IEEE VLSI Circuits and Systems Letter

Volume 8, Issue 2, May 2022

Editorial

Invited articles


Conference spotlight

VLSI Design 2021 – Summary by Program Co-Chairs - Mayam Shojaei (IIT, Bombay), Nele Mentes (KU Leuven), Taher Abbasi (Cadence)


  o VLSID-2022 Nripendra Nath Biswas Best Student Paper Award: “Hardware Accelerator for Capsule Network based Reinforcement Learning”, - Dola Ram, Suraj Panwar and Kuruvilla Varghese ; Indian Institute of Science

  o VLSID-2022: Naresh Malipeddy Honorable Mention Award- “A 5-Gb/s PAM4 Voltage Mode Transmitter with Current Mode Continuous Time Linear Equalizer”, Shraman Mukherjee, Sumanta Seth and Saurabh Saxena ; Texas Instruments, India

Women in VLSI (WiV) series spotlight

- Interview with Dr. Nevine Nassif, Senior Fellow, Intel Corporation

Updates

- Recent relevant news highlights – Ishan Thakkar

- 2021 TCVLSI awards

- Planned 2022 conference sponsorships by TCVLSI
From the Editor-in-Chief’s Desk - Editorial

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

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

This issue of VCaSL features an invited article “The Future of Hardware Technologies for Computing: N3XT 3D MOSAIC, Illusion Scaleup, Co-Design” by Tathagata Srimani, Robert M. Radway, H.-S. Philip Wong and Subhasish Mitra (Stanford University), where the authors describe transformative NanoSystems (computing systems which exploit unique characteristics of emerging nanotechnologies) through synergistic innovations across multiple levels: N3XT 3D MOSAIC (material, device, integration technology), Illusion Scaleup (circuit, architecture), and Co-Design (technology, circuit, architecture, and application, targeting 100-1,000x benefits in system-level Energy Delay Product (EDP) vs. today’s approaches.

The newsletter spotlights one of TCVLSI’s sponsored conferences in 2022, the IEEE International Conference on VLSI Design also known as VLSID. One-page teasers of three best papers awarded at the 2022 VLSID conference are showcased.

In our Women in VLSI (WiV) series, we share an inspiring interview with Dr. Nevine Nassif, Senior Fellow at Intel Corporation.

We have included is a section on relevant recent announcements collated by our Associate Editor, Ishan Thakkar. Thakkar. Additionally, we have announced the winners of our 2021 TCVLSI awards.

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

Mondira (Mandy) Pant, Ph.D
Chair, IEEE Computer Society TCVLSI
Editor-in-Chief of IEEE VCA, TCVLSI
Intel Corporation, USA

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

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

https://tc.computer.org/tcvlsi/
The Future of Hardware Technologies for Computing: 
**N3XT 3D MOSAIC, Illusion Scaleup, Co-Design**

Stanford University, Dept. of Electrical Engineering*, Dept. of Computer Science**

The computation demands of 21st-century abundant-data applications, e.g., AI deep learning, augmented reality / virtual reality (AR/VR), graph analytics, far exceed the capabilities of today’s computing systems. For example, a Dream AI Chip would ideally co-locate all memory and compute on a single chip, quickly accessible at low energy. Such Dream Chips aren’t feasible today. Computing systems instead use large off-chip memory and spend enormous time and energy shuttling data back-and-forth. This memory wall gets worse with growing problem sizes, especially as conventional transistor miniaturization gets increasingly difficult – the miniaturization wall. The next leap in computing performance requires the next leap in integration. Just as integrated circuits brought together discrete components, this next level of integration must seamlessly fuse disparate parts of a system – e.g., compute, memory, inter-chip connections – for large energy and execution time benefits. Here, we describe transformative NanoSystems (computing systems which exploit unique characteristics of emerging nanotechnologies) through synergistic innovations across multiple levels: N3XT 3D MOSAIC (material, device, integration technology), Illusion Scaleup (circuit, architecture), and Co-Design (technology, circuit, architecture, and application). We target 100-1,000× benefits in system-level Energy Delay Product (EDP) vs. today’s approaches.

**N3XT 3D MOSAIC** (Monolithic/Stacked/Assembled IC, Fig 1) [1] builds on N3XT 3D (Nano-Engineered Computing Systems Technology) [2] – which uses ultra-dense (e.g., monolithic) 3D to integrate heterogeneous technologies (logic, memory, sensing) (Fig. 1b). Multiple N3XT 3D chips are integrated via a continuum of chip stacking-/interposer-/wafer-level integration – the N3XT 3D MOSAIC. Ultra-dense 3D (≤ 100 nm pitch vertical connectivity) integration of multiple interleaved layers of memory and memory access logic is important to overcome the memory wall with 100-1,000× EDP benefits [1-2]. Monolithic 3D integration is one technique to achieve ultra-dense 3D by fabricating layers of heterogeneous technologies directly on top of other device layers. Ultra-dense nano-scale inter-layer vias (ILVs) – used for vertical routing in today’s back-end-of-line (BEOL) processes – are used for 3D connectivity. For monolithic fabrication of 3D layers, logic and memory on upper layers must be BEOL-compatible: fabricated at temperatures < 400°C and physically thin to be connected using ILVs (with small aspect ratios).

Fig.1: (a) **N3XT 3D MOSAIC** consisting of multiple N3XT 3D chips. (b) N3XT 3D chip (with some technology examples).

BEOL-compatible logic on 3D layers is required for low-energy and high-speed computing, efficient memory access circuitry, SRAM and 2T/3T gain cells, or selectors for memory cells. Examples include 2D materials, carbon nanotube FETs (CNFETs), oxide semiconducting FETs. CNFET digital circuits are projected to achieve 7× EDP benefits.

---

1 The work cited provides an expanded version of this TCVLSI article.
benefits over silicon CMOS at the 2 nm technology node [3]. BEOL-compatible memory technologies include magnetoresistive RAM (MRAM), Resistive RAM (RRAM), Ferroelectric FETs (FeFETs). Several NanoSystems have been demonstrated using technologies such as CNFETs, RRAM, and their monolithic 3D integration (Fig 2), creating a realistic “lab to fab” path [4] for large-scale 3D NanoSystems. Recently, these technologies (CNFETs, RRAM, monolithic 3D of CNFETs and RRAM) have been established in multiple industrial facilities: Analog Devices and SkyWater Technology foundry [5-6].

New thermal and power delivery solutions are critical for 3D. While existing thermal solutions may be sufficient for memory and memory access circuits on N3XT 3D layers [2], new thermal solutions will be required for dense compute on 3D layers.

Fig. 2: N3XT 3D demonstrations [1]. (a-c) 3D NanoSystem: monolithic 3D integration of CNFET, RRAM and silicon. (d) Hyperdimensional computing exploiting CNFETs, RRAM and their monolithic 3D integration. (e) Non-volatile microcontroller integrating RRAM and silicon. (f) CNFET RISC-V CPU in the BEOL. (g) Deep learning accelerator using foundry RRAM and silicon supporting incremental edge AI training. (h) Multi-3D-layer CNFET and RRAM from SkyWater Technology. (i) CNFETs in industrial manufacturing facilities: Analog Devices and SkyWater Technology Foundry.

Illusion [6] orchestrates workload execution on a system of multiple N3XT 3D chips in N3XT 3D MOSAIC, creating an illusion of a Dream Chip (Fig 3a-b) with near-Dream energy and throughput. Illusion follows the mantra: move computation, not data. With enough local (on-chip) memory (through N3XT 3D) and quick chip ON/OFF (e.g., through non-volatile on-chip memory such as RRAM), Illusion ensures: 1. Computations occur where their data resides to avoid massive inter-chip traffic. 2. Idle energy is eliminated by quickly turning ON/OFF individual chips. Illusion is thus distinct from traditional multi-chip parallel processing. Multiple Illusion hardware prototypes have been demonstrated as well: 6-chip and 8-chip Illusion (Fig 3c) for deep learning inference with additional < 5% energy and < 4% execution time vs. corresponding Dream Chips.

The combination of N3XT 3D MOSAIC and Illusion creates a new scaling path – Illusion Scaleup [1]. By linearly increasing 3D layers in each N3XT 3D chip (e.g., increased monolithic integration) and linearly improving chip-to-chip communication (e.g., moving through the integration continuum toward tighter chip integration), Illusion overheads (vs. the Dream chip) improve quadratically [1]. Such multiplicative benefits allow Illusion Systems on N3XT 3D MOSAICs to maintain near-Dream EDP (e.g., within 1.05x for energy and execution time) as deep learning application sizes grow over several (e.g., 10) generations.
Co-Design through synergistic innovations across abstraction levels (technology, circuit, architecture, software, application) is a key component of N3XT 3D MOSAIC (details in [1]). For example, the imperfection-immune paradigm [8] combines advanced nanofabrication with new circuit design and overcomes inherent imperfections in CNFETs (e.g., mis-positioned or metallic carbon nanotubes, device-to-device variations). This paradigm played a critical role in establishing CNFETs in multiple industrial silicon manufacturing facilities at Analog Devices and the SkyWater Technology Foundry. As another example, a new Low-Rank Training algorithm, combined with a new architecture-level endurance resilience technique (ENDURER), overcomes RRAM write challenges (energy, latency, limited endurance) and enables an RRAM-based system to achieve edge AI incremental training for 10 years at 20 samples/min. (with 340× better EDP vs. traditional Stochastic Gradient Descent).

To minimize the overall cost, new methods that focus on improving 3D fabrication throughput and 3D yield (beyond traditional miniaturization) are necessary [9]. New approaches that ensure robust system operation by overcoming design bugs, manufacturing defects, reliability failures and security attacks are also essential. For example, a new design verification technique called Symbolic QED was demonstrated to achieve drastic productivity benefits (from 6-9 person-months using industrial verification flow to 2 person-days using Symbolic QED) and high degrees of thoroughness in an industrial study by Infineon Technologies for their automotive cores [10].


About the Authors

Tathagata Srimani: Tathagata Srimani received the B.Tech degree in Electronics and Electrical Communication Engineering from IIT Kharagpur, Kharagpur, India in 2016, and the Ph.D. degree in Electrical Engineering and Computer Science from Massachusetts Institute of Technology (MIT) in 2022, under the direction of Professor Max Shulaker. He is currently a postdoctoral research scholar in Stanford Robust Systems Group with Professor Subhasish Mitra. His current research interests involve experimental demonstrations of 3D integrated circuits and systems leveraging emerging nanotechnologies. Mr Srimani was a recipient of the Irwin and Joan Jacobs Presidential Fellowship at MIT in 2016 and Morris Joseph Levin Award—Masterworks (S.M.) thesis presentation at MIT in 2018.
Robert Radway:

Robert M. Radway (Graduate Student Member, IEEE) received a B.S. degree and a M.Eng. degree in electrical engineering and computer science from the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, in 2016 and 2017, respectively. He is currently pursuing a Ph.D. degree in electrical engineering at Stanford University, Stanford, CA, USA, under the advisement of Prof. Subhasish Mitra. His research interests include multi-chip Illusion Systems for application scaleup, 3D MOSAICs (MOntolithic/Stacked/Assembled ICs), and edge AI inference and training using emerging non-volatile memories.

H. -S. Phillip Wong:

H. -S. Philip Wong is the Willard R. and Inez Kerr Bell Professor in the School of Engineering at Stanford University. He joined Stanford University as Professor of Electrical Engineering in September, 2004. From 1988 to 2004, he was with the IBM T.J. Watson Research Center. From 2018 to 2020, he was on leave from Stanford and was the Vice President of Corporate Research at TSMC, the largest semiconductor foundry in the world, and since 2020 remains the Chief Scientist of TSMC. He is a Fellow of the IEEE and received the IEEE Electron Devices Society J.J. Ebers Award, the society’s highest honor to recognize outstanding technical contributions to the field of electron devices that have made a lasting impact. He is the founding Faculty Co-Director of the Stanford SystemX Alliance — an industrial affiliate program focused on building systems, the faculty director of the Stanford Non-Volatile Memory Technology Research Initiative (NMTRI), and the faculty director of the Stanford Nanofabrication Facility.

Subhasish Mitra:

Subhasish Mitra is Professor of Electrical Engineering and of Computer Science at Stanford University. He directs the Stanford Robust Systems Group, leads the Computation Focus Area of the Stanford SystemX Alliance, and is a member of the Wu Tsai Neurosciences Institute. His research ranges across Robust Computing, NanoSystems, Electronic Design Automation (EDA), and Neurosciences. Results from his research group have influenced almost every contemporary electronic system, and have inspired significant government and research initiatives in multiple countries. He has held several international academic appointments — the Carnot Chair of Excellence in NanoSystems at CEA-LETI in France, Invited Professor at EPFL in Switzerland, and Visiting Professor at the University of Tokyo in Japan. Prof. Mitra also has consulted for major technology companies including Cisco, Google, Intel, Samsung, and Xilinx. His honors include the Harry H. Goode Memorial Award (by the IEEE Computer Society for outstanding contributions in the information processing field), Newton Technical Impact Award in EDA (test-of-time honor by ACM SIGDA and IEEE CEDA), the University Researcher Award (by the Semiconductor Industry Association and Semiconductor Research Corporation to recognize lifetime research contributions), the Intel Achievement Award (Intel’s highest honor), and the US Presidential Early Career Award. He and his students have published over 10 award-winning papers across 5 topic areas (technology, circuits, EDA, test, verification) at major venues including the Design Automation Conference, International Solid-State Circuits Conference, International Test Conference, Symposium on VLSI Technology, Symposium on VLSI Circuits, and Formal Methods in Computer-Aided Design. Stanford undergraduates have honored him several times "for being important to them." He is an ACM Fellow and an IEEE Fellow.
The 35th International Conference on VLSI Design & 20th International Conference on Embedded Systems (VLSID 2022), with IEEE/CS-TCVLSI as one of the conference sponsors, was held in Virtual Mode from Feb 26 – Mar 2, 2022 with “Silicon Catalyzing computing, Communication and Cognitive Convergence” as its theme.

The VLSID Conference provides a platform for showcasing R&D, Technology and product development from Academia and Industry in the field of Semiconductors and Embedded Systems.

VLSID 2022 received 2000+ attendees across Academia, Industry, Research & Government dignitaries. The conference had great coverage 150+ press clippings, 1M+ Social Media views. Below are few main attractions of VLSISD 2022.

Tutorials: 16 tutorials on leading edge topics delivered by Industry & Academia experts during first two days of the conference.

Keynote Talks: 14 Visionaries talks delivered by Global Though Leaders : Nick McKeown (Intel), Suk Hwan Lim (Samsung), Joe Sawicki (Siemens), Sanjive Agarwala (Cadence), Shankar Krishnamoorthy (Synopsys), Daisy Chittilapilli (Cisco), Ken Wiseman (Qualcomm), Sonum Wangchuk (Himalayan Institute of Alternatives, Ladakh), Manish Kothari (Silicon Labs), Ramgopal Roa (IIT Delhi), Phillip Wong (Stanford), Jason Cong (UCLA), Andrew B Kahng (UCSD), Subhasis Chaudhuri (IIT Bombay).

Panel Discussion: 3 panel discussion covering major topics : Shaping the Future of Semiconductors in India, Designeering Product Excellence (The Next Edge), Building Semiconductor Research & Talent.

Regular Technical Papers: 278 papers were received, out of which total 53 papers with 11 Invited Talks were presented. A Special Session on Design Automation Conference-DAC (VLSID is a sister conference) was held, in which top four DAC papers were presented including the 57th DAC Best Paper Award winner

Industry Forum (IF): 16 IF talks covering latest developments and offerings from industry leaders.
**Student Research Forum (SRF)**: To promote and encourage the young researchers in the VLSI Design and Embedded Systems domain, 8 Student Research Forum presentation was organized.

**User Designer Track (UDT)**: 10 UDT presentation covering industry & research problems were presented.

**Design Contest**: 2 design contest proposal were selected each for VLSI & Embedded Systems category.

Key contributors of the Conference Organizing committee:

**General Chairs**: Anil Kempanna (Intel), Lakshmy N. Kethamreddy (Samsung)

**Technical Program Chairs**: Mayam Shojaei (IIT, Bombay), Nele Mentes (KU Leuven), Taher Abbasi (Cadence)

**Organizing Chairs**: Chitra Hariharan (Intel), Preet Yadav (NXP Semiconductors)

More details about VLSID 2022 are available on conference website: [https://vlsid.org/](https://vlsid.org/)

For any further information please connect with: Preet Yadav (preet.yadav@ieee.org), Organizing Chair & IEEE Liaison, VLSID 2022

---

<table>
<thead>
<tr>
<th>Award</th>
<th>Paper Title</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.K. Choudhary Best Paper Award</td>
<td>MAPPARAT: A Resource Constrained FPGA-Based Accelerator for Sparse-Dense Matrix Multiplication</td>
<td>M. R. Ashuthosh, Santosh Krishna, Vishvas Sudarshan, Srinivasan Subramaniyan and Madhura Purnaprajna; PES University</td>
</tr>
<tr>
<td>Nripendra Nath Biswas Best Student Paper Award</td>
<td>Hardware Accelerator for Capsule Network based Reinforcement Learning</td>
<td>Dola Ram, Suraj Panwar and Kuruvilla Varghese; Indian Institute of Science</td>
</tr>
<tr>
<td>Naresh Malipeddy Honorable Mention Award</td>
<td>A 5-Gb/s PAM4 Voltage Mode Transmitter with Current Mode Continuous Time Linear Equalizer</td>
<td>Shraman Mukherjee, Sumantra Seth and Saurabh Saxena; Texas Instruments, India</td>
</tr>
</tbody>
</table>

---
**MAPPARAT: A Resource-constrained FPGA-based Accelerator for Sparse-Dense Matrix Multiplication**

M. R. Ashuthosh, Santosh Krishna, Vishvas Sudarshan, Srinivasan Subramaniyan, Madhura Purnaprajna
Centre for Heterogeneous and Intelligent Processing Systems, Dept. of Electronics and Communication Engineering, P.E.S. University, Bangalore, India

In recent times matrix multiplication has risen in importance due to its wide usage in deep-learning applications. The matrices used in such applications are mostly sparse. The energy consuming task in matrix multiplication is fetching the data from external memories. Exploitation of sparsity is crucial to avoid the redundant data transfer, computation and storage space. Matrix compression to exploit sparsity comes with the cost of area and time needed to decode the compressed data. Based on the sparsity, the right choice of compression format is necessary. The algorithm of matrix multiplication determines the number of accesses made to the matrices involved in the multiplication. Depending on the algorithm, the location of the matrices with more accesses is preferred to be kept near the compute device. Considering the target device constraints, questions arise such as, whether to exploit the sparsity or not? Which algorithm is suited for a given input matrix? Where to keep the matrices in a resource constrained device? We try to answer these questions and hence develop MAPPARAT [1], an FPGA-based accelerator for \( \text{sparse} \times \text{dense} \) matrix multiplication.

**Design of MAPPARAT:** Figure 1 shows the high-level organisation of MAPPARAT. We present two variants of MAPPARAT based on the algorithms used for \( \text{sparse} \times \text{dense} \) matrix multiplication, viz., row-wise and column-wise product. The sparse matrix is stored in COO format as it provides a balance between decoding costs (in terms of area and time) and compression. The on-chip memory is used to store the matrix that is more frequently accessed based on the choice of the algorithm. The two variants of MAPPARAT are designed considering the resource constraints of the FPGA device XC7A100TCSG324-1, present on the ARTY A7-100 board.

**Results:** Table 1 shows the implementation reports of MAPPARAT, for both the variants. The multiplication of matrices of the dimension 560 \( \times \) 560 with sparse matrix being 97.2\% sparse has been considered. The final output matrix obtained is the same in both the cases but there is difference in the performance and the amount of resources utilised. Row-wise product algorithm generates dense partial products in \( \text{sparse} \times \text{dense} \) matrix multiplication that can be accumulated without extra logic for matching the partial products. Column-wise product algorithm generates sparse partial products in \( \text{sparse} \times \text{dense} \) matrix multiplication that require extra logic for the accumulation. This has been achieved by implementing an adder tree. The multiplication stage will be idle during the accumulation phase by adder tree, leading to additional cycles and lower frequency in the case of the column-wise variant. We found that the row wise variant is 2.5 \( \times \) faster and consumes 2.5 \( \times \) less energy than the column-wise variant. With MAPPARAT, we demonstrate the inter-play of algorithm, compression formats, resource constraints and the overall impact on the performance of a \( \text{sparse} \times \text{dense} \) matrix multiplier.

**Table 1. MAPPARAT: Resource utilisation on XC7A100TC5SG324-1**

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>BRAM (%)</th>
<th>LUTs (%)</th>
<th>FFs (%)</th>
<th>DSPs (%)</th>
<th>Frequency (MHz)</th>
<th>Latency (cycles)</th>
<th>Energy (mJ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Row-wise</td>
<td>75(55%)</td>
<td>13830(21.81%)</td>
<td>7680(6.08%)</td>
<td>240(100%)</td>
<td>240.44</td>
<td>342720</td>
<td>2.48</td>
</tr>
<tr>
<td>Column-wise</td>
<td>97(71.22%)</td>
<td>9216(14.53%)</td>
<td>12288(9.09%)</td>
<td>192(80%)</td>
<td>226.19</td>
<td>853720</td>
<td>6.57</td>
</tr>
</tbody>
</table>

**References**
Convolutional neural networks are widely used in reinforcement learning. Capsule networks are gaining popularity over traditional neural in many classification tasks due to multidimensional activity vector consisting of neurons that represents the features of the entity. In this paper, we explore the capability of a Capsule Network for reinforcement learning (RL) based applications. We have also proposed a hardware accelerator for deep Q-learning that uses the capsule network as a deep Q-network instead of a convolutional neural network. We have implemented a capsule network-based deep Q-learning architecture for inference on the Xilinx Kintex UltraScale field-programmable gate array. The network has been tested on Pygame based environments. The hardware implementation achieves an overall speedup of 77.45x as compared to the software implementation of the capsule network for deep reinforcement learning on Intel Xeon CPU E5, @3.1GHz and 10.86x as compared to implementation on Nvidia Ge-Force GTX1080 GPU.

Figure 1 shows the hardware accelerator architecture and Figure 2 shows the proposed modified architecture of capsule networks for RL. The main idea is to test the capability of the capsule network in complex environments for RL. Capsule network is used as a deep Q-network instead of a CNN for deep Q-learning. Capsule network uses an iterative routing by agreement mechanism. Instead of using the length of the instantiation vector to represent the probability of digit present in an image, we have additionally added a fully connected layer at the end to the original capsule network for calculating the Q-values for each action to make it work for RL algorithms. The performance comparison between the Capsule network and CNN was carried out for three different game environments, to check the applicability of the networks to different conditions. CNN and capsule network, both works well in complex game environments and the Capsule network exceeds the CNN based network performance. As both CNN and deep Q-capsule network shows smooth improvement in Q-value, the algorithm does not experience any divergence issues. This suggests that RL algorithm can train the capsule network to learn to play complex games. This shows that the capsule network can be used instead of CNN for RL applications.
A 5-Gb/s PAM4 Voltage Mode Transmitter with Current Mode Continuous Time Linear Equalizer

Shraman Mukherjee, Sumantra Seth, and Saurabh Saxena †*

This work presents a 5 Gb/s PAM4 hybrid voltage-mode transmitter with current mode continuous-time linear equalization. The transmitter achieves a large output signal swing with a PAM4 CMOS output driver. CTLE is embedded in the transmitter’s output stage to compensate a range of channel loss without reducing the signal swing at the receiver front end. Fabricated in 65 nm CMOS process, the transmitter dissipates 20.4 mW in the output driver at 5 Gb/s while transmitting 1.1 V peak-to-peak differential output swing across 2 m UTP cable.

Energy-efficient wireline communication requires reducing power consumption in transmitting modulated bitstream, compensating channel loss, and recovering the received bits error-free. Most of the conventional energy-efficient line driver with multi-tap feed-forward equalizer (FFE) are implemented by voltage-mode architecture. TX-FFE suffers from supply-limited output swings. It reduces the received signal swing and constrains signal processing on the receiver (RX) side. In this work, the transmit signal of the proposed line driver architecture is not supply-voltage limited due to equalization, and the received signal swing is large after equalization.

In the line driver design, a current-mode linear equalizer is incorporated with a full-swing voltage-mode PAM4 output driver. Fig. 1 shows a simplified block diagram of the proposed hybrid transmitter. It consists of a PRBS generator, a PAM4 VM output driver, and a CM-CTLE. PRBS generates two pseudorandom bitstreams D0 and D1. D0 and D1 represent LSB and MSB of a PAM4 symbol, respectively. The VM driver provides a large transmit signal swing, but it is limited by VDDVM. The CM output driver amplifies the input signal’s high-frequency content (D0/D1) and adds to the transmitter’s output. VM and CM-CTLE output drivers transmit large signal swings with equalization and are not constrained by supply voltage, unlike FFE embedded in VM or CM output drivers.

![Block diagram of the proposed VM-CM hybrid transmitter.](image)

Figure 1: Block diagram of the proposed VM-CM hybrid transmitter.

References


†S. Mukherjee and S. Seth are with Texas Instruments (India) Pvt. Ltd.
* S. Saxena is with Indian Institute of Technology Madras.
Q1. What does VLSI mean to you?

More and more transistors packed into the same area! I can’t wait to see when and how we integrate over 1 billion transistors on a die: what kind of technology will get us there? what will the transistors look like? What type of innovation will we need?

Q2. Why do you think VLSI is important and how do you see the field of VLSI evolving ahead?

VLSI is key to unlocking the power of technology. It allows us to integrate tens of millions of transistors in a very small area and these transistors in turn are used to implement devices used in all facets of our lives. These devices range from complex processors to communication devices, storage devices, various medical devices as well as simple devices in our homes and community that improve the quality of our lives.

Q3. What is your typical day like as one of the senior most technical leaders in the world’s largest semiconductor company?

I don’t have “typical days” – they vary based on the stage of each of the projects that I am working on. Early on, we have a lot of brainstorming sessions as we discuss the new technologies that we will need to invent, develop, or change in order to deliver a leading-edge product in the future. When we are deep in execution, I could be reviewing a technical issue or poring over data and offering suggestions as to how to proceed or recommending investigations that might be required. I could be called upon to help in understanding a technical issue, for example why the results of a circuit simulation are not as expected as well as present project status to VPs, Senior VPs, and the CEO.

Q4. What motivated you into this field?

I fell into this field by accident and never left!

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

As we lived thru the COVID pandemic, it became clear to me that technology enabled thru VLSI (processors in data centers, networking technology powering the internet, laptops, tablets, phones, etc) helped keep our communities close, enabled us to keep educating our children, allowed us to communicate with each other despite distances and isolation, and many of us kept earning our livelihood while remaining safe at home. I feel proud to be part of a technical community that was able to contribute in such a way to our global family. I don’t know what our next challenge will be but I am confident and excited by the fact that we in VLSI will be able to help surmount it.

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

When I graduated from college in 1979, I fully expected that by 2020 we would be at parity with women making up 50% of engineers in VLSI. I must admit that I am very disappointed with where we are today and, so I have to agree that we in VLSI have not done a good job at welcoming women or students of color. There has been some progress in the last few years, but not enough. We, as a community need to do more to open our doors to women and students of color as I truly believe that the best innovation and creative ideas come about when a team is made up of diverse engineers who bring their varied experiences and perspectives to the table.

Q7. What is your key message to young girls who aspire to be like you someday?

Don’t be afraid to try something new and operate outside your comfort zone. Take risks, be creative, be ready to fail and try again.
Nevine Nassif is one of Intel’s 24 Senior Fellows (the highest technical level in the company). She has been with the company for more than 19 years and has significantly contributed to Intel's business in various facets of the design, convergence, and delivery of key Intel server products. Most recently, as Chief Engineer of Intel’s latest Xeon Server, Nevine has been instrumental in the continued development of our next generation of Xeon® processors for data centers, networks, and edge. Prior to joining Intel, Nevine started her career at Digital Equipment Corporation (later acquired by Compaq which was acquired by HP). At DEC, she worked on the development of VAX and Alpha CPUs. Her focus was design tool development included timing verification tools, power estimation, device modeling for fast simulation methods, and circuit simulation. At that time she worked part time in order to balance her time between work and home when her children were young. Nevine and her husband have 3 adult children who are pursuing careers in physics, architecture, and classics. Her eldest daughter is a post-doctoral researcher in Physics, her younger daughter is an architect focused on project management, and her son is working towards a PhD in ancient Greek poetry. Nevine holds a PhD in Electrical Engineering from McGill University in Montreal. Nevine holds 10 patents and has co-authored several papers.
Announcing the winners of the 2021 IEEE-CS TCVLSI awards.

IEEE-CS TCVLSI Distinguished Research Award : Prof Vijay Narayan (Penn State University)
IEEE-CS TCVLSI Distinguished Research Award : Prof Sri Parameswaran (University of New South Wales, Australia)
IEEE-CS TCVLSI Mid-Career Research Achievement Award : Prof Swaroop Ghosh (Penn State University)

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

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

About the winners:

Vijaykrishnan Narayanan is the Robert Noll Chair Professor of Computer Science and Engineering and Electrical Engineering at The Pennsylvania State University. He received his Bachelor's degree from SVCE, University of Madras in 1993 and his Ph.D. from University of South Florida in 1998. His research focuses on power-aware computing systems, technology-computer architecture interaction and embedded systems. His research has resulted in ~500 publications, several open-source tools and models, and hardware prototypes of embedded vision systems. He received the 2021 IEEE Computer Society Edward McCluskey Technical Achievement Award, 2020 Northeastern Association of Graduate Schools Geoffrey Marshall Mentoring Award and Penn State Engineering Alumni Society Premier Research Award. He is a Fellow of IEEE, ACM and National Academy of Inventors. He served as the general co-chair of the first IEEE CS Annual Symposium on VLSI, the founding co-editor-in-chief of ACM Journal of Emerging Technologies in Computing Systems, editor-in-chief of IEEE transactions on Computer-Aided Design for Integrated Circuits and Systems and the chair of IEEE TCVLSI and ACM SIGDA. He currently serves as the Associate Editor-in-Chief of IEEE Micro.

Sri Parameswaran is a Professor in the School of Computer Science and Engineering at the University of New South Wales. He was Acting Head of School of Computer Science and Engineering at the University of New South Wales from 2019 to 2020 where he served as Program Director for Computer Engineering and the Post Graduate Research coordinator. Prior to the University of New South Wales, he was an academic at the University of Queensland. He was a visiting academic at Kyushu University, Japan, and the University of California, in Irvine. He was a consultant at NEC Research Labs in Princeton and the Asian Development Bank. Sri Parameswaran is a current member of the Australian Research Council College of Experts. Sri Parameswaran served as the Editor in Chief of the IEEE Embedded Systems Letters from 2016 to 2019. He was the General Chair of IEEE/ACM International Conference on Computer-Aided Design (2017) and the Asia South Pacific Design Automation Conference (2012). He was an Executive Committee Member of the Design Automation Conference from 2021-2013. He serves or has served on the editorial boards of IEEE Transactions on Computer-Aided Design, ACM Transactions on Embedded Computing Systems, the EURASIP Journal on Embedded Systems, and the Design Automation of Embedded Systems. He has served on Program Committees at Design Automation Conference (DAC), Design and Test in Europe (DATE), the International Conference on Computer-Aided Design (ICCAD), the International Conference on Hardware/Software Code-sign and System Synthesis (CODES-ISSS), and the International Conference on Compilers, Architectures, and Synthesis for Embedded Systems (CASES). His research interests are in Design Automation, System Level Synthesis, Low power systems, High-Level Synthesis Computer architectures for Bioinformatics, and Network on Chips. Sri Parameswaran received his B.Eng degree from Monash University and his Ph.D. from The University of Queensland.

Swaroop Ghosh received the B.E. from IIT, Roorkee and Ph.D. from Purdue. He is an Associate Professor at Penn State. Prior to that, he was a Senior Research and Development Engineer at Intel. His research interests include quantum computing, low-power circuits and hardware security. His most notable accomplishments include DARPA Young Faculty Award (YFA), ACM SIGDA Outstanding New Faculty Award and YFA Director’s Fellowship. He is a Senior member of the IEEE and the National Academy of Inventors (NAI), and a Distinguished Speaker of the ACM.
Nvidia unveiled its next-generation GPU architecture — named Hopper, alongside the new flagship GPU using the Hopper architecture, the H100. Perhaps surprisingly, Nvidia has not opted to go down the trendy chiplets route favored by Intel and AMD for their mammoth GPUs. While the H100 is the first GPU to use HBM3, its compute die is monolithic, 80 billion transistors in 814mm² built on TSMC’s 4N process. Memory and compute are packaged via TSMC’s CoWoS 2.5D packaging.

Samsung Electronics introduced the industry’s first 512GB CXL memory module. Newly developed CXL memory packs 4x the memory capacity over the previous version, enabling a server to scale to tens of terabytes with only one-fifth of the system latency. Samsung also introduced an upgraded version of its open-source software toolkit that facilitates CXL memory deployment into existing and emerging IT systems.

India prepares to build the nation’s first chip fab. The project would invest about $3 billion in a 65–nm analog chip fab, according to a statement by India’s Next Orbit Ventures, a fund management firm that aims to kickstart the effort. Israel’s Tower Semiconductor said it will be a technology provider for the project. India has also been in talks with Intel, GlobalFoundries, and Taiwan Semiconductor Manufacturing Co. about setting up domestic operations to build more high-tech manufacturing in the country.

The U.S. government and domestic chipmakers strengthened separate efforts this week to pass the $52 billion CHIPS Act of stimulus measures. Separately, U.S. chipmakers Intel, Micron, and Analog Devices joined the Semiconductor Alliance, announcing an agreement to accelerate chip R&D and prototyping to build a more robust domestic industry. This includes advanced manufacturing amid increased global competition.

The recently announced Universal Chiplet Interconnect Express (UCIe) 1.0 specification covers the die–to–die I/O physical layer, die–to–die protocols, and a software stack model leveraging PCI Express (PCIe) and Compute Express Link (CXL) industry standards.

A project at the University of Rochester and the University of Erlangen–Nuremberg (FAU) has taken a step towards faster "lightwave electronics" and ultrafast computers. The group has demonstrated a logic gate that operates at femtosecond timescales, potentially opening the door to information processing at petahertz speeds according to the researchers.

Intel is extending its roadmap for infrastructure processors through 2026. The company’s IPUs (infrastructure processing units) are megachips that are designed to improve datacenter efficiency by offloading functions such as networking control, storage management and security that were traditionally run on a host CPU. Intel is developing IPUs in close collaboration with infrastructure providers that include Google and Microsoft.
TCVLSI Sponsored Conferences for 2022

Financially sponsored/co-sponsored conferences

- ARITH, IEEE Symposium on Computer Arithmetic
- ASAP, IEEE International Conference on Application-specific Systems, Architectures and Processors
  - ASAP 2022: [https://www.asap2022.org/](https://www.asap2022.org/) Virtual conference dates: July 12-14 2022
- ASYNC, IEEE International Symposium on Asynchronous Circuits and Systems
  - ASYNC 2022: [https://asyncsymposium.org/async2022/](https://asyncsymposium.org/async2022/) Virtual conference dates: TBD 2022
- iSES, (formerly IEEE-iNIS) IEEE International Smart Electronic Systems
  - IEEE iSES 2022: [https://ieee-ises.org/2022/](https://ieee-ises.org/2022/) Dec 21-23 2022, NIT Warangal, India
- ISVLSI, IEEE Computer Society Symposium on VLSI
  - ISVLSI 2022: [http://www.eng.ucy.ac.cy/theocharides/isvlsi22/](http://www.eng.ucy.ac.cy/theocharides/isvlsi22/) Virtual conference dates: July 4-6 2022
- IWLS, IEEE International Workshop on Logic & Synthesis – collocated with DAC
  - IWLS 2022: [https://www.iwls.org/iwls2022/](https://www.iwls.org/iwls2022/) Conference dates: July 18 -21, 2022
- SLIP, ACM/IEEE System Level Interconnect Prediction
  - SLIP 2022: [https://dl.acm.org/conference/slip/proceedings](https://dl.acm.org/conference/slip/proceedings) Date TBD

Technically Co-Sponsored Conferences for 2022

- VLSID, International Conference on VLSI Design
  - VLSID 2022: [https://vlsid.org/](https://vlsid.org/) Virtual conference dates: Feb 26 -March 2 2022

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