Siva Satyendra **Sahoo**, PhD

EDA & AI/ML, System-level Design, Computer Architecture, Research and Innovation

s.satyendra.sahoo@gmail.com https://cfaed.tu-dresden.de/pd-staff-siva

Bhubaneswar, India

Siva Satyendra Sahoo is a Senior Staff Engineer at Samsung Advanced Institute of Technology (SAIT), Samsung Semiconductor India Research (SSIR), Bangalore, India, working in the field of AI Computing. He has more than 6 years of experience in electronic system design and Electronic Design Automation.

Satyendra's expertise spans hardware/software co-design, embedded machine learning, system-level design, computer architecture, fault-tolerant computing, approximate computing and digital VLSI. Thanks to his interdisciplinary background, he has a successful record of accomplishments in both industry and academia.

In his current position, Satyendra is working on the efficient modelling of state-of-the-art AI inference engines. The modelling is primarily for functional validation and early-stage architectural exploration.

Prior to joining SAIT, Satyendra was working on the design of low-cost accelerators for AI/ML applications with the Chair for Processor Design at Technische Universität Dresden, Germany. He worked on the applications primarily targeted for embedded and cyber-physical systems. The design involved leveraging approximation and quantization techniques and innovations in system-, architecture- and IP-level for FPGA-based systems. His responsibilities in the research group also involved supervision of research interns and master's students.

Satyendra holds a Ph.D. in Electrical and Computer Engineering from the National University of Singapore, Singapore. He has more than 20 peer-reviewed scientific publications as well as some articles submitted for review.

Satyendra has lived in three countries and speaks English, Odia, and Hindi. His colleagues describe him as analytical, creative, and amicable.

Experience

Senior Staff Engineer, AI Computing, Samsung Advanced Institute of Technology (SAIT), Samsung, Bangalore, India — 2021–Present

Individual contributor for modelling and design space exploration (DSE) of AI computing architectures. Collaborator in the analysis and optimization of mapping AI workloads on modern inference engines.

Post-doctoral Researcher, Chair for Processor Design, TU Dresden, Dresden, Germany — 2019–2021

Subject matter expert and collaborator for various research projects involving Electronic System-level (ESL) design, AI/ML for Electronic Design Automation (EDA) and embedded machine learning design (see Projects below). Co-instructor for Embedded Hardware System Design course. Supervisor (and co-supervisor) for research interns.

Ph.D Scholar, Department of ECE, National University of Singapore, Singapore — 2015–2019

Research Scholar for projects that focused on optimizing the distribution of reliability methods across multiple layers of the system stack with appropriate scheduling of tasks on heterogeneous embedded systems (see Projects below). Expertise on using multiple optimization methods for multi-objective design space exploration (DSE) including AI/ML techniques. Graduate/Teaching Assistant for multiple courses related to electronic system design.

Digital Design Engineer, Intel India, Bangalore, India — 2012–2014

Worked in the domain of ***Physical Design*** of dense VLSI System-on-chip. Was primarily responsible for design quality assessment of full-chip as well as design blocks.

Co-instructor, Computer Science Department, TU Dresden, Germany — 2019-2020

Guest Researcher, Chair for Processor Design, TU Dresden, Germany — 2017

Assistant Systems Engineer, Tata Consultancy Services, Hyderabad — 2008–2010

Projects

Modelling of modern AI compute engines, AI Computing, SAIT, Samsung — 2021–Present

The project involves modeling of state-of-the art AI inference engines for functional validation and architecture exploration.

Application-specific Approximate Computing for FPGA-based Embedded Systems, Chair for Processor Design, TU Dresden — 2020–2021

The project involves designing application-specific approximate arithmetic operators for FPGA-based embedded systems. The project explores the possibility of using AI/ML in the design space exploration for searching novel approximate circuits that can leverage an application's implicit error tolerance to provide a low-cost accelerator designs.
As one of the primary collaborators, Satyendra was involved in the implementation of the DSE framework using Multi-objective Bayesian Optimization (MBO) and Genetic Algorithms (GA). Satyendra was also actively involved in the discussions leading to the novel idea of abstracting approximate operators as a binary string to represent the LUTs utilized in the operator's design.

Accelerator Design for Reinforcement Learning, Chair for Processor Design, TU Dresden 2020–2021
The project involves designing low-cost FPGA-based accelerators for implementing Reinforcement Learning(RL)-based control. The research aspects of the project include the design and implementation of novel techniques at multiple levels –algorithm, micro-architecture, and arithmetic operators. The project focuses on accelerating Q-Learning in the initial stage and delves into more complex Deep RL algorithms in the later stages.
As the **primary researcher** for the project, Satyendra has been contributing to: the survey of state-of-the-art, design, and implementation of novel techniques, determining the testing and characterization of the accelerator IPs, and the technical documentation. Satyendra has also been involved in the hiring and supervision of research interns for the project.

Embedded Machine Learning Design, Chair for Processor Design, TU Dresden 2019–Present
The project involves the design of low-power inference engines for embedded systems. Primarily, the research focuses on an application-specific design and efficient utilization of quantization, precision-scaling, and approximate arithmetic operators.
As a collaborator, Satyendra has contributed to the EDA aspects of the project. This includes application-specific DSE and characterization of the resulting designs. Satyendra has contributed to the related multi-objective optimization using evolutionary computing and Bayesian optimization.

Cross-layer Reliability-aware Embedded System Design, Dep. ECE, NUS Singapore 2015–2019
As part of his doctoral thesis, Satyendra contributed to: a survey of state-of-the-art, defining the scope of research, design, and implementation of novel DSE approaches to cross-layer optimization along with the documentation of the research results. The project focused primarily on system-level design and resulted in novel contributions to modeling and analysis for reliability-aware HW/SW co-design, HW/HW partitioning and Dynamic adaptation to varying operating conditions for embedded applications executing on heterogeneous hardware platforms.

Hardware Accelerator for Training of Support Vector Machine, Department of Electronic Systems Engineering, IISc Bangalore 2011–2012

Major publications

[ACM TECS] S. Ullah, S. S. Sahoo, N. Ahmed, D. Chaudhury, & A. Kumar. *AppAxO: Designing Application-specific Approximate Operators for FPGA-based Embedded Systems*, in **ACM Trans. Embed. Comput. Syst.** Just Accepted (January 2022). DOI:https://doi.org/10.1145/3513262

[IEEE TCAD] B. Ranjbar, A. Hosseinghorban, S. S. Sahoo, A. Ejlali & A. Kumar. *BOT-MICS: Bounding Time Using Analytics in Mixed-Criticality Systems*, in **IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems**, doi: 10.1109/TCAD.2021.3127867.

[IEEE TC] S. Ullah, H. Schmidl, S. S. Sahoo, S. Rehman & A. Kumar, *Area-optimized Accurate and Approximate Softcore Signed Multiplier Architectures*, in **IEEE Transactions on Computers**, doi: 10.1109/TC.2020.2988404.

[Integration] Sahoo, S. S., Nguyen, T. D. A., Veeravalli, B., & Kumar, A, *Multi-objective design space exploration for system partitioning of FPGA-based Dynamic Partially Reconfigurable Systems*, In **Integration, the VLSI Journal, Elsevier**, November 2018.

[VLSI-SoC 2021] S. S. Sahoo & A. Kumar, *Using Monte Carlo Tree Search for EDA – A Case-study with Designing Cross-layer Reliability for Heterogeneous Embedded Systems*, in 2021 IFIP/IEEE 29th International Conference on Very Large Scale Integration (**VLSI-SoC 2021**), 2021, pp. 1-6, doi: 10.1109/VLSI-SoC53125.2021.9606987.

[GLSVLSI 2021] S. S. Sahoo, A. R. Baranwal, S. Ullah & A. Kumar, *MemOReL: A Memory-oriented Optimization Approach to Reinforcement Learning on FPGA-based Embedded Systems*, in Proceedings of the 2021 on Great Lakes Symposium on VLSI (**GLSVLSI 2021**), 339-346.

[DAC 2021] Ullah, S., Sahoo, S. S., & Kumar, A. (2020). *CLAppED: A Design Framework for Implementing Cross-Layer Approximation in FPGA-based Embedded Systems.* In Proceedings of the 58th Annual Design Automation Conference **DAC '21**.

[DFTS 2020] S. S. Sahoo, B. Veeravalli and A. Kumar, ***Markov Chain-based Modeling and Analysis of Checkpointing with Rollback Recovery for Efficient DSE in Soft Real-time Systems***. (2020) IEEE International Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems (**DFT 2020**), Frascati, Italy, 2020, pp. 1-6, doi: 10.1109/DFT50435.2020.9250892.

[DAC 2020] Sahoo, S. S., Veeravalli, B., & Kumar, A. (2020). ***CL(R)Early: An Early-stage DSE Methodology for Cross-Layer Reliability-aware Heterogeneous Embedded Systems.*** In Proceedings of the 57th Annual Design Automation Conference. **DAC '20**. A Virtual Experience.

[DAC 2019] Sahoo, S. S., Veeravalli, B., & Kumar, A. (2019). ***A Hybrid Agent-based Design Methodology for Dynamic Cross-layer Reliability in Heterogeneous Embedded Systems.*** In Proceedings of the 56th Annual Design Automation Conference. **DAC '19**. Las Vegas, NV, USA.

Awards/Grants

HiPEAC Collaboration Grants, HiPEAC4 Network of Excellence — Jan 2017
NUS Graduate Scholarship, National University of Singapore, Singapore. — 2015 – 2018
MHRD (GATE) Scholarship, Government of India, India — 2010 – 2012
CET Merit Scholarship, CET, Bhubaneswar, India. — 2004 – 2008

Education

Doctor of Philosophy (Ph.D.), Dep. of Electrical and Computer Engineering, National University of Singapore, Singapore — 2019
Master of Technology (M.Tech.), Dep. of Electronic Systems Engineering, Indian Institute of Science, Bangalore, India — 2012
Bachelor of Technology (B.Tech.), Dep. of Instrumentation and Electronics Engineering, College of Engineering and Technology, Bhubaneswar, India — 2008

Certifications

Neural Networks and Deep Learning, DeepLearning.AI (Coursera) — 2017
Mathematics for Machine Learning, Imperial College London (Coursera) — 2018
R Programming, Johns Hopkins University (Coursera) — 2017

Skills

Design Skills: Hardware IP (RTL and HLS), Physical Design, EDA Tool Design
Design Languages: VHDL, SystemC, Verilog
Programming Languages: C, C++, Python, MATLAB, Tcl
AI/ML Frameworks: Tensorflow, PyTorch, OpenAI Gym
EDA Tools: Xilinx Vitis & Vivado, Synopsys Design Compiler

Languages

English : Full professional proficiency
Odia: Native proficiency
Hindi: Limited working proficiency

Personal interests

Science, Technology, Cricket, Badminton, Tennis, Traveling, Food, Beer, Movies, Music, Hiking

Publications

20+ publications in peer-reviewed conference proceedings and journals.

References

Akash Kumar, Professor, Technische Universität Dresden, Dresden, Germany.
Bharadwaj Veeravalli, Associate Professor, National University of Singapore, Singapore.
Kuruvilla Varghese, Principal Research Scientist, Indian Institute of Science, Bangalore, India.

List of all publications:

Journals:

[J7] S. Ullah, S. S. Sahoo, N. Ahmed, D. Chaudhury, & A. Kumar. *AppAxO: Designing Application-specific Approximate Operators for FPGA-based Embedded Systems*, in **ACM Trans. Embed. Comput. Syst.** Just Accepted (January 2022). DOI:https://doi.org/10.1145/3513262

[J6] B. Ranjbar, A. Hosseinghorban, S. S. Sahoo, A. Ejlali & A. Kumar. *BOT-MICS: Bounding Time Using Analytics in Mixed-Criticality Systems*, in **IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems**, doi: 10.1109/TCAD.2021.3127867.

[J5] S. S. Sahoo, Ranjbar, B., & Kumar, A, *Reliability-aware Resource Management in Multi-/Many-core Systems: A Perspective Paper*, in **MDPI Journal of Low Power Electronics and Applications**, 2021.

[J4] Nambi, S., Ullah, S., Lohana, A., S. S. Sahoo, Merchant, F., & Kumar, A, *ExPAN(N)D: Exploring Posits for Efficient Artificial Neural Network Design in FPGA-based Edge Processing.* in **IEEE Access**, vol. 9, pp. 103691-103708, 2021, doi: 10.1109/ACCESS.2021.3098730.

[J3] A. R. Baranwal, S. Ullah, S. S. Sahoo & A. Kumar, ReLAccS: *A Multi-level Approach to Accelerator Design for Reinforcement Learning on FPGA-based Systems*, in **IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems**, doi: 10.1109/TCAD.2020.3028350.

[J2] S. Ullah, H. Schmidl, S. S. Sahoo, S. Rehman & A. Kumar, *Area-optimized Accurate and Approximate Softcore Signed Multiplier Architectures*, in **IEEE Transactions on Computers**, doi: 10.1109/TC.2020.2988404.

[J1] Sahoo, S. S., Nguyen, T. D. A., Veeravalli, B., & Kumar, A, *Multi-objective design space exploration for system partitioning of FPGA-based Dynamic Partially Reconfigurable Systems*, In **Integration, the VLSI Journal, Elsevier**, November 2018.

Conference Proceedings

[C15] S. S. Sahoo, A. Kumar, M. Decky, S. C. B. Wong, G. V. Merrett, Y. Zhao, Jiachen Wang, X. Wang & A. K. Singh, *Emergent design challenges for embedded systems and paths forward: mixed-criticality, energy, reliability and security perspectives*, in Proceedings of the 2021 International Conference on Hardware/Software Codesign and System Synthesis (**CODESS, ESWeek 2021**). Association for Computing Machinery, New York, NY, USA, 1–10. DOI:https://doi.org/10.1145/3478684.3479246

[C14] S. S. Sahoo & A. Kumar, *Using Monte Carlo Tree Search for EDA – A Case-study with Designing Cross-layer Reliability for Heterogeneous Embedded Systems*, in 2021 IFIP/IEEE 29th International Conference on Very Large Scale Integration (**VLSI-SoC 2021**), 2021, pp. 1-6, doi: 10.1109/VLSI-SoC53125.2021.9606987.

[C13] S. S. Sahoo & A. Kumar, *CLEO-CoDe: Exploiting Constrained Decoding for Cross-Layer Energy Optimization in Heterogeneous Embedded Systems*, in 2021 IFIP/IEEE 29th International Conference on Very Large Scale Integration (**VLSI-SoC 2021**), 2021, pp. 1-6, doi: 10.1109/VLSI-SoC53125.2021.9606983.

[C12] S. S. Sahoo, A. R. Baranwal, S. Ullah & A. Kumar, *MemOReL: A Memory-oriented Optimization Approach to Reinforcement Learning on FPGA-based Embedded Systems*, in Proceedings of the 2021 on Great Lakes Symposium on VLSI (**GLSVLSI 2021**), 339-346.

[C11] Ullah, S. Sahoo, S. S., & Kumar, A. (2020). *CLAppED: A Design Framework for Implementing Cross-Layer Approximation in FPGA-based Embedded Systems.* In Proceedings of the 58th Annual Design Automation Conference. **DAC '21**(Accepted for publication).

[C10] Ranjbar, B., Hoseinghorban, A., Sahoo, S. S., Ejlali, A. & Kumar, A. (2020). *Improving the Timing Behaviour of Mixed-Criticality Systems Using Chebyshev's Theorem.* In Design, Automation & Test in Europe Conference & Exhibition, **DATE 2021**, Virtual Conference and Exhibition, February 01-05, 2021.

[C9] S. S. Sahoo, B. Veeravalli and A. Kumar, *Markov Chain-based Modeling and Analysis of Checkpointing with Rollback Recovery for Efficient DSE in Soft Real-time Systems*. (2020) IEEE International Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems (**DFT 2020**), Frascati, Italy, 2020, pp. 1-6, doi: 10.1109/DFT50435.2020.9250892.

[C8] Sahoo, S. S., Veeravalli, B., & Kumar, A. (2020). *CL(R)Early: An Early-stage DSE Methodology for Cross-Layer Reliability-aware Heterogeneous Embedded Systems.* In Proceedings of the 57th Annual Design Automation Conference. **DAC '20**. A Virtual Experience.

[C7] Rai, S., Raitza, M., Sahoo, S. S., & Kumar, A. (2020). ***DISCERN: Distilling Standard Cells for Emerging Reconfigurable Nanotechnologies.*** In 2020 Design, Automation & Test in Europe Conference & Exhibition, **DATE 2020**, Virtual Conference and Exhibition, April 21- May 31, 2020.

[C6] Sahoo, S. S., Veeravalli, B., & Kumar, A. (2019). ***A Hybrid Agent-based Design Methodology for Dynamic Cross-layer Reliability in Heterogeneous Embedded Systems.*** In Proceedings of the 56th Annual Design Automation Conference. **DAC '19**. Las Vegas, NV, USA.

[C5] Sahoo, S. S., Nguyen, T. D. A., Veeravalli, B., & Kumar, A. (2018). ***QoS-aware Cross-layer Reliability-integrated FPGA-based Dynamic Partially Reconfigurable System Partitioning.*** In International Conference on Field-Programmable Technology, **ICFPT 2018**, Naha, Okinawa, Japan, December 10-14, 2018.

[C4] Sahoo, S. S., Nguyen, T. D. A., Veeravalli, B., & Kumar, A. (2018). ***Lifetime-aware design methodology for dynamic partially reconfigurable systems.*** In 23rd Asia and South Pacific Design Automation Conference, **ASP-DAC 2018**, Jeju, Korea (South), January 22-25, 2018(pp. 393–398).

[C3] Sahoo, S. S., Veeravalli, B., & Kumar, A. (2018). ***CLRFrame: An analysis framework for designing cross-layer reliability in embedded systems.*** In 31st International Conference on VLSI Design and 17th International Conference on Embedded Systems, **VLSID 2018**, Pune, India, January 6-10, 2018 (pp. 307–312).

[C2] Sahoo, S. S., Kumar, A., & Veeravalli, B. (2016). ***Design and evaluation of reliability-oriented task re-mapping in MPSoCs using time-series analysis of intermittent faults.*** In 2016 Design, Automation & Test in Europe Conference & Exhibition, **DATE 2016**, Dresden, Germany, March 14-18, 2016 (pp. 798–803).

[C1] Sahoo, S. S., Veeravalli, B., & Kumar, A. (2016). ***Cross-layer fault-tolerant design of real-time systems.*** In 2016 IEEE International Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems, **DFT 2016**, Storrs, CT, USA, September 19-20, 2016 (pp. 63–68).

Book Chapters

[BC3] S. Ullah, S. S. Sahoo, & A. Kumar, *Efficient Hardware Arithmetic for Embedded Machine Learning*, in **Embedded Machine Learning for Cyber-Physical, IoT, and Edge Computing, Springer 2022**. *To appear*.

[BC2] Ranjbar, B., Sahoo, S. S., Singh, A., Dziurzanski, P., & Kumar, A., *Power management of Multicore systems*, in **Handbook of Computer Architecture, Springer 2022**. *To appear*.

[BC1] Sahoo, S. S., & Kumar, A., *Fault-tolerant Computer Architectures*, in **Handbook of Computer Architecture, Springer 2022**. *To appear*.