IEEE VLSI Circuits and Systems Letter

Volume 7, Issue 2, Feb 2021

Editorial

Invited articles

- “Semiconductor Microelectronic Research at NSF/CISE: Status and Challenges” – Sankar Basu & Erik Brunvand
- “Exploratory Materials and Devices for Novel Energy Efficient Computation” – Azad Naeemi and Ian Young

Conference spotlight

- VLSI Design 2021 – Summary by Program Co-Chairs - Vijay Raghunathan (Purdue University), Saibal Mukhopadhyay (Georgia Tech.), and V. Kamakoti (IIT Madras)
  - VLSID-2021 Honorable Mention: “Efficient Hierarchical Post-Silicon Validation and Debug” - P. Kalimuthu, K. Basu and B. Carrion Schafer

Women in VLSI (WiV) series spotlight

- Interview with Dr. Tajana Rosing, Professor in the Department of Computer Science and Engineering at UC-San Diego, CA, USA.

Updates

- Recent relevant news highlights – Ishan Thakkar
- 2020 TCVLSI awards
- Planned 2021 conference sponsorships by TCVLSI
From the Editor-in-Chief’s Desk - Editorial

The IEEE VLSI Circuits and Systems Letter (VCaSL) is affiliated with the Technical Committee on VLSI (TCVLSI) under the IEEE Computer Society. It aims to report recent advances in VLSI technology, education, and opportunities and, consequently, grow the research and education activities in the area. The letter, published quarterly (since 2018), highlights snippets from the vast field of VLSI including semiconductor design, digital circuits and systems, analog and radio-frequency circuits, as well as mixed-signal circuits and systems, logic, microarchitecture, architecture and applications of VLSI. TCVLSI aims to encourage efforts around advancing the field of VLSI be it in the device, logic, circuits or systems space, promoting secured computer-aided design, fabrication, application, and business aspects of VLSI while encompassing both hardware and software.

TCVLSI sponsors a number of premium conferences and workshops, including, but not limited to, ASAP, ASYNC, ISVLSI, IWLS, SLIP, and ARITH. Emerging research topics and state-of-the-art advances on VLSI circuits and systems are reported at these events on a regular basis. Best paper awards are selected at these conferences to promote the high-quality research work each year. In addition to these research activities, TCVLSI also supports a variety of educational activities related to TCVLSI. Typically, several student travel grants are sponsored by TCVLSI at the following conferences: ASAP, ISVLSI, IWLS, iSES (formerly iNIS) and SLIP. Funds are typically provided to compensate student travels to these conferences as well as to attract more student participation. The organizing committees of these conferences undertake the task of selecting right candidates for these awards.

This issue of VCaSL features an invited article “Semiconductor Microelectronic Research at NSF/CISE: Status and Challenges” by National Science Foundation (NSF) Computer and Information Sciences (CISE) Program Directors, Sankar Basu and Erik Brunvand, where they describe details of microelectronics research support by the department highlighting key challenges.

The newsletter spotlights one of TCVLSI’s sponsored conferences in 2021, the IEEE International Conference on VLSI Design also known as VLSID. One page teasers of three best papers awarded at the 2021 VLSID conference are showcased: “150nA IQ, Quad Input - Quad Output, Intelligent Integrated Power Management for IoT Applications”; “A Fast Compact Thermal Model for Smartphones” & “Efficient Hierarchical Post-Silicon Validation and Debug”

In our Women in VLSI (WiV) series, we share an inspiring interview with Prof. Tajana Rosing, Professor in the Department of Computer Science and Engineering at UC-San Diego, CA, USA.

We have introduced a section on relevant recent announcements collated by our Associate Editor, Ishan Thakkar. Additionally, we have announced the winners of our 2020 TCVLSI awards.

I’d like to thank Dr. OlivierFranza for designing the cover page of this newsletter. Thank you to the authors of the various articles. I’d like to thank the IEEE CS staff, for their professional services to make the newsletter publicly available. I’d love to hear from the readers on what you would like to see in future newsletters. I welcome recommendations/feedback via email. Happy reading.

From the Editor-in-Chief’s Desk - Editorial

Chair, IEEE Computer Society TCVLSI
Editor-in-Chief of IEEE VCaSL, TCVLSI
Intel Corporation, USA

IEEE CS-TCVLSI: https://www.computer.org/communities/technical-committees/tcvlsi
Email: mondira.pant@ieee.org

TCVLSI has a total of about 1000 active members as of May 2021 and a readership of about 30,000
To join TCVLSI (its free), click here: https://www.ieee.org/membership-catalog/productdetail/showProductDetailPage.html?product=CMYVLSI732

https://tc.computer.org/tcvlsi/
Semiconductor Microelectronic Research at NSF/CISE: Status and Challenges

Sankar Basu, Program Director, CISE/CCF, National Science Foundation
Erik Brunvand, Program Director, CISE/CNF National Science Foundation

Invited article: IEEE Computer Society TCVLSI Newsletter.

1. Introduction:
The Computer and Information Science and Engineering (CISE) Directorate at NSF has been supporting microelectronics research at a significant level through several of its core programs as well as several cross-divisional, cross-directorate, and cross-agency programs for many decades. Indeed, two of its core programs - design automation for micro- and nano-systems, and computer systems architecture – in the Software Hardware Foundations (SHF) cluster of the Computing and Communication Foundations (CCF) Division were the first such NSF programs after the transition to Very Large Scale Integration (VLSI) of electronic chips in the 1980s, heralded as the dawn of modern microelectronics. Since then, these and other CISE programs, have been the primary source of sustained support for generations of research faculty in microelectronics at US universities over the years. The Directorate also has a long history of contributing knowledge and skill sets to the semiconductor industry via educating generations of graduate and undergraduate students in the field, and by co-sponsoring several of its collaborative programs with various industrial entities such as the Semiconductor Research Corporation (SRC) during the past decades, and the Intel Corporation in more recent years.

Much of the research investment of the CISE Directorate ultimately involves dealing with aspects of computing platforms: speed, power efficiency, bandwidth, and other performance metrics are intricately bound to the computing platforms on which applications run. Ultimately, these platforms are almost exclusively built using semiconductor technology. Thus, in a broader sense, many CISE programs leverage from, or are directly tied to, microelectronics. To be more specific, the computing stack is traditionally thought of as a hierarchy of layers with the devices, and circuits in the lowest layers, and with architecture, memory systems, operating systems, support software, algorithms, and applications in progressively higher layers. Lower layers of the stack (e.g., devices, circuits, architectures) more directly involve semiconductor technologies to the extent that researchers may interact with circuit fabrication facilities. Avant-garde designs often involve cross-layer considerations from top to bottom of the stack. Thus, in a broader sense, programs or projects addressing the higher end of the stack are also somewhat relevant to semiconductors, although the connection is often not quite as direct.

Before delving into the details of current CISE support for microelectronics research it may be of some interest to briefly set the historical context of evolution of computer hardware research at NSF. Much of what follows has been handed down to us from our NSF predecessors, the exact chronological details of which are taken from the recently documented history of the CISE directorate [1].

2. Historical context:
Computer and Information Science and Engineering (CISE) was established as one of the NSF Directorates in 1986 with Gordon Bell at its helm as the first Assistant Director for CISE. It had three Divisions: Computer Communications Research (CCR), Intelligent Information Systems (IIS), and the Microelectronics Information Processing (MIPS). The CCR Division was the predecessor of the current Computer Communications Foundations (CCF) Division, whereas the IIS division maintained its identity over the years. The vision and leadership of NSF Director Erich Bloch at the time, who had previously been a celebrated computer engineer at IBM, contributed much to the formation of CISE Directorate, and the MIPS Division in particular. The support for VLSI research at NSF started with Program Director Robert Grafton in the MIPS Division approximately around this time, when it became clear that the Mead-Conway paradigm would become a revolution. The MIPS division was disbanded later in 1997, and microelectronic design automation (along with computer architecture and several other programs having engineering flavor) was moved into the Computer Communications Foundations (CCF) Division, which is where it lives today.

A first set of three grants to Stanford, UC Berkeley, and U of Colorado, Boulder started off the microelectronic design automation, or the VLSI program for short [2]. In view of the current interest in access to semiconductor foundry by US academics (more on it later in Section 5) it may be noted that the MIPS Division provided large support for the Multi Project Wafer (MPW) fabrication service known as the Metal Oxide Semiconductor Implementation Service (MOSIS) program...
during those days [1]. As the field expanded, the VLSI program grew over the next decade, and covered several subareas, e.g., physical design, test and verification, system level design etc. almost exclusively limited to silicon CMOS technologies.

During subsequent years, federal funding for research in Computer Science and Engineering, and perhaps more so, interest in microelectronics research fueled by Moore’s law, grew by leaps and bounds. The field of Electronic Design Automation (EDA) not only made significant progress, but from an intellectual standpoint contributed to Computer Science (CS) research, including CS theory, in general. An article by Alberto Sangiovanni-Vincentelli, published in 2003 [4] states that according to the CiteSeer database, the three most cited papers in ALL of computer science at the time came from the field of Electronic Design Automation. They are, in order: “Optimization by Simulated Annealing,” Kirkpatrick et al., 1983; “Graph-Based Manipulation of Boolean Functions,” Bryant, 1986; and “Statecharts: a Visual Formalism for Complex Systems,” Harel, 1987. The book [14] by recent Turing award winner Jeffrey D. Ullman is one of many pieces of early evidence that VLSI provided much impetus to development of CS theory in general.

NSF funding had indirect influence outside of academic research in EDA as well. An internal assessment (unfortunately, formal publication testifying this is unavailable) during the first decades of 2000s indicated a large number of startups by NSF grantees, that far outnumbered most other subfields supported by CISE. The microelectronic design automation program also maintained close contact with industry sponsors – primarily through the Semiconductor Research Corporation (SRC). The NSF program held many joint workshops, and launched joint programs with SRC, e.g., the Multicore Chip Design and Architecture (MCDA) program in 2008, the Failure Resistant Systems (FRS) program in 2012, the Secure, Trustworthy, Assured and Resilient Semiconductors and Systems (STARSS) program from 2015-17, and the Energy Efficient Computing: Design and Architecture (E2CDA) in 2016-17. The foundational character of the core NSF programs propelled the programmatic agenda of some other mission agencies as well, e.g., DARPA programs on Joint University Microelectronics Program (JUMP), 3D-CNT-SoC, hyperdimensional computing, all of which have their genesis in NSF funding.

It is fair to remark that the enactment of the National Nanotechnology Initiative (NNI) [3] during the Clinton administration around the year 2000 had an eventual impact on the field. Although NNI largely target a broader set of applications other than computing at the beginning, CISE was a beneficiary of NNI. Despite an apparent intellectual gap with chip design efforts, it supported much futuristic research, which 20 years later has now begun to bear fruit in emergence of novel non-silicon devices, e.g., carbon nanotube transistors, spintronic devices and a plethora of other possibilities together with their integration in silicon platform. This trend largely continues today, and is covered by the Foundations of Emerging Technology (EMT) cluster discussed later in Section 3.

During the last decade, Moore’s law slowdown, together with end of Dennard scaling has triggered a natural shift in emphasis and has resulted in new research directions. From a programmatic standpoint two events are largely responsible for this trend: (a) publication of the “Game over” report [5] by the National Academy of Engineering, and (b) announcement in 2015 of a “A Nanotechnology-Inspired Grand Challenge for Future Computing” from the Whitehouse Office of Science and Technology Policy (OSTP) authored by Lloyd Whitman, Principal Assistant Director, Physical Sciences and Engineering at the time, and Randy Bryant, Assistant Director for Information Technology Research and Development on leave from CMU. While [5] rang the alarm bell from the high priests of computer architecture research, [6] laid out a succinctly stated vision, so as to

“create a new type of computer that can proactively interpret and learn from data, solve unfamiliar problems using what it has learned, and operate with the energy efficiency of the human brain.”

Much less is known of the fact that these, in turn, gave rise to multiagency government initiatives [7] that sponsored research on energy efficient computing by using algorithms from Artificial Intelligence (AI) and Machine Learning (ML) that is all pervasive in computer hardware research today.

A note of disclaimer on behalf of the authors is perhaps in order at this point. The specific remarks in the foregoing section should be taken only as a historical note in the context of microelectronic design automation program at NSF. Exactly similar comments, but different in detail, would perhaps also be true of the sister program on Computer Systems Architecture (CSA) as well. The CSA program has almost always been the beneficiary of being served by university rotators. While this aspect benefited the CSA program by maintaining direct contact with the academic research community, unfortunately the specific details of corporate history has been obscured to some extent.
3. **Broad swath of microelectronics in CISE research:**

Over the years as the field of Computer Information Science and Engineering grew, CISE investments also became larger – especially during the 5 year period of the federally funded initiative on Information Technology Research (ITR) beginning in Fiscal Year 2000. Today, with nearly a Billion dollars of total budget, CISE is the 2nd largest among the 8 major directorates in NSF after the Mathematical and Physical Sciences (MPS), and by all measures the largest single discipline directorate at NSF (MPS covers a number of disparate pure science disciplines including physics, chemistry, mathematics/statistics, and materials sciences). CISE funds upwards of 85% of federal support for research in the field, as opposed to a corresponding number of approximately 40% in physical sciences and other forms of engineering. This growth resulted in several reorganizations of the Directorate into Divisions and their constituent Programs, the last major one being in 2003. While the micro-electronic (later called micro- and nano-electronic) design automation program, and the computer architecture program maintained their identities, new programmatic entities came into existence due to new emphasis on system level hardware research as well on other application domains such as the design of intelligent, secure, and cyber-physical systems.

In what follows we describe various pockets where semiconductors or hardware design work resides in the CISE Directorate today. Our coverage may not be fully comprehensive, and we purposely mention only aspects relevant to hardware research that rests on semiconductor electronics in a broad sense.

3.1 **Computing and Communication Foundations (CCF) Division** supports research that advances computing and communication theory, algorithms for computer and computational sciences and architecture, and the design of computers and software. Within the CCF division there are two programs in particular that have direct relevance to semiconductor micro- and nano-electronic research.

The **Software and Hardware Foundations (SHF)** cluster supports research in the design, verification, operation, utilization, and evaluation of computer software and hardware through novel approaches, robust theories, high-leverage tools, and lasting principles. Such advances may involve formal methods, languages, logics, novel hardware artifacts, or algorithms to enable new or enhanced functionality, verification, usability, and scale.

SHF(H), the hardware part of the SHF cluster supports basic research in all topics in computer architecture and design automation, including but not limited to logical, physical, behavioral, and high-level synthesis methods, testing, and verification; pre- and post-silicon validation; and design methodologies for scalable, low-power, and energy-efficient circuits and systems in silicon technologies. This program also supports research on design automation for emerging non-silicon technologies, possibly using non-charge-based state vectors (e.g., electron-spin, micro-electromechanical systems, optics, or phase state of materials), that may have the potential to take computation beyond Moore's Law. Analog circuits and ultra-high-frequency communication circuits and systems within the above context and newer applications are also within scope.

The **Foundations of Emerging Technology (FET)** cluster is a programmatic entity that first came into existence under the name of Emerging Models and Technologies (EMT) shortly after the 2003 reorganization of the CISE Directorate, but was quickly disbanded within a few years. It was recently revamped under its new title supporting overarching fundamental research in disruptive technologies and models in computing and communication. Currently, the program has three areas of emphasis: First, neuromorphic, nanoscale computing, and newer computing paradigms (e.g., probabilistic bits, Ising machines that bridge quantum computing with classical hardware). The latter topics have a distinct overlap with SHF(H) cluster mentioned above, and are evolutions of the traditional micro- and nano- electronics program. The 2nd area of emphasis is bio-inspired computing systems and their design. This area also significantly leverages knowledge of microelectronic design paradigms, but in biological substrates. They are as disparate as molecular programming, DNA memory for storage, and use of design automation principles in systems biology. A third area in the cluster, of recent prominence, is relevant to aspects of quantum information processing.

3.2 **Computer and Network Systems (CNS) Division** came into existence with 2003 reorganization of the CISE Directorate, and was a merger of the previous Advanced Computational Infrastructure and Research (ACIR) Division, and the Advanced Networking Infrastructure and Research (ANIR) Division. Accordingly, the two main clusters in the CNS Division are the **Computer Systems Research (CSR)** and the **Networking Technology and Systems (NeTS) clusters**, with the
intent to encourage cross-fertilization among them.

The *Computer Systems Research* (CSR) cluster is of particular relevance to semiconductor and microelectronic research and supports transformative scientific and engineering research leading to the development of the next generation of highly performant, heterogeneous, power-efficient, environmentally sustainable, and secure computer systems. The scope of the program includes embedded and multicore systems and accelerators; mobile and extensible distributed systems; cloud and data-intensive processing systems; and memory, storage, and file systems. While much of CSR-supported research is at the system level (both hardware and software systems), it also supports research into hardware systems (accelerators, embedded hardware, sensors, etc.) at the systems level that naturally involve semiconductor and microelectronic activities.

The *Networking Technology and Systems* (NeTS) program supports research in helping to make networks, both wired and wireless, more easily controllable and manageable, resource and energy efficient, and secure and resilient to failures and attacks. The NeTS program supports research on the entire range of the networking “stack” including research into physical-layer radio and device technology that supports higher levels of the network communication stack. Here, the next generation 5G/6G networks and accompanying hardware are expected to leverage from basic advances in semiconductor electronics in the larger arena of CISE research.

3.3 **Information and Intelligent Systems (IIS) Division.** in general, supports research that studies the interrelated roles of people, computers and information to increase the ability to understand data, as well as mimic the hallmarks of intelligence in computational systems. While its focus is largely on the higher end of the computational stack, there are several aspects of research supported by the IIS Division that intimately involves and/or eventually depends upon advances in semiconductors and microelectronics. In particular, recent interest in Artificial Intelligence (AI) and Machine Learning (ML) (including the ongoing AI Institutes program) engage semiconductor research as a foundation.

It may be worth mentioning here that the *Computational Neuroscience* program in the IIS Division, has strong intellectual ties and interacts with micro- and nano-electronic design automation program in CCF. Some technical aspects of this will be further elaborated upon in Section 5. Other programs in the IIS Division ultimately rely on systems built on semiconductor electronics e.g., the *Human Centered Computing* (HCC) program supports research in human-computer interaction enabling haptic, tangible, and wearable interfaces; the *Smart and Connected Health* (SCH) program leverages novel electronic devices for healthcare needs, for example.

3.4 **Office of Advanced Cyber-infrastructure (OAC)** acts as a division of CISE but caters to the foundation-wide need for computing infrastructure including those of engineering, physical, biological, social and behavioral sciences at NSF. It supports and coordinates the development, acquisition and provision of state-of-the-art cyberinfrastructure resources, tools and services essential to the advancement and transformation of science and engineering. This includes research and education activities in all aspects of advanced cyberinfrastructure (CI) that lead to deployable, scalable, and sustainable systems capable of transforming science and engineering research and education. Advanced CI includes the spectrum of computational, data, software, networking, and security resources, tools, and services, along with the computational and data skills and expertise, that individually and collectively enable the conduct of science and engineering research and education.

The OAC portfolio includes research in architecture for extreme-scale systems which may include design, benchmarking, and analysis of extreme-scale systems for performance, programmability, and usability; storage, networks, and input/output (I/O); data centers and extreme-scale networked systems; and next-generation architectures. Naturally, many of these system architectures rely on custom microelectronic accelerators and application-specific processing hardware, and thus interacts at a foundational level the other divisions in CISE wherever appropriate.

4. **Cross disciplinary CISE programs**

We have mentioned several core CISE programs at NSF in the previous section. The core programs are longstanding entities with the goal of maintaining continuity and nurturing research communities over extended periods of time. In addition to these, there exist several cross disciplinary programs in CISE Directorate with targeted objectives, but we focus on only a few that are of strong relevance in the context of the present discussion of micro-electronics research.
The Secure and Trustworthy Cyberspace (SaTC) program is inspired by the national need to protect and preserve the growing social and economic benefits of cyber systems while ensuring security and privacy. In today's increasingly networked, distributed, and asynchronous world, cybersecurity involves hardware, software, networks, data, people, and integration with the physical world. Society's overwhelming reliance on this complex cyberspace, however, has exposed its fragility and vulnerabilities that defy existing cyber-defense measures; corporations, agencies, national infrastructure and individuals continue to suffer cyber-attacks.

Within this broad spectrum of research on security and privacy, the SaTC program involves hardware security as well. This includes new ways to design, build, and operate cyber systems, protect existing infrastructure, and motivate and educate individuals about security and privacy. Specific topics of interest include techniques for the development of secure and tamper-resistant hardware; identification, detection, and mitigation of hardware Trojans; watermarking; side channel attacks; reverse engineering of hardware designs; and hardware obfuscation. Also of interest are hardware acceleration of cryptographic algorithms (e.g., homomorphic encryption, post-quantum cryptography etc.), acceleration of security primitives, agile hardware implementations, modeling attacks and countermeasures, proximity verification, security metrics, trusted manufacturing, tamper proofing, and securing the hardware supply chain.

The Cyber-Physical Systems (CPS) program is focused on engineered systems that are built from, and depend upon, the seamless integration of computation and physical components. To wit, the program has been characterized in many ways: as a marriage between the digital and continuous worlds; between Turing's world of digital computation and Newton's laws of classical mechanics, etc. Advances in CPS will enable capability, adaptability, scalability, resiliency, safety, security, and usability that will expand the horizons of these critical systems. The CPS program aims to develop the core research needed to engineer these complex CPS, some of which may also require dependable, high-confidence, or provable behaviors. Core research areas of the program include control, data analytics, and machine learning including real-time learning for control, autonomy, design, Internet of Things (IoT), mixed initiatives including human-in- or human-on-the-loop, networking, privacy, real-time systems, safety, security, and verification.

Of particular relevance in this domain are systems design of intelligent edge devices that interact with the cloud efficiently and securely to accomplish desired tasks. Thus, design of hardware devices, i.e., IoTs, customized for this scenario to be energy efficient, secure, and that optimized to operate in a distributed/federated setting is of critical importance.

Still another cross-divisional program of immediate interest is the Principles and Practice of Scalable Systems (PPoSS), the focus of which is design of high performance, scalable and parallelizable machine architectures that span the entire hardware/software stack.

5. Future Outlook:
Space wouldn’t allow us to espouse on a detailed account of the current vision of the various NSF/CISE programs in the field. The interested reader can consult the various workshops that NSF has organized with the PI community in recent years. A partial listing of such events is provided in [8-11], with most workshop reports available online. A particularly noteworthy and recent NSF workshop from the perspective of micro- and nano- electronic design automation program is [11], a two-part report, elaborating the scientific vision and the needs for foundry access by US academics, is currently in preparation and should be available at the workshop website by the time this publication appears in print. In what follows, we briefly summarize the findings of that workshop roughly grouped into major topical areas.

Traditional EDA of VLSI circuits and systems: This area includes high-level and logic synthesis, physical design, design for manufacturability, testing and verification. These topics represent the core areas of EDA. New challenges introduced by deep submicron device technologies are addressed. With the maturity of CMOS technology and its scaling challenges, more innovations may come from architecture and system levels. These areas explore network-on-chip, system-on-chip designs and design automation, especially for low-power applications. Domain specific design paradigms is an emerging trend.

Machine learning (ML) and artificial intelligence (AI) hardware, neuromorphic hardware: This is triggered by slowdown is scaling, need for domain specific accelerators as announced e.g., in [6], and is a rapidly growing area that deserves extensive considerations.
Hardware paradigms both in conventional CMOS and in emerging technologies accompanied by constraints of their own can be retrofitted to exploit latitudes available in algorithmic considerations, thus giving rise to possibilities of hardware-software-algorithm co-design. Such research can inspire new algorithmic innovations, or could alternatively be driven by empirical or utilitarian considerations. A concomitant issue is how to verify AI/ML hardware, perhaps supplanting known formal or semi-formal methods. New research is also necessary to identify hardware bugs/faults, and methods to mitigate their effect on hardware performance in the context of ML applications.

ML is also emerging as an EDA tool for efficient, reliable and secure hardware design, including better and more efficient validation, verification, and detection of anomalous behavior of the system due to malicious attack. One may also optimize a variety of different system architectures, including for example, those suitable for cloud/edge computing by predicting/characterizing the nature of the computational load, utilize available computing resources, or handle data loss/corruption, and enable federated learning and inference via the use of ML algorithms, thus improving overall system performance. As a concrete example at the lowest level of the computing stack, one can cite current attempts to adapt Google DeepMind’s learning strategy in AlphaGo for EDA, or deep reinforcement learning for logic optimization. Similar design strategies at higher levels of the computing stack can also be envisioned.

Much as simulation has provided a "third leg" to the physical sciences beyond the traditional methods of theory and experiment, computing hardware now provides a new paradigm for the (neuro)scientists working with computer engineers, beyond in-vivo and in-vitro experimentation. Experimentation with new hardware/software now makes it possible for computational (neuro) scientists to test new hypothesis on understanding of the mammalian Brain by building physical prototypes on non-biological substrates. A flip side of the coin is designing neural prosthetics. Circuit design work of this latter type of brain-computer interface is currently supported by SHF(H), and the NSF-wide Integrative Strategies for Understanding Neural and Cognitive Systems (NCS) program.

**Emerging technology-based circuit and architecture design:** Beyond-CMOS technologies, such as carbon nanotubes, devices based on 2D materials (e.g., chalcogenides), phase-change memory, resistive RAM, tunnel FETs, spintronic devices, ferroelectric devices and photonic devices, exhibit unique I-V characteristics, and require combined effort of device scientists, circuit designers, EDA researchers, and computer architects in order to fully exploit their potential. Along with such emerging non-silicon devices, and non-charge based (e.g., electron spin, electro-mechanical phenomena at the nano scale etc.), devices, novel non-von-Neumann architectures can play a role here as well [15].

**Analog and high frequency designs:** Approaches to energy efficiency include selectively slowing down the clock speed, use asynchronous circuits, or at an extreme using analog circuits and systems, wherever possible. While potentially advantageous under certain circumstances, despite the fact that the latter techniques are known to have their own shortcomings, we are witnessing a resurgence of interest in them. Emerging applications such as 5G/6G communications, and imaging at microwave frequencies would depend on design of circuits and systems in the THz frequency regime. Consequently, there is a critical need for design of such systems to interact with the physical world.

**Quantum-inspired circuits and systems:** While Quantum Computing (QC) per se has a life of its own, there are new paradigms and technologies that bear resemblance to QC both from the standpoint of basic principles and applications targeted by them. Examples are: single flux quantum Logic, or computing using probabilistic bits (p-bits). These do not use q-bits as in QC, operate at room temperature, may be used to solve hard computational problems, e.g., integer factorization on the one hand and machine learning on the other. Other computational paradigms such as the Ising machines, implemented in CMOS or emerging technologies, provide yet other examples towards bridging the gap to quantum computing using classical hardware.

**EDA as a methodology:**
It has long been recognized that the sheer scale of digital circuit and systems design involving tens of billions of transistors and associated optimization, validation, verification problems pose an unprecedented methodological challenge in any man-made system design, electronic or not. The extreme-scale design principles developed over the decades in the EDA community have thus begun to be emulated by engineering design methods in other disciplines, e.g., computer networks [13], manufacturing, and even construction industry (cf. various talks in [11]).

**Issue of Semiconductor foundry access by US academics:**
This is a topic of much recent concern in government, industry and academia (both in education and research), which had been exacerbated by unavailability of MOSIS-like entities to the academic community (as noted earlier in the history section
of this document MOSIS originally received heavy support by the NSF/CISE Directorate). The NSF workshop [11] involving members of academic community, industry, the government, and foreign foundries discussed the issue in great detail. While an extensive report exclusively dedicated to this issue is available, their recommendations identifying the needs and a way forward are grouped into three categories:

(a) Access to standard silicon CMOS foundries. While sheer educational needs may not fully justify access to leading edge (e.g., 5nm nodes) foundries, avant-garde circuit design even in academic research labs suffer from lack of access.
(b) A second type, sometimes dubbed as “sauce over pasta”, is integration of a large number of emerging technology devices on a silicon wafer, thus enabling experimental research with novel architectures. This could potentially generate a new line of research as an alternative to aggressive transistor scaling.
(c) Issues surrounding commercial vs open source, licensing and export control of design tools, software/PDKs etc. creating obstacles for researchers from larger academic community.

6. Epilogue:
Chip design is at crossroads, both in the research and in the education arena. Smaller, faster, cheaper has been the mantra in the past; but transistor scaling, including power scaling has stalled. We have witnessed an unprecedented design complexity, unmatched by any human pursuit. Despite 20 years of search for an alternative material, Silicon CMOS remains the centerpiece of the technology, and will probably continue to be so in the foreseeable future. Chip design is an interdisciplinary endeavor requiring skills from engineering, mathematics, computer science and physical sciences. Data from academic institutions indicate that these factors along with lack of access to fabrication facilities are responsible for an apparent decline of student population in the field from undergraduate to doctoral levels, despite high demand from industry workforce. The intellectual challenges ahead are daunting.

One the other hand, information technologies, enabled by semiconductor electronics, have profoundly changed our lives during the last several decades, and continue to do so. Semiconductor chips are fundamental to emerging technological applications such as artificial intelligence, cloud computing, 5G, the Internet-of-Things (IoT), large-scale data analytics, and supercomputing. This is an area of utmost importance for the US, including national security, economic competitiveness, and global prosperity. These sentiments have recently reverberated the chambers of US lawmakers, and concrete actions appear to be in the offing (cf. [18]-[20]).

Sustained NSF/CISE funding (along with industry funding) has supported much of the development of the past decades in the field, enabling the underlying hardware infrastructure on which today’s information technology and scientific computing are built. The challenges facing the field are not just intellectual, but also educational and societal.

References:

Author Biographies:

Sankar Basu is a permanent Program Director at NSF. He started his career working on signals and systems theory, later worked on machine learning problems, and since 2002 have been managing the NSF micro- and nano- electronic design automation program. He is a fellow of the IEEE and the American Association for the Advancement of Science. More details are at his website: https://www.nsf.gov/staff/staff_bio.jsp?lan=sabasu&org=ENG.

Erik Brunvand is a Professor of Computer Science in the School of Computing at the University of Utah. His interests include the design of application-specific computers, graphics processors, physical computing, asynchronous systems, VLSI integrated circuit design, and arts/technology collaboration and integration in both research and education. He is currently serving as a rotating Program Director in the CNS division of the CISE Directorate at NSF. His university website is: https://www.cs.utah.edu/~elb/
Integrated circuit technology for computing will soon reach an inflection point. The power reduction through device scaling that served the semiconductor industry so well for over 30 years is seeing substantial challenges going forward. Exploratory materials, devices, and circuits that were once considered exotic are now being examined seriously because there may be no other alternatives. At the same time, this new-found freedom of breaking out of “CMOS scaling” introduces many new opportunities to do things completely differently—use a different material, invent a new device that operates on a different physical mechanism, and explore a new circuit function for computing that capitalizes on the unique properties of the new devices.

Research in this area is truly multidisciplinary as it brings together the researchers who are focused toward the exploration of a new device and interconnect, or function for a more energy-efficient integrated circuit for computing. Computation based on the emerging devices is not limited to just digital information processing but also encompasses non-Boolean computation, including analog, neuromorphic computing, and novel concepts in computer automata. It is envisioned that these exploratory devices and methods of computing could augment CMOS through their improvement in energy efficiency.

The IEEE Journal on Exploratory Solid-State Computational Devices and Circuits (JXCDC) is an “Open Access” journal for publication of research using exploratory materials and devices aimed at novel energy-efficient computation beyond standard CMOS (Complementary Metal Oxide Semiconductor) transistor technology. The focus of the publication is on the exploration of materials, devices, and computation circuits to enable Moore's Law to continue for computation beyond a 10 to 15 year horizon with the associated density scaling and improvement in energy efficiency. In addition to regular papers that can be on any topic within the scope of the journal, special topic CFPs are issued on a regular basis soliciting original papers covering various aspects of an emerging device technology or computing paradigm. Some examples of the published special topics keenly relevant to the TCVLSI members are presented here. Also presented is the call for new submissions.

Nonvolatile Memory for Efficient Implementation of Neural/Neuromorphic Computing JXCDC 2019, Issue 1 Part-2 (Guest Editor: Prof. Shimeng Yu, Georgia Institute of Technology)

In recent years, artificial intelligence (AI) based on machine/deep learning has shown significantly improved accuracy in large-scale visual/auditory recognition and classification tasks. In particular, deep neural networks (DNNs) and their variants have proved their efficacy in a wide range of image, video, speech, and biomedical applications. To improve accuracy, state-of-the-art deep learning algorithms tend to aggressively increase the depth and size of the network, which imposes ever-increasing computational capacity and storage cost in hardware. Although SRAM technology has been following the CMOS scaling trend well, the SRAM density and on-chip SRAM capacity are insufficient for storing the extremely large number of parameters in deep learning algorithms. Leakage current is undesirable, and parallelism is limited due to row-by-row operation in the digital SRAM array. As an alternative hardware platform, nonvolatile memory (NVM) devices have been proposed for weight storage with higher density and fast parallel analog computing with low power consumption. This Special Topic called for papers on the recent research progress of the NVM-based neuromorphic computing from the device level and array level up to the system level. The interaction and co-optimization between material/device engineering and circuit/architecture design were solicited. After the open submission and rigorous peer review process, six papers were selected for this Special Topic. The topics of these six papers range from neuron device design using emerging spintronic effects, new computation model and learning rule implemented with spintronic devices, to the neural network performance analysis with realistic device properties such as floating-gate memory and phase change memory (PCM).

Ferroelectric Transistors for Advanced Logic, Analog, and Memory Applications JXCDC 2019, Issue 2 Part-2 (Guest Editor: Prof. Azad Naeemi, Georgia Institute of Technology)

With recent advances in the growth and processing of ferroelectric materials and the emergence of CMOS compatible ferroelectrics, major research and development efforts are underway on ferroelectric transistors for logic, analog, and memory applications. Ferroelectric transistors show great promise for lowering the required supply voltage or adding new features and functionalities, such as nonvolatility or reconfigurability. They are also being explored for nontraditional circuits, such as convolutional and spiking neural networks and in-memory computing. Research in this area spans many levels of abstraction: from fundamental physical properties and material processing and characterization to various device concepts and to circuit and system design and benchmarking. This Special Topic in 2019 called for most recent developments in the area of ferroelectric transistors based on experiments and theoretical models.
It solicited original articles on various aspects of this emerging technology, its challenges and opportunities, its intrinsic versus practical limits, and the circuits and systems it may enable. The topics of the articles selected for publication include dense and fast memory arrays, various in-memory compute approaches and synapses and oscillators for convolutional and spiking neural networks.

**Spin–Orbit Coupling Effects for Advanced Logic and Memory**  JXCDC 2019, Issue 2 Part-3 (Guest Editor: Dr. Dmitri Nikonov, Intel Corporation)

In the past few years, fascinating progress has been made in the science of spin–orbit coupling effects, and fundamental research in this field is continuing vigorously. Whether or not the spin–orbit physics is going to make an impact on practical computing applications will depend on the answers to the following questions:

- Is spin–orbit torque (SOT) memory competitive with spin-transfer torque memory?
- Can logic circuits with spin–orbit devices be demonstrated?
- What is the unique advantage of spin–orbit devices for non-traditional computing?

This Special Topic in 2019 called for most recent developments in the area of spin–orbit devices and their applications in memory arrays and logic circuits. The authors were specifically asked to address the questions raised above. The papers selected for publication covered a wide range of topics including SOT memory development, SOT logic devices—proposals and demonstrations, precessional spin–orbit switching, thermal effect in SOT devices, circuits for SOT logic, SOT devices for analog, probabilistic, and neuromorphic computing, and spin–orbit applications to anti-ferromagnets. In summary, the Special Topic painted an optimistic picture for further development of SOC devices. No showstoppers transpired for gradual performance improvement of SOC devices and intriguing circuits and computing applications were presented.

**Coupled Oscillators for Non-von Neumann Computation**  JXCDC 2020, Issue 2 Part-3 (Guest Editor: Prof. Chris Kim, University of Minnesota)

When oscillators are loosely coupled to each other, energy transfer between the individual oscillators causes their frequencies to synchronize. Depending on the strength and time lag of the coupling medium, the phases of the oscillators settle in a way that minimizes the contentions among the oscillating signals. Recent works have shown that the coupled oscillator’s natural ability to resolve to the ground state can be exploited to solve computationally intractable problems, such as the traveling salesman problem, graph coloring, max cut, prime factorization, neural networks, associative memories, and pattern recognition. Here, the problems are first mapped to a coupled oscillator network by configuring the coupling weights, and the phase information is read out once the ground state is found. While resolving to the ground state, the network may get stuck in a local minima state, which can be avoided by a concept called annealing where random noise is added during the early exploration phase to help the oscillators break out of a local minima state.

This Special Topic solicited articles focusing on all aspects of coupled oscillator-based systems for non-von Neumann computing applications. The papers selected for publication covered topics such as novel annealing schemes, the impact of super-harmonic injection locking (SHIL) signal on the oscillator phase dynamics, and noise immunity. In one of the papers, researchers from Intel Co. reported a 22-nm test chip with 26 CMOS ring oscillators operating at gigahertz frequencies. In this chip, the individual oscillator frequencies were programmed using the pixel data while the degree-of-match result was obtained by measuring the amplitude of the common output voltage.

**Emerging Hardware for Cognitive Computing (currently with Open Call for Papers)**  (Guest Editor: Prof. Jean Anne Incorvia, The University of Texas at Austin)

Emerging materials and physics can be leveraged for new device-inherent behavior that can have system-level benefits. Motivation for device, circuit, and system behavior can be drawn from how the human brain processes certain data-intensive tasks adaptively and quickly, such as canonical image recognition. The field of neuromorphic computing has made great strides in implementing multi-weight synaptic behavior, as well as neuronal behavior such as integrate-and-fire and stochastic switching, and implementation of such behaviors in deep neural network (DNN) processing. Using CMOS, emerging resistive memories, and other device types as the basis, neuromorphic computing is innovating vertically from devices, to circuits, to systems to re-define how computation can be done. Looking forward, the realm of “cognitive computing” is inspired by new and continually emerging understanding of advanced brain and neuronal behavior that enables efficient and real-time learning and reaction.
This call for papers is on emerging hardware for cognitive computing. The focus and emphasis of these special topic papers is beyond DNN processing, as well as beyond multi-weight synapses and basic neuron integrating, firing, and stochasticity. Some example topics of interest include coherent reaction to multiple stimuli and input frequencies, hierarchical temporal memories, coupled dynamics between multiple stimuli and between devices or systems, both short-range and long-range connectivity in the devices and circuits, real-time adaptation to the data the computing system is processing, reconfigurability based on inputs, approximate computing that uses sparsity, and many other bio-inspired behaviors. Papers are encouraged that address implementation of advanced cognitive features at all levels (materials, device, circuits, and systems), including showing at a system level how such advanced functions can be useful for computing tasks and understanding the system-level energy efficiency and speed.

Topics of Interest for this Special Topic on Emerging Hardware for Cognitive Computing

- Reconfigurable and tunable materials and their application to computing
- Emerging resistive memories (e.g. MRAM, RRAM, PCM, FeFETs, etc) and taking advantage of their unique physical properties and how those properties can be mapped onto cognitive features
- Emerging materials and devices such as low dimensional, spintronic, ionic, and others that can be used for cognitive computing
- Circuits that make use of new device features or on their own have cognitive computing features such as long-range connectivity
- System-level and algorithm application of the devices and circuits and their benchmarking, e.g. Bayesian confidence propagation neural networks
- System-circuit-device co-design for cognitive computing
- Prototypes, simulation, and theory all welcome

Important Dates for this Special Topic are:

Open for Submission: April 15th, 2021
Submission Deadline: July 15th, 2021
First Notification: Aug 15th, 2021
Revision Submission: Aug 31st, 2021
Final Decision: Sept 15th, 2021
Publication Online: Sept 31st, 2021

Note: JXCDC is an Open Access only Publication: Charge for Authors: $1,350 USD per paper. Paper submissions must be done through the ScholarOne Manuscripts website: https://mc.manuscriptcentral.com/jxcdc. Inquiries for the JXCDC Journal should be sent to: JXCDC@IEEE.ORG

About the authors:

Azad Naeemi is a professor in the School of Electrical and Computer Engineering at the Georgia Institute of Technology. His research crosses the boundaries of materials, devices, circuits and systems, investigating integrated circuits based on conventional and emerging nanoscale devices and interconnects. He serves at the Editor-in-Chief of the IEEE Journal on Exploratory Computational Devices and Circuits (JXCDC) and an Editor of the IEEE Electron Device Letters. He is a recipient of the IEEE Electron Devices Society (EDS) Paul Rappaport Award, an NSF CAREER Award, and an SRC Inventor Recognition Award.

Ian A. Young is a Senior Fellow and director of Exploratory Integrated Circuits in the Technology Development Group of Intel Corporation. He now leads a research group exploring the future options for the integrated circuit in the beyond CMOS era. Ian Young received the PhD in Electrical Engineering from the University of California, Berkeley. He is the recipient of the 2009 International Solid-State Circuits Conference's Jack Raper Award for Outstanding Technology Directions paper. He received the 2018 IEEE Frederik Philips Award “for leadership in research and development on circuits and processes for the evolution of microprocessors”. Ian Young is a Life Fellow of the IEEE. He was the founding editor-in-chief of IEEE Journal on Exploratory Computational Devices and Circuits (JXCDC).
Conference Report - VLSI Design 2021

Vijay Raghunathan (Purdue University), Saibal Mukhopadhyay (Georgia Tech.), and V. Kamakoti (IIT Madras)

Technical Program Co-Chairs, VLSI Design 2021

The 34th IEEE International Conference on VLSI Design and 20th IEEE International Conference on Embedded Systems (VLSID 2021) took place virtually from Feb. 20-24, 2021 with a conference theme of “From the Transistor to Cyber-Physical Systems, For Solving Societal Challenges.” VLSID 2021 was originally scheduled to take place as an in-person event in Guwahati, India, in January 2021. However, the onset of the COVID-19 pandemic forced the organizing committee to quickly pivot to a virtual event with a mix of pre-recorded presentations and live sessions.

The conference received 230 submissions of which 208 went into the review stage. These submissions were rigorously reviewed by a technical program committee of 60 members who were supported by 45 external reviewers. Eventually, 56 papers were accepted, yielding an acceptance rate of 27%. In addition to the peer-reviewed papers, the conference also featured a top-notch invited component to the overall program, with fifteen keynote and plenary talks by titans from academia, industry, and government, including Prof. William Dally (NVIDIA & Stanford University), Prof. Umesh Mishra (UC Santa Barbara), Prof. Suman Datta (University of Notre Dame), Prof. Massimo Alioto (National University of Singapore), Dr. V.K. Saraswat (NITI Aayog, Government of India), Mr. Ravishankar Kuppuswamy (Intel), and senior executives from Cadence, Micron, Analog Devices, Siemens, STMicroelectronics, Xilinx, GlobalFoundries, MosChip, and CircuitSutra. The conference also featured three highly engaging panel discussions on “Decoding the Semiconductor Landscape of 2030”, “Self-Reliance in Semiconductors for AI”, and “The Changing Paradigm of VLSI Education in India.” The conference also had a special session of highlighted papers from the 2020 ACM/IEEE Design Automation Conference (DAC). Rounding out the main conference program were a set of seven tutorials delivered by leading experts from academia and industry on a variety of topics ranging from non-volatile memories to artificial intelligence. The conference also had a student research forum with 12 talks from senior Ph.D. students and a user/designer track that featured 8 short talks. Three papers were selected to receive awards at the conference – a best paper award, a best student paper award, and an honorable mention award. Each of these three award winning papers appears as a 1-page extended abstract in the following pages.

By any measure, the virtual conference was a big success. The conference was attended by 1700 attendees from over 200 organizations around the world. The conference attracted sponsorships from 18 Industry sponsors and technical sponsorship from IEEE TCVLSI, IEEE Circuits and Systems Society (CAS), IEEE Council on Electronic Design Automation (CEDA), ACM SIGDA, and ACM SIGMICRO, among others. Thanks to the sponsorships, the conference was able to provide student fellowships to almost 200 students to attend the conference at no cost. The conference also featured a virtual showcase/exhibit floor of products from 20 leading companies.
Battery operated wireless sensor nodes (WSN) consisting of sensors, microcontroller unit and a low power radio (like BLE) targeting industrial IoT applications, commonly employ heavily duty-cycled operation to minimize the system power consumption and to extend their lifetime. In standby mode the sensor nodes typically consume <1μA, however in active mode, though for short duration, the current consumption can be >10mA. The always-on power management IC regulating such systems must consume ultra-low quiescent current ($I_Q$) to minimize the standby mode current and to maximize the efficiency across a load current span of μA to mA.

Industrial and medical applications demand continuous availability of these sensor nodes which is ensured by augmenting a primary battery with additional energy sources such as solar cells and TEGs. A power management IC should therefore handle multiple energy sources and operate over a wide voltage range. Finally, continuous monitoring of output load currents enables run time data analytics and intelligence to detect malfunctioning loads, malicious code and incorrect device behavior. To accommodate these system requirements today’s sensor nodes are built with 4 to 6 power management chips with multiple passives, adding to the form factor and cost. In this paper, an ultra-low quiescent current ($I_Q$), Quad Input - Quad Output, Intelligent, Integrated power management IC addressing the aforementioned Industrial IoT system requirements is presented.

Figure 1 shows the block level overview of the proposed power management module. It consists of a 4:1 autonomous power-supply multiplexer (Auto-PMUX) which selects an appropriate energy source among the 4 available sources ($V_{IN[3:0]}$) based on the availability or a programmed priority and manages seamless switching between them. The Auto-PMUX output feeds a single-input multiple output (SIMO) DC-DC converter. The SIMO operate in discontinuous conduction mode and regulates 2 output voltages ($V_{OUT[1]}$ & $V_{OUT[2]}$) and 2 other internally used voltages ($V_{OUT[0]}$, 5.3V & $V_{OUT[3]}$, 1.4V) by automatically switching between buck, boost, or buck-boost modes depending on the input voltage and the desired output voltage levels. A load current metering module that tracks the current delivered to each of the SIMO output and provide a load dependent frequency pulse output is also integrated into the power management module.

In conclusion, A quad input, quad output, Integrated power management chip for Industrial IoT sensor nodes comprising of a power multiplexer, SIMO DC-DC converter, and energy metering module is implemented in a 130 nm CMOS process. The power management system consume an ultra-low quiescent current of 150 nA and therefore minimizes the standby mode current consumption of sensor nodes. The Auto-PMUX selects and switches between multiple energy sources based on their availability & priority. The SIMO DC-DC converter supports a wide VIN & VOUT voltage range by autonomously switching between buck/boost/proximity modes and maintains high efficiency across load current range of μA to mA . Energy metering module estimates the load current consumption across the DC-DC converter modes and enables run-time optimization, monitoring for security & defects.

References:
A Fast Compact Thermal Model For Smartphones
Anjali Agrawal, Anand Singh, Ankit Gola, Hameedah Sultan, S. R. Sarangi
IIT Delhi, India

Significant improvements in device performance and function-ality combined with affordable costs have resulted in a huge demand for smartphones and tablets in recent years. The processor speed has increased manifold, but the voltage requirements of transistors haven’t scaled down accordingly. This has resulted in increased power dissipation, which causes many adverse effects such as higher temperatures on the die. This degrades device performance and reliability. As the PCB temperature increases, so does the temperature on the screen of the device, which can degrade the user experience.

Hence, there is a need for thermal-aware system-level design. The structure of the system, its physical layout, material properties, and power dissipation affect the thermal profile. To assess the impact of these design decisions and find an optimal solution, an ultra-fast system-level thermal simulation method is needed. Most thermal simulators use the traditional finite element (FEM) or finite difference (FDM) methods to solve the classical Fourier heat equation and compute the temperature profile. However, these methods are quite slow, especially for complete system-level thermal analysis.

A much faster category of thermal simulators is based on Green’s functions. Here, the impulse response (Green’s function) is first obtained, and then at runtime, this impulse response is convolved with the power map to compute the complete temperature map. The main issue with Green’s functions is that we need to compute them using a slow method: physical measurements or FEM/FDM simulations. To analyze different layouts with different material properties, we need to re-compute the Green’s function for every configuration. This nullifies the performance gains of using Green’s functions. Consequently, in this work, we develop a very fast method to estimate the Green’s functions themselves. We propose a model based on small and simple functions - ultra-compact thermal models in our parlance - to help us calculate the Green’s function for a given layout and configuration (tailored for smartphones).

Figure 1 shows a brief overview of our approach. We employ a novel two-step process to develop an ultra-fast compact thermal model of a complete electronic system. First, we created a novel non-intrusive IR imaging framework. Next, we created an accurate CFD model for the phone using Ansys Icepak by calibrating it against real IR images and thermocouple readings. This three-way cross-validation enables accurate parametric modeling for a wide range of properties. Subsequently, we created a Green’s function-based model for the complete system. Our key contribution is an ultra-compact thermal model based on simple polynomial expressions to compute the Green’s function for different components very quickly.

A much faster category of thermal simulators is based on Green’s functions. Here, the impulse response (Green’s function) is first obtained, and then at runtime, this impulse response is convolved with the power map to compute the complete temperature map. The main issue with Green’s functions is that we need to compute them using a slow method: physical measurements or FEM/FDM simulations. To analyze different layouts with different material properties, we need to re-compute the Green’s function for every configuration. This nullifies the performance gains of using Green’s functions. Consequently, in this work, we develop a very fast method to estimate the Green’s functions themselves. We propose a model based on small and simple functions - ultra-compact thermal models in our parlance - to help us calculate the Green’s function for a given layout and configuration (tailored for smartphones).

Figure 1: Overview of our approach

![Figure 1: Overview of our approach](image)

A much faster category of thermal simulators is based on Green’s functions. Here, the impulse response (Green’s function) is first obtained, and then at runtime, this impulse response is convolved with the power map to compute the complete temperature map. The main issue with Green’s functions is that we need to compute them using a slow method: physical measurements or FEM/FDM simulations. To analyze different layouts with different material properties, we need to re-compute the Green’s function for every configuration. This nullifies the performance gains of using Green’s functions. Consequently, in this work, we develop a very fast method to estimate the Green’s functions themselves. We propose a model based on small and simple functions - ultra-compact thermal models in our parlance - to help us calculate the Green’s function for a given layout and configuration (tailored for smartphones).

Figure 1 shows a brief overview of our approach. We employ a novel two-step process to develop an ultra-fast compact thermal model of a complete electronic system. First, we created a novel non-intrusive IR imaging framework. Next, we created an accurate CFD model for the phone using Ansys Icepak by calibrating it against real IR images and thermocouple readings. This three-way cross-validation enables accurate parametric modeling for a wide range of properties. Subsequently, we created a Green’s function-based model for the complete system. Our key contribution is an ultra-compact thermal model based on simple polynomial expressions to compute the Green’s function for different components very quickly.

A much faster category of thermal simulators is based on Green’s functions. Here, the impulse response (Green’s function) is first obtained, and then at runtime, this impulse response is convolved with the power map to compute the complete temperature map. The main issue with Green’s functions is that we need to compute them using a slow method: physical measurements or FEM/FDM simulations. To analyze different layouts with different material properties, we need to re-compute the Green’s function for every configuration. This nullifies the performance gains of using Green’s functions. Consequently, in this work, we develop a very fast method to estimate the Green’s functions themselves. We propose a model based on small and simple functions - ultra-compact thermal models in our parlance - to help us calculate the Green’s function for a given layout and configuration (tailored for smartphones).

Figure 1 shows a brief overview of our approach. We employ a novel two-step process to develop an ultra-fast compact thermal model of a complete electronic system. First, we created a novel non-intrusive IR imaging framework. Next, we created an accurate CFD model for the phone using Ansys Icepak by calibrating it against real IR images and thermocouple readings. This three-way cross-validation enables accurate parametric modeling for a wide range of properties. Subsequently, we created a Green’s function-based model for the complete system. Our key contribution is an ultra-compact thermal model based on simple polynomial expressions to compute the Green’s function for different components very quickly.

A much faster category of thermal simulators is based on Green’s functions. Here, the impulse response (Green’s function) is first obtained, and then at runtime, this impulse response is convolved with the power map to compute the complete temperature map. The main issue with Green’s functions is that we need to compute them using a slow method: physical measurements or FEM/FDM simulations. To analyze different layouts with different material properties, we need to re-compute the Green’s function for every configuration. This nullifies the performance gains of using Green’s functions. Consequently, in this work, we develop a very fast method to estimate the Green’s functions themselves. We propose a model based on small and simple functions - ultra-compact thermal models in our parlance - to help us calculate the Green’s function for a given layout and configuration (tailored for smartphones).

Figure 1 shows a brief overview of our approach. We employ a novel two-step process to develop an ultra-fast compact thermal model of a complete electronic system. First, we created a novel non-intrusive IR imaging framework. Next, we created an accurate CFD model for the phone using Ansys Icepak by calibrating it against real IR images and thermocouple readings. This three-way cross-validation enables accurate parametric modeling for a wide range of properties. Subsequently, we created a Green’s function-based model for the complete system. Our key contribution is an ultra-compact thermal model based on simple polynomial expressions to compute the Green’s function for different components very quickly.

A much faster category of thermal simulators is based on Green’s functions. Here, the impulse response (Green’s function) is first obtained, and then at runtime, this impulse response is convolved with the power map to compute the complete temperature map. The main issue with Green’s functions is that we need to compute them using a slow method: physical measurements or FEM/FDM simulations. To analyze different layouts with different material properties, we need to re-compute the Green’s function for every configuration. This nullifies the performance gains of using Green’s functions. Consequently, in this work, we develop a very fast method to estimate the Green’s functions themselves. We propose a model based on small and simple functions - ultra-compact thermal models in our parlance - to help us calculate the Green’s function for a given layout and configuration (tailored for smartphones).

Figure 1 shows a brief overview of our approach. We employ a novel two-step process to develop an ultra-fast compact thermal model of a complete electronic system. First, we created a novel non-intrusive IR imaging framework. Next, we created an accurate CFD model for the phone using Ansys Icepak by calibrating it against real IR images and thermocouple readings. This three-way cross-validation enables accurate parametric modeling for a wide range of properties. Subsequently, we created a Green’s function-based model for the complete system. Our key contribution is an ultra-compact thermal model based on simple polynomial expressions to compute the Green’s function for different components very quickly.

We demonstrate that the error using our algorithm is limited to 2.56% while simultaneously being 1300 times faster than the state-of-the-art approaches. Figure 2 shows that our approach computes the location of hotspots and the maximum temperature rise at those locations accurately. This will allow our ultra-fast thermal modeling method to sift through a very large number of configurations to obtain the optimal configuration quickly - something which would be prohibitively expensive with traditional state-of-the-art approaches. Such approximations using simple polynomial and exponential functions for the Green’s functions are disruptively novel and extremely fast. It was hitherto believed that such simple approaches will not be effective. However, we show that such approaches can indeed be very accurate and effective.

References:
Trace Buffer Bandwidth-Aware Post-Silicon Debug
Pandy Kalimuthu, Kanad Basu and Benjamin Carrion Schafer
Department of Electrical and Computer Engineering
The University of Texas at Dallas

Post-silicon validation and debug is an extremely time-consuming process due to various challenges, such as bug localization, signal selection, trace buffer management, and trace data bandwidth. Among all these challenges, the trace data bandwidth requirement is often overlooked, despite being one of the major bottlenecks when it comes to debugging critical problems in real time. To address this important issue, we have proposed a method to validate and debug complex design bugs that require large trace bandwidths by dividing the problem into multiple hierarchies [1]. The proposed method is based on a two-prong approach. The first phase implements a coarse-grain selection of important signals, while the second phase subsequently zooms into the problematic circuit parts. This hierarchical method allows to root-cause the problem with limited trace data bandwidth. This leads to a significant reduction in the trace data volume.

Fig. 1 shows a typical post-silicon validation and debug environment. The process starts by loading an application into the device under test (DUT), enabling an appropriate trace channel, and setting trigger conditions, either through a debugger or through the application itself. Based on the trigger condition, the trace data is generated and exported out of the device and is analyzed offline. This resultant trace data either directly leads to identifying the source of the issue or helps to narrow down the possible causes. If the solution is not found, the trace selection and trigger conditions are updated to produce a different trace data set. This process is repeated till the source of the problem is identified. A large amount of traces is needed in this technique to debug a problem without ambiguity. Most of the existing research fail to appreciate the volume of trace data required, which often becomes the bottleneck for identifying the source of the problem.

In this work, we propose a highly flexible and extensible scheme to debug potentially complex problems that considers the trace data bandwidth available. This is accomplished by partitioning the problem into multiple hierarchical phases. Fig. 2 shows an overview of the complete flow splitting it into online and offline and coarse grain and fine-grain phases:

**Phase 1: Coarse-grain Observability:** Manual selection of coarse-grain signals to be traced before the application is executed. This process is design dependent.

**Phase 2: Fine-grain Observability:** Phase 2 demands high bandwidth and should only be used for detailed analysis of a limited number of transactions. This phase zooms into the problem areas detected in phase 1, so that detailed timing information can be retrieved through the trace buffer.

Experimental results demonstrate the benefits of our proposed approach in performing post-silicon debug, considering the trace bandwidth requirements. We anticipate future post-silicon validation research would consider trace bandwidth along with other challenges like signal selection.

**References:**
WOMEN-IN-VLSI (WiV) SERIES: Dr. Tajana Rosing

Dr. Rosing is a Full Professor of Computer Science and Electrical and Computer Engineering at UCSD, a Fratamico Endowed Chair and an IEEE Fellow. Here, she shares more about her work and the future of her field.

Q1. What VLSI research area do you focus on?

I am currently working on designing chips that can run large scale data workloads at amazing speeds. This includes applications for bioinformatics, where, for example, we have gotten COVID-19 analysis to run 1000x faster than was possible before. We also designed a new chip that mimics how brain operates, and is able to classify and cluster data 100,000x faster than the fastest design available today. It is based on a new type of machine learning, called Hyperdimensional computing, which uses large size vectors to represent the data, similar to what the brain does.

Q2: Why do you think this area is important currently?

In 2018 Turing Award recipients John Hennessy and David Patterson gave a lecture on “A New Golden Age for Computer Architecture: Domain-Specific Hardware/Software Co-Design, Enhanced Security, Open Instruction Sets, and Agile Chip Development.” This talk stressed that developing general purpose CPUs will not be sufficient to meet the needs of the upcoming workloads. In order to do so, we need to design heterogeneous systems that are more efficient and faster, but that are also supported with flexible software infrastructure. My work has focused on design of such novel hardware, and development of libraries needed to use that hardware as a part of a system.

Q3: What is your typical day like as a professor?

Being a professor is a lot of fun, as my days tend to be very varied and filled with people. During a normal week I teach 2-3 times per week, undergraduate and graduate classes, where I get to interact with lots of smart and curious students. I have a team of 20 PhD students who work with me on various research topics. I meet with them in teams that are related to specific projects. We also meet once per week as a whole team to hear a student present a paper they will present at a conference, or to practice the presentation for one of the exams they need to pass on the way to MS or PhD degrees. Many of the projects we work on are funded by industry, so multiple times a week my students and I meet with our collaborators from industry and discuss the next steps in the projects. I also meet with other faculty, both from UCSD and from other universities, to discuss joint funded projects, and to work on proposals for future projects. Normally I do a fair amount of travel, generally at least once per month, to go to conferences, and other meetings with government and companies I work with. I also spend time reading recent publications, helping my students with their research and editing their papers, and thinking about new ideas. I love my job because it is filled with amazing people, lots of new exciting ideas and plenty of variety!

Q4: What motivated you into this field?

When I was 13, I discovered computer games, and soon wanted to have infinite lives to finish one of my games. This led me to learning how to write code in assembly, and eventually to writing my own game. By the time I came to the USA as an exchange student in high school, I really wanted to learn about how computers are built. I took AP Electronics class and built my first computer. It was love at the first sight! My major in college was Electrical Engineering as I wanted to design chips. My dream became reality – in my first job after college I worked on chip design, and have been involved with VLSI ever since!

Q5: What excites you most when you look into the future?

I love seeing that more women and underrepresented minorities are joining the ranks of engineers, and with that are expanding our horizons, and providing new viewpoints. I am also excited to see the new developments in computer architecture, where we are moving away from general purpose CPUs, to a highly heterogeneous world, with different types of accelerators. In some ways the increase in diversity in engineering seems to parallel the changes in computer architecture!

Q6: Many students of color and women worry that the VLSI STEM field won’t welcome them. When you look at the landscape for scientists, what do you see right now? Are there signs of progress?

...
I see lots of signs of progress. When I first started, there were only 2% women in my graduate class. Now at UCSD we have 25% women in CSE. In fact, of 100 largest universities in the USA, UCSD has the highest percentage of female students majoring in STEM (https://cse.ucsd.edu/about/news/uc-san-diego-1-percentage-female-students-stem-majors). I see the same pattern also at other academic institutions and in industry.

Q7: How do you balance your work life and family life?

I have found that being professor has actually made it easier to balance my work and family life than when I was in industry, as I have a lot of flexibility on when and where I work. I am married to the most amazing man ever, have four great children, ranging in ages from 7 to 18, and have two dogs. Two children I had when still in industry, one before tenure, and one prior to becoming a full professor. We love to travel as a family, and also love the outdoors. I find that I am a lot more productive if I ensure that I have a balance between my personal and work life. What has helped me the most is having a partner who, in spite of also having a full time job, is dedicated to our family, and who contributes to our family as much as I do. Prior to COVID-19 we also had live-in AuPairs who helped watch our kids. Finally, it helps a lot to have collaborative and understanding environment at my department, at the School of Engineer and at UCSD.

Q8: What is your key message to young girls who aspire to be like you someday?

Reach for the stars and do not let fear stop you from trying new things. Engineering is a wonderful and exciting profession that women have been and will be very successful in.

Tajana Šimunić Rosing is a Professor, a holder of the Fratamico Endowed Chair, IEEE Fellow, and a director of System Energy Efficiency Lab at UCSD. Her research interests are in energy efficient computing, cyber-physical and distributed systems. She is leading a number of projects, including efforts funded by DARPA/SRC JUMP CRISP program, with focus on design of accelerators for analysis of big data, a project focused on developing AI systems in support of healthy living, SRC funded project on IoT system reliability and maintainability, and NSF funded project on design and calibration of air-quality sensors and others. She recently headed the effort on SmartCities that was a part of DARPA and industry funded TerraSwarm center. Tajana led the energy efficient datacenters theme in MuSyC center, and several large projects funded by both industry and government focused on power and thermal management. Tajana’s work on proactive thermal management and ambient-driven thermal modeling was instrumental in laying the groundwork in this field and has since resulted in a number of industrial implementations of these ideas. Her research on event driven dynamic power management laid the mathematical foundations for the engineering problem, devised a globally optimal solution, and more importantly defined the framework for future researchers to approach these kinds of problems in embedded system design. From 1998 until 2005 she was a full time research scientist at HP Labs while also leading research efforts at Stanford University. She finished her PhD in EE in 2001 at Stanford, concurrently with finishing her Masters in Engineering Management. Her PhD topic was dynamic management of power consumption. Prior to pursuing the PhD, she worked as a senior design engineer at Altera Corporation. She has served at a number of Technical Paper Committees, including being an Associate Editor of IEEE Transactions on Mobile Computing, an Associate Editor of IEEE Transactions
Recent relevant developments:

Captured by Ishan Thakkar

(1) CHIPS for America Act, FY 2021
[https://www.semiconductors.org/chips/]
By enacting the CHIPS for America Act in the FY 2021 National Defense Authorization Act (NDAA), U.S. Congress recognized the critical role the U.S. semiconductor industry plays in America’s future. Now, the administration and Congress must fully fund the semiconductor manufacturing and research provisions authorized by the NDAA — and enact an investment tax credit — to strengthen America’s global leadership in chip technology for years to come.

(2) Intel’s New IDM 2.0 Strategy
The new CEO of Intel, Pat Gelsinger, has outlined his vision for Intel over the coming years. Highlights of Intel’s announcements: Two new fabs in Arizona, $20b investment; New Intel Foundry Services, offering Intel manufacturing to customers; Next generation 7nm chiplets for ‘Meteor Lake’ will finish design in Q2 2021; New research collaboration with IBM in foundational semiconductor design; New Intel Innovation event in Oct 2021, Spirit of IDF.

(3) TSMC to Spend $100 Billion Over Three Years to Grow Capacity
[https://www.bloomberg.com/news/articles/2021-04-01/tsmc-to-invest-100-billion-over-three-years-to-grow-capacity]
Taiwan Semiconductor Manufacturing Co. (TSMC) plans to spend $100 billion over the next three years to expand its chip fabrication capacity, a staggering financial commitment to address booming demand for new technologies.

(4) Arm Unveils Armv9 Architecture for Its Next Generation of Processor Chips
[https://siliconangle.com/2021/03/30/arm-unveils-armv9-architecture-next-generation-processor-chips/]
The British chip designer Arm Holdings Ltd. introduced a roadmap for its next-generation central processing unit architecture. Its new breed of chips will be optimized for specialized workloads such as artificial intelligence and digital signal processing and come with greater security built right into the silicon.

(5) Samsung Develops Industry’s First High Bandwidth Memory with AI Processing Power
Samsung’s paper on the HBM-PIM appeared at the renowned International Solid-State Circuits Virtual Conference (ISSCC) in February 2021. The new architecture will deliver over twice the system performance and reduce energy consumption by more than 70%.

(6) Bandwidth Demand Prompts Micron Transition from 3D Xpoint to CXL
[https://www.eetimes.com/bandwidth-demand-prompts-micron-transition-from-3d-xpoint-to-cxl/]
Micron Technologies is exiting the once-promising 3D Xpoint non-volatile memory market, instead focusing its data center efforts on the emerging Compute Express Link (CXL) interface as bandwidth requirements soar. How long will Intel hang on?

(7) Programmable Optical Quantum Computer Arrives Late, Steals the Show
The researchers, from a startup called Xanadu and the National Institute of Standards and Technology, have pulled together technology developments to produce an integrated optical chip based eight-qubit quantum computer. It is shown that there are no large barriers to scaling this same computer to a bigger number of qubits.

(8) ReRAM Machine Learning Embraces Variability
[https://www.eetimes.com/eram-machine-learning-embraces-variability/]
CEA-Leti researchers implemented a Markov Chain Monte Carlo (MCMC) algorithm in a fabricated chip to actively exploit the ReRAM memristor randomness for use in edge learning systems.

Information collated by Ishan Thakkar, Associate Editor of TCVLSI-VCaSL, Assistant Professor Dept. of Electrical and Computer Engineering
University of Kentucky, Lexington, KY 40508
Announcing the winners of the 2020 IEEE-CS TCVLSI awards.

IEEE-CS TCVLSI Distinguished Research Award : Prof Hai Li (Duke University)
IEEE-CS TCVLSI Distinguished Research Award : Prof Kaushik Roy (Purdue Univ)
IEEE-CS TCVLSI Mid-Career Research Achievement Award : Prof Himanshu Thapliyal (University of Kentucky)

Thank you to those who took the time to nominate and a big congratulations to the winners.

IEEE-CS TCVLSI Awards Committee Chairs: Tajana Rosing (Prof at UCSD) and Ludmila Cherkasova (ARM)

About the winners:

Hai (Helen) Li (M’08-SM’16-F’19) received her bachelor’s and master’s degrees from Tsinghua University, China, and her Ph.D. degree from Purdue University, USA. She is Clare Boothe Luce Professor and Associate Chair of the Electrical and Computer Engineering Department at Duke University. Before that, she was with Qualcomm Inc., San Diego, CA, USA, Intel Corporation, Santa Clara, CA, Seagate Technology, Bloomington, MN, USA, the Polytechnic Institute of New York University, Brooklyn, NY, USA, and the University of Pittsburgh, Pittsburgh, PA, USA. Her research interests include neuromorphic computing systems, machine learning and deep neural networks, memory design and architecture, and cross-layer optimization for low power and high performance. She has authored or co-authored more than 250 technical papers in peer-reviewed journals and conferences and a book entitled Nonvolatile Memory Design: Magnetic, Resistive, and Phase Changing (CRC Press, 2011). She received 9 best paper awards and an additional 9 best paper nominations from international conferences. Dr. Li serves/served as an Associate Editor of a number of IEEE/ACM journals. She was the General Chair or Technical Program Chair of multiple IEEE/ACM conferences. Dr. Li is a Distinguished Lecturer of the IEEE CAS society (2018-2019) and a distinguished speaker of ACM (2017-2020). Dr. Li is a recipient of the NSF Career Award, DARPA Young Faculty Award (YFA), TUM-IAS Hans Fischer Fellowship from Germany, and ELATE Fellowship (2020). Dr. Li is an IEEE fellow and a distinguished member of the ACM.

Kaushik Roy is the Edward G. Tiedemann, Jr., Distinguished Professor of Electrical and Computer Engineering at Purdue University. He received his BTech from Indian Institute of Technology, Kharagpur, PhD from University of Illinois at Urbana-Champaign in 1990 and joined the Semiconductor Process and Design Center of Texas Instruments, Dallas, where he worked for three years on FPGA architecture development and low-power circuit design. His current research focuses on cognitive algorithms, circuits and architecture for energy-efficient neuromorphic computing/machine learning, and neuro-mimetic devices. Kaushik has supervised 91 PhD dissertations and his students are well placed in universities and industry. He is the co-author of two books on Low Power CMOS VLSI Design (John Wiley & McGraw Hill). Kaushik received the National Science Foundation Career Development Award in 1995, IBM faculty partnership award, ATT/Lucent Foundation award, 2005 SRC Technical Excellence Award, SRC Inventors Award, Purdue College of Engineering Research Excellence Award, Humboldt Research Award in 2010, 2010 IEEE Circuits and Systems Society Technical Achievement Award, Distinguished Alumni Award from Indian Institute of Technology (IIT), Kharagpur, Fulbright-Nehru Distinguished Chair, DoD Vannevar Bush Faculty Fellow (2014-2021), Semiconductor Research Corporation Aristotle award in 2015, Arden Bement Research Award (Purdue University's highest research award) in 2020, and best paper awards at over 10 premier conferences. He has been in the editorial board of IEEE Design and Test, IEEE Transactions on Circuits and Systems, IEEE Transactions on VLSI Systems, and IEEE Transactions on Electron Devices. He was Guest Editor for Special Issue on Low-Power VLSI in the IEEE Design and Test (1994) and IEEE Transactions on VLSI Systems (June 2000), IEE Proceedings – Computers and Digital Techniques (July 2002), and IEEE Journal on Emerging and Selected Topics in Circuits and Systems (2011). He is a fellow of IEEE.

Himanshu Thapliyal is an Associate Professor and Endowed Robley D. Evans Faculty Fellow with the Department of Electrical and Computer Engineering, University of Kentucky, Lexington, KY, USA. From 2012-14, he worked as a designer of processor test solutions at Qualcomm, where he received the Qualcomm QualStar Award for contributions to memory built-in self-tests. He has been ranked in the top 50 among scientists throughout the world in the field of ‘Computer Hardware and Architecture’ for the 2019 calendar year (PLOS Biology database). He is the recipient of the 2019 NSF CAREER award. He has authored over 150 journal/conference articles with H-index=38 and received Best Paper awards at the 2021 IEEE International Conference on Consumer Electronics (ICCE), 2020 IEEE World Forum on Internet of Things, 2017 Cyber and Information Security Research Conference (CISR) and 2012 IEEE Computer Society Annual Symposium on VLSI (ISVLSI). He is serving in editorial boards of journals including Section Editor of the Springer Nature Computer Science, Senior Associate Editor of the IEEE Consumer Electronics Magazine, Associate Editor of the IEEE Internet of Things Journal, and the editorial board member of the Microelectronics Journal. His research interests include hardware security of IoT and vehicles, circuit design of emerging technologies including quantum computing, and smart healthcare.
TCVLSI Sponsored Conferences for 2021

Financially sponsored/co-sponsored conferences

- ARITH, IEEE Symposium on Computer Arithmetic
- ASAP, IEEE International Conference on Application-specific Systems, Architectures and Processors
- ASYNC, IEEE International Symposium on Asynchronous Circuits and Systems
  - ASYNC 2021: [https://asyncsymposium.org/async2021/](https://asyncsymposium.org/async2021/) Virtual conference dates: Sept 7-10 2021
- iSES, (formerly IEEE-iNIS) IEEE International Smart Electronic Systems
  - IEEE iSES 2021 Dec 20-22 2021
- ISVLSI, IEEE Computer Society Symposium on VLSI
  - ISVLSI 2021: [http://www.eng.ucy.ac.cy/theocharides/isvlsi21/](http://www.eng.ucy.ac.cy/theocharides/isvlsi21/) Virtual conference dates: July 7-9 2021
- IWLS, IEEE International Workshop on Logic & Synthesis – collocated with DAC
  - IWLS 2021: Virtual conference dates: June 19th-21st, 2021
- SLIP, ACM/IEEE System Level Interconnect Prediction
  - SLIP 2021: [https://dl.acm.org/conference/slip/proceedings](https://dl.acm.org/conference/slip/proceedings) Date TBD

Technically Co-Sponsored Conferences for 2021

- VLSID, International Conference on VLSI Design

Explore conference sponsorship options with TCVLSI here: [https://www.computer.org/conferences/organize-a-conference/sponsorship-options](https://www.computer.org/conferences/organize-a-conference/sponsorship-options)