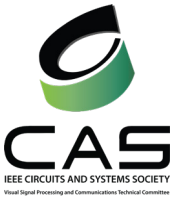
17th EMBEDDED SYSTEMS WEEK

OCTOBER 08-15, 2021 | VIRTUAL CONFERENCE

2021 ESWEEK PROGRAM

SPONSORING SOCIETIES:

Thank You to Our Sponsoring Societies



Thank You to Our Platinum Sponsors

Thank You to Our Gold Sponsors

TABLE OF CONTENTS

WELCOME TO ESWEEK 2021

Andreas Gerstlauer, General Chair
The University of Texas at Austin, USA

Aviral Shrivastava, Vice-General Chair
Arizona State University, USA

Welcome to the Virtual edition of ESWEEK 2021!

Embedded Systems Week (ESWEEK) is the premier event covering all aspects of hardware and software design for smart, intelligent and connected computing systems. By bringing together three leading conferences (CASES, CODES+ISSS, EMSOFT), a symposium (NOCS) and several workshops and tutorials, ESWEEK allows attendees to benefit from a wide range of topics covering the state of the art in embedded systems research and development.

Due to the on-going pandemic, this 17th edition of ESWEEK will remain fully virtual this year. However, ESWEEK 2021 still promises to be a very exciting and engaging event. It consists of tutorials, keynotes, paper sessions with live talks and interactive poster presentations every day that, together with a very low registration fee, provide an opportunity for the community across the globe to come together, engage, interact, and celebrate the latest advances in embedded software and systems. New this year is an Education Track in which top international researchers present short courses on various new/emerging but well established fundamental embedded system topics with “embedded learning” as the theme this year. In addition, we have for the first time established a registration fee waiver program that will allow attendees with demonstrated need, e.g. students and researchers from developing countries with limited financial funds or means (e.g. access to credit cards) to attend ESWEEK for free.

Following the journal-integrated publication model for the three conferences (CASES, CODES+ISSS, and EMSOFT), all regular papers presented are published in a Special Issue of the ACM Transactions on Embedded Computer Systems (TECS). To this end, ESWEEK-related journal submissions to TECS followed a journal-style peer review process conducted in two stages with the opportunity of minor/major revisions before the final decision. In addition, the selected Work-in-Progress (WiP) track papers are published in conference-specific Proceedings.

The technical program on Monday, Tuesday, and Wednesday consists of 4 special sessions and 17 regular sessions from the three conferences. There is a strong emphasis on interaction in the virtual setting. All the accepted journal-track and WiP papers along with prerecorded lightning video presentation for each paper are made available to the ESWEEK participants two weeks before the virtual event. During the conference week, the live technical sessions will feature 15 minute talks for the journal-track papers including Q&A. In addition, all journal-track and WiP papers will be accompanied by interactive virtual poster sessions in which attendees can further interact with presenters and authors.

Tuesday will be a special Industry Day at ESWEEK. The day will be framed by two special industry sessions on on collaborative industry-academia European research projects before the main conference, and an industry pitch session concluding the day in which speakers from industry will outline design challenges aimed at fostering discussion between experts, increase awareness about practical challenges, and allowing academic researchers to put their research into the context of industrial needs.

Wednesday will be a special Edge AI Day focusing on

WELCOME TO ESWEEK 2021 (cont.)

research related to machine learning (ML) and artificial intelligence (AI) on embedded, edge and IoT devices. In addition to relevant journal-track paper sessions, the Wednesday keynote, two special sessions and a final panel concluding the day are devoted to this subject.

Highlights of the ESWEEK program are three distinguished keynote talks by prominent leaders in academia and industry, covering relevant trends for future embedded and cyber physical systems and providing deep insights into technology drivers. Mike Davies from Intel will introduce the latest advances in neuromorphic computing for fast, efficient, and intelligent processing on Monday. On Tuesday, Professor Claire Tomlin from UC Berkeley will talk about safe learning in robotics. Finally, as part of the Edge AI Day on Wednesday, Pete Warden from Google will discuss the needs, challenges and opportunities associated with machine learning on embedded systems.

The main conference program will conclude with a panel on Wednesday as part of the Special Edge AI Day focusing on "Machine Learning on the Edge: How Deep Can We 'Embed' It Into the Cloud-Edge Continuum?". Top experts will share their views on this highly relevant topic.

The main program is preceded by tutorials on the Friday before the conference week, and Education Classes on Saturday and Sunday. These provide an excellent opportunity to get in-depth knowledge on both new trends and hot topics as well as emerging but well-established embedded systems concepts, tools and methods that are not readily available in textbooks. There are six half-day, virtually presented academic and industry tutorials, covering a wide scope from quantum computing, fog computing for IoT and deterministic programming of CPS to edge AI, compute acceleration and GPU code generation products from Intel, Xilinx and MathWorks. The theme for ESWEEK Education Classes this year is "Embedded Learning". Twelve classes will provide an exciting, engaging and hands-on way to learn about topics ranging from neuromorphic computing to edge ML and autonomous driving to students across the globe, especially the ones that do not have access to high quality educational content.

We are looking forward to seeing you virtually at the inspiring, interesting, and interactive ESWEEK 2021!

Thursday and Friday of the conference week are the days for the symposium and workshops. Besides the NOCS (Networks on Chip) symposium, we have three workshops covering a range of important topics in embedded systems: RSP (Rapid System Prototyping), TRAIN (Trustworthy and Reliable Accelerator Design), and MSC (Memory and Storage Computing).

The organization of ESWEEK was only possible with the continuous support and help from the sponsors and many volunteers: The program chairs with their program committee members, the organizers of the workshops, tutorials, and symposia, all members of the organization committee, and, last but not least, the virtual conference chair and the web chair---without their commitment and contributions this virtual event would not exist.

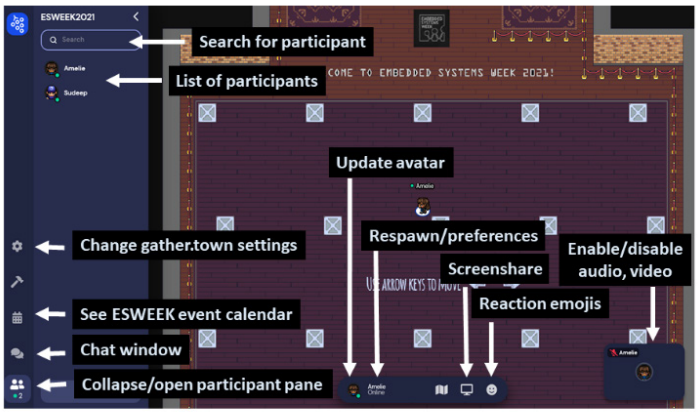
We are looking forward to seeing you virtually at the inspiring, interesting, and interactive ESWEEK 2021!

ESWEEK VIRTUAL CONFERENCE LAYOUT

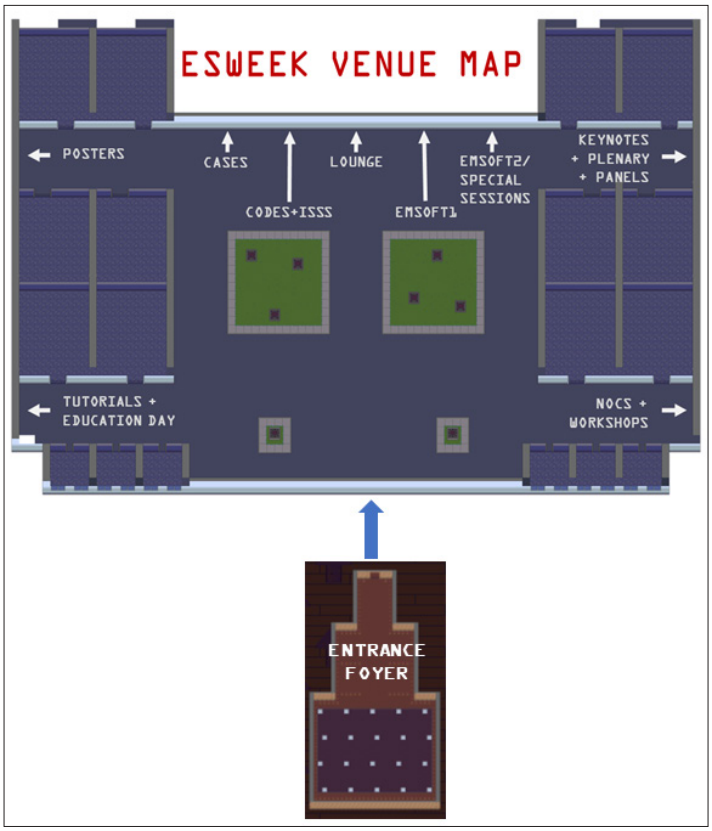
Sudeep Pasricha,
Virtual Conference Chair
Colorado State University, USA

ESWEEK 2021 has a dedicated virtual space, hosted on the Gather.town platform, for conference attendees to interact with other attendees and presenters during the week. In this virtual space, you can attend interactive poster sessions, post-presentation Q&A sessions, and view livestreams of talks for all CASES/CODES+ISSS/EMSOFT technical sessions, keynote and plenary sessions, and co-located symposia (NOCS) and workshops (RSP/TRAIN/MSC). You can visit industry sponsor booths to engage with their representatives. The virtual platform allows you to search for a specific ESWEEK attendee and chat with them via text or video. You can also have a video chat with a person who is close enough with you or with people in the same private space as you.

You will sign into the ESWEEK 2021 virtual space using the link and passcode received from the organizers (a few days before the conference begins). After signing in, the user interface (UI) will have options to edit your avatar, enable/disable your mic and video, see/find other attendees, interact with people and objects, and to quickly teleport to (or get directions to) any scheduled session. The UI image shows annotations to highlight the key UI elements.



The Gather.town space for ESWEEK 2021 consists of 9 areas, as shown in the venue map. Four rooms are dedicated to the four parallel sessions; one each for CASES and CODES+ISSS, and two for EMSOFT and special sessions. Another room is dedicated for plenary sessions, keynotes, and panels. A lounge area provides a dedicated social space. The posters area leads to separate rooms for WIP posters and dedicated interactive spaces for Q&A with authors of all accepted papers after their talks. The tutorials and education day area leads to dedicated rooms for all tutorials and education day sessions. The NOCS and workshop area leads to rooms dedicated to the RSP/TRAIN/MSC workshops and the NOCS symposium. All rooms are connected together and can be reached from anywhere in the map, just like at a real conference venue. You can also use the ESWEEK event calendar in your UI to teleport to any ongoing or upcoming sessions.



ESWEEK 2021 OVERVIEW

Friday, October 8	Saturday, October 9	Sunday October 10
Tutorials	**Education Classes**	**Education Classes**
Monday, October 11	**Tuesday, October 12**	**Wednesday, October 13**
ESWEEK Opening Keynote by Mike Davies CASES, CODES+ISSS, EMSOFT and Special Sessions Work-in-Progress Poster Session	Test of Time Awards Keynote by Claire Tomlin Industry Sessions CASES, CODES+ISSS, EMSOFT and Special Sessions Industry Pitch and Poster Session	Best Paper and Other Awards Keynote by Pete Warden CASES, CODES+ISSS, EMSOFT and Special Sessions Panel: Machine Learning on the Edge: How Deep Can We 'Embed' It Into the Cloud-Edge Continuum?
Thursday, October 14	**Friday, October 15**	
NOCS Opening NOCS Keynote by Keren Bergman NOCS Sessions MSC Workshop TRAIN Workshop RSP Workhop	NOCS Keynote by Gabriel (Gabe) Loh NOCS Sessions	

BEST PAPER CANDIDATES

CASES

1A.1 RiSA: A Reinforced Systolic Array for Depthwise Convolutions and Embedded Tensor Reshaping

Authors:

Hyungmin Cho - *Sungkyunkwan University*

3A.3 Two Birds with One Stone: Boosting Both Read and Write Performance for Tree Indices on Persistent Memory

Authors:

Yongping Luo, Peiquan Jin, Zhou Zhang, Juncheng Zhang - *University of Science and Technology of China*

Qinglin Zhang, Bin Cheng - *Tencent*

3A.4 You Only Traverse Twice: A YOTT Placement, Routing, and Timing Approach for CGRAs

Authors:

Michael Canesche - *Universidade Federal de Viçosa (UFV)*

Westerley Carvalho - *Universidade Federal de Viçosa (UFV)*

Lucas Reis - *Universidade Federal de Viçosa (UFV)*

Matheus Oliveira - *Universidade Federal de Viçosa (UFV)*

Salles Magalhães - *Universidade Federal de Vçosa (UFV)*

Peter Jamieson - *Miami University*

José Augusto Nacif - *Universidade Federal de Viçosa (UFV)*

Ricardo Ferreira - *Universidade Federal de Viçosa (UFV)*

CODES+ISSS

1B.2 Exploring Efficient Architectures for Remote In-Memory NVM over RDMA

Authors:

Qingfeng Zhuge, Hao Zhang, Edwin Hsing-Mean Sha, Rui Xu - *East China Normal University*

Jun Liu, Shengyu Zhang - *Lenovo Ltd.*

2B.2 Intermittent-Aware Neural Architecture Search

Authors:

Hashan Mendis, Chih-Kai Kang, Pi-Cheng Hsiu - *Academia Sinica*

4B.1 MARS: mmWave based Assistive Rehabilitation System for Smart Healthcare

Authors:

Sizhe An and Umit Ogras - *University of Wisconsin-Madison*

EMSOFT

1C.4 Excluding Parallel Execution to Improve Global Fixed Priority Response Time Analysis

Authors:

Quan Zhou, Jianjun Li, Guohui Li - *Huazhong University of Science and Technology*

2D.1 Towards an Integrated Vehicle Management System in DriveOS

Authors:

Soham Sinha and Richard West - *Boston University*

3C.1 Verified Lustre Normalization with Node Subsampling

Authors:

Timothy Bourke, Basile Pesin, Paul Jeanmaire - *Inria / ENS*

Marc Pouzet - *Ecole normale supérieure*

NOCS

I.1 A Novel Network Fabric for Efficient Spatio-Temporal Reduction in Flexible DNN Accelerators

Authors:

Francisco Muñoz-Martinez - Universidad de Murcia

Jose L. Abellan - *Universidad Católica San Antonio de Murcia*

Manuel E. Acacio - *Universidad de Murcia*

Tushar Krishna - *Georgia Institute of Technology*

III.1 PlugSMART: A Pluggable Open-Source Module to Implement Multihop Bypass in Networks-on-Chip

Authors:

Alireza Monemi - *Barcelona Supercomputing Center*

Iván Pérez - *University of Cantabria*

Neiel Israel Leyva - *Barcelona Supercomputing Center*

Enrique Vallejo - *University of Cantabria*

Ramón Beivide - *University of Cantabria*

Miquel Moreto - *Barcelona Supercomputing Center*

TUTORIALS SCHEDULE

Friday, October 08

TUTORIALS

9 am - 1 pm, EDT	**T1. QuantumFlow: A Co-Design Framework of Neural Network and Quantum Circuit towards Quantum Advantage** Weiwen Jiang, *George Mason University, USA* Jinjun Xiong, *IBM T.J. Watson Research Centre, USA* Yiyu Shi, *University of Notre Dame, USA*
9 am - 2 pm, EDT	**T2. Fog Computing for Industrial IoT** Alessandro V. Papadopoulos, *Mälardalen University, Sweden* Wilfried Steiner, *TTTech Computertechnik AG, Austria*
9 am - 1 pm, EDT	**T3. Deterministic Reactive Programming for Cyber-Physical Systems Using Lingua Franca** Soroush Bateni, *University of Texas at Dallas, USA* Hokeun Kim, *Hanyang University, South Korea & UC Berkeley, USA* Edward A. Lee, *UC Berkeley, USA* Shaokai Lin, *UC Berkeley, USA* Marten Lohstroh, *UC Berkeley, USA* Christian Menard, *TU Dresden, Germany* Efsane Soyer, *UC Berkeley, USA*

INDUSTRIAL TUTORIALS

9 am - 1 pm, EDT	**IT1. Scalable SoC Architecture for Edge AI Products** Pawan Chhabra, *Intel* Ajay Upadhyaya, *Intel*
	IT2. Integrating Compute Acceleration Into Embedded System Design Using Vitis Parimal Patel, *Xilinx*
	IT3. GPU Code Generation from MATLAB Jack Ferrari, *Mathworks Inc.*

Tutorial T1, 9 am - 1 pm, EDT

QuantumFlow: A Co-Design Framework of Neural Network and Quantum Circuit towards Quantum Advantage

This tutorial presents a recently proposed co-design framework for neural networks and quantum circuits. It comprises (1) a lecture session on quantum machine learning, including the fundamentals of quantum computing, an overview of quantum machine learning, and the mapping of neural networks onto quantum circuits; (2) a hands-on session using IBM Qiskit to program neuron computation with quantum circuits; (3) a lecture and hands-on hybrid session of co-design neural network and quantum circuit to optimize the quantum implementation with the consideration of noise in qubits. The tutorial targets experts and students in artificial intelligence, machine learning, quantum computing, compiler and architecture optimization, and hardware/software co-design and system synthesis.

Along with the rapid development of quantum computers, e.g., Google's Sycamore with 53 qubits and IBM's Hummingbird with 65 qubits, there is a growing interest in pursuing quantum supremacy or advantages over classical computers in various applications. Machine learning is one of the most promising applications because (1) it encounters the computation-/memory-bound on classical computing and (2) the linear algebra at its core is also central to quantum computing.

Organizers and speakers:
Weiwen Jiang, *George Mason University, USA*
Jinjun Xiong, *IBM T.J. Watson Research Centre, USA*
Yiyu Shi, *University of Notre Dame, USA*

Bio:
Weiwen Jiang will join George Mason University as a Tenure-Track Assistant Professor. He is a Postdoctoral Associate at the University of Notre Dame. He received his Ph.D. degree from Chongqing University in 2019. From 2017 to 2019, he was a research scholar at the University of Pittsburgh. His research interest is on hardware and software co-design; in particular, the co-design of neural networks and different hardware accelerators, including mobile devices, FPGA, and ASIC. Most recently, he moves to the co-design of neural networks and quantum circuits. His work demonstrates the quantum advantages for neural networks for the first time, which has been published at Nature Communications. Dr. Jiang's research works have been published in prestigious journals and conferences, including Nature Electronics, Nature Communications, IEEE/ACM Transactions, DAC, ICCAD, ESWEEK, GLSVLSI, etc. He is the receipt of Best Paper Award in ICCD'17 and Best Paper Nominations in DAC'19, CODES+ISSS'19, ASP-DAC'16, and ASP-DAC'20.

Jinjun Xiong is currently the Program Director for Cognitive Computing Systems Research at the IBM Thomas J. Watson Research Center. He is responsible for defining the scientific agenda and strategic directions for advanced cognitive computing systems research across industries, academia and governmental agencies. In that capacity, he co-directs the IBM-Illinois Center for Cognitive Computing Systems Research (C3SR). Prior to that role, Dr. Xiong was a manager of the Smarter Energy group, responsible for the IBM Research's Big Bet Program on Smarter Energy Research, including its strategies and execution. Dr. Xiong received his Ph.D. degree in Electrical Engineering from University of California, Los Angeles in 2006, and since then has been a Research Staff Member with IBM Thomas J. Watson Research Center. His research interests include quantum computing, cognitive computing, big data analytics, deep learning, smarter energy and application of cognitive computing for industrial solutions.

Yiyu Shi is currently an associate professor in the Department of Computer Science and Engineering at the University of Notre Dame, the site director of NSF I/UCRC Alternative and Sustainable Intelligent Computing, and a visiting scientist at Boston Children's Hospital, the primary pediatric program of Harvard Medical School. His current research interests focus on hardware intelligence with biomedical applications. He has published over 200 peer reviewed papers in premier venues such as Nature research journals, including more than a dozen best papers or nominations in top conferences. He was also the recipient of IBM Invention Achievement Award, Japan Society for the Promotion of Science (JSPS) Faculty Invitation Fellowship, Humboldt Research Fellowship, IEEE St. Louis Section Outstanding Educator Award, Academy of Science (St. Louis) Innovation Award, Missouri S&T Faculty Excellence Award, NSF CAREER Award, the Air Force Summer Faculty Fellowship, IEEE Computer Society Mid-Career Research Achievement Award, and Facebook Research Award. ■

Tutorial T2, 9 am - 2 pm, EDT

Fog Computing for Industrial IoT

The tutorial presents novel architectures, hardware and software platforms, applications, and data-driven techniques for enabling fault tolerance in Industrial IoT and Fog Computing systems. It comprises theoretical and hands-on components with a focus on fault tolerance. Researchers and practitioners will hear about the challenges and opportunities of recent developments and trends. Industrial IoT refers to IoT systems used in the industrial areas to interconnect machines. It provides the infrastructure that underpins Smart Energy Grids, Smart Cities, Smart and Green Mobility, Smart Manufacturing, etc., and provides solutions for several societal challenges. Fog Computing is a system-level architecture that distributes resources and services fro computing, storage, control and networking anywhere along the continuum from Cloud to Things. It is one of the key elements of the IIoT and Industry 4.0. Fog Computing will be enabled by the increased usage of IP-protocols, e.g., standardized Deterministic Ethernet solutions from IEEE Time-Sensitive Networking (TSN)4 Task Group, upcoming 5G wireless standards, and interoperability standards such as OPC Unified Architecture (OPC UA). The resulting integrated Fog Computing platform brings computation, communication and storage closer to the edge of the network.

Organizer(s):
Alessandro V. Papadopoulos, *Mälardalen University, Sweden*
Wilfried Steiner, *TTTech Computertechnik AG, Austria*

Speakers:
Wilfried Steiner, *TTTech Computertechnik AG, Austria*
Karl-Erik Årzén, *Lund University, Sweden*
Ivona Brandic, *TU Wien, Austria*
Atakan Aral, *TU Wien, Austria*

Bio

Alessandro V. Papadopoulos is an Associate Professor of Computer Science Systems and Control at Mälardalen University (MDH), Sweden. From 2020, he is scientific advisor of ABB Industrial Automation, Sweden. He received his Ph.D. degree in systems and control from Politecnico di Milano in 2014. His research is focused on the complex real-time and embedded systems, with a focus on automation systems, fog and cloud computing systems. In this area, he has published over 100 peer-reviewed papers. He was awarded with the Swedish Research Council (VR) Starting grant for the project "PSI: Pervasive Self-Optimizing Computing Infrastructures (VR Swedish Research Council)".

Wilfried Steiner is Director of the TTTech Labs which acts as center for strategic research and the center for IPR management within the TTTech Group. Wilfried Steiner holds a degree of Doctor of Technical Sciences and the Venia Docendi in Computer Science, both from the Vienna University of Technology, Austria. His research is focused on dependable cyber-physical systems, in particular in the following domains: automotive, space, aerospace, as well as new energy and industrial automation. Wilfried Steiner designs algorithms and network protocols with real-time, dependability, and security attributes.

Karl-Erik Årzén received his M.Sc. in Electrical Engineering and Ph.D. in Automatic Control from Lund University in 1981 and 1987, respectively. He was appointed as Full Professor in automatic control in 2000. He has also worked for ABB Corporate Research during 1992-1994. His research interests include cyber-physical systems, real-time systems, real-time and embedded control, control of computer systems, and cloud control. He is the vice-coordinator for the Lund part of the ELLIIT strategic research area on IT and mobile communication, co-director for WASP — the Wallenberg AI, Autonomous Systems and Software Program (the single largest research grant within Engineering Sciences in Sweden ever with a total budget of 550 million EUR), and is an elected member of the Royal Swedish Academy of Engineering Sciences.

Ivona Brandic is Professor for High Performance Computing Systems at the Vienna University of Technology. In 2015 she was awarded the FWF START prize, the highest Austrian award for early career researchers. Since 2016 she is a member of the Young Academy of the Austrian Academy of Sciences. Her interests comprise virtualized HPC systems, energy efficient ultra-scale distributed systems, massive-scale data analytics. She published more than 100 scientific journal, magazine and conference publications and she co-authored a textbook on federated and self-manageable Cloud infrastructures. She has served more than 70 program committees among others Supercomputing, CCGrid, EuroPar, and IPDPS.

Atakan Aral is a Senior Postdoctoral Fellow at the Vienna University of Technology (TU Wien). He received a dual M.Sc. degree in Computer Science and Engineering from Politecnico di Milano in 2011 and Istanbul Technical University (ITU) in 2012, and a Ph.D. degree in Computer Engineering from ITU in 2016. His research interests center around resource and reliability management for geo-distributed and virtualized computing systems such as intercloud and edge computing, as well as optimization of the edge computing architecture for AI services. ■

Tutorial T3, 9 am - 1 pm, EDT

Deterministic Reactive Programming for Cyber-Physical Systems Using Lingua Franca

This tutorial presents Lingua Franca (LF), a polyglot coordination language for writing deterministic reactive programs that incorporate, and compile down to, popular mainstream programming languages such as C, C++, Python, and TypeScript. LF does not replace the languages we love, but instead augments them with a deterministic coordination layer and adds missing abstractions. LF transparently exploits parallelism without burdening the application programmer with the well-documented difficulties of achieving predictable behavior in multi-threaded and/or distributed software.

The tutorial targets attendees with a background in developing software for cyber-physical systems (CPS). We will show the workflow involved in developing a few small CPS applications using LF, and the benefits of determinism that can be achieved out of the box. For example, our custom IDE is capable of diagram synthesis, dependency analysis, and cycle detection that can help CPS developers avoid common sources of potential non-determinism in their software. Our framework also supports real-time capabilities such as deadlines and periodic timers, and includes a deterministic interface to interact with the physical world. We will also show some of the runtime capabilities of LF, including the ability to run a coordinated distributed system with little to no programming effort that conserves a semantic notion of time across different machines.

Organizer(s) and speaker(s):

Soroush Bateni, *University of Texas at Dallas, USA*
Hokeun Kim, *Hanyang University, South Korea & UC Berkeley, USA*
Edward A. Lee, *UC Berkeley, USA*
Shaokai Lin, *UC Berkeley, USA*
Marten Lohstroh, *UC Berkeley, USA*
Christian Menard, *TU Dresden, Germany*
Efsane Soyer, *UC Berkeley, USA*

Bio:

Soroush Bateni is a PhD candidate in the Department of Computer Science at the University of Texas at Dallas. His current research focus is on deterministic distributed embedded systems and predictable AI.

Hokeun Kim is joining Hanyang University, Seoul, Korea, in fall 2021 as an assistant professor in Electronic Engineering. He received his Ph.D. degree in EECS from UC Berkeley in 2017 with a focus on IoT security. He continued research on Internet and computer security at Google. His research interests include the IoT, real-time systems, cyber-physical systems, and computer architecture.

Edward A. Lee has been working on embedded software systems for 40 years, and after detours through Yale, MIT, and Bell Labs, landed at Berkeley, where he is now Professor of the Graduate School in EECS. His research is focused on cyber-physical systems, where he strives to make composable, secure, and verifiable systems, with an emphasis on deterministic concurrency and predictable timing.

Shaokai Lin is a PhD student in the EECS department of UC Berkeley. His research interest is in applying formal methods in the design and verification of cyber-physical systems and timing-predictable hardware.

Marten Lohstroh recently obtained his PhD from UC Berkeley where he wrote his dissertation on reactors. He is still affiliated to Berkeley, but now as a postdoc. His work is focused on broadening the impact of reactors through Lingua Franca.

Christian Menard is a PhD student in the Department of Computer Science at TU Dresden in Germany. His research focuses on applying new models of computation in the field of compiler construction, with a particular focus on creating new tool flows for developing automotive applications.

Efsane Soyer is a junior studying computer science at UC Berkeley. ■

Industry Tutorial IT1, 9 am - 1 pm, EDT

Scalable SoC Architecture for Edge AI Products

Deep neural networks (DNNs) are currently widely used for many artificial intelligence (AI) applications including computer vision, speech recognition, and robotics. While DNNs deliver state-of-the-art accuracy on many AI tasks, it comes at the cost of high computational complexity which presents serious scalability and performance-per-watt challenges for traditional CPU/GPU architectures.

The Intel Edge Inference Product (EIP) family of SoC(s) integrates Vision Processing Unit (VPU) IP targeting computer vision and general Deep Learning inferencing applications. Intel EIP SoC(s) offer scalable performance-per-watt needed for low power connected devices such as drones, intelligent security cameras to mid-high-performance Network Video Recorders, AI in a box and Deep Learning edge server cards.

This tutorial provides an overview of the Scalable EIP Architecture (SEA) framework developed to meet the goal of 5-10 TOPS/Watt for various form-factor EIP devices. Specifically, it will talk about state-of-art techniques adopted in SoC memory technology, general purpose compute, On-chip interconnects, power delivery/mgmt., security/virtualization, High speed IO's and chiplet disaggregation for best-in-class secure, flexible and energy-efficient EIP offering from Intel.

Organizers and speakers:
Pawan Chhabra, *Intel*
Ajay Upadhyaya, *Intel*

Bio:
Pawan Chhabra is a Director Engineering in Intel leading IOTG SOC Architecture. He has led SOC Architecture for the Edge Inference Product family and is currently responsible for leading low-cost client architecture. His knowledge spreads over a wide variety of different domains in the SOC covering Memory Technology, Interconnect, MMU, Cache/Coherency and pm to name a few. Pawan has been contributing over 18 years in the VLSI industry and has held IP and SoC Design lead positions, driving teams to first pass silicon of very successful Snapdragon products. Pawan has a B.Tech degree in Electrical Engineering from IIT Delhi India. He has 6 patents to his credit so far.

Ajay Upadhyaya is a Senior Director Engineering in Intel, India leading IOTG SOC Architecture. He has led IP & SOC Architecture in multiple domains ranging from Smartphones, Entry & Mid-tier Client Laptops, IOTG SoC mainly for Industrial applications & PCIE attach accelerator SoC. His knowledge spreads over multiple SoC domains primarily focusing on Networking offloads, Memory Hierarchies, System Memory Management Unit, Virtualization Technologies, dis-aggregated chiplet based product architectures, Compute sub systems. Ajay has been with the VLSI industry for last 24 areas, primarily working with a couple of startups at the start of his career and with Qualcomm, Intel in the last one decade. Ajay has a B.Tech degree in Electronics Engineering from IIT BHU Varanasi India. He has 2 patents approved. ■

Industry Tutorial IT2, 9 am - 1 pm, EDT

Integrating Compute Acceleration Into Embedded System Design Using Vitis

Important Notes:

- You must register separately for this tutorial through https://www.xilinx.com/support/university/workshops/schedule.html. A confirmation email will be sent upon receiving your registration. Please register ONLY if you are sure to attend the tutorial since the number of seats are limited due to hardware board resources.
- This tutorial will involve the use of online hardware resources located on ECELabs site. The design compilation will occur on AWS instances and hence require good and stable network connection. You will use Remote Desktop (RDP) and Chrome browser. Further instructions will be sent couple of days before the tutorial. The AWS instances will have all necessary tools installed to run the labs.
- It will be hard to catch on with the tutorial in the middle. So, it is highly recommended that you join from the start. Joining in the middle is discouraged or prohibited.

Description:

Xilinx is the inventor of the FPGA, programmable SoCs, and now, the ACAP. Our highly flexible programmable silicon, enabled by a suite of advanced software and tools, drives rapid innovation across a wide span of industries and technologies – from consumer to cars to the cloud.

The goal of this tutorial is to introduce the Xilinx Vitis software development environment for designing accelerators for embedded systems using Vitis. Vitis supports C/C++ and OpenCL.

Attendees will have the opportunity to learn how to use these tools, test the tutorial examples on the target boards, and understand how to build their own custom embedded systems.

Topics to be covered:

- Xilinx Vitis development framework and design flows
- Vitis HLS project creation flow, profiling and design analysis
- Integrating Vitis HLS developed IP using Vivado IPI
- Exporting generated hardware to Vitis embedded software development flow (including Linux and Python)
- Introduction to a compossible pipeline architecture

Organizers and speakers:

Parimal Patel, *Xilinx University Program*

Target Audience:

Embedded hardware and software developers. This is an introductory tutorial, but it is expected that attendees have some programming experience and basic understanding of computer architecture.

Xilinx will provide remote access to cloud instances which will be enabled with Xilinx tools. Attendees will also be provided hardware board access either remotely or locally. Attendees must have their own laptop with reasonable screen size to effectively try the tutorials. (Tablet, and Netbook type devices may not be suitable).

Bio:

Parimal Patel received Ph.D. in Electrical and Computer Engineering from the University of Texas at Austin, Texas in 1986 before joining the University of Texas at San Antonio as an Assistant Professor. He was a Full Professor before joining Xilinx. Parimal has always enjoyed teaching and developing new courses. He has been with Xilinx for over 20 years developing new courses, updating current courses, and delivering workshops worldwide. He is actively engaged in providing training in the areas of High-Level Synthesis, Embedded Systems, DSP Design Flow, Dynamic Partial Reconfiguration, Python Productivity on Zynq (PYNQ), and Accelerated Cloud Computing on AWS with Vitis to list a few. ■

Industry Tutorial IT3, 9 am - 1 pm, EDT

GPU Code Generation from MATLAB – Virtual Lab

Requirements (Who Should Attend):
Engineers experienced in MATLAB with a need to deploy algorithms to a GPU for acceleration in MATLAB or a stand-alone system. Knowledge of CUDA and/ or GPU architectures is not required but may be useful.

Each engineer should have a computer, access to a network, and a sufficiently large monitor. It's recommended to have a resolution of greater than 2048×1536. A simple laptop screen will not be enough. Having a separate screen for the WebEx is ideal, though not necessary. No GPU is required. The virtual lab will be run virtually through WebEx and the tools will be provided with GPU-enabled instances of MATLAB Online. MATLAB Online recommends using the Chrome internet browser.

See this link for recommended browsers: https://www.mathworks.com/support/requirements/browser-requirements.html

Description:
Learn how to generate CUDA code automatically from MATLAB to run on NVIDIA desktop and embedded GPUs. MATLAB is the ideal environment for exploring, developing, and prototyping algorithms. GPUs are the hardware of choice for many applications, such as signal, image processing, and deep learning, that benefit from the parallel processing they offer. GPU Coder offers a direct route to transition from MATLAB development to deployment on GPUs via the generation of CUDA code.

This virtual lab will guide you through hands-on exercises design to ramp you up quickly on GPU Coder. Through these exercises, you will experience a typical workflow that can then be applied to your projects.

- Prototype and accelerate implementations with automatic CUDA code generation
- Enhance performance through code refactoring and design pattern pragmas
- Generate CUDA from deep learning networks for acceleration and implementation
- Deploy generated code to desktop and embedded GPUs

Organizers and speakers:
Jack Ferrari, *The MathWorks, USA*

Bio:
Jack Ferrari is a product marketing engineer focused on supporting a group of code generation tools, including GPU Coder, at MathWorks. Jack holds a bachelor's degree in Mechanical Engineering from Boston University. ■

EDUCATION CLASSES

Saturday & Sunday, October 09-10

SATURDAY CLASSES

Sat, Oct 09, 2021 9 am - 11 am, EDT	A1. **Edge AI Systems,** by Prof. Lin Wang, *VU Amsterdam*	A2. **Memory-Centric Computing,** by Prof. Onur Mutlu, *ETH Zurich and CMU*	A3. **Learn to Drive (and Race!) Autonomous Vehicles,** by Prof. Rahul Mangharam, *University of Pennsylvania* and Dr. Johannes Betz, *University of Pennsylvania*
Sat, Oct 09, 2021 11 am – 1 pm, EDT	B1. **TinyML on Edge,** by Prof. Vijay Janapa Reddi, *Harvard University*	B2. **Neural Networks and Accelerator Co-design,** by Dr. Nicolas J Fraser, *Xilinx*	B3: **Face verification using few-shot deep learning** by Prof. Amit Sethi, and Abhijeet Patil, *IIT Bombay*

SUNDAY CLASSES

Sun, Oct 10, 2021 9 am - 11 am, EDT	C1. **Spiking Neural Networks**, by Prof. Priyadarshini Panda, *Yale*	C2. **Neural Network Accelerator Design,** by Prof. Yu Wang, *Tsinghua University*	C3. **Research Reproducibility in Embedded Learning,** by Dr. Romain Jacob, *ETH Zurich*
Sun, Oct 10, 2021 11 am – 1 pm, EDT	D1. **Introduction to Neuromorphic Computing,** by Prof. Helen Li, *Duke University*	D2. **DNNs on FPGAs,** by Prof. Jaesun Seo, *Arizona State University*	D3. **Machine Learning for Manycore System Design and Optimization,** by Prof. Jana Doppa, *Washington State University* and Dr. Biresh Kumar, *Duke University.*

SATURDAY, OCTOBER 09 - EDUCATION

Education Class A1, 9 am - 11 am, EDT

Edge AI Systems

Instructor:
Lin Wang, *VU Amsterdam*

Abstract:
The past decade has witnessed the proliferation of various intelligent applications, such as augmented reality and autonomous driving, powered by modern AI technologies. Meanwhile, these applications, while being computation intensive, introduce new requirements including ultra-low latency and ultra-high reliability. Such a trend has incentivized a paradigm shift from centralized cloud computing to more distributed computing at the network edge. In this lecture we will talk about edge computing systems for AI-based applications. We will focus on a representative AI task called video stream analytics and perform a case study of how to design an efficient edge computing system to support reliable, real-time video analytics. We will also briefly cover advanced topics like edge-centric programming and resource management for achieving the goal of general-purpose edge computing.

Bio:
Lin Wang is an Assistant Professor in the Computer Systems section at VU Amsterdam, The Netherlands. Before he joined VU Amsterdam, he was an Athene Young Investigator of TU Darmstadt. He obtained his PhD in Computer Science from Institute of Computing Technology, Chinese Academy of Sciences and held positions in IMDEA Networks Institute, Spain, and SnT Luxembourg. He is principal investigator of the DFG Collaborative Research Center MAKI. He is generally interested in distributed systems and networking and his current research is focused on programming models and resource management for modern networked systems including edge AI systems, in-network computing, and cyber-physical systems.

Education Class A2, 9 am - 11 am, EDT

Memory-Centric Computing

Instructor:
Onur Mutlu, *ETH Zurich and CMU*

Abstract:
Computing is bottlenecked by data. Large amounts of application data overwhelm storage capability, communication capability, and computation capability of the modern machines we design today. As a result, many key applications' performance, efficiency and scalability are bottlenecked by data movement. In this lecture, we describe three major shortcomings of modern architectures in terms of 1) dealing with data, 2) taking advantage of the vast amounts of data, and 3) exploiting different semantic properties of application data. We argue that an intelligent architecture should be designed to handle data well. We show that handling data well requires designing architectures based on three key principles: 1) data-centric, 2) data-driven, 3) data-aware. We give several examples for how to exploit each of these principles to design a much more efficient and high performance computing system. We especially discuss recent research that aims to fundamentally reduce memory latency and energy, and practically enable computation close to data, with at least two promising novel directions: 1) processing using memory, which exploits analog operational properties of memory chips to perform massively-parallel operations in memory, with low-cost changes, 2) processing near memory, which integrates sophisticated additional processing capability in memory controllers, the logic layer of 3D-stacked memory technologies, or memory chips to enable high memory bandwidth and low memory latency to near-memory logic. We show both types of architectures can enable orders of magnitude improvements in performance and energy consumption of many important workloads, such as graph analytics, database systems, machine learning, video processing. We discuss how to enable adoption of such fundamentally more intelligent architectures, which we believe are key to efficiency, performance, and sustainability. We conclude with some guiding principles for future computing architecture and system designs.

A short accompanying paper, which appeared in DATE 2021, can be found here and serves as recommended reading: https://people.inf.ethz.ch/omutlu/pub/intelligent-architectures-for-intelligent-computingsystems-invited_paper_DATE21.pdf

Bio:
Onur Mutlu is a Professor of Computer Science at ETH Zurich. He is also a faculty member at Carnegie Mellon University, where he previously held the Strecker Early Career Professorship. His current broader research interests are in computer architecture, systems, hardware security, and bioinformatics. A variety of techniques he, along with his group and collaborators, has invented over the years have influenced industry and have been employed in commercial microprocessors and memory/storage systems. He obtained his PhD and MS in ECE from the University of Texas at Austin and BS degrees in Computer

Engineering and Psychology from the University of Michigan, Ann Arbor. He started the Computer Architecture Group at Microsoft Research (2006-2009), and held various product and research positions at Intel Corporation, Advanced Micro Devices, VMware, and Google. He received the IEEE High Performance Computer Architecture Test of Time Award, the IEEE Computer Society Edward J. McCluskey Technical Achievement Award, ACM SIGARCH Maurice Wilkes Award, the inaugural IEEE Computer Society Young Computer Architect Award, the inaugural Intel Early Career Faculty Award, US National Science Foundation CAREER Award, Carnegie Mellon University Ladd Research Award, faculty partnership awards from various companies, and a healthy number of best paper or "Top Pick" paper recognitions at various computer systems, architecture, and security venues. He is an ACM Fellow "for contributions to computer architecture research, especially in memory systems", IEEE Fellow for "contributions to computer architecture research and practice", and an elected member of the Academy of Europe (Academia Europaea). His computer architecture and digital logic design course lectures and materials are freely available on YouTube (https://www.youtube.com/OnurMutluLectures), and his research group makes a wide variety of software and hardware artifacts freely available online (https://safari.ethz.ch/). For more information, please see his webpage at https://people.inf.ethz.ch/omutlu/.

Education Class A3, 9 am - 11 am, EDT

Learn to Drive (and Race!) Autonomous Vehicles

Instructor:
Dr. Rahul Mangharam, *University of Pennsylvania* and Dr. Johannes Betz, *University of Pennsylvania*

Website:
https://f1tenth.org

Abstract:
If you are interested in learning the foundations of autonomous driving, then this course is for you. But don't be worried that this is another theory-only lecture: In this lecture, we will work hands-on with everyone who is interested in the fields of motion planning, control theory, and applied machine learning. From the very beginning we will introduce you to both a simulation environment and real hardware: Our 1/10th-scale autonomous race car called F1TENTH.

First, we will teach you the fundamentals of autonomous driving in the fields of perception, planning and control and how to program them. We highlight three specific capabilities for streamlined algorithm development, testing and validation: a set of simulators, control and verification, and efficient machine-learning algorithm development. We will also show you how to bring an example autonomous driving algorithm to live in both the simulation environment and on the real F1TENTH hardware, an easy-to-use high-performance autonomous racing platform for developing and deploying autonomous driving algorithms.

F1TENTH has a growing community of over 60 universities, 9 international autonomous racing competitions and hands-on course offerings in over a dozen institutions. In addition, by focusing on the racing environment you will also learn how to develop algorithms that operate on the edge of vehicle dynamics: High accelerations, high velocities and high computation frequencies.

The lecture will be delivered in a manner so that engineering and CS students both on undergraduate and graduate level can understand the material. We will highlight pathways to take your research forward in perception, planning and control for autonomous systems.

Bio:
Rahul Mangharam leads xLAB for Safe Autonomous System in the Dept. of Electrical & Systems Engineering and Dept. of Computer & Information Science at the University of Pennsylvania. His interests are in cyber-physical systems at the intersection of formal methods, machine learning and controls. He is the Penn Director for the Department of Transportation's $14MM Mobility21 National University Transportation Center which focuses on technologies for safe and efficient movement of people and goods. Rahul received the 2016 US Presidential Early Career Award (PECASE) from President Obama for his work on Life-Critical Systems. He also received the 2016 Department of Energy's CleanTech Prize (Regional), the 2014 IEEE Benjamin Franklin Key Award, 2013 NSF CAREER Award, 2012 Intel Early Faculty Career Award and was selected by the National Academy of Engineering for the

2012 and 2017 US Frontiers of Engineering. He has won several ACM and IEEE best paper awards.

Johannes Betz is a postdoctoral researcher at the University of Pennsylvania's xLab for Safe Autonomous Systems. His research focuses on holistic software development for autonomous systems at the dynamic limits in extreme and unknown environments. By using modern algorithms from the field of artificial intelligence he is developing new and advanced methods and intelligent algorithms. Based on his additional studies in philosophy he extends current path and behavior planners to ethical autonomous systems.

Education Class B1, 11 am – 1 pm, EDT

TinyML on Edge

Instructor:

Vijay Janapa Reddi, *Harvard University*

Abstract:

Tiny machine learning (TinyML) is a fast-growing field at the intersection of ML algorithms and low-cost embedded systems. TinyML enables on-device analysis of sensor data (vision, audio, IMU, etc.) at ultra-low-power consumption (<1mW). Processing data close to the sensor allows for an expansive new variety of always-on ML use-cases that preserve bandwidth, latency, and energy while improving responsiveness and maintaining privacy. Yet, there are still numerous challenges to address. Tight memory and storage constraints (KBs not MBs), hardware/software heterogeneity, and a lack of relevant large-scale datasets still pose a substantial barrier to developing TinyML applications. This lecture introduces attendees to the fundamental concepts associated with TinyML. It summarizes state-of-the-art approaches, ranging from model design to deploying them efficiently on embedded hardware, highlighting the challenges and opportunities for things in-between. The lecture presents a macroscopic view of the end-to-end ML workflow, rather than focus on any part in isolation to provide a comprehensive systems-level perspective on TinyML. Attendees walk out of the virtual classroom understanding why the future of ML is tiny and bright. Join the 50,000+ TinyML learners that are a part of this journey and help define its future!

Bio:

I am an Associate Professor in the John A. Paulson School of Engineering and Applied Sciences (SEAS) at Harvard University. Prior to joining Harvard University, I was an Associate Professor at The University of Texas at Austin from 2011 to 2018. I started at Harvard University in the Spring of 2019.

My research is centered on mobile and edge-centric computing systems with a rare taste for cloud computing aspects, mostly as it pertains to edge computing or my students' interests. I direct the Edge Computing Lab. I believe in solving computing problems, rather than associating myself with a particular domain or field of computing (i.e., hardware or software). I take great pride in that, and that reflects in my research groups' training. Having said that, I publish in Computer Architecture and Robotics venues.

My teaching focuses on both hardware and software. At the freshman/sophomore level, I teach topics that are related to embedded systems and help explain their relationship to the rest of the connected world (i.e., the Internet of Things). At the junior and senior undergraduate level, I teach Computer Architecture. Going beyond that, at the graduate level, I teach courses that are focused on several edge computing aspects, which covers a variety of state-of-the-art research challenges and issues. Examples include how we design better systems for robotics, autonomous cars, drones, etc.

Relevant reading:

Widening Access to Applied Machine Learning with TinyML (https://arxiv.org/pdf/2106.04008.pdf). Interested in learning more? Take the HarvardX TinyML course for free: https://www.edx.org/professional-certificate/harvardx-tiny-machine-learning

SATURDAY, OCTOBER 09 - EDUCATION

Education Class B2, 11 am – 1 pm, EDT

Neural Networks and Accelerator Co-Design

Instructor:
Dr. Nicolas J Fraser, *Xilinx*

Abstract:
Machine learning algorithms have been gradually displacing traditional programming techniques across multiple domains, including domains that require low-latency and high-throughput, such as telecommunications and networking. Neural networks designed for these applications may require specialised accelerators in order meet the constraints of their deployment environment. During this talk, we will discuss various forms of specialisations that have been leveraged by the industry with their impact on potential applications, flexibility, performance and efficiency. Furthermore, we will discuss how the specialization in hardware architectures can be automated through end-to-end tool flows.

Bio:
Nicholas J. Fraser received the PhD degree at The University of Sydney, Australia in 2020. Currently he's a research scientist at Xilinx Research Labs, Dublin, Ireland. His main research interests include: training of reduced precision neural networks, software / hardware co-design of neural network topologies / accelerators, and audio signal processing.

Education Class B3, 11 am – 1 pm, EDT

Face Verification using Few-shot Deep Learning

Instructor:
Prof. Amit Sethi, and Abhijeet Patil,
IIT Bombay

Abstract:
In this tutorial, we are going to deploy the popular face recognition algorithm FaceNet on a Jetson Nano Developer Kit. Firstly, we will talk about few-shot learning and the methodology used to train FaceNet. We will go through basic building blocks such as metric learning, triplet loss and the training procedure of FaceNet. After understanding the working of FaceNet, we will jump to hardware deployment of face recognition. We will train FaceNet using pyTorch, then we will convert the trained model to half precision (FP16) using TensorRT. We will also demonstrate camera integration to Jetson Nano. After completing all the steps, we will have a working prototype of a face recognition system ready on a portable Jetson Nano device.

Bio:
Amit Sethi is a Professor of Electrical Engineering at IIT Bombay, and a Visiting Instructor of Pathology at UIC. His research group works on computer vision, deep learning, and medical image analysis. His current research is focused on extracting valuable information, such as for prognosis, using deep learning on inexpensive medical modalities. He obtained his PhD in Electrical and Computer Engineering from University of Illinois at Urbana-Champaign with a focus on computer vision and machine learning, and bachelors in Electrical Engineering from IIT Delhi.

Abhijeet Patil is a PhD student at IIT Bombay, and a deep learning engineer at Griffyn Robotech – a startup working with embedded systems, robotics and deep learning in Pune India. His research interest include deep learning, computational pathology, embedded systems, object detection, tracking, and color normalization. ■

SUNDAY, OCTOBER 10 - EDUCATION

Education Class C1, 9 am - 11 am, EDT

Spiking Neural Networks

Instructor:
Priyadarshini Panda, *Yale*

Abstract:
Spiking Neural Networks (SNNs) have recently emerged as an alternative to deep learning due to their huge energy efficiency benefits on neuromorphic hardware. In this presentation, we suggest important techniques for training SNNs which bring a huge benefit in terms of latency, accuracy, interpretability, and robustness. We will first delve into how training is performed in SNNs. Training SNNs with surrogate gradients presents computational benefits due to short latency and is also considered as a more bio-plausible approach. However, due to the non-differentiable nature of spiking neurons, the training becomes problematic and surrogate methods have thus been limited to shallow networks compared to the conversion method. To address this training issue with surrogate gradients, we will also go over a recently proposed method Batch Normalization Through Time (BNTT) that allows us to target interesting beyond traditional image classification applications like video segmentation. with SNNs. Another critical limitation of SNNs is the lack of interpretability. While a considerable amount of attention has been given to optimizing SNNs, the development of explainability still is at its infancy. I will talk about our recent work on a bio-plausible visualization tool for SNNs, called Spike Activation Map (Sam) compatible with BNTT training. The proposed Sam highlights spikes having short inter-spike interval, containing discriminative information for classification. Finally, with proposed BNTT and Sam, I will highlight the robustness aspect of SNNs with respect to adversarial attacks. In the end, I will talk about interesting prospects of SNNs for non-conventional learning scenarios such as, federated and distributed learning.

Bio:
Priya Panda is an assistant professor in the electrical engineering department at Yale University, USA. She received her B.E. and Master's degree from BITS, Pilani, India in 2013 and her PhD from Purdue University, USA in 2019. During her PhD, she interned in Intel Labs where she developed large scale spiking neural network algorithms for benchmarking the Loihi chip. She is the recipient of the 2019 Amazon Research Award. Her research interests include- neuromorphic computing, deep learning and algorithm-hardware co-design for robust and energy efficient machine intelligence.

Education Class C2, 9 am - 11 am, EDT

Neural Network Accelerator Design

Instructor:
Yu Wang, *Tsinghua University*

Abstract:
We have witnessed the rapid growth of Deep Nerual Networks (DNNs) in the past decade. Deep neural network enabling technology has made a great impact in almost every field of our lives, including automatic driving, health care, smart city, social network, and so on. There are various kinds of DNNs, among which Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) are the most popular ones. CNNs have popularized image classification and object detection, while RNNs are used in natural language processing, time series applications, and sentiment analysis. Recently, Graph Neural Networks (GNNs) excel for their capability to generate high-quality node feature vectors (embeddings) using graph-based deep learning method. GNNs are widely applied in recommendation system, social network, and biomedicine.

However, the high computation and storage complexity of neural network inference poses great difficulty on its application. Besides, the ultra-sparsity of GNNs brings great challenges to the computation on the traditional general-purpose platforms, such as CPUs and GPUs. In the past seven years, both academia and industry have devoted a lot of efforts to design Domain Specific Accelerators (DSAs) for DNN applications, so as to achieve low power and high performance deep neural network inference acceleration.

This talk will first give some basic concepts of DNN models, including CNNs, RNNs, and GNNs, from the perspective of algorithm. Secondly, the basic ideas and methodologies of designing DSAs for DNN applications will be introduced. Specifically, this talk will focus on the FPGA based nerual network accelerator designs. Thirdly, some design principles of accelerating GNNs on GPUs and FPGAs will be discussed. Finally, this talk will outline the history and developing trend of DNN accelerators.

Bio:
Dr. Yu Wang is a tenured professor in the Department of Electronic Engineering at Tsinghua University. He is now the Chair of the Department of Electronic Engineering and the Dean of Institute for Electronics and Information Technology in Tianjin at Tsinghua University. He has published more than 70 journals (51 IEEE/ACM journals) and 200 conference papers (15 DAC, 14 DATE, 7 ICCAD, 31 ASP-DAC, 9 FPGA) in the areas of EDA, FPGA, VLSI Design, and Embedded Systems. He has received Best Paper Award in ASP-DAC19, FPGA17, NVMSA17, ISVLSI12, Best Poster Award in HEART12, and 10 Best Paper Nominations. He is a recipient of ACM/SIGDA Meritorious Service Award in 2020, Under-40 Innovators Award at DAC in 2018 (5 all over the world/year), IBM X10 Faculty Award in 2010 (one of 30 worldwide). He is the co-founder of Deephi Tech (a leading deep learning solution provider), which is acquired by Xilinx in 2018. He served as TPC chair for ISVLSI 2018, Program Co-Chair for ICFPT 2019/2011 and Finance Chair of ISLPED 2012-2016, Track Chair for DATE 2017-2019 and GLSVLSI 2018, General Chair Secretary for ASP-DAC 2020, Executive Committee Member for DAC 2020, PC member for leading conferences in these areas. He also served as Co-Editor-in-Chief for ACM SIGDA E-News in 2017-2019, Associate Editor for Journal of Circuits, Systems, and Computers in 2013-2020, and Special Issue Editor for Microelectronics Journal in 2017-2019. Currently he serves as Associate Editor for IEEE TCAD, IEEE TCSVT, ACM TECS, ACM TODAES, IET Computers and Digital Techniques, IEEE Embedded System Letter. He is an ACM/IEEE Senior Member.

Education Class C3, 9 am - 11 am, EDT

Research Reproducibility in Embedded Learning

Instructor:
Romain Jacob, *ETH Zurich*

Abstract:
When designing their performance evaluations, researchers often encounter questions such as: How long should a run be? How many runs to perform? How to account for the variability across multiple runs? What statistical methods should be used to analyze the data? Despite the best intentions, researchers often answer these questions differently, thus impairing the replicability of evaluations and the confidence in the results.

Improving the standards of replicability in embedded systems has recently gained traction within the community. As an important piece of the puzzle, we have developed a systematic methodology that streamlines the design and analysis of performance evaluations, and we have implemented this methodology into a framework called TriScale.

This lecture introduces the main concepts of the methodology and lets you experiment with TriScale. By the end to the lecture, you will be able to:

- Understand the difference between replicability and reproducibility, and why these notions matter;
- Understand why performance evaluation experiments must be replicable to be meaningful;
- Understand the basics of statistics required to assess replicability;
- Answer questions such as "How many times should I repeat my experiment?" rationally;
- Use the TriScale framework to help you design your next experiments, analyze your data, and report your results in a (more) replicable fashion.

The methods and principles underlying TriScale are broadly applicable to performance evaluations in (embedded) systems and networking, including simulations, emulations, and experiments on physical hardware platforms. Throughout the lecture, we draw on networked embedded systems use cases to illustrate the main concepts. Finally, we conclude with a note on the replicability challenges for machine learning, and how TriScale may help to address those as well.

Bio:
Roman Jacob is postdoctoral researcher at ETH Zurich in the group of Prof. Laurent Vanbever. His current research interests are focused on computer networks, communication protocols, (real-time) scheduling theory, and statistics applied to experimental design. He started to work on TriScale from his own need to design sound performance evaluations for low-power wireless communication protocols, which was the topic of his doctoral dissertation, supervised by Prof. Lothar Thiele. At the time, he has been heavily involved in the IoTBench initiative, which aims at designing benchmarks for low-power wireless; and as he learned then, there can be no proper benchmarking without replicability! So here he is.

SUNDAY, OCTOBER 10 - EDUCATION

Education Class D1, 11 am – 1 pm, EDT

Introduction to Neuromorphic Computing

Instructor:
Helen Li, *Duke University*

Abstract:
The human brain is the most sophisticated organ that nature ever builds. Building a machine that can function like a human brain, indubitably, is the ultimate dream of a computer architect. Although we have not yet fully understood the working mechanism of human brains, the part that we have learned in the past seventy years already guided us to many remarkable successes in computing applications, e.g., artificial neural networks and machine learning. Inspired by the working mechanism of the human brain, neuromorphic system naturally possesses a massively parallel architecture with closely coupled memory, offering a great opportunity to break the "memory wall" in von Neumann architecture. The talk will start with a background introduction of neuromorphic computing, followed by examples of hardware acceleration schemes of learning and neural network algorithms and memristor-based computing engine. I will also share our prospects on the future technology challenges and advances of neuromorphic computing.

Bio:
Hai (Helen) Li 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.

Education Class D2, 11 am – 1 pm, EDT

DNNs on FPGAs

Instructor:
Jaesun Seo, *Arizona State University*

Abstract:
Deep neural networks (DNNs) have been successful in many practical applications including image classification, object detection, speech recognition, etc. GPUs have been a popular hardware platform for DNN workloads, aided by highly parallel computing with a massive number of processing cores. However, due to the lack of reconfigurability and high power consumption, GPU is not an ideal accelerator solution for DNN models especially with high sparsity or customized architectures. Application-specific integrated circuits (ASICs) typically have the highest energy efficiency, but their limited configurability can introduce a significant risk of premature obsolescence. With DNN algorithms evolving at a fast pace, ASIC designs will always lag behind the cutting edge due to the long design cycle. To that end, FPGAs have a unique advantage with potentially higher throughput and efficiency than GPUs, while offering faster time-to-market and potentially longer useful life than ASIC solutions.

In this lecture, we will present FPGA-based DNN accelerator designs and methodologies. We will first introduce the basics of DNN algorithms, and hardware requirements for computation and memory. Next, we will present efficient FPGA acceleration schemes, including loop optimization of iterative DNN operations, parallel computation, dataflow and data re-use, minimization of memory access, low-precision quantization, sparsity and

pruning, etc. The lecture will discuss the key design trade-offs to map DNN algorithms on different FPGAs and how to optimally improve the throughput, power, and energy-efficiency for target applications.

Bio:
Jae-sun Seo received the Ph.D. degree in electrical engineering from the University of Michigan in 2010. From 2010 to 2013, he was with IBM T. J. Watson Research Center, where he worked on cognitive computing chips under the DARPA SyNAPSE project and energy-efficient integrated circuits for high-performance processors. In 2014, he joined ASU in the School of ECEE, where he is now an Associate Professor. During the summer of 2015, he was a visiting faculty at Intel Circuits Research Lab. His research interests include efficient hardware design of machine learning / neuromorphic algorithms and integrated power management. Dr. Seo was a recipient of IBM Outstanding Technical Achievement Award (2012), NSF CAREER Award (2017), and Intel Outstanding Researcher Award (2021).

Education Class D3, 11 am – 1 pm, EDT

Machine Learning for Manycore System Design and Optimization

Instructor:
Biresh Kumar, *Duke University* and Jana Doppa, *Washington State University*

Abstract:
Advanced computing systems have long been enablers for breakthroughs in science and engineering applications including Artificial Intelligence (AI) and Machine Learning (ML) either through sheer computational power or form-factor miniaturization. However, as algorithms become more complex and the size of datasets increase, existing computing platforms are no longer sufficient to bridge the gap between algorithmic innovation and hardware design due to nearing the end of Moore's law. This educational lecture will focus on how manycore systems designed by leveraging the benefits of emerging technologies (e.g., three-dimensional integration, processing-in-memory) and machine learning have the potential to bridge this growing gap. First, we will provide the basics on manycore systems — compute layer, interconnect layer, and memory layer — and discuss the challenges of data movement. Second, we will discuss the advantages of heterogeneous computing, 3D integration, and processing-in-memory in addressing the data movement challenge. Third, we will explain the challenges of application-specific manycore systems design due to large combinatorial spaces and the need to perform expensive simulations to optimize multiple conflicting objectives (e.g., power, performance, and temperature). Fourth, we will describe how machine learning can enable efficient and accurate manycore systems design optimization. Finally, we will discuss hardware and software co-design methodologies for emerging deep learning applications.

The lecture will be delivered in a way that any undergraduate student from ECE or CS can understand the material. No background on manycore systems or machine learning is assumed.

Bio:
Jana Doppa is the George and Joan Berry Distinguished Associate Professor in the School of Electrical Engineering and Computer Science at Washington State University, Pullman. He received his Ph.D. degree in Computer Science from Oregon State University and his M.Tech. degree from Indian Institute of Technology (IIT), Kanpur. His primary research focus is at the intersection of machine learning and electronic design automation by exploring the synergies between these two mutually beneficial areas.

His research has been recognized with a number of awards, honors, and distinctions including the 2019 National Science Foundation CAREER Award; the 2021 Early Career Award in AI by the International Joint Conference on Artificial Intelligence for ML algorithms to accelerate design automation for science and engineering applications including electronic design automation; the 2021 Best Paper Award from ACM Transactions on Design Automation of Electronic Systems; the 2013 Outstanding Paper Award from the AAAI Conference on Artificial Intelligence; the 2018 Best Student Abstract Award from the AAAI Conference on Artificial Intelligence; the 2015 Outstanding PhD Dissertation Award from Oregon State University and was nominated for ACM Doctoral Dissertation Award; a 2015 Google Faculty Research Award; the 2013 Outstanding Graduate Student Award from College of Engineering, Oregon State University; the 2020 Outstanding Junior Faculty Research Award and the 2018 Reid-Miller Teaching Excellence Award from the College of Engineering, Washington State University.

Biresh Kumar Joardar is an NSF sponsored Computing Innovation Fellow (postdoctoral researcher) in Electrical and Computer Engineering at Duke University mentored by Prof. Krishnendu Chakrabarty. He received his Ph.D. degree in Computer Engineering from Washington State University and his BE degree in ECE from Jadavpur University. His primary research is at the intersection of manycore systems design and machine learning with a current focus on reliable ML on unreliable hardware.

His research has been recognized with a number of awards, honors, and distinctions including a 2021 DAAD AInet Fellowship; a 2020 NSF Computing Innovation Fellowship; the 2019 Best Paper Award at NOCS; Nominated for Best Paper Award at DATE-2020 and DATE-2021 conferences; the 2019 Outstanding Graduate Student Researcher Award from the College of Engineering, WSU; and the 2018 Harold and Dianna Frank Electrical Engineering Fellowship, WSU. ■

MONDAY PROGRAM

October 11

9:30 am - 9:55 am	**ESWEEK Opening**			
10:00 am - 10:45 am	**Keynote: *Advances in Neuromorphic Computing for Fast, Efficient, and Intelligent Processing*, Mike Davies, Intel Labs**			
10:45 am - 11:00 am	**Break**			
11:00 am - 12:00 pm	**Session 1A - CASES:** Runtime Efficiency: From Photonic Accelerators to Security and NoC Routing	**Session 1B - CODES+ISSS:** Memory and Storage	**Session 1C - EMSOFT:** Real-Time Scheduling and Analysis	**Special Session 1D:** Towards Scalable, Secure, and Smart Mission-Critical IoT Systems: Review and Vision
12:00 pm - 12:15 am	**Break**			
12:15 pm - 01:00 pm		**Session 2B - CODES+ISSS:** Neural Architecture Design Space Exploration	**Session 2C - EMSOFT:** Cyber-Physical Systems: Synthesis and Verification	**Session 2D - EMSOFT:** Resource Management
01:00 pm - 02:00 pm	**Work-in-Progress Poster Session:** CASES, CODES+ISSS and EMSOFT WiP posters			

MONDAY, OCTOBER 11

9:30-9:55 am, EDT
ESWEEK Opening

Speakers:
Andreas Gerstlauer, *University of Texas at Austin*
Aviral Shrivastava, *Arizona State University*

Welcome and opening remarks.

Monday Keynote, 10-10:45 am, EDT
Advances in Neuromorphic Computing for Fast, Efficient, and Intelligent Processing

Speaker:
Mike Davies, *Intel Labs*

Abstract:
The past fifty years have brought enormous progress in computer architecture, semiconductor scaling, and artificial intelligence, yet our computing technology today still lags far behind biological brains in many respects. While deep artificial neural networks have provided breakthroughs in AI, these gains come with heavy compute and data requirements relative to their biological counterparts. Neuromorphic computing aims to narrow this gap by drawing inspiration from the form and function of biological neural circuits, re-thinking computing from transistors to software informed by biological principles. The pace of neuromorphic research has accelerated in recent years, with chips like Intel's Loihi providing, for the first time, compelling quantitative results over a range of workloads—from sensory perception to data efficient learning to closed-loop adaptive control to combinatorial optimization. This progress suggests the approaching commercial viability of a new class of chips that can autonomously process complex data streams, adapt, plan, behave, and learn in real time at extremely low power levels. This talk surveys some of these recent developments and remaining challenges facing the field.

Bio:
Mike Davies is a Senior Principal Engineer in Intel Labs and the Director of Intel's Neuromorphic Computing Lab. Since 2014 he has been researching neuromorphic architectures, algorithms, software, and systems, and has fabricated several neuromorphic chip prototypes, including the Loihi series. He was a founding employee of Fulcrum Microsystems and Director of its silicon engineering group until Intel's acquisition of Fulcrum in 2011. He led the development of four generations of low latency, highly integrated Ethernet switches using Fulcrum's proprietary asynchronous design methodology. He received B.S. and M.S. degrees from Caltech in 1998 and 2000, respectively.

11:00 am - 12:00 pm, EDT

Session 1A - CASES: Runtime Efficiency: From Photonic Accelerators to Security and NoC Routing

Chair:
Paul Bogdan, *University of Southern California*

Co-Chair:
Sudeep Pasricha, *Colorado State University*

Runtime interventions are critical for squeezing out maximum performance and maintaining system integrity in modern complex systems-on-chip. This session presents three papers that explore different design aspects: systolic arrays, silicon photonics-based DNN acceleration, hardware-assisted control flow integrity checking, and the co-optimization of NoC routing with task mapping.

11:00 am - 11:15 am, EDT, 1A.1
*** RiSA: A Reinforced Systolic Array for Depthwise Convolutions and Embedded Tensor Reshaping**
Hyungmin Cho - *Sungkyunkwan University*

11:15 am - 11:30 am, EDT, 1A.2
ROBIN: A Robust Optical Binary Neural Network Accelerator
Febin Sunny, Mirza Asif, Mahdi Nikdast, Sudeep Pasricha - *Colorado State University*

11:30 am - 11:45 am, EDT, 1A.3
Comparative Analysis and Enhancement of CFG-based Hardware-Assisted CFI Schemes
Mario Telesklav and Stefan Tauner - *TU Wien*

11:45 am - 12:00 pm, EDT, 1A.4
MARCO: A High-performance Task Mapping and Routing Co-optimization Framework for Point-to-Point NoC-based Heterogeneous Computing Systems
Hui Chen, Zihao Zhang, Peng Chen, Xiangzhong Luo, Shiqing Li, Weichen Liu - *Nanyang Technological University*

* Denotes Best Paper Candidate

11:00 am - 12:00 pm, EDT
Session 1B - CODES+ISSS: Memory and Storage

Chair:
Fei Wu, *Huazhong University of Science and Technology*

Co-Chair:
Ishan Thakkar, *University of Kentucky*

Memory has been the bottleneck of modern computing platforms. This session presents recent advances on in-memory computing and reliability enhancement of storage devices.

11:00 am - 11:15 am, EDT, 1B.1
SIam: Chiplet-based Scalable In-Memory Acceleration with Mesh for Deep Neural Networks
Gokul Krishnan - *Arizona State University*
Sumit Mandal - *University of Wisconsin-Madison*
Manvitha Pannala - *Arizona State University*
Chaitali Chakrabarti - *Arizona State University*
Jae-sun Seo - *Arizona State University*
Umit Ogras - *University of Wisconsin-Madison*
Yu Cao - *Arizona State University*

11:15 am - 11:30 am, EDT, 1B.2
*** Exploring Efficient Architectures for Remote In-Memory NVM over RDMA**
Qingfeng Zhuge, Hao Zhang, Edwin Hsing-Mean Sha, Rui Xu - *East China Normal University*
Jun Liu, Shengyu Zhang - *Lenovo Ltd.*

11:30 am - 11:45 am, EDT, 1B.3
Data Pattern Aware Reliability Enhancement Scheme for 3D Solid-State Drives
Shiqiang Nie - *Xi'an Jiaotong University*
Weiguo Wu - *Xi'an Jiaotong University*
Chi Zang - *Xi'an Jiaotong University*

* Denotes Best Paper Candidate

11:00 am - 12:00 pm, EDT
Session 1C - EMSOFT: Real-Time Scheduling and Analysis

Chair:
Enrico Bini, *University of Turin*

Co-Chair:
Bryan Ward, *MIT Lincoln Laboratory*

This session presents new scheduling algorithms and timing analysis methods for real-time systems, with a focus on complex task models and multiprocessor platforms. The first two papers present a hierarchical resource sharing protocol and a federated scheduling algorithm for DAG-based tasks on multiprocessors, respectively. The third presents a schedulability analysis for task automata that integrates demand bound analysis and path abstraction to improve efficiency. The fourth paper proposes a novel response time analysis for global fixed-priority scheduling on multiprocessors that considers parallel execution to more accurately account for resource interference among tasks.

11:00 am - 11:15 am, EDT, 1C.1
A Hierarchical Hybrid Locking Protocol for Parallel Real-Time Tasks
Zewei Chen, Hang Lei, Maolin Yang, Yong Liao - *University of Electronic Science and Technology of China*
Lei Qiao - *Beijing Institute of Control Engineering*

11:15 am - 11:30 am, EDT, 1C.2
Federated scheduling of sporadic DAGs on unrelated multiprocessors
Petros Voudouris, Per Stenström, Risat Mahmud Pathan - *Chalmers University of Technology*

11:30 am - 11:45 am, EDT, 1C.3
Feasibility Analysis for Timed Automata with Tasks
Jinghao Sun - *Dalian University of Technology*
Nan Guan - *City University of Hong Kong*
Rongxiao Shi - *Northeastern University*
Guozhen Tan - *Dalian University of Technology*
Wang Yi - *Uppsala University*

11:45 am - 12:00 pm, EDT, 1C.4
*** Excluding Parallel Execution to Improve Global Fixed Priority Response Time Analysis**
Quan Zhou, Jianjun Li, Guohui Li - *Huazhong University of Science and Technology*

* Denotes Best Paper Candidate

11:00 am - 12:00 pm, EDT

Special Session 1D: Towards Scalable, Secure, and Smart Mission-Critical IoT Systems: Review and Vision

Organizers and Chairs:
Xun Jiao, *Villanova University*
Yier Jin, *University of Florida*
X. Sharon Hu, *University of Notre Dame*

The growing capabilities of sensing, computing and communication devices are leading to an explosion of Internet of Things (IoT). Somewhat orthogonally, emerging technologies such as artificial intelligence have been promising enormous economic and societal benefits. While it is naturally desirable to deploy these technologies in IoT infrastructures, such deployments present daunting challenges for increasingly scaling-up IoT infrastructures in mission-critical applications such as medical, energy, transportation, and industrial automation systems. These challenges pose immediate threat to the performance, efficiency, and dependability of scaling-up IoT infrastructure. The challenges mainly stem from the following sources: (i) the number of IoT devices is growing rapidly, making centralized management infeasible; (ii) the level of heterogeneity in the form of diverse systems, architectures and hardware components is also increasing unprecedentedly, which render one-size-fit-all design techniques less effective or even futile; (iii) as more and more IoT infrastructure deployed in safety-critical applications, dependability including reliability, robustness, and security have been even more challenging to guarantee. This special session aims to shed light on such emerging challenges in achieving a scalable and secure IoT systems and to present recent advances that span the entire software-hardware stack in tackling these challenges. Supported by the National Science Foundation (NSF) PPoSS program, a team of researchers with diverse and complementary backgrounds are formed to address the challenges from a cross-layer perspective. This special session will include four talks that looks into different areas of IoT system including theory & algorithms, computer architecture & systems, resource management, and security & privacy.

11:00 am - 11:15 am, EDT, 1D.1
Data-Driven Modeling and Control of QoS in Safety-Critical IoT Systems
Michael Lemmon - *University of Notre Dame*

11:15 am - 11:30 am, EDT, 1D.2
Dynamic and Distributed Resource Management for Mission-Critical IoT Networks
Song Han - *University of Connecticut*

11:30 am - 11:45 am, EDT, 1D.3
Vulnerability Detection for IoT System through Formal Method
Yier Jin - *University of Florida*

11:45 am - 12:00 pm, EDT, 1D.4
Real-Time Adaptive Sensor Attack Detection for Safety-Critical IoT Systems
Fanxin Kong - *Syracuse University*

12:15 pm - 01:00 pm, EDT
Session 2B - CODES+ISSS: Neural Architecture Design Space Exploration

Chair:
Amit Singh, *University of Essex*

Co-Chair:
Xiang Chen, *George Mason University*

Deep learning and neural networks are becoming a key component of embedded systems across numerous application domains. This session presents design space exploration methods to assist in designing neural network accelerators.

12:15 pm - 12:30 pm, EDT, 2B.1
FLASH: Fast Neural Architecture Search with Hardware Optimization
Guihong Li - *The University of Texas at Austin*
Sumit Mandal - *University of Wisconsin-Madison*
Radu Marculescu - *The University of Texas at Austin*
Umit Ogras - *University of Wisconsin-Madison*

12:30 pm - 12:45 pm, EDT, 2B.2
*** Intermittent-Aware Neural Architecture Search**
Hashan Mendis, Chih-Kai Kang, Pi-Cheng Hsiu - *Academia Sinica*

12:45 pm - 02:00 pm, EDT, 2B.3
Algorithm -Hardware Co-design of Attention Mechanism on FPGA Devices
Xinyi Zhang - *University of Pittsburgh*
Yawen Wu - *University of Pittsburgh*
Peipei Zhou - *University of California, Los Angeles*
Xulong Tang - *University of Pittsburgh*
Jingtong Hu - *University of Pittsburgh*

* Denotes Best Paper Candidate

12:15 pm - 01:00 pm, EDT
Session 2C - EMSOFT: Cyber-Physical Systems: Synthesis and Verification

Chair:
Pavithra Prabhakar, *Kansas State University*

Co-Chair:
Susmit Jha, *SRI International*

This session presents novel methods for the synthesis and verification of cyber-physical systems. The first paper proposes a framework for compositional safety verification and learning of neural network controllers. The second paper presents an approach to synthesizing feedback controllers for CPS based on signal temporal logic specifications. The third paper introduces a gray-box falsification approach for closed-loop systems with simple open-loop plants but complex feedback controllers.

12:15 pm - 12:30 pm, EDT, 2C.1
Compositional Learning and Verification of Neural Network Controllers
Radoslav Ivanov, Kishor Jothimurugan, Steve Hsu, Shaan Vaidya, Rajeev Alur, Osbert Bastani - *University of Pennsylvania*

12:30 pm - 12:45 pm, EDT, 2C.2
Specification Guided Automated Synthesis of Feedback Controllers
Nikhil Kumar Singh and Indranil Saha - *Indian Institute of Technology Kanpur*

12:45 pm - 01:00 pm, EDT, 2C.3
Synthesis-guided Adversarial Scenario Generation for Gray-box Feedback Control Systems with Sensing Imperfections
Liren Yang and Necmiye Ozay - *University of Michigan*

12:15 pm - 01:00 pm, EDT
Session 2D - EMSOFT: Resource Management

Chair:
Nan Guan, *City University of Hong Kong*

Co-Chair:
Renato Mancuso, *Boston University*

This session presents novel resource management methods for heteregeneous embedded systems. The first paper presents a time-predictable, secure system for consolidating safety-critical and non-critical automotive software components onto a shared platform by integrating RTOS and Linux on a partitioning hypervisor. The second paper proposes a procrastination strategy to improve energy saving in embedded architectures with volatile and non-volatile memory. The last one presents a swap-based multitasking scheme to improve user experience and energy consumption in Android devices.

12:15 pm - 12:30 pm, EDT, 2D.1
*** Towards an Integrated Vehicle Management System in DriveOS**
Soham Sinha and Richard West - *Boston University*

12:30 pm - 12:45 pm, EDT, 2D.2
HEART: Hybrid memory and Energy-Aware Real-Time scheduling for multi-processor systems
Mario Guenzel, Christian Hakert, Kuan-Hsun Chen, Jian-Jia Chen - *Technical University of Dortmund*

12:45 pm - 01:00 pm, EDT, 2D.3
Killing Processes or Killing Flash? Escaping from the Dilemma Using Lightweight, Compression-Aware Swap for Mobile Devices
Yong-Xuan Wang, Chung-Hsuan Tsai, Li-Pin Chang - *National Yang Ming Chiao Tung University*

* Denotes Best Paper Candidate

MONDAY, OCTOBER 11

01:00 pm - 02:00 pm, EDT

Work-in-Progress Poster Session

CASES WiP 1

Work-in-Progress: Fast Generation of Optimized Execution Plans for Parameterizable CNN Accelerators

Hyemi Min, Jungyoon Kwon, Bernhard Egger - *Seoul National University*

CASES WiP 2

Work-in-Progress: Multiple Approximate Instances in Neural Processing Units for Energy-Efficient Circuit Synthesis

Tanfer Alan - *Karlsruhe Institute of Technology*
Jorge Castro-Godínez - *Instituto Tecnológico de Costa Rica*
Joerg Henkel - *Karlsruhe Institute of Technology*

CODES+ISSS WiP 1

Work-in-Progress: The RISC-V Instruction Set Architecture Optimization and Fixed-point Math Library Co-design

Liu Meng - *Beijing University of Technology*

CODES+ISSS WiP 2

Work-in-Progress: A Study of Transistor Degradation in Cyber-Physical System Control Devices

Spencer Millican and SueAnne Griffith - *Auburn University*

CODES+ISSS WiP 3

Work-in-Progress: Learned Buffer Management: A New Frontier

Yigui Yuan and Peiquan Jin - *University of Science and Technology of China*

CODES+ISSS WiP 4

Work-in-Progress: Achieving Fast Lane Detection of Autonomous Driving by CNN based Differentiation

Xingzhi Zhou, Jinyu Zhan, Wei Jiang - *School of Information and Software Engineering, University of Electronic Science and Technology of China*

CODES+ISSS WiP 5

Work-in-Progress: Physically Realizable Backdoor Attacks on 3D Point Cloud Deep Learning

Chen Bian, Wei Jiang, Jinyu Zhan, Xiangyu Wen, Ziwei Song, Hong Lei - *School of Information and Software Engineering, University of Electronic Science and Technology of China*

CODES+ISSS WiP 6

Work-in-Progress: Early Power Estimation of CUDA-based CNNs on GPGPUs

Christopher Metz, Mehran Goli, *Rolf Drechsler - University Bremen*

CODES+ISSS WiP 7

Work-in-Progress: Register File Management for PRET Machines

Martin Kostal and Michal Sojka - *Czech Technical University in Prague*

CODES+ISSS WiP 8

Work-in-Progress: Critical-Weight Based Locking Scheme for DNN IP Protection in Edge Computing

Ziwei Song, Wei Jiang, Jinyu Zhan, Xiangyu Wen, Chen Bian - *School of Information and Software Engineering, University of Electronic Science and Technology of China*

CODES+ISSS WiP 9

Work-in-Progress: Improving Security and Maintainability in Modular Embedded Systems with Hardware Support

Leandro Batista Ribeiro, Maja Malenko, Marcel Baunach - *Graz University of Technology*

EMSOFT WiP 1

Work-in-Progress: Generative Strategy based Backdoor Attacks to 3D Point Clouds

Xiangyu Wen, Wei Jiang, Jinyu Zhan, Chen Bian, Ziwei Song - *School of Information and Software Engineering, Universiy of Electronic Science and Technology of China*

MONDAY, OCTOBER 11

EMSOFT WiP 2
Work-in-Progress: OHTLoc: An Online Heterogeneous Transfer Method on WiFi-based Indoor Localization System
Ufei Han and Chen Bian - *School of Information and Software Engineering, University of Electronic Science and Technology of China*

EMSOFT WiP 3
Work-in-Progress: WCET-Aware Reachability for Verified Simplex Design
Ole Lübke and Sibylle Schupp - *Hamburg University of Technology*

EMSOFT WiP 4
Work-in-Progress: Timing Diversity as a Protective Mechanism
Mischa Möstl, Robin Hapka, Anika Christmann, Rolf Ernst - *TU Braunschweig*

EMSOFT WiP 5
Work-in-Progress: Detecting Deepfake Video by Visual-Audio Synchronism
Zhufeng Fan, Jinyu Zhan, Wei Jiang - *School of Information and Software Engineering, University of Electronic Science and Technology of China*

EMSOFT WiP 6
Work-in-Progress: Improving fault tolerance of DNNs through weight remapping based on gaussian distribution
Ruoxu Sun, Jinyu Zhan, Wei Jiang, Yucheng Jiang - *School of Information and Software Engineering, University of Electronic Science and Technology of China*

EMSOFT WiP 7
Work-in-Progress: Large-scale Timer Hardware Analysis for a Flexible Low-level Timer-API Design
Niels Gandraß, Michel Rottleuthner, Thomas Schmidt - *Hamburg University of Applied Sciences*

EMSOFT WiP 8
Work-in-Progress: Performance Analysis and Optimization of Decision Tree Classifiers on Embedded Devices
Anish Krishnakumar and Umit Ogras - *University of Wisconsin-Madison*

EMSOFT WiP 9
Work-in-Progress: Determining MPSoC Layout from Thermal Camera Images
Michal Sojka, Ondřej Benedikt, Zdeněk Hanzálek - *Czech Technical University in Prague*

EMSOFT WiP 10
Work-in-Progress: Towards Assurance Case Evidence Generation through Search Based Testing
Yumeng Cao, Quinn Thibeault, Aniruddh Chandratre, Georgios Fainekos, Giulia Pedrielli - *Arizona State University*
Mauricio Effen - *Lockheed Martin Corporation*

EMSOFT WiP 11
Work-in-Progress: The Cyber-Physical Immune System
Bo Pang, Ashank Verma, Jingchao Zhou, Inigo Incer, Alberto Sangiovanni-Vincentelli - *University of California, Berkeley*

EMSOFT WiP 12
Work-in-Progress: An Energy-Aware Optimization Model for Real-Time Systems Analysis and Design
Suzanne Elashri and Akramul Azim - *University of Ontario Institute of Technology (UOIT)*

EMSOFT WiP 13
Work-in-Progress: Strong APA Scheduling in a Real-Time Operating System
Richi Dubey - *Birla Institute of Technology and Science, Pilani*
Vijay Banerjee, Sena Hounsinou, Gedare Bloom - *University of Colorado Colorado Springs*

TUESDAY PROGRAM

October 12

8:00 am - 9:20 am	**Industry Session 1:** Technology Transfer Experiments in the TETRAMAX and BOWI Projects		**Industry Session 2:** Model-Driven System-Performance Engineering for Cyber-Physical Systems	
9:20 am - 9:30 am	**Break**			
9:30 am - 9:55 am	**Test of Time Awards**			
10:00 am - 10:45 am	**Keynote: *Safe Learning in Robotics*** **Claire Tomlin, University of California, Berkeley**			
10:45 am - 11:00 am	**Break**			
11:00 am - 12:00 pm	**Session 3A - CASES:** Domain-specific Embedded System Architectures and Design Methodologies	**Session 3B - CODES+ISSS:** Embedd System for Next-Generation Automotive	**Session 3C - EMSOFT:** Modeling Languages and Tools	**Special Session 3D:** Emergent Design Challenges for Embedded Systems and Paths Forward: Mixed-criticality, Energy, Reliability and Security Perspectives
12:00 pm - 12:15 am	**Break**			
12:15 pm - 01:15 pm		**Session 4B - CODES+ISSS:** Embedded System Design	**Session 4C - EMSOFT:** Testing and Fault Tolerance	**Session 4D - EMSOFT:** Cyber-Physical Systems: Handling Uncertainty and Attacks
01:15 pm - 01:30 pm	**Break**			
01:30 pm - 02:15 pm	**Industry Pitch Session – Trends and Challenges**			

TUESDAY, OCTOBER 12

08:00 am - 09:00 am, EDT
Industry Session 1: Technology Transfer Experiments in the TETRAMAX and BOWI Projects

Chair:
Timothy Bourke, *INRIA*
Rainer Leupers, *RWTH Aachen University, Germany*
Kristina Karanikolova, *TNO, The Netherlands*

This session showcases technology transfer experiments funded by the European Horizon 2020 innovation actions TETRAMAX ("Technology Transfer via Multinational Application Experiments") and BOWI ("Boosting Widening Digital Innovation Hubs"). The experiments are the result of collaboration between academic researchers and small or medium enterprises in different EU countries.

08:00 am - 09:20 am, EDT
Industry Session 2: Model-Driven System-Performance Engineering for Cyber-Physical Systems

Chair:
Twan Basten, *Eindhoven University of Technology*

System performance often brings the competitive advantage for high-tech cyber-physical systems like semiconductor equipment, analytical instruments, and medical equipment. System-Performance Engineering (SysPE) encompasses modeling formalisms, methods, techniques, and industrial practices to design systems for performance, where performance is taken integrally into account during the whole system life cycle. Due to the rapidly increasing complexity of systems, there is a need to develop and establish model-driven SysPE methods and techniques.
In the first talk of this session, we present (1) industrial challenges motivating the importance of SysPE, (2) scientific challenges that need to be addressed to establish model-driven SysPE, (3) important focus areas for SysPE, (4) SysPE best practices, and (5) validation results from a survey conducted with the international industrial and academic community. The second talk presents why and how ITEC, Nexperia, a world-leading manufacturer of semiconductor equipment, is moving towards model-driven system-level development. The session ends with a moderated Q&A.

08:00 am - 08:40 am, EDT
Model-Driven System-Performance Engineering for CPS – challenges, focus areas, best practices, and community feedback
Twan Basten – *Eindhoven University of Technology & ESI (TNO), Eindhoven, Netherlands*
Bram van der Sanden – *ESI (TNO), Eindhoven, Netherlands*

08:40 am - 09:00 am, EDT
Model-Driven System-Level Design @ ITEC, Nexperia

09:00 am - 09:20 am, EDT
Interactive Q&A
Moderator: Samarjit Chakraborty - *UNC Chapel Hill*

TUESDAY, OCTOBER 12

09:30 am - 09:55 am, EDT
Test of Time Awards
Presentation of the Test-of-Time Awards.

Tuesday Keynote, 10 am -10:45 am, EDT
Safe Learning in Robotics

Speaker:
Claire Tomlin, *University of California at Berkeley*

Abstract:
In many applications of autonomy in robotics, guarantees that constraints are satisfied throughout the learning process are paramount. We present a controller synthesis technique based on the computation of reachable sets, using optimal control and game theory. Then, we present methods for combining reachability with learning-based methods, to enable performance improvement while maintaining safety and to move towards safe robot control with learned models of the dynamics and the environment. We will illustrate these "safe learning" methods on robotic platforms at Berkeley, including demonstrations of motion planning around people, and navigating in a priori unknown environments.

Bio:
Claire Tomlin is the Charles A. Desoer Professor and Chair of Electrical Engineering and Computer Sciences at the University of California at Berkeley. Claire received her BASc in EE from the University of Waterloo in 1992, her M.Sc. in EE from Imperial College, London, in 1993, and her PhD in EECS from Berkeley in 1998. She held the positions of Assistant, Associate, and Full Professor at Stanford from 1998-2007, and in 2005 joined Berkeley. Claire works in hybrid systems and control, and integrates machine learning methods with control theoretic methods in the field of safe learning. She works in the applications of air traffic and unmanned air vehicle systems. Claire is a MacArthur Foundation Fellow, an IEEE Fellow, and an AIMBE Fellow. She was awarded the Donald P. Eckman Award of the American Automatic Control Council in 2003, an Honorary Doctorate from KTH in 2016, and in 2017 she won the IEEE Transportation Technologies Award. In 2019, she was elected to the National Academy of Engineering and the American Academy of Arts and Sciences.

11:00 am - 12:00 pm, EDT

Session 3A - CASES: Domain-specific Embedded System Architectures and Design Methodologies

Chair:
Jingtong Hu, *University of Pittsburgh*

Co-Chair:
Ganapati Bhat, *Washington State University*

Domain-specific systems-on-chip (DSSoCs) can bridge the gap between the power, performance, and energy efficiency of custom chips and general-purpose processors. However, this promise relies on addressing fundamental problems in architectures, compilers, hardware accelerators, and software. To this end, this session presents papers on domain-specific architectures, optimization algorithms, and reconfigurable devices.

11:00 am - 11:15 am, EDT, 3A.1
CICERO: A Domain-Specific Architecture for Efficient Regular Expression Matching
Daniele Parravicini, Davide Conficconi, Emanuele Del Sozzo, Christian Pilato, Marco D. Santambrogio - *Politecnico di Milano*

11:15 am - 11:30 am, EDT, 3A.2
Domain-specific hybrid mapping for energy-efficient baseband processing in wireless networks
Robert Khasanov - *TU Dresden*
Julian Robledo - *TU Dresden*
Christian Menard - *TU Dresden*
Andres Goens - *Barkhausen Institut*
Jeronimo Castrillon - *TU Dresden*

11:30 am - 11:45 am, EDT, 3A.3
*** Two Birds with One Stone: Boosting Both Read and Write Performance for Tree Indices on Persistent Memory**
Yongping Luo, Peiquan Jin, Zhou Zhang, Juncheng Zhang - *University of Science and Technology of China*
Qinglin Zhang, Bin Cheng - *Tencent*

11:45 am - 12:00 pm, EDT, 3A.4
*** You Only Traverse Twice: A YOTT Placement, Routing, and Timing Approach for CGRAs**
Michael Canesche - *Universidade Federal de Viçosa (UFV)*
Westerley Carvalho - *Universidade Federal de Viçosa (UFV)*
Lucas Reis - *Universidade Federal de Viçosa (UFV)*
Matheus Oliveira - *Universidade Federal de Viçosa (UFV)*
Salles Magalhães - Universidade Federal de *Viçosa* (UFV)
Peter Jamieson - *Miami University*
José Augusto Nacif - *Universidade Federal de Viçosa (UFV)*
Ricardo Ferreira - *Universidade Federal de Viçosa (UFV)*

* Denotes Best Paper Candidate

11:00 am - 12:00 pm, EDT
Session 3B - CODES+ISSS: Embedd System for Next-Generation Automotive

Chair:
Fadi Kurdahi, *University of California, Irvine*

Co-Chair:
Robert Dick, *University of Michigan*

Next-generation automotive requires significant embedded system support. This session includes three papers from the perspective of security, efficiency and benchmarking for next-generation automotive embedded systems.

11:00 am - 11:15 am, EDT, 3B.1
LATTE: LSTM Self-Attention based Anomaly Detection in Embedded Automotive Platforms
Vipin Kumar Kukkala, Sooryaa Vignesh Thiruloga Chander, Sudeep Pasricha - *Colorado State University*

11:15 am - 11:30 am, EDT, 3B.2
SAGE: A Split-Architecture Methodology for Efficient End-to-End Autonomous Vehicle Control
Arnav Malawade, Mohanad Odema, Sebastien Lajeunesse-Degroot, Mohammad Al Faruque - *University of California Irvine*

11:30 am - 11:45 am, EDT, 3B.3
Chauffeur: Benchmark Suite for Design and End-to-End Analysis of Self-Driving Vehicles on Embedded Systems
Biswadip Maity - *University of California, Irvine*
Saehanseul Yi - *University of California, Irvine*
Dongjoo Seo - *University of California, Irvine*
Leming Cheng - *University of California, Irvine*
Sung-Soo Lim - *Kookmin University*
Jong-Chan Kim - *Kookmin University*
Bryan Donyanavard - *San Diego State University*
Nikil Dutt - *University of California, Irvine*

11:00 am - 12:00 pm, EDT
Session 3C - EMSOFT: Modeling Languages and Tools

Chair:
Stavros Tripakis, *Northeasten University*

Co-Chair:
Ichiro Hasuo, *National Institute of Informatics*

This session focuses on new modeling languages and tools for embedded systems. The first paper presents extensions of the Velus compiler for the synchronous dataflow language Lustre to support normalization and subsampling of Lustre programs, along with an associated end-to-end correctness proof based on Coq. The second paper proposes algorithms for learning non-deterministic real-time automata through residual real-time automata to preserve the same expressiveness while overcoming the non-uniqueness of minimal representations. The third paper presents a regime inference technique to automatically generate sound and efficient floating-point implementations of a given program. The last paper proposes a modeling language and a tool for synthesizing safe and efficient power sequences for modern computing platforms.

11:00 am - 11:15 am, EDT, 3C.1
*** Verified Lustre Normalization with Node Subsampling**
Timothy Bourke, Basile Pesin, Paul Jeanmaire - *Inria / ENS*
Marc Pouzet - *Ecole normale supérieure*

11:15 am - 11:30 am, EDT, 3C.2
Learning Nondeterministic Real-Time Automata
Jie An - *Max Planck Institute for Software Systems*
Bohua Zhan, Naijun Zhan - *Institute of Software, Chinese Academy of Sciences*
Miaomiao Zhang - *School of Software Engineering, Tongji University*

11:30 am - 11:45 am, EDT, 3C.3
Regime Inference for Sound Floating-Point Optimizations
Robert Rabe, Anastasiia Izycheva - *TU Munich*
Eva Darulova - *MPI-SWS*

11:45 am - 12:00 pm, EDT, 3C.4
Declarative Power Sequencing
Jasmin Schult - *ETH Zürich*
Daniel Schwyn - *ETH Zürich*
Michael Giardino - *ETH Zürich*
David Cock - *ETH Zürich*
Reto Achermann - *University of British Columbia*
Timothy Roscoe - *ETH Zürich*

* Denotes Best Paper Candidate

11:00 am - 12:00 pm, EDT

Special Session 3D: Emergent Design Challenges for Embedded Systems and Paths Forward: Mixed-criticality, Energy, Reliability and Security Perspectives

Organizer and Chair:
Amit Kumar Singh, *University of Essex, UK*

Modern embedded systems need to cater for several needs depending upon the application domain in which they are deployed. For example, mixed-critically needs to be considered for real-time and safety-critical systems and energy for battery-operated systems. At the same time, many of these systems demand for their reliability and security as well. With electronic systems being used for increasingly varying type of applications, novel challenges have emerged. For example, with the use of embedded systems in increasingly complex applications that execute tasks with varying priorities, mixed-criticality systems present unique challenges to designing reliable systems. The large design space involved in implementing cross-layer reliability in heterogeneous systems, particularly for mixed-critical systems, poses new research problems. Further, malicious security attacks on these systems pose additional extraordinary challenges in the system design. In this proposed special hot-topic session, we aim to cover both the industry and academia perspectives of the challenges posed by these emergent aspects of system design towards designing high-performance, energy-efficient, reliable and/or secure embedded systems.

11:00 am - 11:15 am, EDT, 3D.1
Mixed-criticality Systems: Theory vs. Practice
Martin Decky - *Dresden Research Center of Huawei, Germany*

11:15 am - 11:30 am, EDT, 3D.2
Energy-Driven Design
Geoff Merrett - *University of Southampton, UK*

11:30 am - 11:45 am, EDT, 3D.3
Cross-layer Reliability-aware Design
Akash Kumar - *TU Dresden, Germany*

11:45 am - 12:00 pm, EDT, 3D.4
Aging and Covert Channel Attacks and Mitigations
Amit Kumar Singh - *University of Essex, UK*

12:15 pm - 01:15 pm, EDT
Session 4B - CODES+ISSS: Embedded System Design

Chair:
Luciano Lavagno, *Politecnico di Torino*

Co-Chair:
Ann Ramirez, *University of Florida*

The increasing complexity of embedded systems demands efficient system design methodologies. This session presents co-design, scheduling and synchronization methors that can optimize design for different underlying architectures.

12:15 pm - 12:30 pm, EDT, 4B.1
*** MARS: mmWave based Assistive Rehabilitation System for Smart Healthcare**
Sizhe An and Umit Ogras - *University of Wisconsin-Madison*

12:30 pm - 12:45 pm, EDT, 4B.2
Prepare: Power-Aware Approximate Real-time Task Scheduling for Energy-Adaptive QoS Maximization
Shounak Chakraborty - *NTNU*
Sangeet Saha - *University of Essex*
Magnus Själander - *Norwegian University of Science and Technology*
Klaus McDonald-Meier - *University of Essex*

12:45 pm - 01:00 pm, EDT, 4B.3
Heterogeneity-aware Multicore Synchronization for Intermittent Systems
Wei-Ming Chen - *Academia Sinica*
Tei-Wei Kuo - *Academia Sinica & National Taiwan University*
Pi-Cheng Hsiu - *Academia Sinica*

* Denotes Best Paper Candidate

12:15 pm - 01:15 pm, EDT
Session 4C - EMSOFT: Testing and Fault Tolerance

Chair:
Corina Pasareanu, *Carnegie Mellon University/NASA Ames*

Co-Chair:
Sudipta Chattopadhyay, *Singapore University of Technology and Design*

This session focuses on new testing and fault-tolerance techniques for embedded platforms. The first paper presents a new fuzzing approach for finding vulnerabilities in real-time operating systems. The second paper presents an open-source HIL architecture for testing embedded systems with diverse hardware platforms and peripherals. The third paper proposes a complier-assisted scheme for detecting faulty devices based on correlations in measurements between IoT devices. The last paper presents a set of techniques for mitigating the impact of permanent faults in non-volatile memories on the accuracy of DNNs.

12:15 pm - 12:30 pm, EDT, 4C.1
Rtkaller: State-aware Task Generation for RTOS Fuzzing
Yuheng Shen, Hao Sun, Yixiao Yang, Yu Jiang - *Tsinghua University*
Wanli Chang - *University of York*
Heyuan Shi - *Alibaba Group*

12:30 pm - 12:45 pm, EDT, 4C.2
PHiLIP on the HiL: Automated Multi-platform OS Testing with External Reference Devices
Kevin Weiss, Michel Rottleuthner, Thomas Schmidt - *Hamburg University of Applied Sciences*
Matthias Wählisch - *Freie Universität Berlin*

12:45 pm - 01:00 pm, EDT, 4C.3
Precise Correlation Extraction for IoT Fault Detection with Concurrent Activities
Gyeongmin Lee - *Samsung Advanced Institute of Technology*
Bongjun Kim - *POSTECH*
Seungbin Song - *Yonsei University*
Changsu Kim - *POSTECH*
Jong Kim - *POSTECH*
Hanjun Kim - *Yonsei University*

01:00 pm - 01:15 pm, EDT, 4C.4
Tolerating Defects in Low-power Neural Network Accelerators via Retraining-free Weight Approximation
Fateme Sadat Hosseini - *Department of Electrical and Computer Engineering, University of Delaware*
Fanruo Meng - *University of Delaware*
Chengmo Yang - *University of Delaware*
Wujie Wen - *Lehigh Unversity*
Rosario Cammarota - *Intel Labs*

12:15 pm - 01:15 pm, EDT
Session 4D - EMSOFT: Cyber-Physical Systems: Handling Uncertainty and Attacks

Chair:
Miroslav Pajic, *Duke University*

Co-Chair:
Naijun Zhan, *Chinese Academy of Sciences*

This session presents novel techniques for handling uncertainty and security attacks in cyber-physical systems (CPS). The first paper proposes a real-time recovery technique that uses a linear quadratic regulator to bring the system to the safe state in the presence of an attack. The second paper presents a new language and tools for writing and checking structured proofs for differential games, which can be used to prove the correctness of adversarial CPS modeled by hybrid games. The third presents a technique to improve out-of-distribution data detection using Variational Autoencoders for resource-constrained real-time embedded systems. The last paper proposes a new uncertainty analysis that combines Bayesian RNN with a temporal logics-based analysis to estimate uncertainty in predictive monitoring of CPS.

12:15 pm - 12:30 pm, EDT, 4D.1
Real-Time Attack-Recovery for Cyber-Physical Systems using Linear-Quadratic Regulator
Lin Zhang - *Syracuse University*
Pengyuan Lu - *University of Pennsylvania*
Fanxin Kong - *Syracuse University*
Xin Chen - *University of Dayton*
Oleg Sokolsky - *University of Pennsylvania*
Insup Lee - *University of Pennsylvania*

12:30 pm - 12:45 pm, EDT, 4D.2
Structured Proofs for Adversarial Cyber-Physical Systems
Brandon Bohrer and Andre Platzer - *Carnegie Mellon University*

12:45 pm - 01:00 pm, EDT, 4D.3
Improving Variational Autoencoder based Out-of-Distribution Detection for Embedded Real-time Applications
Yeli Feng, Daniel Jun Xian Ng, Arvind Easwaran - *Nanyang Technological University*

01:00 pm - 01:15 pm, EDT, 4D.4
Predictive Monitoring with Logic-Calibrated Uncertainty for Cyber-Physical Systems
Meiyi Ma - *University of Virginia*
John Stankovic - *University of Virginia*
Ezio Bartocci - *TU Wien*
Lu Feng - *University of Virginia*

01:30 pm - 02:15 pm, EDT

Industry Pitch Session – Trends and Challenges

Chair:
Dirk Ziegenbein, *Bosch Research*

Co-Chair:
Selma Saidi, *TU Dortmund*

Welcome to the first ESWEEK Industry Pitch Session where speakers from the embedded systems industry pitch their current or foreseen challenges related to the hardware and software design for smart, intelligent and connected computing systems to the ESWEEK community. The session is organized as follows: a plenary part where the speakers will have 2-3 minutes to outline their design challenge and its engineering context. This is followed by a poster session fostering in-depth discussions and networking between industrial and academic experts.

Data-Driven Development at scale for safe Automated Driving
Marek Jersak - *VP Autonomous Drive Solutions, DXC Luxoft*

OASIS TOSCA – is it possible to bring cloud success stories to the car universe?
Oliver Kopp - *Research Specialist, Mercedes-Benz*

Highest dependability needs continuous in-field debugging
Albrecht Mayer - *Distinguished Engineer, Infineon*

Reliable Distributed Systems & The Fabric
Philipp Mundhenk - *Senior Expert, Bosch Research*

Achieving Certifiability by Increasing Complexity?
Sascha Uhrig - *Roadmap Owner Computer Architecture and Operating Systems, Airbus Defense & Space*

WEDNESDAY PROGRAM

October 13

9:30 am - 9:55 am	**Best Paper and Other Awards**			
10:00 am - 10:45 am	**Keynote: *Why is Machine Learning on Embedded Systems so Important?*** **Pete Warden, Software Engineer, Google**			
10:45 am - 11:00 am	**Break**			
11:00 am - 11:45 am	**Session 5A - CASES:** Architecture and Memory-Aware Optimization of Deep Learning Performance	**Session 5B - CODES+ISSS:** Robust and Trustworthy Systems	**Session 5C - EMSOFT:** Distributed System Coordination and Scheduling	**Special Session 5D:** Automated Edge AI Design – A Jump in Efficiency
11:45 am - 12:00 pm	**Break**			
12:00 pm - 12:45 pm		**Session 6B - CODES+ISSS:** Learning at the Edge	**Session 6C - EMSOFT:** Run-Time Adaptation	**Special Session 6D:** EdgeAI: New Applications, Systems, and Challenges toward Designing New "Things" at the Edge
12:45 pm - 12:55 pm	**Break**			
12:55 pm - 01:45 pm	**Panel: Machine Learning on the Edge: How Deep Can We 'Embed' It Into the Cloud-Edge Continuum?**			

WEDNESDAY, OCTOBER 13

09:30 am - 09:55 am, EDT

Best Paper and Other Awards

Presentation of Best Paper Awards for CASES, CODES+ISSS, EMSOFT, and ACM Transactions on Embedded Computing Systems (TECS), as well as SIGBED Frank Anger Memorial Award.

Wednesday Keynote,
10 am – 10:45 am, EDT
Why is Machine Learning on Embedded Systems so Important?

Speaker:
Pete Warden, *Google*

Abstract:
When I first joined Google in 2014, I was amazed to discover they were using 13 kilobyte neural network models to recognize “OK Google” on tiny embedded chips on Android phones. This felt like deep magic, and it made me wonder how many other problems these kinds of miniscule ML models could solve. Over the past few years I’ve been helping Google ship products using this approach with TensorFlow Lite Micro, and helped external developers create new applications. While it’s still early days for “TinyML”, we’re already seeing interesting impacts on how engineers compose systems, including software-defined sensors, cascades of ML models, air-gapped ambient computing, and ubiquitous on-device voice interfaces. In this talk I’ll cover the past, present, and future of embedded ML systems.

Bio:
Pete Warden is the technical lead of TensorFlow Lite Micro, Google’s open source embedded machine learning framework. He was previously CTO and founder of Jetpac, acquired in 2014, and is the author of the TinyML O’Reilly book. He blogs at petewarden.com, and is @petewarden on Twitter.

11:00 am - 11:45 am, EDT
Session 5A - CASES: Architecture and Memory-Aware Optimization of Deep Learning Performance

Chair:
Jana Doppa, *Washington State University*

Co-Chair:
Jeronimo Castrillon, *TU Dresden*

Runtime interventions are critical for squeezing out maximum performance and maintaining system integrity in modern complex systems-on-chip. This session presents three papers that explore different runtime aspects: silicon photonics-based DNN acceleration, hardware-assisted control flow integrity checking, and the co-optimization of NoC routing with task mapping.

11:00 am - 11:15 am, EDT, 5A.1
Synergistically Exploiting CNN Pruning and HLS Versioning for Adaptive Inference on Multi-FPGAs at the Edge
Guilherme Korol - *Universidade Federal do Rio Grande do Sul*
Michael Jordan - *Universidade Federal do Rio Grande do Sul*
Mateus Beck Rutzig - *UFSM*
Antonio Carlos Schneider Beck - *Universidade Federal do Rio Grande do Sul*

11:15 am - 11:30 am, EDT, 5A.2
On-Device Prior Knowledge Incorporated Learning for Personalized Atrial Fibrillation Detection
Zhenge Jia - *University of Pittsburgh*
Yiyu Shi - *University of Notre Dame*
Samir Saba - *University of Pittsburgh Medical Center*
Jingtong Hu - *University of Pittsburgh*

11:30 am - 11:45 am, EDT, 5A.3
Learning to Train CNNs on Faulty ReRam-based Manycore Accelerators
Biresh Kumar Joardar - *Duke University*
Jana Doppa - *Washington State University*
Hai (Helen) Li - *Duke University*
Krishnendu Chakrabarty - *Duke University*
Partha Pratim Pande - *Washington State University*

11:00 am - 11:45 am, EDT
Session 5B - CODES+ISSS: Robust and Trustworthy Systems

Chair:
Chen Liu, *Intel*

Co-Chair:
Ilia Polian, *University of Stuttgart*

Safety-critical and mission-critical systems require both security and reliability. This sessions presents new findings for bit error resilience, control-flow integrity, and hardware performance counters.

11:00 am - 11:15 am, EDT, 5B.1
SNR: Squeezing Numerical Range Defuses Bit Error Vulnerability Surface in Deep Neural Networks
Elbruz Ozen and Alex Orailoglu - *University of California, San Diego*

11:15 am - 11:30 am, EDT, 5B.2
REPAIR: Control Flow Protection based on Register Pairing Updates for SW-Implemented HW Fault Tolerance
Uzair Sharif, Daniel Mueller-Gritschneder, Ulf Schlichtmann - *Technical University of Munich*

11:30 am - 11:45 am, EDT, 5B.3
Hardware Performance Counters: Ready-Made vs Tailor-Made
Abraham Kuruvila - *University of Texas at Dallas*
Anushree Mahapatra - *New York University*
Ramesh Karri - *New York University*
Kanad Basu - *University of Texas at Dallas*

11:00 am - 11:45 am, EDT
Session 5C - EMSOFT: Distributed System Coordination and Scheduling

Chair:
Jan Reineke, *Saarland University*

Co-Chair:
Marcus Volp, *University of Luxembourg*

This session focuses on new coordinating and scheduling techniques for distributed embedded systems. The first paper presents a real-time extension of the distributed coordination service ZooKeeper's crash fault tolerant agreement protocol to improve response time to node failures. The second paper introduces a formal design and synchronizer for hierarchical multi-rate CPS, which generalizes and combines advantages of both time-triggered architectures (TTA) and single-rate and multirate (PALS) models. The last paper presents techniques for minimizing the makespan of DAG-structured tasks on distributed heterogeneous platforms.

11:00 am - 11:15 am, EDT, 5C.1
RT-ZooKeeper: Taming the Recovery Latency of a Coordination Service
Haoran Li, Chenyang Lu - *Washington University in St. Louis*
Chris Gill - *Washington University*

11:15 am - 11:30 am, EDT, 5C.2
MSYNC: A Generalized Formal Design Pattern for Virtually Synchronous Multirate Cyber-Physical Systems
Kyungmin Bae - *POSTECH*
Peter Olveczky - *University of Oslo*

11:30 am - 11:45 am, EDT, 5C.3
HMDS: A Makespan Minimizing DAG Scheduler for Heterogeneous Distributed Systems
Debabrata Senapati - *Indian Institute of Technology Guwahati*
Arnab Sarkar - *IIT Kharagpur*
Chandan Karfa - *Indian Institute of Technology Guwahati*

11:00 am - 11:45 am, EDT
Special Session 5D: Automated Edge AI Design – A Jump in Efficiency

Organizer and Chair:
Daniel Mueller-Gritschneder, *Technical University of Munich, Germany*

Gigantic rates of data production in the era of Big Data, Internet of Thing (IoT), and Smart Cyber Physical Systems (CPS) pose incessantly escalating demands for massive data processing, storage, and transmission while continuously interacting with the physical world using edge sensors and actuators. For IoT systems, there is now a strong trend to move the intelligence from the cloud to the edge or even the extreme edge, known as TinyML. Yet, this shift requires to design powerful machine learning system under very strict resource constraints, posing a difficult design task that needs to take the complete system stack from machine learning application, ML operator implementations to hardware platforms with ML acceleration into account. This session addresses new approaches to automate this design flow considering all layers to enable a "jump in design efficiency" for systems with tight and loosely-coupled accelerator architectures based on RISC-V as well as for more future brain-inspired computing approaches.

11:00 am - 11:15 am, EDT, 5D.1
Automatic Build of a complete RISC-V HW/SW System with Tightly-coupled ML Accelerator
Wolfgang Ecker - *Infineon Technologies AG, Germany*

11:15 am - 11:30 am, EDT, 5D.2
Automatic generation of application-specific AI accelerators for edge devices
Oliver Bringmann - *University of Tübingen, Germany*

11:30 am - 11:45 am, EDT, 5D.3
Energy-Efficient Edge AI: From Algorithms to Hardware Architectures
Muhammad Shafique - *New York University Abu Dhabi (NYUAD), UAE & Tandon School of Engineering, New York University (NYU), USA*

12:00 pm - 12:45 pm, EDT
Session 6B - CODES+ISSS: Learning at the Edge

Chair:
Youn-Long Lin, *National Tsing Hua University*

Co-Chair:
Gunar Schirner, *Northeatern University*

Advances in Machine Learning (ML) algorithms are key for the successful deployment of IoT and edge devices. This session focuses on the use and optimzation of ML algorithms in different embedded applications.

12:00 pm - 12:15 pm, EDT, 6B.1
MaxTracker: Continuously Tracking the Maximum Computation Progress for Energy Harvesting ReRam-based CNN Accelerators
Keni Qiu - *Capital Normal University*
Nicoholas Jao - *Penn State University*
Kunyu Zhou - *Capital Normal University*
Yongpan Liu - *Tsinghua University*
Jack Sampson - *Penn State University*
Mahmut Taylan - *Penn State University*
Vijay Narayanan - *Penn State University*

12:15 pm - 12:30 pm, EDT, 6B.2
Exploiting Activation Sparsity for Fast CNN Inference on Mobile GPUs
Chanyoung Oh, Junhyuk So, Sumin Kim, Youngmin Yi - *University of Seoul*

12:30 pm - 12:45 pm, EDT, 6B.3
HW-FlowQ: A Multi-Abstraction Level HW-CNN Co-design Quantization Methodology
Nael Fasfous - *Technical University of Munich*
Manoj Rohit Vemparala - *BMW AG*
Alexander Frickenstein - *BMW AG*
Emanuele Valpreda - *Polytechnic University of Turin*
Driton Salihu - *Technical University of Munich*
Nguyen Anh Vu Doan - *Technical University of Munich*
Christian Unger - *BMW AG*
Naveen Shankar Nagaraja - *BMW AG*
Maurizio Martina - *Polytechnic University of Turin*
Walter Stechele - *Technical University of Munich*

12:00 pm - 12:45 pm, EDT
Session 6C - EMSOFT: Run-Time Adaptation

Chair:
Petru Eles, *Linkoping University*

Co-Chair:
Jan Beutel, *University of Innsbruck*

This last session presents novel methods for adapting scheduling and resources to improve real-time performance. The first paper presents a cross-layer approach to save resources or improve performance by allowing proactive skipping of task executions and period adjustment. The second paper proposes a technique for assigning deadlines and periods to update and control tasks to ensure a timely response to changes in CPS. The last paper presents an adaptive thermal-aware scheduling method for real-time tasks on heterogeneous multi-core platforms with dynamic injection of new tasks, which uses a combination of frequency tuning, task migration, or idle slot insertion to reduce tempertature while ensuring timing guarantees.

12:00 pm - 12:15 pm, EDT, 6C.1
Cross-Layer Adaptation with Safety-Assured Proactive Task Job Skipping
Zhilu Wang - *Northwestern University*
Chao Huang - *University of Liverpool, Northwestern University*
Hyoseung Kim - *University of California, Riverside*
Wenchao Li - *Boston University*
Qi Zhu - *Northwestern University*

12:15 pm - 12:30 pm, EDT, 6C.2
Guaranteeing Timely Response to Changes of Monitored Objects by Assigning Deadlines and Periods to Tasks
Quan Zhou, Guohui Li, Qi Chen, Jianjun Li - *Huazhong University of Science and Technology*

12:30 pm - 12:45 pm, EDT, 6C.3
Thermal Aware Adaptive Platform Management for Heterogeneous Embedded Systems
Srijeeta Maity, Anirban Ghose, Soumyajit Dey, Swarnendu Biswas - *Indian Institute of Technology Kanpur*

12:00 pm - 12:45 pm, EDT,
Special Session 6D: EdgeAI: New Applications, Systems, and Challenges toward Designing New "Things" at the Edge

Organizer and Chair:
Radu Marculescu, *University of Texas at Austin*

Traditionally, ESWEEK aimed at becoming the premier event in embedded systems design and benefitted a relatively homogeneous community of people interested in various hardware and/or software aspects. In recent years, however, we have witnessed the raise of new applications and systems which brought inevitably a shift in the overall focus towards the Internet of Things (IoT). By creating smart things connected to each other and the internet, we can enable endless applications that improve our lives. To achieve this vision, we must address the resource and communication constraints of many of the devices along the edge and intelligently utilize the compute resources across the IoT hierarchy. Starting from these overarching ideas, in this special session, we bring together several visions from academia and industry to address these problems and provide a holistic view of the path ahead towards designing new "things" at the edge. The audience will be challenged to leave the comfort zone (i.e., traditional embedded system design), travel the path not taken, see the unseen, and immerse into the joy of designing new "things".

12:00 pm - 12:15 pm, EDT, 6D.1
Efficient Audio-Visual Understanding on AR Devices
Vikas Chandra - *Facebook Reality Labs*

12:15 pm - 12:30 pm, EDT, 6D.2
Truly Wearable Self-Powered Edge Devices for Health Monitoring
Umit Y. Ogras - *The University of Wisconsin-Madison, USA*

12:30 pm - 12:45 pm, EDT, 6D.3
Communication-aware Distributed Training and Inference at the Edge
Radu Marculescu - *The University of Texas, Austin, USA*

12:55 pm - 01:45 pm, EDT

Panel: Machine Learning on the Edge: How Deep Can We 'Embed' It Into the Cloud-Edge Continuum?

Moderator:
Radu Marculescu, *The University of Texas at Austin*

Large amounts of data are generated nowadays on edge devices, such as phones, tablets, wearables. Collectively, such devices produce a data rich environment that can be capitalized on to improve the 'intelligence' of the Internet of Things (IoT). Recently, deep learning models have pushed this intelligence from the cloud closer to the edge in applications such as computer vision, healthcare, autonomous driving, speech processing, to name a few. This panel will address the challenges involved in enabling the widespread deployment of deep learning on networked edge devices (e.g., resource scarcity, energy efficiency, security/privacy etc.), thus making the cloud-edge computing continuum a reality. By scope and structure this panel consists of a multi-angle, highly interactive, discussion about new research, industries, and education opportunities in embedded and cyber-physical systems that can be fueled by these new developments.

Panelists:
Vikas Chandra, *Facebook*
Pete Warden, *Google*
Paul Whatmough, *ARM*
Umit Ogras, *University of Wisconsin-Madison*
Tarek Abdelzaher, *University of Illinois Urbana-Champaign*

WORKSHOPS

Thursday, October 14

9:00 pm (Wed) - 05:00 am	**MSC: International Workshop on Memory and Storage Computing** (MSC Workshops will run 9 am - 5 pm Beijing time.)
09:00 am - 04:00 pm	**TRAIN: Trustworthy and Reliable AI accelerator desigN (TRAIN) Workshop**
09:00 am - 01:00 pm	**RSP: International Workshop on Rapid System Prototyping**

09:00 am (Beijing) - 5:00 pm (Beijing)
MSC: International Workshop on Memory and Storage Computing

Organizers:
Liang Shi, *East China Normal University*
Weichen Liu, *Nanyang Technological University*
Yuan-Hao Chang, *Academia Sinica*

Memory and storage technologies have been developed during the last decades. However, the data transfer cost between CPU and storage or memory becomes the critical challenge for the advanced systems. Storage and memory computing provides a new opportunity to solve this issue by adding computing function beside storage or memory. In this workshop, we are not only interested in the advanced storage and memory computing technologies, but also take a special attention to the advanced storage and memory technologies. The workshop to will bring together scientists and engineers from industry and academia who are working on storage and memory computing architectures and systems.

09:00 am - 04:00 pm, EDT
TRAIN: Trustworthy and Reliable AI accelerator desigN (TRAIN) Workshop

Organizers:
Chengmo Yang, *University of Delaware*
Wujie Wen, *Lehigh University*

Hardware accelerated artificial intelligence (AI) is now becoming ubiquitous, shifting from clouds to resource-limited embedded and IoT platforms. While hardware accelerators facilitate fast and energy-efficient neural network operations that are both memory and computational intensive, they are facing two fundamental challenges in practice. The first is unreliable inference caused by passive hardware faults in traditional CMOS-based accelerator memories, buffers and computation units, as well as the imperfect manufacturing and non-ideal device behaviors of emerging post CMOS processing-in-memory (PIM) accelerators. The second challenge is the violation of AI integrity and confidentiality caused by fault injection attacks and/or side channel attacks targeting these new NN hardware accelerators.

The goal of this workshop is to establish a forum for the discussion on state-of-the-art research in AI accelerator design from the aspects of reliability and security, which are two sides of the same coin – the unexpected accelerator behavior can be induced by either hardware faults or malicious attacks.

09:00 am - 01:00 pm, EDT
RSP: International Workshop on Rapid System Prototyping

Organizers:
Kenneth Kent, *University of New Brunswick*
Frédéric Rousseau, TIMA, *University Grenoble-Alpes*

The International Workshop on Rapid System Prototyping (RSP) emphasizes design experience sharing and collaborative approach between hardware and software research communities from industry and academy. It considers prototyping as an iterative design approach for embedded hardware and software systems. The RSP series of workshop aim at bridging the gaps in embedded system design between applications, architectures, tools, and technologies to achieve rapid system prototyping of emerging software and hardware systems.

Rapid System Prototyping workshop seeks original contributions related to this target, encompassing a wide scope ranging from formal methods for the verification of software and hardware systems to case studies of emerging embedded systems and technologies. The workshop proposes a two-day inspiring international forum for discussing the latest related innovations and research activities. The workshop program will include keynote speeches and technical papers on timely topics.

NOCS SYMPOSIUM

Thursday, October 14

09:00 am - 09:10 am	**Opening Remarks**
09:10 am - 10:10 am	**Keynote I: *Ultra-Energy Efficient, Multi-terabit Photonic Connectivity for Disaggregated Computing***
10:00 am - 10:55 am	**Regular Session I: NoCs for DNN Accelerators**
10:55 am - 11:00 am	**Break**
11:00 am - 11:45 am	**Regular Session II: Security and NoC Routing**
11:45 am - 01:45 pm	**Special Session I: Towards Scalable Multi-core Quantum Computing Architectures: Quantum Networks-On-Chip at Rescue**

Friday, October 15

09:00 am - 10:00 am	**Keynote II: *AMD Chiplet Technologies and Implications on Interconnect Architectures***
10:00 am - 10:55 am	**Regular Session III: NoC Design for Modern Systems**
10:55 am - 11:00 am	**Break**
11:00 am - 11:40 am	**Regular session IV: Secure NoC-based Systems**
11:40 am - 01:00 pm	**Special Session II: Open Source On-Chip Communication from Edge to Cloud: The PULP Experience**
01:00 pm - 01:15 pm	**Concluding Remarks**

THURSDAY, OCTOBER 14 - NOCS

09:00 am - 9:10 am, EDT
NOCS Opening Remarks
Speakers:Sergi Abadal, *Universitat Politècnica de Catalunya*
Joshua San Miguel, *University of Wisconsin-Madison*

9:10 am - 10:10 am, EDT
NOCS Keynote I: Ultra-Energy Efficient, Multi-terabit Photonic Connectivity for Disaggregated Computing

Speaker:
Keren Bergman, *Columbia University*

Bio:
Keren Bergman is the Charles Batchelor Professor of Electrical Engineering at Columbia University where she also serves as the Faculty Director of the Columbia Nano Initiative. Bergman received the B.S. from Bucknell University in 1988, and the M.S. in 1991 and Ph.D. in 1994 from M.I.T. all in Electrical Engineering. At Columbia, Bergman leads the Lightwave Research Laboratory encompassing multiple cross-disciplinary programs at the intersection of computing and photonics. Bergman serves on the Leadership Council of the American Institute of Manufacturing (AIM) Photonics leading projects that support the institute's silicon photonics manufacturing capabilities and Datacom applications. She is the recipient of the 2016 IEEE Photonics Engineering Award and is a Fellow of the Optical Society of America (OSA) and IEEE.

10:10 am - 10:55 am, EDT
NOCS Regular Session I: NoCs for DNN Accelerators
*** (L) A Novel Network Fabric for Efficient Spatio-Temporal Reduction in Flexible DNN Accelerators**
Francisco Muñoz-Martinez - *Universidad de Murcia;* Jose L. Abellan - *Universidad Católica San Antonio de Murcia;* Manuel E. Acacio - *Universidad de Murcia;* Tushar Krishna - *Georgia Institute of Technology*
(S) Analysis of on-chip communication properties in accelerator architectures for Deep Neural Networks
Hana Krichene - *CEA-LIST;* Jean-Marc Philippe - *CEA-LIST*
(S) NewroMap: Mapping CNNs to NoC-interconnected Self-Contained Data-Flow Accelerators for Edge-AI
Jan Moritz Joseph - *RWTH Aachen University;* Sezgin Baloglu - *RWTH Aachen University;* Yue Pan - *Georgia Institute of Technology;* Rainer Leupers - *RWTH Aachen University;* Lennart Bamberg - *GrAi Matter Labs*

11:00 am - 11:45 AM, EDT
NOCS Regular Session II: Security and NoC Routing
(L) Packet Header Attack by Hardware Trojan in NoC based TCMP and its Impact Analysis
Vedika Jitendra Kulkarni - *Indian Institute of Technology Guwahati;* Manju R - *Indian Institute of Technology Guwahati;* Ruchika Gupta - *Indian Institute of Technology Guwahati;* John Jose - *Indian Institute of Technology Guwahati Guwahati ;* Sukumar Nandi - *Indian Institute of Technology Guwahati*
(S) Securing Network-on-Chips via Novel Anonymous Routing
Amin Sarihi - *New Mexico State University;* Ahmad Patooghy - *University of Central Arkansas;* Mahdi Hasanzadeh - *Independent scholar;* Mostafa Abdelrehim - *California State University at Bakersfield;* Abdel Hameed Badawy - *New Mexico State University*
(S) Denial-of-Service Attack Detection using Machine Learning in Network-on-Chip Architectures
Chamika Sudusinghe - *University of Moratuwa;* Subodha Charles - *University of Florida;* Prabhat Mishra - *University of Florida*

* Denotes Best Paper Candidate

11:45 am - 1:45 am, EDT

NOCS Special Session I: Towards Scalable Multi-core Quantum Computing Architectures: Quantum Networks-On-Chip at Rescue

Chairs:

Eduard Alarcón, *Technical University of Catalonia*
Carmen G. Almudéver, *Technical University of Valencia*

Abstract:

The field of quantum computing has experienced a remarkable progress in the last years with the development of intermediate-scale quantum processors. Despite its tremendous potential, it is still unclear how quantum computing systems will scale-up to satisfy the requirements of its most powerful applications. At architectural level quantum multi-core architectures are a firm candidate to unlock the scalability of quantum devices. Nevertheless, a key enabling aspect of these is the development of quantum network on chips (NOCs) for the quantum-coherent communication links among cores. To this purpose, in this workshop we are bringing in experts from the three required disciplines to address this challenge namely: i) quantum computing architectures; ii) quantum communications and networking; iii) state-of-the-art NOCs for conventional computers.

FRIDAY, OCTOBER 15 - NOCS

09:00 am - 10:00 am, EDT

NOCS Keynote II: AMD Chiplet Technologies and Implications on Interconnect Architectures

Speaker:
Gabriel (Gabe) Loh, *AMD Research*

Bio:
Gabe is a Senior Fellow in AMD Research. He received his Ph.D. and M.S. in computer science from Yale University in 2002 and 1999, respectively, and his B.Eng. in electrical engineering from the Cooper Union in 1998. Gabe was also a tenured associate professor in the College of Computing at the Georgia Institute of Technology, a visiting researcher at Microsoft Research, and a senior researcher at Intel Corporation. He is a Fellow of the ACM and IEEE, recipient of ACM SIGARCH's Maurice Wilkes Award, Hall of Fame member for the MICRO, ISCA, and HPCA conferences, (co-)inventor on over one hundred US patent applications and ninety granted patents, and a recipient of a U.S. National Science Foundation CAREER Award.

10:00 am - 10:55 am, EDT

NOCS Regular Session III: NoC Design for Modern Systems

*** (L) PlugSMART: a pluggable open-source module to implement multihop bypass in Networks-on-Chip**
Alireza Monemi - *Barcelona Supercomputing Center;* Iván Pérez - *University of Cantabria;*
Neiel Israel Leyva - *Barcelona Supercomputing Center;* Enrique Vallejo - *University of Cantabria;*
Ramón Beivide - *University of Cantabria;* Miquel Moreto - *Barcelona Supercomputing Center*

(S) DUB: Dynamic Underclocking and Bypassing in Network-on-Chip for Heterogeneous GPU Workloads
Srikant Bharadwaj - *AMD;* Shomit Das - *AMD;* Yasuko Eckert - *AMD;* Mark Oskin - *University of Washington/AMD;*
Tushar Krishna - *Georgia Institute of Technology*

(S) Worst-Case Latency Analysis for the Versal NoC Network Packet Switch
Ian Lang - *University of Waterloo);* Nachiket Kapre - *University of Waterloo;* Rodolfo Pellizzoni - *University of Waterloo*

(S) Synthesis of Predictable Global NoC by Abutment in Synchoros VLSI Design
Jordi Altayo Gonzalez - *KTH Royal Institute of Technology*; Ahmed Hemani - *KTH Royal Institute of Technology;*
Dimitrios Stathis - *KTH Royal Institute of Technology*

11:00 am - 11:40 am, EDT

NOCS Regular Session IV: Secure NoC-based Systems

(L) Sentry-NoC: A Statically Scheduled NoC for Secure SoCs
Ahmed Shalaby - *National University of Singapore;* Yaswanth Tavva - *National University of Singapore;*
Trevor E. Carlson - *National University of Singapore;* Li-Shiuan Peh - *National University of Singapore*

(L) Multilayer NoC Firewall Services: Case-Study on E-Health
Miltos Grammatikakis - *Hellenic Mediterranean University;* Voula Piperaki - *Hellenic Mediterranean University;*
Antonis Papagrigoriou - *Hellenic Mediterranean University*

* Denotes Best Paper Candidate

FRIDAY, OCTOBER 15 - NOCS

11:40 am - 1:00 am, EDT

NOCS Special Session II: Open Source On-Chip Communication from Edge to Cloud: The PULP Experience

Chair:
Davide Rossi, *University of Bologna*

Abstract:
High performance and extreme energy efficiency are growing requirements for wide class of applications ranging from ultra-low power IoT end nodes to high performance computing. While parallelism and heterogeneity are well established techniques to deal with these challenges in the high-performance domain, an increasing amount of embedded systems are embracing the same approach to deal with the increasing complexity of near-sensor analytics applications. However increasing the number and capabilities of compute cores poses a high-pressure on the on-chip communication network, which has to sustain the extreme bandwidth generated by multiple powerful engines. In this special session we propose an open source ecosystem for building computing systems from edge to cloud. The sessions starts with a presentation of the Parallel Ultra-Low-Power project, an open-source platform for near-sensor processing, followed by the presentation of an open-source AXI4 infrastructure and its exploitation heterogeneous application processors and HPC many-core system.

10:00 am - 10:55 am, EDT

NOCS Regular Session III: NoC Design for Modern Systems

(SS) PULP: An Open-Source RISC-V Based Multi-Core Platform for In-Sensor Analytics
Davide Rossi - *University of Bologna*

(SS) An Open-Source Platform for High-Performance Non-Coherent On-Chip Communication
Thomas Benz - *ETH Zürich*

(SS) HERO: A Heterogenous Research Platform to Explore HW/SW Codesign and RISC-V manycore accelerators
Luca Bertaccini - *ETH Zürich*

(SS) Manticore as an NoC Case Study: A 4096 Chiplet-based Architecture for Ultra-Efficient Floating-Point Computing
Florian Zaruba - *ETH Zürich*

1:00 pm - 1:15 pm, EDT

NOCS Concluding Remarks

ESWEEK 2021 COMMITTEE

GENERAL CHAIRS

Andreas Gerstlauer | General Chair
University of Texas at Austin, US

Aviral Shrivastava | Vice General Chair
Arizona State University, US

PAST CHAIR

Tulika Mitra | Past General Chair
National University of Singapore, SG

PROGRAM CHAIRS

Umit Ogras | CASES Chair
Arizona State University, US

Chengmo Yang | CODES+ISSS Co-Chair
University of Delaware, US

Preeti Panda | CASES Co-Chair
IIT Delhi, IN

Linh Thi Xuan Phan | EMSOFT Chair
University of Pennsylvania, US

Jason Xue | CODES+ISSS Chair
City University of Hong Kong, HK

David Broman | EMSOFT Co-Chair
KTH Royal Institute of Technology, SE

ORGANIZATIONAL COMMITTEE

Sudeep Pasricha | Virtual Conference Chair
Colorado State University, US

Andreas Herkersdorf | Industry Liaison Chair
Technical University of Munich, DE

Hoeseok Yang | Virtual Conference Co-Chair
Ajou University, KR

Hussam Amrouch | Finance Chair
Stuttgart University, DE

Paul Bogdan | Special Sessions and Tutorial Chair
University of Southern California, US

Amit Singh | Publication Chair
University of Essex, UK

Timothy Bourke | Special Sessions and Tutorial Chair
INRIA, Paris, FR

Jürgen Teich | Awards Chair
Friedrich-Alexander University Erlangen-Nürnberg, DE

Patricia Derler | Workshop Chair
Kontrol LLC

Selma Saidi | *Young Fellow and Inclusion Chair*
Technical University of Dortmund, DE

Dirk Ziegenbein | Industry Sessions Chair
Robert Bosch GmbH, DE

Anuj Pathania | *Web Chair*
University of Amsterdam, NL

Partha Pande | Industry Liaison Chair
Washington State University, US

Lars Bauer | *Publicity Chair*
Karlsruhe Institute of Technology, DE

CASES 2021 COMMITTEE

PROGRAM CHAIRS

Umit Ogras
University of Wisconsin-Madison

Preeti Ranjan Panda
Indian Institute of Technology - Delhi

PROGRAM COMMITTEE

Ali Akoglu
University of Arizona

Lars Bauer
Karlsruhe Institute of Technology

Ganapati Bhat
Washington State University

Swarup Bhunia
University of Florida

Oliver Bringmann
University of Tuebingen / FZI

Luca Carloni
Columbia University

Jeronimo Castrillon
TU Dresden

Wanli Chang
University of York

Henri-Pierre Charles
CEA

Mainak Chaudhuri
Indian Institute of Technology - Kanpur

Anup Das
Drexel University

Jana Doppa
Washington State University

Lide Duan
Alibaba DamO Academy

Christophe Dubach
McGill University

William Fornaciari
Politecnico di Milano - DEIB

Amlan Ganguly
Rochester Institute of Technology

Dibakar Gope
ARM

Per Gunnar Kjeldsberg
Norwegian University of Science and Technology

Puneet Gupta
University of California, Los Angeles

Yuko Hara-Azumi
Tokyo Institute of Technology

Joerg Henkel
Karlsruhe Institute of Technology

Jingtong Hu
University of Pittsburgh

Paolo Ienne
École Polytechnique Fédérale de Lausanne

Changhee Jung
Purdue University

Ramesh Karri
New York University

Ryan Kim
Colorado State University

Andreas Krall
TU Wien

Akash Kumar
TU Dresden

CASES 2021 COMMITTEE (cont.)

Amit Kumar Singh
University of Essex

Mikko Lipasti
University of Wisconsin-Madison

Scott Mahlke
University of Michigan

Avinash Malik
University of Auckland

Sai Manoj Pudukotai Dinakarrao
George Mason University

Pietro Mercati
Intel

Vijaykrishnan Narayanan
Penn State University

Rupesh Nasre.
Indian Institute of Technology, Chennai

Sri Parameswaran
University of New South Wales

Anuj Pathania
University of Amsterdam

Christian Pilato
Politecnico di Milano

Laura Pozzi
USI Lugano

Partha Pratim Pande
Washington State University

Simon Rokicki
Irisa

Sanghamitra Roy
Utah State University

Sourav Roy
NXP

Nader Sehatbakhsh
University of California, Los Angeles

Karthik Swaminathan
IBM

Hiroyuki Tomiyama
Ritsumeikan University

Hao Wang
Hygon R&D

Weng-Fai Wong
National University of Singapore

Chia-Lin Yang
National Taiwan University

Sungjoo Yoo
Seoul National University

Wei Zhang
Hong Kong University of Science and Technology

CODES + ISSS 2021 COMMITTEE

PROGRAM CHAIRS

Jason Xue
City University of Hong Kong

Chengmo Yang
University of Delaware

PROGRAM COMMITTEE

Tosiron Adegbija
University of Arizona

Siddharth Advani
Samsung Research America

Mohammad Al Faruque
University of California, Irvine

David Atienza
École Polytechnique Fédérale de Lausanne (EPFL)

Lars Bauer
Karlsruhe Institute of Technology (KIT)

Christophe Bobda
University of Florida

Paul Bogdan
University of Southern California

Eli Bozorgzadeh
Univ. of California, Irvine

Oliver Bringmann
University of Tuebingen / FZI

Rosario Cammarota
Intel

Wanli Chang
University of York

Yuan-Hao Chang
Academia Sinica

Xiang Chen
George Mason University

Robert P. Dick
University of Michigan

Rainer Doemer
University of California, Irvine

Nikil Dutt
University of California, Irvine

Petru Eles
Linkoping University

Xin Fang
Qualcomm

Fabrizio Ferrandi
Politecnico di Milano

Swaroop Ghosh
Pennsylvania State University

Dimitris Gizopoulos
University of Athens

Michael Glaß
Ulm University

Soonhoi Ha
Seoul National University

Jie Han
University of Alberta

Frank Hannig
Friedrich-Alexander University Erlangen-Nürnberg

Christian Haubelt
University of Rostock

Jörg Henkel
Karlsruhe Institute of Technology

Pi-Cheng Hsiu
Academic Sinica

Reiley Jeyapaul
ARM Research

Fadi Kurdahi
University of California, Irvine

Luciano Lavagno
Politecnico di Torino

Sébastien Le Beux
Concordia University

Kyoungwoo Lee
Yonsei University

Jenq-Kuen Lee
National Tsing Hua University

Youn-Long Lin
National Tsing Hua University

Chen Liu
Intel

Weichen Liu
Nanyang Technological University

Roman Lysecky
University of Arizona

Enrico Macii
Politecnico di Torino

Grant Martin
Cadence Design Systems

Hiroki Matsutani
Keio University

Nele Mentens
KU Leuven

CODES + ISSS 2021 COMMITTEE (cont.)

Brett Meyer
McGill University

Prabhat Mishra
University of Florida

Wolfgang Mueller
University of Paderborn

Daniel Mueller-Gritschneder
Technical University of Munich

Gabriela Nicolescu
École Polytechnique de Montréal

Mahdi Nikdast
Colorado State University

Hyunok Oh
Hanyang University

Alex Orailoglu
UC San Diego

Gianluca Palermo
Politecnico di Milano

Maurizio Palesi
University of Catania

Partha Pratim Pande
Washington State University

Sri Parameswaran
University of New South Wales

Sudeep Pasricha
Colorado State University

Andy Pimentel
University of Amsterdam

Ilia Polian
University of Stuttgart

Graziano Pravadelli
University of Verona

Ann Ramirez (Gordon-Ross)
University of Florida

Francesco Regazzoni
University of Amsterdam

Gunar Schirner
Northeastern University

Zili Shao
The Chinese University of Hong Kong

Sandeep Shukla
Indian Institute of Technology Kanpur

Amit Kumar Singh
University of Essex

Todor Stefanov
Leiden University

Sander Stuijk
Eindhoven University of Technology

Sara Tehranipoor
Santa Clara University

Jürgen Teich
University of Erlangen-Nuremberg

Ishan Thakkar
University of Kentucky

Hiroyuki Tomiyama
Ritsumeikan University

Eugenio Villar
University of Cantabria

Chao Wang
University of Science and Technology of China

Fei Wu
Huazhong University of Science and Technology

Jiafeng Xie
Villanova University

Jiang Xu
Hong Kong University of Science and Technology

Ming-Chang Yang
The Chinese University of Hong Kong

Mengying Zhao
Shandong University

EMSOFT 2021 COMMITTEE

PROGRAM CHAIRS

Linh Thi Xuan Phan
University of Pennsylvania

David Broman
KTH Royal Institute of Technology

PROGRAM COMMITTEE

Tarek Abdelzaher
University of Illinois at Urbana-Champaign

Mihail Asavoae
CEA, List

Stanley Bak
Stony Brook University

Guillaume Baudart
Inria

Marko Bertogna
University of Modena

Jan Beutel
University of Innsbruck

Enrico Bini
University of Turin

Samarjit Chakraborty
University of North Carolina at Chapel Hill

Li-Pin Chang
National Chiao Tung University

Sudipta Chattopadhyay
Singapore University of Technology and Design

David Cock
ETH Zürich

Thao Dang
CNRS

Eva Darulova
Max Planck Institute for Software Systems (MPI-SWS)

Pallab Dasgupta
IIT Kharagpur

Andre DeHon
University of Pennsylvania

Petru Eles
Linköping University

Rolf Ernst
TU Braunschweig

Chris Gill
Washington University in St. Louis

Gregor Goessler
Inria

Susanne Graf
Grenoble University/Verimag

Nan Guan
Hong Kong Polytechnic University

Ichiro Hasuo
National Institute of Informatics

Susmit Jha
SRI International

Yu Jiang
Tsinghua University

Hokeun Kim
Google

Edward Lee
UC Berkeley

Chung-Wei Lin
National Taiwan University

Cong Liu
University of Texas at Dallas

EMSOFT 2021 COMMITTEE (cont.)

Martina Maggio
Saarland University

Renato Mancuso
Boston University

Diana Marculescu
University of Texas at Austin

Frank Mueller
North Carolina State University

Sam H. Noh
Ulsan National Institute of Science and Technology

Necmiye Ozay
University of Michigan

Claire Pagetti
Onera / ENSEEIHT

Miroslav Pajic
Duke University

Corina Pasareanu
Carnegie Mellon University/ NASA Ames

Hiren Patel
University of Waterloo

Rodolfo Pellizzoni
University of Waterloo

Pavithra Prabhakar
Kansas State University

Jan Reineke
Saarland University

Indranil Saha
IIT Kanpur

Selma Saidi
TU Dortmund

Ingo Sander
KTH Royal Institute of Technology

Wolfgang Schröder-Preikschat
Friedrich–Alexander University Erlangen– Nürnberg

Donatella Sciuto
Polytechnic University of Milan

Roopak Sinha
Auckland University of Technology

Oleg Sokolsky
University of Pennsylvania

Parasara Sridhar Duggirala
University of North Carolina at Chapel Hill

Lothar Thiele
ETH Zürich

Stavros Tripakis
Northeasten University

Marcus Völp
University of Luxembourg

Bryan Ward
MIT Lincoln Laboratory

Richard West
Boston University

Jason Xue
City University of Hong Kong

Bai Xue
Chinese Academy of Sciences

Haibo Zeng
Virginia Tech

Naijun Zhan
Chinese Academy of Sciences

WIP COMMITTEE

Ilge Akkaya
OpenAI

Andrea Bastoni
Technical University of Munich

Matthias Becker
KTH Royal Institute of Technology

Alessandro Biondi
Scuola Superiore Sant'Anna

Marten Lohstroh
UC Berkeley

Soheil Samii
General Motors R&D

Martin Schoeberl
Technical University of Denmark - DTU

Reinhard von Hanxleden
Kiel University

NOCS 2021 COMMITTEE

GENERAL CHAIRS

Tushar Krishna
Georgia Institute of Technology

John Kim
KAIST

TECHNICAL PROGRAM CHAIRS

Sergi Abadal
Universitat Politècnica de Catalunya, abadal@ac.upc.edu

Joshua San Miguel
University of Wisconsin-Madison

PUBLICITY CHAIRS

José Luis Abellán
Catholic University of Murcia

Kun-Chih Chen
National Sun Yat-Sen University

Jieming Yin
Lehigh University

WEB CHAIR

Salvatore Monteleone
Kore University of Enna

INDUSTRY CHAIR

Hao Luan
Futurewei Technologies

PUBLICATION CHAIR

Mahdi Nikdast
Colorado State University

STEERING COMMITTEE CHAIR

Radu Marculescu
University of Texas, Austin

TECHNICAL PROGRAM COMMITTEE

Poona Bahrebar
University of California, Irvine / University of Ghent

Akram Ben Ahmed
AIST

Ramón Beivide Palacio
University of Cantabria

Davide Bertozzi
University of Ferrara

Paul Bogdan
University of Southern California

José Cano
University of Glasgow

Luca Carloni
Columbia University

Shomit Das
AMD Research

Amlan Ganguly
Rochester Institute of Technology

Paul Gratz
Texas A&M University

Syed Minhaj Hassan
Qualcomm

Robert Hesse
Cerebras

Jan Moritz Joseph
RWTH Aachen University

Nachiket Kapre
University of Waterloo

Avinash Karanth
Ohio University

Cédric Killian
IRISA/INRIA

Ryan Kim
Colorado State University

Gwangsun Kim
POSTECH

Eun Jung (EJ) Kim
Texas A&M University

Michihiro Koibuchi
National Institute of Informatics, Japan

Jinho Lee
Yonsei University

Zhonghai Lu
KTH Royal Institute of Technology

Radu Marculescu
University of Texas, Austin

Hiroki Matsutani
Keio University

NOCS 2021 COMMITTEE (cont.)

Chrysostomos Nicopoulos
University of Cyprus

Maurizio Palesi
University of Catania

Mayank Parasar
Samsung

Partha Pratim Pande
Washington State University

Li-Shiuan Peh
National University of Singapore

Hamid Sarbazi-Azad
Sharif University of Technology

Ashif Sikder
Qualcomm

Jens Sparsø
Technical University of Denmark

Ishan Thakkar
University of Kentucky

Jiang Xu
Hong Kong University of Science and Technology

Jieming Yin
AMD Research

EMBEDDED SYSTEMS WEEK

OCTOBER 08-15, 2021 | VIRTUAL CONFERENCE