Postdoctoral Fellow at Harvard University

60 Pinney St, Apt 103, Ellington, CT, 06029 979-739-0421 devashree.tripathy@email.ucr.edu

SUMMARY

My interest lies in Computer Architecture, Domain Specific SoC design, GPGPU Architecture design, Machine Learning applications to SoC Design, High Performance Computing, Fault-Tolerance systems. Recent highlights include:

- Design and modelling of energy-efficient Domain Specific SoCs,
- Building data-dependent programming support for GPUs and Big Data accelerators,
- Boosting GPU performance using the intelligent data-locality analysis,
- Smart power management techniques for Scientific and Big Data applications in high-performance computing systems,
- Integration of the neuromorphic with RISC V architecture.

EDUCATION

Postdoctoral Fellow in Computer Science

HARVARD UNIVERSITY | Cambridge, MA, USA | 2021- current Research Topic: Design and Modeling of Domain Specific SoCs.

Advisor: Dr. David Brooks

PhD in Computer Science (CGPA: 3.9/4.0)

UNIVERSITY OF CALIFORNIA, RIVERSIDE (UCR) | Riverside, CA, USA | 2015-2021 Thesis: Improving Performance and Energy Efficiency of GPUs through Locality Analysis.

Advisor: Dr. Laxmi N. Bhuyan

Master of Technology, Advanced Electronics Systems (CGPA: 8.53/10)

CSIR - CENTRAL ELECTRONICS ENGINEERING RESEARCH INSTITUTE (CSIR-CEERI) | Pilani, India | 2014

Thesis: Patient Assistance System Using Brain Computer Interface.

Advisor: Dr. Jagdish Lal Raheja

Bachelor of Technology, Electronics and Telecommunications Engineering (CGPA: 9.69/10)

VEER SURENDRA SAI UNIVERSITY OF TECHNOLOGY (VSSUT, FORMERLY UCE) | BURLA, India | 2012

• Governor's Gold Medal for being ranked first among all undergraduate students in university. | University Topper

NOTABLE PROJECTS

Postdoctoral Fellow at Harvard University

2021-Present

1. FARSI: FACEBOOK AR SYSTEM INVESTIGATOR FOR AGILE DOMAIN-SPECIFIC SYSTEM-ON-CHIP EXPLORATION (UNDER FINAL REVIEW IN ACM TRANSACTIONS TECS)

- Domain-specific SoCs (DSSoCs) are attractive solutions for domains with stringent power/performance/area constraints; however, they suffer from two fundamental complexities. On the one hand, their many specialized hardware blocks result in complex systems and thus high development effort. On the other, their many system knobs expand the complexity of design space, making the search for the optimal design difficult. Thus to reach prevalence, taming such complexities is necessary.
- This work identifies necessary features of an early-stage design space exploration (DSE) framework that targets the complex design space of DSSoCs and further provides an instance of one called FARSI, (F)acebook (AR) (S)ystem (I)nvestigator. Concretely, FARSI provides an agile system-level simulator with speed up and accuracy of 8,400X and 98.5% comparing to Synopsys Platform Architect.

- FARSI also provides an efficient exploration heuristic and achieves up to 16X improvement in convergence time comparing to naive simulated annealing (SA). This is done by augmenting SA with architectural reasoning such as locality exploitation and bottleneck relaxation. Furthermore, we embed various co-design capabilities and show that on average, they have a 32% impact on the convergence rate.
- Finally, we demonstrate that using simple development-cost-aware policies can lower the system complexity, both in terms of the component count and variation by as much as 150% and 118% (e,g., for Network-on-a-Chip subsystem).

2. PerfSAGE: Generalized Runtime Performance Predictor for Arbitrary Deep Learning Models

- Deep learning models are commonly deployed on various hardware platforms for real time applications. For
 deployment, merely model accuracy or prediction performance are not enough. Run-time performance metrics
 (i.e. latency, memory footprint, energy consumption and etc.) become equally important factors for choosing
 optimal model for deployment. At the same time, it is time-consuming and unscalable to benchmark for those
 metrics. Hence, providing accurate prediction for those metrics, given a arbitrary model, could greatly
 accelerate deployment process.
- In this work, we present a dataset based on benchmarking results from commonly used deployment hardware. We also propose a novel predictor based on graph neural networks for runtime performance prediction. Unlike previous works, this predictor could generalize to models from diverse design space, and provides accurate prediction for various metrics with no higher than 5% Mean Absolute Percentage Error

3. FAST ATTENTION-BASED EARLY-DSE WITH HETEROGENEOUS RESOURCE MODELLING

Uses a GNN-based infrastructure to predict accelerator qualities such as LUT, DSP and Flip Flops resource
consumption as well as the execution time. The paper is focuses on graph classification as profitable or nonprofitable as well as directly predicting these quantities.

4. SSAGE: SSA-BASED MERGED ACCELERATOR GENERATION WITH ML-DRIVEN ENERGY AWARENESS (UNDER REVIEW IN DAC)

- Designing accelerators for resource and power constrained applications is a daunting task, especially in the era of domain specific accelerators which require massive amounts of reuse in Datapath and control logic.
 - The design automation community has been addressing the accelerator merging problem at a low level with resource sharing, usually in the context of RTL-level optimizations. We refer to this set of techniques as Fine-Grained Merging since they usually occur locally at a Basic Block granularity.
 - Coarse grained function merging (CGFM) works at the level of arbitrary function granularities where the control and dataflow abstractions are the focus of the optimization and without the need to deal with complex RTL related concepts such as wires and registers or technology details such as Lookup Tables (LUTs), DSPs, Flip Flops, Pins etc. The main advantage of CGFM is that it views circuitry connected via control and dataflow as a whole as opposed to individual components to reuse, revealing more opportunity for area savings in the resulting larger design space.
- The recent breakthroughs in coarse grained Control and Data Flow graphs (CDFGs) motivate us to adapt an IR driven methodology for accelerator merging. Since function merging optimizes the total number of instructions, in this paper we expand that methodology with an instruction cost model that enables us to drive the function merging code to prioritize merging the power-hungry parts of the CDFG.
- In this paper, we present the novel usage of compiler techniques that efficiently handle the SSA representation of the intermediate representation to fundamentally reduce the amount of circuitry when combining accelerators, bringing power savings with respect to state-of-the-art research and industry-level algorithms respectively in the context of FPGA technology. We further introduce a Machine Learning LASSO power model to predict in high dimensional LLVM-instruction type space the best power-reduction-friendly code generation. As a result, the power savings are of 57.35% with respect to the state of the art.

PhD in Computer Science

2015-2021

1. SEER: ESTIMATING RUNTIME DEPENDENCIES IN GPU APPLICATIONS

• There is a growing trend for execution efficiency of data-dependent workloads on GPUs. Over the years, there have been attempts to provide a framework for the user to express the data dependencies among the GPU

thread blocks for a more fine-grained scheduling process and minimizing unnecessary delays traditionally inflicted on GPU applications, such as kernel launch overheads or redundant inter-kernel synchronizations. Recently, with applications becoming harder for the users to decipher such data for the GPUs, compiler and just-in-time analysis have shown some promise in identifying data dependencies between relevant kernels before their execution, relieving the programmer from this burden to a great extent. However, run-time data dependencies are not uncommon in GPU applications, limiting the scope of the such analyses.

• In this work, we propose SEER, a framework to estimate data dependencies, including run-time dependencies, by predicting memory access patterns for GPU kernels from their intermediate representation using a machine learning model. This can give the GPU an edge for a more efficient thread block scheduling policy for multiple kernels simultaneously. %, especially when unified memory and/or It can also help when transferring data to the host for further processing would not be an easy option.

2. Intelligent Data Locality extraction using Just In Time (JIT) Compiler (NAS 2021)

- Graphics Processing Unit (GPU) has been adopted to process graphs effectively. Recently, multi-GPU systems are also exploited for greater performance boost. To process graphs on multiple GPUs in parallel, input graphs should be partitioned into parts using partitioning schemes. The partitioning schemes can impact the communication overhead, locality of memory accesses, and further improve the overall performance.
- We found that both intra-GPU data sharing and inter-GPU communication can be summarized as inter-TB communication. Based on this key idea, we propose a new graph partitioning scheme by redefining the input graph as a TB Graph with calculated vertex and edge weights, and then partition it to reduce intra \& inter-GPU communication overhead and improve the locality at the granularity of Thread Blocks (TB). We also propose to develop a partitioning and mapping scheme for heterogeneous architectures including physical links with different bandwidths.
- The experimental results on graph partitioning show that our scheme is effective to improve the overall performance of the Breadth First Search (BFS) by up to 33%.

3. PAVER: LOCALITY GRAPH-BASED THREAD BLOCK SCHEDULING FOR GPUS (ACM TRANSACTIONS TACO 2021)

- The massive parallelism present in GPUs comes at the cost of reduced L1 and L2 cache sizes per thread, leading to serious cache contention problems such as thrashing. Hence, the data access locality of an application should be considered during thread scheduling to improve execution time and energy consumption. Recent works have tried to use the locality behavior of regular and structured applications in thread scheduling, but the difficult case of irregular and unstructured parallel applications remains to be explored.
- We present PAVER, a priority-aware vertex scheduler, which takes a graph-theoretic approach towards thread scheduling. We analyze the cache locality behavior among thread blocks (TBs) through a just-in-time (JIT) compilation and represent the problem using a graph representing the TBs and the locality among them. This graph is then partitioned to TB groups that display maximum data sharing, which are then assigned to the same SM by the locality-aware TB scheduler.
- Through exhaustive simulation in Fermi, Pascal and Volta architectures using a number of scheduling techniques, we show that PAVER reduces L2 accesses by 43.3%, 48.5%, 40.21% and increases the average performance benefit by 29%, 49.1%, 41.2% for the benchmarks with high inter-TB locality.

4. GreenMM/ Saou: energy efficient GPU matrix multiplication through Undervolting and Overclocking (ICS, 2019, ISLPED 2020)

- Explored different errors and derived a fault model as a function of undervolting levels, overclocking frequency and data size.
- The current trend of ever-increasing performance in scientific applications comes with tremendous growth in energy consumption. In this paper, we present a framework for GPU applications, which reduces energy consumption in GPUs through Safe Overclocking and Undervolting (SAOU) without sacrificing performance. The idea is to increase the frequency beyond the safe frequency f_{safeMax} and undervolt below V_{safeMin} to get maximum energy saving. Since such overclocking and undervolting may give rise to faults, we employ an enhanced checkpoint-recovery technique to cover the possible errors. Empirically, we explore different errors and derive a fault model that can set the undervolting and overclocking level for maximum energy saving. We target cuBLAS Matrix Multiplication (cuBLAS-MM) kernel for error correction using the checkpoint and

recovery (CR) technique as an example of scientific applications. In case of cuBLAS, SAOU achieves up to 22% energy reduction through undervolting and overclocking without sacrificing the performance.

5. Slumber: Static-Power Management for GPGPU Register Files (ISLPED 2020)

- The leakage power dissipation has become one of the major concerns with technology scaling. The GPGPU register file has grown over last decade to support the parallel execution of thousands of threads. Given that each thread has its own dedicated set of physical registers, these registers remain idle when corresponding threads go for long latency operation.
 - Existing research shows that the leakage energy consumption of the register file can be reduced by under volting the idle registers to a data-retentive low-leakage voltage (Drowsy Voltage) to ensure that the data is not lost while not in use.
- In this paper, we develop a realistic model for determining the wake-up time of registers from various undervolting and power gating modes. Next, we propose a hybrid energy saving technique where a combination of power-gating and under-volting can be used to save optimum energy depending on the idle period of the registers with a negligible performance penalty. Our simulation shows that the hybrid energy-saving technique results in 94% leakage energy savings in register files on an average when compared with the conventional clock gating technique and 9% higher leakage energy saving compared to the state-of-art technique.

6. REDESIGNING EFFICIENT FETCH UNITS IN GPGPUS

• Redesigned the GPGPU Fetch Units to eliminate the stalls caused in the pipeline due to the invalid instructions in the instruction buffers.

7. WIREFRAME: Supporting Data-dependent Parallelism through Dependency Graph Execution in GPUs (MICRO, 2017)

- GPUs lack fundamental support for data-dependent parallelism and synchronization. While CUDA Dynamic Parallelism signals progress in this direction, many limitations and challenges still remain.
- This paper introduces *Wireframe*, a hardware-software solution that enables generalized support for data-dependent parallelism and synchronization. Wireframe enables applications to naturally express execution dependencies across different thread blocks through a *dependency graph* abstraction at run-time, which is sent to the GPU hardware at kernel launch. At run-time, the hardware enforces the dependencies specified in the dependency graph through a dependency-aware thread block scheduler.
- Overall, Wireframe is able to improve total execution time up to 65.20% with an average of 45.07%.

8. LATENT SEMANTIC INDEXING USING NON-NEGATIVE MATRIX FACTORIZATION ON GPU

- Used CUDA to implement the iterative updates of NMF factorization
- Analyzed the effectiveness of NMF for extracting latent semantic information from TREC documents Dataset w.r.t SVD.

9. IMPLEMENTATION OF TWO-LEVEL ROUND ROBIN WARP SCHEDULING AND DYNAMIC WARPS ON GPGPU-SIM

• Used C++ to modify the source code of GPGPU-Sim, specifically its warp scheduling mechanism, for performance improvement.

Master of Technology, Advanced Electronics Systems

2012-2014

1. PATIENT ASSISTANCE SYSTEM USING BRAIN COMPUTER INTERFACE (MTECH THESIS, BOOK IN SPRINGER)

• Used MATLAB to create the BCI Application to detect and quantify the features of the brain signals which indicated the user's intentions and translated these features into device commands to accomplish the user's intent. The project achieved a classification accuracy of 96%.

2. OPTICAL CHARACTER RECOGNITION USING NEURAL NETWORKS

- Processed the scanned images using feature extraction techniques.
- Customized neural network Algorithm was used to obtain a classification accuracy of 90% of the extracted characters.

- 3. AN IMPROVED SOBEL EDGE DETECTION
- Used MATLAB to improve the sobel edge detection of the Noisy images using Wavelet Transform.
 - 4. FINGERPRINT RECOGNITION USING GABOR FILTER
- Used VHDL to implement the fingerprint image enhancement module and Gabor filter on Virtex II Pro FPGA.

Bachelor of Technology, Electronics & Telecommunication Engineering

Sept - Dec 2011

HOME APPLIANCE CONTROL USING SMS (DISSERTATION PROJECT)

• SIM 300 GSM Modem used to send text message(SMS) and enable serial communication with microcontroller (ATMEGA 128).

PEER-REVIEWED PUBLICATIONS [TOTAL: 17]

Detailed list in google scholar.

- International Journals [7 published]
- International Conferences [6 published, 2 Under review]
- 1 SpringerBriefs Book Published
- 1 Springer Book Chapter Published

International Journals [6 published, 1 Under Final review]

 Tripathy, D., A. Abdolrashidi, L. Bhuyan, L. Zhou and D. Wong, "PAVER: Locality Graph-based Thread Block Scheduling for GPUs" ACM Transactions on Architecture and Code Optimization (ACM Transactions TACO, 2021), vol 18, issue 3, article 32, pp 1-26. * Invited to present at European Network on High Performance and Embedded Architecture and Compilation (HiPEAC, 2022) on June 20-22, Budapest, Hungary. https://doi.org/10.1145/3451164

Impact Factor: 1.309(2019), SCImago Journal Rank (SJR): 0.263, H-index: 41, SJR Quartile: Q3

Citation count: 11

- 2. Boroujerdian, B., Y. Jing, **D. Tripathy**, A. Kumar, L. Subramanian, L. Yen, V. Lee, V. Venkatesan, A. Jindal, R. Shearer, V. J. Reddi, "FARSI: Facebook AR System Investigator for Agile Domain-Specific System-on-Chip Exploration" (To Appear in *ACM Transactions TECS 2022- ACM Transactions on Embedded Computing Systems*)
- 3. Patel, D.K., **D. Tripathy**, & C. Tripathy, "An improved load-balancing mechanism based on deadline failure recovery on <u>GridSim"</u>. Engineering with Computers 32, 173–188 (Springer 2016).

https://doi.org/10.1007/s00366-015-0409-y

Impact Factor: 7.963 (2020), H-index: 52, SJR Quartile: Q1

Citation count: 12

- Patel, D.K., D. Tripathy and C. Tripathy, "Survey of load balancing techniques for Grid." Journal of Network and Computer Applications, Vol 65, 103–119 (Elsevier 2016). https://doi.org/10.1016/j.jnca.2016.02.012
 Impact Factor: 6.281 (2020), H-index: 105, SJR Quartile: Q4
 Citation count: 30
- ortation count. 50
- 5. Tripathy, L., **D. Tripathy** and C. Tripathy, "Fault Tolerance in Interconnection Network-a Survey." Research Journal of Applied Sciences, Engineering and Technology 2015, 198–214.

H-index: 27, SJR Quartile: Q1

https://doi.org/10.19026/rjaset.11.1708

Impact Factor: 0.22(2016).

6. Pattanayak, D., **D. Tripathy** and C. Tripathy, "<u>Star-mobius cube: a new interconnection topology for large scale parallel processing"</u>, *International Journal of Emerging Technologies in Computational and Applied Sciences (IJETCAS) 2014*, Issue 1, Vol 7, pp 62-68.

Citations: 3

7. <u>Das, B.J., **D. Tripathy** and B. Mishra, "The Crossed cube-Mesh: A New Fault-Tolerant Interconnection Network Topology for Parallel Systems",</u> *International Journal of Emerging Technologies in Computational and Applied Sciences (IJETCAS) 2014*, Issue 1, Vol 7, pp 211-219.

International Conferences [6 published, 2 Under review]

1. **Tripathy, Devashree,** Hadi Zamani, Debiprasanna Sahoo, Laxmi Narayan Bhuyan, and Manoranjan Satpathy "Slumber: Static-Power Management for GPGPU Register Files" *ACM/IEEE International Symposium on Low Power Electronics and Design (ISLPED, 2020)*, August 10-12, Boston, Massachusetts, USA. (Virtual due to Covid-19). https://doi.org/10.1145/3370748.3406577

Research Impact Score: 3.13, H-index: 61

Citation count: 13

2. **Tripathy, Devashree,** AmirAli Abdolrashidi, Quan Fan, Daniel Wong and And Manoranjan Satpathy, "LocalityGuru: A PTX Analyzer for Extracting Thread Block-level Locality in GPGPUs," *IEEE International Conference on Networking, Architecture and Storage (NAS, 2021)*, October 24-26, Riverside, California, USA.

 $https://doi.org/\underline{10.1109/NAS51552.2021.9605411}$

Citation count: 5

3. Abdolrashidi, AmirAli, **Devashree Tripathy**, Mehmet Esat Belviranli, Laxmi Narayan Bhuyan, and Daniel Wong. "Wireframe: Supporting data-dependent parallelism through dependency graph execution in gpus." In *Proceedings of the 50th Annual IEEE/ACM International Symposium on Microarchitecture* (MICRO 2017), October 14-18, Boston, Massachusetts, USA.

https://doi.org/10.1145/3123939.3123976

Core Ranking: A, H-index: 114

Citation count: 28

4. Zamani, Hadi, **Devashree Tripathy**, Ali Jahanshahi and Daniel Wong, "<u>ICAP: Designing Inrush Current Aware Power Gating Switch for GPGPU</u>," *IEEE International Conference on Networking, Architecture and Storage (NAS, 2021)*, October 24-26, Riverside, California, USA.

https://doi.org/10.1109/NAS51552.2021.9605434

Citation count: 1

5. Zamani, Hadi, **Devashree Tripathy**, Laxmi Bhuyan, and Zizhong Chen. "<u>SAOU: Safe Adaptive Overclocking and Undervolting for Energy-Efficient GPU Computing</u>" *ACM/IEEE International Symposium on Low Power Electronics and Design (ISLPED, 2020)*, August 10-12, Boston, Massachusetts, USA. (Virtual due to Covid-19). https://doi.org/10.1145/3370748.3406553

Research Impact Score: 3.13, H-index: 61

Citation count: 11

6. Zamani, Hadi, Yuanlai Liu, **Devashree Tripathy**, Laxmi Bhuyan, and Zizhong Chen. "<u>GreenMM: energy efficient GPU matrix multiplication through undervolting.</u>" *ACM International Conference on Supercomputing (ICS 2019)*, June 26-28, Phoenix, Arizona, USA. https://doi.org/10.1145/3330345.3330373

Core ranking: A, H-index: 68

Citation count: 22

- 7. Brumar, Iulian, Rodrigo Rocha, Alex Bernat, **Devashree Tripathy**, David Brooks and Gu-Yeon Wei, "SSAge: SSA-Based Merged Accelerator Generation with ML-Driven Energy Awareness", under review in *ICCD 2022- IEEE International Conference on Computer Design, October 23-26, Lake Tahoe, USA.*
- 8. Chai, Yuji, **Devashree Tripathy**, Paul Whatmough, David Brooks and Gu-Yeon Wei, "PerfSAGE: Generalized Runtime Performance Predictor for Arbitrary Deep Learning Models", under review in *ICCAD 2022- IEEE/ACM International Conference on Computer-Aided Design*, October 29-November 3, Sandiego, CA, USA.

Book

Das, Swagata, **Devashree Tripathy**, and Jagdish Lal Raheja. <u>Real-time BCI System Design to Control Arduino Based Speed Controllable Robot Using EEG.</u> Book Published in *Springer Briefs in Applied Sciences and Technology, 2018.*

https://doi.org/10.1007/978-981-13-3098-8

Softcover ISBN: 978-981-13-3097-1

Impact Factor (2020): 0.26, H-index: 15, SJR Quartile: Q3

Citation count: 12

Book Chapter

Tripathy D., Raheja J.L. (2015) "Design and Implementation of Brain Computer Interface Based Robot Motion Control." In: Satapathy S., Biswal B., Udgata S., Mandal J. (eds) Proceedings of the 3rd International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA) 2014. Advances in Intelligent Systems and Computing, vol 328. **Springer**, Cham. https://doi.org/10.1007/978-3-319-12012-6 32,

RESEARCH ADVISING

- 1. Behzad Boroujerdian, Ph.D. in CS at Harvard University, USA (2021- Present)
 - O Paper in **ACM Transactions TECS 2022-** ACM Transactions on Embedded Computing Systems
- 2. Yuji Chai, Ph.D. in CS at Harvard University, USA (2021- Present)
 - o Paper under review in **ICCAD 2022** *IEEE/ACM International Conference on Computer-Aided Design, October 29-November 3, Sandiego, CA, USA.*
- 3. Iulian Valentian Brumar, Ph.D. in CS at Harvard University, USA (2021- Present)
 - o Paper under review *ICCD 2022- IEEE International Conference on Computer Design, October 23-26, Lake Tahoe, USA.*
- 4. Nasrin Imanpour, Ph.D. in CS at University of South Carolina, USA (2021- Present)
- 5. Abhinav Sharma, Research Intern at University of South Carolina, USA (2021- Present)
- 6. Nisarg Shah, Master's in CE at University of California, Riverside, USA (2017-2018)

TEACHING

Worked as a Teaching assistant during PhD at UCR and Master's at CSIR-CEERI. Taught the following courses:

- 1. UCR-UnderGraduate-CS005: Introduction To Computer Programming (Fall 2019)
- 2. UCR-PostGraduate-CS203: Advanced Computer Architecture (Winter 2018, Winter 2019)
- 3. UCR-PostGraduate-CS213: Multiprocessor Architecture and Programming (Spring 2018, Winter 2020)
- 4. CEERI-PostGraduate-2-219: Advanced Signal and Image Processing (2014)

COURSE WORK AT PG LEVEL

PhD (11 courses at UCR: University of California, Riverside, USA)

✓ High Performance Computing	✓ Compiler Construction	✓ Advanced Operating
		Systems
✓ GPU Architecture and Parallel Programming	✓ Advanced Computer Architecture	✓ Reconfigurable Computing
✓ Design And Analysis of	✓ Data Mining Techniques	✓ Multiprocessor Architecture
Algorithms	Butta Finning Teeninques	and Programming
✓ Seminar in Computer Science	✓ Special Topics in Advanced Computer	
	Science	

MTech (8 courses at CSIR: CSIR-CEERI, Pilani)

✓ Advanced Signal and Image	✓ Power Electronics and AC/DC Drives	✓ Intelligent Sensor Systems
Processing		
✓ Real-Time Embedded System	✓ System Modelling and Design	✓ System Design for Process
Design	Languages	Control Applications
✓ Technical Communication	✓ Project Management	

AWARDS & DISTINCTIONS

- 1. Postdoctoral Fellowship Offers from Penn State University, North Carolina State University (NCSU), University of South Carolina and Harvard University. *Accepted the Harvard University Offer (2021).*
- 2. Anita Borg Grace Hopper Scholarship (2020).
- 3. Student Travel Grant for ACM/IEEE ISCA 2019 (\$600 from NSF), ACM HPDC 2019 (\$1739), ACM/IEEE MICRO 2017 (\$474 from ACM/IEEE), CWWMCA 2017 (\$970), NAS 2016 (\$566), NAS 2021 (\$868).
- 4. Student Volunteer Award ISCA 2018 (\$572).
- 5. Invited by Computing Research Association (CRA), USA to attend Grad Cohort for Women 2018 April 13-14, 2018 at San Francisco, CA, USA.
- 6. Award of Excellence as GPU Modelling Intern at Samsung Austin R&D Center, USA, September, 2018
- 7. Dean's Distinguished Fellowship, UCR 2015-2019 (\$171,233)
- 8. CSIR Quick-Hire Fellowship by Government of India (2012 2014)
- 9. **1st rank/University Topper** among all disciplines in B.Tech. (undergraduate) students at VSSUT, Burla (2012)
 - a. Awarded University Gold medal for University Topper, 2012
 - b. Awarded University Silver medal for best Electronics and Telecommunication Engg graduate at VSSUT Burla, 2012
- 10. 1st rank among all girls in first year B.Tech at VSSUT, Burla (2008)
- 11. Top 0.1% among all the students in India, Central Board of Secondary Education (2006)

PROFESSIONAL ACTIVITIES

Editorial Board

• Journal of Systems Research (JSys), Area Chair: computer architecture and hw/sw interface

Program Committee

• European Conference on Computer Systems (EuroSys'22 Artifacts)

Invited Reviewer

- 1. IEEE Computer Architecture Letters (CAL 2022).
- 2. IEEE Transactions on Computers (TC 2021).
- 3. ACM Transactions on Architecture and Code Optimization (TACO 2021).
- 4. IEEE Transactions on Parallel and Distributed Systems (TPDS 2021).
- 5. IEEE International Conference on Networking, Architecture, and Storage (NAS 2021)
- 6. International Journal of Data Warehousing and Mining (IJDWM)
- 7. International Journal of Computational Science and Engineering, Inderscience (IJCSE 2022)
- 8. International Journal of Knowledge Discovery in Bioinformatics (IJKDB 2018)
- 9. International Journal of Rough Sets and Data Analysis (IJRSDA)

External Reviewer

1. Design, Automation and Test in Europe Conference (DATE 2021)

- 2. ACM/IEEE International Symposium on Low Power Electronics and Design (ISLPED 2018-2020)
- 3. International Conference on Parallel Processing (ICPP 2018)
- 4. IEEE International Conference on Computer Design (ICCD 2017, ICCD 2021)

REFERENCES

Dr. Laxmi N. Bhuyan (PhD Advisor)

Distinguished Professor Emeritus Computer Science and Engineering University of California, Riverside, USA bhuyan@cs.ucr.edu

Office: 351 Winston Chung Hall, University of California,

Riverside, CA 92521

Phone: (951) 827-2281, Fax: (951) 827-4643

Dr. Daniel Wong

Assistant Professor Department of Electrical and Computer Engineering University of California, Riverside, USA daniel.wong@ucr.edu

Office: 425 Winston Chung Hall

University of California, Riverside

Riverside, CA 92521 Phone: (951) 827-2986

Dr. Manoranjan Satpathy

Associate Professor School of Electrical Sciences Indian Institute of Technology, Bhubaneswar manoranian@iitbbs.ac.in

Office: SES-221, School of Electrical Sciences,

IIT Bhubaneswar Phone: 674-713-5728

Dr. David Brooks (Postdoc Advisor)

Haley Family Professor Department of Computer Science Harvard University, USA dbrooks@eecs.harvard.edu

Address: Maxwell Dworkin 141 33 Oxford Street Cambridge MA 02138

Phone: (617) 495-3989, Fax: (617) 496-6404

Dr. Gu-Yeon Wei

Robert and Suzanne Case Professor of Electrical Engineering and Computer Science Harvard University, USA guveon@seas.harvard.edu

Office: 150 Western Ave, Sci&Eng 5.413

Phone: (617) 384-8131

Dr. Pradip Bose

Distinguished Research Staff Member & Manager Thomas J. Watson Research Center IBM, USA

pbose@us.ibm.com

Office: IBM T. J. Watson Research Center P. O. Box 218 (Route 134) Yorktown Heights, NY 10598 Tel: (914) 945-3478, Fax: (914) 945-2141

Phone: +1-914-945-3478