System-to-Silicon Performance Modeling and Analysis

Power, Temperature and Reliability

DAC 2015 Workshop
June 7, 2015 - San Francisco, CA, USA

Preliminary Program

Rationale

The integration of heterogeneous electronic systems composed of SW and HW requires not only a proper handling of system functionality, but also an appropriate expression and analysis of various extra-functional properties: timing, energy consumption, thermal behavior, reliability, cost and others as well as performance aspects related to caching, non-determinism, probabilistic effects.

The workshop addresses cross-domain aspects related to the design and verification framework covering methodology, interoperable tools, flows, interfaces and standards that enable formalization, specification, annotation and refinement of functional and extra-functional properties of a system. Special emphasis will be given to formalization and expression of ***power, temperature, reliability, degradation and aging***.

Several research and industry efforts address (parts of) the problem. However, there is a need for community-wide cooperation to establish a holistic vision on extra-functional property treatment, and to agree on research and development directions and further on validation of applicable solutions and standardization.

This event will support collaboration between main actors from system and microelectronics industry, EDA and research.

The workshop is inviting submissions of short abstracts industrial and scientific work in progress and practical solution and experiences.

Workshop Organizers:

- Laurent Maillet-Contoz, STMicroelectronics, France
- Kim Grüttner, OFFIS, Germany
- Gjalt de Jong, ArchWorks, Belgium
- Adam Morawiec, ECSI, France

Session Organizers:

- Andreas Herkersdorf, TU München, Germany
- Jürgen Becker, KIT Karlsruhe, Germany
- Domenik Helms, OFFIS, Germany
- Christoph Sohrmann, Fraunhofer Institute for Integrated Circuits IIS, Germany
- Roland Jancke, Fraunhofer Institute for Integrated Circuits IIS, Germany

Preliminary Agenda

Workshop Program		
9:00	**Intro**	**Welcome & Agenda** *Adam Morawiec (ECSI)*
9:05	**Keynote**	**An Accurate Simulation Framework for Thermal Explorations and Optimizations** *William Fornaciari (PoliMi, Italy)*
9:50	**Session 1**	**System-Level Design for Reliability** Organizers: *Andreas Herkersdorf (TU München, Germany),* *Jürgen Becker (KIT Karlsruhe, Germany)* Abstract: Reliability is a system-level concern, both from the hardware/software architecture as well as design method perspectives. Advanced nanometer CMOS technologies are known to be increasingly vulnerable for radiation induced sporadic soft-errors, device aging and various forms of manufacturing and environmental variations. Another source of reliability exposures for today's and future Systems-on-Chip (SoC) solutions is their inherent complexity, expressed either in Billions of transistors, number of IP cores integrated, and the variety of huge functionality implemented on a single SoC. Today, it is already practically infeasible to validate such SoCs down to clock cycle accuracy under various representative workload scenarios. Both, feature size and complexity induced challenges cannot be addressed at individual, specific abstraction layers with acceptable quality and cost. Senior university and industry researchers from the US and Europe will share their perspectives on crucial design aspects of today's and future embedded and cyber physical systems. Topics span from dependable NoC communication virtualization on MPSoC, to self-aware Cyber Physical Systems-on-Chip, to high-throughput database query acceleration on reconfigurable FPGAs with High-Level Synthesis, to reliability management of 3D stacked wireless baseband SoCs.
9:50	**1.1**	**A Cross Layer Approach for Efficient Reliability Management in 3D Stacked Wireless Baseband SoCs** *Norbert Wehn (Microelectronic System Design Research Group, University of Kaiserslautern, Germany)*
10:15	**1.2**	**Using Roofline Models to Analyze the Performance of Realistic Key Value Store Implementations on FPGAs with High Level Synthesis** *Kees Vissers and Michaela Blott (Xilinx, USA and Ireland)*
10:40	**1.3**	**CyberPhysical-System-On-Chip (CPSoC): A Self-Aware SoC Platform for Cross-Layer Reliability** *Nikil Dutt (Center for Embedded and CyberPhysical Systems, UC Irvine)*
11:05		Coffee Break
11:20	**1.4**	**Intel Euro Labs Presentation (TBC)** *Enno Lübbers, Intel Euro Labs, München, Germany*
11:45	**1.5**	**Enabling Dependable MPSoC Task Migration with On-Chip Interconnect Virtualization** *Andreas Herkersdorf (Integrated Systems Lab, Technische Universität München, Germany)*
12:10	**1.6**	**Dynamic Migration and Performance Optimization of Deterministic Applications Across Platform Components Using Intel® CoFluent™ Studio** *Jérôme Lemaitre, Rocco Le Moigne (Intel Corporation SAS, France)*
12:35		Lunch Break
13:15	**Keynote**	**Towards Parallel Simulation of Multi-Domain System Models** *Rainer Dömer, UC Irvine, USA*
14:00	**Session 2**	**Tools and Methods for Power and Temperature Modeling and Analysis** Organizers: *Kim Grüttner (OFFIS, Germany), Domenik Helms (OFFIS, Germany)* *Laurent Maillet-Contoz (STMicroelectronics, France)* Abstract: With the predicted device, core and multicore scaling, the dark silicon hypothesis predicts the end of multicore scaling, regardless of chip organization and topology, due to power or energy density limitations. For this reason, future system engineers should be able to address power and thermal management as soon as possible in the design flow. Introduction of power and temperature management cannot be done at a single abstraction layer, but must be taken into consideration from the operating system, early system-level models, down to the integration of RTL IP components. For this reason, power and temperature properties need to be modelled across all abstraction layers,

		because they can strongly affect the products overall quality of service or even cause the system to fail meeting its real-time and safety requirements. In this session will discuss breakthrough academic and industrial solutions to control power consumption and heat dissipation at design and run-time. Furthermore, this session will discuss proposed extensions of existing industrial standards and their implications on commercial tool support. The addressed topics are: operating system support for fine-grained system-level energy analysis through orchestration of energy measurements at hardware level; the extensions of IP-XACT and UPF industry standards to support a seamless ESL to RTL low power design methodology; the extension of IP-XACT with verification features including portable stimuli vectors to enable performance and power closure of complex SoCs; tools for architectural level power and thermal modeling; and a new thermal constrained run-time management called "Thermal Safe Power".
14:00	**2.1**	**The FIGAROS Operating System Kernel for Fine-Grained System-Level Energy Analysis** *Timo Hönig, Heiko Janker, Wolfgang Schröder-Preikschat (Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany)*
14:25	**2.2**	**Extending IP-XACT and UPF to Support ESL to RTL Low Power Design Methodology** *Emmanuel Vaumorin, Grégoire Avot (Magillem Design Services; France), Hend Affes, Michel Auguin, Alain Pégatoquet, François Verdier (Univ. Nice Sophia Antipolis, France)*
14:50		Coffee Break
15:05	**2.3**	**IP Configuration for the Right Balance Between Required Performance and Power** *Nick Heaton (Cadence Design Systems, USA), Simon Rance (ARM, UK)*
15:30	**2.4**	**Architectural Level Modeling with Power and Thermal Models** *Sylvian Kaiser (Docea Power, France)*
15:55	**2.5**	**Thermal-Aware Power Budgeting for Dark Silicon Chips** *Santiago Pagani, Muhammad Shafique (Karlsruhe Institute of Technology, Germany), Jian-Jia Cheny, Jörg Henkel (TU Dortmund, Germany)*
16:20	**Session 3**	**Ageing and Variation Prediction from Transistor to RT Level** Organizers: *Christoph Sohrmann & Roland Jancke* *(Fraunhofer Institute for Integrated Circuits IIS, Germany)* Abstract: Even in safety-critical areas such as Automotive, Aviation, Medical, and Industrial the demand for applications having highest performance, lowest energy consumption, and smallest dimensions together with extended service life grows rapidly. Combinations of these requirements can only be provided by extremely scaled technologies. Such safety-critical applications do not tolerate device failure. However, devices from advanced technology nodes are generally more susceptible to parametric deviations, either from process variations, parametric drift over lifetime or a combination thereof. The correct prediction of parametric deviations is similarly important as making them accessible on higher abstraction levels. The focus of this session are thus models and formats which abstract detailed knowledge about parameter variations and reliability from the device level to the circuit or RT level. The solutions presented will enable the designer to take these effects into account early in the design phase and ensure to meet specifications over the entire lifetime and for all application conditions.
16:20	**3.1**	**Impact of Time-dependent Variability on the Yield and Performance of 6T SRAM Cells in an Advanced HK/MG Technology** *Pieter Weckx, Ben Kaczer, Praveen Raghavan, Francky Catthoor, Guido Groeseneken (IMEC, Belgium)*
16:45	**3.2**	**Facilitating Cross-Layer Reliability Management through Universal Reliability Information Exchange** *Enrico Costenaro[1], Domenik Helms[2], Nematollah Bidokhti[3], Adrian Evans[1], Maximilian Glorieux[1] and Dan Alexandrescu[1] ([1]IROC Technologies, France; [2]OFFIS, Germany; [3]OCZ Toshiba, USA)*
17:10	**3.3**	**Statistical Timing Methodology for Low-Power and Multi-Voltage Designs** *Kerim Kalafala, Natesan Venkateswaran, Stephen Shuma, Vladimir Zolotov, Eric A Foreman (IBM Thomas J. Watson Research Center, USA)*
17:35	**3.4**	**Reliability-Driven Analog Circuit Design using gm/Id Method and Cross Layer Modelling of Aging** *Steffen Paul (University Bremen, Germany)*
18:00		**Concluding Remarks & Closing** *Adam Morawiec (ECSI)*

Presentation Abstracts

Keynote 1: An Accurate Simulation Framework for Thermal Explorations and Optimizations

Speaker William Fornaciari, PoliMi, Italy

Abstract:
While technology scaling allows integrating more cores in the same chip, the complexity of current designs requires accurate and fast techniques to explore different trade-offs. Moreover, the increased power densities in current architectures highlight thermal issues as a first class design metric to be addressed. At the same time, the need to access to accurate models for the exploited actuators is of paramount importance, since their overheads can shadow the benefit of the proposed methodologies. This talk proposes a complete simulation framework for the assessment of run-time policies for thermal-performance and power-performance trade-offs optimization with two main improvements over the state of the art. First, it accurately models Dynamic Voltage and Frequency Scaling (DVFS) modules for both cores and NoC routers as well as a complete Globally Asynchronous Locally Synchronous (GALS) design paradigm and power gating support for crossabar and buffers in the NoC. Second, it accounts for the chip thermal dynamics as well as power and performance overheads for the actuators.

Presentation 1.1: A Cross Layer Approach for efficient Reliability Management in 3D Stacked Wireless Baseband SoCs

Speaker: Norbert Wehn, Microelectronic System Design Research Group, University of Kaiserslautern

Abstract:
3D stacking of silicon dies via Through Silicon Vias (TSVs) is an emerging technology to increase performance, energy efficiency and integration density of today's and future System-on-Chips (SoCs). Especially the stacking of Wide I/O DRAMs on top of a logic die is a very promising approach to tackle the memory wall and energy efficiency challenge. The potential of this type of stacking is currently under investigations by many research groups and companies in particular for mobile devices. There, the baseband processing and the application processor are implemented on a logic die. On top of this die one or several Wide I/O DRAMs are stacked. An example of such a SoC is the Wioming 3D Magali chip. However, new challenges emerge especially related to reliability issues.
In this talk we present a cross layer reliability approach for 3D stacked wireless baseband SoCs. Our approach efficiently exploits the probabilistic behaviour of wireless communications. Based on measurements of the Wioming chip we derive a reliability model for Wide I/O DRAMs. We combine the inherent error resilience of communications systems, techniques to mitigate reliability problems on the logic die and the Wide I/O DRAM model to achieve a holistic approach for efficient reliability management.

Presentation 1.2: Using Roofline Models to Analyze the Performance of Realistic Key Value Store Implementations on FPGAs with High Level Synthesis

Speaker: Kees Vissers and Michaela Blott (Xilinx, USA and Ireland)

Abstract:
We will present the results of an actual implementation of Memcached at sustained 10Gbps line rate, with Terabytes of storage. Memcached contains an implementation of a Key – Value store, the basic technology for databases. This implementation has been completely written in C/C++ using Xilinx' Vivado High Level Synthesis. The implementation leverages many memory technologies, including on FPGA BRAM, external DRAM, attached banks of Flash Storage, and external Host memory over PCIe or a Cache Coherent Interface. We will show actual measurements of the implementation, including low-latency, and at least a 38x performance/Watt improvement over the best multicore CPU server implementations. We will introduce the roofline model and analyze the measured performance, under memory Bandwidth limitations, and compute limitations. We will show in detail the benefits of the FPGA implementation and compare this to CPU implementations. We will show the considerations in memory bandwidth and storage for scaling this to guaranteed line rates of 40Gb/s and 100Gb/s with Terabytes of storage on current FPGAs.

Presentation 1.3: CyberPhysical-System-On-Chip (CPSoC): A Self-Aware SoC Platform for Cross-Layer Reliability

Speaker: Nikil Dutt (Center for Embedded and CyberPhysical Systems, UC Irvine)

Abstract:
We present CyberPhysical-Systems-on-Chip (CPSoC), a new class of sensor-actuator rich many-core computing platforms that intrinsically couples on-chip and cross-layer sensing and actuation to enable self-awareness. Unlike traditional MPSoC designs,

CPSoC differs primarily on the co-design of the control, communication, and computing system that interacts with the physical environment in real-time in order to modify the system´s behavior so as to adaptively achieve desired objectives, including reliability and Quality-of-Service (QoS). The CPSoC design paradigm enables self-awareness (i.e., the ability of the system to observe its own internal and external behaviors such that it is capable of making judicious decision) and (opportunistic) adaptation using the concept of cross-layer physical and virtual sensing and actuations applied across different layers of the hardware/software system stack. The closed loop control used for adaptation to dynamic variation – commonly known as observer-decide-act (ODA) loop – is implemented using an adaptive, reflexive middleware layer. The learning abilities of CPSoC provide a unified interface API for sensor and actuator fusion along with the ability to improve autonomy in system management.

Presentation 1.4: Intel Euro Labs Presentation

Speaker: Enno Lübbers, Intel Euro Labs, München, Germany

Presentation 1.5: Enabling Dependable MPSoC Task Migration with On-Chip Interconnect Virtualization

Speaker: Andreas Herkersdorf (Integrated Systems Lab, Technische Universität München, Germany)

Abstract:

Nanometer size CMOS devices are increasingly vulnerable for numerous forms of environmental and manufacturing variability, such as ionizing radiation, transistor aging, and thermal hotspots. These challenges can be tackled at different levels of system abstraction, resulting in different cost, performance and quality of resilience design points.

This talk presents a hardware-assisted on-chip communication virtualization architecture that can be used for QoS-enabled run-time adaptation of communication channels between tasks. Independent of task mapping strategies (duplication, triplication, static or flexible (considering online task migration)), communication channels are seamlessly adapted guaranteeing buffer consistency and in sequence packet/flit flow delivery. Different schemes for channel adaptation (stop & resume, 1:1 protection switching and channel forwarding) are assessed according to latency and complexity overheads. We will show how these virtualization techniques can be reused without additional costs for self-adaptive, run-time instantiation of modular task redundancy (AMR). Without application designer intervention, this ability will improve application resilience under technological, environmental or workload induced challenges.

Presentation 1.6: Dynamic Migration and Performance Optimization of Deterministic Applications Across Platform Components Using Intel® CoFluent™ Studio

Speaker: Jérôme Lemaitre, Rocco Le Moigne (Intel Corporation SAS, France)

Abstract:

Intel Software presents a system-level approach to dynamically optimize the performance of deterministic applications that run in HW/SW multi-processor platforms. This is achieved by simulating transaction-level application models, which include a performance-aware supervisor (PAS), mapped onto platform models that are captured using Intel® CoFluent™ Studio. The PAS model dynamically monitors the evolution of non-functional performance properties (in terms of latency, throughput, resource usage, power, etc…) of application models that run in heterogeneous HW/SW component models. Then the PAS executes at run-time a parameterized performance optimization policy to predict if different mappings of tasks onto HW/SW processing units could lead to better system-level performance. When executing this policy, the PAS also identifies at which point in time the migration of tasks across processors/schedulers should occur to preserve the deterministic behaviour of the application models. Finally, the PAS dynamically stops the execution of those parts of the application models whose mapping must be modified and re-starts the execution of the updated mapping model. In this paper we illustrate how a PAS modifies the mapping of application models onto platform models at run-time when optimizing the performance of a system for low power or for short latencies, without altering the deterministic behaviour of the system.

Keynote 2: Towards Parallel Simulation of Multi-Domain System Models

Speaker: Rainer Dömer, Center for Embedded and Cyber-Physical Systems, UC Irvine, USA

Abstract:

Traditional Discrete Event Simulation (DES) is well-suited for functional validation and performance vs. cost exploration of classic embedded computers, but cannot cope with the additional complexities and extra-functional properties of cyber-physical systems where power consumption, thermal behavior and reliability are critical criteria that also must be considered from the beginning of the design process. Thus, true system simulation must advance to (1) efficiently integrate the dimensions of cyber-physical domains and (2) fully exploit parallel execution to ensure scalability.

In this talk, we outline opportunities for extending cyber-physical system simulation towards additional dimensions that integrate the aspects of energy, temperature, degradation and aging. To cope with the additional complexities, we argue that

advanced Parallel Discrete Event Simulation (PDES) is a key to achieving scalable, fast and accurate system validation and exploration.
As one example, we present an out-of-order PDES technique that can accurately simulate SystemC models with maximum parallelism on many-core platforms. Utilizing a dedicated SystemC compiler with advanced static conflict analysis based on Segment Graphs and a parallel simulator with out-of-order thread scheduling, system models can be evaluated orders of magnitude faster without loss of accuracy.

Presentation 2.1: The FIGAROS Operating System Kernel for Fine-Grained System-Level Energy Analysis

Speaker: Timo Hönig, Heiko Janker, Wolfgang Schröder-Preikschat (Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany)

Abstract:
Energy has become the most important operating resource for computing systems of all sizes—from embedded systems to large-scale high-performance computing systems. However, at system level, engineers remain challenged at efficiently handling energy as first-class operating system resource. The reasons for this are twofold: First, increasingly complex hardware circuits are inherently difficult to model which makes the creation of accurate energy models practically impossible. Second, available energy measurements at system level are coarse-grained and they are insufficient for fine-grained system level energy measurements of the operating system.
The current advent of power constrained many-core systems and the road ahead towards the era of dark silicon requires efficient energy control mechanisms in the system software layer. In this paper, we present FIGAROS, an operating system kernel which implements primitives required for fine-grained system-level energy analysis. Our implementation of FIGAROS orchestrates energy measurements at hardware level by a low-level system software infrastructure at kernel level.

Presentation 2.2 Extending IP-XACT and UPF to Support ESL to RTL Low Power Design Methodology

Speaker : Emmanuel Vaumorin, Grégoire Avot (Magillem Design Services; France), Hend Affes, Michel Auguin, Alain Pégatoquet, François Verdier (Université Nice Sophia Antipolis, France)

Abstract:
Power consumption is a major concern in system-to-silicon design flows for embedded systems. Today, in addition to functional simulation/synthesis features (using e.g. Verilog, SystemC), EDA tools support IP-based and power-oriented design whose specifications can be expressed in advanced standards such as IP-XACT or UPF. IP-XACT is mainly focused on the structural description of IPs and hardware architectures while UPF is devoted to express a power intent. These standards are mainly dedicated to design flows covering the scope from RT level to netlist level although these standards can also be considered at system level. Indeed, extensions of UPF are currently under investigation to address higher abstraction levels (working groups P2415 and P2416) [1] and it has been shown for example in [2] that IP-XACT can be used as an underlying format to support ESL and mixed ESL-RTL design flow.
However, as the IP-XACT-based description of a hardware architecture and the UPF-based description of a power intent for that architecture are two separate specifications, it may be challenging to ensure that the power intent specification is really conformal with the IP-XACT structural functional specification. To enhance the coherence between these descriptions we propose to use Platform Assembly tools from Magillem i), to describe in IP-XACT both the IP-based architecture and the structural view of the power intent (power domains and related components expressed as power attributes) and ii), to generate the UPF-based description of the power intent from the IP-XACT specification. Notice that a hardware implementation of a power intent is a structural assembly of power-oriented components, thus this implementation can be represented using IP-XACT with the support of the vendor extensionsmechanism provided by this standard. This approach is mainly suitable for RTL-based design since UPF is not yet extended to support ESL specification. Furthermore, a complete design specification of a low power system should describe the structural hardware architecture, the power intent, as well as the clock intent (clock tree) whatever the ESL or RTL abstract level considered.
Even if a subset of UPF makes sense at abstract level [3], currently UPF 2.1 is not able to describe a clock intent at RTL, and a fortiori at ESL. To be able to handle both power intent and clock intent at ESL, we propose to extend UPF (and partially IP-XACT) with clock-oriented features such as clock domain, source clock, generated clock and so on. To do so, we have identified the components, and their related behavior, required to describe a clock intent at ESL and which are relevant to enrich a SystemC-TLM virtual architecture simulation model with abstract control mechanisms of clock domains (and power domains).
Our aim is then to apply the same approach as outlined above for the RTL case: using Platform Assembly tools i) to describe in IP-XACT (at ESL) the hardware architecture, a clock intent, a power intent and ii), to generate an extended UPF description of the low power part as well as a full SystemC-TLM simulation model of the complete system. We will show that the proposed extensions in IP-XACT for power intent may not only used for further UPF generation: their graphical representation through multiple virtual hierarchies in Magillem Platform Assembly provides a great assistance in all design phases.

Considering IP-XACT at both ESL and RTL as the underlying format to describe the structural view of the functional architecture and the clock/power intent enables power architecture refinements through a metadata driven methodology based on a common backbone thus enhancing verification and requirements tracing capabilities. An example of an audio hardware platform TLM model extended with a power intent and a clock intent is used to illustrate the approach.

Presentation 2.3 IP Configuration for the Right Balance Between Required Performance and Power

Speaker: Nick Heaton (Cadence Design Systems, USA), Simon Rance (ARM, UK)

Abstract:

Configurable IPs bring many benefits; to the IP vendor: targeting multiple markets, lowering maintenance and support costs, and to the customer: flexibility, reuse and the ability to fine tune the IP's configuration, finding the right balance between required performance and power budget. The downside to configurability is the need to create many similar, yet sufficiently different, test benches for the verification and performance analysis of the IP in question. This is especially true when running tight 'what if' cycles, where you run the same traffic vectors across multiple IP configurations, analyzing and comparing the results to identify the best configuration for your needs.

As configurable IPs are generated by tools and are often described by meta data standards like IP-XACT, it stands to reason to use these same descriptions when automating the process of testbench generation. This paper will review the collaboration between ARM, a leading provider of configurable System IPs and with the acquisition of Duolog in 2014, a strong proponent of IP---XACT, and Cadence, the Verification IP market leader. Key findings include:

• The advantages of configurable IPs

• The need to supplement the IP---XACT standard with meta-data required to describe verification features of the device

• The need for a standard way to define traffic vectors across AMBA interfaces that is reproducible across tools and abstraction levels, enabling consistent performance analysis.

• The development of automated test scenarios aimed at mapping the 'performance envelope' (i.e. minimum latency & maximum bandwidth across all data paths and transaction types)

Presentation 2.4: Architectural Level Modeling with Power and Thermal Models

Speaker: Sylvian Kaiser (Docea Power, France)

Abstract:

System behavior, and associated thermal behavior, directly impact the success of electronic products: cost, reliability, safety, quality of service, user experience, and compliance with industry standards. The design technology trend is yet making power and thermal more and more critical with higher power density, higher system integration and more power-hungry applications. In this context high abstraction level approaches can bring valuable system-wide observation and exploration of power behavior in advance of real physical systems, in order to make appropriate power-aware design trade-offs. In this paper we show how IP power models can be used in an architectural level modeling and simulation environment dedicated to SoC and platform power analysis. The environment allows spreadsheet-like static analysis and dynamic analysis, including early power-performance trade-off. The same IP power models can alternatively be used in high level behavioral and functional modeling platforms, for instance ESL platforms typically relying on SystemC-TLM, in order to augment the system simulation with a power estimation perspective. To help share and reuse the IP models across system modeling environments and across teams within an organization, a framework is set-up that centralizes and manages the models. The complete methodology generally improves a company's ability to address the system power dissipation concerns and launch products meeting or exceeding power targets.

Presentation 3.1: Impact of time-dependent variability on the yield and performance of 6T SRAM cells in an advanced HK/MG technology

Speaker: Praveen Raghavan (IMEC, Belgium)

Abstract:

Stochastic device degradation—due to individual oxide defects—like Random Telegraph Noise (RTN) and Bias Temperature Instability (BTI) causes a threshold voltage drift of transistors resulting in decreased SRAM yield and performance. BTI and RTN have been shown to follow a defect-centric behavior, which can be bimodal in nature for heterogeneous gate oxide stacks. Consequently the tail distribution can significantly deviate from a Gaussian distribution. In this talk we will combine statistical silicon extracted from large transistor arrays (32k) designed and fabricated in an advanced 20nm High-k/Metal Gate process, with current state-of-the-art statistical assessment techniques in order to acquire a realistic impact of BTI degradation on the yield and performance of 6T SRAM cells.

Presentation 3.2: Facilitating Cross-Layer Reliability Management through Universal Reliability Information Exchange

Speaker: Enrico Costenaro (IROC Technologies, France)

Abstract:

Many existing and new applications aim at life-style improvement, high-speed networking, automotive safety, personal health care, wireless communications and consumer electronics. For these applications, reliability, availability and trustability are key factors, requiring careful design to meet the end users' expectations. The complex ASICs, which are now ubiquitous, often embed tens of millions of flip-flops, hundreds of megabits of embedded SRAM, and hundreds of millions of combinatorial cells. These designs integrate IP from multiple providers and are implemented in advanced process technologies, making it challenging to evaluate their reliability. For these reasons, we propose a common, open, reliability information interchange format (RIIF) allowing the formalization, specification and modeling of extra-functional, reliability properties for technology, circuits and systems. The models, which embed this reliability information, can be exploited (assembled together) to build a reliability-centric model of the system, allowing system reliability exploration and optimization using mathematical models and high-level tools. The proposed approach can be combined with performance analysis and management methodologies. As an example, the reliability and variability prediction methodologies at the various abstraction layers (transistor, gate, RTL, system) proposed in the European R&D MoRV project is presented. The RIIF models together with the proposed framework of system-level reliability management will reduce the engineering effort devoted to reliability analysis and will also minimize the amount of silicon area and power allocated to implementing reliability features. Early reliability budgeting will avoid unnecessary overdesign and contribute to the optimal implementation and design choices. Design and integration engineers will be able to use the proposed modelling tools to predict, monitor and manage the reliability of their components and systems through the design and implementation phases and ensure the established targets are met.

Presentation 3.3: Statistical Timing Methodology for Low-Power and Multi-Voltage Designs

Speaker : Kerim Kalafala (IBM Thomas J. Watson Research Center, USA)

Abstract:

The presentation will review a sign-off methodology using a statistical timing tool. The presentation will cover reliability and performance benefits, modeling with higher order terms, technology physics challenges, speed vs accuracy trade-offs, and abstraction/modeling techniques.

Presentation 3.4:

Speaker : Steffen Paul (ITEM, Germany)

Abstract:

High-quality analog design requires circuit insight supported by a suitable design methodology. Over the last years the gm/Id method, which came out as a result of the EKV MOS transistor model with inversion coefficient as one major design parameter, has established as a new design paradigm. In this talk the method is extended to handle device aging during the design phase. This is achieved by finding an optimal dc operating point and the right device dimensions.

Aging is a device specific phenomenon. Typically, analog circuits are part of large systems with given specification and so, device aging of the analog part will change system parameters over time. An efficient evaluation of degradation on system level is achieved by a cross layer approach, relating aging between device and system level. Response surface models are proposed to establish this relation. Simulations show that this approach enables an evaluation of aging at system level with reasonable simulation time.