Curriculum Vitae - Santosh Pandey

PhD Candidate
Electrical and Computer Engineering
Rutgers University, NJ

santosh.pandey@rutgers.edu
Personal Homepage

RESEARCH INTERESTS

My research interests lie at the intersection of systems and ML, focusing on full circle innovation–using ML to design better systems (more current focus) and building optimized systems to advance ML/HPC applications. Notably, I have devised novel techniques for creating accurate and reusable ML-based microarchitecture simulators, while also accelerating and scaling these simulators for performance modeling with hardware software co-design, demonstrating the synergy between system optimization and ML.

EDUCATION

Jan. 2023-Present — **Ph.D.** in **Computer Engineering**
Department of Electrical & Computer Engineering
Rutgers University, USA
Advisor: Hang Liu

Aug. 2019 - Dec. 2022 — **Ph.D.** in **Computer Engineering**
Department of Electrical & Computer Engineering
Stevens Institute of Technology, USA
Advisor: Hang Liu [**Master's degree awarded**]

2012 - 2016 — **B.E.** in **Computer Engineering**
Tribhuvan University, Nepal
Thesis Advisor: Subarna Shakya

PROFESSIONAL EXPERIENCES

May. 2024-Aug. 2024 — **Student Researcher | SystemResearchGroup@Google**
Accelerating explainable machine learning based computer architecture simulation.
Mentors: Victor Lee, Amir Yazdanbakhsh & Mohammad Alizadeh

May. 2020-Aug. 2022 — **Research Intern | Brookhaven National Lab**
Research collaboration on machine learning based computer architecture simulation.
Mentors: Lingda Li & Adolfy Hoisie

May 2019-Aug. 2019 — **Research Internship | Lawrence Berkeley National Lab**
Worked on accelerating graph algorithms with GPUs.
Awarded Graphchallenge Champion.
Mentor: Xiaoye Sherry Li

HONORS & AWARDS

2022 IEEE TCHPC Student Travel Award
2019 Champion of MIT Graph Challenge Competition
2016 Research Grant for HPC, Nepal Academy of Science and Technology (NAST)
2012 Full Academic Scholarship for Undergraduate College

PUBLICATIONS & PATENTS(**Total citations: 535**)

Patent

2024 Hang Liu and **Santosh Pandey**. "Accelerating Microarchitecture Simulation with Machine Learning" (RU Docket 2024-101). U.S. Provisional Application 63/539,950 on April, 2024.

Selected publications

2024 **Santosh Pandey**, Amir Yazdanbakhsh, and Hang Liu. TAO: Re-Thinking DL-based Microarchitecture Simulation. In *the Proceedings of the ACM on Measurement and Analysis of Computing Systems (**SIGMETRICS**)*, 2024. (Acceptance rate = 10.7%)

2022 **Santosh Pandey**, Lingda Li, Thomas Flynn, Adolfy Hoisie, Hang Liu. Scaling Deep Learning-based Microarchitecture Simulation on GPUs. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (**SC**)*. ACM, 2022. (Acceptance rate = 25.3%)

2021 **Santosh Pandey***, Zhibin Wang*, Sheng Zhong, Chen Tian, Lingda Li, Adolfy Hoise, Xiaoye S. Li, Caiwen Ding, Dong Li, Bolong Zheng and Hang Liu. TRUST: Triangle Counting on GPUs. In *the Transactions on Parallel and Distributed Systems (**TPDS**)*. IEEE, 2021.

2020 **Santosh Pandey**, Lingda Li, Adolfy Hoisie, Xiaoye S. Li and Hang Liu. C-SAW: A Framework for Graph Sampling and Random Walk on GPUs. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (**SC**)*. IEEE, 2020. (Acceptance rate = 25.1%)

2019 **Santosh Pandey**, Xiaoye S. Li, Aydin Buluc, Jiejun Xu and Hang Liu. H-INDEX: Hash-Indexing for Parallel Triangle Counting on GPUs. In *the High Performance Extreme Computing **(HPEC)**, Graphchallenge*. IEEE, 2019. (Awarded Champion).

Other publications

2024 **Santosh Pandey**, Amir Yazdanbakhsh, Hang Liu. Fast DL-based Simulation with Microarchitecture Agnostic Traces and Instruction Embeddings. In *Workshop on ML for Computer Architecture and Systems (**ISCA**)*. 2024.

2024 Chengying Huan, Yongchao Liu, Heng Zhang, Shuaiwen Song, **Santosh Pandey**, Shiyang Chen, Xiangfei Fang, Yue Jin, Baptiste Lepers, Yanjun Wu, Hang Liu. TEA+: A Novel Temporal Graph Random Walk Engine with Hybrid Storage Architecture. In *the Proceedings of ACM Transactions on Architecture and Code Optimization (**TACO**)*, 2024.

2023 Chengying Huan, Shuaiwen Leon Song, **Santosh Pandey**, Hang Liu, Yongchao Liu, Baptiste Lepers, Charles He, Kang Chen, Jinlei Jiang, Yongwei Wu. TEA: A General-Purpose

Temporal Graph Random Walk Engine. In *Proceedings of the European Conference on Computer Systems (**Eurosys**)*. ACM, 2023.

2022 Lingda Li, **Santosh Pandey**, Thomas Flynn, Hang Liu, Noel Wheeler, Adolfy Hoisie. SimNet: Computer Architecture Simulation using Machine Learning. In *the Proceedings of the ACM on Measurement and Analysis of Computing Systems (**SIGMETRICS**)*, 2022.

2021 Shiyang Chen, Shaoyi Huang, **Santosh Pandey**, Bingbing Li, Guang Gao, Long Zheng, Caiwen Ding and Hang Liu. E.T.: Rethinking Transformer Models on GPUs. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (**SC**)*. ACM, 2021.

2020 Bingbing Li, **Santosh Pandey**, Haowen Fang, Yanjun Lyv, Ji Li, Jieyang Chen, Mimi Xie, Lipeng Wan, Hang Liu, and Caiwen Ding. FTRANS: Energy-Efficient Acceleration of Transformers using FPGA. In *Proceedings of the International Symposium on Low Power Electronics and Design (**ISLPED**)*. ACM/IEEE, 2020.

2019 Anil Gaihre*, **Santosh Pandey***, Hang Liu. Deanonymizing cryptocurrency with graph learning: The promises and challenges. In *the Conference on Communications and Network Security (**CNS**)*. IEEE, 2019.

2019 **Santosh Pandey**, Gopal Ojha, Bikesh Shrestha, Rohit Kumar. BlockSIM: A Practical Simulation Tool for Optimal Network Design, Stability and Planning. In *the International Conference on Blockchain and Cryptocurrency (**ICBC**)*. IEEE, 2019.

2017 Sadhu Ram Basnet, Ram Sharan Chaulagain, **Santosh Pandey**, Subarna Shakya. Distributed high performance computing in openstack cloud over sdn infrastructure. In *the International Conference on Smart Cloud (**SmartCloud**)*. IEEE, 2017.

2017 Ram Sharan Chaulagain, **Santosh Pandey**, Sadhu Ram Basnet, Subarna Shakya. Cloud based web scraping for big data applications. In *the International Conference on Smart Cloud (**SmartCloud**)*. IEEE, 2017.

Talks and Presentations

Jun. 2024 **ISCA MLArchSys**: Fast DL-based Simulation with Microarchitecture Agnostic Traces and Instruction Embeddings

Mar. 2023 **Princeton University**: Machine Learning for Computer Architecture Design

Mar. 2023 **University of North Texas**: GPU-Accelerated Graph Sampling

Nov. 2022 **IEEE/ACM SC**: Scalable Deep Learning-based Microarchitecture Simulation on GPUs

Nov. 2020 **IEEE/ACM SC**: A Framework for Graph Sampling and Random Walk on GPUs

Sep. 2019 **IEEE HPEC**: H-index: Hash-indexing for Parallel Triangle Counting on GPUs

Jun. 2019 **IEEE CNS**: Deanonymizing Cryptocurrency with Graph Learning: The Promises and Challenges

Nov. 2017 **IEEE SmartCloud**: Cloud-based Web Scraping for Big Data Applications

STUDENT ADVISING & MENTORING

Undergraduate Students

2022 Christian O'Connell (2022.03 - 2022.11)
Topic: Performance prediction of contemporary hardware.

TEACHING ASSISTANT

2024 Digital System Design

2023 Programming Methodology I

2019 Operating Systems

PROFESSIONAL ACTIVITIES

Reviewer

2024 ISCA MLArchSys Workshop

2024 TPDS

2024 OSDI AE, ATC AE

2023 IEEE Micro AE

2022 IEEE Big Data GTA[3]

2022 ACM PPoPP AE

2020 IEEE ICDCS

Program Committee

2023 IEEE Big Data GTA[3]

2022 IEEE Big Data GTA[3]