

MASTER'S STUDENT $\,\cdot\,$ Department of Electrical and Computer Engineering

2501 Speedway, EER 5.860, Austin, TX 78712, United States

□ (+1) 737-217-5056 | snehilv@utexas.edu | \$ snehilverma41@gmail.com | \$ snehilverma41.github.io

Education _

The University of Texas at Austin

4.0*/4.0

M.S. IN ELECTRICAL AND COMPUTER ENGINEERING

TRACK: ARCHITECTURE, COMPUTER SYSTEMS, AND EMBEDDED SYSTEMS (ACSES)

Fall 2018 - Present

Indian Institute of Technology, Kanpur

8.9/10

B.Tech in Electrical Engineering (with DISTINCTION)

Minor in Computer Systems. Computer Science and Engineering

Fall 2014 - Spring 2018
* Calculated at the end of Spring'19

Publications

- Snehil Verma, Qinzhe Wu, Bagus Hanindhito, Gunjan Jha, Eugene John, Ramesh Radhakrishnan, and Lizy Kurian John, "Metrics for Machine Learning Workload Benchmarking," International Workshop on Performance Analysis of Machine Learning Systems (FastPath), In conjunction with International Symposium on Performance Analysis of Systems and Software (ISPASS 2019), Madison, USA, 2019. [Publication] [Presentation]
- Ramesh Radhakrishnan, **Snehil Verma**, Qinzhe Wu, Bagus Hanindhito, Gunjan Jha, Eugene John, and Lizy Kurian John, "Demystifying Hardware Infrastructure Choices for Deep Learning Using MLPerf," GPU Technology Conference (GTC), *NVIDIA's Deep Learning & AI Conference 2019*, Silicon Valley, USA, 2019. [Presentation]
- Snehil Verma, Nayan Deshmukh, Prakhar Agrawal, Biswabandan Panda, and Mainak Chaudhuri, "DFCM++: Augmenting DFCM with Early Update and Data Dependence-driven Value Estimation," 1st Championship Value Prediction (CVP-1), In conjunction with 45th International Symposium on Computer Architecture (ISCA 2018), Los Angeles, USA, 2018. [Publication] [Presentation] [Code]

Experience and Projects ____

Power and Performance Debugging via Emulation

Samsung SARC | ACL

Summer'19

GPU PERFORMANCE INTERN MENTORED BY RAGHAVAN R. SRINIVASA

Qualitative and Quantitative analysis of the MLPerf benchmark suite

UT Austin Fall'18 - Present

GRADUATE RESEARCH ASSISTANT AT LAB FOR COMPUTER ARCHITECTURE UNDER PROF. LIZY K. JOHN

• FastPath, ISPASS'19: Proposed a new metric for benchmarking ML workloads that consider time and accuracy from the perspective of comparing the hardware used for training. Showed that merely taking into account the time for training to multiple thresholds makes the metric less sensitive to the specific threshold chosen and the seed values

- NVIDIA GTC'19: An extensive study on the impact of hardware infrastructure choices on deep learning performance.
 Presented quantitative analysis on various configurations of Dell systems with NVIDIA GPUs using MLPerf [v0.5]
- Under review: Analyzed and characterized the MLPerf [v0.5] benchmark suite exposing various system-level trends

Improving Data Locality by Kernel Fusion in DNNs [PRESENTATION]

UT Austin

COURSE PROJECT FOR COMPARCH: PARALLELISM AND LOCALITY UNDER PROF. MATTAN EREZ

Spring'19

- Studied Convolutional Sequence to Sequence Learning model for translation, a part of Facebook AI Research Sequence to-Sequence Toolkit (FairSeq) implemented using PyTorch (lacks support for explicit memory management)
- Explored various methods of performing kernel fusion involving libraries like CUTLASS, cuBLAS, and cuDNN
- Integrated our C++/CUDA extensions with PyTorch and showed ~2× reduction w.r.t the global memory/L2\$/DRAM writes

Graph Placement Optimization on a HMS [REPORT] [PRESENTATION]

UT Austin

COURSE PROJECT FOR COMPARCH: USER SYSTEM INTERPLAY UNDER PROF. MATTAN EREZ

Fall'18

- Proposed a novel optimization technique that **statically** makes **fine-grain placement** decisions based on the natural properties of a graph: the number of incoming/outgoing edges, topology, frontier composition
- Modified a light-weight shared memory graph processing framework (Ligra) to incorporate the proposed method
- Evaluated the same, demonstrating its good adaptability and up to 2× performance improvement

Value Prediction: DFCM++ [PUBLICATION] [PRESENTATION] [POSTER] [CODE]

IIT Kanpur

Undergraduate Researcher under Prof. B. Panda and Prof. M. Chaudhuri

Spring'18

- Proposed a series of enhancements on top of existing DFCM predictor: Early Update, Value Estimator, PC Blacklister, and Dynamic Context Length. The design achieved an IPC improvement of 28.1% with respect to the baseline, i.e, without any value predictor, and 40.2% in comparison to the base DFCM
- Presented at 1st Championship Value Prediction (CVP-1), ISCA'18 and secured second position in the unlimited track

Perceptron Learning driven Cache Replacement policy [REPORT]

Texas A&M University

VISITING RESEARCH SCHOLAR AT HIGH PERFORMANCE COMPUTING LAB UNDER PROF. EUN J. KIM

Summer'17

- Familiarized myself with various cache performance improvement techniques such as **Reuse Prediction**, **Inclusive**Cache Management and Sharing Awareness Cache Management
- Proposed Coherence-Aware Reuse Prediction that achieved a geometric mean speedup of 20% over LRU and resulted in a 40% drop in average MPKI with respect to LRU, for 4 MB LLC

Emerging Non-Volatile Memory [PRESENTATION] [TERM PAPER]

IIT Kanpur

COURSE PROJECT UNDER PROF. YOGESH S. CHAUHAN AND PROF. BAQUER MAZHARI

Spring'18

[REPORT]

[REPORT]

[REPORT]

[REPORT]

- Studied various emerging flexible non-volatile memory technologies like ReRAM, FeRAM, PCRAM, and Flash
- Delved into the approaches for making flexible NVMs, their operating principles, and some common architectures
- Performed a literature survey on binary metal-oxide resistive switching RAM. The study encompasses the switching mechanism, design and electrical characteristics of various binary metal-oxide ReRAMs

Other Projects IIT Kanpur

- Designed a stable **PLL** with low power linear Current Starved VCO consuming a maximum power of 182 µW
- Studied the noise cancellation techniques and effectively applied them to design a 2.4 GHz Inductorless LNA
- Employed adaptive biasing to design a two-stage folded cascode OTA suitable for large capacitive loads
- Implemented a BSIM4-like model on **Verilog-A** and extracted parameters using **IC-CAP** simulation software
- Designed a H_{\infty} controller and tuned it for desired performances in order to make the system robust
- Built an **all-terrain vehicle** capable of autonomous navigation using Embedded Systems and Google Maps API
- Selected among the **top 5 best ideas** for a game developed using Unity3D Game Engine for **Microsoft Code.Fun.Do**

Technical Skills

Programming languages

Tools / Platforms

C, C++, C#, CUDA, Regent, Java, Python, Bash, Perl, Verilog(-A), HSPICE perf, NVProf, NVVP, CACTI, PAPI, SimPoints, PINTool, Docker, Git, 上下, Cadence Virtuoso, Synopsys, Silvaco (Athena and Atlas), PSPICE, Microcap, Mentor Graphics, Ardupilot, Arduino, Processing, MATLAB, GNU Octave

Selected Coursework

UT Austin

Computer Architecture* Comp Arch: User System Interplay*

Comp Arch: Parallelism and Locality* Superscalar Microprocessor Architecture*

IIT Kanpur

Computer Architecture Microelectronics-I (Circuits), II (Devices)
Modern Memory Systems* Digital Electronics
Principles of Data Base Systems Analog/Digital VLSI Circuits*
Data Structures and Algorithms Compact Modeling*

Introduction To Flexible Electronics* S/C Optical Communication Devices*

Power Electronics
Robust Control Systems*

* indicates Graduate Level Courses

Scholastic Achievements _

- Recipient of the Professional Development Award by UT Austin for presenting research at FastPath, ISPASS'19
- Secured second position in the unlimited track of $\mathbf{1}^{st}$ Championship Value Prediction, ISCA'18
- Awarded with a travel grant of \$1500 by Microsoft Research Labs, India for attending and presenting at CVP-1, ISCA'18
- Recipient of the ISCA 2018 Student Travel Grant Award and the Departmental (E.E.) Travel Grant Award, IIT Kanpur
- Recipient of **TAMU-IITK summer undergraduate research scholarship 2017** (awarded to two students per branch)
- Received Academic Excellence Award for outstanding academic performance (awarded to top 7% students in the institute) for the academic years 2014-15 and 2016-17
- Received A* grade in 3 courses, including Electrical Engineering Lab I (awarded to top 1-2% students in a course)
- Secured **All India Rank 387** in **JEE Advanced 2014**, amongst 120 thousand successful candidates selected from over 1.4 million aspirants who appeared for JEE Mains 2014
- Selected for the **Kishore Vaigyanik Protsahan Yojana (KVPY)** Scholarship in the year 2014, funded by the Department of Science and Technology, Government of India, and secured **All India Rank 236** in the national test
- Awarded Certificate of Merit by HBCSE in International Chemistry Olympiad 2013-14 at the National Level

Teaching Experience

Academic Mentor IIT Kanpur

INSTITUTE COUNSELLING SERVICE

Fall'15 - Spring'16

• Tutored students having difficulties in **Engineering Design and Graphics** by conducting institute level remedial classes and doubt-clearing sessions. Personally mentored academically weaker students to cope with their academic load