the macromodel accuracy based on its specific nominal termination network. The particular focus of this paper is the definition and the inclusion of optimal weights in the passivity enforcement step, which is recognized as the most challenging step. The result is a reliable macromodeling flow, which is able to produce passive, accurate and efficient reduced-order models of general PDN structures for power integrity analysis and verification. |

| 15:00| 3.4.2 | EFFICIENT ANALYSIS OF VARIABILITY IMPACT ON INTERCONNECT LINES AND RESISTOR NETWORKS | Speakers: Jorge Fernandez Villena\(^1\) and Luis Miguel Silvera\(^2\)

\(^1\)INESC-ID, PT; \(^2\)INESC-ID/IST - Lisbon University, PT

Abstract
Continued technology scaling coupled with limited lithographic capabilities is a leading cause of increased design variability. In the nanometer regime lithography tools have failed to keep pace with Moore's Law and printed feature sizes are a small fraction of the wavelength of light used in current processes. Such sub-wavelength printing makes features highly susceptible to perturbations in the lithographic process conditions which leads to printed designs exhibiting increased variability. Such variability directly affects design behavior and performance in multiple ways. One of the areas of concern is power grid (PG) design, where lithographic errors may locally modify the wire widths. These variations, that may affect any and all wires in the grid, have a critical impact on the power distribution across the chip, introducing considerable current fluctuations which are a potential cause for electromigration effects. To analyze and account for the impact of these errors requires a complete extraction of the PG, which generates a large resistive network, potentially with several million elements, whose simulation is computationally challenging. This paper proposes a fast and accurate variability analysis of very large resistor networks, such as PG extracted netlists, that allows estimating the effects of multiple parameter settings in reasonable time. The proposed model can be easily combined with Litho/CMP simulators in order to boost much needed design-aware lithography. |

| 15:30| 3.4.3 | IMPLICIT INDEX-AWARE MODEL ORDER REDUCTION FOR RLC/RC NETWORKS | Speakers: Nicodemus Banagaaya\(^1\), Giuseppe Al'\(\)z, Wil. H. A. Schilders and Caren Tischendorf

\(^1\)Eindhoven University of Technology, NL; \(^2\)University of Calabria and INFN, Gruppo collegato di Cosenza, IT; \(^3\)Institute of Mathematics, Humboldt-Universität zu Berlin, DE

Abstract
This paper introduces the implicit IMOR method for differential algebraic equations. This method is a modification of the Index-aware model order reduction (IMOR) method proposed in our earlier papers which is the explicit-IMOR method. It also allows first splitting the differential-algebraic equations (DAEs) into differential and algebraic parts using a basis of projectors. In contrast with the explicit-IMOR method, the implicit-IMOR method leads to implicit differential and algebraic parts. We demonstrate the implicit-IMOR method using the RLC/RC networks, but it can also be applied to other problems which lead to differential-algebraic equations. |

| 15:45| 3.4.4 | P/G TSV PLANNING FOR IR-DROP REDUCTION IN 3D-ICS | Speakers: Shengcheng Wang\(^1\), Farshad Firooz\(^2\), Fabian Oboni\(^2\) and Mehdi Tahoori\(^1\)

\(^1\)Karlruhe Institute of Technology, DE; \(^2\)Karlruhe Institute of Technology (KIT), DE

Abstract
In recent years, interconnect issues emerged as major performance challenges for Two-Dimensional-Integrated Circuits (2D-ICs). In this context, Three-Dimensional-ICs (3D-ICs), which consist of several active layers stacked above each other, offer a very attractive alternative to conventional 2D-ICs. However, 3D-ICs also face many challenges associated with the Power Distribution Network (PDN) design due to the increasing power density and larger supply current compared to 2D-ICs. As an important part of 3D-IC PDNs, Power/Ground (P/G) Through-Silicon-Vias (TSVs) should be well-managed. Excessive or ill-placed P/G TSVs impact the power integrity (e.g. IR-drop), and also consume a considerable amount of chip real estate. In this work, we propose a Mixed-Integer-Linear-Programming (MILP)-based technique to plan the P/G TSVs. The goal of our approach is to minimize the average IR-drop while satisfying the total area constraint of TSVs by optimizing the P/G TSV placement. Therefore, the locations, sizes and the total number of the P/G TSVs are co-optimized simultaneously. The experimental results show that the average IR-drop can be reduced by 11.8% in average using the proposed method compared to a random placement technique with a much smaller runtime. |

| 16:00| IP1- 15, 69 | PACKAGE GEOMETRIC AWARE THERMAL ANALYSIS BY INFRARED-RADIATION THERMAL IMAGES | Speakers: Jui-Hung Chien\(^1\), Hao Yuz, Ruei-Siang Hsuy, Hsueh-Ju Linn and Shih-Chieh Chang

\(^1\)Industrial Technology Research Institute, TW; \(^2\)Thorne, TW; \(^3\)NTHU, TW

Abstract
Since packages affect the amount of heat transfer, it is important to include package and heat sink in thermal analysis. In this paper, we study the full-chip thermal response with different packages. We first discuss the difficulties of obtaining accurate package models for simulation. To facilitate a designer to perform thermal simulation with different packages, we propose to use a matrix called the package-transfer matrix which can transform a temperature profile of one package to another temperature profile of the desired package. To estimate and verify a package-transfer matrix, we propose an efficient method which uses Infrared Radiation (IR) images from two carefully design test chips with PBGA packages. Our experimental results show that the default package model CBGA in HotSpot can be accurately transferred to any other package through the package-transfer matrix.
design a cost-effective real-time system with improved worst-case latency at reduced area and power consumption. The next paper proposes bus designs for multi-cores that are analyzable for efficient real-time systems.

**3.5 Robust Architectures**

**Date:** Tuesday, March 25, 2014

**Location / Room:** Konferenz 3

**Chair:** Todd Austin, University of Michigan, US; Contact Todd Austin

**Co-Chair:** Muhammad Shafique, Karlsruhe Institute of Technology, DE; Contact

This session presents the design of novel architectures to support real-time and secure systems. The first paper couples a time-division multiplexed NoC with a real-time memory controller to design a cost-effective real-time system with improved worst-case latency at reduced area and power consumption. The next paper proposes bus designs for multi-cores that are analyzable for optimizing different operation frequencies.

**16:02 IP1-17**

**Presentation Title:** CHARACTERIZING POWER DELIVERY SYSTEMS WITH ON/OFF-CHIP VOLTAGE REGULATORS FOR MANY-CORE PROCESSORS

**Authors:** Xuan Wang, Jiang Xu, Zhe Wang, Kevin J. Chen, Xiaowen Wu and Zhehui Wang, HKUST, HK

**Abstract**

Design of power delivery system has great influence on the power management in many-core processor systems. Moving voltage regulators from off-chip to on-chip gains more and more interest in the power delivery system design, because it is able to provide fast voltage scaling and multiple power domains. Previous works are proposed to implement power efficient on-chip regulators. It is also important to analyze the characteristics of the entire power delivery system to explore the tradeoff between the promising properties and costs of employing on-chip regulators. In this work, we develop an analytical model to evaluate important characteristics of the power delivery system, including on-chip/off-chip voltage regulators and the passive on-chip/on-board parasitic. Compared with SPICE simulations, our model achieves a fast system-level evaluation with comparable accuracy. Based on the model, geometric programming is utilized to find the optimal power efficiency of different architectures of power delivery systems under constraints of output voltage stability and area. Experiments show that compared with the conventional architecture using off-chip regulators, the hybrid one using both on-chip and off-chip voltage regulators achieves 1.0% power efficiency improvement and 68% area reduction of voltage regulators on average. We conclude that the hybrid architecture has potential for high power efficiency and small area at heavy workload, but careful account for the overhead of on-chip regulators is needed.

**16:03 IP1-18**

**Presentation Title:** MASK-COST-AWARE ECO ROUTING

**Authors:** Hsi-An Chien, Zhen-Yu Peng, Yun-Ru Wuz, Ting-Hsung Wangz, Hsin-Chang Linz, Chi-Feng Wuza and Ting-Chi Wang1

**Abstract**

In this paper, we study a mask-cost-aware routing problem for engineering change order (ECO). By taking into account old routes for possible reuse, we present an approach for the problem. Encouraging experimental results are reported to demonstrate the effectiveness of our approach.

**16:00 End of session**

**Coffee Break in Exhibition Area**

On Tuesday-Thursday the coffee and lunch breaks will be located in the Exhibition Area (Terrace Level).

**3.5 Robust Architectures**

**Date:** Tuesday, March 25, 2014

**Time:** 14:30 - 16:00

**Location / Room:** Konferenz 3

**Chair:** Todd Austin, University of Michigan, US; Contact Todd Austin

**Co-Chair:** Muhammad Shafique, Karlsruhe Institute of Technology, DE; Contact

This session presents the design of novel architectures to support real-time and secure systems. The first paper couples a time-division multiplexed NoC with a real-time memory controller to design a cost-effective real-time system with improved worst-case latency at reduced area and power consumption. The next paper proposes bus designs for multi-cores that are analyzable for probabilistic timing analysis. The final paper in this session designs a lightweight hardware solution using lockstep shadow thread execution to detect and prevent code injection attacks.

**14:30 3.5.1**

**Presentation Title:** COUPLING TDM NOC AND DRAM CONTROLLER FOR COST AND PERFORMANCE OPTIMIZATION OF REAL-TIME SYSTEMS

**Authors:** Manil Dev Gomony, Benny Axesssonz and Kees Goossens

**Abstract**

Existing memory subsystems and TDM NoCs for real-time systems are optimized independently in terms of cost and performance by configuring their arbiters according to the bandwidth and/or latency requirements of their clients. However, when they are used in conjunction, and run in different clock domains, i.e. they are decoupled, there exists no structured methodology to select the NoC interface width and operating frequency for minimizing area and/or power consumption. Moreover, the multiple arbitration points, one in the NoC and the other in the memory subsystem, introduce additional overhead in the worst-case guaranteed latency. These makes it hard to design cost-efficient real-time systems. The three main contributions in this paper are: (1) We present a novel methodology to couple any existing TDM NoC with a real-time memory controller and compute the different NoC interface width and operating frequency combinations for minimal area and/or power consumption. (2) For two different TDM NoC types, one a packet-switched and the other circuit-switched, we show the tradeoff between area and power consumption with the different NoC configurations, for different DRAM generations. (3) We compare the coupled and decoupled architectures with the two NoCs, in terms of guaranteed worst-case latency, area and power consumption by synthesizing the designs in 40 nm technology. Our experiments show that using a coupled architecture in a system consisting of 16 clients results in savings of over 44% guaranteed latency, 18% and 17% in area, 19% and 11% in power consumption for a packet-switched and a circuit-switched TDM NoC, respectively, with different DRAM types.
3.5.2 Bus Designs for Time-Probabilistic Multicore Processors

Javier Jalle, Leonidas Kosmbris, Jaume Abellàs, Eduardo Quinones and Francisco Cazorla

Abstract

Probabilistic Timing Analysis (PTA) reduces the amount of information needed to provide tight WCET estimates in real-time systems with respect to classic timing analysis. PTA imposes new requirements on hardware design that have been shown implementable for single-core architectures. However, no support has been proposed for multicore so far. In this paper, we propose several probabilistically-analyzable bus designs for multicore processors ranging from 4 cores connected with a single bus, to 16 cores deploying a hierarchical bus design. We derive analytical models of the probabilistic timing behaviour for the different bus designs, show their suitability for PTA and evaluate their hardware cost. Our results show that the proposed bus designs (i) fulfill PTA requirements, (ii) allow deriving WCET estimates with the same cost and complexity as in single-core processors, and (iii) provide higher guaranteed performance than single-core processors, 3.4x and 6.6x on average for an 8-core and a 16-core setup respectively.

15:00 - 15:30

3.5.3 Programmable Decoder and Shadow Threads: Tolerate Remote Code Injection Exploits with Diversified Redundancy

Wenzhen Shi, Ziyi Liu, Shouhual Xu and Zhiqiang Lin

University of Houston, US; University of Texas at San Antonio, US; University of Texas at Dallas, US

Abstract

We present a lightweight hardware framework for providing assurance detection and prevention of code injection attacks using a lockstep diversified shadow execution. Recent studies show that hardware diversification can detect software attacks by checking the consistency of their behavior simultaneously. Unfortunately, the severe performance degradation and extra system costs caused by these methods are unacceptable in many applications. This paper presents a hardware-level, lockstep shadow thread framework to enrich the diversity of the software execution, with the facilitation from programmable hardware decoder and novel CPU support of tightly coupled non-executing shadow thread technique. Specifically, given a piece of (legacy) binary code, we first generate diversified binary versions using an offline binary rewriter and programmable hardware binary translator at runtime. Two diversified binary code images are launched as dual simultaneous threads in the hardware layer with one as the primary thread and the other as shadow thread. Instructions from the shadow thread are not executed but just compared, and thus incur no OS side-effects. The extended CPU is able to decode instructions from both threads, and dispatch them to next stage pipeline for a lockstep comparison. Any mismatch of the decoded instructions from the two threads caused by remotely injected binary code will be detected. Our design provides instruction set randomization (ISR) with minimal cost in performance, when compared with straightforward ISR implementation. The simulation results indicate that our framework incurs very small overheads and provides a protection against code injection attacks.

16:00 - 16:30

16:00 - 16:30

IP1-19

Partial-Set: Write Speedup of PCM Main Memory

Li Bing, Shan Shuchang, Hu Yu and Li Xiaowei

ICT, CAS, CN; ICT, CAS, CN; ICT, CAS, CN

Abstract

Phase change memory (PCM) is a promising nonvolatile memory technology developed as a possible DRAM replacement. Although it offers the read latency close to that of DRAM, PCM generally suffers from the long write latency. Long write request may block the read requests on the critical path of cache/memory access, incurring adverse impact on the system performance. Besides, the write performance of PCM is very asymmetric, i.e., the SET operation (writing '1') is much slower than that of the RESET operation (writing '0'). In this work, we re-examine the resistance transform process during the SET operation of PCM and propose a novel Partial-SET scheme to alleviate the long write latency issue of PCM. During a write access to a memory line, a short Partial-SET pulse is applied first to program the PCM cells to a pre-stable state, then a long Partial-RESET pulse is applied to set the PCM cells to the desired state. The extended CPU is able to decode instructions from both threads, and dispatch them to next stage pipeline for a lockstep comparison. Any mismatch of the decoded instructions from the two threads caused by remotely injected binary code will be detected. Our design provides instruction set randomization (ISR) with minimal cost in performance, when compared with straightforward ISR implementation. The simulation results indicate that our framework incurs very small overheads and provides a protection against code injection attacks.

16:30 - 16:30

IP1-20

Exploiting Narrow-Width Values for Improving Non-Volatile Cache Lifetime

Guangshan Duan and Shuai Wang, Nanjing University, CN

Abstract

Due to the high cell density, low leakage power consumption, and less vulnerability to soft errors, the non-volatile memory technologies are among the most promising alternatives for replacing the traditional DRAM and SRAM technologies used in implementing main memory and caches in the modern microprocessor. However, one of the difficulties is the limited write endurance of most non-volatile memory technologies. In this paper, we propose to exploit the narrow-width values to improve the lifetime of the non-volatile last level caches. Leading zeros masking scheme is first proposed to reduce the write stress to the upper half of the narrow-width data. To balance the write variations between the upper half and the lower half of the narrow-width data, two swap schemes, the swap on write (SW) and swap on replacement (SRepl), are proposed. To further reduce the write stress to the non-volatile cache, we adopt two optimization schemes, the multiple dirty bit (MDB) and read before write (RBW), to improve its lifetime. Our experimental results show that by combining all our proposed schemes, the lifetime of the non-volatile caches can be improved by 245% on average.

16:30 - 16:30

End of session

3.6 Cyber Physical Systems: Security and Co-design

Date: Tuesday, March 25, 2014
Time: 14:30 - 16:00
Location / Room: Konferen 4
Chair: Rolf Ernst, Technische Unversitat Braunschweig, DE, Contact Rolf Ernst
Co-Chair: Anuradha Annaswamy, MIT, US, Contact Anuradha Annaswamy

This session showcases recent results in cybersecurity and codesign in CPS. The first paper analyzes a stealth cyberattack scenario where a distributed sensor system is disturbed by an attacker who tries to reduce the sensor fusion quality and suggests an algorithmic approach to increase robustness against this attack. The second paper addresses the joint design of a feedback controller and a server-based resource reservation mechanism to guarantee closed-loop stability. The third paper describes a codesign approach formally guaranteeing control robustness for a communication channel with a bounded number of frame losses.
This work considers the problem of attack-resilient sensor fusion in an autonomous system where multiple sensors measure the same physical variable. A malicious attacker may corrupt a subset of these sensors and send wrong measurements to the controller on their behalf, potentially compromising the safety of the system. We formalize the goals and constraints of such an attacker who also wants to avoid detection by the system. We argue that the attacker's capabilities depend on the amount of information she has about the correct sensors' measurements. In the presence of a shared bus where messages are broadcast to all components connected to the network, the attacker may consider all other measurements before sending her own in order to achieve maximal impact. Consequently, we investigate effects of communication schedules on sensor fusion performance. We provide worst- and average-case results in support of the Ascending schedule, where sensors send their measurements in a fixed succession based on their precision, starting from the most precise sensors. Finally, we provide a case study to illustrate the use of this approach.

Speakers:
Radoslav Ivanov, Miroslav Pajo and Insup Lee, University of Pennsylvania, US

Abstract
We deal with synthesis of distributed embedded control systems closed over a faulty or severely constrained communication network. Such overloaded communication networks are common in cost-sensitive domains such as automotive. Design of such systems aims to meet all deadlines following the traditional notion of schedulability. In this work, we aim to exploit robustness of the controller and propose a novel implementation approach to achieve a tighter design. Toward this, we answer two research questions: (i) given a distributed architecture, how to characterize and formally verify the bound on deadline misses, (ii) given such a bound, how to design a controller such that desired stability and Quality of Control (QoC) requirements are met. We address question (i) by modeling a distributed embedded architecture as a network of Event Count Automata (ECA), and subsequently introducing and formally verifying a property formulation with reduced complexity. We address question (ii) by introducing a novel fault-tolerant control strategy which adjusts the control input at runtime based on the occurrence of fault or drop. We show that QoC under faulty communication improves significantly using the proposed fault-tolerant strategy.

Speakers:
Matthias Kauer, Damoon Soudabehzahi, Dg Goswami, Samarth Chakrabarty and Anuradha Annaswamy

Abstract
In this paper, we study the important performance issues in using the purging-range query to reclaim old data versions to be free blocks in a flash-based multi-version database. To reduce the overheads for using the purging-range query in garbage collection, the physical block labeling (PBL) scheme is proposed to provide a better estimation on the purging version number to be used for purging old data versions. With the use of the frequency-based placement (FBP) scheme to place data versions in a block, the efficiency in garbage collection can be further enhanced by increasing the deadspans of data versions and reducing reallocation cost especially when the sizes of the flash memory for the databases are limited.

Speakers:
Kam-Yiu Lam1, Jian-Tao Wang1, Yuan-Hao Chang1, Jen-Wei Hsieh1, Po-Chun Huang1, Chung Keung Poorn1 and Chun Jiang Zhu1

Abstract
The next generation of automobiles (also known as cybercars) will increasingly incorporate electronic control units (ECUs) in novel automotive control applications. Recent work has demonstrated vulnerability of modern car control systems to security attacks that directly impact the cybercar's physical safety and dependability. In this paper, we provide an integrated approach for the design of secure and dependable cybercars using a case study: a steer-by-wire (SBW) application over controller area network (CAN). The challenge is to embed both security and dependability over CAN while ensuring that the real-time constraints of the cybercar applications are not violated. Our approach enables early design feasibility analysis by embedding essential security primitives (i.e., confidentiality, integrity, and authentication) over CAN subject to the real-time constraints imposed by the desired quality of service and behavioral reliability. Our method leverages multi-core ECUs for providing fault-tolerance by redundant multi-threading (RMT) and also further enhances RMT for quick error detection. We quantify the error resilience of our approach and evaluate the interplay of performance, fault-tolerance, security, and scalability for our SBW case study.

Speakers:
Pierluigi Nuzzo, John Finn, Antonio Iannopollo and Alberto Sangiovanni-Vincentelli, University of California at Berkeley, US

Abstract
We introduce a platform-based design methodology that addresses the complexity and heterogeneity of cyber-physical systems by using assurance guarantees to formalize the design process and enable realization of control protocols in a hierarchical and compositional manner. Given the architecture of the physical plant to be controlled, the design is carried out as a sequence of refinement steps from an initial specification to a final implementation, including synthesis from requirements and mapping of higher-level functional and non-functional models into a set of candidate solutions built out of a library of components at the lower level. Initial top-level requirements are captured as contracts and expressed using linear temporal logic (LTL) and signal temporal logic (STL) formulas to enable requirement analysis and early detection of inconsistencies. Requirements are then refined into a controller architecture by combining reactive synthesis steps from LTL specifications with simulation-based design space exploration steps. We demonstrate our approach on the design of embedded controllers for aircraft electric power distribution.
### Session Overview

**Title:** Hot Topic: Mission Profile Aware Design - The Solution for Successful Design of Tomorrow's Automotive Electronics

#### 3.7 On line Strategies for Reliability

**Date:** Tuesday, March 25, 2014  
**Time:** 14:30 - 16:00  
**Location / Room:** Konferenz 5

**Chair:** Fabrizio Lombardi, Northwestern University, US  
**Co-Chair:** Jie Han, University of Alberta, CA

This section presents different approaches to improve reliability of circuits and systems by using on-line techniques. It shows different methods that can be applied to caches, processors and multicore architectures.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>14:30</td>
<td>3.7.1</td>
<td>SPATIAL PATTERN PREDICTION BASED MANAGEMENT OF FAULTY DATA CACHES</td>
</tr>
<tr>
<td>15:00</td>
<td>3.7.2</td>
<td>COMBINED DVFS AND MAPPING EXPLORATION FOR LIFETIME AND SOFT-ERROR SUSCEPTIBILITY IMPROVEMENT IN MPSoCs</td>
</tr>
<tr>
<td>15:30</td>
<td>3.7.3</td>
<td>DARP: DYNAMICALLY ADAPTABLE RESILIENT PIPELINE DESIGN IN MICROPROCESSORS</td>
</tr>
<tr>
<td>16:00</td>
<td>IP1-24, 45</td>
<td>A FAULT DETECTION MECHANISM IN A DATA-FLOW SCHEDULED MULTITHREADED PROCESSOR</td>
</tr>
</tbody>
</table>

**Authors**

- 1National University of Singapore, SG; 2Politecnico di Milano, IT
- 1University of Amsterdam, NL; 2National University of Defense Technology, CN

### Presentations

**3.7.1 SPATIAL PATTERN PREDICTION BASED MANAGEMENT OF FAULTY DATA CACHES**  
**Speakers:** Georgios Keramidas1, Michail Mavropoulos1, Anna Karounia1 and Dimitris Nikolos2  
1Researcher, University of Patras, GR; 2University of Patras, GR

**Abstract**  
Technology scaling leads to significant faulty bit rates in on-chip caches. In this work, we propose a methodology to mitigate the impact of defective bits (due to permanent faults) in first-level set-associative data caches. Our technique assumes that faulty caches are enhanced with the ability of disabling their defective parts at cache subblock granularity. Our experimental findings reveal that while the occurrence of hard-errors in faulty caches may have a significant impact in performance, a lot of room for improvement exists, if someone is able to take into account the spatial reuse patterns of the to-be-referenced blocks (not all the data fetched into the cache is accessed).

To this end, we propose frugal PC-indexed spatial predictors (with very small storage requirements) to orchestrate the (re)placement decisions among the fully and partially unusable faulty blocks. Using cycle-accurate simulations, a wide range of scientific applications, and a plethora of cache fault maps, we showcase that our approach is able to offer significant benefits in cache performance.

**3.7.2 COMBINED DVFS AND MAPPING EXPLORATION FOR LIFETIME AND SOFT-ERROR SUSCEPTIBILITY IMPROVEMENT IN MPSoCs**  
**Speakers:** Anup Das1, Akash Kumar1, Bharadwaj Veeravalli1, Cristina Bolchini2 and Antonio Mele2  
1National University of Singapore, SG; 2Politecnico di Milano, IT

**Abstract**  
Energy and reliability optimization are two of the most critical objectives for the synthesis of multiprocessor system-on-chip (MPSoCs). Task mapping has shown significant promise as a low cost solution in achieving these objectives as standalone or in tandem as well. This paper proposes a multi-objective design space exploration to determine the mapping of tasks to a system on a multiprocessor architecture and voltage/frequency level of each task exploiting the DVFS capabilities of modern processors such that the reliability of the platform is improved while fulfilling the energy budget and the performance constraint set by system designers. In this respect, the reliability of a given MPSoC platform incorporates not only the impact of voltage and frequency on the aging of the processors (wear-out effect) but also on the susceptibility to soft-errors -- a joint consideration missing in all existing works in this domain. Further, the proposed exploration also incorporates soft-error tolerance by selective replication of tasks, making the proposed approach an interesting blend of reactive and proactive fault-tolerance. The combined objective of minimizing core aging together with the susceptibility to transient faults under a given performance-energy budget is solved by using a multi-objective genetic algorithm exploiting tasks' mapping, DVFS and selective replication as tuning knobs. Experiments conducted with real-life and synthetic application graphs clearly demonstrate the advantage of the proposed approach.

**3.7.3 DARP: DYNAMICALLY ADAPTABLE RESILIENT PIPELINE DESIGN IN MICROPROCESSORS**  
**Speakers:** Hu Chen, Sanghamitra Roy and Koushik Chakraborty, Utah State University, US

**Abstract**  
In this paper, we demonstrate that the sensitized path delays in various microprocessor pipe stages exhibit intriguing temporal and spatial variations during the execution of real world applications. To effectively exploit these delay variations, we propose Dynamically Adaptable Resilient Pipeline (DARP) -- a series of runtime techniques to boost power performance efficiency and fault tolerance in a pipelined microprocessor. DARP employs early error prediction to avoid a major portion of timing errors. Using a rigorous circuit-architectural infrastructure, we demonstrate substantial improvements in the performance (8.4-20%) and energy efficiency (6.4-27.9%), compared to state-of-the-art techniques.

**End of session**

*On Tuesday-Thursday the coffee and lunch breaks will be located in the Exhibition Area (Terrace Level).*
Additionally, we introduce the mission profile aware design flow and present several application scenarios. 

In order to benefit from modern automotive semiconductor technologies, application robustness must now be considered as a design target. This includes the consequent consideration of environmental stress conditions and functional loads, which are formalized in so-called “mission profiles”. We introduce the motivation to use mission profiles from an OEM and Tier n perspective. Additionally, we introduce the mission profile aware design flow and present several application scenarios.
### UB03.01 LARA: THE LARA COMPILER SUITE

**Authors:**
João Bispo, Pedro Pinto, Ricardo Nobre, Tiago Canhalho and João Cardoso, Universidade do Porto, PT

**Abstract**
LARA is an aspect-oriented programming (AOP) language which allows the description of sophisticated code instrumentation schemes, advanced mapping strategies including conditional decisions, based on hardware/software resources, and of sophisticated sequences of compiler transformations. Furthermore, LARA provides mechanisms for controlling all elements of its toolchain in a consistent and systematic way, using a unified programming interface. We present three compiler tools developed around the LARA technology: MATISS, MANET and ReflectC. MATISS is a compiler which 1) allows analyses and transformations on MATLAB code and 2) generates C code from the MATLAB code. MATISS can be fully controlled through LARA aspects, which can define the type and shape of MATLAB variables, specify code insertion/removal actions, and define specialization directives and other additional information. MATISS can output transformed MATLAB code and specialized C code. The knowledge provided by the LARA aspects allows MATISS to generate C tailored to specific targets (e.g., use statically declared arrays to be compliant with the high-level synthesis tools such as Catapult C). MANET is a source-to-source compiler for ANSI C based on Cetus, and is controlled using LARA aspects. MANET manages to leverage the expressiveness and modularity of LARA to query and manipulate the Cetus AST, providing an easy flow for code instrumentation and code transformations. LARA aspects allow for a simple selection of program elements in the code which can be analyzed or transformed, by either considering their attributes or applying actions. Thus, MANET can be used to provide information reports on compiler analyses, to implement sophisticated code instrumentation strategies, or to perform code optimizations and transformations. ReflectC is a compiler based on CoSy’s compiler framework. CoSy’s configurability and retargetability make ReflectC particularly effective for exploration of compiler transformations and optimizations on possible architecture variations, and it is being used for hardware/software co-design and design space exploration (DSE). We will present demos of the tools and the use of LARA aspects and strategies to guide our suite of compilation tools providing: 1) C code generation from MATLAB code, according to information provided by LARA aspects; 2) Instrumentation of C code to be used for collecting specific compile and runtime information (e.g., execution time, range of values for specific variables, custom profiling); 3) User-controlled compiler optimizations targeting several architectures and DSE of sequences of compiler optimizations in mind performance improvements. In addition to presenting examples for each of the tools of the LARA compilation suite, we show an execution of the complete toolchain, controlled by LARA aspects.

More information ...

### UB03.02 AN AUTOMATED DESIGN FLOW FOR FAST PROTOTYPING OF SIMULINK MODELS ONTO MPSOC

**Authors:**
Francesco Rebino and Johnny Öberg, Royal Institute of Technology, SE

**Abstract**
Simulink is a modelling environment suitable to model embedded systems at system-level. However there is no standard to rapidly prototype Simulink models onto modern multiprocessor system-on-chip (MPSoC). In this demonstration we show how our NoC System Generator tool can be used as part of an automated platform-based design flow to synthesize a Simulink model to a network-on-chip based MPSoC implementation on FPGA. The performance of the generated prototype scales with the number of processors.

More information ...

### UB03.03 PATN: A PERFORMANCE ANALYSIS TOOL FOR NOC

**Authors:**
Yang Chen and Zhonghai Lu, KTH Royal Institute of Technology, SE

**Abstract**
With processors increased onto a single chip, and more and more time sensitive applications added to on-chip systems, performance bound analysis becomes essential for QoS Network-on-Chip (NoC) designs and evaluations. For the purpose of providing the reliable and automated analysis for QoS NoC, we propose PATN (Performance Analysis Tool for NoC), which automatically computes the end-to-end delay bounds of data flows, and backlog bounds of buffers for NoC with arbitrary topology. PATN is designed based on network calculus, which lies on solid mathematical foundations and provides well-guaranteed accuracy of the results. Network Calculus based analysis has been successfully employed for various communications networks, such as SpaceWire, AFDX, etc. For example, Airbus adopted and approved the network calculus based analysis for certification on its aircraft A380. In this demonstration, we give a whole view of PATN through two segments. First, we explain the architecture and main functions; show the working flow and printing log by analysing end-to-end delay bound of a data flow in a simple network. The log shows that the analysis follows the theoretical methodology exactly, hence to obtain the correct and tight results, which as good as that the theory can achieve. Second, we use PATN to analyse the delay bounds and backlog bounds for 3 NoCs with different topologies – binary tree, mesh, and hierarchical topology of binary tree and mesh. The analyses demonstrate computation speed and scalability of PATN. Moreover, comparisons of the delay bound, computed with different configuration parameters of the flows and routers, are conducted. It shows how the delay bound is affected by the parameters.

More information ...

### UB03.04 COMPILER FOR MAPPING STREAM PROCESSING APPLICATIONS ONTO REAL-TIME HETEROGENEOUS MULTIPROCESSOR SYSTEMS

**Authors:**
Stefan Geuns, Berend Dekens, Philip Wilmanns, Joost Hausmans, Guus Kuiper and Marco Bekooij, University of Twente, NL

**Abstract**
Heterogeneous multiprocessors systems are employed for power-efficiency reasons in wearable software defined radios. These systems are hardware cost-effective and deliver a superior performance compared to their homogeneous counterparts. However these systems are notoriously hard to program without tool support, which makes it desirable that programming is simplified with the help of an optimizing multiprocessor compiler for stream processing applications. This demonstration shows how our multiprocessor compiler for mapping real-time stream processing applications onto our real-time heterogeneous multi-core system. The applications are described as sequential programs and are compiled into parallel task graphs. Buffer capacities are computed using dataflow analysis techniques given the real-time constraints of the application. Our multi-core system contains 16 MicroBlaze processor cores as well as two hardware accelerators and is prototyped on a Xilinx Virtex-6 FPGA. A connection-less communication ring is used for inter-processor communication. Our system is equipped with an analog RF front-end, which enables us to demonstrate PAL-video reception and decoding.

More information ...

### UB03.05 HIDE BLUR: DESIGN OF A HIGH PERFORMANCE CORE FOR REMOVING BLUR EFFECT ON IMAGES

**Authors:**
Giovanni Alfonso, Giulio Gambardella, Marco Indaco, Paolo Prinetto, Daniele Rollo and Pascal Trotta, Politecnico di Torino, IT

**Abstract**
This work aims at developing a high performance FPGA-based IP-core able to perform a deblurring algorithm in real-time. Modern approaches to deblurring usually either only handle simple types of blur, or need heavy user interaction. Moreover, they usually require several minutes (or even whole hours) to process a single image. Our purpose is to study the current state-of-the-art and identify the best deblurring algorithms that are suitable for a hardware implementation. The selected algorithm is optimized and implemented in hardware in order to perform the deblurring task with highest possible performances.

More information ...
UB03.06 PHARON: PARALLEL AND HETEROGENEOUS ARCHITECTURES FOR REAL-TIME APPLICATIONS
Authors: Luciano Lavagnino, Mihai Lazarescu, Hector Posadas and Eugenio Villars
1Politecnico di Torino, IT; 2Universidad de Cantabria, ES
Abstract
In this demo, we will present the work-in-progress of the EU FP7 PHARON project, started in September 2011. The first objective of the project is the development of new techniques and tools capable to assist the designer in the development of parallel embedded systems, from executable specifications to target-specific implementation and debugging on a multiprocessor platform. This tool chain offers and implements several parallelization strategies, reflecting the functional and non-functional constraints of the system, and driving the designer into incremental parallelization and adaptation steps. The second objective of the project is to develop monitoring and control techniques in the middleware of the system capable to automatically adapt platform services to application requirements and therefore reduce power consumption transparently. The demo will cover specifically: - the software parallelization tool suite. - the parallel software modeling and code generation suite.
More information ...

UB03.07 COMPSOC: VIRTUAL EXECUTION PLATFORMS FOR MIXED TIME-CRITICALITY APPLICATIONS
Author: Kees Goossens, TU Eindhoven, NL
Abstract
System-on-Chip (SOC) design gets increasingly complex, as a growing number of applications are integrated in such systems. These applications have mixed time-criticality, i.e., some have firm, some soft, and others non-real-time requirements. Executing such a mix of applications on a SOC poses several challenges. First, to reduce cost, platform resources, e.g., processors, interconnect, memories, are shared between applications. However, sharing causes interference between applications, making their behaviors interdependent. This results in two problems for SOC design and verification: 1) accurate system-level simulation and several approaches to formal verification are infeasible, because of the explosion in the number of possible combinations of applications, inputs, and resource states and 2) verification becomes a circular process that must be repeated if an application is added, removed, or modified, making integration and verification of the SOC development, in terms of time and money. The CompSOC platform addresses these problems by executing each application on an independent virtual execution platform (VEP). The VEPs are composable, i.e., cannot affect each other's behaviors. In the temporal domain an applications actual execution never varies by even a single clock cycle. Similarly, the energy and power behaviors of applications are also composable. As a result, applications can be designed, developed, verified, and executed in isolation. The VEPs are also predictable, meaning that all interference is bounded. This makes them virtualized also in terms of performance bounds, which enables firm real-time applications to be verified using formal performance analysis frameworks. The CompSOC platform uses the CoMK microkemel to implement virtual processors on each processor time through temporal partitioning. Each application can use its own operating system (e.g., Compose, µCOS-III) and model of computation (e.g., CSDL, KPN, TT) in its VEP, to suit its level of time critically. As more applications are integrated on a single SOC, the need arises for more dynamic behaviour. The system should be able to start, modify and stop applications at run time without affecting running applications. For this purpose the CompSOC platform has been extended with a predictable and composable resource management framework. It manages application bundles that contain 1) an application in the form of executable files (ELFs on multiple processors), and also 2) the specifications of the (one or more) particular VEPs that the application executes in, consisting of virtual processors, NOC connections, virtualized mem-ories, etc. At run time, the resource management framework can dynamically load and start application bundles by creating a VEP and then loading, booting, and executing an application within it. The VEPs can also be modified, stopped, and deleted at run time. Our University Booth will present virtual execution platform and application bundle concepts using an interactive demonstrator. It will show that the CompSOC has been extended with dynamic functionality, without sacrificing its key strengths: composability and predictability. We will demonstrate this through the use of the resource management framework and application bundles, showing that we can create, modify and delete virtual execution platforms running a mixed time-critcality application dynamically at run-time.
More information ...

UB03.08 A HOLISTIC APPROACH TO POWER MANAGEMENT FOR ENERGY HARVESTING EMBEDDED SYSTEMS
Authors: Kyungsoo Lee, Hideki Takase and Tohru Ishihara, Kyoto University, JP
Abstract
We present a holistic approach to maximizing the energy efficiency of energy harvesting embedded systems which consist of a processor system and an energy harvesting system. A power management program integrated on a real-time OS optimally switches operation mode of the processor and configuration of the energy harvesting system according to the workload of the processor and harvesting situation. The demonstration will show that our prototype system consisting of our processor chip and harvesting system board stably runs using harvested energy only. The processor has multiple cores having a different performance in each to improve the energy efficiency of computation. The energy harvesting board has high transferring efficiency to reduce the power loss. The entire system is controlled efficiently by our power management program implemented on Toppers OS.
More information ...

UB03.09 FAULTIFY: PROBABILISTIC CIRCUIT FAULT EMULATION
Authors: David May and Walter Stechele, TUM, DE
Abstract
We want to demonstrate an FPGA-based probability-aware fault emulator and its corresponding algorithms in the context of a real-time H.264 decoder. The demo will show that reliability constraints can be relaxed inside the circuit without noticeable degradation of the image quality when carefully investigating where the constraints can be relaxed. We will show how this investigation can be done using our emulator and we will show the effect of a relaxed robustness of the circuit in real-time.
More information ...

UB03.10 RTL+: DESIGN ENVIRONMENT: WALK BEFORE YOU RUN.
Authors: Somayeh Sadeghi-Kohan, Behnaz Pourmohseni, Amir Reza Nekooei, Hanie Hashemi, Hamed Najafi Haghi and Zainalabedin Navabi, University of Tehran, IR
Abstract
To enable development of high level designs with hardware correspondence, synthesizability must be satisfied in a top-down manner. Thus in this work, instead of using TLM2.0 which is not established for synthesis, we will start with a level above RT level, "RTL+. "RTL+ is basically using TLM1.0 channels and includes abstract communications and handshakings that are mainly hidden from the designer. We develop a package of SystemC channels with hardware correspondence (synthesizable HDL) for the communication between various cores (with simple interfaces) and standard buses.
More information ...

17:30 End of session
18:30 Exhibition Reception In Several serving points inside the Exhibition Area (Terrace Level)
The Exhibition Reception will take place in the exhibition area (Terrace Level). All exhibitors are welcome to provide drinks and snacks for delegates and visitors.
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>
| IP1-1 | SAFE: SECURITY-AWARE FLEXRAY SCHEDULING ENGINE | Gang Han¹, Haili Deng¹, Yanping Li and Wenhua Dou¹  
¹National University of Defense Technology, CN; ²McGill University, CA; ³The Chinese University of Hong Kong, CN |
| Abstract | In this paper, we propose SAFE (Security Aware FlexRay scheduling Engine), to provide a problem definition and a design framework for FlexRay static segment schedule to address the new challenge on security. From a high level specification of the application, the architecture and communication middleware are synthesized to satisfy security requirements, in addition to extensibility, costs, and end-to-end latencies. The proposed design process is applied to two industrial case studies consisting of a set of active safety functions and an X-by-wire system respectively. |
| IP1-2 | TRANSIENT ERRORS RESILIENCY ANALYSIS TECHNIQUE FOR AUTOMOTIVE SAFETY CRITICAL APPLICATIONS | Sujan Pandey and Bart Vermeulen, NXP Semiconductors, NL |
| Abstract | When a single bit is flipped as a result of a transient error in an electronic circuit, its effect can have a severe impact if the circuit is deployed in safety critical domains such as automotive, aeronautics, and industrial automation. In the design phase it is therefore essential to evaluate, and where necessary improve, the resilience of a circuit to all possible transient errors. In this paper, we present a method to analyze the transient error resiliency of a digital circuit. This method is based on an analytical model. It models a transient error as a random function and finds the vulnerable number of bits for each node. We perform a case study on a circuit implementation of a well-known adaptive filter algorithm. The results from the analytical and simulation models show that the analytical model is accurate enough to estimate the effects of transient errors on the performance of a digital circuit. Our analytical method also reduces the run time significantly in a design phase. |
| IP1-3 | MODEL BASED HIERARCHICAL OPTIMIZATION STRATEGIES FOR ANALOG DESIGN AUTOMATION | Engin Afacan¹, Gunhan Dundar¹, Falk Baskaya¹, Simge Ay¹ and Francisco Fernandez²  
¹Bogazici University, TR; ²Universidad de Sevilla, TR |
| Abstract | The design of complex analog circuits by using flat optimization-based approaches is inefficient, even impossible, due to the high number of design variables and the growth of the cost of performance evaluation with the circuit size. Over the past two decades, top-down hierarchical design approaches have been developed and applied. They are based on hierarchical circuit decomposition and specification transmission from top-level to lower level blocks. However, such specification transmission is usually performed with little knowledge on the feasibility of the specifications, leading, therefore, to costly redesign iterations. Even if the specification transmission is successful, there is no guarantee that it is optimal in terms of e.g., power consumption or area occupation. To palliate this problem, two novel model-based hierarchical synthesis methods are proposed in this paper: Model-Based Hierarchical Optimization (MBHO) and Improved Model-Based Hierarchical Optimization (IMBHO). They are based on the concurrent design at higher and lower hierarchical levels and appropriate communication between the different processes. Experimental results on a filter example comparing the new approaches and the conventional top-down design approach are provided. |
| IP1-4 | A NOVEL LOW POWER 11-BIT HYBRID ADC USING FLASH AND DELAY LINE ARCHITECTURES | Hsun-Cheng Lee and Jacob Abraham, the University of Texas at Austin, US |
| Abstract | This paper presents a novel low power 11-bit hybrid ADC using flash and delay line architectures, where a 4-bit flash ADC is followed by a 7-bit delay-line ADC. This hybrid ADC inherits accuracy and power efficiency from flash ADCs and delay-line ADCs, respectively. Also, in order to reduce the power of the first stage flash ADC, a power-saving technique is adopted by biasing the DC tail current of the pre-amplifier at 5µA instead of the operational current, 47µA in stand-by mode. The hybrid ADC was designed and simulated in a commercial 65nm process. With a 1.1 V supply and 100 Ms/s, the ADC achieves an SNDR of 60 dB and consumes 1.6 mW, which results in a figure of merit (FOM) of 19.4 Uj/conversion-step without any calibration technique. Also, Monte Carlo simulations are performed with a 3o device mismatch for the SNDR estimation, and the SNDR is observed to be better than 58.5 dB. |
| IP1-5 | SEMI-SYMBOLIC ANALYSIS OF MIXED-SIGNAL SYSTEMS INCLUDING DISCONTINUITIES | Cama Radojicic, Christoph Grimm, Javier Moreno and Xiao Pan, TU Kaiserslautern, DE |
| Abstract | The paper describes an approach for semi-symbolic analysis of mixed-signal systems that contain discontinuous functions, e.g. due to modeling comparators. For modeling and semi-symbolic simulation, we use extended Affine Arithmetic. Affine Arithmetic is currently limited to accurate analysis of linear functions and mild non-linear functions, but not yet discontinuities. In this paper we extend the approach to also handle discontinuities. For demonstration, we symbolically analyze a ΣΔ-modulator. |
| IP1-6 | NOVEL CIRCUIT TOPOLOGY SYNTHESIS METHOD USING CIRCUIT FEATURE MINING AND SYMBOLIC COMPARISON | Cristian Ferent and Alex Doboli, Stony Brook University, US |
| Abstract | This paper presents a reasoning-based approach to analog circuit synthesis using ordered node clustering representations (ONCR) to describe alternative circuit features and symbolic circuit comparison to characterize performance trade-offs of synthesized solutions. Case studies illustrate application of the proposed methods to topology selection and refinement. |
IP1-7

AN EMBEDDED OFFSET AND GAIN INSTRUMENT FOR OPAMP IPS

Speakers: Jinbo Wan and Hans Kerkhoff, CAES-TDT, CITT, University of Twente, NL

Abstract
Analog and mixed-signal IPs are increasingly required to use digital fabrication technologies and are deeply embedded into system-on-chips (SoC). These developments append more requirements and challenges on analog testing methodologies. Traditional analog testing methods suffer from less accessibility and control with regard to these embedded analog circuits in SoCs. As an alternative, an embedded instrument for analog OpAmp IP tests is proposed in this paper. It can provide the exact gain and offset values of OpAmps instead of only pass/fail result. What’s more, it is an non-invasive monitor and can work online without isolating the DUT Opamp from its surrounding feedback networks. Nor does it require accurate test stimulations. In addition, the monitor can remove its own offsets without additional complex self-calibration circuits. All self-calibrations are completed in the digital domain after each measurement in real time. Therefore it is also suitable for aging-sensitive applications, in which the monitor may suffer from aging mechanisms and has additional offset drifts as well. The monitor measurement range for offset is from 0.2mV to 70mV, and for gain it is from 0dB to 40dB. The error for offset measurements can be 10% of the measurement value with plus/minus 0.1mV, and -2.5dB for gain measurements.

IP1-8

EVX: VECTOR EXECUTION ON LOW POWER EDGE CORES

Speakers: Milovan Durić, Oscar Palomar, Aaron Smith, Osman Uslu, Adrian Cristal, Mateo Vaknin and Doug Burger
iSupercalibrating Center, ES; iiMicrosoft Research, US

Abstract
In this paper, we present a vector execution model that provides the advantages of vector processors on low-power, general-purpose cores, with limited additional hardware. While accelerating data-level-parallel (DLP) workloads, the vector model increases the efficiency and hardware resources utilization. We use a modest dual issue core based on an Exploit Data Graph Execution (EDGE) architecture to implement our approach, called EVX. Unlike most DLP accelerators which utilize additional hardware and increase the complexity of low-power processors, EVX leverages the available resources of EDGE cores, and with minimal costs allows for specialization of the resources. EVX adds a control logic that increases the core area by 2.1%. We show that EVX yields an average speedup of 3x compared to a scalar baseline and outperforms multimedia SIMD extensions.

IP1-9

PROGRAM AFFINITY PERFORMANCE MODELS FOR PERFORMANCE AND UTILIZATION

Speakers: Ryan Moore and Bruce Childers, University of Pittsburgh, US

Abstract
Multithreaded applications have a wide variety of behavior, causing complex interactions with today’s chip multiprocessor machines. Application threads may have large private working sets, and may compete for cache space and memory bandwidth. These threads benefit from large private caches. Other threads may share data or communicate, and thus, execute more quickly if using shared caches. Many applications fall somewhere in between, requiring careful thread-to-core assignments to maximize performance. Yet because of the large number of thread-to-core assignments on today’s chip multiprocessors, it is time and energy prohibitive to exhaustively try and determine the best assignment. In this paper, we present and demonstrate application performance models which predict application performance given a proposed thread-to-core assignment. We show how these models can be quickly built and used to select thread-to-core assignments for multiple programs and to improve system utilization.

IP1-10

ADVANCED SIMD: EXTENDING THE REACH OF CONTEMPORARY SIMD ARCHITECTURES

Speakers: Matthias Boettcher, Giacomo Gabriele, Mbou Eycle, Alastair Reid and Bashir M. Al-Hashimi
iUniversity of Southampton, GB; iiARM Ltd., GB

Abstract
SIMD extensions have gained widespread acceptance in modern microprocessors as a way to exploit data-level parallelism in general-purpose cores. Popular SIMD architectures (e.g. Intel SSE/AVX) have evolved by adding support for wider registers and datapaths, and advanced features like indexed memory accesses, per-lane predication and inter-lane instructions, at the cost of additional silicon area and design complexity. This paper evaluates the performance impact of such advanced features on a set of workloads considered hard to vectorize for traditional SIMD architectures. Their sensitivity to the most relevant design parameters (e.g. register/datapath width and L1 data cache configuration) is quantified and discussed. We developed an ARMv7 NEON based ISA extension (ARGON), augmented a cycle accurate simulation framework for it, and derived a set of benchmarks from the Berkeley dwarfs. Our analyses demonstrate how ARGCAN can, depending on the structure of an algorithm, achieve speedups of 1.5x to 16x.

IP1-11

A TIGHTLY-COUPLED HARDWARE CONTROLLER TO IMPROVE SCALABILITY AND PROGRAMMABILITY OF SHARED-MEMORY HETEROGENEOUS CLUSTERS

Speakers: Paolo Burgio, Robin Danilo, Andrea Marongiu, Philippe Coussy and Luca Benini
iUniversity of Bologna, Université de Bretagne-Sud, IT; iiUniversity of Bologna, IT; iiiUniversity of Bologna, IT; Lab-STICC, FR; ivUniversità di Bologna, IT

Abstract
Modern designs for embedded many-core systems increasingly include application-specific units to accelerate key computational kernels with orders-of-magnitude higher execution speed and energy efficiency compared to software counterparts. A promising architectural template is based on heterogeneous clusters, where simple RISC cores and specialized HW units (HMPUs) communicate in a tightly-coupled manner via L1 shared memory. Efficiently Integrating processors and a high number of HW Processing Units (HWPUs) in such an system poses two main challenges, namely, architectural scalability and programmability. In this paper, we describe an optimized Data Pump (DP) which connects several accelerators to a restricted set of computation resources, and acts as a virtualization layer for programming, exposing FIFO queues to offload “HW tasks” to them through a set of lightweight APIs. In this work, we aim at optimizing both these mechanisms, for respectively reducing modules area and making programming sequence easier and lighter.

IP1-12

INFORMER: AN INTEGRATED FRAMEWORK FOR EARLY-STAGE MEMORY ROBUSTNESS ANALYSIS

Speakers: Shrikanth Ganapathy, Ramon Canals, Dan Alexandrescu, Enrico Costanaro, Antonio Gonzalez and Antonio Rubio
iUniversitat Politècnica de Catalunya, ES; iiRoC Technologies, FR; iiiTel and Universitat Politècnica de Catalunya, ES

Abstract
With the growing importance of parametric (process and environmental) variations in advanced technologies, it has become a serious challenge to design reliable, fast and low-power embedded memories. Adopting a variation-aware design paradigm requires a holistic perspective of memory-wide metrics such as yield, power and performance. However, accurate estimation of such metrics is largely dependent on circuit implementation styles, technology parameters and architecture-level specifics. In this paper, we propose a fully automated tool - INFORMER that helps high-level designers estimate memory reliability metrics rapidly and accurately. The tool relies on accurate circuit-level simulations of failure mechanisms such as ageing, soft-errors and parametric failures. The obtained statistics can then help couple low-level metrics with higher-level design choices. A new technique for rapid estimation of low-probability failure events is also proposed. We present three use-cases of our prototype tool to demonstrate its diverse capabilities in autonomously guiding large SRAM-based robust memory designs.
**IP1-13 WEAR-OUT ANALYSIS OF ERROR CORRECTION TECHNIQUES IN PHASE-CHANGE MEMORY**

**Speakers:**
Cao Hoffinan, Luiz Ramos, Rodolfo Azevedo and Guido Araújo, University of Campinas, BR

**Abstract**
Phase-Change Memory (PCM) is a new memory technology and a possible replacement for DRAM, whose scaling limitations require new lithography technologies. Despite being promising, PCM has limited endurance (its cells withstand roughly 10^8 bit-flips before failing), which prompted the adoption of Error Correction Techniques (ECTs). However, previous lifetime analyses of ECTs did not consider the difference between the bit-flip frequencies of data and code bits, which may lead to inaccurate wear-out analyses for the ECTs. In this work, we improve the wear-out analysis of PCM by modeling and analyzing the bit-flip probabilities of five ECTs. Our models also enable an accurate estimation of energy consumption and analysis of the endurance-energy trade-off for each ECT.

**IP1-14 APPROXIMATING THE AGE OF RF/ANALOG CIRCUITS THROUGH RE-CHARACTERIZATION AND STATISTICAL ESTIMATION**

**Speakers:**
DooHwang Chang1, Sule Ozev1, Ozgur Sinanoglu1 and Ramesh Kanti1
1Arizona State University, US; 2New York University Abu Dhabi, AE; 3Polytechnic Institute of New York University, US

**Abstract**
Counterfeit ICs have become an issue for semiconductor manufacturers due to impacts on their reputation and lost revenue. Counterfeit ICs are either products that are intentionally mislabeled or legitimate products that are extracted from electronic waste. The former is easier to detect whereas the latter is harder since they are identical to new devices but display degraded performance due to environmental and use stress conditions. Detecting counterfeit ICs that are extracted from electronic waste requires an approach that can approximate the age of manufactured devices based on their parameters. In this paper, we present a methodology that uses information on both fresh and aged ICs and tries to distinguish between the fresh and aged population based on an estimate of the age. Since analog devices age mainly due to their bias stress, input signals play less of a role. Hence, it is possible to use simulation models to approximate the aging process, which would give us access to a large population of aged devices. Using this information, we can construct a statistical model that approximates the age of a given circuit. We use a Low noise amplifier (LNA) and an NMOS LC oscillator to demonstrate that individual aged devices can be accurately classified using the proposed method.

**IP1-15 PACKAGE GEOMETRIC AWARE THERMAL ANALYSIS BY INFRARED-RADIATION THERMAL IMAGES**

**Speakers:**
Jui-Hung Chien1, Hao Yu2, Ruie-Siang Hsu3, Hseuh-Ju Lin2 and Shih-Chieh Chang3
1Industrial Technology Research Institute, TW; 2NTHU, TW

**Abstract**
Since packages affect the amount of heat transfer, it is important to include package and heat sink in thermal analysis. In this paper, we study the full-chip thermal response with different packages. We first discuss the difficulties of obtaining accurate package models for simulation. To facilitate a designer to perform thermal simulation with different packages, we propose to use a matrix called the package-transfer matrix which can transform a temperature profile of one package to another temperature profile of the desired package. To estimate and verify a package-transfer matrix, we propose an efficient method which uses Infrared Radiation (IR) images from two carefully design test chips with PBGA packages. Our experimental results show that the default package model CBGA in HotSpot can be accurately transferred to any other package through the package-transfer matrix.

**IP1-16 COST-EFFECTIVE DECAP SELECTION FOR BEYOND DIE POWER INTENSITY**

**Speakers:**
Yi-En Chen1, Tu-Huang Tsai2, Shi-Hao Chen1 and Hung-Ming Chen1
1Department of Electronics Engineering National Chiao Tung University Hsinchu, Taiwan 300, R.O.C., TW; 2Global Unichip Corp, Hsinchu, Taiwan, TW

**Abstract**
In designing reliable power distribution networks (PDN) for power integrity (PI), it is essential to stabilize voltage supply to devices on chip. We usually employ decoupling capacitor (decap) to suppress the noise generated by the switching of devices. There have been numerous prior works on how to select/insert decaps in chip, package, or board to maintain PI, however optimal decap selection is usually not applicable due to design budget and manufacturability. Moreover, design cost is seldom touched or mentioned. In this research, we propose an efficient methodology "PDCPSO" to automatically optimizing the selection of available decaps. This algorithm not only takes advantage of particle swarm optimization (PSO) to stochastically search the design space, but takes the most effective range of decaps into consideration to outperform the basic PSO. We apply this to three real package designs and the results show that, compared to the original decap selection by rules of thumb, our approach could shorten the design period and we have better combination of decaps at a lower cost. In addition, our methodology can also consider package-board co-design in optimizing different power domains.

**IP1-17 CHARACTERIZING POWER DELIVERY SYSTEMS WITH ON/OFF-CHIP VOLTAGE REGULATORS FOR MANY-CORE PROCESSORS**

**Speakers:**
Xuan Wang, Jiang Xu, Zhe Wang, Kevin J. Chen, Xiaowen Wu and Zhehui Wang, HKUST, HK

**Abstract**
Design of power delivery system has great influence on the power management in many-core processor systems. Moving voltage regulators from off-chip to on-chip gains more and more interest in the power delivery system design, because it is able to provide fast voltage scaling and multiple power domains. Previous works are proposed to implement power efficient on-chip regulators. It is also important to analyze the characteristics of the entire power delivery system to explore the tradeoff between the promising properties and costs of employing on-chip regulators. In this work, we develop an analytical model to evaluate important characteristics of the power delivery system, including on-chip/off-chip voltage regulators and the passive on-chip/on-board parasitic. Compared with SPICE simulations, our model achieves a fast system-level evaluation with comparable accuracy. Based on the model, geometric programming is utilized to find the optimal power efficiency of different architectures of power delivery systems under constraints of output voltage stability and area. Experiments show that compared with the conventional architecture using off-chip regulators, the hybrid one using both on-chip and off-chip voltage regulators achieves 1.0% power efficiency improvement and 68% area reduction of voltage regulators on average. We conclude that the hybrid architecture has potential for high power efficiency and small area at heavy workload, but careful account for the overhead of on-chip regulators is needed.

**IP1-18 MASK-COST-AWARE ECO ROUTING**

**Speakers:**
Hsi-An Chien1, Zhen-Yu Peng1, Yun-Ru Wu1, Ting-Hsiung Wang1, Hsin-Chang Lin1, Chi-Feng Wu1 and Ting-Chi Wang1
1National Tsing Hua University, TW; 2Realtek Semiconductor Corp., TW

**Abstract**
In this paper, we study a mask-cost-aware routing problem for engineering change order (ECO). By taking into account old rules for possible reuse, we present an approach for the problem. Encouraging experimental results are reported to demonstrate the effectiveness of our approach.
4.1 EXECUTIVE SESSION: Addressing Challenges of Reliable Chips

Date: Tuesday, March 25, 2014
Time: 17:00 - 18:30
Location / Room: Saal 1

Organiser:

IP1-19  EXPLOITING NARROW-WIDTH VALUES FOR IMPROVING NON-VOLATILE CACHE LIFETIME
Speakers: Guangshan Duan and Shuai Wang, Nanjing University, CN

Abstract
Due to the high cell density, low leakage power consumption, and less vulnerability to soft errors, the non-volatile memory technologies are among the most promising alternatives for replacing the traditional DRAM and SRAM technologies used in implementing main memory and caches in the modern microprocessor. However, one of the difficulties is the limited write endurance of most non-volatile memory technologies. In this paper, we propose to exploit the narrow-width values to improve the lifetime of the non-volatile last level caches. Leading zeros masking scheme is first proposed to reduce the write stress to the upper half of the narrow-width data. To balance the write variations between the upper half and the lower half of the narrow-width data, two swap schemes, the swap on write (SW) and swap on replacement (SRepl), are proposed. To further reduce the write stress to the non-volatile cache, we adopt two optimization schemes, the multiple dirty bit (MDB) and read before write (RBW), to improve its lifetime. Our experimental results show that by combining all our proposed schemes, the lifetime of the non-volatile caches can be improved by 245% on average.

IP1-20  PARTIAL-SET: WRITE SPEEDUP OF PCM MAIN MEMORY
Speakers: Li Bing, Shan Shuchang, Hu Yuz and Li XiaoWei

1ICT, UCAS, CN; 2CT, CAS, CN; 3ICT, CAS, CN

Abstract
Redundant Multi-Threading (RDMT) is a promising nonvolatile memory technology developed as a possible replacement for DRAM. Although it offers the read latency close to that of DRAM, RDMT generally suffers from the long write latency. Long write request may block the read requests on the critical path of cache/access memory, incurring adverse impact on the system performance. Besides, the write performance of RDMT is very asymmetric, i.e., the SET operation (writing '1') is much slower than that of the RESET operation (writing '0'). In this work, we re-examine the resistance transform process during the SET operation of RDMT and propose a novel Partial-SET scheme to alleviate the long write latency issue of RDMT. During a write access to a memory line, a short Partial-SET pulse is applied first to program the RDMT cells to a pre-stable state, achieving the same write latency as RESET. The partially programmed cells are then fully programmed within the retention window to preserve the data integrity. Experimental results show that our Partial-SET scheme can improve the memory access performance of RDMT by more than 45% averagely with very marginal storage overhead.

IP1-21  GARBAGE COLLECTION FOR MULTI-VERSION INDEX ON FLASH MEMORY
Speakers: Kam Yu Lam, Jian-Tao Wang, Yuan-Hao Chang, Jen-Wei Hsieh, Po-Chun Huang, Chung Keung Poon and Chun-Jiang Zhu

1City University of Hong Kong, HK; 2Academia Sinica, TW; 3National Taiwan University of Science and Technology, TW; 4Academia Sinica, TW; 5City University of Hong Kong, TW

Abstract
Phase change memory (PCM) is a promising nonvolatile memory technology developed as a possible DRAM replacement. Although it offers the read latency close to that of DRAM, PCM generally suffers from the long write latency. Long write request may block the read requests on the critical path of cache/access memory, incurring adverse impact on the system performance. Besides, the write performance of PCM is very asymmetric, i.e., the SET operation (writing '1') is much slower than that of the RESET operation (writing '0'). In this work, we re-examine the resistance transform process during the SET operation of PCM and propose a novel Partial-SET scheme to alleviate the long write latency issue of PCM. During a write access to a memory line, a short Partial-SET pulse is applied first to program the PCM cells to a pre-stable state, achieving the same write latency as RESET. The partially programmed cells are then fully programmed within the retention window to preserve the data integrity. Experimental results show that our Partial-SET scheme can improve the memory access performance of PCM by more than 45% averagely with very marginal storage overhead.

IP1-22  D2CYBER: A DESIGN AUTOMATION TOOL FOR DEPENDABLE CYBERCARS
Speakers: Arslan Munir and Farinaz Koushanfar, Rice University, US

Abstract
The next generation of automobiles (also known as cybercars) will increasingly incorporate electronic control units (ECUs) in novel automotive control applications. Recent work has demonstrated vulnerability of modern car control systems to security attacks that directly impacts the cybercar's physical safety and dependability. In this paper, we provide an integrated approach for the design of secure and dependable cybercars using a case study: a steer-by-wire (SBW) application over controller area network (CAN). The challenge is to embed both security and dependability over CAN while ensuring that the real-time constraints of the cybercar applications are not violated. Our approach enables early design feasibility analysis by embedding essential security primitives (i.e., confidentiality, integrity, and authentication) over CAN subject to the real-time constraints imposed by the desired quality of service and behavioral reliability. Our method leverages multi-core ECUs for providing fault-tolerance by redundant multi-threading (RMT) and also further enhances RMT for quick error detection. We quantify the error resilience of our approach and evaluate the interplay of performance, fault-tolerance, security, and scalability for our SBW case study.

IP1-23  CONTRACT-BASED DESIGN OF CONTROL PROTOCOLS FOR SAFETY-CRITICAL CYBER-PHYSICAL SYSTEMS
Speakers: Pietruigi Nuzzo, John Finn, Antonio Iannapolito and Alberto Sangiovanni-Vincenelli, University of California at Berkeley, US

Abstract
We introduce a platform-based design methodology that addresses the complexity and heterogeneity of cyber-physical systems by using assume-guarantee contracts to formalize the design process and enable realization of control protocols in a hierarchical and compositional manner. Given the architecture of the physical plant to be controlled, the design is carried out as a sequence of refinement steps from an initial specification to a final implementation, including synthesis from requirements and mapping of higher-level functional and non-functional models into a set of candidate solutions built out of a library of components at the lower level. Initial top-level requirements are captured as contracts and expressed using logical temporal logic (LTL) and signal temporal logic (STL) formulas to enable requirement analysis and early detection of inconsistencies. Requirements are then refined into a controller architecture by combining reactive synthesis steps from LTL specifications with simulation-based design space exploration steps. We demonstrate our approach on the design of embedded controllers for aircraft electric power distribution.

IP1-24  A FAULT DETECTION MECHANISM IN A DATA-FLOW SCHEDULED MULTI-THREADED PROCESSOR
Speakers: Jian Fu, Giang Yang, Raphael Poss, Chris Jesshope and Chunyuan Zhang

1University of Amsterdam, NL; 2National University of Defense Technology, CN

Abstract
This paper designs and implements the Redundant Multi-Threading (RMT) in a Data-flow scheduled Multi-Threaded (DMT) multicore processor, called Data-flow scheduled Redundant Multi-Threading (DRMT). Meanwhile, it presents Asynchronous Output Comparison (AOC) for RMT techniques to avoid fault detection related inter-core communication and alleviate the performance and hardware overheads induced by output comparison. Results show that the performance overhead of DRMT is less than 60% even when the number of threads is four times the number of processing elements. Also the performance and hardware overheads of AOC are insignificant.
While today's SOCs systematically use semiconductor production quality assessment and optimization solutions, meeting end-product requirements for reliability and availability augments the need to prepare the SOC design in advance to address such requirements. The speakers in this executive session will address the current trends and challenges in the semiconductor reliability and discuss the level of readiness needed in a chip to meet today's SOC requirements.

### 4.2 Hot Topic: Multicore Systems in Safety Critical Electronic Control Units for Automotive and Avionics

**Date:** Tuesday, March 25, 2014  
**Time:** 17:00 - 18:30  
**Location / Room:** Konferenz 6

**Organisers:**  
Jürgen Becker, KIT, DE,  
Contact Jürgen Becker  
Oliver Sander, KIT, DE,  
Contact Oliver Sander

**Chair:**  
Jürgen Becker, KIT, DE,  
Contact Jürgen Becker  
Oliver Sander, KIT, DE,  
Contact Oliver Sander

Future applications in automotive and avionics show an ever increasing demand of computational processing power. The use of multicore devices is now emerging in embedded electronics. However these solutions are not directly applicable because of technical requirements that come along with the domain of safety critical and mixed critical applications, such as in automotive or avionics. The major challenge for deployment of multicore devices in safety critical applications such as automotive or avionics, is the lack of determinism and support of segregation due to shared resources. The goal of this session is to present the challenges that arise from the use of multicore devices in embedded safety-critical systems and mixed critical systems.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
</table>
| 17:00  | 4.2.1 | AUTOSAR AND MULTICORE                                                              | Stefan Kuntz and Rolf Schneider  
1Continental Automotive GmbH, DE; 2AUDI AG, DE |
|        |       | **Abstract**                                                                      |         |
|        |       | AUTOSAR already supports developing applications for and integrating software components onto multicore based platforms. In addition, these capabilities pave the way for helping to migrate existing applications, originally developed for being executed on single core platforms, to multicore based platforms. This talk provides a brief introduction of the current state of AUTOSARs multicore support and presents some scenarios that draws the attention to multicore specific questions and challenges in the particular context. Possible future directions in improving the AUTOSAR standard with regard to multicore and to gain more benefit from the availability of multiple cores, independent execution units, are sketched out. |
| 17:30  | 4.2.2 | CONCEPTS TO VALIDATE THE SAFE APPLICATION OF MULTICORE ARCHITECTURES IN THE AVIONICS DOMAIN | Ottmar Bender, Airbus Defence and Space, DE |
|        |       | **Abstract**                                                                      |         |
|        |       | This presentation explains how commercially available multicore processors can be applied for safety critical applications in avionics systems. It also describes remaining difficulties which need to be solved for a full exploitation of multicore technology in the avionics domain. Furthermore a concept of an airborne radar application demonstrator built on multicore architecture is shown. This demonstrator shall allow the validation of essential solutions for the specific difficulties emerging from current multicore architectures. |
| 18:00  | 4.2.3 | MONITORING AND WCET ANALYSIS IN COTS MULTI-CORE/SOC-BASED MIXED-CRITICALITY SYSTEMS | Jan Nowotsch, Michael Paulitsch, Ame Henrichsen, Werner Pongratz and Andreas Schacht  
1EADS Innovation Works, DE; 2EADS Innovation Works, DE; 3Cassidian, DE |
|        |       | **Abstract**                                                                      |         |
|        |       | The performance and power efficiency of multi-core processors are attractive features for safety-critical applications, for example in avionics. But the inherent use of shared resources complicates timing analysability. In this paper we discuss a novel approach to compute the Worst-Case Execution Time (WCET) of multiple hard real-time applications scheduled on a Commercial Off-The-Shelf (COTS) multi-core processor. The analysis is closely coupled with mechanisms for temporal partitioning as, for instance, required in ARINC 653 based systems. Based on a discussion of the challenges for temporal partitioning and timing analysis in multi-core systems, we deduce a generic architecture model. Considering the requirements for re-usability and incremental development and certification, we use this model to describe our integrated analysis approach. |
## 4.3 Secure Device Identification

**Date:** Tuesday, March 25, 2014  
**Time:** 17:00 - 18:30  
**Location / Room:** Konferenz 1

**Chair:** Tim Gueneysu, RUB, DE; Contact Tim Gueneysu

**Co-Chair:** Patrick Schaumont, Virginia Tech, US; Contact Patrick Schaumont

Physically Unclonable Functions (PUFs) have received much attention for fingerprinting of electronic devices. This session presents novel constructions and threats on Ring-Oscillator-based and Sense-Amplifier-based PUFs.

### 4.3.1 HARDWARE VIRTUALIZATION SUPPORT FOR SHARED RESOURCES IN MIXED-CRITICALITY MULTICORE SYSTEMS

**Speakers:**  
Oliver Sanden, Timo Sandmannz, Viet Vu Duys, Steffen Bänz, Falco Bapps, Juergen Beckers, Hans Ulrich Micheli, Dirk Kaule, Daniel Adams, Enno Luebben, Jürgen Harbacher, Andre Richter, Christian Herberz, and Andreas Herkersdorf

**Abstract**  
Electric/Electronic architectures in modern automobiles evolve towards an hierarchical approach where functionalities from several ECUs are consolidated into few domain computers. Performance requirements directly lead to multicore solutions but also to a combination of very different requirements on such ECUs. Using virtualization in addition is one promising way of achieving segregation in time and space of shared resources. Based on examples taken from the automotive domain several concepts for efficient hardware extensions of coprocessors and I/O devices are shown in this contribution. These provide mechanisms to ensure quality of service (QoS) levels in terms of execution time, throughput and latency. The resulting infotainment architecture is a feasibility study and is integrated into a vehicle demonstrator as centralised infotainment platform (VCT).

### 4.3.2 ARO-PUF: AN AGING-RESISTANT RING OSCILLATOR PUF DESIGN

**Speakers:**  
Md. Taufikul Rahman, Domenic Forte, Jim Fahmy, and Mohammad Taherzadeh

**Abstract**  
Physically Unclonable Functions (PUFs) have emerged as a security block with the potential to generate chip-specific identifiers and cryptographic keys. However, it has been shown that the stability of these identifiers and keys is heavily impacted by aging and environmental variations. Previous techniques have mostly focused on improving PUF robustness against supply noise and temperature but aging has been largely neglected. In this paper, we propose a new aging resistant design for the popular ring-oscillator (RO) PUF. Simulation results demonstrate that our aging resistant RO-PUF (called ARO-PUF) can produce unique, random, and more reliable keys. Only 7.7% bits get flipped on average over 10 years operation period for an ARO-PUF due to aging where the value is 32% for a conventional RO-PUF. The ARO-PUF shows an average interchip HD of 49.67% (close to ideal value 50%) and better than the conventional RO-PUF (+45%). With lower error, ARO-PUF offers ~24X area reduction for a $128$-bit key because of reduced ECC complexity and smaller PUF footprint.

### 4.3.3 AN EFFICIENT RELIABLE PUF-BASED CRYPTOGRAPHIC KEY GENERATOR IN 65NM CMOS

**Speakers:**  
Mudit Bhargava and Ken Maiz

**Abstract**  
Physical unclonable functions (PUFs) are primitives that generate high-entropy, tamper-resistant bits for use in secure systems. For applications such as cryptographic key generation, the PUF response bits must be highly reliable, consistent across multiple evaluations under voltage and temperature variations. Conventionally, error correcting codes (ECC) have been used to improve response reliability, but these techniques have significant area, power, and delay overheads and are vulnerable to information leakage. In this work, we present a highly-reliable, PUF-based, cryptographic key generator that uses no ECC, but instead uses built-in self-test to determine which PUF bits are reliable and only uses those bits for key generation. We implemented a prototype of the key generator in a 65nm bulk CMOS testchip. The key generator generates 1213 bits in an area of $0.09\mu m^2$ with a measured bit error rate of $<5 \times 10^{-9}$ in both the nominal and worst case corners (100k measurements each). This is equivalent to a 128-bit key failure rate of $<10^{-6}$. The system can generate a 128-bit key in 1.15$\mu s$. Finally, we present a realization of a "strong"-PUF that uses 128 of these highly reliable bits in conjunction with an Advanced Encryption Standard (AES) cryptographic primitive and has a response time of 40ns and is realized in an area of $84\mu m^2$.

### 4.3.4 INCREASING THE EFFICIENCY OF SYNDROME CODING FOR PUFS WITH HELPER DATA COMPRESSION

**Speakers:**  
Matthias Hiller and Georg Sigl, Institute for Security in Information Technology; Technische Universität München, DE

**Abstract**  
Helper data can contain redundancy. We analyze existing schemes and show that data compression can be applied to decrease the size of the helper data of existing implementations. We introduce compressed Differential Sequence Coding (DSC), which is the most efficient syndrome coding scheme known to date for a popular reference scenario. Adding heper data compression to the DSC algorithm leads to an overall decrease of 68% in helper data size compared to other algorithms in a reference scenario. This is achieved without increasing the number of PUF bits and a minimal increase in logic size.
Numerous industrial and academic efforts are targeting 3D integration, and integrated microfluidics promise to have a profound impact on healthcare and other domains. The three papers in this session all address "nearer-term" emerging technologies.

Michael Niemier, University of Notre Dame, US, Co-Chair:
Ian O’Connor, University of Lyon, FR, Chair:

Time:
18:30

End of session
Exhibition Reception in Several serving points inside the Exhibition Area (Terrace Level)
The Exhibition Reception will take place in the exhibition area (Terrace Level). All exhibitors are welcome to provide drinks and snacks for delegates and visitors.

4.4 "Almost there" emerging technologies
Date: Tuesday, March 25, 2014
Time: 17:00 - 18:30
Location / Room: Konferenz 2

Chair:
Iain O’Connor, University of Lyon, FR, Contact Ian O’Connor

Co-Chair:
Michael Niemier, University of Notre Dame, US, Contact Michael Niemier

The three papers in this session all address "nearer-term" emerging technologies. Stochastic computing techniques are becoming increasingly relevant as CMOS becomes more error prone, numerous industrial and academic efforts are targeting 3D integration, and integrated microfluidics promise to have a profound impact on healthcare and other domains.

Abstract
Physically Unclonable Functions (PUFs) are security primitives that exploit the unique manufacturing variations of an integrated circuit (IC). They are mainly used to generate secret keys. Ring oscillator (RO) PUFs are among the most widely researched PUFs. In this work, we claim various RO PUF constructions to be vulnerable against manipulation of their public helper data. Partial/full key-recovery is a threat for the following constructions, in chronological order. (1) Temperature-aware cooperative RO PUFs, proposed at HOST 2009. (2) The sequential pairing algorithm, proposed at HOST 2010. (3) Group-based RO PUFs, proposed at DATE 2013. (4) Or more general, all entropy distiller constructions proposed at DAC 2013.

IIR FILTERS USING STOCHASTIC ARITHMETIC
Speakers:
Naman Saraf, Kia Bazargan, David J Lilja and Marc D Riedel, University of Minnesota, Twin Cities, US

Abstract
We consider the design of IIR filters operating on oversampled sigma-delta modulated bit streams using stochastic arithmetic. Conventional digital filters process multi-bit data at the Nyquist rate using multi-bit multipliers and adders. High resolution ADCs based on the sigma-delta modulation generate random bits at an oversampled rate as intermediate data. We propose to filter the sigma-delta modulated bit streams directly and present first and second order low pass IIR filters based on the stochastic integrator. Experimental results show a significant reduction in hardware area by using stochastic filters.

EFFICIENT THERMAL SIMULATION OF 3D ICS WITH LIQUID-COOLING AND THROUGH SILICON VIAS
Speakers:
Alain Fourmigue, Giovanni Beltrame and Gabriela Nicolescu, Polytechnique Montreal, CA

Abstract
Three-dimensional integrated circuits (3D ICs) with advanced cooling systems are emerging as a viable solution for many-core platforms. These architectures generate a high and rapidly changing thermal flux. Their design requires accurate transient thermal models. Several models have been proposed, either with limited capabilities, or poor simulation performance. This work introduces an efficient algorithm based on the Finite Difference Method to compute the transient temperature in 3D ICs. Our experiments show a 5x speedup versus state-of-the-art models, while maintaining the same level of accuracy, and demonstrate the effect of large through silicon vias arrays on thermal dissipation.

A LOGIC INTEGRATED OPTIMAL PIN-COUNT DESIGN FOR DIGITAL MICROFLUIDIC BIOCHIPS
Speakers:
Trung Anh Dinh, Shigeru Yamashita1 and Tsung-Yi Ho2
1Ritsumeikan University, JP; 2National Cheng Kung University, TW

Abstract
Digital microfluidic biochips have become one of the most promising technologies for biomedical experiments. In modern microfluidic technology, reducing the number of independent control pins that reflects most of the fabrication cost, power consumption and reliability of a microfluidic system, is a key challenge for every digital microfluidic biochip design. However, all the previous chip designs sacrifice the optimality of the problem, and only limited reduction on the number of control pins is observed. Moreover, most existing designs cannot satisfy high-throughput demand for bioassays, and thus inapplicable in practical contexts. In this paper, we propose the first optimal pin-count design scheme for digital microfluidic biochips. By integrating a very simple combinational logic circuit into the original chip, the proposed scheme can provide high-throughput for bioassays with an information-theoretic minimum number of control pins. Furthermore, to cope with the rapid growth of the chip’s scale, we also propose a scalable and efficient heuristic. Experiments demonstrate that the proposed scheme can obtain much fewer number of control pins compared with the previous state-of-the-art works.

FAST AND ACCURATE COMPUTATION USING STOCHASTIC CIRCUITS
Speakers:
Amin Alaghi and John P. Hayes, University of Michigan - Ann Arbor, US

Abstract
Stochastic computing (SC) is a low-cost design technique that has great promise in applications such as image processing. SC enables arithmetic operations to be performed on stochastic bit-streams using ultra-small and low-power circuitry. However, accurate computations tend to require long run-times due to the random fluctuations inherent in stochastic numbers (SNs). We present novel techniques for SN generation that lead to better accuracy/run-time trade-offs. First, we analyze a property called progressive precision (PP) which allows computational accuracy to grow systematically with run-time. Second, borrowing from Monte Carlo methods, we show that SC performance can be greatly improved by replacing the usual pseudo-random number sources by low-discrepancy (LD) sequences that are predictably progressive. Finally, we evaluate the use of LD stochastic numbers in SC, and show they can produce significantly faster and more accurate results than existing stochastic designs.
The memory sub-system plays an increasingly important role in modern multicore systems. Novel solutions are needed in order to deliver the expected performance improvements with minimal energy overheads. In addition, new solutions should be preferably backward compatible with already existing approaches. In this session we have four papers dealing with different aspects of the memory hierarchy in modern computing systems. ALLARM provides a novel, yet power efficient strategy towards cache coherence to simultaneously improve performance and reduce energy. In addition, new solutions should be preferably backward compatible with already existing approaches. The next paper in this session presents a novel packet-based interface and compression, which reduces communication overhead. The third paper deals with prefetcher aggressiveness and proposes a sound solution to reduce overall execution time. The last paper of this session proposes a novel extension of the shared L2 cache memory system, providing a very high aggregated bandwidth with a very low impact on L2 cache design complexity or operating frequency.
**4.6 Code Generation and Optimization for Embedded Platforms**

**Date:** Tuesday, March 25, 2014  
**Time:** 17:00 - 18:30  
**Location / Room:** Konferenz 4

**Chair:** Heiko Falk, Ulm University, DE; Contact Heiko Falk

**Co-Chair:** Florence Maraninchi, Grenoble IMP/VERIMAG, FR; Contact Florence Maraninchi

This session covers the broad spectrum of topics in compilers, code optimization, and validation under consideration of today’s embedded platforms. The first paper addresses the automated validation of binary translators. The second paper focuses on the on-device optimization of apps and system libraries of mobile platforms. The third paper deals with the code generation of Android image processing applications for heterogeneous GPU-based architectures. The session is rounded off by short presentations of work-in-progress ideas on model transformation, energy and wear-leveling optimization, and scheduling/register allocation.
Binary translation makes it convenient to emulate one instruction set by another. Nowadays, it is growing in popularity in various applications, especially the embedded platforms. When it comes to the test of binary translators, traditional methodologies which still mainly rely on manual unit test is costly, labor intensive and often not adequate to test complicated algorithms in the translators. Some standard benchmark suites, like SPEC CPU2006, are compiled with different compilation options for further tests. However, the translation modules still have over 30% of their code unexecuted after such tests, according to our experimental results. Methodologies based on randomization can generate a vast variety of tests, thus improve the code coverage in the translation system. In this paper, we propose such an approach named EATBit. Test binaries are generated with randomly selected instructions and operands. The binaries and a large amount of input data are then refined to exclude invalid ones. Experimental results on a real binary translator demonstrate that EATBit can not only improve code coverage by over 20%, but also find some new bugs in the translator successfully.

Smartphones provide applications that are increasingly similar to those of interactive desktop programs, providing rich graphics and animations. To simplify the creation of these interactive applications, mobile operating systems employ high-level object-oriented programming languages and shared libraries to manipulate the device's peripherals and provide common user-interface frameworks. The presence of dynamic dispatch and polymorphism allows for robust and extensible application coding. Unfortunately, the presence of dynamic dispatch also introduces significant overheads during method calls, which directly impact execution time. Furthermore, since these applications rely heavily on shared libraries and helper routines, the quantity of these method calls is higher than those found in typical desktop-based programs. Optimizing these method calls centrally before consumers download the application onto a given phone is exacerbated due to the large diversity of hardware and operating system versions that the application could run on. This paper proposes a methodology to tailor a given Objective-C application and its associated device-specific shared library codebase using on-device post-compilation code optimization and transformation. In doing so, many polymorphic sites can be resolved statically, improving the overall application performance.

The success of Android is based on its unified Java programming model that allows to write platform-independent programs for a variety of different target platforms. However, this comes at the cost of performance. As a consequence, Google introduced APIs that allow to write native applications and to exploit multiple cores as well as embedded GPUs for compute-intensive parts. This paper presents code generation techniques in order to target the Renderscript and Filerscript APIs. Renderscript harnesses multi-core CPUs and unified shader GPUs, while the more restricted Filerscript also supports GPUs with earlier shader models. Our techniques focus on image processing applications and allow to target these APIs and OpenCL from a common description. We further supersede memory transfers by sharing the same memory region among different processing elements on HSA platforms. As reference, we use an embedded platform hosting a multi-core ARM CPU and an ARM Mali GPU. We show that our generated source code is faster than native implementations in OpenCV as well as the pre-implemented script intrinsics provided by Google for acceleration on the embedded GPU.

An increasingly large number of safety-critical embedded systems rely on software to prevent and mitigate hazards occurring due to design errors and unexpected interactions of the system with its users and the environment. Implementing a safety instrumented function in the way advocated by the traditional software methods requires an intimate understanding and thorough validation of a complex ecosystem of programming languages, compilers, operating systems and hardware. We propose to consider an alternative where a system designer, for each individual problem, creates in a correct-by-construction manner both the design of a system and its compilation infrastructure. This permits an unimpeached chain of a formal correctness argument spanning from formalised requirements all the way to the gate-level characterisation of an execution environment. The past decade of advances in verification technology turned the mechanical verification of large-scale models into a reality while the pressure of certification makes the cost of a formally verified development routine increasingly acceptable. The proposed technique fits the Grand Challenge for Computer Research posed by Hoare in 2003, namely, development of a Verifying Compiler which not only mechanically translates a given program from one language to another but also verifies its correctness according to a formal specification. This allows meeting the most stringent software certification requirements such as SIL 4. We illustrate the idea with a small case study developed using the Event-B modelling notation and tools.
The Exhibition Reception will take place in the exhibition area (Terrace Level). All exhibitors are welcome to provide drinks and snacks for delegates and visitors.

Abstract

Phase Change Memory (PCM) is a promising DRAM replacement in embedded systems due to its attractive characteristics such as extremely low leakage power, high storage density and good scalability. However, PCM's low endurance constrains its practical applications. In this paper, we propose a Wear-Leveling aware dynamic stack to enhance the lifetime of PCM memory.

This paper presents an on-the-fly register allocator which dynamically detects and utilizes lifetime holes for clustered VLIW processors. A lifetime hole is an interval in which a variable does not contain a valid value. A register holding a lifetime hole can be allocated to another variable whose live range fits in the lifetime hole, leading to more efficient utilization of registers. We propose efficient techniques for dynamically utilizing lifetime holes and incorporate these techniques into our on-the-fly register allocator.

We have simulated our register allocator and a linear scan register allocator without considering lifetime holes by using the MediaBench II benchmark suite. Our simulation results show that our register allocator reduces the number of spills by 12.5%, 11.7%, 12.7%, for three different processor models, respectively.

This paper presents an on-the-fly register allocator which dynamically detects and utilizes lifetime holes for clustered VLIW processors. A lifetime hole is an interval in which a variable does not contain a valid value. A register holding a lifetime hole can be allocated to another variable whose live range fits in the lifetime hole, leading to more efficient utilization of registers. We propose efficient techniques for dynamically utilizing lifetime holes and incorporate these techniques into our on-the-fly register allocator.

We have simulated our register allocator and a linear scan register allocator without considering lifetime holes by using the MediaBench II benchmark suite. Our simulation results show that our register allocator reduces the number of spills by 12.5%, 11.7%, 12.7%, for three different processor models, respectively.

VERIFICATION-GUIDED VOTER MINIMIZATION IN TRIPLE-MODULAR REDUNDANT CIRCUITS

Speakers:
Dmitry Burlyaev, Pascal Fradet and Alain Girault, INRIA, FR

Abstract

The use of side-channel measurements and fingerprinting, in conjunction with statistical analysis, has proven to be the most effective method for accurately detecting hardware Trojans in fabricated integrated circuits. However, these past-fabrication trust evaluation methods overlook the capabilities of advanced design skills that attackers can use in designing sophisticated Trojans. To this end, we have designed a Trojan using power-gating techniques and demonstrate that it can be masked from advanced side-channel fingerprinting detection while dormant. We then propose a real-time trust evaluation framework that continuously monitors the on-board global power consumption to monitor chip trustworthiness. The measurements obtained corroborate our frameworks effectiveness for detecting Trojans. Finally, the results presented are experimentally verified by performing measurements on fabricated Trojan-free and Trojan-infected variants of a reconfigurable linear feedback shift register (LFSR) array.

17:30 4.7.2 Verification-guided Voter Minimization in Triple-Modular Redundant Circuits

Speakers:
Dmitry Burlyaev, Pascal Fradet and Alain Girault, INRIA, FR

Abstract

We present a formal approach to minimize the number of voters in triple-modular-redundant sequential circuits. Our technique actually works on a single copy of the circuit and considers a user-defined fault model (under the form "at most 1 bit-flip every k clock cycles"). Verification-based voter minimization guarantees that the resulting circuit (i) is fault tolerant to the soft-errors defined by the fault model and (ii) is functionally equivalent to the initial one. Our approach operates at the logic level and takes into account the input and output interface specifications of the circuit. Its implementation makes use of graph traversal algorithms, fixed-point iterations, and BDDs. Experimental results on the ITC'99 benchmark suite indicate that our method significantly decreases the number of inserted voters which entails a hardware reduction of up to 55% and a clock frequency increase of up to 35% compared to full TMR. We address scalability issues arising from formal verification with approximations and assess their efficiency and precision.
**4.7.3** 18:00

**Presentation Title:** TRADE-OFFS IN EXECUTION SIGNATURE COMPRESSION FOR RELIABLE PROCESSOR SYSTEMS

**Speakers:**
Jonah Caplan, Maria Maras, Peter Miller and Brett Meyer

1\(^{\text{st}}\)McGill University, CA; 2\(^{\text{nd}}\)SUNY Stonybrook, US

**Abstract**
As semiconductor processes scale, making transistors more vulnerable to transient upsets, a wide variety of microarchitectural and system-level strategies are emerging to perform efficient error detection and correction in computer systems. While these approaches often target various application domains and address error detection and correction at different granularities and with different overheads, an emerging trend is the use of state compression, e.g., cyclic redundancy check (CRC), to reduce the cost of redundancy checking. Prior work in literature has shown that Fletcher’s checksum (FC), while less effective where error detection probability is concerned, is less computationally complex when implemented in software than the more-effective CRC. In this paper, we reexamine the suitability of CRC and FC as compression algorithms when implemented in hardware for embedded safety-critical systems. We have developed and evaluated parameterizable implementations of CRC and FC in FPGA, and we observe that what was true for software implementations does not hold in hardware: CRC is more efficient than FC across a wide variety of target input bandwidths and compression strengths.

**4.7.4** 18:15

**Presentation Title:** AN ENERGY-AWARE FAULT TOLERANT SCHEDULING FRAMEWORK FOR SOFT ERROR RESILIENT CLOUD COMPUTING SYSTEMS

**Speakers:**
Yue Gao, Sandeep Gupta, Yanzhi Wang and Massoud Pedram, University of Southern California, CA

**Abstract**
For modern high-performance systems, aggressive technology and voltage scaling has drastically increased their susceptibility to soft errors. At the grand scale of cloud computing, it is clear that soft error induced failures will occur far more frequently, but it is unclear as to how to effectively apply current error detection and fault tolerance techniques in scale. In this paper, we focus on energy-aware fault tolerant scheduling in public, multi-user cloud systems, and explore the three-way tradeoff between reliability (in terms of soft error resiliency), performance and energy. Through a systematically optimized resource allocation, error detection approach selection, virtual machine placement, spatial/temporal redundancy augmentation and task scheduling process, the cloud service provider can achieve high error coverage and fault tolerance confidence while minimizing global energy costs under user deadline constraints. Our scheduling algorithm includes a static scheduling phase that operates on task graph based workload inputs prior to execution, and a light-weight dynamic scheduler that migrates tasks during execution in case of excessive re-executions. All schedules are evaluated on a runtime simulation engine that (1) mimics the performance fluctuations in cloud systems, and (2) supports the injection of arbitrary fault patterns. Compared to current virtual machine or task replication techniques, we are able to reduce overall application failure rates by over 50% with approximately 76% total energy overhead.

**4.8** 18:30

**Presentation Title:** A LOW-POWER, HIGH-PERFORMANCE APPROXIMATE MULTIPLIER WITH CONFIGURABLE PARTIAL ERROR RECOVERY

**Speakers:**
Cong Liu, Jie Han and Fabrizio Lombardi

1 University of Alberta, CA; 2 Northeastern University, US

**Abstract**
Approximate circuits have been considered for error-tolerant applications that can tolerate some loss of accuracy with improved performance and energy efficiency. Multipliers are key arithmetic circuits in many such applications such as digital signal processing (DSP). In this paper, a novel approximate multiplier with a lower power consumption and a shorter critical path than traditional multipliers is proposed for high-performance DSP applications. This multiplier leverages a newly-designed approximate adder that limits its carry propagation to the nearest neighbors for fast partial product accumulation. Different levels of accuracy can be achieved through a configurable error recovery by using different numbers of most significant bits (MSBs) for error reduction. The approximate multiplier has a low mean error distance, i.e., most of the errors are not significant in magnitude. Compared to the Wallace multiplier, a 16-bit approximate multiplier implemented in a 28nm CMOS process shows a reduction in delay and power of 20% and up to 69%, respectively. It is shown that by utilizing an appropriate error recovery, the proposed approximate multiplier achieves similar processing accuracy as traditional exact multipliers but with significant improvements in power and performance.

**19:00** End of session

The Exhibition Reception will take place in the exhibition area (Terrace Level). All exhibitors are welcome to provide drinks and snacks for delegates and visitors.

---

**4.8 State-of-the-art in Verification:**

**European Tertulia IC Design - Enabling AMS Structured Verification / Verification in FPGA & IP design flows**

- **Date:** Tuesday, March 25, 2014
- **Time:** 17:00 - 18:30
- **Location / Room:** Exhibition Theatre

**Organiser:**
Andreas Brüning, Silicon Saxony, DE

---

**4.8.1** 17:00

**Presentation Title:** BRING ASIC-ALIKE VERIFICATION TO YOUR FPGA & IP DESIGN FLOW

**Speaker:**
Scott Calkins, Blue Pearl Software Inc, US

**Abstract**
This talk will highlight how successful design teams and IP firms such as PLDA are able to develop high quality code by using a process to control and optimize the HDL which is developed by different designers in different locations, even those with vastly different skill sets. PLDA designs and sells intellectual property (IP) cores and prototyping tools for ASIC and FPGA that aim to accelerate time-to-market for embedded electronic designs. PLDA specializes in high-speed interface protocols and technologies such as PCIe. Through the use of Blue Pearl Software’s Symmetric Engine that maps code to RTL level then analyzed it for known structures, PLDA is able to generate deterministic results for the handful of synthesizers and target fabrics their customers demand. Analyzing the HDL before it is brought into cycle-based simulators allows designers to run FPGA-centric structural checks for Xilinx and Altera so it helps to detect bugs and specific optimizations earlier in the flow and automatically for the success and satisfaction of our customers' designers: “Blue Pearl Software’s design analysis tool enables integration of formal verification techniques to our design flow, in order for us to detect structural bugs at the very early stage of code integration, and thus to deliver highest quality IP to our customers. On top, we definitely recommend Blue Pearl Software’s solution to anyone who needs to increase design team productivity.” — Hugues Deneux, R&D Director of PLDA
Presentation Title: Enabling AMS Structured Verification

Authors:
Gunter Strube1 and Stefan Getzlaff2
1MunEDA GmbH, DE; 2ZMDI, DE

Abstract
The verification of the robustness of design specifications with respect to all combinations of worst-case parameter conditions not only improves the design confidence, but it is increasingly becoming a requirement for quality assurance and documentation for norms. It is a complex task consuming significant man power and compute power and it tends to be sacrificed under time pressure in the final stage of a project. We present an automated structured approach that differentiates through its thoroughness, its efficiency and most of all its ease-of-use. It enables even novice designers to apply advanced state-of-the-art statistical tools to create a report including a measure of robustness for each specification and for the circuit.

Presentation Title: More information ...

Abstract
We can today design and verify digital hardware and software in a way that deserves the word co-design. Co-design achieves a significantly higher productivity in the design, and better performances of the product. Unfortunately, co-design and co-verification is not yet done in a similar productive way for analog and RF systems. The presentation will give an overview of methodology, tools, and languages that include analog and RF design into a comprehensive co-design methodology. Particular focus is on tool integration and power profiling crossing the discrete-analog border.

Presentation Title: Towards Co-Design and Co-Verification of HW, SW, and Analog Systems

Authors:
Christoph Grimm, TU Kaiserslautern, DE

Abstract
We can today design and verify digital hardware and software in a way that deserves the word co-design. Co-design achieves a significantly higher productivity in the design, and better performances of the product. Unfortunately, co-design and co-verification is not yet done in a similar productive way for analog and RF systems. The presentation will give an overview of methodology, tools, and languages that include analog and RF design into a comprehensive co-design methodology. Particular focus is on tool integration and power profiling crossing the discrete-analog border.

Presentation Title: QUANTUMEDA: A VISUALIZATION AND DESIGN ENVIRONMENT FOR TOPOLOGICAL QUANTUM CIRCUITS

Authors:
Ilia Polan, Wolfgang Walther, and Alexandru Paler, University of Passau, DE

Abstract
Quantum circuits use quantum-mechanical properties of certain physical systems, such as superposition and entanglement, to perform massively parallel calculations. They provide polynomial algorithms for problems for which only inefficient algorithms with asymptotically-exponential running time are known in conventional models of computation. Building a scalable quantum computer that can process a large number of quantum bits (qubits) is one of the grand challenges of modern science. While first small quantum computers have been experimentally demonstrated and a number of implementation technologies have been suggested, all of them encounter difficulties when it comes to scaling. The central difficulty is the high susceptibility of such circuits to noise and decoherence, which necessitates the use of special quantum error correction. Topological quantum computing (TQC) is a paradigm that offers a path to scalability. It strikes a balance between systematic, intuitive methods to design large computations, and relatively loose requirements on the vulnerability of individual qubits to errors. The availability of a platform for implementing large quantum algorithms constitutes the need for methods to manage design complexity, including automatic synthesis, optimization, compaction, verification and visualization of TQC circuits. Topological quantum circuits are based on a three-dimensional cluster of qubits which supports highly efficient topological quantum error-correcting codes. In this way, the circuits can operate even though its individual qubits are subject to relatively high error rates. We will present the first environment for design of TQC circuits. The environment allows the user to graphically enter the structure of a circuit, add, delete and re-shape individual qubits, and perform optimization and compaction (both manually and by global replacement). The circuits are represented on an intermediate technology-independent level where “logical qubits” that consist of a large number of physical qubits perform error-corrected operations. For example, the circuit in Fig. 1 shows an error-corrected CNOT gate implemented by four logical qubits represented by colored structures. The optimized representation can be translated into instruction sequences for a classical computer that operates the actual quantum hardware.

More information ...

Presentation Title: AIDA: Analog IC Design Automation

Authors:
Nuno Horta1, Nuno Lourenço2, Ricardo Martins3, Ricardo Póvoa4, António Canelas2, and Pedro Ventura1
1Instituto de Telecomunicacoes, PT; 2Instituto de Telecomunicacoes / Instituto Superior Técnico, PT

Abstract
This demonstration presents AIDA, an analog integrated circuit (IC) design automation environment. AIDA includes two main modules, namely, AIDA-C and AIDA-L. AIDA-C is a circuit-level synthesis tool which uses state-of-the-art multi-objective multi-constrained optimization kernels, based on evolutionary computation techniques, where the robustness of the solutions is attained by considering a layout-aware approach and, also, extreme process variation by means of PVT corner analysis. The circuit’s performance is measured using Spectre®, ELDO® or HSPICE® electrical simulators as evaluation engines. AIDA-L considers the device sizes and the best floorplan, obtained with AIDA-C, and generates the complete layout by placing and routing the devices, while fulfilling the technology design rules by using built-in design-rule check (DRC) and layout-versus-schematic (LVS) procedures. In order to demonstrate AIDA design environment several analog circuit structures, e.g., OTAs, LNAs, LC-Oscillators, etc., will be synthesized in a 130nm CMOS technology. AIDA-C is demonstrated for circuit-level sizing and optimization by generating a family of Pareto Optimal solutions based on user performance and functional specifications. AIDA-L is demonstrated by generating the layout of a user selected solution from AIDA-C, taking into account electrical currents information to mitigate electromigration and IR-drop effects, and also wiring symmetry for multiphase multi-terminal signal nets of analog ICs.

More information ...
The Exhibition Reception will take place in the exhibition area (Terrace Level). All exhibitors are welcome to provide drinks and snacks for delegates and visitors.

Location / Room: UB04.03

Time: 10:30 - 11:30

More information ...

AUTHORS

David May and Walter Stechek, TUM, DE

Abstract

We want to demonstrate an FPGA-based probability-aware fault emulator and its corresponding algorithms in the context of a real-time H.264 decoder. The demo will show that reliability constraints can be relaxed inside the circuit without noticeable degradation of the image quality when carefully investigating where the constraints can be relaxed. We will show how this investigation can be done using our emulator and we will show the effect of a relaxed robustness of the circuit in real-time.

More information ...

19:30 End of session

Exhibition-Reception Exhibition Reception

Date: Tuesday, March 25, 2014

Time: 18:30 - 19:30

Location / Room: Several serving points inside the Exhibition Area (Terrace Level)

The Exhibition Reception will take place in the exhibition area (Terrace Level). All exhibitors are welcome to provide drinks and snacks for delegates and visitors.
5.1 SPECIAL DAY Hot Topic: Predictable Multi-Core Computing

Date: Wednesday, March 26, 2014
Time: 08:30 - 10:00
Location / Room: Saal 1

Organiser: Jürgen Teich, University of Erlangen-Nuremberg, DE, Contact Jürgen Teich
Chair: Petru Eles, Linköping University, SE, Contact Petru Eles
Co-Chair: Jürgen Teich, University of Erlangen-Nuremberg, DE, Contact Jürgen Teich

The requirement of high performance computing at low power can be met by the parallel execution of an application on a possibly large number of programmable cores. However, the lack of accurate timing properties may prevent parallel execution from being applicable to time-critical applications. This session treats this important problem of time predictability of applications on multi-core platforms by presenting results of the impact of resource sharing on performance, an architecture that has been designed to meet predictability requirements as well as new results on scheduling mixed critical applications on multi-core platforms.

### 5.1.1 IMPACT OF RESOURCE SHARING ON PERFORMANCE AND PERFORMANCE PREDICTION

**Speakers:**
Jan Reineke and Reinhard Wilhelm, Informatik, Universität des Saarlandes, DE

**Abstract**
Multi-core processors are increasingly considered as execution platforms for embedded systems because of their good performance/energy ratio. However, the interference on shared resources poses several problems. It may severely reduce the performance of tasks executed on the cores, and it increases the complexity of timing analysis and/or decreases the precision of its results. Many applications implemented on multi-core platforms are safety- and some also time-critical. A critical issue for these applications is the reduced predictability of such systems resulting from the interference of different applications on shared resources. These interferences can be at least of two kinds: Several applications may request a resource at the same time, but the resource can only admit one access at a time. As a consequence, an arbitration mechanism may delay the request of all but one application, thus slowing down the other applications. This is the case of resources like buses, typically called bandwidth resources. On the other hand, one application may also change the state of a shared resource such that another application using that resource will suffer from a slowdown. This is the case with shared caches, which fall into the class of storage resources. Interference of shared resources makes worst-case execution time (WCET) analysis of applications more difficult since a task or a thread can no longer be analyzed for its timing behavior in isolation. All potential interferences slowing down the task under analysis have to be considered. This leads to a combinatorial explosion of the analysis complexity, as all possible interweavings of different threads have to be analyzed.

### 5.1.2 TIME-CRITICAL COMPUTING ON A SINGLE CHIP MASSIVELY PARALLEL PROCESSOR

**Speaker:**
Benoît Dupont de Dinechin, Kalray, FR

**Abstract**
The requirement of high performance computing at low power can be met by the parallel execution of an application on a possibly large number of programmable cores. However, the lack of accurate timing properties may prevent parallel execution from being applicable to time-critical applications. We illustrate how this problem has been addressed by suitably designing the architecture, implementation, and programming model of the Kalray MPPA-256 single-chip many-core processor. The MPPA-256 (Multi-Purpose Processing Array) processor integrates 256 processing engine (PE) cores and 32 resource management (RM) cores on a single 28nm CMOS chip. These VLIW cores are distributed across 16 compute clusters and 4 I/O subsystems, each with a locally shared memory. On-chip communication and synchronization are supported by an explicitly addressed dual network-on-chip (NoC), with one node per compute cluster and 4 nodes per I/O subsystem. Off-chip interfaces include DDR, PCI and Ethernet, and a direct access to the NoC for low-latency processing of data streams. The key architectural features that support time-critical applications are timing compositional cores, independent memory banks inside the compute clusters, and the data NoC whose guaranteed services are determined by network calculus. The programming model provides a combinatorial explosion of the analysis complexity, as all possible interweavings of different threads have to be analyzed.

### 5.1.3 MAPPING MIXED-CRITICALITY APPLICATIONS ON MULTI-CORE ARCHITECTURES

**Speakers:**
Georgia Giannopoulou1, Nikolay Stoimenov1, Pengcheng Huang2 and Lothar Thiele3
1ETH Zurich, CH; 2ETHZ, CH; 3Swiss Federal Institute of Technology Zurich, CH

**Abstract**
A common trend in real-time embedded systems is to integrate multiple applications on a single platform. Such systems are known as mixed-criticality (MC) systems when the applications are characterized by different criticality levels. Nowadays, multicores platforms are promoted due to cost and performance benefits. However, certification of multicores systems is challenging as concurrent execution of different criticalities may block each other when accessing shared platform resources. Most of the existing research on multicores MC scheduling ignores the effects of resource sharing on the response times of applications. Recently, a MC scheduling strategy was proposed, which explicitly accounts for these effects. This paper discusses how to combine this policy with an optimization method for the partitioning of tasks to cores as well as the static mapping of memory blocks, i.e., task data and communication buffers, to the banks of a shared memory architecture. Optimization is performed at design time targeting at minimizing the worst-case response times of tasks and achieving efficient resource utilization. The proposed optimization method is evaluated using an industrial application.

10:00 End of session

Coffee Break in Exhibition Area
On Tuesday-Thursday the coffee and lunch breaks will be located in the Exhibition Area (Terrace Level).
For this Hot-Topic Session, we will have four leading researchers and experienced speakers from different companies to address both hacking and protecting ICs for chip data. Two speakers will focus on the weaknesses of IC and systems and the ways they can be hacked to retrieve secret data, while the other two will cover smart schemes that can be used to protect ICs from such attacks.

**5.2.1 HARDWARE ATTACKS ON SECURE ICS**

**Speaker:** Gerard van Battum, Brightsight, NL

**Abstract**

He will talk a little bit about the history of attacks and their evolution till today. Thereafter, an overview and a classification of different attacks and their effects will be discussed. Examples will be given of hardware attack techniques on actual secure ICs, such as reverse engineering, mechanical probing, (e-beam) microscopy, etching and polishing, ROM code analysis and Focused Ion Beam modification. This will be put in perspective with commonly applied design practices to protect state-of-the-art secure ICs, which make hardware attacks more difficult.

**5.2.2 ATTACKING SMART PHONES**

**Speaker:** Jean-Luc Danger, Secure IC, FR

**Abstract**

He will address the attack on mobile phones by side-channel. The cryptographic functions executed by the mobile phone processor leak information via the electromagnetic channel. Thus, this non-intrusive observation of the leakage is exploitable at distance to retrieve the secret keys of the cryptographic algorithms. Some attack examples will be given to demonstrate the power of such threats.

**5.2.3 SECURING SYSTEM ON CHIPS**

**Speaker:** Fethulah Smailbegovic, ESCRYPT GmbH – Embedded Security, DE

**Abstract**

He will focus on the challenge of securing generic System on Chip (SoC) architectures. Growing SoC complexity, costs and short time-to-market requirements limit the availability of dedicated hardware security solutions in SoC architectures introducing new potential security risks. Answering the question of how to build secure and reliable SoCs is one of the major challenges in the near future. First, he will briefly talk about current state-of-the-art security solutions in SoCs and afterwards about architectural requirements for future secure System on Chip architectures.

**5.2.4 SILICONAP: A SILICON AUTHENTICATION PLATFORM FOR SECURITY AND ANTI-COUNTERFEITING**

**Speaker:** Mohammad Tehranipoor, TrueLogic, US

**Abstract**

He will talk about design for security and anti-counterfeiting. His talk includes new design techniques for Trojan detection, Trojan prevention, vulnerability analysis, as well as design techniques for preventing counterfeiting of integrated circuits and providing means for easy detection.

10:00 End of session

Coffee Break in Exhibition Area

On Tuesday-Thursday the coffee and lunch breaks will be located in the Exhibition Area (Terrace Level).
<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:30</td>
<td>5.3.1</td>
<td>TEMPORAL MEMOIZATION FOR ENERGY-EFFICIENT TIMING ERROR RECOVERY IN GPGPUS</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Speakers:</strong> Abbas Rahimi¹, Luca Benini¹ and Rajesh Gupta¹</td>
</tr>
<tr>
<td></td>
<td></td>
<td>¹UC San Diego, US; ²Università di Bologna, IT</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Manufacturing and environmental variability lead to timing errors in computing systems that are typically corrected by error detection and correction mechanisms at the circuit level. The cost and speed of recovery can be improved by memoization-based optimization methods that exploit spatial or temporal parallelisms in suitable computing fabrics such as general-purpose graphics processing units (GPGPUs). We propose here a temporal memoization technique for use in floating-point units (FPUs) in GPGPUs that uses value locality inside data-parallel programs. The technique recalls (memorizes) the context of error-free execution of an instruction on a FPU. To enable scalable and independent recovery, a single-cycle lookup table (SUT) is tightly coupled to every FPU to maintain contexts of recent error-free executions. The SUT reuses these memorized contexts to exactly, or approximately, correct errant FP instructions based on application needs. In real-world applications, the temporal memoization technique achieves an average energy saving of 8%-28% for a wide range of timing error rates (0%-4%) and outperforms recent advances in resilient architectures. This technique also enhances robustness in the voltage overscaling regime and achieves relative average energy saving of 66% with 11% voltage overscaling.</td>
</tr>
<tr>
<td>09:00</td>
<td>5.3.2</td>
<td>RELIABILITY-AWARE EXCEPTIONS: TOLERATING INTERMITTENT FAULTS IN MICROPROCESSOR ARRAY STRUCTURES</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Speakers:</strong> Waleed Dweik, Murali Annavararam and Michel Dubois, University of Southern California, US</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>In future technology nodes, reliability is expected to become a first-order design constraint. Faults encountered in a chip can be classified into three categories: transient, intermittent, and permanent. Fault classification allows a chip to take the appropriate corrective action. Mechanisms have been proposed to distinguish transient from non-transient faults where all non-transient faults are handled as permanent. Intermittent faults induced by wearout phenomena have become the dominant reliability concern in nanoscale technology, yet there is no mechanism that provides finer classification of non-transient faults into intermittent and permanent faults. In this paper, we present a new class of exceptions called Reliability-Aware Exceptions (RAEs) which provide the ability to distinguish intermittent faults in microprocessor array structures. The RAE handlers have the ability to manipulate microprocessor array structures to recover from all three categories of faults. Using RAEs, we demonstrate that the reliability of two representative microarchitecture structures, load/store queue and reorder buffer in an out-of-order processor, is improved by average factors of 1.3 and 1.95, respectively.</td>
</tr>
<tr>
<td>09:30</td>
<td>5.3.3</td>
<td>TEMPERATURE AWARE ENERGY-RELIABILITY TRADE-OFFS FOR MAPPING OF THROUGHPUT-CONSTRAINED APPLICATIONS ON MULTIMEDIA MPSOCS</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Speakers:</strong> Anup Das, Akash Kumar and Bharadwaj Veeravalli, National University of Singapore, SG</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>This paper proposes a design-time (offline) analysis technique to determine application task mapping and scheduling on a multiprocessor system and the voltage and frequency levels of each cores (offline DVFS) that minimize application computation and communication energy, simultaneously minimizing processor aging. The proposed technique incorporates (1) the effect of the voltage and frequency on the temperature of a core; (2) the effect of neighboring core voltage and frequency on the temperature (spatial effect); (3) pipelined execution and cyclic dependencies among tasks; and (4) the communication energy component which often constitutes a significant fraction of the total energy for multimedia applications. The temperature model proposed here can be easily integrated in the design space exploration for multiprocessor systems. Experiments conducted with applications modeled as synchronous data-flow graphs in conjunction with the HotSpot tool for temperature modeling clearly demonstrate the quality and the speed-up achieved using the proposed approach. Further, they also show 40% savings in energy consumption with 6% increase in system lifetime.</td>
</tr>
<tr>
<td>09:45</td>
<td>5.3.4</td>
<td>RECOVERY-BASED RESILIENT LATENCY-INSENSITIVE SYSTEMS</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Speakers:</strong> Yuankai Chen¹, Xuan Zeng² and Hai Zhou¹</td>
</tr>
<tr>
<td></td>
<td></td>
<td>¹Northwestern University, US; ²Fudan University, CN</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>As the interconnected delay is becoming a larger fraction of the clock cycle time, the conventional global stalling mechanism, which is used to correct error in general synchronous circuits, would be no longer feasible because of the expensive timing cost for the stalling signal to travel across the circuit. In this paper, we propose recovery-based resilient latency-insensitive systems (RLISs) that efficiently integrate error-recovery techniques with latency-insensitive design to replace the global stalling. We first demonstrate a baseline RLIS as the motivation of our work that uses additional output buffer which guarantees that only correct data can enter the output channel. However this baseline RLIS suffers from performance degradations even when errors do not occur. We propose a novel improved RLIS that allows erroneous data to propagate in the system. Equipped with improved queues that prevent accumulation of erroneous data, the improved RLIS retains the system performance. We provide theoretical studies that analyze the impact of errors on system performance and the queue sizing problem. We also theoretically prove that the improved RLIS performs no worse than the global stalling mechanism. Experimental results show that the improved RLIS has 40.3% and even 3.1% throughput improvements compared to the baseline RLIS and the infeasible global stalling mechanism respectively, with less than 10% hardware overhead.</td>
</tr>
<tr>
<td>10:00</td>
<td>IP2-</td>
<td>A LINUX-GOVERNOR BASED DYNAMIC RELIABILITY MANAGER FOR ANDROID MOBILE DEVICES</td>
</tr>
<tr>
<td></td>
<td>10, 80</td>
<td><strong>Speakers:</strong> Pietro Mercati, Andrea Bartolini, Francesco Paterna, Tajana Simunic Rosang and Luca Beninzi</td>
</tr>
<tr>
<td></td>
<td></td>
<td>¹UCSD, US; ²University of Bologna, IT</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reliability is a major concern in multiprocessors. Dynamic Reliability Management (DRM) aims at trading off processor performance with lifetime. The state-of-the-art publications study only the theory supported by simulation. This paper presents the first complete software implementation, working on a real hardware, of a low-overhead, Android-compatible workload-aware DRM Governor for mobile multiprocessors. We discuss the design challenges and the run-time overhead involved. We show the effectiveness of our governor in guaranteeing the predefined target lifetime and show that it achieves up to 100% of lifetime improvement with respect to traditional governors, while providing comparable performance for critical applications.</td>
</tr>
</tbody>
</table>
The session addresses yield analysis due to timing variations as well as various flip flop design techniques improving timing margins under variability.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>10:01</td>
<td>IP2-</td>
<td>YIELD AND TIMING CONSTRAINED SPARE TSV ASSIGNMENT FOR THREE-DIMENSIONAL INTEGRATED CIRCUITS</td>
<td>Yu-Guang Chen, Kuan-Yu Lai, Ming-Chao Lee, Yiju Shi, Wing-Kai Hon and Shih-Chieh Chang</td>
</tr>
<tr>
<td></td>
<td>11,</td>
<td></td>
<td>National Tsing Hua University, TW; MediaTek Inc., TW; Missouri University of Science and Technology, US</td>
</tr>
<tr>
<td></td>
<td>182</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10:02</td>
<td>IP2-</td>
<td>COMPILER-DRIVEN DYNAMIC RELIABILITY MANAGEMENT FOR ON-CHIP SYSTEMS UNDER VARIABILITIES</td>
<td>Saeemeen Rehman, Florian Kriebel, Muhammad Shafique and Jörg Henkel, Karlsruhe Institute of Technology (KIT), DE</td>
</tr>
<tr>
<td></td>
<td>12,</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>568</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10:00</td>
<td></td>
<td>End of session</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5.4 Prediction and optimization of timing variations

Date: Wednesday, March 26, 2014
Time: 08:30 - 10:00
Location / Room: Konferenz 2

Chair: Antonio Rubio, UPC Barcelona, ES, Contact Antonio Rubio

Co-Chair: Marisa López Vallejo, UPM Madrid, ES, Contact Marisa Lopez Vallejo

The session addresses yield analysis due to timing variations as well as various flip flop design techniques improving timing margins under variability.

<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>5.4.1</td>
<td>EFFICIENT HIGH-SIGMA YIELD ANALYSIS FOR HIGH DIMENSIONAL PROBLEMS</td>
<td>Moning Zhang, Zuochang Ye and Yan Wang, Tsinghua National Laboratory for Information Science and Technology, Institute of Microelectronics, Tsinghua University, CN</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>09:00</td>
<td>5.4.2</td>
<td>SUB-THRESHOLD LOGIC CIRCUIT DESIGN USING FEEDBACK EQUALIZATION</td>
<td>Mahmoud Zangeneh and Ajay Joshi, Boston University, US</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>09:30</td>
<td>5.4.3</td>
<td>STOCHASTIC ANALYSIS OF BUBBLE RAZOR</td>
<td>Guowei Zhang and Peter Beerelz</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Tsinghua University, CN; Khir. of Southern California, US</td>
</tr>
</tbody>
</table>

Authors
5.5 Boosting the Scalability of Formal Verification Technologies

Date: Wednesday, March 26, 2014
Time: 08:30 - 10:00
Location / Room: Konferenz 3

Chair:
Fahim Rahim, Atrenta, FR; Contact Fahim Rahim

Co-Chair:
Bernd Becker, University of Freiburg, DE; Contact Bernd Becker

While the industrial usage of formal methods has proliferated in the past decade, the capacity limitations of these techniques remains a challenge to their applicability. This session introduces a set of novel advances to boost the scalability of numerous state-of-the-art verification core technologies.

Time | Label | Presentation Title |
---|---|---|
08:30 | 5.5.1 | SCALABLE LIVENESS VERIFICATION FOR COMMUNICATION FABRICS |

Authors:
Sebastian Joosten and Julien Schmaltz, Open University, NL

Abstract
In the realm of multi-core processors and systems-on-chip, communication fabrics constitute a key element. A large number of queues and distributed control are two important aspects of this class of designs. These aspects make decomposition and abstraction techniques difficult to apply. For this class of designs, the application of formal methods is a real challenge. In particular, the verification of liveness properties is often intractable. Communication fabrics can be seen as a set of queues and flip-flops interconnected by combinational logic. Based on this simple but powerful observation, we propose a novel method for liveness verification. Our method directly applies to Register Transfer Level designs. The essential aspects of our approach are (1) to abstract away from the details of queue implementations and (2) an efficient encoding of liveness properties in an SMT instance. Experimental results are promising. Designs with hundreds of queues can be analysed for liveness within minutes.

Source: Fahim Rahim, Atrenta, FR; Contact Fahim Rahim
The papers in this session consider new ways to realize both Boolean and non-Boolean logic.

**Marco Ottavi**, University of Rome "Tor Vergata", IT, Co-Chair:

**Mehdi Tahoori**, KIT, DE, Chair:

**Location / Room**: Konferenz 4

**Time** | **Label** | **Presentation Title** | **Authors**
---|---|---|---
09:00 | 5.5.2 | PROPERTY DIRECTED INVARIANT REFINEMENT FOR PROGRAM VERIFICATION | Tobias Welp1 and Andreas Kuehlmann2
1 | UC Berkeley, US; 2Coverity, Inc., US

**Abstract**
We present a novel, sound, and complete algorithm for deciding safety properties in programs with static memory allocation. The new algorithm extends the program verification paradigm using loop invariants with a counterexample guided abstraction refinement (CEGAR) loop where the refinement is achieved by strengthening loop invariants using the QF$_{BV}$, BV generalization of Property Directed Reachability (PDR). We compare the algorithm with other approaches to program verification and report experimental results.

09:30 | 5.5.3 | SIMPLE INTERPOLANTS FOR LINEAR ARITHMETIC | Christoph Scholl1, Florian Pignosc1, Stefan Disch1 and Ernst Althaus2
1 | University Freiburg, DE; 2University Mainz, DE

**Abstract**
Craig interpolation has turned out to be an essential method for many applications in formal verification. In this paper we focus on the computation of simple interpolants for the theory of linear arithmetic with rational coefficients. We successfully minimize the number of linear constraints in the final interpolant by several methods including proof transformations, linear programming, and SMT solving. Experimental results comparing the approach to standard methods from the literature prove the effectiveness of the approach and show reductions of up to 70% in the number of linear constraints.

09:45 | 5.5.4 | TIGHTENING BDD-BASED APPROXIMATE REACHABILITY WITH SAT-BASED CLAUSE GENERALIZATION | Gianpiero Cabodi, Paolo Pasini, Stefano Quer and Danilo Vendraminetti, Politecnico di Torino, IT

**Abstract**
In the framework of symbolic model checking, BDD-based approximate reachability is potentially much more scalable than its exact counterpart. However, its practical applicability is highly limited by its static approach to abstraction, and the intrinsic difficulty to find an acceptable trade-off between accuracy and (memory/time) complexity. In this paper, we explore the use of CNF clauses, and of recent improvements in SAT algorithms, as additional players in BDD-based reachability. Cube generalization, a core step of the IC3 model checking algorithm, is the process of finding a minimal sub-clause, by removing as many literals as possible, such that it over-approximates a set of reachable states while excluding the cube. Generalization is used in IC3 to refine clause-based representations of state sets. We use it, in both the inductive and non-inductive version, in order to strengthen BDD-based representations of state sets, computed by Machine By Machine (MBM) and Frame By Frame (FBF) over-approximate forward traversal algorithms. The resulting approach benefits from the orthogonal power of BDD and CNF representations, and it improves the scalability of BDD-based methods. Preliminary experimental results confirm that this approach can provide tighter representations of reachable state sets. Applications include fully BDD-based engines, as well as using over-approximate state sets as invariants or constraints in SAT-based model checking.

10:00 | IP2-831 | MAKE IT REAL: EFFECTIVE FLOATING-POINT REASONING VIA EXACT ARITHMETIC | Miriam Leeser1, Saoni Mukherjee1, Jaideep Ramachandran1 and Thomas Wahl2
1 | Northeastern University, US; 2Northeastern University, Boston, US

**Abstract**
Floating-point arithmetic is widely used in scientific computing. While many programmers are subliminally aware that floating-point numbers only approximate the reals, few are cognizant of the dangers this entails for programming. Such dangers range from tolerable rounding errors in sequential programs, to unexpected, divergent control flow in parallel code. To address these problems, we present a decision procedure for floating-point arithmetic (FPA) that exploits the proximity to real arithmetic (RA), via a lossless reduction from FPA to RA. Our procedure does not involve any form of bit-blasting or bit-vectorization, and can thus generate much smaller back-end decision problems, albeit in a more complex logic. This tradeoff is beneficial for the exact and reliable analysis of parallel scientific software, which tends to give rise to large but benignly structured formulas. We have implemented a prototype decision engine and present encouraging results analyzing such software for numerical accuracy.

10:00 | End of session | Coffee Break in Exhibition Area
On Tuesday-Thursday the coffee and lunch breaks will be located in the Exhibition Area (Terrace Level).

### 5.6 Emerging logic technologies

**Date**: Wednesday, March 26, 2014

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

**Location / Room**: Konferenz 4

**Chair**: Mehdi Tahoori, KIT, DE, Contact Mehdi Tahoori

**Co-Chair**: Marco Ottavi, University of Rome "Tor Vergata", IT, Contact Marco Ottavi

The papers in this session consider new ways to realize both Boolean and non-Boolean logic. Potential implementations are based on graphene, spin, and resonance energy transfer.
Time | Label | Presentation Title | Authors
--- | --- | --- | ---
08:30 | 5.6.1 | RETLAB: A FAST DESIGN-AUTOMATION FRAMEWORK FOR ARBITRARY RET NETWORKS | Mohammad Mottaghi, Arjun Rallapalli and Chris Dwyer, Duke University, US
Abstract
Resonance energy transfer (RET) circuits are networks of photo-active molecules that can implement arbitrary logic functions. The nanoscale size of these structures can bring high-density computation to new domains, e.g., in vivo sensing and computation. A key challenge in the design of a RET network is to find, among a huge set of configurations (i.e., design space), the optimum choice and arrangement of molecules on a nanostructure. The prohibitively large size of the design space makes it impractical to evaluate every possible configuration, motivating the need for design-space pruning to be integrated into the design flow. To this end, we have developed a computer-aided design framework, called RETLab, that enables structured pruning of the design space to extract a sufficiently small subset, which is fully evaluated and ranked based on user-defined metrics to yield the best configuration. More importantly, we have developed a new RET-simulation algorithm, which is several orders of magnitude (e.g., for a 4-node network, one million times) faster than the conventional Monte-Carlo-based simulation (MCS). This speedup in configuration evaluation enables a significantly more extensive design-space exploration with fewer and less constrained heuristics, compared to existing RET-network design methods which are ad-hoc and rely on MCS for configuration evaluation.

09:00 | 5.6.2 | DESIGN OF 3D NANOMAGNETIC LOGIC CIRCUITS: A FULL-ADDER CASE STUDY | Robert Perricone, X. Sharon Hu, Joe Nahas and Michael Némier, University of Notre Dame, US
Abstract
Nanomagnetic logic (NML) is a "beyond-CMOS" technology that combines logic and memory capabilities through field-coupled interactions between nanoscale magnets. NML is intrinsically non-volatile, low-power, and radiation-hard when compared to CMOS equivalents. Moreover, there have been numerous demonstrations of NML circuit functionality within the last decade. These fabricated structures typically employ devices with in-plane magnetization to move and process data. However, in-plane layouts imply circuits and interconnects in only two dimensions (2D), which makes signal routing -- and hence circuits -- more complex. In this paper, we introduce NML circuits that move and process data in three dimensions (3D). We employ devices with perpendicular magnetic anisotropy (PMA) (i.e., out-of-plane magnetization states) and discuss their behavior when utilized in 3D designs. Furthermore, we provide a systematic design approach for 3D NML circuits using a threshold full adder as a case study. We compare our 3D adder to 2D adders to highlight the benefits of 3D NML circuits, which include simpler signal routing and a smaller area footprint.

09:30 | 5.6.3 | HIGHLY ACCURATE SPICE-COMPATIBLE MODELING FOR SINGLE- AND DOUBLE-GATE GNRFETS WITH STUDIES ON TECHNOLOGY SCALING | Morteza Ghofrani1, Ying Yu Chen2, Amit Sangal2 and Deming Chen2
1:University of Tehran, IR; 2:University of Illinois at Urbana Champaign, US
Abstract
In this paper, we present a highly accurate closed-form compact model for Schottky-Barrier-type Graphene Nano-Ribbon Field Effect Transistors (SB-GNRFETs). This is a physics-based analytical model for the current-voltage (I-V) characteristics of SB-GNRFETs. We carry out accurate approximations of Schottky barrier tunneling, channel charge and current, which provide improved accuracy while maintaining compactness. This SPICE-compatible compact model surpasses the existing model [15] in accuracy, and enables efficient circuit-level simulations of future GNRFET-based circuits. The proposed model considers various design parameters and process variation effects, including graphene-specific edge roughness, which allows complete and thorough exploration and evaluation of SB-GNRFET circuits. We are able to model both single- and double-gate SB-GNRFETs, so we can evaluate and compare these two types of SB-GNRFET. We also compare circuit-level performance of SB-GNRFETs with multi-gate (MG) Si-CMOS for a scalability study in future generation technology. Our circuit simulations indicate that SB-GNRFET has an energy-delay product (EDP) advantage over Si-CMOS; the EDP of the ideal SB-GNRFET (assuming no process variation) is \(1.3\) of that of Si-CMOS, while the EDP of the non-ideal case with process variation is \(13.6\) of that of Si-CMOS. Finally, we study technology scaling with SB-GNRFET and MG Si-CMOS. We show that the EDP of ideal (non-ideal) SB-GNRFET is \(-0.88\%\) (54\%) EDP of that of Si-CMOS as the technology nodes scales down to 7 nm.

09:45 | 5.6.4 | REWRITING FOR THRESHOLD LOGIC CIRCUIT MINIMIZATION | Chia-Chun Lin1, Chun-Yao Wang1, Yuan-Chih Chen2 and Ching-Yi Huang1
1:Dept. of Computer Science, National Tsing Hua University, TW; 2:Dept. of Computer Science and Engineering, Yuan Ze University, TW
Abstract
Recently, there have been many works focusing on synthesis, verification, and testing of threshold circuits due to the rapid development in efficient implementation of threshold logic circuits. To minimize the hardware cost of threshold circuit implementation, this paper proposes a heuristic that consists of rewiring operations and a simplification procedure. Additionally, a subset of input vectors of a gate, called critical-effect vectors, are proved to be complete for formally verifying the equivalence of two threshold logic circuits. To minimize the hardware cost of threshold circuit implementation, this paper proposes a heuristic that consists of rewiring operations and a simplification procedure. Additionally, a subset of input vectors of a gate, called critical-effect vectors, are proved to be complete for formally verifying the equivalence of two threshold logic circuits.

10:00 | IP2-17, 238 | WIDTH MINIMIZATION IN THE SINGLE-ELECTRON TRANSISTOR ARRAY SYNTHESIS | Chian-Wei Liu1, Chang-En Chiang1, Ching-Yi Huang1, Chun-Yao Wang1, Yuan-Chih Chen2, Suman Datta3 and Vijaykrishnan Narayanan4
1:Dept. of Computer Science, National Tsing Hua University, TW; 2:Dept. of Computer Science and Engineering, Yuan Ze University, TW; 3:Department of Electrical Engineering, The Pennsylvania State University, US; 4:Department of Computer Science and Engineering, The Pennsylvania State University, US
Abstract
Power consumption has become one of the primary challenges to meet the Moore's law. For reducing power consumption, Single-Electron Transistor (SET) at room temperature has been demonstrated as a promising device for extending Moore's law due to its ultra-low power consumption during operation. Prior work has proposed an automated mapping approach for SET architecture which focuses on minimizing the number of hexagons in an SET array. However, the area of an SET array is more related to the width. Consequently, in this work, we propose an approach for width minimization of the SET arrays. The experimental results show that the proposed approach saves 26% of width compared with the state-of-the-art for a set of MCNC and IWL5 benchmarks while spending similar CPU time.
### 5.7 Test Generation and Optimization

**Date:** Wednesday, March 26, 2014  
**Time:** 08:30 - 10:00  
**Location / Room:** Konferenz 5

**Chair:**  
Xiaqing Wen, Kyushu Institute of Technology, JP,  
Contact Xiaqing Wen

**Co-Chair:**  
Grzegorz Mrugalski, Mentor Graphics, PL,  
Contact Grzegorz Mrugalski

The session covers generation of tests for different fault models including interconnect opens, interconnect for 3D memories, and small delay faults. Additionally test optimization for SoC designs is presented.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>10:00</td>
<td>IP2-18, 704</td>
<td>AREA MINIMIZATION SYNTHESIS FOR RECONFIGURABLE SINGLE-ELECTRON TRANSISTOR ARRAYS WITH FABRICATION CONSTRAINTS</td>
<td>Yi-Hang Chen, Jan-Yu Chen and Juuni-Dar Huang, Department of Electronics Engineering, National Chiao Tung University, TW</td>
</tr>
<tr>
<td>10:02</td>
<td>IP2-19, 247</td>
<td>SOFTWARE-BASED PAULI TRACKING IN FAULT-TOLERANT QUANTUM CIRCUITS</td>
<td>Alexandru Paler, Simon Devitt, Kae Nemoto and Ila Pollan1</td>
</tr>
<tr>
<td>10:00</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Coffee Break** in Exhibition Area  
On Tuesday-Thursday the coffee and lunch breaks will be located in the Exhibition Area (Terrace Level).

### 5.7.1 EFFICIENT SMT-BASED ATPG FOR INTERCONNECT OPEN DEFECTS

**Speakers:**  
Dominik Erb1, Karsten Schelb1r, Matthias Sauer1 and Bernd Becker2  
1University of Freiburg, Chair of Computer Architecture, DE; 2University of Freiburg, DE

**Abstract**  
Interconnect opens are known to be one of the predominant defects in nanoscale technologies. However, automatic test pattern generation for open faults is challenging, because of their rather unstable behaviour and the numerous electric parameters which need to be considered. Thus, most approaches try to avoid accurate modeling of all constraints and use simplified fault models in order to detect as many faults as possible or make assumptions which decrease both complexity and accuracy. This paper presents a new SMT-based approach which for the first time supports the Robust Enhanced Aggressor Victim model without restrictions and handles oscilations. It is combined with the first open fault simulator fully supporting the Robust Enhanced Aggressor Victim model and thereby accurately considering unknown values.

**Experimental results show the high efficiency of the new method outperforming previous approaches by up to two orders of magnitude.**

### 5.7.2 INTERCONNECT TEST FOR 3D STACKED MEMORY-ON-LOGIC

**Speakers:**  
Mottaqiallah Taouil, Mahmoud Masadeh, Said Hamdioui and Erik Jan Marinissen  
1Delft University of Technology, NL; 2IMEC, BE

**Abstract**  
Three-dimensional stacked IC (3D-SIC) technology based on Through-Silicon Vias (TSVs) provides numerous advantages as compared to traditional 2D-ICs. A potential application is memory stacked on logic, providing enhanced throughput, and reduced latency and power consumption. However, testing the TSV interconnects between the two dies is challenging, as both the memory and the logic die might come from different manufacturers. Currently, no standard exists and the proposed solutions fail to address dynamic and ime-critical faults (at speed testing). In addition, memory vendors have not been in favor to put additional DIT structures such as JTAG for interconnect testing on their memory devices. This paper proposes a new Memory Based Interconnect Test (MBIT) approach for 3D stacked memories. Our test patterns are applied by read and write instructions to the memory and are validated by a case study where a 3D memory is assumed to be stacked on a MIPS64 processor. The main benefits of the MBIT approach are: (1) zero area overhead, (2) the ability to detect both static and dynamic faults and perform at speed testing, (3) flexibility in applying any test pattern, as this can be executed by the CPU on the logic die and (4) extreme short test execution time.
will come to an end due to physical and economic restrictions, the integration of systems (e.g. by stacking dies, or by adding sensor functions) shows a way to maintain the growth in complexity.

System Integration using 3D technology is a very promising way to cope with current and future requirements for electronic systems. Since the pure shrinking of devices (known as “More Moore”) implementations are evaluated on a real case study for which we have production test data from 10,000 devices.

This paper discusses new implementations of the predictive alternate test strategy that exploit model redundancy in order to improve test confidence. The key idea is to build during the training phase, not only one regression model for each specification as in the classical implementation, but several regression models. This redundancy is then used during the testing phase to identify suspect predictions and remove the corresponding devices from the alternate test flow. In this paper, we explore various options for implementing model redundancy, based on the use of different indirect measurement combinations and/or different partitions of the training set. The proposed implementations are evaluated on a real case study for which we have production test data from 10,000 devices.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
</table>
| 09:30 | 5.7.3 | AN EFFECTIVE APPROACH TO AUTOMATIC FUNCTIONAL PROCESSOR TEST GENERATION FOR SMALL-DELAY FAULTS | Andreas Riefert, Lyl Ciganda, Matthias Sauern, Paolo Bernardi, Matteo Sonza Reorda and Bernd Beckers 1
1University of Freiburg, DE; 2Politecnico di Torino, IT; 3Politecnico di Torino - DAUN, IT |

Abstract
Functional microprocessor test methods provide several advantages compared to DFT approaches, like reduced chip cost and at speed execution. However, the automatic generation of functional test patterns is an open issue. In this work we present an approach for the automatic generation of functional microprocessor test sequences for small-delay faults based on Bounded Model Checking. We utilize an ATPG framework for small-delay faults in sequential, non-scan circuits and propose a method for constraining the input space for generating functional test sequences (i.e., test programs). We verify our approach by evaluating the minMAPS microprocessor. In our experiments we were able to reach over 97% fault efficiency. To the best of our knowledge, this is the first fully automated approach to functional microprocessor test for small-delay faults.

09:45 | 5.7.4 | MULTI-SITE TEST OPTIMIZATION FOR MULTI-VDD SOCS USING SPACE- AND TIME-DIVISION MULTIPLEXING | Fotios Vartzios, Chrysovalantis Kavoussianakis, Krishnendu Chakrabarty, Rubin Parekh and Anivid Jain 1
1University of Ioannina, GR; 2Department of Computer Science and Engineering, University of Ioannina, GR; 3Duke University, US; 4Texas Instruments, IN |

Abstract
Even though system-on-chip (SoC) testing at multiple voltage settings significantly increases test complexity, the use of a different shift frequency at each voltage setting offers parallelism that can be exploited by time-division multiplexing (TDM) to reduce test length. We show that TDM is especially effective for small-bitwidth and heavily loaded test-access mechanisms (TAMs), thereby tangibly increasing the effectiveness of multi-site testing. However, TDM suffers from some inherent limitations that do not allow the fullest possible exploitation of TAM bandwidth. To overcome these limitations, we propose space-division multiplexing (SDM), which complements TDM and offers higher multi-site test efficiency. We implement space- and time-division multiplexing (STDM) using a new, scalable test-time minimization method based on a combination of bin packing and simulated annealing. Results for industrial SoCs, highlight the advantages of the proposed optimization method.

10:00 | IP2-20, 50 | AN EFFICIENT TEMPERATURE-GRADIENT BASED BURN-IN TECHNIQUE FOR 3D STACKED ICS | Nima Aghaee, Zebo Peng and Petru Eles, Linköping University, SE |

Abstract
Burn-in is usually carried out with high temperature and elevated voltage. Since some of the early-life failures depend not only on high temperature but also on temperature gradients, simply raising up the temperature of an IC is not sufficient to detect them. This is especially true for 3D stacked ICs, since they have usually very large temperature gradients. The efficient detection of these early-life failures requires that specific temperature gradients are enforced as a part of the burn-in process. This paper presents an efficient method to do so by applying high power stimuli to the cores of the IC under burn-in through the test access mechanism. Therefore, no external heating equipment is required. The scheduling of the heating and cooling intervals to achieve the required temperature gradients is based on thermal simulations and is guided by functions derived from a set of thermal equations. Experimental results demonstrate the efficiency of the proposed method.

10:01 | IP2-21, 17 | TEST AND NON-TEST CUBES FOR DIAGNOSTIC TEST GENERATION BASED ON MERGING OF TEST CUBES | Ith Pomeranz, Purdue University, US |

Abstract
Test generation by merging of test cubes supports test compaction and test data compression. This paper describes a new approach to the use of test cube merging for the generation of compact diagnostic test sets. For this the paper uses the new concept of non-test cubes. While a test cube for a fault f1 detects the fault, a non-test cube for a fault f1 prevents the fault from being detected. Merging a test cube for a fault f1 and a non-test cube for a fault f1 produces a diagnostic test cube that distinguishes the two faults. The paper describes a procedure for diagnostic test generation based on merging of test and non-test cubes. Experimental results demonstrate that compact diagnostic test sets are obtained.

10:02 | IP2-22, 905 | NEW IMPLEMENTATIONS OF PREDICTIVE ALTERNATE ANALOG/RF TEST WITH AUGMENTED MODEL REDUNDANCY | Haithem Ayari, Florence Azais, Serge Bernard, Marlane Comte, Vincent Kerzerho and Michel Renovell, LIRMM, CNRS/Univ. Montpellier 2, FR |

Abstract
In this work we present an approach for the automatic generation of functional microprocessor test sequences for small-delay faults based on Bounded Model Checking. We utilize an ATPG framework for small-delay faults in sequential, non-scan circuits and propose a method for constraining the input space for generating functional test sequences (i.e., test programs). We verify our approach by evaluating the minMAPS microprocessor. In our experiments we were able to reach over 97% fault efficiency. To the best of our knowledge, this is the first fully automated approach to functional microprocessor test for small-delay faults.

10:00 | End of session | Coffee Break In Exhibition Area |

5.8 Hot Topic: System Integration - The Bridge between More than Moore and More Moore

Date: Wednesday, March 26, 2014
Time: 08:30 - 10:00
Location / Room: Exhibition Theatre

Organisers:
Manfred Dietrich, Fraunhofer IIS/EAS Dresden, DE, Contact Manfred Dietrich
Kai Hahn, University Siegen, DE, Contact Kai Hahn

Chair:
Kai Hahn, University Siegen, DE, Contact Kai Hahn

Co-Chair:
Kai Hahn, University Siegen, DE, Contact Kai Hahn

System Integration using 3D technology is a very promising way to cope with current and future requirements for electronic systems. Since the pure shrinking of devices (known as "More Moore") will come to an end due to physical and economic restrictions, the integration of systems (e.g. by stacking dies, or by adding sensor functions) shows a way to maintain the growth in complexity.
System Integration design challenges from different perspectives, ranging from design technology over MEMS product engineering and 3D interconnect to automotive cyber physical systems.

**DESIGN TECHNOLOGY FOR 3-D INTEGRATED SYSTEMS**

**Speaker:** Andy Hening, Fraunhofer IIS/EAS, DE  
**Abstract**
More than Moore technologies (MM) enable the dense integration of different circuits in a package. The short length and smallspacing of wires enable high-speed and highly parallel interconnects between system parts as e.g. processor and memory. In the first part of the presentation we will give an overview on the current status of MM from system-in-package up to 3D stacking with trench silicon vias at interposers or stacked directly. The second part of the presentation is dedicated to the design of MM systems. One challenge is the tight integration of analog and digital dies that requests the consideration of several electrical and multi-physical interactions e.g. thermal management, power distribution and electromagnetic compatibility stronger than in 2D SoC design. The second challenge is the wide design space opened by MM. It request new methods that guides the designer to find the best trade-off between system performance and production costs. By the means of Processor and WideIO memory integration at silicon interposer, that increases the memory bandwidth in future high-end applications we demonstrate new EDA methods for design space exploration, estimation of routing congestion and interposer routing.

**SEMICONDUCTOR PACKAGING IS BACK TO EUROPE - ADVANCES IN SYSTEM INTEGRATION IN WAFER LEVEL PACKAGING**

**Speaker:** Steffen Kreinehr, NANIUM S.A. - Niederfassung Dresden, DE  
**Abstract**
Different market segments from mobile communication and consumer to automotive see the increasing need to focus on system integration on less space instead of single components or functional groups. This drives advanced semiconductor packaging to diversify and become fairly more complex, but at the same time an integrated functional part of the system. The demand for more and more diversified functionality on same or even less space drives the development of "More-than-More" (MM) solutions in the packaging world. The keyword is again "System-in-Package" (SiP). Chip-Package-Board Co-Design and Co-Development are essential key for success. Besides some theory, the paper will show some real product examples where system integration in the package saved up to 4X space on the board for the same functionality with even more performance. While today the majority of SiP is still realized using laminated organic substrate interposers, the need to close the gap to System-on-Chip (SoC) performance is driving closer distances of the single functional elements to each other. This can be realized by Fan-Out Water Level Packaging (FO-WLP) technologies, like eWLB (embedded Wafer Level Ball Grid Array), which overcomes Fan-In Wafer Level Packaging (Fi-WLP) limitations especially in terms of system integration, keeping the advantages of scalability and cost efficient batch processing. In the paper the good progress made to develop eWLB as technology platform will be shown, mainly using FO-WLP as enabler for System-in-Package on Wafer Level (WLSP).

**MEMS AND 3-D PRODUCT ENGINEERING - TECHNOLOGY DESIGN FOR SYSTEM INTEGRATION**

**Speaker:** Kai Hahn, University Siegen, DE  
**Abstract**
Taking into account the diversity of technologies from die manufacturing to packaging it becomes clear that for product engineering of integrated systems such as MEMS or stacked 3D circuits the constraints and inter-dependencies of design and manufacturing are of special interest. The configuration of these technologies is strongly application specific and design methods differ completely from the approach known from the development of conventional two dimensional ICs. The presentation will cover methods and tools for technology design in the area of MEMS as well as for 3D integration.

**3D-TSV-HUB: POTENTIALS AND CHALLENGES FOR VERTICAL INTERCONNECTS IN NETWORKS-ON-CHIPS**

**Speaker:** Andreas Herkersdorf, TU München, DE  
**Abstract**
Sophisticated Network-On-Chips (NoCs) will form the backbone for on-chip communication in future System-On-Chip designs. Already in conventional planar systems the synthesis of application specific NoCs is a complex task. When shifting to a stacked die environment, further degrees of freedom are added and a large design space is created. Through Silicon Vias (TSVs) are deployed for building vertical NoC links in a 3D systems. However, TSVs are cost intensive under several aspects. Area consumption is high due to large TSV diameters (compared to planar metal layer interconnect) and keep out areas in intermediate die layers. Furthermore, mechanical induced stress can lead to runtime failures and an overall low system yield. Therefore, in order to ensure that a minimum number of TSVs are performance and cost efficiently operated to their full capacity, a 3D-TSV-Hub has been proposed to support smart mapping of communication flows onto TSVs during NoC synthesis. Mechanisms to handle production and runtime failures are integrated into the 3D-TSV-Hub concept and also considered during synthesis. Design aspects like compliance to thermal requirements and manufacturing process requirements influence the optimization of 3D-NoCs. In order to consider such factors, we operate the NoC synthesis tool in interplay with Design Space Exploration and 3D-Floorplanner tools of project partners.

**SENSORS AND POWER DRIVERS, BRIDGE BETWEEN SYSTEM ENVIRONMENT AND COMPUTING**

**Speaker:** Jochen Reisinger, Infineon Technologies Austria AG, AT  
**Abstract**
There are many main drivers enabling the most significant innovations in system solutions which are based on, or only supported by, electronics. No doubt, the best known driver is the availability of deep submicron technology nodes for the implementation of computing functions (More Moore). But competitive system architectures and functional system partitionings (technology selection) strongly depend on highly efficient interfaces to the system environment. Those interfaces are supporting many functions as for example: a) The sensing of physical parameters (temperature, pressure, speed, power, ...); b) Providing power and control signals for actuators (drivers for motors, pumps, ...); c) Providing power for the computing system (including safety, power up/down); d) Interfacing to human bodies and/or other system elements and e) Communication of system operational and control data (WiFi, Bluetooth, ...). Those interfaces ask for highly efficient (in terms of space, power, performance, ... and cost) 3D integration technologies and design methodologies. Infineon examples for sensors and drivers will be presented.

**CONCLUSIONS AND DISCUSSION**

**Speaker:** Manfred Dietrich, Fraunhofer, DE

10:00 End of session  
Coffee Break in Exhibition Area  
On Tuesday-Thursday the coffee and lunch breaks will be located in the Exhibition Area (Terrace Level).
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.

### IP2-1

**Title**: FAST AND ACCURATE COMPUTATION USING STOCHASTIC CIRCUITS  
**Speakers**: Armin Alagh and John P. Hayes, University of Michigan - Ann Arbor, US  
**Abstract**  
Stochastic computing (SC) is a low-cost design technique that has great promise in applications such as image processing. SC enables arithmetic operations to be performed on stochastic bit-streams using ultra-small and low-power circuitry. However, accurate computations tend to require long run-times due to the random fluctuations inherent in stochastic numbers (SNs). We present novel techniques for SN generation that lead to better accuracy/run-time trade-offs. First, we analyze a property called progressive precision (PP) which allows computational accuracy to grow systematically with run-time. Second, borrowing from Monte Carlo methods, we show that SC performance can be greatly improved by replacing the usual pseudo-random number sources by low-discrepancy (LD) sequences that are predictably progressive. Finally, we evaluate the use of LD stochastic numbers in SC, and show they can produce significantly better and more accurate results than existing stochastic designs.

### IP2-2

**Title**: DRAM-BASED COHERENT CACHES AND HOW TO TAKE ADVANTAGE OF THE COHERENCE PROTOCOL TO REDUCE THE REFRESH ENERGY  
**Speakers**: Zoran Jakić and Ramon Canal, Universitat Politècnica de Catalunya, ES  
**Abstract**  
Recent technology trends has turned DRAMs into an interesting candidate to substitute traditional SRAM-based on-chip memory structures (i.e. register file, cache memories). Nevertheless, a major problem to introduce these cells is that they lose their state (i.e. value) over time, and they have to be refreshed. This paper proposes the implementation of coherent caches with DRAM cells. Furthermore, we propose to use the coherence state to tune the refresh overhead. According to our analysis, an average of up to 57% of refresh energy can be saved. Also, comparing to the caches implemented in SRAMs total energy savings are on average up to 39% depending of the refresh policy with a performance loss below 8%.

### IP2-3

**Title**: REDUCING SET-ASSOCIATIVE L1 DATA CACHE ENERGY BY EARLY LOAD DATA DEPENDENCE DETECTION (ELD3)  
**Speakers**: Alan Bardubryan, Magnus Själander, David Whalley and Per Larsson-edefors  
**Abstract**  
Fast set-associative level-one data caches (L1-DCs) access all ways in parallel during load operations for reduced access latency. This is required in order to resolve data dependencies as early as possible in the pipeline, which otherwise would suffer from stall cycles. A significant amount of energy is wasted due to this fast access, since the data can only reside in one of the ways. While it is possible to reduce L1 DC energy usage by accessing the tag and data memories sequentially, hence activating only one data way on a tag match, this approach significantly increases execution time due to an increased number of stall cycles. We propose an early load data dependency detection (ELD3) technique for in-order pipelines. This technique makes it possible to detect if a load instruction has a data dependency with a subsequent instruction. If there is no such dependency, then the tag and data access for the load are sequentially performed so that only the data way in which the data resides is accessed. If there is a dependency, then the tag and data arrays are accessed in parallel to avoid introducing additional stall cycles. For the MBench benchmark suite, the ELD3 technique enables about 49% of all load operations to access the L1-DC sequentially. Based on 65-nm data using commercial SRAM blocks, the proposed technique reduces L1-DC energy by 13%.

### IP2-4

**Title**: DISTRIBUTED COOPERATIVE SHARED LAST-LEVEL CACHING IN TILED MULTIPROCESSOR SYSTEM ON CHIP  
**Speakers**: Preethi Parayil Mana Damodaran, Stefan Wallentowitz and Andreas Herkersdorf  
**Abstract**  
In a shared-memory based tiled many-core system-on-chip architecture, memory accesses present a huge performance bottleneck in terms of access latency as well as bandwidth requirements. The best practice approach to address this issue is to provide a multi-level cache hierarchy and a suitable cache-coherency mechanism. This paper presents a method to increase the memory access performance in distributed-directory-coherency protocol based tiled many-core systems. The proposed method introduces an alternate design for the system-wide shared last-level caches (LLC) placed between the memory and the node private caches (NPC). The proposed system-wide shared LLC layer is distributed over the entire network and it interacts with the home directories of specific cache lines. Results from simulating SPEC2000 benchmark applications executed on a SystemC model of the proposed design show a minimum performance improvement of 20-25% when compared to a model without the shared cache layer at the expense of an additional 2% of the total cache memory space (NPC + LLC memory). In addition, the proposed design shows a minimum 7-15% and an average 14-15% improvement in performance in comparison to centralized system-wide shared LLC of equivalent size and dynamic mapped distributed LLC of equivalent size respectively.

### IP2-5

**Title**: DESIGN OF SAFETY CRITICAL SYSTEMS BY REFINEMENT  
**Speakers**: Alex Illasov, Arseniy Alekseyov, Danil Sokolov and Andrey Mokhov  
**Abstract**  
An increasingly large number of safety-critical embedded systems rely on software to prevent and mitigate hazards occurring due to design errors and unexpected interactions of the system with its users and the environment. Implementing a safety instrumented function in the way advocated by the traditional software methods requires an intimate understanding and thorough validation of a complex ecosystem of programming languages, compilers, operating systems and hardware. We propose to consider an alternative where a system designer, for each individual problem, creates in a correct-by-construction manner both the design of a system and its compilation and execution infrastructure. This permits an uninterrupted chain of a formal correctness argument spanning from formalised requirements all the way to the gate-level characterisation of an execution environment. The past decade of advances in verification technology turned the mechanical verification of large-scale models into a reality while the pressure of certification makes the cost of a formally verified development routine increasingly acceptable. The proposed technique fits the Grand Challenge for Computer Research posed by Hoare in 2003, namely, development of a Verifying Compiler which not only mechanically translates a given program from one language to another but also verifies its correctness according to a formal specification. This allows meeting the most stringent software certification requirements such as SIL 4. We illustrate the idea with a small case study developed using the Event-B modelling notation and tools.
IP2-6 ENERGY OPTIMIZATION IN ANDROID APPLICATIONS THROUGH WAKELOCK PLACEMENT

Speakers: Faisal Alam, Preeti Ranjan Panda, Nikhil Tripathi, Namita Sharma and Sanjiv Narayanan

IIT Delhi, IN; Kalypto Design Systems, IN; I Hindian Institute of Technology Delhi, IN

Abstract
Energy efficiency is a critical factor in mobile systems, and a significant body of recent research efforts has focused on reducing the energy dissipation in mobile hardware and applications. The Android OS Power Manager provides programming interface routines called wakelocks for controlling the activation state of devices on a mobile system. An appropriate placement of wakelock acquire and release functions in the application can make a significant difference to the energy consumption. In this paper, we propose a data flow analysis based strategy for determining the placement of wakelock statements corresponding to the uses of devices in an application. Our experimental evaluation on a set of Android applications show significant (up to 32%) energy savings with the proposed optimization strategy.

IP2-7 A WEAR-LEVELING-AWARE DYNAMIC STACK FOR PCM MEMORY IN EMBEDDED SYSTEMS

Speakers: Qingan Li, Yanxiang He, Yong Chen, Chun Xuez, Nan Jiang and Xao Xue

Wuhan University & City University of Hong Kong, CN; Wuhu University, CN; City University of Hong Kong, CN

Abstract
Phase Change Memory (PCM) is a promising DRAM replacement in embedded systems due to its attractive characteristics such as extremely low leakage power, high storage density and good scalability. However, PCM’s low endurance constrains its practical applications. In this paper, we propose a Wear Leveling aware dynamic stack to extend PCM’s lifetime when it is adopted in embedded systems as main memory. Through a dynamic stack, the memory space is circularly allocated to stack objects, and thus an even usage of PCM memory is achieved. The experimental results show that the proposed method can significantly reduce the write variation on PCM cells and enhance the lifetime of PCM memory.

IP2-8 LIFETIME HOLES AWARE REGISTRY ALLOCATION FOR CLUSTERED VLIW PROCESSORS

Speakers: Xuemeng Zhang1, Hui Wu2, Haiyan Sun1 and Jingling Xue3

National University of Defense Technology, CN; The University of New South Wales, AU; UNSW, AU

Abstract
This paper presents an on-the-fly register allocator which dynamically detects and utilizes lifetime holes for clustered VLIW processors. A lifetime hole is an interval in which a variable does not contain a valid value. A register holding a lifetime hole can be allocated to another variable whose live range fits in the lifetime hole, leading to more efficient utilization of registers. We propose efficient techniques for dynamically utilizing lifetime holes and incorporate these techniques into our on-the-fly register allocator. We have simulated our register allocator and a linear scan register allocator without considering lifetime holes by using the MediaBench II benchmark suite. Our simulation results show that our register allocator reduces the number of spills by 12.5%, 11.7%, 12.7%, for three different processor models, respectively.

IP2-9 A LOW-POWER, HIGH-PERFORMANCE APPROXIMATE MULTIPLIER WITH CONFIGURABLE PARTIAL ERROR RECOVERY

Speakers: Cong Liu1, Je Han2 and Fabrizio Lombardi3

University of Alberta, CA; Northeastern University, US

Abstract
Approximate circuits have been considered for error-tolerant applications that can tolerate some loss of accuracy with improved performance and energy efficiency. Multipliers are key arithmetic circuits in many such applications such as digital signal processing (DSP). In this paper, a novel approximate multiplier with a lower power consumption and a shorter critical path than traditional multipliers is proposed for high-performance DSP applications. This multiplier leverages a newly-designed approximate adder that limits its carry propagation to the nearest neighbors for fast partial product accumulation. Different levels of accuracy can be achieved through a configurable error recovery by using different numbers of most significant bits (MSBs) for error reduction. The approximate multiplier has a low mean error distance, i.e., most of the errors are not significant in magnitude. Compared to the Wallace multiplier, a 16-bit approximate multiplier implemented in a 28nm CMOS process shows a reduction in delay and power of 20% and up to 69%, respectively. It is shown that by utilizing an appropriate error recovery, the proposed approximate multiplier achieves similar processing accuracy as traditional exact multipliers but with significant improvements in power and performance.

IP2-10 A LINUX-GOVERNOR BASED DYNAMIC REALIABILITY MANAGER FOR ANDROID MOBILE DEVICES

Speakers: Pietro Mercatini, Andrea Bartolini, Francesco Patema, Tajana Simunic Rosing and Luca Benini

UICSD, US; University of Bologna, IT

Abstract
Reliability is a major concern in multiprocessors. Dynamic Reliability Management (DRM) aims at trading off processor performance with lifetime. The state-of-the-art publications study only the theory supported by simulation. This paper presents the first complete software implementation, working on a real hardware, of a low-overhead, Android-compatible workload-aware DRM Governor for mobile multiprocessors. We design the challenges and the run-time overhead involved. We show the effectiveness of our governor in guaranteeing the predefined target lifetime and show that it achieves up to 100% of lifetime improvement with respect to traditional governors, while providing comparable performance for critical applications.

IP2-11 YIELD AND TIMING CONSTRAINED SPARE TSV ASSIGNMENT FOR THREE-DIMENSIONAL INTEGRATED CIRCUITS

Speakers: Yu-Guang Chen1, Kuan-Yu Lai2, Ming-Chao Lee3, Yuyu Shi1, Wang-Kai Hong1 and Shih-Chieh Chang1

National Tsing Hua University, TW; MediaTek Inc., TW; University of Missouri, US

Abstract
Through Silicon Via (TSV) is a critical enabling technique in three-dimensional integrated circuits (3D ICs). However, it may suffer from many reliability issues. Various fault-tolerance mechanisms have been proposed in literature to improve yield, at the cost of significant area overhead. In this paper, we focus on the structure that uses one spare TSV for a group of original TSVs, and study the optimal assignment of spare TSVs under yield and timing constraints to minimize the total area overhead. We show that such problem can be modeled through constrained graph decomposition. An efficient heuristic is further developed to address this problem. Experimental results show that under the same yield and timing constraints, our heuristic can reduce the area overhead induced by the fault-tolerance mechanisms by up to 38%, compared with a seemingly more intuitive nearest-neighbor based heuristic.

IP2-12 COMPILER-DRIVEN DYNAMIC RELIABILITY MANAGEMENT FOR ON-CHIP SYSTEMS UNDER VARIABILITIES

Speakers: Semeen Rehman, Florian Kriebel, Muhammad Shafique and Jörg Henkel, Karlsruhe Institute of Technology (KIT), DE

Abstract
This paper presents a novel Dynamic Reliability Management System (DiReMS) for on-chip systems that performs resilience-driven resource allocation and mapping. It accounts for both the tasks’ resilience properties and heterogeneous error recovery features of different cores. DiReMS also chooses a reliable task version (out of multiple reliability-aware transformed options) depending upon the reliability level of the allocated core. In case of error detection, rollbacks are performed. Our system provides 70%-87% improved task reliability compared to a timing reliability-optimizing core assignment, i.e. minimizing the probability of deadline misses (with EDF scheduling).
MINIMIZING STATE-OF-HEALTH DEGRADATION IN HYBRID ELECTRICAL ENERGY STORAGE SYSTEMS WITH ARBITRARY SOURCE AND LOAD PROFILES

Speakers:
Yanzhi Wang, Xue Lin, Qing Xie, Naehyuck Chang and Massoud Pedram
1University of Southern California, US; 2Seoul National University, KR

Abstract
Hybrid electrical energy storage (HEES) systems consisting of heterogeneous electrical energy storage (EES) elements are proposed to exploit the strengths of different EES elements and hide their weaknesses. The cycle life of the EES elements is one of the most important metrics. The cycle life is directly related to the state-of-health (SoH), which is defined as the ratio of full charge capacity of an aged EES element to its designed (or nominal) capacity. The SoH degradation models of battery in the previous literature can only be applied to charging/discharging cycles with the same state-of-charge (SoC) swing. To address this shortcoming, this paper derives a novel SoH degradation model of battery for charging/discharging cycles with arbitrary patterns. Based on the proposed model, this paper presents a near-optimal charge management policy focusing on extending the cycle life of battery elements in the HEES systems while simultaneously improving the overall cycle efficiency.

DYNAMIC FLIP-FLOP CONVERSION TO TOLERATE PROCESS VARIATION IN LOW POWER CIRCUITS

Speakers:
Mehrzad Nejat, Bijan Alizadeh and Ali Afsal Kusha, School of Electrical and Computer Eng., College of Eng., University of Tehran, IR

Abstract
A novel time borrowing method called dynamic Flip-Flop conversion is presented in this paper. A timing violation predictor detects the violations halfway in the critical path and dynamically converts the critical Flip-Flop to a latch. This way, time borrowing benefits of latches are utilized in a Flip-Flop based design which is more adaptable with Computer-Aided Design tools. The overhead of this method is smaller than that of similar methods due to the elimination of delay elements. According to the post-synthesis simulations and Monte-Carlo analysis of Spice simulations on some ITC99 benchmark circuits, the power overhead of the proposed method is about 15% and 19% smaller than that of Soft-Edge-Flip-Flop and Dynamic- Clock-Stretching circuits respectively in a simple case of about 40% yield improvement. This overhead would be relatively even smaller for higher performance and yield improvements.

A LOW POWER AND ROBUST CARBON NANOTUBE 6T SRAM DESIGN WITH METALLIC TOLERANCE

Speakers:
Luo Sun, Jimson Mathew, Rishad Shafik, Dhiraj Pradhan and Zhen Lin
1University of Bristol, GB; 2University of Southampton, GB

Abstract
Carbon nanotube field-effect transistor (CNTFET) is envisioned as a promising device to overcome the limitations of traditional CMOS based MOSFETs due to its favourable physical properties. This paper presents a novel six-transistor (6T) static random access memory (SRAM) bitcell design using CNTFETs. Extensive validations and comparative analyses are carried out with the proposed SRAM design using SPICE based simulations. We show that the proposed CNTFET based SRAM has a significantly better static noise margin (SNM) and write ability margin (WAM) compared to a 6T based standard 6T bitcell, equivalent to isolated read-port 6T cell based on CNTFET, while consuming less dynamic power. We further demonstrate that it exhibits higher robustness under process, voltage and temperature (PVT) variations when compared with the traditional CMOS SRAM cell designs. Furthermore, metallic CNTs removal technique is used considering metallic tolerance to make the proposed SRAM design more reliable.

MAKE IT REAL: EFFECTIVE FLOATING-POINT REASONING VIA EXACT ARITHMETIC

Speakers:
Miriam Leeser, Saoni Mukherjee, Jaideep Ramachandran and Thomas Wahl
1Northeastern University, US; 2Northeastern University, Boston, US

Abstract
Floating-point arithmetic is widely used in scientific computing. While many programmers are subliminally aware that floating-point numbers only approximate the reals, few are cognizant of the dangers this entails for programming. Such dangers range from tolerable rounding errors in sequential programs, to unexpected, divergent control flow in parallel code. To address these problems, we present a decision procedure for floating-point arithmetic (FPA) that exploits the proximity to real arithmetic (RA), via a lossless reduction from FPA to RA. Our procedure does not invoke any form of bit-blasting or bit-visualization, and can thus generate much smaller back-end decision problems, albeit in a more complex logic. This tradeoff is beneficial for the exact and reliable analysis of parallel scientific software, which tends to give rise to large but benignly structured formulas. We have implemented a prototype decision engine and present encouraging results analyzing such software for numerical accuracy.

WIDTH MINIMIZATION IN THE SINGLE-ELECTRON TRANSISTOR ARRAY SYNTHESIS

Speakers:
Chian-Wei Liu, Chang-En Chiang, Ching-Yi Huang, Chun-Yao Wang, Yung-Chih Chen, Suman Datta and Vijaykrishnan Narayanan
1Dept. of Computer Science, National Tsing Hua University, TW; 2Dept. of Computer Science and Engineering, Yuan Ze University, TW; 3Department of Electrical Engineering, The Pennsylvania State University, US; 4Department of Computer Science and Engineering, The Pennsylvania State University, US

Abstract
Power consumption has become one of the primary challenges to meet the Moore's law. For reducing power consumption, Single-Electron Transistor (SET) at room temperature has been demonstrated as a promising device for extending Moore's law due to its ultra-low power consumption during operation. Prior work has proposed an automated mapping approach for SET architecture which focuses on minimizing the number of hexagons in an SET array. However, the area of an SET array is more related to the width. Consequently, in this work, we propose an approach for width minimization of the SET arrays. The experimental results show that the proposed approach saves 26% of width compared with the state-of-the-art for a set of MCNC and IWLS 2005 benchmarks while spending similar CPU time.

AREA MINIMIZATION SYNTHESIS FOR RECONFIGURABLE SINGLE-ELECTRON TRANSISTOR ARRAYS WITH FABRICATION CONSTRAINTS

Speakers:
Yi-Hang Chen, Jian-Yu Chen and Jiun-Dar Huang, Department of Electronics Engineering, National Chiao Tung University, TW

Abstract
As fabrication processes exploit even deeper submicron technology, power dissipation has become a crucial issue for most electronic circuit and system designs nowadays. In particular, leakage power is becoming a dominant source of power consumption. Recently, the reconfigurable single-electron transistor (SET) array has been proposed as an emerging circuit design style for continuing Moore's Law due to its ultra-low power consumption. Several automated synthesis approaches have been developed for the reconfigurable SET array in the past few years. Nevertheless, all of those existing methods consider fabrication constraints, which are mandatory, merely in late synthesis stages. In this paper, we propose a synthesis algorithm, featuring both variable reordering and product term reordering, for area minimization. In addition, our algorithm takes those mandatory fabrication constraints into account in early stages for better outcomes. Experimental results show that our new method can achieve an area reduction of up to 24% as compared to current state-of-the-art techniques.
Software-Based Pauli Tracking in Fault-Tolerant Quantum Circuits

Speakers:
Alexandru Paler, Simon Devitt, Kae Nemoto and Ila Polan
\[University of Passau, DE; National Institute of Informatics, JP\]

Abstract

The realization of large-scale quantum computing is no longer simply a hardware question. The rapid development of quantum technology has resulted in dozens of control and programming problems that should be directed towards the classical computer science and engineering community. One such problem is known as Pauli tracking. Methods for implementing quantum algorithms that are compatible with crucial error correction technology utilize extensive quantum teleportation protocols. These protocols are intrinsically probabilistic and result in correction operators that occur as byproducts of teleportation. These byproduct operators do not need to be corrected in the quantum hardware itself, but are tracked through the circuit and output results. This tracking is routinely ignored in quantum information as it is assumed that tracking algorithms will eventually be developed. In this work we help fill this gap and present an algorithm for tracking byproduct operators through a quantum computation.

An Efficient Temperature-Gradient Based Burn-In Technique for 3D Stacked ICs

Speakers:
Nima Aghaee, Zebo Peng and Petru Eles, Linköping University, SE

Abstract

Burn-in is usually carried out with high temperature and elevated voltage. Since some of the early-life failures depend not only on high temperature but also on temperature gradients, simply raising up the temperature of an IC is not sufficient to detect them. This is especially true for 3D stacked ICs, since they have usually very large temperature gradients. The efficient detection of these early-life failures requires that specific temperature gradients are enforced as a part of the burn-in process. This paper presents an efficient method to do so by applying high power stimuli to the cores of the IC under burn-in through the test access mechanism. Therefore, no external heating equipment is required. The scheduling of the heating and cooling intervals to achieve the required temperature gradients is based on thermal simulations and is guided by functions derived from a set of thermal equations. Experimental results demonstrate the efficiency of the proposed method.

Test and Non-Test Cubes for Diagnostic Test Generation Based on Merging of Test Cubes

Speaker:
Irith Pomeranz, Purdue University, US

Abstract

Test generation by merging of test cubes supports test compaction and test data compression. This paper describes a new approach to the use of test cube merging for the generation of compact diagnostic test sets. For this the paper uses the new concept of non-test cubes. While a test cube for a fault f1 detects the fault, a non-test cube for a fault f2 prevents the fault from being detected. Merging a test cube for a fault f1 and a non-test cube for a fault f2 produces a diagnostic test cube that distinguishes the two faults. The paper describes a procedure for diagnostic test generation based on merging of test and non-test cubes. Experimental results demonstrate that compact diagnostic test sets are obtained.

New Implementations of Predictive Alternate Analog/RF Test with Augmented Model Redundancy

Speakers:
Haithem Ayari, Florence Azais, Serge Bernard, Mariane Comte, Vincent Kerzerho and Michel Renovell, LIRMM, CNRS/Univ. Montpellier 2, FR

Abstract

This paper discusses new implementations of the predictive alternate test strategy that exploit model redundancy in order to improve test confidence. The key idea is to build during the training phase, not only one regression model for each specification as in the classical implementation, but several regression models. This redundancy is then used during the testing phase to identify suspect predictions and remove the corresponding devices from the alternate test flow. The proposed implementations are evaluated on a real case study for which we have production test data from 10,000 devices.
KAOLIN: A MODEL-BASED EDA TOOL TO PROGRAM, REUSE OR RETARGET EMBEDDED SYSTEMS ON FPGAS.

Authors:
Yvan Eustache, Dominique Blouin, Mikaël Lanœ, Jean-Philippe Giguert and Philippe Coussey, Lab-STICC, Université de Bretagne-Sud, FR

Abstract
The demonstration presents the Kaolin EDA tool to improve and speed-up embedded systems development on FPGAs. It provides modeling abstractions to shield the user from implementation details and prevent frequent time-consuming errors. It allows you to reuse legacy projects and IP's and retarget them to other platforms with different back-end tools. The Kaolin technology is based on models of components, platforms and FPGA development tools. It allows automating platform-independent system generation including vendor tool files and scripts, verification and high-level analysis, and template-based documentation generation. Kaolin nicely fits in the development flow as a bridge between the user and low level FPGA vendor tools. User appropriation is facilitated: it requires no new language to be learnt; it allows the import of legacy codes (VHDL) and the software (C,C++) to hardware migration with built-in High-Level Synthesis capabilities. Kaolin can be customized to meet domain and user-specific requirements. During the demonstration, Kaolin will be used to quickly implement a control and signal processing system deployed on a FPGA and embedded on a radio-controlled toy car.

More information...

SECURE CLOUD-BASED WORKFLOW-AS-A-SERVICE (WFAAS) ENVIRONMENT WITH ROLE-BASED-ACCESS-CONTROL (RBAC) FOR SOC DESIGN

Authors:
Sai Manoj P. D., Sai Manoj P. D.1, Hao Yu1 and Joseph Lee2
1Nanyang Technological University, SG; 2Silicon Cloud International, US

Abstract
The SoC design process requires multiple EDA tools, custom IP's, and technology design kit from multiple providers. The design environment needs to be secure and collaborative. These requirements can be realized by using an integrated cloud based Workflow-as-a-Service (WFAAS) design environment. We demonstrate a cloud-based design environment for a SoC design with multiple CPU cores and analog IP's. This design environment uses an innovative Role-Based-Access-Control user security model where designers interact through a web portal dashboard to perform the design workflows.

More information...

MOTORRAIN: MODEL-BASED DESIGN AND VIRTUAL INTEGRATION OF AN INTELLIGENT AND SAFE ELECTRICAL POWERTRAIN

Authors:
Sven Rosinger, Mahen Fakh and Jörg Walter, OFFIS - Institut für Informatik, DE

Abstract
Hardware prototypes and hardware in the loop simulations are commonly used during embedded vehicle- and motor-control unit design. This demonstrator presents a platform that is an order of magnitude cheaper than existing systems but still easy to integrate into present workflows: Within an existing model-driven design methodology, a real-time hardware simulation is performed using the Raspberry Pi single-board computer to simulate an e-motor with little development effort and in conjunction with an industrial motor control unit.

More information...

PHARAON: PARALLEL AND HETEROGENEOUS ARCHITECTURES FOR REAL-TIME APPLICATIONS

Authors:
Luciano Lavagno1, Mihai Lazarescu1, Hector Posadas2 and Eugenio Villar2
1Politecnico di Torino, IT; 2Universidad de Cantabria, ES

Abstract
In this demo, we will present the work-in-progress of the EU FP7 PHARAON project, started in September 2011. The first objective of the project is the development of new techniques and tools capable to assist the developer in the development of parallel embedded systems, from executable specifications to target-specific implementation and debugging on a multicores platform. This tool chain offers and implements several parallelization strategies, reflecting the functional and non-functional constraints of the system, and driving the designer into incremental parallelization and adaptation steps. The second objective of the project is to develop monitoring and control techniques in the middleware of the system capable to automatically adapt platform services to application requirements and therefore reduce power consumption transparently. The demo will cover specifically: the software parallelization tool suite, the parallel software modeling and code generation suite.

More information...

LARA: THE LARA COMPILER SUITE

Authors:
Joao Bispo, Pedro Pinto, Ricardo Nobre, Tiago Carvalho and Joao Cardoso, Universidade do Porto, PT

Abstract
LARA is an aspect-oriented programming (AOP) language which allows the description of sophisticated code instrumentation schemes, advanced mapping strategies including conditional decisions, based on hardware/software resources, and of sophisticated sequences of compiler transformations. Furthermore, LARA provides mechanisms for controlling all elements of a toolchain in a consistent and systematic way, using a unified programming interface. We present three compiler tools developed around the LARA technology, MATISSE, MANET and ReflectC. MATISSE is a compiler which 1) allows analyses and transformations on MATLAB code and 2) generates C code from the MATLAB code. MATISSE can be fully controlled through LARA aspects, which can define the type and shape of MATLAB variables, specify code insertion/removal actions, and define specialization directives and other additional information. MATISSE can output transformed MATLAB code and specialized C code. The knowledge provided by the LARA aspects allows MATISSE to generate C tailored to specific targets (e.g., use statically declared arrays to be compliant with the high-level synthesis tools such as Catapult C). MANET is a source-to-source compiler for ANSI C based on Cetus, and is controlled using LARA aspects. MANET manages to leverage the expressiveness and modularity of LARA to query and manipulate the Cetus AST, providing an easy compilation flow with main goal of code instrumentation and code transformations. LARA aspects allow for a simple selection of program elements in the code which can be analyzed or transformed, by either consulting their attributes or applying actions. Thus, MANET can be used to provide information reports based on compiler analyses, to implement sophisticated code instrumentation strategies, or to perform code optimizations and transformations. ReflectC is a C compiler based on CoSy's compiler framework. CoSy's configurability and regetatableity make ReflectC particularly effective for exploration of compiler transformations and optimizations on possible architecture variations, and it is being used for hardware/software co-design and design space exploration (DSE). We will present demos of the tools and the use of LARA aspects and strategies to guide our suite of compilation tools providing: 1) C code generation from MATLAB code, according to information provided by LARA aspects; 2) Instrumentation of C code to be used for collecting specific compile and runtime information (e.g., execution time, range of values for specific variables, custom profiling); 3) User-controlled compiler optimizations targeting several architectures and DSE of sequences of compiler optimizations bearing in mind performance improvements. In addition to presenting examples for each of the tools of the LARA compilation suite, we show an execution of the complete toolchain, controlled by LARA aspects.

More information...
Due to the breakdown of Dennard scaling, the percentage of a silicon chip that can switch at full frequency is dropping exponentially with each process generation. This utilization wall forces designers to ensure that, at any point in time, large fractions of their chips are effectively dark silicon, i.e., significantly underclocked or idle for large periods of time. As exponentially larger fractions of a chip’s transistors become dark, silicon area becomes an exponentially cheaper resource relative to power and energy consumption. This shift is driving a new class of architectural techniques that “spend” area to “buy” energy efficiency. All of these techniques seek to introduce new forms of heterogeneity into the computational stack. This work examines four key approaches—the four horsemen—that have emerged as top contenders for thriving in the dark silicon design regime. In particular, the three talks present the newest trends and developments starting with the problem of Dennard Scaling and how it mandates new design constraints followed by the problem of power delivery and cooling, and concluding with the newest directions in efficient resource management for many-core systems.

**Authors:**
Michael Taylor, University of California, San Diego, US

**Abstract**
Due to the breakdown of Dennard scaling, the percentage of a silicon chip that can switch at full frequency is dropping exponentially with each process generation. This utilization wall forces designers to ensure that, at any point in time, large fractions of their chips are effectively dark silicon, i.e., significantly underclocked or idle for large periods of time. As exponentially larger fractions of a chip’s transistors become dark, silicon area becomes an exponentially cheaper resource relative to power and energy consumption. This shift is driving a new class of architectural techniques that “spend” area to “buy” energy efficiency. All of these techniques seek to introduce new forms of heterogeneity into the computational stack. This work examines four key approaches—the four horsemen—that have emerged as top contenders for thriving in the dark silicon design regime. In particular, the three talks present the newest trends and developments starting with the problem of Dennard Scaling and how it mandates new design constraints followed by the problem of power delivery and cooling, and concluding with the newest directions in efficient resource management for many-core systems.
heterogeneous multi-core architectures and accelerators (where heterogeneity stems from different device technologies) and (b) modeling efforts at all levels of the chip hierarchy (i.e., from undesirable] features of new device technologies; (2) highlight how new transistor technologies could impact von Neumann architectures; a particular emphasis will be placed on (a) This "vertically integrated" session is focused on emerging transistor technologies - particularly devices that operate at low voltages and that have steep slopes.

Michael Niemier, University of Notre Dame, US,
Chair:
Organisers:
Time:
Date:
6.2 Embedded Tutorial: Emerging Transistor Technologies: From Devices to Architectures

12:00 Time
12:30 Time
End of session
Lunch Break in Exhibition Area
Sandwich lunch

6.2 Embedded Tutorial: Emerging Transistor Technologies: From Devices to Architectures

Date: Wednesday, March 26, 2014
Time: 11:00 - 12:30
Location / Room: Konferenz 6

Organisers:
Michael Niemier, University of Notre Dame, US; Contact Michael Niemier
X. Sharon Hu, University of Notre Dame, US; Contact X. Sharon Hu

Chair:
Michael Niemier, University of Notre Dame, US; Contact Michael Niemier

This "vertically integrated" session is focused on emerging transistor technologies - particularly devices that operate at low voltages and that have steep slopes. It will: (1) introduce desirable (and undesirable) features of new device technologies; (2) highlight how new transistor technologies could impact von Neumann architectures; a particular emphasis will be placed on (a) heterogeneous multi-core architectures and accelerators (where heterogeneity stems from different device technologies) and (b) modeling efforts at all levels of the chip hierarchy (i.e., from the device-level to the architectural level); (3) illustrate how new device technologies could lead to significant improvements in the performance/efficiency of non-von Neumann architectures. Notably, talks (2) and (3) will identify roles for new device technologies in analog digital systems with an end goal of improved application-level performance/efficiency.

11:30 6.1.2 INTEGRATED MICROFLUIDIC POWER GENERATION AND COOLING FOR BRIGHT SILICON MPSoCs
Speakers:
Mohamed M. Sabry, Arvind Shidhan, Patrick Ruchu, David Atienza and Bruno Micheo
1EPFL, CH; IBM Research, CH
Abstract
The soaring demand for computing power in our digital information age has produced, as collateral undesirable effect, a surge in power consumption and heat density for Multiprocessors System-on-Chip (MPSoC). Accordingly, significant portion of the energy consumed in state-of-the-art MPSoCs is dissipated in cooling. The remaining energy is used for computation, and causes the temperature ramp-up to operating conditions that already preclude operating all the cores at maximum performance levels, in order to prevent system overheating and failures. This situation is set to worsen as shipments of high-end (i.e., even denser) many-core servers are increasing at a 25% compound annual growth rate. With more power demands, MPSoCs will face a power delivery wall due to the reliability limitations of the underlying power delivery medium. Thus, state-of-the-art worst-case power and cooling delivery solutions are reaching their limits and it will no longer be possible to power up simultaneously all the available on-chip cores (situated as the existence of dark silicon); hence, drastically limiting the benefits of technology scaling. In this paper we propose a disruptive approach to overcome the prevailing worst-case power and cooling provisioning paradigm for MPSoCs. This proposed approach integrates MPSoC with an on-chip microfluidic fuel cell network for joint cooling delivery and power supply (i.e., local power generation and delivery). By providing an alternative mean to power delivery integrated with cooling, MPSoCs are expected to gain in IO connectivity. Thanks to this disruptive technology, we can envision the removal of the current limits of power delivery and heat dissipation in server designs, subsequently avoiding dark silicon in future MPSoCs and enabling new perspectives in future energy-proportional computing architecture designs.

11:30 6.2.1 ENERGY EFFICIENT COMPUTING WITH TUNNEL FETS
Speakers:
Adrian Ionescu, Arnab Biswas, Nilay Dagtekin and Livio Lattanzio, Nanolab, Ecole Polytechnique Fédérale de Lausanne, CH
Abstract
This paper will review the state-of-the-art in energy efficient computing using tunnel FETs from device to circuit level, including digital IC and memory applications. At device level we will particularly discuss the major challenges remaining for tunnel FETs, with particular emphasis on: (i) selection of the most appropriate material systems and band-gap engineering of heterostructure Tunnel FETs to simultaneously offer best performance trade-off: low Ioff, high Ion, high Ion/Ioff, subthermal swing over more than 4 decades of current, and operation below 0.3V; (ii) specifically optimized device design (i.e. field aligned to the tunneling path, avoidance of super-linear onset, minimize Miller effect), (iii) understanding the role of defects for BTBT and providing appropriate control, (iv) understanding and controlling parameter sensitivity and variability, (v) accurate physics-based BTBT modeling of heterojunction tunnel FETs. We will detail the Electron Hole Bilayer Tunnel FET (EHBTFET), as switch candidate for sub-0.1V operation exploiting tunneling through a bias-induced electron-hole bilayer based on a calibrated quantum-mechanical simulator. We will make performance projections for EHBTFET complementary logic compared to CMOS logic of same dimensions and using recent energy benchmarking. Finally, the design and use of Tunnel FETs as capacitorless DRAM cells, implemented as a double-gate (DG) fully-depleted Silicon-On-Insulator (FD-SOI) architecture will be reported and its principle, embodiment and scalability discussed. We will present recent experimental results on Tunnel FET DRAM memory operation schemes and demonstrate its potential for ultra-low power memories. In conclusion, this paper demonstrates that Tunnel FETs stand as the most promising steep slope switch candidates to reduce the supply voltage below 0.3 V and offer significant power dissipation savings for digital computing.
### 6.2.2 MODELING STEEP SLOPE DEVICES: FROM CIRCUITS TO ARCHITECTURES

**Speakers:**
- Kartik Swaminathan, Moon Seok Kim, Nandhini Chandramoorthy, Behnam Sedighi, Robert Penicones, Jack Sampson, Vijaykeshnan Narayanan
  
  1Pennsylvania State University, US; 2The Pennsylvania State University, US; 3University of Notre Dame, US; 4Penn State University, US

**Abstract**
Steep slope devices, with Heterojunction Tunnel FETs (TFETs) in particular, have been proposed as a viable solution to overcome the subthreshold slope limitation in existing CMOS technology and achieve ultra-low voltage operation with acceptable performance. However, state-of-the-art FinFET technologies continue to demonstrate superior performance than steep slope devices in application domains demanding peak single threaded performance. In this context, we examine different computing paradigms where TFET technologies can be used, not just as a drop in replacement, but as an additional parameter to augment the architectural design space. This greatly widens the scope of optimizations for performance and power. We investigate the tradeoffs between device and architectures in general purpose processors when performance, power and temperature are individually constrained. We also synthesize examples of domain-specific accelerators used in computer vision using in-house TFET standard cell libraries to demonstrate the energy benefits of designing TFET-based accelerators. We demonstrate that synthesizing these accelerators using TFETs reduces energy by over 6X in comparison to an equivalent iso-voltage CMOS-based design and by over 30% in comparison to an iso-performance CMOS design.

### 6.3 Management of Micro/Macro Renewable Energy Storage Systems

**Date:** Wednesday, March 26, 2014

**Time:** 11:00 - 12:30

**Location / Room:** Konferenz 1

**Chair:**
- Geoff Merrett, University of Southampton, UK; Contact Geoff Merrett

**Co-Chair:**
- Davide Brunelli, University of Trento, IT; Contact Davide Brunelli

Modern energy storage systems address all areas of power electronics, from micro-power energy harvesting systems to mega-watt Smart Grid systems. Papers in this session address novel approaches for on-chip power electronics operating under variable Vdd, and optimisation approaches to efficient design of smart grid energy storage.

### 6.3.1 ASYNCHRONOUS DESIGN FOR NEW-ON-CHIP WIDE DYNAMIC RANGE POWER ELECTRONICS

**Speakers:**
- Delong Shang, Xuefu Zhang, Fei Xia, and Alex Yakovlev
  
  1School of EEE, Newcastle University, GB; 2School of EEE, Newcastle University, GB; 3School of EEE, GB

**Abstract**
Asynchronous circuits will play an important role in microelectronic systems in the future, especially in energy harvesting and autonomous (EHA) systems where such circuits will be able to offer robustness and deliver high efficiency in a wide range of power-energy conditions. The concept of Capacitor Bank Block (CBB) mechanisms was proposed to form the basis of electronics for powering asynchronous loads. These mechanisms will benefit EHA systems by enabling effective co-scheduling of computational tasks and energy supply. This paper demonstrates how the CBB mechanisms can themselves be controlled by asynchronous circuits, thereby forming a new type of power delivery units (PDU) that will be able to deliver power to intelligent digital logic in future EHA systems. These PDUs are superior to traditional power converters largely because the latter can only regulate sufficiently high power and energy levels (regular and periodic) as well as their controllers require stable power levels themselves. This makes them unsuitable for intermittent and sporadic conditions inherent to EHA systems. In this paper, a novel asynchronous control for the CBB is described. Experiments and analysis of the new PDUs, comprising CBBS and asynchronous control, are presented and discussed in detail.
**REAL-TIME OPTIMIZATION OF THE BATTERY BANKS LIFETIME IN HYBRID RESIDENTIAL ELECTRICAL SYSTEMS**

**Speakers:**
Maurizio Rossi, Alessandro Toppano and Davide Brunei, University of Trento, IT

**Abstract**
We present a real-time optimization framework to manage Hybrid Residential Electrical Systems (HRES) with multiple Energy sources and heterogeneous storage units. HRES represents urban buildings where photovoltaic (PV) or other renewable sources are installed along with the traditional connection to the main grid. In this paper, heterogeneous storage units are used to realize energy buffers for the exceeding energy produced by the renewable when buildings and the grid are not available to accept it. We considered two different battery banks as electric energy storage, in particular lead-acid as the primary one for its low price and low self-discharge rate; while the lithium-ion chemistry is used as secondary bank because of the higher energy density and higher number of cycles. The proposed optimization strategy aims at maximizing the lifetime of the battery banks and to reduce the energy bill by managing the variability of the PV source, in price-varying scenarios. We used a Dynamic Programming (DP) algorithm to schedule off-line the use of the lead-acid bank minimizing the number of cycles and the Depth-of-Discharge (DoD) under given irradiance forecasts and user load profiles. Forecasts of the user loads and of the renewable energy intake are introduced in the optimization. Moreover a Real-Time scheme is introduced to manage the lithium bank and to minimize the need and the purchase of energy from the Grid when the actual demand does not fit the forecast. Our simulation results outperform the state of the art where the efficiency of both banks is not taken into consideration, even if complex approaches based on DP are used.

---

**OPTIMAL DIMENSIONING OF ACTIVE CELL BALANCING ARCHITECTURES**

**Speakers:**
Swaminathan Narayanawamy¹, Sebastian Steinhorst², Martin Lukasiewycz², Matthias Kauer² and Samanāt Chakraborty³

¹TUM CREATE, SG; ²TUM CREATE Singapore, SG; ³TUM CREATE Ltd., SG; ⁴TU Munich, DE

**Abstract**
This paper presents an approach to optimal dimensioning of active cell balancing architectures, which are of increasing relevance in EES for EV or stationary applications such as smart grids. Active cell balancing equalizes the state of charge of cells within a battery pack via charge transfers, increasing the effective capacity and lifetime. While optimization approaches have been introduced into the design of several aspects of EES, active cell balancing architectures have, until now, not been systematically optimized in terms of their components. Therefore, this paper analyzes existing architectures to develop design metrics for energy dissipation, installation volume, and balancing current. Based on these design metrics, a methodology to efficiently obtain Pareto-optimal configurations for a wide range of inductors and transistors at different balancing currents is developed. Our methodology is then applied to a case study, optimizing two state-of-the-art architectures using realistic balancing algorithms. The results give evidence of the applicability of systematic optimization in the domain of cell balancing, leading to higher energy efficiencies with minimized installation space.

---

**OPTIMAL DESIGN AND MANAGEMENT OF A SMART RESIDENTIAL PV AND ENERGY STORAGE SYSTEM**

**Speakers:**
Di Zhu¹, Yanhui Wang¹, Naehyuck Chang² and Massoud Pedram³

¹Univ. of Southern California, US; ²Seoul National University, KR

**Abstract**
Solar photovoltaic (PV) technology has been widely deployed in large power plants operated by utility companies. However, the home owners are not yet convinced of the saving cost benefits of this technology, and consequently, in spite of government subsidies, they have been reluctant to install PV systems in their homes. The main reason for this is the absence of a complete and truthful analysis which could explain to home owners under what conditions spending money on a PV system can actually save them money over a long-term, but known, time horizon. This paper thus presents a design and management mechanism for a smart residential energy system comprising PV modules, electrical energy storage banks, and conversion circuits connected to the power grid. First, we figure out how much savings can be achieved by a system with given PV modules and EES bank capacities by optimally solving the daily energy flow control problem of such a system. Based on the daily optimization results, we come up with the optimal system specifications with a fixed budget. Experiments are conducted for various electricity prices and different profiles of PV output power and load demand. Results show that the designed system breaks even in 6 years and in the system lifetime achieves up to 8% annual profit besides paying back the budget.

---

**DESIGN AND FABRICATION OF A 315 µm BONDWIRE MICRO-TRANSFORMER FOR ULTRA-LOW VOLTAGE ENERGY HARVESTING**

**Speakers:**
Ennio Macrelli¹, Ningning Wang¹, Sabal Roy², Michael Hayes², Rudi Paolo Pagano³, Marcio Tattagni³ and Aldo Romanil¹

¹DEI, University of Bologna, IT; ²Tyndall National Institute, UCC, IE; ³ICON-IEIT, University of Bologna, IT

**Abstract**
This paper presents a design study of a new topology for miniaturized bondwire transformers fabricated and assembled with standard IC bonding wires and toroidal ferrite (Fair-Rite 5975000801) as a magnetic core. The micro-transformer realized on a PCB substrate, enables the build of magnetics on-top-of-chip, thus leading to the design of high power density components. Impedance measurements in a frequency range between 100 kHz to 5 MHz, show that the secondary self-inductance is enhanced from 0.3 µH with an epoxy core to 315 µH with the ferrite core. Moreover, the micro-machined ferrite improves the coupling coefficient from 0.1 to 0.9 and increases the effective turns ratio from 0.5 to 35. Finally, a low-voltage IC DC-DC converter solution, with the transformer mounted on-top, is proposed for energy harvesting applications.

---

**PROVIDING REGULATION SERVICES AND MANAGING DATA CENTER PEAK POWER BUDGETS**

**Speakers:**
Baris Aksanî and Tajana Rosing, University of California San Diego, US

**Abstract**
Data centers are good candidates for providing regulation services in the power markets due to their large power consumption and flexibility. In this paper, we develop a model that explores the feasibility of data center participation in such markets. We use a battery-based design that can not only help with providing ancillary services, but also limit peak power costs without any workload performance degradation. The results of our study using data for a 21MW data center show up to $480,000/year savings can be obtained, corresponding to 1280 more servers providing services.

---

**THE ENERGY BENEFIT OF LEVEL-CROSSING SAMPLING INCLUDING THE ACTUATOR'S ENERGY CONSUMPTION**

**Speakers:**
Burkhard Hensel and Klaus Kabitzsch, Dresden University of Technology, DE

**Abstract**
When using level-crossing (also called send-on-delta) sampling in control loops, messages can be saved compared to periodic sampling without degrading control performance. While it is clear that reducing messages improves also the energy efficiency of battery-powered sensor devices, this can be disadvantageous for the energy efficiency of the actuator device. This paper addresses the question, under which conditions level-crossing sampling is also for the actuator device more energy-efficient than periodic sampling. It is shown that there is an optimum inter-sample interval. Methods for reaching this optimum by appropriate controller and transmission settings are given. The theory is demonstrated using several known, standardized wireless network protocols.
Formal methods are traditionally used to verify the correctness of hardware, software, or protocols. This session introduces a set of applications which extend the use of formal methods into new domains.

### 6.4 Power delivery and distribution

**Date:** Wednesday, March 26, 2014  
**Time:** 11:00 - 12:30  
**Location / Room:** Konferenz 2

**Chair:** Edith Beigné, CEA LETI Grenoble, FR, [Contact Edith Beigne](mailto:edith.beigne@ceatec.japan)  
**Co-Chair:** Dominik Helms, OFFIS Oldenburg, DE, [Contact Dominik Helms](mailto:dominik.helms@offis.de)

This session will present innovative solutions for power delivery in complex SoCs using configurable structures working over large condition range. Configurable DC-DC and LDOs architectures will be considered underlining power efficiency issues of off-chip and on-chip regulators. Fine-grain approaches are also proposed to deal with distributed in-die power generation reducing static and dynamic power in complex SoCs.

#### 6.4.1 DESIGN AND EVALUATION OF FINE-GRAINED POWER-GATING FOR EMBEDDED MICROPROCESSORS

**Speakers:** Masaaki Kondo1, Hiroaki Kobayashi, Ryushi Sakamoto, Motoki Wada, Jun Tsukamoto, Mitao Namikiz, Weihan Wang, Hideharu Amano, Kensaku Matsunaga, Masaru Kudo1, Kimiyoshi Usami1, Toshiya Komoda2 and Hiroshi Nakamura3  
1The University of Electro-Communications, JP; 2Tokyo University of Agriculture and Technology, JP; 3Keio University, JP; 4Sibaura Institute of Technology, JP  
**Abstract**  
Power-performance efficiency is still remaining a primary concern for microprocessor designers. One of the sources of power inefficiency for recent LSI chips is increasing leakage power consumption. Power-gating is a well known technique to reduce leakage power consumption by switching off the power supply to idle logic blocks. Recently, fine-grained power-gating is emerged as a technique to minimize leakage current during the active processor cycles by switching off and on a logic blocks in much finer temporal/spatial granularity. Though fine-grained power-gating is useful, a comprehensive evaluation and analysis has not been conducted on a real LSI chips. In this paper, we evaluate fine-grained run-time power-gating for microprocessors' functional units using a real embedded microprocessor. We also introduce an architecture and compiler co-operative power-gating scheme which mitigates negative power reduction caused by the energy overhead associated with fine-grained power-gating. The experimental results with a fabricated core shows that a hardware-based scheme saves power consumption of functional units by 44% and hardware compiler co-operative scheme further improves power efficiency by 5.9% when core temperature is 25℃.

#### 6.4.2 SUPERRANGE: WIDE OPERATIONAL RANGE POWER DELIVERY DESIGN FOR BOTH STV AND NTV COMPUTING

**Speakers:** Xin He1, Guohai Yan1, Yinhe Han1 and Xiaowei Li3  
1Institute of Computing Technology, Chinese Academy of Sciences; University of Chinese Academy of Sciences, CN; 2Institute of Computing Technology, Chinese Academy of Sciences, CN  
**Abstract**  
The load power range of modern processors is greatly enlarged because many advanced power management techniques like dynamic voltage frequency scaling, Turbo boosting, and Near Threshold Voltage technologies are incorporated. However, the power saving may be offset by power loss in power delivery; moreover, as the efficiency of power delivery varies greatly with different load conditions, conventional power delivery designs cannot maintain high efficiency over the entire voltage range. We propose SuperRange, a wide operational range power delivery scheme. SuperRange complements the power delivery capability of on-chip voltage regulator and off-chip voltage regulator. Experimental results show SuperRange has an average 70% power conversion efficiency over wide operational range which outperforms conventional power delivery schemes. And it also exhibits superior resilience to power-constrained systems.

#### 6.4.3 MODELING AND ANALYSIS OF DIGITAL LDOs WITH ADAPTIVE CONTROL FOR HIGH EFFICIENCY UNDER WIDE DYNAMIC RANGE DIGITAL LOADS

**Speakers:** Samantak Gangopadhyay, Youngtak Lee, Saad Bin Nasir and Arijit Raychowdhury, Georgia Institute of Technology, US  
**Abstract**  
Discrete time digital linear regulators, including low dropout regulators (LDOs) have become competitive in multi-Vcc digital systems for fine-grained spatio-temporal voltage regulation and distribution. However wide dynamic current range of the digital load circuits poses serious problems in maintaining stability and high efficiency at all corners. In this paper we present a control model for discrete time LDOs and demonstrate how online adaptive control can be employed for consistent performance and high efficiency across the load current range.

#### 6.5 Beyond EDA: Extending the Application Domain of Formal Methods

**Date:** Wednesday, March 26, 2014  
**Time:** 11:00 - 12:30  
**Location / Room:** Konferenz 3

**Chair:** Christoph Scholl, University of Freiburg, DE, [Contact Christoph Scholl](mailto:christoph.scholl@uni-freiburg.de)  
**Co-Chair:** Gianpiero Cabodi, Politecnico di Torino, IT, [Contact Gianpiero Cabodi](mailto:gianpiero.cabodi@polito.it)

Formal methods are traditionally used to verify the correctness of hardware, software, or protocols. This session introduces a set of applications which extend the use of formal methods into new domains.
domains. The first three papers demonstrate novel ways to bridge formal verification results into the synthesis domain. The fourth leverages formal reasoning to certify the correctness of photonic systems.

### Using MaxBMC for Pareto-Optimal Circuit Initialization

**Speakers:**
Sven Reimer, Matthias Sauer, Tobias Schubert and Bernd Becker, University of Freiburg, DE

**Abstract**
In this paper we present MaxBMC, a novel formalism for solving optimization problems in sequential systems. Our approach combines techniques from symbolic SAT based Bounded Model Checking (BMC) and incremental MaxSAT, leading to the first MaxBMC solver. In traditional BMC safety and liveness properties are validated. We extend this formalism: in case the required property is satisfied, an optimization problem is defined to maximize the quality of the reached witnesses. Further, we compare its qualities in different depths of the system, leading to Pareto-optimal solutions. We state a sound and complete algorithm that not only tackles the optimization problem but moreover verifies whether a global optimum has been identified by using a complete BMC solver as back-end. As a first reference application we present the problem of circuit initialization. Additionally, we give pointers to other tasks which can be covered by our formalism quite naturally and further demonstrate the efficiency and effectiveness of our approach.

### Partial Witnesses from Preprocessed Quantified Boolean Formulas

**Speakers:**
Martina Seidl and Robert Könighofer, JKU Linz, AT; TU Graz, AT

**Abstract**
For effectively solving quantified Boolean formulas (QBF) in prenex conjunctive normal form, preprocessors have shown to be indispensable. A preprocessor rewrites a formula in such a manner that information valuable for the solver is made explicit and irrelevant information is removed. For this purpose, rewriting techniques, which would be too costly when repeatedly applied during the solving process, are used. Unfortunately, many of these techniques are not model preserving and therefore incompatible with recent certification frameworks. In consequence, the application of a preprocessor prohibits the extraction of witnesses encoding a solution or a counterexample. In this paper, we show how to obtain an assignment for the variables of the outermost quantifier block as partial witness which is sufficient for many practical applications.

### Equivalence Checking for Function Pipelining in Behavioral Synthesis

**Speakers:**
Kecheng Hao, Sandip Roy and Fei Xie, 1Xilinx Inc., US; 2Strategic CAD Labs, Intel Corporation, US; 3Portland State University, US

**Abstract**
Function pipelining is a key transformation in high-level synthesis. However, synthesizing the complex pipeline logic is an error-prone process. Sequential equivalence checking (SEC) support is highly desired to provide confidence in the correctness of synthesized pipelines. However, SEC for function pipelining is challenging due to the significant difference between the behavioral specification and the synthesized RTL. Furthermore, function pipelines include hardware logic for dynamically inserting “bubbles” (pipeline stalls), which bring additional difficulties in equivalence checking. We develop an SEC framework for behaviorally synthesized function pipelines by (1) building a reference pipeline model with a certified function pipelining transformation, which faithfully captures bubble insertion; and (2) checking the equivalence between the reference model and synthesized RTL implementation. We demonstrate the scalability of our approach on industry-strength designs synthesized by a commercial tool.

### Towards the Formal Analysis of Microresonators Based Photonic Systems

**Speakers:**
Umair Siddique, and Sofiene Tahar, 1Concordia University, Montreal, Canada, CA; 2Department of Electrical and Computer Engineering, Concordia University, Canada

**Abstract**
Recent developments in the fabrication technology attracted the attention of optical engineers and physicists in the area of VLSI photonics. Due to the physical nature of light-wave systems and their usage in safety critical domains such as human surgeries and high budget space missions, it is indispensable to build high assurance systems. Traditionally, the analysis of such systems has been carried out by paper-and-pencil based proofs and numerical computations. However, these techniques cannot provide perfectly accurate results due to the risk of human error and inherent approximations of numerical algorithms. In order to overcome these limitations, we propose to use higher-order logic theorem proving to improve the analysis in the domain of integrated optics or VLSI photonics. In particular, this paper provides a higher-order logic formalization of optical microresonators which are the most fundamental building blocks of many photonic devices. In order to illustrate the practical utilization of our work, we present the formal analysis of 2-D microresonator lattice optical filters.
TOWARDS VERIFYING DETERMINISM OF SYSTEMC DESIGNS

Speakers:
Hoang M. Le and Rolf Drechsler, University of Bremen, DE

Abstract
Ensuring the correctness of high-level SystemC designs is an important and challenging problem in today's Electronic System Level (ESL) methodology. Prevalently, a design is checked against a functional specification given by e.g. a testcase with reference output or a user-defined property. Another research direction takes the view of a SystemC design as a piece of concurrent software. The design is then checked for common concurrency problems and thus, a functional specification is not required. Along this line, several methods for deadlock detection and race analysis have been developed. In this work, we propose to consider a new concurrency verification problem, namely input-output determinism, for SystemC designs. That means for each possible input, the design must produce the same output under any valid process schedule.

We argue that determinism verification is stronger than both deadlock detection and race analysis. Beside being an attractive correctness criterion itself, proven determinism helps to accelerate both simulative and formal verification. We also present a preliminary study to show the feasibility of determinism verification for SystemC designs.

6.6 Model-Based Design and Hardware/Software Interfaces

Date: Wednesday, March 26, 2014
Time: 11:00 - 12:30
Location / Room: Konferenz 4

Chair:
Wang Wang Yi, Uppsala University, SE, Contact Wang Yi

Co-Chair:
Wolfgang Nebel, OFFIS, DE, Contact Wolfgang Nebel

This sessions covers multiple abstraction in embedded system design. The first paper proposes a scalable approach to refinement checking of component-based systems using contracts and local refinement assertions. The second paper revisits the paradigm of using a set of communicating asynchronous components for implementation of synchronous models. The third paper presents a hardware scheduling support for OpenMP and the fourth paper proposes an object-aware translation layer for flash memories.

11:00 6.6.1 LIBRARY-BASED SCALABLE REFINEMENT CHECKING FOR CONTRACT-BASED DESIGN

Speakers:
Antonio Iannopollo, Pierluigi Nuzzo, Stavros Tripakis and Alberto Sangiovanni-Vincentelli, University of California, Berkeley, US

Abstract
Given a global specification contract and a system described by a composition of contracts, system verification reduces to checking that the composite contract refines the specification contract, i.e. that any implementation of the composite contract implements the specification contract and is able to operate in any environment admitted by it.

Contracts are captured using high-level declarative languages, for example, linear temporal logic (LTL). In this case, refinement checking reduces to an LTL satisfiability checking problem, which can be very expensive to solve for large composite contracts. This paper proposes a scalable refinement checking approach that relies on a library of contracts and local refinement assertions. We propose an algorithm that, given such a library, breaks down the refinement checking problem into multiple successive refinement checks, each of smaller scale. We illustrate the benefits of the approach on an industrial case study of an aircraft electric power system, with up to two orders of magnitude improvement in terms of execution time.

11:30 6.6.2 ISOCRONOUS NETWORKS BY CONSTRUCTION

Speakers:
Yu Bai and Klaus Schneider, University of Kaiserslautern, DE

Abstract
While synchronous system models have many advantages over asynchronous models concerning verification and validation, many implementation platforms do not provide efficient means for synchronization. For this reason, we consider a design flow that starts with a synchronous system model that is then transformed into an asynchronous one for synthesis. In essence, it partitions the synchronous system into a set of asynchronous components that communicate with each other via FIFO buffers. Of course, the synthesized system still has to behave as the original synchronous model, i.e. for each variable exactly the same flow of data values must be observed and only the membership to synchronous reaction steps is no longer explicitly given. In this paper, we prove that this correctness guarantee is given provided that (1) each component knows which of the input values have to be used for the next reaction (endochrony), (2) each component is able to perform the reaction (constructiveness), and (3) components agree on the clocks of their shared variables (isochrony/lock-consistency).

11:45 6.6.3 TIGHTLY-COUPLLED HARDWARE SUPPORT TO DYNAMIC PARALLELISM ACCELERATION IN EMBEDDED SHARED MEMORY CLUSTERS

Speakers:
Paolo Burgio1, Giuseppe Tagliavini1, Francesco Conti1, Andrea Marongiu2 and Luca Benini1
1University of Bologna, Université de Bretagne-Sud, IT; 2University of Bologna, IT; 3Università di Bologna, IT

Abstract
Modern designs for embedded systems are increasingly embracing cluster-based architectures, where small sets of cores communicate through tightly-coupled shared memory banks and high-performance interconnections. At the same time, the complexity of modern applications requires new programming abstractions to exploit dynamic and/or irregular parallelism on such platforms. Supporting dynamic parallelism in systems which i) are resource-constrained and ii) run applications with small units of work calls for a runtime environment which has minimal overhead for the scheduling of parallel tasks. In this work, we study the major sources of overhead in the implementation of OpenMP dynamic kops, sections and tasks, and propose a hardware implementation of a generic Scheduling Engine (HWSE) which fits the semantics of the three constructs. The HWSE is designed as a tightly-coupled block to the PEs within a multi-core cluster, communicating through a shared-memory interface. This allows very fast programming and synchronization with the controlling PEs, fundamental to achieving fast dynamic scheduling, and ultimately to enable fine-grained parallelism. We prove the effectiveness of our solutions with real applications and synthetic benchmarks, using a cycle-accurate virtual platform.
### 6.7 Hardening Approaches at Different Design Levels

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>11:00</td>
<td>6.6.4</td>
<td><strong>P-OFTL: AN OBJECT-BASED SEMANTIC-AWARE PARALLEL FLASH TRANSLATION LAYER</strong></td>
<td>Wei Wang, Youyou Lu and Jiewu Shu, Tsinghua University, CN</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
<td>With increased density and decreased price, flash memory has been widely used in storage systems for its low latency and low power features. However, traditional storage systems are designed and excessively optimized for magnetic disks, and the potential of flash memory is not harnessed into full play in the form of Solid State Drives (SSDs). In this paper, we propose a P-OFTL, an object-based semantic-aware parallel flash translation layer (FTL). P-OFTL realizes the mapping table in the FTL and directly manages the flash memory in file objects, which enables optimization of data layout in the flash using object semantics. While the removing of the mapping table improves system performance, a challenge remains to exploit the internal parallelism when maintaining the continuity of logical addresses in each object, which is essential for efficient garbage collection. To address this challenge, P-OFTL statically remaps the addresses by shifting the bits in the addresses, which spreads writes to different internal parallel units without another mapping table. Also, P-OFTL employs a semantic-aware data grouping algorithm to group data pages by trading off the hot-cold clustering for the continuity of logical addresses, so as to reduce the page movement in garbage collection. Experiments show that P-OFTL improves system performance by 4.0% - 10.3% and reduces garbage collection overhead by 15.1% - 32.5% in semantic-aware data grouping compared to those in semantic-unaware data grouping algorithms.</td>
</tr>
<tr>
<td>12:30</td>
<td>6.7.2</td>
<td><strong>USING GUIDED LOCAL SEARCH FOR ADAPTIVE RESOURCE RESERVATION IN LARGE-SCALE EMBEDDED SYSTEMS</strong></td>
<td>Timon ter Braak, University of Twente, NL</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
<td>To maintain a predictable execution environment, an embedded system must ensure that applications are, in advance, provided with sufficient resources to process tasks, exchange information and to control peripherals. The problem of assigning tasks to processing elements with limited resources, and routing communication channels through a capacitiated interconnect is combined into an integer linear programming formulation. We describe a guided local search algorithm to solve this problem at runtime. This algorithm allows for a hybrid strategy where configurations computed at design-time may be used as references to lower the computational overhead at runtime. Computational experiments on a dataset with 100 tasks and 20 processing elements show the effectiveness of this algorithm compared to state-of-the-art solvers CPLEX and Gurobi. The guided local search algorithm finds an initial solution within 100 milliseconds, is competitive for small platforms, scales better with the size of the platform, and has lower memory usage (2-18%).</td>
</tr>
<tr>
<td>12:30</td>
<td>6.7.3</td>
<td><strong>ACCELERATING GRAPH COMPUTATION WITH RACETRACK MEMORY AND POINTER-ASSISTED GRAPH REPRESENTATION</strong></td>
<td>Eunhyek Park1, Helen Li, Sungjoo Yoo1 and Sunggu Lee1</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
<td>The poor performance of NAND Flash memory, such as long access latency and large granularity access, is the major bottleneck of graph processing. This paper proposes an intelligent storage for graph processing which is based on fast and low cost racetrack memory and a pointer-assisted graph representation. Our experiments show that the proposed intelligent storage based on racetrack memory reduces total processing time of three representative graph computations by 40.2% - 86.9% compared to the graph processing, GraphChi, which exploits sequential accesses based on normal NAND Flash memory-based SSD. Faster execution also reduces energy consumption by 39.6% - 90.0%. The in-storage processing capability gives additional 10.5% - 16.4% performance improvements and 12.0% - 14.4% reduction of energy consumption.</td>
</tr>
<tr>
<td>13:00</td>
<td>6.7.4</td>
<td><strong>P-OFTL: AN OBJECT-BASED SEMANTIC-AWARE PARALLEL FLASH TRANSLATION LAYER</strong></td>
<td>Wei Wang, Youyou Lu and Jiewu Shu, Tsinghua University, CN</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
<td>With increased density and decreased price, flash memory has been widely used in storage systems for its low latency and low power features. However, traditional storage systems are designed and excessively optimized for magnetic disks, and the potential of flash memory is not harnessed into full play in the form of Solid State Drives (SSDs). In this paper, we propose a P-OFTL, an object-based semantic-aware parallel flash translation layer (FTL). P-OFTL realizes the mapping table in the FTL and directly manages the flash memory in file objects, which enables optimization of data layout in the flash using object semantics. While the removing of the mapping table improves system performance, a challenge remains to exploit the internal parallelism when maintaining the continuity of logical addresses in each object, which is essential for efficient garbage collection. To address this challenge, P-OFTL statically remaps the addresses by shifting the bits in the addresses, which spreads writes to different internal parallel units without another mapping table. Also, P-OFTL employs a semantic-aware data grouping algorithm to group data pages by trading off the hot-cold clustering for the continuity of logical addresses, so as to reduce the page movement in garbage collection. Experiments show that P-OFTL improves system performance by 4.0% - 10.3% and reduces garbage collection overhead by 15.1% - 32.5% in semantic-aware data grouping compared to those in semantic-unaware data grouping algorithms.</td>
</tr>
<tr>
<td>13:00</td>
<td>6.7.5</td>
<td><strong>USING GUIDED LOCAL SEARCH FOR ADAPTIVE RESOURCE RESERVATION IN LARGE-SCALE EMBEDDED SYSTEMS</strong></td>
<td>Timon ter Braak, University of Twente, NL</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
<td>To maintain a predictable execution environment, an embedded system must ensure that applications are, in advance, provided with sufficient resources to process tasks, exchange information and to control peripherals. The problem of assigning tasks to processing elements with limited resources, and routing communication channels through a capacitiated interconnect is combined into an integer linear programming formulation. We describe a guided local search algorithm to solve this problem at runtime. This algorithm allows for a hybrid strategy where configurations computed at design-time may be used as references to lower the computational overhead at runtime. Computational experiments on a dataset with 100 tasks and 20 processing elements show the effectiveness of this algorithm compared to state-of-the-art solvers CPLEX and Gurobi. The guided local search algorithm finds an initial solution within 100 milliseconds, is competitive for small platforms, scales better with the size of the platform, and has lower memory usage (2-18%).</td>
</tr>
<tr>
<td>13:00</td>
<td>6.7.6</td>
<td><strong>ACCELERATING GRAPH COMPUTATION WITH RACETRACK MEMORY AND POINTER-ASSISTED GRAPH REPRESENTATION</strong></td>
<td>Eunhyek Park1, Helen Li, Sungjoo Yoo1 and Sunggu Lee1</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
<td>The poor performance of NAND Flash memory, such as long access latency and large granularity access, is the major bottleneck of graph processing. This paper proposes an intelligent storage for graph processing which is based on fast and low cost racetrack memory and a pointer-assisted graph representation. Our experiments show that the proposed intelligent storage based on racetrack memory reduces total processing time of three representative graph computations by 40.2% - 86.9% compared to the graph processing, GraphChi, which exploits sequential accesses based on normal NAND Flash memory-based SSD. Faster execution also reduces energy consumption by 39.6% - 90.0%. The in-storage processing capability gives additional 10.5% - 16.4% performance improvements and 12.0% - 14.4% reduction of energy consumption.</td>
</tr>
</tbody>
</table>

6.7.1 NOSTRADAMUS: LOW-COST HARDWARE-ONLY ERROR DETECTION FOR PROCESSOR CORES

**Abstract**

We propose a new, low-cost, hardware-only scheme to detect errors in superscalar, out-of-order processor cores. For each instruction decoded, Nostredamus compares what the instruction is expected to do against what the instruction actually does. We implement Nostredamus in RTL on top of a baseline superscalar, out-of-order core, and we experimentally evaluate its ability to detect injected errors. We also evaluate Nostredamus’s area and power overheads.

6.7.2 WORD-LINE POWER SUPPLY SELECTOR FOR STABILITY IMPROVEMENT OF EMBEDDED SRAMs IN HIGH RELIABILITY APPLICATIONS

**Abstract**

Embedded SRAM yield dominates the overall ASIC yield, therefore the methodologies centered on improving SRAM cell stability will be introduced in the design as a mandatory. Word-line voltage modulation has shown that it is possible to improve cell stability during access operations. The high variability of physical and performance parameters introduce the need to adopt adaptable solutions to adequately improve SRAM cell stability. In this work, we present a word-line voltage selector circuit designed to modulate power-supply word-line voltage at each individual embedded SRAM block. The final area overhead is minimal and several strategies can be implemented with the embedded SRAM allowing adjust word-line voltage value during the life of ASIC, taking into account different operation, aging and degradations effects.

---

**Contact Lorena Anghel**

**Contact Cecilia Metra**
### 6.8 First Time Right in Analog Design Enabling New Business Cases

**Date:** Wednesday, March 26, 2014  
**Time:** 11:00 - 12:30  
**Location / Room:** Exhibition Theatre

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>11:00</td>
<td>6.8.1</td>
<td>ACHIEVING FIRST-TIME-RIGHT SILICON IN ANALOG DESIGNS - A FOUNDRY PERSPECTIVE</td>
<td>Jörg Doblaski, X-FAB, DE</td>
</tr>
</tbody>
</table>

**Abstract**

Today’s demanding analog- and mixed-signal applications often do not allow for a "second shot". Due to both schedule- and budget requirements, costly and time-consuming re-spins of all components need to be avoided to be successful. "First-Time-Right" is the goal for these designs. The presentation will outline the challenges involved in achieving first-time-right analog designs. It will highlight what impact the choice of process architecture makes, and will discuss the pros and cons of different process architectures, such as BCD and SOI. Fabless companies rely on their foundry to provide not only the right processes, but also excellent modeling of process and devices, and high-quality, feature-rich process design kits. The influence of the design kit quality will be discussed in a second part of the presentation, as well as choice of the right EDA tools and design flows. Finally, it will be discussed how the relationship between foundry, fabless company and EDA provider needs to be developed in order to better support First-Time-Right in analog- and mixed-signal designs.
UB06 Session 6

Date: Wednesday, March 26, 2014
Time: 12:00 - 14:00
Location / Room: University Booth, Booth 3, Exhibition Area

<table>
<thead>
<tr>
<th>Time</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>11:30</td>
<td><strong>EXPLORING THE DESIGN-SPACE WITH &quot;FAATS&quot; TO ACHIEVE FIRST-TIME-RIGHT SILICON IN ANALOG DESIGNS</strong></td>
<td>Markus Meisner, University of Frankfurt, DE</td>
</tr>
<tr>
<td></td>
<td><strong>Abstract</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>The demand for lower supply voltages, faster processing speeds, smaller technology nodes, the accompanied higher variation impact under constantly reduced product cycles, significantly increases the necessity for automation during the design of analog modules. This presentation demonstrates the recent progress on the “Fully Automated Analog Topology Synthesis Framework” (FAATS) by introducing its unique approach to elevate automated analog circuit design to the next step. How valuable an extensive design-space exploration can support “First-Time-Right” requirements is presented on different (design) case studies and an exclusive peek into an ongoing ASIC development strongly driven by FAATS.</td>
<td></td>
</tr>
<tr>
<td>12:30</td>
<td>End of session</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Lunch Break</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Sandwich lunch</strong></td>
<td></td>
</tr>
</tbody>
</table>

**UB06.01 SOC VERIFICATION: AUTOMATED FUNCTIONAL VERIFICATION OF SYSTEMS-ON-CHIP**

Authors: Zseneik Priekyl, Marcela Simkova and Karel Masarik, Faculty of Information Technology, Brno University of Technology, CZ

**Abstract**
An increase of the complexity of systems-on-chip (SoC) induces an increase of the complexity of their verification as well. The reason is that we must verify not only the functions of separate logic blocks, but we need to check their interconnections, timing and functional collaboration as well. Therefore, there is still a great demand for verification tools, which are time-effective, fast and as automated as possible. Exactly these issues we target in our solution. You are welcome to see the live demonstration at our booth!

More information ...

**UB06.02 HIPACC: AUTOMATIC GPU CODE GENERATION FOR ANDROID**

Authors: Oliver Reichel1, Richard Menbarth1, Frank Hanning1 and Jürgen Teich 1
1University of Erlangen-Nuremberg, DE; 2Saarland University, DE

**Abstract**
We present the Heterogeneous Image Processing Acceleration (HIPAcc) framework. It allows programmers to develop image preprocessing applications while providing high productivity, flexibility, and portability as well as competitive performance. The same algorithm description serves as basis for targeting different GPU accelerators and low-level languages. Hereby, imaging algorithms can be expressed in a compact and productive way by using a domain-specific language (DSL) that is embedded into C++ code. Using the HIPAcc source-to-source compiler, DSL code is compiled to CUDA, OpenCL, C/C++, or even Renderscript code, which targets heterogeneous architectures on recent MPSoCs running Android. Programming those MPSoCs can be challenging, in particular when targeting different architectures (CPU/GPU/DP). HIPAcc lifts this burden from programmers by automatically applying source code transformations based on domain knowledge and a built-in architecture model. This demonstration shows the seamless integration of HIPAcc into the Android Developer Tools and provides a live comparison of generated code to functional identical handwritten naive implementations of image filters on recent MPSoCs running Android.

More information ...

**UB06.03 CUCUMBER-VERILOG: BEHAVIOR DRIVEN DEVELOPMENT FOR CIRCUIT DESIGN AND VERIFICATION**

Authors: Melanie Diepenbeek, Mathias Soeken, Ulrich Kühne and Rolf Drechsler, University of Bremen, DE

**Abstract**
When designing hardware one usually applies a top-down approach in which starting from a natural language specification a design is implemented and afterwards tested and verified for correctness. In contrast, software development is pushed towards agile techniques such as Test Driven Development (TDD), where tests play a central role in driving the implementation. Behavior Driven Development (BDD) extends TDD by using natural language style scenarios to describe the tests. Essentially, in both techniques testing and implementation is interleaved: first, test cases are written, and secondly, the implementation is extended to satisfy them. Since nowadays 70% of the effort to design hardware systems is spent on verification, test and verification should receive more attention and be applied as soon as possible. We present a BDD tool tailored for the Verilog hardware description language which enables a new design flow for hardware design, test, and verification. BDD acceptance tests are readily given by means of the natural language specification. Assigning test code to their sentences yields a testbench which serves as a starting point for the implementation. In the same time, the natural language scenarios form a test documentation that is easily accessible also to non-experts. Furthermore, our tool allows for the generalization of test cases to properties suitable for formal verification. As properties are typically more difficult to formalize than test cases, our approach facilitates the access to formal verification. In our demonstration, we will show how to implement hardware designs using our BDD tool and how properties are generalized from test cases which can then be verified by a model checker automatically.

More information ...

**UB06.04 COMPILER FOR MAPPING STREAM PROCESSING APPLICATIONS INTO REAL-TIME HETEROGENEOUS MULTIPROCESSOR SYSTEMS**

Authors: Stefan Geuns, Berend Dekens, Philip Wilmanns, Joost Haussmans, Guus Kuijer and Marco Bakxkoij, University of Twente, NL

**Abstract**
Heterogeneous multiprocessors system are employed for power-efficiency reasons in wearable software defined radios. These systems are hardware cost-effective and deliver a superior performance compared to their homogeneous counterparts. However these systems are notoriously hard to program without tool support, which makes it desirable that programming is simplified with the help of an optimizing multiprocessor compiler for stream processing applications. This demonstration shows our multiprocessor compiler for mapping real-time stream processing applications onto our real-time heterogeneous multi-core system. The applications are described as sequential programs and are compiled into parallel task graphs. Buffer capacities are computed using dataflow analysis techniques given the real-time constraints of the application. Our multi-core system contains 16 MicroBlaze processor cores as well as two hardware accelerators and is prototyped on a Xilinx Virtex-6 FPGA. A connection-less communication ring is used for inter-processor communication. Our system is equipped with an analog RF front-end, which enables us to demonstrate PAL-video reception and decoding.

More information ...
Abstract
An embedded 24-channel acoustic processing system consisting of an FPGA based front-end and a multi-core microcontroller subsystem is presented here. It is specifically
designed for a smart building solution estimating the occupancy level of rooms and areas solely based on acoustic features and source localization. The overall goal is to use this
occupancy estimate to lower the energy consumption of large buildings. An overview of the hardware and software concept as well as a brief description of the acoustic occupancy
level estimation is given. The APU was developed as part of the EU FP7 project: Sounds for Energy Control of Buildings (S4ECoB).

More information ...

SCOPE: TIME-DECOUPLED PARALLEL SYSTEMIC SIMULATION

Authors:
Jan Weinstock, Christoph Schumacher1, Rainer Leupers1, Gerd Ascheid1 and Laura Tosoratto1

1RWTH Aachen University, DE; 1Institute for Communication Technologies and Embedded Systems, RWTH Aachen University, DE; 1Istituto Nazionale di Fisica Nucleare - Sezione di
Roma, IT

Abstract
With increasing system size and complexity, designers of embedded systems face the challenge of efficiently simulating these systems in order to enable target specific software
development and design space exploration as early as possible. Today's multicore workstations offer enormous computational power, but traditional simulation engines like the OSCI SystemC kernel only operate on a single thread, thereby leaving a lot of computational potential unused. Most modern embedded system designs include multiple processors. This work proposes SCope, a SystemC kernel that aims at exploiting the inherent parallelism of such systems by simulating the processors on different threads. A lookahead mechanism is employed to reduce the required synchronization between the simulation threads, thereby further increasing simulation speed. The virtual prototype of the European FP7 project EURETILE system simulator is used as demonstrator for the proposed work, showing a speedup of 4x on a four core host system compared to sequential simulation. The demo will visualize the simulation of the EURETILE system using an OpenGL based graphical user interface. The simulator will be presented as a regular sequential version based on OSCI SystemC, and as a parallel version based on the new SCope parallel SystemC kernel.

More information ...

COMPSOC: VIRTUAL EXECUTION PLATFORMS FOR MIXED TIME-CRITICALITY APPLICATIONS

Authors:
Kees Goossens, TU Eindhoven, NL

Abstract
System-on-Chip (SOC) design gets increasingly complex, as a growing number of applications are inte- grated in such systems. These applications have mixed time-criticality, i.e.,
some have firm, some soft, and others non-real-time requirements. Executing such a mix of applications on a SOC poses several challenges. First, to reduce cost, platform
resources, e.g., processors, interconnect, memories, are shared between applications. However, sharing causes interference between applications, making their behaviors inter-
dependent. This results in two problems for SOC design and verification: 1) accurate system-level simulation and several approaches to formal verification are infeasible, because of the
explosion in the number of possible combinations of applications, inputs, and resource states and 2) verification becomes a circular process that must be repeated if an
application is added, removed, or modified, making integration and verification dominant parts of SOC development, in terms of time and money. The CompSOC platform addresses these problems by executing each application on an independent virtual execution platform (VEP). The VEPs are composable, i.e., cannot affect each other's behaviors. In the temporal domain an applications actual execution never varies by even a single clock cycle. Similarly, the energy and power behaviors of applications are also composable. As a result, applications can be designed, developed, verified, and executed in isolation. The VEPs are also predictable, meaning that all interference is bounded. This makes them virtualized also in terms of performance bounds, which enables firm real-time applications to be verified using formal performance analysis frameworks. The CompSOC platform uses the CoMK microkernel to implement virtual processors on each processor time through temporal partitioning. Each application can use its own operating system (e.g. Compose,
µCOS-III) and model of computation (e.g. CSDF, KPN, TT) in its VEP, to suit its level of time critically. As more applications are integrated on a single SOC, the need arises for more
dynamic behaviour. The system should be able to start, modify and stop applications at run time without affecting running appli-cations. For this purpose the CompSOC platform has
been extended with a predictable and composable resource management framework. It manages application bundles that contain 1) an application in the form of executables (ELFs
on multiple processors), and also 2) the specifications of the (one or more) particular VEPs that the application executes in, consisting of virtual processors, NOC connections,
virtualised mem-ories, etc. At run time, the resource management framework can dynamically load and start application bundles by creating a VEP and then loading, booting, and
executing an application within it. VEPs can also be modified, stopped, and deleted at run time. Our University Booth will present virtual execution platform and application bundle
concepts using an interactive demonstrator. It will show that the CompSOC has been extended with dynamic functionality, without sacrificing its key strengths: composability and
predictability. We will demonstrate this through the use of the resource management framework and application bundles, showing that we can create, modify and delete virtual
execution platforms running a mixed time-critically application dynamically at run-time.

More information ...

TTOOL/DIPLODOCUSDF: A UML ENVIRONMENT FOR HARDWARE/SOFTWARE CO-DESIGN OF DATA-DOMINATED SYSTEMS-ON-CHIP

Authors:
Andrea Emici, Ludovic Aprville and Renaud Picalet, Telecom ParisTech, FR

Abstract
The development of new Systems on Chip commonly relies on previous products for whom, due to factors such as system complexities, time and cost constraints, little design space
exploration can be performed. Hardware and software are typically composed as if they were separate components, whereas their interactions yield more than the sum of the two
parts. In the scope of the demonstration, we present our enhanced version of TTool/DiplodocusDF, a UML model-driven engineering tool and methodology for the design of
heterogeneous data processing systems. Our contributions enrich the modeling and design space exploration capabilities of TTool/DiplodocusDF to target complex transfer schemes
and control information exchange at different abstraction levels. Our ameliorated methodology is applied to two signal processing applications, showing the analysis of novel
interactions between typically conflicting aspects such as computations vs communications and dataflows vs controlflows.

More information ...

PIGGY’S WEAVER: A DEMONSTRATION FOR FOCUSING ON SEPARATION OF DEBUGGING CONCERNS BASED ON DYNAMIC PROGRAM Rewriting TOOL: PIGGY’S WEAVER

Authors:
Ikuha Tanigawa1, Nobuhiro Ogura2, Midori Sugaya3 and Harumi Watanabe1
1Tokai University, JP; 2Tokyo City University, JP; 3Shibaura Institute of Technology, JP

Abstract
Dynamic program rewriting is needed to continuous work and reduces costs of maintenance. We propose a dynamic rewriting tool “Piggy’s Weaver” for C# program. The tool
attaches and detaches pieces of code to program at any points on each concern. Especially these attachments are focused on debugging concern. In the demonstration, we will
apply the tool to a cloud and embedded system “PiggyNet” which is a cooperating charity pot with SNSs and was awarded 2nd prize on D2C2012 by Microsoft Japan.

More information ...
UNISON: ASSEMBLY CODE GENERATION USING CONSTRAINT PROGRAMMING

Authors:
Roberto Castañeda Lozano, Gabriel Hjort Blindell, Mats Carlsson, Christian Schulte
1Swedish Institute of Computer Science, SE; 2KTH Royal Institute of Technology, SE

Abstract
We demonstrate Unison - a simple, flexible and potentially optimal code generator that solves interdependent code generation tasks together using constraint programming as a modern combinatorial optimization method. We show how Unison takes into account the task interdependencies and their combinatorial nature to improve the speed of the code generated by LLVM (a state-of-the-art compiler) for Hexagon (a digital signal processor ubiquitous in modern mobile platforms).

More information ...

14:00 End of session
16:00 Coffee Break in Exhibition Area
On Tuesday-Thursday the coffee and lunch breaks will be located in the Exhibition Area (Terrace Level).

7.0 Special Day Keynote
Date: Wednesday, March 26, 2014
Time: 13:30 - 14:00
Location / Room: Saal 1

The automotive industry is in a radical change process driven by technology. On the one hand side the proliferation of communication technologies into the car leads to internet connected vehicles. The vehicle will become an integral part of the internet - opening new processing paradigms for the car itself. On the other hand the vehicle itself significantly expands its sensor and processing capabilities by the use of radar, video, ultrasound sensors and usage of state of the art CPU and GPU processor architectures. In our talk we will address both developments and outline foreseen future applications as future driving assistant and infotainment systems as well as highly automated driving. We will discuss major requirements for the future electrical architectures and implications for future automotive chips.

Time | Label | Presentation Title
---|---|---
13:30 | 7.0.1 | SPECIAL DAY KEYNOTE: THE CONNECTED CAR AND ITS IMPLICATION TO THE AUTOMOTIVE CHIP ROADMAP
Speaker:
Dr.-Ing. Michael Bolle, Robert Bosch Gmbh, DE

14:00 | End of session |
16:00 | Coffee Break in Exhibition Area |
On Tuesday-Thursday the coffee and lunch breaks will be located in the Exhibition Area (Terrace Level).

UB07 Session 7
Date: Wednesday, March 26, 2014
Time: 14:00 - 16:00
Location / Room: University Booth, Booth 3, Exhibition Area

Label | Presentation Title
---|---
UB07.01 | VIDEO-BASED ABSOLUTE NAVIGATION APPROACH: A NOVEL APPROACH FOR VIDEO-BASED ABSOLUTE NAVIGATION IN SPACE EXPLORATION MISSIONS
Authors:
Pascal Trotta, Tadewos Getahun Tedewos, Paolo Prinetto, Daniele Rolfo and Pascal Trotta, Politecnico di Torino, IT

Abstract
Nowadays, space agencies have increased their research efforts in order to enhance the success rate of space exploration missions. Future space missions will increasingly adopt Video Based Navigation (VBN) systems to assist the entry, descent and landing (EDL) phase of space modules. This poster will show a preliminary work on a novel approach for Video-based Absolute Navigation (VBAN). Moreover, the poster depicts how a VBAN processing chain can exploit FPGA devices to achieve high throughput. Several visual results will be shown to highlight the peculiarity of the proposed approach.

More information ...

UB07.02 | AIDA: ANALOG IC DESIGN AUTOMATION
Authors:
Nuno Horta, Nuno Lourenço, Ricardo Martins, Ricardo Povoaz, António Canelas and Pedro Ventura
1Instituto de Telecomunicacoes, PT; 2Instituto de Telecomunicacoes / Instituto Superior Técnico, PT

Abstract
This demonstration presents AIDA, an analog integrated circuit (IC) design automation environment. AIDA includes two main modules, namely, AIDA-C and AIDA-L. AIDA-C is a circuit-level synthesis tool which uses state-of-the-art multi-objective multi-constrained optimization kernels, based on evolutionary computation techniques, where the robustness of the solutions is attained by considering a layout-aware approach and, also, extreme process variations by means of PVT corner analysis. The circuit's performance is measured using Spectre®, ELDO® or HSPICE® electrical simulators as evaluation engines. AIDA-L considers the device sizes and the best floorplan, obtained with AIDA-C, and generates the complete layout by placing and routing the devices, while fulfilling the technology design rules by using built-in design-rule check (DRC) and layout-versus-schematic (LVS) procedures. In order to demonstrate AIDA design environment several analog circuit structures, e.g., OTAs, LNAs, LC oscillators, etc., will be synthesized in a 130nm CMOS technology. AIDA-C is demonstrated for circuit-level sizing and optimization by generating a family of Pareto Optimal solutions based on user performance and functional specifications. AIDA-L is demonstrated by generating the layout of a user selected solution from AIDA-C, taking into account electrical currents information to mitigate electromigration and IR-drop effects, and also wiring symmetry for multiport multi-terminal signal nets of analog ICs.

More information ...
BICONDITIONAL BINARY DECISION DIAGRAM MANIPULATION PACKAGE
Authors: Luca Amaru, Alexios Balatsoukas-Stimming, Pierre-Emmanuel Gaillardon, Andreas Burg and Giovanni De Micheli
\textsuperscript{1}\textsuperscript{EPFL, CH; \textsuperscript{2}\textsuperscript{EPFL-TCL, CH; \textsuperscript{3}\textsuperscript{EPFL-LSI, CH

Abstract
In this software demonstration, we present a logic manipulation package based on Biconditional Binary Decision Diagrams (BBDDs). BBDDs are a novel class of canonical binary decision diagrams where the branching condition, and its associated logic expansion, is biconditional on two variables. We show how Verilog files from real life designs can be rapidly read and processed by the BBDD manipulation package, for verification, testing or synthesis purposes. In particular, we demonstrate the benefit deriving from BBDD re-writing of arithmetic circuits in the synthesis of a product code iterative decoder.

GEMINI: A NEW SYNTHESIS AND OPTIMIZATION TOOL FOR GRAPHENE-BASED DIGITAL DEVICES
Authors: Valerio Tenace, Andrea Calmiera, Massimo Poncino and Enrico Macì, Politecnico di Torino, IT

Abstract
Gemini is a synthesis and optimization software for graphene-based digital devices. Given a combinational circuit description through its bookman representation, Gemini produces a SPICE netlist mapped with graphene PN-Junction gates. The software is composed of a parser library to handle input circuit descriptions, a characterization library of graphene gates used in the synthesis process, a Biconditional Binary Decision Diagram library used to manipulate logic networks in Pass-XNOR logic in order to better exploit the intrinsic characteristics of the adopted graphene gates, and a number of optimization algorithms designed to produce better results in terms of area and thus power consumption. As a stand-alone software or as a library easy to integrate into state-of-the-art tools, Gemini represents a first step of an enabling technology for future synthesis and optimization processes for graphene-based devices.

TOMAHAWK2: PERFORMANCE IMPACT OF INSTRUCTION SET ARCHITECTURE EXTENSIONS FOR DYNAMIC TASK SCHEDULING UNITS
Author: Oliver Arnold, Technische Universität Dresden, DE

Abstract
In this demo a heterogeneous MPSoC is controlled by a dynamic task scheduling unit called CoreManager. The instruction set architecture of this unit has been extended to improve performance for dynamic data dependency checking, task scheduling, processing element allocation and data transfer management. The MPSoC as well as the NoC are integrated in a cycle-accurate virtual system prototype. The performance impact of the CoreManager is analyzed on component as well as on system level.

TTOOL/DIPLODOCUSDF: A UML ENVIRONMENT FOR HARDWARE/SOFTWARE CO-DESIGN OF DATA-DOMINATED SYSTEMS-ON-CHIP
Author: Fredrik Jonsson, Royal Institute of Technology, SE

Abstract
Performance of printed devices depends on the geometry, but is also affected by processing steps of other components integrated onto the same substrate. Since different designs use different technologies, process, models and design rules must be dynamically determined. In this work we propose and demonstrate an experimental design flow to allow efficient design of hybrid and printed electronic circuits.

GEMINI: A NEW SYNTHESIS AND OPTIMIZATION TOOL FOR GRAPHENE-BASED DIGITAL DEVICES
Authors: Valerio Tenace, Andrea Calmiera, Massimo Poncino and Enrico Macì, Politecnico di Torino, IT

Abstract
Gemini is a synthesis and optimization software for graphene-based digital devices. Given a combinational circuit description through its bookman representation, Gemini produces a SPICE netlist mapped with graphene PN-Junction gates. The software is composed of a parser library to handle input circuit descriptions, a characterization library of graphene gates used in the synthesis process, a Biconditional Binary Decision Diagram library used to manipulate logic networks in Pass-XNOR logic in order to better exploit the intrinsic characteristics of the adopted graphene gates, and a number of optimization algorithms designed to produce better results in terms of area and thus power consumption. As a stand-alone software or as a library easy to integrate into state-of-the-art tools, Gemini represents a first step of an enabling technology for future synthesis and optimization processes for graphene-based devices.

TOMAHAWK2: PERFORMANCE IMPACT OF INSTRUCTION SET ARCHITECTURE EXTENSIONS FOR DYNAMIC TASK SCHEDULING UNITS
Author: Oliver Arnold, Technische Universität Dresden, DE

Abstract
In this demo a heterogeneous MPSoC is controlled by a dynamic task scheduling unit called CoreManager. The instruction set architecture of this unit has been extended to improve performance for dynamic data dependency checking, task scheduling, processing element allocation and data transfer management. The MPSoC as well as the NoC are integrated in a cycle-accurate virtual system prototype. The performance impact of the CoreManager is analyzed on component as well as on system level.

UVM-SYSTEMC-AMS: UVM STANDARD-COMPLIANT SYSTEMC (AMS)-BASED VERIFICATION FRAMEWORK FOR HETEROGENEOUS SYSTEMS
Authors: Zhi Wang\textsuperscript{1}, Yao Li\textsuperscript{2}, Marie-Mnerv Louerat\textsuperscript{2}, Francois Pechex\textsuperscript{2}, Martin Banascon\textsuperscript{2}, Thilo Vörlte\textsuperscript{2} and Karsten Einwich\textsuperscript{2}
\textsuperscript{1}\textsuperscript{Laboratoire d’informatique de Paris 6, FR; \textsuperscript{2}\textsuperscript{RUPMC-LIP6, FR; \textsuperscript{3}\textsuperscript{INKP, NL; \textsuperscript{4}\textsuperscript{Fraunhofer IIS, DE

Abstract
Today’s societal needs for innovative products in terms of communication, mobility, health, entertainment, and safety directly impact microelectronics design methodologies. The embedded systems are simultaneously software-driven, digitally assisted, complex and heterogeneous, but existing verification methodologies are mostly focused on pure digital devices and are completely decoupled from analog verification. This presentation shows how the principles of the new UVM methodology can be soundly enhanced to offer to the test designer a flexible framework for the virtual prototyping of multi-discipline testbenches that supports both digital and Analog Mixed-Signal (AMS) at the architectural level.

BICONDITIONAL BINARY DECISION DIAGRAM MANIPULATION PACKAGE
Authors: Luca Amaru, Alexios Balatsoukas-Stimming, Pierre-Emmanuel Gaillardon, Andreas Burg and Giovanni De Micheli
\textsuperscript{1}\textsuperscript{EPFL, CH; \textsuperscript{2}\textsuperscript{EPFL-TCL, CH; \textsuperscript{3}\textsuperscript{EPFL-LSI, CH

Abstract
In this software demonstration, we present a logic manipulation package based on Biconditional Binary Decision Diagrams (BBDDs). BBDDs are a novel class of canonical binary decision diagrams where the branching condition, and its associated logic expansion, is biconditional on two variables. We show how Verilog files from real life designs can be rapidly read and processed by the BBDD manipulation package, for verification, testing or synthesis purposes. In particular, we demonstrate the benefit deriving from BBDD re-writing of arithmetic circuits in the synthesis of a product code iterative decoder.

TTOOL/DIPLODOCUSDF: A UML ENVIRONMENT FOR HARDWARE/SOFTWARE CO-DESIGN OF DATA-DOMINATED SYSTEMS-ON-CHIP
Authors: Andrea Enrici, Ludovic Aprille and Renaud Pecalet, Telecom ParisTech, FR

Abstract
The development of new Systems on Chip commonly relies on previous products for whom, due to factors such as system complexities, time and cost constraints, little design space exploration can be performed. Hardware and software are typically composed as if they were separate components, whereas their interactions yield more than the sum of the two parts. In this demo, we present our enhanced version of TTool/DiplodocusDF, a UML model-driven engineering tool and methodology for the design of heterogeneous data processing systems. Our contributions enrich the modeling and design space exploration capabilities of TTool/DiplodocusDF to target complex transfer schemes and control information exchange at different abstraction levels. Our ameliorated methodology is applied to two signal processing applications, showing the analysis of novel interactions between typically conflicting aspects such as computations vs communications and dataflows vs controlflows.

A HOLISTIC APPROACH TO POWER MANAGEMENT FOR ENERGY HARVESTING EMBEDDED SYSTEMS
Authors: Kyungsoo Lee, Hideki Takase and Tohru Ishihara, Kyoto University, JP

Abstract
We present a holistic approach to maximizing the energy efficiency of energy harvesting embedded systems which consist of a processor system and an energy harvesting system. A power management program integrated on a real-time OS optimally switches operation mode of the processor and configuration of the energy harvesting system according to the workload of the processor and harvesting situation. The demonstration will show that our prototype system consisting of our processor chip and harvesting system board stably runs using harvested energy only. The processor has multiple cores having a different performance in each to improve the energy efficiency of computation. The energy harvesting board has high transferring efficiency to reduce the power loss. The entire system is controlled efficiently by our power management program implemented on Toppers OS.
Authors: Nobuhiko Ogura¹, Ikuta Tanigawa², Takuya Todoroki¹, Kenji Arai¹ and Harumi Watanabe²
¹Tokyo City University, JP; ²Tokai University, JP

Abstract
We present a state transition model description programming language. It can be translated to pure standard C programs without any OS or handwritten frameworks, hence it is suitable for developing low level driver software and firmware, unlike many other automatic software generation tools from software models that usually focus on higher level models. We show the language and translator to executable software and visual diagram generator, and analysis tools, with embedded software examples.

More information...

16:00 End of session

Coffee Break in Exhibition Area
On Tuesday-Thursday the coffee and lunch breaks will be located in the Exhibition Area (Terrace Level).

7.1 SPECIAL DAY Panel: HW/SW Co-Development - The Industrial Workflow
Date: Wednesday, March 26, 2014
Time: 14:30 - 16:00
Location / Room: Saal 1

Organiser: Johannes Stahl, Synopsys, US, Contact Johannes Stahl

Chair: Iris Stroh, Markt & Technik, DE, Contact Iris Stroh

This panel brings together the entire supply chain for the use of virtual prototyping starting with the end users at an automotive Tier1, a semiconductor supplier, IP providers and the virtual prototyping and software development tool providers. The panelists will discuss what are the benefits and challenges of accelerating software development using virtual prototyping for deployment in industrial projects.

Panelists:
- Andreas Schwein, Siemens, DE
- Martin Vaupel, Bosch, DE
- Albrecht Mayer, Infineon, DE
- Nick Gatherer, ARM, GB
- Frank Schimmel, Cadence, US
- Stephan Lauterbach, Lauterbach, US
- Colin Walls, Mentor Graphics, US
- Andreas Hoffmann, Synopsys, US
- Andreas Hoffmann, Siemens, DE
- Martin Vaupel, Bosch, DE
- Frank Schimmel, Cadence, US
- Yanjng Li, Intel, US, Contact Yanjing Li
- Ulf Schlichtmann, TUM, DE, Contact Ulf Schlichtmann

Resilience at different design hierarchies will be needed in Complex SoCs to handle failures due to variability, reliability and design errors (logical or electrical). The main reasons for the marginal behavior are sheer design complexity, uncertainties in manufacturing processes, temporal variability and operating conditions. In this session, we will cover the basics of cross layer resiliency and explore the reliability challenges in both embedded processors as well as large scale computing resources.

Time Label Presentation Title Authors
14:30 7.2.1 CROSS-LAYER RESILIENCE EXPLORATION AND OPTIMIZATION
Speaker: Subhasish Mitra, Stanford University, US
Abstract
This talk will discuss systematic methodologies for exploring cross-layer resilience, encompassing error detection, correction and recovery techniques, for complex SoCs. The objective is to address several key questions such as: 1. Given a design, is cross-layer resilience always the best option? 2. What are the right models that link resilience techniques across multiple layers for quick, yet accurate, estimation of coverage and costs? 3. What is the proper framework to explore the large space of existing resilience techniques for error detection, correction, and recovery across various abstraction layers?
signal processing: tuning of sensor usage based on applications and methods to selectively drop computations to save power, without affecting the accuracy.

Architectural methods for bio-signal processing, dealing with synchronisation and innovative memory architecture design. The last two papers focus on low power design of applications for bio-

Achieving low power operation is essential for battery operated mobile health applications. In this session, the papers address this important issue. The first two papers present multicore architectural methods for bio-signal processing, dealing with synchronisation and innovative memory architecture design. The last two papers focus on low power design of applications for bio-

Due to increasing demand for personal devices, high performance computing systems and commercial data centers, microprocessor and main memory designers face numerous challenges in delivering large number of chips at effective cost. While frequency scaling effectively ended, technology scaling continues to provide increasing number of transistors. To effectively utilize these transistors for performance, designers turn to sophisticated and highly integrated chip designs such as multi-core (e.g., Intel i7, IBM POWER7, BlueGene/Q), GPGPU (e.g., NVIDIA Tigra) heterogeneous SoC (e.g., IBM Wirespeed). The increasing demand for chips and transistors presents numerous challenges on reliability, power and manufacturing costs. In large scale HPC systems and data centers, the increasing number of chips also raises per-chip reliability requirement in order to achieve system reliability targets.

Low power methods and multicore architectures for mobile health applications

Date: Wednesday, March 26, 2014
Time: 14:30 - 16:00
Location / Room: Konferenz I
Chair: Giovanni Ansaloni, EPFL, CH, Contact Giovanni Ansaloni
Co-Chair: Andrea Bartolini, University of Bologna, IT, Contact Andrea Bartolini

Achieving low power operation is essential for battery operated mobile health applications. In this session, the papers address this important issue. The first two papers present multicore architectural methods for bio-signal processing, dealing with synchronisation and innovative memory architecture design. The last two papers focus on low power design of applications for bio-

Abstract

HARDWARE/SOFTWARE APPROACH FOR CODE SYNCHRONIZATION IN LOW-POWER MULTI-CORE SENSOR NODES

Speakers:
Rubén Braojos1, Ahmed Dogan2, Ivan Beretta2; Giovanni Ansaloni1 and David Atienza2
1Ecole Polytechnique Fédérale de Lausanne, CH; 2EPFL, CH

Abstract

Latest embedded bio-signal analysis applications, targeting low-power Wireless Body Sensor Nodes (WBSNs), present conflicting requirements. On one hand, bio-signal analysis applications are continuously increasing their demand for high computing capabilities. On the other hand, long-term signal processing in WBSNs must be provided within their highly constrained energy budget. In this context, parallel processing effectively increases the power efficiency of WBSNs, but only if the execution can be properly synchronized among computing elements. To address this challenge, in this work we propose a hardware/software approach to synchronize the execution of bio-

Abstract

HYBRID MEMORY ARCHITECTURE FOR VOLTAGE SCALING IN ULTRA-LOW POWER MULTI-CORE BIOMEDICAL PROCESSORS

Speakers:
Daniele Bortolotti1, Andrea Bartolini2, Christian Weis1, Davide Rossi1 and Luca Benini1
1University of Bologna, IT; 2University of Kaiserslautern, DE

Abstract

Technology scaling enables today the design of sensor-based ultra-low cost chips well suited for emerging applications such as wireless body sensor networks, urban life and environment monitoring. Energy consumption is the key limiting factor of this up-coming revolution and memories are often the energy bottleneck mainly due to leakage power. This paper proposes an ultra-low power multi-core architecture targeting eHealth monitoring systems, where applications involve collection of sequences of slow biomedical signals and highly parallel computations at very low voltage. We propose a hybrid memory architecture that combines 6T-SRAM and 8T-SRAM operating in the same voltage domain and capable of dispatching at high voltage a normal operation and at low voltage a fully reliable small memory partition (8T) while the rest of the memory (6T) is state-retentive. Our architecture offers significant energy savings with a low area overhead in typical eHealth Compressed Sensing-based applications.
The session starts with memory design techniques under PVT variation and ageing for DRAMs and SRAM caches. Afterwards, bus, memory and partitioning techniques for 2D and 3D GPUs and 3D systems are presented.

**Abstract**

Body Area Networks (BANs) are widely used mainly for healthcare and fitness purposes. In both cases, the lifetime of sensor nodes included in the BAN is a key aspect that may affect the functionality of the whole system. Typical approaches to power management are based on a trade-off between the data rate and the monitoring time. Our work introduces a power management layer capable to opportunistically use data sampled by sensors to detect contextual information such as user activity and adapt the node operating point accordingly.

Unfortunately, with technology scaling these design margins have become large and very pessimistic for a majority of the manufactured DRAMs. While run-time variations (temperature, voltage, frequency, etc.) may affect the functionality of the whole system, typical approaches to power management are based on a trade-off between the data rate and the monitoring time. Our work introduces a power management layer capable to opportunistically use data sampled by sensors to detect contextual information such as user activity and adapt the node operating point accordingly.

The use of this layer has been demonstrated on a commercial sensor node, increasing its battery lifetime up to a factor of 5.

**Abstract**

We evaluate this methodology on 48 DDR3 devices (from 12 DIMMs) and verify the derived timings under worst-case operating conditions, showing up to 33.3% and 25.9% reduction in DRAM read and write latencies, respectively. The methodology ascertains the actual impact of process-variations on the particular DRAM device and optimizes its access latencies, thereby improving its overall performance.

**Abstract**

In this paper, we propose a generic post-manufacturing performance characterization methodology for DRAMs that identifies this methodology's applicability and identifies the actual impact of process-variations on the particular DRAM device and optimizes its access latencies, thereby improving its overall performance.

**Abstract**

We evaluate this methodology on 48 DDR3 devices (from 12 DIMMs) and verify the derived timings under worst-case operating conditions, showing up to 33.3% and 25.9% reduction in DRAM read and write latencies, respectively.

**Abstract**

Today there is a growing interest in the integration of health monitoring applications in portable devices necessitating the development of methods that improve the energy efficiency of such systems. In this paper, we present a systematic approach that enables energy-quality trade-offs in spectral analysis systems for bio-signals, which are useful in monitoring various health conditions as those associated with the heart-rate. To enable such trade-offs, the processed signals are expressed initially in a basis in which significant components that carry most of the relevant information can be easily distinguished from the parts that influence the output to a lesser extent. Such a classification allows the pruning of operations associated with the less significant signal components leading to power savings with minor quality loss since only less useful parts are pruned under the given requirements. To exploit the attributes of the modified spectral analysis system, thresholding rules are determined and adopted at design- and run-time, allowing the static or dynamic pruning of less-useful operations based on the accuracy and energy requirements. The proposed algorithm is implemented on a typical sensor node simulator and results show up to 82% energy savings when static pruning is combined with voltage and frequency scaling, compared to the conventional algorithm in which such trade-offs were not available. In addition, experiments with numerous cardiac samples of various patients show that such energy savings come with a 4.9% average accuracy loss, which does not affect the system detection capability of sinus-arrhythmia which was used as a test case.

**Abstract**

Mobile computing has been weaved into everyday lives to a great extent. Their usage is clearly imprinted with user's personal signature. The ability to learn such signature enables immense potential in workplace prediction and resource management. In this work, we investigate the user behavior modeling and apply the model for energy management. Our goal is to maximize the quality of service (QoS) provided by the mobile device (i.e., smartphone), while keep the risk of battery depletion below a given threshold. A Markov Decision Process (MDP) is constructed from history user behavior. The optimal management policy is solved using linear programing. Simulations based on real user traces validate that, compared to existing battery energy management techniques, the stochastic control performs better in boosting the mobile devices' QoS without significantly increasing the chance of battery depletion.

**Abstract**

The work introduces a power management layer capable to opportunistically use data sampled by sensors to detect contextual information such as user activity and adapt the node operating point accordingly. The use of this layer has been demonstrated on a commercial sensor node, increasing its battery lifetime up to a factor of 5.

**Abstract**

In this paper, we propose a generic post-manufacturing performance characterization methodology for DRAMs that identifies this methodology's applicability and identifies the actual impact of process-variations on the particular DRAM device and optimizes its access latencies, thereby improving its overall performance.
Aging of transistors is a limiting factor for long term reliability of devices in sub-100nm technologies. It's a worst-case metric where the lifetime of a device is determined by the earliest failing component. Impact is more serious on memory arrays, where failure of a single SRAM cell would cause the failure of the whole system. Previous works have shown that partitioning based strategies based on power management techniques can effectively control aging effects and can extend lifetime of the cache significantly. However, such a benefit comes as a trade-off with performance which reduces proportionally as the time elapses. To address this problem and provide a single solution to concurrently improve aging, energy and performance of the cache, we propose an architectural solution based on the dynamically re-sizable cache and cache partitioning approaches. By this strategy, cache is dynamically re-sized and reconfigured whenever a cache block becomes unreliable. Coupling such aging mitigation technique along with dynamically re-sizable cache approach provides on average 30% lifetime improvement with less than 0.4x degradation in performance whereas, in previous solutions, performance degradation sometimes goes up to 10x.

The complex buses consume significant power in graphics processing units (GPUs). In this paper, we demonstrate how the power consumption of buses in GPUs can be reduced with 3D IC technologies. Based on layout simulations, we found that partitioning and floorplanning of 3D ICs affect the power benefit amount, as well as the technology setup, target clock frequency, and circuit switching activity. With 3D IC technologies, we achieved the total power reduction of up to 21.5% for our GPU.

High-level programming languages have transformed graphics processing units (GPUs) from domain-restricted devices into powerful compute platforms. Yet many “general-purpose GPU” (GPGPU) applications fail to fully utilize the GPU resources. Executing multiple applications simultaneously on different regions of the GPU (spatial multitasking) thus improves system performance. However, within-die process variations lead to significantly different maximum operating frequencies (Fmax) of the streaming multiprocessors (SMs) within a GPU. As the chip size and number of SMs per chip increase, the frequency variation is also expected to increase, exacerbating the problem. The increased number of SMs also provides a unique opportunity: we can allocate resources to concurrently-executing applications based on how those applications are affected by the different available Fmax values. In this paper, we study the effects of per-SM clocking on spatial multitasking capable GPUs. We demonstrate two factors that affect the performance of simultaneously-running applications: (i) the SM partitioning algorithm that decides how many resources to assign to each application, and (ii) the assignment of SMs to applications based on the operating frequencies of those SMs and the applications characteristics. Our experimental results show that spatial multitasking that partitions SMs based on application characteristics, when combined with per-SM clocking, can greatly improve application performance by up to 46% on average compared to cooperative multitasking with global clocking.
The papers in this session consider ways to improve the energy, performance, and reliability of emerging memory technologies. STT-RAM and PCRAM are addressed.

Brandon Del Bel, Jongyeon Kim, Chris H. Kim and Sachin S. Sapatnekar, University of Minnesota, US

**Abstract**

Data mining, bioinformatics, knowledge discovery, social network analysis, are emerging irregular applications that exploits data structures based on pointers or linked lists, such as graphs, unbalanced trees or unstructured grids. These applications are characterized by unpredictable memory accesses and generally are memory bandwidth bound, but also presents large amounts of inherent dynamic parallelism because they can potentially spawn concurrent activities for each one of the element they are exploring. Hybrid architectures, which integrate general purpose processors with reconfigurable devices, appears promising target platforms for accelerating irregular applications. These systems often connect to distributed and multi-ported memories, potentially enabling parallel memory operations. However, these memory architectures introduce several challenges, such as the necessity to manage concurrency and synchronization to avoid structural conflicts on shared memory locations and to guarantee consistency. In this paper we present an adaptive Memory Interface Controller (MIC) that addresses these issues. The MIC is a general and customizable solution that can target several different memory structures, and is suitable for High Level Synthesis frameworks. It implements a dynamic arbitration scheme, which avoids conflicts on memory resources at runtime, and supports atomic memory operations, commonly exploited for synchronization directives in parallel programming paradigms. The MIC simultaneously maps multiple accesses to different memory ports, allowing fine grained parallelism exploitation and ensuring correctness also in the presence of irregular and statically unpredictable memory access patterns. We evaluated the effectiveness of our approach on a typical irregular kernel, graph Breadth First Search (BFS), exploiting different design alternatives.

Jiayin Li and Kartik Mohanram, University of Pittsburgh, US

**Abstract**

This paper describes a write-once-memory-code phase change memory (WOM-code PCM) architecture for next-generation non-volatile memory applications. Specifically, we address the long latency of the write operation in PCM – attributed to PCM SET – by proposing a novel PCM memory architecture that integrates WOM-codes at the memory organization and memory controller levels. The proposed <2^0><2^1><2^2><2^3> WOM-code PCM architecture is able to reduce memory write (read) latency by 20.1% (10.2%) on average across general-purpose (SPEC CPU2006), embedded (MBench), and high-performance (SPLASH-2) benchmarks. To further improve the write latency of WOM-code PCM, we propose a PCM-refresh approach that uses idle cycles to preemptively set PCM rows to the initial WOM-code state. The results show that WOM-code PCM with PCM-refresh can reduce memory write latency by 54.9% (47.9%) on average across the benchmarks. Finally, to balance write latency improvements against WOM-code PCM-refresh, we propose a PCM-refresh approach that uses idle cycles to preemptively set PCM rows to the initial WOM-code state.

Rajendra Bishnoi, Mojtaba Ebrahimi, Fabian Obonza and Mehdi Tahoori

**Abstract**

Spin Transfer Torque (STT) memory is an emerging and promising non-volatile storage technology. However, the high write current is still a major challenge which leads to a huge power consumption of the memory. Due to an inherent torque asymmetry of the Magnetic Tunnel Junction (MTJ) device employed in STT memories, the switching time between parallel to anti-parallel and anti-parallel to parallel magnetization is significantly different. Hence, the write latencies for writing ‘0’ and ‘1’ are also considerably different. In this paper, we propose a technique called Asynchronous Asymmetrical Write Termination (AAWT) which utilizes this asymmetrical behavior to terminate the write operations asynchronously and as a result significantly reduces the write power consumption. Furthermore, we present two different AAWT implementations to determine the actual write termination times. The first one makes use of a clock signal and the second one employs a self-timing approach based on an internal delay element. As shown by our experimental results, AAWT can reduce the total write energy by 30% in average with a negligible area overhead.

Vito Giovanni Castellani, Antonino Tumeo and Fabrizio Ferrandi

**Abstract**

This paper describes a write-once-memory-code phase change memory (WOM-code PCM) architecture for next-generation non-volatile memory applications. Specifically, we address the long latency of the write operation in PCM – attributed to PCM SET – by proposing a novel PCM memory architecture that integrates WOM-codes at the memory organization and memory controller levels. The proposed <2^0><2^1><2^2><2^3> WOM-code PCM architecture is able to reduce memory write (read) latency by 20.1% (10.2%) on average across general-purpose (SPEC CPU2006), embedded (MBench), and high-performance (SPLASH-2) benchmarks. To further improve the write latency of WOM-code PCM, we propose a PCM-refresh approach that uses idle cycles to preemptively set PCM rows to the initial WOM-code state. Results show that WOM-code PCM with PCM-refresh can reduce memory write (read) latency by 54.9% (47.9%) on average across the benchmarks. Finally, to balance write latency improvements against WOM-code PCM-refresh, we propose a WOM-code cached PCM (WCPM) architecture that uses WOM-code PCM as the cache alongside conventional PCM main memory. For just 4.7% memory overhead, WCPM reduces memory write (read) latency by 47.2% (44.0%) on average across the benchmarks.

Authors
### 7.6 Performance and timing analysis

**Date:** Wednesday, March 26, 2014  
**Time:** 14:30 - 16:00  
**Location / Room:** Konferenz 4

**Chair:**  
Wang Yi, Uppsala University, SE, Contact Wang Yi  

**Co-Chair:**  
Petru Eles, Linköping University, SE, Contact Petru Eles

This session includes three papers. The first uses data mining techniques to detect performance bottlenecks to improve the scalability of multicore platforms for embedded applications. The second proposes to use regular expressions for specifying the patterns of deadline misses and hits to relax schedulability analysis for cyber-physical systems. The third presents an approach to

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
</table>
| 16:00  | IP3-14| ENERGY EFFICIENT IN-MEMORY AES ENCRYPTION BASED ON NONVOLATILE DOMAIN-WALL NANOWIRE| Sofiane Lagraa¹, Pingfan Kong¹, Hao Yu and Dennis Sylvester  
¹Nanyang Technological University, SG; ²University of Michigan, US |

**Abstract**  
The widely applied Advanced Encryption Standard (AES) encryption algorithm is critical in secure big-data storage. Data oriented applications have imposed high throughput and low power; i.e., energy efficiency (J/bit), requirements when applying AES encryption. This paper explores an in-memory AES encryption using the newly introduced domain-wall nanowire. We show that all AES operations can be fully mapped to a logic-in-memory architecture by non-volatile domain-wall nanowire, called DW-AES. The experimental results show that DW-AES can achieve the best energy efficiency of 24.8 µJ/bit, which is 9.3X and 6.5X times better than CMOS ASIC and ReRAM-CMOL implementations, respectively. Under the same area budget, the proposed DW-AES exhibits 6.4X higher throughput and 29% power saving compared to a CMOS ASIC implementation; 1.7X higher throughput and 74% power reduction compared to a ReRAM-CMOL implementation.  

| 16:01  | IP3-15| ICE: INLINE CALIBRATION FOR MEMRISTOR CROSSBAR-BASED COMPUTING ENGINE                | Boxun Li¹, Yu Wang¹, Yiran Chen¹, Helen Li¹ and Huazhong Yang¹  
¹Tsinghua University, CN; ²University of Pittsburgh, US |

**Abstract**  
The emerging neuromorphic computation provides a revolutionary solution to the alternative computing architecture and effectively extends Moore's Law. The discovery of the memristor presents a promising hardware realization of neuromorphic systems with incredible power efficiency, allowing efficiently executing the analog matrix-vector multiplication on the memristor crossbar architecture. However, during computations, the memristor will slowly drift from its initial programmed state, leading to a gradual decline of the computation precision of memristor crossbar-based computing engine (MCE). In this paper, we propose an inline calibration mechanism to guarantee the computation quality of the MCE. The inline calibration mechanism collects the MCE's computation error through interrupt-and-benchmark (I&B) operations and predicts the best calibration time through polynomial fitting of the computation error data. We also develop an adaptive technique to adjust the time interval between two neighbor I&B operations and minimize the negative impact of the I&B operation on system performance. The experiment results demonstrate that the proposed inline calibration mechanism achieves a calibration efficiency of 91.18% on average and negligible performance overhead (i.e., 0.439%)  

| 16:02  | IP3-16| COMPLEMENTARY RESISTIVE SWITCH BASED STATEFUL LOGIC OPERATIONS USING MATERIAL IMPLICATION | Yuanfan Yang¹, Jimon Mathew¹, Driraj K Pradhan¹, Marco Ottavi¹ and Salvatore Pontarelli¹  
¹University of Bristol, GB; ²University of Rome “Tor Vergata”, IT |

**Abstract**  
Memristor based logic and memories are increasingly becoming one of the fundamental building blocks for future system design. Hence, it is important to explore various methodologies for implementing these blocks. In this paper, we present a novel Complementary Resistive Switching (CRS) based stateful logic operations using material implication. The proposed solution benefits from exponential reduction in sneak path current in crossbar implemented logic. We validated the effectiveness of our solution through SPICE simulations on a number of logic circuits. It has been shown that only 4 steps are required for implementing N input NAND gate whereas memristor based stateful logic needs N+1 steps.  

### 7.6.1 Scalability bottlenecks discovery in MPSoC platforms using data mining on simulation traces

**Speakers:**  
Sofiane Lagraa¹, Alexandre Temerier² and Frédéric Pétrot¹  
¹Grenoble institute of Technologie, FR; ²University of Joseph Fourier, FR

**Abstract**  
Nowadays, a challenge faced by many developers is the profiling of parallel applications so that they can scale over more and more cores. This is especially critical for embedded systems powered by Multi-Processor System-on-Chip (MPSoC), where ever demanding applications have to run smoothly on numerous cores, each with modest computational performance. The reasons for the lack of scalability of parallel applications are numerous, and it can be time consuming for a developer to pinpoint the correct one. In this paper, we propose a fully automatic method which detects the instructions of the code which lead to a lack of scalability. The method is based on data mining techniques exploiting low level execution traces produced by MPSoC simulators. Our experiments show the accuracy of the proposed technique on five different kinds of applications, and how the information reported can be exploited by application developers.
16:00 7.6.2 COMPUTING A LANGUAGE-BASED GUARANTEE FOR TIMING PROPERTIES OF CYBER-PHYSICAL SYSTEMS

Speakers: Nell Dhriva, Pratyush Kumar, Georgia Giannopoulos and Lothar Thiele, ETH Zurich, CH

Abstract: Real-time systems are often guaranteed in terms of schedulability, which verifies whether or not all jobs meet their deadlines. However, such a guarantee can be insufficient in certain applications. In this paper, we propose a method to compute a language-based guarantee which provides a more detailed description of the deadline miss patterns of an observed task. The only requirement of our method is that the timing behavior of the real-time system be modelled by a network of timed automata. We compute the language-based guarantee by constructing an equivalent finite state automaton in an iterative manner, using a counter-example guided procedure. We illustrate the language-based guarantee for two applications: design of a networked control system and scheduling in a mixed criticality system. In both cases, we show that the language-based guarantee leads to a more efficient design than the schedulability guarantee.

15:30 7.6.3 RESOURCE OPTIMIZATION FOR CSDF-MODELED STREAMING APPLICATIONS WITH LATENCY CONSTRAINTS

Speakers: Di Liu, Jelena Spasic, Jiali Teddy Zhai, Todor Stefanov and Gang Chen

Abstract: In this paper, we study the problem of minimizing the number of processors required for scheduling latency-constrained streaming applications modeled as CSDF graphs, where the actors of a CSDF are executed as strictly periodic tasks. We formalize the problem and prove that it is an integer convex programming problem, that can be solved efficiently by using an existing convex programming solver. We evaluate our solution approach on a set of 13 real-life streaming applications modeled as CSDF graphs and demonstrate that it can reduce the number of processors in more than 52% of the conducted experiments in comparison to an existing approach.

16:00 IP3-17, 18 A LAYERED APPROACH FOR TESTING TIMING IN THE MODEL-BASED IMPLEMENTATION

Speakers: Jiali Teddy Zhai, Hyeon I Han, Taejoon Park, Sanghyuk Son and Insup Lee

Abstract: The model-based approach is to derive an implementation from a model that has been shown to meet requirements. Even though this approach can be used to guarantee that an implementation satisfies functional requirements that are shown to be correct at the model level, it is still challenging to assure timing requirements at the implementation level. We propose a layered approach in testing timing requirements conformance of implemented systems developed by model-based implementation. In our approach, the abstraction boundary of the implemented system is formally defined using Parnas' four-variables model. Then, the proposed approach tests timing aspects of the interaction between the auto-generated code and the target platform-dependent code based on the four-variables. This approach aims at not only detecting the timing requirement violation, but also at measuring delay-segments that contribute to the timing deviation of the implemented system w.r.t. the model. We show the case study of testing timing requirements of an infusion pump system to illustrate the applicability of the proposed framework.

16:01 IP3-19 MODEL-BASED PROTOCOL LOG GENERATION FOR TESTING A TELECOMMUNICATION TEST HARNESS USING CLP

Speakers: Kenneth Balch, Olga Grinchtein and Justin Pearson

Abstract: Within telecommunications development it is vital to have frameworks and systems to replay complicated scenarios on equipment under test, often there are not enough available scenarios. In this paper we study the problem of testing a test harness, which replays scenarios and analyses protocol logs for the Public Warning System service, which is a part of the Long Term Evolution (LTE) 4G standard. Protocol logs are sequences of messages with timestamps; and are generated by different mobile network entities. In our case study we focus on user equipment protocol logs. In order to test the test harness we require that logs have both incorrect and correct behaviour. It is easy to collect logs from real system runs, but these logs do not show much variation in the behaviour of system under test. We present an approach where we use constraint logic programming (CLP) for both modelling and test generation, where each test case is a protocol log. In this case study, we uncovered previously unknown faults in the test harness.

16:02 IP3-20 TIME-DECOUPLED PARALLEL SYSTEMS SIMULATION

Speakers: Jan Weinstock, Christoph Schumacker, Rainer Leupers, Gerd Ascheid and Laura Tosoratto

Abstract: With increasing system size and complexity, designers of embedded systems face the challenge of efficiently simulating these systems in order to enable target specific software development and design space exploration as early as possible. Today's multicore workstations offer enormous computational power, but traditional simulation engines like the OSi SystemC kernel only operate on a single thread, thereby leaving a lot of computational potential unused. Most modern embedded system designs include multiple processors. This work proposes StOpe, a SystemC kernel that aims at exploiting the inherent parallelism of such systems by simulating the processors on different threads. A loadbalancing mechanism is employed to reduce the required synchronization between the simulation threads, thereby further increasing simulation speed. The virtual prototype of the European FP7 project EUReTiLe system simulator is used as demonstrator for the proposed work, showing a speedup of 4.0x on a four core host system compared to sequential simulation.

16:03 IP3-20 A UNIFIED METHODOLOGY FOR A FAST BENCHMARKING OF PARALLEL ARCHITECTURE

Speakers: Alexandre Guerme, Jean-Thomas Acquaviva and Yves Lhuillier, CEA LIST, FR

Abstract: Benchmarking of architectures is today jeopardized by the explosion of parallel architectures and the dispersion of parallel programming models. Parallel programming requires architecture dependent compilers and languages as well as high programmer expertise. Thus, an objective comparison has become a harder task. This paper presents a novel methodology to evaluate and to compare parallel architectures in order to ease the programmer work. It is based on the usage of micro-benchmarks, code profiling and characterization tools. The main contribution of this methodology is a semi-automatic prediction of the performance for sequential applications on a set of parallel architectures. In addition the performance estimation is correlated with the cost of other criteria such as power or portability. Our methodology prediction was validated on an industrial application. Results are within a range of 20%.
This session covers topics that blend test with fault tolerance, security, and logic placement. As the area of IC test matures the core technology is adapting to the needs of the design and its implementation. As we move forward to advanced nodes in manufacturing the need to tolerate errors could blend with test methods. Test structures provide access to key design IP which is of concern in some situations. Papers in this session address solutions in IC Test for security.

**7.7 Design-for-Test and Test Access**

**Date:** Wednesday, March 26, 2014  
**Time:** 14:30 - 16:00  
**Location / Room:** Konferenz 5

**Chair:**  
Erik Jan Marinissen, IMEC, BE, Contact Erik Jan Marinissen

**Co-Chair:**  
Hans-Joachim Wunderlich, Univ. of Stuttgart, DE, Contact Hans-Joachim Wunderlich

This session covers topics that blend test with fault tolerance, security, and logic placement. As the area of IC test matures the core technology is adapting to the needs of the design and its implementation. As we move forward to advanced nodes in manufacturing the need to tolerate errors could blend with test methods. Test structures provide access to key design IP which is of concern in some situations. Papers in this session address solutions in IC Test for security.

**Time** | **Label** | **Presentation Title** | **Authors**
--- | --- | --- | ---
14:30 | 7.7.1 | BIT-FLIPPING SCAN - A UNIFIED ARCHITECTURE FOR FAULT TOLERANCE AND OFFLINE TEST | Michael Imhof\(^1\) and Hans-Joachim Wunderlich\(^2\)  
\(^1\)Institute of Computer Architecture and Computer Engineering, University of Stuttgart, DE; \(^2\)University of Stuttgart, DE

**Abstract**  
Test is an essential task since the early days of digital circuits. Every produced chip undergoes at least a production test supported by on-chip test infrastructure to reduce test cost. Throughout the technology evolution fault tolerance gained importance and is now necessary in many applications to mitigate soft errors threatening consistent operation. While a variety of effective solutions exists to tackle both areas, test and fault tolerance are often implemented orthogonally, and hence do not exploit the potential synergies of a combined solution. The unified architecture presented here facilitates fault tolerance and test by combining a checksum of the sequential state with the ability to flip arbitrary bits. Experimental results confirm a reduced area overhead compared to an orthogonal combination of classical test and fault tolerance schemes. In combination with heuristically generated test sequences the test application time and test data volume are reduced significantly.

15:00 | 7.7.2 | TESTING PUF-BASED SECURE KEY STORAGE CIRCUITS | Matilda Cortez\(^1\), Gijs Roeelts\(^2\), Said Hamdioui\(^3\) and Giorgio Di Natale\(^4\)  
\(^1\)Delft University of Technology, NL; \(^2\)UVR, FR

**Abstract**  
Design for test is an integral part of any VLSI chip. However, for secure systems extra precautions have to be taken to prevent that the test circuitry could reveal secret information. This paper addresses secure test for Physical Unclonable Function based systems. In particular it provides the testability analysis and a secure Built-In Self-Test (BIST) solution for Fuzzy Extractor (FE) which is the main component of PUF-based systems. The scheme targets high stuck-at-fault (SAF) coverage by performing scan-chain free functional testing, to prevent scan-chain abuse for attacks. The scheme reuses existing FE sub-blocks (for pattern generation and compression) to minimize the area overhead. The scheme is integrated in FE design and simulated; the results show that a SAF fault coverage of 95.1% can be realized with no more than 50k clock cycles at the cost of a negligible area overhead of only 2.2%. Higher fault coverage is possible to realize at extra cost.

15:30 | 7.7.3 | MAKING IT HARDER TO UNLOCK AN LSIB: HONEYTRAPS AND MISDIRECTION IN A P1687 NETWORK | Adam Zygmontowicz, Jennifer Dworak, Al Grouch and John Potterz  
\(^1\)Southern Methodist University, US; \(^2\)ASSET InterTech, US

**Abstract**  
Today’s chips often contain a wealth of embedded instruments and data, including sensors, hardware monitors, built-in self test (BIST) engines, and chip IDs, among others. IEEE P1687 was specifically designed to provide access to such instruments in an efficient manner, and some companies are already implementing the proposed standard on their chips. However, while such instruments provide valuable information and features to authorized users who need to harness them for test, debug, diagnosis, and possibly counterfeet detection, it may be desirable to restrict unauthorized access to them through the P1687 network. Previous work has proposed replacing some of the segment insertion bits (SIBs), which add scan path segments in a P1687 network, withoking SIBs (LSIBs). LSIBs use the data that is naturally scanned through the network as keys to hide instruments from attackers. However, that previous work did not investigate many of the techniques and structures that can be used to significantly increase the time an attacker is likely to unlock LSIBs and gain access to hidden instruments. In this work, we explore some of these techniques and show how simple modifications to a P1687 network protected with LSIBs can significantly increase the difficulty an attacker faces in attempting to access protected instruments.

15:45 | 7.7.4 | CO-OPTIMIZATION OF MEMORY BIST GROUPING, TEST SCHEDULING, AND LOGIC PLACEMENT | Ilgweon Kang and Andrew B. Kahng, UC San Diego, US

**Abstract**  
Built-in self-test (BIST) is a well-known design technique in which part of a circuit is used to test the circuit itself. BIST plays an important role for embedded memories, which do not have pins or pads exposed to the periphery of the chip for testing with automatic test equipment. With the rapidly increasing number of embedded memories in modern SoCs (up to hundreds of memories in each hard macro of the SoC), product designers incur substantial costs of test time (subject to possible power constraints) and BIST logic physical resources (area, routing, power). However, only limited previous work addresses the physical design optimization of BIST logic; notably, Chien et al. [7] optimize BIST design with respect to test time, routing length, and area. In our work, we propose a new three-step heuristic approach to minimize test time as well as test physical layout resources, subject to given upper bounds on power consumption. A key contribution is an integer linear programming (ILP) framework that determines optimal test time for a given cluster of memories using either one or two BIST controllers, subject to test power limits and with full comprehension of available serialization and parallelization. Our heuristic approach integrates (i) generation of a hypergraph over the memories, with test time-aware weighting of hyperedges, along with top-down, FM-style min-cut partitioning; (ii) solution of an ILP that comprehends parallel and serial testing to optimize test scheduling per BIST controller; and (iii) placement of BIST logic to minimize routing and buffering costs. When evaluated on hard macros from a recent industrial 28nm networking SOC, our heuristic solutions reduce test time estimates by up to 11.57% with strictly fewer BIST controllers per hard macro, compared to the industrial solutions.

Date: Wednesday, March 26, 2014
Time: 14:30 - 16:00
Location / Room: Exhibition Theatre

Moderator:
Oliver Bringmann, University of Tübingen, DE, Contact Oliver Bringmann

Fully-Depleted Silicon On Insulator (FD-SOI) is emerging as a promising solution to continue the CMOS scaling roadmap at the 22nm technology node and beyond, especially for low power and System-on-Chip applications. After a short introduction into the FD-SOI technology, this panel discusses the role of FD-SOI as the key enabling technology to tackle the challenges of the major European application domains. This includes the creation of a European ecosystem to provide an easy access for industry and SMEs to a leading-edge semiconductor technology with manageable costs. The panelist take a look at different perspectives and discusses the technology, the SME, the application, the EDA, and the research viewpoint to FD-SOI and its impact to European industry.

16:00 End of session
Coffee Break in Exhibition Area
On Tuesday-Thursday the coffee and lunch breaks will be located in the Exhibition Area (Terrace Level).

IP3 Interactive Presentations

Date: Wednesday, March 26, 2014
Time: 16:00 - 16:30
Location / Room: Conference Level, foyer

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>
<tbody>
<tr>
<td>IP3-1</td>
<td>DESIGN AND FABRICATION OF A 315 µH BONDWIRE MICRO-TRANSFORMER FOR ULTRA-LOW VOLTAGE ENERGY HARVESTING</td>
<td>Enrico Maceri1, Ningning Wang2, Saibal Roy2, Michael Hayes2, Rudi Paolo Paganeli1, Marco Tartagni1 and Aldo Romani1 1iDEI, University of Bologna, Italy; 2Tyndall National Institute, UCC, IE; 3ONR-IEIIT, University of Bologna, IT</td>
</tr>
<tr>
<td>IP3-2</td>
<td>PROVIDING REGULATION SERVICES AND MANAGING DATA CENTER PEAK POWER BUDGETS</td>
<td>Baris Aksanli and Tajana Rosing, University of California San Diego, US</td>
</tr>
<tr>
<td>IP3-3</td>
<td>THE ENERGY BENEFIT OF LEVEL-CROSSING SAMPLING INCLUDING THE ACTUATOR'S ENERGY CONSUMPTION</td>
<td>Burkhard Hensel and Klaus Kabitzsch, Dresden University of Technology, DE</td>
</tr>
<tr>
<td>IP3-4</td>
<td>SKETCHLOG: SKETCHING COMBINATIONAL CIRCUITS</td>
<td>Andrew Becker, David Nova and Paolo Ienne, Ecole Polytechnique Federale de Lausanne, CH</td>
</tr>
</tbody>
</table>
### IP3-5
**Towards Verifying Determinism of SystemC Designs**
**Speakers:** Hicang M. Le and Rolf Drechsler, University of Bremen, DE

**Abstract**
Ensuring the correctness of high-level SystemC designs is an important and challenging problem in today’s Electronic System Level (ESL) methodology. Prevalently, a design is checked against a functional specification given by e.g. a testcase with reference output or a user-defined property. Another research direction takes the view of a SystemC design as a piece of concurrent software. The design is then checked for common concurrency problems and thus, a functional specification is not required. Along this line, several methods for deadlock detection and race analysis have been developed. In this work, we propose to consider a new concurrency verification problem, namely input-output determinism, for SystemC designs. That means for each possible input, the design must produce the same output under any valid process schedule. We argue that determinism verification is stronger than both deadlock detection and race analysis. Beside being an attractive correctness criterion itself, proven determinism helps to accelerate both simulative and formal verification. We also present a preliminary study to show the feasibility of determinism verification for SystemC designs.

### IP3-6
**Using Guided Local Search for Adaptive Resource Reservation in Large-Scale Embedded Systems**
**Speaker:** Timon ter Braak, University of Twente, NL

**Abstract**
To maintain a predictable execution environment, an embedded system must ensure that applications are, in advance, provided with sufficient resources to process tasks, exchange information and to control peripherals. The problem of assigning tasks to processing elements with limited resources, and routing communication channels through a capacitated interconnect is combined into an integer linear programming formulation. We describe a guided local search algorithm to solve this problem at run-time. This algorithm allows for a hybrid strategy where configurations computed at design-time may be used as references to lower the computational overhead at run-time. Computational experiments on a dataset with 100 tasks and 20 processing elements show the effectiveness of this algorithm compared to state-of-the-art solvers CPLEX and Gurobi. The guided local search algorithm finds an initial solution within 100 milliseconds, is competitive for small platforms, scales better with the size of the platform, and has lower memory usage (2-19%).

### IP3-7
**Accelerating Graph Computation with Racetrack Memory and Pointer-Assisted Graph Representation**
**Speakers:** Eunhyek Park¹, Helen Li², Sungjoo Yoo¹ and Sunggu Lee¹
¹POSTECH, KR; ²Univ. of Pittsburgh, US

**Abstract**
The poor performance of NAND Flash memory, such as long access latency and large granularity access, is the major bottleneck of graph processing. This paper proposes an intelligent storage for graph processing which is based on fast and low cost racetrack memory and a pointer-assisted graph representation. Our experiments show that the proposed intelligent storage based on racetrack memory reduces total processing time of three representative graph computations by 40.2%–86.9% compared to the graph processing, GraphChi, which exploits sequential accesses based on normal NAND Flash memory-based SSD. Faster execution also reduces energy consumption by 39.6%–90.0%. The intelligent storage processing capability gives additional 10.5%–16.4% performance improvements and 12.0%–14.4% reduction of energy consumption.

### IP3-8
**PSP-Cache: A Low-Cost Fault-Tolerant Cache Memory Architecture**
**Speakers:** Hamed Farbeh and Seyed Ghassem Miremadi, Sharif University of Technology, IR

**Abstract**
Cache memories constitute a large fraction of processor chip area and are highly vulnerable to soft errors caused by energetic particles. To protect these memories, most of the modern processors employ Error Detection Codes (EDCs) or Error Correction Codes (ECCs). EDCs/ECCs impose significant overheads in terms of area and energy; these overheads increase as a function of interleaving EDCs/ECCs to detect/correct multiple errors. This paper proposes a new cache architecture to minimize the area and energy overheads of EDCs/ECCs in set-associative L1-caches. Simulation results for a 4-way set-associative cache show that the proposed architecture reduces both the area and static power overheads of parity code by about 75% and the dynamic energy overhead by about 73% in comparison to conventional cache architecture. These reduction figures are about 68% and about 66%, respectively, for SEC-DED code. The above reductions are achieved without affecting the error coverage.

### IP3-9
**A Hybrid Non-Volatile SRAM Cell with Concurrent SEU Detection and Correction**
**Speakers:** Pilin Junsang1, Fabrizio Lombardi² and Jie Han²
¹Northeastern University, US; ²University of Alberta, CA

**Abstract**
This paper presents a hybrid non-volatile (NV) SRAM cell with a new scheme for SEU tolerance. The proposed NVSRAM cell consists of a 6T SRAM core and a Resistive RAM (RRAM), made of a 1T and a Programmable Metalization Cell (PMC). The proposed cell has concurrent error detection (CED) and correction capabilities; CED is accomplished using a dual-rail checker, while correction is accomplished by utilizing the restore operation; data from the non-volatile memory element is copied back to the SRAM core. The dual-rail checker utilizes two XOR gates each made of 2 inverters and 2 ambi-polar transistors, hence, it has a hybrid nature. Extensive simulation results are provided. The simulation results show that the proposed scheme is very efficient in terms of numerous figures of merit such as delay and circuit complexity and thus applicable to integrated circuits such as FPGAs requiring secure on-chip non-volatile storage (i.e. LUTs) for multi-context configurability.

### IP3-10
**Battery Aware Stochastic QoS Boosting in Mobile Computing Devices**
**Speakers:** Hao Shen, Qiwen Chen and Qinnu Qiu, Syracuse University, US

**Abstract**
Mobile computing has been woven into everyday lives to a great extent. Their usage is clearly imprinted with user’s personal signature. The ability to learn such signature enables immense potential in workload prediction and resource management. In this work, we investigate the user behavior modeling and apply the model for energy management. Our goal is to maximize the quality of service (QoS) provided by the mobile device (i.e., smartphone), while keeping the risk of battery depletion below a given threshold. A Markov Decision Process (MDP) is constructed from history user behavior. The optimal management policy is solved using linear programming. Simulations based on real user traces validate that, compared to existing battery energy management techniques, the stochastic control performs better in boosting the mobile devices’ QoS without significantly increasing the chance of battery depletion.
A THERMAL RESILIENT INTEGRATION OF MANY-CORE MICROPROCESSORS AND MAIN MEMORY BY 2.5D TSI I/OS

Speakers:
Sih-Gian Wu, Kanwen Wang, Sai Manoj P. D.I, Tsung-Yi Ho and Hao Yu
1Nanyang Technological University, SG; 2National Cheng Kung University, TW

Abstract
One memory-logic-integration design platform is developed in this paper with thermal reliability analysis provided for 2.5D through-silicon interposer (TSI) and 3D through-silicon-via (TSV) based integrations. Temperature-dependent delay and power models have been developed at microarchitecture level for 2.5D and 3D integrations of many-core microprocessors and main memory, respectively. Experiments are performed by general-purpose benchmarks from SPEC CPU2006 and also cloud-oriented benchmarks from Phoenix with the following observations. The memory-logic integration by 3D RC-interconnected TSV I/Os can result in thermal runaway failures due to strong electrical-thermal couplings. On the other hand, the one by 2.5D transmission-line-interconnected TSI I/Os has shown almost the same energy efficiency and better thermal resilience.

LEVERAGING ON-CHIP NETWORKS FOR EFFICIENT PREDICTION ON MULTICORE COHERENCE

Speaker:
Libo Huang, National University of Defense Technology, CN

Abstract
Coherent data prediction is introduced as a promising architectural technique for reducing cache-to-cache accesses in directory protocol. However, limited on-chip resources cause the accuracy of current prediction to be generally low. Low accuracy would result in a large number of unnecessary or incorrect predictions, which would consequently generate excessive network traffic. This leads to large power and performance overhead for coherent memory access. This paper proposes an early abort mechanism (EBT) that leverages NoC design to reduce the negative effect of wrong prediction operations, thus facilitating overall performance improvement and traffic reduction. Using detailed full-system simulations, we conclude that EBT provides a cost-effective solution for designing efficient multicore processors. To the best of our knowledge, this study is the first to leverage on-chip network for the prediction optimization on multicore coherence.

AN ADAPTIVE MEMORY INTERFACE CONTROLLER FOR IMPROVING BANDWIDTH UTILIZATION OF HYBRID AND RECONFIGURABLE SYSTEMS

Speakers:
Vito Giovanni Castellana1, Antonino Tumeo1 and Fabrizio Ferrandi1
1Politecnico di Milano, DEIB, IT; 2Pacific Northwest National Laboratory, US

Abstract
Data mining, bioinformatics, knowledge discovery, social network analysis, are emerging irregular applications that exploit data structures based on pointers or linked lists, such as graphs, unbalanced trees or unstructured grids. These applications are characterized by unpredictable memory accesses and generally are memory bandwidth bound, but also presents large amounts of inherent dynamic parallelism because they can potentially spawn concurrent activities for each one of the element they are exploring. Hybrid architectures, which integrate general purpose processors with reconfigurable accelerators, appear promising target platforms for accelerating irregular applications. These systems often connect to distributed and multi-ported memories, potentially enabling parallel memory operations. However, these memory architectures introduce several challenges, such as the necessity to manage concurrency and synchronization to avoid structural conflicts on shared memory locations and to guarantee consistency. In this paper we present an adaptive Memory Interface Controller (MIC) that addresses these issues. The MIC is a general and customizable solution that can target several different memory structures, and is suitable for High Level Synthesis frameworks. It implements a dynamic arbitration scheme, which avoids conflicts on memory resources at runtime, and supports atomic memory operations, commonly exploited for synchronization directives in parallel programming paradigms. The MIC simultaneously maps multiple accesses to different memory ports, allowing fine grained parallelism exploitation and ensuring correctness also in the presence of irregular and statically unpredictable memory access patterns. We evaluated the effectiveness of our approach on a typical irregular kernel, graph Breadth First Search (BFS), exploring different design alternatives.

ENERGY EFFICIENT IN-MEMORY AES ENCRYPTION BASED ON NONVOLATILE DOMAIN-WALL NANOWIRE

Speakers:
Yuhao Wang1, Pingfan Kong1, Hao Yu1 and Dennis Sylvester2
1Nanyang Technological University, SG; 2University of Michigan, US

Abstract
The widely applied Advanced Encryption Standard (AES) encryption algorithm is critical in secure big-data storage. Data oriented applications have imposed high throughput and low power, i.e., energy efficiency (J/bit), requirements when applying AES encryption. This paper explores an in-memory AES encryption using the newly introduced domain-wall nanowire. We show that all AES operations can be fully mapped to a logic-in-memory architecture by non-volatile domain-wall nanowire, called DW-AES. The experimental results show that DW-AES can achieve the best energy efficiency of 24 pJ/bit, which is 5X and 6.5X times better than CMOS ASIC and ReRAM-CMOL implementations, respectively. Under the same area budget, the proposed DW-AES exhibits 6.4X higher throughput and 29% power saving compared to a CMOS ASIC implementation; 1.7X higher throughput and 74% power reduction compared to a ReRAM-CMOL implementation.

ICE: INLINE CALIBRATION FOR MEMRISTOR CROSSBAR-BASED COMPUTING ENGINE

Speakers:
Boxun Li1, Yu Wang1, Yian Chen1, Helen Li1 and Huazhong Yang1
1Tsinghua University, CN; 2University of Pittsburgh, US

Abstract
The emerging neuromorphic computation provides a revolutionary solution to the alternative computing architecture and effectively extends Moore’s Law. The discovery of the memristor presents a promising hardware realization of neuromorphic systems with incredible power efficiency, allowing efficiently executing the analog matrix-vector multiplication on a nanowire. We show that all AES operations can be fully mapped to a logic-in-memory architecture by non-volatile domain-wall nanowire, called DW-AES. The experimental results show that DW-AES can achieve the best energy efficiency of 24 pJ/bit, which is 5X and 6.5X times better than CMOS ASIC and ReRAM-CMOL implementations, respectively. Under the same area budget, the proposed DW-AES exhibits 6.4X higher throughput and 29% power saving compared to a CMOS ASIC implementation; 1.7X higher throughput and 74% power reduction compared to a ReRAM-CMOL implementation.

COMPLEMENTARY RESISTIVE SWITCH BASED STATEFUL LOGIC OPERATIONS USING MATERIAL IMPLICATION

Speakers:
Yuanfan Yang, Jimson Mathewri, Dhiraj K Pradhan1, Marco Ottavio and Salvatore Pontarelli2
1University of Bristol, GB; 2University of Rome "Tor Vergata", IT

Abstract
Memristor based logic and memories are increasingly becoming one of the fundamental building blocks for future system design. Hence, it is important to explore various methodologies for implementing these blocks. In this paper, we present a novel Complementary Resistive Switching (CRS) based stateful logic operations using material implication. The proposed solution benefits from exponential reduction in sneak path current in crossbar implemented logic. We validated the effectiveness of our solution through SPICE simulations on a number of logic circuits. It has been shown that only 4 steps are required for implementing N input NAND gate whereas memristor based stateful logic needs N+1 steps.
A LAYERED APPROACH FOR TESTING TIMING IN THE MODEL-BASED IMPLEMENTATION

Speakers:
BaekGyu Kim, Hyeon I Hwang, Taejoon Parkz, Sanghyuk Sonz and Insup Leez
1University of Pennsylvania, US; 2Daegu Gyeongbuk Institute of Science & Technology, KR

Abstract
The model-based implementation is to derive an implementation from a model that has been shown to meet requirements. Even though this approach can be used to guarantee that an implementation satisfies functional requirements that are shown to be correct at the model level, it is still challenging to assure timing requirements at the implementation level.

We propose a layered approach to testing timing requirements conformance of implemented systems developed by model-based implementation. In our approach, the abstraction boundary of the implemented system is formally defined using Parnas' four-variables model. Then, the proposed approach tests timing aspects of the interaction between the auto-generated code and the target platform-dependent code based on the four-variables. This approach aims at not only detecting the timing requirement violation, but also at measuring delay-segments that contribute to the timing deviation of the implemented system w.r.t. the model. We show the case study of testing timing requirements of an infusion pump system to illustrate the applicability of the proposed framework.

MODEL-BASED PROTOCOL LOG GENERATION FOR TESTING A TELECOMMUNICATION TEST HARNESS USING CLP

Speakers:
Kenneth Balki, Olga Grinchtein and Justin Pearson
1Ericsson AB, SE; 2Uppsala University, SE

Abstract
Within telecommunications development it is vital to have frameworks and systems to replay complicated scenarios on equipment under test, often there are not enough available scenarios. In this paper we study the problem of testing a test harness, which replays scenarios and analyses protocol logs for the Public Warning System service, which is a part of the Long Term Evolution (LTE) 4G standard. Protocol logs are sequences of messages with timestamps; and are generated by different mobile network entities. In our case study we focus on user equipment protocol logs. In order to test the test harness we require that logs have both incorrect and correct behaviour. It is easy to collect logs from real system runs, but these logs do not show much variation in the behaviour of system under test. We present an approach where we use constraint logic programming (CLP) for both modelling and test generation, where each test case is a protocol log. In this case study, we uncovered previously unknown faults in the test harness.

TIME-DECOUPLLED PARALLEL SYSTEMIC SIMULATION

Speakers:
Jan Weinstock1, Christoph Schumacher1, Rainer Leupers1, Gerd Ascheid1 and Laura Tosoratto2
1RWTH Aachen, DE; 2Istituto Nazionale di Fisica Nucleare, Sezione di Roma, IT

Abstract
With increasing system size and complexity, designers of embedded systems face the challenge of efficiently simulating these systems in order to enable target specific software development and design space exploration as early as possible. Today’s multicore workstations offer enormous computational power, but traditional simulation engines like the OSCI SystemC kernel only operate on a single thread, thereby leaving a lot of computational potential unused. Most modern embedded system designs include multiple processors. This work proposes STope, a SystemC kernel that aims at exploiting the inherent parallelism of such systems by simulating the processors on different threads. A lookahead mechanism is employed to reduce the required synchronization between the simulation threads, thereby further increasing simulation speed. The virtual prototype of the European FP7 project EURETILE system simulator is used as demonstrator for the proposed work, showing a speedup of 4.01x on a four core host system compared to sequential simulation.

A UNIFIED METHODOLOGY FOR A FAST BENCHMARKING OF PARALLEL ARCHITECTURE

Speakers:
Alexandre Guerre, Jean-Thomas Acquaviva and Yves Lhuillier, CEA LIST, FR

Abstract
Benchmarking of architectures is today jeopardized by the explosion of parallel architectures and the dispersion of parallel programming models. Parallel programming requires architecture dependent compilers and languages as well as high programmer expertise. Thus, an objective comparison has become a harder task. This paper presents a novel methodology to evaluate and to compare parallel architectures in order to ease the programmer work. It is based on the usage of micro-benchmarks, code profiling and characterization tools. The main contribution of this methodology is a semi-automatic prediction of the performance for sequential applications on a set of parallel architectures. In addition the performance estimation is correlated with the cost of other criteria such as power or portability. Our methodology prediction was validated on an industrial application.

Results are within a range of 20%.

VIDEO-BASED ABSOLUTE NAVIGATION APPROACH: A NOVEL APPROACH FOR VIDEO-BASED ABSOLUTE NAVIGATION IN SPACE EXPLORATION MISSIONS

Authors:
Pascal Trotta, Tadewos Getahun Tadewos, Paolo Prinetto, Daniele Rolfo and Pascal Trotta, Politecnico di Torino, IT

Abstract
Nowadays, space agencies have increased their research efforts in order to enhance the success rate of space exploration missions. Future space missions will increasingly adopt video-based navigation (VBN) systems to assist the entry, descent and landing (EDL) phase of space modules. This poster will show a preliminary work on a novel approach for Video-based Absolute Navigation (VBA). Moreover, the poster will highlight the peculiarities of the proposed approach.

More information ...
In this session we will review several practical applications of virtual prototyping for architecture design work and software development across different markets such as mobile, industrial and embedded systems. Today’s societal needs for innovative products in terms of communication, mobility, health, entertainment, and safety directly impact microelectronics design methodologies. The performance of printed devices depends on the geometry, but is also affected by processing steps of other components integrated onto the same substrate. Since different designs use different devices, process stack, models and design rules must be dynamically determined. In this work we propose and demonstrate an experimental design flow to allow efficient design of hybrid and printed electronic circuits.

GEMINI: A NEW SYNTHESIS AND OPTIMIZATION TOOL FOR GRAPHENE-BASED DIGITAL DEVICES

Gemini is a synthesis and optimization software for graphene-based digital devices. Given a combinational circuit description through its boolean representation, Gemini produces a SPICE netlist maitled with graphene PN-Junction gates. The software is composed of a parser library to handle input circuit descriptions, a characterization library of graphene gates used in the synthesis process, a Biconditional Binary Decision Diagram library used to manipulate logic networks in Pass-XNOR logic in order to better exploit the intrinsic characteristics of the adopted graphene gates, and a number of optimization algorithms designed to produce better results in terms of area and thus power consumption. As a stand-alone software or as a library easy to integrate into state-of-the-art tools, Gemini represents a first step of an enabling technology for future synthesis and optimization processes for graphene-based devices.

END OF SESSION
Low power consumption of electronic devices has been an important requirement for many cyber-physical systems in field. Today, power dissipation is often estimated by spreadsheet-based power analysis. A leading-edge high-level power analysis method has the objective of providing high confidence levels in early design stages, where power design decisions have severe impact. This work examines and compares three high-level power analysis approaches (spreadsheet-based, Synopsys Platform Architect MCO, and DOCEA Aceptor) by an industrial use case.

SYSTEM-LEVEL DESIGN METHODOLOGY ENABLING FAST DEVELOPMENT OF BASEBAND MP-SOC FOR 4G SMALL CELL BASE STATION

Authors
Shan Tang, Zhu Ziyuan and Yongtao Su, Institute of Computing Technology, Chinese Academy of Sciences, CN

Abstract
“Small Cell” is regarded as the solution to optimize 4G wireless networks with improved coverage and capacity and expected to be deploy in a large number. To meet performance requirements and special constraints on the cost and size, we design a heterogeneous multi-processor SoC for small cell base station, which is composed of ASP (Application Specific Processor) cores, hardware accelerators, general-purpose processor core, and infrastructure and interface blocks. The challenges of developing such a complex chip drive us to employ system-level design methodology in both single core and multi-core architecture optimizations. The paper discusses in detail the LISA (Language for Instruction-Set Architectures)/SystemC based ASP-algorithm joint optimization, and task-graph driven multicore architecture exploration. Finally, the results of silicon implementation on SMIC 55nm technology are presented.

VIRTUAL PROTOTYPE LIFE CYCLE IN AUTOMOTIVE APPLICATIONS

Authors
Manfred Thanner, Freescale, Germany, DE

Abstract
Virtual prototypes for automotive applications see a unique life cycle in the context of the supply chain from semiconductor to Tier 1 to OEs and within the eco-system. The presentation gives an overview of current experiences and finding in this field and challenges observed. The virtual platforms targeting the mid to high end application spaces of chassis, to powertrain and driver information systems. The use cases primarily address today semiconductor internal developments and Tier1 level deployment. Additionally different software vendors use the models in their development cycle which drive model requirements like stimulus and abstraction levels. The development of virtual prototypes often start with the reuse of existing cores, accelerators and IP models. These models had certain use cases to address and were created accordingly. Therefore the models sometimes don't necessarily match fully the requirements of the overall virtual prototype and compromises were made. Further to this, models are often from different design centers, vendors, etc. This can lead to conflicting model features versus the primary use case requirements of the virtual platform for the intended usage. Examples are cycle accuracy vs. functional, correct behavior vs. error behavior and error injection. The virtual platform life cycle is also affected by the availability and integration of 3rd party IP models which adds the commercial terms and license dependency. Further to this, the virtual prototypes need to be integrated or connected to the EDA environments of the “receiving companies”. In the deployment phase of the virtual prototype within the automotive eco system a supply chain needs to be in place. This creates challenges in terms of model interfaces, tool compatibility and integration and support chain.

DATE Party in “Gläserne Manufaktur” of the Volkswagen AG

The DATE Party is again scheduled on the second conference day, Wednesday, March 26, 2014, starting from 19:30 h. This year, it will take place in one of Dresden’s most exciting and modern buildings, the “Gläserne Manufaktur” of the car manufacturer Volkswagen AG (www.glaesernemanufaktur.de/en/). The party will feature a flying buffet style dinner with various catering points and accompanying drinks. Light background music and the possibility of guided visits through the extraordinary premises will round off the evening. It provides a perfect opportunity to meet friends and colleagues in a relaxed atmosphere while enjoying local amenities. Please kindly note that it is no seated dinner. All delegates, exhibitors and their guests are encouraged to attend the party. Please be aware that entrance is only possible with a party ticket. Each full conference registration includes a ticket for the DATE Party. Additional tickets can be purchased on-site at the registration desk (subject to availability of tickets). Ticket price for the full evening Social Programme: 75 € per person.

8.2 Hot Topic: Near Threshold Computing (NTC)

Date: Wednesday, March 26, 2014
Time: 17:00 - 18:30
Location / Room: Konferenz 6

Organiser:
Michael Huebner, Ruhr-University Bochum, Ge, Contact Michael Hübner

Chair:
Michael Huebner, Ruhr-University Bochum, DE, Contact Michael Hübner

To face with the power/utilization wall, Near-Threshold Computing (NTC) has emerged as one of the most promising approach to achieve an order of magnitude improvement or more in energy efficiency of microprocessors and reconfigurable hardware. NTC takes advantage of the quadratic relation between the supply voltage (Vdd) and the dynamic power, by lowering the supply voltage of chips to a value only slightly higher than the threshold voltage.
**Physical Attacks are a major security threat for embedded system applications. This session focuses on several aspects of this problem. The presented papers range from countermeasures against power analysis and fault-based attacks, including electromagnetic and laser injections.**

**8.3 Physical Attacks and countermeasures**

**Date:** Wednesday, March 26, 2014

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

**Location / Room:** Konferenz 1

**Chair:** Francesco Regazzoni, Alari, CH; Contact Francesco Regazzoni

**Co-Chair:** Shivam Bhasin, Telecom Paristech, FR; Contact Shivam Bhasin

Physical Attacks are a major security threat for embedded system applications. This session focuses on several aspects of this problem. The presented papers range from countermeasures against power analysis and fault-based attacks, including electromagnetic and laser injections.
The session presents energy and performance efficient implementations of wireless communication and financial applications.

Amer Baghdadi, Telecom Bretagne, FR,

Co-Chair:
Sergio Saponara, University of Pisa, IT,

Chair:

Location / Room:

Time:

Date:
Wednesday, March 26, 2014

8.3 Efficient Designs for Telecom and Financial Applications

17:00 8.3.1 EFFICIENCY OF A GLITCH DETECTOR AGAINST ELECTROMAGNETIC FAULT INJECTION

Speakers:
Loïc Zussa1, Amine Dehbaoui1, Karim Tobich2, Jean-Max Dutertre1, Philippe Maurinet2, Ludovic Guillaume-Sage3, Jessy Clediere3 and Assia Traia4

1ENSIMSE, FR; 2LIRMM, FR; 3CEA, FR

Abstract
The use of electromagnetic glitches has recently emerged as an effective fault injection technique for the purpose of conducting physical attacks against integrated circuits. First research works have shown that electromagnetic faults are induced by timing constraint violations and that they are also located in the vicinity of the injection probe. This paper reports the study of the efficiency of a glitch detector against EM-injection. This detector was originally designed to detect any attempt of inducing timing violations by means of clock or power glitches. Because electromagnetic disturbances are more local than global, the use of a single detector proved to be inefficient. Our subsequent investigation of the use of several detectors to obtain a full fault detection coverage is reported, it also provides further insights into the properties of electromagnetic injection and into the key role played by the injection probe.

17:30 8.3.2 ANALYZING AND ELIMINATING THE CAUSES OF FAULT SENSITIVITY ANALYSIS

Speakers:
Nahid Farzandy Ghalaty, Aydin Aysu and Patrick Schaumont, Virginia Tech, US

Abstract
Fault Sensitivity Analysis (FSA) is a new type of side-channel attack that exploits the relation between the sensitive data and the faulty behavior of a circuit, the so-called fault sensitivity. This paper analyzes the behavior of different implementations of AES S-box architectures against FSA, and proposes a systematic countermeasure against this attack. This paper has two contributions. First, we study the behavior and structure of several S-box implementations, to understand the causes behind the fault sensitivity. We identify two factors: the timing of fault sensitive paths, and the number of logic levels of fault sensitive gates within the netlist. Next, we propose a systematic countermeasure against FSA. The countermeasure masks the effect of these factors by intelligent insertion of delay elements. We evaluate our methodology by means of an FPGA prototype with built-in timing-measurement. We show that FSA can be thwarted at low hardware overhead. Compared to earlier work, our method operates at the logic-level, is systematic, and can be easily generalized to bigger circuits.

18:00 8.3.3 A SMALLER AND FASTER VARIANT OF RSM

Speakers:
Noritaka Yamashita, Kazuhiro Minematsu, Toshihiko Okamura and Yukiyasu Tsono, NEC, JP

Abstract
Masking is one of the major countermeasures against side-channel attacks to cryptographic modules. Nassar et al. recently proposed a highly efficient masking method, called Rotating S-boxes Masking (RSM), which can be applied to a block cipher based on Substitution-Permutation Network. It arranges multiple masked S-boxes in parallel, which are rotated in each round. This rotation requires remasking process for each round to adjust current masks to those of the S-boxes. In this paper, we propose a method for reducing the complexity of RSM further by omitting the remasking process when the linear diffusion layer of the encryption algorithm has a certain algebraic property. Our method can be applied to AES with a reduced complexity from RSM, while keeping the equivalent security level.

18:30 IP4.1, 140 A MULTIPLE FAULT INJECTION METHODOLOGY BASED ON CONE PARTITIONING TOWARDS RTL MODELING OF LASER ATTACKS

Speakers:
Athanasios Papadimitriou1, David Holy1, Vincent Beroulle1, Paolo Maistri1 and Regis Leveugle3

1LCIS Laboratory - Grenoble INP, FR; 2TIMA Laboratory / Grenoble INP, FR; 3TIMA Laboratory / Grenoble INP, FR

Abstract
Laser attacks, especially on circuits manufactured with recent deep submicron semiconductor technologies, pose a threat to secure integrated circuits due to the multiplicity of errors induced by a single attack. An efficient way to neutralize such effects is the design of appropriate countermeasures, according to the circuit implementation and characteristics. Therefore tools which allow the early evaluation of security implementations are necessary. Our efforts involve the development of an RTL fault injection approach more representative of laser attacks than random multi-bit fault injections and the utilization and evolution of state of the art emulation techniques to reduce the duration of the fault injection campaigns. This will ultimately lead to the design and validation of new countermeasures against laser attacks, on ASICs implementing cryptographic algorithms.

19:00 8.4 Efficient Designs for Telecom and Financial Applications

Date:
Wednesday, March 26, 2014

Time:
17:00 - 18:30

Location / Room:
Konferenz 2

Chair:
Sergio Saponara, University of Pisa, IT, Contact Sergio Saponara

Co-Chair:
Amer Baghdadi, Telecom Bretagne, FR, Contact Amer Baghdadi

The session presents energy and performance efficient implementations of wireless communication and financial applications.
End of session
The presentation proposes a novel formal notation for reactive system requirements in order to reduce translational efforts and thus make specifications both easier and quicker to create.

The second presentation provides an algorithm for automatic aggregation of design blocks based upon their static computation demands. Finally, the last presentation proposes a novel formal notation for reactive system requirements in order to reduce translational efforts and thus make specifications both easier and quicker to create.

The first presentation proposes an analytical model to estimate the contention and the resulting delays on accessing shared components in a multi-core environment. In order to find the right granularity for design space exploration, the second presentation provides an algorithm for automatic aggregation of design blocks based upon their static computation demands. Finally, the last presentation proposes a novel formal notation for reactive system requirements in order to reduce translational efforts and thus make specifications both easier and quicker to create.

8.5 Modeling & Specification

Date: Wednesday, March 26, 2014
Time: 17:00 - 18:30
Location / Room: Konferenz 3

Chair:
Wolfgang Mueller, University of Paderborn, DE, Contact Wolfgang Mueller

Co-Chair:
Francois Pecheux, UPMC, FR, Contact Francois Pecheux

The first presentation proposes an analytical model to estimate the contention and the resulting delays on accessing shared components in a multi-core environment. In order to find the right granularity for design space exploration, the second presentation provides an algorithm for automatic aggregation of design blocks based upon their static computation demands. Finally, the last presentation proposes a novel formal notation for reactive system requirements in order to reduce translational efforts and thus make specifications both easier and quicker to create.

In this paper, we introduce an automatic specification granularity tuning mechanism to determine the granularity in the synthesized specification model hierarchy guided by the computational demands of algorithm blocks. Our granularity selection significantly simplifies the early design space exploration yet sacrificing the mapping flexibility. In this paper, we introduce an automatic specification granularity tuning mechanism to determine the granularity in the synthesized specification model hierarchy guided by the computational demands of algorithm blocks. Our granularity selection significantly simplifies the early design space exploration yet sacrificing the mapping flexibility. In this paper, we introduce an automatic specification granularity tuning mechanism to determine the granularity in the synthesized specification model hierarchy guided by the computational demands of algorithm blocks. Our granularity selection significantly simplifies the early design space exploration yet sacrificing the mapping flexibility.

Algorithm Design Environments (ADE), such as Simulink, have been shown to be efficient for development, analysis, and evaluation of algorithms. Recent tools propose to facilitate algorithm/architecture co-design by bridging the gap from ADE to System-Level Design Environments (SLDE) through automatic synthesis from algorithm models to SLDE specifications. With the wide range of block characteristics (from simple logic functions to complex kernels) in the algorithm model, however, it is challenging to select a suitable compositional granularity for SLDE (SLDL) blocks in the synthesized specification. A high volume of SLDL blocks of little computation will increase the number of mapping possibilities, whereas large blocks with heavy computation on the other hand allow inter-block fusion reducing the computational demands of the overall specification yet sacrificing the mapping flexibility. In this paper, we introduce an automatic specification granularity tuning mechanism to determine the granularity in the synthesized specification model hierarchy guided by the computational demands of algorithm blocks. Our granularity selection significantly simplifies the early design space exploration yet sacrificing the mapping flexibility.

As modern systems are integrating exceeding number of components for better performance and functionality, early full-system simulation tools have become essential for validating complex concurrent system interaction activities. In the past decades, many useful timing-accurate system simulation tools have been developed; however, we find that even for the most efficient techniques, more than 90% of overhead occurs when simulating shared devices, such as buses. Instead of adopting the constant-delay model that compromises accuracy or using the time-consuming precise scheduling approach, we propose in this paper an effective system activity-sensitive contention delay model that can dynamically capture runtime contention situations and system configuration changes. To verify the idea, we construct an analytical bus delay model and integrate that into a system simulation tool. The experimental results show 20 to 80 times performance improvement over the scheduling-based bus model on full-system simulations and the estimated timing difference is less than 3%.

As modern systems are integrating exceeding number of components for better performance and functionality, early full-system simulation tools have become essential for validating complex concurrent system interaction activities. In the past decades, many useful timing-accurate system simulation tools have been developed; however, we find that even for the most efficient techniques, more than 90% of overhead occurs when simulating shared devices, such as buses. Instead of adopting the constant-delay model that compromises accuracy or using the time-consuming precise scheduling approach, we propose in this paper an effective system activity-sensitive contention delay model that can dynamically capture runtime contention situations and system configuration changes. To verify the idea, we construct an analytical bus delay model and integrate that into a system simulation tool. The experimental results show 20 to 80 times performance improvement over the scheduling-based bus model on full-system simulations and the estimated timing difference is less than 3%.

Our approach achieves a significant reduction in overhead by leveraging the block fusion capabilities in the ADE. At the same time our granularity decision ensures that sufficient flexibility remains in the system for exploring heterogeneous mapping of the algorithm. Our results on real world examples show that specification models can be synthesized with 80% efficiency through block fusion with 70-90% fewer but coarser grained blocks.
8.6 Mapping and Scheduling for Many-Core Embedded Systems

Date: Wednesday, March 26, 2014
Time: 17:00 - 18:30
Location / Room: Konferenz 4

Chair:
Marc Geilen, Eindhoven University of Technology, NL, Contact Marc Geilen

Co-Chair:
Sébastien Le Beux, Ecole Centrale de Lyon, FR, Contact Sébastien Le Beux

This session discusses novel ideas for embedded software implementation on many-core architectures. The first presentation deals with an optimized implementation of a H265 video coding algorithm on many-core architectures. A run-time scheduling approach for GPGPU architectures for priority-based systems is presented in the second presentation. The third talk presents an efficient run-time resource manager heuristic for many-core architectures based on a Lagrangian relaxation technique.

Time | Label | Presentation Title | Authors
--- | --- | --- | ---
17:00 | 8.6.1 | SOFTWARE ARCHITECTURE OF HIGH EFFICIENCY VIDEO CODING FOR MANY-CORE SYSTEMS WITH POWER-EFFICIENT WORKLOAD BALANCING | Muhammad Usman Karim Khan, Muhammad Shafique and Jörg Henkel, Karlsruhe Institute of Technology (KIT), DE

Abstract
The High Efficiency Video Coding (HEVC) standard aims at providing ~50% better compression compared to its predecessor (H.264) at the cost of high computational complexity. To enable HEVC video encoding in real-time scenarios, special coding support for parallelization is provided in HEVC that can be exploited by many-core systems. In this work, we present a HEVC software architecture where a video frame is adaptively divided into independent video frame regions (i.e. so-called video tiles) which are processed concurrently on multiple cores. By balancing the workload of each video tile mapped to a particular core, the total power consumption of a system is reduced (through dynamically scaling the operating frequency) under a given frame-rate constraint. We also exploit user tolerance to further curtail the HEVC workload with insignificant video quality degradation. Experimental results illustrate that the proposed approach results in ~43% power savings on a many-core system.
This paper presents an application-driven algorithm for Dynamic Thermal Management (DTM) for the High Efficiency Video Coding (HEVC). For efficient design of such a DTM policy, we perform an offline thermal analysis of an HEVC encoder and demonstrate the impact of different video sequences and different coding configurations on the processor temperature. Our thermal analysis is leveraged to develop an efficient application-driven DTM policy that performs temperature-aware coding along with an application-driven control of DTM knobs (e.g., frequency scaling) in order to meet the temperature constraints while still providing high video quality (i.e., PSNR loss < 0.01dB). For accurate thermal analysis and evaluation, we deploy an infrared camera-based thermal measurement setup that, on the contrary to state-of-the-art setups, does not require adding any extra layer on top of the measured chip, thus allowing the camera to accurately capture the infrared emissions from the die.
8.7 Performance Modeling and Delay Test

Date: Wednesday, March 26, 2014
Time: 17:00 - 18:30
Location / Room: Konferenz 5

Chair: Robert Aitken, ARM, US, Contact Robert Aitken

Co-Chair: Mehdi Tahoori, KIT, DE, Contact Mehdi Tahoori

As technology dimensions shrink and process complexity increases, it becomes vital to accurately model performance limiters such as device and metal variability, as well as to determine when these effects become so critical that delay requirements are exceeded.

---

8.7.2 JOINT VIRTUAL PROBE: JOINT EXPLORATION OF MULTIPLE TEST ITEMS’ SPATIAL PATTERNS FOR EFFICIENT SILICON CHARACTERIZATION AND TEST PREDICTION

Speakers: Shuangyue Zhang1, Fan Lin2, Chun-Kai Hsu2, Kwang-Ting Cheng2 and Hong Wang1
1Department of Automation, Tsinghua University, CN; 2Department of Electrical and Computer Engineering, University of California, Santa Barbara, US

Abstract
Virtual Probe (VP), proposed for characterization of spatial variations and for test time reduction, can effectively reconstruct the spatial pattern of a test item on an entire wafer using measurement values from only a small fraction of dies on the wafer. However, VP calculates the spatial signature of each test item separately, one item at a time, resulting in very long runtime for complex chips which often require hundreds, or even thousands, of test items in production. In this paper, we propose a new method, named Joint Virtual Probe (JVP), which can jointly derive spatial patterns of multiple test items. By simultaneously handling a large group of test items, JVP significantly reduces the overall runtime. And the prediction accuracy can also be improved because of JVP’s implicit use of inter-test-item correlations in predicting spatial patterns. The experimental results on two industrial products, with 277 and 985 parametric test items in the production test programs respectively, demonstrate that JVP achieves an average speedup of 170X and 50X over VP in the pre-test analysis and the test application phases respectively, as well as a slightly higher prediction accuracy than VP.
8.8 Hot Topic: Beyond CMOS Ultra-low-power Computing

Date: Wednesday, March 26, 2014
Time: 17:00 - 18:30
Location / Room: Exhibition Theatre

Organiser:
Saibal Mukhopadhyay, Georgia Institute of Technology, US, Contact Saibal Mukhopadhyay

Chair:
Arijit Raychowdhury, Georgia Institute of Technology, US, Contact Arijit Raychowdhury

Co-Chair:
Saibal Mukhopadhyay, Georgia Institute of Technology, US, Contact Saibal Mukhopadhyay

With conventional CMOS scaling becoming increasingly challenging, the designers wonder what opportunities and challenges exist beyond CMOS for both Boolean and non-Boolean computing. This session will discuss three very different and promising emerging technologies – Tunneling Field-Effect Transistor, Spintronics, and nano-electro-mechanical switches (NEMS) – for low-power electronics. The talks will discuss the need for innovating and evaluating new circuit and system design methods as new-device technologies emerge.

18:30 | End of session
19:30 | DATE Party in “Gläserne Manufaktur” of the Volkswagen AG
The DATE Party is again scheduled on the second conference day, Wednesday, March 26, 2014, starting from 19:30 h. This year, it will take place in one of Dresden’s most exciting and modern buildings, the “Gläserne Manufaktur” of the car manufacturer Volkswagen AG (www.glaesernemanufaktur.de/en/). The party will feature a flying buffet style dinner with various catering points and accompanying drinks. Light background music and the possibility of guided visits through the extraordinary premises will round off the evening. It provides a perfect opportunity to meet friends and colleagues in a relaxed atmosphere while enjoying local amenities. Please kindly note that it is a no seated dinner. All delegates, exhibitors and their guests are encouraged to attend the party. Please be aware that entrance is only possible with a party ticket. Each full conference registration includes a ticket for the DATE Party. Additional tickets can be purchased on-site at the registration desk (subject to availability of tickets). Ticket price for the full Evening Social Programme: 75 € per person.

8.8.1 ULTRA-LOW POWER ELECTRONICS WITH SI/GE TUNNEL FET
Speakers:
Amit Trivedi, Mohammad Faisal Amir and Saibal Mukhopadhyay, Georgia Institute of Technology, US

Abstract
SiGe Tunnel FET (TFET) with its subthermal subthreshold swing is attractive for low power analog and digital designs. Greater Ion/Ioff ratio of TFET can reduce the dynamic power in digital designs, while higher gm/IDS can lower the bias power of analog amplifier. However, the above benefits of TFET are eclipsed by MOSFET at a higher power/performance point. Ultra low power scalability of the key analog and digital circuits, SRAM and operational transconductance amplifier (OTA), with TFET is demonstrated. Analyzing a TFET based cellular neural network, this work shows the feasibility of ultra-low-power neuromorphic computing with TFET.

8.8.2 BRAIN-INSPIRED COMPUTING WITH SPIN TORQUE DEVICES
Speakers:
Kaushik Roy, Mrigank Sharad, Deliang Fan and Karthik Yogendra, Purdue University, US

Abstract
In this paper we discuss the potential of emerging spin-torque devices for computing applications. Recent proposals for spin-based computing schemes may be differentiated as ‘all-spin’ vs. hybrid, programmable vs. fixed, and, Boolean vs. non-Boolean. All-spin logic-styles may offer high area-density due to small form-factor of nano-magnetic devices. However, circuit and system-level design techniques need to be explored that leverage the specific spin-device characteristics to achieve energy-efficiency, performance and reliability comparable to those of CMOS. The non-volatility of nano-magnets can be exploited in the design of energy and area-efficient programmable logic. In such logic-styles, spin-devices may play the dual-role of computing as well as memory-elements that provide field-programmability. Spin-based threshold logic design is presented as an example. Emerging spintronic phenomena may lead to ultra-low-voltage, current-mode, spin-torque switches that can offer attractive computing capabilities, beyond digital switches. Such devices may be suitable for non-Boolean data-processing applications which involve analog processing. Integration of such spin-torque devices with charge-based devices like CMOS and resistive memory can lead to highly energy-efficient information processing hardware for applications like pattern-matching, neuromorphic-computing, image-processing and data-conversion. Finally, we discuss the possibility of using coupled spin-torque nano oscillators for low-power non-Boolean computing.
carbon electronics and the most promising devices in this class.

Growing number of important application areas, including automotive and industrial applications as well as space, avionics, combustion engine, intelligent propulsion systems, and geo-thermal exploration require electronics that can work reliable at extreme conditions - in particular at a temperature > 250 °C and at high radiation (1-30 Mrad), where conventional electronics fail to work reliably. Traditionally, existing wide-band-gap semiconductors, e.g., silicon carbide (SiC) transistor-based electronics have been considered most viable for high temperature and high radiation applications. However, the large-size, high threshold voltage, low switching speed and high leakage current make logic design with these devices unattractive. Additionally, the leakage current markedly increases at high temperature (in the range of 10 µA for a 2-input NAND gate), which induces self-heating effect and makes power delivery at high temperature very challenging. To address these issues, in this paper we present a computing platform for low-power reliable operation at extreme environment using SiC electromechanical switches. We show that a device-circuit-architecture co-design approach can provide reliable long-term operation with virtually zero leakage power.

**Abstract**

Growing number of important application areas, including automotive and industrial applications as well as space, avionics, combustion engine, intelligent propulsion systems, and geo-thermal exploration require electronics that can work reliable at extreme conditions - in particular at a temperature > 250 °C and at high radiation (1-30 Mrad), where conventional electronics fail to work reliably. Traditionally, existing wide-band-gap semiconductors, e.g., silicon carbide (SiC) transistor-based electronics have been considered most viable for high temperature and high radiation applications. However, the large-size, high threshold voltage, low switching speed and high leakage current make logic design with these devices unattractive. Additionally, the leakage current markedly increases at high temperature (in the range of 10 µA for a 2-input NAND gate), which induces self-heating effect and makes power delivery at high temperature very challenging. To address these issues, in this paper we present a computing platform for low-power reliable operation at extreme environment using SiC electromechanical switches. We show that a device-circuit-architecture co-design approach can provide reliable long-term operation with virtually zero leakage power.

**Abstract**

Growing number of important application areas, including automotive and industrial applications as well as space, avionics, combustion engine, intelligent propulsion systems, and geo-thermal exploration require electronics that can work reliable at extreme conditions - in particular at a temperature > 250 °C and at high radiation (1-30 Mrad), where conventional electronics fail to work reliably. Traditionally, existing wide-band-gap semiconductors, e.g., silicon carbide (SiC) transistor-based electronics have been considered most viable for high temperature and high radiation applications. However, the large-size, high threshold voltage, low switching speed and high leakage current make logic design with these devices unattractive. Additionally, the leakage current markedly increases at high temperature (in the range of 10 µA for a 2-input NAND gate), which induces self-heating effect and makes power delivery at high temperature very challenging. To address these issues, in this paper we present a computing platform for low-power reliable operation at extreme environment using SiC electromechanical switches. We show that a device-circuit-architecture co-design approach can provide reliable long-term operation with virtually zero leakage power.

**Abstract**

Growing number of important application areas, including automotive and industrial applications as well as space, avionics, combustion engine, intelligent propulsion systems, and geo-thermal exploration require electronics that can work reliable at extreme conditions - in particular at a temperature > 250 °C and at high radiation (1-30 Mrad), where conventional electronics fail to work reliably. Traditionally, existing wide-band-gap semiconductors, e.g., silicon carbide (SiC) transistor-based electronics have been considered most viable for high temperature and high radiation applications. However, the large-size, high threshold voltage, low switching speed and high leakage current make logic design with these devices unattractive. Additionally, the leakage current markedly increases at high temperature (in the range of 10 µA for a 2-input NAND gate), which induces self-heating effect and makes power delivery at high temperature very challenging. To address these issues, in this paper we present a computing platform for low-power reliable operation at extreme environment using SiC electromechanical switches. We show that a device-circuit-architecture co-design approach can provide reliable long-term operation with virtually zero leakage power.
<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:50</td>
<td>9.1.2</td>
<td>ADVANCED SYSTEM ON A CHIP DESIGN BASED ON CONTROLLABLE-POLARITY FETS</td>
<td>Pierre-Emmanuel Gallardon, Luca Amaru, Jian Zhang and Giovanni De Michiel, Integrated Systems Laboratory – Swiss Federal Institute of Technology, CH</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
<td><strong>Abstract</strong> — Field Effect Transistors (FETs) with on-line controllable polarity are promising candidates to support next generation System-on-Chip (SoC). Thanks to their enhanced functionality, controllable polarity FETs enable a superior design of critical components in a SoC, such as processing units and memories, while also providing native solutions to control power consumption. In this paper, we present the efficient design of a SoC core with controllable polarity FET. Processing units are speeded-up at the datapath level, as arithmetic operations require fewer physical resources than in standard CMOS. Power consumption is decreased via embedded power-gating techniques and tunable high-performance/low-power devices operation. Memory cells are made smaller by merging the access interface with the storage circuitry. We foresee the advantages deriving from these techniques, by evaluating their impact on the design of SoC for a contemporary telecommunication application. Using a 22-nm vertically-stacked silicon nanowire technology, we estimate a delay and power reduction of 20% and 19% respectively, at a cost of a moderate area overhead of 15%, with respect to a state-of-art FinFET technology.</td>
</tr>
<tr>
<td>09:15</td>
<td>9.1.3</td>
<td>RECONFIGURABLE SILICON NANOWIRE DEVICES AND CIRCUITS: OPPORTUNITIES AND CHALLENGES</td>
<td>Walter Weber(^1), André Heinzig(^1), Jens Trommer(^1), Markus König(^2), Matthias Grube(^1) and Thomas Mikolajick(^1), Namlab gGmbH, DE; (\text{zTechnische Universität Dresden, DE})</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
<td>Reconfigurable fine-grain electronics target an increase in the number of integrated logic functions per chip by enhancing the functionality at the device level and by implementing a compact and technologically simple hardware platform. Here we study a promising realization approach by employing reconfigurable nanowire transistors (RFETs) as the multifunctional building blocks to be integrated therein. RFETs merge the electrical characteristics of unipolar n- and p-type FETs into a single universal device. The switch comprises four terminals, where three of them act as the conventional FET electrodes and the fourth acts as an electric select signal to dynamically program the desired switch type. The transistor consists of two independent charge carrier injection valves as represented by two gated Schottky junctions integrated within an intrinsic silicon nanowire. Radial compressive strain applied to the channel is used as a scalable method to adjust n- and p-FET currents to each other, thereby enabling complementary logic circuits. Simple but relevant examples for the reconfiguration of complete gates will be given, demonstrating the potential of this technology.</td>
</tr>
<tr>
<td>09:35</td>
<td>9.1.4</td>
<td>ADVANCING CMOS WITH CARBON ELECTRONICS</td>
<td>Franz Kreupl, TU Munich, DE</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
<td>A fresh look on carbon-based transistor channel materials like single-walled carbon nanotubes (CNT) and graphene nanoribbons (GNR) in future electronic applications is given. Although theoretical predictions initially promised that GNR (which do have a bandgap) would perform equally well as transistors based on CNTs, experimental evidence for the well-behaved transistor action is missing up to now. Possible reasons for the shortcomings as well as possible solutions to overcome the performance gap will be addressed. In contrast to GNR, short channel CNT field effect transistors (FET) demonstrate in the experimental realization almost ideal transistor characteristics down to very low bias voltages. Therefore, CNT-FETs are clear frontrunners in the search of a future CMOS switch, that will enable further voltage and gate length scaling. Essential features which distinguish CNT-FETs from alternative solution will be discussed and benchmarked. Finally, the gap to industrial water-level scale SWCNT integration will be addressed and strategies for achieving highly aligned carbon nanotube fabrics will be discussed. Without such a high yield wafer-scale integration, SWCNT circuits will be an illusional dream.</td>
</tr>
</tbody>
</table>

9.2 Low-Cost, High-Performance NoCs

**Date:** Thursday, March 27, 2014  
**Time:** 08:30 - 10:00  
**Location / Room:** Konferenz 6

**Chair:**  
Kees Goossens, Eindhoven University, NL; Contact Kees Goossens

**Co-Chair:**  
Luca Ramini, University of Ferrara, IT; Contact Luca Ramini

This session pushes the boundaries of NoC performance optimization while at the same time accounting for implementation constraints. The first paper takes a perspective where express channels are added to the topology, and then smart application mapping is performed. The second paper instead chooses the TDM NoC route to provide guaranteed performance, and significantly optimizes the TDM scheduling process. Finally, the last paper reduces buffer sizes, while also providing elasticity, in a router's virtual channel buffers.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:30</td>
<td>9.2.1</td>
<td>APPLICATION MAPPING FOR EXPRESS CHANNEL-BASED NETWORKS-ON-CHIP</td>
<td>Di Zhu(^1), Lizhong Chen(^1), Slyn Yue(^2) and Massoud Pedram(^1), (^1)Univ. of Southern California, US; (^2)University of Southern California, US</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Abstract</strong></td>
<td>With the emergence of many-core multiprocessor system-on-chips (MPSoCs), the on-chip networks are facing serious challenges in providing fast communication for various tasks and cores. One promising solution shown in recent studies is to add express channels to the network as shortcuts to bypass intermediate routers, thereby reducing packet latency. However, this approach also greatly changes the packet delay estimation and traffic behaviors of the network, both of which have not yet been exploited in existing mapping algorithms. In this paper, we explore the opportunities in optimizing application mapping for express channel-based on-chip networks. Specifically, we derive a new delay model for this type of networks, identify their unique characteristics, and propose an efficient heuristic mapping algorithm that increases the bypassing opportunities by reducing unnecessary turns that would otherwise impose the entire router pipeline delay to packets. Simulation results show that the proposed algorithm can achieve a 2<del>4X reduction in the number of turns and 10</del>26% reduction in the average packet delay.</td>
</tr>
</tbody>
</table>
### 9.3 Hardware Implementations for Data Security

**Date:** Thursday, March 27, 2014  
**Time:** 08:30 - 10:00  
**Location / Room:** Konferenz 1

**Chair:**  
Viktor Fischer, St Etienne, FR, Contact Viktor Fischer

**Co-Chair:**  
Tim Gueneysu, RUB, DE, Contact Tim Gueneysu

Hardware features are used as a trust anchor in many secure systems. This includes design obfuscation techniques, encrypted processing, and biometric systems which are discussed in this session.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>09:00</td>
<td>9.2.2</td>
<td>PARALLEL PROBE BASED DYNAMIC CONNECTION SETUP IN TDM NOCS</td>
<td>Shaoting Liu, Axel Jantsch and Zhonghai Lu, KTH, SE</td>
</tr>
<tr>
<td>09:30</td>
<td>9.2.3</td>
<td>ELASTISTORE: AN ELASTIC BUFFER ARCHITECTURE FOR NETWORK-ON-CHIP ROUTERS</td>
<td>Giorgos Dimitrakopoulos, Ioannis Seitandissi, Anastasios Psaras and Chryssostomos Nicopoulos</td>
</tr>
<tr>
<td>10:00</td>
<td>IP-4-12, 581</td>
<td>DYNAMIC CONSTRUCTION OF CIRCUITS FOR REACTIVE TRAFFIC IN HOMOGENEOUS CMPS</td>
<td>Marta Ortín-Olóñiz, Darío Suárez-Gracia Suárez-Gracia, María Villaroya-Gaudó and Víctor Villals-Yúfera</td>
</tr>
<tr>
<td>10:01</td>
<td>IP-4-13, 646</td>
<td>IMPROVING HAMILTONIAN-BASED ROUTING METHODS FOR ON-CHIP NETWORKS: A TURN MODEL APPROACH</td>
<td>Poona Bahrebar and Dirk Stroobandt, Ghent University, BE</td>
</tr>
<tr>
<td>10:00</td>
<td></td>
<td>End of session</td>
<td></td>
</tr>
</tbody>
</table>

**Coffee Break** in Exhibition Area  
On Tuesday-Thursday the coffee and lunch breaks will be located in the Exhibition Area (Terrace Level).
9.4 Timing challenges in validation

09:00  9.3.2 A MINIMALIST APPROACH TO REMOTE ATTESTATION

Speakers:
Aurélien Francillon, Quan Nguyen, Kasper Rasmussen and Gene Tsudik

1EURCOM, FR; 2University of California, Irvine, US

Abstract
Embedded computing devices increasingly permeate many aspects of modern life: from medical to automotive, from building and factory automation to weapons, from critical infrastructures to home entertainment. Despite their specialized nature as well as limited resources and connectivity, these devices are now becoming an increasingly popular and attractive target for attacks, especially, malware infections. A number of approaches have been suggested to detect and/or mitigate such attacks. They vary greatly in terms of applicability generally and underlying assumptions. However, one common theme is the need for Remote Attestation, a distinct security service that allows a trusted party (verifier) to check the internal state of a remote untrusted embedded device (prover). Many prior methods assume some form of trusted hardware on the prover, which is not a good option for small and low-end embedded devices. To this end, we investigate the feasibility of Remote Attestation without trusted hardware. This paper provides a systematic treatment of Remote Attestation, starting with a precise definition of the desired service and proceeding to its systematic deconstruction into necessary and sufficient properties. Next, these are mapped into a minimal collection of hardware and software components that result in secure Remote Attestation. One distinguishing feature of this line of research is the need to prove (or, at least argue for) architectural minimality – an aspect rarely encountered in security research. This work also provides a promising platform for attaining more advanced security services and guarantees.

09:15  9.3.3 MULTI RESOLUTION TOUCH PANEL WITH BUILT-IN FINGERPRINT SENSING SUPPORT

Speakers:
Pranav Koundinya, Sandhya Theri, Tao Feng, Vanun Prakash, Jiming Bao and Weidong Shi, University of Houston, US

Abstract
In today's technology driven world, it is essential to build secure systems with low faulty behavior. Authentication is one of the primary means to gain access to secure systems. Users need to be authenticated in order to gain access to the services and sensitive information contained within the system. Due to the surge in the number of touch based smart devices, there arises a need for a compatible authentication system. Historically, fingerprints have served in its fullest capacity to establish the uniqueness of an individual's identity. It can be detected using capacitive sensing techniques. In this paper we present a novel unified device using transparent electronics for both fingerprint scan and multi-touch interaction. We discuss a high resolution transparent touch sensitive device and a read out circuit that drives the capacitive sensor array for touch interactions at low resolutions and for fingerprint sensing at higher resolutions. Using circuit simulation and custom Verilog-A model for transparent thin-film transistors, we verified that our design can sense fingerprints in 8.25 ms and detect touches in 0.6ms with an efficient power consumption of 1 mW. The results show that such a device can be realized and can serve as a very efficient means of user authentication. Furthermore, from the usability aspect, the proposed device is essential as it provides user transparent and non intrusive authentication.

09:30  9.3.4 HEROIC: HOMOMORPHICALLY ENCRYPTED ONE INSTRUCTION COMPUTER

Speakers:
Nektarios Georgios Tsoutsos1 and Michail Maniatakos2

1NYU Polytechnic School of Engineering, US; 2NYU Abu Dhabi, AE

Abstract
As cloud computing becomes mainstream, the need to ensure the privacy of the data entrusted to third parties keeps rising. Cloud providers resort to numerous security controls and encryption to thwart potential attackers. Still, since the actual computation inside cloud microprocessors remains unencrypted, the opportunity of leakage is theoretically possible. Therefore, in order to address the challenge of protecting the computation inside the microprocessor, we introduce a novel general purpose architecture for secure data processing, called HEROIC (Homomorphically Encrypted One Instruction Computer). This new design utilizes a single instruction architecture and provides native processing of encrypted data at the architecture level. The security of the solution is assured by a variant of Paillier's homomorphic encryption scheme, used to encrypt both instructions and data. Experimental results using our hardware-cognizant software simulator, indicate an average execution overhead between 5 and 45 times for the encrypted computation (depending on the security parameter), compared to the unencrypted variant, for a 16-bit single instruction architecture.

10:00  9.4.1 EDA TOOLS TRUST EVALUATION THROUGH SECURITY PROPERTY PROOFS

Speaker:
Yer Jin, The University of Central Florida, US

Abstract
The security concerns of EDA tools have long been ignored because IC designers and integrators only focus on their functionality and performance. This lack of trusted EDA tools hampers hardware security researchers' efforts to design trusted integrated circuits. To address this concern, a novel EDA tools trust evaluation framework has been proposed to ensure the trustworthiness of EDA tools through its functional operation, rather than scrutinizing the software code. As a result, the newly proposed framework lowers the evaluation cost and is a better fit for hardware security researchers. To support the EDA tools evaluation framework, a new gate-level information assurance scheme is developed for security property checking on any gate-level netlist. Helped by the gate-level scheme, we expand the territory of proof-carrying based IP protection from RT-level designs to gate-level netlist, so that most of the commercially trading third-party IP cores are under the protection of proof-carrying based security properties. Using a sample AES encryption core, we successfully prove the trustworthiness of Synopsys Design Compiler in generating a synthesized netlist.

10:00  End of session

Coffee Break in Exhibition Area
On Tuesday-Thursday the coffee and lunch breaks will be located in the Exhibition Area (Terrace Level).
Accelerated timing simulation is essential for today's chip designs, whether it is performed at the gate-level or at the system-level. This session provides solutions to address the challenges of timing analysis and timing validation performance across multiple levels of design's abstractions.

### Time  Label  Presentation Title  Authors

08:30  9.4.1  **FAST STA PREDICTION-BASED GATE-LEVEL TIMING SIMULATION**  
Speakers:  
Tanq Bashir Ahmed and Maciej Ciesielski, UMASS Amherst, US  
Abstract  
Traditional dynamic simulation with standard delay format (SDF) back-annotation cannot be reliably performed on large designs. The large size of SDF files makes the event-driven timing simulation extremely slow as it has to process an excessive number of events. In order to accelerate gate-level timing simulation we propose an automated fast prediction-based gate-level timing simulation that combines static timing analysis (STA) at the block level with dynamic timing simulation at the I/O interfaces. We demonstrate that the proposed timing simulation can be done earlier in the design cycle in parallel with synthesis.

09:00  9.4.2  **A CROSS-LEVEL VERIFICATION METHODOLOGY FOR DIGITAL IPS AUGMENTED WITH EMBEDDED TIMING MONITORS**  
Speakers:  
Valerio Guarnieri, Massimo Petricca, Alessandro Sassone, Sara Vincent, Nicola Bombieri, Franco Fummi, Enrico Macii and Massimo Porcino  
Abstract  
Smart systems implement the leading technology advances in the context of embedded devices. Current design methodologies are not suitable to deal with tightly interacting subsystems of different technological domains, namely analog, digital, discrete and power devices, MEMS and power sources. The effects of interaction between components and with the environment must be modeled and simulated at system level to achieve high performance. Focusing on the digital domain, additional design constraints have to be considered as a result of the integration of multi-domain subsystems in a single device. The main digital design challenges, combined with those emerging from the heterogeneous nature of the whole system, directly impact on performance and on propagation delay of the digital component. This paper proposes a design approach to enhance the RTL model of a given digital component for the integration in smart systems, and a methodology to verify the added features at system-level. The design approach consists of augmenting the RTL model through the automatic insertion of delay sensors, which can detect and correct timing failures. The augmented model is abstracted to SystemC TLM and, then, mutants (i.e., code mutations for emulating timing failures) are automatically injected into the model. Experimental results demonstrate the applicability of the proposed design and verification methodology and the effectiveness of the simulation performance.

09:30  9.4.3  **EMPOWERING STUDY OF DELAY BOUND TIGHTNESS WITH SIMULATED Annealing**  
Speakers:  
Xueqian Zhao and Zhonghai Lu, KTH Royal Institute of Technology, SE  
Abstract  
Studying the delay bound tightness typically takes a practical approach by comparing simulated results against analytic results. However, this is often a manual process whereas many simulation parameters have to be configured before the simulations run. This is a tedious and time-consuming process. We propose a technique to automate this process by using a simulated annealing approach. We formulate the problem as an online optimization problem, and embed a simulated annealing algorithm in the simulation environment to guide the search of configuration parameters which give good tightness results. This is a fully automated procedure and thus provide a promising path to automatic design space exploration in similar contexts. Experiment results of an all-to-one communication network with large searching space and complicated constraints illustrate the effectiveness of our method.

10:00  IP4-15  **ANALYSIS AND EVALUATION OF PER-FLOW DELAY BOUND FOR MULTIPLEXING MODELS**  
Speakers:  
Yanchen Long1, Zhonghai Lu2 and Xiaolang Yan3  
1Qafejiang Unversity and KTH Royal Institute of Technology, SE; 2KTH Royal Institute of Technology, SE; 3Qafejiang University, CN  
Abstract  
Multiplexing models are common in resource sharing communication media such as buses, crossbars and networks. While sending packets over a multiplexing node, the packet delay bound can be computed using network calculus models. The tightness of such delay bound remains an open problem. This paper studies the multiplexing models for weighted round robin scheduling with different traffic arrival curves, and analyzes per-flow packet delay bounds with different service properties. We empirically evaluate the tightness of the delay bounds. Our results show the quality of different analysis models, and how influential each parameter is to tightness.

10:00  **End of session**  
Coffee Break in Exhibition Area  
On Tuesday-Thursday the coffee and lunch breaks will be located in the Exhibition Area (Terrace Level).

### 9.5 Hot Topic: Connecting Different Worlds - Technology Abstraction for Reliability-Aware Design and Test

Date:  Thursday, March 27, 2014  
Time:  08:30 - 10:00  
Location / Room:  Konferenz 3  

Organisers:  
Ulf Schlichtmann, Technische Universität München, Ge, Contact Ulf Schlichtmann  
Andreas Herkersdorf, Technische Universität München, Ge, Contact Andreas Herkersdorf  

Chair:  
Nikil Dutt, University of California, Irvine, US, Contact Nikil Dutt  

Co-Chair:  
Mehdi Tahoori, Karlsruhe Institute of Technology, DE, Contact Mehdi Tahoori
The rapid shrinking of device geometries in the nanometer regime requires new technology-aware design methodologies. These must be able to evaluate the resilience of the circuit throughout all System on Chip (SoC) abstraction levels. To successfully guide design decisions at the system level, reliability models, which abstract technology information, are required to identify those parts of the system where additional protection in the form of hardware or software countermeasures is most effective. Interfaces such as the presented Resilience Articulation Point (RAP) or the Reliability Interchange Information Format (RIIF) are required to enable EDA-assisted analysis and propagation of reliability information. The models are discussed from different perspectives, such as design and test.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:30</td>
<td>9.5.1</td>
<td>INTRODUCTION TO RAP (RESILIENCE ARTICULATION POINT)</td>
<td>Andreas Herkersdorf, TU München, DE</td>
</tr>
<tr>
<td>08:45</td>
<td>9.5.2</td>
<td>SYSTEM LEVEL DESIGN USING RAP (RESILIENCE ARTICULATION POINT)</td>
<td>Ulf Schlichtmann, Technische Universität München, DE</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abstract</td>
<td>We will demonstrate how technology characteristics can be included in system-level reliability analysis using the RAP (Resilience Articulation Point) model. The specific example of a two-wheeled robot will be used.</td>
</tr>
<tr>
<td>09:00</td>
<td>9.5.3</td>
<td>CROSS-LAYER RELIABILITY IN THE DESIGN OF AN ERROR RESILIENT COMMUNICATION SYSTEM</td>
<td>Norbert Wehn, University of Kaiserslautern, DE</td>
</tr>
<tr>
<td>09:15</td>
<td>9.5.4</td>
<td>RIIF - TOWARD A STANDARD APPROACH FOR CREATING RELIABILITY MODELS FOR COMPLEX SILICON DEVICES</td>
<td>Adrian Evans, IROC Technologies, FR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abstract</td>
<td>Complex silicon devices are increasingly controlling critical systems where safety and reliability are key concerns. Silicon technology is subject to numerous failure modes which can be broadly classified into soft-error effects (due to natural radiation) and life-time effects (e.g. electro-migration, NBTI, HCI). It is necessary to consider all of these failure modes and how they propagate through the system and produce user-visible effects. There are no consistent tools or methodologies to address this problem. Current ad-hoc approaches are not able to cope with the diversity of technology failure modes, increased design sizes and the complex relationships between consumers and suppliers of electronic components. RIIF (Reliability Information Interchange Format), is an initiative to develop a standard modelling language for specifying the failure mechanisms in silicon devices and systems built using these devices. In this session we give a brief overview of RIIF and present an example that highlights some of the challenges in reliability modeling.</td>
</tr>
<tr>
<td>09:30</td>
<td>9.5.5</td>
<td>TEST PERSPECTIVES</td>
<td>Jacob Abrahám, UT Austin, US</td>
</tr>
<tr>
<td>09:45</td>
<td>9.5.6</td>
<td>INDUSTRIAL PERSPECTIVE</td>
<td>Sani Nassif, IBM, US</td>
</tr>
<tr>
<td>10:00</td>
<td></td>
<td>End of session</td>
<td>Coffee Break in Exhibition Area</td>
</tr>
<tr>
<td></td>
<td></td>
<td>On Tuesday-Thursday the coffee and lunch breaks will be located in the Exhibition Area (Terrace Level).</td>
<td></td>
</tr>
</tbody>
</table>

9.6 Schedulability analysis
Date: Thursday, March 27, 2014
Time: 08:30 - 10:00
Location / Room: Konferenz 4
Chair: Giuseppe Lipari, ENS - Cachan, FR, Contact Giuseppe Lipari
Co-Chair: Benny Akesson, CTU Prague, CZ, Contact Benny Akesson

This session deals with scheduling and schedulability analysis of real-time systems. In particular, it presents different models of scheduling, including fixed and dynamic priority, and real-time calculus.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:30</td>
<td>9.6.1</td>
<td>RATE-ADAPTIVE TASKS: MODEL, ANALYSIS, AND DESIGN ISSUES</td>
<td>Giorgio Buttazzo¹, Enrico Bini and Darren Buttle²</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Speakers</td>
<td>¹Scuola Superiore Sant’Anna, IT; ²Lund University, SE; ³ETAS-PGA/PRM-E, DE</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abstract</td>
<td>In automotive systems, some of the engine control tasks are triggered by specific crankshaft rotation angles and are designed to adapt their functionality based on the angular velocity of the engine. This paper proposes a new task model for specifying such a type of real-time activities and presents an approach for analyzing the system feasibility under deadline scheduling for different scenarios. In particular, a feasibility test is derived for tasks under steady-state conditions (constant speed), as well as in dynamic conditions (constant acceleration). A design method is also discussed to determine the most suitable switching speeds for adapting the functionality of tasks without exceeding a desired utilization. Finally, a number of research directions are highlighted to extend the current results to more complex and realistic scenarios.</td>
</tr>
</tbody>
</table>
The papers in this session present static timing techniques and tools for the analysis and synthesis of logic circuits. The papers take into account new aspects of timing analysis like variability.

**9.7.1 FACILITATING TIMING DEBUG BY LOGIC PATH CORRESPONDENCE**

**Speakers:**
Oshri Adler, Eli Arbel, Ilia Averbouch, Ilan Beer and Inna Gitinевич, IBM, IL

**Abstract**
Synthesis tools for high-performance VLSI designs employ aggressive logic optimization techniques in order to meet physical requirements such as area and cycle time. During these optimizations, the original structure of the design, which is usually written in a hardware description language (HDL), is lost. It is difficult, and often impossible, to relate signals after synthesis to the original signals in the HDL code. Some signals only lose their names while for others there are no equivalent counterparts in the design after synthesis.

Debugging timing problems is based on timing reports which are usually represented in terms of the post-synthesis design. Hence, it is difficult to relate critical paths in the timing reports to the relevant paths in the HDL code when a logic fix is needed. In this paper, we propose a different approach for dealing with the correspondence problem: instead of trying to relate signals we relate paths. Given a critical path in a post-synthesis representation, our method is able to find all corresponding paths in the pre-synthesis (HDL) representation. As a result, locating the parts in the HDL which are relevant to the given timing problem becomes trivial.

During these optimizations, the original structure of the design, which is usually written in a hardware description language (HDL), is lost. It is difficult, and often impossible, to relate signals after synthesis to the original signals in the HDL code. Some signals only lose their names while for others there are no equivalent counterparts in the design after synthesis. Debugging timing problems is based on timing reports which are usually represented in terms of the post-synthesis design. Hence, it is difficult to relate critical paths in the timing reports to the relevant paths in the HDL code when a logic fix is needed. In this paper, we propose a different approach for dealing with the correspondence problem: instead of trying to relate signals we relate paths. Given a critical path in a post-synthesis representation, our method is able to find all corresponding paths in the pre-synthesis (HDL) representation. As a result, locating the parts in the HDL which are relevant to the given timing problem becomes trivial.
9:00 9.7.2 STATISTICAL STATIC TIMING ANALYSIS USING A SKEW-NORMAL CANONICAL DELAY MODEL
Speakers: Vijaykumar M and V Vasudevan, Department of Electrical Engineering Indian Institute of Technology Madras, IN
Abstract: In its simplest form, a parameterized block based statistical static timing analysis (SSTA) is performed by assuming that both gate delays and the arrival times at various nodes are Gaussian random variables. These assumptions are not true in many cases. Quadratic models are used for more accurate analysis, but at the cost of increased computational complexity. In this paper, we propose a model based on skew-normal random variables. It can take into account the skewness in the gate delay distribution as well as the nonlinearity of the MAX operation. We derive analytical expressions for the moments of the MAX operator based on the conditional expectations. The computational complexity of using this model is marginally higher than the linear model based on Clark’s approximations. The results obtained using this model match well with Monte-Carlo simulations.

9:30 9.7.3 LEAKAGE-POWER-AWARE CLOCK PERIOD MINIMIZATION
Speakers: Hua-Hsin Yeh1, Shih-Hsi Huang1 and You-Tyng Nieh2
1Chung Yuan Christian University, TW; 2Industrial Technology Research Institute, TW
Abstract: In the design of nonzero clock skew circuits, an increase of the path delay may improve circuit speed and reduce leakage power. However, the impact of increasing path delay on the trade-off between circuit speed and leakage power has not been well studied. In this paper, we propose a two-step approach for leakage-power-aware clock period minimization. Compared with previous works, our approach has the following two significant contributions. First, our approach is the first leakage-power-aware clock skew scheduling that can guarantee working with the lower bound of the clock period. Second, our approach also is the first work that demonstrates the problem of minimizing the number of extra buffers is a polynomial-time problem. Benchmark data show that our approach achieves the best results in terms of the clock period and the leakage power.

9:45 9.7.4 A DEEP LEARNING METHODOLOGY TO PROLIFERATE GOLDEN SIGNOFF TIMING
Speakers: Seung-Soo Han1, Andrew B. Kahnsg, Siddhartha Nathu and Ashok S. Vydyanathan
1Myongji University, Yongin, KR; 2University of California, San Diego, US
Abstract: Signoff timing analysis remains a critical element in the IC design flow. Multiple signoff corners, libraries, design methodologies, and implementation flows make timing closure very complex at advanced technology nodes. Reported timing slack results directly affect chip area and power by forcing additional buffering or sizing (negative slack), or limiting area and power recovery (positive slack). Design teams often wish to ensure that one tool's timing reports are neither optimistic nor pessimistic with respect to another tool's reports. The resulting “correlation” problem is highly complex because tools contain millions of lines of black box and legacy code, licenses prevent any reverse-engineering of algorithms, and the nature of the problem is seemingly “unbounded” across possible designs, timing paths, and electrical parameters. In this work, we apply a “big-data” mindset to approach the timing correlation problem. We develop a machine-learning-based tool, Golden Timer eXtension (GTX), to correct divergence in flip-flop setup time, cell arc delay, wire delay, and path slack at timing endpoints between times. Our models are developed with datasets of >300K data points for cell, wire, and stage delays and >30K data points for path slack and flip-flop setup time. We propose a methodology to apply GTX to two arbitrary timers, and we evaluate scalability of GTX across multiple designs and foundry technologies/libraries, both with and without signal integrity analysis. Our experimental results show reduction in divergence between timing tools from 139.3ps to 21.1ps (i.e., 6.6×) in setup slack, from 25.6ps to 2.4ps (i.e., 10× reduction) in flip-flop setup time, from 454.4ps to 51.9ps (i.e., 6.7× reduction) in cell delay, from 194.8ps to 17.4ps (i.e., 11.2× reduction) in wire delay, and from 117ps to 23.8ps (4.9× reduction) in stage delay. The average (mean) divergence in timing reports after applying GTX is almost zero. We further demonstrate the incremental application of our method so that models can be adapted to any outlier discrepancies when new designs are taped out in the same technology/library. Last, we demonstrate that GTX can also correlate timing reports between signoff and design implementation tools.

10:00 IP-17, 759 AGING-AWARE STANDARD CELL LIBRARY DESIGN
Speakers: Saman Kiamehr1, Farshad Ftouzi2, Mojtaba Ebrahimi1 and Mehdi Tahoori2
1Karlsruhe Institute of Technology (KIT), DE; 2Karlsruhe Institute of Technology, DE
Abstract: Transistor aging, mostly due to Bias Temperature Instability (BTI), is one of the major unreliability sources at nano-scale technology nodes. BTI causes the circuit delay to increase and eventually leads to a decrease in the circuit lifetime. Typically, standard cells in the library are optimized according to the design time delay, however, due to the asymmetric effect of BTI, the rise and fall delays might become significantly imbalanced over the lifetime. In this paper, the BTI effect is mitigated by balancing the rise and fall delays of the standard cells at the excepted lifetime. We find an optimal trade-off between the increase in the size of the library and the lifetime improvement (timing margin reduction) by non-uniform extension of the library cells for various ranges of the input signal probabilities. The simulation results reveal that our technique can prolong the circuit lifetime around 150% with a negligible area overhead.

10:01 IP-18, 279 PASS-XNOR LOGIC: A NEW LOGIC STYLE FOR P-N JUNCTION BASED GRAPHITE CIRCUITS
Speakers: Valerio Tenace, Andrea Calimera, Enrico Macii and Massimo Poncino, Politecnico di Torino, IT
Abstract: In this work we introduce a new logic style for p-n junctions based digital-graphene circuits: the pass-XNOR logic style. The latter enables the realization of compact, energy efficient circuits that better exploit the characteristics of graphene. We first show how a single p-n junction can be conceived as a transmission gate with embedded logic functionality, the XNOR Boolean operator. Second, we propose a smart integration strategy in which series/parallel connections of pass-XNOR gates allow to implement AND/OR logical conjunctions, and, therefore, all possible truth tables. Experimental results conducted on a set of representative logic functions show the superior of pass-XNOR logic circuits w.r.t. standard CMOS circuits and graphene circuits that use p-n junctions in a complementary-like structure.

10:02 IP-19, 365 MIXED ALLOCATION OF Adjustable DELAY BUFFERS COMBINED WITH BUFFER SIZING IN CLOCK TREE SYNTHESIS OF MULTIPLE POWER MODE DESIGNS
Speakers: Ktae Park, Genuho Kim and Taewhan Kim, Seoul National University, KR
Abstract: Recently, many works have shown that adjustable delay buffer (ADB) whose delay is adjustable dynamically can effectively solve the clock skew variation problem in the designs with multiple power modes. However, all the previous works of ADB allocation inherently entail two critical limitations, which are the adjusted delays by ADB are always increments and the low cost buffer sizing has never been or not been primarily taken into account. To demonstrate how much overcoming the two limitations is effective in resolving the clock skew constraint, we characterize the two types of ADBs called CADB (capacitor-based ADB) and IADB (inverter-based ADB) and show that the adjusted delays by IADB can be decremented, and show that the clock skew variation in some clock trees of multiple power modes can be resolved by applying buffer sizing together with using only a small number of IADBs and CADBs.
9.8 Embedded Tutorial: Memcomputing: the Cape of Good Hope

Date: Thursday, March 27, 2014
Time: 08:30 - 10:00
Location / Room: Exhibition Theatre

Organisers:
Yiyu Shi, Missouri University of Science & Technology, US, Contact Yiyu Shi
Hung-Ming Chen, National Chiao Tung University, Taiwan, Ta, Contact Hung-Ming Chen

Chair:
Tsung-Yi Ho, CSIE, NCKU, TW, Contact Tsung-Yi Ho

Co-Chair:
Hung-Ming Chen, National Chiao Tung University, TW, Contact Hung-Ming Chen

Energy efficiency has emerged as a major barrier to performance scalability for modern processors. On the other hand, significant breakthroughs have been achieved in memory technologies recently. As such, the fascinating idea of memcomputing (i.e., use memory for computation purposes) has drawn wide attention from both academia and industry as an effective remedy. Compared with conventional logic computing, memory array provides large set of parallel resources with high bandwidth, which can be configured to perform computing in spatial/temporal manner leading to dramatic reduction in processor-memory traffic. Moreover, memory computing brings the computing engine close to the data, thus drastically minimizing the von Neumann bottleneck. Finally, it exploits the advances in memory technologies and integration approaches (e.g. 3D integration) to achieve better technology scalability. This special session offers a broad-spectrum retreat (devices, processes and systems) on this hot topic to the general CAD community, hoping to inspire more contributions from the design automation perspective.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>08:30</td>
<td>9.8.1</td>
<td>MEMCOMPUTING: A BRAIN-INSPIRED COMPUTING PARADIGM</td>
<td>Yury Pershin, University of South Carolina, US</td>
</tr>
<tr>
<td>09:00</td>
<td>9.8.2</td>
<td>MSIM: A GENERAL CYCLE ACCURATE SIMULATION PLATFORM FOR MEMCOMPUTING STUDIES</td>
<td>Chun Zhang1, Peng Deng2, Hui Geng1, Jianming Liu1, Qi Zhu2, Jinjun Xiong1 and Yiyu Shi2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abstract</td>
<td>The lack of accurate yet open to public simulation infrastructure has puzzled researchers in the memcomputing area for sometime. In this paper, we propose for the first time a full tool chain called MSim that supports the cycle-accurate microarchitecture level simulation for memcomputing studies. With MSim, the performance gains of utilizing memcomputing for arbitrary applications on user-configurable computer system architectures can be evaluated in high accuracy. In addition, MSim provides flexible interfaces with pervasive object-oriented design, which makes it well suited as a good base platform for researchers to explore new memcomputing technologies.</td>
</tr>
<tr>
<td>09:30</td>
<td>9.8.3</td>
<td>ENERGY-EFFICIENT HARDWARE ACCELERATION THROUGH COMPUTING IN THE MEMORY</td>
<td>Somnath Paul1, Robert Karam2, Swarup Bhunia2 and Ruchir Puri1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Abstract</td>
<td>Energy-efficiency has emerged as a major barrier to performance scalability for modern processors. We note that significant part of processor's energy requirement is contributed by processor-memory communication. To address the energy issue in processors, we propose a novel hardware accelerator framework that transforms high-density memory array into a configurable computing resource to accelerate variety of tasks - both compute- and data-intensive. It exploits the block-based architecture of nanoscale memory to create a spatially connected array of lightweight processors, each of which uses a memory block as its local memory. The proposed framework provides some unique advantages for hardware acceleration compared to conventional accelerators: 1) memory array provides large set of parallel resources with high bandwidth, which can be configured to perform computing in spatio-temporal manner leading to dramatic reduction in processor-memory traffic; 2) it brings the computing engine close to the data, thus drastically minimizing the von Neumann bottleneck; 3) finally, it exploits the advances in memory technologies and integration approaches e.g. 3D integration to achieve better technology scalability compared to alternative reconfigurable accelerator platforms. Simulation results for several data-intensive applications show that the proposed computing approach provides significant improvement in energy-efficiency compared to software while achieving significantly lower hardware overhead.</td>
</tr>
</tbody>
</table>

1 Intel Corporation, US; 2 Case Western Reserve University, US; 3 IBM Watson Research Center, US

IP4 Interactive Presentations

Date: Thursday, March 27, 2014
Time: 10:00 - 10:30
Location / Room: Conference Level, foyer

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>
Abstract
Laser attacks, especially on circuits manufactured with recent deep submicron semiconductor technologies, pose a threat to secure integrated circuits due to the multiplicity of errors induced by a single attack. An efficient way to neutralize such effects is the design of appropriate countermeasures, according to the circuit implementation and characteristics. Therefore tools which allow the early evaluation of security implementations are necessary. Our efforts involve the development of an RTL fault injection approach more representative of laser attacks than random multi-bit fault injections and the utilization and evolution of state of the art emulation techniques to reduce the duration of the fault injection campaigns. This will ultimately lead to the design and validation of new countermeasures against laser attacks, on ASICs implementing cryptographic algorithms.

IP4-2
ENERGY EFFICIENT DATA FLOW TRANSFORMATION FOR GIVEN'S ROTATION BASED QR DECOMPOSITION
Speakers:
Namita Sharma¹, Preeti Ranjan Panda¹, Min Li, Prashant Agrawal² and Francky Cattuto²
1Indian Institute of Technology Delhi, IN; ²IMEC, BE

Abstract
QR Decomposition (QRD) is a typical matrix decomposition algorithm that shares many common features with other algorithms such as LU and Cholesky decomposition. The principle can be realized in a large number of valid processing sequences that differ significantly in the number of memory accesses and computations, and hence, the overall implementation energy. With modern low power embedded processors evolving towards register files with wide memory interfaces and vector functional units (FUs), the data flow in matrix decomposition algorithms needs to be carefully devised to achieve energy efficient implementation. In this paper, we present an efficient data flow transformation strategy for the Given's Rotation based QRD that optimizes data memory accesses. We also explore different possible implementations for QRD of multiple matrices using the SIMD feature of the processor. With the proposed data flow transformation, a reduction of up to 36% is achieved in the overall energy over conventional QRD sequences.

IP4-3
MODE-CONTROLLED DATAFLOW BASED MODELING & ANALYSIS OF A 4G-LTE RECEIVER
Speakers:
Harshkesh Salunkhe¹, Orlando Moreira¹ and Kees van Berkel¹
¹PhD Candidate, NL; ²Principal DSP Systems Engineer, NL; ³Prof. Dr., NL

Abstract
Today’s smartphones and tablets contain multiple cellular modems to support 2G/3G/4G standards, including Long Term Evolution (LTE). They run on complex multi-processor hardware platforms and have to meet hard real-time constraints. Dataflow modeling can be used to design an LTE receiver. Static dataflow allows a rich set of analysis techniques, but is too restrictive to model the dynamic behavior in many realistic applications, including LTE receivers. Dynamic dataflow allows modeling of many realistic applications, but does not support rigorous temporal analysis. Mode-Controlled Dataflow (MCDF) is a restricted form of dynamic dataflow, and allows the same analysis techniques as static dataflow, in principle. We prove that MCDF is sufficiently expressive to handle the dynamic behavior of a realistic LTE receiver, by systematically and stepwise developing a complete MCDF model for an LTE receiver.

IP4-4
MODEL-BASED ACTOR MULTIPLEXING WITH APPLICATION TO COMPLEX COMMUNICATION PROTOCOLS
Speakers:
Christian Zebelœvin¹, Christian Haubelt², Joachim Falk², Tobias Schwarze² and Jürgen Teich²
¹University of Rostock, DE; ²University of Erlangen-Nuremberg, DE

Abstract
We propose a dynamic scheduling approach for the concurrent execution of logical actor instances on a single synthesized actor instance. Based on a formal dataflow model of computation, the proposed approach can be applied to a wide range of applications in a model-based design flow. As case-study, we evaluate a bus-cycle-accurate SystemC RTL model based on an InfiniBand network adapter in a PCI Express system.

IP4-5
A NOVEL MODEL FOR SYSTEM-LEVEL DECISION MAKING WITH COMBINED ASP AND SMT SOLVING
Speakers:
Alexander Biewer¹, Jens Glädis¹ and Christian Haubelt²
¹Robert Bosch GmbH, DE; ²University of Rostock, DE

Abstract
In this paper, we present a novel model enabling system-level decision making for time-triggered many-core architectures in automotive systems. The proposed application model includes shared data entities that need to be bound to memories during decision making. As a key enabler to our approach, we explicitly separate computation and shared memory communication over a network-on-chip (NoC). To deal with contention on a NoC, we model the necessary basis to implement a time-triggered schedule that guarantees freedom of interference. We compute fundamental design decisions, namely (a) spatial binding, (b) multi-hop routing, and (c) time-triggered scheduling, by a novel coupling of answer set programming (ASP) with satisfiability modulo theories (SMT) solvers. First results of an automotive case study demonstrate the applicability of our method for complex real-world applications.

IP4-6
DESPERATE: SPEEDING-UP DESIGN SPACE EXPLORATION BY USING PREDICTIVE SIMULATION SCHEDULING
Speakers:
Giovanni Mariani, Gianluca Palermo, Vittorio Zaccaria and Cristina Silvano, Politecnico di Milano, IT

Abstract
Design Space Exploration (DSE) is the problem to find the best architecture configuration in a platform based design problem. To accurately evaluate a configuration, computational expensive simulations are required. A common approach to reduce DSE execution time is to use analytic performance prediction models to approximate some of the required simulations, thus to prune the design space by removing bad configuration candidates. In this paper we will demonstrate that state of the art analytic techniques to speedup the DSE process are not capable to fully exploit the potentialities of a parallel simulation environment. We will demonstrate that, when different simulations can be run in parallel, predicting simulation time to better schedule the simulations on the parallel simulation environment is a more profitable approach with a speedup of more than 2x when compared to state of the art approaches.
The functionality of embedded systems is ever increasing. This has led to mixed time-criticality systems, where applications with a variety of real-time requirements co-exist on the same platform and share resources. Due to the increasing size of interconnects, verifying the real-time requirements of such systems is generally non-trivial. In this paper, we present the CoMik microkernel that provides temporally predictable and composable processor virtualisation. CoMik's virtual processors are cycle-accurately composable, i.e., their timing cannot affect the timing of co-existing virtual processors by even a single cycle. Real-time applications executing on dedicated virtual processors can therefore be verified and executed in isolation, simplifying the verification of mixed time-critical systems. We demonstrate these properties through experimentation on an FPGA prototyped hardware platform.

### IP4-7 COMIK: A PREDICTABLE AND CYCLE-ACCURATELY COMPOSABLE REAL-TIME MICROKERNEL

**Speakers:**
Andrew Nelson, 1, Ashkan Beyranvand Nejadi, 2, Anca Molnos, 1, Martijn Koedama and Kees Goossens 1

**TU Delft, NL; 2IEA Leu, FR; 3TU Eindhoven, NL**

**Abstract**

The functionality of embedded systems is ever increasing. This has led to mixed time-criticality systems, where applications with a variety of real-time requirements co-exist on the same platform and share resources. Due to the increasing size of interconnects, verifying the real-time requirements of such systems is generally non-trivial. In this paper, we present the CoMik microkernel that provides temporally predictable and composable processor virtualisation. CoMik's virtual processors are cycle-accurately composable, i.e., their timing cannot affect the timing of co-existing virtual processors by even a single cycle. Real-time applications executing on dedicated virtual processors can therefore be verified and executed in isolation, simplifying the verification of mixed time-critical systems. We demonstrate these properties through experimentation on an FPGA prototyped hardware platform.

### IP4-8 UTILIZATION-AWARE LOAD BALANCING FOR THE ENERGY EFFICIENT OPERATION ON THE BIG.LITTLE PROCESSOR

**Speakers:**
Myungsun Kim, 1, Kiboom Kim, 2, James Geraci and Seongsoo Hongs 1

1Samsung Electronics, KR; 2SAMsUNG Electronics, KR; 3Seoul National University, KR

**Abstract**

ARM's big.LITTLE architecture introduces the opportunity to optimize power consumption by selecting the core type most suitable for a level of processing demand. To take advantage of this new axis of optimization, we introduce the processor utilization factor into the Linux kernel's load balancing algorithm after carefully analyzing the power management mechanism of the big.LITTLE processor's port of Linux and deriving its state diagram representation. Our mechanism improves the Linux kernel's ability to assign tasks to cores in an energy efficient manner without having to make it directly aware of the available core types. Our experiments with a real test bed show that our algorithm improves energy consumption over the standard Linux scheduler up to 11.35% with almost no corresponding reduction in performance.

### IP4-9 HEVC-DTM: APPLICATION-DRIVEN DYNAMIC THERMAL MANAGEMENT FOR HIGH EFFICIENCY VIDEO CODING

**Speakers:**
Daniel Palomin1, Muhammed Shafique, Hussam Armeesch, Altamir Susín and Jörg Henkelz 1

1Karlsruhe Institute of Technology (KIT), BR; 2Kaisruhe Institute of Technology (KIT), BR; 3Federal University of Rio Grande do Sul, BR

**Abstract**

This paper presents an application-driven algorithm for Dynamic Thermal Management (DTM) for the High Efficiency Video Coding (HEVC). For efficient design of such a DTM policy, we perform an offline thermal analysis of an HEVC encoder and demonstrate the impact of different video sequences and different coding configurations on the processor temperature. Our thermal analysis is leveraged to develop an efficient application-driven DTM policy that performs temperature-aware coding along with an application-driven control of DTM knobs (e.g., frequency scaling) in order to meet the temperature constraints while still providing high video quality (i.e., PSNR loss < 0.01dB). For accurate thermal analysis and evaluation, we deploy an infrared camera-based thermal measurement setup that, on the contrary to state-of-the-art setups, does not require adding any extra layer on top of the measured chip, thus allowing the camera to accurately capture the infrared emissions from the die.

### IP4-10 IMPROVING EFFICIENCY OF EXTENSIBLE PROCESSEORS BY USING APPROXIMATE CUSTOM INSTRUCTIONS

**Speakers:**
Mehdi Kamal1, Amin Ghasem Azar1, Ali Atzali-Kusha1 and Massoud Pedram1

1University of Tehran, IR; 2University of Southern California, US

**Abstract**

In this paper, we propose to move the extensible processor design flow to the approximate computing domain to gain more speedup. In this domain, the instruction set architecture (ISA) design flow selects both exact and approximate custom instructions (CIs). The proposed approach could be used for the applications where imprecise results may be tolerated. In the CI identification phase of the flow, the CIs which do not satisfy the maximum propagation delay but can provide approximate results also may be included in the CI candidate set. Next, in the selection phase, we propose a merit function which selects CIs with higher cycle savings and small error rates. The efficacy of the proposed approximate design flow is investigated using the case studies of the discrete cosine transform (DCT) and inverse DCT (iDCT) of the MPEG2 application. Also, the impact of the process variation on the impreciseness of the results is investigated.

### IP4-11 PROBABLISTIC STANDARD CELL MODELING CONSIDERING NON-GAUSSIAN PARAMETERS AND CORRELATIONS

**Speakers:**
André Lange1, Christoph Sohmann1, Roland Jancke1, Joachim Haase1, Ingolf Lorenz and Ulf Schlichtmann 1

1Fraunhofer Institute for Integrated Circuits (IIS), Design Automation Division (EAS), DE; 2GLOBALFOUNDRIES Inc., DE; 3Technische Universität München, DE

**Abstract**

In this paper, we present a probabilistic approach for variability modeling as an alternative: model parameters are treated as multi-dimensional random variables. Such a fully multivariate statistical description formally accounts for correlations and non-Gaussian random components. Statistical characterization and model application are introduced for standard cells and gate level digital circuits. Example analyses of circuitry in a 28 nm industrial technology illustrate the capabilities of our modeling approach.

### IP4-12 DYNAMIC CONSTRUCTION OF CIRCUITS FOR REACTIVE TRAFFIC IN HOMOGENEOUS CMPS

**Speakers:**
Marta Ortín-Obón1, Dario Suárez-Gracia Suárez-Graciá, María Villamay-Valdés, Cruz Izus and Víctor Víñals-Yúfera1

1University of Zaragoza, ES; 2University of Adelaide, AU

**Abstract**

Networks on Chip (NoC) have a large impact on system performance, area and energy. Considering the characteristics of the memory subsystem while designing the NoC helps identify improvement opportunities and build more efficient designs. Leveraging the frequent request-reply pattern, our proposal dynamically builds the reply path in advance, is able to share circuits between messages, and even removes some implicit replies, significantly reducing NoC latency. A careful implementation of this circuit reservation mechanism achieves an average 17% reduction in router energy consumption, 8% smaller router area and a 2% system performance increase, compared with its baseline counterpart.
Abstract

The overall performance of Multi-Processor System-on-Chip (MPSoC) platforms depends highly on the efficient communication among their cores in the Network-on-Chip (NoC). Routing algorithms are responsible for the on-chip communication and traffic distribution through the network. Hence, designing efficient and high-performance routing algorithms is of significant importance. In this paper, a deadlock-free and highly adaptive path-based routing method is proposed without using virtual channels. This method strives to exploit the maximum number of minimal paths between any source and destination pair. The simulation results in terms of performance and power consumption demonstrate that the proposed method significantly outperforms the other adaptive and non-adaptive schemes. This efficiency is achieved by reducing the number of hotspots and smoothly distributing the traffic across the network.

** Speakers **
Pooja Bahrebar and Dirk Stroobandt, Ghent University, BE

** Authors **
Zhejiang University and KTH Royal Institute of Technology, SE; KTH Royal Institute of Technology, SE; Zhejiang University, CN

---

** Speakers **
Yi Jin, The University of Central Florida, US

** Abstract **
The security concerns of EDA tools have long been ignored because IC designers and integrators only focus on their functionality and performance. This lack of trusted EDA tools hampers hardware security researchers’ efforts to design trusted integrated circuits. To address this concern, a novel EDA tools trust evaluation framework has been proposed to ensure the trustworthiness of EDA tools through its functional operation, rather than scrutinizing the software code. As a result, the newly proposed framework lowers the evaluation cost and is a better fit for hardware security researchers. To support the EDA tools evaluation framework, a new gate-level information assurance scheme is developed for security property checking on any gate-level netlist. Helped by the gate-level scheme, we expand the territory of proof-carrying based IP protection from RT-level designs to gate-level netlist, so that most of the commercially trading third-party IP cores are under the protection of proof-carrying based security properties. Using a sample AES encryption core, we successfully prove the trustworthiness of Synopsys Design Compiler in generating a synthesized netlist.

** Authors **
Valerio Tenace, Andrea Calimera, Enrico Macii and Massimo Poncino, Politecnico di Torino, IT

---

** Speakers **
Yanchen Long¹, Zhonghai Lu² and Xiaolong Yan³

** Abstract **
Multiplexing models are common in resource sharing communication media such as buses, crossbars and networks. While sending packets over a multiplexing node, the packet delay bound can be computed using network calculus models. The tightness of such delay bound remains an open problem. This paper studies the multiplexing models for weighted round robin scheduling with different traffic arrival curves, and analyzes per-flow packet delay bounds with different service properties. We empirically evaluate the tightness of the delay bounds. Our results show the quality of different analysis models, and how influential each parameter is to tightness.

** Authors **
Saman Kiamehr

---

** Speakers **
Transistor aging, mostly due to Bias Temperature Instability (BTI), is one of the major unreliability sources at nano-scale technology nodes. BTI causes the circuit delay to increase and eventually leads to a decrease in the circuit lifetime. Typically, standard cells in the library are optimized according to the design time delay, however, due to the asymmetric effect of BTI, the rise and fall delays might become significantly imbalanced over the lifetime. In this paper, the BTI effect is mitigated by balancing the rise and fall delays of the standard cells at the expected lifetime. We find an optimal trade-off between the increase in the size of the library and the lifetime improvement (marginal reduction) by non-uniform extension of the library cells for various ranges of the input signal probabilities. The simulation results reveal that our technique can prolong the circuit lifetime by around 150% with a negligible area overhead.

** Authors **
Saman Kiamehr, Farshad Firouzi, Mojtaba Ebrahimizadeh and Mehdi Tahoori

---

** Speakers **
Valerio Tenace, Andrea Calimera, Enrico Macii and Massimo Poncino, Politecnico di Torino, IT

** Abstract **
In this work we introduce a new logic style for p-n junctions based digital graphene circuits: the pass-XNOR logic style. The latter enables the realization of compact, energy efficient circuits that better exploit the characteristics of graphene. We first show how a single p-n junction can be conceived as a pass-XNOR gate, i.e., a transmission gate with embedded logic functionality, the XNOR Boolean operator. Secondly, we propose a smart integration strategy in which series/parallel connections of pass-XNOR gates allow to implement AND/OR logical conjunctions, and, therefore, all possible truth tables. Experimental results conducted on a set of representative logic functions show the superior of pass-XNOR logic circuits w.r.t. standard CMOS circuits and graphene circuits that use p-n junctions in a complementary-like structure.

** Authors **
Saman Kiamehr

---

** Speakers **
Kita Park, Geunho Kim and Taewhan Kim, Seoul National University, KR

** Abstract **
Recently, many works have shown that adjustable delay buffer (ADB) whose delay is adjustable dynamically can effectively solve the clock skew variation problem in the designs with multiple power modes. However, all the previous works of ADB allocation inherently entail two critical limitations, which are the adjusted delays by ADB are always increments and the low cost buffer sizing has never been or not been primarily taken into account. To demonstrate how much overcoming the two limitations is effective in resolving the clock skew constraint, we characterize the two types of ADBs called CADB (capacitor based ADB) and IADB (inverter based ADB) and show that the adjusted delays by ADB can be decremented, and show that the clock skew violation in some clock trees of multiple power modes can be resolved by applying buffer sizing together with using only a small number of IADBs and CADBs.

** Authors **
Saman Kiamehr
SOC VERIFICATION: AUTOMATED FUNCTIONAL VERIFICATION OF SYSTEMS-ON-CHIP

Authors:
Zdenek Prikryl, Marcela Simkova and Karel Masaril, Faculty of Information Technology, Bmo University of Technology, CZ

Abstract
An increase of the complexity of systems-on-chip (SoC) induces an increase of the complexity of their verification as well. The reason is that we must verify not only the functions of separate logic blocks, but we need to check their interconnections, timing and functional collaboration as well. Therefore, there is still a great demand for verification tools, which are time-effective, fast and as automated as possible. Exactly these issues we target in our solution. You are welcome to see the live demonstration at our booth!

More information ...

AN AUTOMATED DESIGN FLOW FOR FAST PROTOTYPING OF SIMULINK MODELS INTO MPSOC

Authors:
Francesco Robino and Johnny Öberg, Royal Institute of Technology, SE

Abstract
Simulink is a modeling environment suitable to model embedded systems at system-level. However there is no standard to rapidly prototype Simulink models onto modern multiprocessor system-on-chip (MPSoC). In this demonstration we show how our NoC System Generator tool can be used as part of an automated platform-based design flow to synthesize a Simulink model to a network-on-chip based MPSoC implementation on FPGA. The performance of the generated prototype scales with the number of processors.

More information ...

LARA: THE LARA COMPILER SUITE

Authors:
Joao Bispo, Pedro Pinto, Ricardo Nobre, Tiago Canhalho and Joao Cardoso, Universidade do Porto, PT

Abstract
LARA is an aspect-oriented programming (AOP) language which allows the description of sophisticated code instrumentation schemes, advanced mapping strategies including conditional decisions, based on hardware/software resources, and of sophisticated sequences of compiler transformations. Furthermore, LARA provides mechanisms for controlling all elements of a toolchain in a consistent and systematic way, using a unified programming interface. We present three compiler tools developed around the LARA technology.

More information ...

SECURE CLOUD-BASED WORKFLOW-AS-A-SERVICE (WFaaS) ENVIRONMENT WITH ROLE-BASED-ACCESS-CONTROL (RBAC) FOR SOC DESIGN

Authors:
Sai Manoj P D1, Sai Manoj P. D.1, Hao Yu2 and Joseph Lee2
1Nanyang Technological University, SG; 2Silicon Cloud International, US

Abstract
The SoC design process requires multiple EDA tools, custom IP's, and technology design kit from multiple providers. The design environment needs to be secure and collaborative. These requirements can be realized by using an integrated cloud based Workflow-as-a-Service (WFaaS) design environment. We demonstrate a cloud-based design environment for a SoC design with multiple CPU cores and analog IO's. This design environment uses an innovative Role-Based-Access-Control user security model where designers interact through a web portal dashboard to perform the design workflows.

More information ...

RTL+: DESIGN ENVIRONMENT: WALK BEFORE YOU RUN.

Authors:
Somayeh Sadeghi-Kohan, Behnaz Pourmohseni, Amir Reza Nekooei, Hanieh Hashemi, Hamed Najafi Haghi and Zainalabedin Navabi, University of Tehran, IR

Abstract
To enable development of high level designs with hardware correspondence, synthesizability must be satisfied in a top-down manner. Thus in this work, instead of using TLM-2.0 which is not established for synthesis, we will start with a level above RT level, "RTL+". RTL+ is basically using TLM-1.0 channels and includes abstract communications and handshakings that are mainly hidden from the designer. We develop a package of SystemC channels with hardware correspondence (synthesizable HDL) for the communication between various cores (with simple interfaces) and standard buses.

More information ...

ENERGY-MODULATED COMPUTING

Authors:
Maxim Rykunov, Reza Ramezani, Abdulhak Baz, Xuelu Zhang, Delong Shang, Andrey Mokhov, Danil Sokolov, Fei Xia and Alex Yakovlev, Newcastle University, GB

Abstract
This demo will illustrate the principle of energy-modulated computing according to which the flow of energy entering a computing system determines its computational flow. This principle will be fundamental for building future autonomous systems, such as those powered by energy harvesting sources and aimed for survival in power-deficient conditions. The demo includes a set of experimental circuits (with three VLSI chips and PCBs) to work in variable power supply conditions and software tools for digital and analogue co-design (Workcraft, Petity, MPSAT).

More information ...
System-on-Chip (SOC) design gets increasingly complex, as a growing number of applications are integrated in such systems. These applications have mixed time-criticality, i.e., some have firm, some soft, and others non-real-time requirements. Executing such a mix of applications on a SOC poses several challenges. First, to reduce cost, platform resources, e.g., processors, interconnect, memories, are shared between applications. However, sharing causes interference between applications, making their behaviors inter-dependent. This results in two problems for SOC design and verification: 1) accurate system-level simulation and several approaches to formal verification are infeasible, because of the explosion in the number of possible combinations of applications, inputs, and resource states and 2) verification becomes a circular process that must be repeated if an application is added, removed, or modified, making integration and verification dominant parts of SOC development, in terms of time and money. The CompSOC platform addresses these problems by executing each application on an independent virtual execution platform (VEP). The VEPs are composable, i.e., cannot affect each other’s behaviors. In the temporal domain an applications actual execution never varies by even a single clock cycle. Similarly, the energy and power behaviors of applications are also composable. As a result, applications can be designed, developed, verified, and executed in isolation. The VEPs are also predictable, meaning that all interference is bounded. This makes them virtualized also in terms of performance bounds, which enables firm real-time applications to be verified using formal performance analysis frameworks. The CompSOC platform uses the CoMK microkernel to implement virtual processors on each processor time through temporal partitioning. Each application can use its own operating system (e.g., CompSOC, µOS-II) and model of computation (e.g., CSDF, KPN, TT) in its VEP, to suit its level of time critically. As more applications are integrated on a single SOC, the need arises for more dynamic behavior. The system should be able to start, modify and stop applications at run time without affecting running applications. For this purpose the CompSOC platform has been extended with a predictable and composable resource management framework. It manages application bundles that contain 1) an application in the form of executables (ELFs on multiple processors), and also 2) the specifications of the (one or more) particular VEPs that the application executes in, consisting of virtual processors, NOC connections, virtualised memories, etc. At run time, the resource management framework can dynamically load and start application bundles by creating a VEP and then loading, booting, and executing an application within it. VEPs can also be modified, stopped, and deleted at run time. Our University Booth will present a virtual execution platform and application bundle concepts using an interactive demonstrator. It will show that the CompSOC has been extended with dynamic functionality, without sacrificing its key strengths: composability and predictability. We will demonstrate this through the use of the resource management framework and application bundles, showing that we can create, modify and delete virtual execution platforms running a mixed time-critical application dynamically at run-time.

More information ...

UNISON: ASSEMBLY CODE GENERATION USING CONSTRAINT PROGRAMMING

Authors: Roberto Castera, Lozanno, Gabriel Hor Blinda, Mats Carlsson and Christian Schultz

Abstract: We demonstrate Unison - a simple, flexible and potentially optimal code generator that solves interdependent code generation tasks together using constraint programming as a modern combinatorial optimization method. We show how Unison takes into account the task interdependencies and their combinatorial nature to improve the speed of the code generated by LLVM (a state-of-the-art compiler) for Hexagon (a digital signal processor ubiquitous in modern mobile platforms).

More information ...

Memory devices and technologies have undergone huge transformations in recent years and many industrially viable replacements to conventional technologies are on the brink of entering the market. The first paper in this session gives an overview of alternative memory technologies and how each can contribute or disrupt accepted memory hierarchies. The quest for a universal memory device is still underway, and the other papers in this session focus on various approaches for future memory devices. The second paper examines phase change memories, while magnetic memories are discussed in the third paper, both in terms of standard memory applications but also in terms of how they can improve logic performance. Resistive memories are the topic of the fourth paper, where new applications are considered - in FPGAs, NoCs and crossbars. The fifth paper in this session looks at low-cost memory with a printable manufacturing approach, leading to other applications and market segments.

Time  Label  Presentation Title  Authors

12:00  End of session
12:30  Lunch Break in Exhibition Area  Sandwich lunch

10.1 SPECIAL DAY Hot Topic: Memories today and tomorrow
Date: Thursday, March 27, 2014
Time: 11:00 - 12:30
Location / Room: Saal 1

Organisers: Thomas Mikolajick, NamLab gGmbH, DE, Contact Thomas Mikolajick
Ian O’Connor, Lyon Institute of Nanotechnology, FR, Contact Ian O’Connor

Chair: Tahoori Mehdi, Karlsruhe Institute of Technology, DE, Contact Mehdi Tahoori

Co-Chair: Thomas Mikolajick, NamLab gGmbH, DE, Contact Thomas Mikolajick

Memory devices and technologies have undergone huge transformations in recent years and many industrially viable replacements to conventional technologies are on the brink of entering the market. The first paper in this session gives an overview of alternative memory technologies and how each can contribute or disrupt accepted memory hierarchies. The quest for a universal memory device is still underway, and the other papers in this session focus on various approaches for future memory devices. The second paper examines phase change memories, while magnetic memories are discussed in the third paper, both in terms of standard memory applications but also in terms of how they can improve logic performance. Resistive memories are the topic of the fourth paper, where new applications are considered - in FPGAs, NoCs and crossbars. The fifth paper in this session looks at low-cost memory with a printable manufacturing approach, leading to other applications and market segments.
This session comprises three papers devoted to studying different aspects of wireless NoC design and optimization. The first paper focuses on energy efficiency, by effectively tuning the transmission power of on-chip antennas. The second paper compares the performance and power of different routing algorithms for wireless NoCs, while the third paper explores the adoption of wireless NoCs in 3D chip designs.
Several emerging techniques have been recently proposed for alleviating the communication latency and the energy consumption issues in multi/multi-core architectures. One of such emerging communication techniques, namely, WiNoC replaces the traditional wired links with the use of wireless medium. Unfortunately, the energy consumed by the RF transceiver (i.e., the main building block of a WiNoC), and in particular by its transmitter, accounts for a significant fraction of the overall communication energy. In this paper we propose a runtime tunable transmitting power technique for improving the energy efficiency of the transceiver in wireless NoC architectures. The basic idea is tuning the transmitting power based on the location of the recipient of the current communication. The integration of the proposed technique into two known WiNoC architectures, namely, iWiSe64 and McWiNoC resulted in an energy reduction of 43% and 60%, respectively.

The mm/mLetter (mm)-wave small-world wireless NoC (mSWNoC) is an enabling interconnect architecture to design high performance and low power multicore chips. As the mSWNoC has an overall irregular topology, it is extremely important to design suitable deadlock-free routing mechanisms for it. In this paper we quantify the latency, energy disipation, and thermal profiles of mSWNoC architectures by incorporating irregular network-routing strategies. We demonstrate that the latency, energy dissipation, and thermal profile are affected by the adopted routing methodologies. In presence of the benchmarks considered, the variation in latency and energy dissipation is small. However, the network hotspot temperature can vary considerably depending on the exact routing strategy and the characteristics of the benchmark.

In this paper, we demonstrate that by inducing a certain fraction of randomness into wireless 3D NoCs (where CMOS wireless links are used for vertical inter-chip communication) we can reduce the communication latency when considering the physical constraints of 3D design space. Towards this end, we consider two cases, namely 1) replacing existing horizontal 2D links in a wireless 3D NoC with randomized shortcut NoC links and 2) enabling full connectivity via adding a randomized NoC layer to a wireless 3D system with no or partial horizontal connectivity. Consequently, the packet routing is optimized by exploiting both the existing and the newly added random NoC. Thus, by adding randomly wired shortcut NoCs to a wireless 3D system, one can strike a good balance between the modular design and the minimum randomness needed for achieving low-latency. Experimental results show that by adding a random NoC chip to wireless 3D CMPs without built-in horizontal NoCs we can reduce the communication latency by as much as 26.2% when compared to that of adding 2D mesh NoC. Also, the application execution time and average flit transfer energy can also be improved accordingly.

Network-on-chip (NoC) is a communication paradigm that has emerged to tackle different on-chip challenges and has satisfied different demands in terms of high performance and economical interconnect implementation. However, merely metal based NoC pursuit offers limited scalability with the relentless technology scaling, especially in one-to-many (1-to-M) communication. To meet the scalability demand, this paper proposes a new hybrid architecture empowered by both metal interconnects and Zenneck surface wave interconnects (SWI). This architecture, in conjunction with newly proposed routing and global arbitration schemes, avoids overloading the NoC especially in one-to-many (1-to-M) communication latency by as much as 26.2% when compared to that of adding a 2D mesh NoC. Also, the application execution time and average flit transfer energy can also be improved accordingly.

WiNoC replaces the traditional wired links with the use of wireless medium. Unfortunately, the energy consumed by the RF transceiver (i.e., the main building block of a WiNoC), and in particular by its transmitter, accounts for a significant fraction of the overall communication energy. In this paper we propose a runtime tunable transmitting power technique for improving the energy efficiency of the transceiver in wireless NoC architectures. The basic idea is tuning the transmitting power based on the location of the recipient of the current communication. The integration of the proposed technique into two known WiNoC architectures, namely, iWiSe64 and McWiNoC resulted in an energy reduction of 43% and 60%, respectively.

The mm/mLetter (mm)-wave small-world wireless NoC (mSWNoC) is an enabling interconnect architecture to design high performance and low power multicore chips. As the mSWNoC has an overall irregular topology, it is extremely important to design suitable deadlock-free routing mechanisms for it. In this paper we quantify the latency, energy disipation, and thermal profiles of mSWNoC architectures by incorporating irregular network-routing strategies. We demonstrate that the latency, energy dissipation, and thermal profile are affected by the adopted routing methodologies. In presence of the benchmarks considered, the variation in latency and energy dissipation is small. However, the network hotspot temperature can vary considerably depending on the exact routing strategy and the characteristics of the benchmark.

In this paper, we demonstrate that by inducing a certain fraction of randomness into wireless 3D NoCs (where CMOS wireless links are used for vertical inter-chip communication) we can reduce the communication latency when considering the physical constraints of 3D design space. Towards this end, we consider two cases, namely 1) replacing existing horizontal 2D links in a wireless 3D NoC with randomized shortcut NoC links and 2) enabling full connectivity via adding a randomized NoC layer to a wireless 3D system with no or partial horizontal connectivity. Consequently, the packet routing is optimized by exploiting both the existing and the newly added random NoC. Thus, by adding randomly wired shortcut NoCs to a wireless 3D system, one can strike a good balance between the modular design and the minimum randomness needed for achieving low-latency. Experimental results show that by adding a random NoC chip to wireless 3D CMPs without built-in horizontal NoCs we can reduce the communication latency by as much as 26.2% when compared to that of adding a 2D mesh NoC. Also, the application execution time and average flit transfer energy can also be improved accordingly.

Network-on-chip (NoC) is a communication paradigm that has emerged to tackle different on-chip challenges and has satisfied different demands in terms of high performance and economical interconnect implementation. However, merely metal based NoC pursuit offers limited scalability with the relentless technology scaling, especially in one-to-many (1-to-M) communication. To meet the scalability demand, this paper proposes a new hybrid architecture empowered by both metal interconnects and Zenneck surface wave interconnects (SWI). This architecture, in conjunction with newly proposed routing and global arbitration schemes, avoids overloading the NoC especially in one-to-many (1-to-M) communication latency by as much as 26.2% when compared to that of adding a 2D mesh NoC. Also, the application execution time and average flit transfer energy can also be improved accordingly.

Multi- and many-core architectures using Networks-on-Chip (NoC) are being explored for use in real-time safety-critical applications for their performance and efficiency. Such systems must provide isolation between tasks that may present distinct criticality levels. The NoC is critical to maintain the isolation property as it is a heavily used shared resource. To meet safety-standard requirements, such architectures require a systematic evaluation of the effects of all possible failures such as in the form of a Failure Mode and Effects Analysis (FMEA). We present the results of a detailed system-level analysis of a typical real-time mixed-critical network-on-chip architecture. This comprises an FMEA and error effects classification regarding duration and isolation violation.
This session discusses techniques to improve energy efficiency in large-scale computing systems, many-core systems, servers, and the cloud. The papers in this session particularly emphasize the practical experiences in academia and in industry.

**Time** | **Label** | **Presentation Title** | **Authors**
---|---|---|---
11:00 | 10.3.1 | GLOBAL FAN SPEED CONTROL CONSIDERING NON-IDEAL TEMPERATURE MEASUREMENTS IN ENTERPRISE SERVERS | Jungsoo Kim, Mohamed M. Sabry, David Altenza, Kalyan Vaidyanathan and Kenny Gross
  | | Speakers: | 1EPFL, CH; 2ESL-EPFL, CH; 3Physical Sciences Research Center, Oracle, US |
  | | Abstract | Time lag and quantization in temperature sensors in enterprise servers lead to stability concerns on existing variable fan speed control schemes. Stability challenges become further aggravated when multiple local controllers are running together with the fan control scheme. In this paper, we present a global control scheme which tackles the concerns on the stability of enterprise servers while reducing the performance degradation caused by the variable fan speed control scheme. We first present a stable fan speed control scheme based on the Proportional-Integral-Derivative (PID) controller by adaptively adjusting the PID parameters according to the operating fan speed and eliminating the fan speed oscillation caused by temperature quantization. Then, we present a global control scheme which coordinates control actions among multiple local controllers. In addition, it guarantees the server stability while minimizing the overall performance degradation. We validated the proposed control scheme using a presently shipping commercial enterprise server. Our experimental results show that the proposed fan control scheme is stable under the non-ideal temperature measurement system (10 sec in time lag and 1C in quantization figures). Furthermore, the global control scheme enables to run multiple local controllers in a stable manner while reducing the performance degradation up to 19.2% compared to conventional coordination schemes with 19.1% savings in server power consumption. |
11:30 | 10.3.2 | UNVEILING EURORA - THERMAL AND POWER CHARACTERIZATION OF THE MOST ENERGY-EFFICIENT SUPERCOMPUTER IN THE WORLD | Andrea Bartolini, Matteo Caciari, Carlo Cavazzoni, Giampietro Toghi and Luca Benini
  | | Speakers: | 1University of Bologna, Italy; 2ICINEA, Italy; 3EUROTECH, Italy; 4Università di Bologna, Italy |
  | | Abstract | Eurora (EUropean many integrated CoRe Architecture) is today the most energy efficient supercomputer in the world. Ranked 1st in the Green500 in July 2013, is a prototype built from Eurotech and Cineca toward next-generation Teraflop systems in the PRACE 2IP EU project. Eurora’s outstanding energy-efficiency is achieved by adopting a direct liquid cooling solution and a heterogeneous architecture with best-in-class general purpose HW components (Intel Xeon E5, Intel Xeon Phi and NVIDIA Kepler K20). In this paper we present a novel, low-overhead monitoring infrastructure capable to track in detail and in real-time the thermal and power characteristics of Eurora’s components with fine-grained resolution. Our experiments give insights on Eurora’s thermal/power trade-offs and highlight opportunities for run-time power/thermal management and optimization. |
12:00 | 10.3.3 | CONTENTION AWARE FREQUENCY SCALING ON CMPS WITH GUARANTEED QUALITY OF SERVICE | Hao Shen and Qinru Qiu, Syracuse University, US
  | | Speakers: | 1University of Bologna, Italy; 2ICINEA, Italy; 3EUROTECH, Italy; 4Università di Bologna, Italy |
  | | Abstract | Workload consolidation is usually performed in datacenters to improve server utilization for higher energy efficiency. One of the key issues related to workload consolidation is contention for shared resources such as last level cache, main memory, memory controller, etc. Dynamic voltage and frequency scaling (DVFS) of CPU is another effective technique that has widely been used to trade the performance for power reduction. We have found that the degree of resource contention of a system affects its performance sensitivity to CPU frequency. In this paper, we apply machine learning techniques to construct a model that quantifies runtime performance degradation caused by resource contention and frequency scaling. The inputs of our model are readings from Performance Monitoring Units (PMU) screened using standard feature selection technique. The model is tested on an SMT-enabled chip multiprocessor and it reaches up to 90% accuracy. Experimental results show that, guided by the performance model, runtime power management techniques such as DVFS can achieve more accurate power and performance tradeoff without violating the quality of service (QoS) agreement. The QoS violation of the proposed system is significantly lower than systems that have no performance degradation information. |
12:15 | 10.3.4 | CONCURRENT PLACEMENT, CAPACITY PROVISIONING, AND REQUEST FLOW CONTROL FOR A DISTRIBUTED CLOUD INFRASTRUCTURE | Shuang Chen, Yanzhi Wang and Massoud Pedram, University of Southern California, US
  | | Speakers: | 1University of Bologna, Italy; 2ICINEA, Italy; 3EUROTECH, Italy; 4Università di Bologna, Italy |
  | | Abstract | Cloud computing and storage have attracted a lot of attention due to the ever increasing demand for reliable and cost-effective access to vast resources and services available on the Internet. Cloud services are typically hosted in a Cloud computing and storage have attracted a lot of attention due to the ever increasing demand for reliable and cost-effective access to vast resources and services available on the Internet. Cloud services are typically hosted in a Cloud computing and storage which have widely been used to trade the performance for power reduction. We have found that the degree of resource contention of a system affects its performance sensitivity to CPU frequency. In this paper, we apply machine learning techniques to construct a model that quantifies runtime performance degradation caused by resource contention and frequency scaling. The inputs of our model are readings from Performance Monitoring Units (PMU) screened using standard feature selection technique. The model is tested on an SMT-enabled chip multiprocessor and it reaches up to 90% accuracy. Experimental results show that, guided by the performance model, runtime power management techniques such as DVFS can achieve more accurate power and performance tradeoff without violating the quality of service (QoS) agreement. The QoS violation of the proposed system is significantly lower than systems that have no performance degradation information. |
12:31 | 664 | COOLIP: SIMPLE YET EFFECTIVE JOB ALLOCATION FOR DISTRIBUTED THERMALLY-THROTTLING PROCESSORS | Pratyush Kumar, Hoseok Yang, Iuliana Bacivarov and Lothar Thiele, ETH Zurich, CH
  | | Speakers: | 1University of Bologna, Italy; 2ICINEA, Italy; 3EUROTECH, Italy; 4Università di Bologna, Italy |
  | | Abstract | Thermal constraints limit the time for which a processor can run at high frequency. Such thermal throttling complicates the computation of response times of jobs. For multiple processors, a key decision is where to allocate the next job. For distributed thermally-throttled processors, we present COOLIP with a simple allocation policy: a job is allocated to the earliest available processor, and if there are several available simultaneously, to the coolest one. For Poisson distribution of inter-arrivals times and Gaussian distribution of execution demand of unfinished jobs and thermal models of processors. We argue that COOLIP performs well because it directs the processors into states such that a defined sufficient condition of optimality holds. |
### 10.4 System-level evaluation

**Date:** Thursday, March 27, 2014  
**Time:** 11:00 - 12:30

**Location / Room:** Konferenz 2

**Chair:**  
Pablo Sanchez, University of Cantabria, ES, Contact Pablo Sanchez

**Co-Chair:**  
Florian Letombe, Synopsys, FR, Contact Florian Letombe

The session presents system-level verification and simulation techniques, as well as specific solutions for particular system components. The first paper analyzes how to detect concurrency errors from multi-threaded software on a virtual platform. The second proposes a hybrid simulation platform for cache configuration analysis. The last paper explores SSD verification challenges.

The session is completed by three IPs that introduce novel approaches for parallel simulation and efficient NoC/Smart systems validation.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>12:33</td>
<td>10.4.1</td>
<td>AUTOMATIC DETECTION OF CONCURRENCY BUGS THROUGH EVENT ORDERING CONSTRAINTS</td>
<td>Luis Gabriel Muñío, Simon Wawroschek, Jeronimo Castillón, Rainer Leupers and Gerd Ascheid, RWTH Aachen University, DE</td>
</tr>
<tr>
<td>12:33</td>
<td>10.4.2</td>
<td>HARDWARE-BASED FAST EXPLORATION OF CACHE HIERARCHIES IN APPLICATION SPECIFIC MPSoCs</td>
<td>Isuru Nawinne, Josef Schneider, Haris Javaid and Sri Parameswaran, The University of New South Wales, AU</td>
</tr>
<tr>
<td>12:33</td>
<td>10.4.3</td>
<td>SSDEXPLORER: A VIRTUAL PLATFORM FOR FINE-GRAINED DESIGN SPACE EXPLORATION OF SOLID STATE DRIVES</td>
<td>Lorenzo Zuolo1, Cristian Zambrilli1, Rino Micheloni1, Salvatore Galfano1, Marco Indaco1, Stefano Di Carlo1, Paolo Prinetto1, Prino Olivo1 and Davide Bertozzi3</td>
</tr>
</tbody>
</table>

---

1Department of Electrical, Electronic and Information Engineering (DEI) University of Bologna, IT; 2Microelectronic Systems Design Research Group, University of Kaiserslautern, DE  
2Department of Electrical, Electronic and Information Engineering (DEI) University of Bologna, IT; 3Polytechnic of Torino, IT

---

Lorenzo Zuolo1, Cristian Zambrilli1, Rino Micheloni1, Salvatore Galfano1, Marco Indaco1, Stefano Di Carlo1, Paolo Prinetto1, Prino Olivo1 and Davide Bertozzi3

---

Università degli Studi di Ferrara, IT; 1PMC-Siena, IT; 2Politecnico di Torino, IT

---

Abstract  
Solid State Drives (SSDs) are gaining particular momentum in various frameworks such as multimedia, large data centers and cloud environments. Unfortunately, efficient CAD tools for SSD design space exploration able to assess the optimization of the device microarchitecture w.r.t. the target performance are still missing. This paper tries to close this gap by proposing SSDExplorer, a tool for fine-grained and fast design space exploration of SSD devices. SSDExplorer provides unprecedented insights into the architecture behavior and subcomponent interaction efficiency, while avoiding the need for the actual implementation of an FTL or of key hardware components. This is achieved by the introduction of suitable abstractions of the different components. This is confirmed by the thorough validation of SSDExplorer against a commercial SSD device.
With increasing chip complexity, Networks-on-Chips (NoCs) are becoming a central platform for future on-chip communications. Many regular NoC architectures have been proposed to eliminate the communication bottlenecks of traditional bus-based networks. Non-rectangular and irregular architectures have also been proposed to increase performance. However, the complexity of designing custom non-rectangular networks leads to a rapid increase in design and verification times. To alleviate the conflict between performance and efficiency, this paper proposes a novel method that efficiently constructs virtual non-rectangular topologies on a mesh network by using time-regulated models to emulate irregular patterns. Data routings on virtual hexagonal and two irregular geometries validate the proposed method. An MPEG-4 decoder is used to exemplify the proposed method for media applications. Results analysis shows the virtual topologies emulated by the proposed method can provide precise timing and energy performance.

The first paper proposes a new static analysis approach based on segment graphs that identifies a tight set of potential access conflicts in segments that may-happen-in-parallel (MHP). Our experimental results show that the analysis is complete, accurate and fast to reveal dangerous shared variables in several embedded application models. Compared to earlier work, our approach significantly reduces the number of false conflict reports and thus saves the designer time.

The second paper introduces a technique for latency analysis for shared resource systems. By expressing task enabling and exploration, which is not scalable due to its inherent susceptibility to combinatorial explosion. We propose a scalable timing analysis method on periodically restarted Directed Acyclic Task Graphs, that can provide conservative bounds on task timing properties when shared resources with FCFS scheduling are used. By expressing task enabling and completion times in intervals, denoting best-case and worst-case timing properties, contention on the shared resources can be estimated using conservative approximations.

The third paper proposes a method that improves the tradeoff between simulation speed and accuracy of models. In the second paper, a technique for latency analysis for shared resource systems is introduced. The third paper proposes a method that improves the tradeoff between simulation speed and accuracy of models. In the second paper, a technique for latency analysis for shared resource systems is introduced. The third paper proposes a method that improves the tradeoff between simulation speed and accuracy of models. In the second paper, a technique for latency analysis for shared resource systems is introduced.

### 10.6 Multi-processor and distributed systems

**Date:** Thursday, March 27, 2014  
**Time:** 11:00 - 12:30  
**Location / Room:** Konferenz 4

**Chair:** Orlando Moreira, Ericsson, NL  
**Co-Chair:** Giuseppe Lipari, ENS - Cachan, FR

This session features new results in scheduling, allocation and management of real-time application in multi-core and distributed systems. The first paper presents a control algorithm for managing real-time tasks so to meet thermal constraints in a multi-core chip. The second paper proposes an algorithm for mixed-criticality task allocation in a multiprocessor platform. The third paper proposes a method for generating a schedule for a multi-mode application in a distributed system.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>12:00</td>
<td>10.5.3</td>
<td>A DYNAMIC COMPUTATION METHOD FOR FAST AND ACCURATE PERFORMANCE EVALUATION OF MULTI-CORE Architectures</td>
</tr>
<tr>
<td>12:15</td>
<td>10.5.4</td>
<td>CROSS-CORRELATION OF SPECIFICATION AND RTL FOR SOFT IP ANALYSIS</td>
</tr>
<tr>
<td>12:30</td>
<td></td>
<td>End of session</td>
</tr>
</tbody>
</table>

### Lunch Break

**Lunch Break** in Exhibition Area  
**Sandwich lunch**

**10.6 Multi-processor and distributed systems**

**Abstract**

Early estimation of performance has become necessary to facilitate design of complex multi-core architectures. Performance evaluation based on extensive simulations is time consuming and needs to be improved to allow exploration of different architectures in acceptable time. In this paper, we propose a method that improves the tradeoff between simulation speed and accuracy in performance models of architectures. This method computes during model execution some of the synchronization instants involved in architecture evolution. It allows grouping and abstracting architecture processes and this way significantly reduces the number of simulation events. Experiments show significant benefits from the computation method on the simulation time. Especially, a simulation speed-up by a factor of 4 is achieved in the considered case study, with no loss of accuracy about estimation of processing resource usage. The proposed method has potential to support automatic generation of efficient architecture models.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>11:00</td>
<td>10.6.1</td>
<td>THERMAL-AWARE FREQUENCY SCAILING FOR ADAPTIVE WORKLOADS ON HETEROGENEOUS MPSoCs</td>
</tr>
<tr>
<td>11:30</td>
<td>10.6.2</td>
<td>PARTITIONED MIXED-CRITICALITY SCHEDULING ON MULTIPROCESSOR PLATFORMS</td>
</tr>
</tbody>
</table>

**Abstract**

Semiconductor companies often use third-party IPs in order to improve their design productivity. In practice, there are risks involved in using a third-party IP as bugs may creep in due to versioning issues, poor documentation, and mismatches between specification and RTL. As a result of this, third-party IP specification and RTL must be carefully evaluated. Our methodology addresses this issue, which cross-correlates specification and RTL to discover these discrepancies. The key innovative ideas in our approach are to use prior and trusted experience about designs, which include their specs and RTL code. Also, we have captured this trusted experience into two knowledge bases (KB), Spec-KB and RTL-KB. Finally, knowledge base rules are used to cross-correlate the RTL blocks to the specs. We have tested our approach by analyzing several third-party IPs. We have defined metrics for specification coverage and RTL identification coverage to quantify our results.

Heng Yu, Rizwan Syed and Yajun Ha, National University of Singapore, SG

**Abstract**

For applications featuring adaptive workloads, the quality of their task execution can be dynamically adjusted given the runtime constraints. When mapping them to heterogeneous MPSoCs, it is expected not only to achieve the highest possible execution quality, but also meet the critical thermal challenges from the continuously increasing chip density. Prior thermal management techniques, such as Dynamic Voltage/Frequency Scaling (DVFS) and thread migration, do not take into account the trade-off possibility between execution quality and temperature control. In this paper, we explore the capability of adaptive workloads for effective temperature control, while maximally ensuring the execution Quality-of-Service (QoS). We present a thermal-aware dynamic frequency scaling (DFS) algorithm on heterogeneous MPSoCs, where judicious frequency selection achieves QoS maximization under the temperature threshold, which is converted to the thermal-timing deadline as an additional execution constraint. Results show that our frequency scaling algorithm achieves as large as 31.5% execution cycle/QoS improvement under thermal constraints.

Heng Yu, Rizwan Syed and Yajun Ha, National University of Singapore, SG

**Abstract**

Scheduling mixed-criticality systems that integrate multiple functionalities with different criticality levels into a shared platform appears to be a challenging problem, even on single-processor platforms. Multi-core processors are more and more widely used in embedded systems, which provide great computing capacities for such mixed-criticality systems. In this paper, we propose a partitioned scheduling algorithm MPVD to extend the state-of-the-art single-processor mixed-criticality scheduling algorithm EY to multiprocessor platforms. The key idea of MPVD is to evenly allocate tasks with different criticality levels to different processors, in order to better explore the asymmetry between different criticality levels and improve the system schedulability. Then we propose two enhancements to further improve the schedulability of MPVD. Experiments with randomly generated task sets show significant performance improvement of our proposed approach over existing algorithms.

Chuancai Gu, Nan Guan, Qingxu Deng and Wang YF  
Northeastern University, CN; Uppsala University, SE

**Abstract**

For applications featuring adaptive workloads, the quality of their task execution can be dynamically adjusted given the runtime constraints. When mapping them to heterogeneous MPSoCs, it is expected not only to achieve the highest possible execution quality, but also meet the critical thermal challenges from the continuously increasing chip density. Prior thermal management techniques, such as Dynamic Voltage/Frequency Scaling (DVFS) and thread migration, do not take into account the trade-off possibility between execution quality and temperature control. In this paper, we explore the capability of adaptive workloads for effective temperature control, while maximally ensuring the execution Quality-of-Service (QoS). We present a thermal-aware dynamic frequency scaling (DFS) algorithm on heterogeneous MPSoCs, where judicious frequency selection achieves QoS maximization under the temperature threshold, which is converted to the thermal-timing deadline as an additional execution constraint. Results show that our frequency scaling algorithm achieves as large as 31.5% execution cycle/QoS improvement under thermal constraints.

Heng Yu, Rizwan Syed and Yajun Ha, National University of Singapore, SG
Papers in this session address synthesis algorithms and tools at different levels, targeting power, area and delay minimization.

Kim Taemin, Intel Labs, US,
Co-Chair:
John Hayes, University of Michigan, US,
Chair:

Location / Room: Konferenz 5
Time: Thursday, March 27, 2014
Date:

10.7 Advances in Synthesis

10.7.1 PROVABLY MINIMAL ENERGY USING COORDINATED DVS AND POWER GATING
Speakers:
Nathaniel Conos, Sam Meguerdichian, Foad Dabiri and Miodrag Potkonjak, UCLA, US

Abstract
Both energy and execution speed can be greatly impacted by clock and power gating, nonlinear voltage scaling, and leakage power. We address the problem of coordinated power gating and dynamic voltage scaling (DVS) to minimize the overall energy consumption of an application under user-specified timing constraints. We prove that a solution provided by our convex programming formulation that uses at most two versions of hardware, where each version uses its own constant voltages, is optimal. Comprehensive evaluation of the new approach demonstrates energy improvements over traditional DVS and DVS and power gating techniques by factors of 1.44X-2.82X and 1.44X-2.82X, respectively.

10.7.2 A TREE ARBITER CELL FOR HIGH SPEED RESOURCE SHARING IN ASYNCHRONOUS ENVIRONMENTS
Speakers:
Syed Rameez Naqvi and Andreas Steininger, Vienna University of Technology, AT

Abstract
We present a novel tree arbiter cell that allows a pipelined processing of asynchronous requests. In this way it can achieve significantly lower delay in the critical case of frequent requests coming from different clients. We elaborate the necessary extension to facilitate a cascaded use of this cell in a tree-like fashion, and we show by theoretical analysis that in this configuration our cell provides better fairness than the standard approach. We implement our approach and quantitatively compare its performance properties with related work in a gate-level simulation. In our sample asynchronous Networks-on-Chip application our new cell proves to increase the throughput of three different designs available in literature by approximately 61.28%, 69.24%, and 186.85% respectively.

10.7.3 AN EFFICIENT MANIPULATION PACKAGE FOR BICONDITIONAL BINARY DECISION DIAGRAMS
Speakers:
Luca Amaru, Pierre-Emmanuel Gallatlon and Giovanni De Micheli, EPFL, CH

Abstract
Biconditional Binary Decision Diagrams (BBDDs) are a novel class of binary decision diagrams where the branching condition, and its associated logic expansion, is biconditional on two variables. Reduced and ordered BBDDs are remarkably compact and unique for a given Boolean function. In order to exploit BBDDs in Electronic Design Automation (EDA) applications, efficient manipulation algorithms must be developed and integrated in a software package. In this paper, we present the theory for efficient BBDD manipulation and its practical software implementation. The key features of the proposed approach are (i) strong canonical form pre-conditioning of stored BBDD nodes, (ii) recursive formulation of Boolean operations in terms of biconditional expansions, (iii) performance-oriented memory management and (iv) dedicated BBDD re-ordering techniques. Experimental results show that the developed BBDD package achieves an average node count reduction of 19.48% and a speed-up factor of 1.63x with respect to a state-of-the-art decision diagram manipulation package. Employed in the synthesis of datapath circuits, the BBDD manipulation package is capable to advantageously restructure arithmetic operations producing 11.02% smaller and 32.29% faster circuits as compared to a commercial synthesis flow.

10.7.4 SYNTHESIS ALGORITHM OF PARALLEL INDEX GENERATION UNITS
Speaker:
Yusuke Matsunaga, Kyushu University, JP

Abstract
The index generation function is a multi-valued logic function which checks if the given input vector is a registered or not, and returns its index value if the vector is registered. If the latency of the operation is critical, dedicated hardware is used for implementing the index generation functions. This paper proposes a method implementing the index generation functions using parallel index generation units. A novel and efficient algorithm called 'conflict free partitioning' is proposed to synthesis parallel index generation units. Experimental results show the proposed method outperforms other existing methods.
Today the most powerful innovations in the major industries and the most promising approaches to tackle burning societal challenges are substantially influenced by and depending from the innovations provided by the microelectronics industry. Breakthroughs in manufacturing technologies enable the realization of novel types of devices and of systems, which enable applications with fascinating functionality and enormous performance. However, this innovation chain is not operational without appropriate innovations in design technology: We need an innovation Agenda 2020 for design methodology and EDA tools fueling the innovation chain of electronics.
**UB10 Session 10**

**Date:** Thursday, March 27, 2014  
**Time:** 12:00 - 14:30  
**Location / Room:** University Booth, Booth 3, Exhibition Area

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
</table>
| 11:00 | 10.8.1 | INTRODUCTION | Moderator:  
Ahmed Jerraya, CEA-LETI, FR  

| 12:30 |  | End of session  
Lunch Break in Exhibition Area  
Sandwich lunch |  |

**UB10.01**  
SOC VERIFICATION: AUTOMATED FUNCTIONAL VERIFICATION OF SYSTEMS-ON-CHIP  
**Authors:** Zdenek Prikryl, Marcela Simkova and Karel Masarik, Faculty of Information Technology, Brno University of Technology, CZ  
**Abstract**  
An increase of the complexity of systems-on-chip (SoC) induces an increase of the complexity of their verification as well. The reason is that we must verify not only the functions of separate logic blocks, but we need to check their interconnections, timing and functional collaboration as well. Therefore, there is still a great demand for verification tools, which are time-effective, fast and as automated as possible. Exactly these issues we target in our solution. You are welcome to see the live demonstration at our booth!

**More information...**

**UB10.02**  
BRIDGING MATLAB/SIMULINK AND ESDL DESIGN VIA AUTOMATIC CODE GENERATION  
**Authors:** Liyuan Zhang, Michael Glaß and Jürgen Teich, University of Erlangen-Nuremberg, DE  
**Abstract**  
Matlab/Simulink is today’s de-facto standard for model-based design in domains such as control engineering and signal processing. Commercial tools are available to generate embedded C or HDL code directly from a Simulink model. However, Simulink models are purely functional models and, hence, designers cannot seamlessly consider the architecture that a Simulink model is later implemented on. In particular, it is not possible to explore the different architectural alternatives and investigate the arising interactions and side-effects directly within Simulink. To benefit from Matlab/Simulink’s algorithm exploration capabilities and overcome the outlined drawbacks, we introduce a model transformation framework that converts a Simulink model to an executable specification, written in an actor-oriented modeling language. This specification then serves as the input of an established Electronic System Level (ESL) design flow, enabling Design Space Exploration (DSE) and automatic code generation for both hardware and software. In this demonstration, we will show how to automatically transform Simulink models to an established ESL design flow by means of a code generator. Based on the generated code, we will present a co-simulation approach that combines complex environmental models from Matlab/Simulink with the auto-generated model of a controller. We will use an Anti-lock Braking System (ABS) as an example where we investigate the impact of different controller implementations in the automotive E/E architecture. In detail, the following scientific achievements are included in the proposed demonstration: To bridge Simulink and ESL design flows, we developed an ESL Code-Generator to perform model transformation. The idea is that for any given Simulink models such as a controller in a control system, the designer can simply invoke our Code-Generator to create the ESL model automatically. In our design flow, we use SystemC as a programming language with an extension of actors with a specific Model of Computation (MoC). We guarantee the preservation of the semantics of the generated model by (a) applying a specific 1-to-1 mapping from Simulink basic blocks to an actor library and (b) considering different transformations to capture single-rate and multi-rate Simulink models. After the model transformation is finished, this auto-generated SystemC model serves as the input of a well-established ESL design flow that enables DSE. Besides the Code-Generator we demonstrate also a validation technique that considers the functional correctness by comparing the original Simulink model with the generated SystemC model. The main idea behind this technique is (1) to co-simulate the auto-generated model along with the the original model and (2) to reuse the environment model and the test bench that are originally created in Simulink also for the auto-generated model. Furthermore, the performance of the model can also be measured during co-simulation. In this demonstration, an ABS model will be transformed from Simulink to SystemC by invoking ESL Code-Generator. Then, by applying our validation technique, the correctness and the accuracy of the auto-generated model can be examined. Lastly, to evaluate the performance of the model, application-depended quality of control will be measured, such as the braking distance on an icy road.

**More information...**

**UB10.04**  
GEMINI: A NEW SYNTHESIS AND OPTIMIZATION TOOL FOR GRAPHENE-BASED DIGITAL DEVICES  
**Authors:** Valerio Tenace, Andrea Calimera, Massimo Poncino and Enrico Masi, Politecnico di Torino, IT  
**Abstract**  
Gemini is a synthesis and optimization software for graphene-based digital devices. Given a combinational circuit description through its boolean representation, Gemini produces a SPICE netlist mapped with graphene PN-Junction gates. The software is composed of a parser library to handle input circuit descriptions, a characterization library of graphene gates used in the synthesis process, a Biconditional Binary Decision Diagram library used to manipulate logic networks in PaaX-NOR logic in order to better exploit the intrinsic characteristics of the adopted graphene gates, and a number of optimization algorithms designed to produce better results in terms of area and thus power consumption. As a stand-alone software or as a library easy to integrate into state-of-the-art tools, Gemini represents a first step of an enabling technology for future synthesis and optimization processes for graphene-based devices.

**More information...**

**UB10.05**  
RESCV: RESOURCE-AWARE COMPUTER VISION APPLICATION ON HETEROGENEOUS MULTI-TILE ARCHITECTURE  
**Authors:** Eftdes Sousas, Johny Paul, Vahid Lari, Frank Hannig  
Technische Universität München, DE  
**Abstract**  
We demonstrate the benefits of invasive computing by showing the efficiency and utilization improvements in a resource-aware manner by algorithmic selection of different invasive resources, such as TCPA (tightly-coupled processor array), and RISC processors. More specific we present a dynamic load balancing of a computer vision application between multiple RISC cores and a TCPA, based on invasive mechanisms supported by our operating system and the agent system.

**More information...**
Virtual prototyping and Electronic System Level (ESL) modeling have become valuable approaches to cope with the ever-increasing complexity of embedded systems. Their effectiveness, however, is highly dependent on their quick development time and accuracy both conflicting goals. In this demonstration, we present (a) an ESL methodology [1] [2] for the simulation-based evaluation of power and performance of embedded systems by the use of virtual prototypes. Our methodology permits us to develop ESL models for design space exploration of dynamic power and performance management strategies and hardware/software co-design choices. (b) We present a novel sketch-based tool termed Mahler [3] for the very early design phase of ESL modeling. Mahler provides a playground to quickly model functionality and evaluate performance on different architecture implementations. In Mahler, ESL models are created by literally sketching with a pen or touch interface, e.g. a tablet stylus, or a touchless interface, such as a Leap Motion controller. The application and architecture models are transformed to an executable virtual prototype through sketch recognition. This approach provides a very intuitive and fast way to explore actor-oriented functional modeling and hardware/software partitioning. The output of Mahler is a simulation-ready SystemC-based source-code stub that can be refined for subsequent design iterations. We will show a model of a Voice over LTE (VoLTE) use case, consisting of a heterogeneous cellular SoC platform, together with a wireless channel fading model and a base station network model. State-based [1] and polynomial/equation-based [4] power models are built and co-simulated for the SoC digital module and the RF transceiver module, respectively to abstract their different power consumption characterization accurately. The entire end-to-end modeling enables efficient SoC performance and power simulation with proper network configuration in seconds, which is highly desired in cellular system early design exploration phase and co-optimization with network vendors.

**More information ...**

---

**11.0 SPECIAL DAY Embedded Tutorial: Alternatives to CMOS**

Organic semiconductors with conjugated electron system are currently intensively investigated for optoelectronic applications. This interest is spurred by novel devices such as organic light-emitting diodes (OLED), organic solar cells, and flexible electronics. In this talk, I will discuss some of the recent progress in realizing devices, in particular highly efficient white OLED for lighting and flexible organic solar cells.

**Time** | **Label** | **Presentation Title** | **Authors**
---|---|---|---
13:30 | 11.0.1 | ORGANIC ELECTRONICS - FROM LAB TO MARKETS | Martin Knupfer, IFW Dresden, DE
14:00 | End of session | | |
15:30 | Coffee Break in Exhibition Area | | On Tuesday-Thursday the coffee and lunch breaks will be located in the Exhibition Area (Terrace Level).
11.2 Transitioning NoC Design Techniques to Future Challenges

Date: Thursday, March 27, 2014
Time: 14:00 - 15:30
Location / Room: Konferenz 6

Chair:
Masoud Danesh talab, University of Turku, FI; Contact Masoud Danesh talab

Co-Chair:
Hiroki Matsutani, Keio University, JP; Contact Hiroki Matsutani

The first paper of the session presents an approach to tolerating faults in NoCs through runtime reconfiguration, which is of increasing importance. The second paper focuses on management of thermal behaviour in NoCs to improve the reliability of optical communication given the tight tolerances of silicon photonics. Finally, the third paper also provides an outlook on optical NoCs by questions (Device, Circuit, Architecture, System and CAD) related to this topic to share the most recent results and discuss the future challenges.

11.1.1 SPINTRONICS FOR LOW-POWER COMPUTING

Speakers:
Yue Zhang1, Weisheng Zhao1, Jacques-Olivier Klein1, Wang Kang1, Damen Quekroz1, Youguang Zhang2, Datiné Ravelosona1 and Claude Chappert1

1IEF - Univ. Paris Sud, FR; 2Univ. Beihang, CN

Abstract
Microelectronics has been following Moore’s law for almost 40 years. However this trend tends to run out of steam in recent technology nodes. The continuous improvements in the size of the transistors and in the operating frequencies result in serious power consumption, heat dissipation and reliability issues. Spintronics (Nobel Prize of Physics, 2007 awarded to Prof. Fert from Univ. Paris-Sud and Peter Grünberg from Forschungszentrum Jülich) nanodevices can reduce significantly the power, improve the reliability or allow new functionalities. The 2010 ITRS report on emerging research devices identified Magnetic Tunnel Junction (MTJ) nanopillar (the prevalent spintronics nanodevice) as one of the most promising technologies to be part of the future microelectronics circuits. It provides data non-volatility, hardness to radiation, fast data access and low-power operations. Magnetic memories become the most promising candidate for both low power logic computing and the data storage. This tutorial paper presents multidiscipline operations. Magnetic memories become the most promising candidate for both low power logic computing and the data storage. This tutorial paper presents multidiscipline spintronics research issues (Device, Circuit, Architecture, System and CAD) to share the most recent results and discuss the future challenges.

11.1.2 CHAMELEON: CHANNEL EFFICIENT OPTICAL NETWORK-ON-CHIP

Speakers:
Sébastien Le Beux1, Hui Li1, Ian O’Connor1, Kazem Cheshm1, Xuchen Liu1, Jelena Trajkovic1 and Gabriela Nicolcsou1

1Lyon Institute of Nanotechnology, FR; 2Concordia University, CA; 3Ecole Polytechnique de Montréal, CA

Abstract
The next generation of MPSoC points to the integration of thousands of IP cores, requiring high performance interconnected for high throughput communications. Optical on-chip interconnection enables significantly increased bandwidth and decreased latency in MPSoC. However, the interface between electrical and photonic devices implies strong layout constraints that may impact the system performance and scalability. In this paper, we propose a novel optical interconnection named CHAMELEON. The interface simplifies the layout and allows the bandwidth between IP cores to be adapted according to the communication requirements. Compared to related networks, CHAMELEON demonstrates improved scalability and flexibility at the cost of minor increase in power consumption.

11.1.3 LOW-VOLTAGE ORGANIC TRANSISTORS FOR FLEXIBLE ELECTRONICS

Speakers:
Ute Zschieschang1, Reinhold Rödel1, Ulrike Kraft1, Kazuo Takimiya2, Tarek Zak1, Florian Letzkus3, Jörg Butschke4, Harald Richter4, Joachim Burghart4, Wei Xiong5, Boris Mumann6 and Hagen Klauk1

1Max Planck Institute for Solid State Research, DE; 2Riken Advanced Science Institute, JP; 3University of Stuttgart, DE; 4HMS CHIPS, DE; 5Stanford University, US

Abstract
A process for the fabrication of bottom-gate, top-contact (inverted staggered) organic thin-film transistors (TFTs) with channel lengths as short as 1 µm on flexible plastic substrates has been developed. The TFTs employ vacuum-deposited small-molecule semiconductors and a low-temperature-processed gate dielectric that is sufficiently thin to allow the TFTs to operate with voltages of about 3 V. The p-channel TFTs have an effective field-effect mobility of about 1 cm²/Vs, an on/off ratio of 10⁷, and a signal propagation delay (measured in 11-stage ring oscillators) of 300 ns per stage. For the n-channel TFTs, an effective field-effect mobility of about 0.06 cm²/Vs, an on/off ratio of 10⁶, and a signal propagation delay of 17 µs per stage have been obtained.

15:30 End of session

Coffee Break in Exhibition Area

On Tuesday-Thursday the coffee and lunch breaks will be located in the Exhibition Area (Terrace Level).

11.2 BRISK AND LIMITED-IMPACT NOC ROUTING RECONFIGURATION

Speakers:
Doowon Lee, Ritesh Parikh and Valeria Bertacco, University of Michigan, US

Abstract
The expected low reliability of the silicon substrate at upcoming technology nodes presents a key challenge for digital system designers. Networks-on-chip (NoCs) are especially concerning because they are often the only communication infrastructure for the chips in which they are deployed. Recently, routing reconfiguration solutions have been proposed to address this problem. However, they come at a high silicon cost, and often require suspending the normal network activity while executing a centralized, resource-hungry reconfiguration algorithm. This paper proposes a novel fast and minimalistic routing reconfiguration algorithm, called BLINC. BLINC utilizes precomputed routing metadata to quickly evaluate localized detours upon each fault manifestation. We showcase the efficacy of our algorithm by deploying it in a novel NoC fault detection and reconfiguration solution, where BLINC enables uninterrupted NoC operation during aggressive online testing. If a fault seems likely to occur, we circumvent it in advance with the aid of our BLINC reconfiguration algorithm. Experimental results show an 80% reduction in the average number of routers affected by a reconfiguration event, compared to state-of-the-art techniques. BLINC enables negligible performance degradation in our detection and reconfiguration solution, while solutions based on current techniques suffer a 17-fold latency increase.
This session addresses various aspects of system modeling, synthesis, validation and verification with the strong focus on industrial relevance.

Norbert Wehn, TU Kaiserslautern, DE,
Co-Chair: Emil Matus, Technische Universität Dresden, DE,
Chair: Location / Room: Time: Date: 11.3 Industry relevant research and practice for system design
Time Label Presentation Title Authors
11:22 11.2.2 THERMAL MANAGEMENT OF MANYCORE SYSTEMS WITH SILICON-PHOTONIC NOCS
Speakers: Tiansheng Zhang, José L. Abellán, Ajay Joshi and Ayse K. Coskun, Boston University, US
Abstract Silicon-photonic network-on-chips (NoCs) provide high bandwidth density; therefore, they are promising candidates to replace electrical NoCs in manycore systems. The silicon-photonic NoCs, however, are sensitive to the temperature gradients that typically occur on the chip, and hence, require proactive thermal management. This paper first provides a design space exploration of silicon-photonic networks in manycore systems and quantifies the performance impact of the temperature gradients for various network bandwidths. The paper then introduces a novel job allocation technique that minimizes the temperature gradients among the ring modulators/filters to improve the application performance. Experimental results for a single-chip 256 core system demonstrate that our policy is able to maintain the maximum network bandwidth. Compared to existing workload allocation policies, the proposed policy improves system performance by up to 26.1% when running a single application and 18.3% for multi-program scenarios.
15:30 11.2.3 ASSESSING THE ENERGY BREAK-EVEN POINT BETWEEN AN OPTICAL NOC ARCHITECTURE AND AN AGGRESSIVE ELECTRONIC BASELINE
Speakers: Luca Rami1, Paolo Grani1, Herve Tatenguem Fankem1, Alberto Ghiribaldi2, Sandro Bartolini2 and Davide Bertozzi2
1: Engineering Department of the University of Ferrara, IT; 2: University of Siena, IT
Abstract Many crossbenchmarking results reported in the open literature raise optimistic expectations on the use of optical networks-on-chip (ONoCs) for high-performance and low-power on-chip communication. However, most of those previous works ultimately fail to make a compelling case for chip-level nanophotonic NoCs, especially for the lack of aggressive electronic baselines (ENoC), and the poor accuracy in physical- and architecture-layer analysis of the ONoC. This paper aims at providing the guidelines and minimum requirements so that nanophotonic emerging technology may become of practical relevance. The key differentiating factor of this work consists of contrasting ONoC solutions with an aggressive ENoC architecture with realistic complexity, performance, and power figures, synthesized on an industrial 40nm low-power technology. At the same time, key physical design issues and network interface architecture requirements for the ONoC under test are carefully assessed, thus paving the way for a well-grounded definition of the requirements for the emerging ONoC technology to achieve the energy break-even point with respect to pure electronic interconnect solutions in future multi- and many-core systems.
15:30 IPS-11, 618 DCM: AN IP FOR THE AUTONOMOUS CONTROL OF OPTICAL AND ELECTRICAL RECONFIGURABLE NOCS.
Speakers: Wolfgang Büter1, Christof Osewold1, Daniel Gregorek2 and Alberto Garcia-Ortiz2
1: University of Bremen, DE; 2:ITEM (U.Bremen), DE
Abstract The increasing requirements for bandwidth and quality-of-service motivate the use of parallel interconnect architectures with several degrees of reconfiguration. This paper presents an IP, called Distributed Channel Management (DCM), to extend existing packet-switched NoCs with a reconfigurable point-to-point network seamlessly, i.e., without the need for any modification on the routers. The configuration of the reconfigurable network takes place dynamically and autonomously, so that the topology can be changed at run time. Furthermore, the architecture is scalable due to the autonomous decentralised administration of the links. The paper reports a thorough experimental analysis of the overhead of the approach at the gate level that considers different network parameters such as link size and timing constraints.
15:31 IPS-12, 726 MINIMALLY BUFFERED SINGLE-CYCLE DEFLECTION ROUTER
Speakers: Gnaneswara Rao Jonna1, John Jose1, Rachana Radhakrishnan1 and Madhu Mutyam1
1: Indian Institute of Technology, Madras., IN; 2: Rajagiri School of Engineering & Technology, Kochhi., IN
Abstract With the drift from computation centric designs to communication centric designs in the Chip Multi Processor (CMP) era, the interconnect fabric is gaining more importance. An efficient NoC in terms of power, area and average flit latency has a huge impact on the overall performance of a CMP. In the current work, we propose MinBSD - a minimally buffered, single cycle, deflection router. It incorporates different operations (Injection, Ejection, Preemption, Re-injection) in a single module to handle the traffic effectively and ensures smooth flow of flits through router pipeline. It performs overlapped execution of independent operations. These factors not only make MinBSD to operate in a single cycle but also to reduce the critical path latency resulting in a faster interconnect network. Experimental results show that MinBSD reduces the average flit latency on real work loads, reduces die area and power consumption when compared to the existing state-of-the-art minimally buffered deflection routers.
15:30 End of session
Coffee Break in Exhibition Area
On Tuesday-Thursday the coffee and lunch breaks will be located in the Exhibition Area (Terrace Level).

11.3 Industry relevant research and practice for system design
Date: Thursday, March 27, 2014
Time: 14:00 - 15:30
Location / Room: Konferenz 1
Chair: Emil Matys, Technische Universität Dresden, DE, Contact Emil Matys
Co-Chair: Norbert Wehn, TU Kaiserslautern, DE, Contact Norbert Wehn
This session addresses various aspects of system modeling, synthesis, validation and verification with the strong focus on industrial relevance.
11.4 Enabling validation on fast platforms

Date: Thursday, March 27, 2014
Time: 14:00 - 15:30
Location / Room: Konferenz 2

Chair: Ronny Morad, IBM, IL, Contact Ronny Morad
Co-Chair:
Fast platforms, whether acceleration, post-silicon or virtual prototypes, are key technologies to enabling validation of complex systems. However, they present enormous challenges to become effective. This session presents four papers and two IPs that propose solutions to overcome some of them, thus enabling much higher performance and coverage.

### 11:41 Coverage Evaluation of Post-Silicon Validation Tests with Virtual Prototypes

**Speakers:**
Kalong Cong, Li Lei, Zhunkun Yang and Fei Xie, Portland State University, US

**Abstract**
High-quality tests for post-silicon validation should be ready before a silicon device becomes available in order to save time spent on preparing, debugging and fixing tests after the device is available. Test coverage is an important metric for evaluating the quality and readiness of post-silicon tests. We propose an online-capture offline-replay approach to coverage evaluation of post-silicon validation tests with virtual prototypes for estimating silicon device test coverage. We first capture necessary data from a concrete execution of the virtual prototype within a virtual platform under a given test, and then compute the test coverage by efficiently replaying this execution offline on the virtual prototype itself. Our approach provides early feedback on quality of post-silicon validation tests before silicon is ready. To ensure fidelity of early coverage evaluation, our approach has been further extended to support coverage evaluation and conformance checking in the post-silicon stage. We have applied our approach to evaluate a suite of common tests on virtual prototypes of five network adapters. Our approach was able to reliably estimate that this suite achieves high functional coverage on all five silicon devices.

### 13:00 Archiving: Architectural Checking via Event Diagrams for High Performance Validation

**Speakers:**
Chang-Hong Hsu, Debapriya Chatterjee, Ronny Morad, Raviv Gal and Valeria Bertacco

**Abstract**
Simulation-based techniques play a key role in validating the functional correctness of microprocessor designs. A common approach for validating microprocessors (called instruction-by-instruction, or IBI checking) consists of running a RTL and an architectural simulation in lock-step, while comparing processor architectural state at each instruction retirement. This solution, however, cannot be deployed on long regression tests, because of the limited performance of RTL simulators. Acceleration platforms have the performance power to overcome this issue, but are not amenable to the deployment of an IBI checking methodology. Indeed, validation on these platforms requires logging activity on-platform and then checking it against a golden model off-platform. Unfortunately, an IBI checking approach following this paradigm entails a large slowdown for the acceleration platform, because of the sizable amount of data that must be transferred off-platform for comparison against the golden model. In this work we propose a sequence-by-sequence (SBS) checking approach that is efficient and practical for acceleration platforms. Our solution validates the test execution over sequences of instructions (instead of individual ones), thus greatly reducing the amount of data transferred for off-platform checking. We found that SBS checking delivers the same bug detection accuracy as traditional IBI checking, while reducing the amount of traced data by more than 90%.

### 15:00 Effective Post-Silicon Failure Localization Using Dynamic Program Slicing

**Speakers:**
Ophir Friedler, Wsam Kadry, Arkady Morgenstein, Amr Nahir and Vatsal Sokhi, IBM Research - Haifa, IL

**Abstract**
In post-silicon functional validation, one of the complex and time-consuming processes is the localization of an instruction that exposes a bug detected at system level. The task is especially hard due to the silicon’s limited observability and the long time between the failure’s occurrence and its detection. We propose a novel method that automates the architectural localization of post-silicon test-case failures. The proposed tool analyzes a failing test-case, while leveraging the information derived from executing the test on an Instruction Set software Simulator (ISS), to identify a set of instructions that could lead to the faulty final state. The proposed failure localization process comprises the creation of a resource dependency graph based on the execution of the test-case on the ISS, determining a program slice of instructions that influence the faulty resources, and the reduction of the set of suspicious instructions by leveraging the knowledge of the correct resources. We evaluate our proposed solution through extensive experiments. Experimental results show that, in over 97% of all cases, our method was able to narrow down the list of suspicious instructions to under 2 instructions, on average, out of over 200. In over 59% of all cases, our method correctly reduced a test-case to a single faulty instruction.

### 15:15 Design-for-Debug Routing for FIB Probing

**Speakers:**
Chieh-Yi Lee, Tai-Hung Li and Tai-Chen Chen, National Central University, TW

**Abstract**
To observe internal signals, physical probing is an important step in post-silicon debug. Focused ion beam (FIB) is one of most popular probing technologies. However, an unsuitable layout significantly decreases the percentage of nets which can be observed through FIB probing for advanced process technologies. This paper presents the first design-for-debug routing to increase the FIB observable rate. The proposed algorithm, which adopts three FIB states and costs to enhance the maze routing, keeps at least one FIB candidate for each net while routing. Experimental results demonstrate that the proposed method can significantly increase the FIB observable rate under 100% routability.

### 15:30 Functional Test Generation Guided by Steady-State Probabilities of Abstract Design

**Speakers:**
Jian Wang, Huawei LPP, Tao Lv, Tiancheng Wang and Xiaowei Li

**Abstract**
This paper presents a novel method for functional test generation aiming at exploring control state space of the design. The steady-state probabilities (SPs) of the abstract design’s control FSM are used to guide test generation. The SPs of the states can reflect how hard the states can be reached, and the hard-to-reach states are assigned with high priority to be exercised. Experimental results show that our method has better performance in test generation in comparison with constrained random simulation, and demonstrate that SPs can provide good guidance on traversing hard-to-reach states of the design under validation.

### 15:40 Automated System Testing Using Dynamic and Resource Restricted Clients

**Speakers:**
Miko Caspar, Miko Lippmann and Wolfram Hardt, Technische Universität Chemnitz, DE

**Abstract**
Testing on system level using a static and homogeneous architecture of clients is common practice. This paper introduces a new approach to use a heterogeneous and dynamic set of resource restricted clients for automated testing. Due to changing resources and availability of the clients, the test case distribution needs to be recalculated dynamically during the test execution. All necessary conditions and parameters are represented by a formal model. It is shown that the algorithmic problem of DYNAMIC TESTPARTITIONING can be solved in polynomial time by a heuristic recursive algorithm. A testbench architecture is introduced and by simulation it is shown that the testbench can execute the test requirements within a small variation using a number of several hundred clients. The system can react dynamically on changing resources and availability of the test clients within several seconds. The approach is generic and can be adapted to a huge number of systems.
11.5 Memory Resource Allocation and Scheduling in MPSoC

Date: Thursday, March 27, 2014
Time: 14:00 - 15:30
Location / Room: Konferenz 3

Chair: Andreas Herkersdorf, Technische Universität München, DE, Contact Andreas Herkersdorf
Co-Chair: Donatella Sciuto, Politecnico di Milano, IT, Contact Donatella Sciuto

Low-latency data access and efficient interprocess communication are critical to MPSoC performance and power efficiency. This session introduces innovative approaches for data placement, memory bandwidth allocation and scheduling techniques in MPSoC architectures with heterogeneous 2D/3D memory hierarchies.

---

11.5.1 SCENARIO-AWARE DATA PLACEMENT AND MEMORY AREA ALLOCATION FOR MULTI-PROCESSOR SYSTEM-ON-CHIPS WITH RECONFIGURABLE 3D-STACKED SRAMS

**Speakers:** Meng-Ling Tsai, Yi-Jung Chen, Yi-Ting Chen and Ru-Hua Chang, Department of Computer Science and Information Engineering, National Chi Nan University, TW

**Abstract**

Integrating Multi-Processor System-on-Chips (MPSoCs) with 3D-stacked reconfigurable SRAM tiles has been proposed for embedded systems with high memory demands. At runtime, the SRAM tiles are configured into several memory areas, which can be reconfigured according to the dynamic behavior of the system. Targeting this architecture, in this paper, we propose a data placement and memory area allocation algorithm. The goal of the proposed algorithm is to optimize the performance of the memory system by minimizing the on-chip memory access latency, the number of off-chip memory accesses, and the number of reconfigurations. Since the behavior of an embedded system can be described by a set of scenarios, where each scenario specifies a set of applications that would execute concurrently, the proposed algorithm synthesizes data placements and the memory area allocation for each scenario. Not only the data access patterns within the scenario but also among all scenarios are considered for data placement. We evaluate the proposed algorithm on a set of synthetic and real-world applications. The experimental results show that, compared to the existing data placement method designed for MPSoCs with distributed memory modules, the proposed algorithm achieves up to 11.72% of data access latency reduction.

---

11.5.2 OPTIMIZED BUFFER ALLOCATION IN MULTICORE PLATFORMS

**Speakers:** Maximilian Odenthali, Andres Goens; Rainer Leupers, Gerd Ascheid; Benjamin Res; Berthold Vöcking; and Tomas Henriksson

1RWTH Aachen University, DE; 2Huawei Technologies, SE

**Abstract**

With the availability of advanced MPSoC and emerging Dynamic RAM (DRAM) interface technologies, an optimal allocation of logical data buffers to physical memory cannot be handled manually anymore due to the huge design space. An allocation does not only need to decide between an on- or off-chip memory, but also needs to take an increasing number of available memory channels, different bandwidth capacities and several routing possibilities into account. We formalize this problem and introduce a Mixed Integer Linear Programming (MILP) model based on two different optimization criteria. We implement the MILP model into a retargetable tool and present a case study with representative data of the Long-Term-Evolution (LTE) standard to show the real-life applicability of our approach.

---

11.5.3 MEMORY-CONSTRAINED STATIC RATE-OPTIMAL SCHEDULING OF SYNCHRONOUS DATAFLOW GRAPHS VIA RETIMING

**Speakers:** Xue-Yang Zhu; Marc Geilen, Twan Bastens and Sander Stuijk

1State Key Laboratory of Computer Science, Institute of Software, Chinese Academy of Sciences, CN; 2Department of Electrical Engineering, Eindhoven University of Technology, NL; 3Department of Electrical Engineering, Eindhoven University of Technology, Embedded Systems Institute, NL

**Abstract**

Synchronous dataflow graphs (SDFGs) are widely used to model digital signal processing and streaming media applications. In this paper, we use retiming to optimize SDFGs to achieve a high throughput with low storage requirement. Using a memory constraint as an additional enabling condition, we define a memory constrained self-timed execution of an SDPG. Exploring the state-space generated by the execution, we can check whether a retiming exists that leads to a rate-optimal schedule under the memory constraint. Combining this with a binary search strategy, we present a heuristic method to find a proper retiming and a static scheduling which schedules the retimed SDFG with optimal rate (i.e., maximal throughput) and with as little storage space as possible. Our experiments are carried out on hundreds of synthetic SDFGs and several models of real applications. Differential synthetic graph results and real application results show that, in 78% of the tested models, our method leads to a retimed SDFG whose rate-optimal schedule requires less storage space than the proven minimal storage requirement of the original graph, and in 20% of the cases, the returned storage requirements equal the minimal ones. The average improvement is about 7.3%. The results also show that our method is computationally efficient.

---

11.5.4 A CONSTRAINT-BASED DESIGN SPACE EXPLORATION FRAMEWORK FOR REAL-TIME APPLICATIONS ON MPSoCs

**Speakers:** Kathrin Rosvall and Ingo Sander, KTH Royal Institute of Technology, SE

**Abstract**

Design space exploration (DSE) is a critical step in the design process of real-time multiprocessor systems. Combining a formal base in form of SDF graphs with predictable platforms providing guaranteed QoS, the paper proposes a flexible and extendable DSE framework that can provide performance guarantees for multiple applications implemented on a shared platform. The DSE framework is formulated in a declarative style as interprocess communication-aware constraint programming (CP) model. Apart from mapping and scheduling of application graphs, the model supports design constraints on several cost and performance metrics, as e.g. memory consumption and achievable throughput. Using constraints with different compliance level, the framework introduces support for mixed criticality in the CP model. The potential of the approach is demonstrated by means of experiments using a Sobel filter, a SUSAN filter, a RASTA-PLP application and a JPEG encoder.
Thermal hot spots and unbalanced temperatures between cores on chip can cause either degradation in performance or may have a severe impact on reliability, or both. In this paper, we propose mDTM, a proactive dynamic thermal management technique for on-chip systems. It employs multi-objective management for migrating tasks in order to both prevent the system from hitting an undesirable thermal threshold and to balance the temperatures between the cores. Our evaluation on the Intel SCC platform shows that mDTM can successfully avoid a given thermal threshold and reduce spatial thermal variation by 22%. Compared to state-of-the-art, our mDTM achieves up to 58% performance gain. Additionally, we deploy an FPGA and IR camera based setup to analyze the effectiveness of our technique.
<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
</table>
| 15:00 | 11.6.3 | THERMAL MANAGEMENT OF BATTERIES USING A HYBRID SUPERCAPACITOR ARCHITECTURE | Donghwa Shin\(^1\), Massimo Ponchio\(^1\) and Enrico Macii\(^1\)
\(^1\)Department of Computer Engineering, Yeungnam University, KR; \(^1\)Politecnico di Torino, IT; \(^1\)Dipartimento di Automatica e Informatica, Politecnico di Torino, IT |

**Abstract**

Thermal analysis and management of batteries have been an important research issue for battery-operated systems such as electric vehicles and mobile devices. Nowadays, battery packs are designed considering heat dissipation, and external cooling devices such as a cooling fan are also widely used to ensure the reliability and extend the lifetime of a battery. This type of approaches that target the enhancement of the cooling efficiency via the reduction of the thermal resistance cannot achieve an immediate temperature drop to avoid a thermal emergency situation. Approaches based on removing the heat from the heat sources via idle period insertion (similar to what is done for silicon devices) would allow faster thermal response; however it is not obvious how to implement these schemes in the context of batteries. In this paper, we propose the use of a simple parallel battery-supercapacitor hybrid architecture with a dual-mode discharging strategy that can provide immediate temperature management, in which the supercapacitor is used as an energy buffer during the idle periods of the battery. Simulation results shows that the proposed method can keep the battery temperature within the safe range without external cooling devices while exploiting the advantage of the battery-supercapacitor parallel connection.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
</table>
| 15:31 | IP5-17, 873 | THERMAL ANALYSIS AND MODEL IDENTIFICATION TECHNIQUES FOR A LOGIC + WIDEIO STACKED DRAM TEST CHIP | Francesco Beneventi\(^1\), Andrea Bartolini\(^1\), Pascal Vivel\(^2\), Denis Dutot\(^2\) and Luca Benni\(^2\)
\(^1\)DEI - University of Bologna, IT; \(^2\)CEA-Leti, Grenoble, FR |

**Abstract**

High temperature is one of the limiting factors and major concerns in 3D-chip integration. In this paper we use a 3D test chip (WIDEIO DRAM on top of a logic die) equipped with temperature sensors and heaters to explore thermal effects. We correlated real temperature measurements with the power dissipated by the heaters using model learning techniques. The resulting compact thermal model is able to predict temperatures at chip locations far from the temperature sensors and to infer the power dissipation at any location of the chip. Results are verified by using an off-sample validation technique and show a high accuracy of the compact thermal model when compared with silicon measurements.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
</table>
| 15:32 | IP5-18, 349 | ADAPTIVE POWER ALLOCATION FOR MANY-CORE SYSTEMS INSPIRED FROM MULTIAGENT AUCTION MODEL | Xiaohang Wang\(^1\), Baoxin Zhao\(^1\), Tommaso Makz\(^2\), Mei Yang\(^1\), Yingtao Jiang\(^1\), Masoud Daneshfalab\(^2\) and Maurizio Palesi
\(^1\)Guangzhou Institute of Advanced Technology, CN; \(^2\)The Chinese University of Hong Kong, CN; \(^3\)University of Nevada, Las Vegas, US; \(^4\)University of Turku, FI; \(^5\)University of Enna, Kore, IT |

**Abstract**

Scaling of future many-core chips is hindered by the challenge imposed by ever-escalating power consumption. At its worst, an increasing fraction of the chips will have to be shut down, as power supply is inadequate to simultaneously switch all the transistors. This so-called dark silicon problem brings up a critical issue regarding how to achieve the maximum performance with a given limited power budget. This issue is further complicated by two facts. First, high variation in power budget calls for wide range power control capability, whereas most current frequency/voltage scaling techniques cannot effectively adjust power over such a wide range. Second, as the applications' behavior becomes more complicated, there is a pressing need for scalability and global coordination, rendering heuristic-based centralized or fully distributed control schemes inefficient. To address the aforementioned problems, in this paper, a power allocation method employing multiagent auction models is proposed, referred as Hierarchal MultiAgent based Power allocation (HiMAP). Ties act the role of consumers to bid for power budget and the whole process is modeled by a combinatorial auction, whereas HiMAP finds the Walrasian equilibrium. Experimental results have confirmed that HiMAP can reduce the execution time by as much as 45% compared to three competing methods. The runtime overhead and cost of HiMAP are also small, which makes it suitable for adaptive power allocation in many-core systems.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
</table>
| 15:33 | IP5-19, 815 | UNIFIED, ULTRA COMPACT, QUADRATIC POWER PROXIES FOR MULTI-CORE PROCESSORS | Muhammad Yasin\(^1\), Ibrahim (Abe) Elfadel\(^1\) and Anas Shahnurz\(^2\)
\(^1\)New York University - Abu Dhabi, AE; \(^2\)Masdar Institute of Science and Technology, AE |

**Abstract**

Per-core power proxies for multi-core processors are known to use several dozens of hardware activity monitors to achieve a 2% accuracy on core power estimation. These activity monitors are typically not accessible to the user, and even if they were accessible, there would be a significant overhead in using them at the kernel or OS level for power monitoring or control. Furthermore, when scaled up to hundreds of cores per chip, such power proxies become a computational bottleneck for power management operations such as chip power capping. In this paper, we show that a 4% accuracy or better for per-core power estimation can be achieved using an ultra compact power proxy based on a hybrid set of only four user-accessible parameters, namely core frequency, core temperature, instruction-per-cycle and active state residency. Our proxy is nonlinear, valid across all P and C states, and is based on a randomized power data collection strategy that aims at exercising all the P and C levels of each core. We illustrate the accuracy of the model using the full suite of the SPEC CPU 2006 benchmarks on a 12-core processor.

15:30 End of session

Coffee Break in Exhibition Area

On Tuesday-Thursday the coffee and lunch breaks will be located in the Exhibition Area (Terrace Level).

### 11.7 Power and Emerging Technologies in Reconfigurable Computing

**Date:** Thursday, March 27, 2014

**Time:** 14:00 - 15:30

**Location / Room:** Konferenz Z 5

**Chair:** Diana Goehringer, Ruhr-University Bochum (RUB), DE; Contact Diana Goehringer

**Co-Chair:** Fabrizio Ferrandi, Politecnico di Milano, IT; Contact Fabrizio Ferrandi

The first two papers in this session propose new architectures that take advantage of emerging nonvolatile memory technologies. The third paper proposes a battery cell aware task partitioning and mapping to maximize battery runtime.
14:00 11.7.1 EXPLOITING STT-NV TECHNOLOGY FOR RECONFIGURABLE, HIGH PERFORMANCE, LOW POWER, AND LOW TEMPERATURE FUNCTIONAL UNIT DESIGN

Speakers:
Adarsh Reddy1, Hamid Mahmoodi1 and Houman Homayoun1
1George Mason University, US; 2San Francisco State University, US

Abstract
Unavailability of functional units and their unequal activity makes performance bottlenecks and thermal hot spot units in general-purpose processors. We propose to use reconfigurable functional units to overcome these challenges. A selected set of complex functional units that might be under-utilized, such as a multiplier and divider, are realized in a time multiplexed fashion using a shared programmable Look Up Table (LUT) based fabric. This allows for run-time reconfiguration and migration of their activity. LUT-based implementation also allows under-utilized functional units to be dynamically reconfigured to the functional units that have a performance bottleneck and hence improving performance. The programmable LUTs are realized using Spin Transfer Torque (STT) Magnetic technology (also called STT-NV) due to its zero leakage and CMOS compatibility. The results show significant performance improvement of 16% on average across standard benchmarks, when replacing CMOS multiplier and divider with reconfigurable STT-NV LUT counterpart. In addition, reconfiguration reduces the maximum temperature of functional units by up to 27°C and almost eliminates the thermal variation across them. This comes with small power overhead and no area impact.

14:30 11.7.2 A POWER-EFFICIENT RECONFIGURABLE ARCHITECTURE USING PCM CONFIGURATION TECHNOLOGY

Speakers:
Ali Ahmadi1, Hossein Asadi1, Behnam Khaledi1 and Mehdi Tahoori1
1Sharif University of Technology, IR; 2Karlsruhe Institute of Technology, DE

Abstract
Promising advantages offered by resistive Non-Volatile Memories (NVMs) have brought great attention to replace existing volatile memory technologies. While NVMs were primarily used to be used in the memory hierarchy, they can also provide benefits in Field-Programmable Gate Arrays (FPGAs). One major limitation of employing NVMs in FPGAs is significant power and area overheads imposed by the Peripheral Circuitry (PC) of NVM configuration bits. In this paper, we investigate the applicability of different NVM technologies for configuration bits of FPGAs and propose a power-efficient reconfigurable architecture based on Phase Change Memory (PCM). The proposed PCM-based architecture has been evaluated using different technology nodes and is compared to the SRAM-based FPGA architecture. Power and Power-Delay Product (PDP) estimations of the proposed architecture show up to 37.7% and 35.7% improvements over SRAM-based FPGAs, respectively, with less than 3.2% performance overhead.

15:00 11.7.3 EXTENDING LIFETIME OF BATTERY-POWERED COARSE-GRAINED RECONFIGURABLE COMPUTING PLATFORMS

Speakers:
Shouyi Yin1, Peng Ouyang1, Leibo Liu1 and Shaojun Wei1
1Tsinghua University, CN; Institute of Microelectronics and The National Lab for Information Science and Technology, Tsinghua University, CN

Abstract
The coarse-grained reconfigurable architecture (CGRA) is a promising platform for mobile computing. In this paper, how to prolong the lifetime of battery-powered reconfigurable computing platform is addressed. Considering the nonlinear characteristics of battery, a multi-objective optimization model is built for extending the lifetime of battery. Based on this model, a joint task-battery scheduling algorithm is proposed. The experimental results show that the proposed method achieves 26.22% improvement on battery runtime averagely comparing to the state-of-the-art methods.

15:30 IP5-20, 659 3D FPGA USING HIGH-DENSITY INTERCONNECT MONOLITHIC INTEGRATION

Speakers:
Oguz Turkylmaz1, Gerald Cibrario2, Olivier Rozeau1, Perme Batude1 and Fabien Clermont1
1CEA-LETI, MINATEC Campus, FR; 2CEA, FR

Abstract
New 3D technology, called “MonoLithic Integration”, offers very dense 3D interconnect capabilities. In this paper, we propose a 3D FPGA architecture with logic-on-memory approach based on this technology. The routing and computation blocks are split into two layers where the logic is placed on the top and memory on the bottom. Using extracted values from layout in 14nm FDSOI technology, typical benchmark circuits are evaluated in the VPR toolflow. The results show an area reduction of 55% compared to the 2D FPGA. More importantly, due to the lowered routing congestion, the EDP of the 3D FPGA is improved by 47%.

15:32 IP5-21, 526 JOINT COMMUNICATION SCHEDULING AND INTERCONNECT SYNTHESIS FOR FPGA-BASED MANY-CORE SYSTEMS

Speakers:
Alessandro Cilardo, Edoardo Fusella, Luca Gallo and Antonino Mazzeo, University of Naples Federico II, IT

Abstract
This work proposes an automated methodology for optimizing FPGA-based many-core interconnect architectures. Based on the application communication requirements, the methodology concurrently defines the architecture of the interconnect and the communication task scheduling, taking into account possible dependencies between tasks under given area constraints. The resulting architecture improves the level of communication parallelism that can be exploited while keeping area costs low. The paper thoroughly describes the proposed approach and discusses a few case studies showing the impact of the proposed technique.

15:33 IP5-22, 688 A NOVEL EMBEDDED SYSTEM FOR VISION TRACKING

Speakers:
Antonis Nikitakis1, Theofilos Paganos1 and Ioannis Papaefstathiou2
1Technical University of Crete, Department of Electronic and Computer Engineering Kounoupidiana, Chania, Crete, GR73100, Greece, GR; 2iSynetix Solutions Ltd, Farmakidou 10, Chalkida, GR34100, Greece, GR

Abstract
One of the most important challenges in the field of Computer Vision is the implementation of low-power embedded systems that will execute very accurate, yet real-time, algorithms. In the visual tracking sector one of the most promising approaches is the recently introduced OpenTLD algorithm which uses a random forest classification method. While it is very robust, it cannot be efficiently parallelized in its native form as its memory access pattern has certain characteristics that make it hard to take advantage of the conventional memory hierarchies. In this paper, we present a novel embedded system implementing this algorithm. We accelerate the bottleneck of the algorithm by designing and implementing a high bandwidth distributed memory sub-system which is independent of the various software parameters. We demonstrate the applicability and efficiency of this novel approach by implementing our scheme in a modern FPGA.
The embedded tutorial aims at providing with an updated view on what GPGPUs can provide not only in terms of performance and power, but also in terms of reliability, and how the latter can be evaluated and possibly improved.

**Time | Label | Presentation Title**
--- | --- | ---
14:00 | 11.8.1 | RELIABILITY REQUIREMENTS FOR GPUS IN HPC

**Speakers:**
Nathan A. DeBardeleben¹, Leonardo Bautista Gomez² and Franck Cappello²

¹Los Alamos National Laboratory, US; ²Argonne National Laboratory, US

---
14:30 | 11.8.2 | GPU RELIABILITY ASSESSMENT AND ENHANCEMENT

**Speakers:**
Paolo Rech¹, Luigi Carro¹ and Steve Keckler²

¹UFRGS, BR; ²NVIDIA, US

---
15:00 | 11.8.3 | EVALUATING THE ROBUSTNESS OF GPU APPLICATIONS THROUGH FAULT INJECTION

**Speakers:**
Karthik Pattabiraman¹, Bo Fang¹ and Sudhanva Gurumurthi²

¹UBC, CA; ²AMD, US

---
15:30 | End of session | Coffee Break in Exhibition Area

On Tuesday-Thursday the coffee and lunch breaks will be located in the Exhibition Area (Terrace Level).

**UB11 Session 11**

**Date:** Thursday, March 27, 2014

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

**Location / Room:** University Booth, Booth 3, Exhibition Area

---

**Label | Presentation Title**
--- | ---
UB11.01 | CYCLOSE: DESIGNING CLOUD-BASED SELF-HEALING CYBER-PHYSICAL SYSTEMS

**Authors:**
Giulio Gambardella¹, Silviu Folea², Mihai Hulea², Liviu Miclea², George Mois², Teodora Sanisie², Marco Indaco¹, Paolo Prinetto¹, Daniele Rolfo¹ and Pascal Trotta¹

¹Politecnico di Torino, IT; ²Universitatea Tehnica din Cluj-Napoca Departamentul de Automatica, RO

**Abstract**
Cyber-Physical Systems (CPSs) are a new generation of systems capable to represent more than networking and information technology, information and knowledge being integrated into physical objects. These type of systems are physical and engineered systems whose actions are monitored, controlled, and integrated by a computing and communication kernel. The Cyclose project aims at developing: (1) an infrastructure for designing self-healing Cyber-Physical Systems (CPSs) using cloud computing technology; (2) an experimental model for CPSs using wireless sensor networks (WSNs) for data acquisition, reliable hardware components based on reconfigurable devices - Field Programmable Gate Arrays (FPGAs) and cloud computing technology to store, manage and analyse data in a large context.

More information ...
UB11.02 BRIDGING MATLAB/SIMULINK AND ESL DESIGN VIA AUTOMATIC CODE GENERATION

Authors: Liyuan Zhang, Michael Glaß and Jürgen Teich, University of Erlangen-Nuremberg, DE

Abstract

MATLAB/Simulink is today’s de-facto standard for model-based design in domains such as control engineering and signal processing. Commercial tools are available to generate embedded C or HDL code directly from a Simulink model. However, Simulink models are purely functional models and, hence, designers cannot seamlessly consider the architecture that a Simulink model is later implemented on. In particular, it is not possible to explore the different architectural alternatives and investigate the arising interactions and side-effects directly within Simulink. To benefit from MATLAB/Simulink’s algorithm exploration capabilities and overcome the outlined drawbacks, we introduce a model transformation framework that converts a Simulink model to an executable specification, written in an actor-oriented modeling language. This specification then serves as the input of an established Electronic System Level (ESL) design flow, enabling Design Space Exploration (DSE) and automatic code generation for both hardware and software. In this demonstration, we will show how to automatically transform Simulink models to an established ESL design flow by means of a code generator. Based on the generated code, we present a co-simulation approach that combines complex environmental models from MATLAB/Simulink with the auto-generated model of a controller. We will use an Anti-lock Braking System (ABS) as an example where we investigate the impact of different controller implementations in the automotive E/E architecture. In detail, the following scientific achievements are included in the proposed demonstration: To bridge Simulink and ESL design flows, we developed an ESL code generator to perform model transformation. The idea is that for any given Simulink models such as a controller in a control system, the designer can simply invoke our code generator to create the ESL model automatically. In our design flow, we use SystemC as a programming language with an extension of actors with a specific Model of Computation (MoC). We guarantee the preservation of the semantics of the generated model by (a) applying a specific 1-to-1 mapping from Simulink basic blocks to an actor library and (b) considering different transformations to capture single-rate and multi-rate Simulink models. After the model transformation is finished, this auto-generated SystemC model serves as the input of a well-established ESL design flow that enables DSE. Besides the Code-Generator we demonstrate also a validation technique that considers the functional correctness by comparing the original Simulink model with the generated SystemC model. The main idea behind this technique is (1) to co-simulate the auto-generated model along with the the original model and (2) to reuse the environment model and the test bench that are originally created in Simulink also for the auto-generated model. Furthermore, the performance of the model can also be measured during co-simulation. In this demonstration, an ABS model will be transformed from Simulink to SystemC by invoking ESL code generator. Then, by applying our validation technique, the correctness and the accuracy of the auto-generated model can be examined. Lastly, to evaluate the performance of the model, application-dependent quality of control will be measured, such as the braking distance on an icy road.

More information ...

UB11.03 BICONDITIONAL BINARY DECISION DIAGRAM MANIPULATION PACKAGE

Authors: Luca Amauri, Alexis Balatsoukas-Stimming, Pierre-Emmanuel Gaillardon, Andreas Burg and Giovanni De Micheli, EPFL, CH; EPFL-TCL, CH; EPFL-LSI, CH

Abstract

In this software demonstration, we present a logic manipulation package based on Biconditional Binary Decision Diagrams (BBDDs). BBDDs are a novel class of canonical binary decision diagrams where the branching condition, and its associated logic expansion, is biconditional on two variables. We show how Verilog files from real life designs can be rapidly read and processed by the BBDD manipulation package, for verification, testing or synthesis purposes. In particular, we demonstrate the benefit deriving from BBDD rewriting of arithmetic circuits in the synthesis of a product code iterative decoder.

More information ...

UB11.04 HEROES^2: A SYSTEM FRAMEWORK FOR MODELING, SIMULATION AND TESTING OF HETEROGENEOUS SOFTWARE-INTENSIVE SYSTEMS

Authors: Markus Becker, Wolfgang Mueller, Ulrich Köffler and Joachim Stroop, University of Paderborn/C-LAB, DE; dSPACE GmbH, DE

Abstract

HEROES^2 is a SystemC framework for modeling/simulation of heterogeneous SW-intensive systems. It has 8 abstraction levels for co-refinement of application/environment models from continuous/discrete models to networked embedded SW stacks. Support of various SW/comm. abstractions is achieved by combining AMS MoCs, TLM, HiSi models (MM, RTOS, HAL) and QEMU user mode/system emulator. Interfacing with a commercial AUTOSAR toolchain is provided, i.e., code generators, integration and experimentation tools.

More information ...

UB11.05 RESCY: RESOURCE-AWARE COMPUTER VISION APPLICATION ON HETEROGENEOUS MULTI-TILE ARCHITECTURE

Authors: Erikis Susač, Johny Pauli, Vahid Lari, Frank Hanning, Jürgen Teich and Walter Stechele, University of Erlangen-Nuremberg, DE; Technische Universität München, DE

Abstract

We demonstrate the benefits of invasive computing by showing the efficiency and utilization improvements in a resource-aware manner by algorithmic selection of different invasive resources, such as TCPA (traffic-conditioned processor array), and RISC processors. More specific we present a dynamic load balancing of a computer vision application between multiple RISC cores and a TCPA, based on invasive mechanisms supported by our operating system and the agent system.

More information ...

UB11.06 SKETCH-BASED ESL VIRTUAL PROTOTYPING: SKETCH-BASED DESIGN AND SIMULATION-BASED EVALUATION FOR ESL VIRTUAL PROTOTYPING

Authors: Rafael Rssales, Michael Glaß, Jürgen Teich, Bo Wang, Yang Xu and Ralph Hasholzner, University of Erlangen-Nuremberg, DE; Zinti! Mobile Communications, DE

Abstract

Virtual prototyping and Electronic System Level (ESL) modeling have become valuable approaches to cope with the ever-increasing complexity of embedded systems. Their effectiveness, however, is highly dependent on their quick development time and accuracy both conflicting goals. In this demonstration, we present (a) an ESL methodology [1][2] for the simulation-based evaluation of power and performance of embedded systems by the use of virtual prototyping. Our methodology permits us to develop ESL models for design space exploration of dynamic power and performance management strategies and hardware/software co-design choices. (b) We present a novel sketch-based tool termed Mahler [3] for the very early design phase of ESL modeling. Mahler provides a playground to quickly model functionality and evaluate performance on different architecture implementations. In Mahler, ESL models are created by literally sketching with a pen or touch interface, e.g. a tablet stylus, or a touchless interface, such as a Leap Motion controller. The application and architecture models are transformed to an executable virtual prototype through sketch recognition. This approach provides a very intuitive and fast way to explore actor-oriented functional modeling and hardware/software partitioning. The output of Mahler is a simulation-ready SystemC-based source-code stub that can be refined for subsequent design iterations. We will show a model of a Voice over LTE (VoLTE) use case, consisting of a heterogeneous cellular SoC platform, together with a wireless channel fading model and a base station network model. State-based [1] and polynomial-equation-based [4] power models are built and co-simulated for the SoC digital module and the RF transceiver module, respectively to abstract their different power consumption characterization accurately. The entire end-to-end modeling enables efficient SoC performance and power simulation with proper network configuration in seconds, which is highly desired in cellular system early design exploration phase and co-optimization with network vendors.

More information ...
UB11.07 VERIFIC-MM

Authors: Christoph Kuznik and Wolfgang Müller, University of Paderborn, DE

Abstract

Verific-MM is an approach to systematize and accelerate the coverage plan engineering as well as the verification environment's (functional) metric code generation — usually a time-consuming and error-prone task — in particular by (i) improving automation via assisted model-based approaches, utilizing recent industry standards such as UCIS and (ii) a supporting methodology suitable for various target (functional coverage) languages (IEEE-1800 SystemVerilog, IEEE-1647 e, IEEE-1666 SystemC).

More information ...

16:30 End of session

IP5 Interactive Presentations

Date: Thursday, March 27, 2014
Time: 15:30 - 16:00
Location / Room: Conference Level, foyer

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>
</tr>
</thead>
<tbody>
<tr>
<td>IP5-1</td>
<td>HYBRID WIRE-SURFACE WAVE ARCHITECTURE FOR ONE-TO-MANY COMMUNICATION IN NETWORK-ON-CHIP</td>
</tr>
<tr>
<td>IP5-2</td>
<td>FAILURE ANALYSIS OF A NETWORK-ON-CHIP FOR REAL-TIME MIXED-CRITICAL SYSTEMS</td>
</tr>
<tr>
<td>IP5-3</td>
<td>COOLIP: SIMPLE YET EFFECTIVE JOB ALLOCATION FOR DISTRIBUTED THERMALLY-THROTTLED PROCESSORS</td>
</tr>
<tr>
<td>IP5-4</td>
<td>ENERGY OPTIMIZATION IN 3D MPSOCs WITH WIDE-I/O DRAM USING TEMPERATURE VARIATION AWARE BANK-WISE REFRESH</td>
</tr>
</tbody>
</table>

More information ...
EFFICIENT SIMULATION AND MODELLING OF NON-RECTANGULAR NOC TOPOLOGIES

 Speakers: Ji-Gi and Mark Zwiolski, University of Southampton, GB

 Abstract: With increasing chip complexity, Networks-on-Chips (NoCs) are becoming a central platform for future on-chip communications. Many regular NoC architectures have been proposed to eliminate the communication bottlenecks of traditional bus-based networks. Non-rectangular and irregular architectures have also been proposed to increase performance. However, the complexity of designing custom non-rectangular networks leads to a rapid increase in design and verification times. To alleviate the conflict between performance and efficiency, this paper proposes a novel method that efficiently constructs virtual non-rectangular topologies on a mesh network by using time-regulated models to emulate irregular patterns. Data routings on virtual hexagonal and two irregular geometries validate the proposed method. An MPEG-4 decoder is used to exemplify the proposed method for media applications. Results analysis shows the virtual topologies emulated by the proposed method can provide precise timing and energy performance.

MOVING FROM CO-SIMULATION TO SIMULATION FOR EFFECTIVE SMART SYSTEMS DESIGN

 Speakers: Franco Fummi1, Michele Lora2, Francesco Stefanini, Dimitrios Trachani, Jan Vanhese and Sara Vinco2

1University of Verona, EDALab s.r.l., IT; 2University of Verona, IT; 3EDALab s.r.l., IT; 4Agilent Technologies, BE

 Abstract: Design of smart systems needs to cover a wide variety of domains, ranging from analogue to digital, with power devices, micro-sensors and actuators, up to MEMS. This high level of heterogeneity makes design a very challenging task, as each domain is supported by specific languages, modeling formalisms and simulation frameworks. A major issue is furthermore posed by simulation, that heavily impacts the design and verification loop and that is very hard to be built in such an heterogeneous context. On the other hand, achieving efficient simulation would indeed make smart system design feasible with respect to budget constraints. This work provides a formalization of the typical abstraction levels and design domains of a smart system. This taxonomy allows to identify a precise role in the design flow for co-simulation and simulation scenarios. Moreover, a methodology is proposed to move from the co-simulated heterogeneity to a simulatable homogeneous representation in C++ of the entire smart system. The impact of heterogeneous or homogeneous models of computation is also examined. Experimental results prove the effectiveness of the proposed C++ generation for reaching high-speed simulation.

AUTOMATING DATA REUSE IN HIGH-LEVEL SYNTHESIS

 Speakers: Wim Meeus1 and Dirk Stroobandt2

1imec and Ghent University, BE; 2Ghent University, BE

 Abstract: Current High-Level Synthesis (HLS) tools perform excellently for the synthesis of computation kernels, but they often don’t optimize memory bandwidth. As memory access is a bottleneck in many algorithms, the performance of the generated circuit will benefit substantially from memory access optimization. In this paper we present an automated method and a toolchain to detect reuse of array data in loop nests and to build hardware that exploits this data reuse. This saves memory bandwidth and improves circuit performance. We make use of the polyhedral representation of the source program, which makes our method computationally easy. Our software complements the existing HLS flows. Starting from a loop nest written in C, our tool generates a reuse buffer and a loop controller, and preprocesses the loop body for synthesis with an existing HLS tool. Our automated tool produces designs from optimized source code that are as efficient as those generated by a commercial HLS tool from manually-optimized source code.

A UNIVERSAL SYMMETRY DETECTION ALGORITHM

 Speaker: Peter Maurer, Dept. of Computer Sci., Baylor University, US

 Abstract: Research on symmetry detection focuses on identifying and detecting new types of symmetry. We present an algorithm that is capable of detecting any type of permutation-based symmetry, including many types for which there are no existing algorithms. General symmetry detection is library-based, but symmetries that can be parameterized, (i.e. total, partial, rotational, and dihedral symmetry), can be detected without using libraries. In many cases it is faster than existing techniques. Furthermore, it is simpler than most existing techniques, and can easily be incorporated into existing software.

OPTIMIZATION OF DESIGN COMPLEXITY IN TIME-MULTIPLEXED CONSTANT MULTIPLICATIONS

 Speakers: Levent Aksoy1, Paulo Flores2 and Jose Monteiro3

1INESC-ID, PT; 2INESC-ID/IST, Lisboa, PT; 3INESC-ID / IST, Lisboa, PT

 Abstract: The multiplication of constants by a data input is an essential operation in digital signal processing (DSP) systems. For applications requiring a large number of constant multiplications under stringent hardware constraints, it is generally realized under a folded architecture, where a single constant selected from a set of multiple constants is multiplied by the data input at each time, called time-multiplexed constant multiplication (TMCM). This paper addresses the problem of optimizing the complexity of a TMCM design and introduces an algorithm that finds the least complex TMCM design by sharing the logic operators, i.e., adders, subtractors, adders/subtractors, and multiplexors (MUXes). It includes efficient search methods, yielding better results than existing TMCM algorithms.

HARDWARE PRIMITIVES FOR THE SYNTHESIS OF MULTITHREADED ELASTIC SYSTEMS

 Speakers: Giorgos Dimitrakopoulos1, Seitanidis Ioannis2, Anastasios Psamats1, Konstantinos Tsouris1, Pavlos Mathalaiakis3 and Jordi Cortadella4

1Democritus University of Thrace, GR; 2Democritus University of Thrace, GR; 3Mentor Graphics, FR; 4Universitat Politècnica de Catalunya, ES

 Abstract: Elastic systems operate in a dataflow-like mode using a distributed scalable control and tolerating variable latency computations. At the same time, multithreading increases the utilization of processing units and hides the latency of each operation by time-multiplexing operations of different threads in the datapath. This paper proposes a model to unify multithreading and elasticity. A new multithreaded elastic control protocol is introduced supported by low-cost elastic buffers that minimize the storage requirements without sacrificing performance. To enable the synthesis of multithreaded elastic architectures, new hardware primitives are proposed and utilized in two circuit examples to prove the applicability of the proposed approach.
DCM: AN IP FOR THE AUTONOMOUS CONTROL OF OPTICAL AND ELECTRICAL RECONFIGURABLE NOCS.

Speakers:
Wolfgang Böter, Christof Osewoldt, Daniel Gregorek and Alberto Garcia-Ortiz

Abstract
The increasing requirements for bandwidth and quality-of-service motivate the use of parallel interconnect architectures with several degrees of reconfiguration. This paper presents an IP called Distributed Channel Management (DCM), to extend existing packet-switched NoCs with a reconfigurable point-to-point network seamlessly, i.e., without the need for any modification on the routers. The configuration of the reconfigurable network takes place dynamically and autonomously, so that the topology can be changed at run time. Furthermore, the architecture is scalable due to the autonomous decentralized administration of the links. The Paper reports a thorough experimental analysis of the overhead of the approach at the gate level that considers different network parameters such as flit size and timing constraints.

MINIMALLY BUFFERED SINGLE-CYCLE DEFLECTION ROUTER

Speakers:
Gnaneswar Rao Jonna, John Josei, Rachana Radhakishnan and Madhu Mulyani

Abstract
With the drift from computation centric designs to communication centric designs in the Chip Multi Processor (CMP) era, the interconnect fabric is gaining more importance. An efficient NoC in terms of power, area and average flit latency has a huge impact on the overall performance of a CMP. In the current work, we propose MiniBSD - a minimally buffered, single cycle, deflection router. It incorporates different operations (Injection, Ejection, Preemption, Re-injection) in a single module to handle the traffic effectively and ensures smooth flow of flits through router pipeline. It performs overlapped execution of independent operations. These factors not only make MiniBSD to operate in a single cycle but also reduce the critical path latency in a faster interconnect network. Experimental results show that MiniBSD reduces the average flit latency on real work loads, reduces die area and power consumption when compared to the existing state-of-the-art minimally buffered deflection routers.

FUNCTIONAL TEST GENERATION GUIDED BY STEADY-STATE PROBABILITIES OF ABSTRACT DESIGN

Speakers:
Jian Wang1, Huawei Li1, Tao Lv, Tiancheng Wang1 and Xiaowei Li1

Abstract
This paper presents a novel method for functional test generation aiming at exploring control state space of the design. The steady-state probabilities (SPs) of the abstract design’s control FSM are used to guide test generation. The SPs of the states can reflect how hard the states can be reached, and the hard-to-reach states are assigned with high priority to be exercised. Experimental results show that our method has better performance in test generation in comparison with constrained random simulation, and demonstrate that SPs provide good guidance on traversing hard-to-reach states of the design under validation.

AUTOMATED SYSTEM TESTING USING DYNAMIC AND RESOURCE RESTRICTED CLIENTS

Speakers:
Miiko Caspar, Mikl Lippmann and Wolfram Hardt, Technische Universität Chemnitz, DE

Abstract
Testing on system level using a static and homogeneous architecture of clients is common practice. This paper introduces a new approach to use a heterogeneous and dynamic set of resource restricted test clients for automated testing. Due to changing resources and availability of the clients, the test case distribution needs to be recalculated dynamically during the test execution. All necessary conditions and parameters are represented by a formal model. It is shown that the algorithmic problem of DYNAMIC TESTPARTITIONING can be solved in polynomial time by a heuristic recursive algorithm. A testbench architecture is introduced and by simulation it is shown that the testbench can execute the test requirements within a small variation using a number of several hundred clients. The system can react dynamically on changing resources and availability of the test clients within several seconds. The approach is generic and can be adapted to a huge number of systems.

RELIABILITY-AWARE MAPPING OPTIMIZATION OF MULTI-CORE SYSTEMS WITH MIXED-CRITICALITY

Speakers:
Shin-Haeng Kang1, Hoesek Yong1, Sungchan Kim1, Iuliana Bazivarov1, Soonhoi Ha1 and Lothar Thiele4

Abstract
This paper presents a novel mapping optimization technique for mixed critical multi-core systems with different reliability requirements. For this scope, we derived a quantitative reliability metric and presented a scheduling analysis that certifies given mixed-criticality constraints. Our framework is capable of investigating re-execution, passive replication, and modular redundancy with optimized voter placement, while typical hardening approaches consider only one or two of these techniques. The proposed technique complies with existing safety standards and is power-efficient, as demonstrated by our experiments.

FROM SIMULINK TO NOC-BASED MPSoC ON FPGA

Speakers:
Francesco Robino and Johnny Öberg, KTH Royal Institute of Technology, SE

Abstract
Network-on-chip (NoC) based multi-processor systems are promising candidates for future embedded system platforms. However, because of their complexity, new high level modeling techniques are needed to design, simulate and synthesize embedded systems targeting NoC-based MPSoC. Simulink is a popular modeling environment suitable to model at system level. However, there is no clear standard to synthesize Simulink models into SW and HW towards a NoC-based MPSoC implementation. In addition, many of the proposed solutions require large overhead in terms of SW components and memory requirements, resulting in complex and customized multi-processor platforms. In this paper we present a novel design flow to synthesize Simulink models onto a NoC-based MPSoC running on low-cost FPGAs. Our design flow constrains the MPSoC and the Simulink model to share a common semantics domain. This permits to reduce the need of resource consuming SW components, reducing the memory requirements on the platform. At the same time, performances (throughput) of dataflow applications can increase when the number of processors of the target platform is increased. This is shown through a case study on FPGA.

THERMAL ANALYSIS AND MODEL IDENTIFICATION TECHNIQUES FOR A LOGIC + WIDEIO STACKED DRAM TEST CHIP

Speakers:
Francesco Beneventi1, Andrea Bartolini1, Pascal Vivel1, Denis Dutot1 and Luca Benini1

Abstract
High temperature is one of the limiting factors and major concerns in 3D-chip integration. In this paper we use a 3D test chip (WIDEIO DRAM on top of a logic die) equipped with temperature sensors and heaters to explore thermal effects. We correlated real temperature measurements with the power dissipated by the heaters using model learning techniques. The resulting compact thermal model is able to predict temperatures at chip locations far from the temperature sensors and to infer the power dissipation at any location of the chip. Results are verified by mean of an off-sample validation technique and show a high accuracy of the compact thermal model when compared with silicon measurements.
ADAPTIVE POWER ALLOCATION FOR MANY-CORE SYSTEMS INSPIRED FROM MULTIAGENT AUCTION MODEL

Speakers:
Xiaohang Wang1, Baoxin Zhao1, Terrence Mak2, Mei Yang1, Yingtao Jiang1, Masoud Daneshfalabakhsh2 and Maurizio Palesti2
1Guangzhou Institute of Advanced Technology, CN; 2The Chinese University of Hong Kong, CN; 3University of Nevada, Las Vegas, US; 4University of Turku, FI; 5University of Enna, Italy

Abstract
Scaling of future many-core chips is hindered by the challenge imposed by ever-escalating power consumption. At its worst, an increasing fraction of the chips will have to be shut down, as power supply is inadequate to simultaneously switch all the transistors. This so-called dark silicon problem brings up a critical issue regarding how to achieve the maximum performance with a given limited power budget. This issue is further complicated by two facts. First, high variation in power budget calls for wide range power control capability, whereas most current frequency/voltage scaling techniques cannot effectively adjust power over such a wide range. Second, as the applications' behavior becomes more complicated, there is a pressing need for scalability and global coordination, rendering heuristic-based centralized or fully distributed control schemes inefficient. To address the aforementioned problems, in this paper, a power allocation method employing multiagent auction models is proposed, referred as Hierarchical MultiAgent based Power allocation (HiMAP). Ties act the role of consumers to bid for power budget and the whole process is modeled by a combinatorial auction, whereas HiMAP finds the Walrasian equilibria. Experimental results have confirmed that HiMAP can reduce the execution time by as much as 45% compared to three competing methods. The runtime overhead and cost of HiMAP are also small, which makes it suitable for adaptive power allocation in many-core systems.

UNIFIED, ULTRA COMPACT, QUADRATIC POWER PROXIES FOR MULTI-CORE PROCESSORS

Speakers:
Muhammad Yasini1, Ibahim (Abe) Elfadl2 and Anas Shahrouz2
1New York University - Abu Dhabi, AE; 2Masdar Institute of Science and Technology, AE

Abstract
Per-core power proxies for multi-core processors are known to use several dozens of hardware activity monitors to achieve a 2% accuracy on core power estimation. These activity monitors are typically not accessible to the user, and even if they were accessible, there would be a significant overhead in using them at the kernel or OS level for power monitoring or control. Furthermore, when scaled up to hundreds of cores per chip, such power proxies become a computational bottleneck for power management operations such as chip power capping. In this paper, we show that a 4% accuracy or better for per-core power estimation can be achieved using an ultra compact power proxy based on a hybrid set of only four user-accessible parameters, namely core frequency, core temperature, instruction-per-cycle and active state residency. Our proxy is nonlinear, valid across all P and C states, and is based on a randomized power data collection strategy that aims at exercising all the P and C levels of each core. We illustrate the accuracy of the model using the full suite of the SPEC CPU 2006 benchmarks on a 12-core processor.

3D FPGA USING HIGH-DENSITY INTERCONNECT MONOLITHIC INTEGRATION

Speakers:
Ogun Turkylmez1, Gerald Gibrizo2, Olivier Rezaeul2, Penline Baudieu1 and Fabien Clermidy2
1ICEA-LETI, Minatec Campus, FR; 2ICAE, FR; 2ICEA-LETI, FR

Abstract
New 3D technology, called "Monolithic Integration", offers very dense 3D interconnect capabilities. In this paper, we propose a 3D FPGA architecture with logic-on-memory approach based on this technology. The routing and computation blocks are split into two layers where the logic is placed on the top and memory on the bottom. Using extracted values from layout in 14nm FDSOI technology, typical benchmark circuits are evaluated in the VPR5 tool flow. The results show an area reduction of 55% compared to the 2D FPGA. More importantly, due to the lowered routing congestion, the EDP of the 3D FPGA is improved by 47%.

JOINT COMMUNICATION SCHEDULING AND INTERCONNECT SYNTHESIS FOR FPGA-BASED MANY-CORE SYSTEMS

Speakers:
Alessandro Ciarotto, Edordano Fusella, Luca Gallo and Antonino Mazzeo, University of Naples Federico II, Italy

Abstract
This work proposes an automated methodology for optimizing FPGA-based many-core interconnect architectures. Based on the application communication requirements, the methodology concurrently defines the structure of the interconnect and the communication task scheduling, taking into account possible dependencies between tasks under given area constraints. The resulting architecture improves the level of communication parallelism that can be exploited while keeping area costs low. The paper thoroughly describes the proposed approach and discusses a few case studies showing the impact of the proposed technique.

A NOVEL EMBEDDED SYSTEM FOR VISION TRACKING

Speakers:
Antonis Nikitakis1, Theofilos Paganos1 and Ioannis Papaefstathiou2
1Technical University of Crete, Department of Electronic and Computer Engineering Kounoupidiana, Chania, Crete, Greece, GR; 2Synelixis Solutions Ltd, Farmakidou 10, Chaskida, GR34100, Greece, GR

Abstract
One of the most important challenges in the field of Computer Vision is the implementation of low-power embedded systems that will execute very accurate, yet real-time, algorithms. In the visual tracking sector one of the most promising approaches is the recently introduced OpenTLD algorithm which uses a random forest classification method. While it is very accurate, implementing our scheme in a modern FPGA.

12.1 SPECIAL DAY Hot Topic: The future of interfacing to the natural world

Date: Thursday, March 27, 2014
Time: 16:00 - 17:30
Location / Room: Saal 1

Organisers:
Ian O'Connor, Lyon Institute of Nanotechnology, FR; Contact Ian O'Connor
Thomas Mikolajick, NanLab gGmbH, DE; Contact Thomas Mikolajick

Chair:
Michael Huebner, Ruhr Universitaet Bochum, DE; Contact Michael Huebner

Co-Chair:
Ian O'Connor, Lyon Institute of Nanotechnology, FR; Contact Ian O'Connor

Challenges for acquiring and processing data from the real world includes the development of interfaces capable of extracting relevant information from massive sensor networks or from living
organisms, sitting through the wealth of data to arrive systematically at a meaningful conclusion, and building hardware platforms suited to carry out these operations in an energy-efficient way.

The first paper in this session looks at the necessarily complex processing of chemical information with hardware components that are capable of responding to various chemical conditions. Interfaces to living organisms are examined in the second paper, which discusses challenges and approaches for efficient detection of disease. In the third paper, novel hardware devices and architectures are explored for use in energy-efficient video analysis applications such as movement detection and face recognition. The fourth paper discusses handling of complex data with large-scale GPU-based recurrent networks, exploiting specific features of the data to improve energy efficiency.

### 12.2 Hot topic: How Secure are PUFs Really? On the Reach and Limits of Recent PUF Attacks

**Date:** Thursday, March 27, 2014  
**Time:** 16:00 - 17:30  
**Location / Room:** Konferenz 6

**Organiser:**  
Ulrich Rührmair, TU München, DE; Contact Ulrich Rührmair

**Chair:**  
Ulf Schlichtmann, TU München, DE; Contact Ulf Schlichtmann

PUFs are an emerging and promising security primitive. However, some strong attacks on their core security features have been reported recently, for example on their unclonability. We discuss the reach, but also the limits of these attacks, aiming at a well-balanced treatment, and also evaluate the future perspectives of the field.

<table>
<thead>
<tr>
<th>Time</th>
<th>Label</th>
<th>Presentation Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>16:00</td>
<td>12.2.1</td>
<td>INTEGRATED CIRCUITS PROCESSING CHEMICAL INFORMATION: PROSPECTS AND CHALLENGES</td>
<td>Andreas Richter, Axel Voigt, René Schüffny, Stephan Henker and Marcus Völpi, Technische Universität Dresden, DE</td>
</tr>
<tr>
<td>16:00</td>
<td>12.2.2</td>
<td>INTERFACE TO LIVING CELLS</td>
<td>Rudy Lauwereins, IMEC, BE</td>
</tr>
<tr>
<td>16:00</td>
<td>12.2.3</td>
<td>VIDEO ANALYTICS USING BEYOND CMOS DEVICES</td>
<td>Vijaykrishnan Narayanan1, Gert Gauwenberghs, Donald Chiarulli, Suman Datta, Steve Levitan and Philip Wengs</td>
</tr>
<tr>
<td>16:00</td>
<td>12.2.4</td>
<td>ENERGY EFFICIENT NEURAL NETWORKS FOR BIG DATA ANALYTICS</td>
<td>Wang Yu, Boxun Li, Rong Luo, Yiran Chen, Ningyi Xu and Huazhong Yang, Tsinghua University, CN</td>
</tr>
</tbody>
</table>

---

**Abstract**

The unbelievable properties of our information processing capabilities regarding the processing of big data, resilience, and energy efficiency are inspiration sources for the optimization and the rethinking of the principles of electronic information processing. Here, we present an approach of integrated circuits intended to solve chemical problems by active processing of chemical information.
### 12.2.2 PUF MODELING ATTACKS: AN INTRODUCTION AND OVERVIEW

**Speakers:**
Utecht Röhrmair and Jan Söltzer  
TU München, DE; Freie Universität Berlin, DE

**Abstract**
Machine learning (ML) based modeling attacks are the currently most relevant and effective attack form for so-called Strong Physical Unclonable Functions (Strong PUFs). We provide an overview of this method in this paper: We discuss the basic conditions under which it is applicable; the ML algorithms that have been used in this context; the latest and most advanced results on simulated and silicon data; the right interpretation of existing results; and possible future research directions.

### 12.2.3 HYBRID SIDE-CHANNEL / MACHINE-LEARNING ATTACKS ON PUFs: A NEW THREAT?

**Speakers:**
Xiaolin Xu and Wayne Burleson, Umass, Amherst, US

**Abstract**
Machine Learning (ML) is a well-studied strategy in modeling Physical Unclonable Functions (PUFs) but reaches its limits while applied on instances of high complexity. To address this issue, side-channel attack is combined to help reduce the computational workload of ML modeling attacks and make it more applicable. In this work, we present the currently known hybrid side-channel attacks on PUFs. A taxonomy is proposed based on the characteristics of different side-channel attacks. The practical reach of some published side channel attacks is discussed. Both challenges and opportunities for PUF attackers are introduced. Countermeasures against some certain side-channel attacks are also analyzed. To better understand the side-channel attacks on PUFs, three different methodologies of implementing side-channel attacks are compared. At the end of this paper, we bring forward some open problems for this research area.

### 12.2.4 PHYSICAL VULNERABILITIES OF PHYSICALLY UNCLONABLE FUNCTIONS

**Speakers:**
Clemens Helfmeier, Dmitry Nedospasov, Shahin Tajik, Christian Bolt and Jean-Pierre Seifert, Technische Universität Berlin, DE

**Abstract**
In recent years one of the most popular areas of research in hardware security has been Physically Unclonable Functions (PUF). PUFs provide primitives for implementing tamper detection, encryption and device fingerprinting. One particularly common application is replacing Non-volatile Memory (NVM) as key storage in embedded devices like smart cards and secure microcontrollers. Though a wide array of PUF have been demonstrated in the academic literature, vendors have only begun to roll out PUFs in their end-user products. Moreover, the improvement to overall system security provided by PUFs is still the subject of much debate. This work reviews the state of the art of PUFs in general, and as a replacement for key storage in particular. We review also techniques and methodologies which make the physical response characterization and physical/digital cloning of PUFs possible.

### 12.2.5 PROTOCOL ATTACKS ON ADVANCED PUF PROTOCOLS AND COUNTERMEASURES

**Speakers:**
Marten van Dijk and Ulrich Röhrmair  
University of Connecticut, US; TU München, DE

**Abstract**
In recent years, PUF-based schemes have not only been suggested for the basic security tasks of tamper sensitive key storage or system identification, but also for more complex cryptographic protocols like oblivious transfer (OT), bit commitment (BC), or key exchange (KE). These more complex protocols are secure against adversaries in the stand-alone, good PUF model. In this survey, a shortened version of [17], we explain the stronger bad PUF model and PUF re-use model. We argue why these stronger attack models are realistic, and that existing protocols, if used in practice, will need to face these. One consequence is that the design of advanced cryptographic PUF protocols needs to be strongly reconsidered. It suggests that Strong PUFs require additional hardware properties in order to be broadly usable in such protocols: Firstly, they should ideally be erasable, meaning that single PUF-responses can be erased without affecting other responses. If the area efficient implementation of this feature turns out to be difficult, new forms of Controlled PUFs [3] (such as Logically Erasable and Logically Reconfigurable PUFs [6]) may suffice in certain applications. Secondly, PUFs should be certifiable, meaning that one can verify that the PUF has been produced faithfully and has not been manipulated in any way afterwards. The combined implementation of these features represents a pressing and challenging problem for the PUF hardware community.

### 12.2.6 QUO VADIS, PUF? TRENDS AND CHALLENGES OF EMERGING PHYSICAL-DISORDER BASED SECURITY

**Speakers:**
Masoud Rostami, Farinaz Koushanfar, James Wendt and Modrag Potkonjak
Rice University, US; UCLA, US

**Abstract**
Physical unclonable Function (PUF) has emerged as a popular and widely studied security primitive based on the randomness of the underlying physical medium. To date, most of the research emphasis has been placed on finding new ways to measure randomness, hardware realization and analysis of a few initially proposed structures, and conventional secret-key based protocols. In this work, we suggest our subjective analysis of the emerging and future trends in this area that aim to change the scope, widen the application domain, and make lasting impact. We emphasize on development of new PUF-based primitives and paradigms, robust protocols, public-key protocols, digital PUF, new technologies, implementation, metrics and tests for evaluation/validation, as well as relevant attacks and countermeasures.

### 12.3 Multimedia Systems

**Date:** Thursday, March 27, 2014  
**Time:** 16:00 - 17:30  
**Location / Room:** Konferenz 1

**Chair:**  
Theocharides Theocharis, University of Cyprus, CY, Contact Theocharis Theocharides

**Co-Chair:**  
Cristiana Bolchini, Politecnico di Milano, IT, Contact Cristiana Bolchini

The session presents designs for energy efficient and flexible implementations of advanced video coders or image acquisition/processing systems.
**16:00** 12.3.1 **FLEXIBLE AND SCALABLE IMPLEMENTATION OF H.264/AVC ENCODER FOR MULTIPLE RESOLUTIONS USING ASIPS**

**Speakers:**
Hong Chinh Doan, Haris Javadi and Parameswaran, School of Computer Science and Engineering, University of New South Wales, Sydney, Australia, AU

**Abstract**
Real-time encoding of video streams is computationally intensive and rarely carried out at high resolutions. In this paper, for the first time, we propose a platform for H.264 encoder which is both flexible (allows software upgrades) and scalable (supports multiple resolutions), and supports high video quality (by using both intraprediction and interprediction) and high throughput (by exploiting slice-level and pixel-level parallelism). Our platform uses multiple Application Specific Instruction Set Processors (ASIPs) with local and shared memories, and hardware accelerators (in the form of custom instructions). Our platform can be configured to use a particular number of ASIPs (slices per video frame) for a specific video resolution at design-time. The MPSoC architecture is automatically generated by our platform and the H.264 software does not need any modification, which enables quick design space exploration. We implemented the proposed platform in a commercial design environment, and illustrated its utility by creating systems with up to 170 ASIPs supporting resolutions up to HD1080. We further show how power gating can be used in our platform to save energy consumption.

**16:30** 12.3.2 **A FLEXIBLE ASIP ARCHITECTURE FOR CONNECTED COMPONENTS LABELING IN EMBEDDED VISION APPLICATIONS**

**Speakers:**
Juan Fernando Usse1, Rainer Laupers1, Gerd Ascheid1, Patrick Sudowe1, Bastian Labe1 and Tamon Sadasue2

1RWTH Aachen University, DE; 2RICOH Company LTD., JP

**Abstract**
Real-time identification of connected regions of pixels in large (e.g. FullHD) frames is a mandatory and expensive step in many computer vision applications that are becoming increasingly popular in embedded mobile devices such as smartphones, tablets and head mounted devices. Standard off-the-shelf embedded processors are not yet able to cope with the performance/flexibility trade-offs required by such applications. Therefore, in this work we present an Application Specific Instruction Set Processor (ASIP) tailored to concurrently execute thresholding, connected components labeling and basic feature extraction of image frames. The proposed architecture is capable to cope with frame complexities ranging from QCIF to FullHD frames with 1 to 4 bytes-per-pixel formats, while achieving an average frame rate of 30 frames-per-second (fps). Synthesis was performed for a standard 65nm CMOS library, obtaining an operating frequency of 350MHz and 2.1mm² area. Moreover, evaluations were conducted both on typical and synthetic data sets, in order to thoroughly assess the achievable performance. Finally, an entire planar-marker based augmented reality application was developed and simulated for the ASIP.

**17:00** 12.3.3 **IMAGE PROGRESSIVE ACQUISITION FOR HARDWARE SYSTEMS**

**Speakers:**
Jianxiong Liu, Christos Bouganis and Peter Y.K. Cheung, Imperial College London, GB

**Abstract**
As the resolution of digital images increases, accessing raw image data from memory has become a major consideration during the design of image/video processing systems. This is due to the fact that the bandwidth requirement and energy consumption of such image accessing process has increased. Inspired by the successful application of progressive image sampling techniques in many image processing tasks, this work proposes to apply similar concepts within hardware systems to efficiently trade image quality for reduced memory bandwidth requirement and lower energy consumption. Based on this idea, a hardware system is proposed that is placed between the memory subsystem and the processing core of the design. The proposed system alters the conventional memory access pattern to progressively and adaptively access pixels from a target memory external to the system. The sampled pixels are used to reconstruct an approximation to the ground truth, which is stored in an internal image buffer for further processing. The system is prototyped on FPGA and its performance evaluation shows that a saving of up to 85% of memory accessing time and 33%/45% of image acquisition time/energy is achieved on the benchmark image “lena” while maintaining a PSNR of about 30 dB.

**17:15** 12.3.4 **HIGH-QUALITY REAL-TIME HARDWARE STEREO MATCHING BASED ON GUIDED IMAGE FILTERING**

**Speakers:**
Christos Totsis and Theocarhs Theocharides, University of Cyprus, CY

**Abstract**
Stereo matching is a vital task in several emerging embedded vision applications requiring high-quality depth computation and real-time frame-rate. Although several stereo matching dedicated hardware systems have been proposed in recent years, only few of them focus on balancing accuracy and speed. This paper proposes a hardware-based stereo matching architecture that aims to provide high accuracy and concurrently high performance in embedded vision applications. The proposed architecture integrates a compact and efficient design of the recently proposed guided image filter, an edge-preserving filter that reduces the hardware complexity of the implemented stereo algorithm, while at the same time maintains high-quality results. A prototype of the architecture has been implemented on a Kintex-7 FPGA board, achieving 60 fps for 720p resolution images. Moreover, the proposed design delivers leading accuracy when compared to state-of-the-art hardware implementations.

**17:30** End of session

### 12.4 Physical Aspects

**Date:** Thursday, March 27, 2014

**Time:** 16:00 - 17:30

**Location / Room:** Konferenz 2

**Chair:**
Carl Sechen, University of Texas at Dallas, US, Contact Carl Sechen

**Co-Chair:**
Jens Lienig, Technische Universität Dresden, DE, Contact Jens Lienig

This session focuses on contemporary issues in physical design. The first paper concerns detailed placement for sub-20nm technologies. Then pattern matching for more efficient hotspot detection is introduced. Finally, a flow for minimizing the number of wiring layers on multichip interposers is presented.
12.4.1 Optimizing Standard Cell Based Detailed Placement for 16 nm FinFET Process

**Speakers:**
Yue Lin and Martin D. F. Wong, University of Illinois at Urbana-Champaign, US

**Abstract**
FinFET transistors have great advantages over traditional planar MOSFET transistors in high performance and low power applications. Major foundries are adopting the FinFET technology for CMOS semiconductor device fabrication in the 16 nm technology node and beyond. Edge device degradation is among the major challenges for the FinFET process. To avoid such degradation, dummy gates are needed on device edges, and the dummy gates have to be tied to power rails in order not to introduce unconnected parasitic transistors. This requires that each dummy gate must abut at least one source node after standard cell placement. If the drain nodes at two adjacent cell boundaries abut each other, additional source nodes must be inserted in between for dummy gate power tying, which costs more placement area. Usually there is some flexibility during detailed placement to horizontally flip the cells or switch the positions of adjacent cells, which has little impact on the global placement objectives, such as timing conditions and net congestion. This paper proposes a detailed placement optimization strategy for the standard cell based designs. By flipping a subset of cells in a standard cell row and switching pairs of adjacent cells, the number of drain to drain abutments between adjacent cell boundaries can be optimally minimized, which saves additional source node insertion and reduces the length of the standard cell row. In addition, the proposed graph model can be easily modified to consider more complicated design rules. The experimental results show that the optimization of 100k cells is completed within 0.1 second, verifying the efficiency of the proposed algorithm.

12.4.2 Signature Indexing of Design Layouts for Hotspot Detection

**Speakers:**
Cristian Andrades1, Andrea Rodriguez2, and Charles Chiango2

1Universidad de Concepcion, CL; 2Synopsys Inc., US

**Abstract**
This work presents a new signature for 2D spatial configurations that is useful for the optimization of a hotspot detection process. The signature is a string of numbers representing changes along the horizontal and vertical slices of a configuration, which serves as the key of an inverted index that groups layout windows with the same signature. The method extracts signatures from a compact specification of similar exact patterns with a fixed size. Then, these signatures are used as search keys of the inverted index to retrieve candidate windows that can match the pattern. Experimental results show that this simple type of signature has 100% recall and, in average, over 85% of precision in terms of the area effectively covered by the pattern and the retrieved area of the layout. In addition, the signature shows a good discriminate quality, since around 99% of the extracted signatures match each of them with a single pattern.

12.4.3 Metal Layer Planning for Silicon Interposers with Consideration of Routability and Manufacturing Cost

**Speakers:**
Wen-Hao Liu, Tzu-Kai Chien, and Ting-Chi Wang, National Tsing Hua University, TW

**Abstract**
A 2.5D IC provides a silicon interposer to integrate multiple dies into a package, which not only offers better performance than 2D ICs but also has lower manufacturing complexity than true 3D ICs. In an interposer, routing wires connect signals between dies or route signals from dies to the package substrate. The number of metal layers in an interposer is one of the critical factors to affect the routability and manufacturing cost of the 2.5D IC. Thus, how to achieve 100% routing completion rate in an interposer using a minimum number of metal layers plays a key role for the success of a 2.5D IC. This paper presents a global-routing-based metal layer planner called VGR to identify a minimal number of metal layers for an interposer with consideration of routability and manufacturing cost. Also, VGR can identify a good stacking order of the horizontal and vertical layers in an interposer such that the routing solution in the interposer costs fewer vias. To our best knowledge, this paper is the first study to solve the metal layer planning problem for silicon interposers.

17.00 End of Session
This session addresses the trade-off between accuracy and power consumption and the management of multi-core/multi systems. The power management is addressed at several abstraction levels from circuit and performance counters up to the system level (operating system).
Applications from several important domains exhibit intrinsic resilience to approximations or inexactness in their underlying computations. Approximate circuits are commonly used to realize highly efficient hardware implementations of such applications. A wide range of manual and automated techniques for the design of approximate circuits have been proposed. However, all of them target combinational circuits, leaving a gap between these techniques and the natural granularity at which quality is specified. In practice, the designer is concerned with quality or accuracy at the output of a sequential circuit after several cycles of computation, and not at the output of an embedded combinational block. We propose ASLAN (Automatic methodology for Sequential Logic ApproximateN), the first effort towards the synthesis of approximate sequential circuits. Given an RTL or gate-level description of a sequential circuit and a quality constraint at its output, ASLAN automatically synthesizes an approximate version that guarantees the specified quality bound. The key challenges in approximating sequential circuits are (i) to model how errors due to approximations are generated, propagate through multiple cycles of operation, and eventually impact quality of the final output, and (ii) to select the most beneficial approximations, i.e., those that result in higher energy savings for smaller impact on output quality. We address the first challenge by constructing a virtual Sequential Quality Constraint Circuit (SQC) utilizing formal verification to ensure that a given approximation satisfies the quality constraint during synthesis. To address the second challenge, we identify combinational blocks in the sequential circuit that are amenable to approximation, generate local quality-energy trade-off curves for them, and use a gradient-descent approach to iteratively approximate the sequential circuit. We used ASLAN to automatically synthesize approximate versions for several sequential benchmarks (DCT, FIR, IIR, etc.). Our experiments demonstrate energy reductions of 1.20X-2.44X for tight error constraints, and 1.32X-4.42X for relaxed error constraints. We also present case studies of using the approximate circuits generated by ASLAN in two well known applications — MPEG Encoding and K-Means Clustering. We obtain energy savings of 1.32X with 0.5% average degradation in PSNR for the MPEG Encoder and 1.26X with 0.8% quality loss in case of KMeans Clustering.
In general, we would like to determine the importance of system-level verification and its unique needs—whether generators, checking, coverage, or teams. Does it exist? Is it sufficient to extend UVM capabilities (e.g., SystemC, TLM) or are dedicated tools and methodology needed? Are formal methods ready to play a significant role in SoC-level verification methodology? Is the gap in today's SoC verification methodology significant? Is it growing? Or perhaps it does not exist? What is the right way to close the gap, if one exists?

Today, common practice for verification is based on Universal Verification Methodology (UVM), which, at the system level, is built on reusing and combining unit-level environments, followed by growing complexity of SoCs, which support diverse IO interfaces and devices, and have complex architectures including multiple heterogeneous cores, multi-level caches, and multiple IO bridges. Smartphones are the most notable example of this, but we are also seeing this with TV chips, in-car controllers, network routers, and more. This trend is occurring in parallel to the constantly increasing complexity of SoCs and the need for high performance, low power, and low cost.

SoCs are becoming more similar to servers. Many SoCs today are no longer tied to a single application and look more like general purpose PCs and high-end servers. This is a recent trend that SoCs are becoming more similar to servers. Many SoCs today are no longer tied to a single application and look more like general purpose PCs and high-end servers. This is a recent trend that SoCs are becoming more similar to servers.
Panelists:

- Lyes Benalycherif, STMicroelectronics, FR
- Franco Fummi, University of Verona, IT
- Alan J. Hu, University of British Columbia, Vancouver, CA
- Ronny Morad, IBM Research - Haifa, IL
- Frank Schirmeister, Cadence Design Systems, US

17:30 End of session